Mechanistic and inhibition studies on a nucleophilic cysteine protease and transpeptidases



Tika Ratna Malla
Schofield Group

Exeter College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

September 2022

To my parents
Jamuna and Bidur

जब म चढेर भावको डुंगा
सयर गर्छु स्मृतिको प्रिय गंगा
स्वप्न गगनबाट सुशीतल जल
सुख-स्मृतिको बर्सिन्छ रिमझिम
मेरो प्यारो ओखलढुंगा !

कुसुमाकरको आँचल पक्री
मलाई कति खोज्दी हुन् प्रकृति ;
मृग सावकहरू उफ्री-उफ्री
तरुतलिकाहरु हल्ली हल्ली,
मेरो प्यारो ओखलढुंगा !

ती अविरल परिश्रमी पतकार
द्रुतगामी निर्झर तुल्य चपल,
बहँदै वनवनमा चारैतिर
खोज्दा हुन् मेरो पथको पथ !
मेरो प्यारो ओखलढुंगा !

पर्दावाल सुन्दर पल्लवित्
वृक्ष खोपिमा, कालो कोकिल्
उनको यो वियोग प्रिय कविको
कति मृदु विरह गीत रच्दो हो !
मेरो प्यारो ओखलढुंगा !

त्यो ढुङ्गा, त्यो खोला, त्यो रुख्,
त्यो बन्, त्यो लतिका, त्यो चिर मुख्
यो हृदय केन्द्र स्नेह तडित्
चम्काइरहेछ आज क्षण क्षण
मेरो प्यारो ओखलढुंगा !

तामाकोशी, त्यो सुनकोसी
त्यो गिरि फोरी बहने रोशी
मेरो हृदयको तिम्रो चित्र
पारिरहन्छ हर्दम पवित्र
मेरो प्यारो ओखलढुंगा !

भाग्य-लहरामा लहरी लहरी
पुगें म यस मरूस्थलमा कसरी !
तर खेद छैन तिम्रो आकृति
लेखिएको छ यो हृदयभरि ;
मेरो प्यारो ओखलढुंगा !

सिद्धिचरण श्रेष्ठ – मेरो प्यारो ओखलढुंगा

A vow to embrace lifelong learning and perseverance, acquire new skills and experiences, explore and travel, question and challenge, reflect and adapt, and change.

Acknowledgements

I am immensely grateful to Prof. Chris J. Schofield and Dr. Chris Lohans for their supervision and for providing me with an opportunity to be involved in a diverse set of collaborations throughout this project. I shall always remain thankful for the support and guidance I received and the mentoring that I continue to receive. For the main protease COVID project, I am bound to miss thanking many of the collaborators I had the pleasure of working with so I shall start with an apology for it being unintentional. Prof. Chris J. Schofield, looking back I shall always feel honoured to have embarked on the COVID project in the middle of my DPhil and for sharing a common vision that we could attempt to contribute to the community by re-purposing β-lactam antibiotics and existing drugs. This project has been one of the greatest highlights and a much-appreciated learning experience. Dr. Tony Tumber, thank you for letting me learn from you and gain an invaluable experience in high throughput assays and mass spectrometric techniques. I also cannot not mention the many RapidFire and Mass Spectrometer troubleshooting adventures. One cannot help but learn from your attributes of patience and courteousness. I am also greatly thankful to Toby for undertaking the peptide synthesis work and for his willingness to indiscriminately offer help at all times. The assay development work could not have been pioneered without your generous help. Thank you to Dr. Lennart Brewitz for being a wonderful collaborator on synthesising the β-lactam compounds and being a great support to bounce off many wacky ideas for this project and beyond. It has been great to find a mentor in you. I should also like to thank Dorian Munetean, Hiba Aslam, Ziggy, Tim and Dr. Katherine England and Eidarus Salah for compound synthesis and protein production, respectively.

I would also like to express my gratitude towards the team at Diamond Light Source: Prof. Martin Walsh, Prof. Frank Von Delft, Anthony Aimon, David Owen, Claire Strain-Damerell, Petra Lukacik and Lizbe for collaborating on the COVID Moonshot Consortium. For virtual collaborative work with the Modelling team, I'd like to thank Prof. Fernanda Duarte, Prof. Garett Morris, Dr. Deborah Shoemacker, Henry, Marc, Rebecca, Tristan and the rest of the team. A large part of my gratitude is owed to Prof. Hiroaki Suga group. I have been very fortunate to work with international peers, Marina Kawai, Takashi Miura and Wenyu Liu who relentlessly applied their newly developed protocol to synthesize macrocyclic inhibitors. This was a wonderful introduction to peptide chemistry. I am also thankful to Prof. Akane Kawamura and her team for the extended collaboration on macrocyles.

Though a lot of collaboration on the Ldt project has not been enlisted in this thesis, regardless of it I would like to thank my wonderful collaborators and colleagues: I am thankful to Dr. Mike McDonough for his supervision on the crystallography aspects of my project. Alen Krajnc and Tim Suits for undertaking synthesis of natural substrates of Ldts and lactivicin, respectively. I am also thankful to my collaborator Prof. Cynthia Dacquah from the University of Ghana for the synthesis of thiol compounds, and preparation and transport of plant extract to test activity against Ldts. I am also thankful to our collaborators Dr. Bina Singh from the Indian Institute of Technology and Dr. Amit K. Patra for producing and purifying an oxygenase enzyme to produce sulfonic acids, for investigation with Ldts. I would also like to thank our lab manager Dr. Adam Hardy for the smooth operation of our lab. Thanks to the Department of Chemistry for their efforts to remain operational during all times during the pandemic and for supporting our work; my utmost gratitude to the cleaning staff that selflessly ensured a safe working environment.

I also remain grateful to Prof. Tim Claridge and Dr. Nader Amin for their technical support with NMR experiments. I also thank Mr. Manso Jubier for generously preparing media for protein production and disposing of the biohazard wastes safely, and the running of the fermentation lab. I am also thankful to all my colleagues, present and past who all kindly shared their expertise with me in the lab, especially on long weekends. Pratheesh, Xiao, Toby, Raf, Klem, Sam, Noelle, Lennart and other members of CJS group, thanks for making our lab a wonderful community to work and celebrate in. Pratheesh, thank you for sharing your friendship and your meals during long days in the lab and for introducing me to Sri Lankan spices. I shall endeavour to use them in my cooking. Thank you to Razia, Ellen, Rachel, Shihong, Tati, Sofi, Josh, Andre, and other friends from MCR and DTP for being such a joyous company. Thomas, thank you for your unfailing invites to your socials. I am also thankful to Exeter College. Being part of the MCR was a great way to acclimatise to Oxford so thanks for being such a welcoming college and being home to the magnificent view of Oxford from the mount. I shall also like to thank my second college of affiliation, Worcester College including the catering team and the gardening team for the nourishment and the scenic college grounds. It has been a privilege to live in this college for the majority of my degree. Thank you to the Worcester College Welfare team for being such an excellent team to work with. Neva, Georgie, Shazeaa, Kris, and Max, I am thankful for your friendship. Linacre Ladies that Lift, thank you for being such a great community to power-lift and uplift each other. Shez, thank you immensely for your coaching. Prof. Esther Becker and Prof. Gail Preston, thank you for leading a supportive and collaborative Doctoral Training Partnership programme. I feel honoured and extremely fortunate to have been a part of this cohort and to benefit from your sound guidance throughout my DPhil journey.

I remain most grateful to my family and friends who are the pillar of support and strength in everything I attempt. My parents and my late grandmother will always be my biggest inspiration. My dear siblings, dijju, and daji, thank you for the food drops, your support, and your kind conversations. Nita, Jess, Oumie, Alina, Reeya, Mina, and Matt, thank you for being there for me.

Acknowledgements: Funding

I'd like to express my gratitude to all the funding sources that have enabled me to undertake my DPhil. Notably, I'd like to thank the Oxford Interdisciplinary Bioscience DTP and BBSRC for the studentship and the grants and scholarships from the Eurofins, the Department of Chemistry in Oxford and the Exeter College.

Declaration of Authorship

This body of work was carried out by myself unless otherwise stated. My work has been a highly collaborative project, especially on work on M^{pro}. Collaborators have been duly credited. I, Tika Malla, hereby confirm that this thesis has not been submitted in whole or in part for any other academic degree or professional qualification.

Tika Malla
Oxford, September 2022

Abstract

Enzymes with a nucleophilic cysteine residue catalyse important reactions in biology, including proteolysis, redox reactions, and radical transfer reactions. Nucleophilic cysteine enzymes are validated medicinal chemistry targets, e.g. thioredoxin reductase and glutathione reductase. The work presented in this thesis describes biochemical and inhibition studies on viral and bacterial nucleophilic cysteine enzymes. Chapters 2-6 focus on inhibitors of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (M^{pro}), which is a therapeutic target to treat coronavirus disease 2019 (COVID-19). Chapter 7 describes studies focusing on the mycobacterial L,D-transpeptidases, which are investigational targets for tuberculosis treatment.

Antiviral therapies to treat infections caused by SARS-CoV-2 were lacking at the onset of the COVID-19 pandemic in December 2019. In March 2020, I initiated a collaborative research program in the Schofield group to identify efficient and selective inhibitors of SARS-CoV-2 M^{pro}, which is essential for viral replication and is structurally distinct from human proteases, making it a potential target for antiviral therapies. Note that at the end of 2021, a small molecule M^{pro} inhibitor, i.e. nirmatrelvir, was approved for clinical use in humans.

Initially, a mass spectrometry (MS)-based label-free SARS-CoV-2 M^{pro} assay was developed to investigate the effect of small molecules on catalysis and to identify efficient inhibitors. Screens of commercial small molecules and compounds available in the Schofield group revealed that β-lactams can be effective M^{pro} inhibitors, a finding of interest considering the established safety profiles of β-lactams. Structure-activity relationship (SAR) studies, in combination with MS and crystallographic studies, helped to improve the potency of the β-lactams as M^{pro} inhibitors and provided evidence for their covalent reaction with the active site cysteine. Apart from β-lactams, compounds containing reactive sulfur (e.g., ebsulfur derivatives), substrate-competitive linear peptides, and heterocyclic molecules (characterised as part of the COVID Moonshot Consortium) were identified as promising M^{pro} inhibitors.

Macrocyclic peptide M^{pro} inhibitors, which tightly bind the M^{pro} active site (e.g., TMGM4-05), were developed in collaboration with Prof. Suga's group (University of Tokyo). SAR studies led to the identification of residues and intrapeptide interactions crucial for potent M^{pro} inhibition. Efforts to develop cell-permeable macrocyclic peptides are ongoing and are expected to enable their use as potential antivirals for the treatment of COVID-19.

In addition to the SARS-CoV-2 M^{pro} inhibition work, the research investigated the fragmentation of β-lactam antibiotics by mycobacterial L,D-transpeptidase nucleophilic cysteine enzymes, and analysed the extent to which these pathways resemble those catalysed by serine-β-lactamases and metallo-β-lactamases. In the case of Ldt_{Mt2} and aminopenicillins, contrary to the expected hydrolytic mechanism, non-hydrolytic C5-C6 fragmentation and cyclisation via rearrangement to give a diketopiperazine were observed. Faropenem undergoes C5-C6 fragmentation on reaction with Ldt_{Mt}s, whereas an intact hydrolysed product is formed upon reaction of faropenem with metallo-β-lactamases (MBLs) and serine-β-lactamases. Interestingly, a degradation product in which the faropenem tetrahydrofuran ring is opened is observed with serine-β-lactamases; the THF ring remains closed in the cases of the metallo-β-lactamases, VIM-1 and VIM-2, but not NDM-1 and L1.

In summary, the results have identified new types of inhibitors of nucleophilic cysteine enzymes and have provided insight into their mechanism of action.

Publications

This thesis is based in part on the following publications:

1. Malla, T.R. *et al.* Mass spectrometry reveals potential of beta-lactams as SARS-CoV-2 Mpro inhibitors.*Chem. Com.*, **2021**, *57(12)*, 1430-1433.

   Contribution: Conceptualisation, methodology development, data generation and analysis, visualisation, drafting.

2. Malla, T.R.[1], Brewitz, L.[1] *et al.* Penicillin Derivatives Inhibit the SARS-CoV-2 Main Protease by Reaction with Its Nucleophilic Cysteine. *J. Med. Chem.*, **2022**, *65(11)*, 7682–7696.

   Contribution: Conceptualisation, methodology development, data generation, and analysis, visualisation, review, and editing.
   Reproduced with permission from the American Chemical Society.

3. Chan, H.H.[1], Moesser, M.A.[1], Walters, R.K.[1], Malla, T.R.[1] *et al.* Discovery of SARS-CoV-2 Mpro Peptide Inhibitors from Modelling Substrate and Ligand Binding. *Chem. Sci.*, **2021**, *12(41)*, 13686-13703.

   Contribution: Conceptualisation, methodology development, data generation, and analysis, visualisation, drafting.
   Reproduced with permission from the Royal Society of Chemistry.

4. Miura, T., Malla, T.R. *et al.*, In vitro selection of macrocyclic peptide inhibitors containing cyclic γ-amino acids targeting SARS-CoV-2 main protease. *Nat. Chem.* **2023**, 1-8

   Contribution: Methodology development, data generation, and analysis, visualisation, drafting.

5. Thun-Hohenstein, S. T.[1], Suits, T. F.[1], Malla, T.R. *et al.* Structure-Activity Studies Reveal Scope for Optimisation of Ebselen-Type Inhibition of SARS-CoV-2 Main Protease *ChemMedChem*, **2022**, *17(4)*, e202100582.

   Contribution: Conceptualisation, methodology development, Data generation and analysis, investigation, visualisation, drafting

6. Jose, A., The Covid Moonshot Consortium[2] *et al.*Expanding the Repertoire of Low-Molecular-Weight Pentafluorosulfanyl-Substituted Scaffolds. *ChemMedChem*, **2022**, *17(7)*, e202100641.

   Contribution: Data generation and analysis, investigation, visualisation, review, and editing

7. The Covid Moonshot Consortium[2] *et al.*COVID moonshot: open science discovery of SARS-CoV-2 main protease inhibitors by combining crowdsourcing, high-throughput experiments, computational simulations, and machine learning. BioRxiv (2020)

   Contribution: Data generation and analysis, investigation, review and editing

8. Lucic, A.[1], Hinchliffe, P.[1], Malla, T.R.[1] *et al.* Faropenem reacts with serine and metallo-beta-lactamases to give multiple products. *Eur. J. Med. Chem.*, **2021**, *58(7)*, e202100641*215*, 113257

   Contribution: Data generation and analysis, investigation, visualisation, drafting

9. Lucic, A., Malla, T.R. *et al.* Studies on the Reactions of Biapenem with VIM Metallo β-Lactamases and the Serine β-Lactamase KPC-2.*Antibiotics*, **2022**, *11(3)*, 396.

   Contribution: Data generation and analysis, investigation, visualisation, review, and editing

1 2

[1]= These authors contributed equally

[2]= Contributed as a part of the COVID Moonshot Consortium

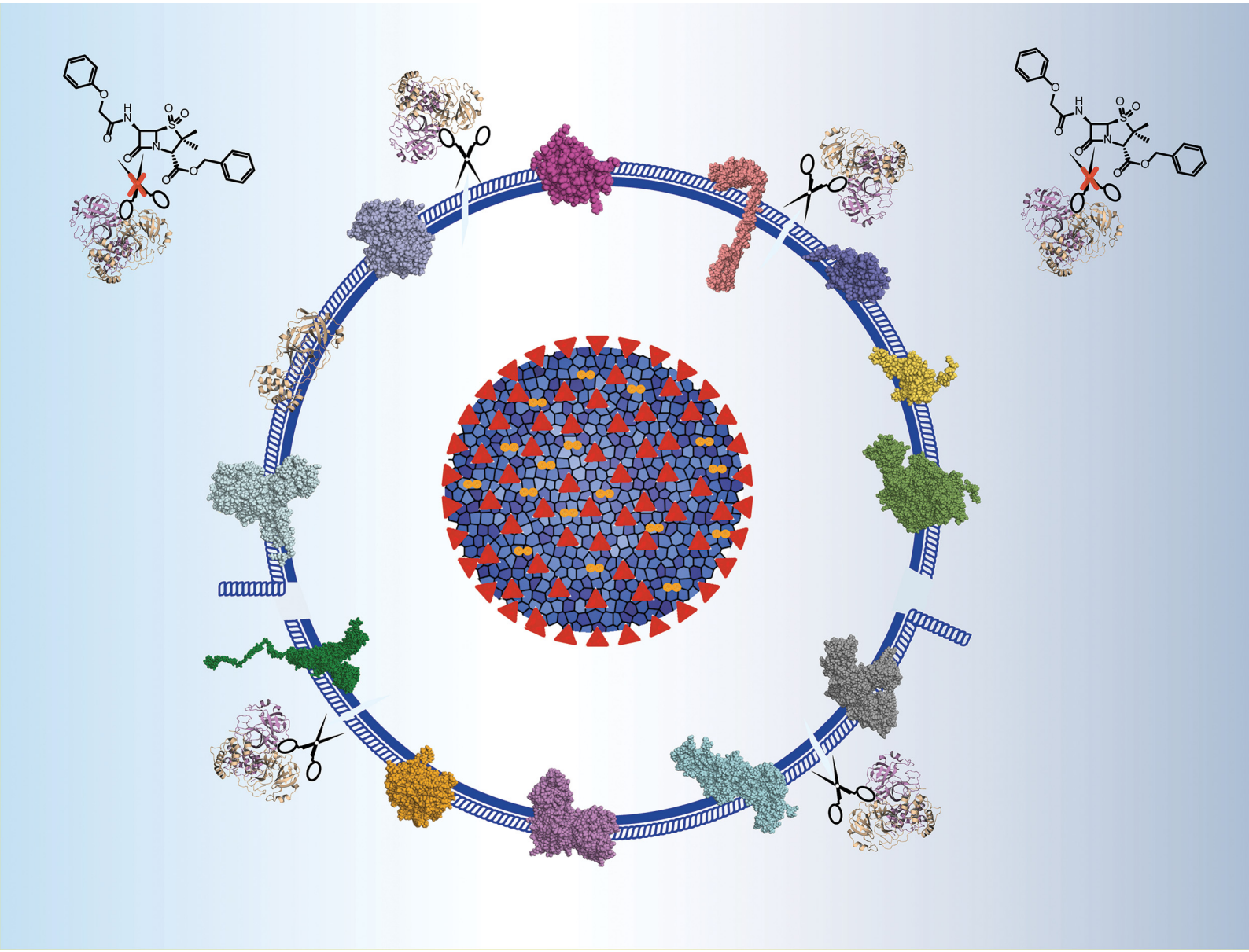
Manuscripts 1-9 are attached herein. Manuscript 3 is reproduced with permission from Royal Society of Chemistry. Manuscripts 4 and 7 are currently under review. All other manuscripts are open access articles with permission to reuse with proper citations.

Volume 57
Number 12
11 February 2021
Pages 1403–1528

ChemComm

Chemical Communications

rsc.li/chemcomm



ISSN 1359-7345

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

Christopher J. Schofield *et al.*

Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 M^{pro} inhibitors

ChemComm

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Commun.*, 2021, **57**, 1430

Received 14th October 2020,
Accepted 23rd November 2020

DOI: 10.1039/d0cc06870e

rsc.li/chemcomm

Open Access Article. Published on 19 January 2021. Downloaded on 7/6/2023 11:02:41 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 M^{pro} inhibitors†

Tika R. Malla, [a] Anthony Tumber,[a] Tobias John, [a] Lennart Brewitz, [a] Claire Strain-Damerell, [bc] C David Owen, [bc] Petra Lukacik, [bc] H. T. Henry Chan, [a] Pratheesh Maheswaran, [a] Eidarus Salah,[a] Fernanda Duarte, [a] Haitao Yang,[d] Zihe Rao,[d] Martin A. Walsh [bc] and Christopher J. Schofield *[a]

The main viral protease (M^{pro}) of SARS-CoV-2 is a nucleophilic cysteine hydrolase and a current target for anti-viral chemotherapy. We describe a high-throughput solid phase extraction coupled to mass spectrometry M^{pro} assay. The results reveal some β-lactams, including penicillin esters, are active site reacting M^{pro} inhibitors, thus highlighting the potential of acylating agents for M^{pro} inhibition.

The main viral protease (M^{pro}) of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1] is a COVID-19 treatment target.[2] M^{pro} along with the papain-like protease (PL^{pro}), processes initially translated viral polyproteins to give cleaved proteins with biological functions essential for viral replication in cells.[3] Following formation of a non-covalent enzyme-substrate complex, M^{pro} catalysis proceeds *via* His-41 enabled reaction of Cys-145 with a scissile peptide bond forming a hydrolytically labile thioester. M^{pro} cleaves after glutamine-residues with a preference for small-residues on the C-terminal side of the cleaved amide (Fig. 1A and B).[4]

Most reported M^{pro} assays measure fluorescence, as precedented for other protease assays.[1,2b,5] Whilst efficient, such methods do not simultaneously monitor substrate depletion/product formation and some compounds interfere with fluorescence.[6] We were thus interested in establishing an alternative mass spectrometry (MS)-based high-throughput M^{pro} assay for identifying new inhibitors and testing known drugs.

Solid phase extraction coupled with MS (SPE-MS) has been applied to high-throughput screens of enzymes.[7] We envisaged it could simultaneously monitor both M^{pro} substrate depletion/product formation and covalent modification. The latter is of interest because many reported inhibitors of nucleophilic cysteine enzymes work by covalent reaction.[8] Here we report how such an assay enabled identification of new M^{pro} inhibitors, including β-lactams, the most important antibacterial class.[9]

We developed conditions for an SPE-MS based SARS-CoV-2 M^{pro} assay (0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2, 20 mM HEPES, pH 7.5, 50 mM NaCl) using protein prepared as reported.[1,10] Isolated M^{pro} was found to be active when monitoring turnover of peptide substrates, including TSAVLQ/SGFRK-NH_2 which was cleaved to give TSAVLQ and SGFRK-NH_2 fragments (Fig. 1C and D). Kinetic parameters were determined for the 11-mer substrate (K_m = 14.4 μM; k_{cat} = 2.7 min^{-1}), both by monitoring substrate depletion and N-terminal product fragment formation (Fig. S1, ESI†). The efficiency (k_{cat}/K_m) of M^{pro} determined by SPE-MS (28 500 M^{-1} s^{-1}) is comparable to that observed for a similar substrate Mca–AVLQ/SGFRK(Dnp)K using a fluorescence resonance energy transfer (FRET) assay (27 000 M^{-1} s^{-1}, as reported and in our hands).[1] Steady state kinetics for a 37-mer substrate were also investigated; a 2-fold increase in k_{cat}/K_m (60 026 M^{-1} s^{-1}) was observed (Fig. S2, ESI†). Comparison of kinetic parameters for the SARS-CoV-2 M^{pro} and the related SARS-CoV M^{pro} reveal similar k_{cat}/K_m values (though the values for SARS-CoV were somewhat lower when using shorter substrates in an HPLC assay) (Table S1, ESI†). Note, the interconversion between monomeric/dimeric forms of M^{pro} has the potential to introduce complexity in kinetic analyses.[11]

Next, the SPE-MS assay was validated for inhibition studies with ebselen,[1] N3,[1] disulfiram,[1] and boceprevir[12] using the 11-mer TSAVLQ/SGFRK-NH_2 substrate (Table 1 and Fig. S3, ESI†). The ebselen IC_{50} was ~0.09 μM under standard conditions (0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2 *i.e.* [S] < K_m, 20 mM HEPES, pH 7.5, 50 mM NaCl at ambient temperature) compared to an IC_{50} of ~0.67 μM[1] using a FRET assay (0.2 μM M^{pro}, 20 μM Mca–AVLQ/SGFRK(Dnp)K *i.e.* [S] ≈ K_m, 50 mM Tris–HCl, pH 7.3, 1 mM EDTA, 30 °C)[1] (Table 1, entry 1; Fig. S3D, ESI†).

[a] *Chemistry Research Laboratory, Department of Chemistry, 12 Mansfield Road, Oxford, OX1 3TA, UK. E-mail: christopher.schofield@chem.ox.ac.uk*
[b] *Diamond Light Source, Harwell Science & Innovation Campus, Didcot, Oxfordshire OX11 0DE, UK*
[c] *Research Complex at Harwell, Harwell Science & Innovation Campus, Didcot, Oxfordshire OX11 0FA, UK*
[d] *Shanghai Institute for Advanced Immunochemical Studies and School of Life Science and Technology, ShanghaiTech University, Shanghai, China*

† Electronic supplementary information (ESI) available: Experimental details and inhibition data. See DOI: 10.1039/d0cc06870e

Open Access Article. Published on 19 January 2021. Downloaded on 7/6/2023 11:02:41 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

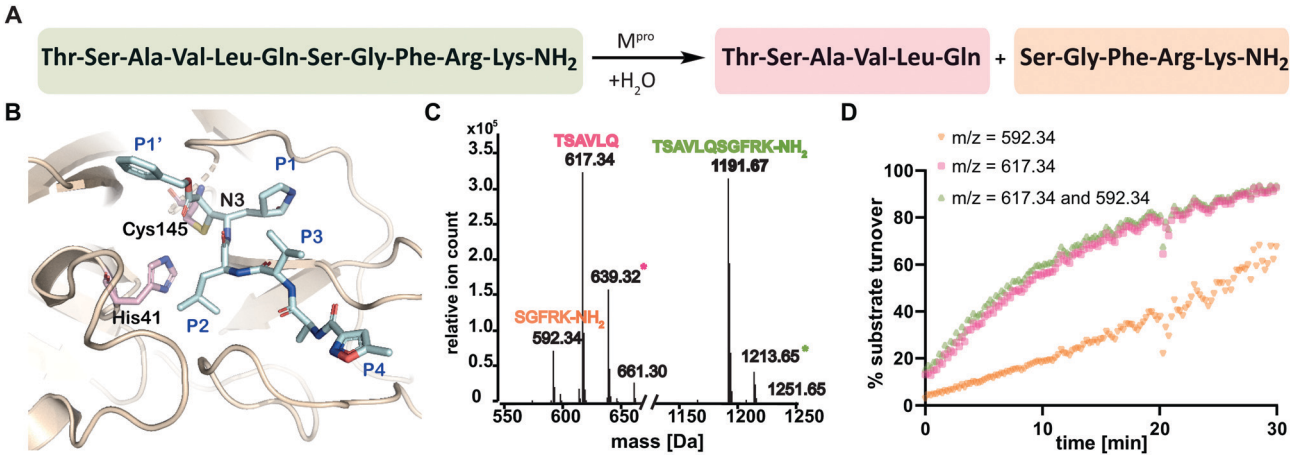


Fig. 1 SPE-MS assay monitoring M^{pro} catalyzed cleavage of the TSAVLQ/SGFRK-NH_2 substrate. (A) M^{pro} catalyzed hydrolysis of TSAVLQ/SGFRK-NH_2. (B) View from a structure of Cys-145 linked M^{pro}-N3 complex (**PDB ID: 6LU7**);[1] the Cys–His dyad is in pink; substrate binding sites are labelled in blue. (C) Deconvoluted mass spectrum of substrate/cleaved products after 10 minute incubation with M^{pro}. Note, the C-terminal product was not efficiently retained by the SPE cartridge resulting in a low abundance compared to the N-terminal cleavage product. Sodium ion adducts (+23 Da) for the TSAVLQ product (639 Da) and substrate (1214 Da) are labelled with magenta and green asterisks, respectively. (D) % substrate turnover based on integration of the total abundance of cleaved products (TSAVLQ or SGFRK-NH_2) or the individual products. Conditions: 0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2 (1192 Da) (20 mM HEPES, pH 7.5, 50 mM NaCl).

Table 1 IC_{50}s of selected M^{pro} inhibitors determined using SPE-MS assays compared to those obtained using FRET assays

Inhibitor	IC_{50} (SPE-MS) [μM][ab]	IC_{50} (SPE-MS)[ac] [μM]	IC_{50} (FRET) [μM]
Ebselen	0.09 ± 0.07	0.09 ± 0.07	0.67 ± 0.09[1]
N3	0.04 ± 0.01	0.03 ± 0.01	n.d.
Disulfiram	0.60 ± 0.01	0.46 ± 0.02	9.35 ± 0.18[1]
Boceprevir	11.0 ± 4.8	9.2 ± 5.5	2.70 ± 0.05[12]

[a] Mean of two independent replicates each performed in technical duplicate (n = 2 ± standard deviation, SD). Conditions: 0.15 μM M^{pro} and 2.0 μM TSAVLQ/SGFRK-NH_2 substrate in 20 mM HEPES, pH 7.5, 50 mM NaCl. [b] 30 min inhibitor preincubation. [c] 60 min inhibitor preincubation.

We optimized the assay for studying covalent modifications with a higher M^{pro} concentration being used to enable robust analyses (1 μM M^{pro}), though IC_{50} and preliminary covalent modification data can be accumulated from the same experiment. SPE purification is denaturing, so monomer modification was observed. Assay validation used N3;[1] predominantly (but not exclusively) a single N3 adduct was observed (Fig. 2A, B and Fig. S4, ESI†), consistent with structural work revealing Cys-145 reaction[1] (Fig. 1B and Fig. S5, Table S2, ESI†). We exploited selective reaction of N3 to test selectivity of other inhibitors as exemplified with ebselen, comparing results for N3 treated/untreated M^{pro} (Fig. 2C, E and Fig, S6, ESI†). By contrast with N3, we saw time dependent modification of multiple residues with ebselen with or without N3 pre-treatment (Fig. 2C and E), implying reaction of some of the 11 non-active site cysteines (Fig. S5, ESI†). Ebselen was used as a readily available positive inhibition control in subsequent studies.

The assay was used to screen the Library of Pharmacologically Active Compounds (LOPAC) and a library of 1600 small-molecule active pharmaceutical ingredients (API) at 20 μM compound (Fig. S7, ESI†). Excellent Z′-factors[13] and signal to noise ratios reveal the assay robustness (Fig. S7, ESI†). In addition to ebselen (identification of which validates the method), diverse inhibitors (≥80% at a fixed 20 μM inhibitor concentration) were identified, some (related to) known inhibitors,[1,14] including auranofin, cisplatin, IPA-3, bismuth subsalicylate, thioguanine, carmustine, and disulfiram (Tables S3 and S4, ESI†).

IC_{50}s were determined for compounds with ≥80% inhibition at 20 μM, excluding known interference compounds.[15] Auranofin (IC_{50} ~ 1.5 μM; reported IC_{50} ~ 0.5 μM[14]), an α-chloroketone (TPCK) (IC_{50} ~ 0.8 μM), IPA-3 (IC_{50} ~ 0.1 μM), and 5-thioguanine (IC_{50} ~ 13.5 μM) are some of the more potent inhibitors (Fig. S8, ESI†). Some of these covalently modified M^{pro}, sometimes with more than one reaction being observed (Fig. S9–S15, ESI†). Active site selectivity was investigated using N3 treated and untreated M^{pro}. Following N3 treatment, in some cases, *e.g.* TPCK and N_α-*p*-toluenesulfonyl-L-lysine chloromethyl ketone, substantial covalent modification was no longer observed, implying selective Cys-145 reaction (Fig. 2D, F and Fig. S12, S13, ESI†). Although further validation is required, with BAY 11-7082 and IPA-3 the multiple adducts observed with unmodified M^{pro} were diminished when the active site was N3 blocked, suggesting reaction with Cys-145 might alter the M^{pro} conformation (Fig. S9 and S11, ESI†).

The screen identified β-lactam drugs as potential M^{pro} inhibitors, including penicillins and cephalosporins (Table S5, ESI†). This was of interest, as in preliminary work we observed some β-lactams react covalently (data not shown). β-Lactam antibiotics form stable acyl–enzyme complexes with bacterial nucleophilic serine enzymes; they inhibit other nucleophilic serine enzymes including proteases and β-lactamases[16] and nucleophilic cysteine enzymes.[17]

Studies on cephalosporins identified as potential inhibitors from the screen revealed no substantial covalent M^{pro} modification,

 This journal is © The Royal Society of Chemistry 2021

Open Access Article. Published on 19 January 2021. Downloaded on 7/6/2023 11:02:41 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

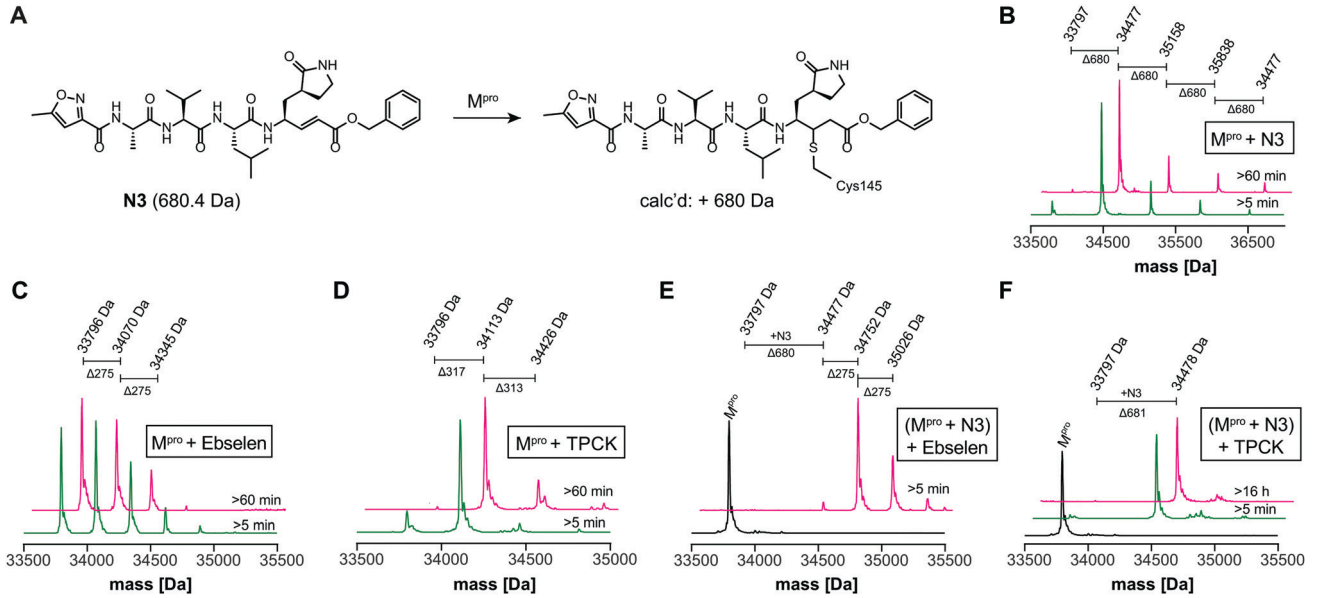


Fig. 2 N3 dependent reaction monitoring the active site selectivity of inhibitors. (A) Reaction of N3 with the M^{pro} active site Cys-145.[1] (B) N3, (C) ebselen and (D) TPCK modifies M^{pro} in a covalent manner. (E) Ebselen covalently modifies multiple M^{pro} cysteine residues in the presence of N3. (F) TPCK does not covalently modify M^{pro} in the presence of N3, suggesting that it selectively reacts with active site Cys-145. Black spectra: wild-type M^{pro} (33796 Da).

though cephalosporin C Zn(II) salt and cephalosporin C Na(I) salt inhibited. However, the IC_{50}s for cephalosporin C Zn(II) salt and $ZnCl_2$ were similar, indicating much of the inhibition is due to Zn(II) ions (Fig. S16, ESI†), as observed for cephalosporin C Zn(II) salt inhibition of other enzymes.[7b]

We further investigated β-lactam reactions with M^{pro} using a diverse set of β-lactams (Fig. S17, ESI†). Though most β-lactams were inactive (IC_{50} > 100 μM), two penicillin esters manifested IC_{50}s < 5 μM, *i.e.* **1**: a penicillin V sulfone C3 benzyl ester (IC_{50} ~ 1.5 μM), and **2**: a derivative of penicillin G sulfoxide C3

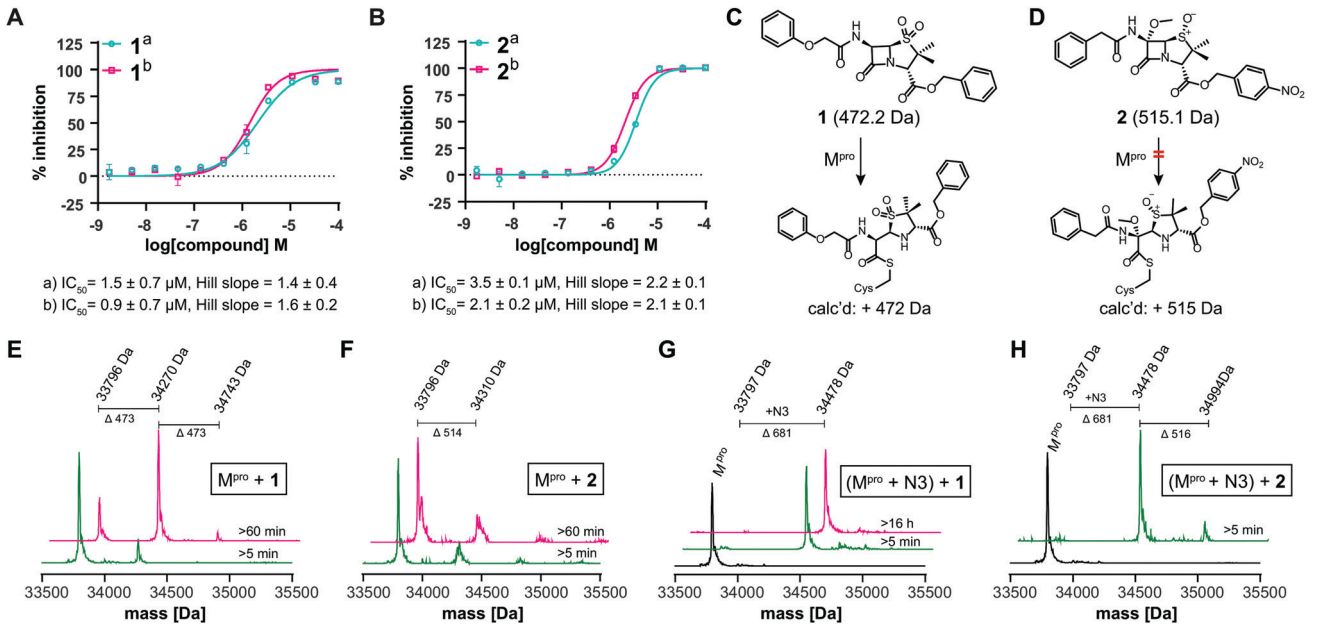


Fig. 3 β-Lactams inhibit M^{pro}. IC_{50}s for (A) **1** (penicillin V sulfone C3 benzyl ester) and (B) **2** (C6-methoxy penicillin G sulfoxide C3 p-nitrobenzyl ester) determined using SPE-MS; data are a mean of technical duplicates with (a) 30 min and (b) 60 min preincubation. IC_{50}s are means of two independent repeats each composed of technical duplicates ($n = 2 \pm$ SD). Proposed reaction of **1** (C) and **2** (D) with M^{pro}. (E) A single molecule **1** covalently modifies M^{pro}. (F) **2** does not efficiently modify M^{pro} through covalent reaction. Conditions: 1 μM M^{pro}, 20 μM β-lactam, 20 mM HEPES, pH 7.5, 50 mM NaCl. (G) **1** does not covalently modify M^{pro} preincubated with N3, suggesting **1** reacts with Cys-145. (H) **2** does not efficiently react with M^{pro} preincubated N3. Conditions: 1 μM M^{pro} preincubated with 3 μM N3, 20 μM **1** or **2**, 20 mM HEPES, pH 7.5, 50 mM NaCl. Black spectra: wild-type M^{pro} (33796 Da).

Open Access Article. Published on 19 January 2021. Downloaded on 7/6/2023 11:02:41 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

p-nitrobenzyl ester ($IC_{50} \sim 3.5$ μM), both with similar potency with either 30 or 60 min preincubation (Fig. 3A and B). Other β-lactams inhibited, though more weakly (Fig. S17, ESI†). The inhibition by the penicillin benzyl esters may, in part, reflect binding of the N3 benzyl ester, likely binding in the P1′ or P2 pocket (Fig. 1B).[1] Structures of M^{pro} complexed with a β-lactam were not obtained; however, docking studies reveal potential of **1** and **2** to bind favourably at the active site (Fig. S18, ESI†), in the case of **1** in a manner enabling Cys-145 reaction.

Evidence for covalent reaction was observed with representatives of the penem, carbapenem prodrug, penicillin, penicillin sulfone, clavam, cephem, and monobactam β-lactam sub-families (Fig. S19–S21, ESI†). In some cases, *e.g.* clavulanate (Fig. S19H, ESI†) and moxalactam (Fig. S21F, ESI†), (partial) inhibitor fragmentation was observed. There was no clear correlation between a propensity to react covalently and M^{pro} inhibition; in some cases evidence for partial covalent modification, but no inhibition was observed (Fig. S17, ESI†). Covalent modification was observed with **1**, but only to a small (<10%) extent with **2**, suggesting the latter likely inhibits principally by a non-covalent interaction (Fig. 3C–H). After Cys-145 blocking with N3, no reaction with **1** was observed (Fig. 3G). Minor further modification of Cys-145 reacted M^{pro} was observed with **2** (Fig. 3H), suggesting the low levels of covalent modification by **2** do not solely involve Cys-145.

In summary, SPE-MS is a useful method for M^{pro} assays enabling analysis of inhibition by both substrate depletion/product formation. The method complements reported *in vitro* M^{pro} assays and compares favourably to those in terms of its robustness and ability to enable efficient high-throughput screening/repurposing efforts. The SPE-MS assay also enables ready analysis of covalent M^{pro} modification and use of M^{pro} reacted with a selective inhibitor such as N3/TPCK informs on whether covalent reaction of a test inhibitor occur at the active site or not.

Although the available evidence implies that β-lactams can inhibit M^{pro} non-covalently, the observation that some react with and inhibit M^{pro} by covalent active site modification should promote interest in the development of inhibitors for M^{pro} and other thiol proteases working *via* cysteinyl *S*-acylation. By contrast with *S*-alkylating inhibitors, which can have toxicity issues, *S*-acylation has not been widely explored for nucleophilic cysteine proteases.

The identification of β-lactam containing M^{pro} inhibitors with structures closely related to drugs should promote work on the development of related compounds for progression towards clinical use for treatment of COVID-19 and viral diseases.

We thank the Oxford COVID fund and its generous donors, the BBSRC, MRC, the Wellcome Trust, Cancer Research UK and GSK for funding, and D. Ebner (TDI, Oxford) for supplying the compound libraries. We thank present and past group members including T. Suits and J. Brem and support staff including A. Hardy, Diamond and RCaH support staff who have supported our COVID-19 work, including those who made compounds for initial screening. T. R. M. was supported by the BBSRC (BB/M011224/1). T. J. thanks the EPSRC Oxford-GSK-Crick Program (EP/R512060/1) and GSK. H. T. H. C. thanks the EPSRC SBM CDT (EP/L015838/1) and the Clarendon Fund. We thank J. E. Baldwin, R. M. Adlington and R.D.G. Cooper for encouragement and inspiration.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 Z. Jin, X. Du, Y. Xu, Y. Deng and M. Liu, *et al.*, *Nature*, 2020, **582**, 289.
2 (*a*) K. Anand, G. J. Palm, J. R. Mesters, S. G. Siddell and J. Ziebuhr, *et al.*, *EMBO J.*, 2002, **21**, 3213; (*b*) H. Yang, M. Yang, Y. Ding, Y. Liu and Z. Lou, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13190; (*c*) A. K. Gosh, M. Brindisi, D. Shahabi, M. E. Chapman and A. D. Mesecar, *ChemMedChem*, 2020, **15**, 907.
3 J. Ziebuhr, E. J. Snijder and A. E. Gorbalenya, *J. Gen. Virol.*, 2000, **81**, 853.
4 L. Zhang, D. Lin, X. Sun, U. Curth and C. Drosten, *et al.*, *Science*, 2020, **368**, 409.
5 K. Fan, P. Wei and Q. Feng, *et al.*, *J. Biol. Chem.*, 2004, **279**, 1637.
6 (*a*) A. Simeonov and M. I. Davis, *Interference with Fluorescence and Absorbance*, 2004; (*b*) D. S. Auld, N. T. Southall, A. Jadhav, R. L. Johnson and D. J. Diller, *et al.*, *J. Med. Chem.*, 2008, **51**, 2372; (*c*) A. Simeonov, A. Jadhav, C. J. Thomas, Y. Wang and R. Huang, *et al.*, *J. Med. Chem.*, 2008, **51**, 2363.
7 (*a*) J. Meng, M. T. Lai, V. Munshi, J. Grobler and J. McCauley, *et al.*, *J. Biomol. Screen.*, 2015, **20**, 606; (*b*) L. Brewitz, A. Tumber, I. Pfeffer, M. A. McDonough and C. J. Schofield, *Sci. Rep.*, 2020, **10**, 8650; (*c*) L. Brewitz, A. Tumber, X. Zhang and C. J. Schofield, *Bioorg. Med. Chem.*, 2020, **28**, 115675.
8 A. K. Ghosh, I. Samanta, A. Mondal and W. R. Liu, *ChemMedChem*, 2019, **14**, 889.
9 E. Y. Klein, T. P. Van Boeckel, E. M. Martinez, S. Pant and S. Gandra, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E3463.
10 A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer and P. Lukacik, *et al.*, *Nat. Commun.*, 2020, **11**, 5047.
11 T. J. El-Baba, C. A. Lutomski, A. L. Kantsadi, T. R. Malla and T. John, *et al.*, *Angew. Chem., Int. Ed.*, 2020, DOI: 10.1002/anie.202010316.
12 C. Ma, M. D. Sacco, B. Hurst, J. A. Townsend and Y. Hu, *et al.*, *Cell Res.*, 2020, **30**, 678.
13 J. H. Zhang, T. D. Chung and K. R. Oldenburg, *J. Biomol. Screen.*, 1999, **4**, 67.
14 Z. He, W. Zhao, W. Niu, X. Gao and X. Gao, *et al.*, *bioRxiv*, 2020, DOI: 10.1101/2020.05.28.120642.
15 J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719.
16 (*a*) C. L. Tooke, P. Hinchliffe, E. C. Bragginton, C. K. Colenso and V. H. A. Hirvonen, *et al.*, *J. Mol. Biol.*, 2019, **431**, 3472; (*b*) W. B. Knight, A. L. Maycock and B. G. Green, *et al.*, *Biochemistry*, 1992, **31**, 4980.
17 M. A. Gun, B. Bozdogan and A. Y. Coban, *Future Microbiol.*, 2020, **15**, 937.

 This journal is © The Royal Society of Chemistry 2021

Journal of Medicinal Chemistry

pubs.acs.org/jmc

Article

Penicillin Derivatives Inhibit the SARS-CoV-2 Main Protease by Reaction with Its Nucleophilic Cysteine

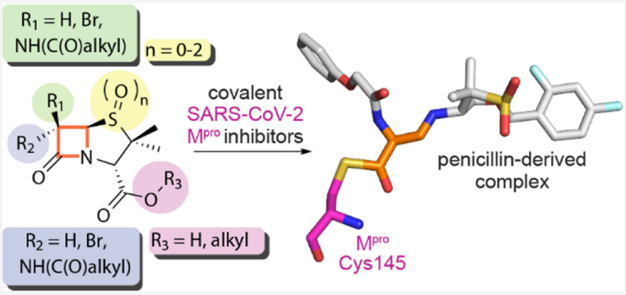
Tika R. Malla,[‖] Lennart Brewitz,*[,‖] Dorian-Gabriel Muntean, Hiba Aslam, C. David Owen, Eidarus Salah, Anthony Tumber, Petra Lukacik, Claire Strain-Damerell, Halina Mikolajek, Martin A. Walsh, and Christopher J. Schofield*

Cite This: *J. Med. Chem.* 2022, 65, 7682−7696

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The SARS-CoV-2 main protease (M^{pro}) is a medicinal chemistry target for COVID-19 treatment. Given the clinical efficacy of β-lactams as inhibitors of bacterial nucleophilic enzymes, they are of interest as inhibitors of viral nucleophilic serine and cysteine proteases. We describe the synthesis of penicillin derivatives which are potent M^{pro} inhibitors and investigate their mechanism of inhibition using mass spectrometric and crystallographic analyses. The results suggest that β-lactams have considerable potential as M^{pro} inhibitors via a mechanism involving reaction with the nucleophilic cysteine to form a stable acyl−enzyme complex as shown by crystallographic analysis. The results highlight the potential for inhibition of viral proteases employing nucleophilic catalysis by β-lactams and related acylating agents.



Downloaded via UNIV OF OXFORD on July 6, 2023 at 10:03:30 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

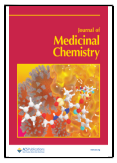
The inhibition of proteases that hydrolyze viral polyproteins to give functional proteins is a validated mechanism for antiviral chemotherapy, as exemplified by pioneering work on human immunodeficiency virus (HIV) protease and, more recently, hepatitis C virus (HCV) protease inhibitors.[1] Thus, both the severe acute respiratory disease coronavirus-2 (SARS-CoV-2)[2] main protease (M^{pro} or 3C-like protease, $3CL^{pro}$) and the papain-like protease (PL^{pro}) are targets for the treatment and, possibly, prevention of coronavirus disease 2019 (COVID-19).[3−8] M^{pro} is a particularly attractive drug target because (i) M^{pro} is vital in the SARS-CoV-2 life cycle, (ii) M^{pro} is tractable from a small-molecule inhibition perspective as a nucleophilic cysteine protease, and (iii) the structure and substrate selectivities of M^{pro} are different from human proteases,[9,10] suggesting clinically useful selective M^{pro} inhibition should be possible.

To enable the identification of small-molecule M^{pro} inhibitors for development as human therapeutics, high-throughput *in vitro* inhibition assays using recombinant viral M^{pro} have been developed.[4,9−12] Most reported M^{pro} inhibition assays employ fluorescence-based methods, though label-free assays, which directly monitor product formation/substrate depletion using mass spectrometry (MS) and SARS-CoV-2 polyprotein peptide fragments, have been reported.[13−15] The availability of efficient high-throughput M^{pro} inhibition assays and libraries of bioactive and safety-assessed small molecules has enabled the identification of multiple lead M^{pro} inhibitors, such as boceprevir (**1**),[11,16] an HCV serine protease inhibitor,[17,18] SDZ-224015 (**2**),[19] an investigational caspase-1 inhibitor,[20] and GC-376 (**3**),[11,16,21] for (partially) selective inhibition of M^{pro} (Figure 1A−C).

To date, drug repurposing efforts have not yielded safe and efficient M^{pro} inhibitors for approved human clinical use. Thus, *de novo* M^{pro} inhibitor development programs have been initiated based on the structural information gained from the identified lead structures in the SARS-CoV-2 M^{pro} screening campaigns as well as from structure−activity relationship (SAR) studies with reported SARS-CoV and MERS-CoV M^{pro} inhibitors.[5,27−30] Compounds arising from such efforts include PF-00835231 (**4**) and PF-07304814 (**5**),[23,24] N3 (**6**),[9,31] and the α-ketoamide **7**, which are all potent SARS-CoV-2 M^{pro} inhibitors displaying high *in vitro* and *in vivo* potency (Figure 1D−F). Novel SARS-CoV-2 M^{pro} inhibitors include compounds PF-07321332 (**8**, nirmatrelvir),[25] which is in clinical use,[32] and MI-09 (**9**)[26] and structurally related molecules (Figure 1G,H).[6,7,33−37]

Most M^{pro} inhibitors work by covalent modification, in part, likely because of well-precedented mechanisms for inhibiting

Received: December 26, 2021
Published: May 12, 2022



ACS Publications

© 2022 The Authors. Published by American Chemical Society

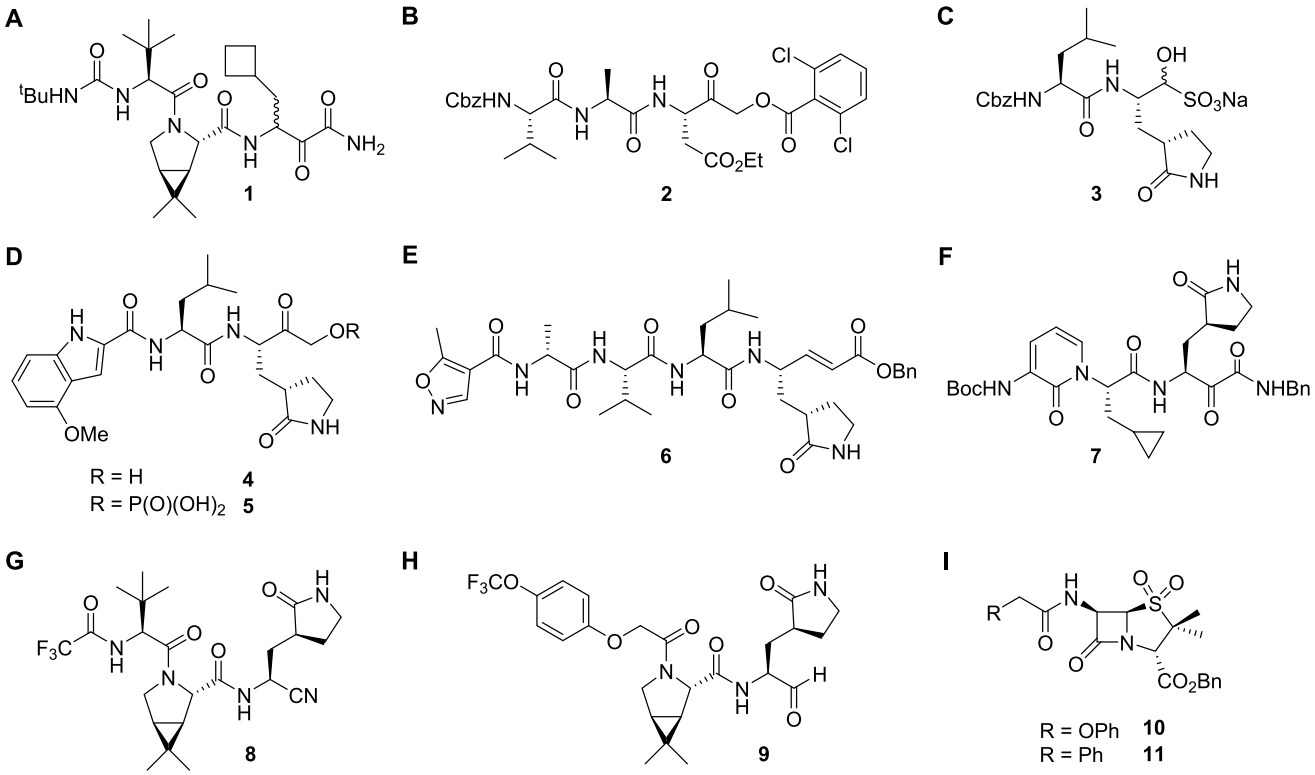


Figure 1. Examples of reported SARS-CoV-2 M^{pro} small-molecule inhibitors. (A) Boceprevir (**1**);[11,16] (B) SDZ-224015 (**2**);[22] (C) GC-376 (**3**);[11,16,21] (D) PF-00835231 (**4**) and its prodrug PF-07304814 (**5**);[23,24] (E) N3 (**6**);[9] (F) α-ketoamide **7**;[10] (G) PF-07321332 (**8**, nirmatrelvir);[25] (H) MI-09 (**9**);[26] and (I) penicillin V and G sulfone benzyl esters **10** and **11**.[15]

proteases and related enzymes by covalent reaction with nucleophilic serine or cysteine residues,[1] although noncovalent M^{pro} inhibitors have also been reported.[38−41] Electrophiles employed in covalently reacting M^{pro} inhibitors include, for example, nitrile, α-ketoamide, α-acyloxymethylketone, aldehyde, and Michael acceptor, amongst other functional groups (Figure 1).[42−49] By contrast with the extensive work on alkylating agents such as SDZ-224015 (Figure 1B), work on acylating agents, such as β-lactams, which generally have good safety profiles as antibacterials, has been limited. Because of their demonstrated efficacy and safety records,[50,51] we are particularly interested in optimizing the potential of β-lactams and related acylating agents as inhibitors of nucleophilic cysteine enzymes, in particular, M^{pro}.

Recently, we reported a solid-phase extraction coupled to MS (SPE-MS) M^{pro} assay, which enabled the identification of a certain penicillin V derivative, i.e., **10**, that inhibits M^{pro} by reaction with the active site cysteine residue; by contrast, the corresponding penicillin G derivative **11** was inactive (Figure 1I).[15] Here, we report SAR studies with penicillin derivatives, leading to the identification of efficient M^{pro} inhibitors with a penicillin scaffold; their mechanism of inhibition was investigated using MS and crystallography. The results highlight the potential of β-lactams for use as M^{pro} inhibitors working by acylation of the nucleophilic cysteine.

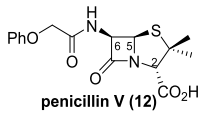
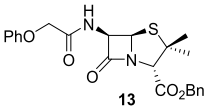
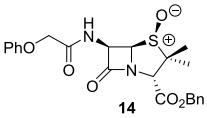
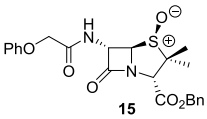
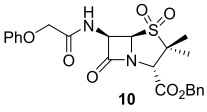
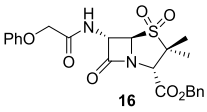
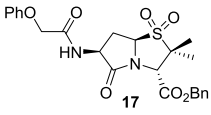
RESULTS

Penicillin Stereochemistry Affects M^{pro} Inhibition. To enable SAR studies, an initial set of penicillin V derivatives was synthesized based on the identified penicillin V sulfone benzyl ester M^{pro} inhibitor (**10**; Figure 1I).[15] Half-maximum inhibitory concentrations (IC_{50} values) were determined using the reported SPE-MS inhibition assay, monitoring M^{pro}-catalyzed hydrolysis of an 11mer substrate peptide (TSAVLQ/SGFRK-NH_2, "/" indicates the M^{pro} cleavage site), the sequence of which is based on the N-terminal self-cleavage site of M^{pro}. However, some of the initially investigated penicillin V derivatives appeared to suppress product peptide ionization at high inhibitor concentrations, perturbing the reliability of the inhibition results. Therefore, the M^{pro} inhibition assays were performed using an extended 37mer peptide substrate based on the same M^{pro} self-cleavage site (ALNDFSNSGSDV-LYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2),[15] which was less susceptible to penicillin inhibitor-induced ion suppression. Substituting the 11mer peptide with the 37mer peptide in the SPE-MS M^{pro} inhibition assays did not affect the IC_{50} values of reported selected M^{pro} inhibitors (Supporting Information Table S1); thus, the 37mer substrate was used for subsequent IC_{50} determinations. The observed high *Z*-factors (>0.5 for each inhibition plate) indicate excellent SPE-MS assay quality using the 37mer substrate peptide (Supporting Information Figure S1).

The modified SPE-MS M^{pro} inhibition assay was used to investigate the influence of structural features of the penicillin V sulfone benzyl ester (**10**) on potency. Unlike **10**, neither commercially sourced penicillin V (**12**) nor its benzyl ester (**13**)[52] inhibited M^{pro} efficiently, in agreement with previous results using the 11mer M^{pro} substrate[15] (Table 1, entries 1 and 2). Stereoselective oxidation of **13** with *meta*-chloroperbenzoic acid (mCPBA) afforded the reported penicillin V (*S*)-sulfoxide benzyl ester (**14**),[52] which is a less efficient inhibitor than the corresponding sulfone **10** (IC_{50} ~ 22.9 μM, Table 1, entry 3),

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

Table 1. Inhibition of SARS-CoV-2 M^{pro} by Penicillin V Derivatives

	Penicillin V derivative	[a]IC_{50} [μM]
1	 penicillin V (12)	>50
2	 13	>50
3	 14	22.9 ± 13.3
4[b]	 15	>50
5	 10	6.6 ± 2.7
6	 16	>50
7	 17	26.1 ± 1.2

[a]M^{pro} inhibition assays were performed using SPE-MS as described in the Experimental Section employing SARS-CoV-2 M^{pro} (0.15 μM) and a substrate (2.0 μM). Results are means of at least two independent runs, each composed of technical duplicates ($n \geq 2$; mean ± standard deviation, SD). Representative dose−response curves are shown in Supporting Information Figure S4. [b]Contains minor amounts of a decomposition product, as reported.[53] Bn: $-CH_2Ph$.

highlighting the importance of an additional *pro-R*-sulfone oxygen of **10** for M^{pro} inhibition.

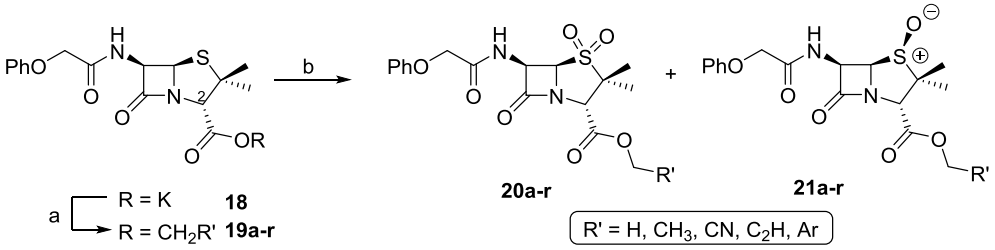
Next, the effect of the stereochemistry at the C6 stereocenter on inhibition was investigated by inversion of the (*R*)-configuration of **10** using a reported protocol.[53] The resultant (6*S*)-penicillin V (*S*)-sulfoxide benzyl ester **15**[53] showed reduced M^{pro} inhibition compared with **10** (IC_{50} > 50 μM, Table 1, entry 4). The corresponding (6*S*)-penicillin V sulfone benzyl ester (**16**), which was obtained from **15** using $KMnO_4$ as an oxidant, was also less efficient in inhibiting M^{pro} than the (6*R*)-isomer **10** (IC_{50} > 50 μM, Table 1, entry 6). Thus, the (6*R*)-configuration at the penicillin V C6 stereocenter appears to be preferred for efficient M^{pro} inhibition.

The importance of the β-lactam ring for efficient M^{pro} inhibition was investigated by preparing the corresponding γ-lactam **17**, which was synthesized using a modified literature procedure (Supporting Information Figure S2). The potency of **17** was reduced compared to the β-lactam **10** (IC_{50} ~ 26.1 μM, Table 1, entry 7) but was not ablated. This observation may in part reflect the enhanced and/or less reversible reaction of β-lactams with nucleophiles compared to γ-lactams.[54,55]

Penicillin Ester Group Fine-Tunes Inhibitor Potency. Having identified important structural features of penicillin V sulfone benzyl ester (**10**) for efficient M^{pro} inhibition, the impact of its C2 ester group on potency was investigated. To obtain a set of varied penicillin V sulfone esters, the commercially sourced penicillin V potassium salt (**18**) was initially reacted with different alkylhalides (Scheme 1). The resultant penicillin V esters **19a−r** were oxidized using mCPBA to afford a chromatographically separable mixture of both the sulfones **20a−r** and (*S*)-sulfoxides **21a−r**.

The penicillin V sulfone esters **20a−r** were investigated for M^{pro} inhibition using the SPE-MS assay with the 37mer peptide substrate (Table 2). The preferred phenyl ring substitution pattern of the benzyl ester group for M^{pro} inhibition was investigated by fluorine atom substitution. The results reveal that a single fluorine substituent in the *ortho-*, *meta-*, or *para-*position does not substantially alter the potency (Table 2, entries 1−3). However, the presence of two fluorine *meta-*substituents, as in **20d**, appears to reduce the potency (IC_{50} ~ 16.9 μM, Table 2, entry 4) while the corresponding isomer **20e**, which bears a fluorine substituent at both the *ortho-* and *para-*positions, is slightly more potent than the benzyl ester derivative **10** (IC_{50} ~ 3.6 μM, Table 2, entry 5). This trend was clearer when comparing the trifluorinated benzyl ester derivatives **20f** and **20g** (Table 2, entries 6 and 7). The derivative **20f** with two fluorine substituents as the *meta-*positions and one at the *para-*position is less potent than the originally identified benzyl ester

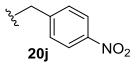
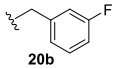
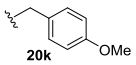
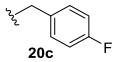
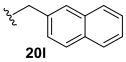
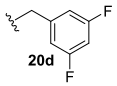
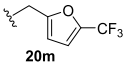
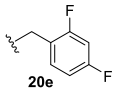
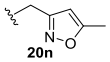
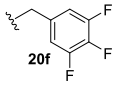
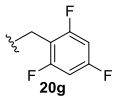
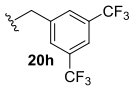
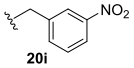
Scheme 1. Synthesis of Penicillin V Ester Derivatives[a]



[a]Reagents and conditions: (a) alkylhalide (1.2 equiv), DMF, rt, 59−93%; (b) mCPBA, CH_2Cl_2, 0 °C to rt, 89−94%. Note that the sulfoxide stereochemistry was tentatively assigned the (*S*)-configuration based on reported mCPBA-mediated penicillin ester oxidations to sulfoxides.[52,56,57]

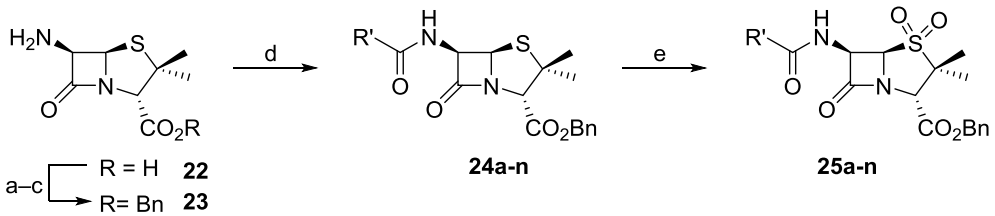
https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

Table 2. Inhibition of SARS-CoV-2 M^{pro} by Penicillin V Sulfone Esters

	R:	[a]IC_{50} [μM]		R:	[a]IC_{50} [μM]
1	 20a	5.8 ± 2.0	10	 20j	>50
2	 20b	9.2 ± 3.7	11	 20k	31.1 ± 13.9
3	 20c	7.4 ± 2.3	12	 20l	>50
4	 20d	16.9 ± 11.3	13	 20m	10.2 ± 2.2
5	 20e	3.6 ± 1.0	14	 20n	9.0 ± 1.3
6	 20f	>50	15	 20o	7.9 ± 1.4
7	 20g	1.6 ± 0.6	16	 20p	5.2 ± 0.6
8	 20h	35.3 ± 2.2	17	 20q	3.0 ± 0.4
9	 20i	13.3 ± 5.3	18	 20r	9.7 ± 3.3

[a]M^{pro} inhibition assays were performed using SPE-MS as described in the Experimental Section employing SARS-CoV-2 M^{pro} (0.15 μM) and a substrate (2.0 μM). Results are means of at least two independent runs, each composed of technical duplicates ($n \geq 2$; mean ± SD). Representative dose−response curves are shown in Supporting Information Figure S4.

Scheme 2. Synthesis of Penicillin Sulfone Benzyl Esters from (+)-6-Aminopenicillanic Acid (6-APA, 22)[a]



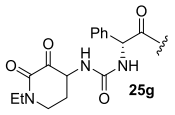
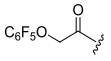
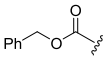
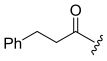
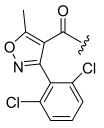
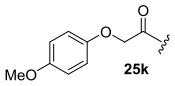
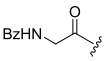
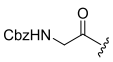
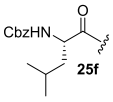
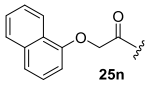
[a]Reagents and conditions: (a) benzylbromide, triethyl amine, CH_2Cl_2, 0 °C; (b) *para*-toluenesulfonic acid, acetone, rt; (c) $NaHCO_3$, ethyl acetate/H_2O, rt, 33% over three steps; (d) carboxylic acid, COMU,[58] DMF, 0 °C to rt, 44−82%; (e) mCPBA, CH_2Cl_2, 0 °C to rt, 11−71%. Note that the C6 NHCbz penicillin benzyl ester **24a** was synthesized by a different sequence, as described in the Supporting Information.

10 (IC_{50} > 50 μM), while the derivative **20g** with two fluorine substituents as the *ortho*-positions and one at the *para*-position is more potent (IC_{50} ∼ 1.6 μM). It is unclear whether these observations reflect alterations in (hydrophobic) interactions of the fluorinated ester groups with M^{pro} and/or (less likely) enhanced β-lactam reactivity due to remote electronic effects.

The benzyl esters **20h** and **20i**, which bear a relatively bulky and electron-withdrawing trifluoromethyl or nitro *meta*-

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

Table 3. Inhibition of SARS-CoV-2 M^{pro} by Penicillin Sulfone Benzyl C6 Derivatives

	R:	[a]IC_{50} [μM]		R:	[a]IC_{50} [μM]
1	 **10**	6.6 ± 2.7	9	 **25g**	>50
2	 **11**	>50	10	 **25h**	5.3 ± 3.3
3	 **25a**	>50	11[b]	 **25i**	>50
4	 **25b**	45.7 ± 0.9	12[b,c]	 **25j**	>50
5	 **25c**	>50	13	 **25k**	2.8 ± 1.2
6	 **25d**	>50	14	 **25l**	2.3 ± 1.2
7	 **25e**	>50	15	 **25m**	4.1 ± 0.1
8	 **25f**	>50	16	 **25n**	42.1 ± 10.6

[a]M^{pro} inhibition assays were performed using SPE-MS as described in the Experimental Section employing SARS-CoV-2 M^{pro} (0.15 μM) and a substrate (2.0 μM). Results are means of at least two independent runs, each composed of technical duplicates ($n \geq 2$; mean ± SD). Representative dose−response curves are shown in Supporting Information Figure S4. [b]Used as a 1:1 mixture of diastereomers. [c]Used as the 4-fluorobenzyl ester. Bn: $-CH_2Ph$, Cy: $-C_6H_{11}$, Bz: $-C(O)C_6H_5$, Cbz: $-C(O)OCH_2Ph$.

substituent, inhibited M^{pro} (Table 2, entries 8 and 9), while the *para*-substituted benzyl esters **20j** and **k**, as well as the 2-naphthylmethyl ester **20l**, do not, at least efficiently, inhibit M^{pro} (Table 2, entries 10−12).

The presence of a penicillin C2 benzyl ester derivative is not required for efficient M^{pro} inhibition, as apparent by the penicillin V sulfone esters **20m** and **20n** with a heteroaromatic ester group, which inhibit M^{pro}, albeit with slightly reduced potency compared to **10** (Table 2, entries 13 and 14). By contrast, the corresponding alkyl esters **20o−r** inhibit with similar potency as the penicillin V sulfone benzyl ester **10** (Table 2, entries 15−18). However, in general, the corresponding penicillin V esters **19a−r** and the penicillin V (*S*)-sulfoxide esters **21a−r** did not manifest substantial levels of M^{pro} inhibition (Supporting Information Table S2), in accord with the initial SAR results (Table 1).

Penicillin C6 Side Chain Modulates M^{pro} Inhibition. A set of penicillin sulfone benzyl esters with different C6 amido groups were synthesized in three steps from commercially sourced (+)-6-aminopenicillanic acid (6-APA, **22**) to investigate the impact of the C6 side chain on M^{pro} inhibition (Scheme 2). Initially, 6-APA was transformed into its benzyl ester (**23**), which was then used in amide bond-forming reactions with an appropriate carboxylic acid using COMU[58] as a coupling reagent. The resultant penicillin derivatives **24a−n** were oxidized with mCPBA to the penicillin sulfones **25a−n**.

The results of using the SPE-MS assay to test the penicillin sulfone benzyl esters **25a−n** for M^{pro} inhibition reveal that the presence and position of a C6 phenoxyacetyl ether oxygen is important in enabling efficient M^{pro} inhibition by the tested compounds (Table 3), in agreement with the observation that penicillin G sulfone benzyl ester **11** did not inhibit M^{pro} (Table

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

3, entry 2).[15] Swapping the C6 phenoxyacetyl ether oxygen from the amide β-position to the α-position, as in urethane **25a**, abolished M^{pro} inhibition (Table 3, entry 3). The substitution of the C6 phenoxyacetyl ether oxygen for a methylene group substantially diminished M^{pro} inhibition ($IC_{50} \sim 45.7\ \mu M$, Table 3, entry 4), whereas the substitution of the penicillin V C6 side chain for the dicloxacillin C6 side chain, which does not bear an oxygen atom at the same position, abolished inhibition completely (Table 3, entry 5). Additionally, substitution of the C6 phenoxyacetyl ether oxygen for an NH group, as present in penicillin sulfone benzyl esters **25d−g**, results in loss of inhibition (Table 3, entries 6−9).

These results suggest that the ability of the C6 phenoxyacetyl ether oxygen to function as a hydrogen bond acceptor/Lewis acid/conformation restrictor may be important for efficient M^{pro} inhibition, as supported by preliminary molecular docking studies.[15] The effect of the Lewis acidity of the C6 phenoxyacetyl ether oxygen was probed by substituting its phenyl substituent for an electron-withdrawing pentafluorophenyl substituent (i.e. **25h**); however, this substitution did not alter M^{pro} inhibition substantially (Table 3, entry 10). By contrast, decreasing the accessibility of the C6 phenoxyacetyl ether oxygen by introducing an alkyl-substituent α to the ether oxygen, as in **25i** and **25j**, resulted in substantially reduced inhibition (Table 3, entries 11 and 12). Increasing the electron-donating capability of the phenyl ether by introducing a methoxy substituent at its *para*-position (**25k**) appeared to improve inhibition ($IC_{50} \sim 2.8\ \mu M$, Table 3, entry 13).

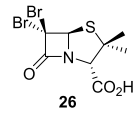
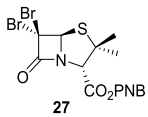
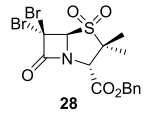
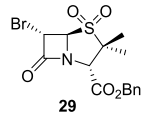
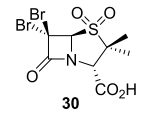
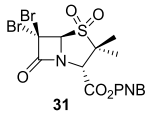
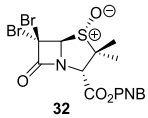
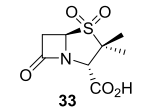
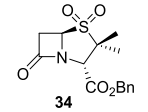
Modifying the steric bulk of the C6 phenoxyacetyl ether by substituting the phenyl group for a benzyl group (**25l**) appeared to improve inhibition ($IC_{50} \sim 2.3\ \mu M$, Table 3, entry 14), while its substitution by a cyclohexyl group (**25m**) did not substantially affect M^{pro} inhibition ($IC_{50} \sim 4.1\ \mu M$, Table 3, entry 15). By contrast, its substitution by a more bulky and rigid 1-naphthyl group (**25n**) resulted in substantially decreased inhibition ($IC_{50} \sim 42.1\ \mu M$, Table 3, entry 16).

C6 Dibromo-Penicillins Are Efficient M^{pro} Inhibitors. Considering the importance of the penicillin V sulfone benzyl ester C6 side chain on inhibitor potency, the corresponding C6 mono- and dibromo-substituted penicillin sulfones **26−32** were prepared because such substitutions alter the reaction outcome of β-lactams with nucleophilic serine β-lactamases.[59−62] The C6 mono- and dibrominated penicillins were synthesized from 6-APA (**22**) as reported[63−65] and investigated for M^{pro} inhibition using the SPE-MS assay (Table 4).

In agreement with the previous SAR studies (Table 1), neither 6,6-dibromopenicillanic acid **26** nor its *para*-nitrobenzyl ester derivative **27** inhibited M^{pro} (Table 4, entries 1 and 2). By contrast, the 6,6-dibromopenicillanic acid sulfone benzyl ester **28** was the most efficient penicillin M^{pro} inhibitor identified so far ($IC_{50} \sim 0.7\ \mu M$, Table 4, entry 3), while the corresponding 6-bromopenicillanic acid sulfone benzyl ester **29** was substantially less efficient in inhibiting M^{pro} ($IC_{50} \sim 10.1\ \mu M$, Table 4, entry 4). Surprisingly, however, the 6,6-dibromopenicillanic acid sulfone **30**, which is an intermediate in the synthesis of **28**, also showed inhibition ($IC_{50} \sim 24.2\ \mu M$, Table 4, entry 5); thus, for the first time in our SAR studies, inhibition was observed for a penicillanic acid, which is structurally closely related to sulbactam (**33**), a penam sulfone that is clinically used as a serine β-lactamase inhibitor (Table 4, entry 8).[66,67] The importance of the C6 dibromo substituents for efficient M^{pro} inhibition is further highlighted by the observation that neither sulbactam (**33**) nor its benzyl ester derivative (**34**) displayed notable inhibition (Table 4, entries 8 and 9), in accord with the reported inability of sulbactam to inhibit M^{pro}.[15]

By contrast with the previous SAR studies that showed less efficient M^{pro} inhibition of the penicillin V sulfone *para*-nitrobenzyl ester **20j** compared to the benzyl ester **10** (Table 2),

Table 4. Inhibition of SARS-CoV-2 M^{pro} by C6 Mono- and Dibromo-Penicillin Derivatives

	Penicillin derivative	[a]IC_{50} [μM]
1	 **26**	>50
2	 **27**	>50
3	 **28**	0.7 ± 0.1
4	 **29**	10.1 ± 0.1
5	 **30**	24.2 ± 6.5
6	 **31**	0.6 ± 0.1
7	 **32**	0.5 ± 0.1
8	 **33**	>50[15]
9	 **34**	>50

[a]M^{pro} inhibition assays were performed using SPE-MS as described in the Experimental Section employing SARS-CoV-2 M^{pro} (0.15 μM) and a substrate (2.0 μM). Results are means of at least two independent runs, each composed of technical duplicates ($n \geq 2$; mean ± SD). Representative dose−response curves are shown in Supporting Information Figure S4. Bn: $-CH_2Ph$; PNB: $-CH_2C_6H_4(4\text{-}NO_2)$.

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

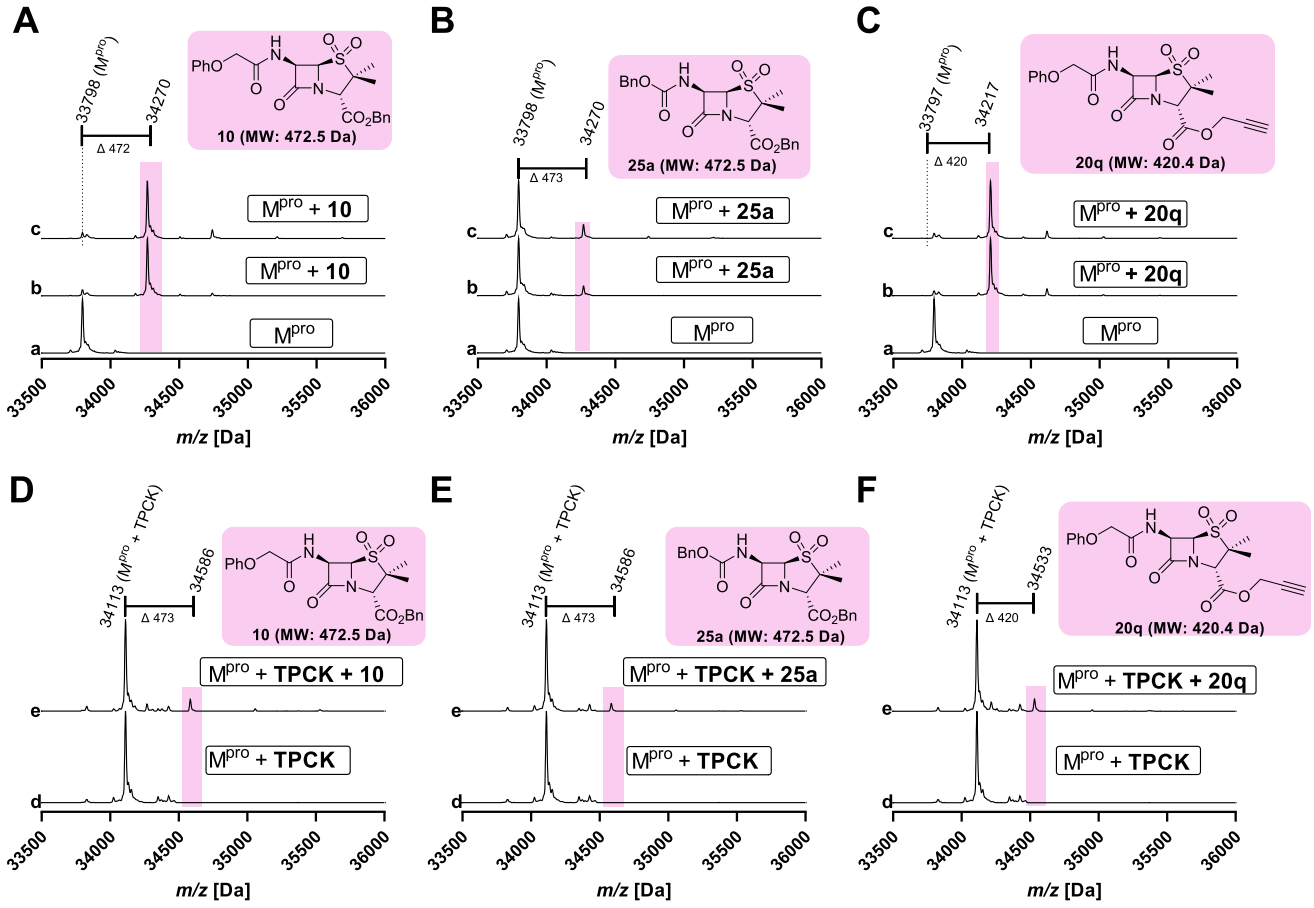


Figure 2. Evidence that penicillin V sulfone ester derivatives inhibit M^{pro} by selective active site cysteine covalent modification. M^{pro} assays with penicillin sulfone ester derivatives **10** (A, D), **25a** (B, E), and **20q** (C, F) were performed in the absence (A−C) or presence (D−F) of TPCK using SPE-MS as described in the Experimental Section employing SARS-CoV-2 M^{pro} (2.0 μM), penicillin sulfone ester derivatives (11 μM for b, and 33 μM for c and e), and TPCK (10 μM for d and e). The reactions were incubated for either 45 min (a−c) or 180 min (with TPCK, d and e) followed by additional 60 min (with a penicillin sulfone ester derivative, e) prior to analysis by SPE-MS. The reactions were performed in technical duplicates (Supporting Information Figure S5). Note (i) the clear evidence for the covalent reaction of **10** and **20q** but not **25a**, and (ii) that reaction is ablated by pretreatment of M^{pro} with the active site binding inhibitor TPCK.

6,6-dibromopenicillanic acid sulfone *para*-nitrobenzyl ester **31** had a similar potency as the corresponding benzyl ester **28** (IC_{50} ~ 0.6 μM, Table 4, entry 6); the 6,6-dibromopenicillaic acid (*R*)-sulfoxide *para*-nitrobenzyl ester **32** also inhibited with similar efficiency (IC_{50} ~ 0.5 μM, Table 4, entry 7). Note that mCPBA oxidation of **27** occurs from the least hindered side to afford a (*R*)-configured sulfoxide in the absence of a C6 amido directing group. The combined results suggest that the binding mode and/or mechanism of inhibition of the C6 mono- and dibromo-penicillin derivatives **28**−**32** differs compared to those of the C6 amido penicillin V derivatives previously investigated (Tables 1−3). Note that studies with β-lactamases imply the modes of inhibition by the C6 dibromo penicillin derivatives, which are presently under investigation, might be complex.[59−62]

Mass Spectrometric Evidence That Selected Penicillin Derivatives Inhibit by Covalent M^{pro} Modification. Selected compounds were tested for reaction with M^{pro} using protein-observed MS to inform on their mechanism of inhibition (Figure 2). Initially, M^{pro} was incubated with a ~6-fold excess of the penicillin sulfone ester for 45 min and then analyzed by MS. The results imply that penicillins **10** and **20q** modify M^{pro} by, at least predominantly, covalent reaction with a single nucleophilic protein residue, likely the active site Cys145 (Figure 2A,C). By contrast, penicillin **25a** with a C6 CbzNH group does not modify M^{pro} covalently (Figure 2B). When the inhibitor concentration was increased to ~17-fold to account for the possibility of modifying all 12 M^{pro} cysteine or other residues, only low levels of a second covalent M^{pro} modification by the inhibitors were observed (Figure 2A−C). No evidence for fragmentation of the inhibitors once bound to M^{pro} was accrued.

To investigate the site of covalent modification, M^{pro} was first preincubated with *N*-*para*-toluenesulfonyl-L-phenylalanine chloromethyl ketone (TPCK), which is reported to selectively alkylate the active site cysteine (Cys145) versus the other 11 M^{pro} cysteine residues,[15] then incubated with selected penicillin sulfone ester derivatives, i.e., **10**, **20q**, and **25a** (Figure 2D−F). The results reveal that the penicillin sulfone ester derivatives, in particular **10** and **20q**, do not efficiently react covalently with the active site (Cys145) TPCK-blocked M^{pro}, in agreement with a mechanism involving a covalent reaction with the nucleophilic active site cysteine residue Cys145. This mode of inhibition is consistent with the results obtained for the other penicillin sulfone ester derivatives investigated in this study (Supporting Information Figure S5), while γ-lactam **17** appears to inhibit

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

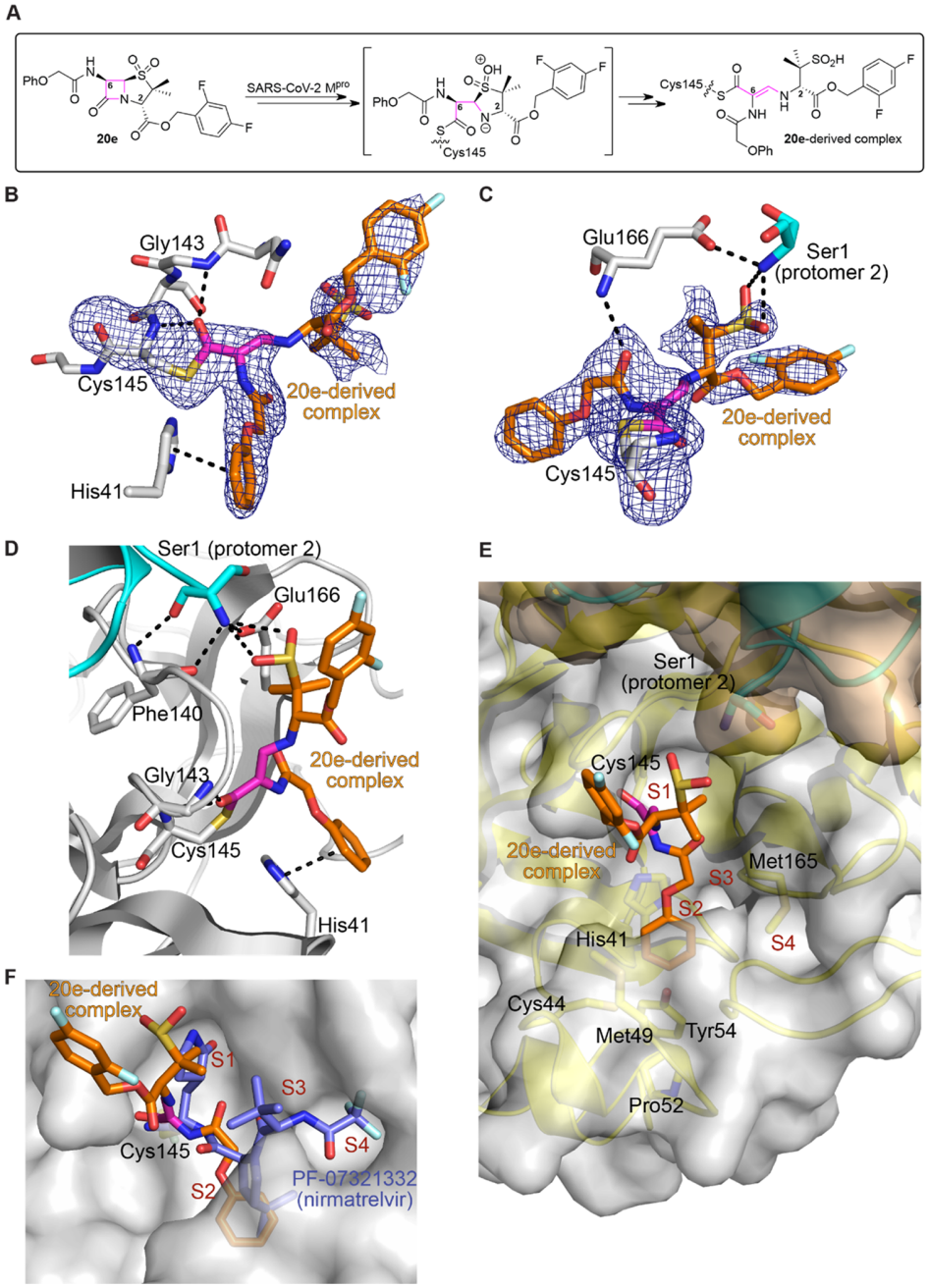


Figure 3. Crystallographic evidence that penicillin V sulfone ester derivatives inhibit M^{pro} by active site cysteine covalent modification. Color code: M^{pro}: gray (protomer 1) and cyan (protomer 2); carbon-backbone of the **20e**-derived complex is in orange, with the β-lactam ring-derived carbon-backbone in magenta; oxygen: red; nitrogen: blue; sulfur: yellow; and fluorine: light blue. (A) Reaction of penicillin sulfone **20e** with SARS-CoV-2 M^{pro}. (B, C) Representative OMIT electron density map (mF_o-DF_c) contoured to 2.5σ around Cys145 and the **20e**-derived complex showing clear evidence for (B) β-lactam ring opening by the active site Cys145 leading to thioester formation and (C) positioning of the SO_2H group of the **20e**-derived complex formed by opening of the thiazolidine sulfone ring to enable interactions with the main chain amino group of Ser1 of the second M^{pro} protomer. (D) Phe140, Glu166, and the SO_2H group of the **20e**-derived complex are positioned to interact with Ser1 of the second M^{pro} protomer. (E) C6 amido penicillin-derived side chain of the **20e**-derived complex binds in the hydrophobic S2 M^{pro} binding site. (F) Superimposition of active sites' views of the M^{pro}:**20e**-derived complex and the M^{pro}:PF-07321332 (slate blue: carbon-backbone of **8**, nirmatrelvir; PDB ID: 7VH8[68]) structures.

M^{pro} via a different mechanism (Supporting Information Figure S6), potentially involving noncovalent binding.

Crystallographic Evidence That Penicillin Derivatives Selectively Inhibit M^{pro} by *S*-Acylation of Cys145. To investigate the mode of M^{pro} inhibition by penicillin derivatives, we carried out crystallographic studies and obtained a structure of M^{pro} complexed with a penicillin sulfone **20e**-derived ligand following cocrystallization (*C*2 space group, 2.0 Å resolution; Supporting Information Figure S7). The structure was solved by molecular replacement using a reported M^{pro} structure (PDB ID: 6YB7[29]) as a search model. The overall fold of the structure is similar to those previously reported for M^{pro} (RMSD = 0.41 Å for M^{pro} complexed with PF-07321332 (**8**, nirmatrelvir; PDB ID: 7VH8);[68] Supporting Information Figure S7).

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

Consistent with our MS data (Figure 2), analysis of the electron density at the M^{pro}-ligand complex active site provides clear evidence for active site Cys145 *S*-acylation via a reaction with the β-lactam ring of penicillin **20e** (Figure 3A), in a manner reminiscent of the covalent reaction of β-lactamases with penicillins and L,D-transpeptidases employing a nucleophilic cysteine with carbapenems.[69,70] The thioester carbonyl of the penicillin sulfone **20e**-derived complex (corresponding to the β-lactam carbonyl of **20e**) is positioned to interact with the main chain amino group of Cys145 and Gly143 (3.0 and 2.8 Å, respectively) (Figure 3B). Note that no evidence for acylation of other M^{pro} residues by **20e** was observed in the crystal structure, indicating a selective covalent reaction, at least under the cocrystallization conditions, consistent with the MS studies.

The amide carbonyl of the C6 penicillin-derived side chain is positioned to interact with the main chain amino group of Glu166 (3.1 Å) (Figure 3C). Further, the C6 amido penicillin-derived side chain binds in the hydrophobic S2 binding site formed *inter alia* by the side chains of His41, Cys44, Met49, Pro52, Tyr54, and Met165. The C6 side chain phenyl group is positioned to π-stack in an offset manner with the imidazole side chain of His41 (3.8 Å), which is part of the catalytic dyad; analogous interactions have been observed with other M^{pro} inhibitors.[33,71] It could be that the C6 phenoxy ether oxygen, which is important for efficient inhibition (Table 3), helps position the phenyl group of the C6 side chain to productively interact with His41, though it may also be important in binding prior to covalent reactions leading to the crystallographically observed complex. Notably, the S1 binding pocket is not occupied in the M^{pro}:**20e**-derived complex structure; substitution at the penicillin sulfone C6 position is of interest in this regard and will be explored in future work. The penicillin C2-derived ester projects out the active site, rationalizing the relatively flat SAR at this position; however, as with C6 ether oxygen, the C2 ester may be important in initial inhibitor binding.

Interestingly, the crystallographic data imply that the opening of the penicillin thiazolidine sulfone ring of **20e** via C5−S bond cleavage follows an initial covalent reaction of Cys145 with the β-lactam, to give an acyclic enamine/imine. An analogous reaction occurs during serine β-lactamase inhibition by sulbactam and tazobactam.[72,73] We carried out trial refinements with both the enamine and imine complexes; analysis of the electron density implies the presence of a planar C5−C6 bond (penicillin numbering), suggesting the presence of the enamine, but we cannot rule out the additional partial presence of the imine tautomer. The SO_2H group formed by opening of the thiazolidine sulfone ring projects toward the side chain of Glu166 (3.8 Å) and the main chain amino group of Ser1 (3.0 and 3.3 Å), the latter being the N-terminus of the second protomer making up the functional M^{pro} dimer (Figure 3C−E).

DISCUSSION

Antibacterial drugs containing a β-lactam ring are among the most successful of all small-molecule therapeutics. Their mechanism of action involves reaction with a nucleophilic serine in bacterial transpeptidases to afford stable acyl−enzyme complexes.[74] They are also important inhibitors of serine β-lactamases, which are mechanistically related to transpeptidases.[75] Despite the widespread use of β-lactams as antibacterials and work showing they have the potential to inhibit other classes of nucleophilic enzymes,[76,77] including human[78−83] and viral[84−87] serine proteases, they have found limited utility in other therapeutic fields. The reasons for this are unclear but, at least in the case of bicyclic β-lactams such as penicillins, may in part reflect synthetic challenges and/or long-term stability issues. β-Lactams also have potential as useful inhibitors of nucleophilic cysteine enzymes, as shown, for example, by the inhibition of (i) human cathepsins by monocyclic β-lactams,[88,89] (ii) viral cysteine proteases by spirocyclic β-lactams,[90] and (iii) mycobacterial L,D-transpeptidases by bicyclic β-lactams.[69,70,91,92] Although there is considerable scope for further optimization, our results highlight the potential of β-lactams as covalently reacting inhibitors of SARS-CoV-2 M^{pro} and, by implication, other (viral) nucleophilic cysteine proteases, including SARS-CoV-2 PL^{pro}.

Recently, we reported MS-based SARS-CoV-2 M^{pro} and PL^{pro} assays, which monitor protease-catalyzed substrate hydrolysis and/or protease modification and which are suitable for inhibition studies.[15,93] The MS-based M^{pro} assays enabled the identification of certain β-lactams, notably the penicillin V sulfone benzyl ester **10**, as covalently reacting M^{pro} inhibitors.[15] In the current study, we report SAR studies that show the potency of **10** can be optimized by about 10-fold, i.e., from $IC_{50} \sim 6.5\ \mu M$ for **10** to $IC_{50} \sim 0.6\ \mu M$ for **28**, **31**, and **32** (Table 4). In general, for efficient M^{pro} inhibition, the penicillin sulfone oxidation state is preferred over the sulfoxide and sulfide oxidation states (Table 1); however, while the (*S*)-configured penicillin sulfoxides do not inhibit efficiently (Table 1, entries 3 and 4), the corresponding (*R*)-configured sulfoxides do inhibit (Table 4, entry 9). The importance of the oxidized sulfur in inhibition is consistent with the ring opening of the thiazolidine ring during inhibition, as supported by crystallographic studies with **20e** (Figure 3). Further, for penicillins bearing a C6 amido side chain, the (6*R*)-configuration is preferred over the (6*S*)-configuration (Table 1).

Protein-observed MS and crystallographic studies imply that, at least, some of the penam sulfones selectively react with the active site Cys145 thiol to give a stable acyl−enzyme complex (Figures 2 and 3). Crystallographic analysis revealed that, at least in the case of penicillin **20e**, β-lactam opening is followed by opening of the five-membered penicillin ring to give, at least predominantly, an acyclic enamine, a reaction precedented in the inhibition mechanisms of nucleophilic serine β-lactamases by clinically used drugs sulbactam and tazobactam.[72,73] Although care should be taken in assuming crystallographically observed complexes necessarily reflect those relevant in solution, the structure nonetheless highlights the potential for M^{pro} inhibition via cysteine-acylation and for subsequent reaction, leading to a stable acyl−enzyme complex.

The possibility of reactions subsequent to initial noncovalent binding/acylation may contribute to the rather complex SAR, including for the C2 ester derivatives, with both small alkyl (e.g., **20o** and **20p**) and benzyl esters being potent inhibitors (Table 2). In the M^{pro}:**20e**-derived complex structure, the C2 ester projects away from the active site; the interaction of the SO_2H group formed by opening of the thiazolidine sulfone ring with the main chain amino group of Ser1 of the second M^{pro} protomer may stabilize this conformation. It is likely that the C2 ester group occupies a different conformation prior to opening of the five-membered penicillin ring; thus, it may be important in initial M^{pro} binding. Modification of the C2 ester group may enable tuning of pharmacokinetic properties, for example, to optimize cell permeability. Crystal structure analysis suggests that the penicillin C6 amido penicillin side chain binds in the P3 binding pocket (Figure 3). The C6 phenoxyacetyl ether oxygen of

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

penicillin V sulfone derivatives are potent M^{pro} inhibitors in contrast to the penicillin G derivatives, which lack the ether oxygen (Table 3); the structure suggests the C6 phenoxy ether oxygen may help position the phenyl group of the C6 side chain to productively interact with the His41 imidazole ring, though it may also be important in binding prior to covalent reaction, leading to the crystallographically observed complex, potentially by interaction with the Asn142 side chain, as indicated by docking studies.[15]

It should be noted that a covalent reaction is not necessarily a prerequisite for useful inhibition of M^{pro} by a β-lactam.[15] Although the efficient reaction of penicillin and related bicyclic β-lactams with transpeptidases/β-lactamases is often proposed to reflect the reactive nature of the β-lactam ring, the bicyclic β-lactam ring system is also a mimic of a strained conformation of the scissile substrate peptide bond,[74] thus β-lactams have the potential as noncovalent M^{pro} inhibitors. Notably, a γ-lactam derivative (i.e., **17**) was less active than the analogous penicillin **10** (Table 1) but still clearly showed inhibition, suggesting lactams (and related acylating agents) other than β-lactams have potential as active site binding M^{pro} inhibitors, as reported to be the case for transpeptidases/β-lactamases[94,95] and a viral serine protease.[96]

Interestingly, among the most potent compounds identified in our work were the C6 dibromo-penicillin sulfones **28**, **31**, and **32** (Table 4); ongoing mechanistic studies on these compounds involving protein-observed MS suggest initial covalent M^{pro} binding is followed by rapid subsequent reaction to give new species. Although the precise mechanisms of action of these compounds remain to be determined, work on penicillin C6 bromo derivatives and β-lactamase inhibition has shown that related compounds can react to give acyl−enzyme complexes that undergo subsequent rearrangements.[59−62] These results suggest the unexploited potential for "mechanism-based" inhibition of M^{pro} and related nucleophilic cysteine enzymes, which may complement drug development efforts on mechanistically distinct substrate mimics, such as PF-07321332 (**8**, nirmatrelvir, Figures 1G and 3F).

CONCLUSIONS

The combined results highlight the potential of β-lactams, including penicillin derivatives prepared by semisynthesis from natural products, as covalently reacting M^{pro} inhibitors, though noncovalent inhibition by them is also possible. Given the proven efficacy of β-lactams and related covalently reacting groups as antibacterials and β-lactamase inhibitors, we suggest that they should be explored as antiviral drugs.

EXPERIMENTAL SECTION

The syntheses and characterizations of the penicillin derivatives used in this work are disclosed in the associated Supporting Information. All compounds are ≥95% pure by NMR and HPLC analysis unless stated otherwise. NMR spectra and HPLC traces are shown in the associated Supporting Information.

M^{pro} Inhibition Assays. SPE-MS M^{pro} inhibition assays were performed as reported,[15] however, using a 37mer peptide (ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2) as a substrate rather than an 11mer peptide (TSAVLQ/SGFRK-NH_2). The 37mer peptide was synthesized by solid-phase peptide synthesis as a C-terminal amide and purified by GL Biochem (Shanghai) Ltd. (Shanghai, China). Recombinant SARS-CoV-2 M^{pro} was prepared according to established procedures;[15] note that fresh aliquots, which were not frozen more than once, were used for inhibition assays. Solutions of the inhibitors (100% DMSO) were dry-dispensed across 384-well polypropylene assay plates (Greiner) in an approximately 3-fold and 11-point dilution series (100 μM top concentration) using an ECHO 550 acoustic dispenser (Labcyte). DMSO and formic acid were used as negative and positive inhibition controls, respectively. The final DMSO concentration was kept constant at 0.5%$_{v/v}$ throughout all experiments (using the DMSO backfill option of the acoustic dispenser). Each reaction was performed in technical duplicates in adjacent wells of the assay plates, and assays were performed in at least two independent duplicates.

In brief, the Enzyme Mixture (25 μL/well), containing M^{pro} (0.3 μM) in buffer (20 mM HEPES, pH 7.5, 50 mM NaCl), was dispensed across the inhibitor-containing 384-well assay plates with a multidrop dispenser (Thermo Fischer Scientific) at 20 °C under an ambient atmosphere. The plates were subsequently centrifuged (1000 rpm, 10 s) and incubated for 15 min at 20 °C. Note that we previously incubated M^{pro} with inhibitors for 30 or 60 min, resulting in more efficient inhibition.[15] The substrate mixture (25 μL/well), containing ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2 (4.0 μM) in buffer (20 mM HEPES, pH 7.5, 50 mM NaCl), was added using the multidrop dispenser. The plates were centrifuged (1000 rpm, 10 s), and after incubating for 6 min, the reaction was stopped by addition of 10%$_{v/v}$ aqueous formic acid (5 μL/well). The plates were then centrifuged (1000 rpm, 30 s) and analyzed by MS.

MS analyses were performed using a RapidFire RF 365 high-throughput sampling robot (Agilent) attached to an iFunnel Agilent 6550 accurate mass quadrupole time-of-flight (Q-TOF) mass spectrometer operated in the positive ionization mode. Assay samples were aspirated under vacuum for 0.6 s and loaded onto a C4 solid-phase extraction (SPE) cartridge. After loading, the C4 SPE cartridge was washed with 0.1%$_{v/v}$ aqueous formic acid to remove nonvolatile buffer salts (5.5 s, 1.5 mL/min). The peptide was eluted from the SPE cartridge with 0.1%$_{v/v}$ aqueous formic acid in 85/15$_{v/v}$ acetonitrile/water into the mass spectrometer (5.5 s, 1.5 mL/min), and the SPE cartridge was re-equilibrated with 0.1%$_{v/v}$ aqueous formic acid (0.5 s, 1.25 mL/min). The mass spectrometer parameters were as follows: capillary voltage (4000 V), nozzle voltage (1000 V), fragmentor voltage (365 V), gas temperature (280 °C), gas flow (13 L/min), sheath gas temperature (350 °C), and sheath gas flow (12 L/min). For data analysis, the m/z +3 charge state of the 37mer peptide (substrate) and the m/z +1 charge state of the SGFRKMAFPS-NH_2 C-terminal product peptide were used to extract and integrate ion chromatogram data using RapidFire Integrator software (Agilent). Data were exported into Microsoft Excel and used to calculate the % conversion using the equation: % conversion = 100 × (integral C-terminal product peptide)/(integral C-terminal product peptide + integral 37mer substrate peptide). Normalized dose−response curves (formic acid and DMSO controls) were obtained from the raw data by nonlinear regression (GraphPad Prism 9) and used to determine IC_{50}-values. For compounds **15**, **16**, **25g**, **25c**, **29**, and **30**, the 11mer peptide was used as the substrate.

Protein-Observed M^{pro} Assays. Solutions of the inhibitors (100% DMSO) were dry-dispensed across 384-well polypropylene assay plates (Greiner) (for 11 or 33 μM top concentrations) using an ECHO 550 acoustic dispenser (Labcyte). DMSO was used as a negative control. Each reaction was performed in technical duplicates. The enzyme mixture (50 μL/well), containing M^{pro} (2.0 μM) in buffer (20 mM HEPES, pH 7.5), was dispensed across the penicillin-containing 384-well assay plates with a multidrop dispenser (Thermo Fischer Scientific). The reaction mixture was incubated for 45 min at 20 °C under an ambient atmosphere prior to analysis by SPE-MS.

To investigate the importance of the covalent modification of the active site Cys145 for M^{pro} inhibition, M^{pro} (2.0 μM) was incubated with the selective Cys145-alkylating agent TPCK[15] (10 μM) in buffer (20 mM HEPES, pH 7.5) for 3 h at 0 °C. The mixture was then dispensed across the penicillin-containing 384-well assay plates with a multidrop dispenser (Thermo Fischer Scientific) and incubated for 1 h at 20 °C under an ambient atmosphere prior to analysis by SPE-MS.

MS analyses were performed using a RapidFire RF 365 high-throughput sampling robot (Agilent) attached to an iFunnel Agilent 6550 accurate mass Q-TOF mass spectrometer using a C4 cartridge and

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

the same parameters as described above, with the exception of the gas temperature that was reduced to 225 °C. Protein spectra were deconvoluted for the m/z range 850−1350 Da, with a resolution of 2 Da and with a 10−60 kDa cutoff using the MaxEnt1 function in Agilent MassHunter Version 7. The deconvoluted files were extracted as csv files, sorted using Enthought Canopy GUI, and normalized and plotted using GraphPad Prism 9.

Crystallization. A frozen SARS-CoV-2 M^{pro} solution was thawed and diluted to 6 mg/mL (using 20 mM HEPES, pH 7.5, 50 mM NaCl). β-Lactam **20e** was added to the protein solution to a final concentration of 10 mM; the mixture was incubated for 2 h at ambient temperature prior to dispensing the plates. The drop composition was: 0.15 μL protein−ligand solution, 0.3 μL 11%$_{v/v}$ PEG 4000, 0.1 M MES, pH 6.5, and 0.05 μL M^{pro} crystal seed stock. A M^{pro} crystal seed stock was prepared by crushing M^{pro} crystals with a pipette tip, suspending them in 30% PEG 4000, 5%$_{v/v}$ DMSO, 0.1 M MES pH 6.5, and vortexing for 60 s with approximately 10 glass beads (1.0 mm diameter, BioSpec products). The reservoir solution was: 11%$_{v/v}$ PEG 4K, 5%$_{v/v}$ DMSO, 0.1 M MES, pH 6.5. Crystals were grown using the sitting drop vapor diffusion method at 20 °C and appeared within 24 h, reaching full size within 36 h. Crystals were looped after 1 week.

Data Collection and Structure Determination. Diffraction data were collected on beamline I0-3 at the Diamond Light Source at 100 K using a wavelength of 0.9762 Å. Data were processed using Dials[97] via Xia2[98] and Aimless[99] within CCP4i2.[100] The datasets were phased using Molrep[101] and the M^{pro} apo structure (PDB ID: 6YB7). Ligand restraints were generated using AceDRG.[102] Typically, 97% of residues are in the favored regions of the Ramachandran plot, 2% in the allowed region, and 1% in high-energy conformations (2 residues). Crystal structures were manually rebuilt in Coot and refined using Refmac,[103] Buster,[104] and PDB_Redo (Supporting Information Table S3).[105]

The crystal structure data for SARS-CoV-2 M^{pro}:**20e**-derived complex have been deposited in the Protein Data Bank (PDB) with accession code 7Z59.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jmedchem.1c02214.

Inhibition assays, synthesis, NMR analysis, representative dose-response curves, and crystallographic analysis; the use of 37mer, data collection, and reference statistics; general synthetic procedures; and experimental procedures (PDF)

Molecular formula strings (CSV)

■ AUTHOR INFORMATION

Corresponding Authors

Lennart Brewitz − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*; Email: lennart.brewitz@chem.ox.ac.uk

Christopher J. Schofield − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*; orcid.org/0000-0002-0290-6565; Email: christopher.schofield@chem.ox.ac.uk

Authors

Tika R. Malla − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*

Dorian-Gabriel Muntean − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*

Hiba Aslam − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*

C. David Owen − *Diamond Light Source Ltd., OX11 0DE Didcot, United Kingdom; Research Complex at Harwell, OX11 0FA Didcot, United Kingdom*; orcid.org/0000-0001-5774-8202

Eidarus Salah − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*

Anthony Tumber − *Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, OX1 3TA Oxford, United Kingdom*

Petra Lukacik − *Diamond Light Source Ltd., OX11 0DE Didcot, United Kingdom; Research Complex at Harwell, OX11 0FA Didcot, United Kingdom*

Claire Strain-Damerell − *Diamond Light Source Ltd., OX11 0DE Didcot, United Kingdom; Research Complex at Harwell, OX11 0FA Didcot, United Kingdom*

Halina Mikolajek − *Diamond Light Source Ltd., OX11 0DE Didcot, United Kingdom; Research Complex at Harwell, OX11 0FA Didcot, United Kingdom*

Martin A. Walsh − *Diamond Light Source Ltd., OX11 0DE Didcot, United Kingdom; Research Complex at Harwell, OX11 0FA Didcot, United Kingdom*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.1c02214

Author Contributions

‖T.R.M. and L.B. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The investigators acknowledge the philanthropic support of the donors to the University of Oxford's COVID-19 Research Response Fund and King Abdulaziz University, Saudi Arabia, for funding. This research was funded in part by the Wellcome Trust (106244/Z/14/Z). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The authors acknowledge Diamond for the award of beamtime through the COVID-19 dedicated call (proposal ID MX27088) and thank the Diamond MX group for their support and expertise. The authors thank Cancer Research UK (C8717/A18245) and the Biotechnology and Biological Sciences Research Council (BB/J003018/1 and BB/R000344/1) for funding. T.R.M. was supported by the BBSRC (BB/M011224/1).

■ REFERENCES

(1) De Clercq, E.; Li, G. Approved antiviral drugs over the past 50 years. *Clin. Microbiol. Rev.* **2016**, *29*, 695−747.

(2) Gorbalenya, A. E.; Baker, S. C.; Baric, R. S.; de Groot, R. J.; Drosten, C.; Gulyaeva, A. A.; Haagmans, B. L.; Lauber, C.; Leontovich, A. M.; Neuman, B. W.; Penzar, D.; Perlman, S.; Poon, L. L. M.; Samborskiy, D. V.; Sidorov, I. A.; Sola, I.; Ziebuhr, J. The species severe

acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536−544.
(3) Cannalire, R.; Cerchia, C.; Beccari, A. R.; Di Leva, F. S.; Summa, V. Targeting SARS-CoV-2 proteases and polymerase for COVID-19 treatment: state of the art and future opportunities. *J. Med. Chem.* **2022**, *65*, 2716−2746.
(4) Zhu, W.; Shyr, Z.; Lo, D. C.; Zheng, W. Viral proteases as targets for coronavirus disease 2019 drug development. *J. Pharmacol. Exp. Ther.* **2021**, *378*, 166−172.
(5) Yang, H.; Yang, J. A review of the latest research on Mpro targeting SARS-COV inhibitors. *RSC Med. Chem.* **2021**, *12*, 1026−1036.
(6) Citarella, A.; Scala, A.; Piperno, A.; Micale, N. SARS-CoV-2 Mpro: a potential target for peptidomimetics and small-molecule inhibitors. *Biomolecules* **2021**, *11*, 607.
(7) Gao, K.; Wang, R.; Chen, J.; Tepe, J. J.; Huang, F.; Wei, G.-W. Perspectives on SARS-CoV-2 main protease inhibitors. *J. Med. Chem.* **2021**, *64*, 16922−16955.
(8) Anirudhan, V.; Lee, H.; Cheng, H.; Cooper, L.; Rong, L. Targeting SARS-CoV-2 viral proteases as a therapeutic strategy to treat COVID-19. *J. Med. Virol.* **2021**, *93*, 2722−2734.
(9) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289−293.
(10) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **2020**, *368*, 409−412.
(11) Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M. T.; Chen, Y.; Wang, J. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **2020**, *30*, 678−692.
(12) Dražić, T.; Kühl, N.; Leuthold, M. M.; Behnam, M. A. M.; Klein, C. D. Efficiency improvements and discovery of new substrates for a SARS-CoV-2 main protease FRET assay. *SLAS Discovery* **2021**, *26*, 1189−1199.
(13) Gurard-Levin, Z. A.; Liu, C.; Jekle, A.; Jaisinghani, R.; Ren, S.; Vandyck, K.; Jochmans, D.; Leyssen, P.; Neyts, J.; Blatt, L. M.; Beigelman, L.; Symons, J. A.; Raboisson, P.; Scholle, M. D.; Deval, J. Evaluation of SARS-CoV-2 3C-like protease inhibitors using self-assembled monolayer desorption ionization mass spectrometry. *Antiviral Res.* **2020**, *182*, 104924.
(14) Liu, C.; Boland, S.; Scholle, M. D.; Bardiot, D.; Marchand, A.; Chaltin, P.; Blatt, L. M.; Beigelman, L.; Symons, J. A.; Raboisson, P.; Gurard-Levin, Z. A.; Vandyck, K.; Deval, J. Dual inhibition of SARS-CoV-2 and human rhinovirus with protease inhibitors in clinical development. *Antiviral Res.* **2021**, *187*, 105020.
(15) Malla, T. R.; Tumber, A.; John, T.; Brewitz, L.; Strain-Damerell, C.; Owen, C. D.; Lukacik, P.; Chan, H. T. H.; Maheswaran, P.; Salah, E.; Duarte, F.; Yang, H.; Rao, Z.; Walsh, M. A.; Schofield, C. J. Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 Mpro inhibitors. *Chem. Commun.* **2021**, *57*, 1430−1433.
(16) Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; Song, H.; Shi, Y.; Li, X.; Tan, W.; Qi, J.; Gao, G. F. Both boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **2020**, *11*, 4417.
(17) Venkatraman, S.; Bogen, S. L.; Arasappan, A.; Bennett, F.; Chen, K.; Jao, E.; Liu, Y.-T.; Lovey, R.; Hendrata, S.; Huang, Y.; Pan, W.; Parekh, T.; Pinto, P.; Popov, V.; Pike, R.; Ruan, S.; Santhanam, B.; Vibulbhan, B.; Wu, W.; Yang, W.; Kong, J.; Liang, X.; Wong, J.; Liu, R.; Butkiewicz, N.; Chase, R.; Hart, A.; Agrawal, S.; Ingravallo, P.; Pichardo, J.; Kong, R.; Baroudy, B.; Malcolm, B.; Guo, Z.; Prongay, A.; Madison, V.; Broske, L.; Cui, X.; Cheng, K.-C.; Hsieh, Y.; Brisson, J.-M.; Prelusky, D.; Korfmacher, W.; White, R.; Bogdanowich-Knipp, S.; Pavlovsky, A.; Bradley, P.; Saksena, A. K.; Ganguly, A.; Piwinski, J.; Girijavallabhan, V.; Njoroge, F. G. Discovery of (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-zabicyclo[3.1.0]hexan-2(S)-carboxamide (SCH 503034), a selective, potent, orally bioavailable hepatitis C virus NS3 protease inhibitor: a potential therapeutic agent for the treatment of hepatitis C infection. *J. Med. Chem.* **2006**, *49*, 6074−6086.
(18) Njoroge, F. G.; Chen, K. X.; Shih, N.-Y.; Piwinski, J. J. Challenges in modern drug discovery: a case study of boceprevir, an HCV protease inhibitor for the treatment of hepatitis C virus infection. *Acc. Chem. Res.* **2008**, *41*, 50−59.
(19) Redhead, M. A.; Owen, C. D.; Brewitz, L.; Collette, A. H.; Lukacik, P.; Strain-Damerell, C.; Robinson, S. W.; Collins, P. M.; Schäfer, P.; Swindells, M.; Radoux, C. J.; Hopkins, I. N.; Fearon, D.; Douangamath, A.; von Delft, F.; Malla, T. R.; Vangeel, L.; Vercruysse, T.; Thibaut, J.; Leyssen, P.; Nguyen, T.-T.; Hull, M.; Tumber, A.; Hallett, D. J.; Schofield, C. J.; Stuart, D. I.; Hopkins, A. L.; Walsh, M. A. Bispecific repurposed medicines targeting the viral and immunological arms of COVID-19. *Sci. Rep.* **2021**, *11*, 13208.
(20) Elford, P. R.; Heng, R.; Révész, L.; MacKenzie, A. R. Reduction of inflammation and pyrexia in the rat by oral administration of SDZ 224-015, an inhibitor of the interleukin-1β converting enzyme. *Br. J. Pharmacol.* **1995**, *115*, 601−606.
(21) Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H. A.; McKay, R. T.; van Belkum, M. J.; Joyce, M. A.; Young, H. S.; Tyrrell, D. L.; Vederas, J. C.; Lemieux, M. J. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **2020**, *11*, 4282.
(22) Revesz, L.; Briswalter, C.; Heng, R.; Leutwiler, A.; Mueller, R.; Wuethrich, H.-J. Synthesis of P1 aspartate-based peptide acyloxymethyl and fluoromethyl ketones as inhibitors of interleukin-1β-converting enzyme. *Tetrahedron Lett.* **1994**, *35*, 9693−9696.
(23) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; Lockner, J. W.; Lou, J.; Marra, M. T.; Mitchell, L. J.; Murray, B. W.; Nieman, J. A.; Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J.; Solowiej, J. E.; Steppan, C. M.; Taggart, B. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J. Med. Chem.* **2020**, *63*, 12725−12747.
(24) Boras, B.; Jones, R. M.; Anson, B. J.; Arenson, D.; Aschenbrenner, L.; Bakowski, M. A.; Beutler, N.; Binder, J.; Chen, E.; Eng, H.; Hammond, H.; Hammond, J.; Haupt, R. E.; Hoffman, R.; Kadar, E. P.; Kania, R.; Kimoto, E.; Kirkpatrick, M. G.; Lanyon, L.; Lendy, E. K.; Lillis, J. R.; Logue, J.; Luthra, S. A.; Ma, C.; Mason, S. W.; McGrath, M. E.; Noell, S.; Obach, R. S.; O'Brien, M. N.; O'Connor, R.; Ogilvie, K.; Owen, D.; Pettersson, M.; Reese, M. R.; Rogers, T. F.; Rossulek, M. I.; Sathish, J. G.; Shirai, N.; Steppan, C.; Ticehurst, M.; Updyke, L. W.; Weston, S.; Zhu, Y.; Wang, J.; Chatterjee, A. K.; Mesecar, A. D.; Frieman, M. B.; Anderson, A. S.; Allerton, C. Discovery of a novel inhibitor of coronavirus 3CL protease for the potential treatment of COVID-19. *bioRxiv* **2021**, DOI: 10.1101/2020.09.12.293498.
(25) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science* **2021**, *374*, 1586−1593.
(26) Qiao, J.; Li, Y.-S.; Zeng, R.; Liu, F.-L.; Luo, R.-H.; Huang, C.; Wang, Y.-F.; Zhang, J.; Quan, B.; Shen, C.; Mao, X.; Liu, X.; Sun, W.; Yang, W.; Ni, X.; Wang, K.; Xu, L.; Duan, Z.-L.; Zou, Q.-C.; Zhang, H.-L.; Qu, W.; Long, Y.-H.-P.; Li, M.-H.; Yang, R.-C.; Liu, X.; You, J.; Zhou, Y.; Yao, R.; Li, W.-P.; Liu, J.-M.; Chen, P.; Liu, Y.; Lin, G.-F.; Yang, X.; Zou, J.; Li, L.; Hu, Y.; Lu, G.-W.; Li, W.-M.; Wei, Y.-Q.; Zheng, Y.-T.; Lei, J.; Yang, S. SARS-CoV-2 Mpro inhibitors with

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

antiviral activity in a transgenic mouse model. *Science* **2021**, *371*, 1374−1378.
(27) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An overview of severe acute respiratory syndrome−coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. *J. Med. Chem.* **2016**, *59*, 6595−6628.
(28) Ghosh, A. K.; Brindisi, M.; Shahabi, D.; Chapman, M. E.; Mesecar, A. D. Drug development and medicinal chemistry efforts toward SARS-coronavirus and covid-19 therapeutics. *ChemMedChem* **2020**, *15*, 907−932.
(29) Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Ábrányi-Balogh, P.; Brandão-Neto, J.; Carbery, A.; Davison, G.; Dias, A.; Downes, T. D.; Dunnett, L.; Fairhead, M.; Firth, J. D.; Jones, S. P.; Keeley, A.; Keserü, G. M.; Klein, H. F.; Martin, M. P.; Noble, M. E. M.; O'Brien, P.; Powell, A.; Reddi, R. N.; Skyner, R.; Snee, M.; Waring, M. J.; Wild, C.; London, N.; von Delft, F.; Walsh, M. A. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat. Commun.* **2020**, *11*, 5047.
(30) Günther, S.; Reinke, P. Y. A.; Fernández-García, Y.; Lieske, J.; Lane, T. J.; Ginn, H. M.; Koua, F. H. M.; Ehrt, C.; Ewert, W.; Oberthuer, D.; Yefanov, O.; Meier, S.; Lorenzen, K.; Krichel, B.; Kopicki, J.-D.; Gelisio, L.; Brehm, W.; Dunkel, I.; Seychell, B.; Gieseler, H.; Norton-Baker, B.; Escudero-Pérez, B.; Domaracky, M.; Saouane, S.; Tolstikova, A.; White, T. A.; Hänle, A.; Groessler, M.; Fleckenstein, H.; Trost, F.; Galchenkova, M.; Gevorkov, Y.; Li, C.; Awel, S.; Peck, A.; Barthelmess, M.; Schlünzen, F.; Lourdu, X. P.; Werner, N.; Andaleeb, H.; Ullah, N.; Falke, S.; Srinivasan, V.; França Bruno, A.; Schwinzer, M.; Brognaro, H.; Rogers, C.; Melo, D.; Zaitseva-Doyle, J. J.; Knoska, J.; Peña-Murillo, G. E.; Mashhour Aida, R.; Hennicke, V.; Fischer, P.; Hakanpää, J.; Meyer, J.; Gribbon, P.; Ellinger, B.; Kuzikov, M.; Wolf, M.; Beccari, A. R.; Bourenkov, G.; von Stetten, D.; Pompidor, G.; Bento, I.; Panneerselvam, S.; Karpics, I.; Schneider, T. R.; Garcia-Alai Maria, M.; Niebling, S.; Günther, C.; Schmidt, C.; Schubert, R.; Han, H.; Boger, J.; Monteiro, D. C. F.; Zhang, L.; Sun, X.; Pletzer-Zelgert, J.; Wollenhaupt, J.; Feiler, C. G.; Weiss, M. S.; Schulz, E.-C.; Mehrabi, P.; Karničar, K.; Usenik, A.; Loboda, J.; Tidow, H.; Chari, A.; Hilgenfeld, R.; Uetrecht, C.; Cox, R.; Zaliani, A.; Beck, T.; Rarey, M.; Günther, S.; Turk, D.; Hinrichs, W.; Chapman, H. N.; Pearson, A. R.; Betzel, C.; Meents, A. X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. *Science* **2021**, *372*, 642−646.
(31) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; Hilgenfeld, R.; Yuen, K. Y.; Wong, L.; Gao, G.; Chen, S.; Chen, Z.; Ma, D.; Bartlam, M.; Rao, Z. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* **2005**, *3*, e324.
(32) Hammond, J.; Leister-Tebbe, H.; Gardner, A.; Abreu, P.; Bao, W.; Wisemandle, W.; Baniecki, M.; Hendrick, V. M.; Damle, B.; Simón-Campos, A.; Pypstra, R.; Rusnak, J. M. Oral nirmatrelvir for high-risk, nonhospitalized adults with covid-19. *N. Engl. J. Med.* **2022**, *386*, 1397−1408.
(33) Yang, K. S.; Ma, X. R.; Ma, Y.; Alugubelli, Y. R.; Scott, D. A.; Vatansever, E. C.; Drelich, A. K.; Sankaran, B.; Geng, Z. Z.; Blankenship, L. R.; Ward, H. E.; Sheng, Y. J.; Hsu, J. C.; Kratch, K. C.; Zhao, B.; Hayatshahi, H. S.; Liu, J.; Li, P.; Fierke, C. A.; Tseng, C.-T. K.; Xu, S.; Liu, W. R. A quick route to multiple highly potent SARS-CoV-2 main protease inhibitors. *ChemMedChem* **2021**, *16*, 942−948.
(34) Dai, W.; Jochmans, D.; Xie, H.; Yang, H.; Li, J.; Su, H.; Chang, D.; Wang, J.; Peng, J.; Zhu, L.; Nian, Y.; Hilgenfeld, R.; Jiang, H.; Chen, K.; Zhang, L.; Xu, Y.; Neyts, J.; Liu, H. Design, synthesis, and biological evaluation of peptidomimetic aldehydes as broad-spectrum inhibitors against enterovirus and SARS-CoV-2. *J. Med. Chem.* **2022**, *65*, 2794−2808.
(35) Dampalla, C. S.; Kim, Y.; Bickmeier, N.; Rathnayake, A. D.; Nguyen, H. N.; Zheng, J.; Kashipathy, M. M.; Baird, M. A.; Battaile, K. P.; Lovell, S.; Perlman, S.; Chang, K.-O.; Groutas, W. C. Structure-guided design of conformationally constrained cyclohexane inhibitors of severe acute respiratory syndrome coronavirus-2 3CL protease. *J. Med. Chem.* **2021**, *64*, 10047−10058.
(36) Bai, B.; Belovodskiy, A.; Hena, M.; Kandadai, A. S.; Joyce, M. A.; Saffran, H. A.; Shields, J. A.; Khan, M. B.; Arutyunova, E.; Lu, J.; Bajwa, S. K.; Hockman, D.; Fischer, C.; Lamer, T.; Vuong, W.; van Belkum, M. J.; Gu, Z.; Lin, F.; Du, Y.; Xu, J.; Rahim, M.; Young, H. S.; Vederas, J. C.; Tyrrell, D. L.; Lemieux, M. J.; Nieman, J. A. Peptidomimetic α-acyloxymethylketone warheads with six-membered lactam P1 glutamine mimic: SARS-CoV-2 3CL protease inhibition, coronavirus antiviral activity, and in vitro biological stability. *J. Med. Chem.* **2022**, *65*, 2905−2925.
(37) Breidenbach, J.; Lemke, C.; Pillaiyar, T.; Schäkel, L.; Al Hamwi, G.; Diett, M.; Gedschold, R.; Geiger, N.; Lopez, V.; Mirza, S.; Namasivayam, V.; Schiedel, A. C.; Sylvester, K.; Thimm, D.; Vielmuth, C.; Phuong Vu, L.; Zyulina, M.; Bodem, J.; Gütschow, M.; Müller, C. E. Targeting the main protease of SARS-CoV-2: from the establishment of high throughput screening to the design of tailored inhibitors. *Angew. Chem., Int. Ed.* **2021**, *60*, 10423−10429.
(38) Shcherbakov, D.; Baev, D.; Kalinin, M.; Dalinger, A.; Chirkova, V.; Belenkaya, S.; Khvostov, A.; Krut'ko, D.; Medved'ko, A.; Volosnikova, E.; Sharlaeva, E.; Shanshin, D.; Tolstikova, T.; Yarovaya, O.; Maksyutov, R.; Salakhutdinov, N.; Vatsadze, S. Design and evaluation of bispidine-based SARS-CoV-2 main protease inhibitors. *ACS Med. Chem. Lett.* **2022**, *13*, 140−147.
(39) Han, S. H.; Goins, C. M.; Arya, T.; Shin, W.-J.; Maw, J.; Hooper, A.; Sonawane, D. P.; Porter, M. R.; Bannister, B. E.; Crouch, R. D.; Lindsey, A. A.; Lakatos, G.; Martinez, S. R.; Alvarado, J.; Akers, W. S.; Wang, N. S.; Jung, J. U.; Macdonald, J. D.; Stauffer, S. R. Structure-based optimization of ML300-derived, noncovalent inhibitors targeting the severe acute respiratory syndrome coronavirus 3CL protease (SARS-CoV-2 3CLpro). *J. Med. Chem.* **2022**, *65*, 2880−2904.
(40) Deshmukh, M. G.; Ippolito, J. A.; Zhang, C.-H.; Stone, E. A.; Reilly, R. A.; Miller, S. J.; Jorgensen, W. L.; Anderson, K. S. Structure-guided design of a perampanel-derived pharmacophore targeting the SARS-CoV-2 main protease. *Structure* **2021**, *29*, 823−833.
(41) Lockbaum, G. J.; Reyes, A. C.; Lee, J. M.; Tilvawala, R.; Nalivaika, E. A.; Ali, A.; Kurt Yilmaz, N.; Thompson, P. R.; Schiffer, C. A. Crystal structure of SARS-CoV-2 main protease in complex with the non-covalent inhibitor ML188. *Viruses* **2021**, *13*, 174.
(42) Sun, L.-Y.; Chen, C.; Su, J.; Li, J.-Q.; Jiang, Z.; Gao, H.; Chigan, J.-Z.; Ding, H.-H.; Zhai, L.; Yang, K.-W. Ebsulfur and ebselen as highly potent scaffolds for the development of potential SARS-CoV-2 antivirals. *Bioorg. Chem.* **2021**, *112*, 104889.
(43) Thun-Hohenstein, S. T. D.; Suits, T. F.; Malla, T. R.; Tumber, A.; Brewitz, L.; Salah, E.; Choudhry, H.; Schofield, C. J. Structure-activity studies reveal scope for optimisation of ebselen-type inhibition of SARS-CoV-2 main protease. *ChemMedChem* **2022**, *17*, e202100582.
(44) Konno, S.; Kobayashi, K.; Senda, M.; Funai, Y.; Seki, Y.; Tamai, I.; Schäkel, L.; Sakata, K.; Pillaiyar, T.; Taguchi, A.; Taniguchi, A.; Gütschow, M.; Müller, C. E.; Takeuchi, K.; Hirohama, M.; Kawaguchi, A.; Kojima, M.; Senda, T.; Shirasaka, Y.; Kamitani, W.; Hayashi, Y. 3CL protease inhibitors with an electrophilic arylketone moiety as anti-SARS-CoV-2 agents. *J. Med. Chem.* **2022**, *65*, 2926−2939.
(45) Stille, J. K.; Tjutrins, J.; Wang, G.; Venegas, F. A.; Hennecker, C.; Rueda, A. M.; Sharon, I.; Blaine, N.; Miron, C. E.; Pinus, S.; Labarre, A.; Plescia, J.; Patrascu, M. B.; Zhang, X.; Wahba, A. S.; Vlaho, D.; Huot, M. J.; Schmeing, T. M.; Mittermaier, A. K.; Moitessier, N. Design, synthesis and in vitro evaluation of novel SARS-CoV-2 3CLpro covalent inhibitors. *Eur. J. Med. Chem.* **2022**, *229*, 114046.
(46) Chamakuri, S.; Lu, S.; Ucisik, M. N.; Bohren, K. M.; Chen, Y.-C.; Du, H.-C.; Faver, J. C.; Jimmidi, R.; Li, F.; Li, J.-Y.; Nyshadham, P.; Palmer, S. S.; Pollet, J.; Qin, X.; Ronca, S. E.; Sankaran, B.; Sharma, K. L.; Tan, Z.; Versteeg, L.; Yu, Z.; Matzuk, M. M.; Palzkill, T.; Young, D. W. DNA-encoded chemistry technology yields expedient access to SARS-CoV-2 Mpro inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, e2111172118.
(47) Zhu, W.; Xu, M.; Chen, C. Z.; Guo, H.; Shen, M.; Hu, X.; Shinn, P.; Klumpp-Thomas, C.; Michael, S. G.; Zheng, W. Identification of SARS-CoV-2 3CL protease inhibitors by a quantitative high-throughput screening. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 1008−1016.

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696

(48) Ghosh, A. K.; Raghavaiah, J.; Shahabi, D.; Yadav, M.; Anson, B. J.; Lendy, E. K.; Hattori, S.-i.; Higashi-Kuwata, N.; Mitsuya, H.; Mesecar, A. D. Indole chloropyridinyl ester-derived SARS-CoV-2 3CLpro inhibitors: enzyme inhibition, antiviral efficacy, structure−activity relationship, and X-ray structural studies. *J. Med. Chem.* **2021**, *64*, 14702−14714.
(49) Ma, C.; Xia, Z.; Sacco, M. D.; Hu, Y.; Townsend, J. A.; Meng, X.; Choza, J.; Tan, H.; Jang, J.; Gongora, M. V.; Zhang, X.; Zhang, F.; Xiang, Y.; Marty, M. T.; Chen, Y.; Wang, J. Discovery of di- and trihaloacetamides as covalent SARS-CoV-2 main protease inhibitors with high target specificity. *J. Am. Chem. Soc.* **2021**, *143*, 20697−20709.
(50) Watkins, R. R.; Bonomo, R. A. 140 - β-Lactam Antibiotics. In *Infectious Diseases*, 4th ed.; Cohen, J.; Powderly, W. G.; Opal, S. M., Eds.; Elsevier, 2017; pp 1203−1216.
(51) *Antibiotic and Chemotherapy*, 9th ed.; Finch, R. G.; Greenwood, D.; Norrby, S. R.; Whitley, R. J., Eds.; W.B. Saunders: London, 2010.
(52) Baldwin, J. E.; Chakravarti, B.; Field, L. D.; Murphy, J. A.; Whitten, K. R.; Abraham, E. P.; Jayatilake, G. The synthesis of L-α-aminoadipyl-L-cysteinyl-D-3,4-didehydrovaline, a potent inhibitor of isopenicillin synthetase. *Tetrahedron* **1982**, *38*, 2773−2776.
(53) Claes, P.; Vlietinck, A.; Roets, E.; Vanderhaeghe, H.; Toppet, S. Preparation and deoxygenation of 6-epi-penicillin S-oxides. *J. Chem. Soc., Perkin Trans. 1* **1973**, 932−937.
(54) Westwood, N. J.; Claridge, T. D. W.; Edwards, P. N.; Schofield, C. J. Reversible acylation of elastase by γ-lactam analogues of β-lactam inhibitors. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2973−2978.
(55) Imming, P.; Klar, B.; Dix, D. Hydrolytic stability versus ring size in lactams: implications for the development of lactam antibiotics and other serine protease inhibitors. *J. Med. Chem.* **2000**, *43*, 4328−4331.
(56) Cooper, R. D. G.; DeMarco, P. V.; Cheng, J. C.; Jones, N. D. Structural studies on penicillin derivatives. I. Configuration of phenoxymethylpenicillin sulfoxide. *J. Am. Chem. Soc.* **1969**, *91*, 1408−1415.
(57) Shin, W.; Kim, J.; Kim, J. Structure of penicillin V benzyl ester sulfoxide. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **1992**, *48*, 1449−1451.
(58) El-Faham, A.; Albericio, F. COMU: a third generation of uronium-type coupling reagents. *J. Pept. Sci.* **2010**, *16*, 6−9.
(59) Cohen, S. A.; Pratt, R. F. Inactivation of bacillus cereus β-lactamase I by 6β-bromopenicillanic acid: mechanism. *Biochemistry* **1980**, *19*, 3996−4003.
(60) Kemp, J. E. G.; Closier, M. D.; Narayanaswami, S.; Stefaniak, M. H. Nucleophilic SN2 displacements on penicillin-6- and cephalosporin-7- triflates; 6β-iodopenicillanic acid, a new β-lactamase inhibitor. *Tetrahedron Lett.* **1980**, *21*, 2991−2994.
(61) Orlek, B. S.; Sammes, P. G.; Knott-Hunziker, V.; Waley, S. G. On the chemical binding of 6β-bromopenicillanic acid to β-lactamase I. *J. Chem. Soc., Chem. Commun.* **1979**, *21*, 962−963.
(62) Sauvage, E.; Zervosen, A.; Dive, G.; Herman, R.; Amoroso, A.; Joris, B.; Fonzé, E.; Pratt, R. F.; Luxen, A.; Charlier, P.; Kerff, F. Structural basis of the inhibition of class A β-lactamases and penicillin-binding proteins by 6-β-iodopenicillanate. *J. Am. Chem. Soc.* **2009**, *131*, 15262−15269.
(63) Brennan, J.; Hussain, F. H. S. The synthesis of penicillanate esters from ultrasonically formed organozinc intermediates. *Synthesis* **1985**, *1985*, 749−751.
(64) Zhang, H. L.; Zhang, Y. M.; Liu, P.; Wang, Y. Q. Synthesis and crystal of methyl 6,6-dihydropenicillanate S,S-dioxide. *Synth. React. Inorg. Met.-Org. Chem.* **2012**, *42*, 1083−1086.
(65) Miyashita, K.; Massova, I.; Taibi, P.; Mobashery, S. Design, synthesis, and evaluation of a potent mechanism-based inhibitor for the TEM β-lactamase with implications for the enzyme mechanism. *J. Am. Chem. Soc.* **1995**, *117*, 11055−11059.
(66) Carcione, D.; Siracusa, C.; Sulejmani, A.; Leoni, V.; Intra, J. Old and new beta-lactamase inhibitors: molecular structure, mechanism of action, and clinical use. *Antibiotics* **2021**, *10*, 995.
(67) Rafailidis, P. I.; Ioannidou, E. N.; Falagas, M. E. Ampicillin/sulbactam. *Drugs* **2007**, *67*, 1829−1849.
(68) Zhao, Y.; Fang, C.; Zhang, Q.; Zhang, R.; Zhao, X.; Duan, Y.; Wang, H.; Zhu, Y.; Feng, L.; Zhao, J.; Shao, M.; Yang, X.; Zhang, L.; Peng, C.; Yang, K.; Ma, D.; Rao, Z.; Yang, H. Crystal structure of SARS-CoV-2 main protease in complex with protease inhibitor PF-07321332. *Protein Cell* **2021**, DOI: 10.1007/s13238-021-00883-2.
(69) Cordillot, M.; Dubée, V.; Triboulet, S.; Dubost, L.; Marie, A.; Hugonnet, J.-E.; Arthur, M.; Mainardi, J.-L. In vitro cross-linking of *Mycobacterium tuberculosis* peptidoglycan by L,D-transpeptidases and inactivation of these enzymes by carbapenems. *Antimicrob. Agents Chemother.* **2013**, *57*, 5940−5945.
(70) Kumar, P.; Kaushik, A.; Lloyd, E. P.; Li, S.-G.; Mattoo, R.; Ammerman, N. C.; Bell, D. T.; Perryman, A. L.; Zandi, T. A.; Ekins, S.; Ginell, S. L.; Townsend, C. A.; Freundlich, J. S.; Lamichhane, G. Non-classical transpeptidases yield insight into new antibacterials. *Nat. Chem. Biol.* **2017**, *13*, 54−61.
(71) Sacco Michael, D.; Ma, C.; Lagarias, P.; Gao, A.; Townsend Julia, A.; Meng, X.; Dube, P.; Zhang, X.; Hu, Y.; Kitamura, N.; Hurst, B.; Tarbet, B.; Marty Michael, T.; Kolocouris, A.; Xiang, Y.; Chen, Y.; Wang, J. Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci. Adv.* **2020**, *6*, eabe0751.
(72) Padayatti, P. S.; Helfand, M. S.; Totir, M. A.; Carey, M. P.; Carey, P. R.; Bonomo, R. A.; van den Akker, F. High resolution crystal structures of the trans-enamine intermediates formed by sulbactam and clavulanic acid and E166A SHV-1 β-lactamase. *J. Biol. Chem.* **2005**, *280*, 34900−34907.
(73) Padayatti, P. S.; Helfand, M. S.; Totir, M. A.; Carey, M. P.; Hujer, A. M.; Carey, P. R.; Bonomo, R. A.; van den Akker, F. Tazobactam forms a stoichiometric trans-enamine intermediate in the E166A variant of SHV-1 β-lactamase: 1.63 Å crystal structure. *Biochemistry* **2004**, *43*, 843−848.
(74) Zapun, A.; Contreras-Martel, C.; Vernet, T. Penicillin-binding proteins and β-lactam resistance. *FEMS Microbiol. Rev.* **2008**, *32*, 361−385.
(75) González-Bello, C.; Rodríguez, D.; Pernas, M.; Rodríguez, Á.; Colchón, E. β-Lactamase inhibitors to restore the efficacy of antibiotics against superbugs. *J. Med. Chem.* **2020**, *63*, 1859−1881.
(76) Veinberg, G.; Potorocina, I.; Vorona, M. Recent trends in the design, synthesis and biological exploration of β-lactams. *Curr. Med. Chem.* **2013**, *21*, 393−416.
(77) Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **2002**, *102*, 4639−4750.
(78) Doherty, J. B.; Ashe, B. M.; Argenbright, L. W.; Barker, P. L.; Bonney, R. J.; Chandler, G. O.; Dahlgren, M. E.; Dorn, C. P.; Finke, P. E.; Firestone, R. A.; Fletcher, D.; Hagmann, W. K.; Mumford, R.; O'Grady, L.; Maycock, A. L.; Pisano, J. M.; Shah, S. K.; Thompson, K. R.; Zimmerman, M. Cephalosporin antibiotics can be modified to inhibit human leukocyte elastase. *Nature* **1986**, *322*, 192−194.
(79) Han, W. T.; Trehan, A. K.; Kim Wright, J. J.; Federici, M. E.; Seiler, S. M.; Meanwell, N. A. Azetidin-2-one derivatives as inhibitors of thrombin. *Bioorg. Med. Chem.* **1995**, *3*, 1123−1143.
(80) Adlington, R. M.; Baldwin, J. E.; Chen, B.; Cooper, S. L.; McCoull, W.; Pritchard, G. J.; Howe, T. J.; Becker, G. W.; Hermann, R. B.; McNulty, A. M.; Neubauer, B. L. Design and synthesis of novel monocyclic β-lactam inhibitors of prostate specific antigen. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1689−1694.
(81) Thompson, K. R.; Finke, P. E.; Shah, S. K.; Ashe, B. M.; Dahlgren, M. E.; Maycock, A. L.; Doherty, J. B. Inhibition of human leukocyte elastase. 6. Inhibition by 6-substituted penicillin esters. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 2283−2288.
(82) Sutton, J. C.; Bolton, S. A.; Davis, M. E.; Hartl, K. S.; Jacobson, B.; Mathur, A.; Ogletree, M. L.; Slusarchyk, W. A.; Zahler, R.; Seiler, S. M.; Bisacchi, G. S. Solid-phase synthesis and SAR of 4-carboxy-2-azetidinone mechanism-based tryptase inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2233−2239.
(83) Aoyama, Y.; Uenaka, M.; Konoike, T.; Iso, Y.; Nishitani, Y.; Kanda, A.; Naya, N.; Nakajima, M. Synthesis and structure−activity

relationships of a new class of 1-oxacephem-based human chymase inhibitors. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2397−2401.
(84) Yoakim, C.; Ogilvie, W. W.; Cameron, D. R.; Chabot, C.; Grand-Maître, C.; Guse, I.; Haché, B.; Kawai, S.; Naud, J.; O'Meara, J. A.; Plante, R.; Déziel, R. Potent β-lactam inhibitors of human cytomegalovirus protease. *Antiviral Chem. Chemother.* **1998**, *9*, 379−387.
(85) Borthwick, A. D.; Weingarten, G.; Haley, T. M.; Tomaszewski, M.; Wang, W.; Hu, Z.; Bedard, J.; Jin, H.; Yuen, L.; Mansour, T. S. Design and synthesis of monocyclic β-lactams as mechanism-based inhibitors of human cytomegalovirus protease. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 365−370.
(86) Sperka, T.; Pitlik, J.; Bagossi, P.; Tözsér, J. Beta-lactam compounds as apparently uncompetitive inhibitors of HIV-1 protease. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3086−3090.
(87) Dražić, T.; Kopf, S.; Corridan, J.; Leuthold, M. M.; Bertoša, B.; Klein, C. D. Peptide-β-lactam inhibitors of dengue and west nile virus NS2B-NS3 protease display two distinct binding modes. *J. Med. Chem.* **2020**, *63*, 140−156.
(88) Setti, E. L.; Davis, D.; Chung, T.; McCarter, J. 3,4-Disubstituted azetidinones as selective inhibitors of the cysteine protease cathepsin K. Exploring P2 elements for selectivity. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2051−2053.
(89) Zhou, N. E.; Guo, D.; Thomas, G.; Reddy, A. V. N.; Kaleta, J.; Purisima, E.; Menard, R.; Micetich, R. G.; Singh, R. 3-Acylamino-azetidin-2-one as a novel class of cysteine proteases inhibitors. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 139−141.
(90) Skiles, J. W.; McNeil, D. Spiro indolinone beta-lactams, inhibitors of poliovirus and rhinovirus 3C-proteinases. *Tetrahedron Lett.* **1990**, *31*, 7277−7280.
(91) Munnik, M.; Lohans, C. T.; Langley, G. W.; Bon, C.; Brem, J.; Schofield, C. J. A fluorescence-based assay for screening β-lactams targeting the *Mycobacterium tuberculosis* transpeptidase LdtMt2. *ChemBioChem* **2020**, *21*, 368−372.
(92) Steiner, E. M.; Schneider, G.; Schnell, R. Binding and processing of β-lactam antibiotics by the transpeptidase LdtMt2 from *Mycobacterium tuberculosis*. *FEBS J.* **2017**, *284*, 725−741.
(93) Brewitz, L.; Kamps, J. J. A. G.; Lukacik, P.; Strain-Damerell, C.; Zhao, Y.; Tumber, A.; Malla, T. R.; Orville, A. M.; Walsh, M. A.; Schofield, C. J. Mass spectrometric assays reveal discrepancies in inhibition profiles for the SARS-CoV-2 papain-like protease. *ChemMedChem* **2022**, *17*, e202200016.
(94) Durand-Reville, T. F.; Miller, A. A.; O'Donnell, J. P.; Wu, X.; Sylvester, M. A.; Guler, S.; Iyer, R.; Shapiro, A. B.; Carter, N. M.; Velez-Vega, C.; Moussa, S. H.; McLeod, S. M.; Chen, A.; Tanudra, A. M.; Zhang, J.; Comita-Prevoir, J.; Romero, J. A.; Huynh, H.; Ferguson, A. D.; Horanyi, P. S.; Mayclin, S. J.; Heine, H. S.; Drusano, G. L.; Cummings, J. E.; Slayden, R. A.; Tommasi, R. A. Rational design of a new antibiotic class for drug-resistant infections. *Nature* **2021**, *597*, 698−702.
(95) Wang, D. Y.; Abboud, M. I.; Markoulides, M. S.; Brem, J.; Schofield, C. J. The road to avibactam: the first clinically useful non-β-lactam working somewhat like a β-lactam. *Future Med. Chem.* **2016**, *8*, 1063−1084.
(96) Borthwick, A. D.; Davies, D. E.; Ertl, P. F.; Exall, A. M.; Haley, T. M.; Hart, G. J.; Jackson, D. L.; Parry, N. R.; Patikis, A.; Trivedi, N.; Weingarten, G. G.; Woolven, J. M. Design and synthesis of pyrrolidine-5,5′-trans-lactams (5-oxo-hexahydropyrrolo[3,2-b]pyrroles) as novel mechanism-based inhibitors of human cytomegalovirus protease. 4. Antiviral activity and plasma stability. *J. Med. Chem.* **2003**, *46*, 4428−4449.
(97) Winter, G.; Beilsten-Edmands, J.; Devenish, N.; Gerstel, M.; Gildea, R. J.; McDonagh, D.; Pascal, E.; Waterman, D. G.; Williams, B. H.; Evans, G. DIALS as a toolkit. *Protein Sci.* **2022**, *31*, 232−250.
(98) Winter, G.; Lobley, C. M. C.; Prince, S. M. Decision making in xia2. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1260−1273.
(99) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204−1214.
(100) Potterton, L.; Agirre, J.; Ballard, C.; Cowtan, K.; Dodson, E.; Evans, P. R.; Jenkins, H. T.; Keegan, R.; Krissinel, E.; Stevenson, K.; Lebedev, A.; McNicholas, S. J.; Nicholls, R. A.; Noble, M.; Pannu, N. S.; Roth, C.; Sheldrick, G.; Skubak, P.; Turkenburg, J.; Uski, V.; von Delft, F.; Waterman, D.; Wilson, K.; Winn, M.; Wojdyr, M. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Struct. Biol.* **2018**, *74*, 68−84.
(101) Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *D66*, 22−25.
(102) Long, F.; Nicholls, R. A.; Emsley, P.; Grazulis, S.; Merkys, A.; Vaitkus, A.; Murshudov, G. N. AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr., Sect. D: Struct. Biol.* **2017**, *73*, 112−122.
(103) Winn, M. D.; Murshudov, G. N.; Papiz, M. Z. Macromolecular TLS Refinement in REFMAC at Moderate Resolutions. In *Methods in Enzymology*; Academic Press, 2003; Vol. *374*, pp 300−321.
(104) Smart, O. S.; Womack, T. O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 368−380.
(105) Joosten, R. P.; Joosten, K.; Murshudov, G. N.; Perrakis, A. PDB_REDO: constructive validation, more than just looking for errors. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 484−496.

Recommended by ACS

Targeting SARS-CoV-2 Main Protease: A Successful Story Guided by an *In Silico* Drug Repurposing Approach
Francesca Alessandra Ambrosio, Anna Artese, *et al.*
MAY 25, 2023
JOURNAL OF CHEMICAL INFORMATION AND MODELING
READ

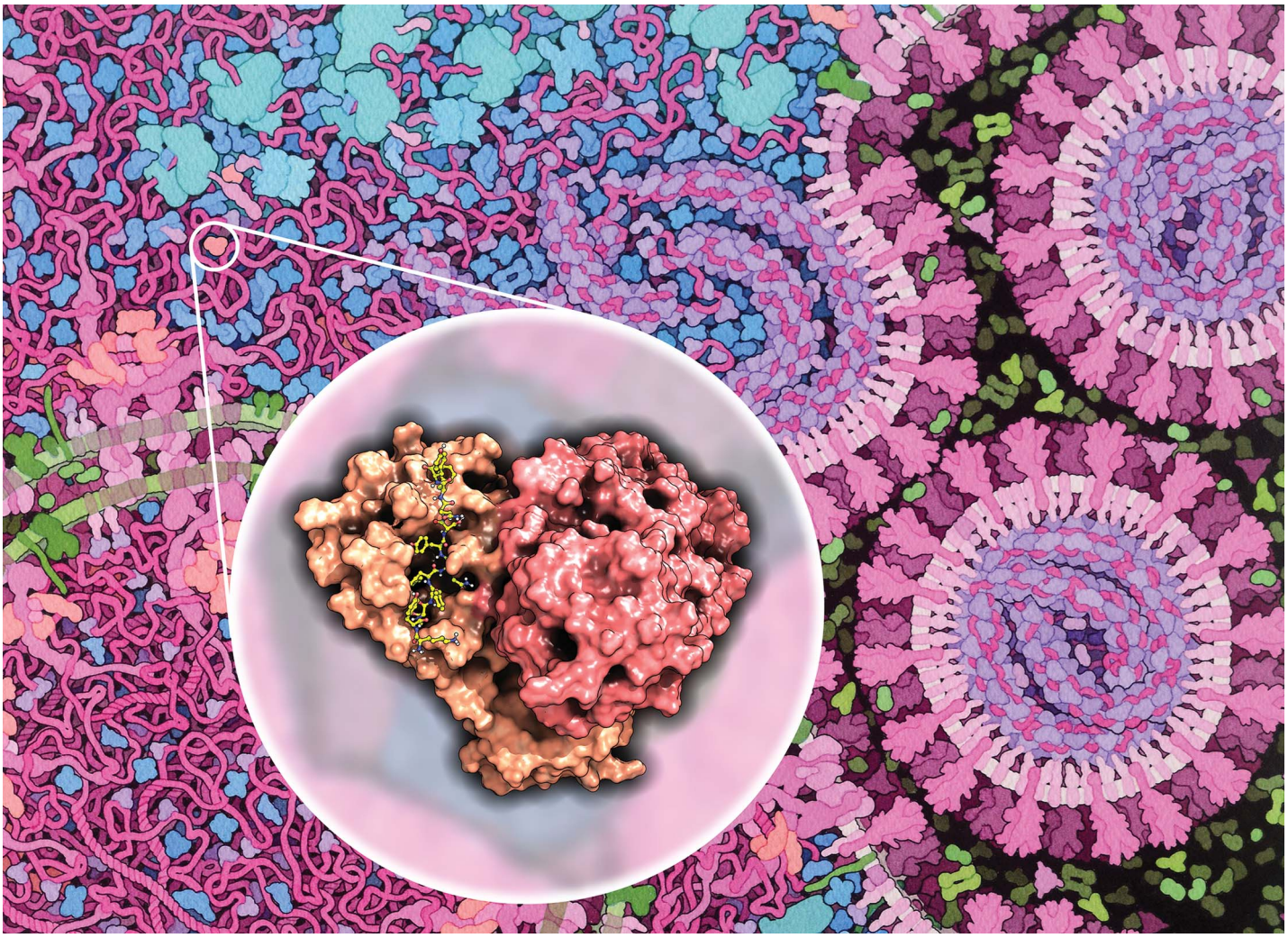
Detailed Insights into the Inhibitory Mechanism of New Ebselen Derivatives against Main Protease (M^{pro}) of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)
Pritiranjan Sahoo, Atul Kumar, *et al.*
DECEMBER 23, 2022
ACS PHARMACOLOGY & TRANSLATIONAL SCIENCE
READ

Dual Inhibitors of Main Protease (M^{Pro}) and Cathepsin L as Potent Antivirals against SARS-CoV2
Santanu Mondal, Paul R. Thompson, *et al.*
NOVEMBER 10, 2022
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
READ

Discovery and Crystallographic Studies of Nonpeptidic Piperazine Derivatives as Covalent SARS-CoV-2 Main Protease Inhibitors
Shenghua Gao, Xinyong Liu, *et al.*
DECEMBER 07, 2022
JOURNAL OF MEDICINAL CHEMISTRY
READ

Get More Suggestions >

https://doi.org/10.1021/acs.jmedchem.1c02214
J. Med. Chem. 2022, 65, 7682−7696



Showcasing research from a collaboration between the laboratories of Prof. Lodola (Parma, Italy); Prof. Moliner (Castello, Spain); Dr Genovese (Grenoble, France); Prof. Glowacki, Prof. Spencer, Dr Shoemark, and Prof. Mulholland (University of Bristol, U.K.); Prof. Walsh (Diamond Light Source, U.K.); and Prof. Schofield, Prof. Duarte, and Prof. Morris (University of Oxford, U.K.).

Discovery of SARS-CoV-2 M^{pro} peptide inhibitors from modelling substrate and ligand binding

An array of molecular simulation techniques, combined with experiments and analysis of X-ray crystal structures, reveal a detailed picture of the essential interactions and dynamics of the SARS-CoV-2 main protease (M^{pro}), showing how natural substrate peptides and small molecules bind. This underpinned the design and testing of novel peptides that competitively inhibit M^{pro}, giving insights for the development of targeted COVID-19 therapeutics. Image credit: inset: Garrett M. Morris; background: David S. Goodsell.

As featured in:



See Adrian J. Mulholland, Garrett M. Morris *et al.*, *Chem. Sci.*, 2021, **12**, 13686.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

Registered charity number: 207890

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2021, **12**, 13686

All publication charges for this article have been paid for by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Discovery of SARS-CoV-2 M^{pro} peptide inhibitors from modelling substrate and ligand binding†

H. T. Henry Chan, ‡[a] Marc A. Moesser, ‡[b] Rebecca K. Walters, ‡[cd] Tika R. Malla, ‡[a] Rebecca M. Twidale, [c] Tobias John, [a] Helen M. Deeks, [cd] Tristan Johnston-Wood, [a] Victor Mikhailov, [a] Richard B. Sessions, [e] William Dawson, [f] Eidarus Salah, [a] Petra Lukacik, [gh] Claire Strain-Damerell, [gh] C. David Owen, [gh] Takahito Nakajima, [f] Katarzyna Świderek, [i] Alessio Lodola, [j] Vicent Moliner, [i] David R. Glowacki, [d] James Spencer, [d] Martin A. Walsh, [gh] Christopher J. Schofield, *[a] Luigi Genovese, *[k] Deborah K. Shoemark, *[e] Adrian J. Mulholland, *[c] Fernanda Duarte *[a] and Garrett M. Morris *[b]

The main protease (M^{pro}) of SARS-CoV-2 is central to viral maturation and is a promising drug target, but little is known about structural aspects of how it binds to its 11 natural cleavage sites. We used biophysical and crystallographic data and an array of biomolecular simulation techniques, including automated docking, molecular dynamics (MD) and interactive MD in virtual reality, QM/MM, and linear-scaling DFT, to investigate the molecular features underlying recognition of the natural M^{pro} substrates. We extensively analysed the subsite interactions of modelled 11-residue cleavage site peptides, crystallographic ligands, and docked COVID Moonshot-designed covalent inhibitors. Our modelling studies reveal remarkable consistency in the hydrogen bonding patterns of the natural M^{pro} substrates, particularly on the N-terminal side of the scissile bond. They highlight the critical role of interactions beyond the immediate active site in recognition and catalysis, in particular plasticity at the S2 site. Building on our initial M^{pro}-substrate models, we used predictive saturation variation scanning (PreSaVS) to design peptides with improved affinity. Non-denaturing mass spectrometry and other biophysical analyses confirm these new and effective 'peptibitors' inhibit M^{pro} competitively. Our combined results provide new insights and highlight opportunities for the development of M^{pro} inhibitors as anti-COVID-19 drugs.

Received 3rd July 2021
Accepted 5th September 2021

DOI: 10.1039/d1sc03628a

rsc.li/chemical-science

1. Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the etiological agent of coronavirus disease 2019 (COVID-19) that caused the World Health Organization to declare a global pandemic in March 2020. At the time of writing, >210 million COVID-19 cases and >4.4 million deaths have been reported worldwide.[1] A key step in maturation of SARS-CoV-2, a single-stranded positive-sense RNA virus, is hydrolysis of its polyproteins pp1a and pp1ab. Most of the cleavage events—at 11 sites—are performed by the SARS-CoV-2 main protease

[a]*Chemistry Research Laboratory, Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, 12 Mansfield Road, Oxford, OX1 3TA, UK. E-mail: fernanda.duartegonzalez@chem.ox.ac.uk; christopher.schofield@chem.ox.ac.uk*

[b]*Department of Statistics, University of Oxford, 24-29 St Giles', Oxford, OX1 3LB, UK. E-mail: garrett.morris@stats.ox.ac.uk*

[c]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK. E-mail: adrian.mulholland@bristol.ac.uk*

[d]*Intangible Realities Laboratory, School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK*

[e]*School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol, BS8 1TD, UK. E-mail: deb.shoemark@bristol.ac.uk*

[f]*RIKEN Center for Computational Science, 7-1-26 Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan*

[g]*Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK*

[h]*Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot, OX11 0FA, UK*

[i]*Biocomp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castello, Spain*

[j]*Food and Drug Department, University of Parma, Parco Area delle Scienze, 27/A, 43124 Parma, Italy*

[k]*Univ. Grenoble Alpes, CEA, IRIG-MEM-L_Sim, 38000 Grenoble, France. E-mail: luigi.genovese@cea.fr*

† Electronic supplementary information (ESI) available: *via* GitHub https://github.com/gmm/SARS-CoV-2-Modelling. See DOI: 10.1039/d1sc03628a

‡ First author equal.

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

(M^{pro}; 3 chymotrypsin-like or 3CL proteinase, 3C-like protease, $3CL^{pro}$; or non-structural protein 5, Nsp5).

M^{pro} is a nucleophilic cysteine protease, which in solution is predominantly homodimeric. Each protomer consists of three domains, and the active site contains a cysteine–histidine catalytic dyad, Cys-145 and His-41, located near the dimer interface.[2] SARS-CoV-2 M^{pro} is 96% identical to the M^{pro} of SARS-CoV, which causes SARS.[3] Dimerisation of M^{pro} is proposed as a prerequisite for catalysis: the N-terminus of one protomer contributes part of the active site of the other.[4] Indeed, the monomeric form of SARS-CoV M^{pro} is reported to be inactive.[5] Evidence from non-denaturing mass spectrometry (MS)-based assays indicates that M^{pro} monomers are not only inactive (at least with tested substrates), but do not bind 11-mer substrates with high affinity.[6]

SARS-CoV-2 M^{pro} and SARS-CoV M^{pro} have similar substrate specificities, both recognizing the motif: [P4:Small] [P3:X] [P2:Leu/Phe/Val/Met] [P1:Gln]↓[P1′:Gly/Ala/Ser/Asn], "Small" denoting Ala, Val, Pro or Thr; "X" any residue; and "↓" the scissile amide (Fig. 1).[7,8] In part, because such sequences are not known to be recognised by any human protease, M^{pro} represents an attractive drug target.[4] Although no clinically approved M^{pro} drugs are available, small molecule inhibitors and peptidomimetics have been designed to inhibit SARS-CoV M^{pro} and, more recently, SARS-CoV-2 M^{pro}.[11,12] Indeed, a covalent M^{pro} inhibitor from Pfizer has recently entered clinical trials.[13,14]

Multiple crystallographic and computational modelling studies concerning the M^{pro} mechanism[15–18] and inhibition are available;[19–24] the CORD-19 database[25] documents many such studies. It is proposed that during M^{pro} catalysis, His-41 deprotonates the Cys-145 thiol, which reacts with the carbonyl of the scissile amide to give an acyl-enzyme intermediate. This intermediate is stabilised by a hydrogen bond network that holds the scissile amide carbonyl in an 'oxyanion hole'. The C-terminal part of the product likely leaves the active site at this stage. The acyl-enzyme intermediate is subsequently hydrolysed with loss of the N-terminal product regenerating active M^{pro}. Computational and mechanistic studies on SARS-CoV M^{pro}[26–29] and SARS-CoV-2 M^{pro}[30–32] suggest that in the resting state His-41 and Cys-145 are likely neutral and that the protonation states of nearby histidines (*e.g.* His-163, 164, and 172) affect the structure of the catalytic machinery—although it has been suggested in SARS-CoV M^{pro} that the protonation state of the catalytic dyad may change in the presence of an inhibitor or substrate.[33] A different picture has been obtained from neutron crystallographic studies, which indicate that an ion pair form of the dyad is favoured at pH 6.6.[34] While neutron crystallography, in principle, enables the direct determination of hydrogen atom positions, questions remain about how pH and the presence of active site-bound ligands influence the precise—and likely dynamic—protonation state(s) of the dyad.

Important questions remain regarding M^{pro} catalysis, including to what extent the active site protonation state, solvent accessibility, induced fit, and substrate sequence influence activity. The lack of this knowledge makes it difficult to carry out effective computational studies on M^{pro} catalysis and inhibition.

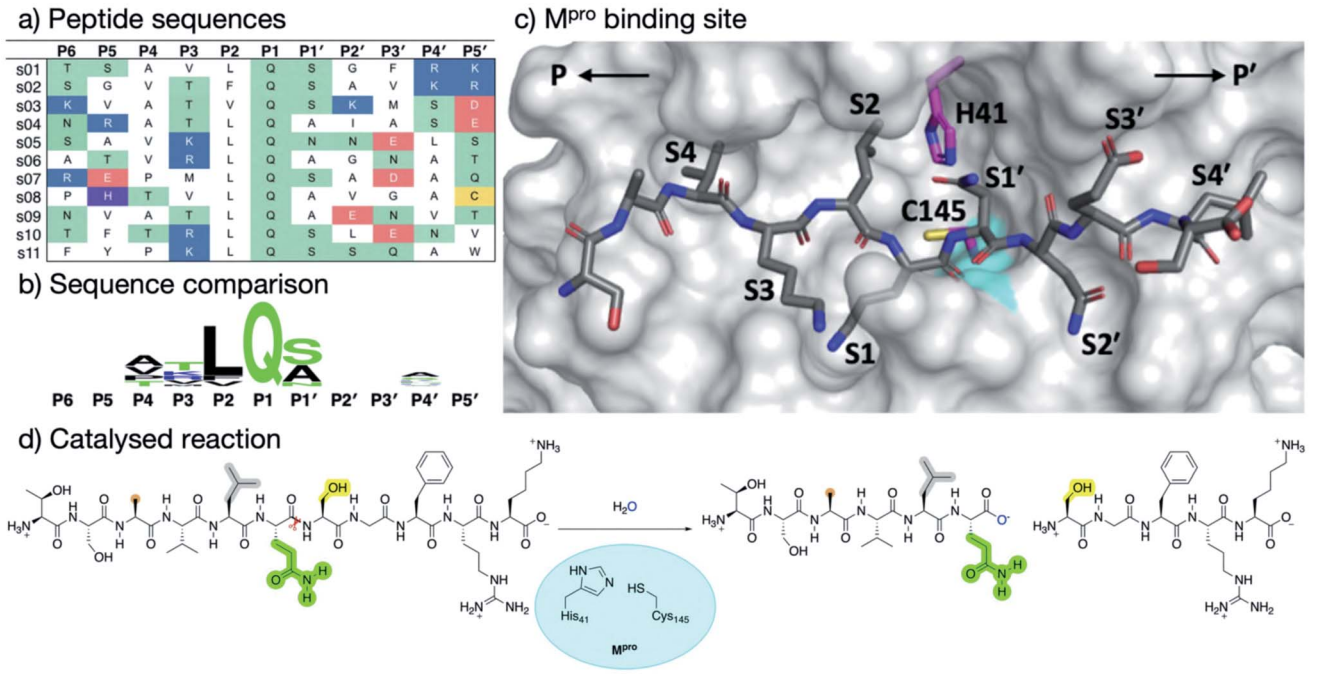


Fig. 1 Substrates processed by SARS-CoV-2 M^{pro}. (a) The 11 SARS-CoV-2 M^{pro} cleavage sites, and the corresponding 11-residue peptides, s01–s11; positively/negatively charged residues are blue/red, respectively; histidine is purple; residues with polar sidechains are green; and cysteine is yellow. (b) Comparison between the 11 substrate sequences (generated by WebLogo)[9] highlighting the completely conserved Gln at P1 and the highly conserved Leu at P2. (c) View of an energy minimised model, built using *apo* M^{pro} (PDB: 6yb7, light grey surface),[10] of M^{pro} complexed with s05 (dark grey sticks); subsites S4–S4′ are labelled. The oxyanion hole formed by the M^{pro} backbone NHs of Gly-143, Ser-144 and Cys-145 is cyan. (d) The reaction catalysed by M^{pro} exemplified by s01. Substrate residues important in recognition (see main text) are highlighted.

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

With the aim of helping to combat COVID-19, in April 2020 we embarked on a collaborative effort involving weekly virtual meetings, initially to investigate the relationship between M^{pro} substrate selectivity and activity. We employed an array of classical molecular mechanics (MM) and quantum mechanical (QM) techniques, including non-covalent and covalent automated docking, molecular dynamics (MD) simulations, density functional theory (DFT), combined quantum mechanics/molecular mechanics (QM/MM) modelling, and interactive MD in virtual reality (iMD-VR). Our results provide consensus atomic-level insights into the interactions of M^{pro} with 11-residue peptides derived from the 11 natural cleavage sites (named "s01" to "s11", in order of occurrence in the viral polyprotein, Fig. 1a). The identification of key interactions between M^{pro} and its substrates, together with analysis of fragment/inhibitor structures,[35,36] led to the design of peptides proposed to bind more tightly than the natural substrates, several of which inhibit M^{pro}. The results are freely available *via* GitHub (https://github.com/gmm/SARS-CoV-2-Modelling).

2. Results and discussion: understanding substrate binding and recognition

2.1 Protonation state of the catalytic dyad

The protonation state of the dyad following substrate binding was studied with s01-bound (Fig. 1a) M^{pro}[37] using QM/MM umbrella sampling simulations at the DFTB3/MM and ωB97X-D/6-31G(d)/MM levels of theory (Section S1.1†). The protonation states of nearby histidines were also evaluated (Table S2.1 and Fig. S2.5, S2.6†).

His-41 was treated as Nδ-protonated in the neutral state of the dyad, as reported by Pavlova *et al.* to be preferred for both uncomplexed and N3 inhibitor-bound M^{pro} based on MD studies.[30] The protonation state of the dyad-neighbouring His-163, which interacts with Tyr-161, Phe-140 and the substrate P1 Gln sidechain, was also studied. Three His-163 protonation states were considered: (i) Nδ-protonated, neutral ("HID"); (ii) Nε-protonated, neutral ("HIE"); and (iii) Nδ and Nε-protonated, positively charged ("HIP").[38] For all three His-163 protonation states, the forward trajectories (neutral, N to ion pair, IP) showed the anionic Cys-145 form to be stabilised solely by interaction with positively charged His-41 (Fig. 2a and S2.2A–C†), and the thiolate in an unreactive conformation for nucleophilic attack. This suggests that such a zwitterionic state is transient, with concerted proton transfer and simultaneous nucleophilic attack of the thiolate onto the scissile amide carbon being more likely than a stepwise mechanism.[15] By contrast, the backwards PT trajectories (from IP to N) showed stabilisation of the Cys-145 thiolate by His-41 and the P1 Gln backbone N–H, in the case of HID-163 and HIP-163 (Fig. S2.2E, F†). For HIE-163, additional stabilisation comes from interactions with the backbone N–H and the hydroxyl of the P1′ Ser, and an additional water that diffuses into the active site (Fig. 2b and S2.2D†).

The zwitterionic state with HID-163 was less stable than the neutral state and the zwitterionic states with HIE-163 and HIP-163 (Fig. S2.1C†). This was due to perturbation of the interaction network with Tyr-161 and Phe-140, suggesting that a Nδ-protonated His-163 is unlikely. Double protonation of His-163 results in a loss of both interactions in the forwards and backwards PT trajectories. Despite both HIP-163 and HIE-163 giving similar PT free energy profiles, the loss of these interactions suggests HIP-163 is unfavourable for productive catalysis. These QM/MM results therefore suggest that an Nε-protonated neutral His-163 is most likely. Along with conserving interactions with Tyr-161 and Phe-140, an Nε-protonated His-163 also formed a hydrogen bond with the P1 Gln side chain (Fig. 2a and b), an interaction not observed in PT trajectories with HID-163 and HIP-163.

Considering that DFTB3 overestimates the proton affinity of methylimidazole, it is expected that this method will overstabilise the zwitterionic state relative to the neutral state.[39] To account for this, the backwards PT reaction with a Nδ-protonated His-163 was modelled at the ωB97X-D/6-31G(d)/MM level of theory. This showed the zwitterionic state was 24.3 kJ mol^{-1} above the neutral state, an increase of 26.4 kJ mol^{-1} compared to DFTB3/MM (Fig. S2.3†). Applying the free energy difference between ωB97X-D/6-31G(d) and DFTB3 at each reaction coordinate value as a correction to the combined QM/MM free energy profile in the case of HIE-163, the neutral catalytic dyad is preferred, with the ion pair being 28.5 kJ mol^{-1} higher in energy than the neutral state (Fig. 2c). Similar results were obtained with a different QM approach (Fig. S2.4 and associated ESI Movie†).

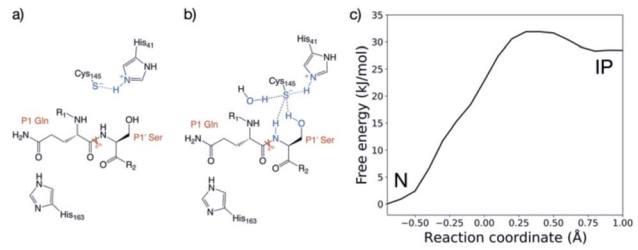


Fig. 2 QM/MM umbrella sampling of the proton transfer in the catalytic dyad. Representation of the interactions of the Cys-145 thiolate in the ion pair (IP) state from the (a) forwards and (b) backwards simulations. (c) Free energy profile for interconversion of the neutral (N) and ion pair states of the dyad in the HIE-163 system, from the combined forwards and backwards QM/MM umbrella sampling MD simulations, corrected to the ωB97X-D/6-31G(d)/MM level of theory.

2.2 Models of SARS-CoV-2 M^{pro}–substrate peptide complexes

To understand substrate specificity and to assess their relative binding affinities, we constructed 11 models of SARS-CoV-2 M^{pro} complexed with its cleavage site-derived substrates[40] as 11-amino acid peptides, from P6 to P5′ (Section S1.2; Fig. 1a and S2.7†). We refer to these peptides as 'substrates' as their hydrolytic sites are all cleaved by M^{pro} (*vide infra*). The substrates were modelled in crystallographic chain A of the M^{pro} dimer (PDB: 6yb7)[10] with a neutral dyad; unless otherwise stated, all M^{pro} residue numbers and names in the following discussions refer to chain A. Initial

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

models were subjected to three independent explicit-solvent MD simulations each of 200 ns.

Three of the 11 cleavage site-derived peptides (s01, s02, and s05) were also modelled with interactive MD using virtual reality (iMD-VR), as an alternative to comparative modelling and traditional MD. iMD-VR provides an immersive 3D environment for users to interact with physically rigorous MD simulations.[41–43] The three substrates were chosen because s01 and s02 have the highest relative efficiencies (of SARS-CoV M^{pro}) of all substrates; while s05 has the second-lowest catalytic efficiency but the same P2 and P1 residues as s01.[44]

Throughout both the explicit-solvent MD and implicit-solvent iMD-VR simulations, all the substrates remained tightly bound in the active site (Fig. S2.8–S2.11†). Substrate backbone stability was maintained especially in the central region, with only the N- and C-terminal residues showing substantial flexibility. Local sidechain fluctuations were present, notably at the solvent-facing P3 residue (Fig. 1c and S2.9†). C-terminal P′-residues consistently fluctuated more than the N-terminal P-side residues, likely in part because of fewer protein–substrate hydrogen bonds on the P′ side (*vide infra*).

2.2.1 Conserved hydrogen bond interactions. Crystallographic studies on SARS-CoV M^{pro} revealed the importance of hydrogen bonds in binding of substrate s01,[37] which has the same cleavage site sequence in SARS-CoV-2 (Fig. 1a). To investigate whether this was true for SARS-CoV-2 M^{pro} complexed with its 11 substrates, we analysed the hydrogen bond (HB) prevalence at each subsite in both explicit-solvent MD and implicit-solvent iMD-VR simulations. Twelve HBs were consistently identified (Fig. 3) and their distance and angular distributions analysed (Fig. S2.12†).

In both explicit-solvent MD and iMD-VR simulations, all 11 substrates are primarily held in place by four consistently formed backbone–backbone HBs: Glu-166 at S3 (2 & 3) and Thr-26 at S2′ (10 & 11; Fig. 3). Backbone HBs 1 and 12 further from the cleavage site show greater variation between the MD and iMD-VR studies. Although P1 Gln is conserved in all 11 M^{pro} cleavage sites (Fig. 1a and b), HBs 5–9 formed in the S1 site are observed less often than 2, 3, 10 and 11, but outnumber other sites. In both types of simulations, HB 8 from the Cys-145 backbone amide is consistently formed, suggesting that this could play a fundamental role in catalysis. HB 8 forms part of the oxyanion hole, which stabilises the tetrahedral intermediate formed upon nucleophilic attack by Cys-145 Sγ on the scissile amide. M^{pro}'s exquisite specificity for Gln at P1 is likely due to formation of HB 6 with His-163, and to a lesser extent HB 7,

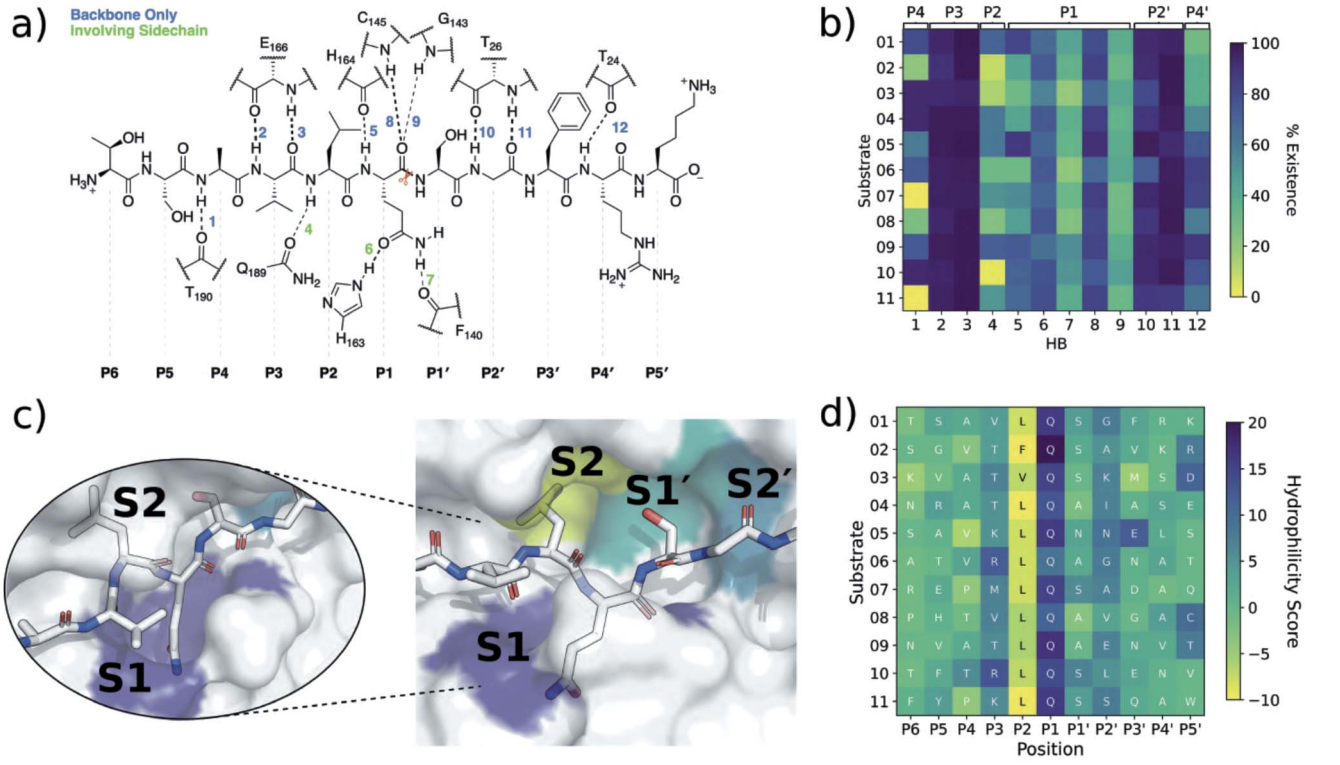


Fig. 3 Interactions between SARS-CoV-2 M^{pro} and its substrates. (a) M^{pro}–substrate hydrogen bonds (HBs) exemplified by substrate s01. (b) An annotated heat-map showing the frequency of each HB, with blue indicating highest frequency. Frames were extracted every ns from 600 ns of cumulative explicit-solvent MD conducted per system. (c) Close-up of the MD-generated binding mode of SARS-CoV-2 M^{pro}-substrate s01 with subsites S1, S2, S1′ and S2′ labelled. Different views of the S1 subsite are shown, emphasizing the deep S1 pocket that accommodates the P1 Gln sidechain. Subsite surface colour corresponds to the hydrophilicity score, with hydrophobic subsites shown in yellow, hydrophilic subsites in dark blue, and amphiphilic subsites in turquoise. (d) Hydrophilicity map for the 11 substrates calculated as the sum of hydrophilic interactions (dark blue) subtracted from the sum of hydrophobic interactions (yellow). Interactions were identified from substrate:M^{pro} frames using Arpeggio[45] (Methods Section S1.5†).

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

along with the narrowness of the S1 pocket accommodating the Gln sidechain in an extended conformation.

2.2.2 Hydrophobicity analysis. Hydrophilicity analysis shows that the S1 subsite is substantially hydrophilic in all 11 substrate–M^{pro} complexes, while S2 is consistently hydrophobic (Fig. 3d). In accord with the MD-HB analysis (Fig. 3b), the conservation of subsite interactions decreases further from the cleavage site. Although S3 and S2′ show a slight bias towards hydrophilic interactions, none of the other subsites show a consistent pattern.

2.2.3 Other non-covalent interactions. Beyond HBs, several other interaction types are conserved across the substrates; these were identified by running Arpeggio[45] on snapshots extracted from the explicit-solvent MD simulations (Fig. 4). Six of the eight most common P1 interactions are present in most (*i.e.*, ≥9/11) substrates. This includes the previously described HBs between the P1 backbone oxygen and the backbone NHs of Cys-145 (HB 8) and Gly-143 (HB 9) that constitute the oxyanion hole, as well as HB 6 (His-163) and HB 7 (Phe-140) that stabilise the P1 Gln sidechain. In addition, interactions with Ser-144, His-163, His-164, Glu-166 and, to a lesser extent, Phe-140 were prevalent at P1. None of the other subsites show this level of consistency in residue-level contacts, although some interactions such as hydrophobic contacts with Met-49 and Met-165 were always present at P2. Furthermore, important stabilising backbone HBs (HBs 2–3 between Glu-166 and P3; and HBs 10–11 between Thr-26 and P2′) were conserved in all substrates. Finally, P′ interactions are less common than those on the P side (Fig. 4). The same trend was found when docking s01, s02, and s05 using iMD-VR, where P′ residues tended to be more flexible than P-side residues.

We further analysed the energetic contributions of each M^{pro} residue using the Molecular Mechanics-Generalised Born Surface Area (MM-GBSA) method (Section S2.3.1†),[46–49] which highlighted hotspot residues that were also recognised by Arpeggio as conserved contacts.

2.2.4 Density functional theory analysis of the interaction network. We performed linear-scaling DFT (BigDFT[50]) calculations using representative snapshots extracted from explicitly solvated MD trajectories of M^{pro} complexed with s01, s02, and s05. By automatically decomposing large molecular systems into coarse-grained subsets of atoms (or 'fragments') in an unbiased manner,[51] quantities like inter-fragment bond order and interaction strengths, E_{cont}, can be easily calculated (Fig. 5a, S2.18–S2.21 and Section S2.4†). This analysis supports the essential roles of Glu-166 and Thr-26, with interactions observed in all three peptides s01, s02, and s05, consistent with the HB analysis described earlier (Fig. 3). Gln-189 consistently hydrogen-bonds with P2 (HB-4) in s01 and s05, but rarely in s02. This weakening of HB-4 in P2 may be due the greater bulk of Phe in s02 (Fig. 5b).

Conserved contacts are present in the three substrates between Cys-145 and both P1 and P1′ residues. Interactions between His-41 and P2/P1′ are observed for s01 and s05, and

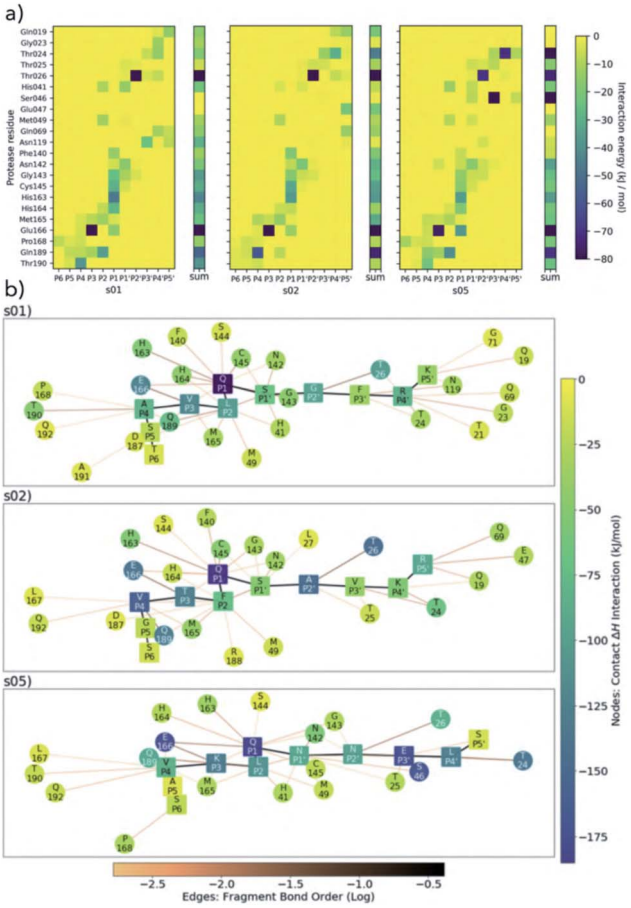


Fig. 5 BigDFT analysis of M^{pro}–substrate interactions. (a) Heatmaps showing QM interaction energies between 22 selected residues of M^{pro} and s01, s02 and s05. (b) QM interaction networks where node colour indicates interaction strength, from dark blue (strongest) through green to yellow (weakest). Square nodes denote substrate, while circular ones denote M^{pro}. The thickness and colour of the edges show the fragment bond order between residues, a unitless measure associated with bond strength and analogous to bond order; black is strongest, orange is weakest.[51] Interaction energies and bond orders were computed using BigDFT and ensemble-averaged results of MD snapshots.

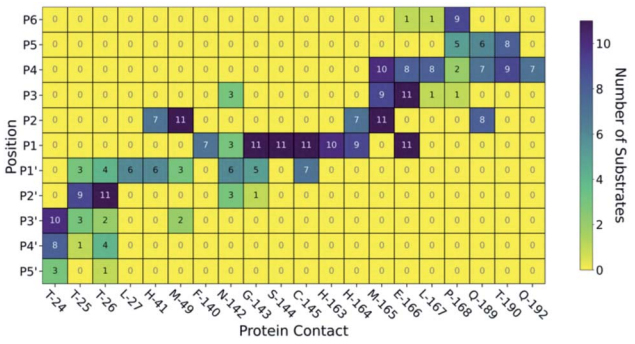


Fig. 4 M^{pro}–substrate contacts. Map of HBs and other non-covalent interactions between the 11 substrates and M^{pro} from Arpeggio analysis of the most representative pose for each substrate generated by MD. Dark blue indicates the interaction is formed by all 11 substrates at that substrate position, while yellow indicates no substrates form this interaction.

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

between Glu-166 and P1/P3 (and, to some extent, to P4) for all three substrates. This analysis singles out the character of s02, which is dominated by the bulky character of its P2 Phe. Substitutions at P2 may have a substantial effect on the interaction network close to the catalytic site. While the P side exhibits an interconnected character especially from P1 to P4 (Fig. 5b), the network on the P′ side has a more linear character, once again indicating that hotspot residues responsible for binding are present on the P side. Distributions of E_{cont} are shown in Fig. S2.20.†

The following trends emerge from our studies on M^{pro} in complex with models of its 11 substrates: (i) binding stability is partly conferred by a series of HBs from P4 to P4′, in particular between the backbones of M^{pro} Glu-166 and Thr-26 and substrate positions P3 and P2′ respectively, as well as HBs involving the conserved P1 Gln sidechain; (ii) substrate residues N-terminal of the cleavage site (P side) form more, and more consistent, contact interactions with M^{pro} compared to the P′ side, with interactions at Met-49, Gly-143, Ser-144, Cys-145, His-163, His-164, Met-165 and Glu-166 being most conserved. We conclude that the S1 and S2 pockets are prime targets for active site substrate-competing inhibitor design due to their well-defined hydrophilic character, large energy contributions to substrate binding, and vital conserved hydrogen bonds in S1 for substrate recognition.

2.2.5 Conformational plasticity in M^{pro} crystal structures. Previous studies have compared the dynamics of ligand binding sites across SARS-CoV-2, SARS-CoV and MERS-CoV M^{pro}.[52] Here, we investigated the conformational plasticity of the SARS-CoV-2 M^{pro} active site upon binding by comparing 333 M^{pro}:ligand co-crystal structures obtained from Fragalysis[53] to a reference *apo* structure of M^{pro}, PDB entry 6yb7 (ref. 10) (Fig. 6). A high degree of plasticity was observed at residues Thr-24, Thr-25, His-41, Thr-45, Ser-46, Met-49, Asn-142, Met-165, Glu-166, Arg-188, Gln-189 and Ala-191. Unsurprisingly, the S1 subsite is particularly rigid, with almost no change in residue conformations across all 333 crystal structures.

In all cases, P1 Gln recognition is mainly driven by interactions with Gly-143, Ser-144, Cys-145 (oxyanion hole) and His-163 and Phe-140. The S2 subsite, however, is highly flexible, especially at Thr-45, Ser-46 and Met-49. Although P2 is conserved in terms of hydrophobicity (Leu, Phe, Val), the S2 pocket is highly flexible and can adapt to accommodate functional groups of varying sizes, including aliphatic and aromatic groups. The outer regions of the active site (S3–S6 and S2′–S5′) vary in flexibility, echoing our MD simulations.

2.3 Monitoring of substrate sequence hydrolysis by mass spectrometry

To rank the SARS-CoV-2 M^{pro} preferences for hydrolysis of the 11 cleavage sites, we monitored turnover of 11-mer peptides by solid-phase extraction (SPE) coupled to mass spectrometry (MS). Interestingly, after the N-terminal autocleavage site s01, s11 was the next preferred substrate for catalysis (Fig. S2.22†). Peptides s06, s02, and s10 were hydrolysed less efficiently than s11. Slow turnover was observed for s07 and s09. Evidence for low turnover of s05 was obtained after prolonged incubation with M^{pro} (9.56%) (Fig. S2.23†). Under our standard conditions, no evidence for cleavage was observed for s03, s04, and s08.

We then examined turnover under non-denaturing MS conditions using ammonium acetate buffer (Fig. 7a). Peptides s01, s06, s08, s10 and s11 evidenced fast turnover. The level of substrate ion depletion was >70% after 1 min and >90% after 6 min incubation. Peptides s02, s04 and s09 showed substrate ion depletion from 35 to 45% after 1 min incubation, >70% depletion after 6 min, and >90% depletion after 12 min. Peptides s03, s05 and s07 demonstrated slow turnover that was below 50% after 12 min incubation.

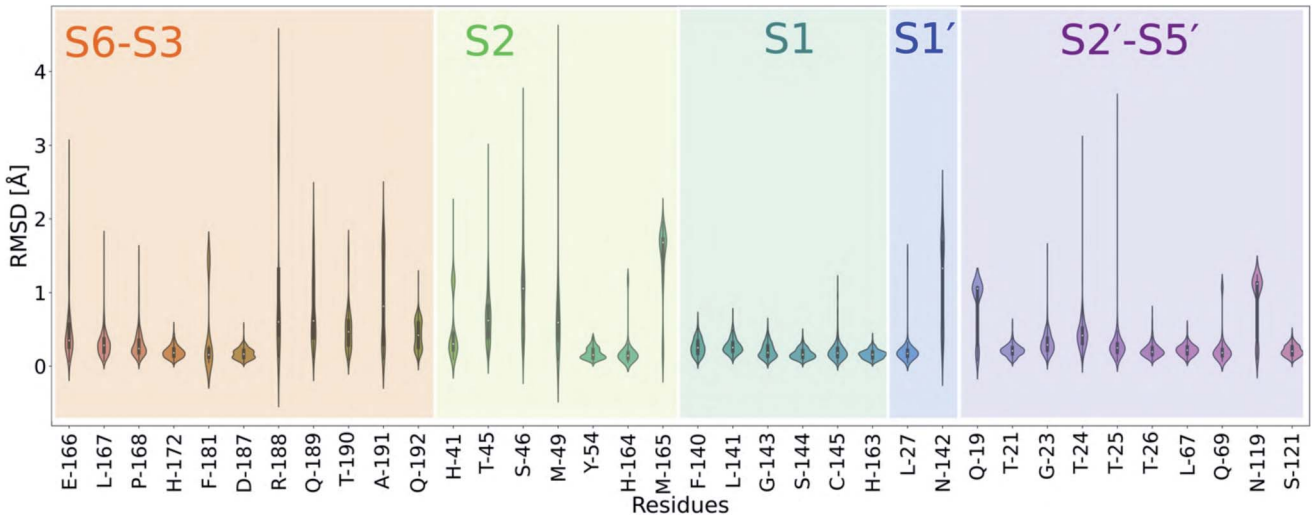


Fig. 6 Analysis of the active site plasticity of 333 M^{pro} co-crystal structures. Active site residues (residues 19, 21, 23–27, 41, 45, 46, 49, 54, 67, 69, 119, 121, 140–145, 163–168, 172, 181, 187–192) were chosen based on the MD analysis of the 11 substrate–M^{pro} models and correspond to all M^{pro} residues that contact any substrate. The violin plots show the distributions of per-residue heavy atom RMSD values between the 333 M^{pro}–ligand co-crystal structures[53] and a reference uncomplexed structure (PDB 6yb7).[10] Each M^{pro} subsite is colour-coded.

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

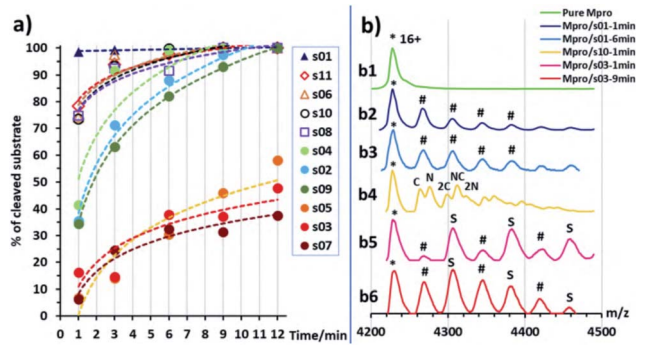


Fig. 7 Non-denaturing mass spectrometry of M^{pro} substrate turnover. (a) Substrate turnover *versus* incubation time as measured by non-denaturing MS. Trend lines are given for visual guidance only. (b) Examples of mass spectra showing normalized intensity in the *m*/*z* region around the 16+ charge state of M^{pro} dimer (asterisk, *): (b1) pure M^{pro} solution (5 μM); (b2) M^{pro} and s01 solution after 1 min incubation, hashes (#) indicate the mass peaks corresponding to the s01 cleaved fragments sequentially attached to the M^{pro} dimer; note: the resolution is not sufficient to distinguish between the N- or C-terminal fragments (mass shifts of 617 and 593 Da, respectively); (b3) same solution as (b2) after 6 min incubation; (b4) M^{pro} and s10 solution after 1 min incubation, 'N' labels N-terminal fragment(s) attached (765 Da), 'C' labels C-terminal fragment(s) attached (560 Da); (b5) M^{pro} and s03 solution after 3 min incubation, 'S' labels intact substrate(s) attached, hashes (#) label attached substrate fragments, but the N- and C-terminal fragments cannot be distinguished (mass shift 644 and 566 Da, respectively); (b6) same solution as (b5) after 9 min incubation.

In the protein region of the mass spectra, complexes of the M^{pro} dimer and the cleavage products were observed after 1 min of incubation for the fast-turnover substrates s01, s06, s08, s10 and s11, and also s02, s04 and s09 (Fig. 7b1–b4). For the slow-turnover substrates s03, s05 and s07, only M^{pro} complexes with intact substrates were observed after 1 min incubation. For longer incubation times, complexes between M^{pro} and the products from these substrates emerged and increased in abundance (Fig. 7b5 and b6).

The rank order of the substrates in part depends on the MS method used, likely due to the differences in the buffers and concentrations used: *i.e.*, non-denaturing MS used ammonium acetate buffer and an M^{pro} concentration of 5 μM, which is higher than the 0.15 μM used in denaturing MS assays. Higher concentrations of both enzyme and substrate in the non-denaturing MS experiments explain the faster substrate turnover than seen with denaturing MS, especially as the concentration of catalytically active M^{pro} dimer would be higher at higher enzyme concentrations.[6,54]

Regardless of the MS method used, a clear trend is observed in the catalytic turnover of the cleavage site-derived peptides. The rank order of substrate preference under denaturing MS conditions was s01 > s11 > s06 > s02 > s10 > s07 > s09 > s05 (Fig. S2.22 and S2.23†). Under non-denaturing conditions (Fig. 7) turnover was: fast (s01, s11, s06, s10, and s08), medium (s04, s02, and s09), and slow (s05, s03, and s07). Substrates s01, s11 and s06 turned over fastest; while s07, s05 and s03 were slow as measured by both methods. This is in broad agreement with the reversed phase high performance liquid chromatography analysis of substrate turnover by SARS-CoV M^{pro}, where s01 and s02 display fast turnover; s10, s11 and s06 manifest medium turnover; and the rest (s09, s08, s04, s03, s05, s07) show slow turnover.[44] Both of our MS studies on SARS-CoV-2 M^{pro} indicate that s02 consistently displayed slower turnover than s11. Previous reports on SARS-CoV M^{pro} have shown evidence for cooperativity between subsites during substrate binding, in particular during autocleavage of the M^{pro} C-terminal site (s02), where the Phe at P2 induces formation of the S3′ subsite to accommodate the P3′ Phe residue.[55] SARS-CoV-2 M^{pro} substrate s02 has a Phe at P2, but not at P3′ (Fig. 1a). The absence of a Phe at the P3′ position may in part explain the reduced activity of SARS-CoV-2 M^{pro} for s02 relative to s01, compared to the same pair in SARS-CoV M^{pro}.[44]

The observed turnover of all 11 SARS-CoV-2 cleavage-site-derived peptides by M^{pro} is consistent with our atomistic models, where the peptides remain bound in the active site during MD simulations and where the scissile amide carbonyl remains well-positioned in the oxyanion hole (*e.g.*, HB 8 in Fig. 3) for reaction initiation. The stability of the M^{pro}–peptide interactions involving the S2 and S1 subsites, as well as backbone–backbone HBs 2, 3, 10 and 11, could explain the observation using non-denaturing MS of complexes of M^{pro} with products—because of slow product dissociation. Nevertheless, we envisage that the order of substrate turnover rates is likely determined by various factors, including peptide conformations, the influence of the P2 and P1′ residues on the catalytic dyad (as highlighted by the BigDFT analysis), entropic effects, and rates of product dissociation, all of which prompt ongoing experimental and computational investigations.

3. *In silico* mutational analysis of substrate peptides enables peptide inhibitor design

Building on insights gained from our binding studies of SARS-CoV-2 M^{pro} and the 11 SARS-CoV-2 polypeptide substrate sequences, we designed peptides that could bind more tightly than the native substrates. We quantified the per-residue energetic contributions of these sequences to the overall binding in the M^{pro} active site and proposed substitutions that would increase affinity. We hypothesised these peptides would: (a) behave as competitive inhibitors, and (b) provide counterpoints for comparison with natural substrates, shedding light on requirements for M^{pro} binding and, perhaps, turnover.

3.1 *In silico* alanine scanning and predictive saturation variation scanning

We used the interactive web application BAlaS to perform Computational Alanine-Scanning mutagenesis (CAS) using BudeAlaScan[56] and the BUDE_SM algorithm[57] for Predictive Saturation Variation Scanning (PreSaVS).[58] Both are built on the docking algorithm BUDE,[59] which uses a semi-empirical free energy force-field to calculate binding energies.[60] To identify key binding interactions of the natural substrate peptides to M^{pro}, the 11 substrate:M^{pro} complexes were first subjected to CAS

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

using BAlaS. By sequentially substituting for alanine, the energetic contribution of each substrate residue to the overall interaction energy between the singly mutated peptide and M^{pro} is calculated using:

$$\Delta\Delta G = \Delta G_{Ala} - \Delta G_{wt}$$

where ΔG_{wt} is the interaction energy between the peptide and M^{pro}, and ΔG_{Ala} is the interaction energy for the peptide with a single alanine mutation at a given position. The more positive the value for each residue, the greater the contribution from that substrate residue to binding. This method was used later to evaluate potential inhibitor peptides.

Having identified residues contributing most to the binding energy of the natural M^{pro} substrates, each of the sequences was subjected to PreSaVS using the BUDE_SM algorithm. This sequentially substitutes each substrate residue with a range of residues (D, E, F, H, I, K, L, M, N, Q, R, S, T, V, W and Y). BUDE_SM calculates the $\Delta\Delta G = \Delta G_{wt} - \Delta G_{mut}$ for the binding interaction of each, entire, singly mutated peptide with M^{pro}. Substitutions predicted to improve binding over wildtype sequences have a positive $\Delta\Delta G$. Fig. 8 shows an example of the BUDE_SM PreSaVS results for all the P2 substitutions for the 11 substrate peptides (normally Leu, Phe, or Val in the 11 substrates). The most positive results suggest that Phe, Trp and Tyr favour increased predicted affinity at the P2 position (Fig. 8b). However, although Tyr generally increased the predicted binding affinity ($\Delta\Delta G_{sum}$ = 68.8 kJ mol^{-1}), it was not considered for substitution at P2 due to its negative effect at this position in s11 (scoring −18.9, Fig. 8a). Candidate residues for each position, from P6 to P5′, were shortlisted similarly based on those with the best total, and the fewest unfavourable, scores.

In addition to the computed $\Delta\Delta G$ values, we considered the propensity of each residue to promote an extended conformation. All bound substrates are largely extended, so entropic penalties may be avoided if inherently extended conformations could be favoured in the designed peptide. Thus, the best β-forming (and therefore least α-forming) residues from the first triage were selected (Fig. 9a).[61] We also considered solubility. This was achieved by limiting the number of hydrophobic residues in each designed peptide and ensuring a net positive charge (except p14, which was neutral).

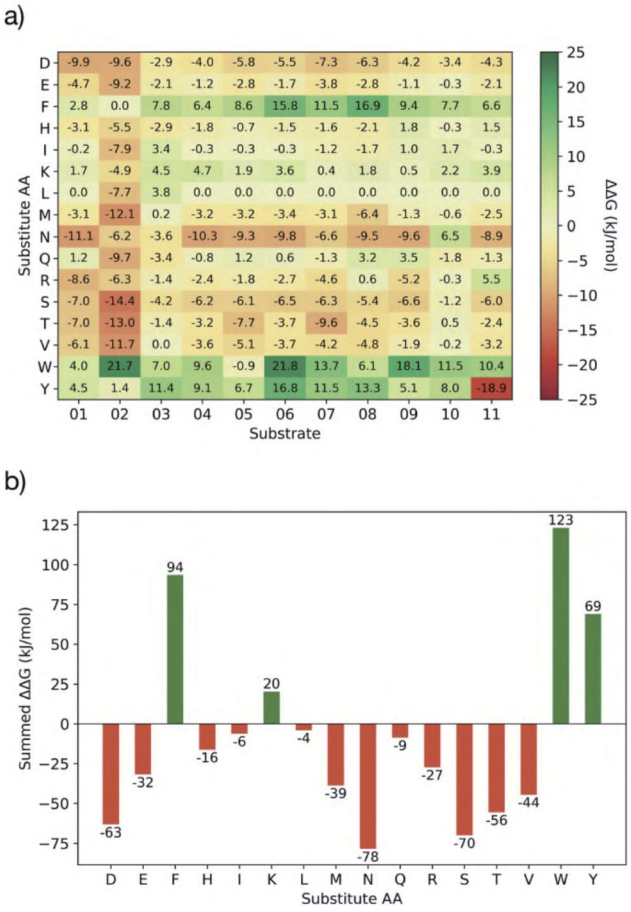


Fig. 8 BUDE_SM PreSaVS for the P2 position. (a) Heat-map for BUDE_SM PreSaVS saturation mutagenesis at P2, showing the $\Delta\Delta G = \Delta G_{wt} - \Delta G_{mut}$ value calculated for each substitution and each M^{pro} substrate. Mutations predicted to improve peptide binding have a positive $\Delta\Delta G$ and are greener; those disfavouring binding are in red. (b) The summed $\Delta\Delta G$ values for each residue type substituted at P2.

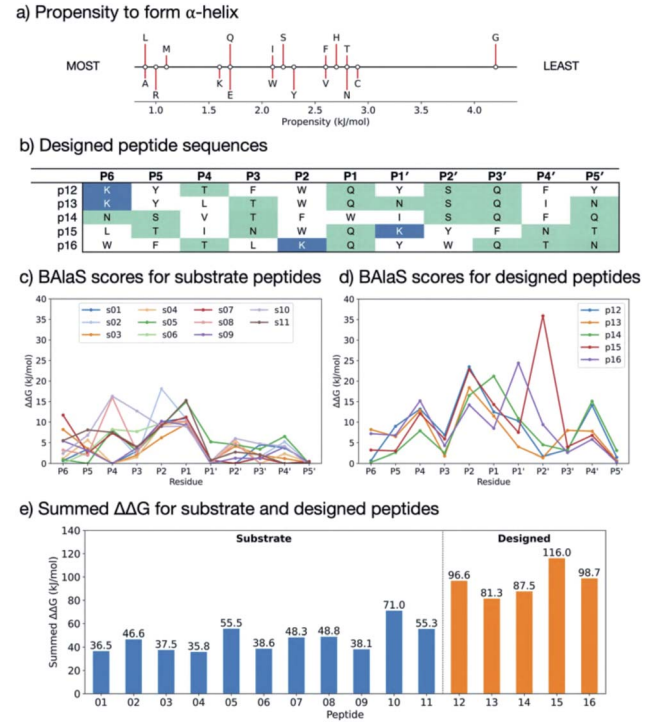


Fig. 9 BAlaS-guided design of tight-binding peptides. (a) Propensity scale of each amino acid to form an α-helical peptide conformation. (b) Sequences of designed peptides p12–p16. Scatter plots with predicted BAlaS $\Delta\Delta G = \Delta G_{Ala} - \Delta G_{wt}$ values on substitution to alanine for each residue of (c) the 11 M^{pro} natural substrates and (d) designed peptides based on these. The more positive the value, the greater the contribution made by the sidechain to the overall binding energy. (e) The BAlaS $\Delta\Delta G_{Sum}$ comparing values between complexes of M^{pro} with substrate and designed peptides as a proxy for predicting relative binding affinity (larger score = tighter binder).

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

3.2 Designed peptide sequences

Employing the criteria described above, five new peptides, p12–p16, were designed (Fig. 9b). Comparison of the computed $\Delta\Delta G$ values for s01–s11 (Fig. 9c) and p12–p16 (Fig. 9d) reveals that substitutions at the P sites provide only occasional, moderate improvements to binding energy over the corresponding substrate P sites, with the notable exception of P2, which can accommodate Trp, Phe or Lys. These results agree with the HB analysis, which predicts that the sidechains of residues that are on the N-terminal side of the cleavage site (P sites) contribute more to binding than C-terminal, P′ sites. The most striking difference between substrates and designed peptides is in this P′ region, where the predicted binding energy contributions for the designed peptides exceed those of the substrates, an advantage that is distributed over most of the designed P′ positions.

The final step in design was to assess the relative binding affinities of the substrates and designed peptides. Hence the summed $\Delta\Delta G$s (Fig. 9e) provide a proxy for the binding energies (BAlaS)[62] for the substrates and designed peptides with M^{pro}. The substrate:M^{pro} complexes are stabilised by an average of 46.5 kJ mol^{-1}, whereas the designed-peptide:M^{pro} complexes are predicted to have, in some cases, double the interaction stability of the substrates, with an average of 96.0 kJ mol^{-1}. The full analysis is in the ESI file SI_BAlaS_BUDE_SM_12-04-2021.xlsx.†

3.3 Synthesis and analysis of designed peptides

To test the designed sequences, p12, p13, p15 and p16 were synthesised with a carboxyl-amide C-terminus by solid phase synthesis. Their M^{pro} inhibitory activity was determined by dose–response analysis (Table 1) using SPE MS, monitoring both substrate s01 (1191.68 Da) depletion and N-terminally cleaved product (617 Da) formation. Ebselen which reacts multiple times with M^{pro}[63] was used as a standard ($IC_{50} = 0.14 \pm 0.04$ μM; Fig. 10).

All four designed peptides manifested similar potency with IC_{50} values ranging from 3.11 μM to 5.36 μM (Table 1 and Fig. S3.1†). Strikingly, despite the presence of Gln at P1 in all the designed peptides assayed, no evidence for hydrolysis was observed by SPE MS. This observation was supported by LCMS of the peptides incubated overnight with M^{pro} (Fig. S3.2†). We probed the inhibition mode of the designed peptides by monitoring changes in IC_{50} while varying the substrate concentration (2 μM, 10 μM, 20 μM and 40 μM TSAVLQ↓SGFRK-NH_2 s01; $K_m \sim 14.4$ μM).[63] The results indicated a linear dependency between substrate concentration and IC_{50} values (Fig. 10a–d). This was not observed with a control 15-mer peptide or ebselen (Fig. 10e and f). Analysis of the data by the procedure of Wei *et al.*[64] implies competitive inhibition (Fig. S3.3 and Tables S3.1, S3.2†). By contrast, the same analysis for ebselen did not support competitive inhibition, consistent with MS studies showing it has a complex mode of inhibition.[63]

Table 1 Designed peptides inhibit M^{pro} in a dose-dependent manner. The assay conditions were 0.15 μM M^{pro}, 2 μM s01 in 20 mM HEPES, pH 7.5, and 50 mM NaCl

Peptide	IC_{50}/μM	Hill slope
p12	5.36 ± 2.17	1.25 ± 0.06
p13	3.11 ± 1.80	0.94 ± 0.09
p15	5.31 ± 1.08	1.17 ± 0.16
p16	3.76 ± 0.51	1.19 ± 0.16

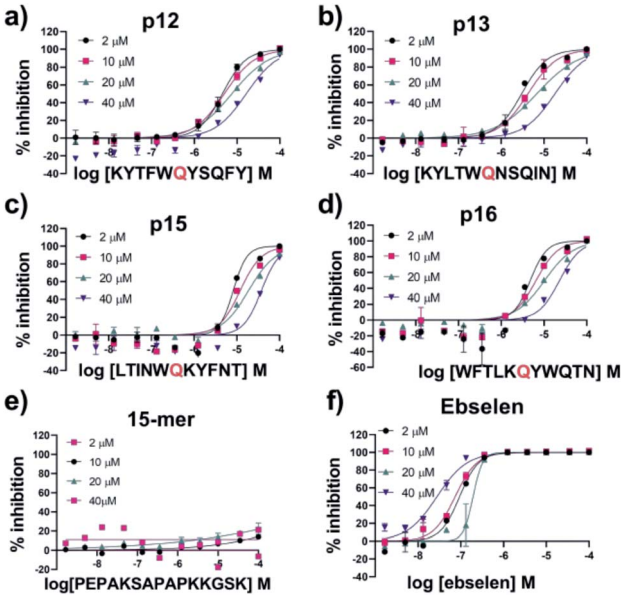


Fig. 10 IC_{50} of designed peptides against M^{pro} with varied substrate concentrations. IC_{50}s for (a) p12, (b) p13, (c) p15, (d) p16, (e) 15-mer control peptide and (f) ebselen with 2 μM, 10 μM, 20 μM and 40 μM of substrate peptide s01. IC_{50} values were calculated from technical duplicates (Table S3.2†). See Experimental Section S1.8† for assay details.

Three of the synthesised peptides—p12, p13, and p15—have a Trp at P2 (Fig. 9b) while the other, p16, has a Lys at P2. The 11 M^{pro} substrates all have hydrophobic residues (Leu, Val or Phe) at P2 (Fig. 1a). To investigate if the nature of the hydrophobic P2 residue, or the hydrophilic nature of the Gln at P1, alters the interaction of the peptide and hence its reactivity at the active site, we synthesised p13-WP2L, s01-LP2W, and s01-QP1W. There was no evidence for cleavage of p13-WP2L or s01-QP1W. However, s01-LP2W underwent partial cleavage (12.6 ± 4.5)% after overnight incubation. These results suggest that the presence of a Trp at P2 hinders catalytically productive binding, at least with these peptides, and that other residues (including the P1′ and P2′ residues) play roles in orienting the substrates for cleavage (*vide infra*).

We then used non-denaturing protein MS to study enzyme–substrate/product/inhibitor complexes simultaneously with turnover. Complexes between M^{pro} dimer and p12 and p13 were observed, together with the uncomplexed M^{pro} dimer in the protein region of the mass spectra. No binding was observed for p15 and p16, due to relatively high noise in that m/z region. None of the designed peptides were cleaved by M^{pro}, as recorded in the peptide region. As a control, s01 was

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

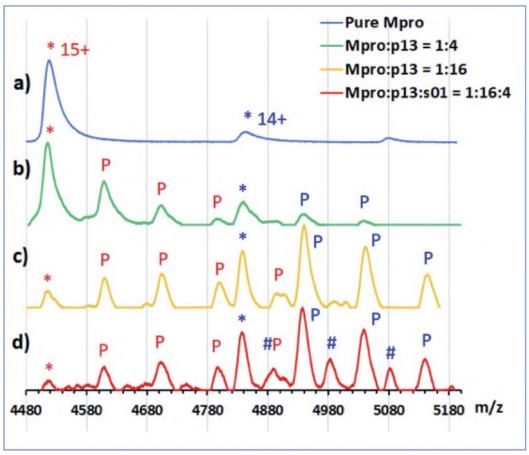


Fig. 11 Non-denaturing MS analysis of designed peptides binding to the M^{pro} dimer. Inhibitor binding from non-denaturing MS showing normalized intensity in the *m/z* region around the 14+ and 15+ charge states of the M^{pro} dimer. (*) indicates unbound M^{pro} dimer. (a) 5 μM M^{pro} solution; (b) 4-fold excess of p13 relative to the M^{pro} dimer; 'P' indicates sequential binding of p13 peptides to M^{pro} in the 15+ charge state (red) and 14+ state (blue); (c) 16-fold excess of p13; (d) 16-fold excess of p13 and 4-fold excess of s01; hash (#) indicates sequential binding of s01-cleavage products (note: the resolution is not sufficient to distinguish between the N- and C-terminal fragments; some non-specific binding of p13 is also observed in (c) and (d) due to the high concentration of the peptide).

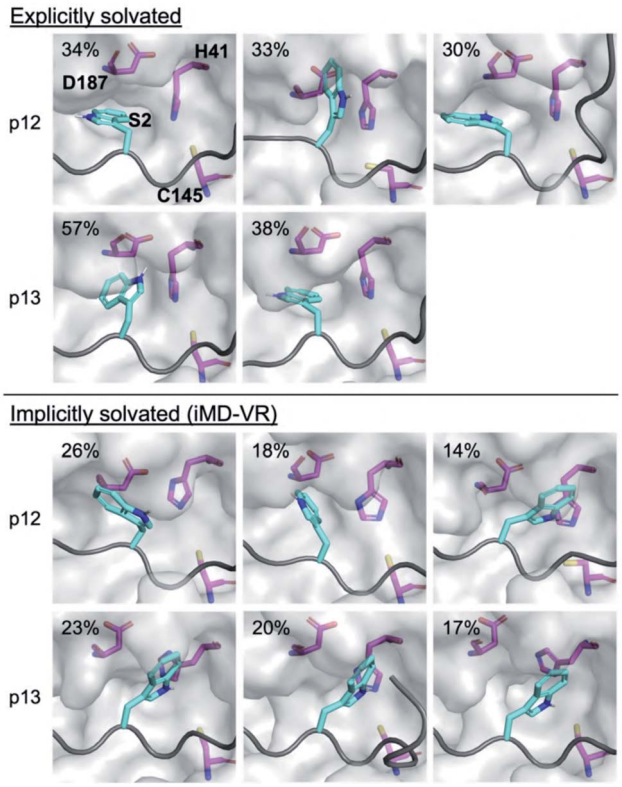


Fig. 12 Binding of P2 Trp in the designed peptides. Conformations adopted by the P2 Trp sidechain (cyan sticks; non-polar hydrogens omitted for clarity) in p12 and p13 (grey ribbon) observed during explicit and implicit solvent MD simulations, showing representative structures obtained by RMSD clustering. His-41, Cys-145 and Asp-187 are shown in magenta. See Fig. S3.13† for cluster populations formed during MD. For each peptide, conformations are displayed in decreasing order of occurrence (above 10%).

added to the protein/inhibitor mixtures; for all the inhibitors, turnover of s01 was observed after 3 min incubation. Depletion of s01 was 95%, 91%, 70% and 78% in the presence of p12, p13, p15 and p16, respectively, with an 8-fold excess of inhibitor over M^{pro}, *versus* >98% depletion for the M^{pro}/s01 mixture without the inhibitor. In the protein region of the mass spectra, complexes between M^{pro} dimers and the s01-cleavage products were observed in the presence of p13, but the abundance of these complexes was lower than the abundance of M^{pro}/p13 complexes (Fig. 11). These results validate the above-described evidence that the peptide inhibitors both bind and competitively inhibit M^{pro}.

3.4 Understanding the basis of SARS-CoV-2 M^{pro} inhibition by the designed peptides

3.4.1 Modelling of the designed peptides. Modelling of p12 and p13 shows that both bind stably at the active site during MD simulation (Fig. S3.4–S3.8†). Like the natural substrates, key HBs form with Glu-166, Thr-26, Thr-24, and the oxyanion hole-contributing Cys-145 (Fig. S3.9–S3.11†). However, HBs involving the P1-Gln sidechain of p12 and p13 showed greater variability. The favourability of the P2 Trp mutation, as predicted by the BAlaS scores, prompted us to investigate its binding. In line with the plasticity observed at S2, a variety of conformations are observed during MD simulations at this position, showing varying degrees of immersion in S2 (Fig. 12 and S3.12, S3.13†). Similar results were obtained using iMD-VR (Fig. 12).

Analysis of the conformations of the most populated cluster from MD using Arpeggio-generated hydrophilicity maps (Fig. S3.14†) reveals that the P2 Trp is more deeply buried within S2 than the native P2 residues in the natural substrates, forming more than double the number of hydrophobic contacts in the cases of p12 and p13. Some conformations involve the indole ring π–π-stacking, or hydrogen-bonding *via* its indole N–H, with the catalytic His-41 sidechain, forming an extended HB network (Fig. 12). It is possible that these interactions may hamper the ability of His-41 to deprotonate Cys-145 at the start of peptide hydrolysis, which could be tested using QM/MM

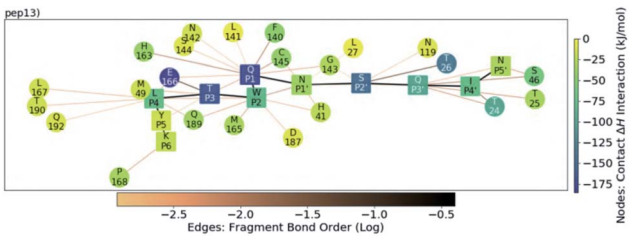


Fig. 13 QM contact interaction graph for p13 and M^{pro}. Interactions are computed using ensemble-averaged results of MD snapshots with the BigDFT code.[50]

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

calculations. Interestingly, DFT-based interaction analysis reveals that one of the slowest turnover substrates, s05 (Fig. 5), and inhibitor p13 (Fig. 13), share similar short-range interaction networks.

3.4.2 Comparative peptide docking. To investigate the ability of the M^{pro} subsites to recognise residues in the designed sequences, AutoDock CrankPep (ADCP) was used (Table S3.3†).[65] A trial was performed by redocking s01 into the H41A SARS-CoV M^{pro} structure originally complexed with s01 (PDB entry 2q6g).[37] ADCP successfully positioned the peptide mostly correctly in its top solution, with the C_α positions from P5 to P1′ deviating by less than 1 Å (Table S3.4 and Fig. S3.15†). Deviations increased up to 16 Å at P5′ as the peptide coiled up in the P′ positions, but this is deemed acceptable since the S′ subsites are less well defined, as discussed earlier.

Following the promising redocking results with ADCP, s01–s11, p12, p13, p15, and p16 were docked with an M^{pro} structure originally complexed with the N3 inhibitor (PDB entry 7bqy; 1.7 Å resolution).[2] Substrate-docked structures were found to have the P4 and P2 residues correctly positioned in their corresponding S4 and S2 pockets (Table S3.5†). From P1 to P5′ the poses were more variable, with some peptide backbones turning through S1 rather than continuing an extended conformation, likely due to the less well-defined S′ subsites (Fig. S3.16†). For the designed peptides, by contrast, docking appeared less successful (except p16), with none of the top 10 solutions positioning the peptide in the manner observed in our MD simulations (Fig. S3.17†). The S2 pocket in 7bqy binds the Leu sidechain of N3 and is probably too shallow to accommodate the larger Trp sidechain, given the assumption of a rigid receptor in ADCP docking. Hence, the four designed peptides were also docked to the C145A M^{pro} structure in complex with the s02 cleaved product (PDB entry 7joy; 2 Å resolution),[66] which has a deeper S2 pocket that binds the P2 Phe sidechain in s02. Interestingly, for both p12 and p16, the top docked solution matched our design more closely (Table S3.6 and Fig. S3.18†). Docking of p13 and p15 was challenging, possibly due to the difficulty of recognising a larger Leu (p13) or Ile (p15) residue in the S4 pocket, which originally accommodated a Val sidechain. This highlights the ability of the M^{pro} active site to adapt when binding to different substrates or inhibitors.

3.5 Summary – designed peptides

We used *in silico* Predictive Saturation Variation Scanning to design peptides that were shown *in vitro* to inhibit M^{pro} competitively. Structures of p12 and p13 generated by both iMD-VR docking and comparative modelling behaved similarly, in terms of HB formation and peptide backbone RMSD and RMSF, when performing MD. These studies highlight how the S2 subsite can adapt its size and interaction network *via* induced fit to accommodate different substrate or inhibitor P2 residues.

Notably, while these models suggested similarly stable binding modes as seen with the natural substrates, turnover of the inhibitor peptides by M^{pro} was not detected. This may relate to the more favourable predicted binding affinity of the designed peptide–M^{pro} complexes, both in terms of higher overall interaction energies, and greater contribution of the P′ residues than in the natural substrates. Our MD simulations suggest it is also possible that the larger P2 residue prevents the catalytically vital His-41 from adopting a reactive conformation (Fig. 12).

4. M^{pro}–ligand interaction analysis

Having elucidated how M^{pro} recognises its substrates and our designed peptide inhibitors, we hypothesised that this might be reflected in the extensive small-molecule inhibitor work on M^{pro} and could, in turn, be exploited for the design of novel small-molecule inhibitors and peptidomimetics. We explored whether ligands sharing the same contacts as the natural substrates could lead to better inhibitory activity. We analysed all 91 X-ray structures of small molecule fragments complexed

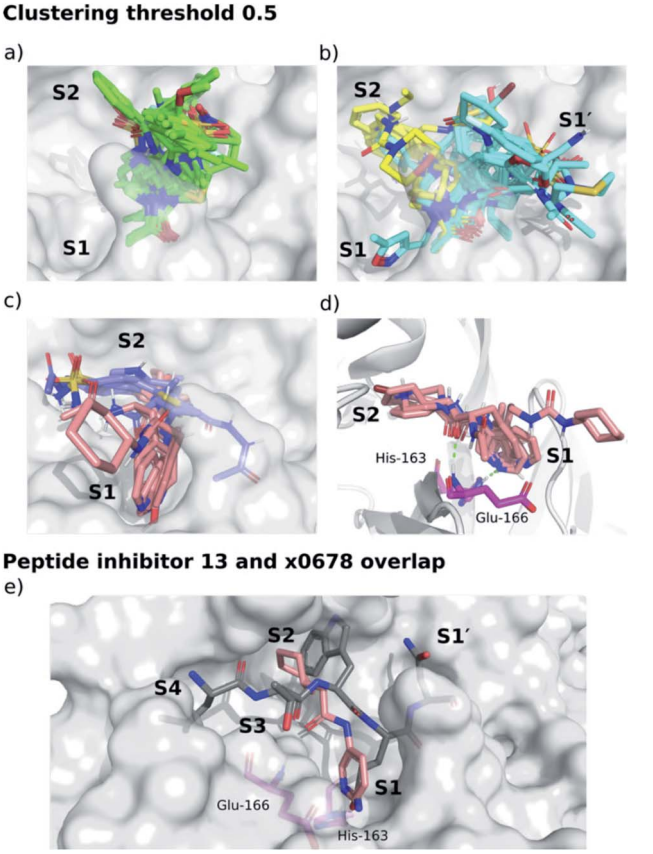


Fig. 14 Clustering of XChem active site-binding fragments. Surface of the x0830-bound M^{pro} structure (white surface) and the top 5 most populated fragment clusters using a clustering threshold of 0.5. (a) Cluster 1 fragments tend to occupy S1′ (green); (b) clusters 2 (cyan) and 3 (yellow) tend to span S1′ and S2; (c) clusters 4 (lilac) and 5 (pink) tend to occupy S2 and S1. (d) Close-up of cluster 5. Green dotted lines indicate the two key HBs between the fragment carbonyl oxygen and the backbone nitrogen of Glu-166 (HB 3, Fig. 3), and between the His-163 Nε and the heterocyclic nitrogen of the fragment (HB 6, Fig. 3). (e) Overlay of the P4–P1′-truncated structure of peptide inhibitor p13 (grey) from an MD snapshot and cluster 5 binder x0678 (pink), with the x0678 co-crystal structure (white surface).

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

with M^{pro} obtained by high-throughput crystallographic screening at Diamond's XChem facility,[35] as well as the dataset of 798 designed inhibitors and 245 crystal structures obtained from the COVID Moonshot project.[36] We analysed them by investigating their protein–ligand interaction patterns.

4.1 Interaction analysis of XChem fragments

We separated fragments into non-active-site binders (25 fragments) and active-site binding/likely-substrate competing molecules (66 fragments; Fig. S4.1†). A fingerprint bit-vector was constructed for every active-site binding fragment, with each bit denoting the presence or absence of a given interaction with M^{pro} residues, and used for clustering fragments by their interaction fingerprint Tanimoto similarity,[67] with 1 corresponding to identical contacts, and 0 to no shared contacts (Fig. 14, S4.2–S4.5 and Table S4.1†). All the fragments and ligands in clusters 1 and 2 (except x0397, x0978 and x0981) are covalently bound to Cys-145. As a result, a highly conserved binding mode is observed for the carbonyl-containing covalent warheads (*e.g.*, chloroacetamides), where the carbonyl oxygen binds into the oxyanion hole between residues Gly-143 and Cys-145, mimicking substrate HBs 8 and 9 (Fig. 3). Cluster 5 stands out as the only major cluster with fragments that bind deeply into S1, one of the main conserved contacts identified in all substrates. Cluster 5 shows a distinct binding motif primarily driven by: (i) hydrogen bonding between a carbonyl oxygen on the fragment and the Glu-166 backbone NH-group; and (ii) a strong polar interaction between His-163 and the fragment. Notably, the protonation of the imidazole of His-163 appears to depend on the fragment.

Overall, the primary functionality that facilitates interaction with His-163 is the nitrogen-containing heterocycle present in almost all ligands in cluster 5 (Fig. 15); the exception is x0967, which forms the His-163 HB *via* its phenol oxygen. Such heterocycles are well suited to replace the substrate P1 Gln sidechain by mimicking its HB donor/acceptor abilities. In addition, most cluster 5 binders also extend into the hydrophobic S2 pocket, although there is no clear preference in functional group at S2. This agrees with our plasticity analysis, which shows that S2 can accommodate a large variety of functional groups (Fig. S4.4†). As seen in the overlap of peptide inhibitor p13 and cluster 5 representative x0678 (Fig. 14e), the binding modes of both inhibitors in the S1 and S2 subsites are very similar, with both forming HBs to His-163 (HB-6) and Glu-166 (HBs 2 & 3) and binding deep in the S2 pocket. In addition, all cluster 5 ligands (Fig. 15) contain an amide or urea linker between the P1 and P2 binding groups making them interesting building blocks for the development of peptidomimetics.

The interactions between the fragments, substrates, and peptide inhibitors with M^{pro-} were analysed by employing linear scaling DFT. Using short-range (E_{cont}) DFT interactions with M^{pro} as a "descriptor" for clustering, a cluster containing both the substrates and the cluster 5 compounds (x0107, x0434, x0540, x0678, x0967 and x1093) was identified (Fig. S4.6†). This cluster also included compounds x0426, x0946, x0195, x0995, x0104, x0874, x1077, x0161 and x0397. This agnostic analysis, based on quantum mechanical descriptors, provides further confirmation and a powerful alternative to evaluate compounds of differing sizes in biomolecular complexes.

To test whether cluster 5 inhibitors are promising building blocks for optimization, we identified all assayed and crystallized cluster 5 binders in the COVID Moonshot project database[36] as of 11th Jan 2021 and analysed them using Arpeggio. Compounds were deemed cluster 5 inhibitors if they shared at least 70% of the contacts identified in fragment cluster 5. We observed that cluster 5 inhibitors have a significantly higher proportion of "strong" binders, classified as IC_{50} < 99 μM (85% of cluster 5 compounds), unlike the rest of the Moonshot project database (67%). Closer analysis can be found in Section S4.2 and Fig. S4.7.† In summary, based on Arpeggio and BigDFT contact analysis and reported assay data, cluster 5 binders are promising building blocks for substrate-competing inhibitor design.

4.2 Covalent docking of COVID Moonshot compounds

To accommodate induced fit and create high quality poses of covalent inhibitors for future optimisation, we selected 540 covalently-reacting compounds from 10 001 Moonshot-designed compounds and docked them using AutoDock4 [68] into the M^{pro} structure of the corresponding covalent "inspiration" fragments.[36] We generated an interaction Tanimoto

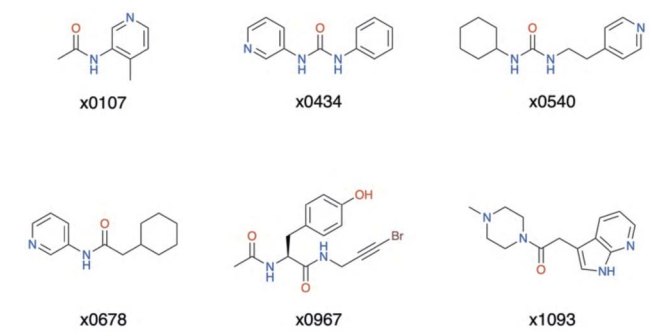


Fig. 15 Structures of the cluster 5 XChem compounds. Note the prevalence of nitrogen-containing heterocycles, and the phenol-containing derivative x0967.

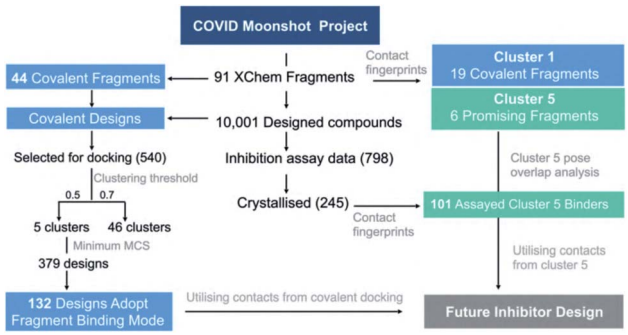


Fig. 16 Analysis of fragment and designed compounds from the Moonshot project. Workflow used to identify promising fragments and guide novel designs.

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

distance matrix as described earlier, and analysed the ability of the procedure to recapitulate the binding mode of the parent fragment. The normalized shape and pharmacophoric overlap (SuCOS[69]) of the lowest energy pose of the highest populated cluster for each Moonshot compound was compared with the inspiration covalent XChem fragment (Fig. S4.8†). When controlling for the smallest maximum common substructure (MCS) that encompasses the covalent warhead and one additional atom in the compound, 379 designs remain, 132 (34.8%) of which adopted the binding mode of the inspiration fragment. Given the high similarity between the fragments and docked designs, it is likely that these binding modes are more representative of the actual binding mode. A summary of the workflow is shown in Fig. 16.

In summary, our covalent docking method is more likely to identify the correct binding mode when substantial overlap exists between the inspiration fragment and designed compound beyond the covalent warhead (Section S4.3 and Fig. S4.8, S4.9†). This generated 132 high quality docked poses which serve as inspiration for future inhibitor design and were used in our proposals for compound derivatisation in Section 4.3. All poses of the 540 docking runs are available at https://github.com/gmm/SARS-CoV-2-Modelling.

4.3 Implications for future inhibitor design

We compared the interactions of the cluster 5 binders with those in the substrates, peptide inhibitors, and XChem fragments. Interestingly, unlike the peptides, almost none of the cluster 5 binders interact with the oxyanion hole. The only cluster 5 compounds where this contact is made are a series of covalent inhibitors, none of which showed promising potency (Fig. S4.11†). An exhaustive search of Moonshot structures showed that at the time of the analysis, no non-covalent inhibitor has ever been tested that includes both the typical cluster 5 binding mode while also being able to interact with the oxyanion hole.

We compared the structures of the top 10 compounds in cluster 5 (part of the dataset analysed in Section 4.1) to the docked structures of covalent Moonshot designs (Section 4.2).

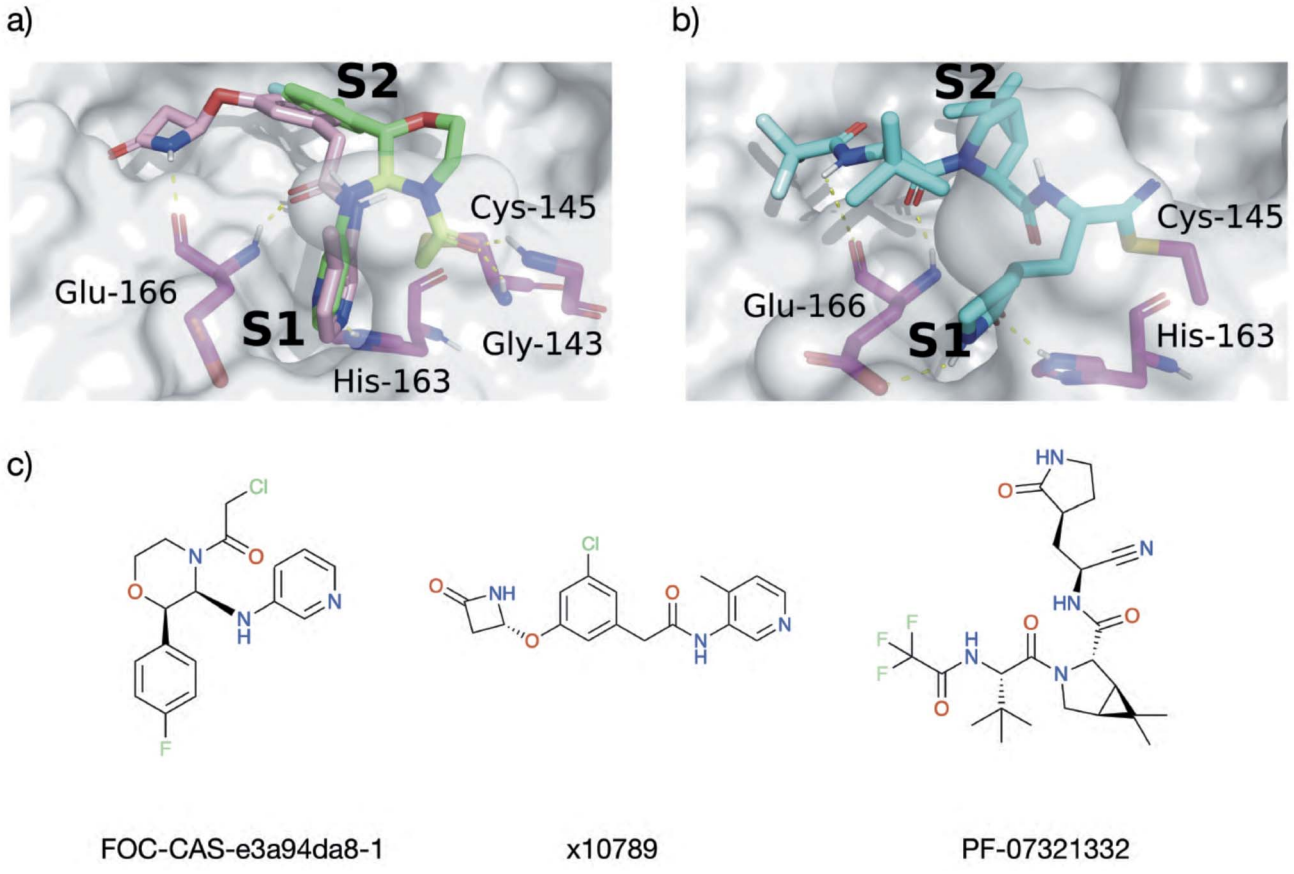


Fig. 17 Docking informs novel inhibitor design. HBs between M^{pro} (magenta) and the ligands are shown as dotted yellow lines. (a) Overlay of the docked pose of FOC-CAS-e3a94da8-1 (green and greenish-yellow) with the crystal structure of x10789 (pink) on the M^{pro} surface (PDB entry 5RER; 1.88 Å resolution).[35] Derivatisation of x10789 into the oxyanion hole could be achieved by attaching a methylene amide group present in x0830 (highlighted greenish-yellow). (b) Docked pose of Pfizer's Phase I covalent inhibitor PF-07321332, covalently docked into M^{pro} (PDB entry 6XHM; 1.41 Å resolution).[71] PF-07321332 (cyan) is covalently attached to Cys-145. The docked PF-07321332 adopts the same major contacts as the 'combination' of x10789 and x0830, namely the double HB to the backbone of Glu-166, the HB to His-163 in the S1 subsite, and a series of hydrophobic interactions in the S2 subsite. (c) Structures of Moonshot designed compound FOC-CAS-e3a94da8-1, crystallographic fragment x10789, and inhibitor PF-07321332.

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

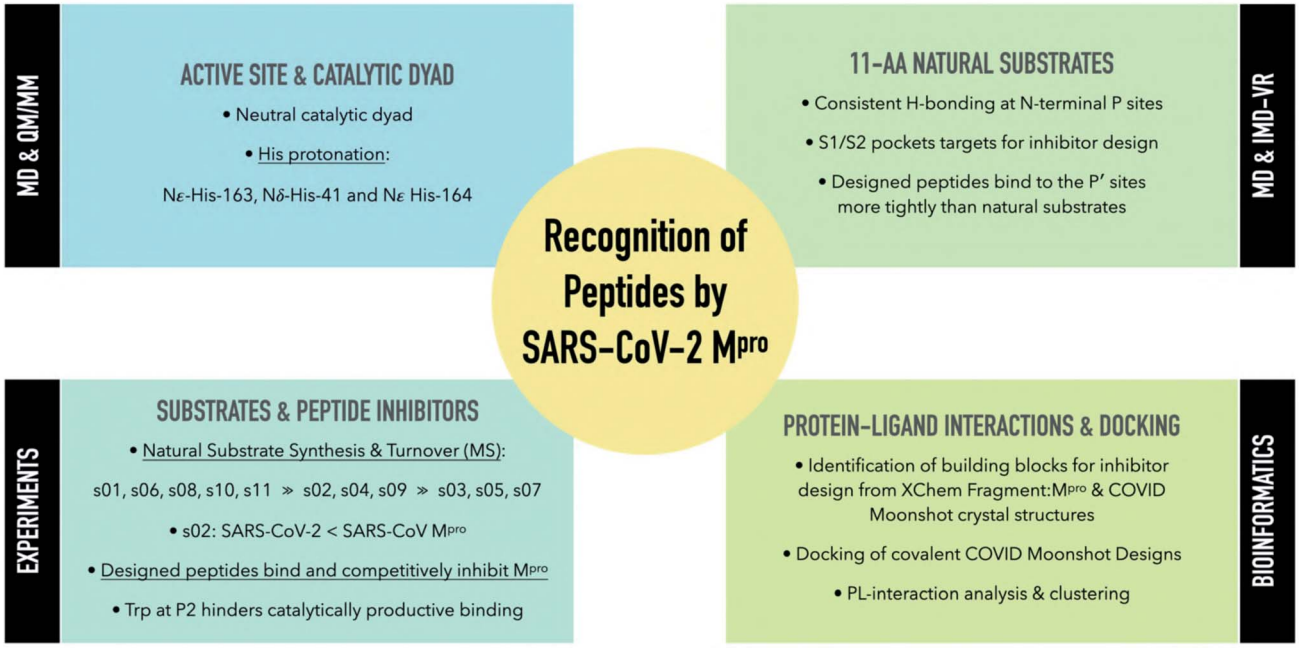


Fig. 18 Summary of the research and results obtained in this work.

Two compounds—FOC-CAS-e3a94da8-1 and MIH-UNI-e573136b-3—were selected based on their high normalized SuCOS overlap with their inspiration fragments, strongly suggesting that their docked binding modes reflect the actual poses.[70] Both compounds bind into the oxyanion hole as well as into S1 and S2, providing a clear opportunity for extension of the cluster 5 binders (Fig. S4.12†).

Most cluster 5 binders place the aromatic heterocycle into the S1 site and the carbonyl oxygen of the amide linker bonds to Glu-166 (Fig. 17). The position of this amide nitrogen overlays perfectly with the ring amine present in the docked compound FOC-CAS-e3a94da8-1. Thus, extension of cluster 5 binders into the oxyanion hole could be achieved by adding a substituent at the amide nitrogen. A promising candidate for extension is x10789, which makes a HB with the backbone oxygen of Glu-166 (Fig. 17a) and mimics the non-prime side binding mode of peptide inhibitor p13 (Fig. 14e), even binding into the S4 site *via* its β-lactam ring (Fig. 17a). Additional expansion into the oxyanion hole and S1′ through the amide linker could yield a powerful peptidomimetic inhibitor, combining protein interactions observed for the substrates, peptide inhibitors and small molecule fragments.

When comparing interactions exhibited by cluster 5 binders (Glu-166, His-163) or covalent fragments (Gly-143, Cys-145) with the contacts present in the docked structure of the recently published Phase 1 clinical trial candidate PF-07321332 by Pfizer (Fig. 17b),[13,14] a nearly identical interaction pattern to the cluster 5 binding motif is observed. However, note that for reacted PF-07321332, AutoDock4 was unable to place the negatively charged azanide nitrogen in the oxyanion hole, which is the expected position given its similarity to related warheads previously docked (Fig. S4.9†).

5. Conclusions

A wealth of crystal structures of SARS-CoV-2 M^{pro} is available, including hundreds with ligands. There is thus the question of how best to use this static information to help develop M^{pro} inhibitors optimised in terms of efficacy and safety for COVID-19 treatment. The dimeric nature of M^{pro}, coupled with its multiple substrates, makes it challenging to understand the structural and dynamic features underpinning selectivity and catalysis, as is the case with many other proteases. Such an understanding is, of course, not essential to develop medicines, as shown by work with other viral proteases. However, it may help improve the quality of such medicines and the efficiency with which they are developed. It will also lay the foundation for tackling anti-COVID-19 drug resistance—a challenge we will likely encounter as experience with the HIV global pandemic implies. The scale of global efforts on M^{pro} makes this system an excellent model for collaborative efforts linking experimental biophysics, modelling, and drug development (Fig. 18).

The results of our combined computational studies, employing classical molecular mechanics and quantum mechanical techniques, ranging from automated docking and MD simulations to linear-scaling DFT, QM/MM, and iMD-VR, provide consistent insights into key binding and mechanistic questions. One such question concerns the protonation state of the 'catalytic' His-41/Cys-145 dyad, an important consideration in the rates of reaction of covalently linked M^{pro} inhibitors which ultimately relates to their selectivity and potency. Our results indicate that a neutral catalytic dyad is thermodynamically preferred in M^{pro} complexed with an unreacted substrate, justifying the neutral state for MD simulations. A more reactive thiolate anion may be deleterious to the virus, as it will be

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

susceptible to reaction with electrophiles. Importantly, analysis of the active site suggests that the precise mechanism of proton transfer in the His-41/Cys-145 dyad involves dynamic interactions with other residues, including His-163, His-164, Asp-187, and a water hydrogen bonded to the latter two residues and His-41. Proton transfer may be considered a relatively simple part of the overall catalytic cycle—these results thus highlight how M^{pro} catalysis is likely a property of (at least) the entire active site region, with a future challenge being to understand motions during substrate binding, covalent reaction, and product release.

The models we have developed of M^{pro} in complex with its 11 natural substrates provided a basis for analysis of key interactions involved in substrate recognition and for comparison with (potential) inhibitor binding modes. Notably, the P′ (C-terminal) side of substrates appears to be much less tightly bound than the P (N-terminal) side, where there is remarkable consistency in the hydrogen bonding patterns across the substrates. This difference may in part reflect the need for the P′ side to leave (at least from the immediate active site region) after acyl–enzyme complex formation and prior to acyl–enzyme hydrolysis. The tighter binding of the N-terminal P-side residues suggests these are likely more important in substrate recognition by M^{pro}. This is also reflected in potent inhibitors, such as N3 and peptidomimetic ketoamides,[2,4] which predominantly bind in these non-prime S subsites. The development of S-site-binding inhibitors may also reflect the nature of the substrates used in screens leading to them, which typically comprise an S-site binding peptide with a C-terminal group enabling fluorescence-based measurement. Our results imply that there is considerable scope for developing inhibitors exploiting the S′ subsites, or both S and S′ subsites, though relatively more effort may be required to obtain tight binders compared to targeting the S subsites.

Consistent with prior studies, our work highlights the critical role of the completely conserved P1 Gln residue in productive substrate binding and analogously in inhibitor binding. However, the nature of the P2/S2 interaction is also important in catalysis. In the natural substrates (Fig. 1), the P2 position is Leu in 9 of the 11 substrates, Phe in s02 (which displays medium turnover efficiency), and Val in s03 (which is a poor substrate). Our results show that the S2 subsite plays a critical role in recognition and inhibition. S2 is highly plastic (Fig. 6 and S4.4†) and can accommodate a range of different sidechains, including larger groups, though not necessarily in a productive manner. The observation that substrates with a P2 Leu vary in efficiency reveals that interactions beyond those involving P1 and P2 are important, reinforcing the notion that (likely dynamic) interactions beyond the immediate active site are important in determining selectivity both in terms of binding and rates of reaction of enzyme–substrate complexes.

Notably, the results of computational alanine scanning mutagenesis followed by design, aimed at identifying peptides that would bind more tightly than the natural substrates, led to the finding that substitution of a Trp at P2 ablates hydrolysis creating an inhibitor. The observations with peptide inhibitors of M^{pro} have precedence in studies with other nucleophilic proteases, including the serine protease elastase, showing that substrate substitutions away from the scissile P1/P′ residues can cause inhibition.[72,73] There is thus scope for the extensive development of tight binding peptidic and peptidomimetic M^{pro} inhibitors for use in inhibition and mechanistic/biophysical studies, with the Trp at P2 of the peptide inhibitors being a good point for SAR exploration, potentially by (i) replacement of the indole hydrogen with suitable alkyl or aryl substituents; (ii) introduction of substituents with different stereoelectronic properties at C-2 or C-5 of the indole ring; or (iii) cyclization by the insertion of a methylene group linking position 2 of the indole ring to the α-nitrogen of Trp itself.[74]

Finally, the combined analysis of interactions involved in substrate binding and extensive structural information on inhibitor/fragment binding to M^{pro} enabled us to identify a cluster of inhibitors whose interactions relate to those conserved in substrate binding (*e.g.*, involving the Glu-166 backbone, His-163 sidechain, and/or the oxyanion hole formed by the Cys-145 and Gly-143 backbones). Building out from these 'privileged' interactions (Fig. 17) might be a useful path for inhibitor discovery. Indeed, an M^{pro} inhibitor now in clinical trials[13,14] exploits the same 'privileged' interactions that we identified. We hope the methods and results that have emerged from our collaborative efforts will help accelerate the development of drugs for treatment of viral infections, and particularly COVID-19.

6. Methods

A detailed description of the experimental and computational methods employed in this work is provided in the ESI.†

Data availability

Structures, input files, underlying data, and source code are publicly available on GitHub at https://github.com/gmm/SARS-CoV-2-Modelling.

Author contributions

G. M. M. generated the initial 11 natural substrate-SARS-CoV-2 M^{pro} models. R. M. T. carried out QM/MM studies with subsequent contributions from K. Ś. and V. Mo.; H. T. H. C. carried out classical MD simulations and ADCP peptide docking. R. K. W. and H. M. D. performed iMD-VR simulations. M. A. M. performed protein–substrate, protein–peptide inhibitor and protein–ligand contact analysis, fragment clustering, COVID Moonshot covalent docking and future inhibitor suggestions. M. A. M. and L. G. performed bioinformatic study of fragment binding to M^{pro}. L. G. performed linear-scaling QM-DFT calculations, and W. D., T. N. and T. J. W. contributed with its setup and analysis. D. K. S. carried out mutagenesis analysis and designed novel peptides (using software devised by R. B. S.). E. S., P. L., C. S. D., C. D. O. and M. A. W. carried out M^{pro} production and purification. T. R. M. and T. J. synthesised and purified natural and designed peptides. T. R. M. also performed kinetic analyses. V. M. undertook non-denaturing MS

© 2021 The Author(s). Published by the Royal Society of Chemistry

analyses. V. Mo. and A. L. contributed to discussions. C. J. S., A. J. M., D. K. S., F. D. and G. M. M. conceptualised and supervised the study. H. T. H. C., R. K. W., M. A. M., T. R. M., R. M. T., V. M., C. J. S., D. K. S., L. G., A. J. M., F. D. and G. M. M. wrote the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Prof. Michel Sanner for helpful discussions about AutoDock CrankPep. H. T. H. C. thanks the EPSRC Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1) and the Clarendon Scholarship. M. A. M. thanks the EPSRC University of Oxford Mathematics, Physical, and Life Sciences Division (MPLS) Doctoral Training Partnership (DTP) Grant Number EP/R513295/1 and GlaxoSmithKline. R. K. W. thanks the EPSRC for a PhD studentship. T. R. M. thanks the BBSRC *via* BB/M011224/1, and Dr Anthony Aimon for dispensing peptibitors using ECHO 550 for the mode of inhibition studies. R. M. T. and T. J. W. acknowledge the EPSRC Centre for Doctoral Training in Theory and Modelling in Chemical Sciences (EP/L015722/1). T. J. was supported by the Oxford-GSK-Crick Doctoral Programme in Chemical Biology, EPSRC (EP/R512060/1) and GlaxoSmithKline. E. S. thanks Anastasia Kantsadi and Prof. Ioannis Vakonakis for providing the M^{pro} plasmid in a pFLOAT vector. D. R. G. acknowledges funding from the Royal Society (URF\R\180033). L. G. thanks Michel Masella for useful discussions and for the force field comparison. L. G., W. D. and T. N. gratefully acknowledge the CEA-RIKEN collaboration. A. J. M. thanks EPSRC (EP/M022609/1). R. B. S. and D. K. S. thank EPSRC Poppi Programme Grant (EP/N013573/1); and BrisSynBio, the BBSRC and EPSRC Synthetic Biology Research Centre (BB/L01386X/1). A. J. M., J. S. and H. M. D. thank the British Society for Antimicrobial Chemotherapy for support (BSAC-COVID-30). V. M. thanks Barcelona Supercomputing Center (QSB-2021-1-0007). G. M. M. thanks the EPSRC and MRC for their indirect support *via* EP/S024093/1 and EP/L016044/1. This project made use of time on JADE granted *via* the UK High-End Computing Consortium for Biomolecular Simulation (HECBioSim, http://www.hecbiosim.ac.uk), supported by the EPSRC *via* EP/P020275/1. MD simulations were also carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol (http://www.bris.ac.uk/acrc) under an award for COVID-19 research. D. K. S. and A. J. M. thank EPSRC *via* HECBioSim for providing ARCHER/ARCHER2 time through a COVID-19 rapid response call; and Oracle Research for Oracle Public Cloud Infrastructure time under an award for COVID-19 research (http://cloud.oracle.com/en_US/iaas). L. G., W. D. and T. N. acknowledge RIKEN through the HPCI System Research Project (Project ID: hp200179 and hp210011) for computer time at the Fugaku supercomputer facility. L. G. is also grateful to the TGCC of CEA for granting of compute hours (gch0429 and gen12047 projects), and to the MaX Center of Excellence. C. J. S. thanks the Wellcome Trust, Cancer Research UK and the Biotechnology and Biological Sciences Research Council for funding. This research was funded in whole, or in part, by the Wellcome Trust (grant no. 106244/Z/14/Z). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

References

1 Johns Hopkins University of Medicine, *COVID-19 Map*, https://coronavirus.jhu.edu/map.html, (accessed 2021).
2 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289–293.
3 J. S. Morse, T. Lalonde, S. Xu and W. R. Liu, *ChemBioChem*, 2020, **21**, 730–738.
4 L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409–412.
5 B. Xia and X. Kang, *Protein Cell*, 2011, **2**, 282–290.
6 T. J. El-Baba, C. A. Lutomski, A. L. Kantsadi, T. R. Malla, T. John, V. Mikhailov, J. R. Bolla, C. J. Schofield, N. Zitzmann, I. Vakonakis and C. V. Robinson, *Angew. Chem., Int. Ed.*, 2020, **59**, 23544–23548.
7 W. Rut, K. Groborz, L. Zhang, X. Sun, M. Zmudzinski, B. Pawlik, X. Wang, D. Jochmans, J. Neyts, W. Młynarski, R. Hilgenfeld and M. Drag, *Nat. Chem. Biol.*, 2021, **17**, 222–228.
8 L. Zhu, S. George, M. F. Schmidt, S. I. Al-Gharabli, J. Rademann and R. Hilgenfeld, *Antiviral Res.*, 2011, **92**, 204–212.
9 G. E. Crooks, G. Hon, J. M. Chandonia and S. E. Brenner, *Genome Res.*, 2004, **14**, 1188–1190.
10 C. D. Owen, P. Lukacik, C. M. Strain-Damerell, A. Douangamath, A. J. Powell, D. Fearon, J. Brandao-Neto, A. D. Crawshaw, D. Aragao, M. Williams, R. Flaig, D. Hall, K. McAauley, D. I. Stuart, F. von Delft and M. A. Walsh, *PDB 6YB7*, 2020, DOI: 10.2210/pdb6yb7/pdb.
11 T. Pillaiyar, M. Manickam, V. Namasivayam, Y. Hayashi and S.-H. Jung, *J. Med. Chem.*, 2016, **59**, 6595–6628.
12 H. M. Mengist, T. Dilnessa and T. Jin, *Front. Chem.*, 2021, **9**, 622898.
13 D. Owen, *presented in part at the American Chemical Society Spring 2021 Meeting*, 2021.
14 B. Halford, *C&EN: Pfizer unveils its oral SARS-CoV-2 inhibitor*, https://cen.acs.org/acs-news/acs-meeting-news/Pfizer-unveils-oral-SARS-CoV/99/i13, (accessed April 2021).
15 K. Świderek and V. Moliner, *Chem. Sci.*, 2020, **11**, 10626–10630.
16 K. Arafet, N. Serrano-Aparicio, A. Lodola, A. J. Mulholland, F. V. González, K. Świderek and V. Moliner, *Chem. Sci.*, 2021, **12**, 1433–1444.
17 C. A. Ramos-Guzmán, J. J. Ruiz-Pernía and I. Tuñón, *ACS Catal.*, 2020, **10**, 12544–12554.

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

18 D. Mondal and A. Warshel, *Biochemistry*, 2020, **59**, 4601–4608.
19 A. Acharya, R. Agarwal, M. B. Baker, J. Baudry, D. Bhowmik, S. Boehm, K. G. Byler, S. Y. Chen, L. Coates, C. J. Cooper, O. Demerdash, I. Daidone, J. D. Eblen, S. Ellingson, S. Forli, J. Glaser, J. C. Gumbart, J. Gunnels, O. Hernandez, S. Irle, D. W. Kneller, A. Kovalevsky, J. Larkin, T. J. Lawrence, S. LeGrand, S. H. Liu, J. C. Mitchell, G. Park, J. M. Parks, A. Pavlova, L. Petridis, D. Poole, L. Pouchard, A. Ramanathan, D. M. Rogers, D. Santos-Martins, A. Scheinberg, A. Sedova, Y. Shen, J. C. Smith, M. D. Smith, C. Soto, A. Tsaris, M. Thavappiragasam, A. F. Tillack, J. V. Vermaas, V. Q. Vuong, J. Yin, S. Yoo, M. Zahran and L. Zanetti-Polzi, *J. Chem. Inf. Model.*, 2020, **60**, 5832–5852.
20 J. Chodera, A. A. Lee, N. London and F. von Delft, *Nat. Chem.*, 2020, **12**, 581.
21 J. Loschwitz, A. Jäckering, M. Keutmann, M. Olagunju, R. J. Eberle, M. A. Coronado, O. O. Olubiyi and B. Strodel, *Bioorg. Chem.*, 2021, **111**, 104862.
22 R. Abel, M. Paredes Ramos, Q. Chen, H. Pérez-Sánchez, F. Coluzzi, M. Rocco, P. Marchetti, C. Mura, M. Simmaco, P. E. Bourne, R. Preissner and P. Banerjee, *Front. Chem.*, 2020, **8**, 590263.
23 M. M. Ghahremanpour, J. Tirado-Rives, M. Deshmukh, J. A. Ippolito, C.-H. Zhang, I. Cabeza de Vaca, M.-E. Liosi, K. S. Anderson and W. L. Jorgensen, *ACS Med. Chem. Lett.*, 2020, **11**, 2526–2533.
24 C.-H. Zhang, E. A. Stone, M. Deshmukh, J. A. Ippolito, M. M. Ghahremanpour, J. Tirado-Rives, K. A. Spasov, S. Zhang, Y. Takeo, S. N. Kudalkar, Z. Liang, F. Isaacs, B. Lindenbach, S. J. Miller, K. S. Anderson and W. L. Jorgensen, *ACS Cent. Sci.*, 2021, **7**, 467–475.
25 L. L. Wang, K. Lo, Y. Chandrasekhar, R. Reas, J. Yang, D. Eide, K. Funk, R. Kinney, Z. Liu, W. Merrill, P. Mooney, D. Murdick, D. Rishi, J. Sheehan, Z. Shen, B. Stilson, A. D. Wade, K. Wang, C. Wilhelm, B. Xie, D. Raymond, D. S. Weld, O. Etzioni and S. Kohlmeier, arXiv:2004.10706v2, 2020.
26 H. Yang, M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld and Z. Rao, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13190.
27 C. Huang, P. Wei, K. Fan, Y. Liu and L. Lai, *Biochemistry*, 2004, **43**, 4568–4574.
28 J. Tan, K. H. G. Verschueren, K. Anand, J. Shen, M. Yang, Y. Xu, Z. Rao, J. Bigalke, B. Heisen, J. R. Mesters, K. Chen, X. Shen, H. Jiang and R. Hilgenfeld, *J. Mol. Biol.*, 2005, **354**, 25–40.
29 A. Paasche, A. Zipper, S. Schafer, J. Ziebuhr, T. Schirmeister and B. Engels, *Biochemistry*, 2014, **53**, 5930–5946.
30 A. Pavlova, D. L. Lynch, I. Daidone, L. Zanetti-Polzi, M. D. Smith, C. Chipot, D. W. Kneller, A. Kovalevsky, L. Coates, A. A. Golosov, C. J. Dickson, C. Velez-Vega, J. S. Duca, J. V. Vermaas, Y. T. Pang, A. Acharya, J. M. Parks, J. C. Smith and J. C. Gumbart, *Chem. Sci.*, 2021, **12**, 1513–1527.
31 N. Verma, J. A. Henderson and J. Shen, *J. Am. Chem. Soc.*, 2020, **142**, 21883–21890.
32 N. Díaz and D. Suárez, *Chem. Commun.*, 2021, **57**, 5314–5317.
33 A. Paasche, PhD thesis, Julius-Maximilians-Universität Würzburg, 2013.
34 D. W. Kneller, G. Phillips, K. L. Weiss, S. Pant, Q. Zhang, H. M. O'Neill, L. Coates and A. Kovalevsky, *J. Biol. Chem.*, 2020, **295**, 17365–17373.
35 A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandão-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft and M. A. Walsh, *Nat. Commun.*, 2020, **11**, 5047.
36 H. Achdout, A. Aimon, E. Bar-David, H. Barr, A. Ben-Shmuel, J. Bennett, M. L. Bobby, J. Brun, B. Sarma, M. Calmiano, A. Carbery, E. Cattermole, J. D. Chodera, A. Clyde, J. E. Coffland, G. Cohen, J. Cole, A. Contini, L. Cox, M. Cvitkovic, A. Dias, A. Douangamath, S. Duberstein, T. Dudgeon, L. Dunnett, P. K. Eastman, N. Erez, M. Fairhead, D. Fearon, O. Fedorov, M. Ferla, H. Foster, R. Foster, R. Gabizon, P. Gehrtz, C. Gileadi, C. Giroud, W. G. Glass, R. Glen, I. Glinert, M. Gorichko, T. Gorrie-Stone, E. J. Griffen, J. Heer, M. Hill, S. Horrell, M. F. D. Hurley, T. Israely, A. Jajack, E. Jnoff, T. John, A. L. Kantsadi, P. W. Kenny, J. L. Kiappes, L. Koekemoer, B. Kovar, T. Krojer, A. A. Lee, B. A. Lefker, H. Levy, N. London, P. Lukacik, H. B. Macdonald, B. MacLean, T. R. Malla, T. Matviiuk, W. McCorkindale, S. Melamed, O. Michurin, H. Mikolajek, A. Morris, G. M. Morris, M. J. Morwitzer, D. Moustakas, J. B. Neto, V. Oleinikovas, G. J. Overheul, D. Owen, R. Pai, J. Pan, N. Paran, B. Perry, M. Pingle, J. Pinjari, B. Politi, A. Powell, V. Psenak, R. Puni, V. L. Rangel, R. N. Reddi, S. P. Reid, E. Resnick, M. C. Robinson, R. P. Robinson, D. Rufa, C. Schofield, A. Shaikh, J. Shi, K. Shurrush, A. Sittner, R. Skyner, A. Smalley, M. D. Smilova, J. Spencer, C. Strain-Damerell, V. Swamy, H. Tamir, R. Tennant, A. Thompson, W. Thompson, S. Tomasio, A. Tumber, I. Vakonakis, R. P. van Rij, F. S. Varghese, M. Vaschetto, E. B. Vitner, V. Voelz, A. von Delft, F. von Delft, M. Walsh, W. Ward, C. Weatherall, S. Weiss, C. F. Wild, M. Wittmann, N. Wright, Y. Yahalom-Ronen, D. Zaidmann, H. Zidane and N. Zitzmann, *bioRxiv*, 2020, DOI: 10.1101/2020.10.29.339317.
37 X. Xue, H. Yu, H. Yang, F. Xue, Z. Wu, W. Shen, J. Li, Z. Zhou, Y. Ding, Q. Zhao, X. C. Zhang, M. Liao, M. Bartlam and Z. Rao, *J. Virol.*, 2008, **82**, 2515–2527.
38 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
39 M. Gaus, A. Goez and M. Elstner, *J. Chem. Theory Comput.*, 2013, **9**, 338–354.

© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 September 2021. Downloaded on 7/6/2023 11:04:31 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

40 F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes and Y.-Z. Zhang, *Nature*, 2020, **579**, 265–269.

41 M. B. O'Connor, S. J. Bennie, H. M. Deeks, A. Jamieson-Binnie, A. J. Jones, R. J. Shannon, R. Walters, T. J. Mitchell, A. J. Mulholland and D. R. Glowacki, *J. Chem. Phys.*, 2019, **150**, 220901.

42 H. M. Deeks, R. K. Walters, S. R. Hare, M. B. O'Connor, A. J. Mulholland and D. R. Glowacki, *PLoS One*, 2020, **15**, e0228461.

43 H. M. Deeks, R. K. Walters, J. Barnoud, D. R. Glowacki and A. J. Mulholland, *J. Chem. Inf. Model.*, 2020, **60**, 5803–5814.

44 K. Fan, P. Wei, Q. Feng, S. Chen, C. Huang, L. Ma, B. Lai, J. Pei, Y. Liu, J. Chen and L. Lai, *J. Biol. Chem.*, 2004, **279**, 1637–1642.

45 H. C. Jubb, A. P. Higueruelo, B. Ochoa-Montaño, W. R. Pitt, D. B. Ascher and T. L. Blundell, *J. Mol. Biol.*, 2017, **429**, 365–371.

46 J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman and D. A. Case, *J. Am. Chem. Soc.*, 1998, **120**, 9401–9409.

47 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.

48 H. Gohlke, C. Kiel and D. A. Case, *J. Mol. Biol.*, 2003, **330**, 891–913.

49 A. Metz, C. Pfleger, H. Kopitz, S. Pfeiffer-Marek, K.-H. Baringhaus and H. Gohlke, *J. Chem. Inf. Model.*, 2012, **52**, 120–133.

50 L. E. Ratcliff, W. Dawson, G. Fisicaro, D. Caliste, S. Mohr, A. Degomme, B. Videau, V. Cristiglio, M. Stella, M. D'Alessandro, S. Goedecker, T. Nakajima, T. Deutsch and L. Genovese, *J. Chem. Phys.*, 2020, **152**, 194110.

51 W. Dawson, S. Mohr, L. E. Ratcliff, T. Nakajima and L. Genovese, *J. Chem. Theory Comput.*, 2020, **16**, 2952–2964.

52 E. Cho, M. Rosa, R. Anjum, S. Mehmood, M. Soban, M. Mujtaba, K. Bux, S. T. Moin, M. Tanweer, S. Dantu, A. Pandini, J. Yin, H. Ma, A. Ramanathan, B. Islam, A. S. J. S. Mey, D. Bhowmik and S. Haider, *J. Chem. Inf. Model.*, 2021, **61**, 3058–3073.

53 *Diamond, Fragalysis*, https://fragalysis.diamond.ac.uk/, (accessed January 2021).

54 L. Silvestrini, N. Belhaj, L. Comez, Y. Gerelli, A. Lauria, V. Libera, P. Mariani, P. Marzullo, M. G. Ortore, A. Palumbo Piccionello, C. Petrillo, L. Savini, A. Paciaroni and F. Spinozzi, *Sci. Rep.*, 2021, **11**, 9283.

55 T. Muramatsu, C. Takemoto, Y.-T. Kim, H. Wang, W. Nishii, T. Terada, M. Shirouzu and S. Yokoyama, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 12997–13002.

56 C. W. Wood, A. A. Ibarra, G. J. Bartlett, A. J. Wilson, D. N. Woolfson and R. B. Sessions, *Bioinformatics*, 2020, **36**, 2917–2919.

57 R. B. Sessions, *BUDE_SM Algorithm*, https://github.com/richardbsessions, (accessed April 2021).

58 K. Hetherington, S. Dutt, A. A. Ibarra, E. E. Cawood, F. Hobor, D. N. Woolfson, T. A. Edwards, A. Nelson, R. B. Sessions and A. J. Wilson, *RSC Chem. Biol.*, 2021, DOI: 10.1039/d1cb00137j.

59 S. McIntosh-Smith, J. Price, R. B. Sessions and A. A. Ibarra, *Int. J. High Perform. Comput. Appl.*, 2015, **29**, 119–134.

60 S. McIntosh-Smith, T. Wilson, A. Á. Ibarra, J. Crisp and R. B. Sessions, *Comput. J.*, 2012, **55**, 192–205.

61 C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422–427.

62 A. A. Ibarra, G. J. Bartlett, Z. Hegedüs, S. Dutt, F. Hobor, K. A. Horner, K. Hetherington, K. Spence, A. Nelson, T. A. Edwards, D. N. Woolfson, R. B. Sessions and A. J. Wilson, *ACS Chem. Biol.*, 2019, **14**, 2252–2263.

63 T. R. Malla, A. Tumber, T. John, L. Brewitz, C. Strain-Damerell, C. D. Owen, P. Lukacik, H. T. H. Chan, P. Maheswaran, E. Salah, F. Duarte, H. Yang, Z. Rao, M. A. Walsh and C. J. Schofield, *Chem. Commun.*, 2021, **57**, 1430–1433.

64 M. Wei, R. Wynn, G. Hollis, B. Liao, A. Margulis, B. G. Reid, R. Klabe, P. C. Liu, M. Becker-Pasha, M. Rupar, T. C. Burn, D. E. McCall and Y. Li, *J. Biomol. Screen.*, 2007, **12**, 220–228.

65 Y. Zhang and M. F. Sanner, *Bioinformatics*, 2019, **35**, 5121–5127.

66 J. Lee, L. J. Worrall, M. Vuckovic, F. I. Rosell, F. Gentile, A.-T. Ton, N. A. Caveney, F. Ban, A. Cherkasov, M. Paetzel and N. C. J. Strynadka, *Nat. Commun.*, 2020, **11**, 5877.

67 P. Jaccard, *New Phytol.*, 1912, **11**, 37–50.

68 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.

69 S. Leung, M. Bodkin, F. von Delft, P. Brennan and G. M. Morris, *ChemRxiv*, 2019, DOI: 10.26434/chemrxiv.8100203.v1.

70 S. Malhotra and J. Karanicolas, *J. Med. Chem.*, 2017, **60**, 128–145.

71 R. L. Hoffman, R. S. Kania, M. A. Brothers, J. F. Davies, R. A. Ferre, K. S. Gajiwala, M. He, R. J. Hogan, K. Kozminski, L. Y. Li, J. W. Lockner, J. Lou, M. T. Marra, L. J. Mitchell, B. W. Murray, J. A. Nieman, S. Noell, S. P. Planken, T. Rowe, K. Ryan, G. J. Smith, J. E. Solowiej, C. M. Steppan and B. Taggart, *J. Med. Chem.*, 2020, **63**, 12725–12747.

72 P. A. Wright, R. C. Wilmouth, I. J. Clifton and C. J. Schofield, *Eur. J. Biochem.*, 2001, **268**, 2969–2974.

73 R. C. Wilmouth, I. J. Clifton, C. V. Robinson, P. L. Roach, R. T. Aplin, N. J. Westwood, J. Hajdu and C. J. Schofield, *Nat. Struct. Mol. Biol.*, 1997, **4**, 456–462.

74 R. Castelli, M. Tognolini, F. Vacondio, M. Incerti, D. Pala, D. Callegari, S. Bertoni, C. Giorgio, I. Hassan-Mohamed, I. Zanotti, A. Bugatti, M. Rusnati, C. Festuccia, S. Rivara, E. Barocelli, M. Mor and A. Lodola, *Eur. J. Med. Chem.*, 2015, **103**, 312–324.

© 2021 The Author(s). Published by the Royal Society of Chemistry

nature chemistry

Article

https://doi.org/10.1038/s41557-023-01205-1

In vitro selection of macrocyclic peptide inhibitors containing cyclic γ^{2,4}-amino acids targeting the SARS-CoV-2 main protease

Received: 5 April 2022

Accepted: 14 April 2023

Published online: 22 May 2023

Check for updates

Takashi Miura[1], Tika R. Malla[2], C. David Owen[3,4], Anthony Tumber[2], Lennart Brewitz[2], Michael A. McDonough[2], Eidarus Salah[2], Naohiro Terasaka[1], Takayuki Katoh[1], Petra Lukacik[3,4], Claire Strain-Damerell[3,4], Halina Mikolajek[3,4], Martin A. Walsh[3,4], Akane Kawamura[2,5], Christopher J. Schofield[2] & Hiroaki Suga[1]

γ-Amino acids can play important roles in the biological activities of natural products; however, the ribosomal incorporation of γ-amino acids into peptides is challenging. Here we report how a selection campaign employing a non-canonical peptide library containing cyclic γ^{2,4}-amino acids resulted in the discovery of very potent inhibitors of the SARS-CoV-2 main protease (M^{pro}). Two kinds of cyclic γ^{2,4}-amino acids, *cis*-3-aminocyclobutane carboxylic acid (γ^{1}) and (1*R*,3*S*)-3-aminocyclopentane carboxylic acid (γ^{2}), were ribosomally introduced into a library of thioether-macrocyclic peptides. One resultant potent M^{pro} inhibitor (half-maximal inhibitory concentration = 50 nM), GM4, comprising 13 residues with γ^{1} at the fourth position, manifests a 5.2 nM dissociation constant. An M^{pro}:GM4 complex crystal structure reveals the intact inhibitor spans the substrate binding cleft. The γ^{1} interacts with the S1′ catalytic subsite and contributes to a 12-fold increase in proteolytic stability compared to its alanine-substituted variant. Knowledge of interactions between GM4 and M^{pro} enabled production of a variant with a 5-fold increase in potency.

γ-Amino acid residues can induce unique peptide conformations, such as C_{14}-helix, $C_{12/10}$-helix and C_{12}-turn secondary structures[1–10]. The γ-amino acids present in natural products play important roles in their biological activities and enable improved proteolytic stability and cell permeability[11,12]. Pepstatin, isolated from *Actinomyces*, has two γ-amino acid residues, which contribute to its protease inhibitory activity by enhancing binding to the active sites[13–15]. Didemnin, an antiviral/antitumor compound, has a γ-amino acid and a macrocyclic scaffold[16]. Given their remarkable bioactivities and conformational rigidity, macrocyclic peptides containing γ-amino acids are an attractive option for incorporation into novel therapeutic peptides.

Over the past two decades, the ribosomal synthesis and selection-based screening of peptides containing nonproteinogenic amino acids, including *N*-methyl-L-α-amino, D-amino and β-amino acids, have been accomplished[17–26]. However, the ribosomal incorporation of γ-amino acids into peptides has been challenging for the following three reasons: (1) efficient self-deacylation of γ-aminoacyl-tRNA via stereoelectronically favoured 5-oxo-trig cyclization, (2) slow accommodation of γ-aminoacyl-tRNAs into the ribosome A site and (3) slow

[1]Department of Chemistry, Graduate School of Science, The University of Tokyo, Tokyo, Japan. [2]Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, Chemistry Research Laboratory, University of Oxford, Oxford, UK. [3]Diamond Light Source, Harwell Science & Innovation Campus, Didcot, UK. [4]Research Complex at Harwell, Harwell Science & Innovation Campus, Didcot, UK. [5]Chemistry – School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK. e-mail: hsuga@chem.s.u-tokyo.ac.jp

Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1

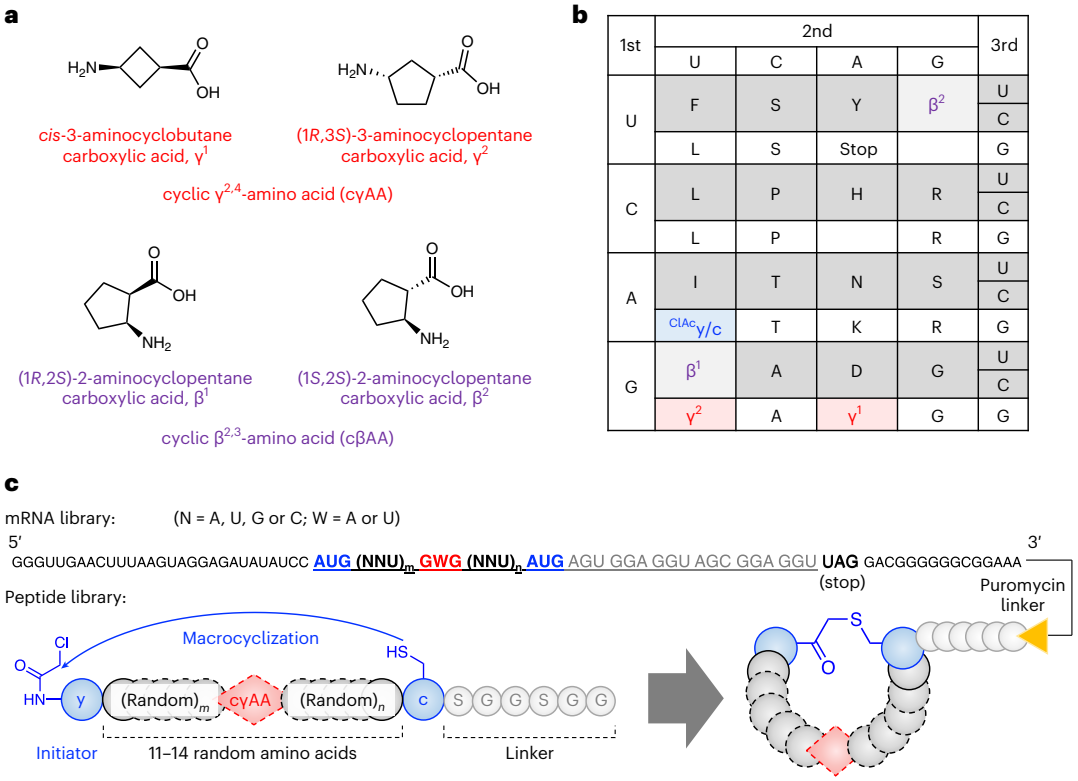


Fig. 1 | Ribosomal incorporation of cγAAs into a macrocyclic peptide library. **a**, Structures of the cγAAs and cβAAs used in this study. **b**, The reprogrammed codon table that contains ^{ClAc}y at the initiator AUG codon, and c, γ^1, γ^2, β^1 and β^2 at the elongator AUG, GAG, GUG, GUU and UGU codons, respectively. NNA codons are omitted because they were not used in the mRNA library. **c**, Construction of the random mRNA library and the corresponding peptide library. Peptides spontaneously macrocyclized between ^{ClAc}y and c via a thioether bond. The mRNA and peptide were covalently linked via a puromycin linker.

peptidyl transfer of the P-site peptidyl-tRNA onto the γ-amino group of the A-site γ-aminoacyl-tRNA due to poor compatibility with the peptidyl transferase centre of the ribosome[27,28]. We recently reported that the self-deacylation issue can be circumvented by using structurally constrained cyclic $\gamma^{2,4}$-amino acids (cγAAs)[29]. The cγAAs were charged on an engineered tRNA, referred to as $tRNA^{Pro1E2}$ (Supplementary Fig. 1), the T-stem and D-arm of which were optimized for efficient binding to elongation factor Tu (EF-Tu) and EF-P, thereby accelerating accommodation and peptidyl transfer, respectively[30–32]. Combination of the flexizyme-assisted[33,34] synthesis of cγAA–$tRNA^{Pro1E2}$ and the reconstituted *Escherichia coli* translation system, referred to as the Flexible In vitro Translation (FIT) system[23], that contains EF-Tu and EF-P with their adjusted concentrations enabled the incorporation of cγAAs into nascent peptide chains[29,35,36].

Here we report the ribosomal synthesis of a macrocyclic peptide library containing cγAAs and its application to an in vitro display, referred to as the Random nonstandard Peptides Integrated Discovery (RaPID) system[23]. The library contains two types of cγAAs, *cis*-3-aminocyclobutane carboxylic acid (γ^1) and (1*R*,3*S*)-3-aminocyclopentane carboxylic acid (γ^2) (Fig. 1a). Owing to the constrained rigid conformations of these cγAAs, they can be exquisite building blocks for preorganization of peptide secondary structures. Therefore, screening of macrocyclic peptide libraries containing cγAAs can lead to molecules with potent binding affinities and peptidase resistance. As a biomedicinally important test target protein, we chose M^{pro}, which plays an essential role in processing the polyproteins translated from the viral RNA[37]. Peptidomimetic inhibitors of M^{pro} have been shown to be therapeutically valuable for the treatment of coronavirus disease 2019 (COVID-19), with one approved for clinical use (nirmatrelvir, PF-07321332)[38]. Importantly, however, most of these inhibitors contain a reactive warhead (a nitrile in the case of nirmatrelvir) that reacts to form a covalent bond with a Cys residue in the M^{pro} active site. We took up the challenge of discovering potent 'non-covalent' active site binding M^{pro} inhibitors by applying the RaPID system with cγAA-containing macrocyclic peptide libraries.

Results

Selection of macrocyclic peptide binders to SARS-CoV-2 M^{pro}

To construct a cγAA-containing macrocyclic peptide library, γ^1 and γ^2 (Fig. 1a) were assigned to GAG and GUG codons using $tRNA^{Pro1E2}_{CUC}$ and $tRNA^{Pro1E2}_{CAC}$, respectively (Fig. 1b), employing the FIT system. Two cyclic $\beta^{2,3}$-amino acids (cβAAs), (1*R*,2*S*)-2-aminocyclopentane carboxylic acid (β^1) and (1*S*,2*S*)-2-aminocyclopentane carboxylic acid (β^2), were introduced at the GUU and UGU codons using $tRNA^{Pro1E2}_{GAC}$ and $tRNA^{GluE2}_{GCA}$, respectively, because we considered that the incorporation not only of cγAA, but also of cβAA, would give more diverse folding possibilities[25]. For the macrocyclization of the library via thioether bonds, we introduced *N*-chloroacetyl-D-tyrosine (^{ClAc}y) and D-cysteine (c) at the initiator AUG codon using $tRNA^{fMet}_{CAU}$ and elongator AUG codon using $tRNA^{Pro1E2}_{CAU}$, respectively. The thioether bond spontaneously forms between the *N*-terminal chloroacetyl group on ^{ClAc}y and the thiol group of the c residue (Fig. 1c). These amino acids were precharged onto the respective tRNAs using flexizymes. The peptide library comprised a repeat of 11–14 random residues encoded by GWG and NNU codons (W = A or U; N = A, U, G or C) flanked by the cyclizing ^{ClAc}y and c residues. Since multiple cγAA incorporations could not be efficiently achieved, as reported in our previous study[29], the library was designed to have only a single γ^1 or γ^2 appear at the GWG codon. The two cβAAs and the thirteen proteinogenic L-α-amino acids (A, D, F, G, H, I, L, N, P, R, S, T and Y) were assigned at NNU codons. The C-terminal SGGSGG sequence following

Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1

Table 1 | Binding affinities, inhibitory activities and serum stabilities of the macrocyclic peptides

Peptide name	Sequence	Read (%)	k_a (10^4 M^{-1} s^{-1})	k_d (10^{-3} s^{-1})	K_D (nM)	IC_{50} (nM)	$t_{1/2}$ (h)
GM1	^{Ac}yFHγ^{1}FPPHTYTLc-NH_2	7.0	41	0.94	2.3	40	90
GM1γ^{1}4A	^{Ac}yFHA FPPHTYTLc-NH_2	Mutant	32	3.2	10	>5,000	8.3
GM2	^{Ac}yFHγ^{1}FSIFRYRLc-NH_2	5.0	21	15	71	1,750	1.1
GM3	^{Ac}yLHYASTYSγ^{2}YRYAAc-NH_2	4.6	20	13	62	2,060	2.5
GM3γ^{2}10A	^{Ac}yLHYASTYSA YRYAAc-NH_2	Mutant	28	2.1	7.5	120	0.59
GM4	^{Ac}yFHγ^{1}LNLGYRPGc-NH_2	3.2	10	5.2	5.2	50	126
GM4γ^{1}4A	^{Ac}yFHA LNLGYRPGc-NH_2	Mutant	110	2.6	2.4	10	11
GM4γ^{1}4N	^{Ac}yFHN LNLGYRPGc-NH_2	Mutant	2.0	9.5	480	1,000	Not tested
GM4H3Q	^{Ac}yFQγ^{1}LNLGYRPGc-NH_2	Mutant	120	1.0	0.86	10	82
GM4H3A	^{Ac}yFAγ^{1}LNLGYRPGc-NH_2	Mutant	—	—	>10,000	>25,000	Not tested
GM4H3E	^{Ac}yFEγ^{1}LNLGYRPGc-NH_2	Mutant	2.0	7.1	360	>25,000	Not tested
GM5	^{Ac}yFHγ^{2}FRDLYILGc-NH_2	1.5	24	1.2	5.2	40	32
GM5γ^{2}4A	^{Ac}yFHA FRDLYILGc-NH_2	Mutant	14	8.7	61	50	21
GM6	^{Ac}yFIHTIFHDPHγ^{1}LPIc-NH_2	1.2	7.8	6.4	82	2,280	Not tested
GM7	^{Ac}yLHFYTSγ^{2}ATYIYPc-NH_2	0.67	4.8	23	490	7,930	Not tested

The thioether-macrocyclic peptides selected by the RaPID system, GM1–GM7, and their mutants are shown. The thioether bond is shown as a blue line where the sulfide atom is omitted. The γ^1 and γ^2 residues are in red. The table shows sequences, read (%) at the fourth round, kinetic association (k_a), kinetic dissociation (k_d), K_D, IC_{50} and half-life in human serum ($t_{1/2}$). Supplementary Fig. 5 shows the sensorgrams of SPR analysis; Fig. 2a and Extended Data Fig. 2 show the results of the M^{pro} inhibition assay; Fig. 2b and Extended Data Fig. 3 show the results of serum stability assays. –, the kinetic values could not be accurately determined due to low affinity.

the c residue was a linker peptide connected to the 3′ end of mRNA via a puromycin linker. To demonstrate the fidelity of translation, model macrocyclic peptides containing any of cγAA and cβAA (γ^1, γ^2, β^1 and β^2 assigned at GAG, GUG, GUU and UGU codons, respectively) were synthesized, and their identities were confirmed by matrix-assisted laser desorption/ionization coupled with time of flight mass spectrometry (MALDI-TOF MS; Supplementary Fig. 2).

The macrocyclic peptide library was then applied to the RaPID selection against recombinant SARS-CoV-2 M^{pro}. Translation of the random mRNA library into the peptide library, conjugation of the peptide with the parent mRNA via a puromycin linker and reverse transcription of the mRNA into the cDNA yielded peptide/mRNA/cDNA conjugates (Extended Data Fig. 1a). The library was first subjected to naked magnetic bead treatment to remove bead binders, and then applied to M^{pro}-immobilized magnetic beads to recover M^{pro} binders. The cDNA moiety of the recovered peptide/mRNA/cDNA conjugate was amplified by the polymerase chain reaction (PCR) and transcribed into the mRNA library for the next selection round (details in Methods). By repeating this affinity selection, the recovery rate of the M^{pro} binders was greatly increased at the third round of selection and, even more obviously, in the fourth round, while the recovery of the bead binders did not increase (Extended Data Fig. 1b). Deep sequencing of the cDNA library at the fourth round revealed that the library was enriched with several families of peptides bearing cγAA (Supplementary Table 1 shows the top 100 sequences). Among these, seven macrocyclic peptides containing cγAA were selected for further analysis of their binding affinities, inhibitory activity and proteolytic stability (Table 1, GM1–GM7). GM1, GM2, GM4 and GM6 contain a γ^1 residue in their sequences, while GM3, GM5 and GM7 contain γ^2. Note that the cβAA-containing peptides were not selected because they were not included in the major peptide families with high read numbers. GM1–GM7 were chemically synthesized on a large scale using the standard solid-phase method without the C-terminal SGGSGG linker, and their purities and identities were confirmed by ultra-performance liquid chromatography (UPLC) and MALDI-TOF MS, respectively (Supplementary Figs. 3 and 4).

Biological activities and stabilities of peptide inhibitors

We first evaluated the binding affinity of GM1–GM7 to M^{pro} by surface plasmon resonance (SPR). All the peptides exhibited low-to-moderate nanomolar kinetic equilibrium (K_D) values (Table 1 and Supplementary Fig. 5). Notably, GM1 and GM4, which contain γ^1, and GM5, which contains γ^2, have a conserved four residue motif, yFHγ^X (γ^X = γ^1 or γ^2), at their N termini and showed potent affinities, with K_D values of 2.3, 5.2 and 5.2 nM, respectively; GM2 has the same motif, but has a higher K_D (71 nM). GM6 and GM7, which had relatively low read frequencies, exhibited weak binding compared to GM1–GM5, indicating selection for tight binding peptides. To evaluate contributions of the cγAA residues to potency, we synthesized peptides where the cγAA was substituted with alanine (Table 1, GM1γ^{1}4A, GM3γ^{2}10A, GM4γ^{1}4A and GM5γ^{2}4A).

Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1

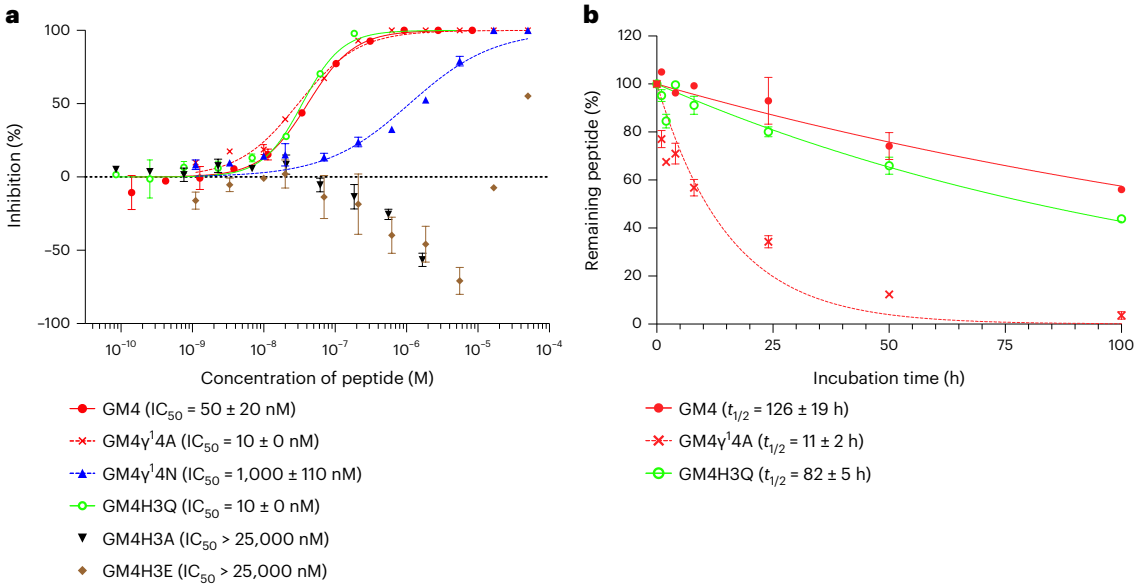


Fig. 2 | M^{pro} inhibitory activity and serum stability of GM4 and its mutants. **a**, Dose response analysis of peptides against M^{pro}. The M^{pro} inhibitory activities of peptides were investigated by solid-phase extraction purification coupled to MS analysis using a RapidFire 365 high-throughput sampling robot (Agilent) connected to an iFunnel Agilent 6550 accurate mass quadrupole TOF mass spectrometer[39]. IC_{50} values were determined by the mean of three or five independent replicates each performed in technical duplicate. Data are presented as mean values ± standard deviation, s.d. ($n = 5$ for GM4, GM4γ^{1}4A and GM4H3Q; $n = 3$ for GM4γ^{1}4N, GM4H3A and GM4H3E). Extended Data Fig. 2 shows other peptides. **b**, Serum stability assay of macrocyclic peptides. GM4, GM4γ^{1}4A and GM4H3Q were co-incubated with an internal standard peptide in human serum (37 °C). At each time point, the relative intensity of each peptide to the standard peptide was estimated by LC/MS. The relative intensity at 0 h was defined as 100%. Half-lives ($t_{1/2}$) were determined by analysing the mean of three technical replicates of each sample by nonlinear regression using GraphPad Prism 9. Data are presented as mean values ± s.d. ($n = 3$). Extended Data Fig. 3 shows results for other peptides.

In addition to GM1, GM4 and GM5, which share the conserved yFHγ^X motif, GM3 was selected for alanine substitution because of its high read percentage among peptides without the yFHγ^X motif. SPR measurements of the alanine mutants revealed that the K_D values of GM1γ^{1}4A, GM4γ^{1}4A and GM5γ^{2}4A were comparable or one order higher than those of the original peptides, indicating that the cγAA residue in the conserved yFHγ^X motif is important for binding. To our surprise, GM3γ^{2}10A exhibited an eightfold stronger binding affinity ($K_D = 7.5$ nM), revealing that the contribution of cγAA to the binding affinity is sequence context dependent.

We next evaluated the inhibitory activity of GM1–GM7 against the hydrolytic activity of SARS-CoV-2 M^{pro} using a reported MS-based method[39]. GM1–GM7 exhibited inhibition of M^{pro}; in particular, the yFHγ^X motif containing GM1, GM4 and GM5, which manifest single-digit nanomolar K_D values, showed particularly potent inhibition, with half-maximal inhibitory concentration (IC_{50}) values of 40, 50 and 40 nM, respectively (Table 1, Fig. 2a and Extended Data Fig. 2). The peptides with weaker binding affinity, that is, GM2, GM3, GM6 and GM7, had IC_{50} values of 1,750–7,930 nM, implying correlation between IC_{50} and K_D values. The inhibitory activities of the alanine mutants were then determined. Notably, despite GM1γ^{1}4A having only a fourfold weaker binding affinity compared with GM1, inhibition was ablated, revealing the importance of the cγAA. By contrast, GM3γ^{2}10A, GM4γ^{1}4A and GM5γ^{2}4A retained inhibition activity, showing the context-dependent effects of the cγAA ($IC_{50} = 120$, 10 and 50 nM, respectively; Table 1, Fig. 2a and Extended Data Fig. 2).

We evaluated the half-life of potent cγAA-containing peptides and their mutants in human serum, because in vivo stability is a critical factor in the development of therapeutic peptides. Each peptide and an uncleavable internal standard peptide were co-incubated in human serum at 37 °C, and the relative amount of the remaining sample peptide was estimated by liquid chromatography/MS (LC/MS). The potent inhibitors GM1 and GM4, containing γ^1, and GM5, containing γ^2, exhibit high peptidase resistance with half-lives ($t_{1/2}$) of 90 h, 126 h and 32 h, respectively (Table 1, Fig. 2b and Extended Data Fig. 3a). Importantly, by contrast, their alanine mutants show substantially shorter half-lives ($t_{1/2} = 8.3$ h, 11 h and 21 h for GM1γ^{1}4A, GM4γ^{1}4A and GM5γ^{2}4A, respectively; Table 1). The enhancement of serum resistance observed for GM4, for instance, was 12-fold. Notably, despite the substitution of γ^1 to an Ala residue not diminishing inhibition relative to GM4, the enhancement in serum stability by the γ^1 residue is substantial. Moreover, the less potent inhibitors GM2 and GM3 have more than nearly three orders of magnitude shorter half-lives (1.1 h and 2.5 h, respectively) compared to GM4, suggesting that the higher (binding and inhibitory) activity of GM4 is reflected in its higher serum stability, possibly due to its more conformationally constrained fold.

To identify peptidase cleavage sites in serum, the products of GM4 and GM4γ^{1}4A after 24 h incubation were analysed by LC/MS. In the case of GM4, six fragments (GM4-f1–GM4-f6, Extended Data Fig. 3b) containing the c-(thioether)-^{Ac}yFHγ^{1}L motif were detected, but no fragmentation of the motif itself was apparent. With GM4γ^{1}4A, three shorter fragments containing c-(thioether)-^{Ac}yF were detected (GM4γ^{1}4A-f1–GM4γ^{1}4A-f3; Extended Data Fig. 3c), indicating cleavage between F and H of the motif. These results show that the presence of non-canonical residues (one cγAA (γ^1), two D-amino acids and a thioether bond) can enable high peptidase resistance by preventing the proteolysis of flanking residues, leading to a remarkable improvement in serum stability.

X-ray crystallographic analysis of the GM4:M^{pro} complex

To gain structural insights into how GM4, including its γ^1 residue, interacts with M^{pro}, we obtained an X-ray crystal structure of their complex to 1.7 Å resolution (Fig. 3a and Supplementary Table 2). The structure reveals GM4 bound with high occupancy and in the same manner at

Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1

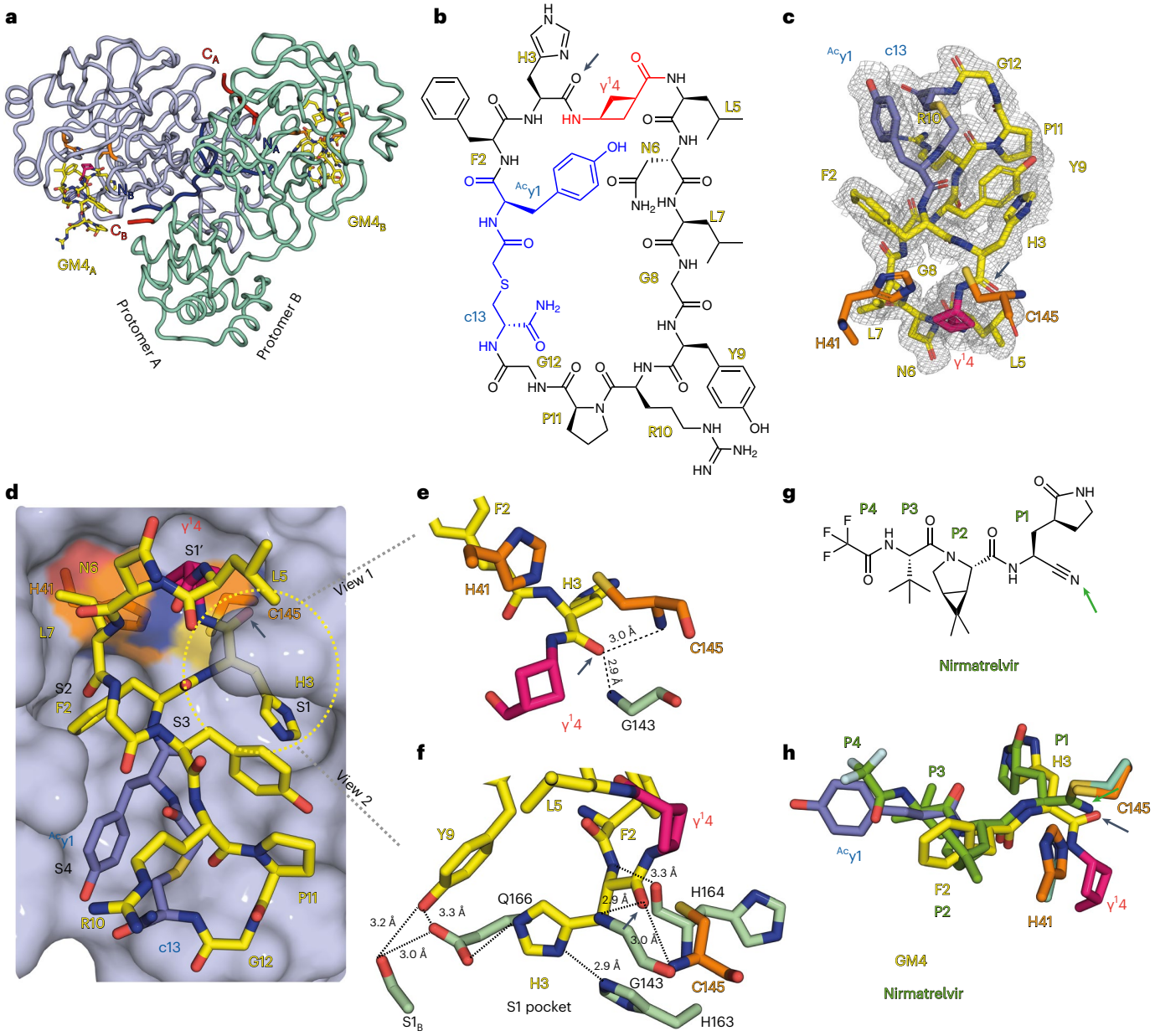


Fig. 3 | Crystallographic studies reveal the binding mode of the GM4 macrocyclic peptide at the M^{pro} active site. a, GM4 binds in the substrate binding cleft of both protomers A and B in the M^{pro} dimer. **b**, Structure of GM4; the H3$_{GM4}$ carbonyl O is indicated with an arrow. **c**, Polder omit map of GM4 contoured at a level of ±1.5 standard deviation (σ). **d**, View of GM4 at the active site. The H41 and C145 catalytic dyad is in orange. γ^{1}4$_{GM4}$ (magenta) occupies the S1′ pocket; the side chains of H3$_{GM4}$, F2$_{GM4}$ and ^{Ac}y1$_{GM4}$ occupy the S1, S2 and S4 pockets, respectively. **e**,**f**, Close-ups of the S1 pocket, showing the H3$_{GM4}$ backbone carbonyl in the oxyanion hole (C145, G143 backbone amides). The backbone NH of H3$_{GM4}$ is positioned to interact with the H164 backbone CO. The H3$_{GM4}$ imidazole is positioned to hydrogen bond with the H163 side chain and to interact with the E166 side chain, which interacts with S1 of protomer B, as observed in apo-M^{pro} structures. Y9$_{GM4}$ is 3.1 Å and 3.3 Å from the E166 and S1$_B$ side chains, respectively. **g**, Nirmatrelvir with the residues occupying the S1–S4 subsites labelled P1–P4 (the reactive nitrile is indicated with an arrow). **h**, Superimposition of views from crystal structures of M^{pro} with GM4 and nirmatrelvir (Protein Data Bank 7VH8 (ref. 60)). GM4 non-covalently interacts at the active site, while the nitrile of nirmatrelvir reacts with C145. Note the similar locations of the nirmatrelvir nitrile-derived N and the H3$_{GM4}$ carbonyl O.

the active site regions of both monomers in the M^{pro} dimer. Excellent electron density was observed for all residues in the GM4 macrocycle (Fig. 3c). The overall conformations of M^{pro} in the GM4 and nirmatrelvir[38] (used for comparison) complex structures are very similar (backbone root mean square deviation, 0.34).

The four residues in the yFHγ^{1} motif of GM4 occupy the S4, S2, S1 and S1′ substrate binding subsites, respectively[40] (Fig. 3d). H3$_{GM4}$ is accommodated in the S1 subsite, forming hydrogen bonds with M^{pro} C145, H164 and H163 (GM4 residues are identified by a subscript; Fig. 3e,f). The F2$_{GM4}$ residue binds in the S2 subsite, making hydrophobic interactions with M49, M165 and H41. The y1$_{GM4}$ side chain forms hydrogen bonds with T190 and Q192 in the S4 subsite (Extended Data Fig. 4a).

The c13$_{GM4}$–thioether link, which enables macrocyclization, and the adjacent G12$_{GM4}$ enable a turn connected to P11$_{GM4}$, R10$_{GM4}$ and Y9$_{GM4}$; none of these residues are positioned to make direct contacts with M^{pro} in the structure, though these cannot be ruled out in solution. The backbone of GM4 forms three intramolecular hydrogen bonds, that is between γ^{1}4$_{GM4}$ and L7$_{GM4}$, ^{Ac}y1$_{GM4}$ and R10$_{GM4}$, and P11$_{GM4}$ and

Content courtesy of Springer Nature, terms of use apply. Rights reserved

c13$_{GM4}$ (Extended Data Fig. 4b). The side chains of P11$_{GM4}$ and Y9$_{GM4}$ are arranged such that they 'stack' with each other, an arrangement positioning the phenolic side chain of Y9$_{GM4}$ adjacent to the imidazole ring of H3$_{GM4}$, appearing to lock it into the S1 pocket; the phenolic side chain of Y9$_{GM4}$ also forms interactions with Ser1′ from the other protomer of the M^{pro} dimer (Fig. 3f).

G8$_{GM4}$, L7$_{GM4}$, N6$_{GM4}$ and L5$_{GM4}$ form a loop leading to γ^{1}4, which is linked to H3$_{GM4}$. γ^{1}4$_{GM4}$ is positioned to make hydrophobic contact with T24 and T25; the side chains of L5$_{GM4}$ and N6$_{GM4}$ are positioned close to the surface of M^{pro}, with L7$_{GM4}$ projecting towards H41, C44 and T45. The observation of the unnatural γ^{1}4$_{GM4}$ residue, the cyclobutane ring of which was refined in a near planar conformation, in the P1′ position is striking and implies it contributes to the inhibition or lack of efficient GM4 hydrolysis.

At the active site region, the sulfur of the nucleophilic C145 is positioned close (3.4 Å) to the imidazole ε-nitrogen of H41, which acts as a general base/acid in M^{pro} catalysis, and to the carbonyl carbon (4.4 Å) of the amide linking H3$_{GM4}$ and γ^{1}4$_{GM4}$; this carbonyl carbon is positioned in a similar manner to the imine (anion) derived by reaction of the nirmatrelvir nitrile with C145 (Fig. 3g,h).

Like the binding modes of substrates (predicted) and inhibitors (as observed by crystallography), the carbonyl of the amide linking H3$_{GM4}$ and γ^{1}4$_{GM4}$ is positioned in the oxyanion hole, making hydrogen bonds with the C145 NH (2.9 Å) and with the G143 NH (2.9 Å). The ~100° angle between the C145 sulfur and the carbon and oxygen of the amide linking H3 and γ^{1}4 is close to that predicted for nucleophilic attack onto an amide carbonyl as observed in studies on serine proteases[41]. Despite this apparently catalytically productive arrangement, analysis of the electron density map implies a lack of substantial covalent reaction, consistent with solution studies showing a lack of efficient GM4 hydrolysis by M^{pro}. Lack of reaction appears to be a consequence of multiple interactions made between the peptide macrocycle and M^{pro}, likely involving P11$_{GM4}$, Y9$_{GM4}$ and H3$_{GM4}$, resulting in tight binding in a catalytically non-productive manner.

SARS-CoV-2 M^{pro} cleaves after a conserved glutamine residue (P1) in substrates with the following sequence motif: (P2, L/F/V/M (hydrophobic))(P1, Q)↓(P1′, S/A/G/N (small)), where ↓ represents the cleavage site[40,42]. Interestingly, rather than a glutamine (Q), GM4 has a histidine (H3) at the analogous 'P1' position. To investigate the contribution of the H3/P1 residue for activity, we evaluated peptides bearing Q, A or E residues at the analogous position. Among these, GM4H3Q exhibited a sixfold stronger binding affinity (K_D = 0.86 nM) and fivefold more potent inhibitory activity (IC_{50} = 10 nM; Table 1, Fig. 2 and Supplementary Fig. 5). By contrast, GM4H3A manifested no inhibition, and inhibition of GM4H3E was observed only with the highest tested concentration (25 μM). These results are consistent with the known substrate selectivity of M^{pro} at the P1 position[40].

In the case of the P1′ residue, M^{pro} prefers substrates with small residues, such as G, A and S, although N is also tolerated[40]. Strikingly, GM4 has γ^{1}4 at the P1′ position, likely reflecting the compact size of the γ^1 four-membered-ring main-chain residue. In support of this, GM4γ^{1}4A (γ^1 → A) exhibited a comparable inhibitory activity to that of GM4 (IC_{50} of GM4γ^{1}4A = 10 nM), and GM4γ^{1}4N was 20-fold less active (IC_{50} = 1,000 nM). Despite the almost equivalent M^{pro} inhibition activities of GM4 and GM4γ^{1}4A, the serum stability of GM4 is 12-fold higher ($t_{1/2}$ = 126 h) than that of GM4γ^{1}4A (Fig. 2b). Furthermore, the most potent mutant GM4H3Q bearing γ^{1}4 at P1 also exhibited a comparable serum stability ($t_{1/2}$ = 82 h). These results show how the introduction of a nonstandard amino acid at the P1′ position can prolong the serum stability to non-targeted protease-catalysed hydrolysis while maintaining potency against the targeted protease (M^{pro} in our case).

Discussion

Our results reveal GM4 and GM4H3Q as highly potent cyclic peptides targeting M^{pro}. Other peptides and peptidomimetic inhibitors of SARS-CoV-2 M^{pro} have been reported with varying potencies (Extended Data Fig. 5). Johansen-Leete et al. recently used the RaPID system to screen thioether-macrocyclic peptides containing standard proteinogenic amino acids in the random region between the cyclizing ^{ClAc}y and the downstream L-Cys (ref. 43). A macrocyclic peptide, with yLQY residues at the P3–P1′ equivalent positions, exhibiting an inhibition constant (K_i) value of 14 nM against M^{pro} was identified. The P1 equivalent Q and P1′ equivalent Y residues are on either side of the scissile amide in this peptide, resulting in slow cleavage by M^{pro}.

In the case of the M^{pro}:GM4 complex crystal structure, well-defined electron density was observed for the complete intact peptide in the active site pocket despite the carbonyl of the amide linking H3$_{GM4}$ and γ^{1}4$_{GM4}$ being positioned in the oxyanion hole. This observation both highlights the advantage of introducing a cγAA into the library and implies that GM4 and GM4H3Q in part resist hydrolytic cleavage due to the presence of γ^1 at the P1′ position. Other groups have also reported a substrate-derived cyclic peptide inhibitor of M^{pro} (UCI-1 (ref. 44)) and linear peptide inhibitors of M^{pro} (p13 (ref. 40) and compound 21 (ref. 45)), but the potency is limited to the micromolar range. Peptidomimetic inhibitors, including nirmatrelvir (PF-07321332) and PF-00835231, where the C-terminal residue reacts covalently with the nucleophilic cysteine have been developed[38,42,46–49] (Extended Data Fig. 5a,b). X^Q, which is present in these inhibitors, is an analogue of the substrate Q/P1 residue wherein the side chain δ-nitrogen is covalently linked to the γ-carbon (Extended Data Fig. 5b, indicated by the blue line) to give a conformationally constrained (*S*)-γ-lactam, which is well accommodated in the S1 subsite, resulting in improved potency relative to Q in inhibitors[46,50]. GM4 and GM4H3Q contain a related cγAA ring but are non-covalently binding inhibitors, but notably still exhibit low-nanomolar IC_{50} values.

Recently, non-covalent small molecule inhibitors against M^{pro} have also been developed by means of virtual screening and optimization of compounds based on structure–activity relationships[51–54]. Small molecular inhibitors often have an advantage of better cell membrane permeability than macrocycles and thus are able to inhibit M^{pro} in cells. In fact, an impressive M^{pro} non-covalent inhibitor, which is a current clinical candidate, S-217622, exhibited antiviral activity in cells (half-maximal effective concentration (EC_{50}) = 0.37 μM, in Extended Data Fig. 5c)[51]. Another example of a non-covalent inhibitor, compound 23, is reported to have potent in vitro inhibitory activity (IC_{50} = 20 nM), but exhibited cellular cytotoxicity (half-maximal cytotoxic concentration (CC_{50}) in Vero E6 cells = 1.15 μM)[52]. This contrasts with S-217622, for which cell cytotoxicity was not reported, presumably because of substantial medicinal chemistry efforts to optimize in vivo properties, including from a toxicity perspective. However, classical small molecule drugs often suffer from insufficient target specificity and unwanted cell cytotoxicity. By contrast, macrocyclic peptide inhibitors generally have remarkably high target specificities, leading to low cytotoxicity; however, at least for intracellular targets, relatively low membrane permeability can be an issue for peptide macrocycles. Our GM4 and related macrocycles show no cytotoxicity at concentrations as high as 100 μM, but as yet do not manifest obvious antiviral activity in cells, likely due to their poor membrane permeability. Johansen-Leete et al. have reported an M^{pro}-inhibiting thioether-macrocyclic peptide using the same technology as that used by us, except without elongation reprogramming (compound 1 in Extended Data Fig. 5a)[43]. The originally discovered peptide showed a potent in vitro efficacy (K_i = 14 nM) but no antiviral activity. However, an increase in cellular uptake, by attaching a cell-penetrating peptide (CPP) at its C-terminal region, resulted in modest but observable antiviral activity. It is known that the addition of CPP to a macrocycle can not only increase cell permeability but also increase cytotoxicity depending on the CPP sequence and cell types used for the study[55,56]. Since we have witnessed improvements in in vivo potency by reengineering the macrocyclic peptide backbone structures, such as by *N*-methylation, in our previous works[57,58],

Content courtesy of Springer Nature, terms of use apply. Rights reserved

our ultimate goal is to generate a non-CPP-modified macrocycle with better cell membrane permeability and potent antiviral activity. In the case of the GM4 peptides, our current efforts are directed at reducing molecular weight while maintaining the critical yFHγ[1] binding motif[59].

Most importantly, our strategy for the construction of a cγAA-containing macrocyclic peptide library and its application for the RaPID system has demonstrated that unprecedented protease inhibitors containing cγAAs can be generated. It is striking that crystallography reveals that the γ[1] residue of GM4 sits deeply in the M^{pro} active site, and that the compactness of γ[1] plays a critical role in organizing a unique turn conformation, resulting in resistance to proteolytic cleavage of GM4. The general ability of cγAAs in the preorganization of turn conformations thus appears to be advantageous in obtaining potent, compact and stable macrocycles with drug-like properties.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41557-023-01205-1.

References

1. Hintermann, T., Gademann, K., Jaun, B. & Seebach, D. γ-Peptides forming more stable secondary structures than α-peptides: synthesis and helical NMR-solution structure of the γ-hexapeptide analog of H-(Val-Ala-Leu)$_2$-OH. *Helv. Chim. Acta* **81**, 983–1002 (1998).
2. Seebach, D., Brenner, M., Rueping, M. & Jaun, B. $γ^2$-, $γ^3$-, and $γ^{2,3,4}$-amino acids, coupling to γ-hexapeptides: CD spectra, NMR solution and X-ray crystal structures of γ-peptides. *Chem. Eur. J.* **8**, 573–584 (2002).
3. Chatterjee, S. et al. Expanding the peptide β-turn in αγ hybrid sequences: 12 atom hydrogen bonded helical and hairpin turns. *J. Am. Chem. Soc.* **131**, 5956–5965 (2009).
4. Vasudev, P. G., Chatterjee, S., Shamala, N. & Balaram, P. Gabapentin: a stereochemically constrained γ amino acid residue in hybrid peptide design. *Acc. Chem. Res.* **42**, 1628–1639 (2009).
5. Basuroy, K. et al. Unconstrained homooligomeric γ-peptides show high propensity for C_{14} helix formation. *Org. Lett.* **15**, 4866–4869 (2013).
6. Giuliano, M. W. et al. Evaluation of a cyclopentane-based γ-amino acid for the ability to promote α/γ-peptide secondary structure. *J. Org. Chem.* **78**, 12351–12361 (2013).
7. Giuliano, M. W. et al. A γ-amino acid that favors 12/10-helical secondary structure in α/γ-peptides. *J. Am. Chem. Soc.* **136**, 15046–15053 (2014).
8. Konda, M., Kauffmann, B., Rasale, D. B. & Das, A. K. Structural and morphological diversity of self-assembled synthetic γ-amino acid containing peptides. *Org. Biomol. Chem.* **14**, 4089–4102 (2016).
9. Misra, R. et al. Structural dimorphism of achiral α,γ-hybrid peptide foldamers: coexistence of 12- and 15/17-helices. *Chem. Eur. J.* **23**, 3764–3772 (2017).
10. Debnath, S., Ghosh, S., Pandit, G., Satpati, P. & Chatterjee, S. Effect of differential geminal substitution of γ amino acid residues at the (*i* + 2) position of αγ turn segments on the conformation of template β-hairpin peptides. *J. Org. Chem.* **86**, 11310–11323 (2021).
11. Frackenpohl, J., Arvidsson, P. I., Schreiber, J. V. & Seebach, D. The outstanding biological stability of β- and γ-peptides toward proteolytic enzymes: an in vitro investigation with fifteen peptidases. *ChemBioChem* **2**, 445–455 (2001).
12. Bockus, A. T. et al. Probing the physicochemical boundaries of cell permeability and oral bioavailability in lipophilic macrocycles inspired by natural products. *J. Med. Chem.* **58**, 4581–4589 (2015).
13. Umezawa, H., Aoyagi, T., Morishima, H., Matsuzaki, M. & Hamada, M. Pepstatin, a new pepsin inhibitor produced by actinomycetes. *J. Antibiot. Tokyo* **23**, 259–262 (1970).
14. Marks, N., Grynbaum, A. & Lajtha, A. Pentapeptide (pepstatin) inhibition of brain acid proteinase. *Science* **181**, 949–951 (1973).
15. Katoh, I., Yasunaga, T., Ikawa, Y. & Yoshinaka, Y. Inhibition of retroviral protease activity by an aspartyl proteinase inhibitor. *Nature* **329**, 654–656 (1987).
16. Rinehart, K. L. Jr. et al. Didemnins: antiviral and antitumor depsipeptides from a Caribbean tunicate. *Science* **212**, 933–935 (1981).
17. Wang, L., Xie, J. & Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 225–249 (2006).
18. Dedkova, L. M., Fahmi, N. E., Golovine, S. Y. & Hecht, S. M. Enhanced D-amino acid incorporation into protein by modified ribosomes. *J. Am. Chem. Soc.* **125**, 6616–6617 (2003).
19. Fujino, T., Goto, Y., Suga, H. & Murakami, H. Reevaluation of the D-amino acid compatibility with the elongation event in translation. *J. Am. Chem. Soc.* **135**, 1830–1837 (2013).
20. Fujino, T., Goto, Y., Suga, H. & Murakami, H. Ribosomal synthesis of peptides with multiple β-amino acids. *J. Am. Chem. Soc.* **138**, 1962–1969 (2016).
21. Guillen Schlippe, Y. V., Hartman, M. C. T., Josephson, K. & Szostak, J. W. *In vitro* selection of highly modified cyclic peptides that act as tight binding inhibitors. *J. Am. Chem. Soc.* **134**, 10469–10477 (2012).
22. Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu. Rev. Biochem.* **83**, 727–752 (2014).
23. Yamagishi, Y. et al. Natural product-like macrocyclic *N*-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol.* **18**, 1562–1570 (2011).
24. Imanishi, S. et al. *In vitro* selection of macrocyclic D/L-hybrid peptides against human EGFR. *J. Am. Chem. Soc.* **143**, 5680–5684 (2021).
25. Katoh, T., Sengoku, T., Hirata, K., Ogata, K. & Suga, H. Ribosomal synthesis and de novo discovery of bioactive foldamer peptides containing cyclic β-amino acids. *Nat. Chem.* **12**, 1081–1088 (2020).
26. Katoh, T. & Suga, H. *In vitro* selection of foldamer-like macrocyclic peptides containing 2-aminobenzoic acid and 3-aminothiophene-2-carboxylic acid. *J. Am. Chem. Soc.* **144**, 2069–2072 (2022).
27. Trobro, S. & Åqvist, J. Mechanism of peptide bond synthesis on the ribosome. *Proc. Natl. Acad. Sci. USA* **102**, 12395–12400 (2005).
28. Ohshiro, Y. et al. Ribosomal synthesis of backbone-macrocyclic peptides containing γ-amino acids. *ChemBioChem* **12**, 1183–1187 (2011).
29. Katoh, T. & Suga, H. Ribosomal elongation of cyclic γ-amino acids using a reprogrammed genetic code. *J. Am. Chem. Soc.* **142**, 4965–4969 (2020).
30. Dale, T., Sanderson, L. E. & Uhlenbeck, O. C. The affinity of elongation factor Tu for an aminoacyl-tRNA is modulated by the esterified amino acid. *Biochemistry* **43**, 6159–6166 (2004).
31. Katoh, T., Wohlgemuth, I., Nagano, M., Rodnina, M. V. & Suga, H. Essential structural elements in $tRNA^{pro}$ for EF-P-mediated alleviation of translation stalling. *Nat. Commun.* **7**, 11657 (2016).
32. Katoh, T., Iwane, Y. & Suga, H. Logical engineering of D-arm and T-stem of tRNA that enhances D-amino acid incorporation. *Nucleic Acids Res.* **45**, 12601–12610 (2017).
33. Murakami, H., Ohta, A., Ashigai, H. & Suga, H. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nat. Methods* **3**, 357–359 (2006).

Content courtesy of Springer Nature, terms of use apply. Rights reserved

34. Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779–790 (2011).
35. Lee, J., Schwarz, K. J., Kim, D. S., Moore, J. S. & Jewett, M. C. Ribosome-mediated polymerization of long chain carbon and cyclic amino acids into peptides in vitro. *Nat. Commun.* **11**, 4304 (2020).
36. Adaligil, E., Song, A., Cunningham, C. N. & Fairbrother, W. J. Ribosomal synthesis of macrocyclic peptides with linear γ^4- and β-hydroxy-γ^4-amino acids. *ACS Chem. Biol.* **16**, 1325–1331 (2021).
37. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
38. Owen, D. R. et al. An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586–1593 (2021).
39. Malla, T. R. et al. Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 M^{pro} inhibitors. *Chem. Commun.* **57**, 1430–1433 (2021).
40. Chan, H. T. H. et al. Discovery of SARS-CoV-2 M^{pro} peptide inhibitors from modelling substrate and ligand binding. *Chem. Sci.* **12**, 13686–13703 (2021).
41. Wilmouth, R. C. et al. Structure of a specific acyl-enzyme complex formed between β-casomorphin-7 and porcine pancreatic elastase. *Nat. Struct. Biol.* **4**, 456–462 (1997).
42. Zhang, L. et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).
43. Johansen-Leete, J. et al. Antiviral cyclic peptides targeting the main protease of SARS-CoV-2. *Chem. Sci.* **13**, 3826–3836 (2022).
44. Kreutzer, A. G. et al. A cyclic peptide inhibitor of the SARS-CoV-2 main protease. *Eur. J. Med. Chem.* **221**, 113530 (2021).
45. Ullrich, S. et al. Challenges of short substrate analogues as SARS-CoV-2 main protease inhibitors. *Bioorg. Med. Chem. Lett.* **50**, 128333 (2021).
46. Hoffman, R. L. et al. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J. Med. Chem.* **63**, 12725–12747 (2020).
47. Citarella, A., Scala, A., Piperno, A. & Micale, N. SARS-CoV-2 M^{pro}: a potential target for peptidomimetics and small-molecule inhibitors. *Biomolecules* **11**, 607 (2021).
48. Jin, Z. et al. Structure of M^{pro} from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
49. Dai, W. et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **368**, 1331–1335 (2020).
50. Dragovich, P. S. et al. Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 4. Incorporation of P1 lactam moieties as L-glutamine replacements. *J. Med. Chem.* **42**, 1213–1224 (1999).
51. Unoh, Y. et al. Discovery of S-217622, a noncovalent oral SARS-CoV-2 3CL protease inhibitor clinical candidate for treating COVID-19. *J. Med. Chem.* **65**, 6499–6512 (2022).
52. Zhang, C. H. et al. Potent noncovalent inhibitors of the main protease of SARS-CoV-2 from molecular sculpting of the drug perampanel guided by free energy perturbation calculations. *ACS Cent. Sci.* **7**, 467–475 (2021).
53. Rossetti, G. G. et al. Non-covalent SARS-CoV-2 M^{pro} inhibitors developed from in silico screen hits. *Sci. Rep.* **12**, 2505 (2022).
54. Glaser, J. et al. Hit expansion of a noncovalent SARS-CoV-2 main protease inhibitor. *ACS Pharmacol. Transl. Sci.* **5**, 255–265 (2022).
55. Cardozo, A. K. et al. Cell-permeable peptides induce dose- and length-dependent cytotoxic effects. *Biochim. Biophys. Acta Biomembr.* **1768**, 2222–2234 (2007).
56. El-Andaloussi, S., Järver, P., Johansson, Henrik, J. & Langel, Ü. Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study. *Biochem. J.* **407**, 285–292 (2007).
57. Kawamura, A. et al. Highly selective inhibition of histone demethylases by de novo macrocyclic peptides. *Nat. Commun.* **8**, 14773 (2017).
58. Rogers, J. M. et al. *In vivo* modulation of ubiquitin chains by *N*-methylated non-proteinogenic cyclic peptides. *RSC Chem. Biol.* **2**, 513–522 (2021).
59. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic peptides as drug candidates: recent progress and remaining challenges. *J. Am. Chem. Soc.* **141**, 4167–4181 (2019).
60. Zhao, Y. et al. Crystal structure of SARS-CoV-2 main protease in complex with protease inhibitor PF-07321332. *Protein Cell* **13**, 689–693 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Methods

Preparation of flexizymes and tRNAs

Flexizymes (dFx and eFx) and tRNAs for charging cγAAs, cβAAs and D-amino acids were transcribed in vitro using the T7 RNA polymerase from the corresponding template DNAs and prepared by extension and PCR (Supplementary Table 3 for primer sequences). PCR products were purified by phenol/chloroform extraction and ethanol precipitation. The transcription reaction was then carried out using the following mixture: 40 mM Tris-HCl buffer (pH 8.0), 22.5 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM spermidine, 0.01% Triton X-100, 3.75 mM NTP mix, 0.04 U μl^{-1} RNasin RNase inhibitor (Promega, N2615) and 120 nM T7 RNA polymerase (37 °C, 16 h). For transcription of tRNAs, 5 mM GMP was added to the above solution to introduce a monophosphate at the 5′ end of tRNAs. The resulting RNA transcripts were then treated with RQ1 DNase (Promega, M6101) at 37 °C for 30 min and purified by 12% (flexizymes) or 8% (tRNAs) denaturing polyacrylamide gel electrophoresis (PAGE) containing 6 M urea.

Preparation of aminoacyl-tRNAs

Preparation of aminoacyl-tRNA employed the flexizyme method[25,33,29]. The cγAAs, cβAAs and D-cysteine (c) were pre-activated as their 3,5-dinitrobenzyl esters; *N*-chloroacetyl-D-tyrosine (^{ClAc}y) was activated as its cyanomethyl ester. The activated amino acids were charged onto the respective tRNAs using flexizymes (dFx for 3,5-dinitrobenzyl ester or eFx for cyanomethyl ester). Aminoacylation was carried out at 4 °C for 16 h for cγAA and cβAA, 6 h for c or 2 h for ^{ClAc}y in the following mixture: 600 mM $MgCl_2$, 20% DMSO solvent, 25 μM dFx or eFx, 25 μM tRNA and 5 mM activated amino acid. The reaction pH was adjusted by bicine-KOH (pH 8.7) for cγAAs and cβAAs, or HEPES-KOH buffer (pH 7.5) for c and ^{ClAc}y. The reaction was stopped by addition of ×4 volume of 0.3 M sodium acetate (pH 5.2) and ×10 volume of ethanol. The resulting aminoacyl-tRNAs were purified by ethanol precipitation.

Translation of peptides

Ribosomal synthesis of the model peptides and the peptide library was carried out using the modified FIT system (37 °C, 40 min)[32,34], which employs 3 μM initiation factor 2 (IF2), 20 μM EF-Tu, 5 μM EF-P and 0.1 μM elongation factor G (EF-G), and the following components: 50 mM HEPES-KOH (pH 7.6), 100 mM KOAc, 12.3 mM $Mg(OAc)_2$, 2 mM ATP, 2 mM GTP, 1 mM CTP, 1 mM UTP, 20 mM creatine phosphate, 2 mM spermidine, 1 mM dithiothreitol, 1.5 mg ml^{-1} *E. coli* total tRNA, 1.2 μM *E. coli* ribosome, 2.7 μM IF1, 1.5 μM IF3, 0.25 μM release factor 2 (RF2), 0.17 μM RF3, 0.5 μM ribosome recycling factor (RRF), 4 μg ml^{-1} creatine kinase, 0.1 μM T7 RNA polymerase, 3 μg ml^{-1} myokinase, 0.1 μM inorganic pyrophosphatase, 0.1 μM nucleotide diphosphate kinase, L-α-amino acids (500 μM each G/H/N/P/R/T/Y, 250 μM each D/F/I/L/S/K and 125 μM A), 0.73 μM AlaRS, 0.03 μM ArgRS, 0.38 μM AsnRS, 0.13 μM AspRS, 0.09 μM GlyRS, 0.02 μM HisRS, 0.4 μM IleRS, 0.04 μM LeuRS, 0.11 μM LysRS, 0.68 μM PheRS, 0.16 μM ProRS, 0.04 μM SerRS, 0.09 μM ThrRS, 0.02 μM TyrRS, 50 μM γ^1-$tRNA^{Pro1E2}_{CUC}$, 100 μM γ^2-$tRNA^{Pro1E2}_{CAC}$, 20 μM β^1-$tRNA^{Pro1E2}_{GAC}$, 20 μM β^2-$tRNA^{GluE2}_{GCA}$, 20 μM ^{ClAc}y-$tRNA^{fMet}_{CAU}$ and 20 μM c-$tRNA^{Pro1E2}_{CAU}$. For translation of model peptides, a 0.04 μM DNA template was added to the above solution for transcription/translation coupled reactions. The DNA templates were prepared by extension and PCR (Supplementary Table 4 for the primer sequences). For the translation of the peptide library, a 1.5 μM mRNA library conjugated to puromycin linker was used.

MALDI-TOF MS of translated model peptides

The model peptides with a C-terminal FLAG-tag (DYKDDDDK) were translated and diluted with an equal volume of ×2 TBS buffer (100 mM Tris-HCl (pH 7.6), 300 mM NaCl), and then incubated with 10 μl of ANTI-FLAG M2 affinity gel (Sigma, A2220) at 25 °C for 30 min. The gel beads were washed with 50 μl TBS buffer (50 mM Tris-HCl (pH 7.6), 150 mM NaCl); the peptides were eluted from the beads by adding 20 μl of 0.2% (v/v) trifluoroacetic acid (TFA). The peptides were desalted with SPE C-tip (Nikkyo Technos) and eluted with 1.2 μl of 80% (v/v) acetonitrile and 0.5% (v/v) acetic acid solution containing 50%-saturated α-cyano-4-hydroxycinnamic acid. MALDI-TOF MS was carried out using an ultrafleXtreme instrument (Bruker Daltonics) in reflector/positive mode. Peptide calibration standard II (Bruker Daltonics, 8222570) was used for the external mass calibration.

Production of recombinant SARS-CoV-2 M^{pro}

M^{pro} was prepared as reported, and assays with it were performed exclusively using freshly purified recombinant M^{pro} solution[39,61]. Refrozen M^{pro} samples exhibited reduced activity and were not used.

RaPID selection of peptides against SARS-CoV-2 M^{pro}

The random mRNA library was ligated with a puromycin linker at the 3′ end, and then added to the above described translation mixture (Extended Data Fig. 1a, step 1). The peptide library was translated at 37 °C for 40 min in 150 μl (for the first round of selection) or 10 μl (for the second to fourth rounds) of the FIT system and incubated at 25 °C for 5 min to conjugate the translated peptide with the corresponding mRNA–puromycin (step 2). A ×0.04 volume of 500 mM ethylenediaminetetraacetic acid (EDTA; pH 8.0) was then added and incubated (37 °C, 10 min) to dissociate ribosomes from the mRNA–peptide conjugates. Reverse transcription (42 °C, 30 min) used the NNUGWGAUG.R38 primer (5′-TTTCCGCCCCCCGTCCTAACCTCCGCTACCTCCACTCA-3′) and M-MLV reverse transcriptase lacking RNase H activity (Promega, M3682; step 3). The resulting cDNA/mRNA/peptide conjugates were subjected to naked Dynabead streptavidin (Thermo Fisher, DB11206) treatment (4 °C, 15 min) three times to remove bead-binding peptides; the supernatant was then applied to M^{pro}-immobilized Dynabeads (4 °C, 15 min; step 4). The beads were washed with 100 μl ice-cold TBS-T buffer (50 mM Tris-HCl (pH 7.6), 150 mM NaCl, 0.05% (v/v) Tween 20) three times. Note that the removal of bead-binding peptides was not performed for the first selection round. Then, 100 μl of ×1 PCR buffer (10 mM Tris-HCl (pH 9.0), 50 mM KCl, 0.1% (v/v) Triton X-100, 0.25 mM dNTP, 2.5 mM $MgCl_2$, 0.25 μM T7.F53 primer (5′-GGCGTAATACGACTC ACTATAGGGTTGAACTTTAAGTAGGAGATATATCCATG-3′) and 0.25 μM NNUGWGAUG.R38 primer) was added to the beads; the cDNAs were eluted at 95 °C for 5 min and PCR amplified to make a cDNA library (step 5). To estimate the recovery rate of cDNA, 1 μl of the elute was mixed with 19 μl of ×1 PCR buffer containing SYBR Green I (Lonza, 50513) and *Taq* DNA polymerase; amounts of cDNA were quantified by real-time PCR.

Solid-phase peptide synthesis

Macrocyclic peptides were synthesized by standard Fmoc solid-phase peptide synthesis using a Syro I automated peptide synthesizer (Biotage). Fmoc-protected amino acids and coupling reagents were from Merck, Watanabe Chemical Industries or Enamine. NovaPEG Rink Amide Resin (54 mg, 25 μmol) was incubated with *N*,*N*-dimethylformamide (DMF; room temperature, 1 h). Each Fmoc-protected amino acid was coupled at 30 °C for 40 min on the resin in a DMF solution containing 0.2 M Fmoc-protected amino acid (6 equiv.), 0.2 M 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (5 equiv.), 0.2 M 1-hydroxybenzotriazole (5 equiv.) and 0.1 M *N*,*N*-diisopropylethylamine (12 equiv.). After washing the resin five times with 600 μl DMF, the Fmoc group was deprotected with 600 μl of 40% (v/v) piperidine in DMF at 30 °C for 12 min. Coupling of the Fmoc-protected amino acid and Fmoc deprotection were repeated accordingly. After automated peptide synthesis, 0.2 M chloroacetyl *N*-hydroxysuccinimide ester (8 equiv.) in *N*-methylpyrrolidone was added to the resin; the mixture was incubated at room temperature for 1 h with rotation. After washing the resin subsequently with DMF five times and with dichloromethane five times, the resin-bound peptides were treated with 2 ml of a solution of 92.5% (v/v) TFA, 2.5% (v/v) water, 2.5% triisopropylsilane and 2.5% 3,6-dioxa-1,8-octanedithiol at

Content courtesy of Springer Nature, terms of use apply. Rights reserved

room temperature for 3 h with rotation to deprotect and cleave off from the resin. The resulting linear peptides were precipitated with diethyl ether and dissolved in 10 ml of 80% (v/v) DMSO and 0.1% TFA in water. Following the addition of 200 µl of 0.5 M tris(2-carboxyethyl) phosphine and triethylamine to adjust the pH to 8, the peptide mixture was incubated with rotation at room temperature for 10 h to form a thioether bond between the N-terminal chloroacetamide and thiol group of the downstream cysteine. Macrocyclization of the peptides was confirmed by MALDI-TOF MS; the crude peptides were purified by reverse-phase high-performance LC (Shimadzu). The purities of the peptides were evaluated by UPLC (Shimadzu) using a reverse-phase column (ACQUITY UPLC BEH C18, 1.7 µm, 2.1 × 150 mm; Waters) with a linear gradient from 10% buffer B to 70% buffer B. Buffer A was water with 0.1% (v/v) TFA; buffer B was acetonitrile with 0.1% (v/v) TFA.

Binding kinetics analysis of peptides by SPR

The binding affinities of selected peptides and M^{pro} were analysed by SPR using a Biacore T200 instrument (Cytiva) at 25 °C. The composition of the running buffer was 10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.05% (v/v) Tween 20 and 0.1% (v/v) DMSO. Biotin-tagged M^{pro} was immobilized on a Biacore sensor chip CAP (Cytiva) to a surface density of 1,500–2,000 response units following the immobilization protocols of the Biotin CAPture Kit (Cytiva). The kinetic constants were determined by a single-cycle kinetics method by the injection of five different concentrations (twofold dilution series) of each peptide at a flow rate of 30 µl min^{-1}. Binding sensorgrams were fitted to the standard 1:1 interaction model and analysed using Biacore evaluation software.

Solid-phase extraction coupled to MS inhibition assays

Inhibition of M^{pro} was measured by solid-phase extraction purification coupled to MS analysis using a RapidFire 365 high-throughput sampling robot (Agilent) connected to an iFunnel Agilent 6550 accurate mass quadrupole TOF spectrometer as reported[39]. In brief, the cyclic peptides were dispensed in an 11-point, threefold dilution series across 384-well plates using an acoustic Echo Dispenser machine (LabCyte). Formic acid and DMSO were used as positive and negative controls, respectively. The assay was adapted for a reaction volume of 20 µl with a final condition of 75 nM M^{pro} and 4 µM 37-mer substrate (ALNDF-SNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2). Reactions were incubated (15 min) and then quenched by the addition of 10% (v/v) aqueous formic acid (5 µl per well). C-terminal product peptide-derived SGFRKMAF-NH_2 (10 µl per well) was employed as the internal standard; data were extracted and processed as reported.

Serum stability assays

A synthetic macrocyclic peptide (10 µM) and an internal standard peptide (10 µM) (Supplementary Fig. 6a for its structure) were mixed and incubated in human serum (Cosmo Bio, 12181201) at 37 °C for up to 100 h. We have previously shown that the standard peptide is peptidase resistant, with no degradation occurring under these experimental conditions[25]. At each time point (0, 0.5, 1, 2, 4, 8, 24, 50 and 100 h), 4 µl of the mixture was removed and quenched by adding 12 µl methanol. Following centrifugation (15,000*g*, 25 °C, 3 min), 10 µl of the supernatant was mixed with 40 µl of 1% (v/v) TFA. Following centrifugation (15,000*g*, 25 °C, 3 min), the supernatant was collected for LC/MS analysis, which used a reverse-phase column (ACQUITY UPLC BEH C18, 1.7 µm, 2.1 × 150 mm; Waters) and a Xevo G2-XS QTof system (Waters) with a linear gradient from 1% B to 60% B. Buffer A was water with 0.1% (v/v) formic acid; buffer B was acetonitrile with 0.1% (v/v) formic acid. The percentages of remaining peptides were determined by the peak area integration of the chromatograms.

Crystallization, data collection and structure solution

M^{pro} was thawed and diluted to 6 mg ml^{-1} using 20 mM HEPES (pH 7.5) and 50 mM NaCl. GM4 was diluted into the protein solution to a final concentration of 10 mM, which was incubated for two hours at room temperature prior to dispensing plates. The drop composition was 0.15 µl protein ligand solution, 0.3 µl 11% (v/v) polyethylene glycol (PEG) 4K, 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.5 and 0.05 µl M^{pro} crystal seed stock. The M^{pro} crystal seed stock was prepared by crushing M^{pro} crystals with a pipette tip; suspending them in 30% PEG 4K, 5% (v/v) DMSO and 0.1 M MES pH 6.5; and vortexing for 60 s with ~10 glass beads (1.0 mm diameter, BioSpec Products). The reservoir solution was 11% (v/v) PEG 4K, 5% (v/v) DMSO and 0.1 M MES pH 6.5. Crystals were grown by the sitting drop vapour diffusion method (20 °C) and appeared within 24 h, reaching full size within 36 h; they were harvested after one week.

Data collection and structure determination. All diffraction data were collected at 100 K with a wavelength of 0.9762 Å on beamline I03 at the Diamond Light Source. Data were processed using Dials[62] via Xia2[63] and Aimless[64] within CCP4i2 (ref. 65). The datasets were phased using Molrep[66] using the M^{pro} apo structure (Protein Data Bank 6YB7). Ligand restraints were generated using AceDRG[67]. Crystal structures were manually rebuilt in Coot and refined using Refmac[68] and PDBredo[69].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession code 7Z4S. Other data supporting this study are available in the Supplementary Information. Source data are provided with this paper.

References

61. Douangamath, A. et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat. Commun.* **11**, 5047 (2020).
62. Winter, G. et al. DIALS as a toolkit. *Protein Sci.* **31**, 232–250 (2022).
63. Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in *xia*2. *Acta Crystallogr. D* **69**, 1260–1273 (2013).
64. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).
65. Potterton, L. et al. CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **74**, 68–84 (2018).
66. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010).
67. Long, F. et al. AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr. D* **73**, 112–122 (2017).
68. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **374**, 300–321 (2003).
69. Joosten, R. P., Joosten, K., Murshudov, G. N. & Perrakis, A. *PDB_REDO*: constructive validation, more than just looking for errors. *Acta Crystallogr. D* **68**, 484–496 (2012).

Acknowledgements

This work was supported by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for JSPS Fellows (JP22J12466) to T.M.; a Grant-in-Aid for Scientific Research (B) (JP18H02080), a Grant-in-Aid for Scientific Research (A) (22H00439), and a Grant-in-Aid for Challenging Research (Pioneering) (JP21K18233) to T.K.; and a Grant-in-Aid for Specially Promoted Research (JP20H05618) to H.S. The Oxford researchers thank the COVID-19 Research Response Fund and King Abdulaziz University, Saudi Arabia, for funding. This research was funded in part by the Wellcome Trust (106244/Z/14/Z) and the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme (101003111). For

Content courtesy of Springer Nature, terms of use apply. Rights reserved

the purpose of open access, the authors have applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. We thank Diamond for beamtime through the COVID-19 dedicated call (proposal ID MX27088) and the Diamond MX group for their support and expertise. T.R.M. was supported by the Biotechnology and Biological Sciences Research Council (BB/M011224/1).

Author contributions

T.M., T.R.M., C.D.O., L.B., T.K., M.A.W., C.J.S. and H.S. wrote the manuscript. T.M., T.R.M., A.T., C.D.O., L.B., M.A.M., E.S., P.L., C.S.-D. and H.M. carried out the experiments and analysed the data. N.T., T.K., M.A.W., A.K., C.J.S. and H.S. supervised the programme.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41557-023-01205-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41557-023-01205-1.

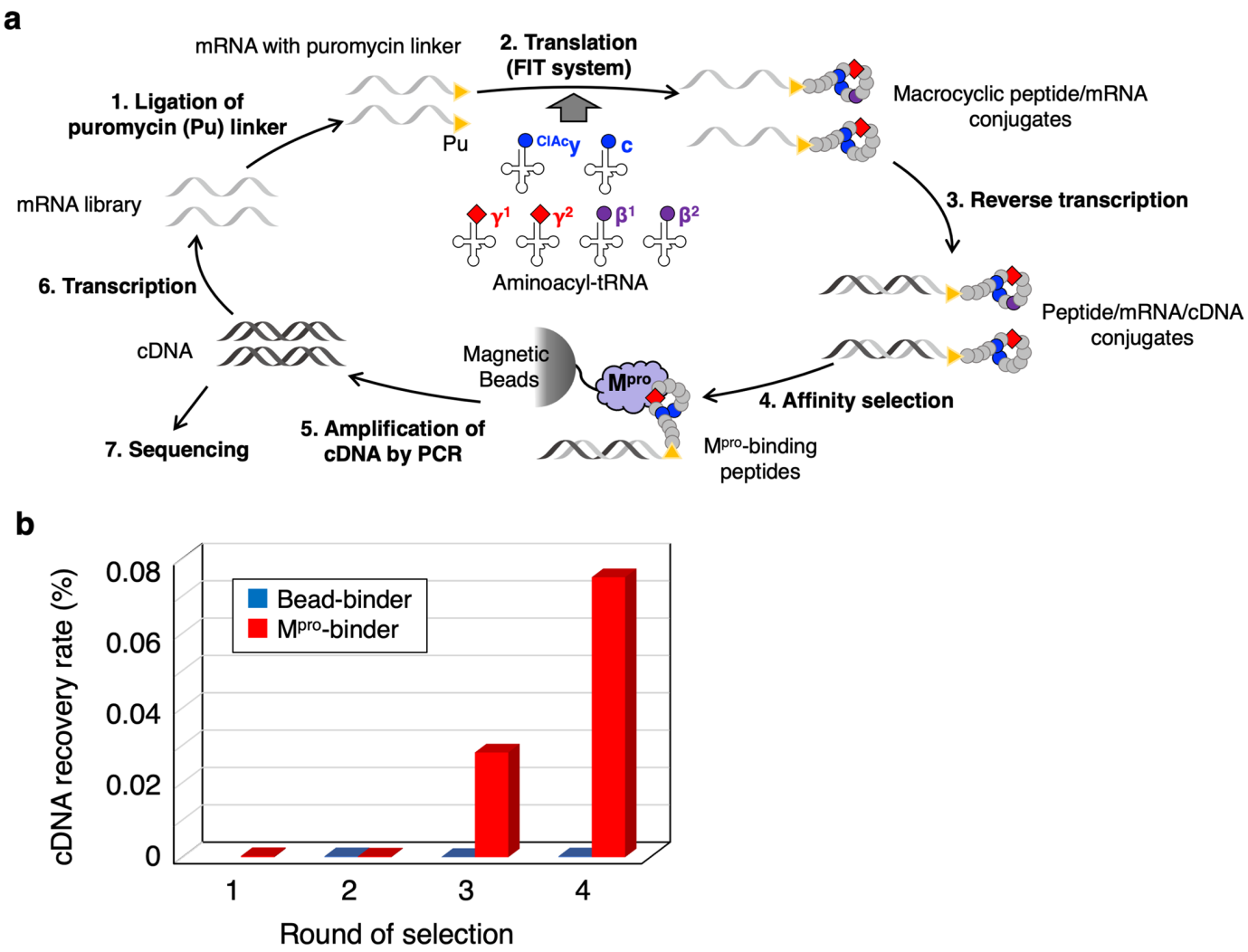
Correspondence and requests for materials should be addressed to Hiroaki Suga.

Peer review information *Nature Chemistry* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

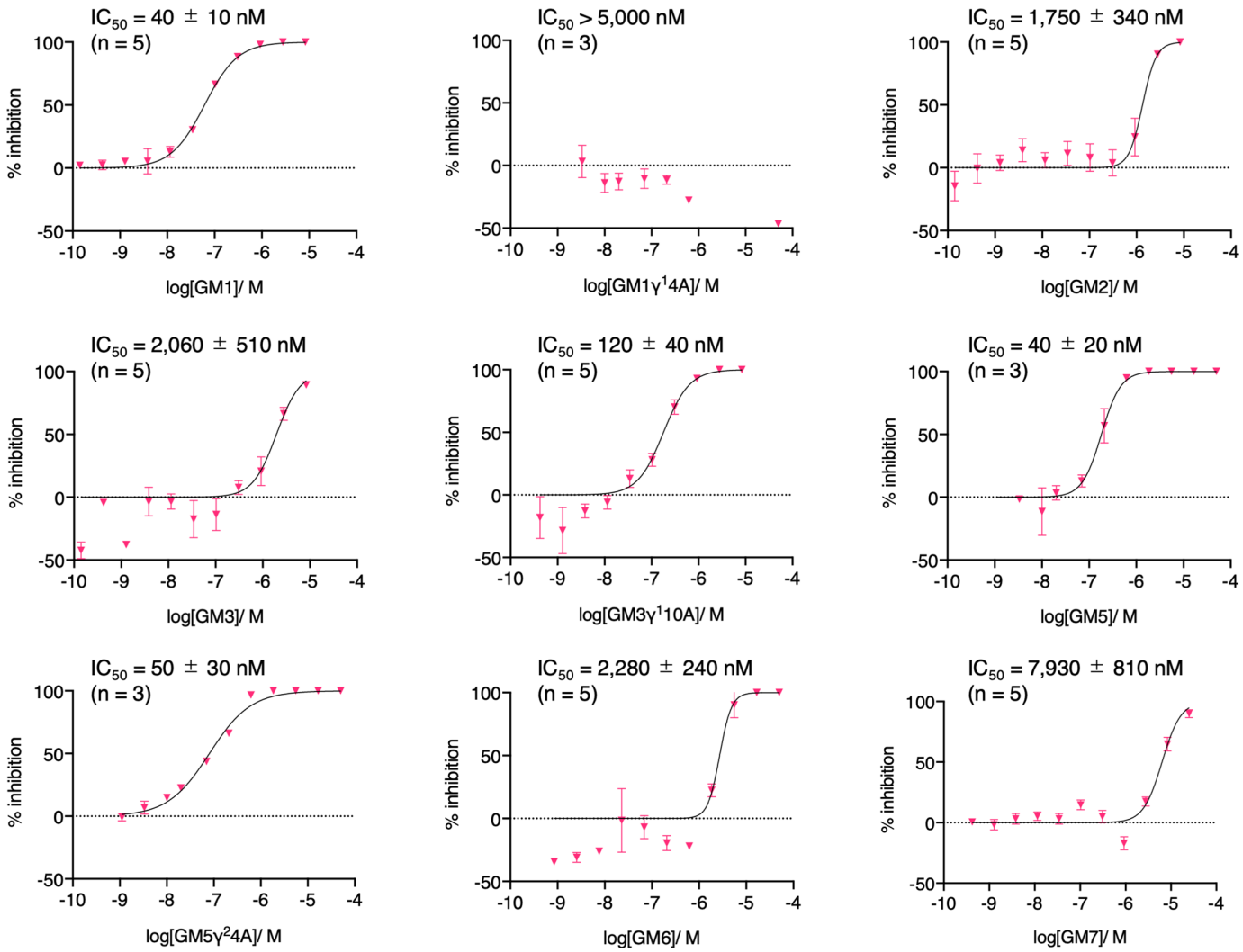
https://doi.org/10.1038/s41557-023-01205-1



Extended Data Fig. 1 | RaPID selection against M^{pro} using a macrocyclic peptide library containing cγAA. a, Schematic depiction of RaPID display. 1) Puromycin linker ligation to the 3′-end of the mRNA library. 2) Translation of peptides using the reprogrammed genetic code, followed by spontaneous macrocyclization of peptide via a thioether bond. 3) Reverse transcription of mRNA into cDNA. 4) Binding selection of peptides against M^{pro} immobilized on magnetic beads. 5) Recovery of the bound fraction and amplification of cDNA by PCR. 6) Transcription of cDNA library into mRNA library. 7) Deep sequencing analysis of cDNA library. **b**, Recovery rate of cDNA after the binding selection at each round. Red and blue bars indicate the recovery rate of M^{pro} binders and magnetic beads binders, respectively. Bead-binder selection was not performed in the first round.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

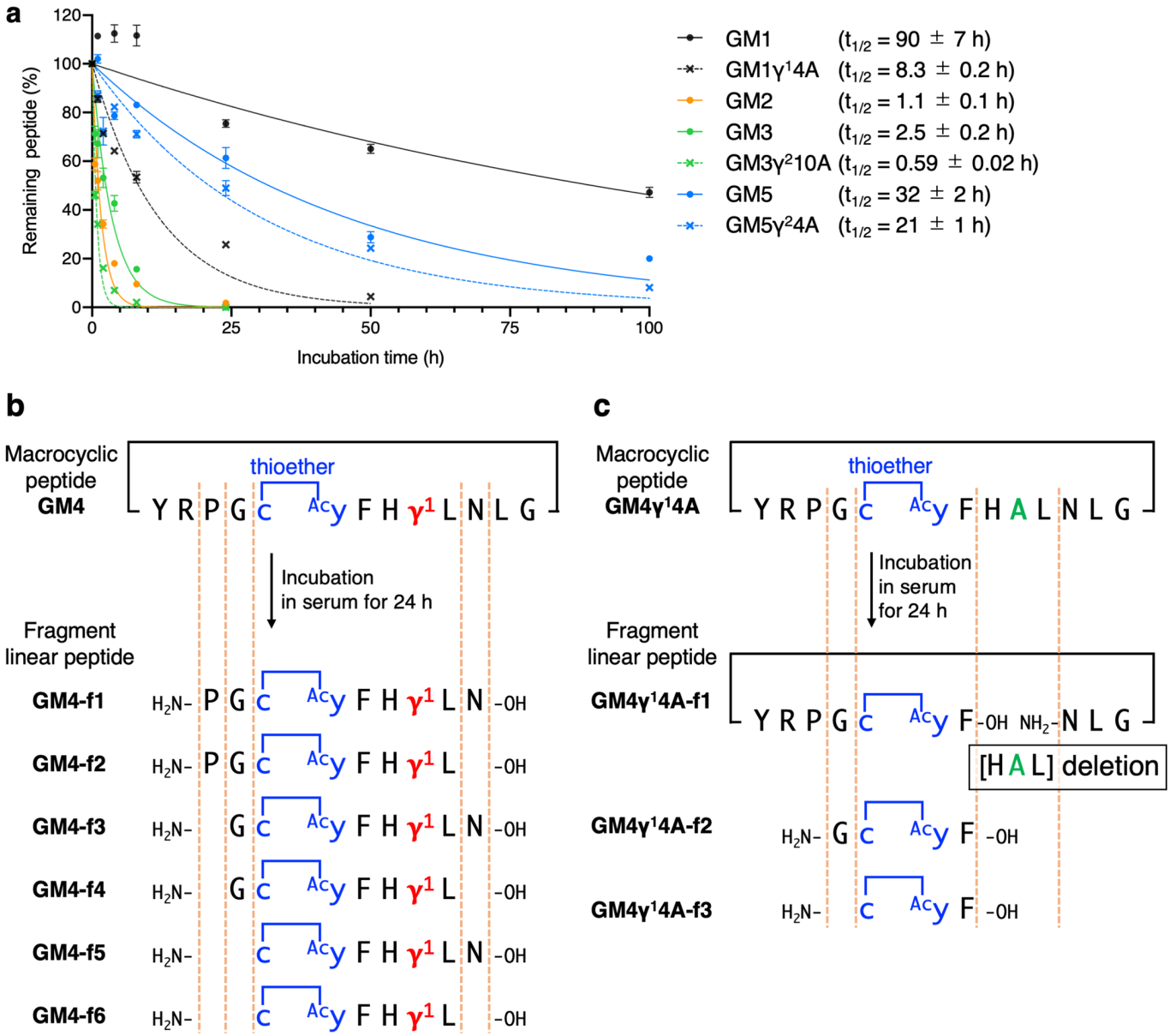
https://doi.org/10.1038/s41557-023-01205-1



Extended Data Fig. 2 | Representative dose response analyses of inhibition of M^{pro} by macrocycles. Conditions: 75 nM M^{pro}, 4 μM substrate peptide, macrocyclic peptides in 20 mM HEPES, pH 7.5, 50 mM NaCl. See Methods for details. Table 1 gives the recorded IC_{50} values as the mean of 3 or 5 independent replicates each performed in technical duplicate. Data are presented as mean values ± SD. Number of replicates is indicated above the graph. See Fig. 2 for GM4 and its variants.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

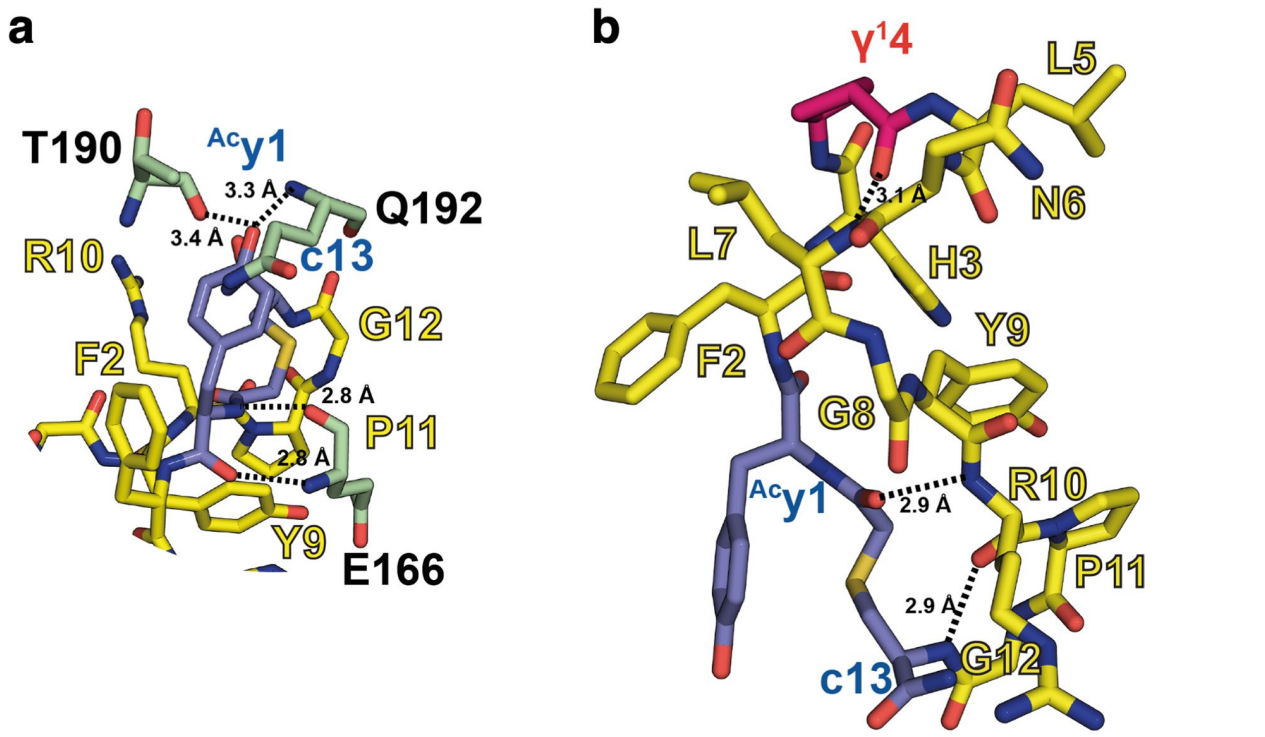
https://doi.org/10.1038/s41557-023-01205-1



Extended Data Fig. 3 | Serum stability assay of macrocyclic peptides. a, Each of GM1, GM1γ14A, GM2, GM3, GM3γ210A, GM5, and GM5γ24A were co-incubated with an internal standard peptide in human serum at 37 °C. At each time point (0, 0.5, 1, 2, 4, 8, 24, 50, and 100 h), the relative intensity of each peptide with respect to the standard peptide was estimated by LC/MS. The relative intensity at 0 h was defined as 100%. Half-lives ($t_{1/2}$) were determined by analyzing the mean of 3 technical replicates of each sample by non-linear regression using GraphPad Prism 9. Data are presented as mean values ± SD (n = 3). **b**,**c**, Fragments of GM4 (GM4-f1–6) and GM4γ14A (GM4γ14A-f1–3), respectively, after 24 h incubation in human serum as analyzed by LC/MS.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1



Extended Data Fig. 4 | Views from crystallographic studies. a, Interaction of the $y1_{GM4}$ residue at the M^{pro} active site. The phenolic sidechain of $y1_{GM4}$ forms polar interactions with T190 and Q192. The backbone NH of $y1_{GM4}$ is positioned to form H-bonds with the mainchain carbonyl and NH groups of E166, as is typical of P3 residues in peptidomimetic M^{pro} inhibitors (Ref. 40). **b**, Intramolecular hydrogen bonds in GM4; H-bonds are formed between $\gamma^1 4_{GM4}$ and $L7_{GM4}$, $^{Ac}y1_{GM4}$ and $R10_{GM4}$, and $P11_{GM4}$ and $c13_{GM4}$. H-bonds are indicated with dashed black lines; associated distances are in Ångström.

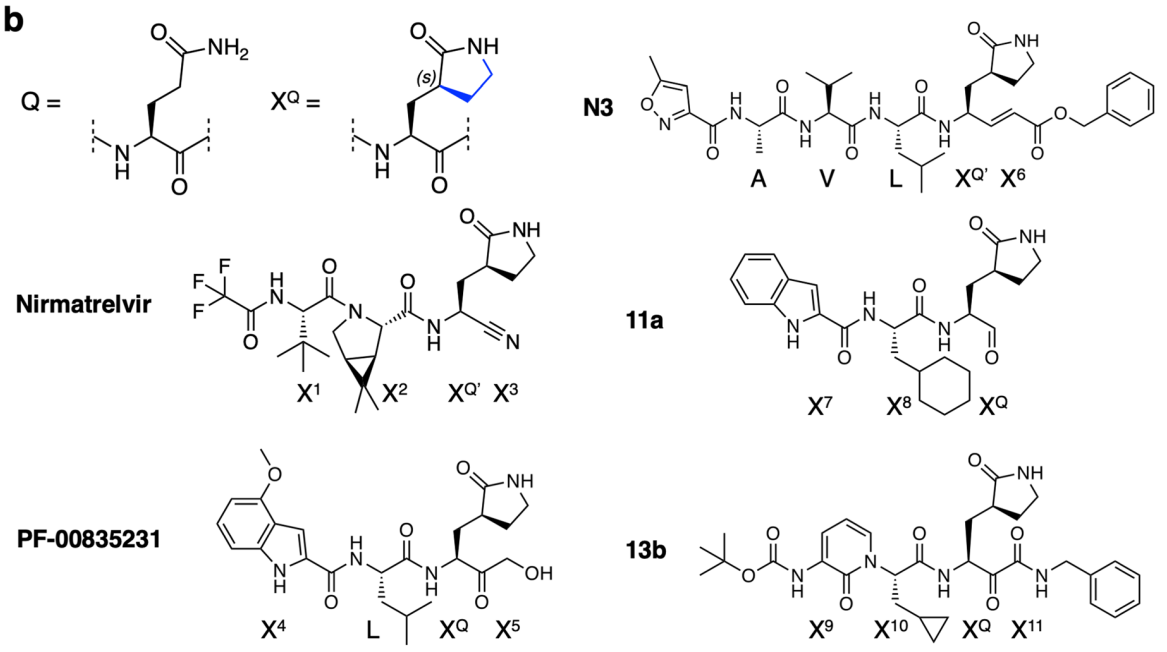
Content courtesy of Springer Nature, terms of use apply. Rights reserved

https://doi.org/10.1038/s41557-023-01205-1

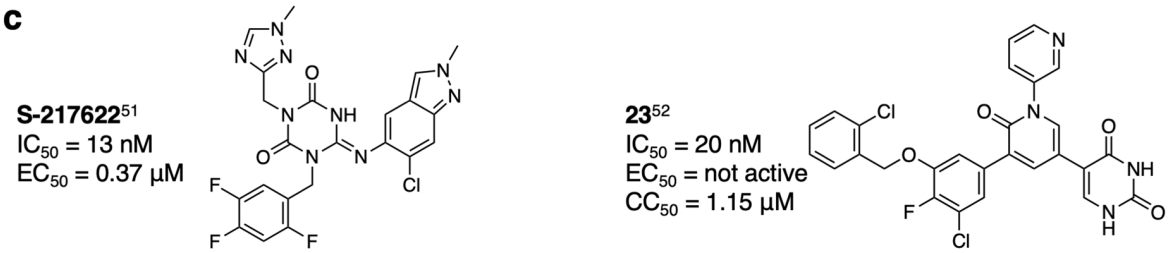
a

		P4	P3	P2	P1	P1′
Substrate		A/V/P/T	Any residue	L/F/V/M	Q	S/A/G/N
Inhibitors	Inhibitory activity					
GM4 (This work)	IC_{50} = 50 nM	c(thioether)	^{Ac}y	F	H	γ[1]
GM4γ[1]4A (This work)	IC_{50} = 10 nM	c(thioether)	^{Ac}y	F	H	A
GM4H3Q (This work)	IC_{50} = 10 nM	c(thioether)	^{Ac}y	F	Q	γ[1]
1 (Cyclic peptide)[43]	K_i = 14 nM	C(thioether)	^{Ac}y	L	Q	Y
UCI-1 (Cyclic peptide)[44]	IC_{50} = 160,000 nM			F	Q	S
p13 (Linear peptide)[40]	IC_{50} = 3,110 nM	L	T	W	Q	N
21 (Linear peptide)[45]	K_i = 57,000 nM	A	V	L	Q	S
Nirmatrelvir, (PF-07321332) (Peptidomimetic)[38]	K_i = 3.11 nM		X^1	X^2	$X^{Q'}$	X^3
PF-00835231 (Peptidomimetic)[46]	K_i = 0.27 nM		X^4	L	X^Q	X^5
N3 (Peptidomimetic)[48]	$k_{obs}/[I]$ = 11,300 $M^{-1}S^{-1}$	A	V	L	$X^{Q'}$	X^6
11a (Peptidomimetic)[49]	IC_{50} = 53 nM		X^7	X^8	X^Q	
13b (Peptidomimetic)[42]	IC_{50} = 670 nM		X^9	X^{10}	X^Q	X^{11}

b



c



Extended Data Fig. 5 | Comparison of GM4 and its variants with M^{pro} substrates and reported M^{pro} inhibitors. a, Amino acid sequences around the cleavage site of the M^{pro} substrate and inhibitors. Superscript numbers indicate references. **b**, Structures of glutamine analog X^Q and peptidomimetics. **c**, Structures and activities of non-covalent small molecule inhibitors of M^{pro}.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

nature portfolio

Corresponding author(s): Hiroaki Suga

Last updated by author(s): Jan 30, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	MALDI-TOF MS data was collected on Bruker flexControl 3.4. Next generation sequencing data was collected on Illumina MiSeq Control Software 2.4.1.3. HPLC/UPLC data was collected on Shimadzu LabSolutions 5.99. SPR data was collected on Cytiva Biacore T200 Control Software 2.0.1. LC/MS data was collected on Waters MassLynx 4.2.
Data analysis	MALDI-TOF MS data was analyzed on Bruker flexAnalysis 3.4. HPLC/UPLC data was analyzed on Shimadzu LabSolutions 5.99. SPR data was analyzed on Cytiva Biacore T200 Evaluation Software 3.0. LC/MS data was analyzed on MassLynx 4.2. Kinetic data fitting was performed using GraphPad Prism 9.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Coordinates and structure factors have been deposited in the Protein Data Bank, under accession code 7Z4S. Other data supporting this study are available in the Supplementary Information. Source data are provided with this paper. Other results are available from the corresponding author upon reasonable request.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Next generation sequencing (NGS) after selection was performed by selecting appropriate sequencing kit (MiSeq Reagent kit v3, 150-cycle) to read more than 10,000 sequences to confirm that the sequences had converged. Seven peptides with high read numbers (>200) were selected from the NGS results. Top 100 peptide sequences obtained by the NGS are listed in Supplementary Table 1.
Data exclusions	No data were excluded.
Replication	All assays were performed at least three times with the same results.
Randomization	There were no experiments for randomization. We selected peptide sequences containing γ-amino acids with high read numbers in the NGS results for subsequent assays.
Blinding	Not required for this study because there is no data selection in the experimental approach that could introduce bias.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com

Check for updates

ChemMedChem

Communications
doi.org/10.1002/cmdc.202100582

Chemistry Europe
European Chemical Societies Publishing

www.chemmedchem.org

VIP *Very Important Paper*

Structure-Activity Studies Reveal Scope for Optimisation of Ebselen-Type Inhibition of SARS-CoV-2 Main Protease

Siegfried T. D. Thun-Hohenstein[+],[a] Timothy F. Suits[+],[a] Tika R. Malla,[a] Anthony Tumber,[a] Lennart Brewitz,[a] Hani Choudhry,[b] Eidarus Salah,[a] and Christopher J. Schofield*[a]

The reactive organoselenium compound ebselen is being investigated for treatment of coronavirus disease 2019 (COVID-19) and other diseases. We report structure-activity studies on sulfur analogues of ebselen with the Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) main protease (M^{pro}), employing turnover and protein-observed mass spectrometry-based assays. The results reveal scope for optimisation of ebselen/ebselen derivative- mediated inhibition of M^{pro}, particularly with respect to improved selectivity.

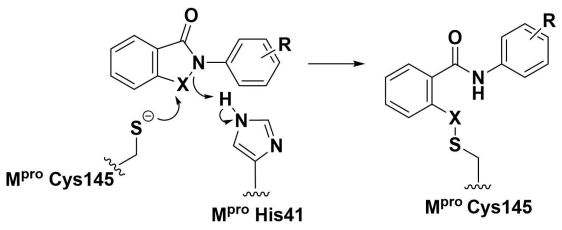


Figure 1. Initial reaction of ebselen/ebsulfur derivatives with the cysteine protease M^{pro}. X = S or Se. Ebselen, X = Se, R = H. Note reaction may also occur with non-catalytic cysteines - M^{pro} has a total of 12 cysteine residues.[7]

Ebselen is a synthetic organoselenium compound which has been under investigation for the treatment of diseases including stroke, hearing loss, and neurodegenerative disorders.[1–3] Ebselen has also been proposed as a treatment for bacterial, viral, and fungal infections.[4–6] Recently, ebselen has attracted attention because of its potential to treat COVID-19, as shown in cell-based assays and by its ability to inhibit the SARS-CoV-2 main protease (M^{pro}), the action of which is essential for the virus life cycle.[7] Ebselen is a potent inhibitor of M^{pro}, with an IC_{50} of <100 nM.[7] It is proposed to inhibit M^{pro} via covalent reaction of its N–Se bond, in particular with the M^{pro} nucleophilic active site cysteine (Cys145).

Crystallographic studies on the inhibition of catalytic cysteine enzymes, such as Ldt_{Mt2}, by ebselen reveal that an active site nucleophilic cysteine can react with the inhibitor N–Se group to form a catalytically inactive thioselenide, without other fragmentation of the inhibitor.[8] As M^{pro} has a catalytic cysteine, it is possible that a similar process is responsible for the inhibition of M^{pro} (Figure 1). However, a crystal structure with intact ebselen bound to M^{pro} has not been reported and there is crystallographic evidence that ebselen can react with M^{pro} to transfer a selenium atom to the active site cysteine.[9] Mass spectrometric (MS) studies indicate that ebselen can react with multiple nucleophilic residues on M^{pro}.[10] The reactivity of the N–Se ebselen bond is likely important in all its reported biological activities, raising questions about its selectivity and potential toxicity.[11–13]

Here we report initial structure-activity relationship (SAR) studies on the inhibition of M^{pro} by derivatives of ebselen with a sulfur substituted for the selenium (ebsulfur compounds).[14] The MS analyses reveal ebsulfur reacts more than once with M^{pro}, but that some ebsulfur derivatives retain potency against M^{pro} while showing reduced rates of reaction with non-active site nucleophilic residues. Together with other studies reported during the course of our work,[14–15] the results reveal the potential for improvement of ebselen-type inhibition of M^{pro}, perhaps most importantly in terms of selectivity.

The ebsulfur derivatives were synthesised with the aim of investigating their potency and selectivity for the active site Cys145 of M^{pro}. The changes made to ebselen included replacing the selenium for a sulfur atom to decrease reactivity, varying the *N*-linked sidechain, and adding substituents at the C6 and C8 positions (see Table 1) to alter the fit at the M^{pro} active site. The ebsulfur analogues were synthesised (Figure 2) via amide formation between the appropriate acyl chloride and primary amine to give amides **13**–**29**. Subsequent copper-mediated oxidative sulfur insertion and ring closure as described by Bhakuni et al.[16] gave the cyclised products **3a**–**c** and **3e**–**11**. The bromide substituent of **3c** enabled functionalization by Suzuki-Miyaura cross-coupling to give **3d**. In one case cyclisation of **29** gave the expected product **12i** as well as the oxidised product **12**, after purification.

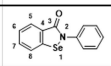
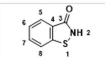
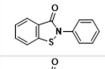
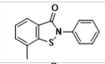
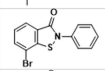
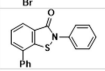
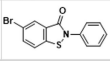
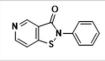
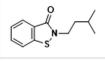
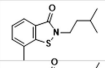
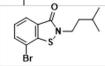
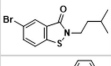
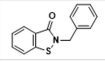
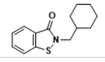
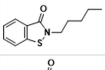
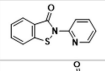
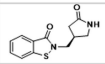
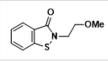
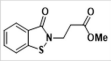
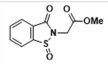
[a] *S. T. D. Thun-Hohenstein,[+] T. F. Suits,[+] T. R. Malla, Dr. A. Tumber, Dr. L. Brewitz, E. Salah, Prof. C. J. Schofield*
Department of Chemistry, University of Oxford
Chemistry Research Laboratory and the Ineos Oxford Institute for Antimicrobial Research
12 Mansfield Road, Oxford, OX1 3TA (UK)
E-mail: christopher.schofield@chem.ox.ac.uk
[b] *Dr. H. Choudhry*
Department of Biochemistry
Center for Artificial Intelligence in Precision Medicines
King Abdulaziz University, Jeddah (Saudi Arabia)
[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cmdc.202100582

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH

Table 1. Ebsulfur derivative inhibition and reactivity with M^{pro} via FRET and SPE-MS assay. See Supporting Information for assay conditions. Note the extent and the number of covalent reactions is likely condition/concentration dependent. See figures S1, S2 and S3 for dose response curves and corresponding covalent modification studies.

Compound	Structure	Apparent max. no. of reactions / most dominant no. of reactions	FRET IC_{50} (μM) / pIC_{50}	SPE-MS IC_{50} (μM) / pIC_{50}	FRET pIC_{50} / apparent max. no. of reactions
1		12/12	0.04 ± 0.01/ 7.40 ± 0.03	0.26 ± 0.02/ 7.42 ± 0.15	0.6
2		6/2	1.51 ± 0.09/ 5.82 ± 0.03	0.99 ± 0.15/ 6.00 ± 0.06	1.0
3a		11/11	0.13 ± 0.04/ 6.89 ± 0.14	0.34 ± 0.06/ 6.47 ± 0.07	0.6
3b		1/0	>100	>100	N/A
3c		0	>100	10.88 ± 0.82/ 6.28 ± 0.04	N/A
3d		0	>100	>100	N/A
3e		2/1	0.03 ± 0.01/ 7.51 ± 0.10	0.34 ± 0.01/ 6.47 ± 0.02	3.8
3f		9/5	0.48 ± 0.09/ 6.32 ± 0.08	0.43 ± 0.01/ 6.37 ± 0.01	0.7
4a		6/4	0.32 ± 0.04/ 6.50 ± 0.06	0.42 ± 0.03/ 6.38 ± 0.04	1.1
4b		0	>100	>100	N/A
4c		0	94.5 ± 8.22/ 4.03 ± 0.04	>100	N/A
4d		2/0	0.21 ± 0.13/ 6.68 ± 0.08	0.43± 0.01/ 6.47 ± 0.02	3.3
5		6/4	0.19 ± 0.04/ 6.75 ± 0.09	0.43 ± 0.03/ 6.36 ± 0.03	1.1
6		6/4	1.00 ± 0.5/ 6.03 ± 0.23	0.53± 0.05/ 6.28 ± 0.04	1.0
7		5/4	0.41 ± 0.12/ 6.39 ± 0.13	0.47 ± 0.00/ 6.33 ± 0.00	1.3
8		6/2	0.22 ± 0.10/ 6.67± 0.21	0.79 ± 0.09/ 6.10 ± 0.05	1.1
9		3/1	0.51± 0.02/ 7.51 ± 0.10	2.55 ± 0.57/ 5.60 ± 0.1	2.1
10		5/2	0.95 ± 0.04/ 6.11 ± 0.10	0.78 ± 0.18/ 7.51 ± 0.10	1.2
11		6/2	0.25 ± 0.02/ 6.61 ± 0.04	0.39 ± 0.05/ 6.41 ± 0.05	1.1
12		0	2.83 ± 0.34/ 5.55 ± 0.05	20.12 ± 1.12/ 4.70 ± 0.02	N/A

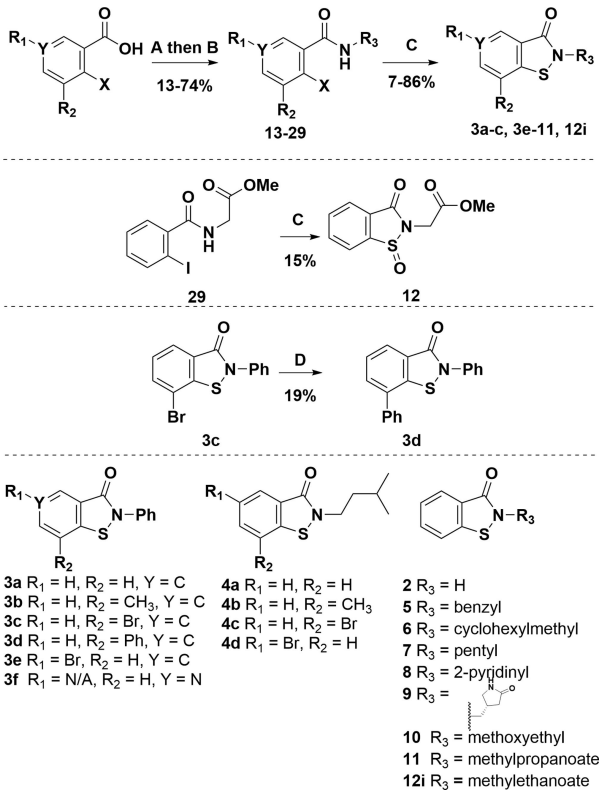


Figure 2. Synthesis of ebsulfur derivatives **3a–f**, **4a–d**, **5–12**. Reagents and conditions: **A**: oxalyl chloride, CH_2Cl_2, DMF, 0 C→rt. **B**: Et_3N, R_3NH_2, CH_2Cl_2, rt. **C**: CuI, 1,10-phenanthroline, S_8, K_2CO_3, DMF, 110 C. **D**: phenylboronic acid, K_2CO_3, $Pd(dppf)Cl_2$, 1,4-dioxane, 80 C.

The ebsulfur compounds were initially screened for M^{pro} inhibition activity using a fluorescence-based turnover (FRET) assay and a solid-phase extraction coupled to mass spectrometry (SPE-MS) assay, and for selectivity in terms of reaction with the catalytic Cys145 of M^{pro} by protein-observed mass spectrometry (POMS) under denaturing conditions.[10] M^{pro} contains a total of twelve cysteines,[7] hence the number of covalent reactions observed with M^{pro} gives potential insight into the degree of active site selectivity of the ebsulfur compounds.

The turnover assay results (Table 1) show that the direct ebsulfur analogue (**3a**) of ebselen (**1**) has a comparable IC_{50} to ebselen (**1**), confirming that an S–N bond is sufficiently reactive to inhibit M^{pro} at nanomolar concentrations. The unsubstituted ebsulfur benzoisothiazolinone (BIT) core (**2**) was much less active (IC_{50} ~ 1.6 μM). Ebsulfur derivatives **3a**, **4a**, **5–7**, varying solely in the *N*-linked hydrophobic sidechain, were synthesised since it was envisaged that these sidechains might occupy the same binding pocket as the sidechain of the P1′ residue of M^{pro} substrates, as anticipated based on the structures of substrate-like inhibitors.[7] These compounds were the most potent M^{pro} inhibitors described here, i.e. **3a** (*N*-phenyl) ~0.13 μM; **4a** (isopentyl) ~0.32 μM; **5** (*N*-benzyl) ~0.19 μM; **6** (*N*-cyclohexylmethyl) ~1.00 μM; and **7** (n-pentyl) ~0.41 μM. The variations in IC_{50}s for the compounds, together with the results of Sun *et al.*[14] show clear scope for further optimisation of the *N*-substituent of the ebsulfur derivatives (see supplementary Figures S1 and S2).

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 4, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100582 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

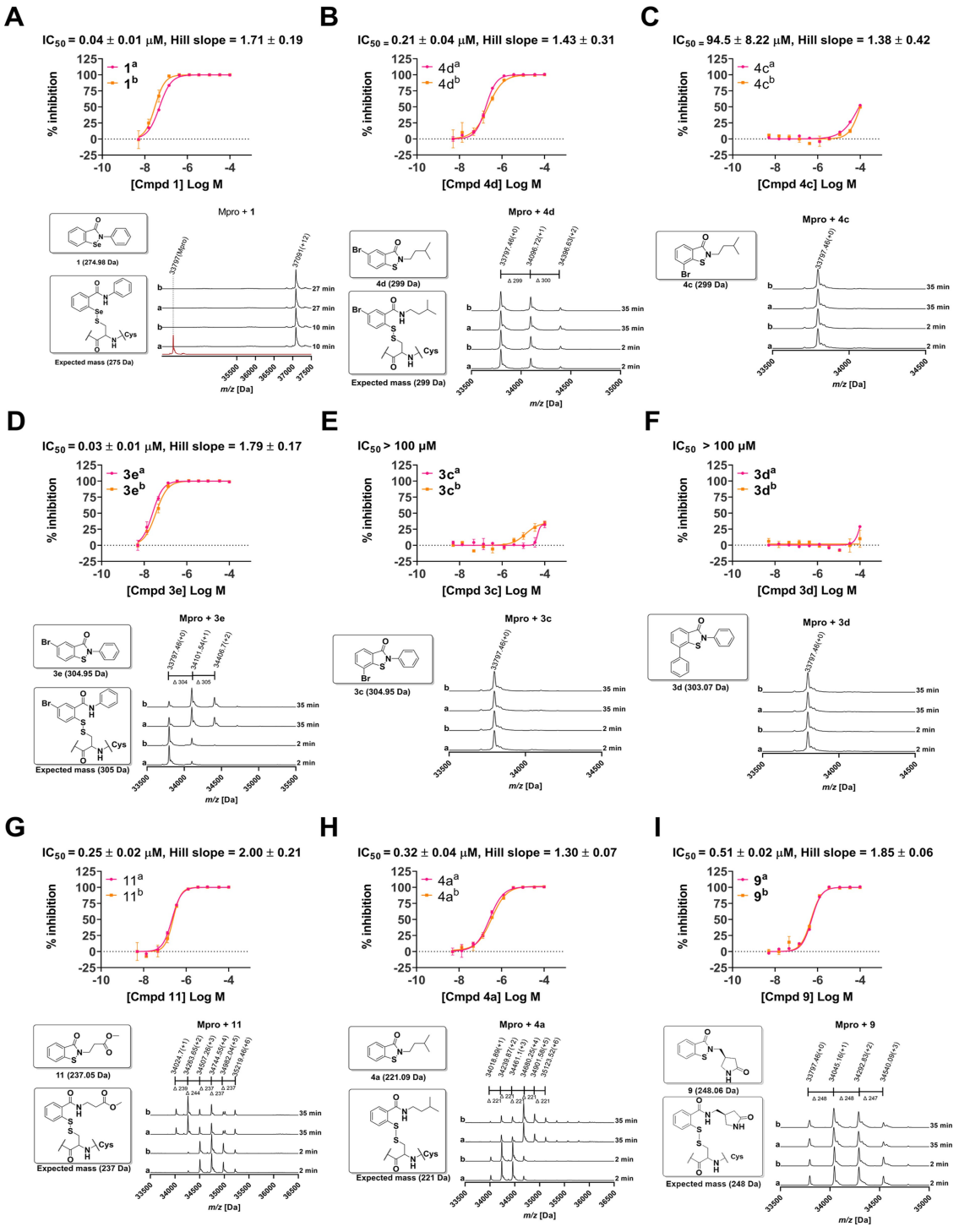


Figure 3. Dose response curves obtained using FRET assay and protein observed mass spectrometry. **A**) 1 (Ebselen) inhibits M^{pro} and reacts at multiple sites. **B**) 4d inhibits M^{pro}, but reacts covalently at less sites than ebselen. **C**) 4c (an isomer of 4d) apparently does not inhibit or react with M^{pro}. **D**) 3e manifests good inhibition and apparently reacts twice, but predominantly once. **E**) 3c apparently does not inhibit or react with M^{pro}. **F**) 3d shows no inhibition and does not react with M^{pro}, potentially due to steric hindrance by its C-8 phenyl group. **G**) 11 inhibits and reacts. **H**) 4a inhibits and reacts, predominantly 4 times. **I**) 9 inhibits M^{pro} and apparently reacts with up to 3 times, but predominantly once. Conditions: 40 μM compound, 2 μM M^{pro}, 20 mM HEPES, pH 7.5, 50 mM NaCl. ([a]) and ([b]) represent technical duplicates for POMS studies. See Supporting Information for assay conditions.

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 4, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100582 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18607187, 2022, 4, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100582 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Several compounds with more polar *N*-substituents (**8**–**11**) were synthesised to increase solubility and/or selectivity, including **9**, the γ-lactam of which was envisaged may bind in an analogous manner to the P1 glutamine sidechain in M^{pro} substrates.[7] None of these compounds were as potent as **1** or **3a**, perhaps reflecting the hydrophobic nature of regions of the M^{pro} active site, including part of the P1 residue binding subsite.[7]

Increasing the oxidation state of the sulfur substantially increased the IC_{50} (**12**, ~2.9 μM cf. **11** 0.25 μM), suggesting steric hindrance may be one factor in determining reactivity (Figure 3). We thus made **3b–d**, **4b,c** with methyl-, phenyl-, or bromo-derivatives at the C8 position.

Such derivatisation substantially reduced potency: **4c** has an IC_{50} of ~94.5 μM, and **3b–d**, and **4b** have IC_{50}s > 100 μM. Conversely, C6-substitution had the opposite effect; addition of a bromine at the C6 position of **3a** and **4a**, to give **3e** and **4d**, improved potency relative to values for the unsubstituted compounds (**3e**, ~0.03 μM cf. **3a** ~0.13 μM, and **4d**, ~0.21 μM cf. **4a**, ~0.32 μM). Introduction of a nitrogen at the C6 position to give the pyridine (**3f**) showed a slight decrease in potency (IC_{50} ~0.5 μM).

It is important to note that it cannot be assumed that inhibition is mediated via reaction with Cys145 alone and derivatisation may alter the patterns of reactivity with the 12 M^{pro} cysteines. We thus carried out protein-observed mass spectrometry (MS) under denaturing conditions to compare the number of (irreversible) covalent reactions by the ebsulfur derivatives. Although we did not carry out time-course studies (except for ebselen and ebsulfur[10]), and the nature and the extent of the reactions will likely vary depending on the conditions, the results give initial insight into the relative reactivity of the compounds.

The results show that there is a general, but clearly imperfect, correlation between the apparent extent of covalent reaction and potency, consistent with the proposed mode of reaction of the ebsulfur derivatives with cysteines to form disulfide bonds (Figure 1). However, it should be noted that alternative modes of covalent reaction are possible, e.g. via reaction of the ebsulfur derivative carbonyl group. Secondly, introduction of substituents at the C6 and C8 positions of ebsulfur decreases the number of reactions. Thus the presence of bromo- or phenyl- substituents at C8 apparently completely prevents irreversible covalent reaction (**3c**, **3d**, **4c**). C8 methyl substitution shows a reduction in the apparent maximum number of observed reactions from 6 (**4a**) to 0 (**4b**) for the *N*-isopentyl sidechain, and from 11 (**3a**) to 1 (**3b**) for the *N*-phenyl sidechain; increasing the oxidation state of the sulfur also apparently eliminates observed covalent reaction (**12**). Importantly, the presence of a more polar *N*-substituent (compare **9**, **11**/1 and 2 reactions, with **6**, **7**/both 4 reactions) or C6 bromine (compare **3e**/1 reaction with **3a**/11 reactions, and **4d**/no reaction with **4a**/4 reactions) reduces the dominant number of observed reactions. The combined IC_{50} and MS results for **9**, **4d**, and, particularly, for **3e**, are notable as they suggest that optimised substitution of the ebsulfur core to achieve enhanced selectivity in terms of its reaction with M^{pro} should be possible. We suggest that optimising the pIC_{50}/number of reactions observed by MS as a useful, readily obtainable parameter for initial optimisation of reaction selectivity (Table 1).

The reactivity of ebselen or ebselen/ebsulfur derivatives with cysteines, and maybe other biomolecules, means that there are likely substantial challenges in developing these compounds as medicines. While the activities of the ebsulfur compounds reported here do not improve on the potency of ebselen, they clearly demonstrate the potential for achieving compounds that are more selective in their reaction with M^{pro} and, by implication, other biomolecules including both proteins and small-molecule thiols such as glutathione/cysteine. This is clearly exemplified by comparing the MS results for ebselen, which, under our standard conditions, covalently reacts up to 12 times with M^{pro}, with those for **3e**, which was observed to only react twice under our conditions, yet retains substantial potency (~0.03 μM).

Under our conditions, we saw no evidence for the fragmentation of the ebsulfur derivatives or ebselen, as has been provided on the basis of crystallographic and other evidence for ebselen itself,[9] potentially in part reflecting differences in reaction conditions/methods of analysis. We suggest that if such improvements in selectivity can be achieved by simple sulfur for selenium substitution, then there is considerable scope for more sophisticated 'caging' of the reactive N–Se/S bonds of ebselen/ebsulfur type compounds. Induced-fit binding to M^{pro} or another target could result in exposure of a reactive N–S/Se bond, which is otherwise protected from nucleophilic attack by other endogenous cysteines. It is also likely the potency and selectivity of the compounds can be improved by taking advantage of interactions observed with other types of M^{pro} inhibitor, e.g. incorporation of a benzisothiazolinone core into a larger structure that mimics the binding of the natural substrates.

Acknowledgements

This work was supported by funding from the Biotechnology and Biological Research Council, Cancer Research UK, the Wellcome Trust, the Oxford Covid Development Fund and from King Abdulaziz University, Saudi Arabia. This research was funded in whole, or in part, by the Wellcome Trust [grant no. 106244/Z/14/Z]. TRM thanks BBSRC (BB/M011224/1) for doctoratal funding. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: M^{pro} inhibition · SARS-CoV-2 · COVID-19 · ebselen · ebsulfur · nucleophilic cysteine protease.

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH

[1] G. K. Azad, R. S. Tomar, *Mol. Biol. Rep.* **2014**, *41*(8), 4865–4879.
[2] M. J. Parnham, H. Sies, *Biochem. Pharmacol.* **2013**, *86*(9), 1248–1253.
[3] J. Kil, E. Lobarinas, C. Spankovich, S. K. Griffiths, P. J. Antonelli, E. D. Lynch, C. G. Le Prell, *Lancet* **2017**, *390*(10098), 969–979.
[4] H. Sies, M. J. Parnham, *Free Radical Biol. Med.* **2020**, *156*, 107–112.
[5] S. Thangamani, W. Younis, M. N. Seleem, *Sci. Rep.* **2015**, *5*, 11596.
[6] S. Thangamani, H. E. Eldesouky, H. Mohammad, P. E. Pascuzzi, L. Avramova, T. R. Hazbun, M. N. Seleem, *Biochim. Biophys. Acta Gen. Subj.* **2017**, *1861*(1), 3002–3010.
[7] Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, *Natur* **2020**, *582*(7811), 289–293.
[8] M. de Munnik, C. T. Lohans, P. A. Lang, G. W. Langley, T. R. Malla, A. Tumber, C. J. Schofield, J. Brem, *Chem. Commun.* **2019**, *55*(69), 10214–10217.
[9] K. Amporndanai, X. Meng, W. Shang, Z. Jin, M. Rogers, Y. Zhao, Z. Rao, Z.-J. Liu, H. Yang, L. Zhang, P. M. O'Neill, S. Samar Hasnain, *Nat. Commun.* **2021**, *12*(1), 3061.
[10] T. R. Malla, A. Tumber, T. John, L. Brewitz, C. Strain-Damerell, C. D. Owen, P. Lukacik, H. T. H. Chan, P. Maheswaran, E. Salah, F. Duarte, H. Yang, Z. Rao, M. A. Walsh, C. J. Schofield, *Chem. Commun.* **2021**, *57*(12), 1430–1433.
[11] F. C. Meotti, V. C. Borges, G. Zeni, J. B. T. Rocha, C. W. Nogueira, *Toxicol. Lett.* **2003**, *143*(1), 9–16.
[12] C. Jacob, W. Maret, B. L. Vallee, *Biochem. Biophys. Res. Commun.* **1998**, *248*(3), 569–573.
[13] R. Sekirnik, N. R. Rose, A. Thalhammer, P. T. Seden, J. Mecinović, C. J. Schofield, *Chem. Commun.* **2009**, 42, 6376–6378.
[14] L.-Y. Sun, C. Chen, J. Su, J.-Q. Li, Z. Jiang, H. Gao, J.-Z. Chigan, H.-H. Ding, L. Zhai, K.-W. Yang, *Bioorg. Chem.* **2021**, *112*, 104889.
[15] M. Zmudzinski, W. Rut, K. Olech, J. Granda, M. Giurg, M. Burda-Grabowska, L. Zhang, X. Sun, Z. Lv, D. Nayak, M. Kesik-Brodacka, S. K. Olsen, R. Hilgenfeld, M. Drag, *bioRxiv* **2020**. doi: https://doi.org/10.1101/2020.08.30.273979.
[16] B. S. Bhakuni, S. J. Balkrishna, A. Kumar, S. J. T. L. Kumar, *Tetrahedron Lett.* **2012**, *53*(11), 1354–1357.

Manuscript received: September 3, 2021
Accepted manuscript online: December 1, 2021
Version of record online: December 27, 2021

© 2021 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 4, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100582 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Check for updates

ChemMedChem

Full Papers
doi.org/10.1002/cmdc.202100641

Chemistry Europe
European Chemical Societies Publishing

www.chemmedchem.org

Expanding the Repertoire of Low-Molecular-Weight Pentafluorosulfanyl-Substituted Scaffolds

Arathy Jose,[a] Daniel Guest,[a] Remi LeGay,[b] Graham J. Tizzard,[c] Simon J. Coles,[c] Mariliza Derveni,[d] Edward Wright,[d] Lester Marrison,[e] Alpha A. Lee,[f] Aaron Morris,[f] Matt Robinson,[f] Frank von Delft,[g, h, i] Daren Fearon,[g] Lizbé Koekemoer,[h] Tetiana Matviuk,[j] Anthony Aimon,[g] Christopher J. Schofield,[k] Tika R. Malla,[k] Nir London,[l] Barnaby W. Greenland,[a] Mark C. Bagley,[a] John Spencer[+],*[a] and The Covid Moonshot Consortium[m]

The pentafluorosulfanyl (-SF_5) functional group is of increasing interest as a bioisostere in medicinal chemistry. A library of SF_5-containing compounds, including amide, isoxazole, and oxindole derivatives, was synthesised using a range of solution-based and solventless methods, including microwave and ball-mill techniques. The library was tested against targets including human dihydroorotate dehydrogenase (HDHODH). A subsequent focused approach led to synthesis of analogues of the clinically used disease modifying anti-rheumatic drugs (DMARDs), Teriflunomide and Leflunomide, considered for potential COVID-19 use, where SF_5 bioisostere deployment led to improved inhibition of HDHODH compared with the parent drugs. The results demonstrate the utility of the SF_5 group in medicinal chemistry.

Introduction

Small-molecule organic compounds are often used as tool compounds and chemical probes for functional validation in biological systems and as therapeutics.[1] Halogen containing fragments and higher molecular weight derivatives form a large proportion of drug-like molecules.[2–5] The pentafluorosulfanyl group is gaining popularity as a bioisostere in bioactive compounds[6,7] and in materials[8] as it is considered to be relatively stable, electronegative and lipophilic alternative to a CF_3 group. Recent years have seen a rise in SF_5-substituted compounds as direct access to aryl- and alkyl-SF_5 building blocks has been achieved.[9,10] We have recently incorporated the SF_5 group in benzodiazepine and oxindole analogues (Figure 1).[11,12] In the former example, significant activity was lost, likely due to the steric size of the SF_5 group compared to a Cl substituent.

Owing to the relative dearth of bioactive pentafluorosulfanyl containing compounds, yet commercial availability of a number of attractive building blocks, we set out to synthesise libraries of SF_5-phenyl derivatives endowed with further functionality. This work was intended to serve two purposes: the synthesis of novel small libraries to show synthetic scope and possible interest for screening programmes and the

[a] Dr. A. Jose, Dr. D. Guest, Dr. B. W. Greenland, Prof. M. C. Bagley, Prof. J. Spencer[+]
Chemistry Department, School of Life Sciences, Falmer, Brighton, BN1 9QJ (UK)
E-mail: J.Spencer@sussex.ac.uk
[b] Dr. R. LeGay
Normandie Université, Laboratoire de Chimie Moléculaire et Thioorganique LCMT UMR 6507
ENSICAEN, UNICAEN, CNRS, 6 Bd. Du Marechal Juin, 14050, Caen (France)
[c] Dr. G. J. Tizzard, Prof. S. J. Coles
National Crystallography Service, School of Chemistry
University of Southampton, Southampton, SO17 1BJ (UK)
[d] Dr. M. Derveni, Dr. E. Wright
Biochemistry, School of Life Sciences, Falmer, Brighton, BN1 9QG (UK)
[e] Dr. L. Marrison
eMolecules, 3430, Carmel Mountain Road, Suite 250, San Diego, CA 92121 (USA)
[f] A. A. Lee, A. Morris, M. Robinson
PostEra Inc., 2 Embarcadero Centre, San Franciso, CA 94111 (USA)
[g] Prof. F. von Delft, D. Fearon, A. Aimon
Diamond Light Source (DLS), Harwell Science and Innovation Campus, Didcot OX11 0DE (UK)
[h] Prof. F. von Delft, Dr. L. Koekemoer
Centre of Medicines Discovery (CMD), University of Oxford, Department of Biochemistry, Oxford OX1 3QU (UK)
[i] Prof. F. von Delft
Department of Biochemistry, University of Johannesburg, Auckland Park 2006 (South Africa)
[j] Dr. T. Matviuk
Enamine, Chervonotkatska St, 67, Kyiv, 02094 (Ukraine)
[k] Prof. C. J. Schofield, T. R. Malla
Chemistry Research Laboratory, The Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, 12 Mansfield Road, OX1 3TA, Oxford (UK)
[l] Prof. N. London
Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot, 76100 (Israel)
[m] The Covid Moonshot Consortium
Members list: https://tinyurl.com/y3r7redd
[+] Current address: Sussex Drug Discovery Centre, School of Life Sciences, Falmer, Brighton, BN1 9QG (UK)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cmdc.202100641

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

directed synthesis of SF_5-analogues as a comparison with known, electron withdrawing, CF_3, Cl– and NO_2-substituted bioactive molecules.

Results and Discussion

Initially, SF_5-containing small molecules were constructed via a simple amide bond-forming reaction, using trimethylamine in dichloromethane (conditions "a"); in general, were satisfactory (Scheme 1). Purifications varied according to the reaction, but typically involved flash silica chromatography or the use of the nucleophilic scavenger, MP-*Trisamine* (macroporous polystyrene-bound nucleophilic scavenger), to remove unreacted acid or acid chloride. Attempted reactions with 1-methylpiperazine led to poor yields, e.g., **3b**, and the free benzoic acid was detected in the crude reaction mixture, suggesting competing acid chloride hydrolysis and poor yields for the amide coupling. Hence, the coupling reagent, HATU (hexafluorophosphate azabenzotriazole tetramethyl uronium), was added (conditions "b") to mitigate for any benzoic acid formed *in situ*; this led to mainly improved yields, e.g. **4b** (83% vs 10%). The *Boc*-piperazine analogues **3d** and **4d** were successfully deprotected and further functionalised as their amide and sulphonamide derivatives **3e**, **4e** and **3f**, **4f** respectively. In total, a dozen new SF_5-containing amide analogues were made, which may have useful applications as halogen-rich screening library compounds, for example.[13]

Next, we focussed our attention on to oxindole derivatives, having recently reported SF_5-containing analogues with kinase inhibitory properties. Microwave heating was employed as in our previous work[12] and, for acetone-derived products, this acted both as a solvent and a reagent. In the synthesis of **8b**, soon after combining the oxindole, acetone and piperidine, product formation was observed, which precipitated out of solution. Nevertheless, the reaction mixture was heated under microwave irradiation until the reaction was complete and pure product was obtained, in 90% yield. Crystallisation from acetone afforded single crystals of **8b** suitable for X-ray structure determination (Scheme 2). In essence, these Knoevenagel condensations utilized a variety of solvents and nitro analogue, **10a**, was prepared to demonstrate the applicability of this reaction to other electron-withdrawing systems. (Note; to avoid overpressure, ensure the microwave tube is less than half-full).

Mechanochemical synthesis is an attractive method that is an environmentally friendly alternative to traditional routes[14] which has been applied to Knoevenagel condensations.[15,16] We attempted the synthesis of **10a** in a vibratory ball mill (VBM) in steel jars (SST) and zirconium oxide jars (ZrO_2); with and without catalyst, and for various reaction times. Zirconia jars were best

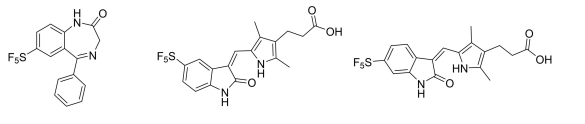
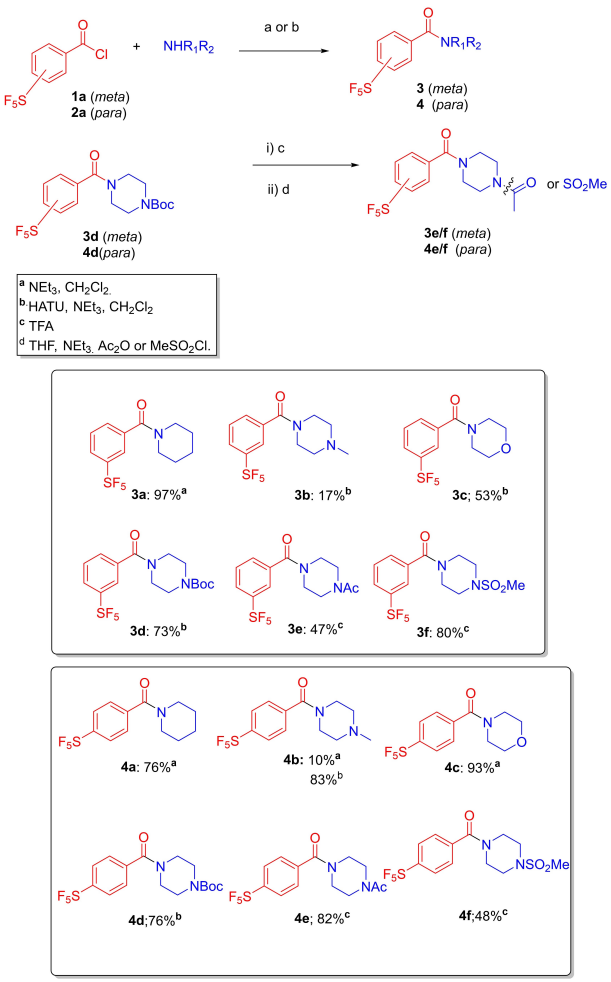
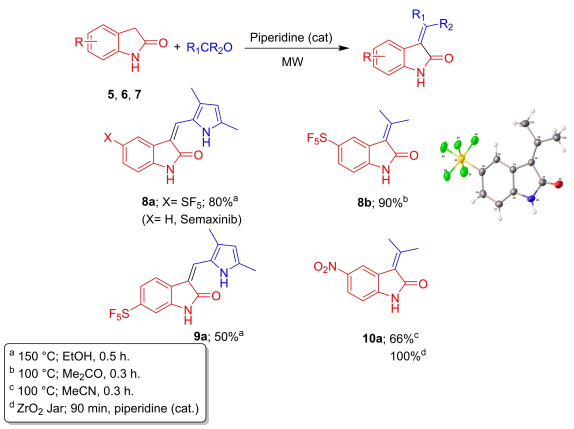


Figure 1. SF_5-benzodiazepine and oxindole analogues previously made in our group.



Scheme 1. Synthesis of a small SF_5-subsituted amide library.



Scheme 2. Oxindole synthesis via a Knoevenagel condensation.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

suited for mechanochemical condensation as opposed to steel jars. [17][18] Reactions performed in the absence of catalyst did not yield any product.

We compared **8a** and **9a** with Semaxinib (Scheme 2), a selective inhibitor of VEGFR[19], using both molecular docking and in vitro analysis in vascular endothelial growth factor receptor-2 VEGFR2 (Figure 2). Semaxinib inhibits VEGFR2 by binding in the ATP binding site, where the pyrrole ring occupies a hydrophobic pocket and forms van der Waals interactions. The oxindole ring binds in a similar fashion to the ATP's purine and provides a hydrogen bond donor and an acceptor to form hydrogen bond interactions. The bulk of the SF_5 group seemingly pushes the oxindole ring into the hydrophobic pocket and the pyrrole into the solvent region, thus the oxindole is no longer able to occupy the adenine pocket. **9a** retains a key aromatic π–π interaction, between Phe1047 (DFG motif) and the oxindole group. However, **8a** does not form an interaction between the oxindole ring and the adenine pocket except for a hydrogen bond to an aromatic hydrogen. The SF_5 moiety is likely to form van der Waals interactions with Lys868 and Ala866 in the hydrophobic binding pocket. A biochemical assay showed Semaxinib to inhibit VEGFR2 with an IC_{50} of 1.5 μM (lit. 1.04 ± 0.53 μM).[20] Compounds **8a** and **9a** did not manifest sufficient inhibition to generate IC_{50} values. At 10 μM, the former exhibited 27% inhibition of VEGFR2 while **9a** gave 23% inhibition at the same concentration (Table 1).

A spirocyclisation reaction led to the product **11a** in poor yield, whose structure was determined by X-ray crystal analysis. **11a** was then methylated to give **11b** (Scheme 3). Such high Csp^3 content compounds may offer more diversity to compound libraries and improved physiochemical properties.

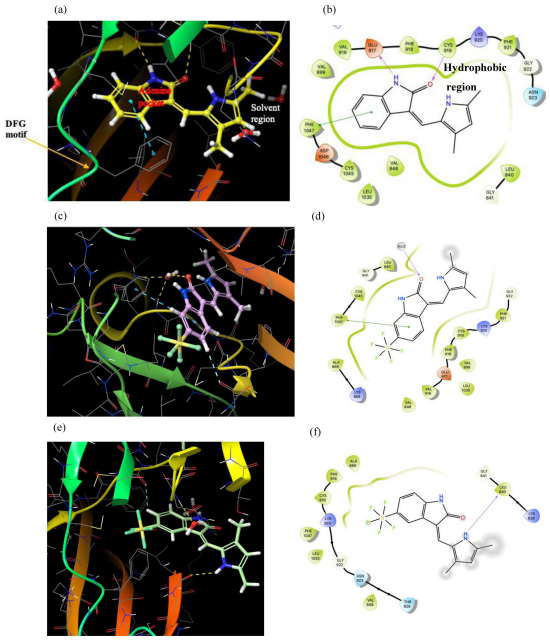


Figure 2. a,b) Docking studies on VEGFR2. (Schrodinger Maestro) show Semaxinib (yellow) in complex with VEGFR2, adenine pocket, hydrophobic pocket and solvent region; c,d) **9a** (pink) in complex with VEGFR2; e) **8a** (green) in complex with VEGFR2; f) lost π–π interactions due to displaced pose. Blue dashed lines indicate aromatic π–π interactions with the DFG motif and hydrogen bonds to Tyr phenol Trp. Hydrogen bonds are shown as yellow dashes.

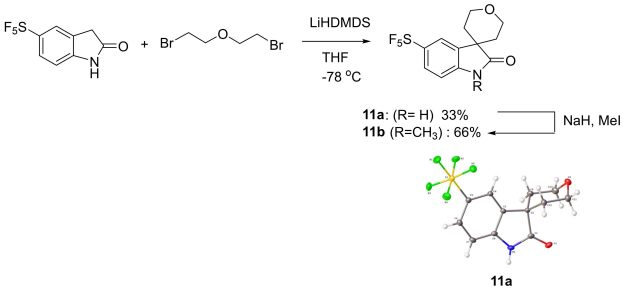
Table 1. Biochemical kinase assays of semaxanib and its SF_5 counterparts.

Entry	Compound	IC_{50} [μM][a]
1	semaxanib	1.5
2	**8a**	—[b]
3	**9a**	—[b]

[a] $n = 1$; 10-dose IC_{50} mode with twofold serial dilutions, starting at 10 μM. [b] Percent inhibition at 10 μM; **8a** (27%), **9a** (23%).

Biological studies

A recently initiated open source effort, termed the Covid Moonshot Consortium, has led to the design of nM-potent inhibitors of M^{Pro} (or CLPro), a vital enzyme in SARS- CoV-2 replication and transcription.[21] M^{pro} releases the functional polypeptides from the polyproteins by extensive proteolytic processing while digesting the polyproteins at 11 cleavage sites, starting with autolytic cleavage[22] and exclusively cleaves polypeptides after a glutamine residue. Its functional importance and the absence of closely related homologues in humans, renders M^{pro} an attractive target for antiviral drug design as an example of a direct acting antiviral agent.[23] Pyridine substituted 3-Cl-phenylacetamide analogues **I** and **II** were found to be moderate to active analogues and we wished to explore both alternatives to the Cl substituent and more water solubilising groups such as CO_2H as opposed to fused aryl or methyl groups. (Figure 3 and Scheme 4).



Scheme 3. Spirocycle formation.

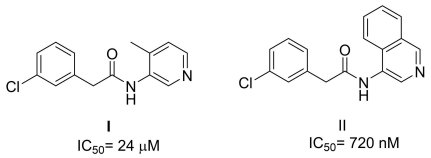
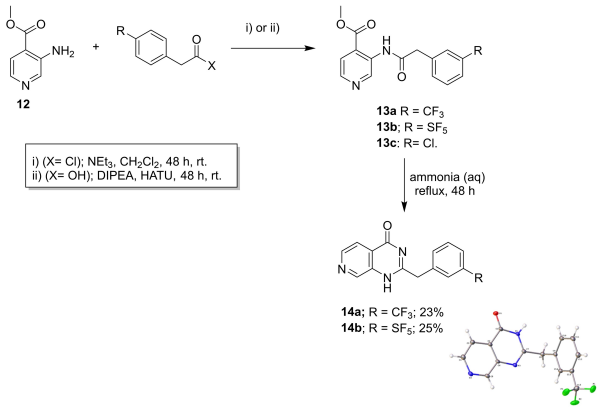


Figure 3. M^{Pro} inhibitors based on 3-chlorophenylacetamides.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Scheme 4. Synthesis of a small-molecule library with CF_3/SF_5 groups.

The aminopyridine **12** was coupled with acid chlorides or an acid to yield the ester analogues **13**, substituted with CF_3, SF_5, and Cl groups (Scheme 4). Surprisingly, reaction of the latter with ammonia led to the cyclised compounds **14**, one of which, **14a**, was crystallised, by a diffusion method, using hexane and dichloromethane, to obtain clear crystals which enabled confirmation of its structure by X-ray analysis.[24] The reaction likely involves the formation of a carboxamide, from the methyl ester, which cyclises onto the amide carbonyl group, followed by water elimination.

Compounds **13–14** were tested for inhibition versus M^{Pro}, but none had any appreciable enzyme activity ($IC_{50} > 99$ μM).

Finally, we focused our efforts on Leflunomide (Arava®), a DMARD (disease modifying antirheumatic drug) with potential COVID-19 use[25] and its active metabolite Teriflunomide (Aubagio®), which is used for multiple sclerosis, (Figure 4). Teriflunomide inhibits human DHODH in the low μM range and binds in the same region as ubiquinone, a redox cofactor. Both Teriflunomide and ubiquinone occupy a narrow cleft near a flavin molecule (another redox cofactor), which leads to the active site (Figure 5).[26] As Teriflunomide competes with ubiquinone,[27] it is regarded as a redox silent coenzyme Q antagonist of DHODH.[28] Teriflunomide has a polar head consisting of one hydrogen bond donor; an enol and two hydrogen bond acceptors, a nitrile and a carbonyl group, while the tail of Teriflunomide, which occupies the entrance of the tunnel, is a hydrophobic CF_3-substituted aromatic group.

We first prepared the SF_5 derivative of Leflunomide **15**, which, when treated with sodium hydroxide, provided its

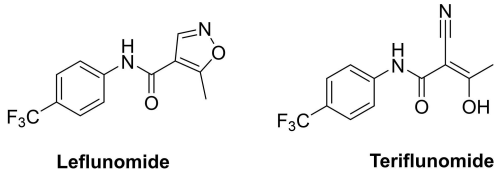


Figure 4. Examples of CF_3-containing DMARDs.

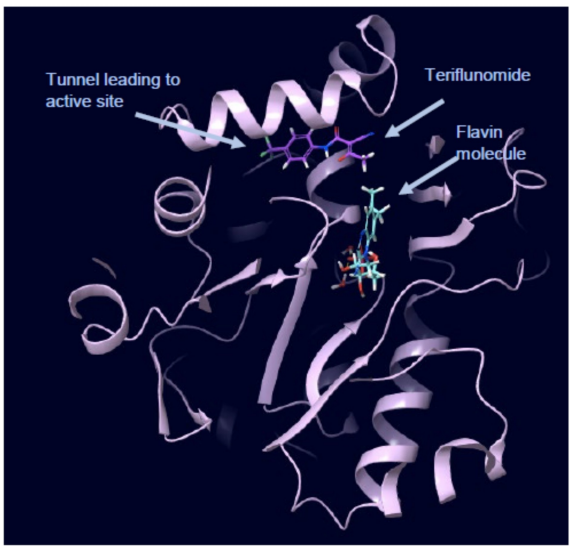
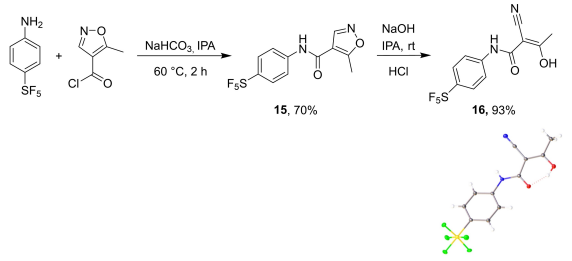


Figure 5. View from a crystal structure of Teriflunomide and FMN in complex with HDHODH. PDB ID: 1D3H. FMN: teal, Teriflunomide: purple.

Teriflunomide equivalent **16** in good yields (Scheme 5).[29] The latter was further characterised in the solid state by crystallography, proving its molecular structure and regiochemistry.

Docking of SF_5-Teriflunomide **16** in HDHODH was performed using Schrodinger Maestro. Interactions predicted between the ligand and the binding pocket are similar to those for Teriflunomide in the same binding site (Figure 6). The hydroxyl group of **16** is hydrogen bonded to a water molecule, which, in turn is bound to Gln47 and Thr360. The hydroxyl also makes a polar interaction with Arg136. The nitrile group interacts with Tyr356 via a H-bond. Finally, the carbonyl of **16** is positioned to form a H-bond with a water molecule that H-bonds with Thr360. Several hydrophobic interactions are formed between the aromatic rings and amino acid residues lining the hydrophobic pocket. The electrostatic potential diagram (Figure 7) shows the small difference in the size of CF_3 and SF_5 groups, and that SF_5-Teriflunomide is slightly larger than its parent analogue.

[Following our docking studies, we tested SF_5-Teriflunomide and SF_5-Leflunomide against human HDHODH using Terifluno-



Scheme 5. Synthesis of potential SF_5-containing DMARDs.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

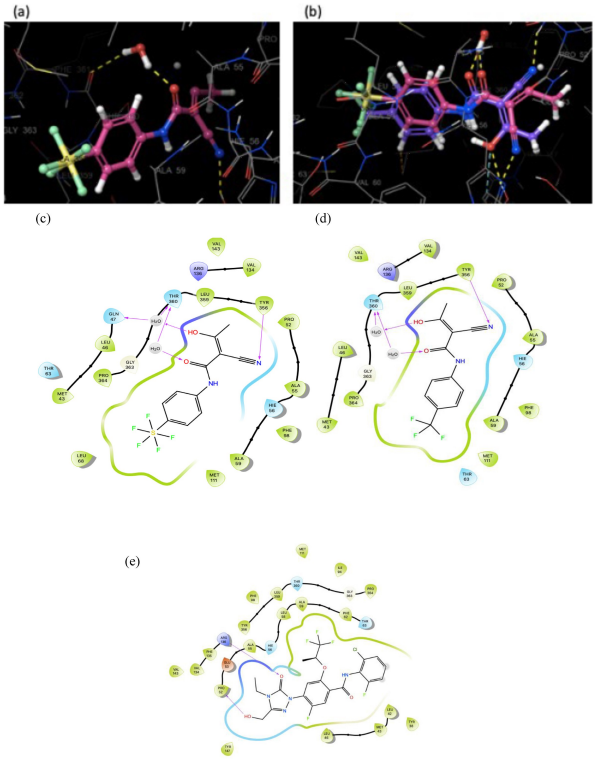


Figure 6. a) SF_5-substituted Teriflunomide in complex with HDHODH, docked using Schrödinger Maestro. b) Teriflunomide and SF_5-Teriflunomide located in the HDHODH binding pocket. Purple: Teriflunomide, salmon pink: SF_5-Teriflunomide Ligand interaction diagram, comparing c) SF_5-Teriflunomide and d) Teriflunomide; showing that the bulkier SF_5 is able to fill the binding pocket better than a CF_3 group. e) **BAY-2402234**, a Bayer clinical trial candidate that inhibits HDHODH with an IC_{50} value of 1.2 nM.

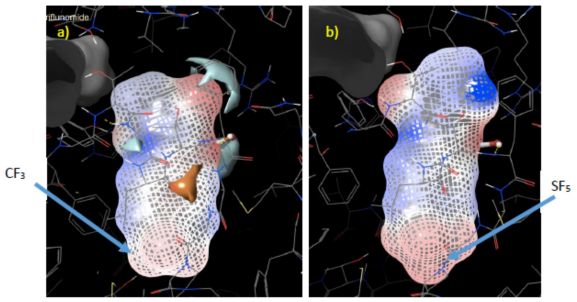
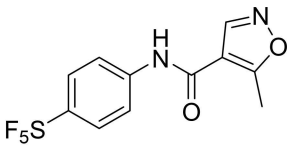
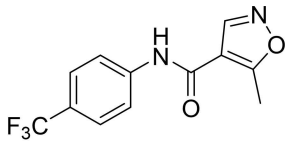
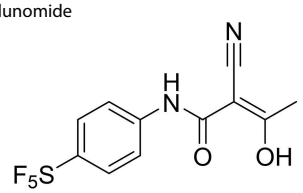
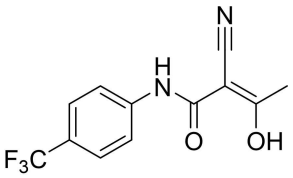
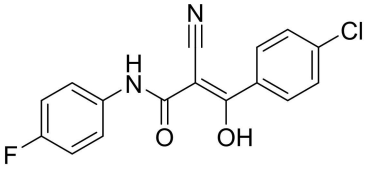
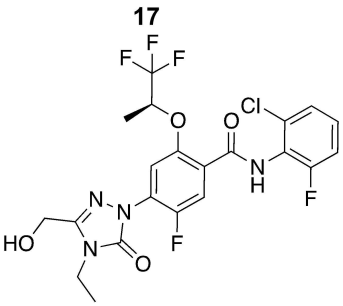


Figure 7. a) Electrostatic potential of Teriflunomide; b) Electrostatic potential of SF_5-Teriflunomide.

Table 2. HDHODH inhibition; comparing SF_5-Teriflunomide, SF_5-Leflunomide and Leflunomide vs. Teriflunomide and **BAY-2402234**.

Entry	Structure	IC_{50} [nM]	pIC_{50}
1	 **15**	365	6.4
2	 Leflunomide	982	6.0
3	 **16**	27[a]	7.6
4	 Teriflunomide	50[b]	7.3
5		_[c]	
6	 BAY-2402234	1.8[d]	8.7

[a] Mean ($n=2$); 29 nM and 25 nM. [b] Unless stated otherwise; in vitro fluorescence-based assays run by Reaction Biology. [c] Percent inhibition at 10 μM < 10%. [d] In vitro control ($n=1$).

mide and **BAY-2402234** (a DHODH inhibitor in phase I clinical trials) as positive controls. A comparison of IC_{50} and pIC_{50} values of SF_5-Teriflunomide, SF_5-Leflunomide, Leflunomide against Teriflunomide and **BAY-2402234** (Table 2) was undertaken. As anticipated, SF_5-Leflunomide is more potent than Leflunomide. SF_5-Teriflunomide is approximately twice as active as Teriflunomide. The simpler analogue, **17**, tested in this assay, due to its MPro affinity (*vide infra*), was inactive. Compound, **BAY-2402234**, gave the best potency with an IC_{50} of 1.8 nM, which is comparable to its literature value (1.2 nM).[30] From similar docking, as well as the hydrophilic interactions, **BAY-2402234** makes hydrophobic interactions with a large set of non-polar residues such as Leu42, Met43, Leu46, etc. and, with its bigger

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

size, and the combination of hydrophobic and hydrophilic groups, it is able to fill the HDHODH active site to a greater degree.

We envisioned Teriflunomide and its SF_5 analogues such as **16** to have a covalent warhead capable of potentially acting as a Michael acceptor with the Sγ atom of Cys145 of M^{Pro}, as precedented with other M^{Pro} inhibitors; other modes of covalent reaction are also possible including with the nitrile, as with an Mpro inhibitor in clinical trials,[31] or the carbonyl of the keto form of **16**. In addition to Teriflunomide and **16**, fourteen other potential candidates were assayed. Out of the sixteen compounds tested, a single hit (**17**, Table 2) was identified which gave a promising IC_{50} of 0.23 μM using the fluorescence-based turnover assay (ESI, Table S1).[32] Docking studies imply ligand **17**, (Figure 8, shown in orange) could covalently react with Sγ of the nucleophilic Cys145 residue (Sγ in yellow, Figure 8), resulting in Michael addition. Besides the covalent bond, the modelling suggests formation of two hydrogen bonds (yellow dashes) are formed between the O–H of **17** and Cys145, as well as the amide carbonyl of **17** and Ser144. In addition, several hydrophobic interactions are predicted.

Although **17** has structural similarities to Teriflunomide, it was found to be inactive versus HDHODH (Table 2). Docking in HDHODH suggests that **17** occupies the binding pocket differently to Teriflunomide (Figure 9). Evidently, the hydrogen bond donors and acceptors on **17** lie in a hydrophobic tunnel of HDHODH.

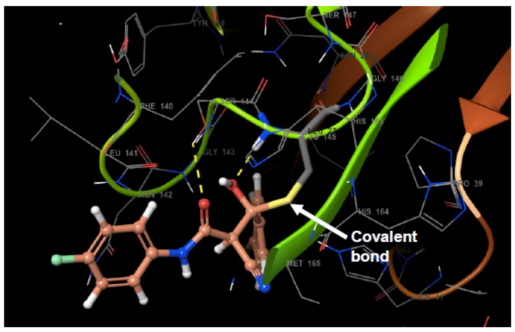


Figure 8. Compound **17** docked in SARS-CoV-2 M^{pro}.

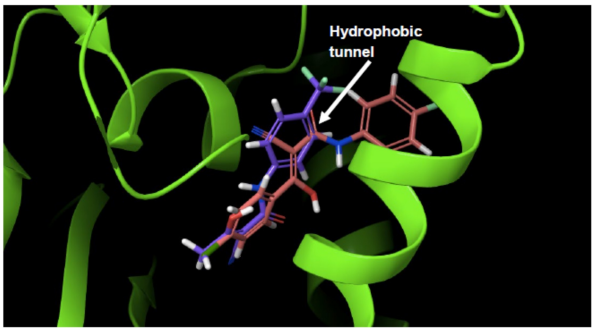


Figure 9. HDHODH in complex with Teriflunomide (purple) and **17** (orange).

The outer surface of coronaviruses has a spike (S) protein which enables its entry into host cells. The S1 subunit of S protein attaches to ACE2 receptor, which is found on the surface of target cells. In addition to this, the transmembrane protease, serine 2 (TMPRSS2) processes S protein into its constituent subunits, S1 and S2, thus allowing the virus to fuse into the plasma membrane of the host cell. As HDHODH is a druggable target against SARS-CoV-2, we were interested in probing the inhibitory effect of Teriflunomide and compounds **16** and **17** against SARS-CoV-2. Teriflunomide was the most cytostatic with effects at concentrations above 0.78 μM. **16** had cytostatic effects at concentrations >3.1 μM and DMSO had no measurable effect. An infection inhibition assay showed that none of the compounds had an inhibitory effect against SARS-CoV-2 infection in the infected cells. The normalised data show that Teriflunomide and **16** have no effect on the efficiency of infection (Fig S2).

Conclusion

SF_5-analogues of Leflunomide and Teriflunomide, **15** and **16** respectively, have been synthesised and tested for affinity towards HDHODH. Biophysical assays revealed an approximately two-fold greater affinity for SF_5-Teriflunomide (**15**) towards HDHODH compared with Teriflunomide. Molecular binding studies revealed that the bulky SF_5 group fills the binding pocket better than the CF_3 group. A singleton hit **17** with structural similarities to Teriflunomide was identified as inhibiting SARS-CoV-2 M^{pro} the low micromolar IC_{50} range whereas a commercial library of similar analogues as well as an in-house library showed no affinity. Testing **17** against HDHODH, however, revealed no affinity. Neither Teriflunomide nor compounds **15** and **16** gave satisfactory inhibition of SARS-CoV-2 infection. Indeed, they were found to be cytostatic in HEK293T/17 cells. Nevertheless, the SF_5 analogue of Teriflunomide, **16**, might be a useful DHODH probe molecule for future investigation.

Experimental Section

4-(Pentafluorosulfanyl)aniline was obtained from Fluorochem and 5-methyl-4-isoxazolecarbonyl chloride was purchased from Apollo Scientific. Magnesium sulphate and sodium bicarbonate were obtained from Fisher Scientific. Preparative TLC plates were obtained from Analtech. Solvents and reagents were purchased from commercial suppliers and were used without purification. All reactions were performed in a fume hood. NMR spectra were recorded on Varian 500 MHz or 400 MHz spectrometers and chemical shifts are reported in ppm, usually referenced to TMS as an internal standard. LCMS measurements were performed on a Shimadzu LCMS-2020 equipped with a Gemini® 5 μm C18 110 Å column and percentage purity measurements were run over 30 minutes in water/acetonitrile with 0.1% formic acid (5 min at 5%, 5%–95% over 20 min, 5 min at 95%) with the UV detector at 254 nm. Mass spectrometry: ESI mass spectra were obtained using a Bruker Daltonics Apex III, using Apollo ESI as the ESI source. For EI mass spectra, a Fissions VG Autospec instrument was used at 70 eV. All analyses were run by Dr. Alla K. Abdul Sada. Analyses are for the

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

molecular ion peak $[M]^+$ and are given in m/z, mass to charge ratio. Melting points were determined using a Stanford Research Systems Optimelt and are uncorrected.

[3-(Pentafluoro-λ⁶-sulfanyl)phenyl](piperidin-1-yl)methanone (3 a)

Triethylamine (247.0 mg, 2.44 mmol), was added dropwise to piperidine (175.8 mg, 2.07 mmol) dissolved in dichloromethane (2 mL) followed by the dropwise addition of 3-(pentafluorosulfanyl)benzoyl chloride (500.0 mg, 1.88 mmol). The reaction mixture was stirred at rt overnight. It was then diluted with dichloromethane (5 mL) and washed with 2 M HCl (10 mL × 3). The aqueous layers were extracted with dichloromethane (15 mL × 3), dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude was purified over a column of silica (dichloromethane:methanol, 9:1) to give **3 a** as a colourless solid (572.0 mg, 97%).^{1}H NMR (600 MHz, $CDCl_3$) δ 7.72 (m, 2H, ArH), 7.47 (m, 2H, ArH), 3.46 (m, 4H, $2CH_2$), 1.63 (m, 4H, $2CH_2$), 1.47 (m, 2H, CH_2). ^{13}C NMR (151 MHz, $CDCl_3$) δ 167.9 (C=O), 153.7 (C-SF_5, p, $^2J_{F,C}$ = 17.5 Hz), 137.3 (ArC), 129.8 (ArC), 128.9 (ArC), 126.6 (ArC-CSF_5, m), 124.6 (ArC-CSF_5, m), 48.6 (N-CH_2), 43.2 (N-CH_2), 26.3 ($2CH_2$), 25.4 (CH_2); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.60 (p, *J* = 150.0 Hz), 62.44 (d, *J* = 150.0 Hz); LCMS Purity (UV) = 96%, t_R 21.0 min; HRMS- ESI (*m/z*) found 316.0785, calc. for $[C_{12}H_{14}F_5NOS][H]^+$: 316.0789; IR (neat) ν_{max}/cm^{-1}: 3680, 1633, 820; mp 88–90 C.

(3-(Pentafluoro-λ⁶-sulfanyl)phenyl](4-methylpiperazin-1-yl)methanone (3 b)

To *N*-methylpiperazine (207.0 mg, 2.07 mmol), dissolved in dichloromethane (2 mL) was added triethylamine (247.4 mg, 2.44 mmol) together with HATU (787.0 mg, 2.07 mmol) and, finally, 3-(pentafluorosulfanyl)benzoyl chloride (500.0 mg, 1.88 mmol) was added dropwise. The reaction mixture was stirred at room temperature overnight, then filtered and the filtrate was stirred with MP Trisamine (3 equiv.) for 3 hours to remove the unreacted acid chloride. The crude product was purified over a column of silica (dichloromethane:methanol; 9:1) to give **3 b** as a colourless gum (107 mg, 17%).^{1}H NMR (600 MHz, $CDCl_3$) δ 7.79 (ArH, m, 2H), 7.53 (ArH, m, 2H), 3.80 (CH_2, bs, 2H), 3.41 (CH_2, bs, 2H), 2.50 (CH_2, bs, 2H), 2.36 (CH_2, bs, 2H), 2.32 (CH_3, s, 3H); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.1 (C=O), 153.9 (ArC-SF_5, p, $^2J_{FC}$ = 17.8 Hz), 136.6 (ArC), 130.1 (ArC), 129.1 (ArC), 127.1 (ArC, p, $^2J_{FC}$ = 4.4 Hz), 124.9 (ArC, p, $^2J_{FC}$ = 4.4 Hz), 55.0 (C), 54.5 (C), 47.6 (C), 45.8 (C), 42.2 (C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.45 (1F, p, *J* = 150.1 Hz), 62.83 (4F, d, *J* = 150.1 Hz); LCMS Purity (UV) = 97%, t_R 7.1 min; HRMS-ESI (*m/z*) found:331.0896, calc. for $[C_{12}H_{15}F_5N_2OS][H]^+$:331.0898; IR (neat) ν_{max}/cm^{-1} 3499 (C–H), 1614 (C=O), 823 (S–F).

[3-(Pentafluoro-λ⁶-sulfanyl)phenyl](morpholino)methanone (3 c)

To morpholine (180.30 mg, 2.07 mmol), dissolved in dichloromethane (2 mL), was added triethylamine (246.90 mg, 2.44 mmol) and the mixture was stirred for a few minutes. HATU (787.10 mg, 2.07 mmol) was added to this solution followed by dropwise addition of 3-(pentafluorosulfanyl)benzoyl chloride (500.0 mg, 1.88 mmol). The reaction mixture was stirred at room temperature overnight then diluted with dichloromethane (5 mL) and the precipitate formed was discarded after gravity filtration. MP-Trisamine (520.0 mg, 3.00 mmol) was added to the filtrate and stirred for 3 hours after which the scavenger was removed by gravity filtration. The crude was purified over a column of silica (dichloromethane:methanol, 9:1) to give **3 c** as a colourless solid (317.0 mg, 53%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.75–7.71 (m, 2H, ArH), 7.50–7.43 (m, 2H, ArH), 3.75–3.63 (m, 4H, $(CH_2)_2$), 3.57–3.26 (m, 4H, $(CH_2)_2$); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.0 (C=O), 153.8 (p, *J* = 17.7 Hz, $ArCSF_5$), 136.2 (ArC), 130.1(ArC), 129.1 (ArC), 127.2 (p, *J* = 4.5 Hz, (ArC-CSF_5)), 124.9 (p, *J* = 4.5 Hz, ArC-CSF_5), 66.6 (4 C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 84.29–82.21 (m), 62.66 (d, *J* = 150.4 Hz); LCMS Purity (UV) = 97%, t_R 18.3 min; HRMS-ESI (*m/z*) found 340.0384, calc. for $[C_{11}H_{12}F_5NO_2S][Na]^+$: 340.0401; IR (neat) ν_{max}/cm^{-1} 2858, 1644, 821; mp 63–64 C.

tert-Butyl 4-(3-pentafluoro-λ⁶-sulfanyl benzoyl)piperazine-1-carboxylate (3 d)

To *N*-Boc piperazine (385.5 mg, 2.07 mmol), dissolved in dichloromethane (3 mL) was added triethylamine (247.0 mg, 2.44 mmol) and the mixture was stirred for a few minutes. HATU (787.2 mg, 2.07 mmol) was added to this solution followed by dropwise addition of 3-(pentafluorosulfanyl)benzoyl chloride (500.0 mg, 1.88 mmol) and the reaction mixture was stirred at room temperature overnight. After completion of the reaction, it was diluted with dichloromethane (10 mL) and the precipitate formed was removed by filtration. MP Trisamine (520 mg, 3 mmol) was added to the filtrate then removed by gravity filtration after 3 hours of stirring. The crude material was purified over a column of silica (dichloromethane:methanol, 9:1) to obtain **3 d** as a colourless solid (573.0 mg, 73%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.81(m, 2H, ArH), 7.53 (m, 2H, ArH), 3.51 (m, 4H, $(CH_2)_2$), 3.39 (m, 4H, $(CH_2)_2$), 1.45 (s, 9H, $(CH_3)_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.4 (C=O), 154.5 (C=O), 154.0 (ArC-SF_5), 136.3 (ArC), 130.1 (ArC), 129.2 (ArC), 127.4 (ArC-CSF_5), 125.0 (ArC-CSF_5), 80.5 (C), 47.5 (2 C, m), 42.9 (2 C, m), 28.3 (3 C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.20 (m, F), 62.67 (d, *J* = 150.1 Hz, 4F); LCMS Purity (UV) = 99%, t_R 20.4 min. HRMS-ESI (*m/z*) found 439.1078, calc. for $[C_{16}H_{21}F_5N_2O_3S][Na]^+$: 439.1085; IR (neat) ν_{max}/cm^{-1} 2977, 1685, 1626, 1243, 828; mp 153–155 C.

1-[4-[3-(Pentafluoro-λ⁶-sulfanyl)benzoyl]piperazin-1-yl] ethanone (3 e)

Compound **3 d** (573 mg, 1.38 mmol) in dichloromethane (2 mL) was treated with HCl in dioxane (4 M, 5.0 equiv.) overnight. The reaction mixture was concentrated in *vacuo* to afford the HCl salt, **3 d'** (4-[3-(pentafluoro-λ⁶-sulfanyl)benzoyl]piperazinyl hydrochloride) as a colourless solid. To crude **3 d'** (0.100 g, 0.28 mmol), dissolved in THF (3 mL) was added triethylamine (0.028 g, 0.28 mmol) and stirred for 30 minutes. Acetic anhydride (0.028 g, 0.28 mmol) was added to the mixture and stirred overnight at room temperature. The reaction mixture was concentrated in *vacuo* and purified by flash chromatography (dichloromethane:methanol; 9:1) to obtain the purified **3 e** as a colourless oil (47.0 mg, 47%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.84–7.81 (m, 1H, ArH), 7.80–7.79 (m, 1H, ArH), 7.57–7.51 (m, 2H, ArH), 3.88–3.28 (m, 8H, ArH), 2.11 (s, 3H, CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 169.3 (C=O, MeCO), 168.4 (C=O), 154.0 (t, $^2J_{FC}$ = 18.0 Hz, ArC), 136.0 (ArC), 130.2 (ArC), 129.3 (ArC), 127.7–127.4 (m, ArC), 125.0 (ArC), 47.4 (bs, CH_2), 45.9 (bs, CH_2), 42.4 (bs, CH_2), 41.4 (bs, CH_2), 21.3 (CH_3); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.19 (p, *J* = 150.3 Hz, F), 62.78 (d, *J* = 150.3 Hz, 4F); LCMS Purity (UV) = 96%, t_R 13.5 min; HRMS-ESI (*m/z*) found 381.0662, calc. for $[C_{13}H_{15}F_5N_2O_2S][Na]^+$: 381.0667; IR (neat) ν_{max}/cm^{-1} 2917 (C–H), 1633 (C=O), 1322 (S=O), 821 (S–F).

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(3-Pentafluoro-λ^6-sulfanyl phenyl)(4-(methylsulfonyl)piperazin-1-yl)methanone (3 f)

To crude **3 d'** (75 mg, 0.21 mmol), dissolved in THF (3 mL) was added triethylamine (21 mg, 0.21 mmol) and stirred for 30 min. Methylsulfonyl chloride (24 mg, 0.21 mmol) was added to the reaction mixture and stirred overnight at room temperature. The reaction was followed by TLC. After completion of the reaction, it was separated over dichloromethane (5 mL) and water (5 mL). The biphasic layers were passed through a hydrophobic frit to obtain the organic phase which was concentrated in *vacuo*. The crude was then purified over a column of silica, (dichloromethane:methanol; 7:3) to obtain **3 f** as a colourless solid (66 mg, 80%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.86–7.82 (m, 1H, ArH), 7.82–7.80 (m, 1H, ArH), 7.57–7.54 (m, 2H, ArH), 4.03–3.47 (m, 4H, $(CH_2)_2$), 3.38–3.14 (m, 4H, $(CH_2)_2$), 2.81 (s, 3H, CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.4 (C=O), 154.2–153.8 (m, C-SF_5), 135.7 (ArC), 130.2 (ArC), 129.4 (ArC), 127.8–127.6 (m, ArC), 125.2–125.0 (m, ArC), 47.4 (CH_2), 45.5 (2 C, $(CH_2)_2$), 42.0 (CH_2), 35.0 (CH_3); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.32 (p, J = 150.4 Hz, 1F), 62.96 (d, J = 150.4 Hz, 4F); LCMS Purity (UV) = 98%, t_R 15.6 min; HRMS-ESI (*m/z*) found 417.0349, calc. for $[C_{12}H_{15}F_5N_2O_2S][Na]^+$: 417.0342; IR (neat) ν_{max}/cm^{-1} 2007 (C–H), 1616 (C=O), 1342 (S=O), 826 (S–F); mp 159–160 C.

[4-(Pentafluoro-λ^6-sulfanyl)phenyl](piperidin-1-yl)methanone (4 a)

Triethylamine (247.4 mg, 2.44 mmol) was added to piperidine (176.0 mg, 2.07 mmol) dissolved in dichloromethane (2 mL) followed by dropwise addition of 4-(pentafluorosulfanyl) benzoyl chloride (500.0 mg, 1.88 mmol). The reaction mixture was stirred at room temperature overnight. It was then diluted with dichloromethane (5 mL) and washed with 2 M HCl (10 mL×3). The aqueous layers were extracted with dichloromethane (15 mL×3), dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude was purified over a column of silica (dichloromethane:methanol, 9:1) to give **4 a** as a colourless solid (451.0 mg, 76%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.74 (d, J = 8.4 Hz, 2H, ArH), 7.43 (d, J = 8.4 Hz, 2H, ArH), 3.66 (m, 2H, CH_2), 3.24 (m, 2H, CH_2), 1.63 (m, 4H, $(CH_2)_2$), 1.47(m, 2H, CH_2). ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.0 (C=O), 154.1(p, J = 17.8 Hz, ArC-SF_5), 139.9 (ArC), 127.1 (2ArC), 126.2 (m, 2ArC-CSF_5), 48.6 (C), 43.1 (C), 26.4 (C), 25.4 (C), 24.3 (C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.75 (m, F), 62.53 (d, J = 150.0 Hz, 4F); LCMS Purity (UV) = 95%, t_R 19.5 min; HRMS-ESI (*m/z*) found 316.0789, calc. for $[C_{12}H_{14}F_5NOS][H]^+$:316.0789; IR (neat) ν_{max}/cm^{-1}: 2945, 1627, 817; mp 84.5 - 86.7 C.

(4-Pentafluoro-λ^6-sulfanylphenyl)(4-methylpiperazin-1-yl)methanone) (4 b)

To 1-methyl piperazine (188.0 mg, 1.88 mmol), dissolved in dichloromethane (2 mL) was added triethylamine (190.3 mg, 1.88 mmol), followed by dropwise addition of 4-(pentafluorosulfanyl) benzoyl chloride (250.0 mg, 0.94 mmol) and HATU (391.2 mg, 0.94 mmol). The reaction mixture was stirred at room temperature overnight and then diluted with dichloromethane (10 mL). The reaction mixture was filtered and the organic filtrate was washed with water (10 mL×3) and the aqueous layer extracted with dichloromethane (10 mL×3). The organic layer was then dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude product was purified over a column of silica (dichloromethane:methanol, 9:1) to afford a yellow solid as the title compound (257 mg, 83%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.78 (m, 2H, ArH), 7.52 (m, 2H, ArH), 3.78 (m, 2H, CH_2), 3.39 (m, 2H, CH_2), 2.50 (m, 4H, $(CH_2)_2$), 2.30 (s, 3H, CH_3). ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.0 (C=O), 153.9 (m, ArC-SF_5), 136.6 (ArC), 130.1 (ArC), 129.1 (ArC), 127.1 (m, ArC-CSF_5), 124.9 (m, ArC-CSF_5), 54.8 (2 C), 47.6 (C), 45.9 (CH_3), 42.2 (C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.50 (m, F), 62.67 (d, J = 150.3 Hz, 4F); LCMS Purity (UV) = 98%, t_R 11.6 min. HRMS-ESI (*m/z*) found 331.0901, calc. for $[C_{12}H_{15}F_5N_2OS][H]^+$:331.0898. IR (neat) ν_{max}/cm^{-1} 2811.3, 1613.5, 822.4; mp 159.7–160.2 C.

[4-(Pentafluorol-λ^6-sulfanyl)phenyl)](morpholino)methanone (4 c)

To morpholine (180.3 mg, 2.07 mmol), dissolved in dichloromethane (2 mL) was added triethylamine (247.0 mg, 2.44 mmol), followed by dropwise addition of 4-(pentafluorosulfanyl) benzoyl chloride (500.0 mg, 1.88 mmol). The reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with dichloromethane (5 mL) and the organic layer was washed with 2 M HCl (10 mL×3). The combined aqueous layers were extracted with dichloromethane (15 mL×3). The organic layers were dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude material was purified over a column of silica (dichloromethane:methanol, 9:1) to afford a colourless solid as the title compound (556.0 mg, 93%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.58 (d, J = 8.1 Hz, 2H, ArH), 7.30 (d, J = 8.1 Hz, 2H, ArH), 3.33 (m, 8H, $((CH)_2)_4$); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.1 (C=O), 154.2 (p, J = 17.7 Hz, ArC-SF_5), 138.7 (ArC), 127.4 (2ArC), 126.3 (m, 2ArC), 66.5 (2 C, N-$(CH_2)_2$, 47.9 (CH_2), 42.4 (CH_2); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.40 (p, J = 150.3, F), 62.39 (d, J = 150.3 Hz, 4F); LCMS Purity (UV) = 95%, t_R 18.3 min; HRMS-ESI (*m/z*) found 340.0394, calc. for $[C_{11}H_{12}F_5NO_2S]$ $[Na]^+$: 340.0401. IR (neat) ν_{max}/cm^{-1}: 3734 (CH), 1626 (s, C=O), 818 (SF_5); mp 72–73 C.

tert-Butyl 4-(4-pentafluoro-λ^6-sulfanyl) benzoyl piperazine-1-carboxylate (4 d)

To *N*-Boc piperazine (385.5 mg, 2.07 mmol), dissolved in dichloromethane (3 mL) was added triethylamine (247.0 mg, 2.44 mmol) and the mixture was stirred for a few minutes. HATU (787.1 mg, 2.07 mmol) was added to this solution followed by dropwise addition of 4-(pentafluorosulfanyl) benzoyl chloride (500.0 mg, 1.88 mmol). The reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with dichloromethane (5 mL) and the precipitate formed was discarded after gravity filtration. MP-Trisamine scavenger resin (520.0 mg, 3 mmol) was added to the filtrate and after 3 hours, the scavenger was removed by gravity filtration. The crude material was purified over a column of silica (dichloromethane:methanol, 9:1) to obtain **4 d** as a white solid (352.0 mg, 76%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.67 (m, 2H, ArH), 7.37(m, 2H, ArH), 3.78–3.49 (m, 4H), 3.40–3.32 (m, 4H), 1.32 (s, 9H, $(CH_3)_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.4 (C=O), 154.6–153.9 (m, 2 C, C=O, ArC-SF_5), 138.8 (ArC), 127.4 (2ArC), 126.5 (2ArC), 80.5 (C), 47.4 (2 C), 42.1 (2 C), 28.3 (3 C); ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.32 (p, J = 150.2 Hz, F), 62.47 (d, J = 150.2 Hz, 4F); LCMS Purity (UV) = 97%, t_R 21.5 min; HRMS - ESI (*m/z*) found: 439.1078, calc. for $[C_{16}H_{21}F_5N_2O_3S][Na]^+$: 439.1085. IR (neat) ν_{max}/cm^{-1} 2920, 1637, 1279, 816; mp: 151–153 C.

1-[4-[4-(Pentafluoro-λ^6-sulfanyl)benzoyl]piperazin-1-yl] ethanone (4 e)

Compound **4 d** was dissolved in dichloromethane (2 mL) and treated with HCl in dioxane (4 M, 6.0 equiv.) overnight to remove the Boc group. The reaction was monitored by TLC. Following the completion of the reaction, it was concentrated in *vacuo* to obtain the HCl salt (4-[4-(pentafluoro-λ^6-sulfanyl)benzoyl]piperazinyl hydrochloride, **4 d'**), as a colourless solid. Compound **4 d'** was used without purification in the following step.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To **4d′** (80.0 mg, 0.2 mmol), dissolved in THF (1.0 mL), was added triethylamine (20.0 mg, 0.2 mmol). The reaction mixture was stirred for 30 minutes, followed by which, acetic anhydride (20 mg, 0.2 mmol) was added to the mixture and this was stirred at room temperature overnight. The crude was dissolved in dichloromethane (5 mL) and passed through a hydrophobic frit, to collect the organic layer. Afterwards, the crude material was purified over a column of silica (dichloromethane:methanol; 7:3) to obtain the pure product as a colourless solid (59 mg, 82%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.81 (d, J = 8.4 Hz, 2ArH), 7.49 (d, J = 8.4 Hz, 2ArH), 3.89–3.64 (m, 4H, piperazine), 3.56 (s, 3H, CH_3), 3.47–3.30 (m, 4H, piperazine); ^{13}C NMR (151 MHz, $CDCl_3$) δ 169.2 (Ar-C=O), 168.5 (C=O), 154.6 (m, ArC-SF_5), 138.4 (ArC), 127.5 (2ArC), 126.6 (2ArC), 47.3 (bs, C), 46.1 (bs, C), 42.1 (bs, C), 41.3 (bs, C), 21.4 (CH_3) ^{19}F NMR (400 MHz, $CDCl_3$) δ 83.32 (p, J = 150.4 Hz, 1F), 62.96 (d, J = 150.4 Hz, 4F); HPLC Purity (UV) = 95%, t_R 13.4 min; HRMS-ESI (*m/z*) found 381.0665, calc. for $[C_{13}H_{15}F_5N_2O_2S][Na]^+$:381.0667; IR (neat) ν_{max}/cm^{-1} 2917(C–H), 1633 (C=O), 1625 (C=O), 820 (S–F); mp 158–160 C.

(4-Pentafluoro-λ^{6}sulfanyl phenyl)(4-(methylsulfonyl)piperazin-1-yl)methanone (4f)

To crude **4d′** (82.0 mg, 0.23 mmol) dissolved in THF (3.0 mL) was added triethylamine (23.0 mg, 0.23 mmol) and the mixture was stirred for 30 minutes. Methanesulfonyl chloride (32.0 mg, 0.28 mmol) was added to the reaction mixture, which was stirred overnight at room temperature. The reaction mixture was quenched with dichloromethane (5 mL) and water (5 mL) then passed through a hydrophobic frit. The crude was purified over column of silica (dichloromethane:methanol; 7:3) to get the pure product as a colourless solid (43 mg, 48%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.83 (d, J = 8.2 Hz, ArH, 2H), 7.50 (d, J = 8.2 Hz, ArH, 2H), 3.90 (m, CH_2, 2H), 3.52 (m, CH_2, 2H), 3.34 (m, CH_2, 2H), 3.18 (m, CH_2, 2H), 2.81 (s, CH_3, 3H); ^{13}C NMR (151 MHz, $CDCl_3$) δ 168.4 (C=O), 154.8 (m, C-SF_5), 138.1 (ArC), 127.5 (ArC, 2 C), 126.7 (ArC, 2 C), 47.2 (CH_2), 46.0 (CH_2), 45.5 (CH_2), 41.8 (CH_2), 35.1 (CH_3); ^{19}F NMR (400 MHz, $CDCl_3$) δ 82.99 (p, J = 150.3 Hz, 1F), 62.51 (d, J = 150.3 Hz, 4F); HRMS - ESI (*m/z*) found 417.0342, calc. for $[C_{12}H_{15}F_5N_2O_3S_2][Na]^+$:417.0349; Anal. calc (%) for $C_{12}H_{15}F_5N_2O_3S_2$: C, 36.55; H, 3.83; N, 7.10; found (%): C, 36.65; H, 3.69; N, 6.98. IR (neat) ν_{max}/cm^{-1}2187 (C–H), 1637 (C=O), 1322 (S=O), 821 (S–F); mp 163–164 C.

(3Z)-3-[(3,5-Dimethyl-1H-pyrrol-2-yl)methylidene]-5-(pentafluoro-λ^6-sulfanyl)-2,3-dihydro-1H-indol-2-one (8a)

In a 10 mL microwave vial equipped with a stirrer bar was added 5-(pentafluorosulfanyl)-1,3-dihydro-indol-2-one (0.096 g, 0.37 mmol), 3,5-dimethyl-1H-pyrrole-2-carbaldehyde, ethanol (2.5 mL) and 3 drops of piperidine. The vessel was then sealed using a rubber microwave septum and placed into the microwave cavity. The reaction mixture was subjected to microwave irradiation by ramping to 150 C with 200 W of power. It was held at that temperature for 30 min. The vessel was then cooled to ambient temperature and solvent removed in *vacuo*. The crude was purified over a column of C18 silica (100% acetonitrile) to obtain purified **8a** as a red solid (0.137 g, 80%). ^{1}H NMR (600 MHz, $(CD_3)_2CO$) δ 13.46 (1H, s, NH), 10.15 (1H, s, NH), 8.21 (1H, s, ArH), 7.91 (1H, s, ArH), 7.60 (1H, d, J = 8.6 Hz, ArH), 7.10 (1H, d, J = 8.6 Hz, ArH), 6.06 (1H, s, C=H), 2.38 (3H, s, CH_3), 2.35 (3H, s, CH_3); ^{13}C NMR (151 MHz, $(CD_3)_2CO$) δ 169.8 (C=O), 147.7 (m, ArC-SF_5), 140.0 (ArC=C), 138.0 (ArC), 134.4 (ArC), 127.4 (ArC), 126.6 (ArC), 125.4 (m, ArC-CSF_5), 123.0 (ArC), 115.5 (ArC), 113.3 (m, ArC-CSF_5), 110.5 (C=C), 108.5 (ArC), 12.9 (CH_3), 10.7 (CH_3); ^{19}F NMR (400 MHz, $(CD_3)_2CO$) δ 87.2 (1F, m, equatorial), 63.74 (1F, d, J = 148.1 Hz, axial); LCMS Purity (UV) = 100%, t_R 23.6 min; HRMS-ESI (*m/z*) found 365.0737, calc. for $[C_{15}H_{13}F_5N_2OS][H]^+$: 365.0742; IR (neat) ν_{max}/cm^{-1} 3060 (N–H), 1659 (C=O), 820 (S–F); mp 203–205 C.

5-(Pentafluoro-λ^6-sulfanyl)-3-(propan-2-ylidene)-2,3-dihydro-1H-indol-2-one (8b)

In a 10 mL microwave vial equipped with a stirrer bar was added 5-(pentafluorosulfanyl)-1,3-dihydro-indol-2-one (0.096 g, 0.37 mmol) in acetone (2 mL) followed by 3 drops of piperidine. The vessel was then sealed using a rubber microwave septum and placed into the microwave cavity. The reaction mixture was irradiated with 200 W of power. It was maintained at 100 C by moderation of power for 20 min. The vessel was then cooled to room temperature. The reaction mixture was extracted with ethyl acetate (3×20 mL), washed with water (20 mL), dried over $MgSO_4$, filtered under gravity and concentrated in *vacuo*. The obtained orange solid was used without further purification (0.100 g, 90%). Crystallization by mixed solvents, dichloromethane and hexane, provided orange crystals of **8b**. ^{1}H NMR, (600 MHz, (CD_3OD) δ N–H not observed, 7.87 (1H, s, ArH), 7.65 (1H, d, J = 8.5 Hz, ArH), 6.94 (1H, d, J = 8.5 Hz, ArH),2.59 (3H, s, CH_3), 2.39 (3H, s, CH_3), ^{13}C NMR (600 MHz, CD_3OD) δ 169.6 (C=O), 159.0 (C=C), 147.6 (ArC-SF_5, m), 142.5 (ArC), 125.6 (ArC-CSF_5, m), 123.6 (ArC), 121.8 (ArC), 120.7 (C, ArC-CSF_5, m), 108.2 (C=C), 24.0 (CH_3), 21.9 (CH_3); ^{19}F NMR, (400 MHz, CD_3OD) δ 62.9 (4F, d, J = 147.8 Hz, equatorial), 85.3(1F, p, J = 147.8, axial); LCMS Purity (UV) = 99%, t_R 18.5 min; HRMS-ESI (*m/z*) found 322.0285, calc. for $[C_{11}H_{10}F_5NOS][Na]^+$:322.0295; IR (neat) ν_{max}/cm^{-1} 3062 (N–H), 2925 (C–H), 1694 (C=O), 1618 (C=C), 813 (S–F); mp 191–194 C.

(3Z)-3-[(3,5-Dimethyl-1H-pyrrol-2-yl)methylidene]-6-(pentafluoro-λ^6-sulfanyl)-2,3-dihydro-1H-indol-2-one (9a)

In a 10 mL microwave vial equipped with a stirrer bar was added 6-(pentafluorosulfanyl)-1,3-dihydro-indol-2-one(0.100 g, 0.4 mmol), 3,5-dimethyl-1H-pyrrole-2-carbaldehyde (0.062 g, 0.5 mmol), ethanol (2.5 mL) and 3 drops of piperidine. The vessel was then sealed using a rubber microwave septum and placed into the microwave cavity. The reaction mixture was irradiated with 200 W of power. It was maintained at 150 C by moderation of power for 30 min. The vessel was then cooled to room temperature and solvent removed in *vacuo*. The aqueous layer was extracted using ethyl acetate (3×20 mL), washed with brine (3×20 mL), dried over $MgSO_4$, filtered under gravity and concentrated in *vacuo*. The crude was purified over a column of C18 silica (100% acetonitrile) (70 mg, 50%) to obtain an orange solid. ^{1}H NMR (600 MHz, $(CD_3)_2CO$) δ 13.52 (1H, s, NH), 10.02 (1H, s, NH), 7.82 (1H, d, J = 8.5 Hz, ArH), 7.79 (1H, s, ArH), 7.48 (1H, d, J = 8.5 Hz, ArH), 7.41 (1H, s, ArH), 6.08 (1H, s, C=H), 2.36 (3H, s, CH_3), 2.38 (3H, s, CH_3). ^{13}C NMR (151 MHz, $(CD_3)_2CO$) δ 169.6 (C=O), 138.7 (ArC-SF_5), 137.3 (ArC=C), 135.0 (ArC), 130.2 (ArC), 127.7 (ArC), 126.0 (ArC-CSF_5), 118.6 (ArC), 117.0 (ArC), 113.6 (ArC-CSF_5), 110.0 (ArC), 106.8. (ArC), 54.1 (C=C), 12.9 (CH_3), 10.7 (CH_3). ^{19}F NMR (400 MHz, $(CD_3)_2CO$) δ 86.32 (1F, m), 63.74 (4F, d, J = 148.1 Hz); LCMS Purity (UV) = 100%, t_R 23.7 min; HRMS-ESI (*m/z*) found 365.0742, calc. for $[C_{15}H_{13}F_5N_2O_3S][Na]^+$: 365.0742. IR (neat) ν_{max}/cm^{-1} 3559, 1650, 814; mp 199–201 C.

5-Nitro-3-(propan-2-ylidene)-2,3-dihydro-1H-indol-2-one (10a)

Microwave reaction procedure

In a 10 mL microwave vial equipped with a stirrer bar was added 5-nitro-2-oxindole (0.50 g, 2.85 mmol), acetone (0.25 mL, 3.42 mmol),

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

acetonitrile (5 mL) and 3 drops of piperidine. The vessel was then sealed using a rubber microwave septum and placed into the microwave cavity. The reaction mixture was irradiated with 200 W of power and heated to 100 C in a CEM Explorer. It was maintained at 100 C by moderation of power for 20 min. The vessel was then cooled to ambient temperature and the reaction mixture was concentrated in *vacuo*. The crude material was recrystallized in acetonitrile to obtain a brown solid as pure product (0.411 g, 66%).

Solventless reaction procedure

5-Nitro-2-oxindole (242.0 mg, 1.36 mmol), acetone (200.0 mL, 2.71 mmol) and piperidine (40.2 mL, 0.41 mmol) were mechanically activated in a Retsch MM400 vibratory ball-mill for 90 minutes at 30 Hz in 25 mL zirconia jar with one, 1.2 cm dia. ball. The product was suspended in acetone (10 mL) and filtered to obtain the pure product as a brown solid (229.0 mg, 78%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 10.96 (1H, s, NH), 8.21 (1H, s, ArH), 8.11 (1H, dd, J = 8.6, 2.0 Hz, ArH), 6.94 (1H, d, J = 8.6 Hz, ArH), 2.52 (3H, s, CH_3), 2.38 (3H, s, ArH).^{13}C NMR (151 MHz, $(CD_3)_2SO$) δ 169.0 (C=O), 160.0 (ArC), 146.6 (ArC), 141.9 (ArC), 125.0 (ArC–H), 124.1 (ArC-NH), 121.5 (C=ArC), 118.9 (ArC–H), 109.4 (ArC–H), 25.5 (CH_3), 23.2 (CH_3); HRMS-ESI (m/z) found 241.0584, calc. for $[C_{11}H_{10}N_2O_3][Na]^+$: 241.0584; IR (neat) ν_{max}/cm^{-1}1694 (C=O), 1509 (N–O stretch); LCMS Purity (UV) = 100%, t_R = 14.6 min; mp 228–230 C.

5-(Pentafluoro-λ^6-sulfanyl)-1,2-dihydrospiro[indole-3,4′-oxan]-2-one (11 a)

5-(Pentafluorosulfanyl)-1,3-dihydro-indol-2-one (0.500 g, 1.93 mmol) dissolved in tetrahydrofuran (10 mL) was added to LiHMDS (1 M) in tetrahydrofuran (5.80 mL, 5.80 mmol) at −78 C. After an hour of stirring, 1-bromo-2-(2-bromoethoxy)ethane (0.448 g, 1.93 mmol) was added to the mixture which was allowed to warm to ambient temperature. The reaction was stirred for 46 hours, then quenched with water. The reaction mixture was extracted with ethyl acetate (3×20 mL), washed with brine (3×20 mL), dried over $MgSO_4$, filtered under gravity and concentrated in *vacuo*. The crude material was purified over a column of silica (hexane:ethyl acetate; 4:1) to afford a pink solid as the title compound (210 mg, 33%). ^{1}H NMR (600 MHz, CD_3OD) δ 7.82 (1H, s, ArH), 7.71 (1H, d, J = 8.6 Hz, ArH), 7.01 (1H, d, J = 8.6 Hz, ArH), 4.22 (2H, m, OCH_2), 3.88 (2H, m, OCH_2), 1.95 (2H, m, CH_2), 1.82 (2H, m, CH_2); ^{13}C NMR (151 MHz, $(CD_3)_2CO$) δ 180.7 (C=O), 147.8 (m, ArC-SF_5), 144.3 (ArC), 135.2 (ArC), 126.7 (m, ArC-CSF_5), 121.2 (m, ArC-CSF_5), 109.0 (ArC), 62.0 (2 C), 44.6 (C), 32.5 (2 C);^{19}F NMR (400 MHz, $(CD_3)_2CO$) δ 86.41 (4F, m), 64.09 (F, d, J = 148.1 Hz); LCMS Purity (UV) = 98%, t_R = 20.2 min; HRMS-ESI (*m/z*) found 329.0535, calc. for $[C_{12}H_{11}F_5NO_2S][H]^+$: 329.0509; IR (neat) ν_{max}/cm^{-1} 3303, 1690, 1101, 818; mp 247–249 C.

1-Methyl-5-(pentafluoro-λ^6-sulfanyl)-1,2-dihydrospiro[indole-3,4′-oxan]-2-one (11 b)

Sodium hydride (11 mg, 0.45 mmol) was added to a round bottomed flask purged with argon. Dry dimethyl formamide (15 mL) was then added to the flask followed by compound **11 a** (0.1 g, 0.30 mmol) and stirred for 1 hour. To this mixture was added methyl iodide (0.213 g, 1.50 mmol) and stirred overnight at room temperature. The reaction mixture was quenched with water (10 mL), extracted with ethyl acetate (10 mL×3), washed with brine (20 mL), dried over $MgSO_4$, filtered, and concentrated in *vacuo*. The crude material was purified over a column of silica (hexane:ethyl acetate;4:1) to give **11 b**, as a yellow solid (68 mg, 66%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.71 (ArH, dd, J = 8.6, 1.9 Hz, 1H), 7.68 (ArH, d, J = 1.9 Hz, 1H), 6.86 (ArH, d, J = 8.6 Hz, 1H), 4.25 (CH_2, ddd, J = 12.0, 9.5, 3.0 Hz, 2H), 3.89 (CH, dt, J = 12.0, 4.3 Hz, 2H), 3.22 (s, 3H), 1.91 (ddd, J = 13.4, 9.5, 4.3 Hz, 2H), 1.82 (dt, J = 13.4, 3.0 Hz, 2H).^{13}C NMR (151 MHz, $CDCl_3$) δ 179.5 (C=O), 148.8 (ArC, p, $^2J_{FC}$ = 17.7, 17.2 Hz), 145.1 (ArC), 134.0 (ArC), 126.7 (ArC–H, p, $^3J_{FC}$ = 4.7 Hz), 120.9 (ArC–H, p, $^3J_{FC}$ = 4.7 Hz), 107.3 (ArC–H), 62.6 (2 C), ((CH_2)$_2$), 44.4 (C), 32.7 (2 C, (CH_2)$_2$), 26.3 (CH_3).^{19}F NMR (400 MHz, $CDCl_3$) δ 86.61–84.84 (p, J = 150.1 Hz), 64.42 (d, J = 150.1 Hz); HRMS-ESI (m/z) found 344.0743, calc. for $[C_{13}H_{14}F_5NO_2S][H]^+$: 344.0744; IR (neat) ν_{max}/cm^{-1} 2947 (C–H, st), 1727 (C=O), 1258 (N–C, methyl), 828 (S–F); LCMS Purity (UV) = 95%, t_R 18.7 min; mp: 158–158 C.

Methyl 3-{2-[3-(trifluoromethyl)phenyl]acetamido} pyridine-4-carboxylate (13 a)

To methyl 3-aminopyridine-4-carboxylate (100.0 mg, 0.66 mmol) in dichloromethane (5 mL) was added triethylamine (134.0 mg, 1.32 mmol) followed by dropwise addition of 3-(trifluoromethyl)phenyl acetyl chloride (220.0 mg, 0.99 mmol). The reaction mixture was stirred at room temperature for 48 h and monitored by TLC. After 48 h, it was diluted by dichloromethane (10 mL), washed with 2 M HCl (10 mL×3) and the aqueous layer was extracted with dichloromethane (25 mL). The combined organic layers were washed with sodium bicarbonate (25 mL), brine (15 mL×3), dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude material was purified over a column of silica (dichloromethane:methanol; 19:1) to give **13 a**, as a colourless solid (144 mg, 65%), which was used as such for the next step. ^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 10.51 (s, 1H, NH), 8.98 (s, 1H, ArH), 8.47 (d, J = 5.0 Hz, 1H, ArH), 7.73 (s, 1H, ArH), 7.67–7.63 (m, 4H, ArH), 3.87 (s, 2H, CH_2), 3.67 (s, 3H).

Methyl 3-[2-(3-chlorophenyl)acetamido]pyridine-4-carboxylate (13 c)

N,N-Diisopropylethylamine (1.28 g, 9.92 mmol) was added to a stirred solution of 2-(3-chlorophenyl)acetic acid (0.56 g, 3.29 mmol) and hexafluorophosphate azabenzotrizole tetramethyl uronium (HATU, 1.5 g, 3.95 mmol), in DMF (10 mL). The solution was stirred for 30 min at room temperature and turned green. Methyl 3-aminoisonicotinate (0.5 g, 3.29 mmol) was then added to the mixture, which was stirred at room temperature overnight. Once the reaction was completed, water (10 mL) was added to the mixture and the product extracted with ethyl acetate (3×10 mL). This extract was then washed with brine (3×10 mL) dried over $MgSO_4$ and the solvent removed in vacuo. The resultant beige solid was purified by flash chromatography ethyl acetate, hexane (1:1) to yield the target compound (0.66 g, 66%). ^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 10.47 (s, 1H), 9.04 (s, 1H), 8.46 (d, J = 4.9 Hz, 1H), 7.65 (d, J = 4.9 Hz, 1H), 7.44 (s, 1H), 7.40–7.30 (m, 3H), 3.78 (s, 2H), 3.71 (s, 3H). ^{13}C NMR (151 MHz, DMSO-d_6) δ 169.1, 165.8, 145.3, 144.8, 137.5, 132.9, 132.5, 130.2, 129.3, 128.6, 128., 126., 122.3, 52.5, 42.4.

2-{[3-(Trifluoromethyl)phenyl]methyl}-3H,4H-pyrido[3,4-d]pyrimidin-4-one (14 a)

Product **13 a** (144 mg, from above) was refluxed with 7 M ammonia in methanol (10 mL) for 48 hours. After cooling, the crude product was diluted with dichloromethane (15 mL) and washed with 2 M HCl (10 mL×2). The combined organic layers were dried over $MgSO_4$, filtered, and concentrated in *vacuo*. The crude mixture was purified over a column of silica (dichloromethane:methanol; 97:3) to obtain the title compound as a colourless solid (30 mg, 23%). Crystallization of **14 a** was achieved by dissolving it in minimum amount of dichloromethane. ^{1}H NMR (600 MHz, $(CD_3)_2SO$) δ 12.79 (s, 1H, NH), 8.94 (s, 1H, ArH), 8.59 (d, J = 5.1 Hz, 1H, ArH), 7.89 (d, J = 5.1 Hz, 1H, ArH), 7.77 (s, 1H, ArH), 7.67 (d, J = 7.8 Hz, 1H, ArH), 7.62

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

(d, J = 7.8 Hz, 1H, ArH), 7.58–7.52 (m, 1H, ArH), 4.08 (s, 2H, CH_2); ^{13}C NMR (151 MHz, CD_3OD) δ 161.5 (C=O), 158.1 (ArC), 149.8 (ArC), 145.1 (ArC), 143.9 (ArC), 136.7 (ArC), 132.5 (ArC), 130.6 (q, $^2J_{FC}$ = 32.0 Hz, ArC-CF_3), 129.2 (ArC), 126.4 (ArC), 125.5 (q, $^3J_{FC}$ = 4.0 Hz, ArC), 124.2 (q, $^1J_{FC}$ = 271.4 Hz, ArC), 123.7 (q, $^3J_{FC}$ = 4.0 Hz, ArC), 118.4 (ArC), 40.5 (CH_2); ^{19}F NMR (400 MHz, $CDCl_3$) δ −62.71 (s, 3F); HRMS-ESI (*m/z*) found 306.0845, calc. for $[C_{15}H_{10}F_3N_2O][H]^+$:306.0869; LCMS Purity (UV) = 90 %, t_R 14.9 min; IR (neat) ν_{max}/cm^{-1}: 2769 (NH), 1704 (C=O), 1329 (CF_3).

2-{[3-(Pentafluoro-λ⁶-sulfanyl)phenyl]methyl}-3H,4H-pyrido[3,4-]-pyrimidin-4-one (14 b)

To 2-[3-(pentafluoro-λ⁶-sulfanyl)phenyl]acetic acid (120.0 mg, 0.46 mmol) was added thionyl chloride (201.0 μL, 2.76 mmol) and 2 drops of DMF. The reaction was refluxed under argon overnight. The reaction was followed by LCMS and following completion of the reaction, it was concentrated in *vacuo* to obtain 2-[3-(pentafluoro-λ⁶-sulfanyl)phenyl]acetyl chloride. The crude was taken through to the following steps without purification (115 mg).

To methyl 3-aminopyridine-4-carboxylate **12** (41.0 mg, 0.27 mmol) in dichloromethane (5 mL) was added triethylamine (55.0 mg, 0.54 mmol) followed by dropwise addition of 2-[3-(pentafluoro-λ⁶-sulfanyl)phenyl]acetyl chloride ((made as above) 115.0 mg, 0.40 mmol). The reaction mixture was stirred at room temperature for 48 h and monitored by TLC. After 48 h, it was diluted with dichloromethane (10 mL), washed with 2 M HCl (10 mL×3) and aqueous layer extracted with dichloromethane (25 mL). The combined organic layers were washed with sodium bicarbonate (25 mL), brine (15 mL×3), dried over $MgSO_4$, filtered and concentrated in *vacuo*. The crude material was purified using a column of silica (dichloromethane:methanol; 19:1) to obtain crude **13 b**, as a colourless solid (144 mg, 65 %). The latter was used as such and was refluxed with 7 M ammonia in methanol (10 mL) for 48 hours. After cooling, the crude mixture was diluted with dichloromethane (15 mL) and washed with 2 M HCl (10 mL×2). The combined organic layers were dried over $MgSO_4$, filtered, and concentrated in *vacuo*. The crude was purified over a column of silica (dichloromethane:methanol; 97:3) to obtain the title compound as a colourless solid (9 mg, 25 %). ^1H NMR (600 MHz, $CDCl_3$) δ 11.85 (s, 1H, NH), 9.20 (s, 1H, ArH), 8.76 (d, J = 5.1 Hz, 1H, ArH), 8.05 (d, J = 5.1, 1H, ArH), 7.97 (dd, J = 8.0 Hz, 1.9 Hz, 1H, ArH), 7.71–7.67 (m, 1H, ArH), 7.64 (d, J = 7.7 Hz, 1H, ArH), 7.49–7.44 (m, 1H, ArH), 4.19 (s, 2H, CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ 161.7 (C=O), 156.0 (ArC), 154.3–153.7 (m, ArC-SF_5), 151.1 (ArC), 146.3 (ArC), 143.5 (ArC), 136.2 (ArC), 132.3 (ArC), 129.2 (ArC), 126.8 (m, ArC), 126.0 (ArC), 125.1 (m, ArC), 118.1 (ArC), 41.5 (C). ^{19}F NMR (400 MHz, $CDCl_3$) δ 87.64–85.94 (p, J = 148.4 Hz, 1F), 65.21 (d, J = 148.4 Hz, 4F); HRMS-ESI (*m/z*) found 364.0543, calc. for $[C_{14}H_{10}F_5N_3OS][H]^+$:364.0527; LCMS Purity (UV) = 92 %, t_R 16.2 min; IR (neat) ν_{max}/cm^{-1}: 2854 (NH), 1672 (C=O), 834 (SF_5).

5-Methyl-N-[4-(pentafluoro-λ⁶-sulfanyl)phenyl]-1,2-oxazole-4-carboxamide (15)

4-(Pentafluorosulfanyl)aniline (500 mg, 2.28 mmol) was dissolved in isopropanol (IPA) (4 mL) followed by the addition of $NaHCO_3$ (201 mg, 2.39 mmol). The mixture was heated to 60 C and 5-methyl-4-isoxazolecarbonyl chloride (0.2 mL, 2.28 mmol) was added dropwise. The reaction was followed by TLC and after 2 hours of stirring, the reaction was allowed to cool to ambient temperature. The white precipitate formed was filtered under vacuum and washed with water and toluene, then dried to give the product (530 mg, 70 %). ^1H NMR (600 MHz, $CDCl_3$) δ 8.47 (s, 1H, NH), 7.77–7.72 (m, 2H, ArH), 7.67 (d, J = 8.8 Hz, 2H, ArH), 7.55 (s, 1H, ArH), 2.77 (s, 3H, CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ 174.3 (C=O), 159.2 (C=N), 149.7 (C-SF_5), 147.3 (ArC), 139.8 (ArC), 127.2 (2 C, m, ArC), 119.5 (2 C, ArC), 111.6 (ArC), 12.7 (CH_3); ^{19}F NMR (400 MHz, $CDCl_3$) δ 84.62 (p, J = 150 Hz), 63.34 (d, J = 150 Hz). LCMS Purity (UV) = 97 %, tR 20.12 min; HRMS-ESI (m/z) found 329.0385, calc. for $[C_{11}H_9F_5N_2O_2S][H]^+$: 329.0383; IR (neat) νmax/cm^{-1}: 3366 (N–H), 1663 (C=O), 824 (S–F); mp = 167–168 C.

(2Z)-2-Cyano-3-hydroxy-N-[4-(pentafluoro-λ⁶-sulfanyl)phenyl] but-2-enamide (16)

Compound **15** (304 mg, 0.93 mmol) was suspended in a mixture of IPA (2 mL) and water (2 mL). To the stirred mixture was added dropwise an aqueous solution of NaOH (39 % in water) until the solution reached a pH of *ca.* 12, at which point, all of the starting material dissolved. The solution was then filtered and concentrated. Concentrated HCl was added to the filtrate until pH 3. The resultant precipitate was filtered, washed with water and dried under vacuum to obtain the product as a colourless powder (283 mg, 93 %). ^1H NMR (600 MHz, $CDCl_3$) δ 7.79–7.74 (m, 2H, ArH), 7.62 (d, J = 8.8 Hz, 2H, ArH), 2.38 (s, 3H, CH_3); N–H and O–H not found; ^{13}C NMR (151 MHz $CDCl_3$) δ 189.2 (C=O), 167.5 (CN), 150.2 (C-SF_5), 138.6 (ArC), 127.3 (2 C, ArC, m), 120.2 (2 C, ArC), 116.1(C), 80.6 (C), 22.1 (CH_3); ^{19}F NMR (400 MHz, $CDCl_3$) δ 84.82 (p, J = 150.4 Hz), 63.34 (d, J = 150.4 Hz). HPLC Purity (UV) = 99 %, t_R 17.91 min; HRMS-ESI (*m/z*) found 327.0247, calc. for $[C_{11}H_9F_5N_2O_2S][H]^+$: 327.0227; IR (neat) ν_{max}/cm^{-1}: 3312 (N–H), 2219 (CN), 1641 (C=O), 1589 (C=C), 825 (S–F); mp = 167–170 C.

Acknowledgements

The authors acknowledge financial support from the ERDF (LabFact: InterReg V project 121), HEIF Covid-19 Emergency Funds (University of Sussex), the Ewart Bequest Fund (University of Sussex), EPSRC (EP/P026990/1), Oxford COVID development fund, and Wellcome Trust for Institutional Strategic Support Fund (ISSF) (University of Sussex). This work was partly conducted as part of the Covid Moonshot Consortium (https://postera.ai/moonshot/) and we are grateful for their support. The authors also thank the Diamond Light Source for beam time (proposal mx19301) and the staff of beamlines I03 for assistance with crystal testing and data collection. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KgaA Darmstadt Germany, MSD, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, FAPDF, CAPES, CNPq, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome [106169/ZZ14/Z]. CJS thanks the Biotechnology and Biological Research Council, the Wellcome Trust, and Cancer Research UK. This research was funded in whole, or in part, by the Wellcome Trust [grant no. 106244/Z/14/Z]. For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. T. R. M. thanks the BBSRC (BB/M011224/1) for her doctoral funding. N.L. received support from the Weizmann Institute Corona Response Fund.

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conflict of Interest

The authors declare no conflict of interest.

Keywords: SF_5 group · DMARDs · COVID-19, SARS-COV-2 main protease (Mpro)

[1] B. R. Stockwell, *Nature* **2004**, *432*, 846–854.
[2] M. R. Bauer, R. N. Jones, M. G. J. Baud, R. Wilcken, F. M. Boeckler, A. R. Fersht, A. C. Joerger, J. Spencer, *ACS Chem. Biol.* **2016**, *11*, 2265–2274.
[3] S. Purser, P. R. Moore, S. Swallow, V. Gouverneur, *Chem. Soc. Rev.* **2008**, *37*, 320–330.
[4] E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 8315–8359.
[5] E. Pujol, N. Blanco-Cabra, E. Julián, R. Leiva, E. Torrents, S. Vázquez, *Molecules* **2018**, *23*, 1–17.
[6] M. S. Owaileh, R. A. Hazlitt, D. A. Colby, *ChemMedChem* **2017**, 1481–1490.
[7] P. R. Savoie, J. T. Welch, *Chem. Rev.* **2015**, *115*, 1130–1190.
[8] J. M. W. Chan, *J. Mater. Chem. C* **2019**, *7*, 12822–12834.
[9] A. Gilbert, P. Langowski, M. Delgado, L. Chabaud, M. Pucheault, J. F. Paquin, *Beilstein J. Org. Chem.* **2020**, *16*, 3069–3077.
[10] C. R. Pitts, D. Bornemann, P. Liebing, N. Santschi, A. Togni, *Angew. Chem. Int. Ed.* **2019**, *58*, 1950–1954; *Angew. Chem.* **2019**, *131*, 1970–1974.
[11] A. Jose, R. K. Tareque, M. Mortensen, R. Legay, S. J. Coles, G. J. Tizzard, B. W. Greenland, T. G. Smart, M. C. Bagley, J. Spencer, *Tetrahedron* **2021**, 132020.
[12] S. Sansook, C. A. Ocasio, I. J. Day, G. J. Tizzard, S. J. Coles, O. Fedorov, J. M. Bennett, J. M. Elkins, J. Spencer, *Org. Biomol. Chem.* **2017**, *15*, 8655.
[13] D. J. Wood, J. D. Lopez-Fernandez, L. E. Knight, I. Al-Khawaldeh, C. Gai, S. Lin, M. P. Martin, D. C. Miller, C. Cano, J. A. Endicott, et al., *J. Med. Chem.* **2019**, *62*, 3741–3752.
[14] C. Xu, S. De, A. M. Balu, M. Ojeda, R. Luque, *Chem. Commun.* **2015**, *51*, 6698–6713.
[15] S. Haferkamp, A. Paul, A. A. L. Michalchuk, F. Emmerling, *Beilstein J. Org. Chem.* **2019**, *15*, 1141–1148.
[16] S. Haferkamp, F. Fischer, W. Kraus, F. Emmerling, *Beilstein J. Org. Chem.* **2017**, *13*, 2010–2014.
[17] O. Eguaogie, J. S. Vyle, P. F. Conlon, M. A. Gîlea, Y. Liang, *Beilstein J. Org. Chem.* **2018**, *14*, 955–970.
[18] "Ceramic Grinding and Milling Media," can be found under https://zircoa.com/product.fine.grain/grinding.media.html.
[19] L. Sun, N. Tran, F. Tang, H. App, P. Hirth, G. Mcmahon, C. Tang, *J. Med. Chem.* **1998**, *41*, 2588–2603.
[20] T. A. T. Fong, L. K. Shawver, L. Sun, C. Tang, H. App, T. J. Powell, Y. H. Kim, R. Schreck, X. Wang, W. Risau, et al., *Cancer Res.* **1999**, *59*, 99–106.
[21] The COVID Moonshot Consortium, H. Achdout, A. Aimon, E. Bar-David, H. Barr, A. Ben-Shmuel, J. Bennett, M. L. Bobby, J. Brun, B. Sarma, et al., *bioRxiv* **2020**, DOI 10.1101/2020.10.29.339317.
[22] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox, R. Hilgenfeld, *Science* **2020**, *368*, 409–412.
[23] L. Zhang, D. Lin, Y. Kusov, Y. Nian, Q. Ma, J. Wang, A. Von Brunn, P. Leyssen, K. Lanko, J. Neyts, et al., *J. Med. Chem.* **2020**, *63*, 4562–4578.
[24] S. J. Coles, P. A. Gale, *Chem. Sci.* **2012**, *3*, 683–689.
[25] K. Hu, M. Wang, Y. Zhao, Y. Zhang, T. Wang, Z. Zheng, X. Li, S. Zeng, D. Zhao, H. Li, et al., *Virol. Sin.* **2020**, *35*, 725–733.
[26] S. Liu, E. A. Neidhardt, T. H. Grossman, T. Ocain, J. Clardy, *Structure* **2000**, *8*, 25–33.
[27] D. Cowen, P. Bedingfield, G. A. McConkey, C. W. G. Fishwick, A. P. Johnson, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1284–1287.
[28] N. Hail, P. Chen, J. J. Kepa, L. R. Bushman, C. Shearn, *Free Radical Biol. Med.* **2010**, *49*, 109–116.
[29] A. Ilya, N. Gershon, A. Liberman, *US2002022646A1.Pdf*, **2002**, US 2002/0022646A1.
[30] S. Christian, C. Merz, L. Evans, S. Gradl, H. Seidel, A. Friberg, A. Eheim, P. Lejeune, K. Brzezinka, K. Zimmermann, et al., *Leukemia* **2019**, *33*, 2403–2415.
[31] B. Halford, **2021**; DOI: 10.47287/cen-09913-scicon3.
[32] T. R. Malla, A. Tumber, T. John, L. Brewitz, C. Strain-Damerell, C. D. Owen, P. Lukacik, H. T. H. Chan, P. Maheswaran, E. Salah, et al., *Chem. Commun.* **2021**, *57*, 1430–1433.

Manuscript received: October 1, 2021
Revised manuscript received: November 24, 2021
Version of record online: February 22, 2022

© 2022 The Authors. ChemMedChem published by Wiley-VCH GmbH

18607187, 2022, 7, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmdc.202100641 by Oxford University, Wiley Online Library on [06/07/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

COVID Moonshot: Open Science Discovery of SARS-CoV-2 Main Protease Inhibitors by Combining Crowdsourcing, High-Throughput Experiments, Computational Simulations, and Machine Learning

The COVID Moonshot Consortium[1]

[1]Full list of authors: https://tinyurl.com/y3r7redd

***For correspondence:**
john.chodera@choderalab.org (John Chodera); alpha.lee@postera.ai (Alpha Lee); nir.london@weizmann.ac.il (Nir London); frank.von-delft@diamond.ac.uk (Frank von Delft)

Abstract Herein we provide a living summary of the data generated during the COVID Moonshot project focused on the development of SARS-CoV-2 main protease (Mpro) inhibitors. Our approach uniquely combines crowdsourced medicinal chemistry insights with high throughput crystallography, exascale computational chemistry infrastructure for simulations, and machine learning in triaging designs and predicting synthetic routes. This manuscript describes our methodologies leading to both covalent and non-covalent inhibitors displaying protease IC50 values under 150 nM and viral inhibition under 5 uM in multiple different viral replication assays. Furthermore, we provide over 200 crystal structures of fragment-like and lead-like molecules in complex with the main protease. Over 1000 synthesized and ordered compounds are also reported with the corresponding activity in Mpro enzymatic assays using two different experimental setups. The data referenced in this document will be continually updated to reflect the current experimental progress of the COVID Moonshot project, and serves as a citable reference for ensuing publications. All of the generated data is open to other researchers who may find it of use.

Since early 2020, when the coronavirus SARS-CoV-2 was implicated as the cause of the pandemic soon to engulf the world, the need for targeted therapeutic development has been evident due to the dearth of existing antiviral treatments. Realizing this need, researchers worked rapidly to determine the structure of the novel SARS-CoV-2 main protease (MPro) [1], a promising antiviral drug target involved in SARS-CoV-2 viral replication. Other researchers quickly followed with a crystallographic and electrophilic fragment-screen of the structure [2].

Recognizing the necessity for a large-scale, international effort to further develop antiviral compounds, the COVID Moonshot project was created to build on these early results and advance the development of Mpro inhibitors [3]. The project brought together academic and industrial partners from across the world to combine experimental and computational expertise towards this urgent goal.

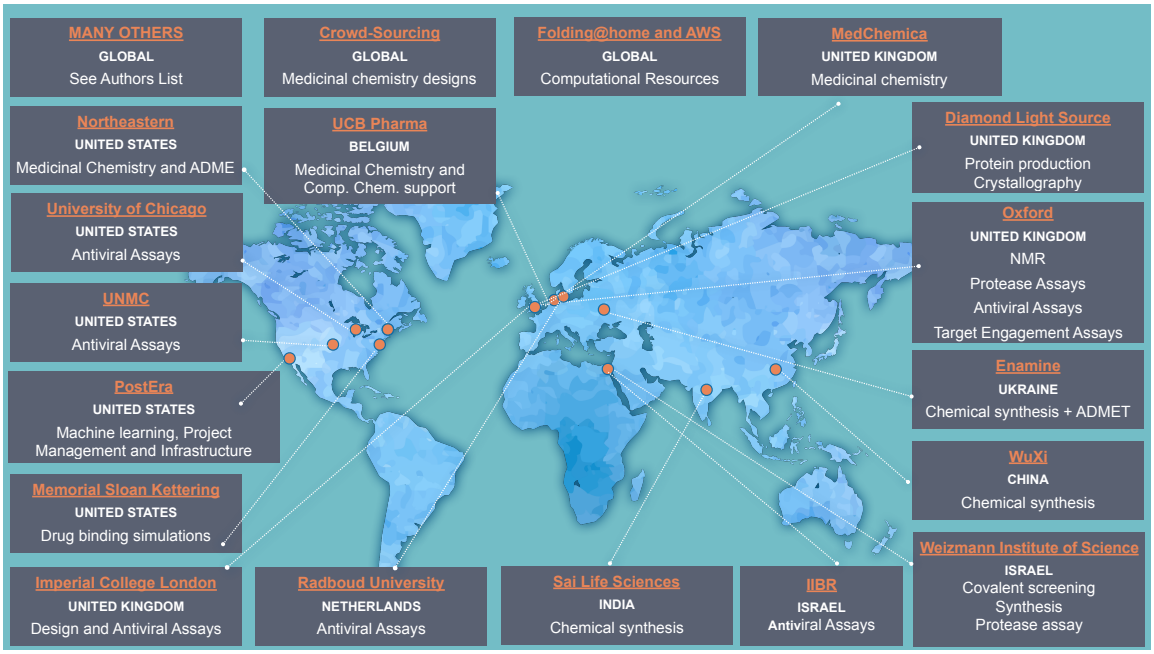


Figure 1. COVID Moonshot is a global open-science effort to develop a patent-free antiviral. The project is an international effort involving over 30 groups (see author list), many of whom have provided work pro-bono or at-cost. The project is IP-free and has relied on philanthropic donations up to this point.

The COVID Moonshot project has focused on progressing early fragment-screening results into potent compounds with activity against both the main protease and the virus. Design of compounds was done through a novel crowd sourcing mechanism, harnessing the expertise of drug-designers across the world. Building on their individual sets of expertise, these designers have entered over 1000 submissions for a total of over 10,000 individual molecule designs to date. The rationales for each design include docking-based approaches, by-eye structure-based designs, machine learning approaches, crawling of the past literature on SARS and MERS compounds, and other general medicinal-chemistry insights that can be visualised at https://postera.ai/covid. In the background, a High-Throughput Screen of 185,000 compounds has been carried out to further aid with lead generation. The results of this screen will be reported in a future manuscript.

Out of these myriad designs, over 1000 compounds have been either ordered or synthesized and tested for activity against MPro in assays at two different sites. The enzymatic activity has been measured using two complementary biochemical assays assessing inhibition of the SARS-CoV-2 main protease: a fluorescence based assay (Weizmann Institute) and RapidFire Mass Spectrometry assay (University of Oxford). Furthermore, structural characterization of selected compounds has resulted in over 200 crystal structures, when combined with the fragment screening data.

Particularly promising compounds have been sent through a further ADMET assay cascade, and sufficiently suitable compounds have been assessed in cellular antiviral screening assays performed by collaborators at the University of Oxford, the University of Nebraska Medical Centre, Israel Institute for Biological Research and the Radboutd University Medical Centre.

All design and experimental data can be found on the COVID Moonshot website at https://postera.ai/covid, which also links to many crystal structures found at https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro. The complete raw data in csv format with detailed description is located at the Moonshot GitHub repository, https://github.com/postera-ai/COVID_moonshot_submissions. The data is frequently updated and we kindly ask that all questions regarding the data be directed at the active COVID Moonshot forum https://discuss.postera.ai/c/covid/general/10.

This is an ongoing project, and this manuscript is very much a "living document". The reason for the publication of this document is to provide readers and ensuing publications a central point to reference

data generated from COVID Moonshot.

Figures 2 and 3 exemplify our progress – from crowdsourced fragment merges (TRY-UNI-714a760b-6) to antiviral activity that is comparable to that of Remdesivir.

1 Methodology

1.1 Crowdsourced design campaign

Our approach combines the collective creativity of the scientific community with algorithms that aid with synthesis planning and free energy calculations for prioritizing crowdsourced designs. Specifically, as we started with a dense set of fragment hits with over 70 protein-fragment structures, a crowdsourcing approach allows us to sample different creative merges and expansions of the different fragments.

Those crowdsourced designs are then triaged based on a lax docking-based threshold. Docking was performed by several groups pursuing several different approaches—of both covalent and non-covalent compounds. Consensus docking results were used to filter out designs which demonstrably do not fit the binding site, a difficult task to do by-eye given the sheer number of designs. We then triage compounds based on synthetic accessibility to ensure short cycle times. As the total number of commercially available virtual compounds exceeds 10 billion, keeping track of synthetic accessibility is a non-trivial task. Complex-looking molecules might be synthesizable in a single step, while simple functionalizations could require a complex multi-step synthesis. During the project, consortium member PostEra built a new tool (https://postera.ai/manifold) in order to accelerate this difficult task. Additional custom synthesis tools from PostEra allow for rapid prioritization of compounds and generation of synthetic routes. Furthermore, the algorithms are tailored to the building blocks available at each of our partner CROs, such that we can algorithmically allocate load.

The resulting compounds are synthesized at partner CROs, and all compounds are subjected to: (1) MPro RapidFire mass spectrometry assay, (2) MPro fluorescence assay, (3) High throughput X-ray crystallography. Non-covalent compounds are also analysed by NMR. Early on the use of two complementary protease assays was deemed important due to the rapid development of procedures and the need for verification. Agreement between the label free MS based approaches and fluorophore-based approaches have been reassuring, while discrepancies between assays have led to discovery of false-positives. High throughput crystallography has allowed for rapid generation of design ideas that merge early binding fragments with more recent actives generated during the project.

As soon as the data is cleaned and standardized, the results of all assays are released as on the Moonshot platform (http://postera.ai/covid and https://discuss.postera.ai/c/covid/general/10), ready for the crowd to submit followup compounds.

We received over 13,000 diverse designs from more than 400 contributors. Our data-driven approach enabled the synthesis and testing of over 1,000 compounds in less than 6 months. Some members of the community informally organised themselves into a medicinal chemistry design team, with weekly meetings to go through the data and suggest new designs. Moreover, exascale computing resources donated by Folding@Home enable Moonshot to use Free Energy Perturbation to prioritise compounds.

1.2 Fluorescence MPro inhibition assay

Compounds were seeded into assay-ready plates (Greiner 384 low volume 784900) using an Echo 555 acoustic dispenser, and DMSO was back-filled for a uniform concentration in assay plates (maximum 1%). Screening assays were performed in duplicate at 20 µM and 50 µM. Hits of greater than 50% inhibition at 50 µM were confirmed by dose response assays. Reagents for Mpro assay reagents were dispensed into the assay plate in 10μl volumes for a final volume of 20 µL. Final reaction concentrations were 20 mM HEPES pH 7.3, 1 mM TCEP, 50 mM NaCl, 0.01% Tween-20, 10% glycerol, 5 nM Mpro, 375 nM fluorogenic peptide substrate ([5-FAM]-AVLQSGFR-[Lys(Dabcyl)]-K-amide). Mpro was pre-incubated for 15 minutes at room temperature with compound before addition of substrate. Protease reaction was measured continuously in a BMG Pherastar FS with a 480/520 ex/em filter set. Data analysis was performed with Collaborative Drug Discovery (CDD).

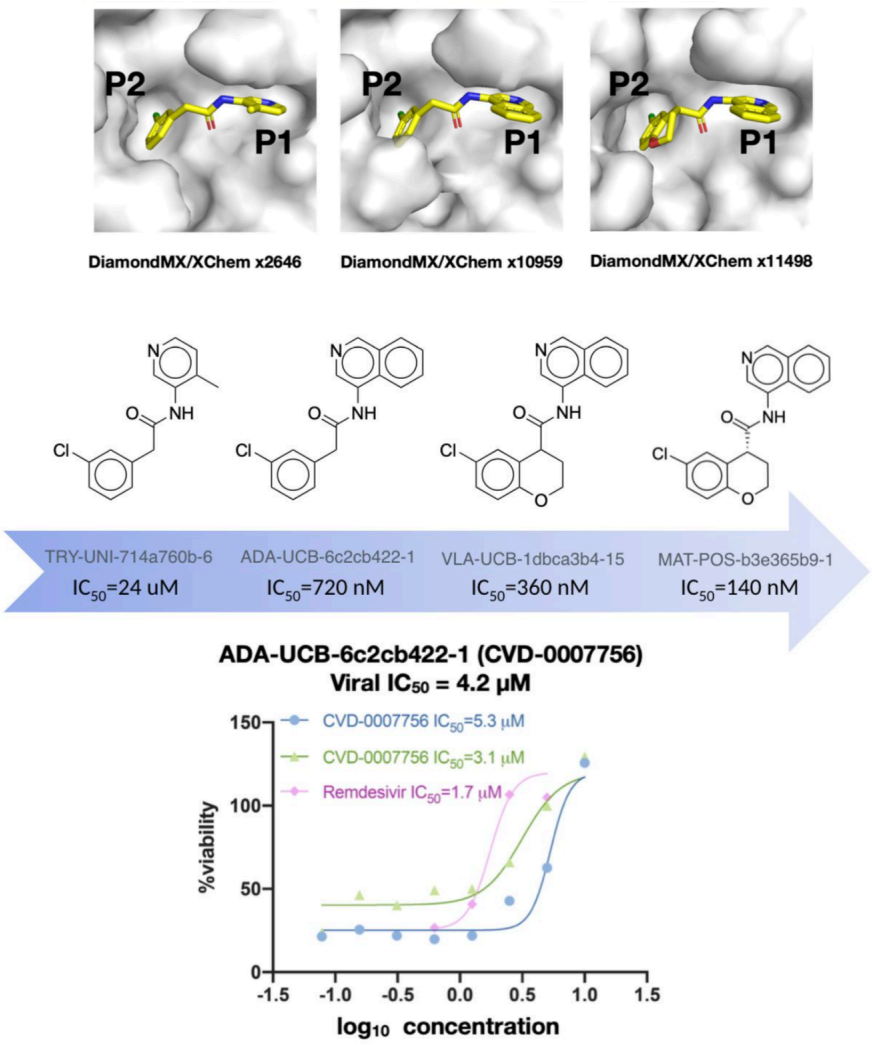


Figure 2. The progression of a fragment merge (TRY-UNI-714a760b-6) into potent, non-covalent MPro inhibitors. One of these compounds ADA-UCB-6c2cb422-1, has also been tested in an antiviral assay and showed promising activity. Antiviral activity was measured at the Israel Institute for Biological Research (IIBR) with Remdesivir serving as the positive control. The more potent analogs of ADA-UCB-6c2cb422-1 have not yet been tested for antiviral activity. The development of these series will be further detailed in a future publication.

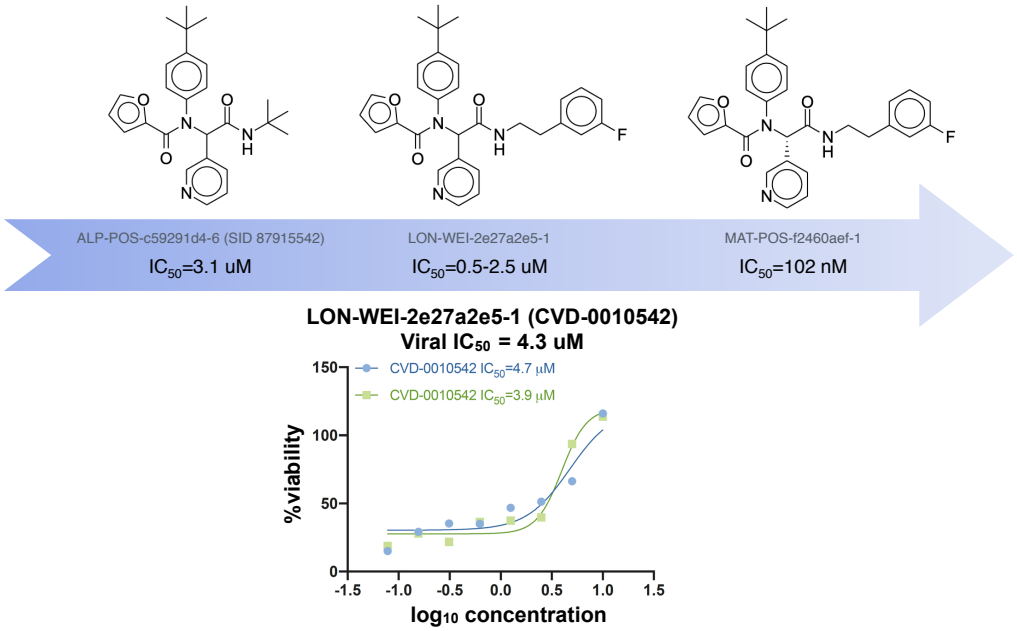


Figure 3. The development of a SARS-CoV-1 non-covalent MPro inhibitor (Jacobs et al. [4]) into a potent SARS-CoV-2 MPro inhibitor with antiviral activity. LON-WEI-2e27a2e5-1 has also been tested in the IIBR antiviral assay and showed activity. The development of these series will be further detailed in a future publication.

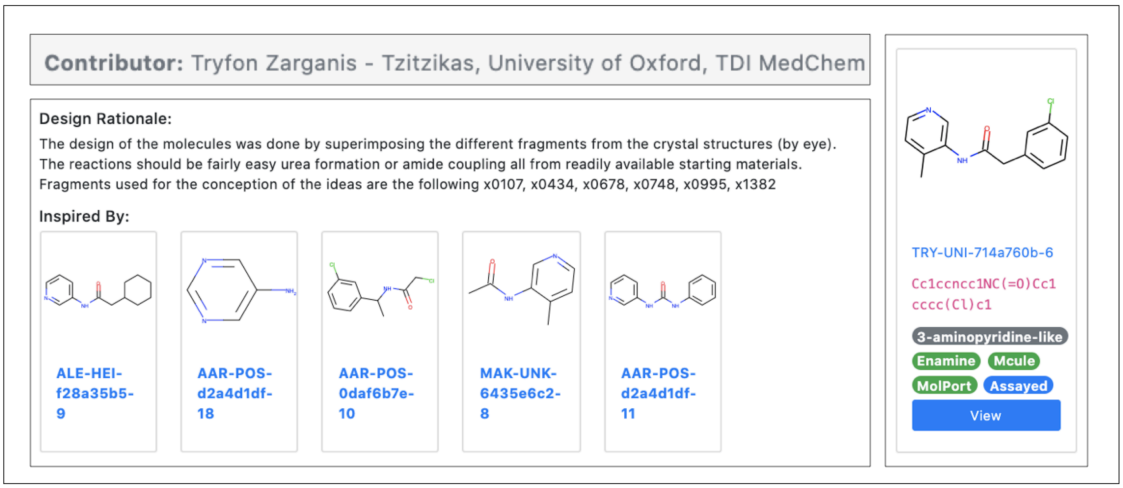


Figure 4. The original crowdsourced design of the fragment merge (TRY-UNI-714a760b-6) that led to the development of the lead series shown in Figure 2. The system to enter compounds with rationales, and add attribution to the original fragment inspirations was built by PostEra specifically for this project. All linked biochemical and structural results on this specific molecule can be viewed on the detail page: https://postera.ai/covid/submissions/714a760b-0e02-4b09-8736-f27f854f8c22/6

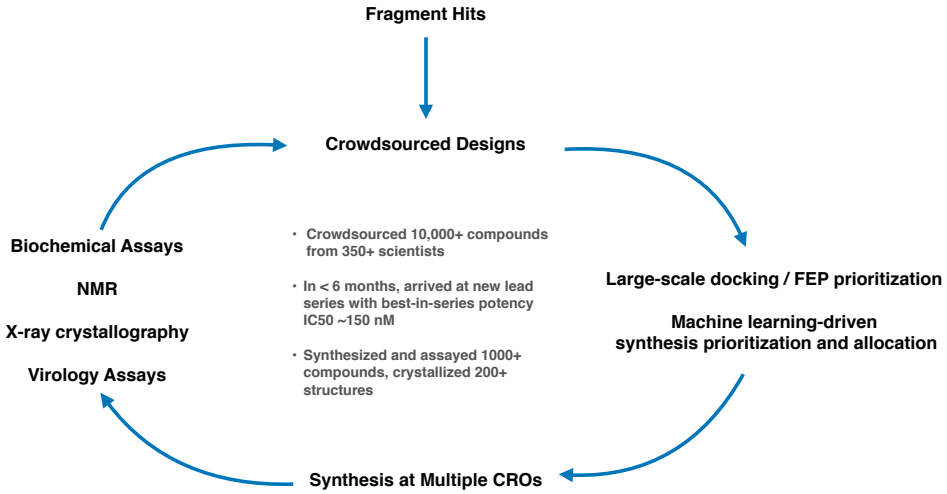


Figure 5. Overall workflow for the COVID Moonshot project. These interconnected steps take place among teams distributed internationally among numerous researchers who have never met in person. Coordination of these synergistic efforts has taken a significant amount of logistical planning amid lockdowns and other complications resulting from the COVID pandemic.

1.3 RapidFire MPro inhibition assay

Inhibitor compounds at 20mM in DMSO are dispensed into 384-well plates (Greiner 384PP 781280) using an ECHO 650T dispenser (DMSO concentration < 1%, final volume = 500 nL.). A 15 µM enzyme stock solution is prepared in 20 mM HEPES, pH 7.5 and 300 mM NaCl, and subsequently diluted to a working solution of 300 nM Mpro in assay buffer (20 mM HEPES, pH 7.5 and 50 mM NaCl) before the addition of 25 µL to each well using a Multidrop Combi (Thermo Scientific). After a quick centrifugation step (1 000 rpm, 15 s) the plate is incubated for 15 min at room temperature. The reaction is initiated with the addition of 25 µL of 4 µM substrate (TSAVLQSGFRK-NH2, initially custom synthesized by the Schofield group, GLBiochem) dissolved in assay buffer. After centrifugation (1000 rpm, 14 s) the reaction is incubated for 10 min at room temperature before quenching with 10% formic acid. The reactions are analysed with MS using RapidFire (RF) 365 high-throughput sampling robot (Agilent) connected to an iFunnel Agilent 6550 accurate mass quadrupole time-of-flight (Q-TOF) mass spectrometer using electrospray. All compounds are triaged by testing the % inhibition at 5 and 50 µM final concentration. Dose response curves uses an 11-point range of 100–0.0017 µM inhibitor concentrations. RapidFire integrator software (Agilent) was used to extract the m/z (+1) charge states of both the substrate (1191.67 Da) and cleaved N-terminal product TSAVLQ (617.34 Da) from the total ion chromatogram data followed by peak integration. Percentage conversion (product peak integral / (product peak integral + substrate peak integral))*100) and percentage inhibitions were calculated and normalised against DMSO control with deduction of any background signal in Microsoft Excel. IC_{50}s were calculated using Levenberg–Marquardt algorithm used to fit a restrained Hill equation to the dose-response data with both GraphPad PRISM and CDD.

1.4 Saturation Transfer Difference NMR (STD-NMR) assay

Method described in Kantsadi and Vakonakis [5].

1.5 High throughput x-ray crystallography

Method described in Douangamath et al. [2].

1.6 High Throughput Screening (HTS) of SARS-CoV-2 Mpro

184,631 compounds from the G-INCPM screening collection, as well as a 3,000 Mpro targeted compound library from Enamine, were screened against Mpro using a fluorogenic protease assay. A screening cascade was established as follows: 1) hits were defined as showing a greater than 50% inhibition were selected, filtered for obvious structural alerts and promiscuity in unrelated assays 2) 885 hits were re-tested in three independent Mpro assay runs to confirm activity 3) 128 confirmed hits in 5 concentrations were tested in two independent Mpro assay runs to determine preliminary potency as IC_{50}. Promising compounds were re-sourced, and are awaiting validation.

HTS was performed in both 384-well (20 µL) and 1536-well (10 µM, Greiner 782076) formats. Compounds were pre-plated into barcoded assay ready plates for a final concentration of 10 µM (0.1% DMSO). The assay was performed using Thermo combi dispensers with 1 µL tubing integrated with plate storage and BMG Pherastar FS plate reader and Spinnaker plate mover driven by Momentum software (Thermo). For 1536-well screening, plates were unloaded and spun for 1 min at 1500 RPM prior to plate reading. Data was loaded to Screener software (Genedata) for mapping, normalization and annotation. Curated data was then loaded to CDD Vault for merging with compound annotations.

The results of the HTS screen can be viewed publicly at https://github.com/postera-ai/COVID_moonshot_submissions

1.7 High-throughput solubility threshold measurement

Nephelometry-based solubility assay to define threshold compound solubility at the concentrations of 20 µM and 100 µM in PBS solution. The compound solubilities are normalized to Deoxyfluorouridine (100% soluble) and ondansetron (0% solubility). To put the numbers in perspective, the expected relative solubility values are:

- High >0.8
- Mid 0.6 - 0.8
- Low <0.6

To perform the screening, 245 μL aliquots of PBS buffer were added to each well of 96-well microplates with clear flat bottom. Plates with the buffer were subjected to optical integrity inspection using Nephelostar. Total scanning time for one 96-well plate was 3 min. The pass criterion was set as the background signal in any of the scanned wells below 25 RNU, thus making optical quality of the plates satisfactory for the assay. In case of an excessively high background signal in any wells of the test-plates, those plates/wells were excluded from the study (data not shown). Compounds were obtained from Enamine repository as solids formatted in polypropylene, round bottom blank tubes, in latch boxes (Matrix 4271). Compounds were dissolved in DMSO at 50 mM, incubated at 24–26°C for 8 hours, shaken for 1 hour at 1800 rpm using high-speed microplate shaker Illumina, then incubated at 24–26°C for 14 hours. Intermediate DMSO solutions of the tested compounds were prepared to 2 mM (for the tested concentration 20 µM) and 8 mM (for the tested concentration 100 µM). Ondansetron and DOFU were added to the final plates in 100 mM concentration to get 1 mM and 2 mM final concentrations in the assay plates. To prepare the test plates, 2.5 µL aliquots of DMSO solutions of the tested compounds, reference compounds or pure DMSO were transferred from the polypropylene tubes to the corresponding wells of 96-well plates with PBS buffer using Plate Mate, according to the plate map (Figure 1). Thus, final volumes in the test plates were brought to 247.5 µL, resulting in concentrations of the compounds in test wells of 20 µM and 1% DMSO, correspondingly. Then, turbidity of the solutions was immediately scanned for each well.

1.8 Synthesis prediction

We employ an approach based on the Molecular Transformer technology [6, 7]. Our algorithm uses natural language processing to predict the outcomes of chemical reactions and design retrosynthetic routes starting from commercially available building blocks. This proprietary platform is provided free of charge by PostEra Inc (http://postera.ai). Additionally, Manifold (https://postera.ai/manifold) was built by PostEra Inc. during the

project to search the entire space of purchasable molecules, and automatically find the optimal building blocks.

1.9 Viral screening assays

A variety of antiviral replication assays were performed in collaborating laboratories, including cytopathic effect (CPE) inhibition assays at the IIBR, Israel, viral qPCR at Radboud, Netherlands, immunofluorescence assays at University of Nebraska Medical centre, USA, and plaque assays and focus forming unit assays at University of Oxford, UK.

1.9.1 Antiviral assays Cytopathic Effect CPE (IIBR, Ness-Ziona, Israel)

SARS-CoV-2 (GISAID accession EPI_ISL_406862) was kindly provided by Bundeswehr Institute of Microbiology, Munich, Germany. Virus stocks were propagated (4 passages) and tittered on Vero E6 cells. Handling and working with SARS-CoV-2 virus was conducted in a BSL3 facility in accordance with the biosafety guidelines of the Israel Institute for Biological Research (IIBR). Vero E6 were plated in 96-well plates and treated with compounds in medium containing 2% fetal bovine serum. The assay plates containing compound dilutions and cells were incubated for 1 hour at 37^{o}C temperature prior to adding Multiplicity of infection (MOI) 0.01 of viruses. Viruses were added to the entire plate, including virus control wells that did not contain test compound and Remdesivir drug used as positive control. After 72h incubation viral cytopathic effect (CPE) inhibition assay was measured with XTT reagent. Three replicate plates were used.

1.9.2 Antiviral assays qPCR (Radboud University Medical Center, Nijmegen, Netherlands)

Cells. African green monkey Vero E6 kidney cells (ATCC CRL-1586) and Vero FM kidney cells (ATCC CCL-81) were cultured in Dulbecco's modified Eagle medium (DMEM) with 4.5g/L glucose and L-glutamine (Gibco), supplemented with 10% Fetal calf serum (FCS, Sigma Aldrich), 100 µg/ml streptomycin and 100 U/ml penicillin (Gibco). Cells were maintained at 37^{o}C with 5% CO2.

Virus. SARS-CoV-2 (isolate BetaCoV/Munich/BavPat1/2020) was kindly provided by Prof. C. Drosten (Charité-Universitätsmedizin Berlin Institute of Virology, Berlin, Germany) and was initially cultured in Vero E6 cells up to three passages in the laboratory of Prof. Bart Haagmans (Viroscience Department, Erasmus Medical Center, Rotterdam, The Netherlands). Vero FM cells were infected with passage 3 stock at an MOI of 0.01 in infection medium (DMEM containing L-glutamine, 2% FCS, 20 mM HEPES buffer, 100 µg/ml streptomycin and 100 U/ml penicillin). Cell culture supernatant containing virus was harvested at 48 hours post-infection (hpi), centrifuged to remove cellular debris, filtered using a 0.2 μm syringe filter (Whatman), and stored in 100 μl aliquots at -80^{o}C.

Compounds. 10mM compound stocks dissolved in DMSO were provided by Enamine (Ukraine).

Virus stock titration. Vero E6 cells were seeded in 12-well plates at a density of 500,000 cells/well. Cell culture medium was discarded at 24 h post-seeding, cells were washed twice with PBS and infected with 10-fold dilutions of the virus stock. At 1 hpi, cells were washed with PBS and replaced with overlay medium containing Minimum Essential medium (MEM), 2% FCS, 20mM HEPES buffer, 100μg/ml streptomycin, 100 U/ml penicillin and 0.75% carboxymethyl cellulose (Sigma Aldrich). At 48 hpi, the overlay medium was discarded, cells were washed with PBS and stained with 0.25% crystal violet solution containing 4% formaldehyde for 30 minutes. Thereafter, staining solution was discarded, plates washed with double-distilled water, dried and plaques were counted.

Antiviral assay. Vero E6 cells were seeded onto 24-well plates at a density of 150,000 cells/well. At 24 h post-seeding, cell culture medium was discarded, cells were washed twice with PBS and infected with SARS-CoV-2 at an MOI of 0.01 in the presence of six concentrations of the inhibitors (25μM – 0.06μM). At 1 hpi, the inoculum was discarded, cells were washed with PBS, and infection medium containing the same concentration of the inhibitors was added to the wells. SARS-CoV-2 infection in the presence of 0.1% DMSO was used as a negative control. At 24 hpi, 100 μl of the cell culture supernatant was added to RNA-Solv reagent (Omega Bio-Tek) and RNA was isolated and precipitated in the presence of glycogen according to manufacturer's instructions. qRT-PCR. TaqMan Reverse Transcription reagent and random hexamers (Applied Biosystems) were used for cDNA synthesis. Semi-quantitative real-time PCR was performed using GoTaq

qPCR (Promega) BRYT Green Dye-based kit using primers targeting the SARS-CoV-2 E protein gene (Forward primer, 5-ACAGGTACGTTAATAGTTAATAGCGT-3'; Reverse primer, 5-ACAGGTACGTTAATAGTTAATAGCGT-3'). A standard curve of a plasmid containing the E gene qPCR amplicon was used to convert Ct values relative genome copy numbers.

Cell viability. Vero E6 cells were seeded in 96-well white-bottom culture plates (Perkin Elmer) at a density of 30,000 cells per well. At 24 h post-seeding, cells were treated with the same concentrations of compounds as used for the antiviral assay. Cells treated with 0.1% DMSO were used as a negative control. At 24 h post-treatment, cell viability was assessed using the Cell Titer Glo 2.0 kit (Promega) using the Victor Multilabel Plate Reader (Perkin Elmer) to measure luminescence signal.

1.9.3 Antiviral assays Immunoflourescence (Pathology and Microbiology, University of Nebraska Medical Centre, USA, St Patrick Reid)

Vero E6 cells were pretreated with 20 uM of the Moonshot compounds for around 2h. Cells were then infected with SARS-CoV-2 at a MOI of 0.1 for 24h. Virus infection was terminated by 4% PFA fixation. Cells were stained using a Rabbit SARS-CoV-2 antibody (Sino Biological40150-R007) as a primary antibody, and Alexa-488, Hoechst and Cell Mask (Thermo Fisher) as a secondary antibody. Images were collected on the Operetta system imaging system, and analysed using the Harmony software.

1.9.4 Antiviral assays Plaque assay and Focus Forming Unit Assay (University of Oxford, UK)

Cell culture. The African green monkey Vero E6 cell line (ATCC CRL-1586) was cultured in Dulbecco's modified Eagle medium (DMEM) with Glutamax supplemented with 100 µg/mL streptomycin, 100 U/mL penicillin, and 10% heat-inactivated fetal calf serum (FCS). The human lung cancer cell line Calu-3 (Anderson Ryan, Department of Oncology, Medical Science Division, University of Oxford) was cultured in a 1:1 mixture of DMEM with Glutamax and Ham's F-12 medium supplemented with 100 µg/mL streptomycin, 100 U/mL penicillin, and 10% heat-inactivated FCS. All cells were maintained as mycoplasma free, with regular verifications by polymerase chain reaction (PCR).

Virus propagation. SARS-CoV-2 England/2/2020 was provided at passage 1 from Public Health England, Collindale. Passage 2 submaster and passage 3 working stocks were produced by infecting Vero E6 cells at a multiplicity of infection of 0.01 in virus propagation medium (DMEM with Glutamax supplemented with 2% FCS) and incubating until cytopathic effect was visible. The cell supernatant was then centrifuged at 500 g for 5 minutes, aliquoted and stored at -80°C. The titre of viral stocks was determined by plaque assay. All subsequent assays were performed using a passage 3 stock.

Cell viability. Cell viability was was measured using the CellTiter 96 R AQueous One Solution Cell Proliferation MTA (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H - 15 tetrazolium, inner salt) Assay (Promega) according to the manufacturer's instruction after treatment with compound. Briefly, Calu 3 cells were treated with compounds in quadruplicate for 3 days. Wells with 200 µL growth medium with and without cells were included as controls in quadruplicate. After the incubation, 100 µL of growth medium was removed and 20 µL of MTS reagent was added to the remaining medium in each well. After a further one to two hour incubation, the absorbance at 490 nm was measured on a Molecular Devices SpectraMax M5 microplate reader.

Antiviral assays. Calu 3 cells were seeded into 96 well plates (VWR) at a cell density of 8×10^4 cells/well. After 48 hours, the wells were inoculated for 90 minutes with 50 L SARS-CoV-2 England 2/2020 at a M.O.I. of 0.5. The inoculum was replaced with 200 L growth medium containing the appropriate compound dilutions or matched controls and incubated for 3 days. The supernatant was harvested and stored at -80°C prior to analysis by plaque or focus forming unit (FFU) assay. For plaque assays, four 10-fold serial dilutions of each supernatant to be analysed were prepared in virus propagation medium. 100l of each dilution was inoculated in triplicate into wells of a 24-well plate followed by 0.5 mL of Vero E6 cells at 5×10^5 cells/mL in virus propagation medium. The plates were incubated for 2 hours before the addition of 0.5 mL of 1.8% CMC overlay (1:1 mix of 3.6% carboxymethylcellulose in H_2O and virus propagation medium). The plates were incubated for a further 4 days before staining with Napthol blue black. Plaques were counted by eye. For Focus forming unit assays, a SARS-CoV-2 Microneutralization assay from the W James lab (Dunn School

of Pathology, University of Oxford) was adapted for use as a FFU assay. Briefly, 3 half log dilutions of each supernatant to be analyzed were prepared in virus propagation medium. 20 µL of each dilution was inoculated into wells of a 96-well plate in quadruplicate followed by 100 L Vero E6 cells at 4.5×10^5 cells/mL in virus propagation medium. The plates were incubated for 2 hours prior to the addition of 100 L of 1.8% CMC overlay. The plates were incubated for a further 24 hours. After 24 hours the overlay was carefully removed and the cells washed once with PBS before fixing with 50 µL of 4% paraformaldehyde, after 30 minutes the paraformaldehyde was removed and replaced with 100 µL of 1% ethanolamine in PBS. The cells were permeabilized by replacing the ethanolamine with 2% Triton X100 in PBS and incubating at 37°C for 30 minutes. The plates were then washed 3 times with wash buffer (0.1% Tween 20 in PBS) inverted and gently tapped onto tissue to dry before the addition of 50 l of EY2A anti-N human mAb (Arthur Huang (Taiwan)/Alain Townsend (Weatherall Institute of Molecular Medicine, University of Oxford)) at 10 pmol in wash buffer. The plates were rocked at room temperature for 1 hour, washed and incubated with 100l of secondary antibody Anti-Human IgG (Fc-specific)-peroxidase-conjugate produced in Goat diluted 1:5000 at room temperature for a further hour. 50 µL of TrueBlue peroxidase substrate was added to the wells and incubated at RT for 10 min on the rocker, after 10 minutes the substrate was removed and the plates washed with ddH_2O for 10 minutes. The H2O was removed and the plates allowed to air dry. The foci were then counted using an ELISPOT classic reader system (AID GmbH).

References

[1] **Jin Z**, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020; p. 1–5.

[2] **Douangamath A**, Fearon D, Gehrtz P, Krojer T, Lukacik P, Owen CD, Resnick E, Strain-Damerell C, Ábrányi-Balogh P, Brandaõ-Neto J, et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. Nature Communications. 2020; 11:5047.

[3] **Chodera J**, Lee AA, London N, von Delft F. Crowdsourcing drug discovery for pandemics. Nature Chemistry. 2020; 12(7):581–581.

[4] **Jacobs J**, Grum-Tokars V, Zhou Y, Turlington M, Saldanha SA, Chase P, Eggler A, Dawson ES, Baez-Santos YM, Tomar S, et al. Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. Journal of medicinal chemistry. 2013; 56(2):534–546.

[5] **Kantsadi AL**, Vakonakis I. Rapid assessment of ligand binding to the SARS-CoV-2 main protease by saturation transfer difference NMR spectroscopy. bioRxiv. 2020; .

[6] **Lee AA**, Yang Q, Sresht V, Bolgar P, Hou X, Klug-McLeod JL, Butler CR, et al. Molecular transformer unifies reaction prediction and retrosynthesis across pharma chemical space. Chemical Communications. 2019; 55(81):12152–12155.

[7] **Schwaller P**, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, Lee AA. Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. ACS central science. 2019; 5(9):1572–1583.

Europe PMC Funders Group
Author Manuscript
Eur J Med Chem. Author manuscript; available in PMC 2023 July 04.

Published in final edited form as:
Eur J Med Chem. 2021 April 05; 215: 113257. doi:10.1016/j.ejmech.2021.113257.

Faropenem reacts with serine and metallo-β-lactamases to give multiple products

Anka Lucic[#a], **Philip Hinchliffe**[#b], **Tika R. Malla**[#a], **Catherine L. Tooke**[b], **Jürgen Brem**[a], **Karina Calvopiña**[a], **Christopher T. Lohans**[c], **Patrick Rabe**[a], **Michael A. McDonough**[a], **Timothy Armistead**[b], **Allen M. Orville**[d], **James Spencer**[b,+], **Christopher J. Schofield**[a,+]

Christopher T. Lohans: christopher.lohans@queensu.ca; Allen M. Orville: allen.orville@diamond.ac.uk

[a]Chemistry Research Laboratory, The Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, 12 Mansfield Road, Oxford OX1 3TA, United Kingdom

[b]Cellular and Molecular Medicine, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, United Kingdom

[c]Botterell Hall, Kingston, Ontario, Canada, K7L 2V7

[d]Diamond House, Harwell Science and Innovation Campus, Didcot, OX11 0DE

[#] These authors contributed equally to this work.

Abstract

Penems have demonstrated potential as antibacterials and β-lactamase inhibitors; however, their clinical use has been limited, especially in comparison with the structurally related carbapenems. Faropenem is an orally active antibiotic with a C2 tetrahydrofuran (THF) ring, which is resistant to hydrolysis by some β-lactamases. We report studies on the reactions of faropenem with carbapenem-hydrolysing β-lactamases, focusing on the class A serine β-lactamase KPC-2 and the metallo β-lactamases (MBLs) VIM-2 (a subclass B1 MBL) and L1 (a B3 MBL). Kinetic studies show that faropenem is a substrate for all three β-lactamases, though it is less efficiently hydrolysed by KPC-2. Crystallographic analyses on faropenem-derived complexes reveal the opening of the β-lactam ring with formation of an imine with KPC-2, VIM-2, and L1. In the cases of the KPC-2 and VIM-2 structures, the THF ring is opened to give an alkene, but with L1 the THF ring remains intact. Solution state studies, employing NMR, were performed on L1, KPC-2, VIM-2, VIM-1, NDM-1, OXA-23, OXA-10, and OXA-48. The solution results reveal, in all cases, formation of imine products in which the THF ring is opened; formation of a THF ring-closed imine product was only observed with VIM-1 and VIM-2. An enamine product with a closed THF ring was also observed in all cases, at varying levels. Combined with previous reports, the results exemplify the potential for different outcomes in the reactions of penems with MBLs and SBLs

[+]Correspondence to be sent to these authors. christopher.schofield@chem.ox.ac.uk, jim.spencer@bristol.ac.uk.

Author Contributions

A.L., P.H. and C.L.T. crystallised proteins and solved and refined the X-ray data with help from M.A.M and J.B. T.R.M measured the NMR data and analysed with help from C.T.L. A.L., P.H., C.L.T, P.R., K.C and C.T.L produced enzymes. P.H., C.L.T and T.A. collected kinetic data. A.L., P.H., T.R.M., and C.J.S. drafted the manuscript. P.H., J.S. and C.J.S. designed and conceived the research. All authors edited the manuscript.

Conflicts of Interest

The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

and imply further structure-activity relationship studies are worthwhile to optimise the interactions of penems with β-lactamases. They also exemplify how crystal structures of β-lactamase substrate/inhibitor complexes do not always reflect reaction outcomes in solution.

Keywords

antimicrobial resistance; β-lactams; penems; carbapenems; β-lactams; serine-β-lactamases; metallo-β-lactamases

1.0 Introduction

β-Lactams are presently the most prescribed antibacterials, though their efficacy is being eroded by antimicrobial resistance[1]. The best characterised of such mechanisms involves β-lactamase-catalysed hydrolysis (Figure 1). β-Lactamases of Ambler classes A, C and D[2] are nucleophilic serine enzymes (serine β-lactamases, SBLs). Class B comprises of zinc-dependent metallo-β-lactamases (MBLs)[3,4], which are divided into sub classes B1, B2 and B3, based on sequence, structure and zinc ion utilisation. The B1 and B3 MBLs employ two zinc ions, whereas the B2 MBLs employ a single zinc ion[5].

Clinically-available carbapenems are synthetic antibiotics that are used intravenously to treat serious bacterial infections[7]. Although once considered last-resort drugs, in part because of their inhibition of some SBLs, extensive carbapenem use has resulted in the dissemination of carbapenem-hydrolysing β-lactamases, such as the class A SBL *Klebsiella pneumoniae* carbapenemase-2 (KPC-2), which is now present in multiple Gram-negative pathogens, including *Klebsiella pneumoniae*, *Escherichia coli* and *Acinetobacter spp.*[8,9]. Some class D SBLs (including OXA-48, OXA-23 and OXA-10[10]) can also manifest carbapenemase activity. The role of MBLs in carbapenem resistance is growing; indeed, B1 (e.g. VIM and NDM MBLs) and B3 (e.g. L1) MBLs can efficiently hydrolyse almost all β-lactam antibiotics, excluding monobactams (Figure 1)[3,11].

One factor in determining the β-lactamase susceptibility of β-lactam substrates and inhibitors is the stability of the complexes formed between the enzyme and the substrate / inhibitor. The ability of carbapenems to form relatively stable complexes with both their antibacterial transpeptidase targets and SBLs is an important factor in their efficacy relative to other β-lactams[12]. However, the precise factors that determine the stability of such complexes are incompletely understood, as exemplified by a recent report that carbapenems can react with class D SBLs to produce β-lactones as well as hydrolysed products (Figure 2)[10,13]. Studies on the products of carbapenem hydrolysis by representative MBLs and SBLs suggest that the initial hydrolysis products are predominantly in the enamine tautomeric form, and can undergo subsequent non-enzymatic tautomerisation to the more stable imine isomer[10,14].

Studies on carbapenems[11,12] led us to investigate the hydrolysis mechanism of penems[15] by selected SBLs and MBLs. Penems are closely related to the carbapenems by structure but contain a thiazoline rather than a pyrroline ring (Figure 1). The penems were developed in the 1970s in attempts to combine the structures of the penicillins and

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

cephalosporins to give more stable and efficacious drugs[16]; they have a broad spectrum of activity against both Gram negative and Gram positive bacteria[17] although they have little efficacy against *Pseudomonas aeruginosa*[18]. Penems also have the potential to act as potent SBL inhibitors, as exemplified by BRL 42715[19], which potently inhibits SBLs, including those from the OXA, SHV and TEM subfamilies[20]. BRL 42715 reacts with the nucleophilic serine to form an acyl-enzyme complex which subsequently rearranges to form a cyclic amino-acrylate derivative (Figure S1)[19].

Faropenem is an orally active penem antibiotic with a tetrahydrofuran (THF) C-2 side chain; it is used extensively in Asia to treat respiratory and urinary tract infections[21,22]. Faropenem is currently the only penem in clinical use, with sulopenem presently in clinical trials[23]. Like the carbapenems, faropenem inhibits some SBLs, including the class C enzyme AmpC and extended spectrum SBLs such as BlaC and members of the CTX-M family[24,25]. Faropenem is, however, susceptible to hydrolysis by carbapenemases such as the B1 MBLs, e.g. NDM-1. There are reported crystal structures of complexes formed by the reaction of faropenem with transpeptidases (PBP3 and LdtMt2)[26,27] and SBLs (KPC-2, BlaC, OXA-48 and SED-G238C)[28], though these crystal structures were not correlated with the reaction products in solution (Figure 3).

Here we report studies using steady-state kinetics, X-ray crystallography, and NMR on the reaction of faropenem with SBLs and MBLs. We initially focused on the class A SBL KPC-2 and the class B1 and B3 MBLs VIM-2 and L1. VIM-2 is a widely disseminated, clinically important, broad spectrum B1 MBL[29]. The B3 MBL L1 is a broad spectrum MBL, chromosomally encoded by *Stenotrophomonas maltophilia* and which has the same overall fold as class B1 MBLs, though it differs in its Zn2 ion coordination[30]. The combined production of L1 and a class A SBL (L2) makes infections by *S. maltophilia* difficult to treat. Crystallographic analyses of faropenem-derived complexes suggest that with KPC-2 and VIM-2, the THF ring opens to give an alkene product, but with L1 the ring remains intact. Subsequent NMR solution studies on the faropenem derived reaction products with multiple β-lactamases reveal the formation of both imine and enamine products The THF ring was observed in both the opened and closed forms, with the precise product profiles being enzyme specific. Note in contrast to crystallographic findings, the THF ring intact hydrolysis product was not observed for L1.

2 Results

2.1 Faropenem hydrolysis by β-lactamases in solution

We initially determined steady state kinetic parameters for faropenem hydrolysis by the carbapenemases KPC-2, VIM-2, and L1 by monitoring the hydrolysis of the faropenem β-lactam ring at λ 306 nm (with 10 nM enzyme) (Table 1). Comparison of turnover numbers (k_{cat}) reveals that, under saturating conditions, VIM-2 and L1 hydrolyse faropenem at a rate ~10 times greater than KPC-2, though the L1 K_M is higher than that for KPC-2 or VIM-2. The catalytic efficiency (k_{cat}/K_M) of KPC-2 with faropenem is ~5- and 33-fold lower than for L1 or VIM-2, respectively. We also observed weak substrate (faropenem) inhibition of KPC-2 (K_i ~44 μM).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

2.2 Crystallographic determination of faropenem products

We next sought to investigate how the faropenem-derived products interact with the active sites of VIM-2, KPC-2, and L1. To obtain a crystal structure of the faropenem derived complex with KPC-2, we substituted Glu166, which is important for deacylation of the acyl-enzyme intermediate (Figure 1), for the isosteric Gln (KPC-2^{E166Q}) that prevents antibiotic turnover[33]. Pre-formed KPC-2^{E166Q}, VIM-2, and L1 crystals were then soaked with a faropenem solution (see methods for details) to obtain X-ray diffraction datasets with resolutions extending to 1.25 Å, 1.29 Å and 1.57 Å, respectively (Table S1). VIM-2 crystallised in two different space groups, *I*222 and *C*2 [32,33] with one or two protein molecules in the asymmetric unit, respectively. KPC-2 and L1 crystallised in space groups $P2_12_12$[34] and $P6_422$[35], both with one molecule in the asymmetric unit (Figure S2). In all four datasets, electron density in the active site is consistent with the presence of a faropenem-derived product with an opened β-lactam ring (Figure 4).

2.2.1 KPC-2 Structure—With KPC-2, the faropenem-derived species is covalently bound to Ser70; the C2 atom is sp^2 hybridized and its THF ring is opened to give an imine alkene product (Figures 4D and S3). The exocyclic C-2 – C-11 alkene (Figure S3 and Figure S4A) is present as the Z-isomer, a feature that could not be resolved in our NMR studies (see below). An analogous complex was also observed with VIM-2 (see below). Electron density for THF-derived atoms is more poorly defined than for the rest of the ligand in both KPC-2 and VIM-2, manifesting as higher B-factors, likely due to a lack of defined interactions with protein (Figure S4B-D). A prior crystallographic study of KPC-2 with faropenem resulted in the observation of a non-covalently bound hydrolysed product at the active site, which was modelled and refined as the enamine form with the THF ring intact[36] (Figure 3); however, electron density is notably absent for the entire THF side chain, so the exact nature of the ligand is uncertain.

The β-lactam derived carbonyl oxygen is positioned in the oxyanion hole, forming hydrogen bonds with the backbone amides of Ser70 and Thr237, consistent with observations for acyl-enzyme complexes of other SBLs (Figure 4D and Figure S3A)[6]. Additional polar interactions include the C-10 carboxylate with Thr235 and the C-6 hydroxyethyl oxygen with Asn132. A hydrophobic interaction of the fragmented THF ring is made with Trp105 (Figure S3B), a residue involved in interactions with substrates and SBL inhibitors, e.g. relebactam [34]. The interactions of hydrolysed faropenem with KPC-2 resemble those of meropenem with SFC-1 (PDB 4EV4[37]), a class A SBL that, like KPC-2, can hydrolyse carbapenems (Figure S5). Notably, the hydroxyethyl chain is orientated away from the deacylating water in both the KPC-2: faropenem and SFC-1: meropenem-derived structures, being positioned to form a hydrogen bond with Asn132. This orientation is suggested to be favourable for hydrolysis, whereas in carbapenem-inhibited class A enzymes, the deacylating water is thought to be deactivated by interaction with the hydroxyethyl group[37].

In our KPC-2: faropenem complex, there is a conformational change in the catalytically important Ω-loop, relative to its location in the unreacted enzyme (Figure S6). This results in two conformations of Gln166, (which is the general base for the deacylation step [38]).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

In one conformation Gln166 is directed into the active site in a manner that enables it to interact with the deacylating water; in the other conformation, it is directed out of the active site (Figure S6). We have previously observed a similar movement of Glu166 in molecular dynamics simulations of KPC-2 with hydrolysed ceftazidime in the active site, which we attributed as the major factor for the poor turnover of ceftazidime by KPC-2 [33]. Therefore, this conformational dynamic may also contribute to the relatively slow turnover rate for KPC-2 with faropenem (Table S1).

2.2.2 VIM-2 Structure—Interactions of the faropenem-derived hydrolysis product with VIM-2 are equivalent in both the *I*222 (soaking time: 30 mins) and *C*2 (soaking time: 5 hours) structures (Figure 4A, 4B and Figure S7; standard MBL numbering is used throughout[39]). As with KPC-2, faropenem reacts with VIM-2 to give an imine-alkene, as crystallographically observed (Figure 1). One of the oxygen molecules of the newly formed C-7 carboxylate displaces the Zn-bridging water molecule (Figures 4 and 5). This results in an increase in the Zn1-Zn2 distance (4.2 Å) compared to the un-complexed VIM-2 (PDB 4BZ3, 3.5 Å), as has been observed on the binding of some MBL inhibitors, e.g. bicyclic boronates, to VIM-2 (Figure S8)[39]. The C-7 carboxylate hydrogen bonds with the Asn233 sidechain amide group (Figure 5) and the C-6 hydroxyethyl oxygen hydrogen bonds with the Asp120 backbone amide. Binding is further stabilised by hydrophobic interactions between faropenem-derived elements and Phe61 and Tyr67 on loop L3 (a region of the active site often involved in interactions with substrates and inhibitors[40–42]) and Trp87 in the active site (Figure S7B and S7C).

2.2.3 L1 Structure—Crystallographic analysis reveals that the reaction of faropenem with L1 results in an imine with the THF ring remaining closed; this occurs *via* protonation at C-2, which is clearly defined by the electron density as being sp^3 hybridized and in the (*S*)-configuration (Figure 4C and Figure S9). The C-7 carboxylate and C-6 hydroxyethyl sidechain are rotated in comparison with the orientations observed in the KPC-2 and VIM-2 faropenem complexes (Figure S10). Furthermore, in contrast to binding to VIM-2, in the L1 complex the Zn-bridging water is present and is stabilised by interaction with the rotated hydroxyethyl alcohol. By contrast with what has normally been reported for interactions of carbapenems with MBLs, the C-7 carboxylate does not interact with Zn1, but instead with Tyr32. Interactions of the faropenem-derived product with the L1 metal centre thus exclusively involve Zn2, which is ligated by the nitrogen and carboxylate of the thiazoline ring, as in the VIM-2 complexes. The C-7 carboxylate also forms hydrogen bonds with Ser221 (3.1 Å) and Ser223 (2.7 Å), residues previously reported to be involved in L1 binding to inhibitors and substrates [43,44].

The faropenem binding mode to L1 is similar to that observed for imipenem binding to the B3 MBL SMB-1 (PDB 5B1U), in which the hydrolysed carbapenem substrate is observed in the imine form (C-2 is sp^3 hybridized) and the C-7 carboxylate does not displace the Zn-bridging water (Figure S11). Further, in both structures, C-2 of the antibiotic-derived product is in the (*S*)-configuration. Comparison of the VIM-2:faropenem complex with the complex formed by reaction of meropenem with NDM-1[45], in which the deprotonated Δ^2-pyrroline (enamine) form was observed (C-2 is sp^2 hybridized) at 2.12 Å

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

resolution (Figure S12A), reveals similar binding modes in which the zinc bridging active site water is displaced. However, comparison with meropenem binding to VIM-1 shows differences in the positioning of the C-6 hydroxyethyl group that results in the presence of the zinc bridging active site water in the VIM-1:meropenem complex, but not in the VIM-2:faropenem complex (Figure S12B)[42].

2.3 Faropenem hydrolysis studied by NMR

We then investigated the relevance of the crystallographic observations to the β-lactamase-catalysed hydrolysis of faropenem in solution by using 1D and 2D NMR methods, initially with L1, KPC-2, and VIM-2 (Figure 6, S13-28). For comparison, we subsequently studied the reaction of faropenem with other SBLs/MBLs, i.e. the class D SBLs OXA-10, OXA-23 and OXA-48 and the B1 MBLs NDM-1 and VIM-1. Figures S13-28 show individual spectra; peaks were assigned based on the values in Table S2. The ^{1}H-NMR studies for KPC-2 (Figure S14), VIM-2 (Figure S16), L1 (Figure S13) and NDM-1 (Figure S15) were consistent with the prior kinetic analyses (Table 1), including with respect to KPC-2 manifesting the poorest rate of faropenem hydrolysis. In total, five faropenem-derived product structures were assigned, though low levels of other unassigned species were also observed (Figures 6, S13 and S15). Compounds **1** and **2** (major and minor, respectively) were observed in all cases and were assigned as the C-5 epimeric forms of the THF ring-closed enamines, with the (5*R*) minor product (**2**) arising from epimerisation of the major (5*S*) enamine (**1**). **1** is likely to be the major enzymatic product, consistent with studies on carbapenems [11,12]. The amounts of **1** and **2** decreased with time, due to their conversion to the imine-alkenes **3** and **4**, respectively, as confirmed by time-dependent analyses of evolution of faropenem-derived products on reaction with VIM-2 (Figure S28A).

Compounds **3** and **4** were observed in all cases and were assigned as (5*R*)-**3** and (5*S*)-**4** imines with an opened THF ring. The Z-alkene geometry was assigned on the basis of crystallographic analyses (see above) – it is likely but not certain that this is the alkene geometry in solution (Figures 6 and 7). **3** and **4** were observed to be formed from **1** and **2** non-enzymatically when using NMR to study the hydrolysis of faropenem alone, but at least in some cases may be direct enzyme products. The imine with a closed THF ring, **5,** was only observed with VIM-1 (minor product) and VIM-2 (a major product) (Figures 6, S16 and S17). NMR experiments following the time course of faropenem reaction with VIM-2 confirm this, with an initial increase in the amounts of compounds **1** and **2.** The presence of **1** and **2** decreases with time however, and there is a gradual increase of **3, 4** and **5** with time (Figure S28A).

The stereochemistry of **5** at C-2 was not assigned but is likely to be (2*S*), based upon our high-resolution crystallographic studies with L1 (Figures 4, 5 and S13). We did not obtain evidence for the C-5 epimer of **5**, i.e. **6** (Figure S8), though we cannot rule **6** out as a minor product. Note that C-5 epimerisation should be more facile for the enamines (**1**, **2**) than the imine tautomers (**3**-**6**). Thus, **5** is likely produced as direct result of VIM-1/VIM-2 catalysis; notably we saw no evidence for conversion of **1/2** to **5/6** (Figures 6 and 7).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Overall, the solution studies provide important information additional to the crystallographic observations (Figure 7). In the case of the L1 complex, the imine with the closed THF ring **5*** (*= crystallographically observed, Figure 7) was observed; however, **5** was not detected in the solution studies with L1 (Figures 6 and S13). It is possible that, in the case of L1, the crystallographically observed **5*** results from tautomerisation of L1 complexed enamine **1*** under the crystallographic analysis conditions. It is also possible that the crystallographically observed product **5*** is not released from the L1 active site in solution, though this seems less likely. Further, in the L1 crystal structures the C-2 configuration is clearly defined as (*S*), which could not be resolved by solution NMR. This observation contrasts with the recent crystal structure of a faropenem-derived product in OXA-48 in which the imine was modelled with C-2 in the (*R*)-configuration[46]. However, we note that the lower resolution of these OXA-48 crystallographic data [46], compared to L1 (2.10 Å vs 1.57 Å), makes assignment of the C-2 configuration in the OXA-48 structure uncertain. Although the same (2*S*)-**5*** product was modelled in crystal structure of the class A BlaC enzyme (PDB 4EBL), the electron density for the side chain in that case is poorly defined and inspection of the electron density is indicative of an sp^2 hybridized C-2 atom, ruling out the **5*** complex. In the KPC-2/VIM-2 crystal structures there was no evidence for formation of the imine with a closed THF ring (**5***); instead the THF ring-opened imine-alkene **3***, was modelled (Figures 6 and 7) and was observed in solution (**3**). However, whilst **3** could be produced as a direct enzyme product arising from **3***, the combined results indicate that it is more likely **3** that is at least predominantly produced from enamine **1** *via* **1*** (Figure 6); however it should be noted that in the case of VIM-1 **4** is a major product (Figures S17 and S28).

3 Discussion

From a chemical perspective, the penems are important in the history of β-lactams because, as hybrid penicillin-cephalosporin molecules, they were predicted to have improved properties as drugs. As yet, this has not been the case, at least in terms of major clinical impact. The penems are not natural products, unlike (as far as is known) the subsequently-identified naturally occurring but synthetically produced carbapenems that have undergone significant development to now include clinically-important antibacterials[12]. The reasons for the relative lack of development of the penems are unclear; they may reflect difficulties in synthesis (these should be surmountable), stability (this should also be surmountable), and/or toxicity. Despite these potential issues, given that faropenem is orally active and all carbapenems are injected medicines (carbapenem prodrugs are reported), it would seem more research on penems is merited. Only one penem has been approved for clinical use, i.e., faropenem, the subject of our study.

Our results, together with those previously reported[10,13], further highlight the potential of penems as β-lactamase inhibitors. They imply that, once faropenem has reacted with the β-lactamase, the acyl-enzyme (SBL) or analogous Zn(II) complexed intermediate (MBL) can undergo various fates depending on the precise nature of the active site. One possibility is protonation of the β-lactam derived nitrogen followed by release of a hydrolysed enamine (**1**, Figures 6 and 7), in a manner analogous to that proposed for the release of hydrolysed carbapenems[47]. Alternatively, at least in the case of some, but not all, MBLs (VIM-1 and

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

VIM-2, but not L1 or NDM-1 in our studies), the enamine can tautomerise with protonation at C-2 to give an imine product with a closed THF ring (**5**). The nascent product (**1**) can epimerise at C-5 to give **2**; **1** and **2** can also rearrange to give imines with the THF ring opened (**3**,**4**), though in some cases we cannot rule out that **3** and/or **4** are direct enzyme products (Figures S17 and S28B). The results also show that subtle differences between active sites can alter reaction outcomes and kinetics, and that static crystallographic studies do not necessarily reflect solution observations, hence inhibition profiles. Further experiments should be done to see if any of these compounds display inhibitory properties against MBLs and/or SBLs.

Previous studies have also shown that when reacted with nucleophilic cysteine enzymes, e.g. the LdtMT2 transpeptidase from *M. tuberculosis*, faropenem can react to give a thioester linked complex that can undergo C-5-C-6 fragmentation[14,47]. Other modes of penem reactivity can also be envisaged, for example by the conjugate addition to the alkene in the intact penem. An alternative mode of penem reactivity is manifested by C-6 exomethylene penems such as BRL-42715, a potent inhibitor of some SBLs, which reacts to give an acyl-enzyme complex that converts to an aminoacrylate[19,20].

The combined results presented here and previously thus demonstrate that initially formed enzyme-penem complexes can have many potential fates – a property that might be useful in identifying a single compound that potently inhibits multiple different classes of enzymes (e.g. SBLs/MBLs/transpeptidases), by reacting in multiple ways to give stable enzyme-inhibitor complexes. Such an ability is manifest in certain other β-lactam containing β-lactamase inhibitors, e.g. some clavams / oxapenems[19,20]. Although other classes of non β-lactam containing β-lactam inhibitors might have this potential, very few (if any) have the demonstrated clinical utility of bicyclic β-lactams. We thus suggest there is considerable reason for renewed efforts to synthesise new penems and related β-lactam containing compounds for testing as β-lactamase inhibitors and antibiotics. Drug discovery efforts on penems and related compounds should perhaps involve extensive synthetic efforts combined with screening against multiple clinically relevant bacteria and enzyme targets, and rigorous characterisation of the resulting products and inhibitory species, in order to investigate the full potential of compounds that can react *via* alternative reaction pathways.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Medical Research Council MRC (MR/T016035/1), EPSRC, National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH, R01AI100560 to J.S.) and Wellcome Trust for funding our research and Diamond Light Source and their beamline scientists for aiding with data collection. TRM is funded by the BBSRC (BB/M011224/1). PR thanks the Deutsche Akademie für Naturforscher Leopoldina, Germany, for funding a postdoctoral fellowship. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

References

[1]. Fair RJ, Tor Y. Antibiotics and Bacterial Resistance in the 21st Century. Perspect Medicin Chem. 2014; 6 doi: 10.4137/PMC.S14459

[2]. Bush K, Jacoby GA. Updated functional classification of beta-lactamases. Antimicrob Agents Chemother. 2010; 54: 969–76. DOI: 10.1128/AAC.01009-09 [PubMed: 19995920]

[3]. Bush K. Metallo-beta-lactamases: a class apart. Clin Infect Dis. 1998; 27 (Suppl 1) S48–53. accessed May 29, 2018 [PubMed: 9710671]

[4]. Bebrone C. Metallo-β-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. Biochem Pharmacol. 2007; 74: 1686–1701. DOI: 10.1016/j.bcp.2007.05.021 [PubMed: 17597585]

[5]. Lisa M-N, Palacios AR, Aitha M, González MM, Moreno DM, Crowder MW, Bonomo RA, Spencer J, Tierney DL, Llarrull LI, Vila AJ. A general reaction mechanism for carbapenem hydrolysis by mononuclear and binuclear metallo-β-lactamases. Nat Commun. 2017; 8: 538. doi: 10.1038/s41467-017-00601-9 [PubMed: 28912448]

[6]. Tooke CL, Hinchliffe P, Bragginton EC, Colenso CK, Hirvonen VHA, Takebayashi Y. β-Lactamases and β-Lactamase Inhibitors in the 21st Century. J Mol Biol. 2019; 431: 3472–3500. DOI: 10.1016/J.JMB.2019.04.002 [PubMed: 30959050]

[7]. Bush K, Bradford PA. β-Lactams and β-Lactamase Inhibitors: An Overview. Cold Spring Harb Perspect Med. 2016; 6 doi: 10.1101/cshperspect.a025247

[8]. Leavitt A, Navon-Venezia S, Chmelnitsky I, Schwaber MJ, Carmeli Y. Emergence of KPC-2 and KPC-3 in Carbapenem-Resistant Klebsiella pneumoniae Strains in an Israeli Hospital. Antimicrob Agents Chemother. 2007; 51: 3026–3029. DOI: 10.1128/AAC.00299-07 [PubMed: 17562800]

[9]. Shen P, Wei Z, Jiang Y, Du X, Ji S, Yu Y, Li L. Novel genetic environment of the carbapenem-hydrolyzing beta-lactamase KPC-2 among Enterobacteriaceae in China. Antimicrob Agents Chemother. 2009; 53: 4333–8. DOI: 10.1128/AAC.00260-09 [PubMed: 19620332]

[10]. Lohans CT, van Groesen E, Kumar K, Tooke CL, Spencer J, Paton RS, Brem J, Schofield CJ. A New Mechanism for β-Lactamases: Class D Enzymes Degrade 1β-Methyl Carbapenems through Lactone Formation. Angew Chemie Int. 2018; 57: 1282–1285. DOI: 10.1002/anie.201711308

[11]. Walsh TR, Toleman MA, Poirel L, Nordmann P. Metallo-beta-lactamases: the quiet before the storm? Clin Microbiol Rev. 2005; 18: 306–25. DOI: 10.1128/CMR.18.2.306-325.2005 [PubMed: 15831827]

[12]. Papp-Wallace KM, Endimiani A, Taracila MA, Bonomo RA. Carbapenems: Past, Present, and Future. Antimicrob Agents Chemother. 2011; 55: 4943. doi: 10.1128/AAC.00296-11 [PubMed: 21859938]

[13]. van Groesen E, Lohans CT, Brem J, Aertker KM, Claridge TD, Schofield C. 19F-NMR Monitoring of Reversible Protein Post-Translational Modifications: Class D β-Lactamase Carbamylation and Inhibition. Chem – A Eur J. 2019; doi: 10.1002/chem.201902529

[14]. Lohans CT, Chan HTH, Malla TR, Kumar K, Kamps JJAG, McArdle DJB, van Groesen E, de Munnik M, Tooke CL, Spencer J, Paton RS, et al. Non-Hydrolytic β-Lactam Antibiotic Fragmentation by l,d -Transpeptidases and Serine β-Lactamase Cysteine Variants. Angew Chemie Int. 2019; 58: 1990–1994. DOI: 10.1002/anie.201809424

[15]. Dalhoff A, Janjic N, Echols R. Redefining penems. Biochem Pharmacol. 2006; 71: 1085–1095. DOI: 10.1016/J.BCP.2005.12.003 [PubMed: 16413506]

[16]. Ernest I, Gosteli J, Greengrass CW, Holick W, Pfaendler HR, Woodward RB. The penems, a new class of .beta.-lactam antibiotics: 6-acylaminopenem-3-carboxylic acids. J Am Chem Soc. 1978; 100: 8214–8222. DOI: 10.1021/ja00494a032

[17]. Hamilton-Miller JMT. Chemical and Microbiologic Aspects of Penems, a Distinct Class of β-Lactams: Focus on Faropenem. Pharmacotherapy. 2003; 23: 1497–1507. DOI: 10.1592/phco.23.14.1497.31937 [PubMed: 14620395]

[18]. Okamoto K, Gotoh N, Nishino T. Pseudomonas aeruginosa reveals high intrinsic resistance to penem antibiotics: penem resistance mechanisms and their interplay. Antimicrob Agents Chemother. 2001; 45: 1964–71. DOI: 10.1128/AAC.45.7.1964-1971.2001 [PubMed: 11408209]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

[19]. Michaux, Catherine; Charlier, Paulette; Frère, Jean-Marie; Wouters, Johan. Crystal Structure of BRL 42715, C6-(N1-Methyl-1,2,3-triazolylmethylene)penem. Complex with Enterobacter cloacae 908R β-Lactamase: Evidence for a Stereoselective Mechanism from Docking Studies. 2005; doi: 10.1021/JA0426241

[20]. Coleman K, Griffin DR, Page JW, Upshon PA. In vitro evaluation of BRL 42715, a novel beta-lactamase inhibitor. Antimicrob Agents Chemother. 1989; 33: 1580–7. DOI: 10.1128/aac.33.9.1580 [PubMed: 2817854]

[21]. Qu X, Yin C, Sun X, Huang S, Li C, Dong P, Lu X, Zhang Z, Yin A. Consumption of antibiotics in Chinese public general tertiary hospitals (2011-2014): Trends, pattern changes and regional differences. PLoS One. 2018; 13 doi: 10.1371/JOURNAL.PONE.0196668

[22]. Gandra S, Klein EY, Pant S, Malhotra-Kumar S, Laxminarayan R. Faropenem Consumption is Increasing in India. Clin Infect Dis. 2016; 62: 1050.2–1052. DOI: 10.1093/cid/ciw055

[23]. Karlowsky JA, Adam HJ, Baxter MR, Denisuik AJ, Lagacé-Wiens PRS, Walkty AJ, Puttagunta S, Dunne MW, Zhanel GG. In Vitro Activity of Sulopenem, an Oral Penem, against Urinary Isolates of Escherichia coli. Antimicrob Agents Chemother. 2019; 63 doi: 10.1128/AAC.01832-18

[24]. Dhar N, Dubée V, Ballell L, Cuinet G, Hugonnet J-E, Signorino-Gelo F, Barros D, Arthur M, McKinney JD. Rapid cytolysis of Mycobacterium tuberculosis by faropenem, an orally bioavailable β-lactam antibiotic. Antimicrob Agents Chemother. 2015; 59: 1308–19. DOI: 10.1128/AAC.03461-14 [PubMed: 25421469]

[25]. Schneider I, Queenan AM, Markovska R, Markova B, Keuleyan E, Bauernfeind A. New variant of CTX-M-type extended-spectrum beta-lactamases, CTX-M-71, with a Gly238Cys substitution in a Klebsiella pneumoniae isolate from Bulgaria. Antimicrob Agents Chemother. 2009; 53: 4518–21. DOI: 10.1128/AAC.00461-09 [PubMed: 19620330]

[26]. Lu Z, Wang H, Zhang A, Liu X, Zhou W, Yang C, Guddat L, Yang H, Schofield CJ, Rao Z. Structures of Mycobacterium tuberculosis Penicillin-Binding Protein 3 in Complex with Five β-Lactam Antibiotics Reveal Mechanism of Inactivation. Mol Pharmacol. 2020; 97: 287–294. DOI: 10.1124/mol.119.118042 [PubMed: 32086254]

[27]. EM S, G S, R S. Binding and processing of β-lactam antibiotics by the transpeptidase Ldt Mt2 from Mycobacterium tuberculosis. FEBS J. 2017; 284 doi: 10.1111/FEBS.14010

[28]. P OA, Z X, C Y. Molecular Basis of Substrate Recognition and Product Release by the Klebsiella Pneumoniae Carbapenemase (KPC-2). J Med Chem. 2017; 60 doi: 10.1021/ACS.JMEDCHEM.7B00158

[29]. Aitha M, Marts AR, Bergstrom A, Møller AJ, Moritz L, Turner L, Nix JC, Bonomo RA, Page RC, Tierney DL, Crowder MW. Biochemical, mechanistic, and spectroscopic characterization of metallo-β-lactamase VIM-2. Biochemistry. 2014; 53: 7321–31. DOI: 10.1021/bi500916y [PubMed: 25356958]

[30]. Spencer J, Clarke AR, Walsh TR. Novel Mechanism of Hydrolysis of Therapeutic β-Lactams by Stenotrophomonas maltophilia L1 Metallo-β-lactamase. J Biol Chem. 2001; doi: 10.1074/jbc.M105550200

[31]. Kim Y, Tesar C, Mire J, Jedrzejczak R, Binkowski A, Babnigg G, Sacchettini J, Joachimiak A. Structure of apo- and monometalated forms of NDM-1--a highly potent carbapenem-hydrolyzing metallo-β-lactamase. PLoS One. 2011; 6 e24621 doi: 10.1371/journal.pone.0024621 [PubMed: 21931780]

[32]. Stoczko M, Frère J-M, Rossolini GM, Docquier J-D. Postgenomic scan of metallo-beta-lactamase homologues in rhizobacteria: identification and characterization of BJP-1, a subclass B3 ortholog from Bradyrhizobium japonicum. Antimicrob Agents Chemother. 2006; 50: 1973–81. DOI: 10.1128/AAC.01551-05 [PubMed: 16723554]

[33]. Tooke CL, Hinchliffe P, Bonomo RA, Schofield CJ, Mulholland AJ, Spencer J. Natural variants modify Klebsiella pneumoniae carbapenemase (KPC) acyl-enzyme conformational dynamics to extend antibiotic resistance. J Biol Chem. 2020; doi: 10.1074/JBC.RA120.016461

[34]. Tooke CL, Hinchliffe P, Lang PA, Mulholland AJ, Brem J, Schofield CJ, Spencer J. Molecular Basis of Class A β-Lactamase Inhibition by Relebactam. Antimicrob Agents Chemother. 2019; 63 doi: 10.1128/AAC.00564-19

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

[35]. Ullah JH, Walsh TR, Taylor IA, Emery DC, Verma CS, Gamblin SJ, Spencer J. The crystal structure of the L1 metallo-β-lactamase from Stenotrophomonas maltophilia at 1.7 å resolution 1. Nagai K. J Mol Biol. 1998; 284: 125–136. DOI: 10.1006/jmbi.1998.2148 [PubMed: 9811546]

[36]. Pemberton OA, Zhang X, Chen Y. Molecular Basis of Substrate Recognition and Product Release by the Klebsiella pneumoniae Carbapenemase (KPC-2). J Med Chem. 2017; 60: 3525–3530. DOI: 10.1021/acs.jmedchem.7b00158 [PubMed: 28388065]

[37]. Fonseca F, Chudyk EI, Van Der Kamp MW, Correia A, Mulholland AJ, Spencer J. The basis for carbapenem hydrolysis by class a β-lactamases: A combined investigation using crystallography and simulations. J Am Chem Soc. 2012; 134: 18275–18285. DOI: 10.1021/ja304460j [PubMed: 23030300]

[38]. Fisher, Jed F; Meroueh, Samy O; Mobashery, S. Bacterial Resistance to β-Lactam Antibiotics: Compelling Opportunism. Compelling Opportunity. 2005; doi: 10.1021/CR030102I

[39]. Brem J, Cain R, Cahill S, McDonough MA, Clifton IJ, Jiménez-Castellanos J-C, Avison MB, Spencer J, Fishwick CWG, Schofield CJ. Structural basis of metallo-β-lactamase, serine-β-lactamase and penicillin-binding protein inhibition by cyclic boronates. Nat Commun. 2016; 7 12406 doi: 10.1038/ncomms12406 [PubMed: 27499424]

[40]. Hinchliffe P, González MM, Mojica MF, González JM, Castillo V, Saiz C, Kosmopoulou M, Tooke CL, Llarrull LI, Mahler G, Bonomo RA, et al. Cross-class metallo-β-lactamase inhibition by bisthiazolidines reveals multiple binding modes. Proc Natl Acad Sci U S A. 2016; 113: E3745–E3754. DOI: 10.1073/pnas.1601368113 [PubMed: 27303030]

[41]. Hinchliffe P, Tanner CA, Krismanich AP, Labbé G, Goodfellow VJ, Marrone L, Desoky AY, Calvopiña K, Whittle EE, Zeng F, Avison MB, et al. Structural and Kinetic Studies of the Potent Inhibition of Metallo-β-lactamases by 6-Phosphonomethylpyridine-2-carboxylates. Biochemistry. 2018; 57: 1880–1892. DOI: 10.1021/acs.biochem.7b01299 [PubMed: 29485857]

[42]. Salimraj R, Hinchliffe P, Kosmopoulou M, Tyrrell JM, Brem J, van Berkel SS, Verma A, Owens RJ, McDonough MA, Walsh TR, Schofield CJ, et al. Crystal structures of VIM-1 complexes explain active site heterogeneity in VIM-class metallo-β-lactamases. FEBS J. 2019; 286: 169–183. DOI: 10.1111/febs.14695 [PubMed: 30430727]

[43]. Hinchliffe P, González MM, Mojica MF, González JM, Castillo V, Saiz C, Kosmopoulou M, Tooke CL, Llarrull LI, Mahler G, Bonomo RA, et al. Cross-class metallo-β-lactamase inhibition by bisthiazolidines reveals multiple binding modes. Proc Natl Acad Sci. 2016; 113: E3745–E3754. DOI: 10.1073/PNAS.1601368113 [PubMed: 27303030]

[44]. Spencer J, Read J, Sessions RB, Howell S, Blackburn GM, Gamblin SJ. Antibiotic recognition by binuclear metallo-beta-lactamases revealed by X-ray crystallography. J Am Chem Soc. 2005; 127: 14439–44. DOI: 10.1021/ja0536062 [PubMed: 16218639]

[45]. King DT, Worrall LJ, Gruninger R, Strynadka NCJ. New Delhi metallo-β-lactamase: structural insights into β-lactam recognition and inhibition. J Am Chem Soc. 2012; 134: 11362–5. DOI: 10.1021/ja303579d [PubMed: 22713171]

[46]. Akhtar A, Pemberton OA, Chen Y. Structural Basis for Substrate Specificity and Carbapenemase Activity of OXA-48 Class D β-Lactamase. ACS Infect Dis. 2020; 6: 261–271. DOI: 10.1021/acsinfecdis.9b00304 [PubMed: 31872762]

[47]. Lohans CT, Freeman EI, van Groesen E, Tooke CL, Hinchliffe P, Spencer J, Brem J, Schofield CJ. Mechanistic Insights into β-Lactamase-Catalysed Carbapenem Degradation Through Product Characterisation. Sci Rep. 2019; 9 13608 doi: 10.1038/s41598-019-49264-0 [PubMed: 31541180]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

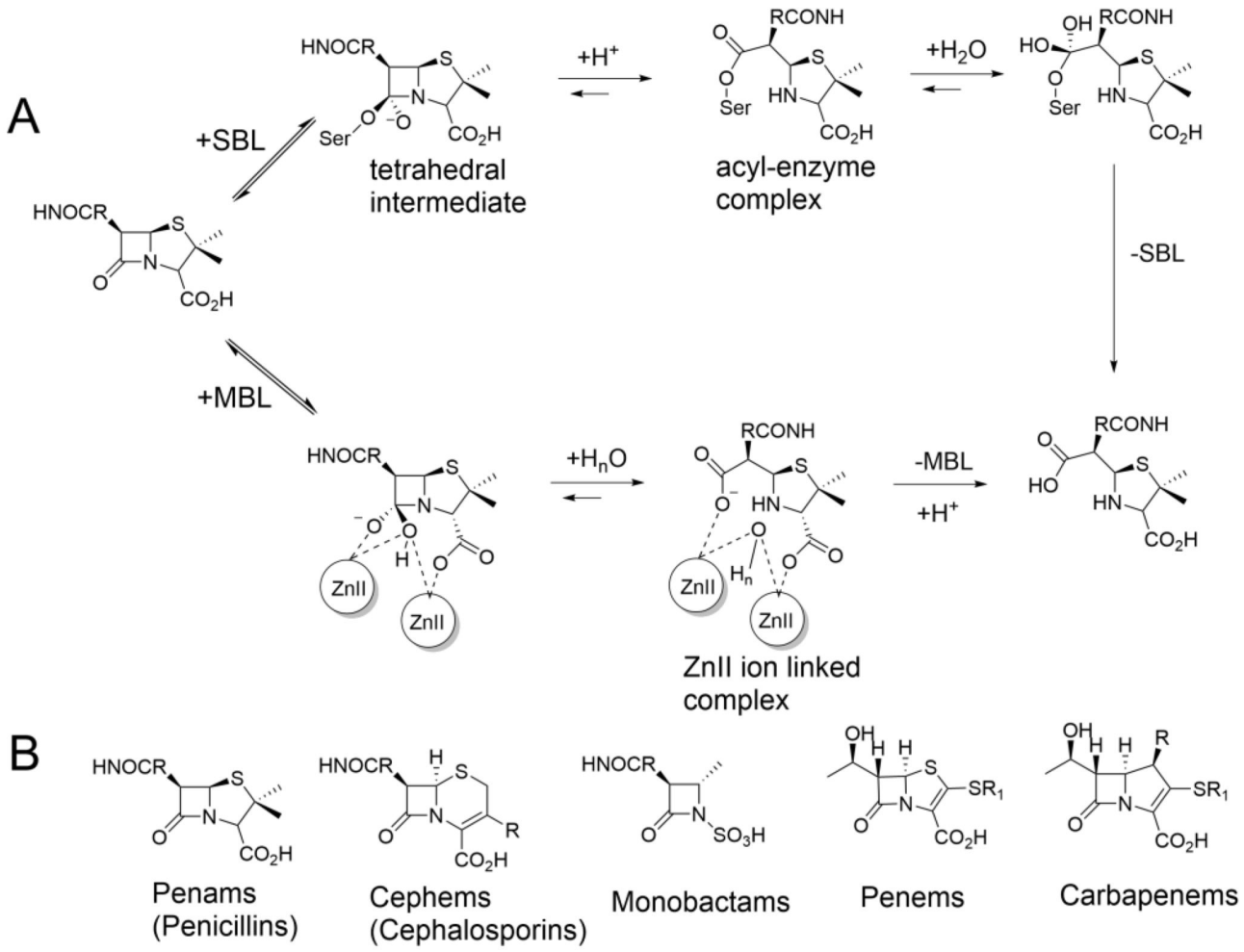


Figure 1. β-Lactam hydrolysis by serine and metallo-β-lactamase.

(A) SBL catalysis proceeds via an acyl-enzyme complex formed after being hydrolysed via a transient tetrahedral intermediate. In MBL catalysis, a zinc ion activated water molecule (that bridges the two Zn ions in the resting states of B1 and B3 MBLs), reacts with the β-lactam ring. Hydrolysis also proceeds via a transient tetrahedral intermediate and a hydrolysed intermediate stabilised through interactions with the Zn ions[6]. (B) Classes of clinically used β-lactam antibiotics.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

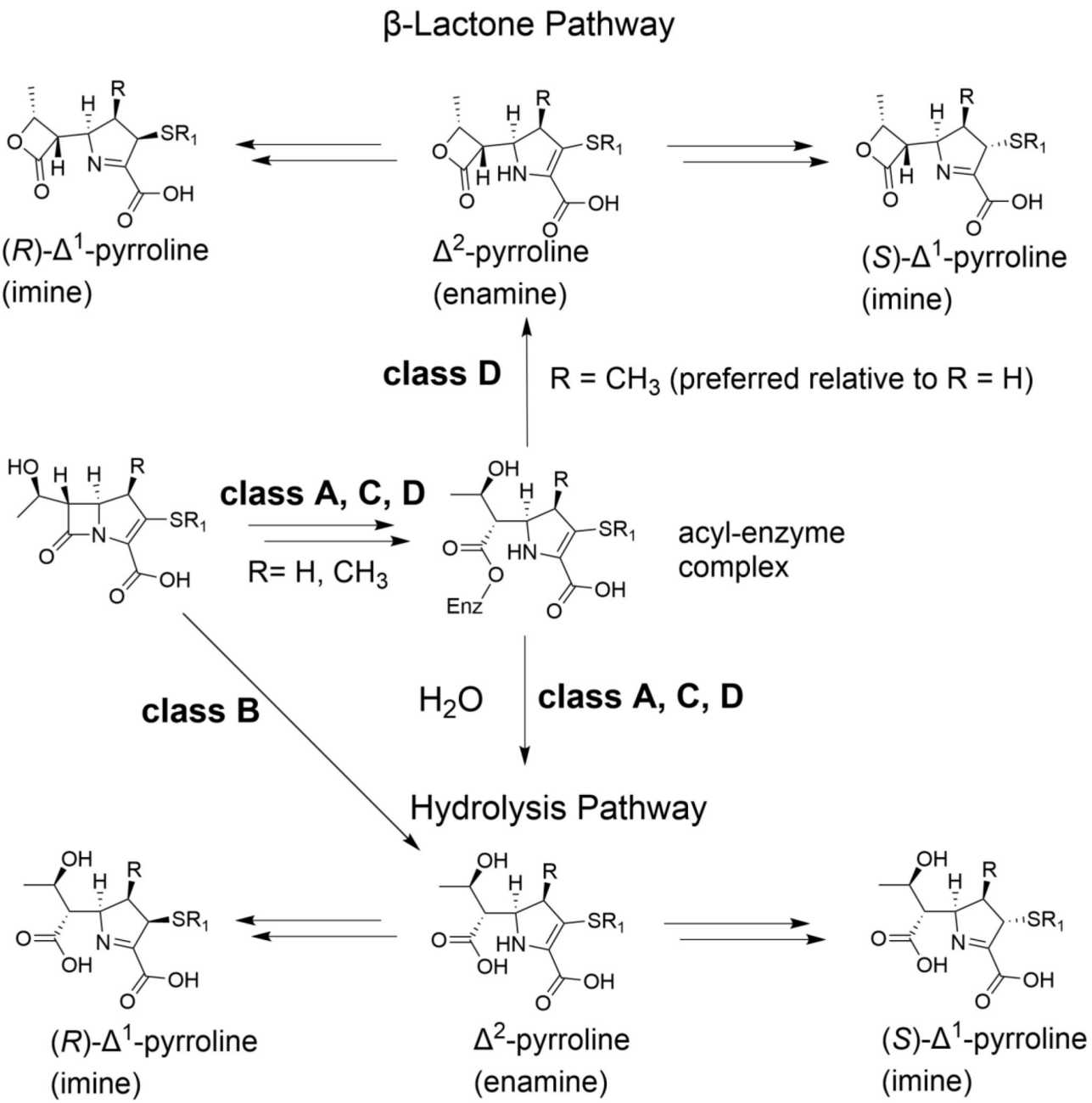


Figure 2. Carbapenem hydrolysis by β-lactamases.

SBL and MBL catalysed carbapenem hydrolysis is proposed to result (predominantly) in the initial formation of an enamine (Δ^2-pyrroline) that isomerises to give [(2*R*)- and (2*S*)-Δ^1-imine] products[10,13]. Various isomeric imine / enamine forms of reacted carbapenems have been observed crystallographically at transpeptidase/SBL/MBL active sites. In the case of the class D SBLs, the acyl-enzyme intermediate can also react to give lactone products.

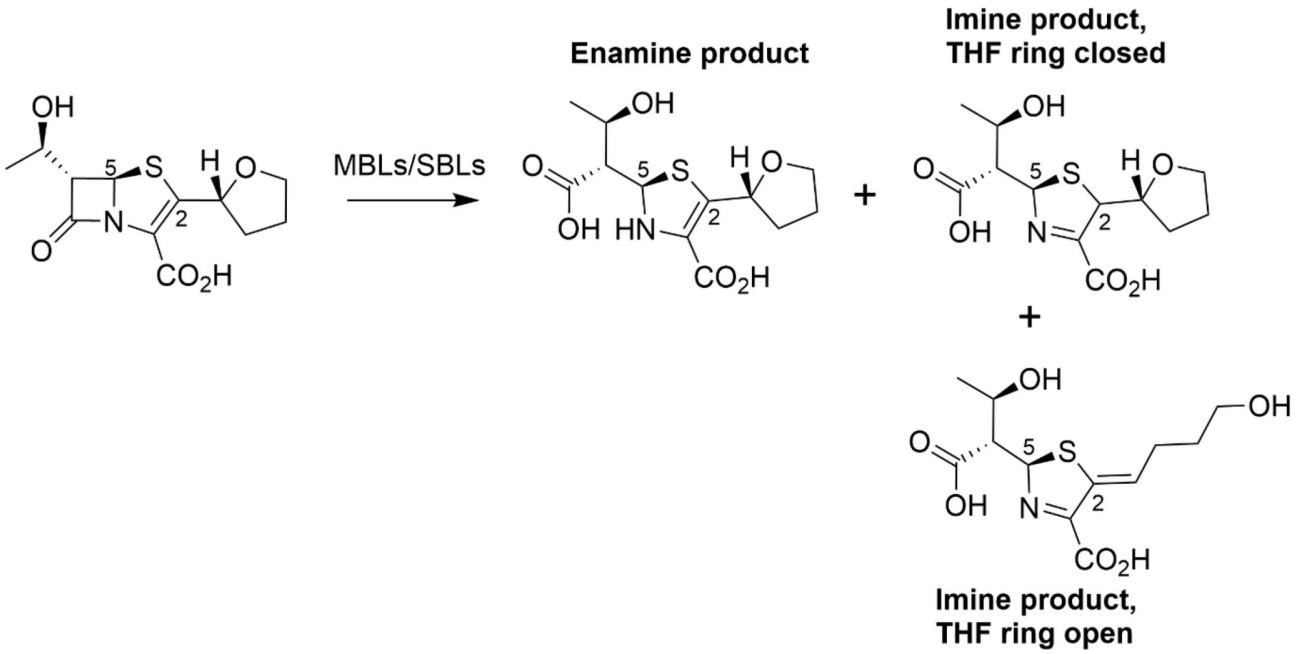


Figure 3. Potential reaction products of β-lactamase catalysed faropenem hydrolysis.
Note the possibility of isomer formation, e.g. epimerisation at C-5.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

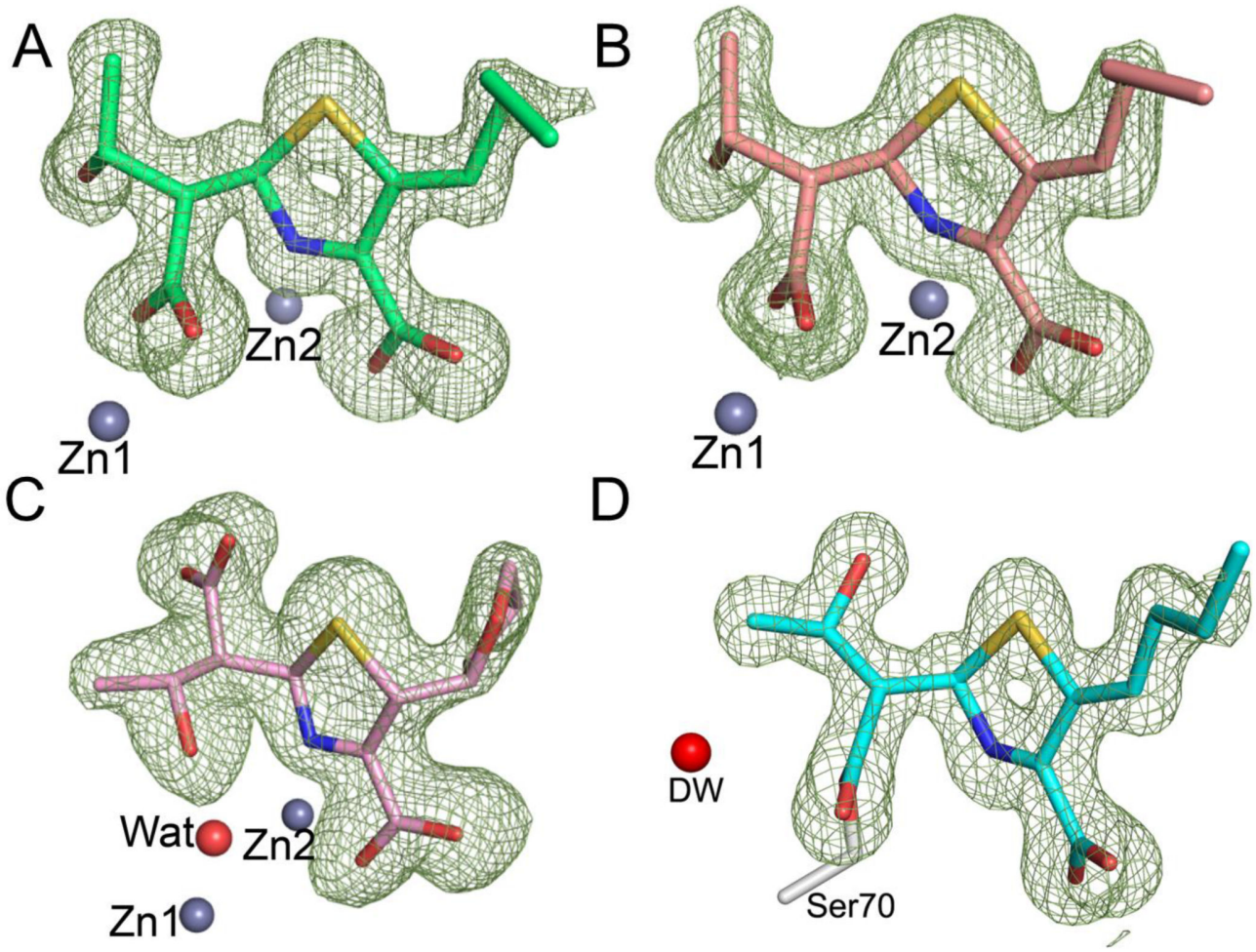


Figure 4. Faropenem derived complexes in the β-lactamase active sites as defined by electron density maps.

Views from the β-lactamase active sites, with the F_o-F_c electron density (green mesh, contoured at 3σ) calculated from the final model after removal of the ligand. (A) VIM-2, in the *I*222 space group (green, PDB code 7A5Z). (B) VIM-2 in the *C*2 space group (light red, PDB code 7A60). (C) L1 (pink, PDB code 7A63) and (D) KPC-2 (cyan, PDB code 7A61). The deacylating water (DW) is shown as a red sphere. Note the different conformations of the β-lactam derived carboxylate and hydroxyethyl groups in L1 compared to VIM-2 and the presence of the zinc bridging water with L1 (Wat, red sphere), but not VIM-2. No density was observed for the hydroxyl of the fragmented THF ring so it was omitted from the final models of the VIM-2 and KPC-2 complexes.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

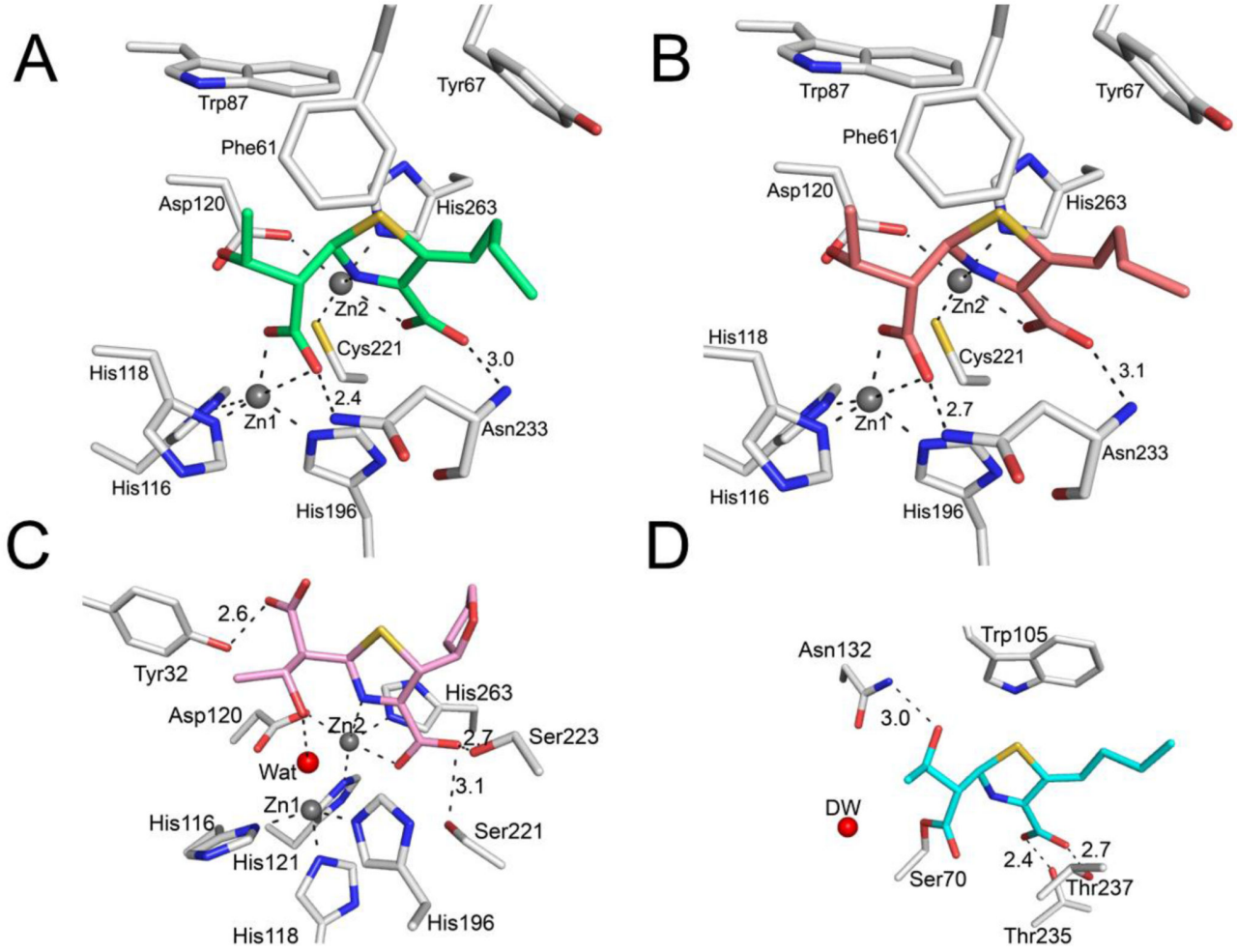


Figure 5. Interactions of the faropenem-derived complexes with serine- and metallo-β-lactamase active sites.

Protein residues are in light grey; faropenem-derived atoms are coloured as in Figure 4. Zinc ions are shown as grey spheres and the catalytic water (Wat) as a red sphere. Metal coordination and hydrogen bonds are shown as black dashes, with the distances labelled in Å. (A) VIM-2 (*I*222 space group), PDB 7A5Z; (B) VIM-2 (*C*2 space group) PDB 7A60; (C) L1 PDB 7A63; (D) KPC-2 PDB 7A61. Note the different conformations of the hydroxyethyl group in L1 (C) compared to VIM-2, leading to a lack of interaction with Zn1. KPC-2 (D) forms an acyl-enzyme complex with the faropenem derived complex.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

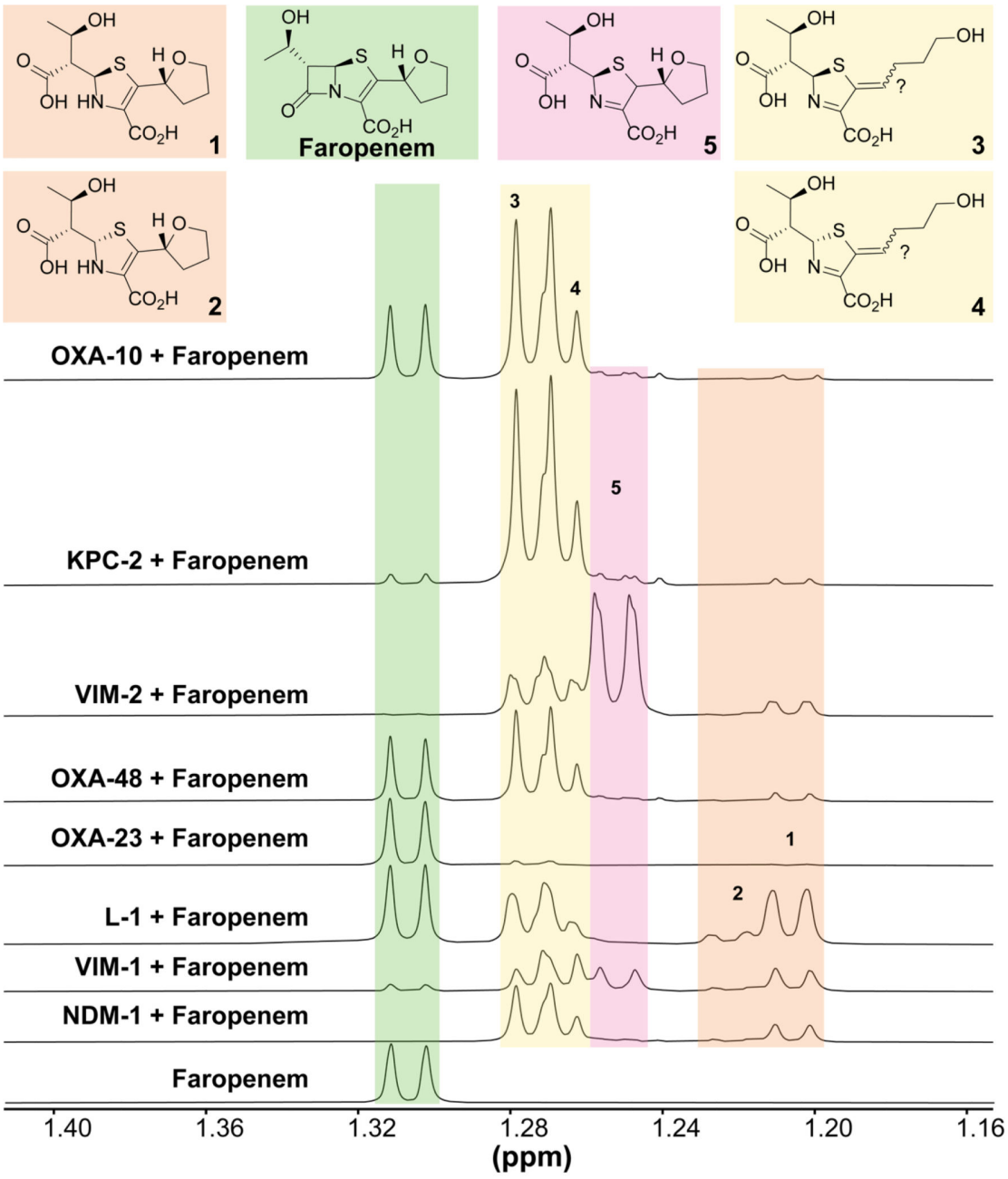


Figure 6. ^{1}H NMR (AVIII 700 MHz) spectra displaying methyl group resonances of the faropenem derived products.

Enamines **1** and **2** ((5*S*) and (5*R*) enamine THF ring-closed products, orange), THF ring-opened (5*R*) and (5*S*) imines **3** and **4** (yellow) and THF ring closed imine **5** (pink) degradation products were formed after incubation of faropenem (green) with OXA-10 (5 μM, 16 h), OXA-23 (5 μM, 15 min), OXA-48 (5 μM, 55 min), VIM-1 (280 nM, 5 min), VIM-2 (280 nM, 5 min), NDM-1 (5 μM, 5 min), KPC-2 (125 nM, 72 min), L1 (125 nM, 5 min). All enzymes were incubated with 5 mM faropenem apart from OXA-23, which was

incubated with 2 mM faropenem. Reactions were in 50 mM sodium phosphate, pH 7.5, 10% (*v/v*) D_2O. Note, **5** is only observed with VIM-1 and VIM-2. Colour coding corresponds to the structures shown and associated resonances. Individual spectra are provided in Figures S13-19. Characterisation data for compounds **2** and **5** are available in Figure 21 and Figures S22-27 respectively.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

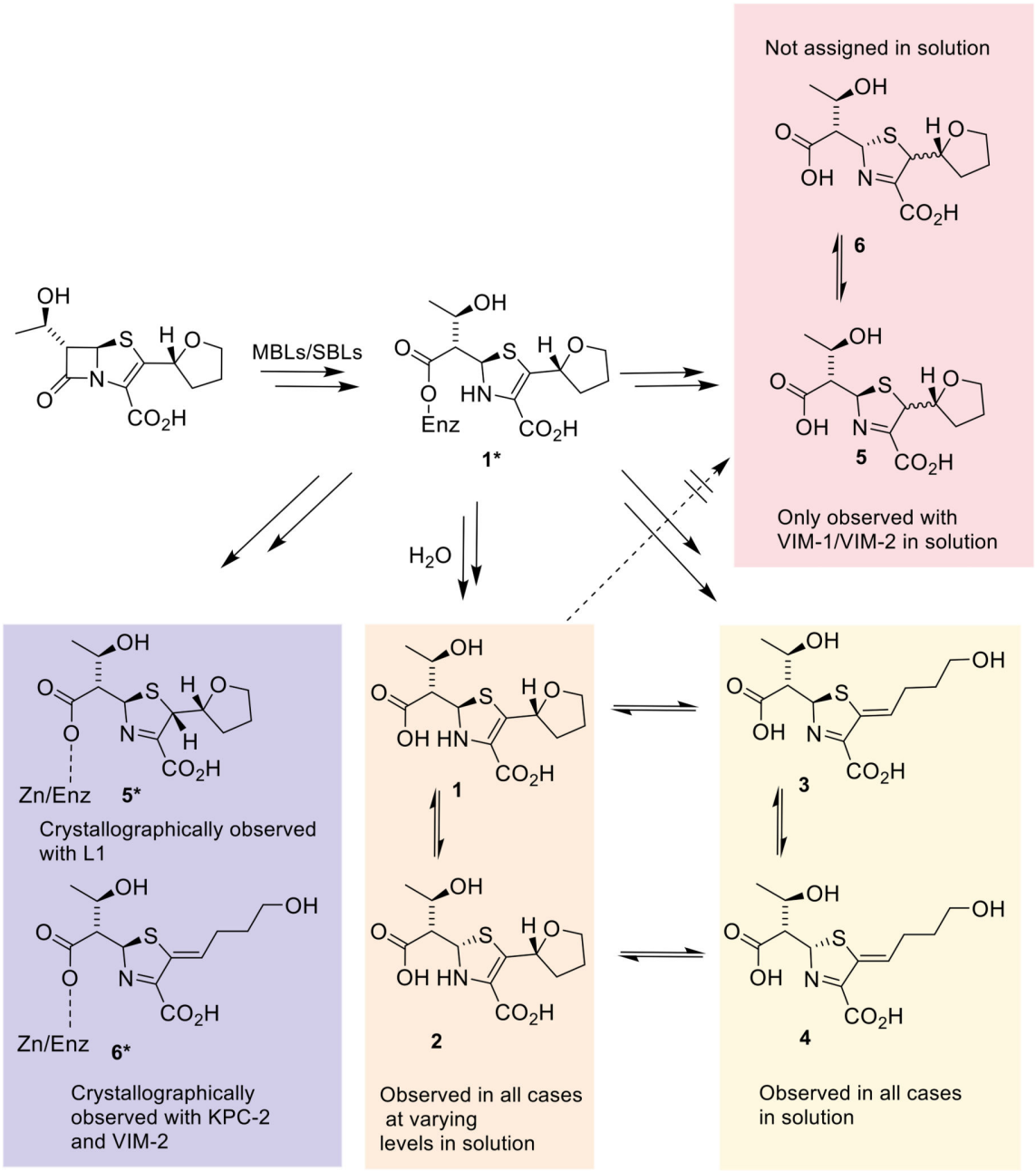


Figure 7. Summary of solution and crystallographic observations for faropenem reactions with SBLs/MBLs.

*Denotes crystallographically observed enzyme bound complexes. Note that **5** and **6** were only observed in the cases of solution studies with VIM-1 and VIM-2.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Steady state kinetic parameters for faropenem turnover including reported values.

Enzyme	Type	Class	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (s^{-1} $μM^{-1}$)	K_i (μM)
KPC-2	SBL	A	3.71 (0.21)	16.6 (3.84)	0.22	44.1 (25.0)
VIM-2	MBL	B1	36.7 (1.03)	5.11 (0.63)	7.17	-
L1	MBL	B3	43.7 (2.70)	36.3 (8.80)	1.20	-
BlaC[24]	SBL	A	0.91 (0.03)	23 (3)	0.040	-
NDM-1[31]	MBL	B1	14.7 (0.08)	99 (0.16)	0.15	-
BJP-1[32]	MBL	B3	2 (<0.2)	245 (<25)	0.008	-

Standard errors of the mean in parenthesis (n=3). See Experimental Section for KPC-2, VIM-2 and L1 assay details.

antibiotics

MDPI

Article

Studies on the Reactions of Biapenem with VIM Metallo β-Lactamases and the Serine β-Lactamase KPC-2

Anka Lucic [1], Tika R. Malla [1], Karina Calvopiña [1], Catherine L. Tooke [2], Jürgen Brem [1], Michael A. McDonough [1], James Spencer [2] and Christopher J. Schofield [1,*]

[1] Chemistry Research Laboratory, The Department of Chemistry and the Ineos Oxford Institute for Antimicrobial Research, University of Oxford, Oxford OX1 3TA, UK; anka.lucic@strubi.ox.ac.uk (A.L.); tika.malla@chem.ox.ac.uk (T.R.M.); karina.calvopinatapia@chem.ox.ac.uk (K.C.); jurgen.brem@chem.ox.ac.uk (J.B.); michael.mcdonough@chem.ox.ac.uk (M.A.M.)

[2] Biomedical Sciences Building, School of Cellular and Molecular Medicine, Faculty of Life Sciences, University of Bristol, University Walk, Bristol BS8 1TD, UK; ct12425@bristol.ac.uk (C.L.T.); jim.spencer@bristol.ac.uk (J.S.)

* Correspondence: christopher.schofield@chem.ox.ac.uk

Abstract: Carbapenems are important antibacterials and are both substrates and inhibitors of some β-lactamases. We report studies on the reaction of the unusual carbapenem biapenem, with the subclass B1 metallo-β-lactamases VIM-1 and VIM-2 and the class A serine-β-lactamase KPC-2. X-ray diffraction studies with VIM-2 crystals treated with biapenem reveal the opening of the β-lactam ring to form a mixture of the (2*S*)-imine and enamine complexed at the active site. NMR studies on the reactions of biapenem with VIM-1, VIM-2, and KPC-2 reveal the formation of hydrolysed enamine and (2*R*)- and (2*S*)-imine products. The combined results support the proposal that SBL/MBL-mediated carbapenem hydrolysis results in a mixture of tautomerizing enamine and (2*R*)- and (2*S*)-imine products, with the thermodynamically favoured (2*S*)-imine being the major observed species over a relatively long-time scale. The results suggest that prolonging the lifetimes of β-lactamase carbapenem complexes by optimising tautomerisation of the nascently formed enamine to the (2*R*)-imine and likely more stable (2*S*)-imine tautomer is of interest in developing improved carbapenems.

Keywords: carbapenems; antimicrobial resistance; metallo-β-lactamases; serine-β-lactamases; biapenem

check for updates

Citation: Lucic, A.; Malla, T.R.; Calvopiña, K.; Tooke, C.L.; Brem, J.; McDonough, M.A.; Spencer, J.; Schofield, C.J. Studies on the Reactions of Biapenem with VIM Metallo β-Lactamases and the Serine β-Lactamase KPC-2. *Antibiotics* **2022**, *11*, 396. https://doi.org/10.3390/antibiotics11030396

Academic Editor: Krisztina M. Papp-Wallace

Received: 25 January 2022
Accepted: 9 March 2022
Published: 16 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

β-Lactam-containing drugs are the most clinically important antibacterial class [1], though their use is increasingly impaired by resistance due to nucleophilic serine-β-lactamases (SBLs, Ambler classes A, C and D) and zinc-dependent metallo-β-lactamases (MBLs, Ambler class B) (Figure 1) [2–4]. Carbapenems were once considered antibiotics of 'last resort', in part because they are resistant to hydrolysis by some SBLs [1]. However, the increasingly widespread use of carbapenems is correlated with an increase in extended-spectrum SBLs (ESBLs) in both Gram-positive and Gram-negative bacteria [5]. The role of MBLs, such as the Verona integron MBLs, which often efficiently hydrolyse carbapenems, in antimicrobial resistance is also rapidly growing and is now endemic in some regions [6].

The stability of the complexes formed by β-lactams with SBLs/MBLs is an important factor in determining their susceptibility to β-lactamases and hence efficacy. The ability of carbapenems to form relatively stable complexes both with their transpeptidase targets and certain SBLs has contributed to their greater efficacy compared to other β-lactams against some resistant strains. However, there is incomplete knowledge of the molecular factors involved in determining enzyme-carbapenem complex stabilities, as exemplified by the recent observation that carbapenems react with Class D SBLs to produce β-lactones in addition to the established enamine and imine hydrolysed products (Figure 1) [7,8].

Studies on the products of carbapenem hydrolysis by several MBLs and SBLs have led to the proposal that, at least in some cases, during efficient catalysis the major nascent product is in the enamine tautomeric form, which undergoes isomerisation to give the (2*R*)- and (2*S*)-C2 imines (Figure 1) [8–11].

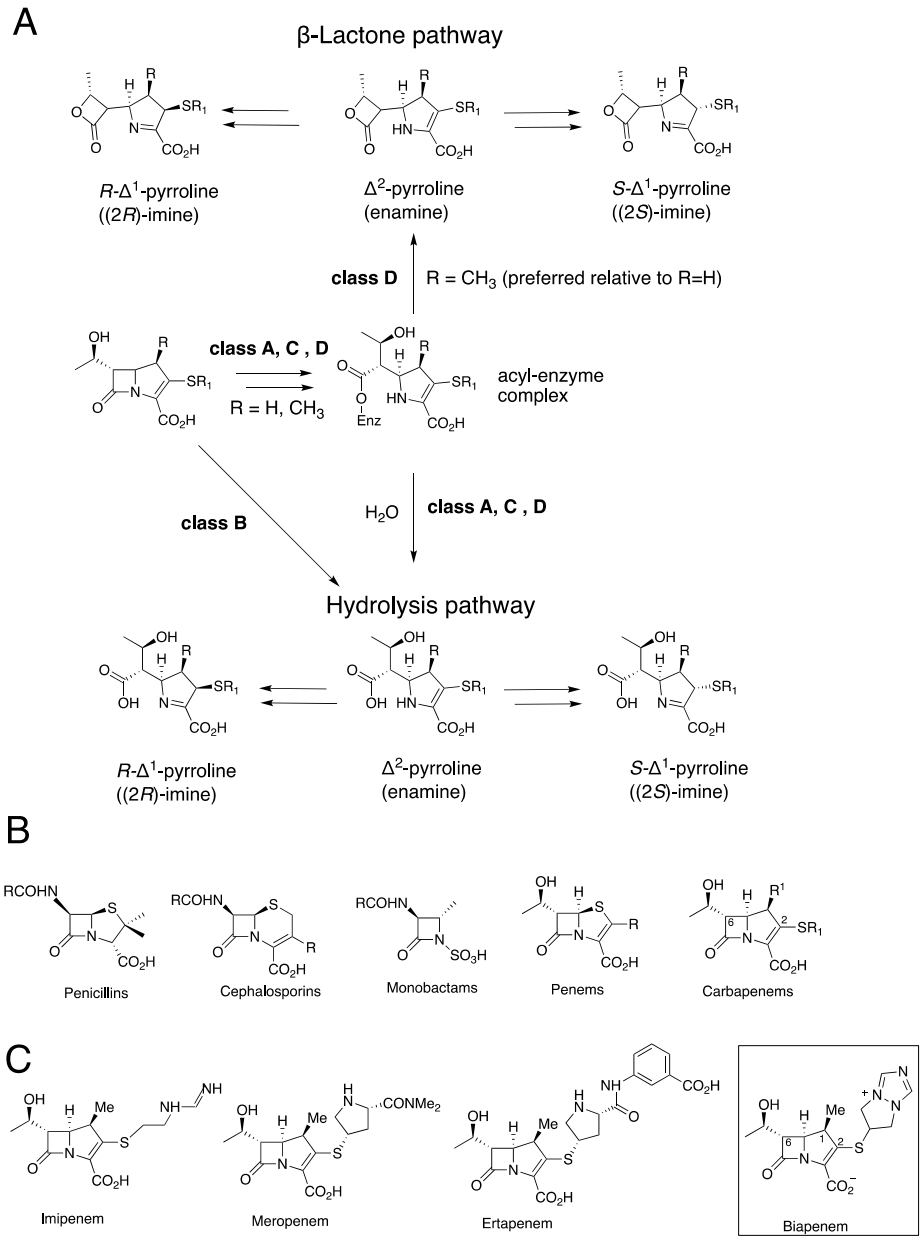


Figure 1. Carbapenem hydrolysis by β-lactamases. (**A**) It is proposed that SBLs and MBLs hydrolyse carbapenems to produce an enamine (Δ^2-pyrroline), which during efficient catalysis in solution isomerises to give (2*R*)- and (2*S*)-Δ^1-imine products. Both enamine and (2*S*)- and (2*R*)-imine carbapenem-derived ligands have been observed by crystallography at MBL/SBL/transpeptidase active sites. Lactones can also be formed, at least, in the case of class D SBLs. (**B**) Classes of clinically used β-lactam antibiotics (for carbapenems R^1 = H or Me). (**C**) Examples of clinically used carbapenems; note the unusual positively charged bicyclic bicyclotriazolium C2 side chain of biapenem.

The intravenously administered carbapenem biapenem has a broad spectrum of activity and is used to treat serious infections [12–14]. Biapenem is reported to have less adverse side effects than other carbapenems and does not need to be administered with a dihydropeptidase 1 inhibitor [13]. Biapenem is an inhibitor of some SBLs but is subject to MBL hydrolysis [12–14]. Since biapenem has the same C6 hydroxyethyl side chain as other carbapenems, its distinctive properties must be due to the presence of its unusual positively charged 'bicyclotriazolium' sulfur-linked C2 bicyclic ring system (Figure 1) [12–14]. The unusual C2 side chain of biapenem compared to the more typical C2 side chains of carbapenems such as meropenem makes it of particular interest with respect to defining its tautomeric enamine and imine product profile.

Here, we report studies on the reactions of biapenem with the clinically important VIM-1, VIM-2, and KPC-2 β-lactamases using NMR spectroscopy and, in the case of VIM-2, X-ray crystallographic analysis. After the pioneering identification of the first Verona integron MBL (VIM) in 1996, multiple VIM variants have been observed in Gram-negative bacteria, including in *Klebsiella* and *Eneterobacteriaceae* [15,16]. The VIM MBLs are broad-spectrum β-lactamases that catalyse the hydrolysis of penicillins, cephalosporins, and carbapenems, but not monobactams [17–20] (Figure 1C). VIMs have the typical αβ/βα MBL fold and are di-zinc utilising B1 subfamily MBLs [17,19,20]. The *Klebsiella pneumoniae* carbapenemase 2 (KPC-2) is a Class A SBL of increasing clinical importance and has broad-spectrum activity, including with respect to the particularly efficient hydrolysis of carbapenems [21,22].

2. Results

We initially investigated how biapenem reacts with VIM-1, VIM-2, and KPC-2 using ^{1}H NMR (600 MHz) spectroscopy. In all three cases, enamine and both (2*R*)-imine and (2*S*)-imine products were observed, as evidenced by the analysis of the methyl group resonances (C6 hydroxyethyl methyl and C1 methyl groups) in the high-field region (0.8–1.3 ppm) of the ^{1}H NMR spectrum. The major product observed was the (2*S*)-imine, consistent with previous work indicating that this is the thermodynamically favoured species of the three observed major tautomeric products. [7–9,20]. Lower levels of the (2*R*)-imine and enamine products were observed, along with low levels of at least one unassigned product with methyl group resonances at ~1.00 and 1.02 ppm (Figures 2 and S1–S4, Table S1). No evidence for the formation of lactone products was accrued, contrasting with the results of reaction of carbapenems with class D SBLs, but consistent with prior work on the reactions of carbapenems with MBLs (Figure 1), [7–9,20]. Overall, our NMR results support the proposal of the (2*S*)-imine as the major product (on the timescale of the NMR assays), though they do not preclude the possibility that the enamine or (2*R*)-imine are nascent products.

Crystal structures have been reported for biapenem-derived complexes with the class B2 MBL CphA and certain transpeptidases [23–25]. However, there is no reported structure for biapenem with a representative of the most clinically relevant class B1 MBLs. Thus, we performed studies to investigate how biapenem interacts with the B1 MBL Verona integron metallo β-lactamase (VIM-2), on which studies with inhibitors have previously been carried out [26,27]. Crystals of recombinant VIM-2 were produced as previously described [20,26] and were soaked with a 50 mM solution of biapenem for 1, 5, 10, and 30 min. A diffraction dataset with a crystal soaked in biapenem for 10 min was obtained with the Diamond Light Source at a 1.3 Å resolution (space group *I*222 with one molecule in the asymmetric unit) (Figure S5, Table S2). Due to low completeness in the high-resolution shell, the resolution for data processing was cut to 1.5 Å.

Analysis of the electron density at the VIM-2 active site led to trial refinements with either the (2*S*)-imine, the (2*R*)-imine or the enamine (independently), the two imines together at 50% occupancy, or the enamine with either the (2*S*)- or (2*R*)-imine at 50% occupancy (Figures 3 and 4). The best fit was obtained with the (2*S*)-imine and the enamine at 50% occupancy. Note that the occupancy level should be regarded as an estimation and we cannot rule out low levels of the presence of the (2*R*)-imine and other ligands; in the

NMR studies, there were some unassigned peaks, consistent with minor product formation (Figures S1 and S2).

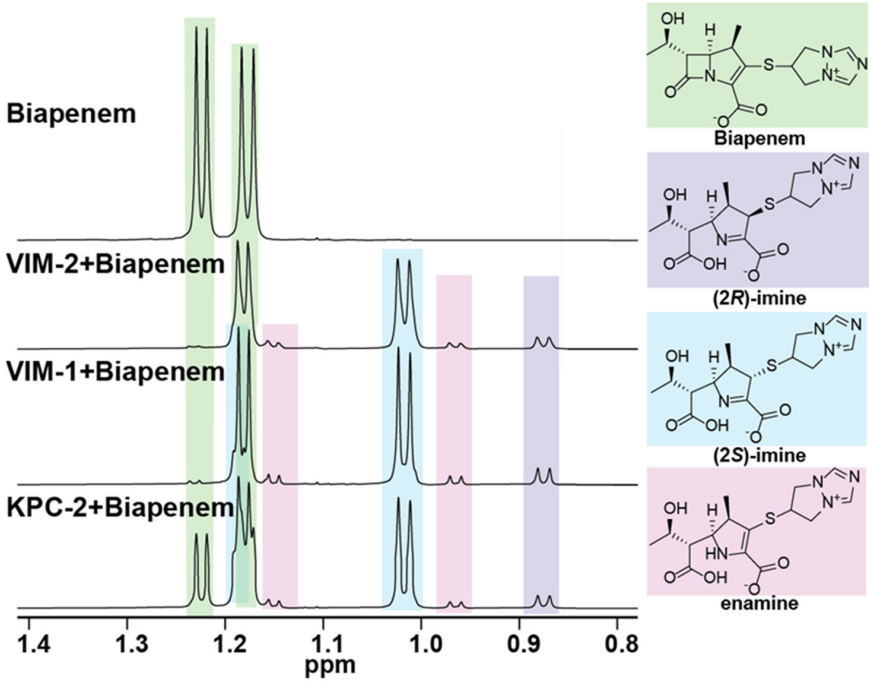


Figure 2. High-field regions of ^{1}H NMR spectra (600 MHz) for reaction of biapenem with the MBLs VIM-1 or VIM-2, or the SBL KPC-2. Biapenem (5 mM, green) was treated with the purified β-lactamase (280 nM, 30 min) in 50 mM of sodium phosphate at pH 7.6 (with 10% *v*/*v* D_2O). Formation of enamine (pink) and (2*R*)-imine and (2*S*)-imine (purple and blue, respectively) products was observed in all three cases. Time-course analyses imply the (2*S*)-imine is the major and 'thermodynamic' product (Figures S1–S4). The slowest rate of biapenem turnover was observed with KPC-2 (Figure S1), with increasing rates observed for VIM-1 (Figure S2), and VIM-2 (Figure S3).

The apparent electron density for the bicyclic side chains of both the (2*S*)-imine and the enamine ligands was weak, likely due to the lack of protein contacts causing conformational mobility (Figure 3). In general, the interactions made by both the enamine and (2*S*)-imine are analogous to those observed in other VIM-1: carbapenem-derived structures [19], i.e., in both cases the newly formed β-lactam derived carboxylates displace the water/hydroxide ion bridging the two zinc ions, and the C3 carboxylate is positioned to interact with the Arg228 sidechain (Figures S6 and S7). An increase of 0.8 Å in the distance between the two active site zinc ions (4.3 Å) compared to the unliganded structure was also observed, as previously reported in MBL carbapenem and faropenem derived complexes [19,20].

Notably, the thioether of the C2 biapenem-derived sidechain adopts different conformations for the enamine and the (2*S*)-imine ligands, with the sulfur being positioned close to the primary amide of Asn233 in the case of the enamine, but not the (2*S*)-imine (Figures 4 and S6). The C6 hydroxyethyl group is positioned similarly for both the enamine and (2*S*)-imine ligands, in a manner similar to that observed in previous VIM-2 carbapenem/faropenem-derived complexes (Figure S7) [19,20].

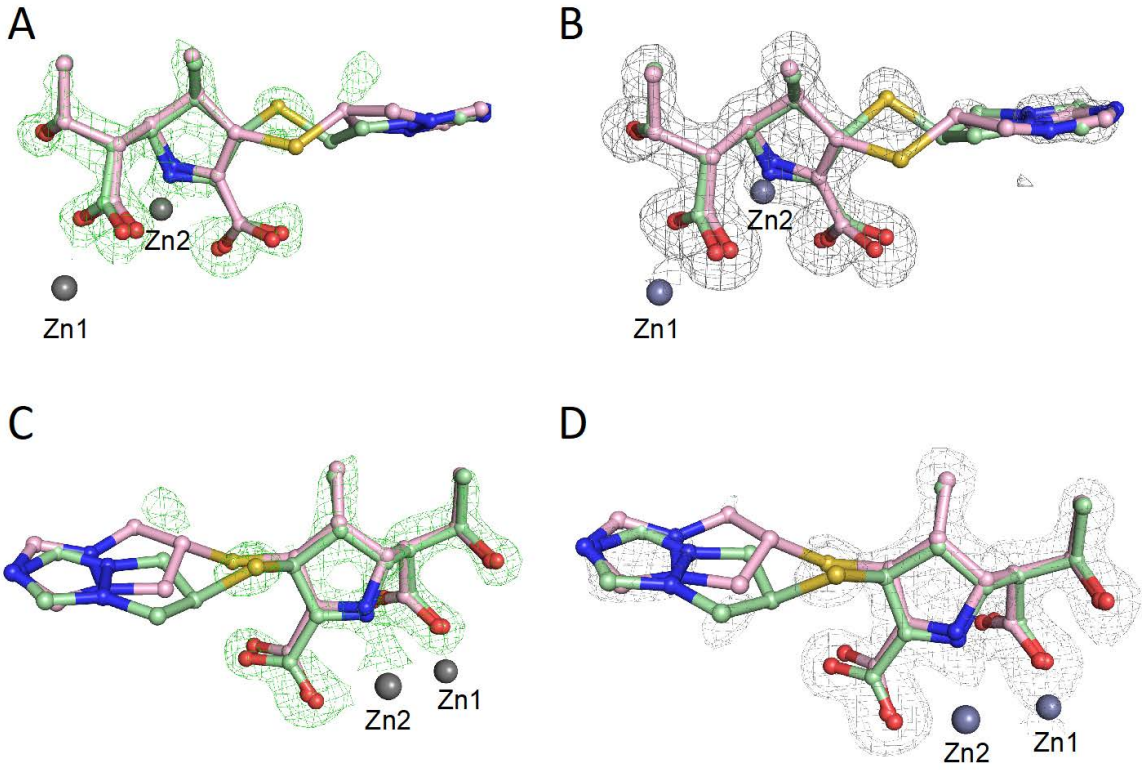


Figure 3. Electron density maps for biapenem-derived products in complex with VIM-2. (2*S*)-Imine: pale green balls and sticks; enamine: pink balls and sticks. (**A**,**C**) show an OMIT map (green mesh) of mFo-DFc contoured to 3σ (PDB: 6Y6J). (**B**,**D**) show the 2mFo-DFc difference map in grey mesh, contoured to 1σ. Note the weak density for the biapenem-derived bicyclotriazolium sidechain, indicating the presence of multiple conformations.

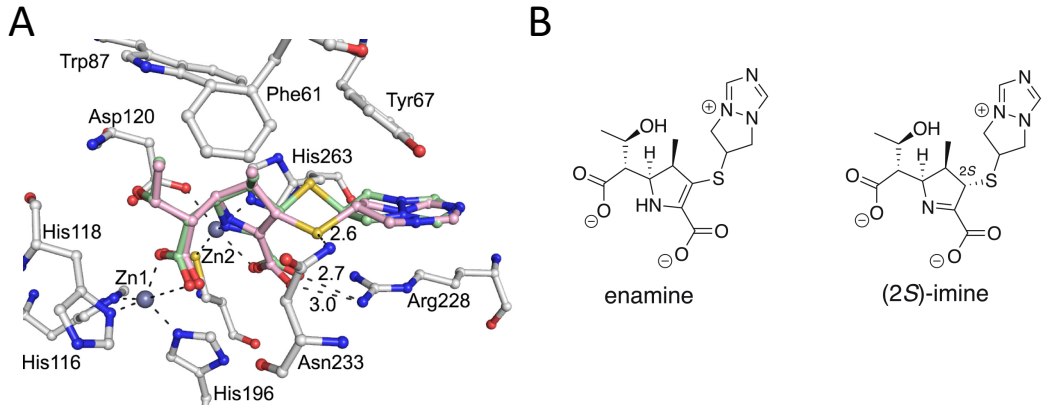


Figure 4. Views of a crystal structure of VIM-2 complexed with biapenem-derived ligands (PDB 6Y6J). (**A**) Interactions occurring between the (2*S*)-imine (green) and enamine (pink) products at the VIM-2 active site. Note that both products interact via their C3 carboxylate with Arg228 and that the sulfur of the thioether sidechain of the enamine is positioned adjacent to the primary amide of Asn233. (**B**) The (2*S*)-imine and enamine tautomers were each modelled at 50% occupancy.

When compared with a VIM-1: meropenem-derived (2*S*)-imine structure (PDB: 5N5I) [19], the binding modes in the VIM-2: biapenem-derived complex structure manifest some differences (Figure S7). The C6 hydroxyethyl group and newly formed carboxylate in the VIM-1: meropenem (2*S*)-imine product complex structure [19] have different orientations compared to VIM-2: biapenem products complex for both the enamine and the (2*S*)-imine. Notably, the binding modes of the VIM-2: biapenem-derived ligand complex and the VIM-2: faropenem Z-alkene imine-derived ligand complex appear more similar, with the analogous positioning of the C2 carboxylate (Figures S7 and S8) [20]. There are also differences in the binding modes of biapenem derived ligands when reacted with penicillin binding proteins and the class B2 MBL CphA (Figure S8). It is also notable that the different tautomeric complexes formed by carbapenem-derived ligands with MBLs are mimicked to different extents by various heterocyclic MBL inhibitors binding at the di-Zn(II) containing active site, as observed by crystallography [28,29].

3. Discussion

Previous studies have shown that the SBL/MBL catalysis of carbapenem hydrolysis results in the formation of enamine and (2*R*) and or (2*S*)-imines as the major observed hydrolysed products observed in the NMR analyses, with the analogous β-lactones being additionally observed in the case of class D SBLs [7–9,11]. Consistent with these findings, in the results reported here with biapenem, the (2*S*)-imine was observed as the major product with the B1 metallo-β-lactamases VIM-1, VIM-2 and the class A serine-β-lactamase KPC-2, at least on the relatively long timescale of the NMR assays. This observation indicates that the unusual biapenem bicyclotriazolium side chain does not substantially perturb the enamine/imine equilibrium position in the product enamine and imine tautomers (Figures 2 and S1–S4).

We appreciate that care should be taken in assigning crystallographically observed complexes, at least in terms of the details of binding, as being relevant to catalytically relevant intermediate structures in solution. However, consistent with the solution studies employing NMR, the crystallographic analysis of VIM-2 crystals soaked with biapenem revealed evidence for the enamine and the (2*S*)-imine as the major products bound, with no evidence for the (2*R*)-imine. We cannot rule out that the (2*R*)-imine, along with other ligands, may be formed at the active site at low levels, either in crystallo or in solution. In solution, at least for some carbapenems, the (2*R*)-imine is the kinetic product of enamine tautomerisation [7,9,10]. Thus, it is possible that we did not observe the (2*R*)-imine in part because of the relatively long timescale of the crystallographic analyses, which involved a 10 min crystal-soaking procedure. However, our combined results are also consistent with the proposal that during efficient catalysis, the enamine might be the major nascent product, with subsequent optimisation to the (2*R*)-imine and then (2*S*)-imines occurring in solution [7,9]. The formation of the (2*S*)-imine and, maybe, (2*R*)-imine can also occur at the active site in crystallo, at least as observed for the (2*S*)-imine in our VIM-2: biapenem work and with previous VIM-1: meropenem work [19].

Detailed spectroscopic analyses are required, both in solution and in crystals, to understand the precise nature of the reactions, including tautomerisations, that carbapenems undergo when bound to SBL and MBL active sites. It should also be noted that the observed product ratios are dependent on the efficiency of β-lactamase catalysis relative to that of tautomerisation, either in solution or in crystallo. Nonetheless, the combined results suggest that work aiming to prolong the lifetimes of the β-lactamase-bound carbapenem products is of interest with respect to developing improved carbapenems. This might be achieved either by the modification of the C2 carbapenem side chain or other group, in order to optimise the tautomerisation of the bound enamine tautomer to the (2*R*)-imine and potentially more stable (2*S*)-imine tautomer. The combined mechanistic studies with current clinically used carbapenems, including biapenem, however, suggest that the synthesis of new types of carbapenems may be required to achieve this.

4. Materials and Methods

4.1. Protein Purification and X-ray Crystallography

Recombinant forms of VIM-1, VIM-2, and KPC-2 were purified to near homogeneity (by SDS-PAGE analysis), as described [20,30,31]. VIM-2 (8.5 mg/mL) was crystallised using the vapour diffusion method under previously described conditions (Art Robbins low profile 96 well Intelli-plates) [20,30,31]. Crystallisation plates were set up using a PhoenixRE (ArtRobbins, Sunnyvale, CA, USA) robot with droplets of volume 200–300 nL with a 1:1, 1:2, and 2:1 protein: reservoir ratio equilibrated against 80 µL reservoir solution. VIM-2 crystals appeared after 24–48 h and were soaked with ~200–300 nL of 50 mM biapenem aqueous stock solution for 10 min. The crystal was harvested, then cryo-cooled using liquid nitrogen and sent for data collection at the Diamond Light Source.

4.2. Data Collection and Processing

Diffraction data were collected on a single crystal at 100K at Diamond Light Source beamline I04-1. Images were indexed and integrated using FastDP [32]; the structure was solved by molecular replacement with Phaser using a previously reported VIM-2 structure as a model (PDB file 4BZ3) [30,31,33,34]. A dataset was collected on a crystal soaked with biapenem for 10 min. After the first round of refinement, the mF_o–DF_c map was consistent with the presence of (a) biapenem-derived ligand(s). The ligand was manually fitted using WinCOOT before the second round of refinement. Subsequent refinements were carried out using the program Phenix, with manual model building in WinCOOT [35–37]. After further refinement, the electron density maps implied the presence of two different imposed products. Trial refinements with the (2*S*)-imine, the (2*R*)-imine, or the enamine (independently); the two imines together at 50% occupancy; and the enamine with either the (2*S*)- or (2*R*)-imine at 50% occupancy were carried out. The best fit was obtained with the (2*S*)-imine and the enamine with 50% occupancy. The geometry restraints for the biapenem (2*S*)-imine and enamine-derived products were calculated using electronic ligand builder and optimisation workbench (eLBOW), and the electron density maps were calculated using PHENIX [35].

4.3. ^{1}H NMR Studies

NMR-monitored assays were performed with VIM-1 (0.28 µM, final concentration) VIM-2 (0.28 µM), and KPC-2 (0.28 µM) mixed with 5 mM of biapenem in a 50 mM sodium phosphate solution at pH 7.6 with 10% (*v*/*v*) D_2O, using trimethylsilylpropanoic acid as an internal standard (2 µL of a 3 mg/mL solution). Experiments were conducted using a Bruker AVIIIHD 600 MHz spectrometer equipped with a Prodigy broadband cryoprobe. Water signal suppression was conducted using excitation sculpting with perfect echo. Chemical shift assignments were based on reported values; data were processed using MestReNova [38]. Time-courses are presented with earliest time point at the bottom and last time point (40 min) at the top (Figure 2 and Figures S1–S3). The spectrum of biapenem (5 mM) is displayed at the bottom. Each spectrum was acquired at 2.5 min intervals (16 scans). Note that the presence of the enamine product was assigned based on previous studies.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/antibiotics11030396/s1. Table S1. Chemical shift assignments for the major biapenem-derived hydrolysis product. The chemical shifts and the C2 coupling constants (J 3.6, 1.3 Hz) of this hydrolysis product suggest that it is likely the (2*S*)-imine based on previous studies [8,9]; Table S2. Data collection and refinement statistics. High resolution statistics (1.36–1.33) are in parentheses. Due to poor completeness, the resolution for data processing was cut of to 1.5 Å, with a high resolution shell of 1.556–1.5 Å; Figure S1. ^{1}H NMR time course analysis of methyl group resonances of species formed on reaction of biapenem with KPC-2. Biapenem (5 mM, green) was treated with KPC-2 (280 nM) in 50 mM sodium phosphate at pH 7.5, 10% (*v*/*v*) D_2O. Evidence was accrued for formation of (2*R*)-imine, (2*S*)-imine and enamine products (blue, purple,

and pink, respectively). 15 spectra were acquired (16 scans, every 2.5 min) over 37.5 min (stacked from bottom/first spectrum to top). The bottom spectrum is of biapenem (5 mM) without added enzyme; Figure S2. ^{1}H NMR time course analysis of methyl group resonances of species formed on reaction of biapenem with VIM-1. Biapenem (5 mM, green) was treated with VIM-1 (280 nM) in 50 mM sodium phosphate at pH 7.5, 10% (*v*/*v*) D_2O. Unassigned product peaks are labelled with asterisks. Evidence was accrued for formation of (2*R*)-imine, (2*S*)-imine and enamine products (blue, purple, and pink, respectively). 15 spectra were acquired (16 scans, every 2.5 min) over 37.5 min (stacked from bottom/first spectrum to top). The bottom spectrum is biapenem (5 mM) without added enzyme; Figure S3. ^{1}H NMR time course analysis of methyl group resonances of products formed on reaction of biapenem with VIM-2. Biapenem (5 mM, green) was treated with VIM-2 (280 nM) in 50 mM sodium phosphate at pH 7.5, 10% (*v*/*v*) D_2O. Evidence was accrued for formation of (2*R*)-imine and (2*S*)-imine and enamine products (blue, purple, and pink, respectively). 8 spectra were acquired (16 scans, every 2.5 min) over 20 min (stacked from bottom/first spectrum to top). The bottom spectrum is biapenem (5 mM) without added enzyme; Figure S4. ^{1}H-NMR time course analyses for reaction of biapenem with VIM-1 and VIM-2 (MBLs) and KPC-2 (an SBL). Methyl group resonance peak integrations from time course data were analysed to monitor reaction of biapenem to give (2*R*)-imine, (2*S*)-imine and enamine (blue, pink, and purple, respectively) products using a Bruker AVIIIHD 600 MHz machine. Biapenem (5 mM) was treated with (A) VIM-1 (280 nM), (B) VIM-2 (280 nM), (C) KPC-2 (280 nM) and (D) VIM-2 (100 nM) in 50 mM sodium phosphate at pH 7.5, 10% (*v*/*v*) D_2O. Integration of methyl resonance of biapenem was normalised ((integration of biapenem) / (integration of biapenem + (2*R*)-imine + (2*S*)-imine + enamine) $\times$ 100 was used to calculate the percentage of biapenem turnover; Figure S5. View of the overall fold of the VIM-2: biapenem derived complex. (A) The structure was processed in the space group *I*222 (PDB 6Y6J). Zinc ions are shown as grey spheres and chloride ions are in green. (B) Active site view of VIM-2 complexed with biapenem derived products. Relevant loops for interaction are labelled. The biapenem derived enamine ligand is in pale pink, and the biapenem derived imine ligand is in pale green (Figure S6); Figure S6. (A) Interactions between the biapenem derived enamine ligand (3) and the VIM-2 active site. (B) Interactions between the biapenem derived (2*S*)-imine (1) derived ligand and the VIM-2 active site; Figure S7. Comparisons of conformations of carbapenem and penem derived ligands with VIM-1 and VIM-2. (A) View of VIM-1 complexed with (i) the meropenem derived (2S)-imine ligand in slate blue (PDB 5N5I [19]); (ii) the VIM-2 biapenem derived (2*S*)-imine in green; and (iii) the VIM-2 biapenem derived enamine in pink. (PDBs: 6Y6J). (B) View of VIM-2 complexed with the (i) faropenem derived product (purple) (PDBs: 7A5Z [20]); (ii) the biapenem derived (2*S*)-imine in green and (iii) the biapenem derived enamine in pink. (PDBs: 6Y6J). Note that the binding modes of the faropenem derived imine (*Z*)-alkene (dark purple) and biapenem (2*S*)-imine derived ligands are similar; Figure S8. Overlay of the conformations of biapenem derived enamine and imine ligands with β-lactamases and penicillin binding proteins. Colour code: VIM-2 (class B1 MBL) complexed with the biapenem derived enamine—light pink; VIM-2 complexed with the biapenem derived (2*S*)-imine—light green (PDB 6Y6J); PBP2X complexed with the biapenem derived enamine—slate blue (PDB 2ZC3), PBP1A complexed with the biapenem substrate—orange) (PDB 2ZC5) [24], and CphA (class B2 MBL) complexed with the biapenem derived enamine—purple (PDB 1X8I) [23].

Author Contributions: Conceptualisation, J.B. and C.J.S.; methodology, A.L.; provision of materials, C.L.T., K.C. and T.R.M.; formal analysis, A.L., M.A.M. and T.R.M.; investigation, C.J.S. and J.S.; data curation, A.L.—original draft preparation, A.L.; writing, analysis, review, and editing, A.L. and C.J.S.; project administration, C.J.S.; funding acquisition, C.J.S. and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the Medical Research Council (EP/L016044/1, MR/T016035/1), the Biotechnology and Biological Research Council (BB/S50676X/1), and the Ineos Oxford Institute for Antimicrobial Research for funding. T.R.M. was supported by the BBSRC (BB/M011224/1). This research was funded in whole, or in part, by the Wellcome Trust (grant no. 106244/Z/14/Z and no. 099141/Z/12/Z).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: PDB available at https://www.rcsb.org/structure/6Y6J (accessed on 20 January 2022).

Acknowledgments: We thank the beamline scientists at the Diamond Light Source and the NMR team at the Chemistry Research Laboratory for their technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bush, K.; Bradford, P.A. β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, 1–22. [CrossRef] [PubMed]
2. Fair, R.J.; Tor, Y. Antibiotics and Bacterial Resistance in the 21st Century. *Perspect. Med. Chem.* **2014**, *6*, 25–64. [CrossRef] [PubMed]
3. Bush, K. Bench-to-Bedside Review: The Role of Beta-Lactamases in Antibiotic-Resistant Gram-Negative Infections. *Crit. Care.* **2010**, *14*, 1–8. [CrossRef] [PubMed]
4. Bush, K. Metallo-Beta-Lactamases: A Class Apart. *Clin. Infect. Dis.* **1998**, *27*, S48–S53. [CrossRef]
5. Paterson, D.L.; Bonomo, R.A. Extended-Spectrum β-Lactamases: A Clinical Update. *Clin. Microbiol. Rev.* **2005**, *18*, 657–686. [CrossRef] [PubMed]
6. Meini, M.-R.; Llarrull, L.I.; Vila, A.J. Evolution of Metallo-β-lactamases: Trends Revealed by Natural Diversity and in vitro Evolution. *Antibiotics* **2014**, *3*, 285–316. [CrossRef] [PubMed]
7. Lohans, C.T.; Van Groesen, E.; Kumar, K.; Tooke, C.L.; Spencer, J.; Paton, R.S.; Brem, J.; Schofield, C.J. A New Mechanism for β-Lactamases: Class D Enzymes Degrade 1β-Methyl Carbapenems through Lactone Formation. *Angew. Chem.* **2018**, *130*, 1296–1299. [CrossRef]
8. Aertker, K.M.J.; Chan, H.T.H.; Lohans, C.T.; Schofield, C.J. Analysis of β-Lactone Formation by Clinically Observed Carbapenemases Informs on a Novel Antibiotic Resistance Mechanism. *J. Biol. Chem.* **2020**, *295*, 16604–16613. [CrossRef] [PubMed]
9. Lohans, C.T.; Freeman, E.I.; van Groesen, E.; Tooke, C.L.; Hinchliffe, P.; Spencer, J.; Brem, J.; Schofield, C.J. Mechanistic Insights into β-Lactamase-Catalysed Carbapenem Degradation Through Product Characterisation. *Sci. Rep.* **2019**, *9*, 1–9. [CrossRef]
10. van Groesen, E.; Lohans, C.T.; Brem, J.; Aertker, K.M.; Claridge, T.D.W.; Schofield, C.J. 19F-NMR Monitoring of Reversible Protein Post-Translational Modifications: Class D B-Lactamase Carbamylation and Inhibition. *Chem. Eur. J.* **2019**, *25*, 11837–11841. [CrossRef] [PubMed]
11. Lisa, M.-N.; Palacios, A.R.; Aitha, M.; González, M.M.; Moreno, D.M.; Crowder, M.W.; Bonomo, R.A.; Spencer, J.; Tierney, D.L.; Llarrull, L.I.; et al. A General Reaction Mechanism for Carbapenem Hydrolysis by Mononuclear and Binuclear Metallo-β-Lactamases. *Nat. Comm.* **2017**, *8*, 538. [CrossRef] [PubMed]
12. Bonfiglio, G.; Maccarone, G.; Mezzatesta, M.L.; Privitera, A.; Carciotto, V.; Santagati, M.; Stefani, S.; Nicoletti, G. In vitro Activity of Biapenem against Recent Gram-Negative and Gram-Positive Clinical Isolates. *Chemotherapy* **1997**, *43*, 393–399. [CrossRef] [PubMed]
13. Kumagai, T.; Tamai, S.; Abe, T.; Matsunaga, H.; Hayashi, K.; Kishi, I.; Shiro, M.; Nagao, Y. New Straightforward Synthesis and Characterization of a Unique 1β-Methylcarbapenem Antibiotic Biapenem Bearing a σ-Symmetric Bicyclotriazoliumthio Group as the Pendant Moiety. *J. Org. Chem.* **1998**, *63*, 8145–8149. [CrossRef]
14. Perry, C.M.; Ibbotson, T. Biapenem. *Drugs* **2002**, *62*, 2221–2234. [CrossRef] [PubMed]
15. Lepsanovic, Z.; Libisch, B.; Tomanovic, B.; Nonkovic, Z.; Balogh, B.; Füzi, M. Characterisation of the First VIM Metallo- *β* -Lactamase-Producing *Pseudomonas Aeruginosa* Clinical Isolate in Serbia. *Acta Microbiol. Immunol. Hung.* **2008**, *55*, 447–454. [CrossRef] [PubMed]
16. Mavroidi, A.; Tsakris, A.; Tzelepi, E.; Pournaras, S.; Loukova, V.; Tzouvelekis, L.S. Carbapenem-hydrolysing VIM-2 metallo-beta-lactamase in Pseudomonas aeruginosa from Greece. *J. Antimicrob. Chemother.* **2000**, *46*, 1041–1043. [CrossRef]
17. Docquier, J.-D.; Lamotte-Brasseur, J.; Galleni, M.; Amicosante, G.; Frère, J.-M.; Rossolini, G.M. On functional and structural heterogeneity of VIM-type metallo-beta-lactamases. *J. Antimicrob. Chemother.* **2003**, *51*, 257–266. [CrossRef]
18. Makena, A.; Düzgün, A.Ö.; Brem, J.; Mcdonough, M.A.; Rydzik, A.M.; Abboud, M.I.; Saral, A.; Çiçek, Ç.A.; Sandalli, C.; Schofield, C.J. Comparison of Verona Integron-Borne Metallo-Lactamase (VIM) Variants Reveals Differences in Stability and Inhibition Profiles. *Antimicrob. Agents Chemother.* **2016**, *60*, 1377–1384. [CrossRef]
19. Salimraj, R.; Hinchliffe, P.; Kosmopoulou, M.; Tyrrell, J.M.; Brem, J.; van Berkel, S.S.; Verma, A.; Owens, R.J.; McDonough, M.A.; Walsh, T.R.; et al. Crystal Structures of VIM-1 Complexes Explain Active Site Heterogeneity in VIM-Class Metallo-β-Lactamases. *FEBS J.* **2019**, *286*, 169–183. [CrossRef] [PubMed]
20. Lucic, A.; Hinchliffe, P.; Malla, T.R.; Tooke, C.L.; Brem, J.; Calvopiña, K.; Lohans, C.T.; Rabe, P.; McDonough, M.A.; Armistead, T.; et al. Faropenem Reacts with Serine and Metallo-β-Lactamases to Give Multiple Products. *Eur. J. Med. Chem.* **2021**, *215*, 1–10. [CrossRef]
21. Mehta, S.C.; Furey, I.M.; Pemberton, O.A.; Boragine, D.M.; Chen, Y.; Palzkill, T. KPC-2 β-lactamase enables carbapenem antibiotic resistance through fast deacylation of the covalent intermediate. *J. Biol. Chem.* **2021**, *296*, 1–15. [CrossRef] [PubMed]
22. Galdadas, I.; Lovera, S.; Pérez-Hernández, G.; Barnes, M.D.; Healy, J.; Afsharikho, H.; Woodford, N.; Bonomo, R.A.; Gervasio, F.L.; Haider, S. Defining the Architecture of KPC-2 Carbapenemase: Identifying Allosteric Networks to Fight Antibiotics Resistance. *Sci. Rep.* **2018**, *8*, 1–13. [CrossRef]

23. Garau, G.; Bebrone, C.; Anne, C.; Galleni, M.; Frère, J.-M.; Dideberg, O. A Metallo-β-lactamase Enzyme in Action: Crystal Structures of the Monozinc Carbapenemase CphA and its Complex with Biapenem. *J. Mol. Biol.* **2005**, *345*, 785–795. [CrossRef] [PubMed]
24. Yamada, M.; Watanabe, T.; Baba, N.; Takeuchi, Y.; Ohsawa, F.; Gomi, S. Crystal Structures of Biapenem and Tebipenem Complexed with Penicillin-Binding Proteins 2X and 1A from *Streptococcus Pneumoniae*. *Antimicrob. Agents Chemother.* **2008**, *52*, 2053–2060. [CrossRef] [PubMed]
25. Bianchet, M.A.; Pan, Y.H.; Basta, L.A.B.; Saavedra, H.; Lloyd, E.P.; Kumar, P.; Mattoo, R.; Townsend, C.A.; Lamichhane, G. Structural insight into the inactivation of Mycobacterium tuberculosis non-classical transpeptidase LdtMt2 by biapenem and tebipenem. *BMC Biochem.* **2017**, *18*, 1–14. [CrossRef] [PubMed]
26. Yamaguchi, Y.; Jin, W.; Matsunaga, K.; Ikemizu, S.; Yamagata, Y.; Wachino, J.; Shibata, N.; Arakawa, Y.; Kurosaki, H. Crystallographic Investigation of the Inhibition Mode of a VIM-2 Metallo-β-Lactamase from *Pseudomonas Aeruginosa* by a Mercaptocarboxylate Inhibitor. *J. Med. Chem.* **2007**, *50*, 6647–6653. [CrossRef] [PubMed]
27. Tooke, C.L.; Hinchliffe, P.; Bragginton, E.C.; Colenso, C.K.; Hirvonen, V.H.A.; Takebayashi, Y. Spencer, J. β-Lactamases and β-Lactamase Inhibitors in the 21st Century. *J. Mol. Biol.* **2019**, *431*, 3472–3500. [CrossRef] [PubMed]
28. Brem, J.; Panduwawala, T.; Hansen, J.U.; Hewitt, J.; Liepins, E.; Donets, P.; Espina, L.; Farley, A.J.M.; Shubin, K.; Campillos, G.G.; et al. Imitation of β-lactam binding enables broad-spectrum metallo-β-lactamase inhibitors. *Nat. Chem.* **2022**, *14*, 15–24. [CrossRef]
29. Davies, D.T.; Leiris, S.; Sprynski, N.; Castandet, J.; Lozano, C.; Bousquet, J.; Zalacain, M.; Vasa, S.; Dasari, P.K.; Pattipati, R.; et al. ANT2681: SAR Studies Leading to the Identification of a Metallo-β-Lactamase Inhibitor with Potential for Clinical Use in Combination with Meropenem for the Treatment of Infections Caused by NDM-Producing *Enterobacteriaceae*. *ACS Infect. Dis.* **2020**, *6*, 2419–2430. [CrossRef]
30. Cahill, S.; Tarhonskaya, H.; Rydzik, A.; Flashman, E.; McDonough, M.A.; Schofield, C.J.; Brem, J. Use of ferrous iron by metallo-β-lactamases. *J. Inorg. Biochem.* **2016**, *163*, 185–193. [CrossRef] [PubMed]
31. Parkova, A.; Lucic, A.; Krajnc, A.; Brem, J.; Calvopiña, K.; Langley, G.W.; McDonough, M.A.; Trapencieris, P.; Schofield, C.J. Broad Spectrum β-Lactamase Inhibition by a Thioether Substituted Bicyclic Boronate. *ACS Infect. Dis.* **2020**, *6*, 1398–1404. [CrossRef] [PubMed]
32. Winter, G.; Waterman, D.G.; Parkhurst, J.M.; Brewster, A.S.; Gildea, R.J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I.; et al. DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 85–97. [CrossRef] [PubMed]
33. Brem, J.; van Berkel, S.S.; Zollman, D.; Lee, S.Y.; Gileadi, O.; McHugh, P.J.; Walsh, T.R.; McDonough, M.A.; Schofield, C.J. Structural Basis of Metallo-β-Lactamase Inhibition by Captopril Stereoisomers. *Antimicrob. Agents Chemother.* **2016**, *60*, 142–150. [CrossRef] [PubMed]
34. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. *Phaser* Crystallographic Software. *J. App. Crystallogr.* **2007**, *40*, 658–674. [CrossRef]
35. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213–221. [CrossRef]
36. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 352–367. [CrossRef]
37. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Struct. Biol.* **2004**, *D60*, 2126–2132. [CrossRef]
38. Willcott, M.R. MestRe Nova. *J. Am. Chem. Soc.* **2009**, *131*, 13180. [CrossRef]

Contents

List of Abbreviations **cxxxiv**

1 Introduction **1**

1.1 Proteases . . . 1

1.1.1 Protease classification . . . 2

1.2 SARS-CoV-2 Protease . . . 4

1.2.1 M^{pro} substrates . . . 5

1.2.2 M^{pro} mechanism . . . 8

1.2.3 Structure of M^{pro} . . . 9

1.2.4 Suitability of M^{pro} as a target for drug development . . . 11

1.3 Viral protease inhibitors . . . 12

1.3.1 HIV protease inhibitors . . . 12

1.3.2 HCV protease inhibitors . . . 14

1.3.3 SARS-CoV-2 M^{pro} peptidomimetic covalent inhibitors . . . 16

1.3.3.1 Inhibitor substituents at P4-P1′ positions to target M^{pro} S4-S1′ subsites 19

1.3.4 Peptide inhibitors of clinical interest . . . 24

1.3.4.1 Natural macrocyclic peptide inhibitors . . . 24

1.3.4.2 Library based screens to identify macrocyclic inhibitors . . . 29

1.3.4.3 Macrocyclic peptide inhibitors of M^{pro} . . . 31

1.4 L,D-transpeptidases . . . 32

1.4.1 Bacterial peptidoglycan layer . . . 33

1.4.2 Mycobacterial peptidoglycan layer . . . 34

1.4.3 Mycobacterial L,D-transpeptidases . . . 36

1.4.3.1 Structure of Ldt_{Mt2} and other paralogs . . . 37

1.4.3.2 Functions of the L,D-transpeptidases . . . 40

1.4.4 β-Lactams as potential drugs for *M. tuberculosis* therapy . . . 41

1.4.4.1 Carbapenems and penems . . . 42

1.4.4.2 Interaction of Ldt_{Mt2} with β-lactams . . . 45

1.5 Objectives . . . 52

2 Assay development, validation and application towards identification of β-lactam inhibitors against M^{pro} **53**

2.1 Abstract . . . 53

2.2 Introduction . . . 54

2.3 Results . . . 58

2.3.1 Small scale preparation of M^{pro} protein from pGEX-6p1-M^{pro} plasmid . . . 58

2.3.2 Activity of different C terminal constructs of M^{pro}. . . . 59

2.3.3 RapidFire Mass Spectrometry based assay development . . . 61

2.3.4 Solid Phase extraction based MS assay validation . . . 67

2.3.4.1 Assay validation with reported M^{pro} inhibitors . . . 73

2.3.4.2 Determination of catalytic coefficients of the 11-mer M^{pro} substrate peptide. . . . 74
2.3.4.3 Determination of equilibrium constants of M^{pro} monomer/dimer. . 75
2.3.4.4 Mode of M^{pro} inhibition of selected compounds identified from the commercial libraries . . . 76
2.3.5 Screening of commercial or synthetic β-lactam and non-β-lactam compounds for M^{pro} inhibition . . . 86
2.3.5.1 Grouping of screened compounds based on percentage inhibition . 92
2.3.5.2 Mode of interaction of β-lactam compounds . . . 95
2.3.5.3 Revalidation of selected screened in-house compounds via dry dispensing to assess reproducibility . . . 99
2.3.5.4 Validation of metal ion based M^{pro} inhibitors . . . 103
2.3.5.5 M^{pro} inhibition from cephalosporin C zinc salt arises from zinc metal ion . . . 104
2.3.5.6 Investigating the interaction of tazobactam and aztreonam with M^{pro}. 106
2.3.5.7 Investigating the inhibition of M^{pro} via monobactams. . . . 107
2.3.5.8 Investigating the inhibition of M^{pro} by penicillin derivatives. . . . 110
2.3.5.9 Validating the M^{pro} inhibition by the penicillin *S*-oxidised derivatives **4** and **5** . . . 112
2.3.6 C-Terminal M^{pro} constructs and their activity . . . 113
2.3.7 Assay development and validation . . . 114
2.3.8 β-Lactams inhibit M^{pro} . . . 117
2.4 Methods . . . 119
2.4.0.1 Reagents . . . 119
2.4.0.2 Small scale protein production and purification . . . 119
2.4.0.3 Protein production and purification . . . 120
2.4.0.4 SDS-PAGE Gel Electrophoresis . . . 122
2.4.1 Peptide synthesis and purification . . . 122
2.4.2 General information on the M^{pro} SPE-MS assays . . . 123
2.4.3 Kinetic parameters of M^{pro} . . . 124
2.4.3.1 Turnover assay with different M^{pro} constructs . . . 125
2.4.4 Single concentration M^{pro} inhibition assays . . . 125
2.4.4.1 In-house single concentration M^{pro} inhibition assays . . . 125
2.4.5 IC_{50} value determinations . . . 126
2.4.6 Protein Observed Mass Spectrometry Assay . . . 126

3 SAR studies with esters of β-lactam penicillin sulfones and γ-lactam derivatives against M^{pro} 128
3.1 Abstract . . . 128
3.2 Introduction . . . 130
3.2.1 Penicillin sulfones as M^{pro} inhibitors . . . 130

3.2.2 γ-Lactams as M^{pro} inhibitors . . . 131
3.2.3 Contributions to manuscript 2 . . . 133
3.2.4 Manuscript 2: Penicillin Derivatives Inhibit the SARS-CoV-2 Main Protease by Reaction with its Nucleophilic Cysteine . . . 134
3.3 Results . . . 134
3.3.1 Investigating the impact of the 11-mer and the 37-mer substrates on MS assays with β-lactam derivatives . . . 134
3.3.2 Comparison of SPE-MS assay results obtained with 11-mer and 37-mer substrate peptides . . . 142
3.3.3 Investigating the mode of interaction of bromopenicillins with M^{pro} using a denaturing MS assay . . . 144
3.3.4 Investigating the effect of γ-lactams and representative sulfoximines on M^{pro} activity . . . 147
3.4 Discussion . . . 152
3.5 Future directions . . . 153
3.6 Methods . . . 153
3.6.1 FRET assays . . . 153
3.7 Appendix . . . 154

4 Studies on M^{pro} substrates and peptide inhibitors of M^{pro} 155
4.1 Abstract . . . 155
4.2 Introduction . . . 156
4.2.1 Contributions to the manuscript: Discovery of SARS-CoV-2 M^{pro} peptide inhibitors based on substrate modelling and ligand binding experiments . . . 158
4.3 Results . . . 159
4.3.1 M^{pro} catalysis with natural peptide substrates . . . 159
4.3.2 Product inhibition of M^{pro} . . . 161
4.3.3 Investigation of truncated s01 substrate peptides as M^{pro} substrates and inhibitors 163
4.3.4 Investigation of a 20-mer substrate variant at the P1 position . . . 164
4.3.5 Validating the inhibitory effect of the peptide inhibitors on M^{pro} catalysis using assays with a 37-mer substrate . . . 165
4.4 Future directions . . . 166

5 Macrocyclic inhibitors of M^{pro} 167
5.1 Abstract . . . 167
5.2 Introduction . . . 168
5.3 Contributions to the manuscript . . . 168
5.3.1 Manuscript 4 (published prior to thesis deposition) . . . 168
5.4 Results . . . 169
5.4.1 Optimisation of the M^{pro} assay with a 37-mer substrate and an internal peptide standard . . . 169

5.4.2 Comparison of pIC_{50} values obtained for cyclic peptide M^{pro} inhibitors with or without normalisation to an internal standard . . . 172
5.4.3 Inhibition studies of peptide variants derived from TMGM4-05 . . . 174
5.4.3.1 Significance of the unnatural amino acid at the P1′ position towards the potency of TMGM4-05 and related macrocyclic peptides . . . 175
5.4.3.2 Significance of the solvent exposed residues on the potency of TMGM4-05 . . . 177
5.4.3.3 Investigating the significance of the crystallographically observed H3-Y9 intrapeptide interaction for TMGM4-05 inhibitory potency 178
5.4.3.4 Alanine scan of the TMGM4-05A peptide . . . 179
5.4.4 M^{pro} inhibition studies with shorter cyclic peptides . . . 180
5.4.5 Screening of tricyclic disulfide rich cyclic peptides as M^{pro} inhibitors . . . 182
5.4.5.1 Results . . . 183
5.5 Discussion . . . 185
5.6 Future directions . . . 187

6 Non-peptidomimetic inhibitors of M^{pro} 189
6.1 Abstract . . . 189
6.2 Introduction . . . 190
6.3 Contributions to the manuscripts 5-7 . . . 191
6.3.1 Manuscript 5: Structure-Activity Studies Reveal Scope for Optimisation of Ebselen-Type Inhibitors of SARS-CoV-2 Main Protease . . . 192
6.3.2 Manuscripts 6 and 7 describing the work carried out as part of the COVID Moonshot Consortium . . . 192
6.4 Future directions . . . 193

7 Mechanistic insights into interactions of Ldts with β-lactam antibiotics 194
7.1 Abstract . . . 194
7.2 Introduction . . . 195
7.3 Results and Discussion . . . 198
7.3.1 Protein production and purification . . . 198
7.3.2 Interaction of mycobacterial and enterococcal L,D-transpeptidases with representative β-lactams . . . 200
7.3.2.1 Interaction of Ldts with faropenem . . . 200
7.3.2.2 Interaction of Ldts with representative carbapenems . . . 201
7.3.2.3 Interaction of Ldts with a representative cephalosporin . . . 203
7.3.2.4 Interaction of Ldts with representative penicillin . . . 204
7.3.2.5 The nucleophilic serine variant of Ldt_{Mt2} does not react with β-lactams 206
7.3.3 Mechanistic insights into the interaction of penicillins with Ldts . . . 208
7.3.3.1 Application of NMR spectroscopy to monitor C5-C6 fragmentation of penicillins upon reaction with Ldts . . . 208

7.3.3.2 Degradation of aminopenicillins via recyclisation to form diketopiperazines . . . 209
7.3.4 Mechanistic insights into interaction and degradation of carbapenems . . . 217
7.3.4.1 Lactone formation . . . 218
7.3.4.2 Recyclisation of AECs to give β-lactams . . . 220
7.3.5 Mechanistic insights into interaction and degradation of cephalopsorins . . . 225
7.3.5.1 Recyclisation . . . 225
7.3.6 Mechanistic insights into interaction and degradation of faropenem, a penem subclass of β-lactam antibiotics . . . 230
7.3.6.1 D-amino acid mediated deacylation . . . 231
7.3.7 D-amino acid mediated deacylation of penicillin derived AECs . . . 235
7.4 Conclusions . . . 237
7.5 Future directions . . . 239
7.6 Methods . . . 240
7.6.1 Cloning, protein production and purification . . . 240
7.6.2 HPLC purification of diketopiperatizine **3** . . . 246
7.6.3 NMR Spectroscopy . . . 246
7.6.3.1 Monitoring enzymatic hydrolytic products of penicillins . . . 246
7.6.3.2 D-amino acid mediated deacylation of C5-C6 fragment bound Ldt_{Mt2} acyl-enzyme. . . . 246
7.6.3.3 Recyclisation of carbapenems and nitrocefin. . . . 247
7.7 Appendix . . . 247

8 Mechanistic studies employing NMR on β-lactamase catalysed hydrolysis of carbapenems and a penem 290
8.1 Abstract . . . 291
8.2 Introduction . . . 291
8.3 Contributions to the manuscripts . . . 292
8.3.1 Manuscript 8: Faropenem reacts with serine and metallo-β-lactamases to give multiple products . . . 293
8.3.2 Manuscript 9: Studies on the Reactions of Biapenem with VIM metallo-β-lactamases and the serine-β-lactamase KPC-2 . . . 293
8.4 Additional work to that described in manuscript 9 . . . 293
8.5 Future directions . . . 298

9 Discussion 299

10 Conclusions 306

Bibliography 315

List of Abbreviations

aa . . . . . Amino acids

6,7-ADTN . . . (±)-2-Amino-6,7-dihydroxy-1,2,3,4-tetrahydronaphthalene hydrobromide

FRET . . . . . Fluorescence Resonance Energy Transfer

LOPAC^{1280} . . . Library of Pharmacologically Active Compounds

FDA . . . . . Food and Drug Administration

2TY . . . . . 2-Tryptone/yeast extract

3CLpro . . . . . 3-Chymotrypsin-like cysteine protease

A280 . . . . . Absorbance at 280 nm

AMC 7-Amino-4-methylcoumarin

AEPA . . . . . [4-(2-Aminoethyl)phenyl]-acetic acid

BSA . . . . . Bovine Serum Albumin

Cmpd . . . . . Compound

COSY . . . . . Correlation Spectroscopy

CoV-2 . . . . . Coronavirus type 2

CV . . . . . Column volume

d . . . . . Doublet

Da . . . . . Dalton (mass unit)

DAP . . . . . Diaminopimelic acid

DHP-I . . . . . Dehydropeptidase I

DMSO-d_6 . . . Deuterated dimethylsulfoxide

Dnp 2,4-Dinitrophenyl

DpaA . . . . . *Meso*-diaminopimelic acid protein amidase A

DRPs . . . . . Disulfide rich cyclic peptides

DTT . . . . . Dithiothreitol

E. coli . . . . . *Escherichia coli*

FDA . . . . . Food and Drug Administration

FPLC . . . . . Fast Protein Liquid Chromatography

GABA . . . . . γ-aminobutyric acid

GlcNaC . . . . . *N*-acetylglucosamine

GST . . . . . Glutathione S-transferase

HCV . . . . . Hepatitis C virus

HEPES . . . . . 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid

HIV . . . . . Human Immunodeficiency Virus

HMBC . . . . . Heteronuclear Multiple-bond Correlation Spectroscopy

HPLC . . . . . High Performance Liquid Chromatography

HSQC . . . . . Heteronuclear Single-Quantum Correlation Spectroscopy

HT . . . . . High throughput

IMAC . . . . . Immobilized metal affinity chromatography

IPA-3 . . . . . 1,1'-Disulfanediyldinaphthalen-2-ol

IPTG . . . . . Isopropyl β-D-1-thiogalactopyranoside

LC-MS . . . . . Liquid Chromatography-Mass Spectrometry

Ldts . . . . . L,D-transpeptidases

Ldt_{Mt} . . . . . L,D-transpeptidases from *Mycobacterium tuberculosis*

LIC . . . . . Ligation Independent Cloning

m . . . . . Multiplet

M . . . . . Membrane

MS . . . . . Mass spectrometry (MS)

Mca . . . . . 7-Methoxycoumarin-4-acetic acid

MALDI . . . . . Matrix-Assisted Laser Desorption/Ionisation

MERS . . . . . Middle East Respiratory Syndrome

M^{pro} . . . . . Main protease

MurNGlyc . . . N-Glycolyl muramic acid

N . . . . . Nucleocapsid

NMR . . . . . Nuclear Magnetic Resonance

NEMO . . . . . Nuclear factor essential modulator

NOESY . . . . . Nuclear Overhauser Effect Spectroscopy

nsps . . . . . Non-structural proteins

ORF . . . . . Open Reading Frame

PAINS . . . . . Pan Assay Interference Compounds

PBPs . . . . . Penicillin binding proteins

PCR . . . . . Polymerase Chain Reaction

PG . . . . . Peptidoglycan

PLPro . . . . . Papain-like protease

ppm . . . . . Parts per million

q . . . . . Quartet

QM/MM . . . . Quantum Mechanics/Molecular Mechanics

RaPID . . . . . Random nonstandard Peptide Integrated Discovery

RF . . . . . RapidFire

s . . . . . Singlet

S . . . . . Spike

SAMDI . . . . . Self-assembled monolayer desorption ionization

SARS . . . . . Severe Acute Respiratory Syndrome

SDS-PAGE . . . Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis

SPE-MS . . . . Solid Phase Extraction coupled to Mass Spectrometry

SPPS . . . . . Solid Phase Peptide Synthesis

TCEP . . . . . Tris(2-carboxyethyl)phosphine

TFA . . . . . Trifluoroacetic acid

THF . . . . . Tetrahydrofuran

TLCK . . . . . N_α-*p*-Toluenesulfonyl-L-lysine chloromethyl ketone hydrochloride

TOCSY . . . . . Total Correlation Spectroscopy

TOF . . . . . Time-of-Flight

Tosyl . . . . . Toluenesulfonyl

TPCK . . . . . *N*-*p*-Toluenesulfonyl-L-phenylalanine chloromethyl ketone

Tris . . . . . Tris(hydroxymethyl)aminomethane

UV . . . . . Ultraviolet

Vis . . . . . Visible

Z-L-Phe . . . . *N*-Carbobenzyloxy-L-phenylanlanyl

WHO . . . . . World Health Organisation

1 Introduction

1.1 Proteases

Proteases catalyse the hydrolysis of amide bonds in peptides and proteins. About 6% of all proteins in living organisms are estimated to be proteases [1]; proteases are ubiquitously encoded in the genomes of viruses, bacteria, fungi, and eukaryotes. It is therefore perhaps unsurprising that ~2% of the human genome encodes for approximately 570 proteases [2]. Protease catalysis can alter the function of target substrates as a means of regulating expression at a post transcriptional level; they play roles in regulating the localisation, activity, and stability of proteins, and in the synthesis of new biomolecules. Proteases directly regulate important biological processes such as DNA replication, the cell cycle, apoptosis, the unfolded protein response, homeostasis, senescence, ovulation and necrosis [3].

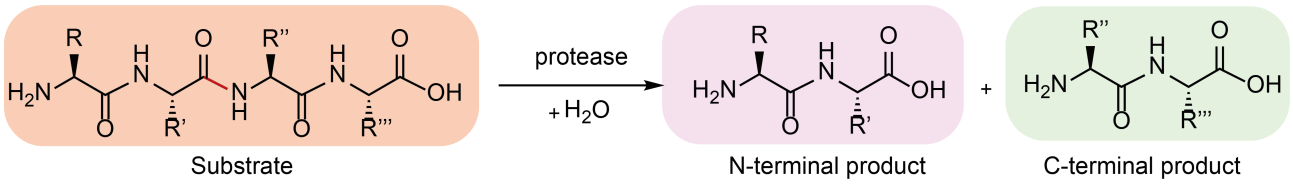


Figure 1.1 | Protease catalysed hydrolysis of a peptide substrate yields N- and C- terminus cleaved products.

The dysregulation of protease activity can alter biological processes leading to pathological conditions in humans [3]. The activity and localisation of proteases is thus tightly regulated *in vivo*. Many proteases are synthesised in an inactive precursor form (zymogen) that can be activated by the auto-catalysed cleavage of the protease or by catalysis mediated by another protease. The activation of some zymogens may require additional factors e.g. inactive pro-caspases are activated by the apoptosome [4]. Alternate modes of protease precursor activation have been reported, such as allosteric

mechanisms (e.g. matrix metallo-proteases [5]) or changes in physiological conditions (e.g. the pH dependent activation of pepsinogen [6]). Altering protease activity by using either small molecules or biomolecules is an attractive and validated target for drug development to treat diseases [7, 8].

1.1.1 Protease classification

Proteases can be grouped on the basis of their mechanisms, their biological functions, their amino acid sequences, and their homology to other proteases. Classifying proteases according to conservation at their catalytic active site is desirable as it correlates with their mechanism of action. The MOREPS protease database [9] classifies proteases into families with catalytically important serine, cysteine, threonine, aspartate, glutamate, and asparagine residues. Metallo-proteases are a separate family. Some protease families use more than one residue for the same role. For example, pantetheinyl hydrolase from *Streptomyces cattleya* employs Thr282 for autocleavage and activity [10, 11]; however, the homologous protein in *Ochrobactrum anthropi* uses serine in catalysis [9]. A selection of proteases grouped according to their key residues at the active site is depicted in Figure 1.2.

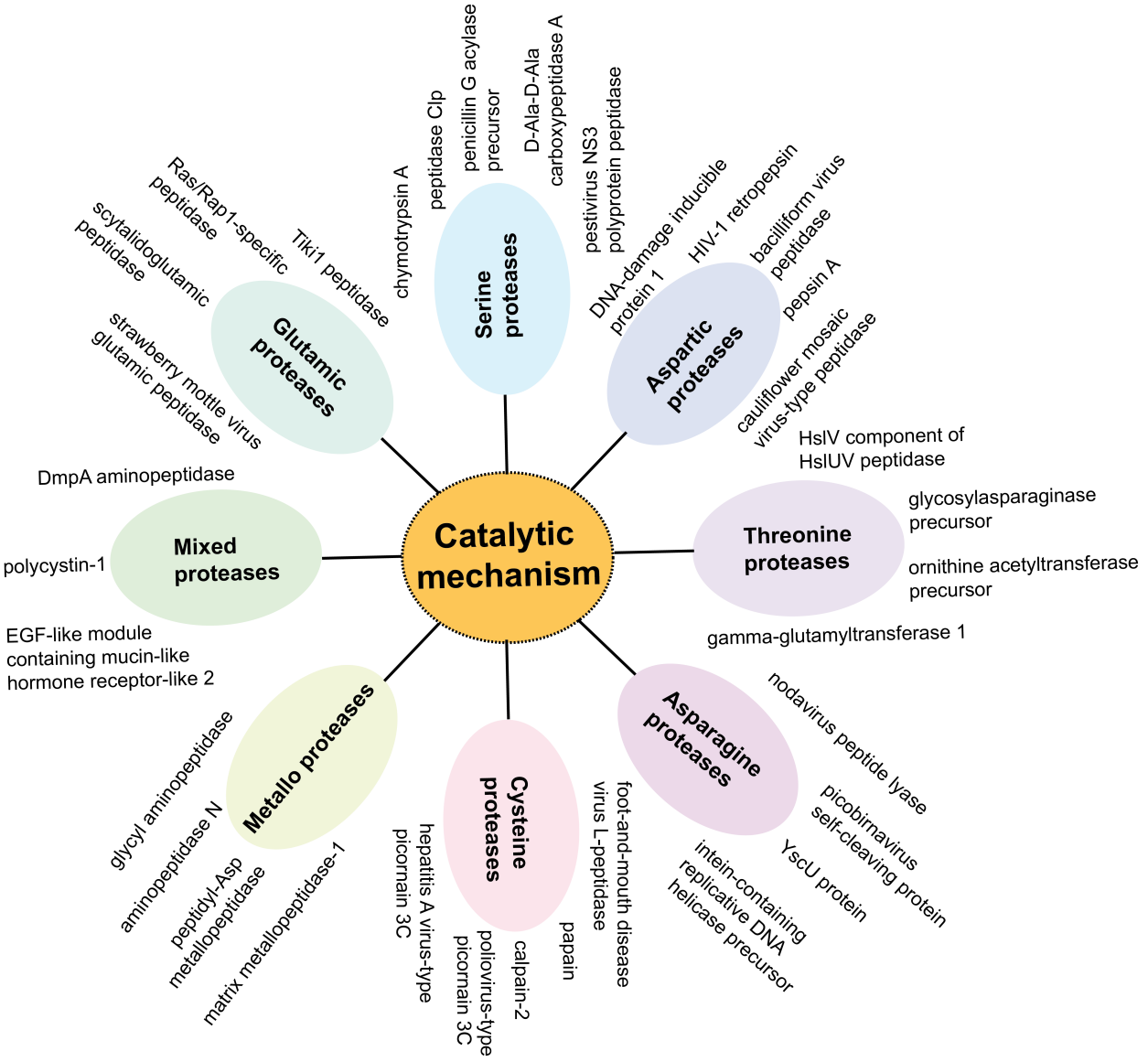


Figure 1.2 | Classification of proteases based on important active site residues. [9]

For proteases utilising Ser, Thr, or Cys as a nucleophile, their reactions normally proceed via the formation of an acyl-enzyme complex. Such proteases are important in virus biology. For example, catalysis by the hepatitis C virus (HCV) NS3-4a protease (Figure 1.3A) proceeds via nucleophilic attack of Ser139 (activated by proton abstraction by His57) onto the scissile amide bond. The subsequently formed tetrahedral intermediate is stabilised by binding in an oxyanion hole, comprised of two backbone NH groups [12]. Following the collapse of the tetrahedral intermediate, the C-terminal product is released, and an acyl-enzyme complex is formed. His57 then abstracts a proton from a water molecule which attacks the acyl-enzyme intermediate to form a second tetrahedral intermediate,

which collapses to release the N-terminal product [13].

In the case of the aspartyl proteases such as the human immunodeficiency virus (HIV) protease, catalysis involves activation of a water molecule that attacks the scissile substrate amide forming a tetrahedral intermediate, as depicted in Figure 1.3B. Protonation of the NH group of the scissile bond involving proton shuffle between the initially protonated aspartate leads to hydrolysis of the scissile bond [14].

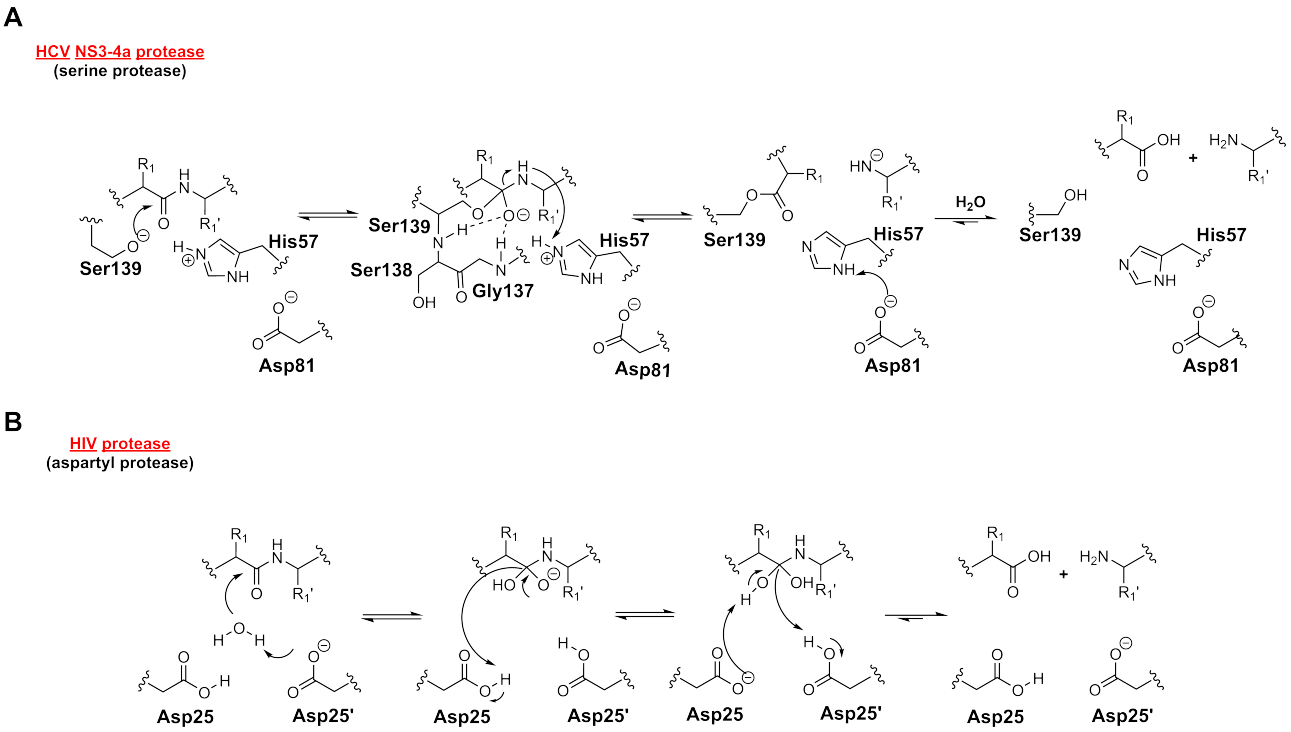


Figure 1.3 | Outline mechanism for A. a hepatitis C virus (HCV) nucleophilic serine protease and B. a human immunodeficiency virus (HIV) aspartyl protease.

Proteases belonging to the same family share sequence homology in their catalytic domains and hence use common catalytic residues (apart from mixed proteases). Protease families that share a similar tertiary structure may not appear related when only considering the primary structure [3,9]. Therefore, grouping based on function, catalytic residues, and structures i.e. protein-species [9] can enable the useful categorisation of proteases, including with respect to inhibition studies.

1.2 SARS-CoV-2 Protease

The severe acute respiratory disease coronavirus-2 (SARS-CoV-2) is the etiological agent of the Coronavirus disease 2019 (COVID-19) pandemic which manifests similar respiratory symptoms to

that of the severe acute respiratory syndrome (SARS) [15]. As of June 2022, COVID-19 has resulted in more than 5.5 million global deaths [16], i.e., approximately 7000 fold more deaths than from the SARS epidemic caused by SARS-CoV-1 virus in 2002 [17]. The genome of SARS-CoV-2 encodes for 27 different proteins [18] (Figure 1.4). The SARS-CoV-2 polyproteins include two proteases:the main protease (M^{pro}) and papain-like protease (PL^{pro}), which catalyse the proteolytic processing of the viral polyproteins into functional nsps [19–22].

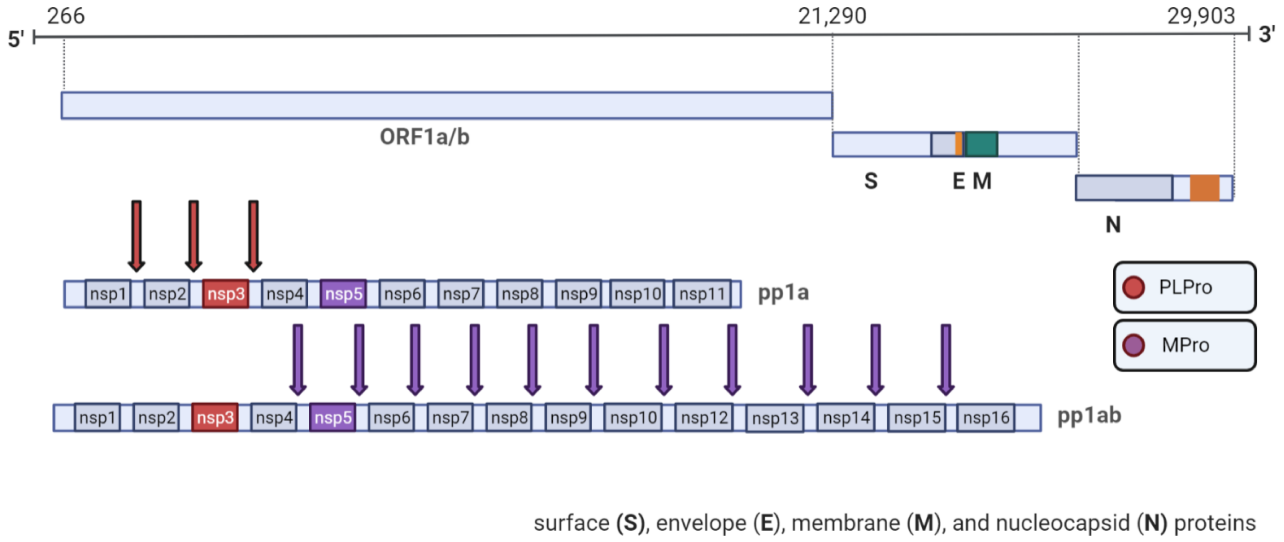


Figure 1.4 | M^{pro} and PL^{pro} mediated processing of the orf1a/b polyprotein. PL^{pro} catalyses the hydrolysis of the viral polyprotein pp1a at three sites, two of which releases PL^{pro} from the polyprotein. The catalytic activity of M^{pro} on the polyprotein pp1ab releases itself (nsp5) and nsps6-16 which are essential for viral replication. Arrows represent the cleavage sites of PLpro and M_{pro} in polyproteins pp1a in pp1ab, respectively.

1.2.1 M^{pro} substrates

According to the Schechter-Berger [23] nomenclature for proteases, the residues N- and C- terminus to the substrate scissile bond are referred as P1 and P1′, respectively. The adjacent substrate residues are labelled in incremental order as P1, P2, P3, P4 or P1′, P2′, P3′, P4′, respectively (Figure 1.5). The M^{pro} subsites that accommodate the corresponding substrate residues are denoted as S1, S2, S3, S4 or S1′, S2′, S3′, S4′, etc.

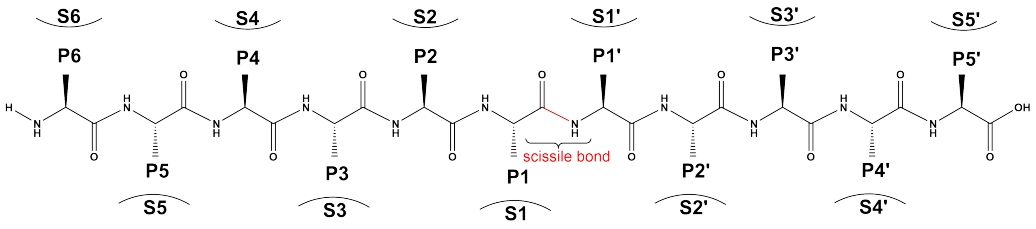
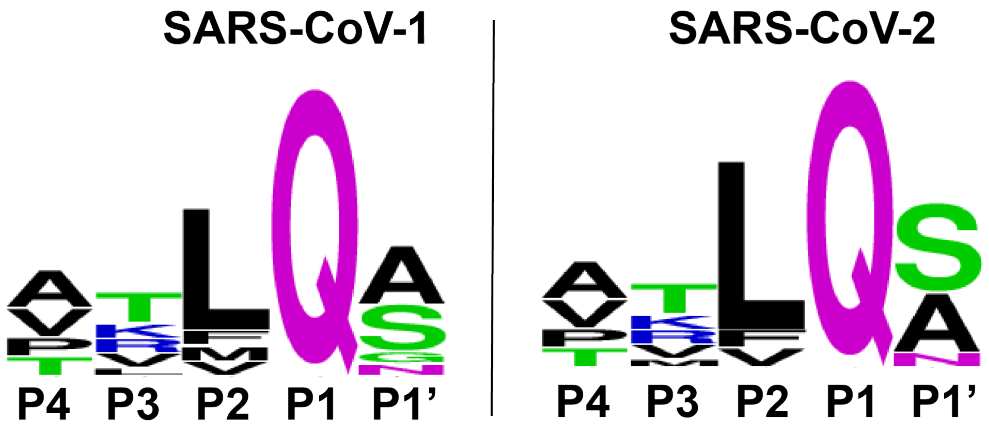


Figure 1.5 | Schechter-Berger nomenclature [23] for protease subsites and substrate residues. The prime and non-prime sites which represent N- and C- terminal residues with respect to the scissile bond (red) are labeled according to their relative position.

M^{pro} is a nucleophilic cysteine protease that catalyses hydrolysis of the orf1a/b viral polyprotein at 11 different sites. M^{pro} (nsp5) autocatalyses cleavage of itself from the polyprotein (i.e. at the nsp4/5 and nsp5/6 sites, respectively). The autocleavage event of M^{pro} from orf1a/b is followed by trans-cleavage at distal sites in the orf1a/b polyprotein to release nsp6-16, which are essential for viral replication [24]. Analysis of the M^{pro} cleavage sequences (nsp4-16, Table 1.1) reveals that Gln is ubiquitously present at the P1 position and that Leu and Ser are the most conserved residues at the P2 and P1′ positions, respectively, in both the SARS-CoV-1 and the SARS-CoV-2 genomes. Less conservation of the M^{pro} substrate residues is observed between the SARS-CoV-1 and the SARS-CoV-2 M^{pro} cleavage site, that is distal to the P2-P1′ positions, potentially reflecting the high structural plasticity of the sub-pockets recognising these residues (Table 1.1). Notably, Ser followed by Ala are the most prevalent residues at the P1′ position of SARS-CoV-1 M^{pro} cleavage sites; Ala and Ser are conserved in the analogous SARS-CoV-2 M^{pro} cleavage sites.

Table 1.1: Frequency logo map of the conserved residues at P4-P1′ positions of SARS-CoV-1 and SARS-CoV-2 M^{pro} substrates. The size of the amino acid letter corresponds to the conserved frequency of that residue at the stated position i.e., the bigger the letter, the higher the abundance of the amino acid in M^{pro} substrate cleavage site.



	Cleavage site	SARS-CoV-1	SARS-CoV-2
s01	nsp4/nsp5	TSAVLQ/SGFRK	TSAVLQ/SGFRK
s02	nsp5/nsp6	SGVTFQ/GKFKK	SGVTFQ/SAVKR
s03	nsp6/nsp7	KVATVQ/SKMSD	KVATVQ/SKMSD
s04	nsp7/nsp8	NRATLQ/AIASE	NRATLQ/AIASE
s05	nsp8/nsp9	SAVKLQ/NNELS	SAVKLQ/NNELS
s06	nsp9/nsp10	ATVRLQ/AGNAT	ATVRLQ/AGNAT
s07	nsp10/nsp11	REPLMQ/SADAS	REPMLQ/SADAQ
s08	nsp12/nsp13	PHTVLQ/AVGAC	PHTVLQ/AVGAC
s09	nsp13/nsp14	NVATLQ/AENVT	NVATLQ/AENVT
s10	nsp14/nsp15	TFTRLQ/SLENV	TFTRLQ/SLENV
s11	nsp15/nsp16	FYPKLQ/ASQAW	FYPKLQ/SSQAW

M^{pro} is reported to display proteolytic activity against human host substrates with a preference for Gln at the P1 position and for a small residue at P1′ position; these reactions may assist the virus in evading the innate immune response upon infection [25,26]. For instance, nuclear factor essential modulator (NEMO), an immune signalling protein, is reported to be cleaved between Gln231 and Val232 by SARS-CoV-2 M^{pro} [25]. Proteolytic cleavage by SARS-CoV-1 and SARS-Cov-2 M^{pro} has been reported with optineurin, which plays a significant role in the inactivation of an innate immune response upon viral infection [26,27]. Additionally, cleavage of human substrates with His at P1 has been observed, although less frequently than substrates with Gln at P1 [26]. As a P1 Gln

residue is generally considered to be important for M^{pro} cleavage activity, the observations for M^{pro} catalysed cleavage of substrates with His at P1 oppose this paradigm and indicate that M^{pro} may be able to access a wider pool of substrates *in vivo* [26].

1.2.2 M^{pro} mechanism

The M^{pro} catalytic dyad is composed of a Cys145-His41 pair of residues (Figure 1.6). Similarly to the mechanism of the HCV serine protease (Figure 1.3), the thiol proton of Cys145 is deprotonated by His41. This enables the S_γ of Cys145 to attack the scissile amide bond. A tetrahedral complex is then formed, which is stabilised by the coordination of the negatively charged substrate amide carbonyl-derived oxygen with the NH groups from the main chain amides of Cys145 and Gly142. Breakdown of the tetrahedral complex releases the C-terminal product peptide to form a thioester with Cys145 (acyl-enzyme complex). Subsequently, an activated water molecule attacks the thioester of the acyl-enzyme complex, forming a second tetrahedral intermediate. The collapse of this intermediate complex releases the N-terminal product returning the enzyme to its ground state.

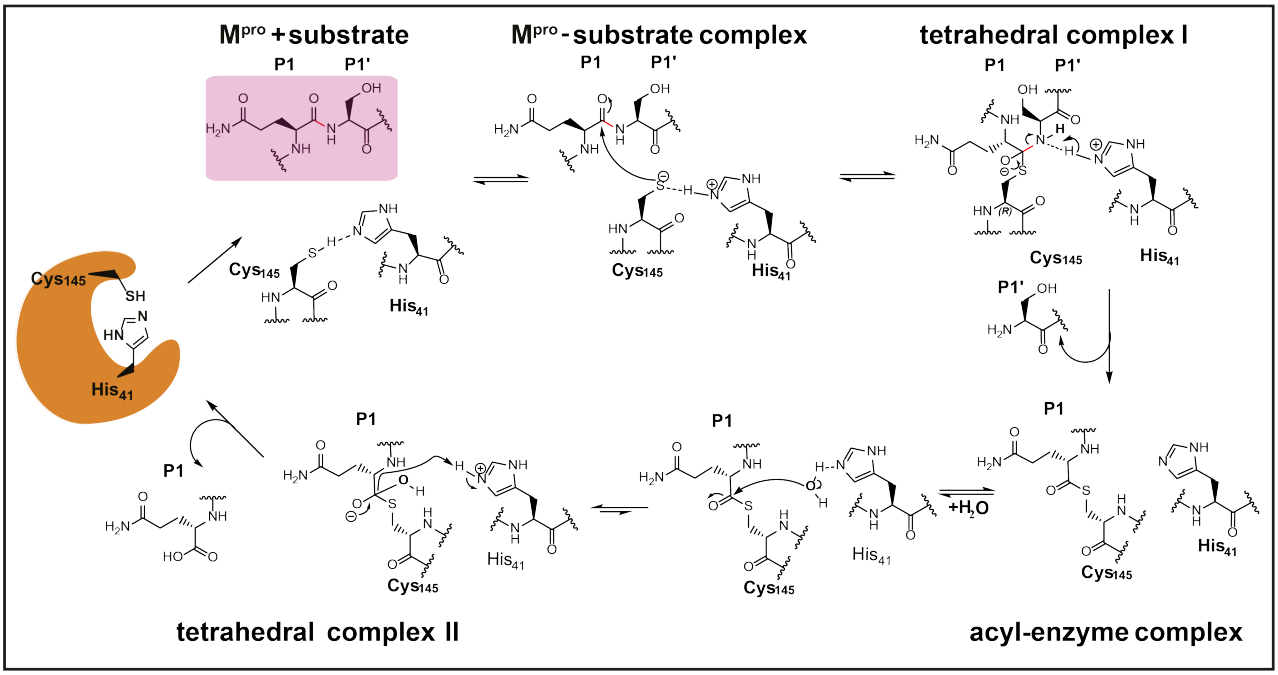


Figure 1.6 | Proposed outline mechanism for M^{pro} catalysed substrate proteolysis. M^{pro} catalysed peptide cleavage occurs between P1 (conserved Gln) and P1′ hydrophobic residues.

1.2.3 Structure of M^{pro}

M^{pro} is comprised of three distinct domains: six stranded antiparallel β-sheets form domains I (residues 10-99) and II (residues 100-182); domain III is α-helical in nature (residues 198-303) which provides the surface for dimerisation of M^{pro} (Figure 1.7A) [28]. Eight N-terminal residues (the N-finger) flank domain I and the C-terminal tail flanks domain III (Figure 1.7B). Two M^{pro} protomers (A and B) dimerise to form catalytically active M^{pro} (Figure 1.8A) [28,29]. Hydrophobic interactions between domains III of the two protomers [30] and a salt-bridge between Glu290 of protomer A and Arg4 of protomer B stabilise the dimer. The N-finger of protomer B protrudes into the active site of protomer A to help form the S1 subsite by positioning its Ser1 to interact with the side chain of Glu166 of protomer A (Figure 1.8C) and vice versa. However, the volumes of the two active sites in the dimer likely differ [31,32] and only one active site is considered to have a sufficiently large volume to efficiently bind the substrates in a catalytically productive manner [33]. The catalytic dyad His41 and Cys145 at the active site is flanked by residues from domains I and II (Figure 1.8B).

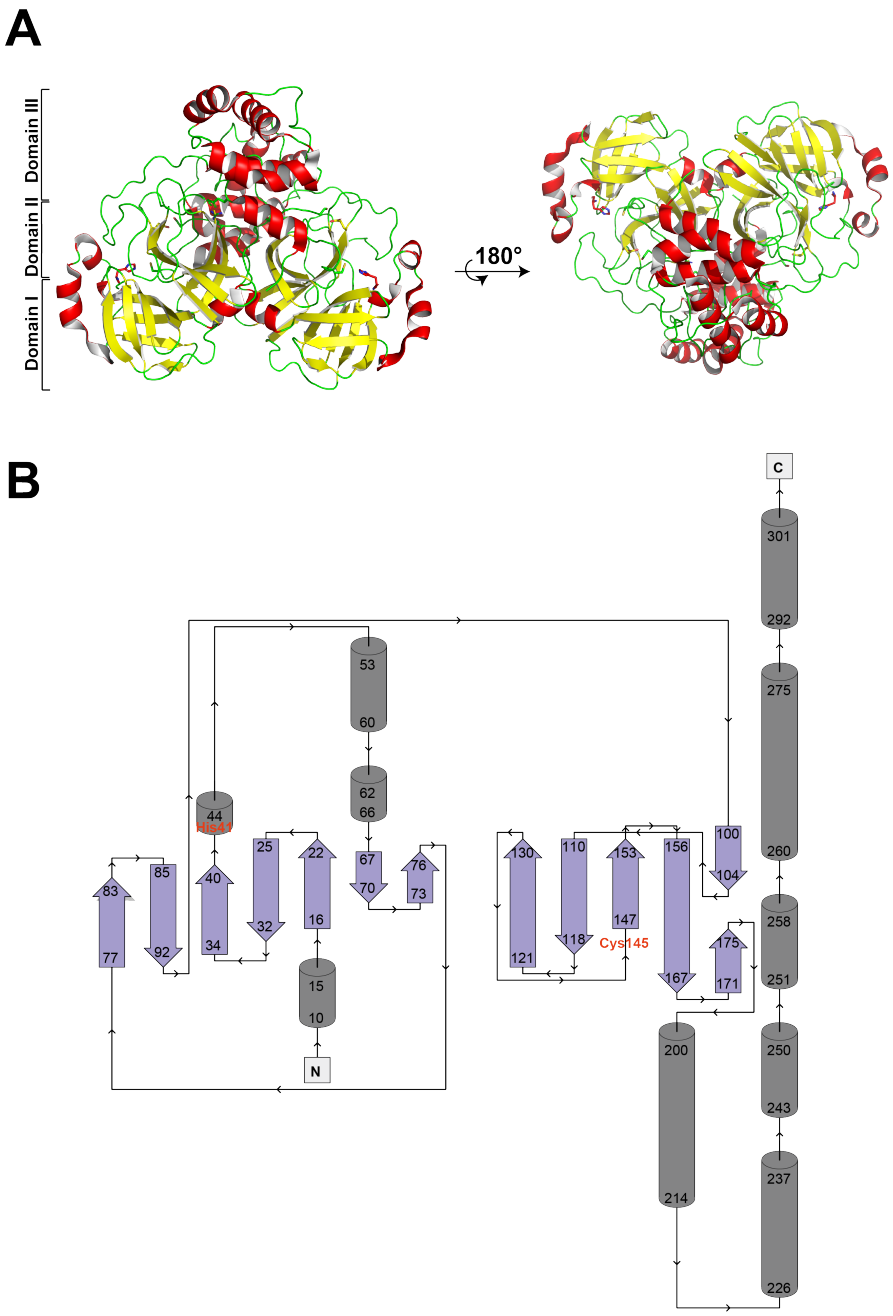


Figure 1.7 | Views from a M^{pro} crystal structure and domain topology of M^{pro}. A. Cartoon representation of the the structure of M^{pro} (PDB:6LU7). Domain I and II are comprised of β-barrel sheets (yellow). α-Helical domain III (red) is involved in homodimerisation. **B.** Domain topology of M^{pro}.

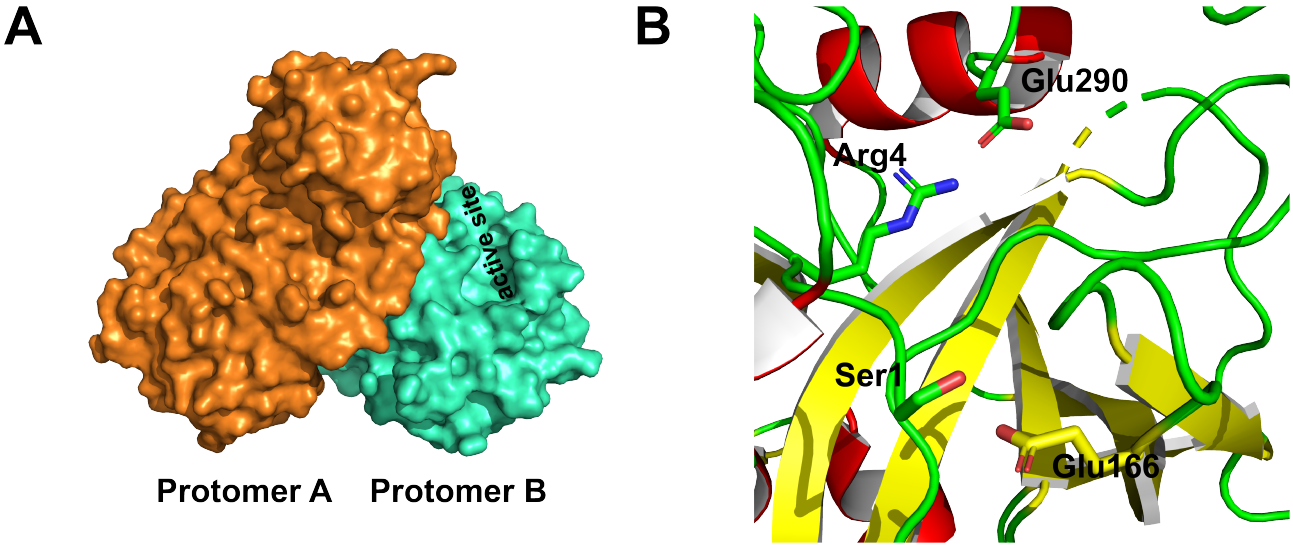


Figure 1.8 | Dimerisation of M^{pro} and view of its active site. **A.** Surface representation of protomers A and B of M^{pro} (PDB:6LU7). **B.** Interaction between the two protomers of M^{pro}. Arg4 interacts with Glu290 of another promoter to form a salt bridge and this helps to stabilise the dimer. Ser1 from the N-finger of protomer A forms part of the active site at the promoter B via hydrogen bonding to the side chain of Glu166.

1.2.4 Suitability of M^{pro} as a target for drug development

PL^{pro} cleaves after an LxGG motif [34] which is present in the substrates of human deubiquinitases such as ZUP1 [35]. By contrast, the substrate specificity of M^{pro} is proposed to be non-overlapping with human proteases [28,36,37], making it attractive from a medicinal chemistry perspective to be pursued as a target for the development of selective M^{pro} protease inhibitors.

Drug discovery efforts subsequent to the advent of the pandemic and in parallel with the work described in this thesis, have led to the development of breakthrough COVID-19 antivirals including nirmatrelvir which targets M^{pro} [38,39]. The Medicines and Healthcare products Regulatory Agency in the UK has approved two additional therapies, molnupiravir [40]which targets RNA-dependent RNA polymerase [41] and a prophylactic monoclonal antibody based treatment with Evusheld [42,43] which targets the spike protein of SARS-CoV-2. Although currently approved SARS-CoV-2 therapies may be life-saving, the approved drugs have safety concerns. Molnuprevir can induce mutations in mammalian cells [44], while ritonavir, which is currently co-administered with nirmatrelvir, may cause apparent hepatological injury [45,46]. Therefore, despite the clinical efficacy of both drugs, there is a need for

new antivirals to improve safety profiles and/or combat potential antiviral resistance [38].

1.3 Viral protease inhibitors

Proteases are crucial for viral maturation and replication; in general they are believed to not share similar substrate profiles to human proteases [47,48]. Thus, they are attractive therapeutic targets [49]. Inhibitors of viral proteases are important therapeutics as exemplified by inhibitors for the treatment of HIV and HCV viral infections. Protease inhibitors are included in the the top 200 drugs by sales [50–52], demonstrating their importance to human health.

The World Health Organisation (WHO) has reported that 37.7 million [7] people are living with HIV globally and an estimated 58 million [53] people suffer from chronic HCV infections. Sixty-five percent of HIV patients are undergoing anti-retroviral therapy with medications that only manage the infection [7]. Protease inhibitors constitute a major component of HIV antiviral therapy. By contrast to HIV protease inhibitors, it is reported that HCV protease inhibitors can cure >95% of HCV infections [53].

Medicinally useful protease inhibitors are often peptidomimetic type molecules that, at least initially, were designed based on resemblance to the substrates cleaved by their respective proteases [54,55]. Thus, these inhibitors bind to the substrate binding cleft of the protease which spans both the prime and non-prime sites [47]. Inhibitor binding can be enhanced via covalent reaction by an electrophilic warhead such as an aldehyde, an α-ketoamide, or a nitrile. However, peptidomimetic inhibitors frequently suffer from the same challenges in drug development as peptides themselves, i.e. low absorption, high metabolic clearance, and short half-lives [55,56].

1.3.1 HIV protease inhibitors

Saquinavir, a selective peptidomimetic inhibitor of the HIV protease, was the first FDA-approved protease inhibitor for HIV treatment [57]. Subsequently, nine other HIV protease inhibitors have been marketed for managing HIV infections (Figure 1.9) [58]. All are peptidomimetic inhibitors of the homodimer HIV protease, and bind non-covalently including via a hydroxy group in hydroxyethylamine or phosphonooxyethylamine (fosamprenavir prodrug) groups (apart from tipranavir) (Figure 1.9); such interactions

are proposed to help the inhibitor mimic the transition state of proteolytic catalysis [59]. Proline is preferred at the P1′ position in HIV protease substrates [54]. Substitution of the proline by a decahydro-isoquinoline derivative was shown to improve potency in saquinavir and nelfinavir (Figure 1.9) [54]. Toxicity, low potency, and poor oral bio-availability have resulted in reduced use of the first generation protease inhibitors, such as saquinavir, nelfinavir, indinavir, and ritonavir as antiviral therapies [60].

Increasing antiviral resistance has stimulated the development of second generation HIV protease inhibitors [47,61]. Structural analogues of first generation drugs (ritonavir → lopinavir) or non-peptidic inhibitors (tipranavir) (Figure 1.9) have been developed to improve potency and oral bioavailability [60]. Darunaprevir, a derivative of amprenavir/fosamprenavir (Figure 1.9) has a bis-tetrahydrofuran ring instead of the tetrahydrofuran ring present in the the earlier inhibitors, such as amprenavir and fosamprenavir, enabling darunaprevir to overcome antiviral resistance [62].

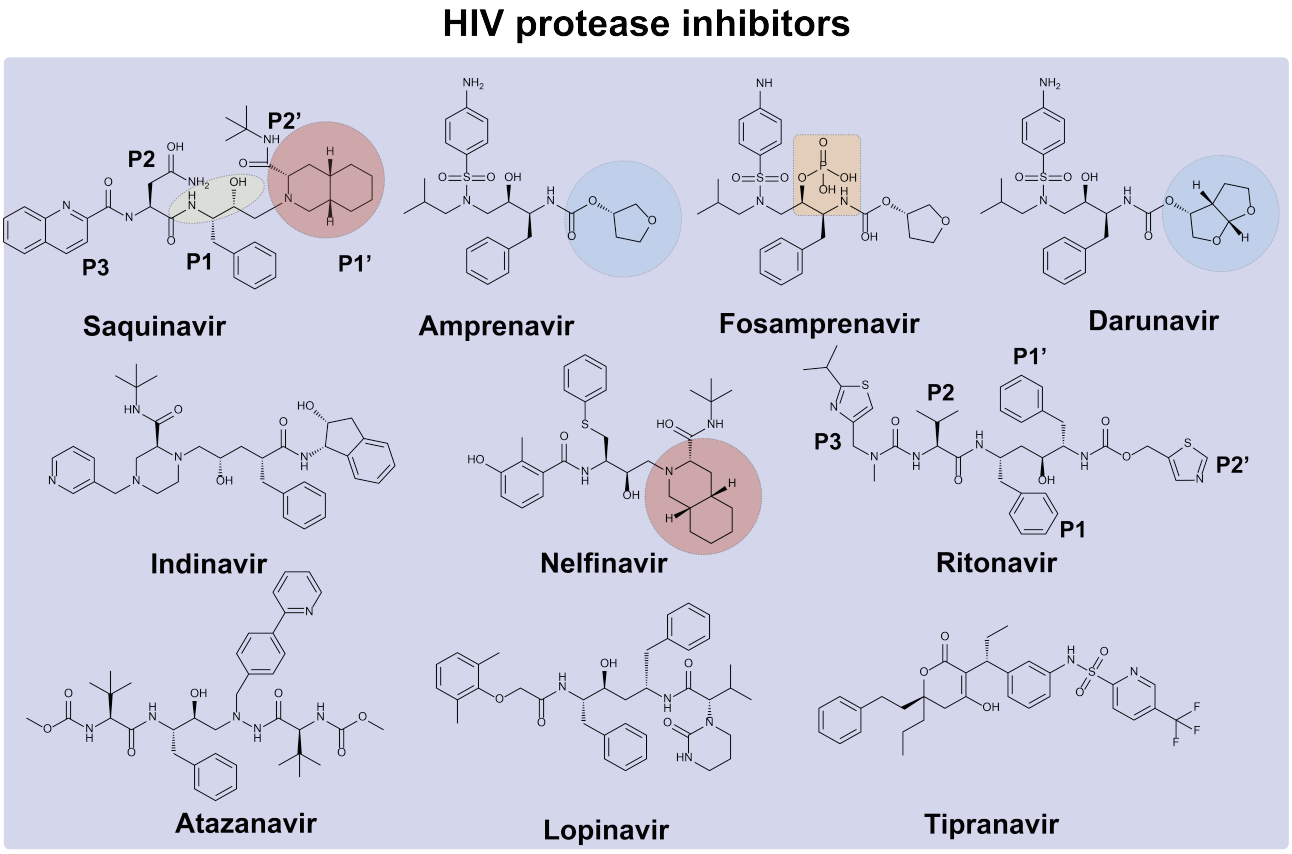


Figure 1.9 | Viral protease inhibitors approved for clinical use against HIV infections. First generation HIV antivirals such as saquinavir, indinavir, nelfinavir and ritonavir have been withdrawn from the market due to toxicity or have superseded by superior second-generation inhibitors [47,60]. Functional groups discussed in the text are highlighted (blue: bis-tetrahydrofuran/monotetrahydrofuran, red:decahydro-isoquinoline, yellow: ethylamine isostere, orange: phosphonooxyethylamine).

1.3.2 HCV protease inhibitors

Three protease inhibitors (glecaprevir, grazoprevir, and voxilaprevir) [63] (Figure 1.10) are approved for treatment of HCV infections in the USA; an additional protease inhibitor, paritaprevir, has been approved in the UK [64] (Figure 1.10). Boceprevir and telaprevir were first approved for clinical use in 2011 [8], but their use has been superseded by superior second-generation inhibitors. The lead compounds for boceprevir and telaprevir (Figure 1.10) were a substrate-derived undecapeptide and a decapeptide bearing dicarbonyl electrophiles (α-keto amides) spanning P6-P5′ and P6-P4′ respectively. The peptidic nature of the lead compounds was reduced via a series of ′depeptidization′ steps which led to potent HCV protease inhibitors, boceprevir (K_i = 14 nM, lead compound K_i = 1.9 nM) [65] and telaprevir (K_i = 7.0 nM) [66].

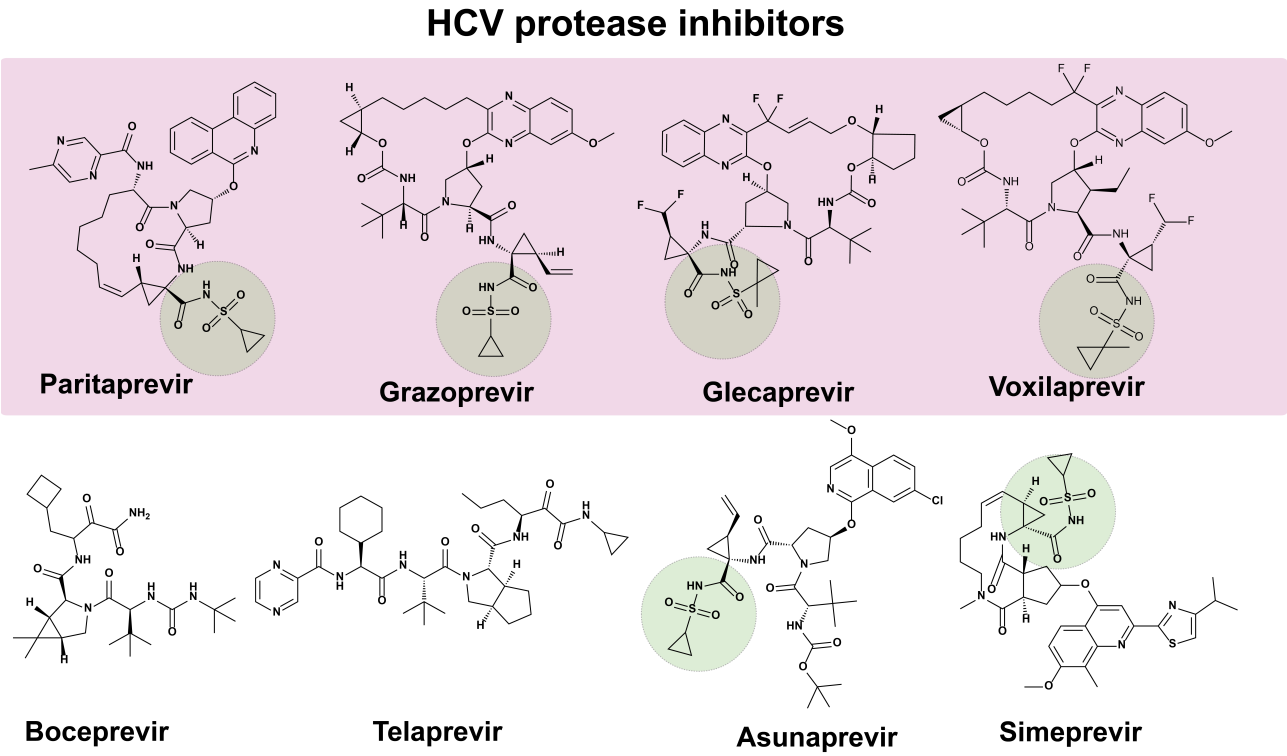


Figure 1.10 | Viral protease inhibitors approved for clinical use against HCV infections. First-generation HCV antivirals such as boceprevir and teleprevir, have been withdrawn due to toxicity or availability of superior second generation inhibitors [47,60]. *N*-Acylsulfonamide, a functional group discussed in the text is highlighted in green. Drugs which are currently used for HCV treatment are highlighted in pink.

The high doses of boceprevir/teleprevir required for HCV treatment have likely contributed to increased antiviral resistance against the first generation HCV protease inhibitors, leading to their replacement with second generation protease inhibitors [67]. Simeprevir was optimised from an initial lead compound, BILN 2061, which had poor toxological profile in animal studies, by the introduction of an *N*-acylsulfonamide group at P1, a tris-substituted cyclopentane-dicarbonyl moiety at P2, and a cyclopentyloxy amide group at P4 (Figure 1.11). An *N*-acylsulfonamide group at P1 is now present in all clinically used protease inhibitors (Figure 1.10) which bind tightly to the shallow pocket in S1′ subsite of HCV protease [68,69].

Genomic analyses have enabled broader spectrum classification of HCV genomes into six genotypes based on sequence homology [70]. However, not all HCV protease inhibitors exhibit pangenotypic HCV activity. For instance, simeprevir has a narrow genotype spectrum activity (genotype 1 [71] or, in some cases, used with genotype 4 [72]). Consequently, since 2018, Janssen

has discontinued production of simeprevir as other protease inhibitors manifest broader spectrum activity (e.g. glecaprevir, voxilaprevir). Notably, all of the currently prescribed protease inhibitors for HCV treatment are macrocycles (Figure 1.10).

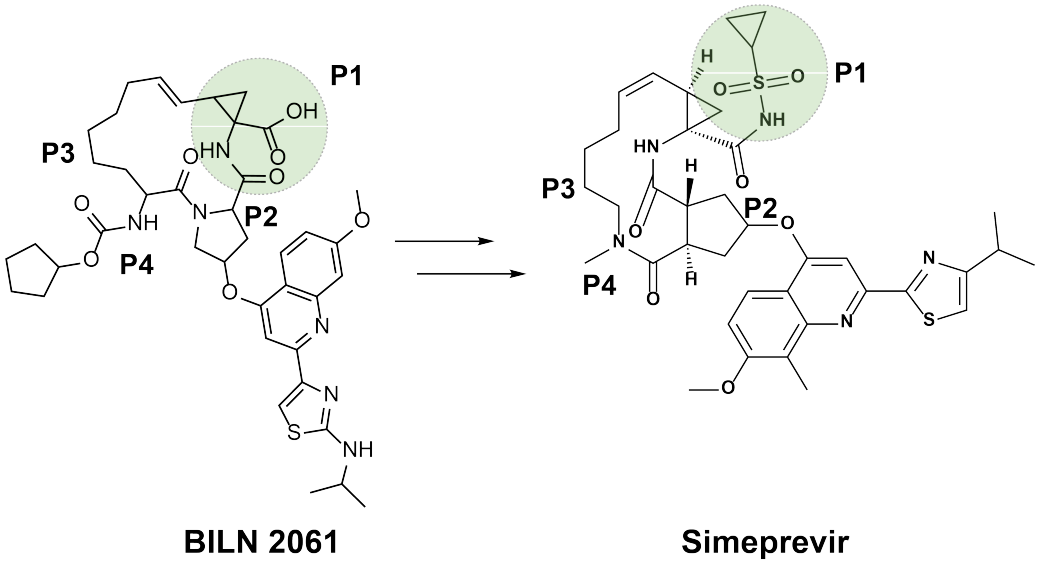


Figure 1.11 | Optimisation of the HCV inhibitor BILN 2061 via substitutions at P1, P2 and P4 positions to give to simeprevir with improved oral bioavailability and toxicological profile.

The clinical advances in HIV/HCV treatment via the development of protease inhibitors suggest that other viral diseases can also be targeted in a similar manner. Herpes virus, which can lead to complications in immuno-compromised patients [73], relies on proteases for viral maturation via hydrolysis of the polyproteins. The cleavage activity of the protease on its polyprotein enables assembly of the scaffold proteins and RNA polymerases which are essential for viral replication [74]. Furthermore, retroviruses implicated in immunodeficiency, leukaemia, mammary tumour, and infectious anaemia encode retroviral homodimeric aspartic proteases essential for viral maturation [48]. Other viruses like human coronaviruses employ cysteine proteases for viral maturation and for evading innate host immune response [75].

1.3.3 SARS-CoV-2 M^{pro} peptidomimetic covalent inhibitors

To date, most reported SARS-CoV-2 M^{pro} peptidomimetic inhibitors are based on the nsp4/5 cleavage site (SAVLQ↓SGFRK), in part, because of the high catalytic turnover observed for the nsp4/5 derived

peptide substrate at this site [28] and their efficiency to bind M^{pro}. However, when the SARS-CoV-2 related work described in this thesis was commenced, N3, a peptidomimetic inhibitor was the only reported selective M^{pro} protease inhibitor [28]. N3 possesses a covalent warhead (a Michael acceptor), which reacts with the active site Cys145 [28], a benzyl ester at the P1′ position, a γ-lactam derivative at the P1, Leu, Val and Gly at the P2-P4 positions interact with S1′, S1, S2, S3 and S4 subsites of M^{pro}, respectively [76]. The isoxazole at the P5 position is solvent exposed in a structure for the complex of N3 with M^{pro} (Figure 1.12A) [28].

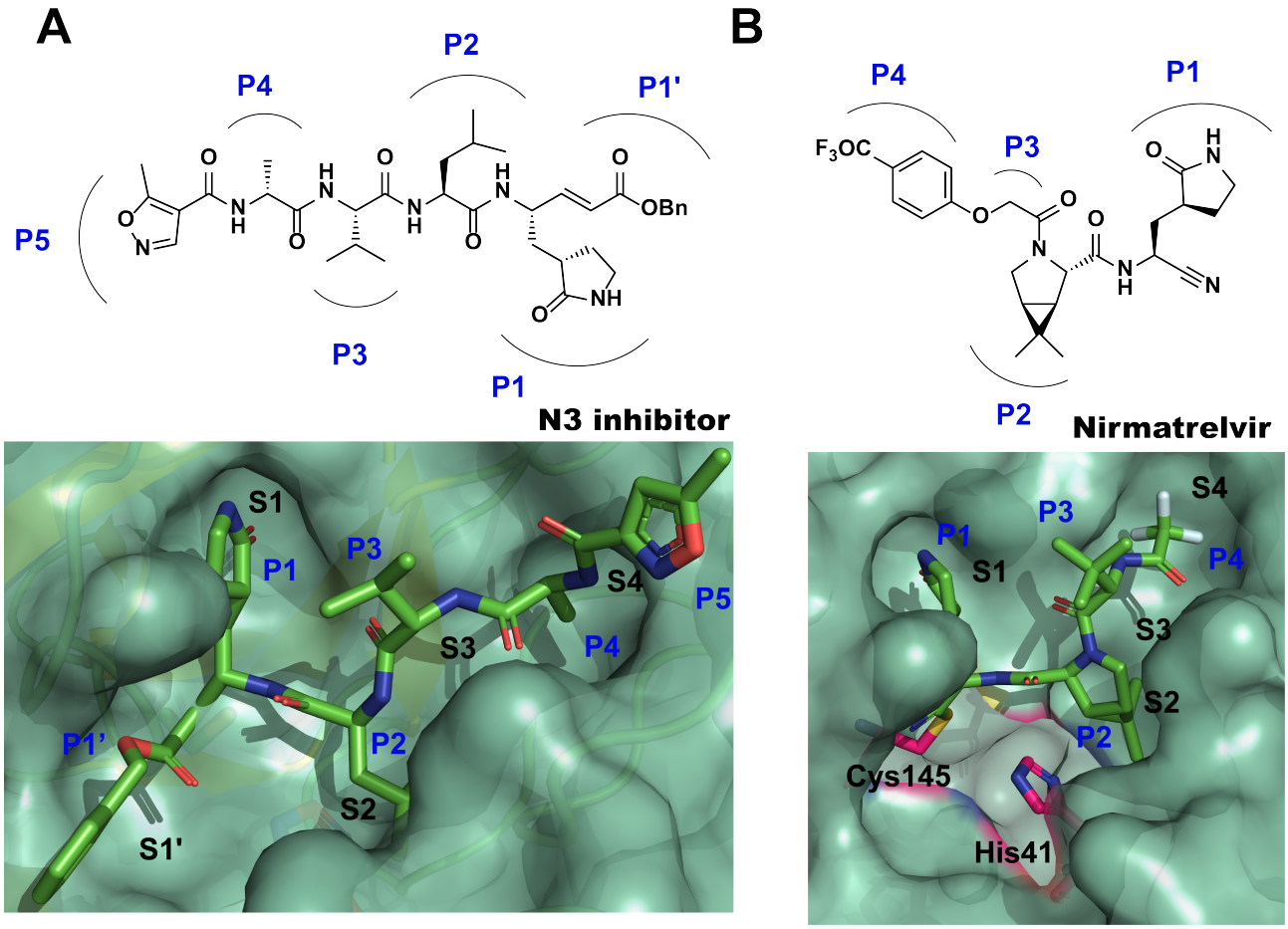


Figure 1.12 | M^{pro} inhibitors. **A.** View of N3 inhibitor at the active site of M^{pro}(PDB: 6LU7). Structure of N3 inhibitor is shown. **B.** View of Cys145 linked nirmatrelvir at the active site of M^{pro}(PDB: 7VH8). The structure of nirmatrelvir is shown.

Subsequent to the development of N3, known protease inhibitor compounds including the HCV protease inhibitor boceprevir [77], the feline antiviral GC376 (a prodrug of an aldehyde inhibitor GC373) [78], and the caspase-1 inhibitor SDZ 224015 [79] have been identified as M^{pro} inhibitors (Figure

1.13). In addition, the ketone-based phosphate prodrug inhibitor, PF-00835231 [80] and multiple α-ketoamides [29,37] (originally developed for SARS-CoV-1 M^{pro} inhibition) have been optimised for SARS-CoV-2 M^{pro} inhibition. Previous inhibition studies with SARS-CoV-1 M^{pro}, which shares 96% sequence similarity [81,82] and 100% similarity at the active site with SARS-CoV-2 M^{pro}, have facilitated the rapid development of potent and selective M^{pro} inhibitors, including the aldehyde-based covalently reacting inhibitors (designed based on boceprevir and nirmatrelvir) [83,84], e.g. MI-09, that display potent inhibition against M^{pro}. Likewise, nirmatrelvir (Pfizer), which was approved in December 2019, binds tightly at the active site and inhibits M^{pro} (Figure 1.12B).

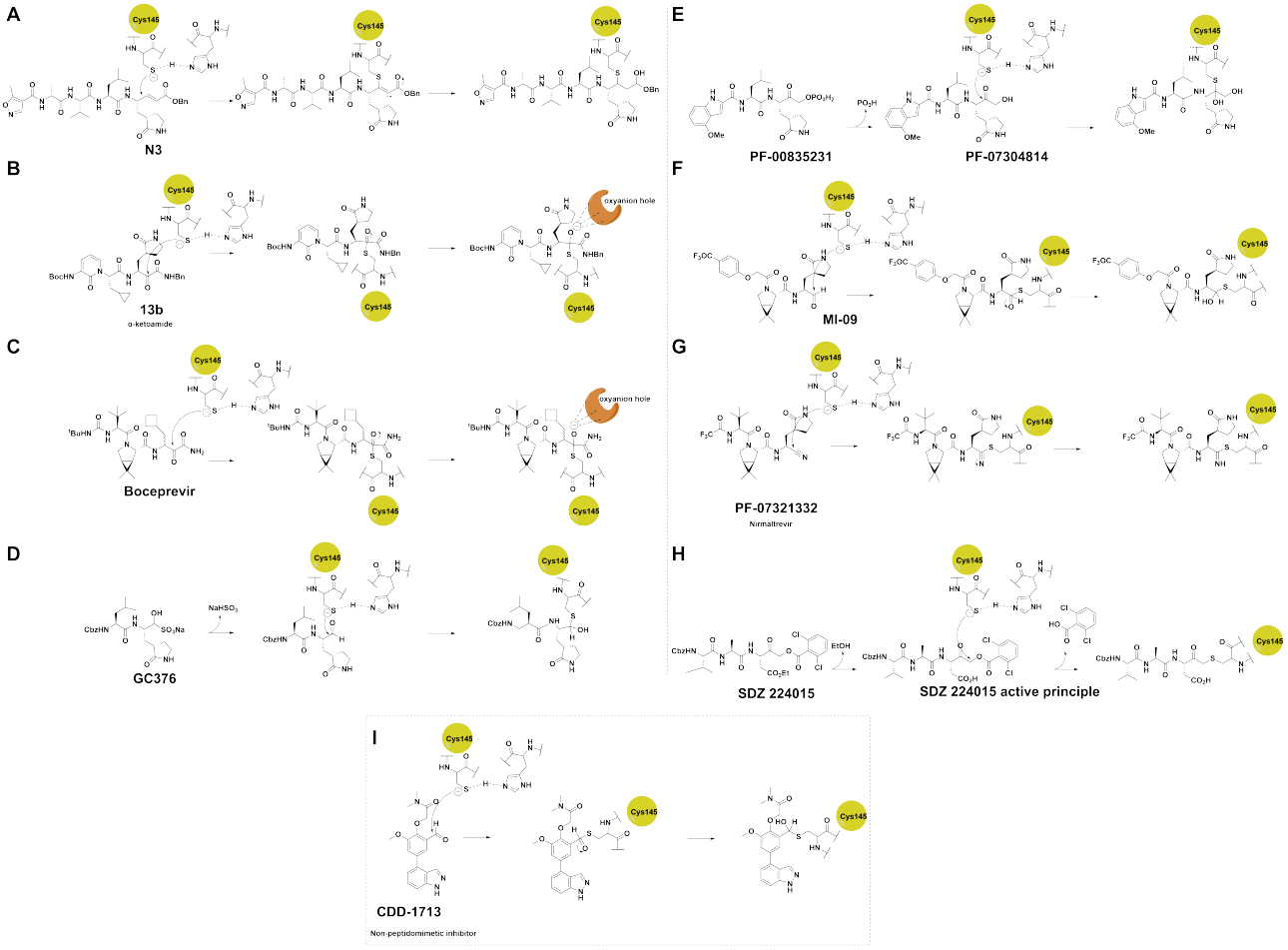


Figure 1.13 | Selected reported covalently reacting inhibitors of M^{pro}. **A.** N3 [28], **B.** 13b, an α-ketoamide [37], **C.** boceprevir [77], **D.** GC376 [78], **E.** PF-00835231 and itS active principle PF-07304814 [80], **H.** PF-073212332 [38], **F.** MI-09 [83,84], and **G.** repurposed HCV protease inhibitor SDZ-224015 [79] react covalently/reversibly with M^{pro}. **I.** A non-peptidomimetic inhibitor covalently reacts via aldehyde acylation moiety to M^{pro} [85].

1.3.3.1 Inhibitor substituents at P4-P1′ positions to target M^{pro} S4-S1′ subsites

Nirmatrelvir binds at the S4-S1 subsites at the active site of M^{pro} (Figure 1.12B); the S4-S1′ subsites are targeted by earlier reported inhibitors, including N3 (Figure 1.8D) [28,37]. The S1′ subsite has a preference for aromatic or small hydrophobic amino acids. Substitution at the P1′ position with an aromatic group in α-ketoamides (13b) or abenzothiazol-2-yl ketones (3) did not impact the inhibitory potential [29,37,38]. Pyridine and cyclopropyl groups have been observed to bind at the S1′ subsite of M^{pro}, e.g. in the case of the calpain inhibitor XII [86] and 13a [37]. A structure of GC376 in complex with M^{pro} has demonstrated that high potency could be achieved without targeting the S1′ subsite [78]. Subsequent studies with compounds developed by Pfizer and others have optimised inhibitors only targeting the non-prime subsites.

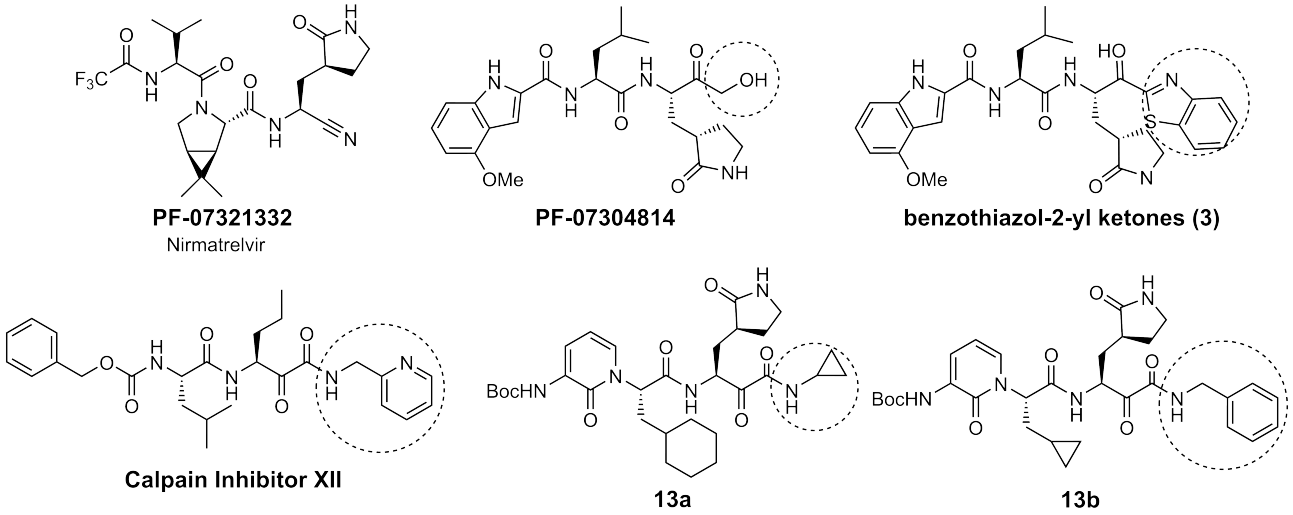


Figure 1.14 | P1′ equivalent groups in reported M^{pro} inhibitors. Structures of PF-07321332 (nirmatrelvir) [38], PF-07304814 [38], benzothiazol-2-yl ketones (3) [38], calpain inhibitor XII [86], 13a [29] and 13b [29] (α-ketoamide) with the P1′ equivalent groups circled. Note that the oxyanion hole is formed by the backbone amide groups of Gly143, Ser144, and Cys145 [86].

P1 equivalent groups in M^{pro} inhibitors

P1 equivalent groups of M^{pro} inhibitors have commonly included γ-lactam derivatives that mimic the substrate Gln (Figure 1.13), and which reduce the number of hydrogen bond donors at that position, as an increased number of hydrogen bonds has been correlated to poor bio-availability of drug candidates [87]. Furthermore, a 10-fold increase in potency was reported with the cyclised

γ-lactam moiety compared to Gln [88], possibly by reducing the loss of entropy upon binding at the S1 subsite [89,90]. Alternatively, boceprevir, an M^{pro} inhibitor discovered via drug repurposing efforts, has a cyclobutane moiety in the P1 position and displays micromolar (μM) inhibition against M^{pro} [77]. The carbonyl and NH groups of the γ-lactam moiety form a hydrogen bond with the Nϵ atom of His163 and polar interactions with Glu166 and/or backbone carbonyl of Phe140, respectively. The calpain inhibitor II, unusually, has a hydrophobic Met residue present in the P1 equivalent position. Hydrogen bond interactions at the S1 subsite are also absent with the cyclobutyl P1 equivalent moiety of boceprevir. Interestingly, the indazole moiety present in CDD-1713, a non-peptidomimetic inhibitor of M^{pro}, forms a hydrogen bond with His163 and the backbone carbonyl of Phe140 [85]. Thus, both hydrophilic and hydrophobic substituents can be tolerated at the S1 subsite.

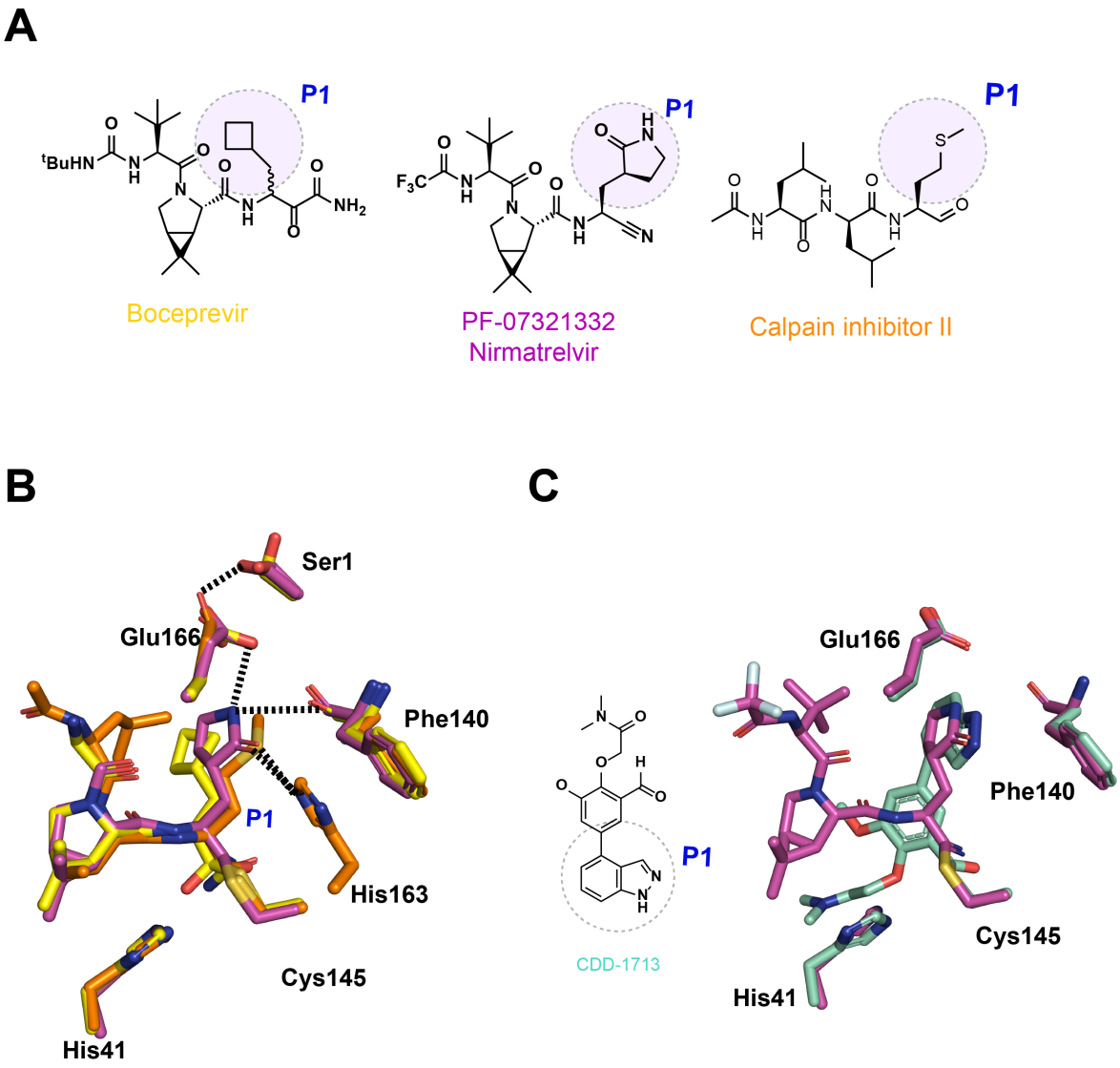


Figure 1.15 | Interaction of P1 equivalent groups at the M^{pro} active site. **A.** Structures of boceprevir, nirmatrelvir and calpain inhibitor II. The P1 equivalent group is highlighted. **B.** Views from the superimposition of the crystal structures of M^{pro} bound to boceprevir ((PDB: 6ZRU, yellow), nirmatrelvir (PDB: 7VH8, purple) and calpain inhibitor II (PDB: 6XA4, orange). **C.** The indazole moiety in CDD-1713 (PDB: 7LTN, cyan) is poised to form a hydrogen bond with the side chain of His163 and the main chain carbonyl of Phe140, in a manner identical to the pyrrolidone ring in nirmatrelvir.

P2 equivalent groups in M^{pro} inhibitors

His41, Tyr54, Met49, Gln189 and Met165 of M^{pro} form the hydrophobic S2 subsite. Leu, the most prevalent P2 residue in M^{pro} substrates, has also been exploited for binding at the S2 pocket in inhibitor design. A bicycloproline group, which is a leucine bioisostere (Figure 1.16), is present in boceprevir, and is a common motif in Pfizer M^{pro} inhibitors, including nirmatrelvir where it fits

tightly in the S2 subsite [38,39]. MI-09, also a potent M^{pro} inhibitor [83], contains a bicycloproline derivative (Figure 1.16) that binds in the S2 subsite.

Zhang *et al.* [37] have conducted extensive structure-activity relationship studies on the P2 equivalent substituent with SARS-CoV-1 M^{pro} inhibitors (Figure 1.16). Their results suggest that binding in the S2 subsite is critical in potent peptidometic M^{pro} inhibitors. Leu (**8**), ethylcyclopropaneleucine (**2**) and ethylcyclohexane S2 binding groups (**5**) were the most potent substitutions, whereas Phe (**6**) and the ethylcyclopropane (**1**), cyclobutane (**3**) and cyclopentane (**4**) containing P2 equivalent derivatives were slightly less efficient inhibitors (IC_{50} 1-2 μM) [37]. Alkyl substituted derivatives were less active e.g. those containing pentyne (**9**) or but-3-yne (**10**) (IC_{50} 8-11 μM) group. By contrast, removal of the side chain at the P2 equivalent position or addition of a fluorobenzyl group (**7**) caused a complete loss of inhibitory activity, indicating the selectivity of P2 equivalent sidechains for potent M^{pro} inhibition.

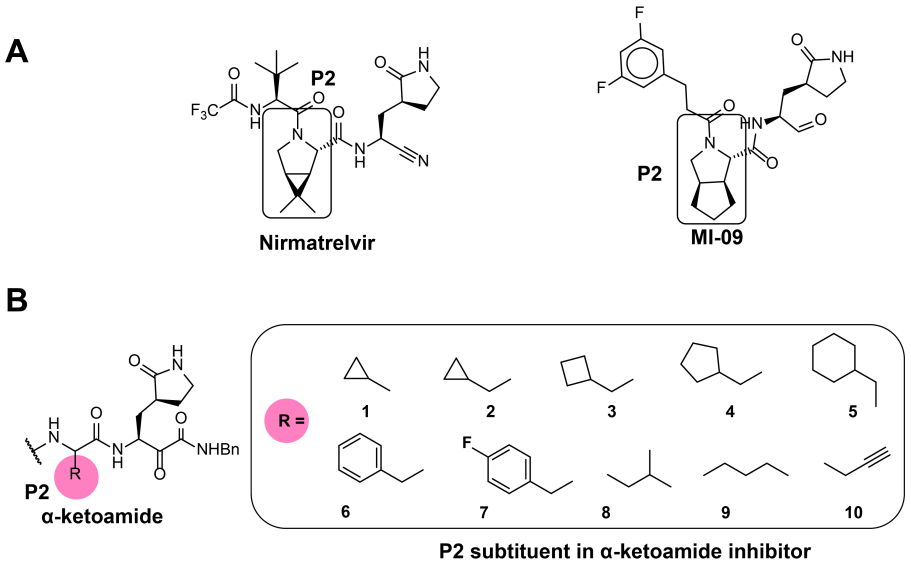


Figure 1.16 | P2 equivalent substituent of known M^{pro} inhibitors. **A.** Structures highlighting the P2 equivalent group in nirmatrelvir and MI-09. **B.** Zhang *et al.* [37] observed high inhibitory effects of alphaketoamides with Leu or ethylcyclopropane, ethylcyclohexane at the P2 equivalent position. Substitution with phenylalanine or cyclopentane and alkyl derivatives also retained the inhibitory activity. Decreasing the aromaticity of the benzyl ring with a fluoro group (**7**) also depleted the activity of inhibitors. Bicycloproline derivatives at the P2 equivalent position of the potent M^{pro} inhibitors, nirmatrelvir, and MI-09, are highlighted.

P3 and P4 equivalent groups in M^{pro} inhibitors

In an M^{pro}:N3 complex crystal structure, the side chain of the valine residue at the P3 equivalent

position of N3 was observed to be solvent exposed [28]. A *tert*-butyl group is also tolerated at P3, as is the case with boceprevir [77] and nirmatrelvir [38] (Figure 1.17). The main chain NH and carbonyl groups of the P3 equivalent group are positioned to form hydrogen bonds with the main chain carbonyl and NH groups of M^{pro} Glu166, respectively. The side chain of Glu166 forms a hydrogen bond with the S1 binding group. Inhibitors have been designed by conserving at least one hydrogen bond of the S3 binding group with the main chain Glu166 [37]. In the case of dipeptide derivatives, GC376 and MI-09 [83], only one hydrogen bond is apparently formed with the backbone of Glu166 (Figure 1.17); instead, the aromatic groups at the P3 equivalent position extend into the S4 subsite. Met165, Leu167, Phe185, Gln192 flank the small hydrophobic S4 subsite which can accommodate aromatic groups e.g. the methyoxy-indole of PF-00835231 [38], or small side chains e.g. the *tert*-butyl group of boceprevir [77] or the trifluoro-group of nirmatrelvir [38]. Cyclization of the P3-P2 junction with pyridone in 13b has also been reported to retain activity and to contribute to better cellular efficacy due to its reduced hydrogen bond donor capability [29].

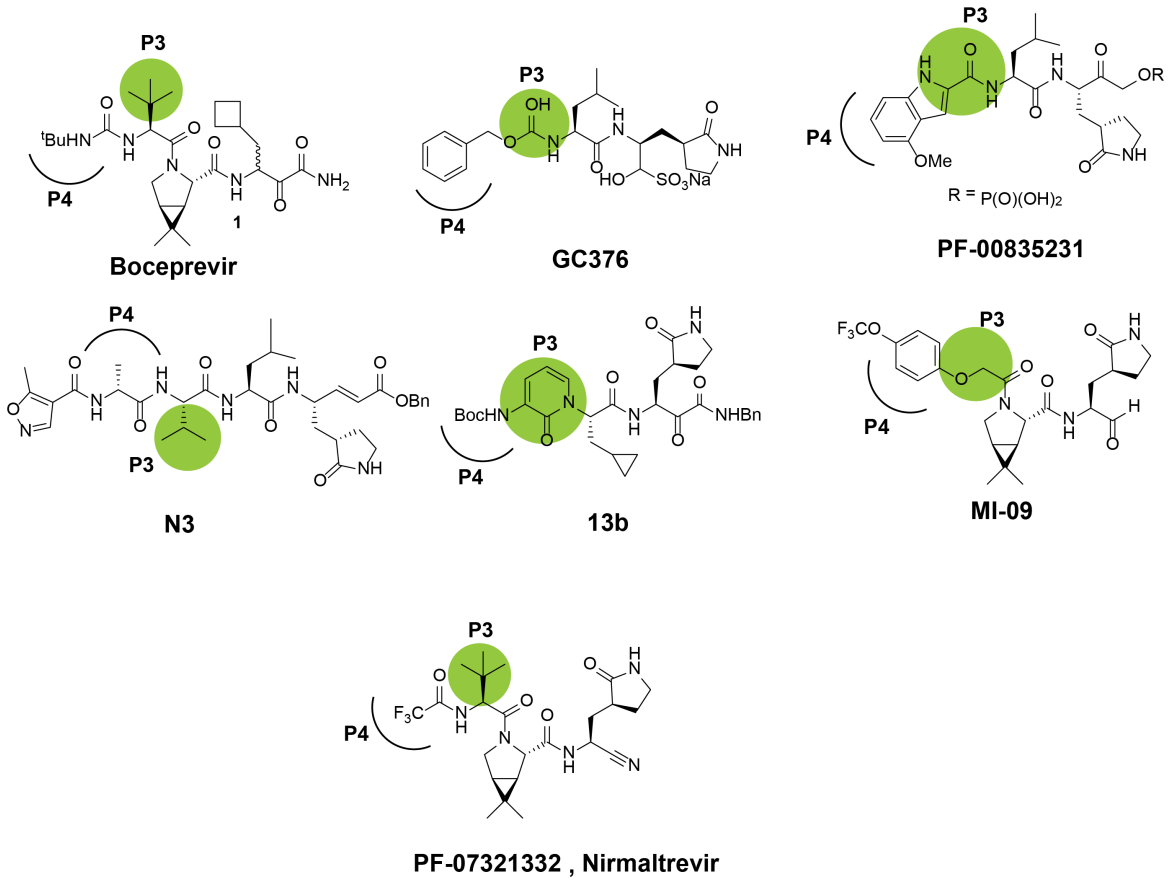


Figure 1.17 | The P3 equivalent substituent (highlighted in green) in various reported [28,29,38,77,78,83] covalently reacting M^{pro} inhibitors.

1.3.4 Peptide inhibitors of clinical interest

Peptides have *in vivo* functions in cell signalling and protein localisation as well as roles as neurotransmitters, hormones, and regulators of enzyme activity [91,92]. However, the reduced stability and cell permeability of linear peptides often limit their use in pharmaceutical applications. To address these issues, the cyclisation of peptides to generate macrocyclic peptides is one promising strategy.

1.3.4.1 Natural macrocyclic peptide inhibitors

Peptides cyclised by amide bonds are common in natural products which have biological functions, e.g. immuno-suppression, antibacterial, antiviral, and antifungal properties [93,94]. Peptide cyclisation can occur via peptide bond formation between the N- and C-terminal residues, which flank a polypeptide composed of canonical and non-canonical residues, or via peptide bond formation between side chains or between side chain and terminal residues (Figure 1.18). Cyclisation renders

peptides more resistant to exopeptidases [93,95] and may increase their cell permeability by reducing the number of available hydrogen bond donors and hydrogen bond acceptors due to intrapeptide interactions in a macrocyclic conformation.

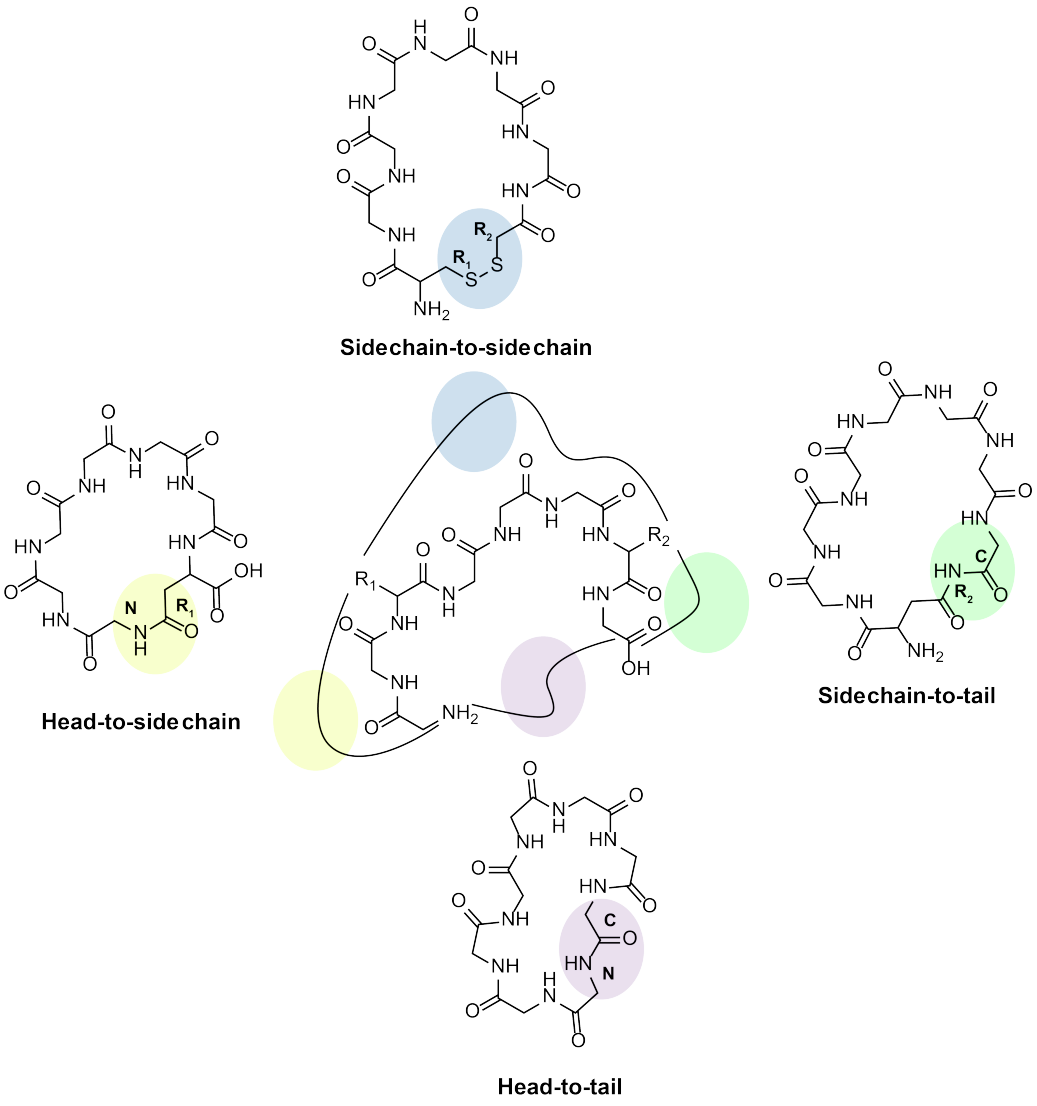


Figure 1.18 | Some possible ways to constrain a linear peptide into a macrocyle.

Like linear peptides, cyclic peptides often possess low toxicity, in part because their hydrolysis products are simple amino acid residues. Therefore, improved cell permeability, stability and good target affinity associated with cyclic peptides may make them more useful for therapeutic applications. In fact, 40 cyclic peptides are currently approved for clinical use, of which 18 were approved within the last two decades [96,97].

Most clinically approved cyclic peptides occur naturally, and are derived from plants [98,99],

bacteria [100], fungi, mammals [101], sponges and algae [97]. They possess therapeutic properties, including for use in antibacterial, antiviral, antifungal, and anticancer treatments, and can also be harnessed for biochemical applications e.g. RNA regulation and bioimaging tools. Examples of selected antibacterial, antifungal and antiviral cyclic peptides approved for clinical use or currently in clinical trials are provided in the section below.

Antibacterial, antifungal, antitumour and antiviral cyclic peptides

Cyclic peptides have found important applications as immunosuppressants [102]. Cyclosporin A (Figure 1.19) was isolated from the fungus *Beauveria nivea*, and its approval for clinical use as immunosuppressant in 1983, has made organ transplantation safer [103]. Gramicidin S, an antibacterial peptide, is one of the oldest cyclic peptide-based drugs on the market [93]. Gramicidin S is a cationic cyclic decapeptide (Figure 1.19) with antimicrobial activity against Gram positive and Gram negative bacteria and fungal targets [93]. However, hemolytic issues reduced its wider application. Gramicidin S forms a pore at the cell membrane, causing cell death by triggering membrane lysis via potassium ion efflux [104]. Similarly, vancomycin and Dipi-van are cyclic glycopeptidic antibiotics (Figure 1.19) effective against methicillin-resistant *Staphylococcus aureus* (MRSA) bacteremia [105]. Cyclic griselimycin and cyclohexylgriselimycin (Figure 1.19) display anti-tuberculosis activities [106, 107]. Simple modifications of vancomycin and griselimycin enabled them to overcome resistance and improve their potency, respectively. This sets a precedent for other approved macrocyclic peptide inhibitors to be potentially optimised to overcome future resistance without undertaking major redesign of the macrocyclic peptide structure.

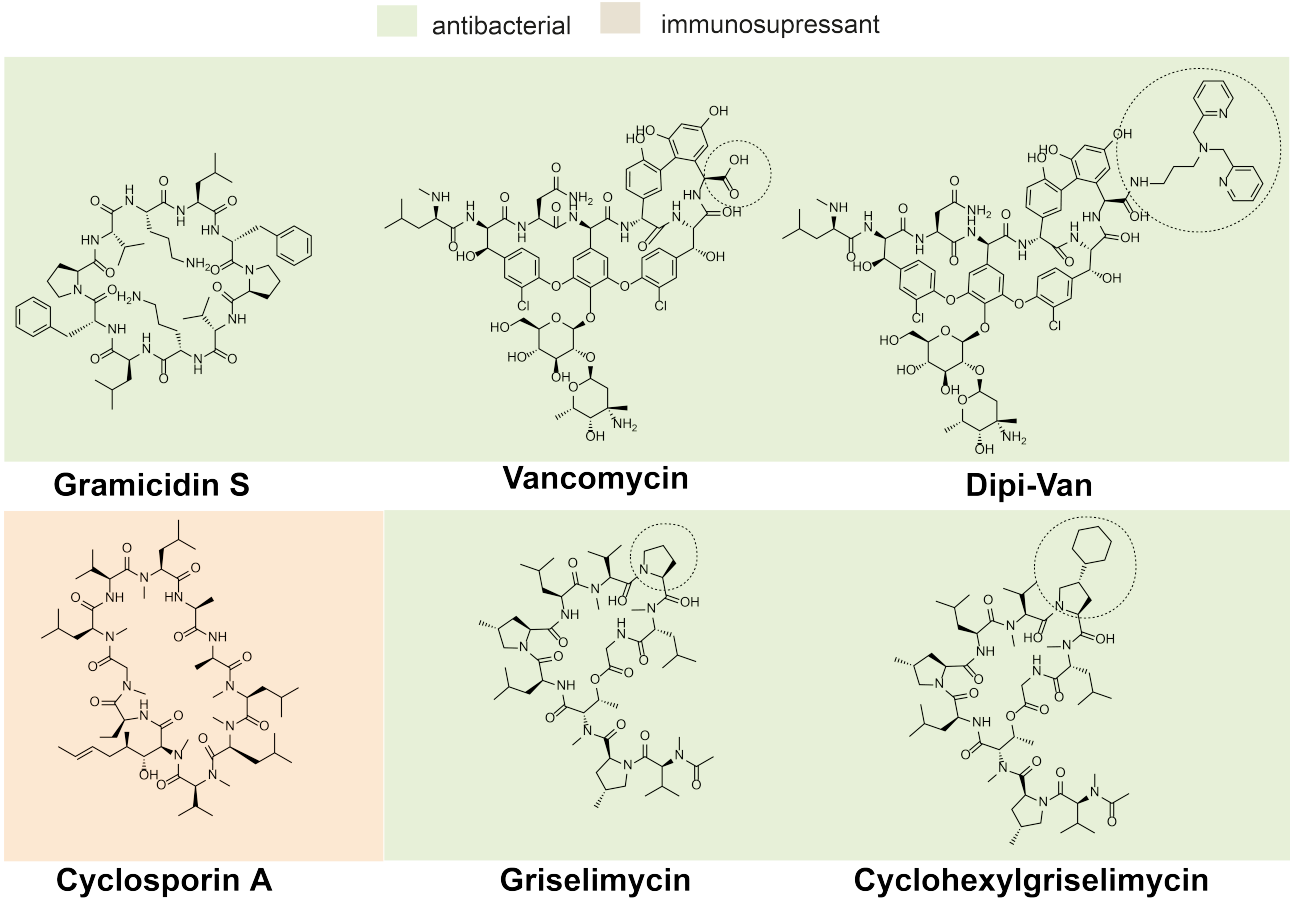


Figure 1.19 | Macrocyles of clinical interest with antibacterial and immunosuppresive properties. Modified groups for improved activity are circled.

Macrocycles with antifungal activities, e.g. amphotericin B, and semi-synthetic caspofunghin (Figure 1.20) are used to treat life-threatening fungal infections. Other natural cyclic peptides, such as ASP2397 which harbour activities against pathogenic fungus, have been recently isolated [97]. Macrocycles such as Aspergillipeptide A, D, and simplicilliumtide K (Figure 1.20) exhibit inhibitory activities against herpes simplex virus-1 [108–110]. Likewise, N-methylated octadecapeptides gymnopeptide A and B, isolated from mushroom species, demonstrate biological activity in anti-tumour studies [111]. Cyclic peptides are therefore a validated class of therapeutics for the treatment of a plethora of pathological conditions.

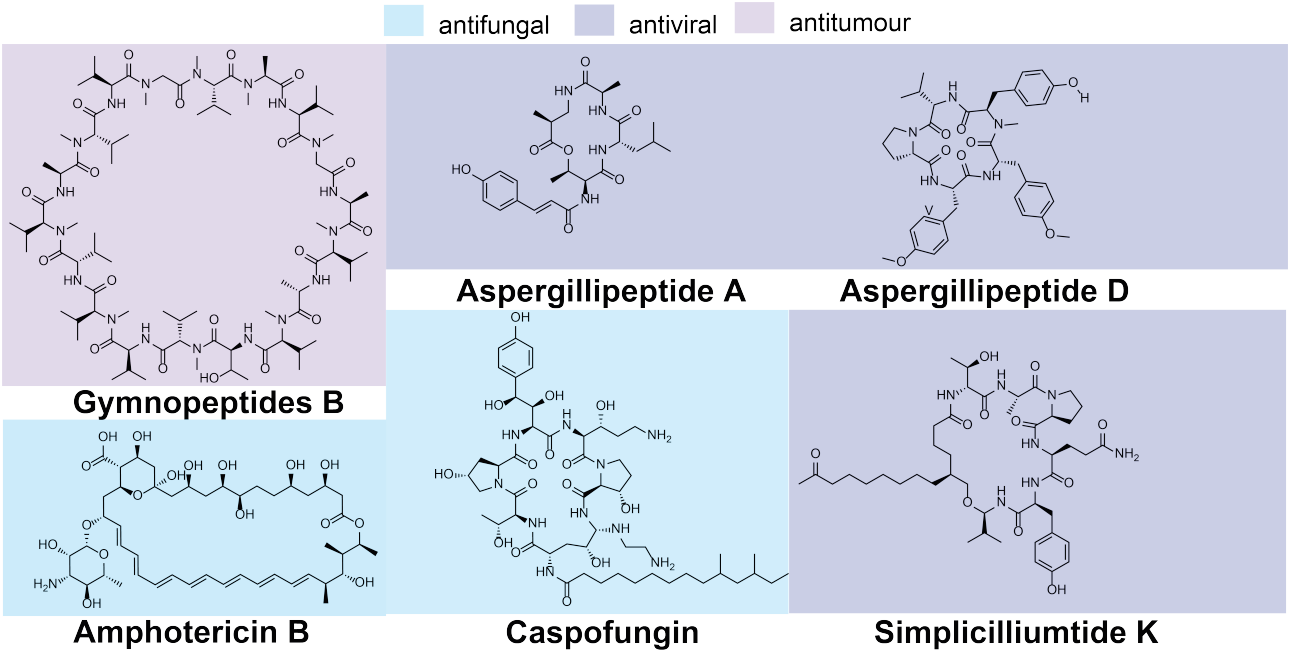


Figure 1.20 | Natural and synthetic derivatives of macrocycles with antiviral, antiproliferative and antifungal activities.

Natural disulfide rich peptides

Aside from the backbone cyclised macrocycles, cyclic disulfide rich peptide scaffolds are commonly present in plants and animals, and often play roles in host defense mechanisms [112,113]. For example, the backbone cyclised highly selective sunflower-derived trypsin inhibitor (STF1) (K_i = 0.1 nM) [114], consists of 12-14 amino acid residues with a single intrapeptide disulfide bond (Figure 1.21) [114,115]. In addition to trypsin, STF1 also inhibits other serine proteases such as chymotrypsin and cathepsin G [116]. Disulfide rich cyclic peptides (DRPs) identified from animal sources include the 18 amino acid long rhesus θ-defensin-1 (RTD-1) and conotoxins. RTD-1 has three disulfide bonds with potential applications for anti-tumour, and anti-cancer therapies and also in the treatment of autoimmune diseases [115].

Unlike RTD-1, the disulfide rich conotoxin peptides are linear and identified from venom of marine cone snail [117,118]. α-Conotoxins, a subgroup of conotoxins, inhibit nicotinic acetylcholine receptors [119] e.g. α-conotoxin identified from venom of *Conus victoriae*, Vc1.1 [120]. Vc1.1 is a 16 residue, α-helical peptide connected by two disulfide links and loops. Currently, Vc1.1 is being investigated as a drug for neuropathic pain treatment [121] due to promising results exhibiting potent

analgesic effects in rat models [122]. Cyclisation of N- and C-termini residues led to the development of the orally active synthetic cyclic peptide Vc1.1 (cVc1.1) [121]. The three dimensional structure and activity of linear Vc1.1 is conserved in cVc1.1 [121, 123]. Similarly, cyclisation and chemical modification of another disulfide rich peptide, chlorotoxin [46 amino acid residues (aa) long] with four disulphide bonds has enabled its use in tumour-imaging and targeted tumour cell therapy [124, 125].

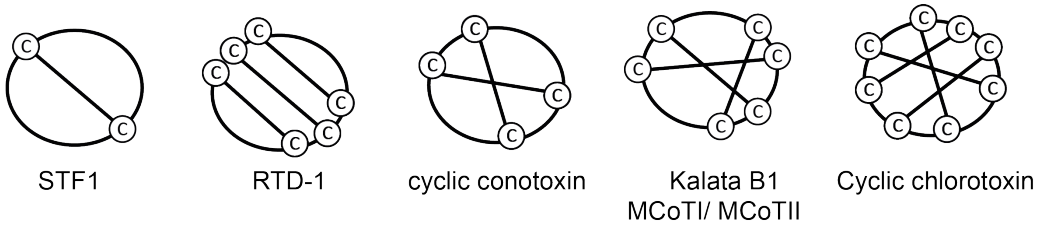


Figure 1.21 | Disulfide rich peptide topologies. The number of disulfide bonds are shown. In the scheme, only synthetically cyclised peptides are prefixed with cyclic. Cysteine scaffolds are conserved and only the backbone loop regions are modified for molecular grafting for target adaptation.

Other notable natural disulfide rich peptides include the plant-derived cyclotides which possess three conserved disulfide bonds in a cyclic cysteine knot motif and a head-to-tail backbone cyclisation [126]. Cyclotides typically compromise 27-30 amino acid acids, and are less prone to proteolytic cleavage than the shorter DRPs with lower numbers of disulfide cross-links; they are highly stable at high temperature, low pH, and in the presence of chaotropic agents (all of which would be detrimental to linear peptides), and manifest high serum stability [127]. Moreover, some cyclotides like Kalata B1 are cell permeable, comparable to protein derived cell penetrating peptide, TAT protein [128]. However, their proteolytic stability depends both on the macrocyclic and disulfide bond connectivity. Examples of cyclotides include MCoTI-I and II [129] which contain three disulfide bonds are cell permeable and have been engineered to target proteases of clinical significance [129–131].

1.3.4.2 Library based screens to identify macrocyclic inhibitors

Molecular grafting has enabled the application of natural cyclic peptides across different targets in the treatment of tumour, cancer, cardiovascular, autoimmune, and other diseases [132–137]. Selective linear peptides are grafted onto the loops of disulfide scaffolds of natural peptides to generate a stable cyclic disulfide rich peptide. Loop regions in disulfide rich peptides have high sequence variability and thus

are more likely to be tolerant of substitutions. To achieve this, two strategies have been implemented: 1) mutagenesis of the loop region, and 2) application of approaches such as phage display, synthetic peptide library, bacterial display, and yeast display [138]. Library based grafting methods are preferred as these methods can rapidly identify *de novo* tight binders of the target protein.

Phage display enables production of large peptide libraries, ($\sim 10^{10}$ molecules) [139] compared to chemical peptide libraries ($\sim 10^{7}$ molecules). Phage display is performed by displaying semi-randomised peptide sequences on the bacteriophage capsid, such that each virus particle has a cognate genetic material to its surface peptide [140,141]. The libraries can be screened for binding against a target enzyme, and the genetic sequences associated with the tightly bound peptides can be determined by PCR followed by DNA sequencing. The introduction of engineered tRNAs can allow for the incorporation of non-canonical amino acids to increase the diversity of the library.

Libraries of cyclic peptides can also be made via cell free translation employing mRNA technology; this enables diverse library production ($\sim 10^{12}$ molecules) [139]. Application of the Flexible In vitro Translation (FIT) method which is enabled by flexizymes that catalyze tRNAs acylation with non-canonical amino acids, coupled with mRNA display can be used to incorporate non-canonical amino acids into the cyclic peptide library [142]. When combined with mRNA display, the FIT method of cell-free translation is termed RaPID (Random nonstandard Peptide Integrated Discovery) [143]. Incorporation of non-canonical amino acids in the RaPID method is desired as it may contribute to improved proteolytic stability, bioavailability, cell permeability [144] and in some cases, the affinity of the macrocycle to the target enzyme. RaPID selection has been applied to identify disulfide rich peptide inhibitors of a human serine protease [145]. A RaPID library based approach was applied to loop regions in Kalata and McoTI-II scaffolds to identify DRP inhibitors of M^{pro} (Figure 1.22) as described in Chapter 5.

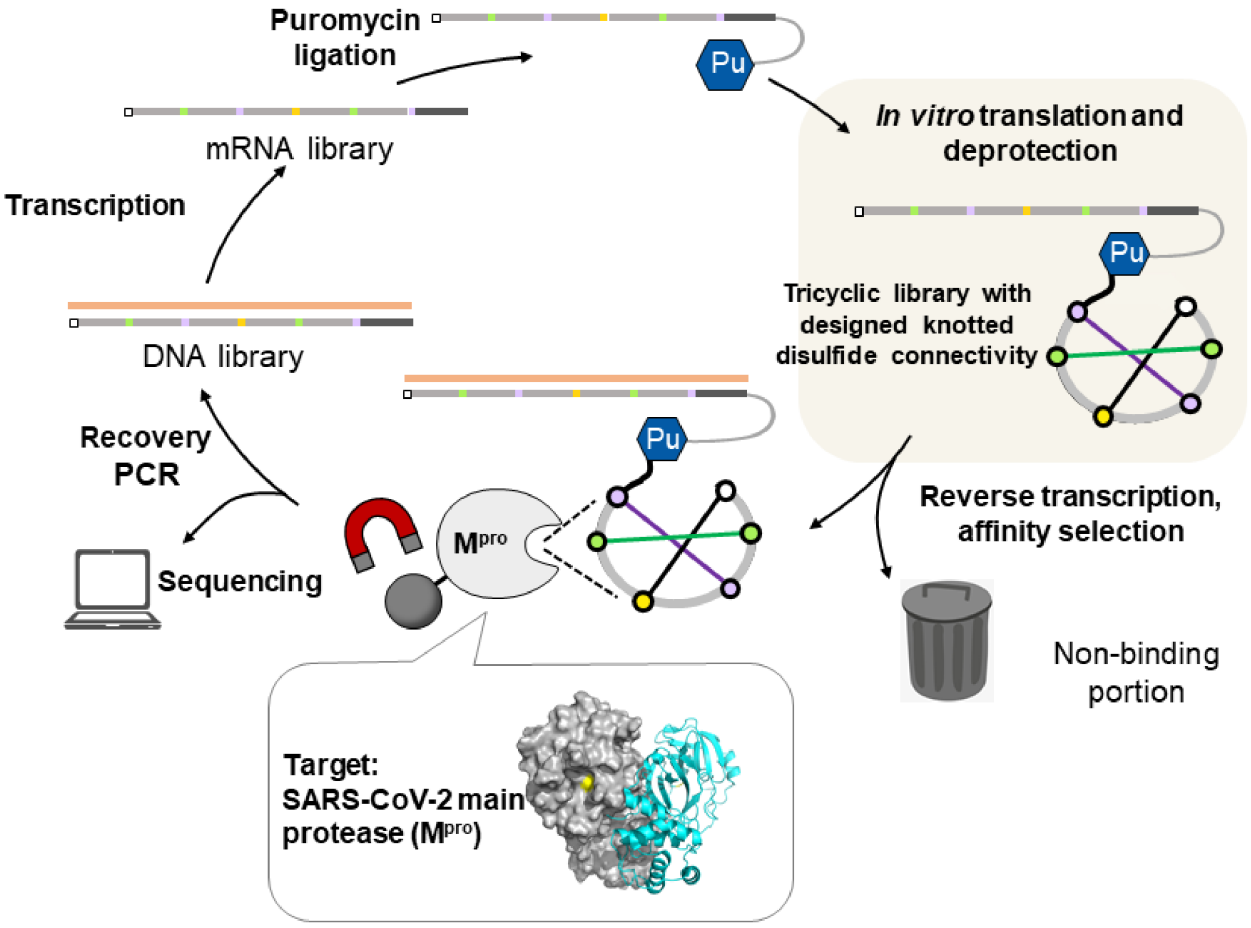


Figure 1.22 | Overview of the RaPID selection using tricyclic DRP library. In vitro translation of Puromycin libated mRNA library via FIT system and the stepwise deprotection of translated macrocyclic peptide leads to the generation of DRP macrocyclic library. Affinity selection of the DRP library is performed with M^{pro} after reverse transcription of its cognate mRNA. Non-binding macrocyclic peptides are washed away and the mRNA of the binding macrocycle is recovered by PCR amplification. A DNA library is generated and transcribed to form an mRNA library. This is performed in an iterative manner. The figure was adapted from an image produced by Dr. Wenyu Liu (University of Tokyo).

1.3.4.3 Macrocyclic peptide inhibitors of M^{pro}

In parallel with the work described in Chapter 5, Johansen-Leete *et al.* [146] have recently applied the RaPID technology to identify cyclic peptides which are potent inhibitors of M^{pro} *in vitro* (Figure 1.23); these peptides were cyclised through thioether formation. However, the lead peptide with Leu at the P2 equivalent position and Gln at P1 equivalent position did not manifest inhibitory activity in cells. Incorporating the cell-penetrating peptide penetratin at the C-terminus of the lead peptide resulted in improved cellular activity [146].

Macrocyclic peptides have also been designed based on the C-terminal auto-cleavage site of M^{pro} [147], e.g. UCI-1 (Figure 1.23). However, these peptides did not display potent M^{pro} inhibition. Al-Wahaibi *et al.* have identified six membered oxazole-based macrocyclic peptide mimetics which display moderate inhibition (2.6 - 15.3 µM) [148].

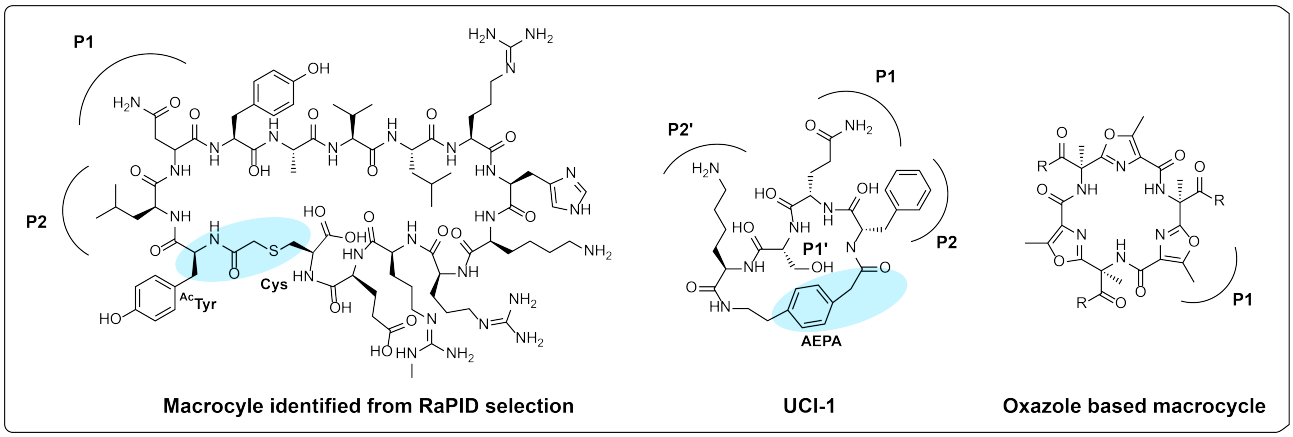


Figure 1.23 | Reported macrocyclic peptide inhibitors of M^{pro}.

1.4 L,D-transpeptidases

Many proteins with similar folds and mechanisms resembling proteases are major therapeutic targets in other disease contexts. Bacterial penicillin-binding proteins (PBPs) and L,D-transpeptidases (Ldts), transpeptidases involved in cell wall biosynthesis, are antibacterial targets. The Ldts utilise a Cys-His catalytic dyad similar to M^{pro}, whereas the PBPs utilise a serine nucleophile as part of a catalytic triad, as present in some serine proteases. Furthermore, the catalytic mechanisms of the PBPs and Ldts are similar to those of the cysteine and serine proteases, and both involve the formation of an acyl-enzyme intermediate. However, this complex is deacylated by an amine nucleophile during PBP and Ldt catalysis, forming a new amide bond and liberating the nucleophilic residue; this notably differs from the use of an activated water molecule as a nucleophile by proteases. Whilst PBPs have long been targeted in drug discovery and antibacterial therapy, the Ldts are relatively recently explored targets of potential importance for the treatment of tuberculosis [149, 150].

As stated by the WHO, tuberculosis (TB) is the infectious disease associated with the highest mortality rates in adult populations after COVID-19 [151]. The current drug regimen for tuberculosis

treatment spans 6 months and is ineffective against some drug resistant mycobacteria, therefore increasing the urgency for the development of novel antimycobacterial therapeutics.

1.4.1 Bacterial peptidoglycan layer

The cell wall is an integral component of fungi, plants, protists, algae and bacteria [152], and provides mechanistic and structural support, allowing these cells to withstand high osmotic pressure [153]. The cell wall can also act as a semi-selective permeability barrier that regulates the diffusion of ions and molecules in and out of cells. Loss of cell wall integrity is detrimental to cells, and generally leads to cell death via lysis [154, 155]. This has been exploited for many years through the use of β-lactam antibiotics, which target essential enzymes such as PBPs that maintain the bacterial cell wall.

A major component of the bacterial cell wall is the peptidoglycan layer, which is a complex macromolecule present in most eubacteria [156–158]. Peptidoglycan is composed of a repeating disaccharide backbone which is covalently linked via polypeptide chains [157]. A thick peptidoglycan layer enriched with teichoic acids and lipoteichoic acids is present in Gram positive bacteria, which are generally commensal in nature (Figure 1.24) [159]. By contrast, an additional membrane surrounds the thin peptidoglycan layer in Gram negative bacteria, as found in many pathogens (e.g. *E. coli*, *Salmonella* spp., *Vibrio cholerae*, and *Pseudomonas aeruginosa*) [160]. The plasma membrane, which is embedded with proteins that regulate peptidoglycan maturation or lipid anchoring, forms a barrier between the periplasmic space and the cytosolic space in all bacteria. The outermost membrane, which consists of lipopolysaccharides and porins, renders Gram negative bacteria less susceptible compared to Gram positive bacteria to many antibiotics [161].

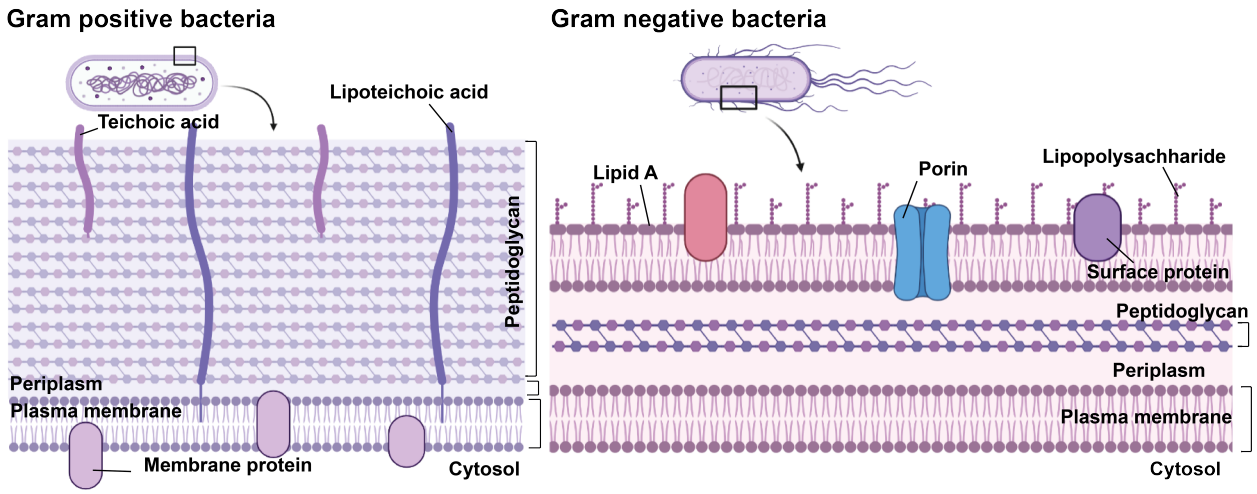


Figure 1.24 | The compositions of different bacterial cell walls. Gram positive bacteria possess a relatively thick peptidoglycan layer compared to Gram negative bacteria, and the Gram positive cell wall is modified with teichoic and lipoteichoic acids. The outer membrane in Gram negative bacteria is embedded with porins, surface proteins, and lipopolysaccharides.

1.4.2 Mycobacterial peptidoglycan layer

The causative agent of tuberculosis, *Mycobacterium tuberculosis*, and other *Mycobacterium* spp., have cell walls that are distinct from most other prokaryotes (Figure 1.24). Their relatively thick peptidoglycan layer has led to the classification of mycobacteria as Gram positive bacteria. However, mycobacteria share notable features with Gram negative bacteria, including a distinguishable pseudo outer membrane (Figure 1.25) [158]. Mycobacteria possess a peptidoglycan (PG) layer which is covalently linked with arabinogalactans, giving the mycobacterial cell wall a distinct appearance compared to the cell walls of Gram negative bacteria [158, 162]. Alternating β(1→4) linked *N*-acetylglucosamine (GlcNAc) and *N*-acetyl/*N*-glycolyl muramic acid derivatives form repeating heterosaccharides that make up the backbone of the mycobacterial peptidoglycan [163]. The incorporation of *N*-glycolyl muramic acid (MurNGlyc), obtained via hydroxylation of *N*-acetylmuramic acid [164], is uncommon in bacterial PG. This *N*-glycolyl-derivative is thought to form additional hydrogen bonds and increase the strength of the PG in the mycobacterial cell wall [165].

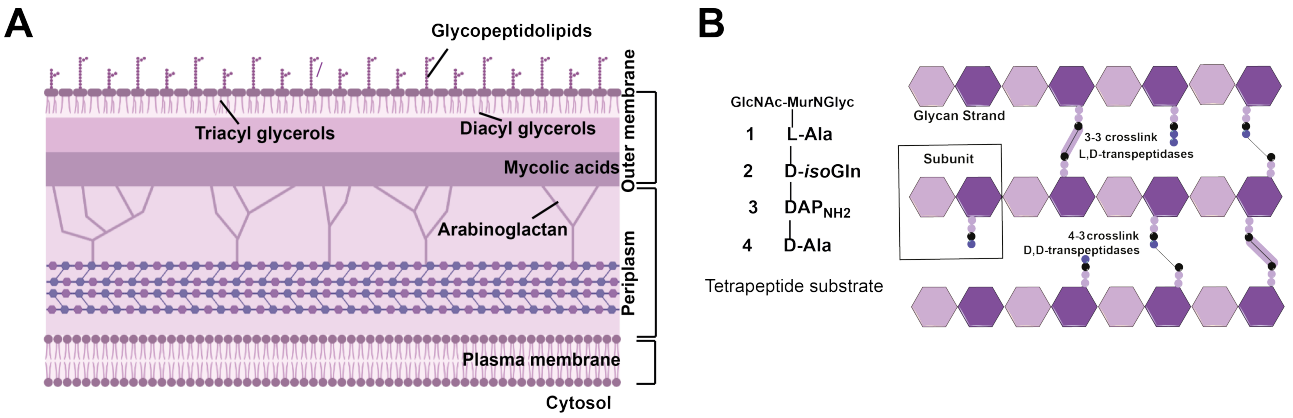


Figure 1.25 | Mycobacterial cell wall and the cross-linking of peptidoglycan tetrapeptide substrates by Ldts. A. Mycobacteria share features of both Gram positive bacteria (e.g., a thick PG layer) and Gram negative bacteria (e.g., the presence of an outer membrane). **B.** The tetrapeptide chains in PG are cross-linked by the transpeptidation activity of the Ldts.

The notable feature of the mycobacterial peptidoglycan layer is the composition and pattern of the covalent cross-links that are formed between the peptide chains in adjacent PG strands. A tetrapeptide stem composed of L-Ala-D-γGlu-*meso*-diaminopimelate(Dap)-D-Ala is linked to the MurNGlyc sugars of the polysaccharide backbone in mycobacterial PG [166, 167]. The formation of a 3→3 tetrapeptide crosslink from the tetrapeptide substrate is catalysed by the activity of L,D-transpeptidase in mycobacteria. The peptide chains in the PG of most bacteria are predominantly cross-linked through 4→3 linkages, contributing to the mesh-like appearance of PG layer [158, 168].

Penicillin-binding proteins (PBPs) catalyse the formation of these covalent 4→3 cross-links. This process begins with the reaction of the PBP serine nucleophile with a pentapeptide stem, releasing the terminal D-Ala residue [169]. The ester carbonyl of the resulting acyl-enzyme complex is then attacked by the amino group of Lys/*meso*-Dap, the third residue of a nearby peptide chain, forming a 4→3 cross-link and concomitantly releasing the enzyme to catalyse further transpeptidation reactions (Figure 1.25).

By contrast, 3→3 cross-links represent about 80% of the total cross-links occurring in non-replicating dormant mycobacteria [170]. The amidation of D-Glu and *meso*-Dap in the second and third position of the tetrapeptide is unique and essential for mycobacterial PG cross-linking, compared to other bacteria e.g. *E. coli* and *Bacillus subtilis* [171]. The total level of cell wall cross-linking is also

greater in *Mycobacterium* spp., with 70-80% of peptidoglycan peptide chains cross-linked, as opposed to the 50% peptide cross-linking that occurs in the peptidoglycan layer of *E. coli* [162, 172].

1.4.3 Mycobacterial L,D-transpeptidases

In the *M. tuberculosis* H37Rv genome, five genes [Rv0116c (Ldt_{Mt1}), Rv2518c (Ldt_{Mt2}), Rv0192 (Ldt_{Mt3}), RV0483 (Ldt_{Mt4}) and Rv1433 (Ldt_{Mt5})] encode for L,D-transpeptidases [172, 173]. Functional assays with purified proteins and peptidoglycan precursors as substrates suggest that only Ldt_{Mt2} and Ldt_{Mt1} possess transpeptidase activity [172]. However, this observation has been challenged, and a different research group has reported 3→3 cross-linking activity for Ldt_{Mt1}, Ldt_{Mt2}, Ldt_{Mt4}, and Ldt_{Mt5} with PG fragments [174]. The different functional assignments of these L,D-transpeptidases could be due to the different composition of the precursor disaccharide-tetrapeptide substrates used: GlcNAc–MurNAc-L-Ala-D-γGlu-*meso*-Dap-D-Ala and GlcNAc–MurNGlyc-L-Ala-D-γGlu-*meso*-Dap-D-Ala isolated from *Corynebacterium jeikeium* strain CIP103337 [172] and *Mycobacterium abscessus* strain CIP104536T [174], respectively. Quantitative Real-time-PCR has been used to confirm that Ldt_{Mt2} expression levels are at least five-fold greater than the combined expression of the other four Ldt_{Mt} paralogs and that Ldt_{Mt2} expression is at least 10-fold greater than that of Ldt_{Mt1} during all phases of growth [172]. Gene knock down assays with Ldt_{Mt2} paralogs reveal that the mutant strains have the same phenotype as wild-type strains [172], suggesting they are not essential for mycobacterial growth. Microarray analyses, however, indicate that Ldt_{Mt1} along with Ldt_{Mt2} may have roles in the adaptive response of mycobacteria during the dormant stage, as their expression is upregulated during nutritional starvation [175]. *In vitro* studies and tests with TB mouse models indicate that loss of Ldt_{Mt2} results in a loss of virulence, altered cell wall morphology, and increased susceptibility to the antibiotic amoxicillin-clavulanate of mycobacteria [172]. Ldt_{Mt2} was also shown to be essential for infectivity in mouse models for TB [172]. The low expression levels of Ldt_{Mt5} and the other paralogs may render them unable to compensate for the loss of Ldt_{Mt2} activity [172, 174]. These results demonstrate the importance of Ldt_{Mt2}–catalysed formation of 3→3-cross-links in the peptidoglycan

layer of *M. tuberculosis*. Hence, inhibition of Ldt_{Mt2} to destabilise the peptidoglycan layer and cause cell lysis holds great potential for the treatment of TB.

1.4.3.1 Structure of Ldt_{Mt2} and other paralogs

The Ldt_{Mt2} fold is composed of two consecutive N-terminal β-barrel bacterial immunoglobulin-like (IgG-like) domains spanning from three-stranded (residues 60-148) and four-stranded (residues 149-250) antiparallel β-sheets (Figure 1.26A) [176,177]. The C-terminal $\alpha\beta$ mixed catalytic domain (residues 251-408) has a distinct ErfK/YbiS/YhnG fold with a conserved SHGC motif [172,173]. Transpeptidation catalysis involves a catalytic dyad composed of Cys354 and His336 (Figure 1.26B). The backbones of Cys, Ser, and Gly residues (from the SHGC motif) form an oxyanion hole [176,178]. The insertion of a β-hairpin loop (residues 300-323) and a C-terminal helical tail (residues 379-408) in Ldt_{Mt2} is not observed in the Ldt from *Bacillus subtilis*; this loop is disordered and shorter in the Ldt from *Enterococcus faecium* (Ldt_{fm}) [176,179] (Figure 1.27B). The presence of this β-hairpin loop in Ldt_{Mt2} is notable as it is absent in PBPs; this loop decreases the solvent exposed volume of the Ldt_{Mt2} active site [180] potentially giving it different substrate selectivity when compared to PBPs.

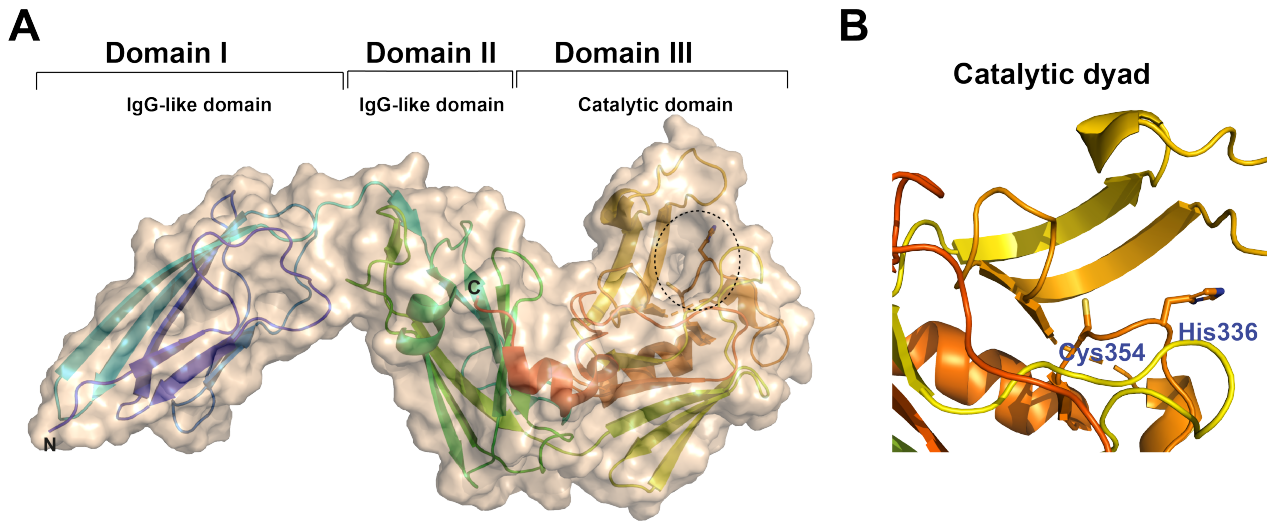


Figure 1.26 | Domain topology of Ldt_{Mt2}. **A.** The two IgG-like domains and the catalytic domain of Ldt_{Mt2} are shown [173]. N- and C-termini are indicated, and the catalytic residues are represented as sticks (PDB: 6RLG). **B.** Close-up view of Cys354 and His336, which form the catalytic dyad required for Ldt_{Mt2} catalysis.

Low overall sequence conservation (30%-50%) is observed [174,181] amongst the mycobacterial

L,D-transpeptidases. The structures of these Ldts are semicircular in shape and are comprised of IgG-like domains and an α,β-mix catalytic domain [173, 174, 181–183]. The domain topology of Ldt_{Mt5} is identical to Ldt_{Mt2} (31% sequence similarity) with two sequential IgG-like domains and a catalytic domain. However, the His residue of the conserved SHGC motif is substituted with an Asn in Ldt_{Mt5}, a difference with all other Ldt_{Mt}s [182, 184]. In the case of both Ldt_{Mt5} and Ldt_{Mt4}, the serine in the SHGC motif is substituted with a Thr. A crystal structure of Ldt_{Mt4} has not yet been solved; however, domain analysis indicates that it resembles Ldt_{Mt1} and Ldt_{Mt3}, which possess a single IgG-like domain and a catalytic domain [185]. The β-hairpin lid present near the active site of Ldt_{Mt2} is present in all other Ldts, and all of these enzymes employ a Cys-His dyad for catalysis. The length of the β-hairpin loop in the active site varies, and Ldt_{Mt5} and Ldt_{fm} possess smaller lids than $Ldt_{Mt1\text{-}4}$ (Figure 1.27B). The structural and functional diversity of the Ldts from different species and within mycobacterial Ldt_{Mt}s indicate that the Ldt_{Mt}s are likely to be functionally non-redundant in *M. tuberculosis*.

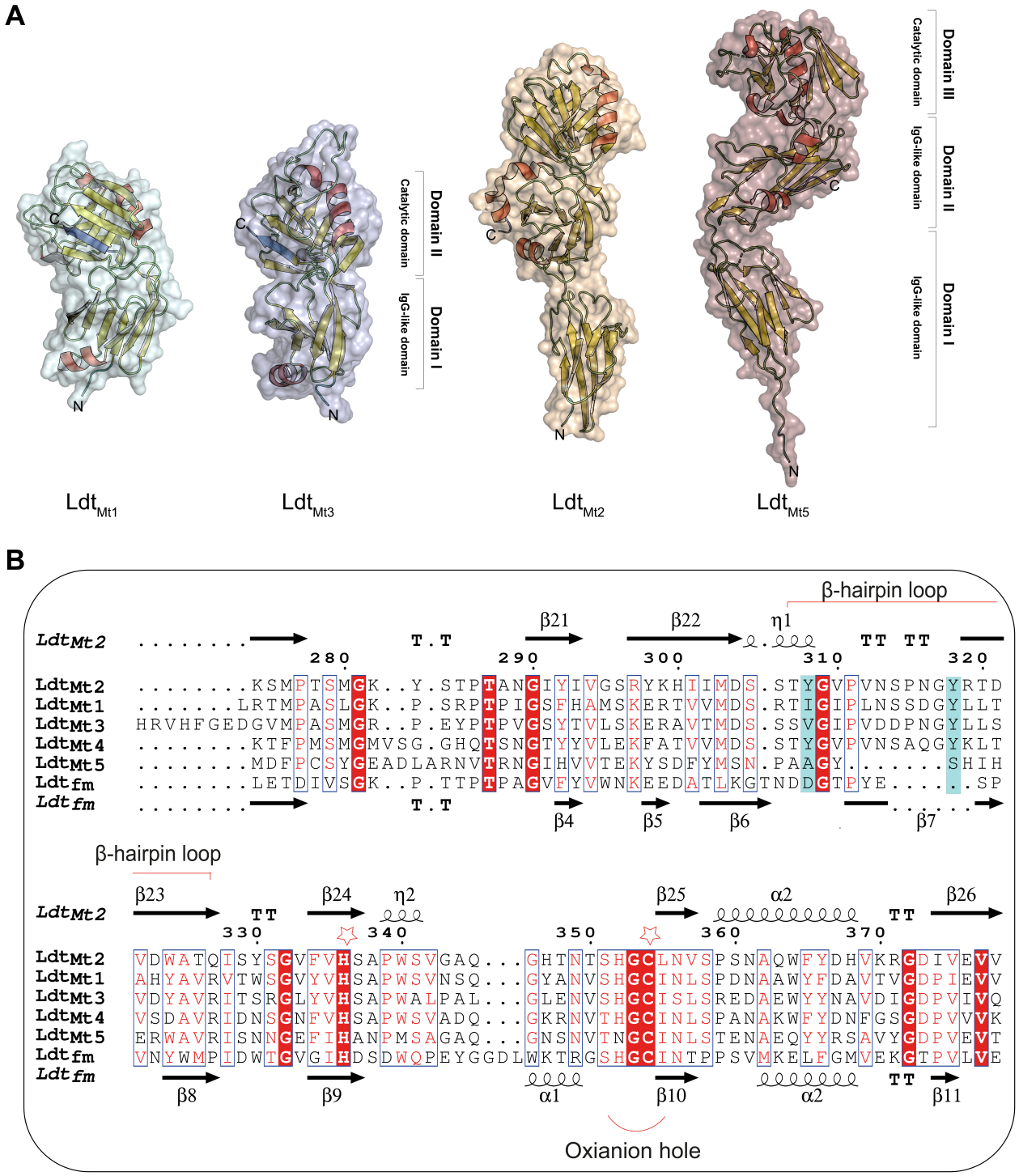


Figure 1.27 | Comparison of the domains of Ldts. A. Ldt$_{Mt1}$ (PDB: 4JMN) and Ldt$_{Mt3}$ (PDB: 6D4K) possess a single bacterial immunoglobulin-like domain and a catalytic domain. Ldt$_{Mt2}$ (PDB: 5DU7) and Ldt$_{Mt5}$ (PDB: 6D5A) possess sequential repeating immunoglobulin-like domains which precede the catalytic domain. The N- and C-termini are indicated. **B.** Sequence alignment indicates that the β-hairpin lid is nearly identical in all Ldt$_{Mt}$s apart from Ldt$_{Mt5}$, which has a deletion in the loop. Ldt$_{fm}$ also has a shorter β-hairpin lid at the active site. The mostly invariant SHGC motif which forms the oxyanion hole is labelled and the catalytic residues are represented with a star. The aromatic residues Tyr308 and Tyr318 present in the β-hairpin loop are highlighted in blue.

1.4.3.2 Functions of the L,D-transpeptidases

The L,D-transpeptidases catalyse the formation of 3→3 cross-links in mycobacterial peptidoglycan [186] using tetrapeptides (i.e., L-Ala-D-isoGlu-*meso*-DAP-D-Ala) as donor peptides (Figure 1.28). This tetrapeptide stem is linked to the *N*-acetyl muramic acid sugars in the PG glycan backbone. The transpeptidation reaction proceeds via the cleavage of the peptide bond between *meso*-DAP and D-Ala of the donor tetrapeptide, leading to the formation of an intermediate thioester acyl-enzyme complex and releasing D-Ala (Figure 1.28) [178, 182]. The amidated *meso*-DAP from an adjacent tetrapeptide stem attacks the thioester of the acyl-enzyme complex, resulting in the transfer of the donor peptide to the acceptor stem [186, 187] (Figure 1.28).

The Ldts belong to a distinct family from the PBPs, the classically studied cell wall transpeptidases. PBPs are bifunctional enzymes with carboxypeptidase/transpeptidase and transglycosylation activities, and generate tetrapeptide stems (i.e., the Ldt substrates) from pentapeptide stems using their D, D-carboxypeptidase activity [174, 188]. Both the Ldts and PBPs are crucial for the maintenance of the mycobacterial cell wall [188]. The transpeptidase function of the PBPs can be rendered redundant by the activity of mycobacterial Ldts, decreasing the susceptibility of mycobacteria to some β-lactam drugs. In *E. coli*, the abundance of 3→3 cross-links increases upon entering stationary phase from 1.8%-5% [189]. An increased abundance of 3→3 cross-links has also been reported in penicillin-resistant *E. faecium*, suggesting a potential adaptive response to β-lactam antibiotics involving the bypassing of PBP D,D-transpeptidation activity (4→3 cross-link formation) by upregulating Ldt production or activity [187].

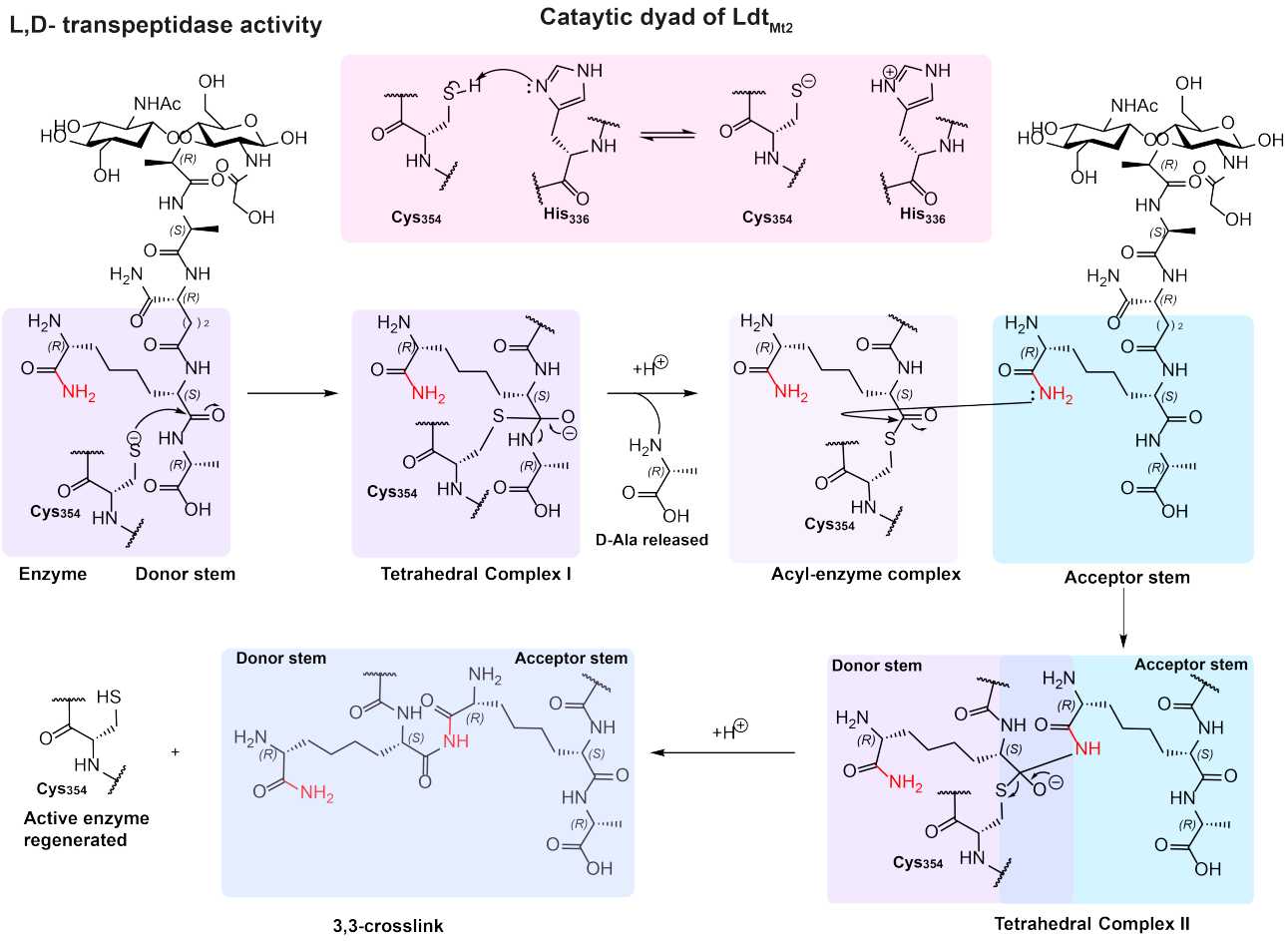


Figure 1.28 | Ldt$_{Mt2}$-catalysed transpeptidation reaction.

1.4.4 β-Lactams as potential drugs for *M. tuberculosis* therapy

β-Lactam antibiotics are covalent inhibitors of bacterial penicillin-binding proteins. The conformationally constrained four-membered β-lactam ring fused to a thiazolidine ring in the penicillin core resembles the D-Ala-D-Ala termini of the pentapeptide stems in PG which interacts with the PBP active site (Figure 1.29) [190]. The most prominent β-lactams are the penicillins and cephalosporins, which were first purified at the Sir William Dunn School in Oxford [191] and which remain the most widely consumed β-lactam antibiotics by sales (Figure 1.29) [192]. However, use of these antibiotics to treat infections caused by pathogenic bacteria has become increasingly difficult, because many bacterial pathogens have become resistant due to the horizontal gene transfer of mobile elements that encode β-lactamase enzymes which degrade β-lactams [191,193,194]. Thus, the development of novel β-lactam-containing antibiotics is an ongoing area of research. More recent work has led to the

identification and optimisation of the carbapenems (Figure 1.29), which display broad spectrum activity against Gram positive, Gram negative, and anaerobic bacteria [195,196], and which are considered amongst the most effective antibiotics for the treatment of bacterial infections [197].

β-Lactam antibiotics are not currently used as a standard treatment for mycobacterial infections [198,199], in part, because the genome of *Mycobacterium tuberculosis* encodes for the β-lactamase BlaC [200,201]. Approved antimicrobial TB therapies primarily target RNA polymerases, ATP synthases, transmembrane transport proteins, and other metabolic enzymes [202] in *M. tuberculosis*. New therapeutic drugs for TB are in clinical trials, such as the synthetic macrocyclic peptide inhibitor cyclohexylgriselimycin which inhibits the mycobacterial DNA polymerase sliding clamp [107]. However, recent studies show that the carbapenems, which are approved antimicrobial drugs, are effective in treating TB in mouse models when combined with the β-lactamase inhibitor clavulanate [203]. Interestingly, the class A serine β-lactamase BlaC manifests limited activity against carbapenems [204–206]. Thus, carbapenems hold the potential to be effective drugs against *M. tuberculosis*.

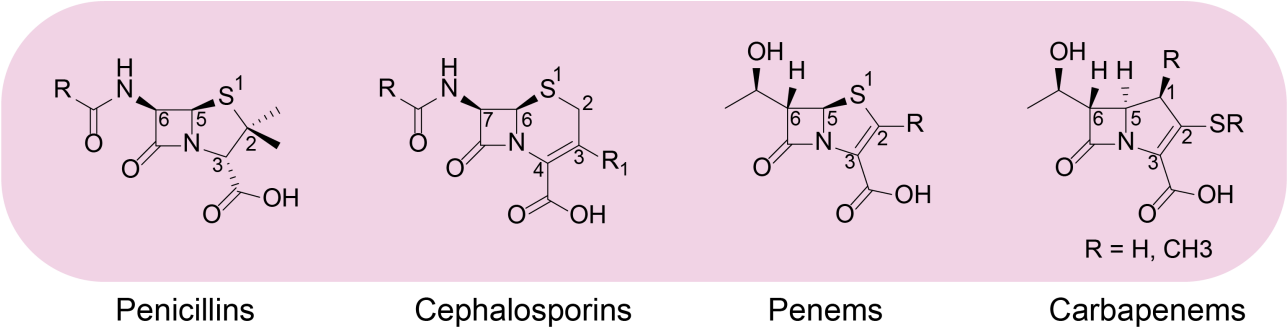


Figure 1.29 | Scheme of the major subclasses of β-lactam antibiotics. The four membered β-lactam ring is fused to a five-membered thiazolidine ring in the penicillins, to a dihydrothiazine ring in the cephalosporins, to a dihydrothiazole ring in the penems, and to a pyrroline ring in the carbapenems.

1.4.4.1 Carbapenems and penems

The work described in this thesis involves biochemical and inhibition studies on SARS-CoV-2 M^{pro} and Ldts, as well as biochemical studies on the interactions of serine-β lactamases and metallo-β-lactamases with the carbapenem and penem β-lactam subclasses. Thus, these β-lactam subclasses are described in greater detail here.

The bicyclic carbapenem core bears a thioether substituent at C2 and an α-hydroxyethyl substituent at C6. The latter group contrasts with the penicillins and cephalosporins, which possess an acylamino substituent at the corresponding C6 position. Although early carbapenems (e.g., the olivanic acids; Figure 1.30) demonstrated the ability to inhibit β-lactamases, they manifested poor chemical stability and cell permeability [207]. The search for β-lactamase inhibitors also led to the discovery of thienamycin (a carbapenem) [195], the latter of which also displayed broad spectrum activity against *Pseudomonas aeruginosa* (Gram negative), *Bacteroides fragilis* (anaerobe) and methicillin/oxacillin-susceptible *Staphylococcus aureus* [208,209]. Imipenem, a semi-synthetic thienamycin derivative is relatively stable, and became the first carbapenem approved for clinical use for the treatment of complex microbial infections [195]. Panipenem, another thienamycin derivative, also displayed high affinity to PBPs and was less susceptible to β-lactamases compared to the penicillins [210].

The generally greater stability of carbapenems compared to penicillins and cephalosporins in the presence of β-lactamases may arise in part from the trans configuration of the C5-C6 bond, and the presence of a carbon atom in the pyrroline ring instead of the analogous sulfur atom in the penicillin thiazolidine ring (Figure 1.29) [211]. Imipenem and panipenem could therefore be administered without a β-lactamase inhibitor, although they are susceptible to degradation by human renal dehydropeptidase I (DHP-I) [212]. Co-administration of imipenem or panipenem with the DHP-I inhibitor cilastatin [213] prolonged the half-lives of these early carbapenems. The addition of a 1-β methyl substituent to C1 of the carbapenem scaffold yielded drugs such as ertapenem, biapenem, meropenem, and doripenem, which are not susceptible to DHP-I-mediated hydrolysis [212]. The stability and spectrum of activity of these carbapenems varies; Gram negative bacteria are more susceptible to meropenem, biapenem, ertapenem, and doripenem, whilst imipenem, panipenem, and doripenem are more effective against Gram-positive bacteria [214,215].

Treatment of bacterial infections by combining a β-lactam with a β-lactamase inhibitor, such as clavulanate, has been used to improve the antibacterial efficacy of β-lactam antibiotics against resistant bacteria [214]. However, carbapenem-resistant *Enterobacterales* have become a major problem worldwide due to the production of serine and metallo-β-lactamases with carbapenemase

activity [216,217]. Genes encoding these serine β-lactamase (SBL) and metallo-β-lactamase (MBL) have been widely disseminated on mobile genetic elements, which often also encode resistance mechanisms against aminoglycosides, fluoroquinolones, and other antibiotics [218].

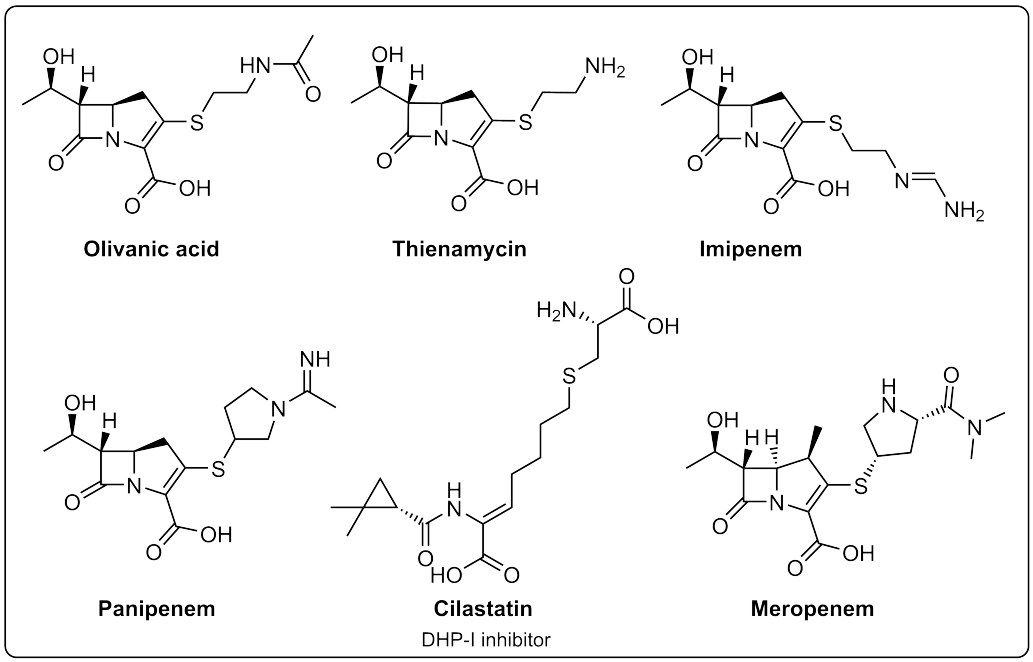


Figure 1.30 | Chemical structures of carbapenems and β-lactamase inhibitors. Some carbapenems are co-administered with cilastatin to prevent their hydrolysis by DHP-I, a hydrolytic enzyme present in the gut.

Despite preliminary evidence showing the efficacy of the carbapenem meropenem for the treatment of tuberculosis in *in vitro* studies [173,219], mouse models of tuberculosis [203] and in clinical trials [220,221], the administration of carbapenems is inherently challenging as intravenous injection is required owing to their short half-lives and poor bioavailability [195,214]. Faropenem is an orally administrated penem which, like meropenem, is inefficiently hydrolysed by BlaC [222]; faropenem is also more efficient at inhibiting Ldt_{Mt2} than meropenem [222]. The structural core of faropenem contains a 4:5 fused bicyclic ring system that resembles to that of the carbapenems, but contains a sulfur atom instead of a carbon at position 1 in the five-membered ring in carbapenems (Figure 1.29); faropenem also has a tetrahydrofuran substituent on C2. Similar to the carbapenems, faropenem possesses a double bond between C2 and C3, and a C6 α-hydroxyethyl side chain with a trans configuration relative to the ring junction at C5. Faropenem is currently in clinical trials for anti-tuberculosis therapy [223]. In non-growing metabolically active mycobacterial cells, faropenem manifested a 2-fold higher

rate of cytolysis of mycobacterial cells compared to meropenem-clavulanate, and 20-fold higher rate of cytolysis compared to isoniazid, which is currently prescribed as a part of anti-tuberculosis treatment [222]. Thus, the application of β-lactams for TB treatment is a promising strategy to fight multiple drug resistant and extensively drug resistant *M. tuberculosis*.

1.4.4.2 Interaction of Ldt_{Mt2} with β-lactams

Acylation of the Ldt_{Mt2} cysteine by β-lactams is analogous to the acylation of this cysteine nucleophile by the donor tetrapeptide substrate in PG. The amide bond in the β-lactam ring opens upon reaction with Cys, forming a stable thioester acyl enzyme complex; the formation of this thioester intermediate with β-lactams inhibits the transpeptidation reaction catalysed by Ldts. The stability of these acyl-enzyme complexes may vary based on affinity, hydrolysis, and interactions with the β-hairpin lid in the catalytic domain [178, 224].

Interaction of Ldt_{Mt2} with penicillins

Penicillins are poor inhibitors of Ldts, and are far more effective against the D, D-transpeptidases [225, 226]. Penicillins target PBPs by acylating the nucleophilic serine, inhibiting their cross-linking activity (except for enzymes like PBP5 which are resistant to penicillins [227]). Upon reaction of penicillin with Ldt_{Mt2}, a thioester adduct is formed. Evidence for the formation of the intact acyl-enzyme complex and degradation of this complex via C5-C6 fragmentation have been reported [including my previous work [228] (manuscript not submitted as a part of this thesis), Figure 1.31] with Ldt_{Mt2} [173] and Ldt_{fm} [229].

Interaction of Ldt_{Mt2} with cephalosporins

Cephalosporins contain a bicyclic 4:6 fused ring system (Figure 1.29) and a C6-amido group analogous to that of penicillins. It is reported that most cephalosporins do not display substantial inhibitory activity against L,D-transpeptidases [226]. Some cephalosporins can form covalent acyl-enzyme complexes with Ldt_{Mt2} but do not undergo fragmentation of the C6-C7 bond to release the six-membered dihydrothiazine ring, in an analogous manner as observed with the penicillins and penems [228, 230, 231]. A concerted

mechanism of fragmentation involving the loss of the C3' leaving group has been proposed for the acyl-enzyme complexes formed with cephalosporins such as ceftriaxone [232].

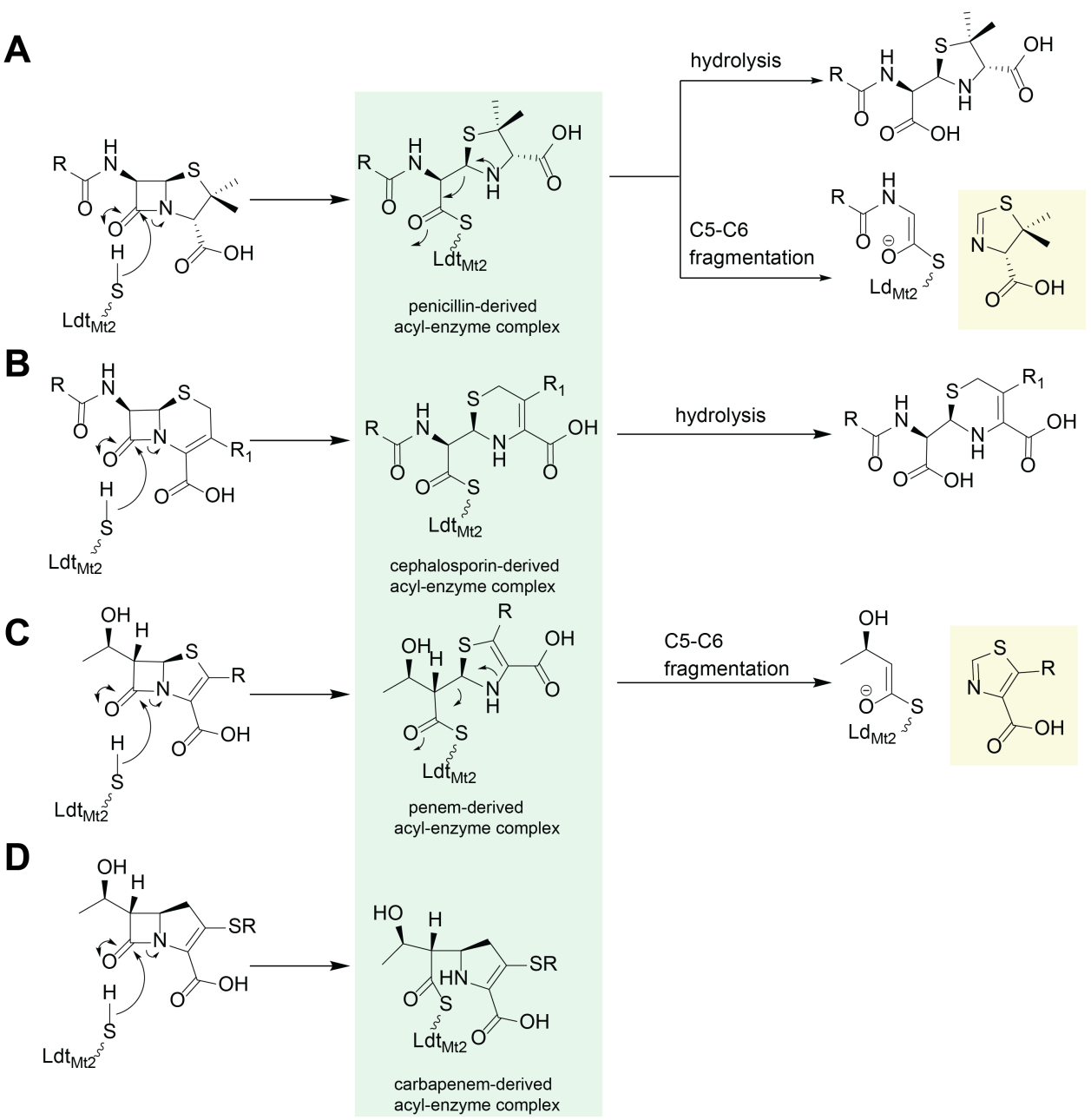


Figure 1.31 | Reactions of different subclasses of β-lactam antibiotics with Ldt$_{Mt2}$. Mechanisms of interaction of **A.** penicillin, **B.** cephalosporin, **C.** penem, and **D.** carbapenem β-lactam antibiotics with Ldt$_{Mt2}$ via formation of an acyl-enzyme complex involving the nucleophilic cysteine. C5-C6 fragmentation was observed with the penicillin and penem subclasses; the degraded products released from the acyl-enzyme complex are highlighted in yellow. Hydrolysis of penicillins and cephalosporins was also observed, whereas the complexes derived from penems and carbapenems did not undergo hydrolysis after more than 16 hours [228].

Interaction of Ldt_{Mt2} with carbapenems

Compared to the penicillins and cephalosporins, the carbapenem subclass of β-lactam antibiotics is capable of efficiently inhibiting both L,D-transpeptidases and PBPs [226,233]. Carbapenems interact with L,D-transpeptidases by acylating the active site cysteine residue [173,174], although not all Ldts are inhibited by carbapenems (e.g., Ldt_{Mt5} is not inhibited) [182,222].

The acylation of L,D-transpeptidases by carbapenems (such as meropenem, imipenem, panipenem, and doripenem) [174,178,225,234,235] occurs with the opening of the β-lactam ring. In the case of the Ldt acyl-enzyme complexes derived from meropenem, ertapenem, and imipenem, the pyrroline ring has been observed in two conformations [183,233,236]. In conformation I, the nitrogen of the *2S*-pyrroline ring is close to the thioester carbonyl group, while in the second conformation, the nitrogen is rotated by a 180° and is farther from the thioester carbonyl group. In addition to the interaction of the β-lactam carbonyl group in the oxyanion hole, the C6-hydroxyethyl side chain of the carbapenems forms additional contacts with Tyr318 and, upon flipping of the pyrroline ring, hydrogen bonds are formed with Tyr308 [176,224] (Figure 1.33C). Whilst most carbapenems form stable acyl-enzyme complexes with most Ldts, degradation of the acyl-enzyme complexes derived from biapenem and tebipenem have been proposed based on crystallographic studies [225]. However, the reported mass spectrometric evidence is not in accord with this proposed degradation of the intact acyl-enzyme complexes [173,225].

Binding studies with different carbapenems suggest that the structure of the thioether sidechain and the presence or absence of the 1-β-methyl group impact the affinity and inhibition of these carbapenems with Ldt_{Mt2}. For instance, imipenem, which has a higher binding affinity to Ldt_{Mt2} than ertapenem and meropenem, exhibits higher antibacterial efficacy than meropenem in *in vivo* studies [203,236,237]. Crystallographic studies have provided insights related to the higher binding affinity of imipenem [236]. In the case of the imipenem-Ldt_{Mt2} complex in conformation II, the pyrroline ring is sandwiched between Tyr308 in the β-hairpin loop and Gly353. It is proposed that the 1-β-methyl group makes the pyrroline ring bulkier in 1-β-methyl-containing carbapenems, such that they do not fit between Tyr308 and Gly353. As a consequence, Tyr308 is pushed into an alternative conformation by 1-β-methyl containing carbapenems, which results in the movement of the β-hairpin loop away from the active site

(Figure 1.32) [178, 224, 236]. This loop movement may expose the thioester carbonyl of the acyl-enzyme complex in 1-β-methyl containing carbapenems, increasing its likelihood to undergo hydrolysis.

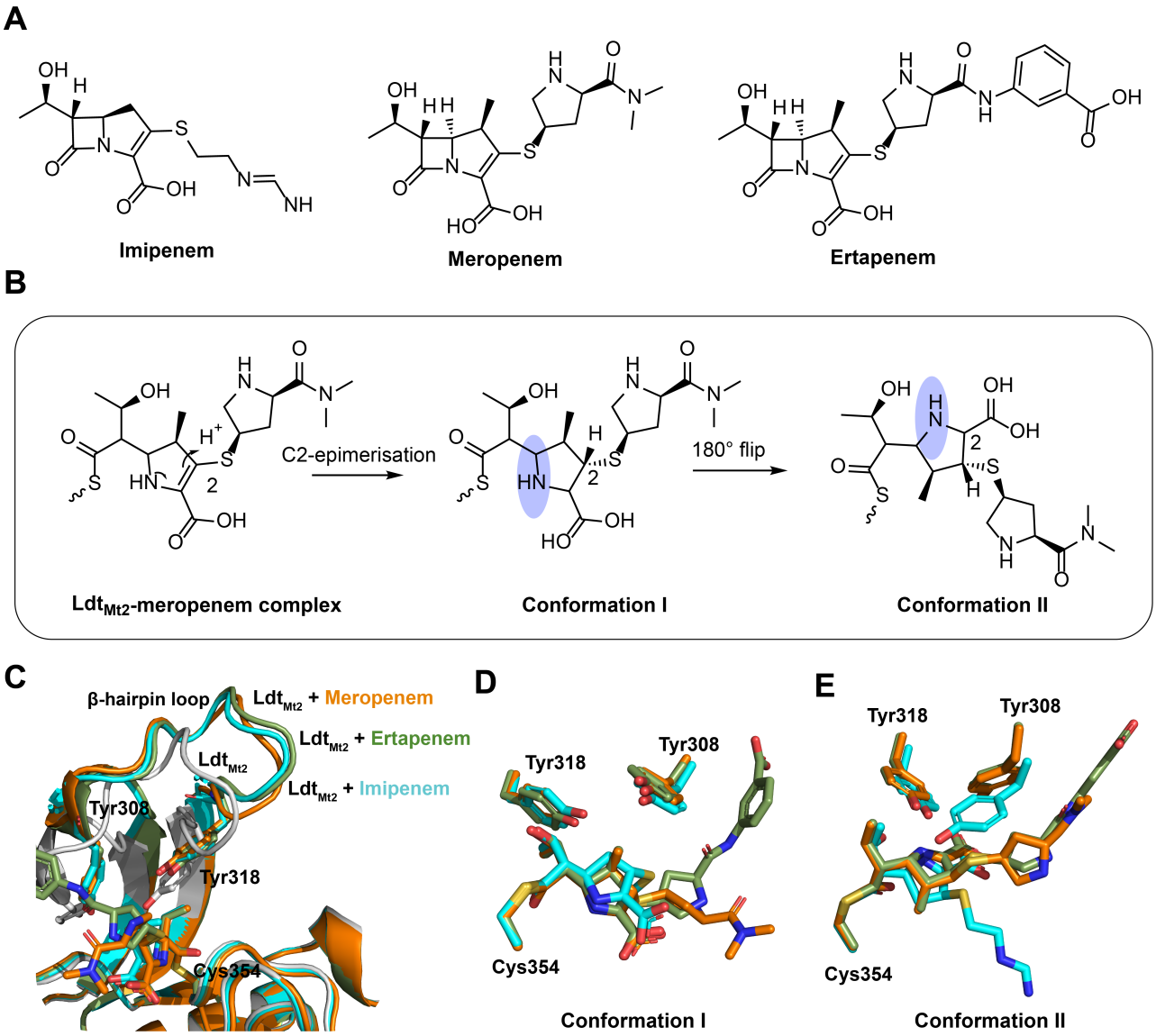


Figure 1.32 | Carbapenem-derived complexes adopt two different conformations following acylation of Ldt$_{Mt2}$. **A.** Structures of representative carbapenems. Imipenem lacks the 1-β-methyl group that is present in meropenem and ertapenem. **B.** Proposed reaction mechanism of the Ldt$_{Mt2}$-meropenem complex under crystallographic conditions [176, 224]. **C.** The β-hairpin loop moves away from the active site upon the covalent reaction of the active site Cys354 with carbapenems imipenem (PDB: 6IYW), ertapenem (PDB: 6IYV), and meropenem (PDB: 3VYP). **D.** Conformation I adopted by the carbapenem-Ldt$_{Mt2}$ complex. **E.** Conformation II adopted by the carbapenem-Ldt$_{Mt2}$ complex. The pyrroline ring in conformation I rotated by a 180° to adopt conformation II which allows both Tyr308 and Tyr318 to form hydrogen bonds with the C3-carboxylic acid and the C6-hydroxyethyl side chain. In the case of imipenem, the absence of 1-β-methyl group in the pyyroline ring allows the pyyroline ring to be sandwiched between Tyr308 and Gly353. The presence of 1-βmethyl in the pyrroline ring of meropenem/ertapenem results in rotation of Tyr308 and subsequent movement of the β-hairpin loop.

Interaction of Ldt_{Mt2} with penems, and comparison with interactions of Ldt_{Mt2} with carbapenems

Faropenem, the most effective inhibitor of Ldt_{Mt2}, covalently modifies the nucleophilic Cys354 [178] of Ldt_{Mt2}. Following acylation, the resulting acyl-enzyme complex has been proposed to undergo scission of the C5-C6 bond, leading to the release of a dihydrothiazole ring (Figure 1.31C) [178]. Like the carbapenems, faropenem has a C6-hydroxyethyl group which interacts with the aromatic residues in the β-hairpin lid (i.e., Tyr308 and Tyr318). This interaction with the aromatic residues is absent in the faropenem-derived 3-hydroxybutyryl fragmentation product bound to Ldt_{Mt2} (Figure 1.33A, C). The loss of the dihydrothiazole ring from the faropenem adduct allows the lid to adopt a closed conformation identical to that seen in the apo Ldt_{Mt2} structure [178]; this loop is pushed further away from the active site in all Ldt_{Mt2}-carbapenem complexes in conformation II, including with non-1-β-methyl-containing carbapenems such as imipenem (Figure 1.33A), where the loop is relatively closer to the active site compared to other carbapenems [176, 178, 224, 236].

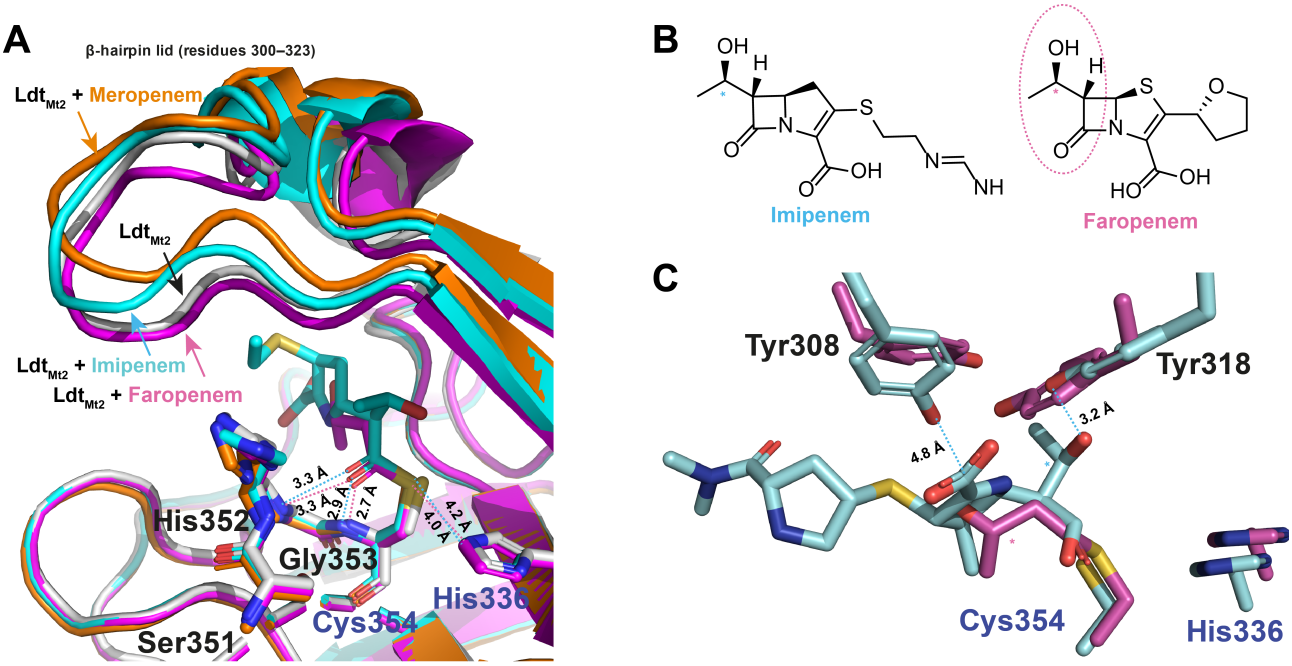


Figure 1.33 | View from a crystal structure of the Ldt$_{Mt2}$-faropenem complex superimposed with the Ldt$_{Mt2}$-imipenem complex. A. The backbone NH groups of Cys354 and Gly353 are positioned to form hydrogen bonds with the β-lactam derived carbonyl in the inhibitor complexes. **B.** Comparison of the chemical structures of imipenem (carbapenem) and faropenem (penem). Faropenem undergoes C5-C6 fragmentation following reaction with the nucleophilic cysteine. The resultant bound 3-hydroxybutyryl group is circled (magenta). The hydroxyethyl groups in faropenem (magenta asterisk) and imipenem (blue asterisk) bind to different regions in the active site [178, 236]. **C.** The C6-hydroxyethyl group is present in both carbapenems and penems. However, only the carbapenem C6-hydroxy group is oriented to form polar interactions with Tyr308 and Tyr318 in the active site lid.

1.5 Objectives

The biochemical and mechanistic studies undertaken in this thesis involved the identification of potent inhibitors of cysteine proteases and the non-hydrolytic mechanism of reaction with cysteine transpeptidases.

The specific aims were:

- SPE-MS based assay development and application for screening new and repurposed drugs for M^{pro} inhibition.
- SAR studies with β-lactam penicillin sulfones for improved M^{pro} inhibition.
- Screening of acyclic and cyclic peptides, small molecules and heterocycles with novel scaffolds against M^{pro}.
- Mechanistic studies with β-lactams with cysteine transpeptidases, SBLs and MBLs

2

Assay development, validation and application towards identification of β-lactam inhibitors against M^{pro}

Contents

2.1 Abstract **53**
2.2 Introduction **54**
2.3 Results **58**
2.4 Methods **119**

2.1 Abstract

SARS-CoV-2 M^{pro} (M^{pro}) is now an established target for antiviral chemotherapeutics to treat COVID-19. Huge efforts have been undertaken by the scientific community to identify potent drugs that target M^{pro}. Some M^{pro} inhibitors, such as N3 and ebselen have been identified using fluorescence resonance energy transfer (FRET) based assays. In March 2020, work to develop a label-free mass spectrometry assay was initiated to validate reported M^{pro} inhibitors and to identify novel M^{pro} inhibitors, for example by investigating commercial sets of small molecules. High throughput assay with excellent assay parameters ($Z' \geq 0.8$) was developed and more than >2500 commercial compounds were screened. Application of the assay to investigate β-lactam antibiotics led to the identification of β-lactams as potent M^{pro} inhibitors as part of these studies, suggesting they hold the potential to be developed as antivirals.

2.2 Introduction

At the onset of the COVID-19 pandemic, no viable chemotherapeutics were available to treat COVID-19. One of the means of achieving antivirals would be via targeting the essential viral protein M^{pro}, which catalyses the hydrolysis of the viral polyprotein to release the functional viral proteins essential for viral transcription-replication machinery [19–22]. A robust assay that can monitor M^{pro} activity, is crucial for identifying selective inhibitors of M^{pro}, information that would eventually help the development of antiviral COVID-19 therapeutics. Most reported assays for M^{pro} use FRET probes (Figure 2.1A) [28,238]. An alternative fluorescence based M^{pro} assay employing a rhodamine-bearing substrate (Figure 2.1B) has been reported for application in SARS-CoV-1 M^{pro} assay [239]. The bifunctional rhodamine substrate of fluorescence based M^{pro} assay is unique; this substrate undergoes cleavage at two sites to release rhodamine 110 in a manner such that each cleavage can be monitored separately (Figure 2.1B). The initial cleavage can be monitored via absorbance at 469 or 492 nm and the latter event can be monitored by absorbance at 496 nm or fluorescence at 533 nm [239]. The quantum yield for rhodamine 110 is high, making it one of the most sensitive probes to be used in M^{pro} assays. However, the catalytic efficiency of M^{pro} with this substrate is low (k_{cat}/K_m = 20 ± 3.3 $M^{-1}s^{-1}$) [239] compared to the FRET substrates (k_{cat}/K_m = 26000 $M^{-1}s^{-1}$) [28], making it a less attractive substrate for M^{pro} assays. Furthermore, the two reactions involved in the rhodamine probe assay introduce complications. Therefore, despite their high sensitivity and potential utility for high throughput (HT) screening, rhodamine-based probes have not been widely used in M^{pro} screens. Accordingly, to date, only FRET-based assays, originally developed for SARS-CoV-1 M^{pro}, have been translated to HT screening campaigns with SARS-CoV-2 M^{pro}.

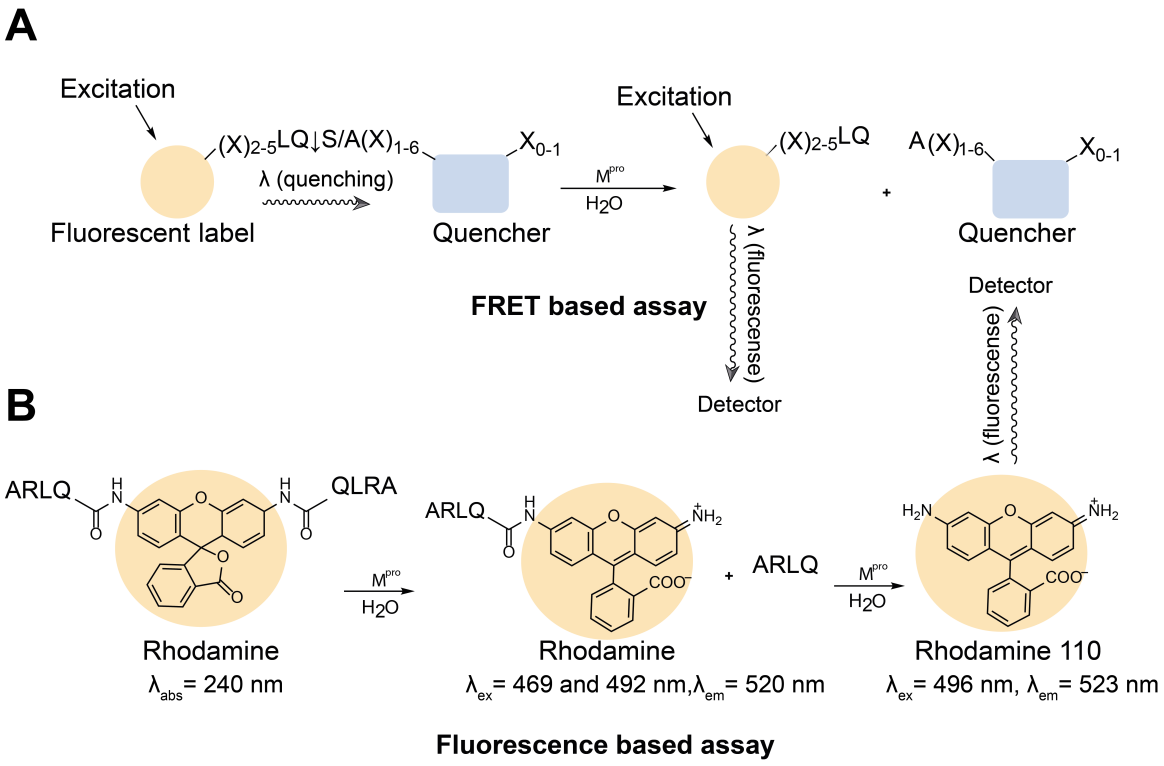


Figure 2.1 | SARS-CoV-1 M^{pro} fluorescence-based assays. A. The FRET-based M^{pro} assay. M^{pro} catalyses the hydrolysis of the fluorogenic substrate peptide, releasing the quencher. The fluorescence from the cleaved peptide is then detected by a plate reader. **B.** A fluorescence based M^{pro} assay. Two cleavage sites are present in the bis-substituted C2-symmetric rhodamine probe; M^{pro} releases the fluorogenically active rhodamine 110 dye via hydrolysis of the monosubstituted dye.

FRET based M^{pro} assays have found widespread applications for inhibitor screening. For instance, the M^{pro} inhibitor N3, which was reported within 3 months of the onset of COVID-19 pandemic, was identified by a FRET-based assay which uses 7-methoxycoumarin-4-acetic acid (Mca) and 2,4-dinitrophenyl (Dnp) FRET-pairs in a Mca–AVLQ/SGFRK(Dnp)K substrate peptide (Figure 2.2C) [240]. Other FRET acceptor-quencher pairs, such as the (i) 2-aminobenzoyl (Abz) and nitro-tyrosine pair in an Abz(-SVTLQ/SG-Tyr(NO_2)R [78,241] substrate (Figure 2.2A) and (ii) 4-([4-(dimethylamino)phenyl]azo)benzoic acid (Dabcyl) and 5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid (Edans) pair in Dabcyl-SAVLQ/SGFRKM-Edans [32] and Dabcyl-KTSAVLQ/SGFRKM-E(Edans) [29] substrates (Figure 2.2B), have been applied to screen compounds against SARS-CoV-1 M^{pro}. In all cases, the quencher is released from the FRET substrate through M^{pro} catalysis, resulting in a fluorescent signal from the product bearing the fluorescent probe (Figure 2.2A-C). SARS-CoV-1

M^{pro} FRET assays have also employed other fluorophores, e.g. AlexaFluor® 488/AlexaFluor® 594 coupled with quencher pairs QSY®-7/QSY®-21 in their substrates [242]. The catalytic efficiency of M^{pro} is 10 times slower with these fluorophores than other FRET pairs, possibly due to the steric hindrance resulting from the relatively larger size of the fluorophore dyes [242]. In general, the FRET based assays are efficient, and suitable for high-throughput applications, but may incur interference by them from autofluorescence or quenching from ligands [243]. Kinetic studies with FRET substrates have not demonstrated consistent catalytic efficiencies, potentially resulting from the use of different (unnatural) FRET pairs [32, 78, 240, 241, 244, 245]. Additionally, when designing longer peptide substrates, FRET probes may require considerable optimisation, in part due to the requirement for close proximity of the FRET pairs [246].

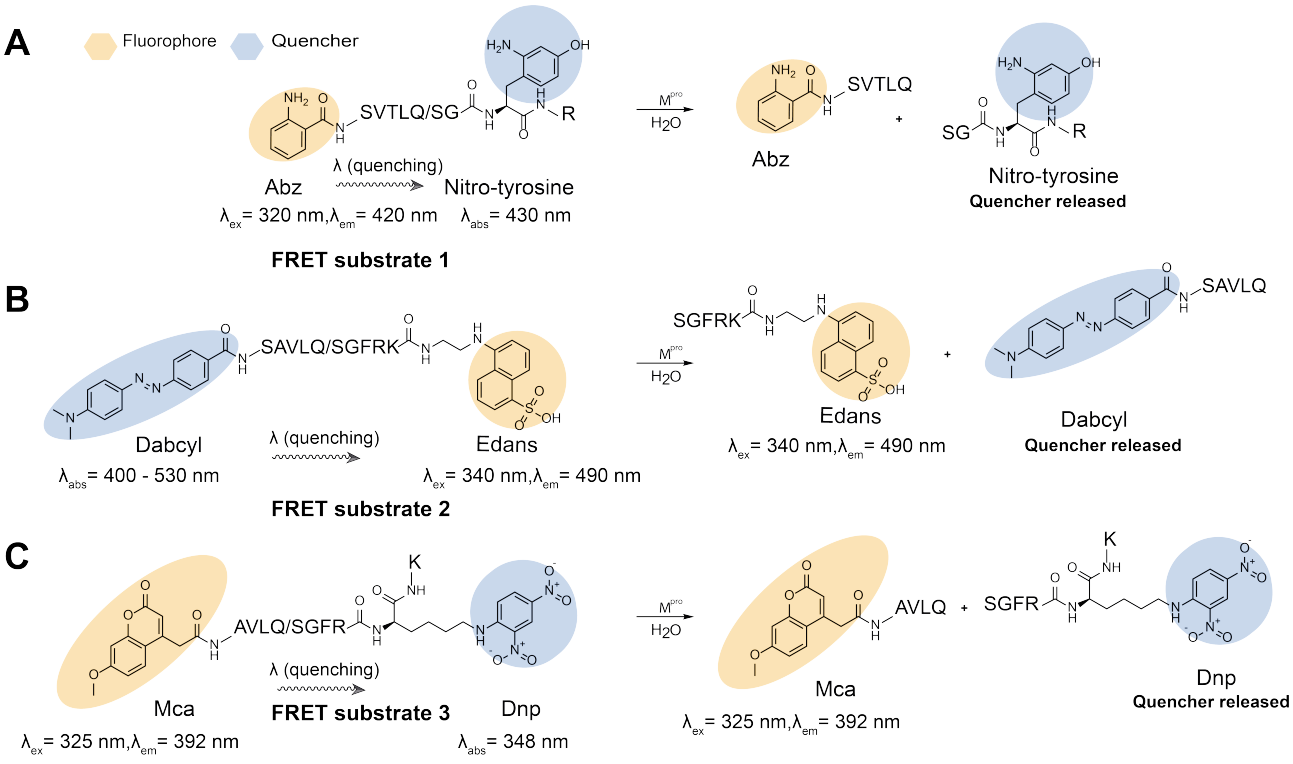


Figure 2.2 | Fluorogenic peptide substrates of M^{pro} with different FRET-pairs. A. Cleavage of a Abz-NO_2(Tyr) FRET pair-containing M^{pro} substrate. **B.** Cleavage of an Edans-Dabcyl FRET pair-containing M^{pro} substrate. **C.** Cleavage of an Mca-Dnp FRET pair-containing M^{pro} substrate.

In addition to FRET and fluorescence-based assays, label-free peptide substrates have been employed in high-performance liquid chromatography (HPLC) based M^{pro} assays. These assays are typically low throughput, require large amounts of sample, and thus are inapplicable to large-scale inhibitor

screening efforts. Therefore, they have only been exploited to probe substrate kinetics which revealed discrepancies amongst different reports [31,244] (Table 2.1).

This chapter details the development of a solid phase extraction coupled to mass spectrometry (SPE-MS) based M^{pro} assay, its application for inhibition, kinetic studies, and mechanistic studies with M^{pro}. The chapter also entails detailed studies on assay validation using commercial small-molecule libraries of FDA approved and pharmacologically active compounds, and commercial and in-house β-lactams and protease inhibitors. The M^{pro} assay guided medicinal chemistry efforts to develop the identified penicillin β-lactam inhibitors into potent M^{pro} inhibitors.

Table 2.1: Reported methods and substrates for investigating the kinetic parameters of M^{pro}. Reported kinetic parameters of SARS-CoV-1 and SARS-CoV-2 M^{pro}.

SARS-CoV	Substrate	k_{cat}/k_m ($M^{-1}s^{-1}$)	Method
1	[a]Abz-SVTLQ/SG-Tyr(NO_2)R	20 ± 3.3	FRET [241]
1	TSAVLQ/SGFRK-NH_2	176 ± 10	RP-HPLC [31]
1	[b]Dabcyl-LAQ/AVRSSSR-Edans	3666	RP-HPLC [245]
1	[c]Mca-AVLQ/SGFR-Lys(Dnp)-K-NH_2	26500	FRET [240]
1	[b]Dabcyl-SAVLQ/SGFRK-Edans	29000	FRET [244]
1	TSAVLQ/SGFRKW	2400	RP-HPLC [244]
1	[a]Abz-SVTLQSG-Tyr(NO_2)-R	10000 ± 3333	FRET [78]
1	(ARLG-$NH)_2$-Rhodamine	50	FRET [239]
1	[b]Dabcyl-KTSAVLQ/SGFRKM-E(Edans)	3011.3 ± 294.6	FRET [32]
2	[b]Dabcyl-KTSAVLQ/SGFRKM-E(Edans)	3426.1 ± 416.9	FRET [32]
2	[a]Abz-SVTLQSG-Tyr(NO_2)-R	30000 ± 6667	FRET [78]
2	[c]Mca–AVLQ/SGFRK(Dnp)K	28500	FRET [28]

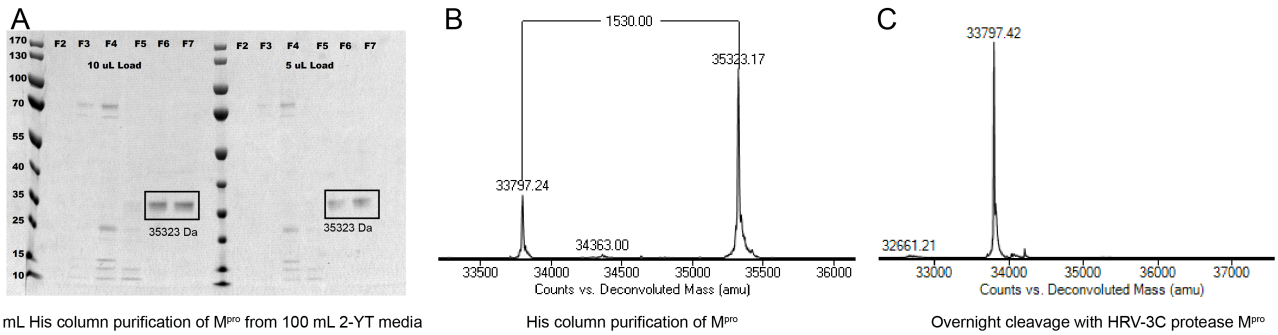
[a] Abz: 2-aminobenzoyl. [b] Dabcyl: 4-([4-(dimethylamino)phenyl]azo)benzoic acid; Edans: 5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid. [c] Mca: 7-methoxycoumarin-4-acetic acid; Dnp: 2,4-dinitrophenyl.

2.3 Results

2.3.1 Small scale preparation of M^{pro} protein from pGEX-6p1-M^{pro} plasmid

M^{pro} was kindly provided by the group of Prof. Martin Walsh for assay development and validation work. The M^{pro} gene inserted in pFLOAT vector was also made available by the group of Prof. John Vakonakis [247]. However, the protein produced with the latter plasmid required an extra purification step for Glutathione S-transferase (GST) tag removal. Therefore, a commercial plasmid pGEX-6p1, based on reported literature [28,240], was designed and production of the M^{pro} protein using it was assessed (Figure 2.3). After immobilised metal affinity chromatography (IMAC) purification of the clarified filtered cell lysates, a single protein band was visible by SDS-PAGE analysis. Analysis of the pooled fraction by mass spectrometry (MS) revealed the presence of a peak corresponding to the C-terminal hexaHis tag cleaved M^{pro} (33797 Da) as well as the expected C-terminal hexaHis tagged M^{pro} (Figure 2.3 B). Overnight incubation (>10 h, 8 °C) of the C-terminal hexaHis tagged M^{pro} with the HRV 3C protease (100:1 molar ratio) successfully removed the C-terminal hexaHis tag to produce wildtype (WT) M^{pro} (Figure 2.3C). A low yield (0.5 mg/ 100 mL) was obtained after IMAC purification of the hexaHis-tag cleaved M^{pro} to separate the His tag cleaved M^{pro} from the cleaved His tag and the His-tagged HRV 3C protease.

M^{pro} was subsequently obtained from the Walsh group [248] or was produced by Eidarus Salah (Department of Chemistry, Oxford).



1 mL His column purification of M^{pro} from 100 mL 2-YT media | His column purification of M^{pro} | Overnight cleavage with HRV-3C protease M^{pro}

Figure 2.3 | His column purification of M^{pro}. A. SDS PAGE gel loaded with fractions obtained from His column elutes was stained using InstantBlue® Coomassie Protein Stain. The bands corresponding to the expected MW (35323 Da) of hexaHis tagged M^{pro} was observed in fractions 6 and 7. **B.** Mass spectrometry analysis of fraction 6 reveals the presence of a mixture of His cleaved and His tagged M^{pro} protein. **C.** Treatment with HRV 3C protease cleaves the C-terminal hexaHis tag of M^{pro}. A single peak at 33797 Da was observed which corresponds to the expected mass of WT M^{pro} (33796.6 Da).

2.3.2 Activity of different C terminal constructs of M^{pro}.

The presence of the authentic (i.e. natural) N-terminus of M^{pro} is reportedly crucial for its activity [240]. By contrast, the presence of additional residues at the C-terminus of M^{pro} are reported to not affect the activity of SARS-CoV-1 M^{pro} [240]. Due to the high sequence and structural similarity between SARS-CoV-1 M^{pro} and SARS-CoV-2 M^{pro}, a similar outcome was expected for SARS-CoV-2 M^{pro} (hereafter referred to as M^{pro}). The activities of four C-terminally affinity-tagged M^{pro} constructs: C-terminally decaHis tagged M^{pro} (M^{pro} GP-10 His, M^{pro} GP-(SG)2-10 His, and M^{pro} GP-SG-10), and Avi-tagged M^{pro} (See Methods for full sequences), were compared that of with WT M^{pro} using the SPE-MS assay which simultaneously monitors substrate depletion and product formation (see below).

A single time-point of 10 min was selected to calculate activity based on both product formation and substrate depletion. The percentage activity was calculated as (abundance of product/(abundance of product + abundance of substrate))*100). Similar activities were observed for all the M^{pro} samples, with the exception of Avi-tagged M^{pro}. WT M^{pro} demonstrated the highest activity (59 ± 3 %), followed by M^{pro} GP-10 His (44 ± 2.3 %), M^{pro} GP-(SG)2-10 His (40 ± 2.3 %) and M^{pro} GP-SG-10 His (36 ± 0.6 %). WT M^{pro} manifested the highest activity at all times compared to all the C-terminally His-tagged constructs; no turnover was observed with M^{pro} with a C-terminal Avi-tag. The M^{pro}

variants which were decaHis tagged displayed identical activity profiles (Figure 2.4) suggesting that modification of the WT C-terminus does not substantially impact on M^{pro} activity. However, the lack of activity with the C-terminal Avi-tagged M^{pro} shows that C-terminal modification can ablate activity. In the case of M^{pro} that had undergone repeated freeze thaw cycles, a 6-fold lower activity was observed compared to the activity of the freshly thawed WT M^{pro} (Figure 2.4B). Therefore, all following experiments were performed with freshly thawed aliquots of WT M^{pro}.

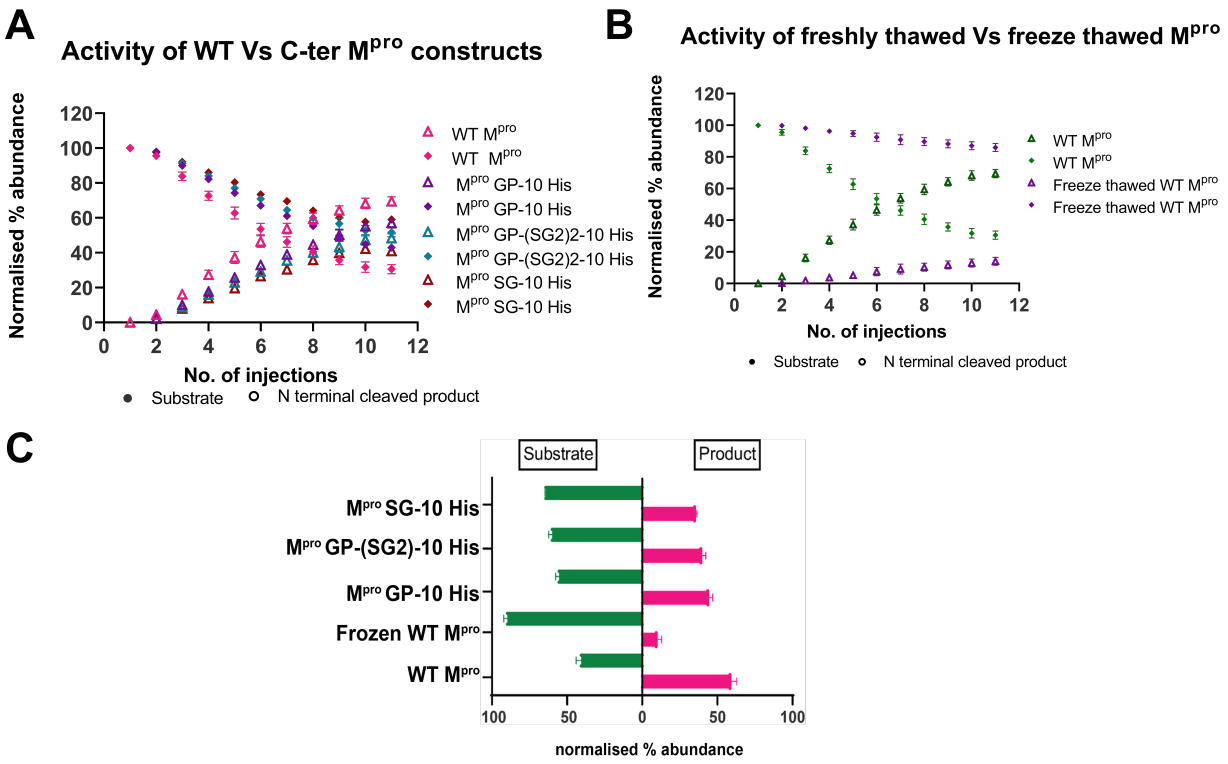


Figure 2.4 | Comparison of the activity of different C terminus constructs of M^{pro} with WT M^{pro} using an MS based turnover assay. The assay monitors the abundance of the substrate and the N-terminal product peptide. **A.** Activity of WT and three different C terminus M^{pro} constructs. WT M^{pro} has the highest activity. All three constructs of M^{pro} with a decaHis tag at the C-terminus are active and manifest similar activities. **B.** Freeze thaw cycles reduce M^{pro} activity. Substrates and products are represented as filled symbols and open symbols, respectively. **C.** The normalised percentage abundance of substrate and peptide with different constructs and freeze-thawed M^{pro} sample at 10 min. The number of sample injections indicates the consecutive sample aspiration by the RapidFire liquid handling robot from a reaction well consisting of M^{pro}: peptide mixture into the QTOF. Each injection was ∼1 min apart.

2.3.3 RapidFire Mass Spectrometry based assay development

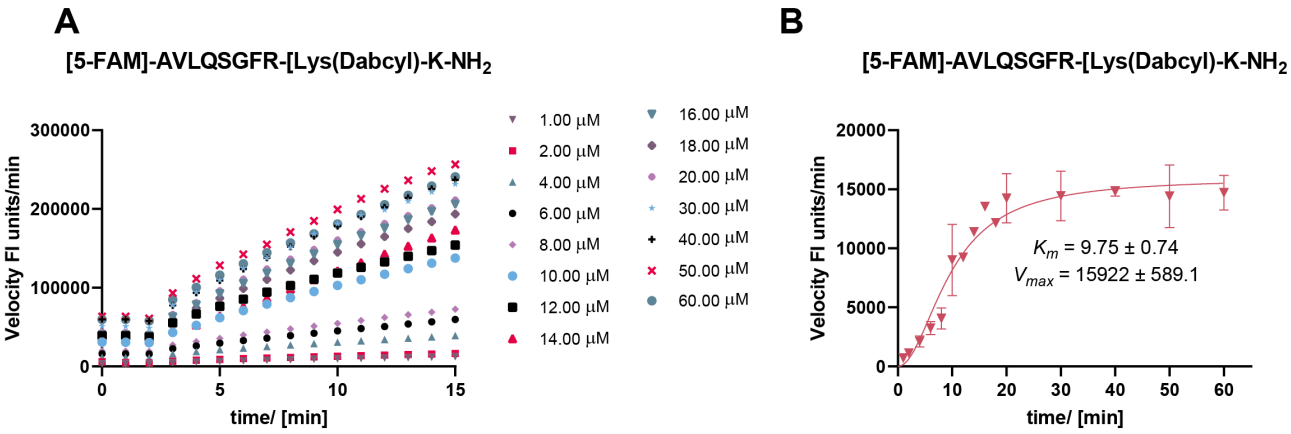


Figure 2.5 | K_m calculation of FRET substrate [5-FAM]-AVLQSGFR-[Lys(Dabcyl)-K-NH_2 of M^{pro} A. Titration of FRET substrate with M^{pro}. **B.** K_m determination of FRET substrate. Condition: 0.01 µM M^{pro}, 1 µM-60 µM [5-FAM]-AVLQSGFR-[Lys(Dabcyl)-K-NH_2, 20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 0.01% Tween-20

The optimal pH of 7.3-8.5 for SARS-CoV-1 M^{pro} activity has been investigated using FRET substrate peptides containing P4-P5′ residues based on the N-terminal autocatalytic nsp4/5 site of M^{pro} [249]. A commercially available FRET substrate [5-FAM]-AVLQ→SGFR-[Lys(Dabcyl)-K-NH_2 comprising of a green fluorescent tag (5-carboxyfluorescein) and a fluorescence quencher (Dabcyl) pair, was used to assess M^{pro} activity [250,251]. The assay conditions were adapted from those reported for SARS-CoV-1 M^{pro} [240].

Varying the substrate concentration (1 µM - 60 µM, Figure 2.5A) enabled determination of the K_m of M^{pro} as 9.75 ± 0.74 µM (Figure 2.5B). M^{pro} catalysed hydrolysis of the same substrate was then monitored using the SPE-MS paired to a RapidFire 365 liquid handling robot (Agilent) which aspirates the sample and loads it into a selected SPE C4-cartridge for purification [252]. Following an aqueous wash, the sample is directly eluted, with aqueous 85%$_{(v/v)}$ acetonitrile, into a QTOF mass spectrometer, enabling the parallel detection of the mass of substrate (1741.84 Da) and product peptides (787.3 Da, 971.5 Da) (Figure 2.6) [253].

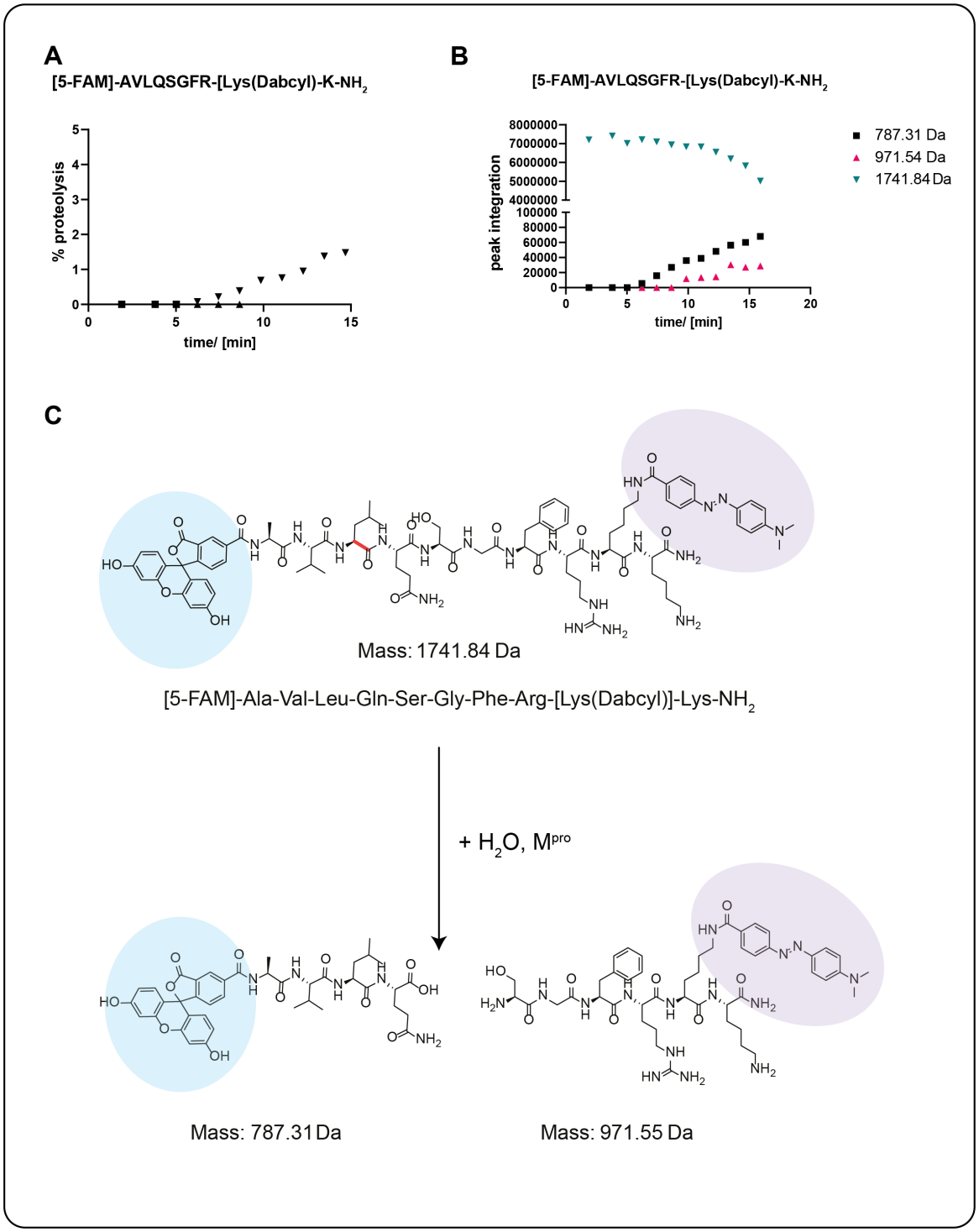


Figure 2.6 | The SPE-MS based assay detects both cleaved products of the FRET substrate [5-FAM]-AVLQSGFR-[Lys(Dabcyl)-K-NH_2 of M^{pro} A. M^{pro} catalysed turnover of FRET substrate. **B.** Abundance of the FRET substrate peptide (mass: 1741.84 Da), the N-terminal product peptide (mass: 787.31 Da), and the C-terminal released product (mass: 971.55 Da). **C.** Scheme depicting the substrate hydrolysis by M^{pro}. Conditions: 0.02 µM M^{pro}, 5.0 µM [5-FAM]-AVLQSGFR-[Lys(Dabcyl)-K-NH_2, 20 mM HEPES pH 7.5, 50 mM NaCl.

A microwave assisted peptide synthesizer was used to synthesize an M^{pro} substrate peptide (with a C-terminal amide) based on the N-terminal autocatalytic site of M^{pro}. Initially, a 20-mer peptide substrate (QTSITSAVLQ/SGFRKMAFPS-NH_2), was screened for activity 0.1 μM M^{pro}, 2.0 μM QTSITSAVLQ/SGFRKMAFPS-NH_2, 20 mM HEPES pH 7.5, 50 mM NaCl. N- and C-terminal product peptides, with masses of 918.49 Da and 1125.57 Da, respectively were detected upon incubation with M^{pro}. The observed N-terminal product peptide was smaller than the expected N-terminal product peptide of 1046.6 Da; the mass difference of 128 Da between the expected and observed N-terminal product peptide corresponds to the loss of the N-terminal Gln residue, likely due to M^{pro} catalysed hydrolysis (Figure 2.7).

The observed shorter N-terminal product from the 20-mer peptide hydrolysis may be a consequence of M^{pro} cleavage of the 20-mer peptide at two sites, i.e. cleavage after the N-terminal Gln and at the expected P1 site between Q/S. However, as the 20-mer peptide was purified as a mixture of 20-mer and contaminant peptides (>85% purity; the mass of the contaminant corresponds to N-terminal Gln cleaved 19-mer peptide), it cannot be ruled out that the 19-mer contaminant peptide without the N-terminal Gln is preferentially hydrolysed (Figure 2.8). The presence of 19-mer peptide was detected in the 20-mer peptide even after repeated rounds of HPLC purification. Furthermore, slow turnover (<25%) of the 20-mer substrate was observed (12 min incubation). Hence, a shorter 11-mer peptide, s01 (also based on nsp4/5 M^{pro} cleavage site, Table 1.1) was synthesized to mimic the 10-mer FRET substrate to test as a M^{pro} substrate in SPE-MS assays.

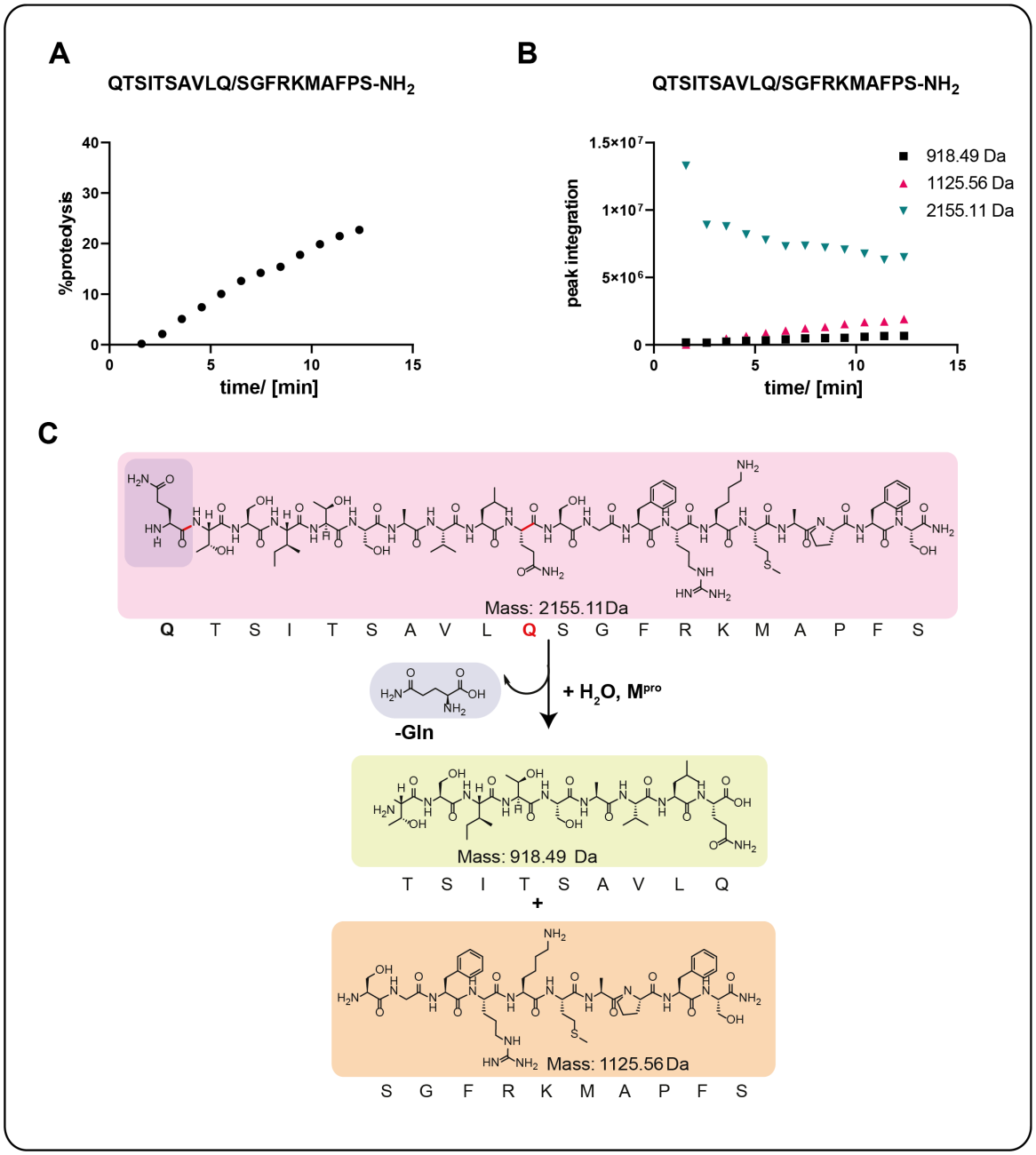


Figure 2.7 | The SPE-MS based assay detects both cleaved products of 20-mer substrate of M^{pro}. A. % Peptide turnover of 20-mer peptide substrate. Note that the N-terminal product peptide had also lost its N-terminal Glu residue. **B.** Abundance of 20-mer substrate peptide (mass: 2155.11 Da), N-terminal product peptide (mass: 918.49 Da), and C-terminal released product (mass: 1125.56 Da). **C.** Scheme depicting the M^{pro} catalysed hydrolysis of the 20-mer peptide.

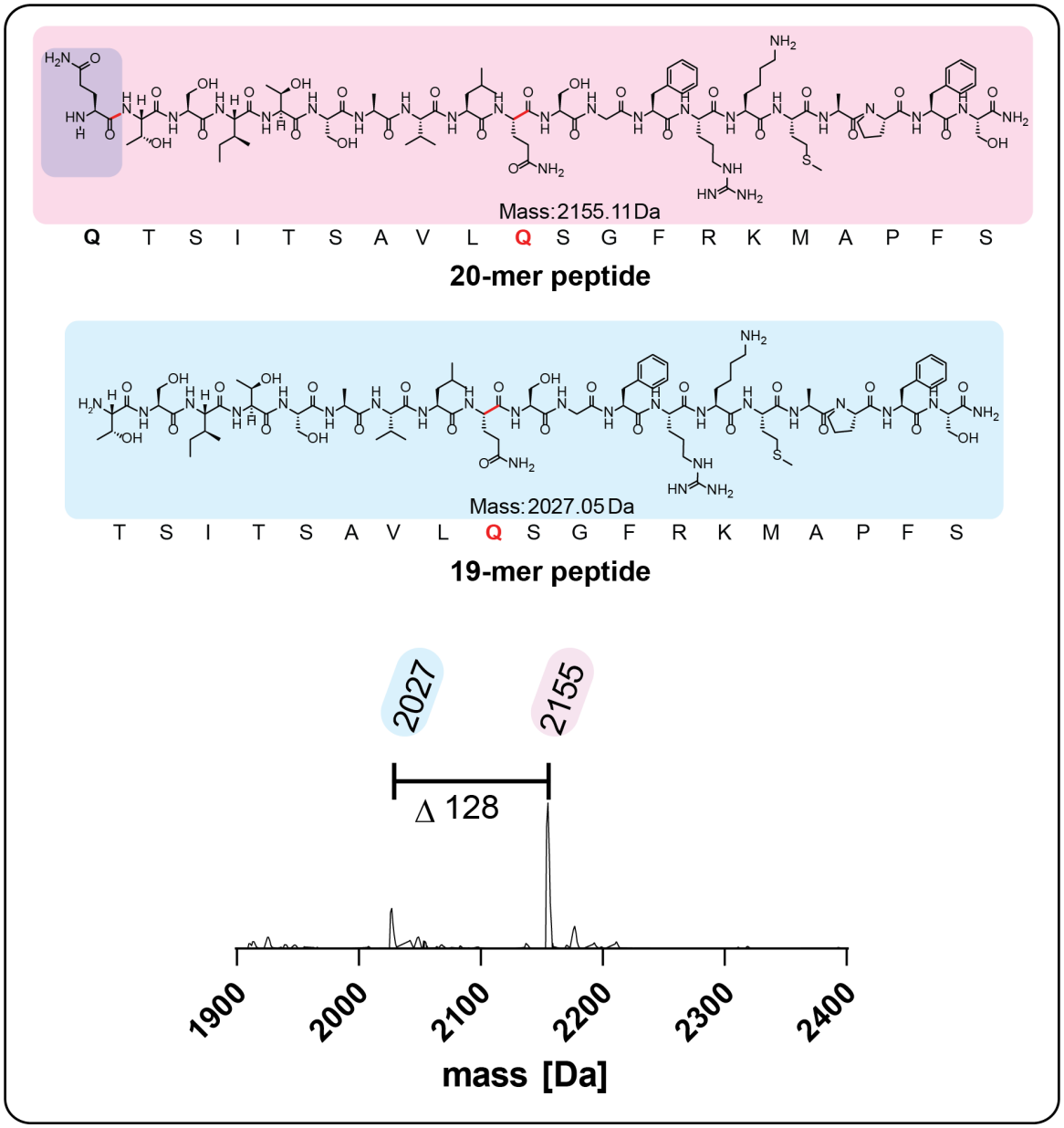


Figure 2.8 | Mass spectrum of 20-mer QTSITSAVLQ/SGFRKMAFPS-NH_2 M^{pro} substrate. The N-terminal product peptide lacking the N-terminal Gln (mass: 2027 Da) is observed in the HPLC purified 20-mer peptide (mass: 2155 Da).

Under the same conditions that were used to assess the 20-mer substrate (i.e. 0.10 µM M^{pro}, 4.0 µM substrate, 20 mM HEPES pH 7.5, 50 mM NaCl), relatively high turnover of the 11-mer peptide s01 (TSAVLQ/SGFRK-NH_2, 1191 Da, Table 1.1) was observed. As M^{pro} contains 12 cysteines, reducing agents were added to hinder the oxidation of the active site cysteine which may alter the activity profile of M^{pro}. Glycerol was added to increase the stability of M^{pro}. The effects of a reducing agent (TCEP) and of glycerol were investigated using the SPE-MS assay. Neither glycerol nor TCEP alters M^{pro}

activity over 3 h (Figure 2.9). These reagents were not included in subsequent assays.

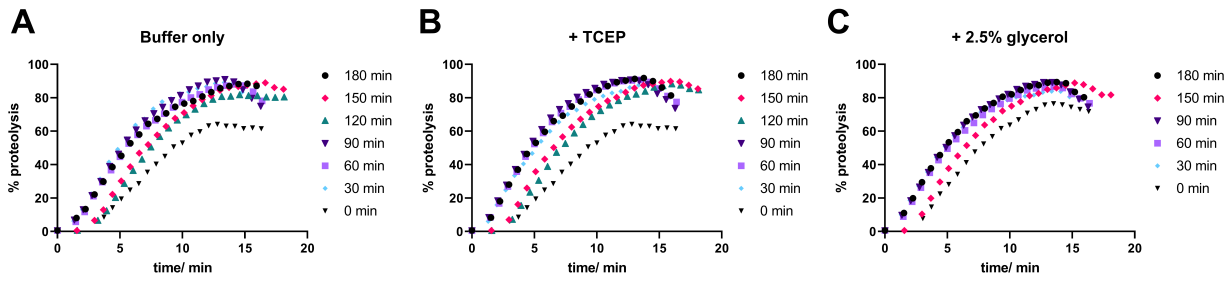


Figure 2.9 | Reducing agent and glycerol do not impact on M^{pro} activity.

The effect of varying the M^{pro} (50 nM - 500 nM) and substrate peptide (2 µM - 10 µM) concentrations was next investigated. Following, titration studies, 150 nM M^{pro} and 2 µM substrate concentrations were selected; these resulted in 30% substrate conversion after 10 min incubation at ambient temperature (Figure 2.10). The relative abundance of the N-terminal product (TSAVLQ, 617 Da) was higher than the C-terminal product peptide (SGFRK-NH_2, 592 Da) at all times studied. The C-terminal cleaved product peptide SGFRK-NH_2 has a basic composition (33%); this may result in poor retention by the C4-solid phase cartridge compared to the N-terminal cleaved product peptide and substrate peptide. Therefore, to calculate percentage product formation or percentage catalysis, only the N-terminal cleaved product peptide was taken into account. Note, the inclusion of the C-terminal product peptide to calculate M^{pro} activity did not significantly differ from M^{pro} activity calculated from the abundance of the N-terminal product peptide alone, as represented in green and magenta in Figure 2.10B, respectively. Therefore, M^{pro} activity was calculated based on the abundance of the s01 substrate and the N-terminal product peptide. The selected reaction conditions for the subsequent SPE-MS assays were 0.15 µM M^{pro}, 2.0 µM substrate, 20 mM HEPES pH 7.5, 50 mM NaCl.

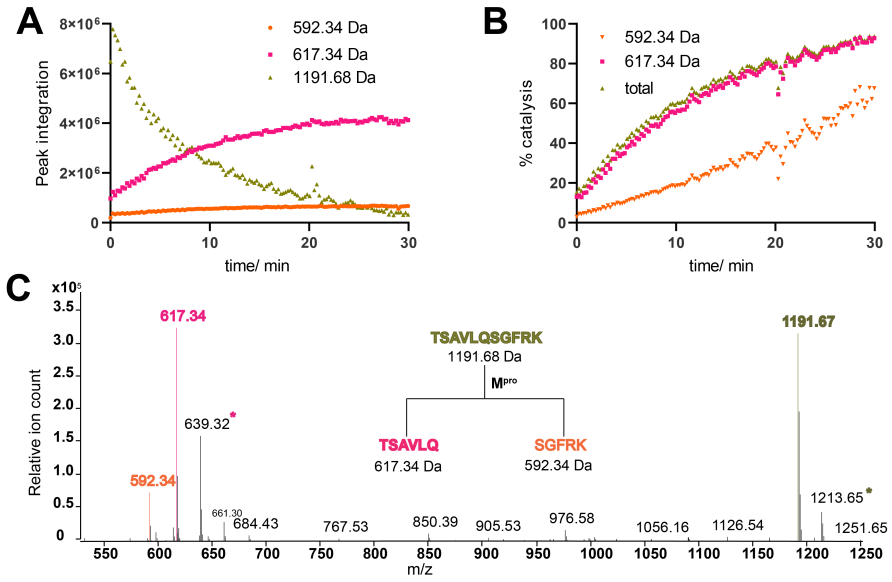


Figure 2.10 | The SPE-MS based assay detects both cleaved products of 11-mer s01 peptide substrate of M^{pro}. **A.**The N-terminal (617.34 Da) and C-terminal fragments (592.34 Da) are observed to increase over time whereas the peptide corresponding to the substrate peptide (1191.68 Da) decreases. **B.** Percentage turnover calculated based on N-terminal product and substrate only (pink), C-terminal product and substrate only (orange), and both the N- and C- terminal products and substrate indicate increased catalysis of the substrate in a time-dependent manner. Linear conversion is observed until ∼10 minutes. **C.**Plot depicting the total ion counts i.e. abundance of product and substrate peptide ions at 10 minutes after reaction initiation.

2.3.4 Solid Phase extraction based MS assay validation

The SPE-MS M^{pro} assay with the 11-mer substrate with $Z' \geq 0.8$ was considered to be suitable for high-throughput applications to screen drugs in repurposing efforts. The assay setup is shown in Figure 2.11A. Two commercial libraries of compounds were selected: 1600 FDA approved compounds and 1280 compounds in the library of pharmacologically active compounds ($LOPAC^{1280}$). The FDA approved compounds include drugs used for antidiabetic, antiseizure, antipsychotic, antibiotic, and antiviral treatments, as well as non-steroidal anti-inflammatory drugs (NSAIDs), and chemotherapeutics [254]. The $LOPAC^{1280}$ library includes drugs for all major enzyme classes including G protein-coupled receptors (GPCRs) and kinases [255].

Compounds were available as 10 mM DMSO stock solutions which were dry dispensed to 384 polypropylene plates using an Echo 550 acoustic dispenser at a fixed concentration of 20 µM. Using a Multidrop Combi liquid handling robot, both the substrates and peptides were dispensed into 384 well

polypropylene plates. Compounds were incubated with the M^{pro} for 30 min, followed by the addition of the substrate and further incubation of 10 min. The reaction was then quenched with 10%$_{(v/v)}$ aqueous formic acid and the samples were analysed by SPE-MS. M^{pro} inhibitors were discovered in both libraries (Figures 2.10B, 2.10C). The assay quality parameter Z′ [256] and the signal-to-noise (S/N) ratio were determined for the five plates that consisted of the compounds from the commercial libraries. Excellent Z′-factors and S/N ratios implied high robustness of the SPE-MS M^{pro} assay (Figures 2.11D, 2.11E).

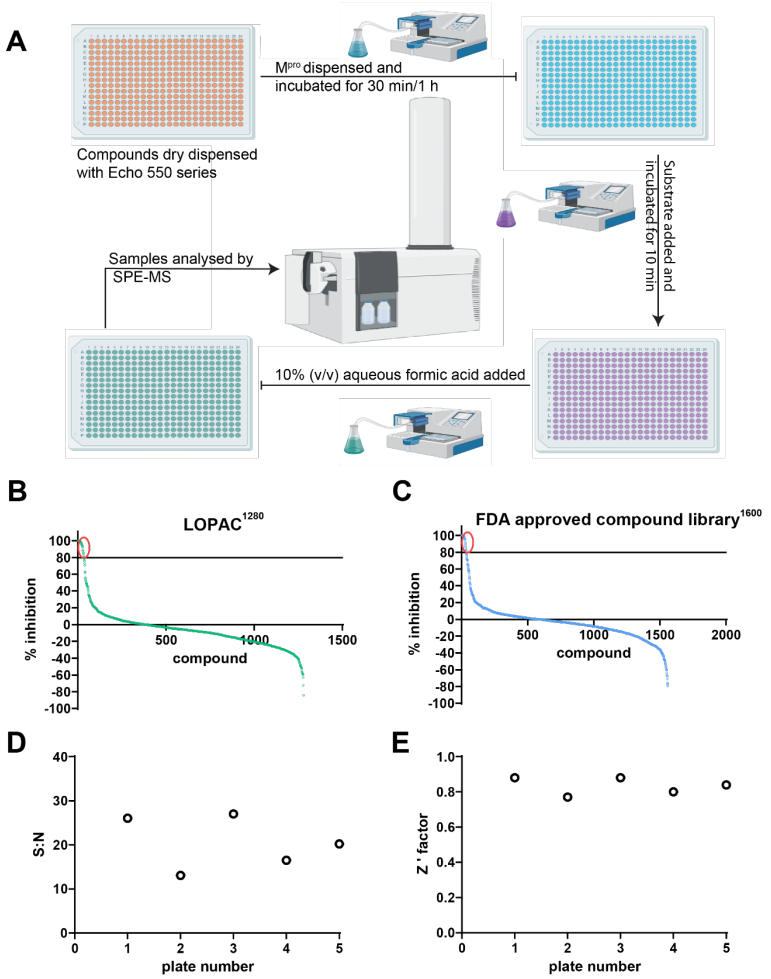


Figure 2.11 | The SPE-MS based M^{pro} assay setup and validation via screening of the FDA approved compounds and the LOPAC library. **A.** The experimental setup of the SPE-MS M^{pro} assay. M^{pro} was profiled against the **B.** library of pharmacologically active compounds ($LOPAC^{1280}$, Sigma-Aldrich), which contains 1280 bioactive small molecules and **C.** library of small molecules approved by the FDA (1600 compounds, Pharmacon) for therapeutic use in humans (provided by the Target Discovery Institute, Oxford). As a measure of assay robustness both **D.** signal/noise (S/N) ratios and **E.**, the assay Z′ factors were calculated for all 5 assay plates according to the literature [256]. Both values indicate that the SPE-MS assay is of high quality and robustness. Full list of compounds that exhibited ≥80% inhibition in $LOPAC^{1280}$ library and library of FDA approved small molecules are shown in Table 2.2 and Table 2.3.

51 compounds from the the $LOPAC^{1280}$ and the FDA approved compound libraries showed ≥80% M^{pro} inhibition (Tables 2.2 and 2.3). In addition to previously reported M^{pro} inhibitors, such as ebselen and disulfiram [28], metal based inhibitors e.g. auranofin, cisplatin, bismuth subsalicylate, and aurothioglucose, the protease inhibitors *N*-carbobenzyloxy-L-phenylanlanyl (Z-L-Phe) chloromethyl ketone, TPCK, TLCK, BAY11-7085 and BAY-11 7082, phosphatase inhibitor NSC 95397 and kinase

inhibitor IPA-3 were identified as potential M^{pro} inhibitors [257–261]. Interestingly, NSC 95937 is also reported to inhibit the activity of another SARS-CoV-2 target, nsp15, an endoribonuclease [262]. Pan assay interference compounds such as pyrocatechol and myricetin with flavonoid groups also inhibited M^{pro} (Tables 2.2, 2.3) [263].

Selected compounds that displayed ≥80% M^{pro} inhibition were then investigated in dose response studies. IC_{50} values ranging from 0.07 µM - 41.5 µM were obtained; disulphide IPA-3, naphthoquinone-based inhibitor NSC95937, protease inhibitors TPCK and TLCK and metal based inhibitor cisplatin, auranofin and bismuth subsalicylate were the most M^{pro} potent inhibitors whereas 5-thioguanine and 8-hydroxy-5-nitroquinoline displayed mild potency in 14.7 µM - 41.5 µM range (Table 2.4, Figure 2.12).

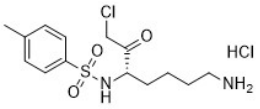
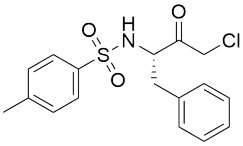
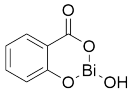
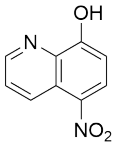
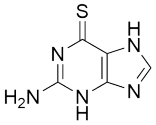
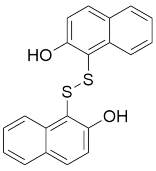
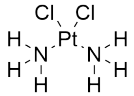
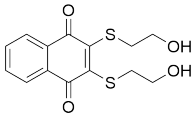
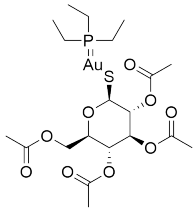
Table 2.2: FDA-approved small molecules for therapeutic use in humans which inhibit ≥80% M^{pro} activity at a fixed concentration as observed by the SPE-MS assay. Conditions: 20 µM compounds, 0.15 µM M^{pro}, 2.0 µM TSAVLQ/SGFRK-NH_2 in the reaction buffer. See the Methods Section for assay details.

Compound	∼% M^{pro} inhibition at 20 µM
Cisplatin	100.0
Sanguinarium chloride	100.0
Pipenzolate bromide	100.0
Oltipraz	99.6
Disulfiram [28]	99.1
Ebselen [28]	99.1
Apomorphine hydrochloride	97.0
Carmustine	96.0
Nitroxoline	96.0
Benzyl isothiocyanate	95.8
Colistin sulfate	95.8
Bismuth subsalicylate	92.9
Thioguanine	90.5
Netilmicin sulfate	89.1
Erythrosine sodium	86.5
Chlorophyllide copper complex sodium salt	86.5
Methazolamide	82.9

Table 2.3: Small molecules from LOPAC^{1280} library which inhibit ⩾80% M^{pro} activity at a fixed concentration as observed by the SPE-MS assay. See Methods Section for assay details.

Compound	∼% M^{pro} inhibition at 20 µM
Auranofin [264]	100.0
Sch-202676 hydrobromide	100.0
(*R*)-(-)-Propylnorapomorphine hydrochloride	100.0
Apomorphine hydrochloride hemihydrate	100.0
N-Carbobenzyloxy-L-phenylanlanyl chloromethyl ketone	100.0
Guanidinyl-naltrindole di-trifluoroacetate	99.6
(*R*)-(-)-2,10,11-Trihydroxyaporphine hybrobromide	99.6
cis-Diammine(pyridine) chloroplatinum(II) chloride	99.6
6-Nitroso-1,2-benzopyrone	99.5
Iodoacetamide	99.3
1,4-Benzoquinone	99.3
Cisplatin	99.2
1,1'-Disulfanediyldinaphthalen-2-ol (IPA-3)	99.1
NSC 95397 [265, 266]	99.0
Sanguinarine chloride	98.7
Myricetin	98.5
(*R*)-(-)-2,10,11-Trihydroxy-n-propylnoraporphine hydrobromide	98.3
Nordihydroguaiaretic acid from *Larrea divaricata*	97.8
Tetramethylthiuram disulphide (disulfiram)	97.5
ZM 39923 hydrochloride	96.9
Aurothioglucose [265]	96.6
Ammonium pyrrolidinedithiocarbamate	95.1
Bay 11-7085	95.0
Capsazepine	94.5
TPCK	94.1
Piceatannol	93.8
Cephalosporin C zinc salt	92.3
Bay 11-7082	92.0
6,7-ADTN hydrobromide	88.3
TLCK	88.3
Pyrocatechol	85.4
Chelerythrine chloride	83.0
Carmustine	86.5
3,4-Dichloroisocoumarin	80.4

Table 2.4: IC_{50} values for selected compounds initially identified as potent M^{pro} inhibitors from $LOPAC^{1280}$ and FDA approved compound libraries. Compounds were preincubated for **A.** 30 min or **B.** 60 min with M^{pro}. Assays were carried out with 150 nM M^{pro} and 2 µM TSAVLQ/SGFRK-NH_2 peptide.

M^{pro} inhibitor	Structure	∼%Inhibition at 20 µM	$\sim IC_{50}(\mu M)^A$	$\sim IC_{50}(\mu M)^B$
TLCK		88.3	1.79	0.94
TPCK		94.1	0.74	0.49
Bismuth(*III*) subsalicylate		92.9	0.29	0.28
8-Hydroxy-5-nitroquinoline		95.9	41.5	17.6
5-Thioguanine		90.5	19.0	14.7
IPA-3		99.1	0.10	0.07
Cisplatin		100.0	1.05	0.66
NSC95397		99.0	0.55	0.38
Auranofin		100.0	1.62	1.58

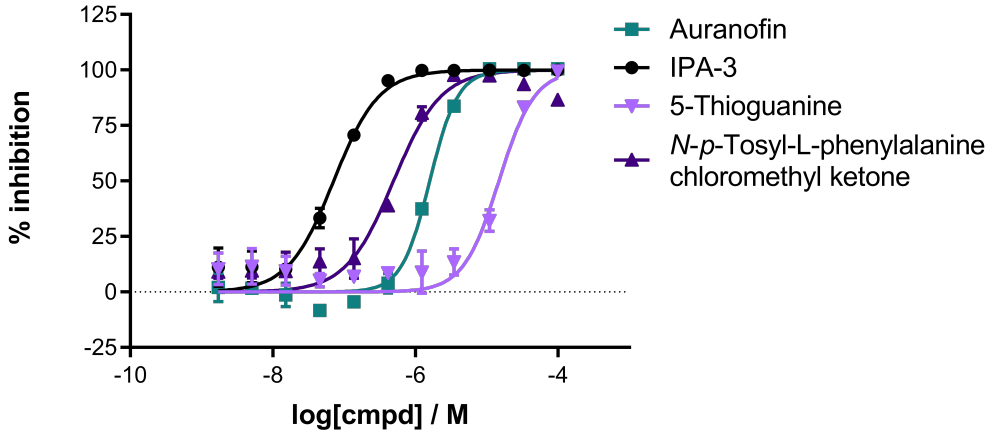


Figure 2.12 | Dose response curves for selected hits from the LOPAC^{1280} library and FDA approved compounds. Dose response curves were generated to determine IC_{50} values for selected hits. The data manifested a range of potencies from sub-micromolar to double-digit micromolar IC_{50} values. The selective p21 activated kinase (PAK) inhibitor (IPA-3), the anti-rheumatoid arthritis agent (auranofin), the purine analog (5-thioguanine), and the cysteine proteinase inhibitor, *N*-p-tosyl-L-phenylalanine chloromethyl ketone are potent M^{pro} inhibitors. See Table 2.4 for IC_{50} values.

2.3.4.1 Assay validation with reported M^{pro} inhibitors

In addition to the hits from the library compounds, reported M^{pro} inhibitors that were commercially available at the time were purchased and screened using the SPE-MS assay. N3 (IC_{50} 0.83 ± 0.47 µM) and ebselen (IC_{50} 0.09 ± 0.07 µM, FRET IC_{50} 0.67 ± 0.09 µM [28]) had comparable potency as determined by SPE-MS and the reported FRET based assay utilising the Mca–AVLQ/SGFRK(Dnp)K substrate (Figure 2.13). By contrast, the SPE-MS assays derived IC_{50} values for boceprevir (IC_{50} 11.0 ± 4.8, FRET IC_{50} 2.70 ± 0.05 µM [77]) and disulfiram (IC_{50} 0.60 ± 0.01 µM, FRET IC_{50} 9.35 ± 0.18 µM [28]) reveal discrepancies between the SPE-MS and the FRET based assay, potentially arising from different substrates in the assays employed.

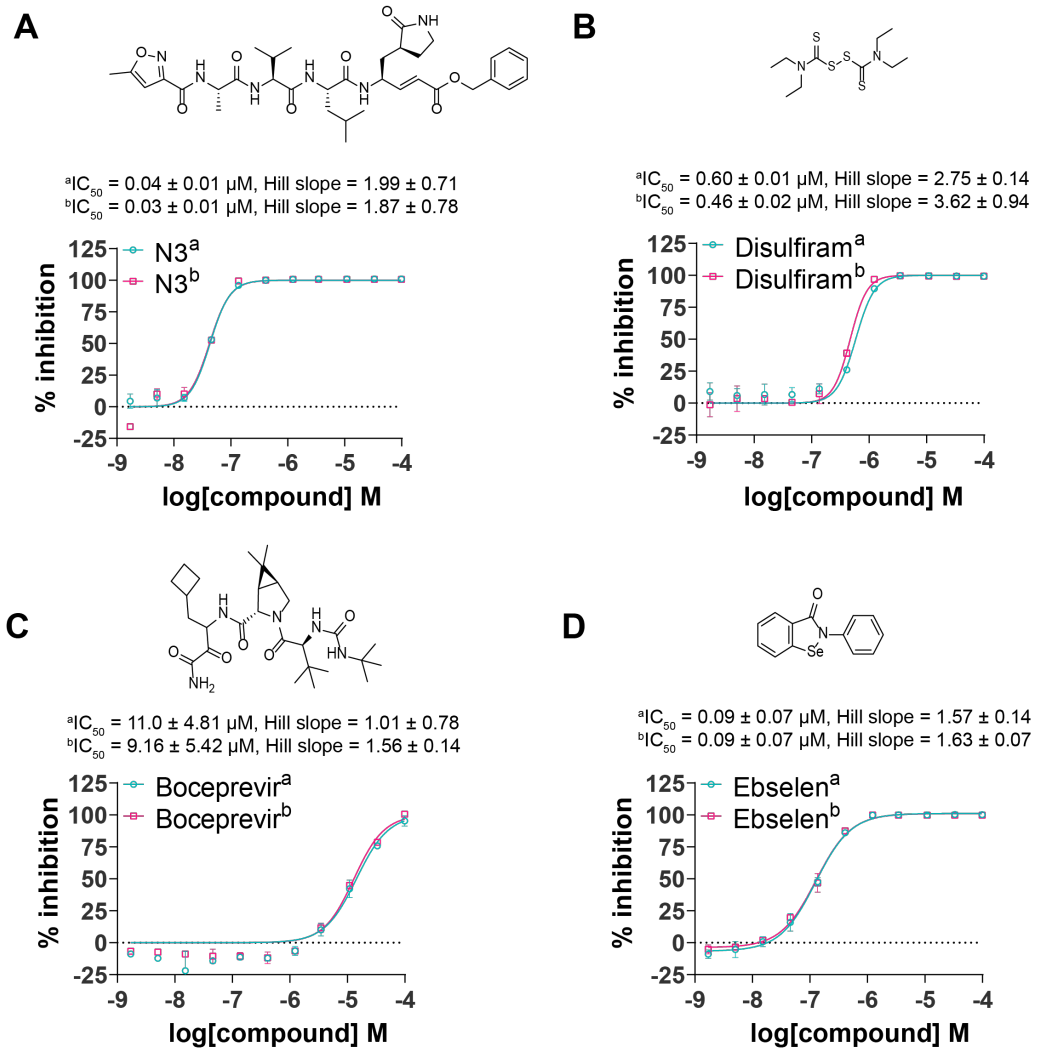


Figure 2.13 | Dose response curves of reported M^{pro} inhibitors determined using SPE-MS. A. N3, **B.** disulfiram, **C.** boceprevir, **D.** ebselen. Each data point represents technical duplicates n = 2 ± SD). Conditions: 0.15 µM M^{pro} and 2.0 µM TSAVLQSGFRK-NH_2 substrate peptide in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl) with ([a]) 30 min compound preincubation (open square, turquoise) and ([b]) 60 min preincubation (open circle, magenta). The structures of the inhibitors are shown.

2.3.4.2 Determination of catalytic coefficients of the 11-mer M^{pro} substrate peptide.

The reported catalytic efficiency (k_{cat}/K_m) for the 10-mer FRET substrate (Mca–AVLQ/SGFR-K(Dnp)K) with SARS-CoV-2 M^{pro} is 28500 $M^{-1}s^{-1}$ [28]. K_m and k_{cat} values were obtained for the 11-mer substrate s01 using the SPE-MS assay; a similar M^{pro} catalytic efficiency, k_{cat}/K_m of 27000 $M^{-1}s^{-1}$ was determined. Non-linear curve fitting using the Michaelis-Menten equation with a lower concentration range(<30 µM) of peptide concentration was poor; 16 points ranging from 0.5 µM to 40 µM were thus taken to improve curve fitting and for the accurate determination of K_m values. Note that M^{pro}

is (partially) monomeric at lower concentrations, and is reported to be inactive as a monomer [240]. The monomer/dimer equilibrium may cause complexity in its kinetics.

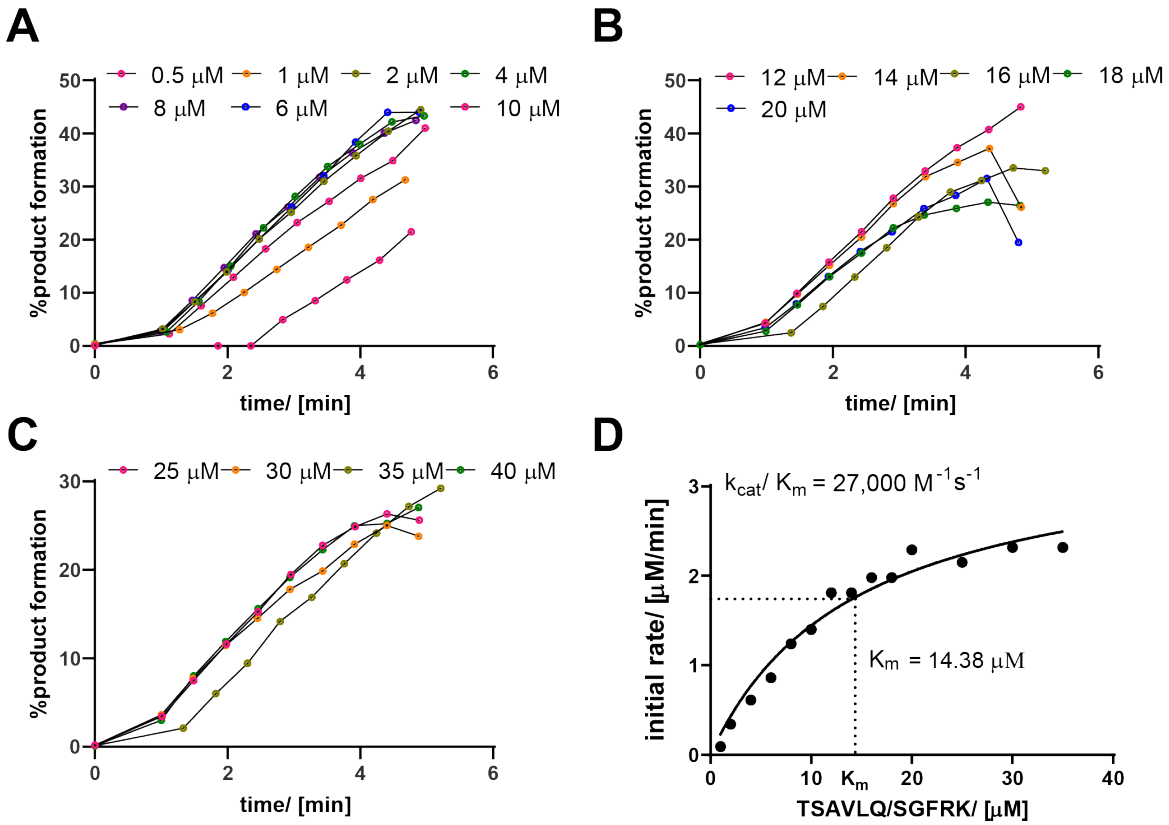


Figure 2.14 | Steady state kinetic analyses for M^{pro} substrate s01 determined using SPE-MS. M^{pro} (0.15 μM) was incubated with different concentrations of the 11-mer TSAVLQ/SGFRK-NH_2 substrate: **A.** 0.5-10 μM, **B.** 12-20 μM, and **C.** 25-40 μM, and the product formation monitored by SPE-MS. **D.** Plot of initial rate versus all substrate concentrations fitted using non-linear regression. Fitting of the data to the Michaelis-Menten equation gave k_{cat} ~23.47 min^{-1} and K_m ~14.38 μM.

2.3.4.3 Determination of equilibrium constants of M^{pro} monomer/dimer.

Using the SPE-MS derived catalytic co-efficient of the M^{pro} for the 11-mer substrate and the optimised M^{pro}: substrate ratio for turnover studies, non-denaturing MS studies were performed in a collaborative effort by the group of Prof. Carol Robinson to investigate the M^{pro} monomer/dimer equilibrium constant [267]. The 11-mer substrate and varying M^{pro} concentrations (0.325 μM - 10 μM) were used to identify the charge states corresponding to monomeric and dimeric forms of M^{pro}. The mole fraction of each species was plotted as a function of total M^{pro} concentration and a K_d was measured as 0.14 ± 0.03 μM (Figure 2.15). Comparison of the measured K_d with a value obtained from a monomer/dimer

equilibrium binding model suggested excellent agreement. However, M^{pro} activity was detected with the FRET substrate at 5, 10, and 50 nM (data not shown), the concentrations at which majority of the M^{pro} would be expected to be monomeric and hence inactive. The results suggest M^{pro} is able to catalyse the hydrolysis of the nsp4/5 derived substrate even with relatively low fractions of the dimeric form.

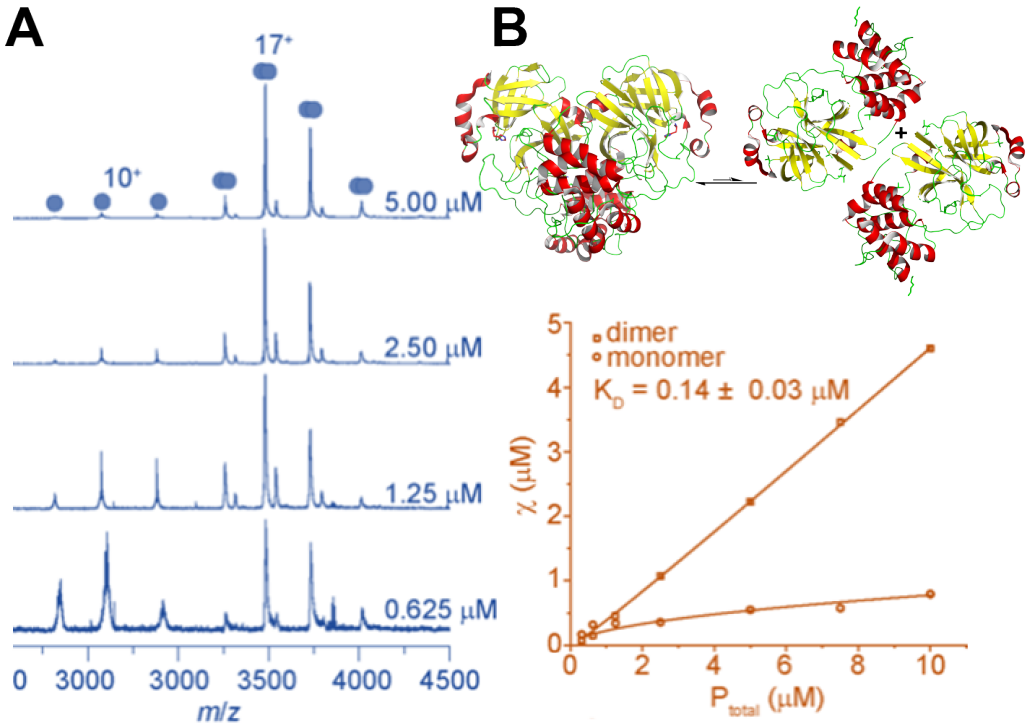


Figure 2.15 | Monomer/ dimer equilibrium constant derivation of M^{pro} using non-denaturing MS. A. Non-denaturing mass spectra for M^{pro} at 0.625 µM - 5 µM, **B.** Representative plot of mole fraction versus M^{pro} concentration to quantify the dissociation constant, K_d. The experiment was performed by Dr. Tarick J El-Baba [267].

2.3.4.4 Mode of M^{pro} inhibition of selected compounds identified from the commercial libraries

MS analyses were used to monitor the covalent modification of M^{pro} by selected hits from the compound libraries. Compounds were incubated with M^{pro} for the indicated time, and samples were then analysed by SPE-MS without quenching the reaction. As M^{pro} has 12 cysteines including the active site Cys145, it was important to investigate if the modification observed were active site-specific or occurred at non-active site cysteines. N3, which is an active site-selective M^{pro} inhibitor was used to investigate the specificity of different drugs via initial incubation with N3, which results in the covalent reaction of the S_γ of the active site Cys145 with the (α, β) unsaturated carbon of the Michael acceptor of N3 (Figure

2.16). On incubation with N3, a dominant peak at 34447 Da was observed, which corresponds to a single +680 Da modification of M^{pro}. If an inhibitor is M^{pro} active site-specific, no major secondary modification of the N3 reacted M^{pro} is expected. In support of the proposal, no further reaction of the N3 reacted M^{pro} (3-fold excess) upon the addition of additional N3 (20-fold excess) was observed as its reaction is selective with the active site Cys145 in Figure 2.16.

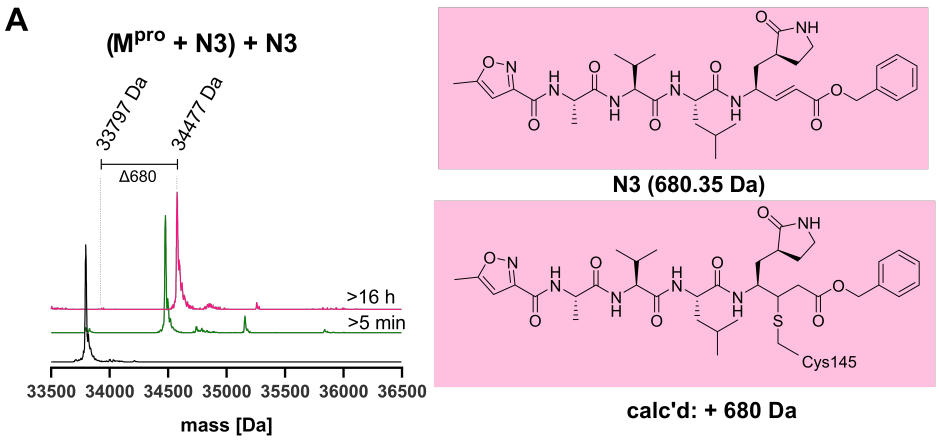


Figure 2.16 | N3 binds M^{pro} through covalent modification of Cys145. SPE-MS binding assays were performed as described in the Methods Section. The reaction of a preincubated M^{pro}: N3 mixture (1:3 ratio on ice) with excess N3 (20 µM) results in one major binding event, likely at the M^{pro} active site Cys-145 as independently shown by crystallographic analysis [28]. The black spectrum is control M^{pro} (33796 ± 1 Da) without an inhibitor. See the Experimental Section for assay details. The structure of the inhibitor and the outcome of the reaction with a nucleophilic cysteine are shown.

By contrast with N3, ebselen which is a known non-specific cysteine-targeting compound, modifies multiple M^{pro} residues (Figure 2.17 A) [177, 268–270]. An ebselen solution was prepared by dilution of the fresh DMSO stock solution in the assay buffer; mixing of this with M^{pro} solution resulted in incomplete modification of M^{pro} by ebselen (Figure 2.17A). The addition of ebselen to the N3:M^{pro} complex resulted in further modification of N3:M^{pro} complex; the mass difference suggested modification of the N3:M^{pro} complex by reaction with two ebselen molecules (Figure 2.17B).

Excess ebselen was expected to modify all 12 M^{pro} cysteine residues. The lack of modification of all M^{pro} cysteines upon treatment with a 20-fold excess of ebselen was possibly due to incomplete solubility of ebselen in assay buffer (Figure 2.17A). Hence, the experiment was repeated by directly

adding a DMSO stock solution of ebselen to the reaction mixture; a complete reaction of ebselen, apparently with all 12 cysteines of M^{pro} was observed (Figure 2.17C). Upon treatment of the N3:M^{pro} complex with ebselen, the mass difference suggested ~11 reactions with ebselen, consistent with the reaction of the active site and other cysteines with N3 (Figure 2.17D).

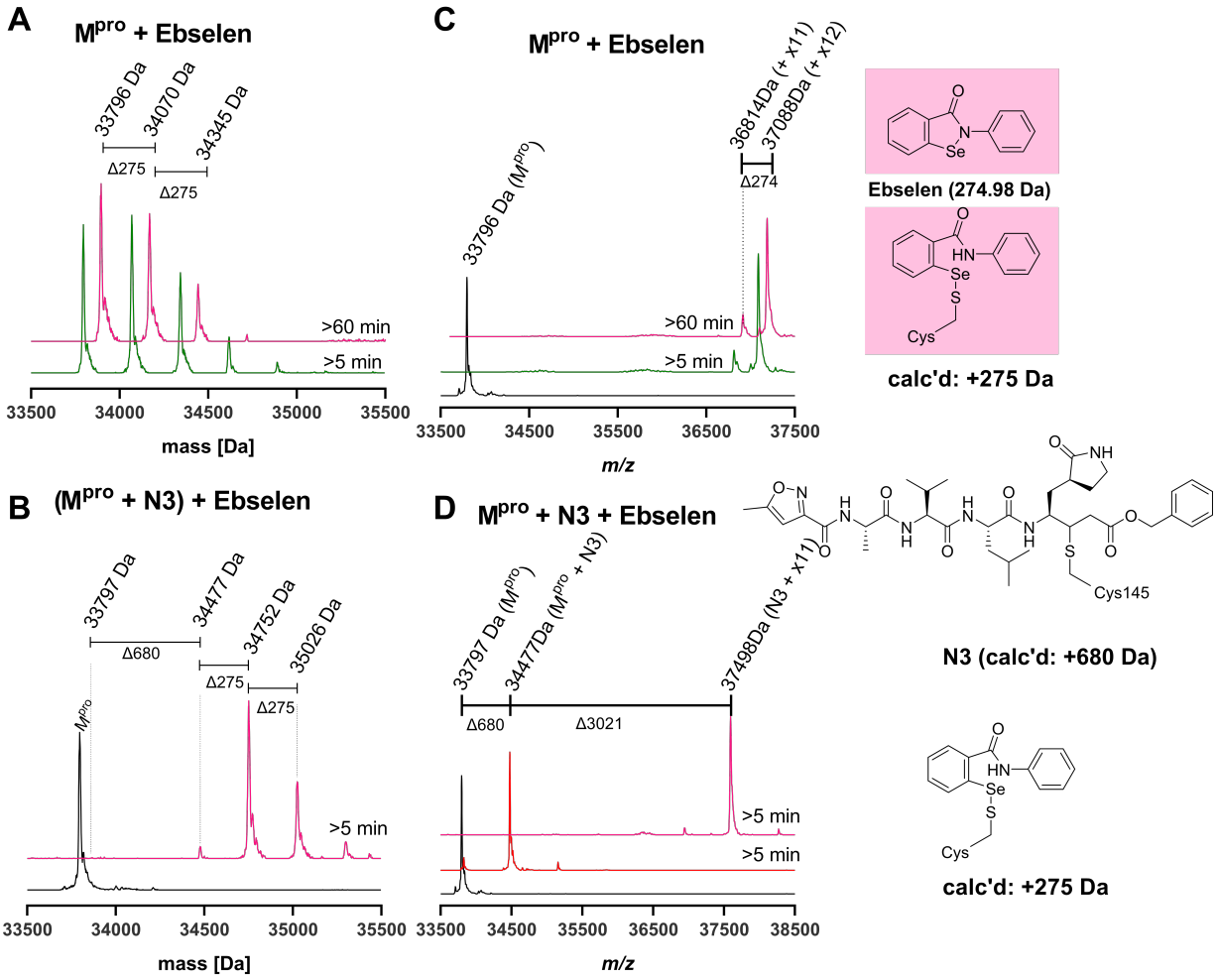


Figure 2.17 | Ebselen reacts covalently with M^{pro}. SPE-MS binding assays were performed as described in the Methods Section using M^{pro} (1 μM) and ebselen (3 or 20 μM). **A.** Reaction with excess ebselen (1:20, M^{pro}:inhibitor) results in more than one reaction with the M^{pro} monomer. **B.** Reaction of a preincubated M^{pro}: N3 complex (1:3 ratio) with excess ebselen (20 μM) results in two major binding event, likely at the non-active site Cys residues. The black spectrum is a control M^{pro} (33796 ± 1 Da) without an inhibitor. **C.** and **D.** are repeats of **A.** and **B.**, respectively, conducted without predilution of the DMSO stock of ebselen with the assay buffer. See the Experimental Section for assay details. Note, the structure of the inhibitor and the outcome of the reaction with the nucleophilic cysteine are shown. Note that the top spectrum in panels A, and B and the top two spectra in panels C, and D were shifted in the X, Y directions.

The reaction of the α-chloroketone protease inhibitors TPCK and TLCK yielded a M^{pro} complex with a mass shift of 315 Da (Figure 2.18A) or 296 Da (Figure 2.18B), respectively, suggesting the S_γ of Cys145 reacts covalently via nucleophilic substitution of the chlorine from the chloroketone warhead. No mass shift of N3: M^{pro} complex peak was observed upon addition of TPCK/TLCK suggesting that TPCK/TLCK are relatively selective inhibitors of the M^{pro} active site Cys145 (Figure 2.18C, 2.18D). By contrast, BAY-11 70852 and NSC 95397 react with M^{pro} at multiple sites (Figure 2.19A, 2.20A). BAY-11 7082 has both vinyl sulfone and nitrile potential warheads. The mass difference of 51 Da between M^{pro} and BAY-11 7082 reacted M^{pro} suggests that the reaction proceeds via initial nucleophilic attack by S_γ at the α-β unsaturated carbon 1 of BAY11-7082 (Figure 2.19A) with M^{pro}, followed by elimination of the methyl-sulfonyl benzyl group. With BAY-11 7082, 11 peaks corresponding to 51 Da incremental mass shifts were observed (Figure 2.19A). Note that in contrast to the results with M^{pro}, a Michael reaction at the C2 position of BAY 11-7082 is reported for tyrosine phosphatases inhibition [260]. The multiple reactions of BAY 11-7082 suggest poor selectivity for the M^{pro} Cys145, and potentially other nucleophilic cysteine proteases.

Unlike the multiple peaks obtained upon the reaction of excess BAY-11 7082 with M^{pro} (Figure 2.19A), the reaction of excess BAY-11 7082 with N3 reacted M^{pro}, results in relatively fewer BAY-11 7082 modified N3:M^{pro} peaks with Δ52 Da mass shifts(Figure 2.18B). Although a preliminary observation, this suggests the initial reaction with active site Cys145 may be important in promoting subsequent reactions of BAY-11 7082 (Figure 2.19).

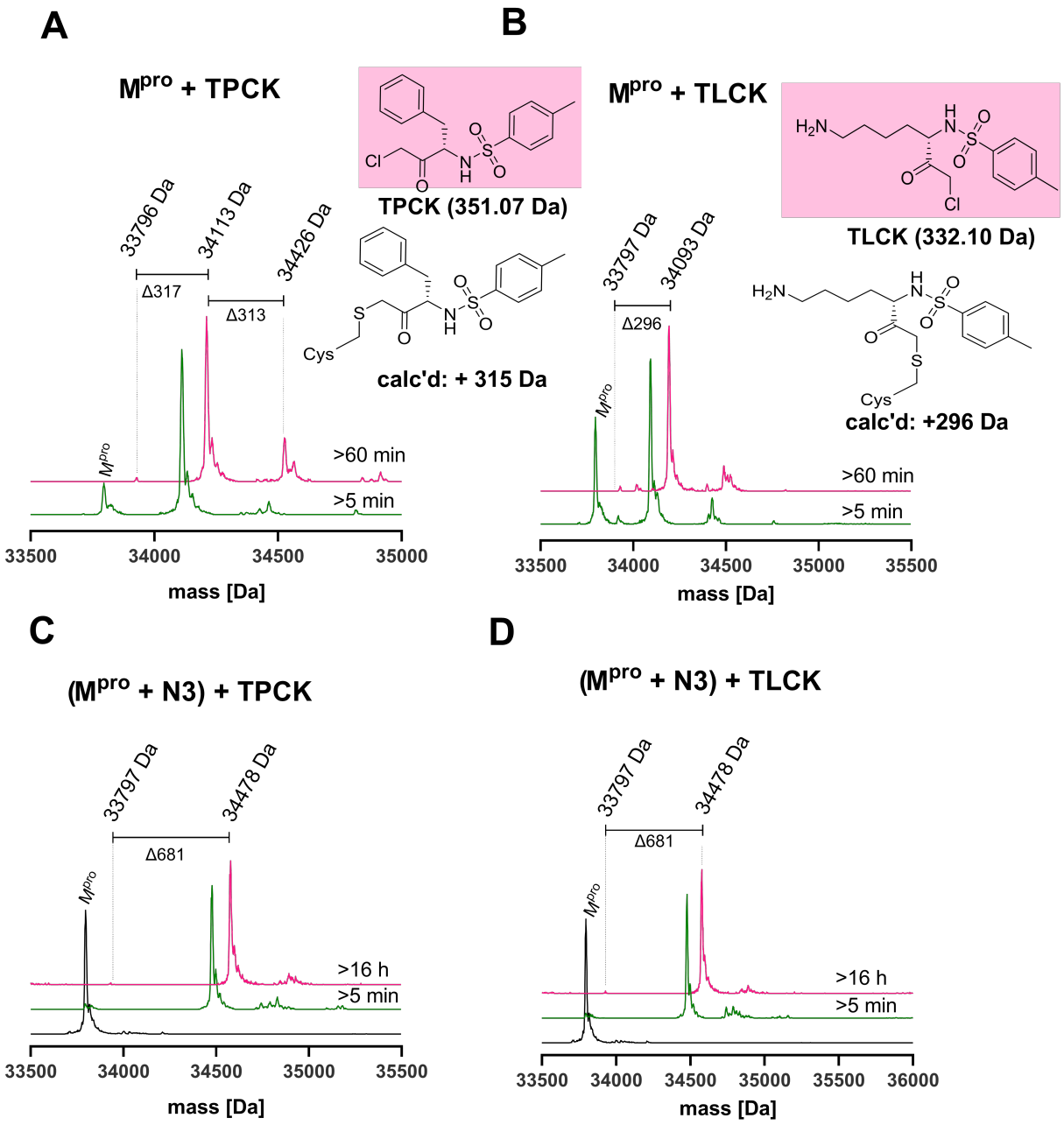


Figure 2.18 | *N*-*p*-Toluenesulfonyl-L-phenylalanine chloromethyl ketone hydrochloride or N_α-*p*-toluenesulfonyl-L-lysine chloromethyl ketone hydrochloride covalently reacts with M^{pro}. Reaction with excess (1:20, M^{pro}:inhibitor) **A.** *N*-*p*-toluenesulfonyl-L-phenylalanine chloromethyl ketone hydrochloride (TPCK) or **B.** N_α-*p*-toluenesulfonyl-L-lysine chloromethyl ketone hydrochloride (TPCK) results in a single modification with the M^{pro} monomer (TPCK, Δ= 315 Da, TLCK, Δ= 298 Da). **C.** Reaction of a preincubated M^{pro}:N3 mixture (1:3 ratio) with excess **C.** *N*-*p*-toluenesulfonyl-L-phenylalanine chloromethyl ketone hydrochloride or **D.** N_α-*p*-toluenesulfonyl-L-lysine chloromethyl ketone hydrochloride (20 μM) results in no further modification of the N3:M^{pro} complex. The results indicate that TPCK binds selectively to the M^{pro} active site. The black spectrum is control M^{pro} (33796 ± 1 Da) without an inhibitor. The structure of the inhibitor and the proposed outcome of the reaction with a nucleophilic cysteine are shown. Note that top spectrum in panels A,B and the top two spectra in panels C, and D were shifted in the X,Y directions.

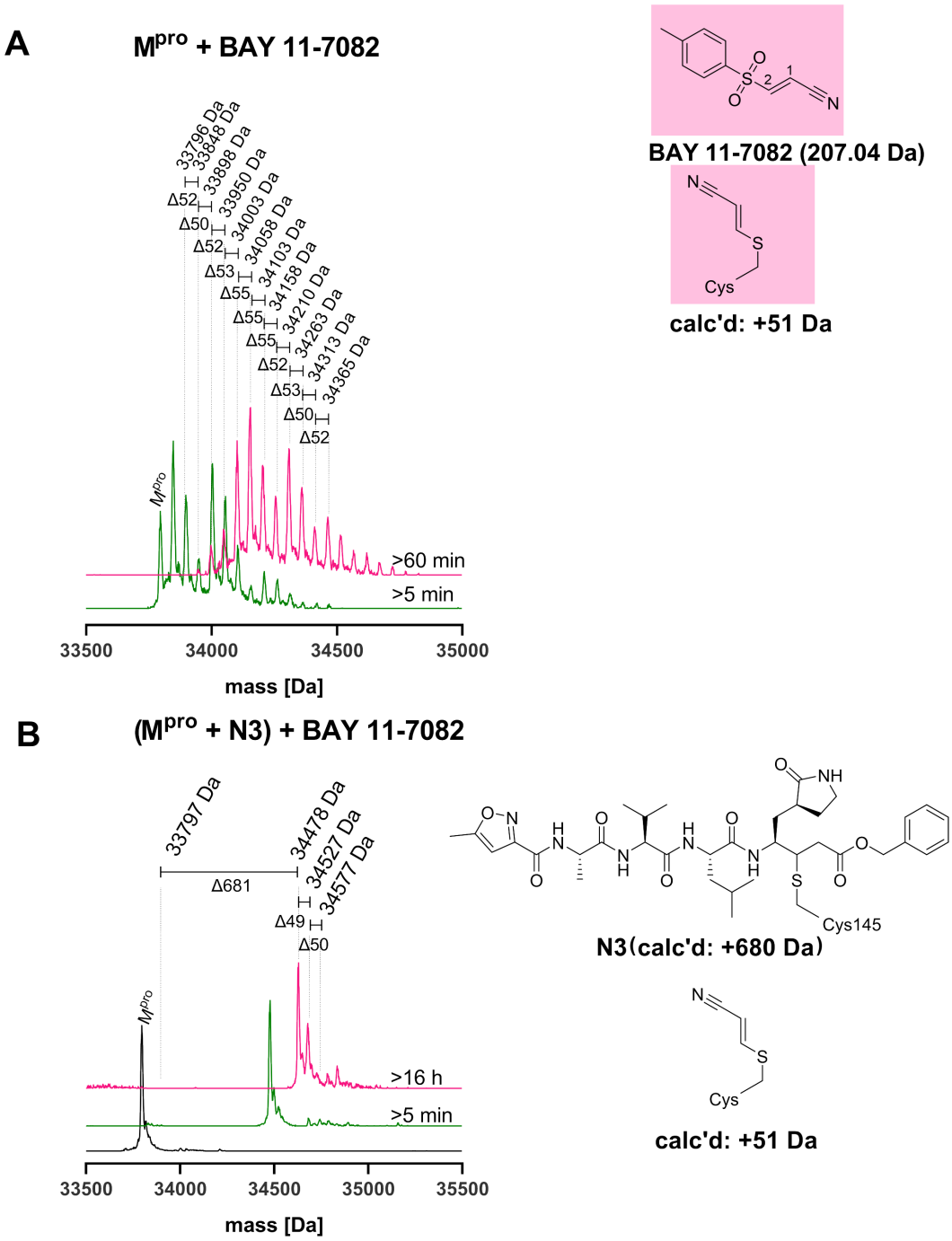


Figure 2.19 | BAY 11-7082 reacts covalently with M^{pro}. BAY 11-7082 reacts with M^{pro} at multiple positions, possibly by Michael addition of the α-β -unsaturated carbons at position 1 followed by an elimination reaction, although in a non-specific manner. Conditions: 1 µM M^{pro}, 20 µM BAY 11-7082 in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). **B.** BAY 11-7082 (in part) reacts with M^{pro} in the presence of N3, although to a lesser extent compared to active site free M^{pro} suggesting it reacts at the active site. Conditions: 1 µM M^{pro} preincubated with 3 µM N3, 20 µM BAY 11-7082 in the reaction buffer. The black spectrum is M^{pro} (33796 ± 1 Da) without an inhibitor. The structure of the inhibitor and a proposed outcome of the reaction with the M^{pro} nucleophilic cysteine are shown.

NSC 95397 apparently reacts with M^{pro} at multiple positions, possibly by Michael addition-elimination reactions (232 Da, Figure 2.20A). In addition to the 232 Da mass shift, peaks corresponding to 152 Da mass shifts were also observed with NSC 95397 reacted M^{pro}. The latter may correspond to the formation of disulfides with 2-mercaptoethanols. Disulfide formation with Cys residues of M^{pro} was also observed with IPA-3. Thiol disulfide exchange between IPA-3 and M^{pro} results in multiple modifications of M^{pro} with mass shifts of 174 Da, possibly corresponding to the disulfide adduct (Figure 2.21A).

As with the observation of reduced modifications of N3:M^{pro} complex with excess BAY-11 7082, treatment with excess NSC 95397 or IPA-3 resulted in a lower number of peaks than with M^{pro} alone (Figures 2.20B, 2.21B), suggesting that the formation of multiple modifications may be promoted by conformational changes following initial reaction with the active site Cys145.

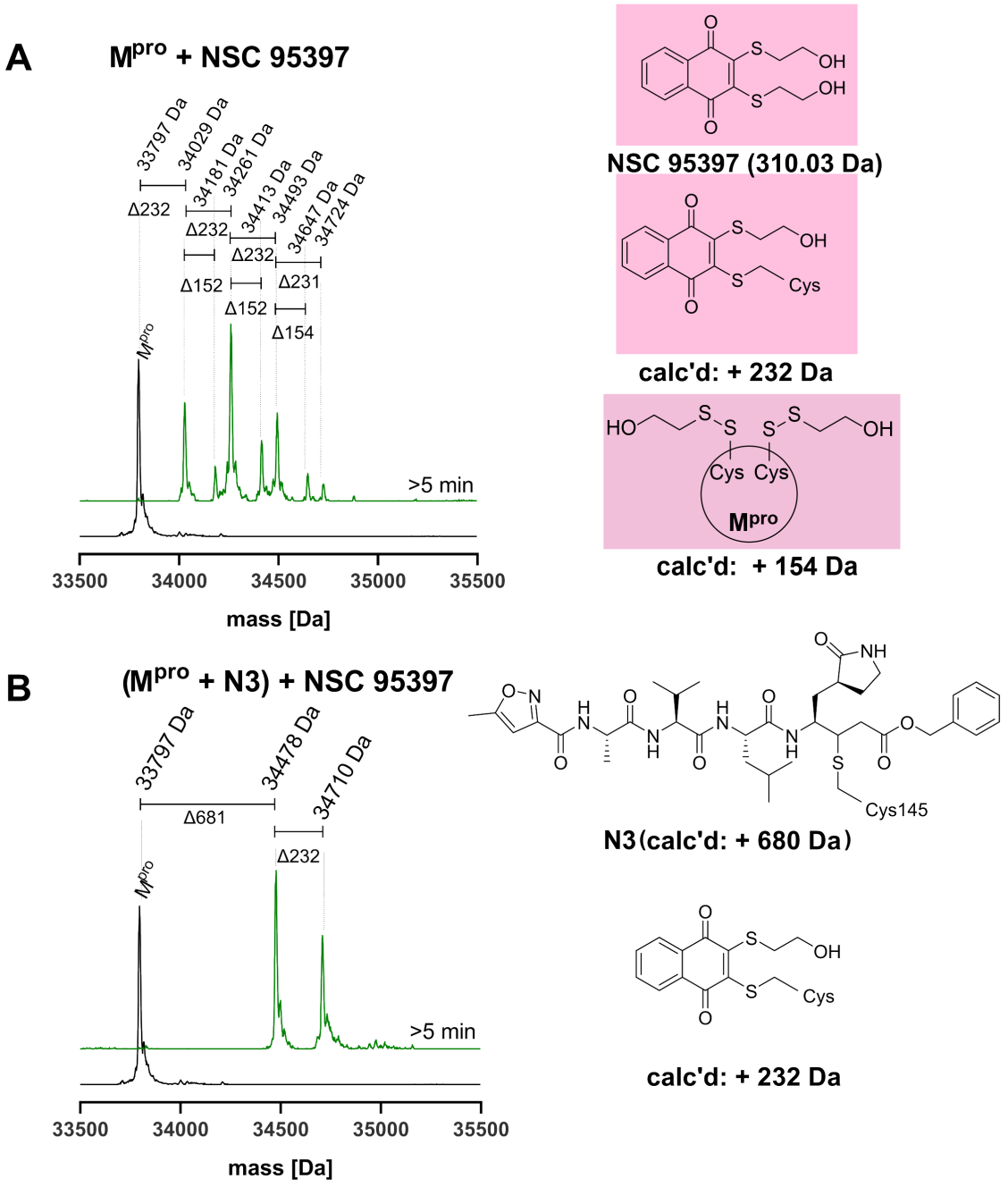


Figure 2.20 | NSC 95397 reacts covalently with M^{pro}. A. NSC 95397 reacts with M^{pro} at multiple positions. Conditions: 1 µM M^{pro}, 20 µM NSC 95397 in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). **B.** NSC 95397 reacts with M^{pro} in the presence of N3 suggesting NSC 95397 does not selectively bind to the M^{pro} active site Cys-145 residue. Conditions: 1 µM M^{pro} preincubated with 3 µM N3, 20 µM NSC 95397 in the reaction buffer. The black spectrum is M^{pro} (33796 ± 1 Da) without an inhibitor. The structure of the inhibitor and a proposed outcome of the reaction with the M^{pro} nucleophilic cysteine are shown.

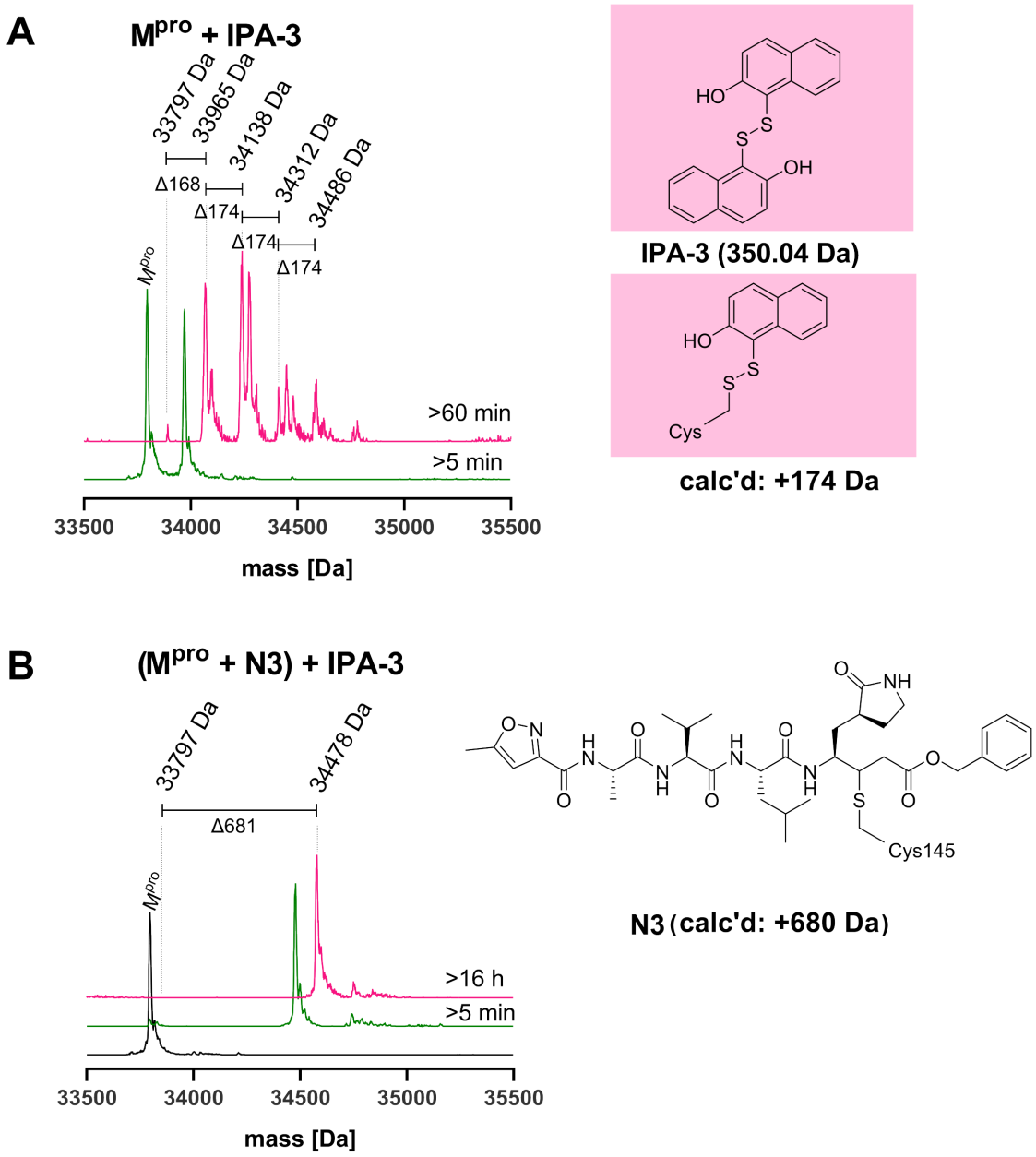


Figure 2.21 | IPA-3 reacts covalently with M^{pro}. A. IPA-3 reacts with M^{pro} at multiple positions. Conditions: 1 µM M^{pro}, 20 µM IPA-3 in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). **B.** IPA-3 does not (at least substantially) covalently modify M^{pro} in the presence of N3 suggesting that IPA-3 modifies the active site. Conditions: 1 µM M^{pro} preincubated with 3 µM N3, 20 µM IPA-3 in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). The black spectrum is M^{pro} (33796 ± 1 Da) without an inhibitor. The structure of the inhibitor and a proposed outcome of the reaction with the M^{pro} nucleophilic cysteine are shown.

Bismuth subsalicylate and 5-thioguanine were not observed to react covalently with M^{pro} to give a stable complex (Figures 2.22A, 2.22C). So, the inhibition by these compounds is likely non-covalent

or reversible in nature as shown by the absence of modification of M^{pro} only peak at 33797 Da. Excess treatment of the N3:M^{pro} complex with these drugs also did not afford any new peaks (Figures 2.22C, 2.22D). Notably, non-specific interaction with excess 5-thioguanine was observed with the N3:M^{pro} complex after prolonged incubation of >16 h (Figure 2.22D).

Overall, screening of the commercial libraries identified covalent and non-covalent, and selective and non-selective M^{pro} inhibitors.

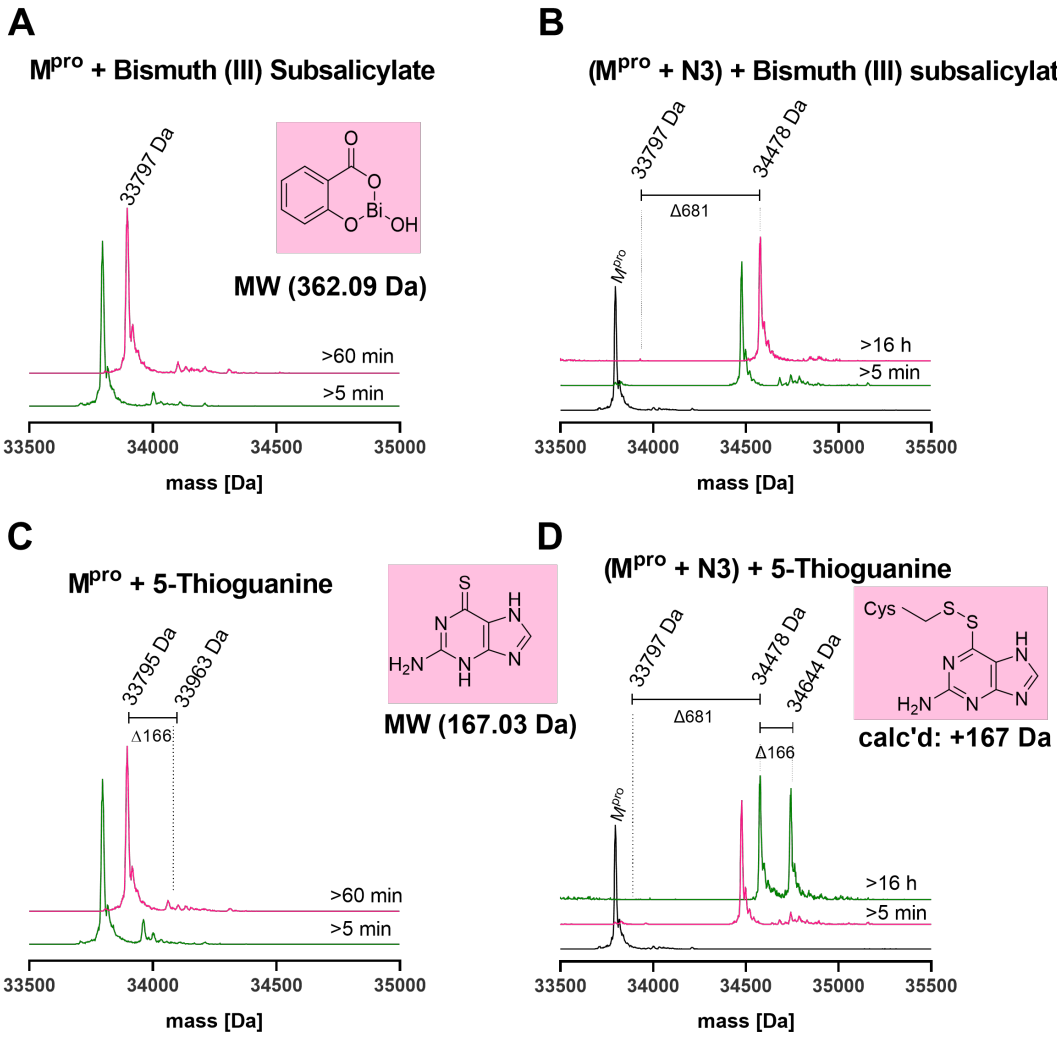


Figure 2.22 | Evidence for non-covalent inhibitors of M^{pro}. A. Bismuth subsalicylate and **C.** 5-thioguanine are non-covalent M^{pro} inhibitors. Conditions: 1 µM M^{pro}, 20 µM IPA-3 in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). **B.** Bismuth(III) subsalicylate does not modify N3 reacted M^{pro}. **D.** 5-Thioguanine does not modify N3 reacted M^{pro}. A 166 Da adduct corresponding to the m/z difference of 5-thioguanine was observed after >16 h incubation with N3 reacted M^{pro} suggesting non-specific covalent labelling of M^{pro} cysteines by 5-thioguanine after prolonged incubation. Conditions: 1 µM M^{pro} preincubated with 3 µM N3, 20 µM IPA-3 in the reaction buffer. The black spectrum is M^{pro} (33796 ± 1 Da) without an inhibitor. The structure of the inhibitor and a proposed outcome of the reaction with the M^{pro} nucleophilic cysteine are shown.

2.3.5 Screening of commercial or synthetic β-lactam and non-β-lactam compounds for M^{pro} inhibition

After validating the SPE-MS M^{pro} assay via screening of library compounds, screening of in-house and commercial compounds was undertaken with the aim of identifying novel covalent warheads

for M^{pro} inhibitor development. Due to the restricted access to an Echo 550 acoustic dispenser (LabCyte) for acoustic dry dispensing of compounds during the lockdown period, efforts were made to set up the assay in-house via the use of a CyBio-liquid handing robot for compound transfer of DMSO stock solutions to 384 well plates (see Methods Section for details). Compounds belonging to different classes such as protease inhibitors, peptides, antibiotics, and others were included at three fixed concentrations (20 µM, 50 µM, and 100 µM) (Figure 2.23).

A number of β-lactams were observed to inhibit M^{pro} activity (Figure 2.23A). The cephalosporin C zinc salt was the most potent identified inhibitor with >98% M^{pro} inhibition at 20 µM followed by cefaclor which manifested ∼80% M^{pro} inhibition at 20 µM. Cephalothin, cefdinir-C1, cefbuperazone, cephalexin, and cefoxitin inhibited M^{pro} at 20 µM and 50 µM concentrations. Tebipenem inhibited M^{pro} at all the tested fixed concentrations, i.e. >50% M^{pro} inhibition at 20 µM, >77% M^{pro} inhibition at 50 µM and >98% M^{pro} inhibition at 100 µM. Avibactam, cefmetazole, cefditoren, penicillin G, sulbactam, metampicillin, and ezetimibe inhibited M^{pro}, at least at one of the high concentrations. By contrast, no significant M^{pro} inhibition was detected from the set of γ-lactam compounds tested. At 100 µM, γ-lactam **11**, *N*-toluenesulfonyl (tosyl)-γ-lactam **12** and **51** and γ-lactam ester derivative **18** showed mild inhibition (Figures 2.23B, 2.24). The screened synthetic penicillin C3-esters **32**-**44** and penicillin G benzyl ester **45** were also poor inhibitors of M^{pro} (Figure 2.23B, 2.24). However, notable inhibition of M^{pro} activity was observed with commercially available penicillin esters, i.e. bacampicillin, pivampicillin and talampicillin. Likewise, the C3-ester of penicillin V (**46^1**) also displayed 80% M^{pro} inhibition at 100 µM. Penicillin V ester (**46^2**) from a newly synthesized batch was also screened (Figure 2.23C); these did not display any inhibitory activity, this discrepancy will be discussed in Section 2.3.5.8.

Protease inhibitors such as antipain which is reported to inhibit both serine and cysteine proteases, and has an Arg residue at the P1, Val residue at the P2 equivalent position manifested some M^{pro} inhibition [271]. The HIV protease inhibitors, nelfinavir, and ritonavir have M^{pro} inhibitory activity [60], but indinavir, only manifested weak M^{pro} inhibition. In this regard, it is of interest that one of the substrates of aspartyl HIV proteases is composed of RPGNF/LQSRP residues, where the P1′-P3′ residues (i.e. LQS) correspond to the P2-P1′ nsp4/5 cleavage site of M^{pro}, respectively [272]. The

presence of the apparently overlapping substrate preference between HIV aspartyl protease and M^{pro} may reflect the activity of some of the HIV protease inhibitors with M^{pro}. By contrast, inhibitors such as simprenavir and asunaprevir of the HCV serine protease (which has specificity for a P1 Cys and a P3 Val), did not inhibit M^{pro} activity [273]. The side chain of Met is accommodated in the S1 subsites of both HCV protease and M^{pro} (though less preferentially than Cys and Gln, respectively) [86, 273]. Like the HCV protease, Val is also the preferred P3 equivalent residue in M^{pro} indicating tolerance of similar residues in the P3 and P1 substrate positions of the HCV protease and M^{pro}. The overlapping selectivity, at least at the P1 and P3 equivalent positions, of HCV protease and M^{pro} suggest that repurposing of inhibitors of HCV proteases for M^{pro} inhibition may be possible. As such, boceprevir, which is a potent HCV protease inhibitor, also manifests good inhibition of M^{pro}.

A trypsin protease inhibitor isolated from soya bean, which is a disulfide rich 181 aa residues long peptide, did not inhibit M^{pro} (Figure 2.23D). The P1 residue of trypsin protease inhibitors is usually a positively charged residue to interact with the negatively charged residue (Asp189) in the S1 subsite of trypsin [274], consistent with the lack of M^{pro} inhibition from this inhibitor. However, the aldehyde-containing protease inhibitor antipain, which has a positively charged Arg at the P1 position moderately inhibits M^{pro}, so it is likely that, unlike antipain, the bulky trypsin protease inhibitor may not be able to access the active site of M^{pro}.

Overall, these results suggest that careful consideration of substrate specificity between proteases to select peptidomimetic inhibitors containing the overlapping substrate residues/scaffolds between the proteases may enable successful repurposing of viral protease inhibitors or the identification of lead compounds to optimise as M^{pro} inhibitors.

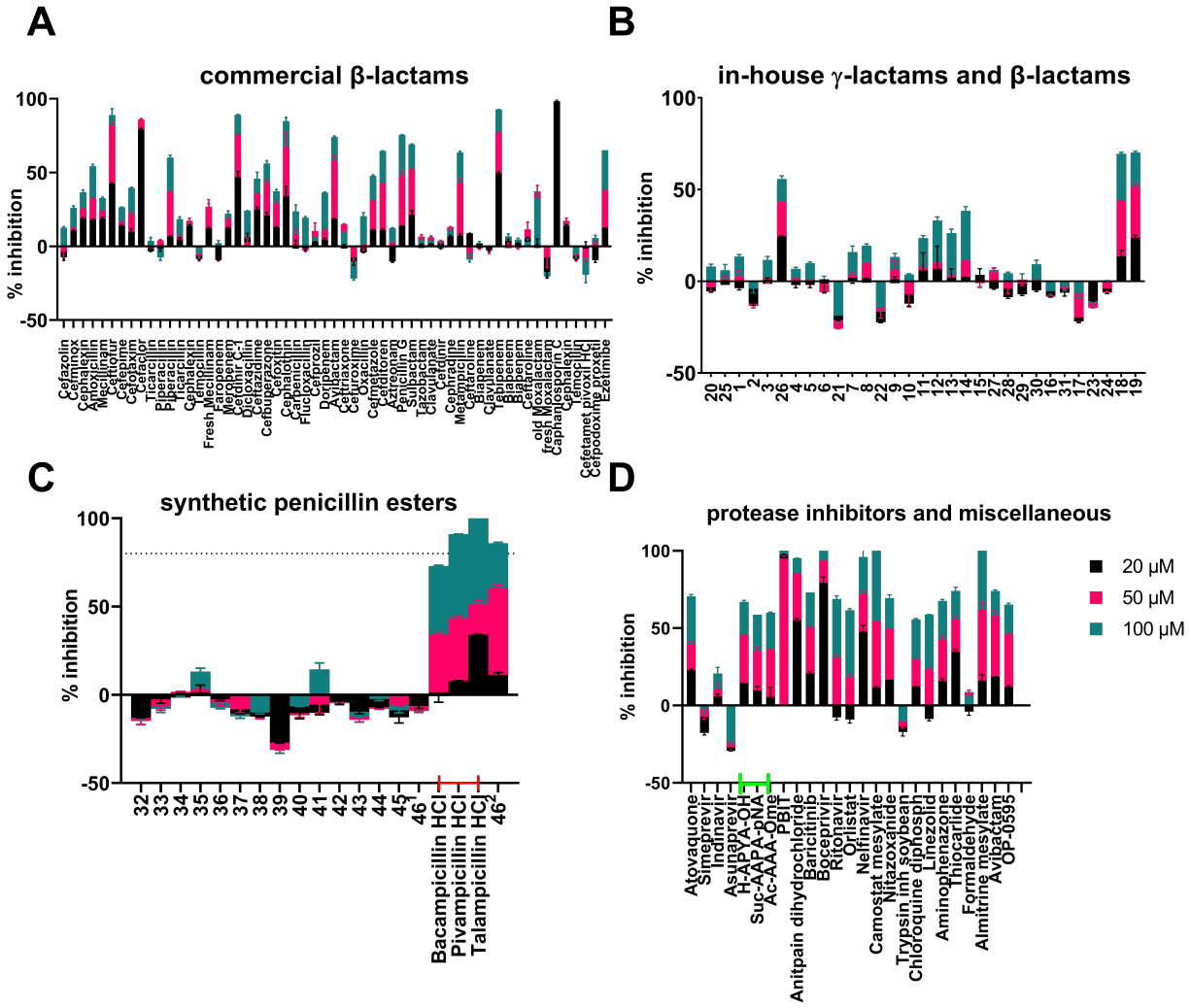


Figure 2.23 | Assaying selected compounds for M^{pro} inhibition. A. β-lactams, **B.** γ-lactams and other β-lactam compounds, **C.** β-lactam penicillin C3-esters and **D.** protease inhibitors and miscellaneous compounds were screened at fixed concentrations of 20 μM, 50 μM and 100 μM. Structures of synthetic compounds belonging to panels B and C are in Figure 2.24 and 2.25, respectively. Penicillin esters (red) and short peptides (green) were included in the screen. [1] and [2] represent the different synthetic batches of penicillin V ester (**46**). Inhibition values were calculated from the mean of technical duplicates (n=2 ± SD). The dashed line in panel C represents the 80% M^{pro} inhibition. Note penicillin V esters [1,2] from different synthetic batches were screened.

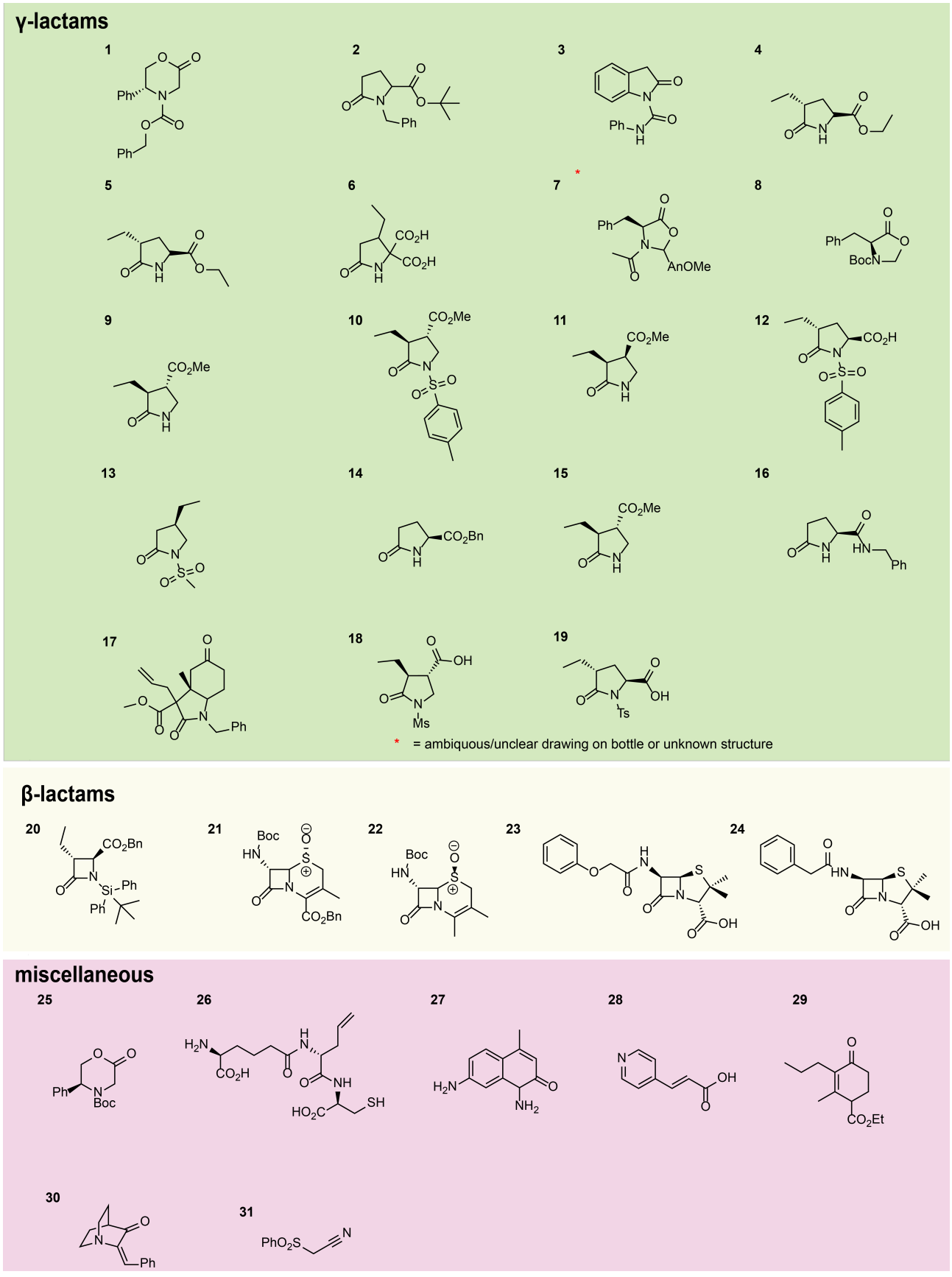


Figure 2.24 | Structure of compounds relating to Figure 2.23B. Compounds belong to γ-lactam (green), β-lactam (yellow), and miscellaneous group (pink): nitrile or reactive vinyl warhead containing or non-covalently reacting compounds.

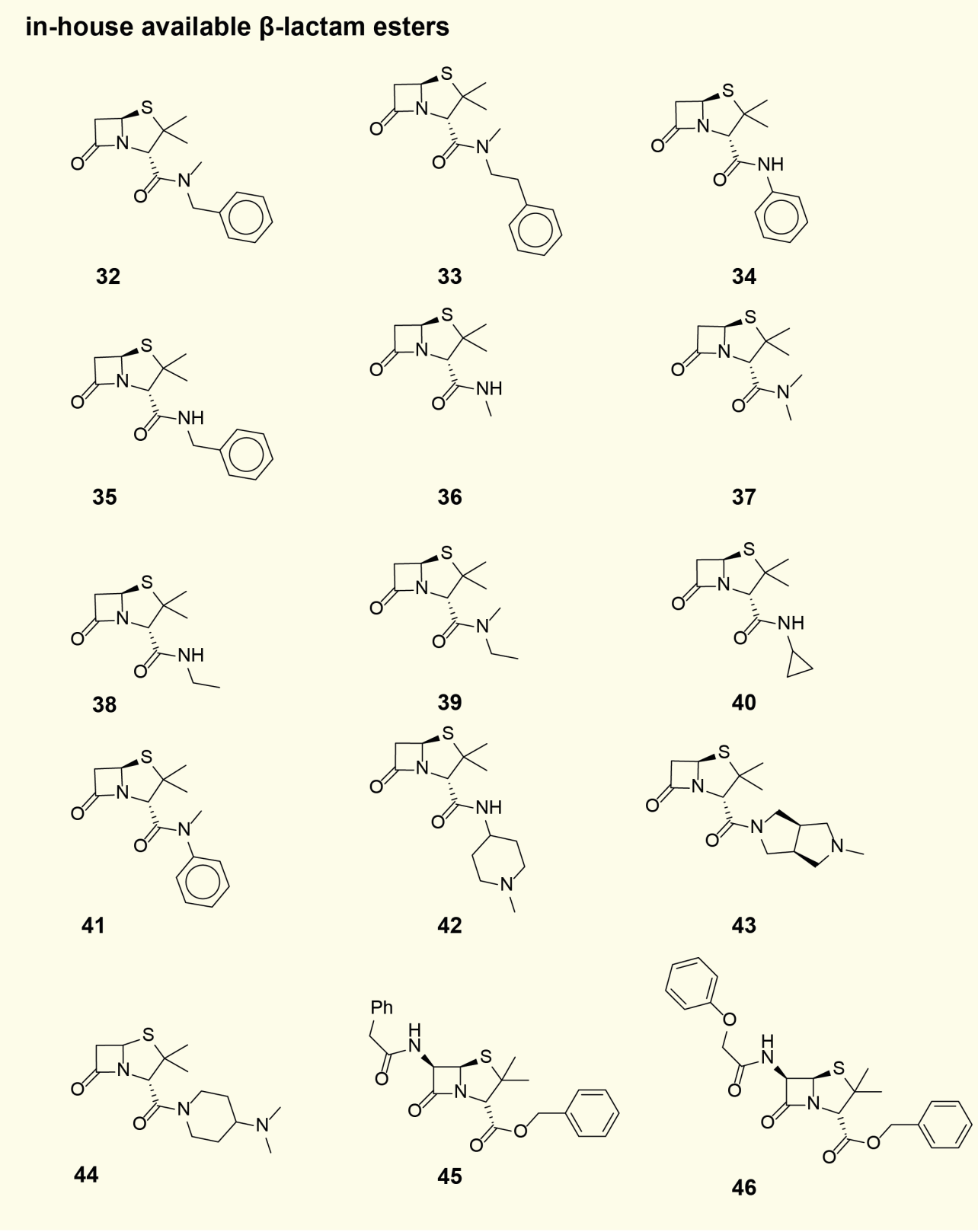
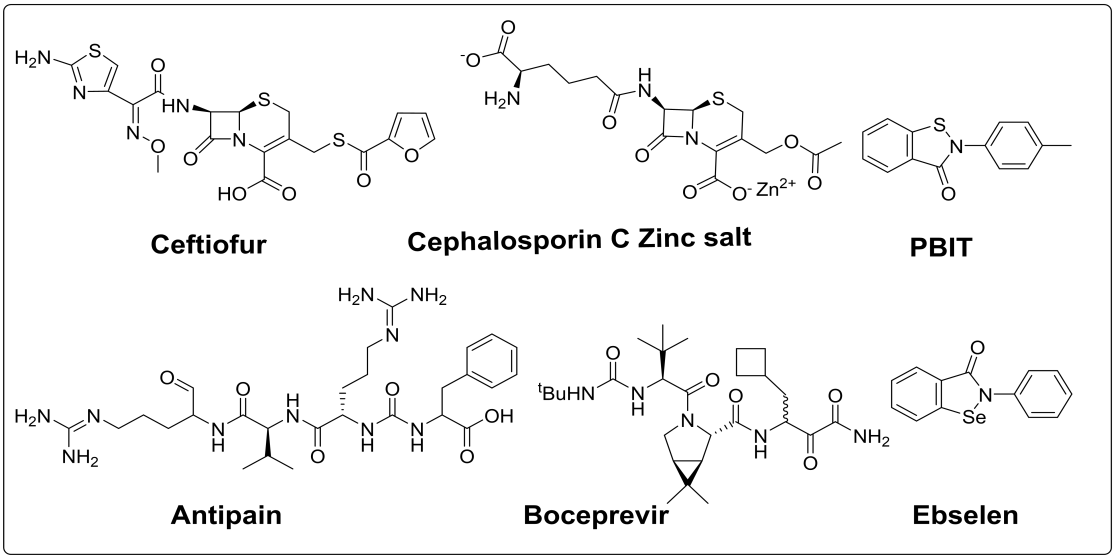


Figure 2.25 | Structures of compounds relating to Figure 2.23C. C3-Esters of β-lactam compounds are shown.

2.3.5.1 Grouping of screened compounds based on percentage inhibition

PBIT, ebselen, cephalosporin C zinc salt, and boceprevir displayed >80% M^{pro} inhibition at 20 µM. ∼80% inhibition of M^{pro} activity was observed with cefaclor at 20 µM. Other cephalosporins, such as ceftiofur, cefdinir C-1, and cephalothin also displayed >80% inhibition, at least at one of the high concentrations tested. Penicillins such as pivampicillin, talampicillin, and penicillin V benzyl ester, i.e. **46** also displayed relatively high inhibition at 100 µM. Compounds that displayed >80% inhibition at any one of the concentrations are shown in Figure 2.26.



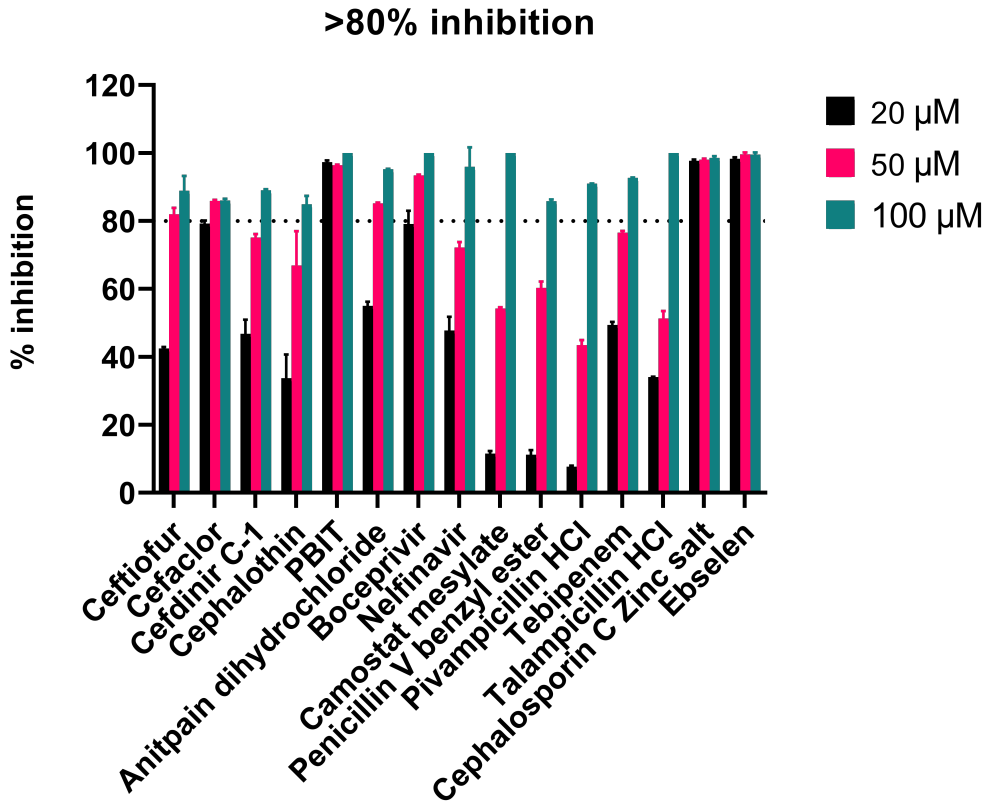


Figure 2.26 | Compounds identified with >80% inhibition of M^{pro} from the in-house screening campaign. Compounds were screened at fixed concentrations of 20 µM, 50 µM, and 100 µM. The structure of some of the inhibitors is shown.

Other penicillins (amoxicillin, penicillin G, metampicllin, and bacampicillin), cephalosporins (cefbuperazone, ceftidoren), monobactam (sulbactam), and protease inhibitors (ritonavir, baricitinib) displayed weak M^{pro} inhibition (Figure 2.27). Surprisingly, an N-terminal acetylated tripeptide (AAA

peptide) with O-methoxy C-terminal residue (Ac-AAA-Ome) and two tetrapeptides (N-terminal succinyl capped AAPA-*p*-nitroanilide (Suc-AAPA-pNA) and H-APYA-OH) also displayed weak apparent M^{pro} inhibition (Figure 2.27). Compound **10**, with an allyl group and free thiol group, also displayed >50% M^{pro} inhibition. The diazobicyclooctanes avibactam and its derivative OP-0595 were active at 50 μM and 100 μM concentrations (Figure 2.23D, 2.27). γ-Lactams, **18** and **19** only inhibited M^{pro} activity at the highest concentration of 100 μM (Figure 2.23B, 2.27, 2.24).

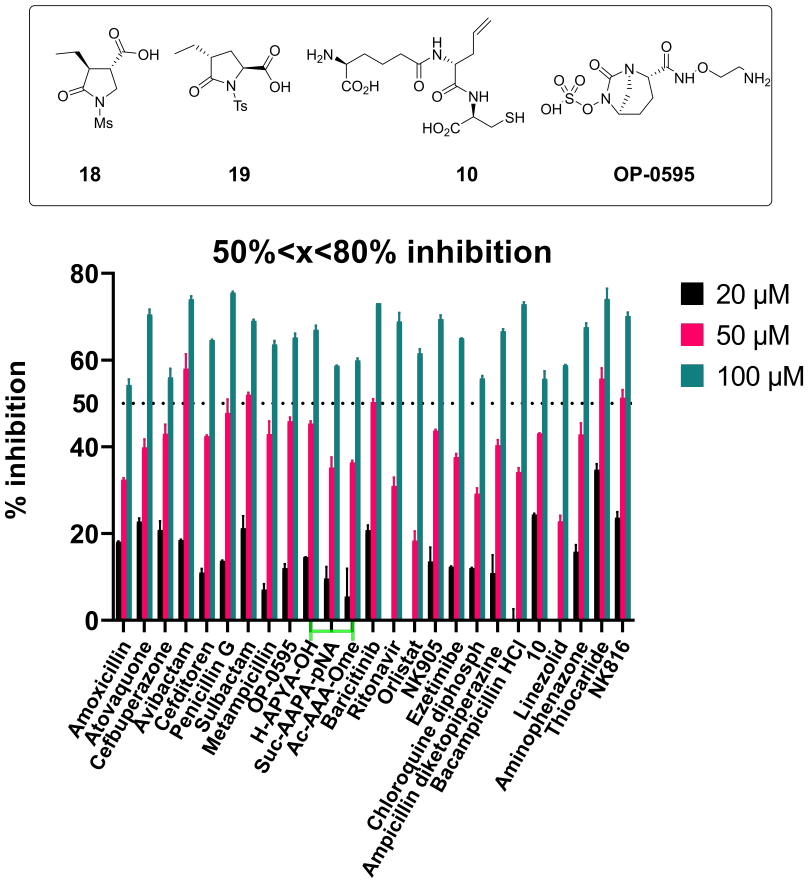


Figure 2.27 | Compounds identified with >50% and <80% inhibition of M^{pro} from the in-house screening campaign. Penicillins, cephalosporins, and protease inhibitors displayed weak M^{pro} inhibition. The short linear peptides (green) and γ-lactams inhibit >50% M^{pro} activity at 100 μM. The structure of some of the inhibitors is shown.

The results from the in-house screen suggested β-lactams might be optimisable for potent M^{pro} inhibition. Thus, their mode of inhibition via denaturing MS was investigated to probe the nature of the interaction of these compounds with M^{pro}.

2.3.5.2 Mode of interaction of β-lactam compounds

Multiple β-lactams were analysed to investigate their mode of M^{pro} inhibition. Cephalosporins and penicillins were the primary focus due to the positive inhibition results accrued as described in the preceding section. Carbapenems (ertapenem and tebipenem), monobactam (β-lactamase inhibitors: aztreonam, clavulanate, tazobactam and sulbactam) were also screened. Compounds were incubated with M^{pro} and their interaction was analysed after at least 25 min, 5 h, 9 h, and 24 h to potentially monitor the increasing levels of M^{pro} modification.

Cephalosporins

All the investigated cephalosporins apparently covalently modified M^{pro}, albeit at varying levels. Notably, complete M^{pro} modification was not observed even after prolonged incubation with any of the tested cephalosporins (Figure 2.28). The highest levels of M^{pro} modification were observed with cefotetan, cefbuperazone, and cefoxitin. For cefotetan (+576 Da adducts), cefbuperazone (+627 Da adducts), and cefoxitin (+427 Da adducts), evidence for more than one reaction was accrued. For cefazolin, an additional mass shift of 338 Da adduct was observed. For M^{pro} incubated with moxalactam, an additional peak (+476 Da adduct) to the intact moxalactam-M^{pro} complex (+520 Da) with a mass difference of 45 Da was observed, possibly due to decarboxylation of moxalactam from the intact moxalactam-M^{pro} complex.

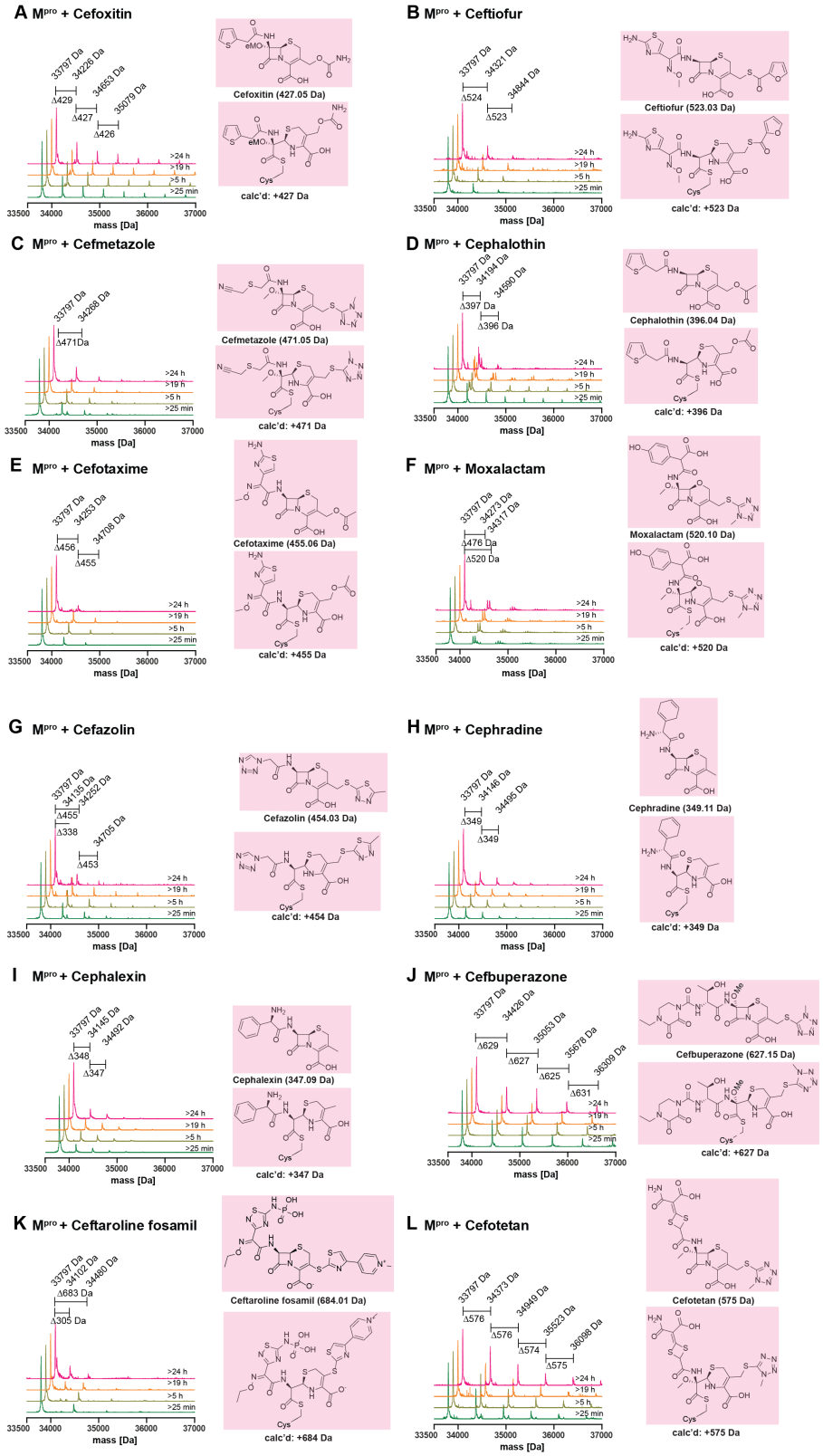


Figure 2.28 | Several cephalosporins covalently react with M^{pro}. M^{pro} (1 µM) was incubated with a 100-fold excess of **A.** cefoxitin, **B.** ceftiofur, **C.** cefmetazole, **D.** cephalothin, **E.** cefotaxime, **F.** moxalactam, **G.** cefazolin, **H.** cephradine, **I.** cephalexin, **J.** cefbuperazone, **K.** ceftaroline fosoamil, or **L.** cefotetan for (at least) 25 min, 5 h, 19 h and 24 h in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). The structures of the inhibitors and possible outcomes of the reaction with the M^{pro} nucleophilic cysteine are shown. See the Methods Section for assay details.

Penicillins

All the tested penicillins investigated apparently covalently modified M^pro, with the exception of mecillinam. As observed with cephalosporins, in no case was M^pro modification complete (Figure 2.29). The mass discrepancy of +19 Da adduct for M^pro incubated with oxacillin may correspond to a salt adduct or to hydration.

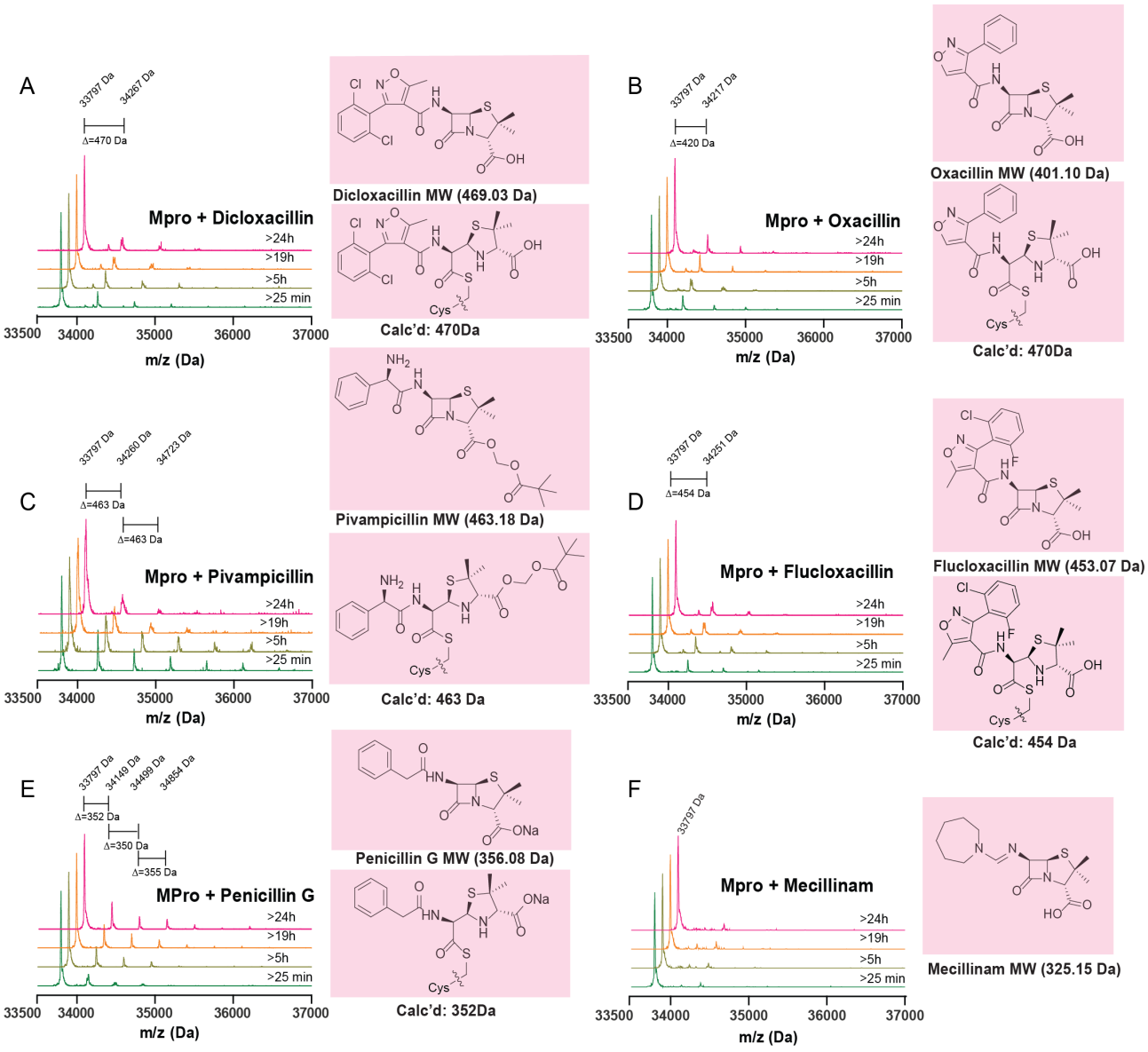


Figure 2.29 | Several penicillins covalently react with M^pro. M^pro (1 µM) was incubated with a 100-fold excess of **A.** dicloxacillin, **B.** oxacillin, **C.** pivampicillin, **D.** flucloxacillin, **E.** penicillin G, or **F.** mecillinam for (at least) 25 min, 5 h, 19 h, and 24 h in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). See the Methods Section for assay details. The structures of the inhibitors and possible outcomes of the reaction with the M^pro nucleophilic cysteine are shown.

Other β-lactams

Of the tested compounds, which are representatives of clavam, penem, carbapenem, and penicillin sulfone subclasses, only tazobactam, a penicillin sulfone, manifested increasing acylation of M^{pro} overtime (mass shift of 301 Da, expected adduct: 301 Da) (Figure 2.30). None of the carbapenems or penem or other penicillin sulfones reacted substantially with M^{pro} (Figure 2.30). Evidence for the occurrence of relatively low levels or multiple M^{pro} modifications were observed (ertapenem: mass shift of 474 Da, expected adduct: 475 Da, tebipenem pivoxil: mass shift of 498 Da, expected adduct: 497 Da, faropenem: mass shift of 285 Da, expected adduct: 285 Da, aztreonam: mass shift of 435 Da, expected adduct: 435 Da, clavulanate: mass shift of 158 Da, expected adduct: 156 Da, sulbactam: mass shift of 233 Da, expected adduct: 233 Da).

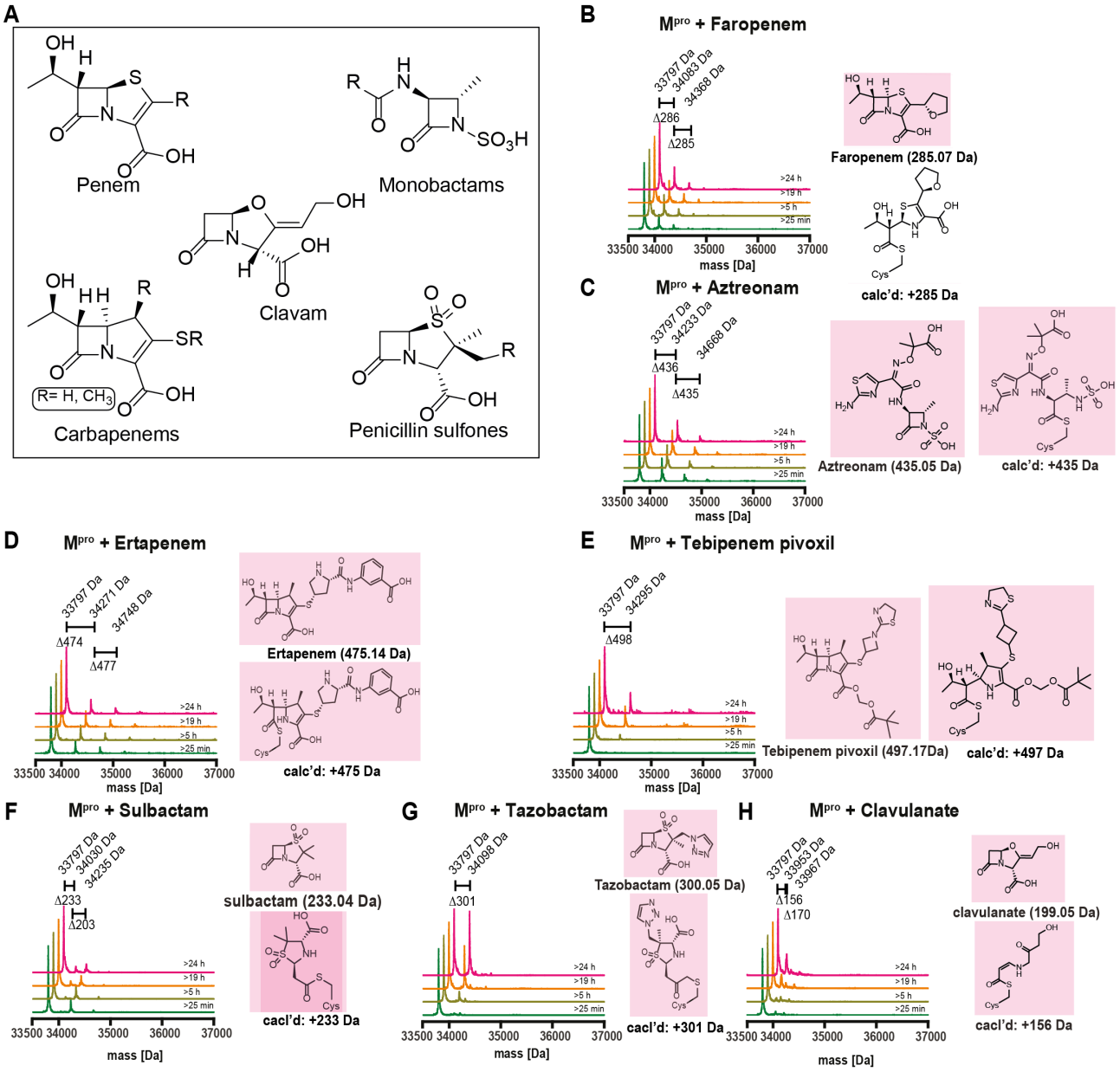


Figure 2.30 | Some β-lactams react covalently with M^{pro}. **A.** Different subclasses of β-lactams investigated. M^{pro} (1 µM) was incubated with a 100 fold excess of: **B.** faropenem (penem), **C.** aztreonam (monobactam), **D.** ertapenem (carbapenem), **E.** tebipenem (C3 protected carbapenem prodrug), **F.** sulbactam (a penicillin sulfone), **G.** tazobactam (a penicillin sulfone), or **H.** clavulanate (a clavam) for (at least) 25 min, 5 h, 19 h, and 24 h in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). See the Methods Section for assay details. The structures of the inhibitors and the possible outcomes of their reactions with the M^{pro} nucleophilic cysteine are shown.

2.3.5.3 Revalidation of selected screened in-house compounds via dry dispensing to assess reproducibility

The library of compounds for assay validation was dry dispensed using an Echo acoustic liquid dispenser; in-house screening was undertaken by dispensing the compounds with a Cybio liquid

handling robot. To investigate any differences arising in compound activity from these two methods of compound dispensations, some of the hits identified from in-house screening were rescreened. Additional in-house compounds **1-3** and the control inhibitor, N3 were also included. PBIT, disulfiram, boceprevir, and ebselen inhibited M^{pro} activity in assays employing both methods of compound dispensations. Although some inhibitory activity was still observed with cefdinir, ceftiofur, cefaclor, and cefprozil, all tested cephalosporins manifested lower percentage inhibition with the dry dispensing method. With 30 min of preincubation of compounds with M^{pro}, some of the hits identified before, such as penicillin esters (**46**, bacampicillin, and pivampicillin), did not inhibit M^{pro} activity (Figure 2.31A). Inhibition of >50% had been previously observed with 50 μM avibactam, which is a non-β-lactam β-lactamase inhibitor (Figure 2.23D). Avibactam did not display inhibitory activity at any of the fixed concentrations in the revalidation screen (Figure 2.31).

β-Lactams likely inhibit M^{pro} by covalent modification, therefore prolonged incubation with M^{pro} would be expected to improve their inhibitory activity. Lower IC_{50} values were obtained for cephalosporins and penicillin with an incubation time of 60 min (Figures 2.31A, 2.31B). However, compounds that did not display any inhibition with 30 min preincubation, also did not show any inhibition with 60 min incubation (Figure 2.31B).

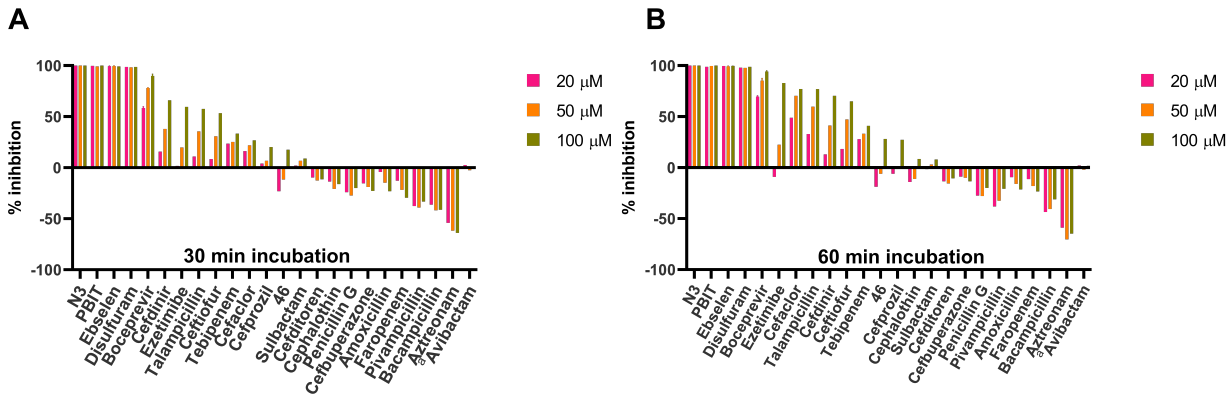


Figure 2.31 | Assaying M^{pro} inhibition of dry-dispensed β-lactams (antibiotics). SPE-MS M^{pro} inhibition assays with β-lactams were performed at 20 μM (pink), 50 μM (orange), and 100 μM (green) single β-lactams concentrations with either **A.** 30 min or **B.** 60 min preincubation times. Conditions: 0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2 peptide substrate in the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl).

To further investigate the observed inhibition by β-lactams at fixed concentration screens (Figures 2.31A, 2.31B), dose response curves were obtained with both (a) 30 min and (b) 60 min preincubation times (Table 2.5). Most cephalosporins that showed activity in the in-house screen were not active apart from ceftiofur (IC_{50} >100 μM, 30 min, IC_{50} ~75.7 μM, 60 min), cefaclor (IC_{50} >100 μM, 30 min, IC_{50} ~30.76 μM, 60 min) and cefdinir (IC_{50} >100 μM, 30 min, IC_{50} ~88.97 μM, 60 min), which only displayed inhibitory activity with 60 min preincubation. Talampicillin also displayed better inhibition with 60 min preincubation (IC_{50} ~90.24 μM, 30 min, IC_{50} 55.77 μM, 60 min).

Overall, good consistency in the extent of $\%M^{pro}$ inhibition was observed, when using either the in-house or dry dispensation method for the potent compounds, e.g. ebselen, cephalosporin C zinc salt, boceprevir, and PBIT. However, for compounds with mild potency, better corroboration between the fixed concentrations screen and IC_{50} values was observed when using the dry dispensation, e.g. bacampicillin, pivampicillin, and talampicillin. The findings summarised in Table 2.5 suggest some of the β-lactams may covalently inhibit M^{pro}; time dependency was observed for some compounds, including talampicillin, ezetimibe, ceftiofur, cefdinir C1 and cefaclor in IC_{50} studies (though only mild inhibition was detected).

Table 2.5: IC_{50} values for selected compounds. Compounds were preincubated for **A.** 30 min or **B.** 60 min with M^{pro}. Assays were carried out with 150 nM M^{pro} and 2 µM TSAVLQ/SGFRK-NH_2 peptide in assay buffer. See the Methods Section for assay details.

M^{pro} inhibitor	**$\sim IC_{50}(\mu M)^A$**	**$\sim IC_{50}(\mu M)^B$**
N3	0.04	0.04
Ebselen	0.14	0.13
Disulfiram	0.59	0.47
Boceprevir	13.4	13.1
Cephalosporin C zinc salt	16.3	18.2
Antipain	24.6	35.9
Thiocarlide	38.6	87.1
Nelfinavir	63.6	41.9
Talampicillin	90.2	55.8
Ezetimibe	100	82.9
Bacampicillin	>100	>100
Ceftiofur	>100	75.7
Cephalothin	>100	>100
Tebipenem	>100	>100
Avibactam	>100	>100
Aztreonam	>100	>100
Camostat Mesylate	>100	>100
Pivampicillin	>100	>100
Cefprozil	>100	>100
Cefdinir C1	>100	89.0
Cefaclor	>100	30.8
Chloroquine Diphosph	>100	>100
Linezolid	>100	>100
Faropenem	>100	>100
Cefbuperazone	>100	>100
Sulbactam	>100	>100
Cefditoren	>100	>100
Penicillin V benzyl ester (**46**)	>100	>100
Amoxicillin	>100	>100
Penicillin G	>100	>100

2.3.5.4 Validation of metal ion based M^{pro} inhibitors

The cephalosporin C zinc salt was identified as a hit inhibitor from both the compound library screen and the in-house screening. As a zinc metal ion is present in cephalosporin C zinc salt, the potential of metal ion based M^{pro} inhibition rather than via its β-lactam warhead, was investigated. Metal based M^{pro} inhibitors such as auranofin, cisplatin, and bismuth subsalicylate have been identified with nanomolar/sub-micromolar half-maximal inhibitory concentration in the LOPAC^{1280} and FDA approved compound libraries screen (Table 2.4). Metal salts were therefore assayed to investigate their effect on M^{pro} catalysis.

Lithium, magnesium, manganese, calcium, chromium, and nickel salts did not inhibit M^{pro} (Figure 2.32). Mild inhibition was observed with ferric gluconate (IC_{50} ∼85 µM) and ferric sulfate (IC_{50} ∼44 µM). Copper (IC_{50} 0.16 µM) and cadmium (IC_{50} 0.35 µM) ions showed potent M^{pro} inhibition (Figure 2.32). The results indicate only specific transition metal ions can inhibit M^{pro} activity. Results with zinc ions are described in the next section.

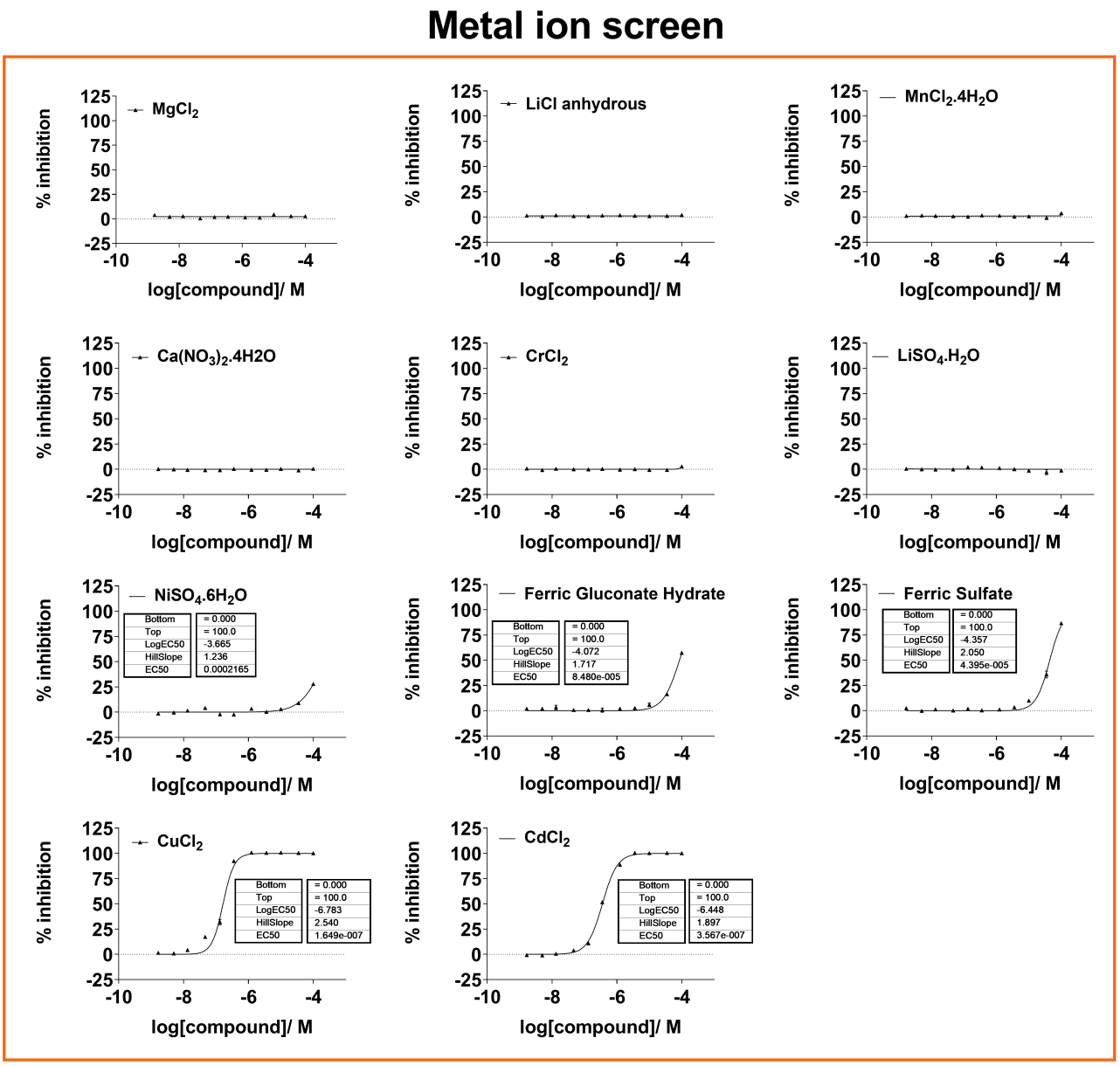


Figure 2.32 | Inhibition of M^{pro} by metal salts. SPE-MS M^{pro} inhibition assays with metal ions were performed as described in the Methods section. Li^{+}, Mg^{2+}, Mn^{2+}, Ca^{2+}, Cr^{2+} and Ni^{2+} are poor inhibitors of M^{pro}. Fe^{2+} shows moderate M^{pro} inhibition.

2.3.5.5 M^{pro} inhibition from cephalosporin C zinc salt arises from zinc metal ion

In an initial assay (Table 2.5), the cephalosporin C zinc salt was solubilized in DMSO via sonication with heat. The sample was then centrifuged (14k rpm, 5 min) to remove precipitation for dose response analysis. By contrast, for the subsequent screen, the cephalosporin C zinc salt and metal ion salts were prepared in 20 mM HCl solution (400 mM stock solutions) for consistent sample preparation. The compounds were diluted in DMSO to 20 mM stock solutions, then assayed. This procedure

gave a higher potency for the cephalosporin C zinc salt, i.e. an IC_{50} of ∼1.3 µM compared to the IC_{50} of ∼16.3 µM, obtained with 30 min preincubation. $ZnCl_2$ (IC_{50} 0.18 µM), and $ZnSO_4$ (IC_{50} 0.32 µM) were both relatively potent inhibitors of M^{pro}. The potencies of the cephalosporin C zinc salt and $ZnCl_2$ are thus similar, implying that most of the inhibition potencies of the cephalosporin C zinc salt is from the zinc ions (Figure 2.33A-2.33C).

Protein observed MS was performed with cephalosporin C and M^{pro} and N3 reacted M^{pro} (Figure 2.33D, 2.33E). No evidence for a peak corresponding to the reaction with M^{pro} with the cephalosporin C zinc salt was observed. A minor unidentified peak corresponding to a mass increase of 314 Da was observed. This peak was still observed with N3:M^{pro} complex suggesting, any covalent reaction with cephalosporin C is not efficient (if occurrence at the active site Cys145), implying that inhibition by the cephalosporin C zinc salt, likely arises from the zinc ions.

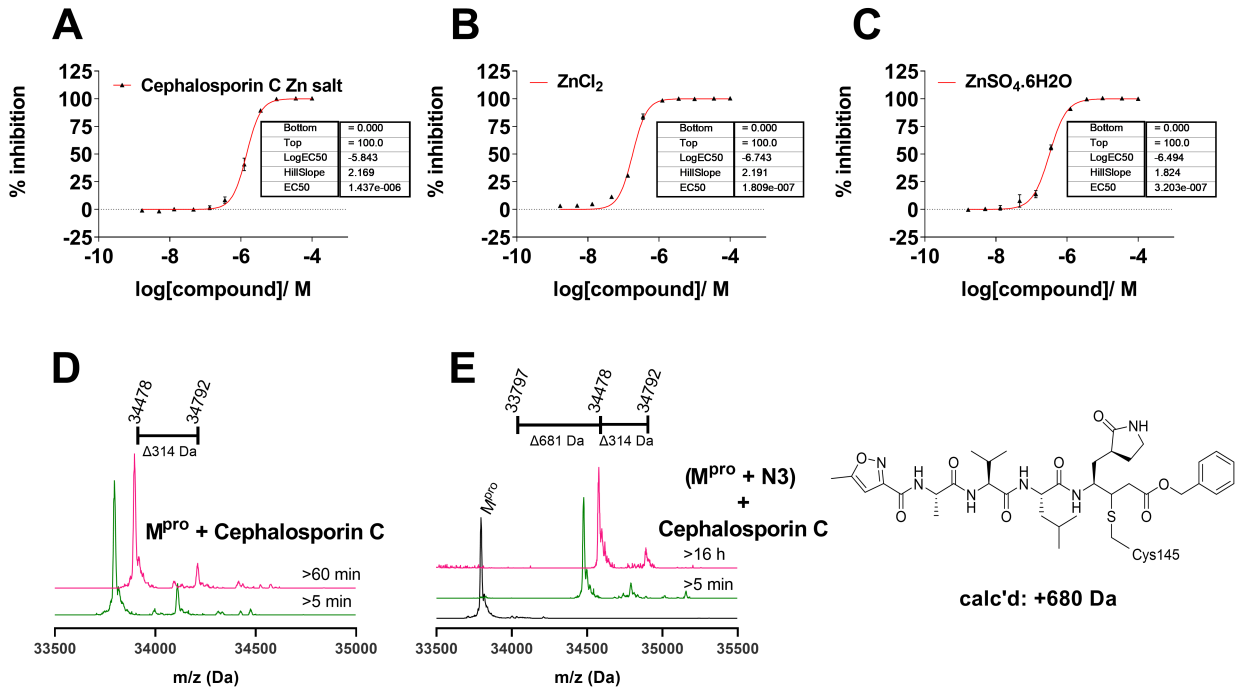


Figure 2.33 | The cephalosporin C zinc salt is an M^{pro} inhibitor. A. The cephalosporin C zinc salt inhibits M^{pro}. **B.** and **C.** Zinc ions inhibit M^{pro}. **D.** Cephalosporin C does not react covalently with M^{pro}. **E.** The cephalosporin C does not react with the M^{pro}:N3 complex. The results support that most, if not all, inhibition from the cephalosporin C zinc salt may come from the zinc ions.

2.3.5.6 Investigating the interaction of tazobactam and aztreonam with M^{pro}.

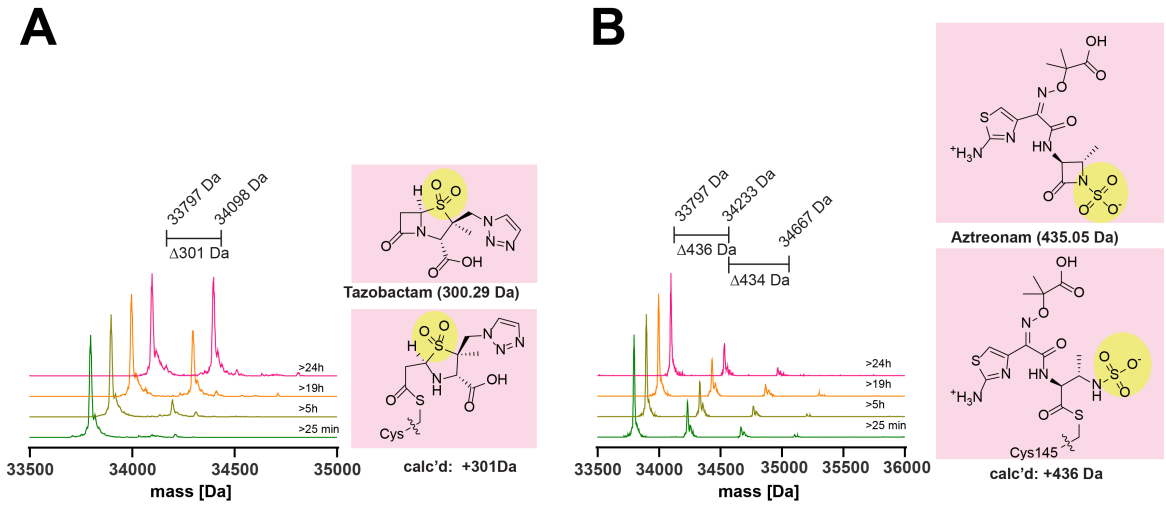


Figure 2.34 | A. *N*-sulfonyl containing tazobactam and B.*N*-Sulfonate containing aztreonam react with M^{pro}. SPE-MS M^{pro} assays were performed as described in Methods Section. Possible modes of reaction are shown.

Mode of interaction studies with representatives of different β-lactam subclasses revealed interesting findings. With tazobactam, a time dependent increase in an apparent M^{pro}: tazobactam complex was observed (Figure 2.34B). Aztreonam, which has an *N*-sulfonyl group, also showed evidence for covalent reaction with M^{pro} (Figure 2.34A). However, the level of the peak at 34233 Da corresponding to modification of M^{pro} by aztreonam did not alter over time. The interaction of the aztreonam in the active site M^{pro} pocket was probed by docking studies using the program Gold [275], experiments that were performed by Henry Chan (Department of Chemistry, University of Oxford). The top three poses for aztreonam show the *N*-sulfonate group is positioned in the S1 M^{pro} subsite. With tazobactam, the sulfone was located in the S1′ pocket in one of the poses (Figure 2.35).

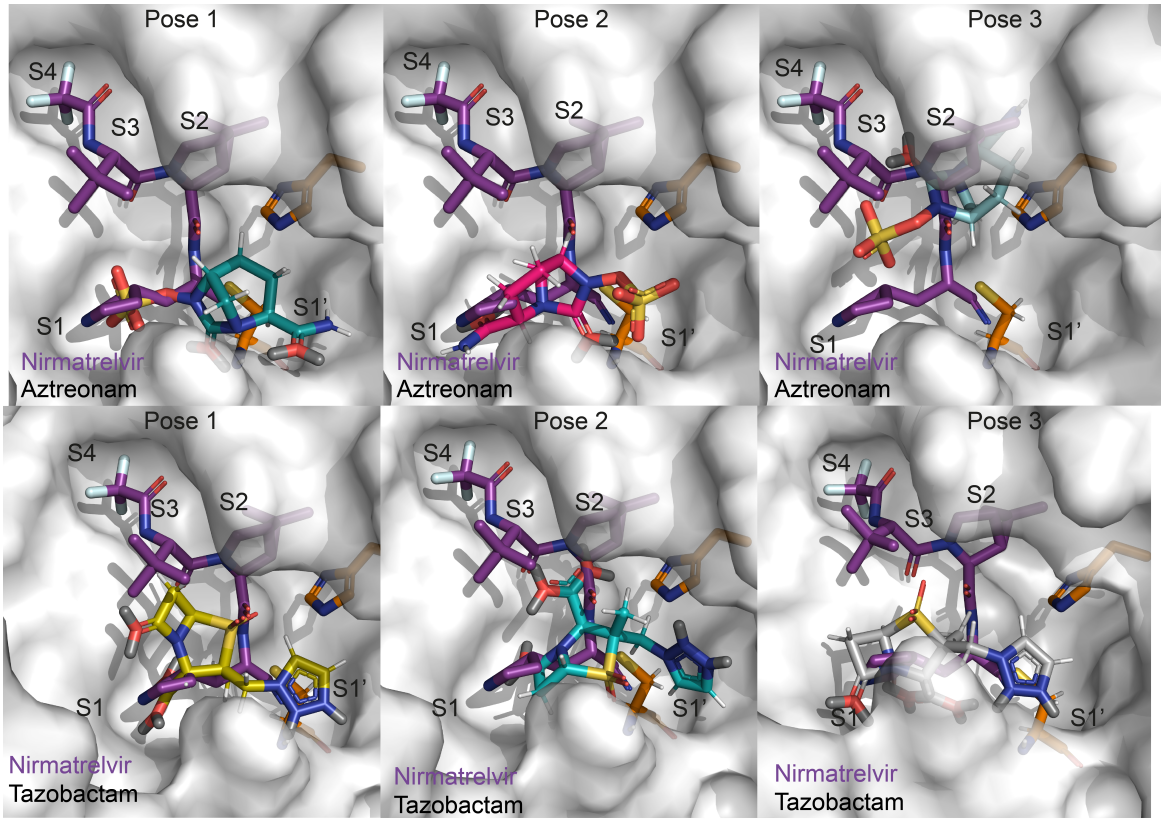


Figure 2.35 | Docking poses of aztreonam and tazobactam with M^{pro}. Docking studies were performed by Henry Chan using the program Gold [275].

2.3.5.7 Investigating the inhibition of M^{pro} via monobactams.

The results with tazobactam and aztreonam suggested that sulfone/sulfonate groups are of interest with respect to M^{pro} inhibition. A small library of synthetic *N*-sulfonate and related *N*-functionalised derivatives of monobactams was obtained by collaboration with Prof. Stanislav Gobec (Faculty of Pharmacy, University of Ljubljana), and was screened against M^{pro}. (This collaboration was established with the help of Dr. Alen Kranjc, who is currently working with Prof. Gobec.)

N-Sulfonate derivatives **50** and **51** displayed mild M^{pro} inhibition. **61** , which has a C4 thalidomide group similar to **50** and **51**, also weakly inhibited M^{pro} (Table 2.6). Interestingly, **50** also has a nitrile moiety; the nucleophilic cysteine of M^{pro} could therefore react at either the nitrile or with the β-lactam ring. As the monocyclic β-lactams inhibit M^{pro}, the results of the initial screening of commercial small-molecule libraries (Section 2.3.4) were reanalysed for β-lactam containing molecules. Apart from cephalosporin C, other β-lactams were identified as M^{pro} inhibitors (Table 2.7).

Table 2.6: IC_{50} values for selected monobactams as potenital M^{pro} inhibitors. Compounds were preincubated for (a) 30 min.

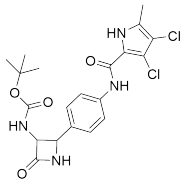
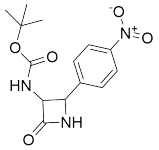
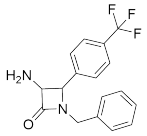
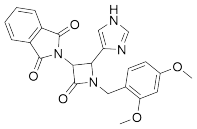
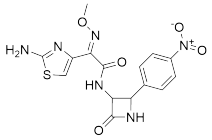
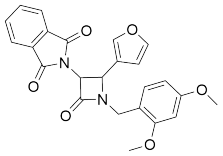
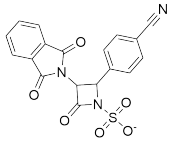
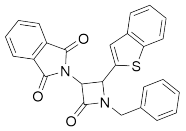
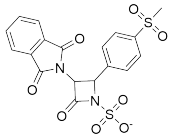
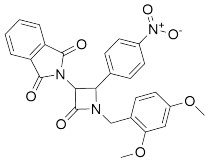
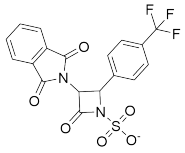
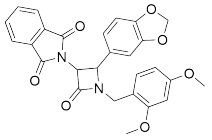
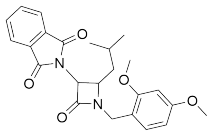
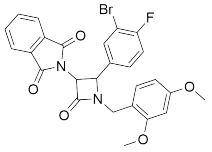
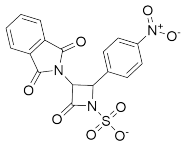
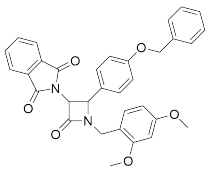
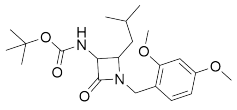
Monobactams	Structure	**IC_{50}(µM)[a]**	**Monobactams**	Structure	**IC_{50}(µM)[a]**
47		>100	56		>100
48		>100	57		>100
49		>100	58		>100
50		∼36.6	59		>100
51		∼47.7	60		>100
52		>100	61		∼36.3
53		>100	62		>100
54		>100	63		>100
55		>100			

Table 2.7: β-Lactam M^{pro} inhibitors were identified from screening libraries of FDA approved molecules or LOPAC (Figures 2.11B, 2.11C, Tables 2.2,2.3). The observed inhibition of M^{pro} by the cephalosporin C zinc salt is likely due to zinc ions (at least predominantly). Compounds were preincubated for 30 min.

$LOPAC^{1280}$	**∼% inhibition at 20 (μM)**
Cephalosporin C zinc salt	92.3
BLI-489 hydrate	77.3
Imipenem monohydrate	12.4
Cefaclor	8.7
FDA approved molecules	**∼% inhibition at 20 (μM)**
Cephalosporin C sodium	78.3
Ceftiofur hydrochloride	65.0
Cefprozil	57.5
Cefepime hydrochloride [276]	20.6
Imipenem	12.4
Ceforanide	9.2
Ceforanide	10.2
Cefalonium	9.2
Sulbactam	6.6

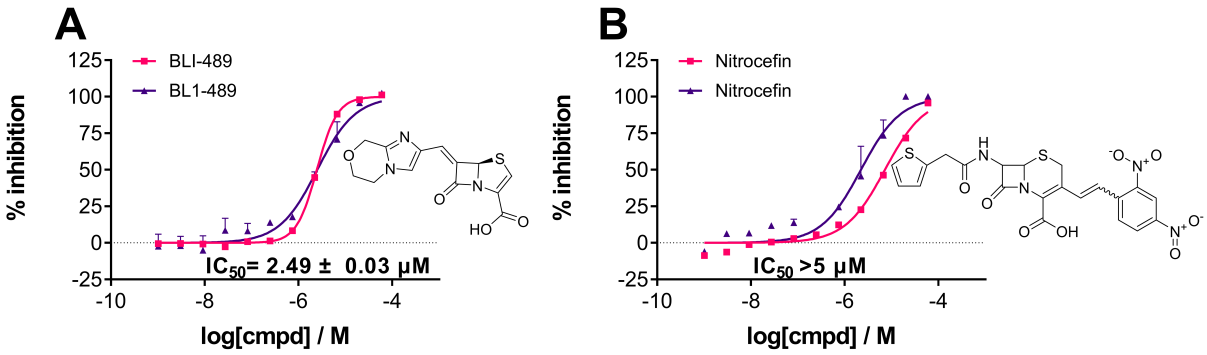


Figure 2.36 | Dose response curve studies of BLI-489 and nitrocefin inhibition of M^{pro}. A. Penem (BLI-489) and **B.** cephalosporin (nitrocefin) both inhibit M^{pro}. Independent repeats each composed of technical duplicates were used to derive the IC_{50} values (n= 2 ± SD).

Most of the active β-lactams in the commercial library screen belonged to the cephalosporin subclass of β-lactams. Investigations of cephalosporins as M^{pro} inhibitors had previously been undertaken (Section 2.3.5); in this initial screening, apart from cephalosporin C zinc salt, no other

potent cephalosporin M^{pro} inhibitor was identified. Nitrocefin, a cephalosporin that is used as a reporter substrate for β-lactamase was investigated [277], which was observed to manifest moderate M^{pro} inhibition (IC_{50} ~5 μM). Ceftazidime and cefepime cephalosporins have also been reported by others to inhibit M^{pro} with comparable potency to nitrocefin [276]. This report, along with the observation with nitrocefin suggests that the cephalosporins core could be further optimised for potent M^{pro} inhibition. BLI-489, a penem subclass of β-lactams, manifested 77% M^{pro} inhibition at 20 μM. Dose response studies showed BLI 489 is a moderately potent M^{pro} inhibitor (IC_{50} 2.49 ± 0.03 μM, Figure 2.36). Faropenem, a medicinally used penem, did not inhibit M^{pro} (Table 2.5). Both BLI-489 and nitrocefin have a vinyl group, a potential site for reactivity in addition to the β-lactam carbonyl, which may undergo nucleophilic addition. Rearrangement of the BRL 42715 penem acycl enzyme complexes with SBLs via the opening of the thiazole ring where the nucleophilic sulfur attacks the aforementioned olefin group to form a cyclic amino-acrylate derivative has also been reported [278]. Therefore, although moderate M^{pro} inhibition was observed with BLI-489 and nitrocefin, they were not pursued further as β-lactam containing M^{pro} inhibitors.

2.3.5.8 Investigating the inhibition of M^{pro} by penicillin derivatives.

Screening of compounds using acoustic dry dispensing or in-house dispensation (CyBio-Liquid handling robot) revealed inconsistencies with the activities of penicillin esters such as bacampicillin, talampicillin, and pivampicillin (Figure 2.37A, 2.37B). These inconsistencies may reflect degradation. Two batches of penicillin V esters also gave inconsistent results (Figure 2.23C, 2.31, 2.34). Penicillin V esters and their oxidised derivatives were screened further with aim of addressing the inconsistencies as well as identifying penicillin β-lactams with potent M^{pro} inhibition.

β-Lactam derivatives synthesized by previous members in the Schofield lab were catalogued; this included a C6-methoxy, a C5-epimer of penicillin G ester (**65**), *para*-nitrobenzyl ester of penicillin G sulfoxide (**68**). Cephalosporin derivatives, i.e. a C6-epimer of cephalosporin ester with C7 thalidomide side chain (**64**) and the biphenyl ester of cephalosporin sulfoxide (**66**) were also included to investigate if alternatives of commercially available cephalosporins manifest M^{pro}

inhibition. However, the C6-epimer of cephalosporin sulfoxide did not inhibit M^{pro}. Of the in-house catalogued compounds with *R* configuration at the C5 position, potent M^{pro} inhibition was only observed with **68**, a C6-methoxy penicillin sulfoxide.

Dr. Lennart Brewitz (Department of Chemistry, University Of Oxford) synthesized the penicillin G esters (sulfoxide **69**) and penicillin V ester derivatives (sulfoxide **70**, sulfone **67**). However, no M^{pro} inhibition was observed with the sulfoxide ester derivatives. Only penicillin V sulfone derivative **67** showed notable inhibition, which had been included in both methods of compound dispensations as a control and displayed inhibition under both experimental setups.

Docking studies were performed by Henry Chan (Department of Chemistry, University Of Oxford) to investigate if the β-lactam is poised to react with active site Cys145. The docking studies indicate that the β-lactam in the penicillin esters bind in a manner that may enable reaction with Cys145 [253].

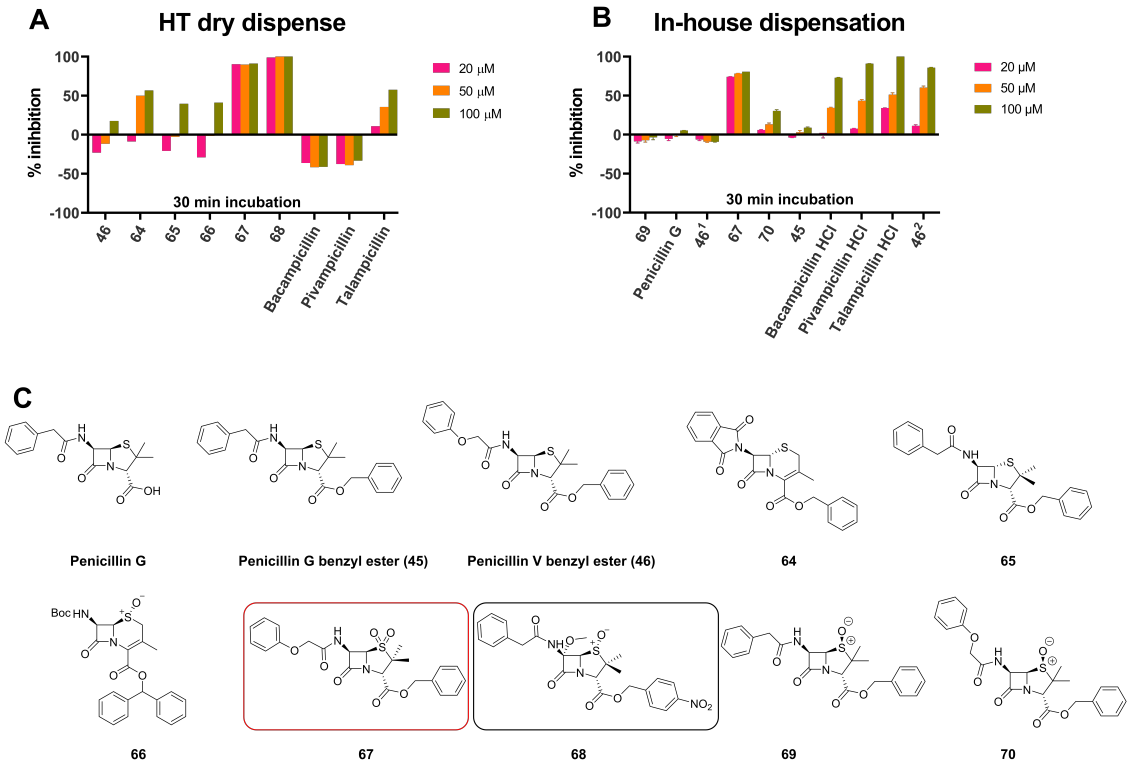


Figure 2.37 | M^{pro} inhibition is mediated by penicillin esters. **A.** Compound screened via high throughput dry dispensing method shows penicillin esters are not active against M^{pro}. **B.** Compound screened via in-house setup shows penicillin esters are active against M^{pro}. M^{pro} inhibition is observed with penicillin sulfone (**4**) under both experimental setups. Data is composed of technical duplicates (n= 2 ± SD).

2.3.5.9 Validating the M^{pro} inhibition by the penicillin *S*-oxidised derivatives 4 and 5

A penicillin sulfone (**67**) and a penicillin sulfoxide (**68**) that manifested M^{pro} inhibition at the fixed concentrations screen were investigated in dose response studies. Note that, to maintain consistency, the compound numbering identical to manuscript 1 for (**67**→**4**) and (**68**→**5**) is henceforth used in this thesis for these compounds [253]. **4** and **5** both manifest moderate inhibition (IC_{50} 1.5 ± 0.7 μM) and (IC_{50} 3.5 ± 0.1 μM), respectively with 30 min preincubation. No substantial time dependency was observed with dose response studies conducted with a longer preincubation period of 60 min [**4** (IC_{50} 0.9 ± 0.7 μM) and **5** (IC_{50} 2.1 ± 0.2 μM)]. **4** was observed to covalently react with M^{pro} by SPE-MS, but **5** was not observed to do so. Incubation of **4** with N3 reacted M^{pro}, does not result in further modification of N3:M^{pro} complex, suggesting **4** is an active site Cys145 selective covalent inhibitor of M^{pro}. Only low level modification of M^{pro} or the M^{pro}:N3 complex was observed with **5**. These results suggest that β-lactam penicillin ester sulfones or C6-methoxy sulfoxides could be optimised for potent M^{pro} inhibition (Chapter 3).

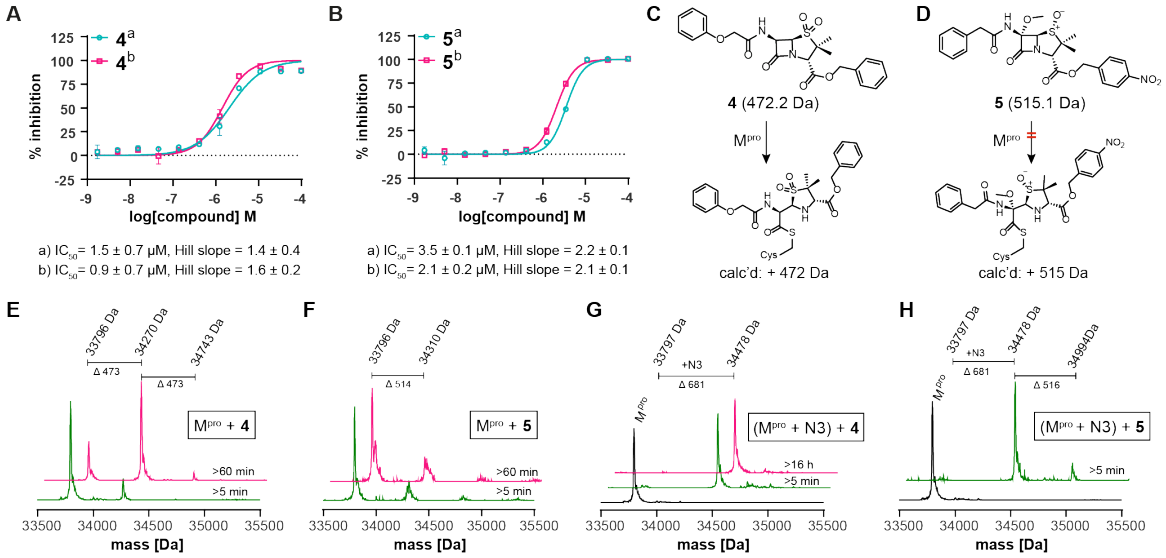


Figure 2.38 | Penicillins 4 and 5 inhibit M^{pro}. IC_{50}s for **A. 4** (C3-benzyl ester of penicillin V sulfone) and **B. 5** (C3-*para*-nitrobenzyl ester of C6-methoxy penicillin G sulfoxide) determined using SPE-MS; data are a mean of technical duplicates with (a) 30 min and (b) 60 min preincubation. IC_{50}s are means of two independent repeats each composed of technical duplicates (n=2 ± SD). Proposed reaction of **C. 4** and **D. 5** with M^{pro}. **E.** A single molecule **4** covalently modifies M^{pro}. **F. 5** does not efficiently modify M^{pro} through covalent reaction. Conditions: 1 µM M^{pro}, 20 µM β-lactam, 20 mM HEPES pH 7.5, 50 mM NaCl. **G. 4** does not covalently modify M^{pro} preincubated with N3, suggesting **4** reacts with Cys-145. **H. 5** does not efficiently react with M^{pro} preincubated N3. Conditions: 1 µM M^{pro} preincubated with 3 µM N3, 20 µM 1 or 2, 20 mM HEPES pH 7.5, 50 mM NaCl. Black spectra: wild-type M^{pro} (33796 Da).

2.3.6 C-Terminal M^{pro} constructs and their activity

The substrate turnover studies with M^{pro} constructs with a C-terminal decaHis-tag suggest that extended residues at C-terminus can be tolerated by M^{pro}, although there is a slight reduction in efficiency of substrate conversion compared to M^{pro} with native N- and C-termini. Reduced catalytic efficiency has also been reported for C-terminally His-tagged constructs of SARS-CoV-1 M^{pro} (k_{cat}/K^m 600 $M^{-1}s^{-1}$) [28]. The results with C-terminally decaHis-tagged constructs of M^{pro} suggest that the composition of the C-terminal residues is critical for M^{pro} activity, as C-terminally Avi-tagged M^{pro} is inactive.

Apparent autocatalysis of the C-terminal His-tag was observed during MS characterization of M^{pro} purified by IMAC; this may, in part, be explained by the similarities between the altered HRV 3C protease cleavage sequence (SGVTFQ↓GP) to the C-terminal autocatalytic M^{pro} cleavage site, i.e. (SGVTFQ↓GK) [240]. The preferred cleavage sequence of HRV 3C protease, i.e. (LEVLFQ↓GP),

was altered to generate M^{pro} with a natural C-terminus [240]. This change was made based on the observation that SARS-CoV-1 M^{pro} did not tolerate proline at the P2′ position. However, the P2′ substrate residues are not highly conserved among the different cleavage sites in SARS-CoV-2 M^{pro}, and to a lesser extent in the P2′ substrate residue in cleavage sites of SARS-CoV-1 and SARS-CoV-2 M^{pro}. For instance, Lys is present at the P2′ position in the nsp5/6 and nsp6/7 SARS-CoV-1 M^{pro} cleavage sites [279], whereas a Lys at the P2′ position is only present at the nsp5/6 cleavage site in SARS-CoV-2 M^{pro} [280]. Further studies to identify the optimal P2′ residue (i.e., X in SGVTFQ↓GX) for which SARS-CoV-2 M^{pro} does not display any activity (even after prolonged incubation) could lead to increased protein yields following purification.

2.3.7 Assay development and validation

The excellent assay parameter (Z′ >0.8) suggests that the robust SPE-MS based M^{pro} assays are suitable for high throughput applications to screen novel and repurposed M^{pro} inhibitors. The SPE-MS assay is an attractive alternative to FRET-based assays as it not only allows the inhibitory effect of small molecules on M^{pro} activity to be assessed but also simultaneously informs on the mechanism of inhibition. The absence of fluorescent tags in the SPE-MS M^{pro} assay substrate also eliminates potential interference on M^{pro} activity and specificity [281]. FRET probes are also susceptible to inner filter effects at concentrations higher than 20 μM, impacting their application in kinetic studies [242, 282–286]. However, SPE-MS assay may be limited by the different levels of ionisation observed for product and substrate peptides; hence, the optimisation of SPE-MS conditions requires an internal standard peptide (Chapter 5).

An alternative MS M^{pro} assay employing self-assembled monolayer desorption ionization based mass spectrometry (SAMDI-MS) has been reported [281]. Both the SPE-MS and SAMDI MS assays have excellent parameters [253, 281]. Whilst the SAMDI-MS assay requires sample immobilisation and the addition of matrix for sample analysis, the SPE-MS assay allows samples to be directly loaded into SPE cartridges without further treatment after the reaction is quenched [281, 287]. Compared to the SPE-MS assay, a biotinylated substrate is required for sample immobilisation in the SAMDI-MS assay. The relatively high K_m for this substrate (245 μM) may impact the screening of substrate-competitive

inhibitors [281]. EDTA and detergents were included in the SAMDI-MS assay conditions to remove trace metals and prevent identification of false hits from compound aggregates, respectively [281,288]. The inclusion of EDTA and detergent is feasible in the SAMDI-MS assay as they are washed away from the immobilised substrate in SAMDI biochip arrays, whereas they could cause ion suppression in MS studies [281,289]. The protocol for the SPE-MS M^{pro} assay includes an aqueous wash step of the sample loaded cartridge to remove non-volatile salts. However, EDTA and detergent cannot be efficiently washed off of the SPE cartridge, and residual levels may remain present in the tubing of the coupled RapidFire 365 liquid handling robot which could be detrimental to the SPE-MS assay. Therefore, despite the advantage of eliminating false negatives by adding EDTA and detergent to the enzymatic assays, this is less feasible in SPE-MS M^{pro} assays.

Validation of the initial hits identified from the fixed concentration screen of LOPAC^{1280} and library of FDA-approved compounds suggest that the SPE-MS assay is capable of accurately identifying hits using a fixed concentration screen. The majority of the identified hits from the LOPAC^{1280} screen were not reported in the LOPAC^{1280} SPE-MS screen employing FRET assays. Only two hits (NSC 95937 and aurothioglucose) identified from the LOPAC^{1280} screen using the SPE-MS assay were also identified from the LOPAC^{1280} screen using FRET assays [265,266]. Some of the differences in the results obtained from the LOPAC^{1280} screen employing the SPE-MS assay (Table 2.3) and FRET assays could arise from the differences in assay conditions, such as presence of a chelating agent (EDTA) and a reducing agent, dithiothreitol (DTT) in the FRET assays [265,266]. For instance, disulfides (e.g., IPA-3, disulfiram) are less likely to show potent M^{pro} inhibition under reducing conditions. Indeed, disulfiram was not identified as a hit from LOPAC^{1280} screening employing FRET assays in the presence of reducing agents [265,266]. By contrast, and in agreement with the SPE-MS assays, FRET assays performed without the reducing agent identified disulfide-containing molecules, e.g. PX-12 and ebselen, as hits during a screen of FDA-approved and other small molecule libraries [28,240]. However, Barker *et al.* [265] did not observe M^{pro} inhibition by ebselen when employing the FRET assay conditions with the reducing agent DTT. In a separate study of IPA-3 mediated inhibition of Pak1 kinase, inhibition from

IPA-3 was shown to be reversible under reducing conditions [290]. Thus, the presence of a reducing agent may influence the nature of the hits identified from these small molecule libraries.

These results emphasise the importance of screening inhibitors under both reducing and non-reducing conditions to increase the probability of identifying compounds that may display activity in cells. Note that this may not be sufficient to increase the corroboration between the results of *in vitro* enzymatic assays and cell-based assays. For instance, NSC 95937 manifested M^{pro} inhibition under both reducing and non-reducing conditions but did not manifest activity in SARS-CoV-2 infected Vero cells [262].

Screening of the LOPAC library led to the identification of chloromethyl ketone based protease inhibitors, i.e. Z-L-Phe chloromethyl ketone, TPCK, and TLCK. Z-L-Phe chloromethyl ketone and TPCK inhibit chymotrypsin, whereas and TLCK inhibits trypsin-like serine proteases by alkylating the active site His residue [257]. Furthermore, cysteine proteases can be inhibited by the reaction of nucleophilic cysteine residue with TPCK and TLCK [259,291]. This observation is in agreement with the findings that the chymotrypsin-like protease inhibitor Z-L-Phe chloromethyl ketone (IC_{50} 50 nM, data not shown) and TPCK have higher potency against the chymotrypsin-like protease M^{pro} than TLCK (Tables 2.3, 2.4). Assays with N3-treated M^{pro} support the selective modification of Cys residue of M^{pro}. However, TPCK, TLCK, and chloromethyl-containing inhibitors may also interact with and modify targets in human cells [258]. Improved selectivity of these inhibitors against M^{pro} rather than host proteins could be achieved via the use of Gln analogues at P1 instead of the Phe and Lys residues in TPCK and TLCK, respectively. Interestingly, inhibition by chloromethyl ketone inhibitors was not reported in the results from a LOPAC screen employing FRET-based M^{pro} assays [266]. The group that reported the latter work [266] also screened other compound libraries from the National Center for Advancing Translational Sciences (NCATS) Pharmaceutical Collection and anti-infective, MIPE5.0. From these screens, they report potent M^{pro} inhibition by fluoromethyl ketones (FMK) such as Z-DEVD-FMK and Z-FA-FMK, which react with M^{pro} cysteine by similar mechanisms as the chloromethyl ketones [253,266,292].

Further corroboration between literature and the in-house screening efforts has been established by the identification of nelfinavir, an HIV protease inhibitor that displayed mild M^{pro} inhibition (IC_{50} ~63.6 μM, 30 min) under the assay conditions (Table 2.5). In agreement with this observation,

nelfinavir has been reported to inhibit M^{pro} (36% inhibition at 100 µM [293]). It is of note that when the sigmoid curve is not fully sampled, the IC_{50} values cannot be predicted accurately. Hence approximate IC_{50} values as in the case of nelfinavir have been reported in such instances.

2.3.8 β-Lactams inhibit M^{pro}

β-Lactam drugs are an extremely important class of medicines; their utilisation over 70 years establishes them as one of the safest drugs for human use. Some β-lactam compounds were shown to inhibit M^{pro} from screening of LOPAC^{1280} and FDA approved molecule libraries, in addition to in-house screening efforts. In the literature, talampicillin has been reported to be a poor inhibitor of M^{pro} (estimated IC_{50} 250-600 µM) [293]. Similarly, poor inhibition was observed with talampicillin in the SPE-MS studies using acoustic dry dispensation of compounds (IC_{50} ∼90 µM, 30 min). Dose response studies indicated that talampicillin is a weak M^{pro} inhibitor; experiments with longer incubation period show that the dose response is time-dependent (IC_{50} ∼55 µM, 60 min). However, the screening results with in-house compound dispensation using CyBio robot showed that talampicillin was active at ≥50 µM. The discrepancy observed between the in-house assay setup and the dry dispensing method may in part be explained by contamination of the tips used in CyBio robot by PEG, as PEGs can cause ion suppression of peptides [294, 295].

The cephalopsorin C zinc salt was one of the most potent β-lactams to be identified from the initial library screening (Table 2.3). However, metals, including zinc are known to inhibit cysteine proteases [296, 297]. Furthermore, inhibition arising due to the zinc ion of the cephalosporin C zinc salt has been reported in studies of dNTPase [298] and aspartase [299] proteins. Crystallographic structures with zinc ions bound to the active site of M^{pro} have been reported [300]. SPE-MS assays with the cephalsporin C zinc salt, auranofin, and cisplatin show that compounds containing Zn^{+2}, Au^{+1}, and Pt^{+2}, as well as metal ions such as zinc, copper, cobalt and iron can inhibit M^{pro}. *In silico* studies with metal ions such as Co^{+2} and Cu^{+2} indicate that these metal ions alone can inhibit M^{pro} activity [301, 302]. Though my identification of metal complexes and metal ions as M^{pro} inhibitors preceded these publications, Cd^{+2} has still not been reported elsewhere as metal ion inhibitors. Although metal ions

alone may not be sufficient for therapeutically useful inhibition of proteases in cells [300], metal ions that inhibit M^{pro} could be incorporated into the design of novel metal complexed M^{pro} ligands.

Some β-lactams belonging to the monobactam subclass are weak inhibitors of M^{pro}. The β-lactam *N*-sulfonates (**50** and **51**) and C6-thalidomide (**61**) showed some M^{pro} inhibition (IC_{50}36-48 μM). This observation raises interest in the potential activity of *S*-oxidised compounds as M^{pro} inhibitors. A penam sulfone (tazobactam) was observed to covalently modify M^{pro}, albeit the reaction was slow as observed by SPE-MS. *In silico* modelling studies could be utilised to identify C6 and C3 side chains which could improve the reactivity of the β-lactams, whilst retaining the sulfonyl/sulfone moiety. The observed inhibition by the penicillin benzyl esters could, in part, resemble the binding of the N3 benzyl ester in the S1′ subsite [28]. In the case of penicillin trifluoro-benzyl ester, the ester group was solvent exposed and positioned between S1 and S1′ subsite (Manuscript 3 [303]). The identification of active site selective *S*-oxidised penicillin esters **4** and **5** suggests that there is scope for optimising β-lactams as potent M^{pro} inhibitors.

In conclusion, the SPE-MS assay is an excellent M^{pro} assay that employs label-free peptide substrates. Assay validation with commercial library screens identified several different types of inhibitors as hits which could be further optimised to enhance their selectivity and potency. In-house screening with β-lactams shows that the β-lactam warhead can also be optimised for more potent M^{pro} inhibition. Cephalosporins covalently react with M^{pro}, although significant inhibition of M^{pro} by these β-lactams was not observed. There was no clear correlation between the propensity of a β-lactam to react covalently and M^{pro} inhibition, possibly due to reaction with non-active site cysteines. In some cases, although there was evidence of a covalent reaction, no inhibition was observed (e.g. cefbuperazone and cefotetan). Cephalosporin C zinc salt was the most potent β-lactam inhibitor identified. However, the inhibition arises from the zinc metal ion. The results with *S*-oxidised penicillin β-lactam derivatives **4** and **5** reveal that some β-lactams can selectively and covalently react with the active site Cys145 to inhibit M^{pro}, and could thus be further optimised as potential chemotherapeutics for COVID-19 antiviral therapy.

2.4 Methods

2.4.0.1 Reagents

All compounds were purchased from Sigma-Aldrich, Apollo Scientific or MedChemExpress unless otherwise stated.

2.4.0.2 Small scale protein production and purification

For recombinant protein production, a starter culture of BL21(DE3) cells transformed with the pGEX-6p1-M^{pro} plasmid [28,248] was grown in 2TY medium supplemented with 50 μg/mL ampicillin overnight at 37 °C. This culture was used to inoculate (1%$_{v/v}$) 100 mL 2-Ty culture, which was grown at 37 °C, 180 rpm to an OD of 0.5-0.8, induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubated overnight at 18 °C. The induced cultures were harvested using an Avanti J-26S XP centrifuge and JLA25.50 rotor (20000 rpm, 10 min). The cell pellets were frozen on dry ice and stored at -80 °C. For protein purification, the cell pellets were first resuspended in HisTrap Buffer A (50 mL, 50 mM Tris pH 8, 300 mM NaCl, 20 mM imidazole) with added DNase I for 15 min using a magnetic stirrer. The cells were then lysed on ice using a SONIC Vibra-Cell sonicator at 60% amplitude with a 3 s on cycle and 3 s off cycle for 10 min. The lysed cells were centrifuged using an Avanti J-26S XP centrifuge and JLA 16.250 rotor for 30 min at 16,000 rpm, 4 °C. The supernatant was filtered using 0.45 μm syringe-tip filters and loaded onto a 1 mL HisTrap column (GE Life Sciences) pre-equilibrated in HisTrap Buffer A using an AKTA fast liquid protein chromatography (FPLC) system. The loaded column was washed with 100 mL of HisTrap Buffer A, then protein was eluted at 2 mL/min using a gradient from 20 mM - 500 mM imidazole over the course of 20 min. The eluted protein was concentrated using a 10 KDa MWCO Amicon centrifugal filter at 4000 rpm, 4 °C using an Allegra X-30R centrifuge equipped with an SX4750 rotor to a final volume of approximately 2 mL. N-terminal His tagged HRV 3C protease was added to the eluted M^{pro} fractions (1:100 molar excess protease:M^{pro}). The protein mixture was dialysed overnight (50 mM Tris pH 8, 300 mM NaCl, 0.5 mM TCEP, 8 °C). The eluted fractions were pooled together as described above and the concentration

was measured using a ND1000 NanoDrop spectrometer ($\mathcal{E}$ = 32890 M^{-1} cm^{-1}, M^{pro} MW = 33796.64 Da (predicted using exPASy – compute pI/MW tool [304]) as input parameters.

2.4.0.3 Protein production and purification

The reported pGEX-6p1-M^{pro} plasmid was used for M^{pro} protein production [248]. In brief, the pGFX-6p1-M^{pro} plasmid was transformed into a competent *Escherichia coli* expression cell line based on the BL21(DE3)-R3-pRARE. Multiple colonies were used to inoculate an LB media starter culture supplied with 100 µg/mL carbenicillin. The starter culture was incubated (37 °C, 200 rpm) until OD600 $\leq$2 was reached (∼8 h). 1% $_{(v/v)}$ of the starter culture was used to inoculate pre-warmed autoinduction medium (Formedium, Terrific Broth base including trace elements) supplemented with 1% $_{(v/v)}$ glycerol and 50 µg/mL carbenicillin. Cells were grown at 37 °C (200 rpm) for 5 h and at 18°C overnight (200 rpm). Cells were harvested by centrifugation (10 krpm, 30 min, 4 °C). The resulting cell pellet was suspended in lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 10 mM imidazole, 0.03 µg/ml benzonase) at a ratio of 1:3 $_{(w/v)}$. Cells were lysed using an Emulsiflex high-pressure homogeniser (Avestin Inc) (3 passes, 30 kpsi, 4 °C). The lysates were centrifuged (50000 g, 30 min-60 min, 4 °C) and then filtered using a 0.22 µm pore size membrane filter (MF-Millipore^{TM}). The filtrate was incubated with 10 mL 50% $_{(v/v)}$ pre-equilibrated His60 Ni Superflow resin (Clontech Laboratories) (with agitation, 1 h, 4 °C) and then applied to a gravity flow column. The column was washed (50 mM Tris pH 8, 300 mM NaCl, 25 mM imidazole), and M^{pro} was eluted with elution buffer (50 mM Tris pH 8, 300 mM NaCl, 500 mM imidazole). To remove the M^{pro} poly-histidine tag, the N-terminal His tagged HRV 3C protease was added to the eluted M^{pro} fractions (1:10$_{(w/w)}$ protease:M^{pro}). The mixture was dialysed overnight into 50 mM Tris pH 8, 300 mM NaCl, 0.5 mM TCEP at 4 °C. M^{pro} was purified by reverse His purification using a 5 mL HisTrap FF column (GE Healthcare) eluting with wash buffer (50 mM Tris pH 8, 300 mM NaCl, 25 mM imidazole). Purified M^{pro} fractions were pooled and concentrated (<5 mL) using an Amicon® Ultra 10 MWCO centrifugal filter unit (Merck Millipore). The concentrate was loaded onto an S200 pg 16/60 size exclusion column (GE Healthcare), eluting at a flow rate of 1 mL/min (50 mM Tris pH 8, 300 mM NaCl). M^{pro} fractions were concentrated to 36 mg/mL using a 10

kDa MWCO centrifugal filter. M^{pro} was >95% pure by SDS PAGE analysis and had the anticipated mass of 33786 ± 1 Da (calculated: 33796.64 Da). C-terminal His/Avi tagged M^{pro} was produced in a similar manner via omission of His cleavage and reverse His column purification steps.

The sequences of M^{pro} constructs used were:

WT M^{pro}: SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSED MLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRI QPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHM ELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRF TTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRTIL GSALLEDEFTPFDVVRQCSGVTFQ

M^{pro} GP-10 His: SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICT SEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFV RIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHH MELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNR FTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRTIL GSALLEDEFTPFDVVRQCGVTFQGPHHHHHHHHHH

M^{pro} SG-10 His: SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICT SEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFV RIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHH MELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNR FTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRTI LGSALLEDEFTPFDVVRQCSHHHHHHHHH

M^{pro} GP(SG2)-10 His : SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPR HVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPK YKFVRIQPGQFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYM HHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFL

NRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGR TILGSALLEDEFTPFDVVRQCSGVTFQGPSGGSGGHHHHHHHHHH

M^{pro} Avi: SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDM LNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRIQP GQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELP TGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRFTTTL NDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNGMNGRTILGSAL LEDEFTPFDVVRQCSGVTFQGPGLNDIFEAQKIEWHE

2.4.0.4 SDS-PAGE Gel Electrophoresis

NuPAGE™ Bis-Tris, 4 to 12% gradient gel (ThermoFisher), NuPAGE™ MOPS SDS Running Buffer (20X) (ThermoFisher) and PageRuler™ Prestained Protein Ladder (10 – 170 kDa) (Thermo Scientific™) were used. 5-10 µL sample was loaded into the gel and run at 160 V for 50 min. Gels were visualised and imaged using Image Lab (Bio-Rad Laboratories).

2.4.1 Peptide synthesis and purification

The 11-mer M^{pro} substrate peptide TSAVLQ/SGFRK-NH_2, corresponding to the N-terminal self-cleavage site of M^{pro} on pp1a/b, was synthesised as a C-terminal amide by microwave-assisted solid-phase peptide synthesis (SPPS) as reported [253, 267]. In brief, the M^{pro} substrate peptide was synthesized on a 0.1 mmol scale from the C- to N-terminus using Rink amide-MBHA resin (100–200 mesh, 0.6–0.8 $mmol.g^{-1}$ loading, AGTC Bioproducts) using a LibertyBlue peptide synthesizer (CEM) and N-Fmoc protected α-amino acids (CS Bio, Novabiochem, Sigma-Aldrich, TCI, Alfa Aesar, Merck or AGTC Bioproducts). A mixture of *N*,*N*′-diisopropylcarbodiimide (TCI Europe) and Oxyma Pure (Merck) in DMF was used for coupling and 20% $_{(v/v)}$ piperidine in DMF (peptide synthesis grade, AGTC Bioproducts) for Fmoc deprotection. Iterative cycles of peptide coupling and deprotection were performed using the manufacturer's protocol. After the final Fmoc-deprotection step, the resin was washed with CH_2Cl_2, dried in air, then treated with 5 mL of the deprotection

solution (2.5%:2.5%:2.5%:92.5%)$_{(v/v)}$ of 1,3-dimethoxybenzene, triisopropylsilane, MilliQ water and trifluoroacetic acid) (4 h at ambient temperature). The resulting mixture was filtered; the filtrate was diluted with cold Et_2O (45 mL) to precipitate the peptide. The resulting suspension was centrifuged (4255 g, 10 min, 4.0 °C) and the supernatant was decanted. The solid peptide was dried in air, dissolved in MilliQ water, then lyophilised. The peptide was purified by reverse phase HPLC (20 mL/min; linear gradient over 10 min: 2%→20% acetonitrile in water, each containing 0.1%$_{(v/v)}$ formic acid; t_R = 8 min) using a Shimadzu HPLC purification system (composed of DGU-20A, 2 LC-20AR, CBM-20A, SPD-20A, and FRC-10A units) equipped with a NX-C18 LC column (250 × 21.2 mm, 110 Å; Phenomenex Gemini). The purified peptide was dissolved in DMSO and its concentration was determined by ^{1}H NMR using an AVIII 700 machine (Bruker) equipped with a 5 mm ^{1}H(^{13}C/^{15}N) inverse cryoprobe: 16 µL of the DMSO peptide solution were added to 143 µL DMSO-*d6* and 1 µL of 1 mg/mL 3-(trimethylsilyl)propionic-2,2,3,3-*d4* acid sodium salt (TSP; Apollo Scientific) (Acquisition parameters: d1 = 30 s, ns = 256). Peaks were assigned in MestreNova using the global spectrum deconvolution algorithm. The peptide concentration was then calculated using the following equation: Mx/My=Ix/Iy×Ny/Nx, where, Mx/My is the molar ratio of TSP and the 11-mer substrate peptide respectively Ix/Iy is the ratio of signal intensity corresponding to the methyl protons of TSP and the 11-mer substrate peptide (averaged for the Ala, Val, and Leu residues) respectively, and Ny/Nx is the number of nuclei responsible for the signal intensity for 11-mer substrate (averaged for the Ala, Val, and Leu residues) and TSP respectively. Peptide aliquots were prepared and stored at -20°C.

2.4.2 General information on the M^{pro} SPE-MS assays

MS assays were always performed using freshly thawed M^{pro}, which was diluted to 15 µM in freshly prepared 20 mM HEPES (Gibco) buffer pH 7.5, 300 mM NaCl for all MS assays. An exception was for protein observed MS assays, where M^{pro} was diluted to 1 µM in freshly prepared reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). LC-MS (Merck, Supelco) grade solvents were used for MS analyses and to prepare buffers. All assays were performed in 384 well polypropylene plates (Greiner). All compounds were purchased from Sigma Aldrich, Apollo, and Fluorochem. All compounds were

dry dispensed using an Echo 550 acoustic dispenser (LabCyte) unless otherwise stated. The substrate peptide and the enzyme were diluted in the reaction buffer to prepare stock solutions, 2-fold concentrated relative to the final concentration in the assay. M^{pro}, peptide substrate, and formic acid solutions were all dispensed across 384 well plates using a Multidrop Combi dispenser (Thermo Scientific), the plates were then centrifuged (500 x g, 15s, Axygen Plate Spinner Centrifuge, Corning).

M^{pro} activity was assayed using a RapidFire (RF) 365 high-throughput sampling robot (Agilent) connected to an iFunnel Agilent 6550 accurate mass quadrupole time-of-flight (Q-TOF) mass spectrometer. The spectrometer was operated in the positive ion ionisation mode with the following operating parameters for all assays: capillary voltage (4000 V), nozzle voltage (1000 V), fragmentor voltage (365 V), drying gas temperature (280 °C), gas flow (13 L/min), sheath gas temperature (350 °C), sheath gas flow (12 L/min); except for protein observe MS assays, for which the gas temperature was reduced to 225 °C. The peptide/protein sample was loaded onto a solid-phase extraction (SPE) C4-cartridge, which was then washed with 0.1% $_{(v/v)}$ aqueous formic acid to remove non-volatile buffer salts (5.5 s, 1.5 mL/min) and with aqueous 85% $_{(v/v)}$ acetonitrile containing 0.1% $_{(v/v)}$ formic acid (5.5 s, 1.25 mL/min) to elute the peptides/protein. The cartridge was re-equilibrated with 0.1% $_{(v/v)}$ aqueous formic acid (0.5 s, 1.25 mL/min). Each sample aspiration step was followed by an aqueous wash before the next protein sample was injected onto the SPE cartridge.

2.4.3 Kinetic parameters of M^{pro}

The kinetic parameters of the M^{pro} were determined by monitoring M^{pro}-catalysed cleavage of TSAVLQ/SGFRK-NH_2 over time using 16 different peptide concentrations (the final DMSO concentration was >1%$_{(v/v)}$in each experiment, Figure 2.14). The peptide was diluted (1:100) with the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl) to 0.5 mL reaction volume. Initially, a no-enzyme control was analysed using SPE-MS before M^{pro} was added to the mixture (0.15 μM final concentration, based on the M^{pro} monomer mass) using SPE-MS.

The RapidFire integrator software (Agilent) was used to extract and integrate the m/z (+1) charge states of both the substrate peptide (1191.67 Da) and the N-terminal product peptide (TSAVLQ; 617.34

Da). The basic C-terminal cleavage product (SGFRK-NH_2) was not detected, likely due to poor retention on the SPE cartridge under these conditions, and hence was not included in subsequent calculations. The percentage conversion (product peak integral/(product peak integral + substrate peak integral)*100) was calculated in Microsoft Excel. The slopes of the initial reaction rate were fitted with the Michaelis-Menten equation [$Y = V_{max}*X/(K_m + X)$, E_t= 0.15 µM] using non-linear regression (GraphPad Prism 8) to obtain the K_m and k_{cat} values (Figure 2.14). The experiments were performed in independent duplicates (n = 2; mean ± standard deviation, SD).

2.4.3.1 Turnover assay with different M^{pro} constructs

The DMSO stock solution of the peptide was diluted to 2 µM with the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl) to 0.5 mL reaction volume. Initially, a no-enzyme control was analysed using SPE-MS before M^{pro} was added to the mixture (0.10 µM final concentration, based on the M^{pro} monomer mass) using SPE-MS. Data analysed as described above.

2.4.4 Single concentration M^{pro} inhibition assays

The final compound assay concentrations were 20 µM, 50 µM, or 100 µM; DMSO and ebselen were used as negative and positive inhibition controls, respectively. M^{pro} (0.30 µM) was dispensed across the plate (25 µL/well); the resulting mixture was incubated for 30 or 60 min at ambient temperature. TSAVLQ/SGFRK-NH_2 peptide (4 µM) was dispensed to the mixture (25 µL/well), which was incubated (10 min), then quenched by addition of 10% $_{(v/v)}$ aqueous formic acid (5 µL/well). The individual reactions were analysed using SPE-MS. Data were extracted and processed as described for the kinetic M^{pro} assays above and the normalized percentage inhibition was determined in Microsoft Excel with respect to the positive and negative inhibition controls. Errors values were determined with respect to technical duplicates (n = 2; mean ± SD).

2.4.4.1 In-house single concentration M^{pro} inhibition assays

The final compound assay concentrations were 20 µM, 50 µM, or 100 µM. 1 mM, 2.5 mM, or 5 mM DMSO stocks from 96 well plates were transferred to 384 well plates in technical duplicates

using a CyBio-Liquid handling robot (Analytik Jena) (DMSO 1%). Assay setup and analysis was performed as described above.

2.4.5 IC_{50} value determinations

M^{pro} inhibitors were dry dispensed in an 11-point 3-fold dilution series (100 µM top concentration). Ebselen and DMSO were used as positive and negative inhibition controls, respectively (DMSO <1%). For in-house screening work, plates with 11-point 3-fold dilution series were manually prepared in 384 well plates from 2.5 mM stocks. Compounds were transferred to a destination 384 well plate (2 µL/well in technical duplicates) using E1-ClipTip™ Electronic Multichannel Pipettes. M^{pro} (0.30 µM) was dispensed across the plate (25 µL/well) and the resulting mixture was incubated (30 or 60 min) at ambient temperature. TSAVLQ/SGFRK-NH_2 peptide solution (4 µM) was then dispensed to the mixture (25 µL/well). Reactions were incubated (10 min), then quenched by the addition of 10% *(v/v)* aqueous formic acid (5 µL/well). Data were extracted and processed as described for the kinetic M^{pro} assays above. IC_{50} curves were generated using the four-parameter sigmoidal curve equation Y = Bottom + (Top-Bottom)/(1+$10^{X-LogIC_{50}}$), where the top and bottom values were constrained at 100% and 0%, respectively and X represents the log concentration of the inhibitor. Curves were fitted using a least square non-linear regression method(GraphPad Prism 8). IC_{50}-values are reported as the mean of two independently determined IC_{50}-curves each composed of technical duplicates (n = 2; mean ± SD). Signal to noise (S/N) and Z′-factor was calculated using the equations i) S/N = $\frac{\mu_p - \mu_n}{\sqrt{\sigma_p^2 - \sigma_n^2}}$ and ii) $Z - factor = 1 - \frac{3(\sigma p + \sigma n)}{|\mu p - \mu n|}$ in Microsoft Excel for each plate analysed [256].

2.4.6 Protein Observed Mass Spectrometry Assay

For the initial screening of β-lactams, β-lactam solutions (10 mM in DMSO) were transferred in quadruplicate from a 96 well plate to a 384 well plate (1 µL/well) using a CyBio Liquid Handler (Analytik Jena AG). M^{pro} (1.0 µM) in the reaction was dispensed across the plate (100 µL/well) using a Multidrop Combi dispenser (Thermo Scientific). The resulting mixture was incubated for a minimum of >25 min, 5 h, 19 h, and 24 h at ambient temperature before analysis using SPE-MS.

For selected M^{pro} inhibitors identified from the initial screening of β-lactams or from the screening of small-molecule libraries [LOPAC^{1280} (Sigma Aldrich) or the small-molecule library containing FDA approved compounds for human use (Pharmacon)], inhibitor solutions (20 mM in DMSO) were diluted to 1 mM with the reaction buffer and transferred to a 384 well plate (2 µL/well). M^{pro} (1.0 µM) was dispensed across the plate (100 µL/well) using a Multidrop Combi dispenser (Thermo Scientific) and the resulting mixture was incubated for a minimum of (5 min and 1 h) at ambient temperature before analysis using SPE-MS. To investigate whether the identified M^{pro} inhibitors selectively modify the M^{pro} active site Cys-145, M^{pro} (1 µM) was incubated with the reported Cys-145-selective M^{pro} inhibitor N3 [28] (3 µM) on ice for 3 h in 7.5 mL of the reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). The inhibitor of interest (20 mM in DMSO) was diluted to 1 mM with the reaction buffer and transferred to a 384 well plate (2 µL/well). M^{pro}/N3 mixture was dispensed to the inhibitor of interest on a 384 well plate (100 µL/well) and the resulting mixture was incubated at ambient temperature for 5 min and 16 h, respectively, before analysis using SPE-MS.

Protein observed mass spectra were acquired (at least) in independent duplicates using SPE-MS with the instrument parameters as stated in Section 2.4.2. Only one representative of the data is shown as all duplicates were apparently identical. The reactions were not quenched. Thus, each subsequent sample was injected with a delay from the primary injection time (stated in the individual figures). Protein spectra were deconvoluted (m/z range: 10-60 kDa, m/z range 850-1350 Da, mass step: 1 Da) using the MaxEnt1 function in Agilent MassHunter Version 7. The deconvoluted files were extracted, normalised, and plotted (GraphPad Prism 8).

3

SAR studies with esters of β-lactam penicillin sulfones and γ-lactam derivatives against M^{pro}

Contents

3.1 Abstract **128**
3.2 Introduction **130**
3.3 Results **134**
3.4 Discussion **152**
3.5 Future directions **153**
3.6 Methods **153**
3.7 Appendix **154**

3.1 Abstract

The work described in this chapter was carried out following the initial observation that the penicillin V sulfone benzyl ester inhibits M^{pro} (Chapter 2). The objectives of these studies were to investigate the effect of the penicillin C3 ester group, the C6 side chain, and the *S*-oxidation state in the thiazolidine ring on M^{pro} inhibition. The results indicate that of the tested compounds, the C3 ester group does not substantially affect the inhibitory activity. Modification of the penicillin C6 amido side chain showed that the presence of the penicillin V phenoxy ether oxygen is important in maintaining a high level of M^{pro} inhibition. An analogous γ-lactam derivative also moderately inhibits M^{pro}. Studies with penicillins bearing 6-Bromo substituents led to the identification of potent M^{pro} inhibitors. A crystal structure of M^{pro} complexed with a penicillin sulfone ester derivative revealed that β-lactams

covalently modify the Cys145 thiol of M^{pro}; in the crystal structure, the five membered thiazolidine ring was observed to open to form an enamine and a sulfinic acid.

3.2 Introduction

Clavulanic acid and semi-synthetic penicillin sulfones (tazobactam and sulbactam) are widely used broad-spectrum inhibitors of serine β-lactamases [305]. Whilst ineffective antibiotics by themselves, in the presence of serine β-lactamases they show synergistic effects when used in combination with penicillin antibiotics for inhibition of PBPs. C2-substituted 6-alkylidene penicillin sulfones (e.g. LN-I-255) have also been pursued as class D serine-β-lactamase inhibitors [306, 307]. However, they have not been widely investigated for the inhibition of cysteine proteases.

3.2.1 Penicillin sulfones as M^{pro} inhibitors

Penicillin sulfone β-lactamase inhibitors did not display any inhibitory activity against M^{pro} (as described in Chapter 2, Figure 2.23,Table 2.7). The C3-ester prodrugs of ampicillin, mecillinam, carbenicillin, bacampicillin, talampicillin, and pivampicillin also did not potently inhibit M^{pro} (Table 2.5, Figure 3.1). However, oxidation of the thioether group of the penicillin V benzyl ester to give sulfone **4** did result in M^{pro} inhibition [253], likely by the covalent modification of the active site Cys145. Penicillin sulfones have been successfully optimised to inhibit other cysteine proteases with high potency [308]. The latter observation suggested that the potency of penicillin sulfone inhibitors against M^{pro} could be improved by conducting SAR studies.

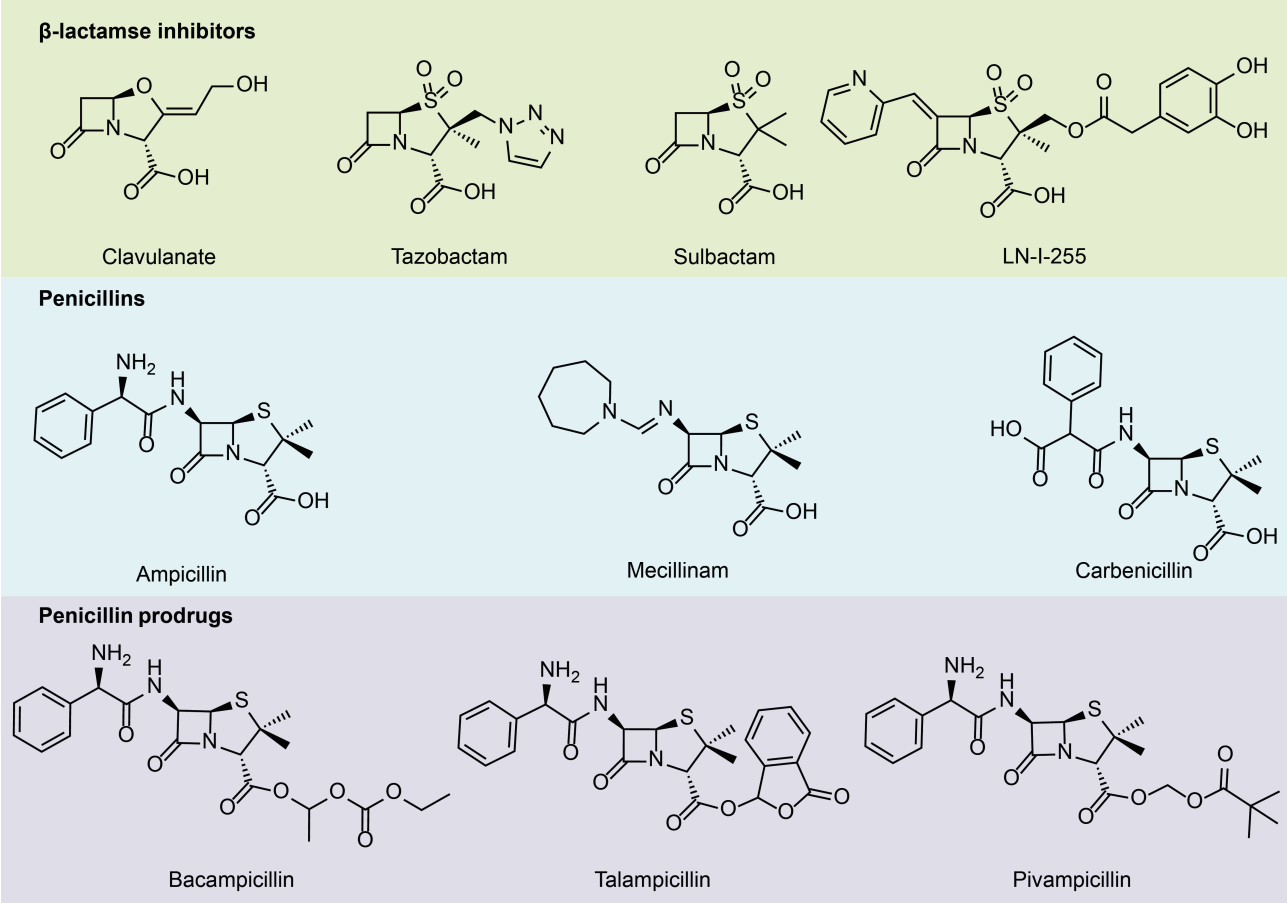


Figure 3.1 | Penicillin sulfone β-lactamase inhibitors and C3 ester penicillin prodrugs.

3.2.2 γ-Lactams as M^{pro} inhibitors

Chapter 2 describes studies on M^{pro} inhibition by monocyclic γ-lactams which displayed weak inhibitory activity (Figure 2.23, 2.24). γ-Lactams have been exploited as serine protease inhibitors [309] but they are yet to be widely explored as electrophilic warheads for cysteine protease inhibitors [310]. Monocyclic γ-lactams are excellent mimics of Gln [37,311], and have been employed at the P1 position in peptidomimetic inhibitors of M^{pro} (Figure 1.12A). However, studies with monocyclic γ-lactams showed them to be poor inhibitors of M^{pro} (Figures 2.23B, 2.24). Therefore, bicyclic compounds, in which a γ-lactam ring was fused to a 5-membered thiazolidine ring to mimic the bicyclic penicillin *S*-oxidised inhibitors, were screened for activity against M^{pro}. Similar to the oxidised penicillin esters tested previously, the thioethers of γ-lactam esters were oxidised to sulfoxides and sulfones to afford a set of bicyclic γ-lactams with different sulfur oxidation states (Figure 3.2).

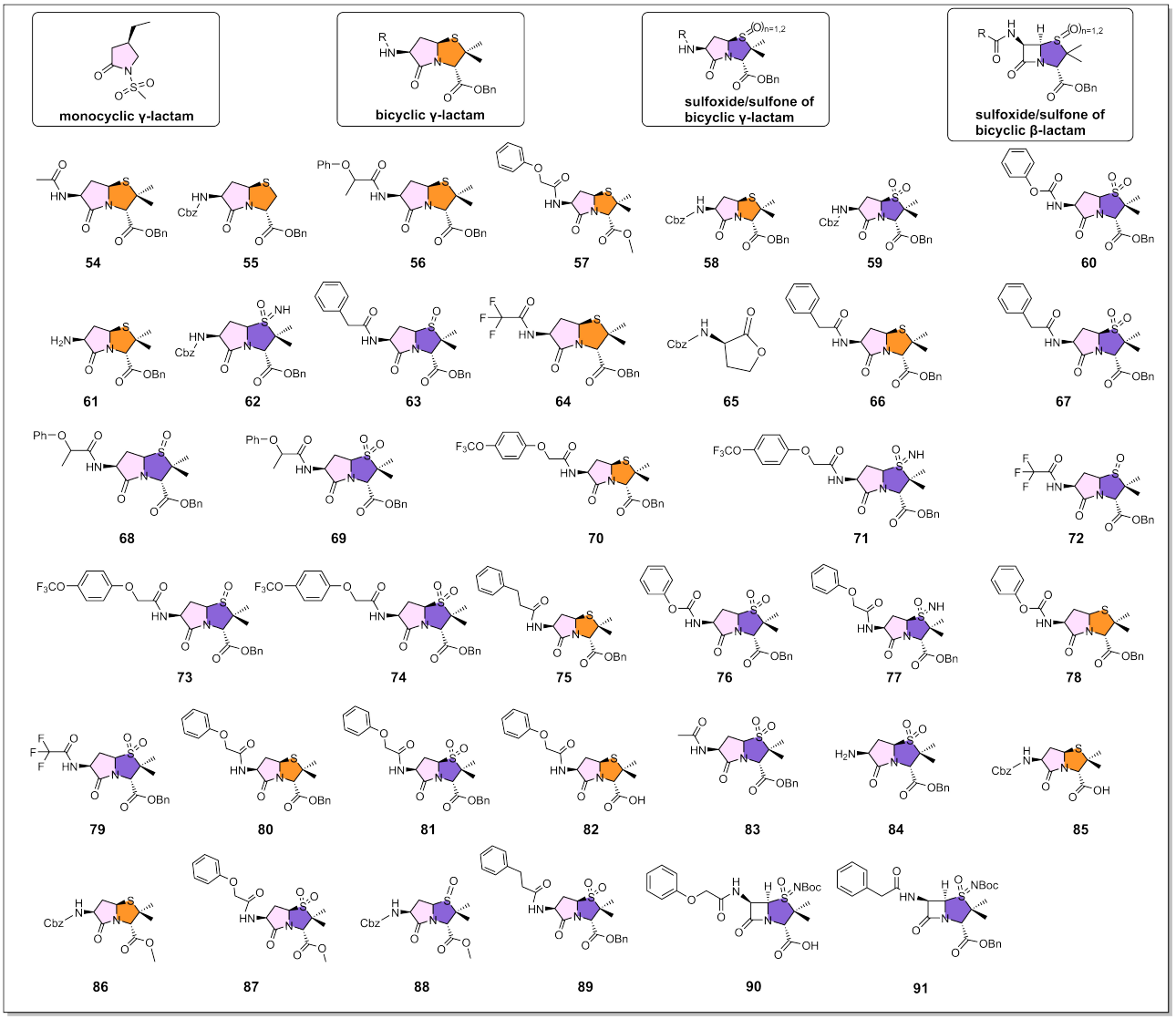


Figure 3.2 | Bicyclic γ-lactam derivatives with a γ-lactam ring (pink) fused with a 5-membered thiazolidine ring (orange). The thioether of the thiazolidine ring was oxidised to a sulfoxide or sulfone (purple) to mimic the penicillin sulfoxide and sulfone M^{pro} inhibitors (purple).

Sulfoximine scaffolds as M^{pro} inhibitors

Sulfoximines, which are monoaza analogues of sulfones and sulfonamides, have emerged as isosteres of these sulfur-bearing functional groups in medicinal chemistry (Figure 3.2) [312]. Asymmetric substitution at the tetrahedral sulfur heteroatom can give rise to a stereocenter with a basic nitrogen atom [pKa (NH_2^+) = 2.7], improving the solubility and chemical stability of sulfoximines over the corresponding sulfones [313]. Substitution at the nitrogen atom can allow for modulation of the hydrogen bond donor/acceptor capabilities of sulfoximines. As sulfones were crucial for the activity of penicillin

inhibitors against M^{pro}, sulfoximine analogues of γ-lactams and penicillin were synthesised by Hiba Aslam under the supervision of Dr. Lennart Brewitz and Prof. Christopher J. Schofield.

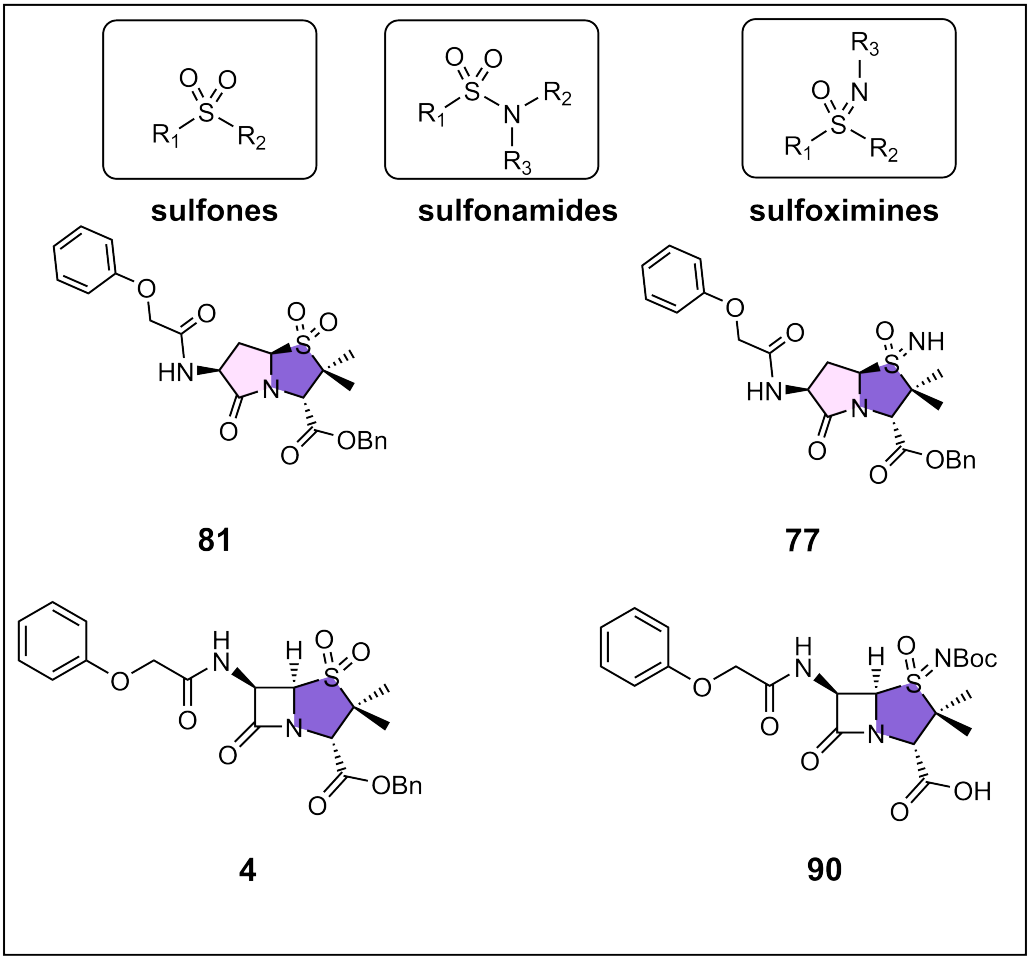


Figure 3.3 | Sulfoximine derivatives of γ-lactam (pink) and β-lactam compounds.

3.2.3 Contributions to manuscript 2

This chapter describes work performed to expand on the penicillin V sulfone benzyl ester (**4**), which was identified as a M^{pro} inhibitor in chapter 2 to develop and optimise it as a potent M^{pro} inhibitor. Derivatives of compound (**4**) were synthesised by Dr. Brewitz (Department of Chemistry, University of Oxford). At the peak of the pandemic, C5/C6-stereoisomers and C6-halo penicillin derivatives and some available γ-lactams were investigated for M^{pro} inhibition. The library of C5/C6-stereoisomers and C6-halo derivatives was expanded by Dorian Muntean (Department of Chemistry, University of Oxford), and analogues of γ-lactams were synthesised by Hiba Aslam (Department of Chemistry, University of Oxford).

3.2.4 Manuscript 2: Penicillin Derivatives Inhibit the SARS-CoV-2 Main Protease by Reaction with its Nucleophilic Cysteine

My contributions to manuscript 2 [303] included:

- Development of a M^{pro} assay with a 37-mer substrate, and validation of this assay with reported M^{pro} inhibitors.
- Screening of penicillin derivatives for M^{pro} inhibition with the 37-mer substrate peptide.
- Determination of the mode of inhibition and the active site selectivity using protein-observed MS assays under denaturing conditions.

In addition to the published experiments described in manuscript 2, I also carried out the following work:

- Screening penicillin derivatives for M^{pro} inhibition with the 11-mer substrate peptide.
- Screening γ-lactams as M^{pro} inhibitors using SPE-MS and FRET assays.
- Further screening of synthetic β-lactams, which led to the identification of C6 bromopenicillin sulfones as the most potent β-lactam inhibitors of M^{pro}.
- Studies on the mode of M^{pro} inhibition by bromopenicillins using protein-observed MS under denaturing conditions.

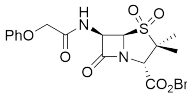
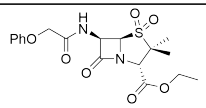
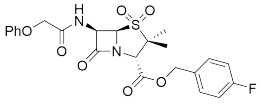
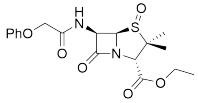
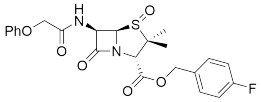
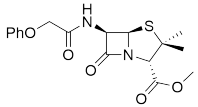
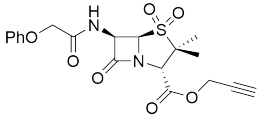
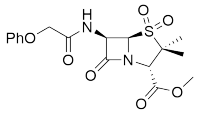
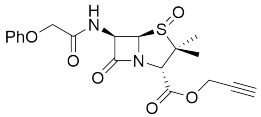
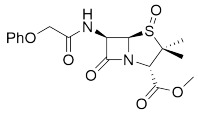
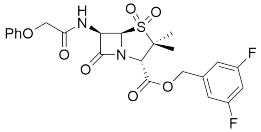
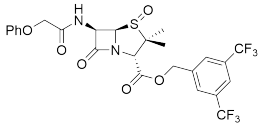
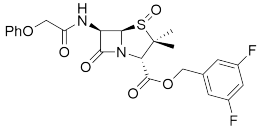
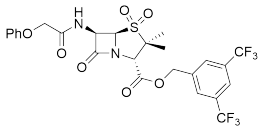
3.3 Results

3.3.1 Investigating the impact of the 11-mer and the 37-mer substrates on MS assays with β-lactam derivatives

β-Lactams were initially investigated for M^{pro} inhibition using the 11-mer substrate peptide (TSAVLQ /SGFRK-NH_2). However, a flat structure activity relationship was observed with weak M^{pro} inhibitors when using the 11-mer substrate. Therefore, a M^{pro} SPE-MS assay employing a 37-mer substrate

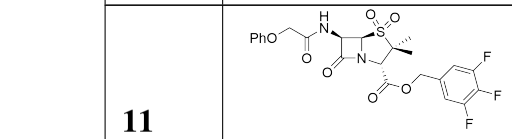
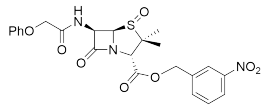
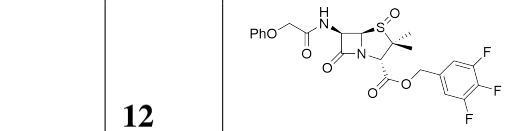
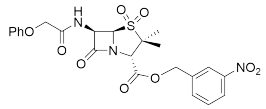
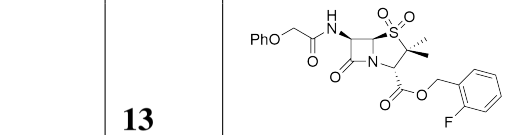
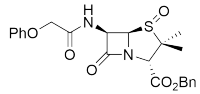
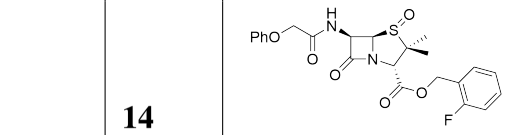
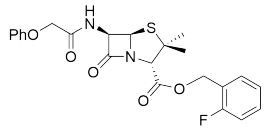
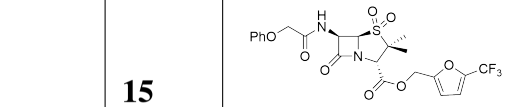
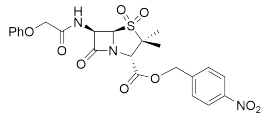
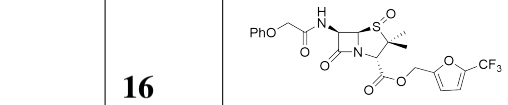
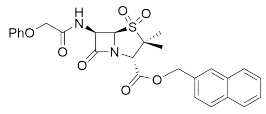
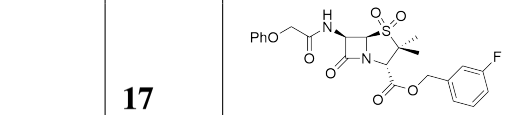
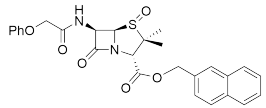
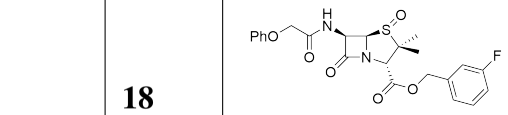
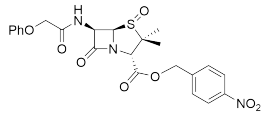
(ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2) (Figure 3.4) was developed to improve substrate and product ionisation. IC_{50} values obtained for selected penicillin sulfones with both 11-mer and 37-mer substrates are given in Tables 3.1, 3.2 and 3.3.

Table 3.1: IC_{50}s of penicillin sulfone/sulfoxide C3-ester derivatives. [a] Represents IC_{50}s determined using the 11-mer substrate and [b] represents IC_{50}s determined using the 37-mer substrate. Assays were performed as described [253,303,314]. Conditions: 0.15 µM M^{pro}, 2.0 µM 11-mer (TSAVLQ/SGFRK-NH_2) or 37-mer (ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2) in the reaction buffer. Mean of three independent replicates (n = 3; mean ± SD).

Cmpd	Structure	IC_{50}[a,b] (µM)	Cmpd	Structure	IC_{50}[a,b] (µM)
4		6.0 ± 3.1, 6.6 ± 2.7	**22**		6.7 ± 4.6, 5.2 ± 0.6
5		4.9 ± 0.2, 7.43 ± 2.3	**23**		>50, 28.5 ± 15.2
6		>50, >50	**24**		>50, >50
7		3.0 ± 1.9, 3.0 ± 0.4	**25**		4.6 ± 3.6, 7.9 ± 1.4
8		>50, 42.9 ± 5.8	**26**		>50, 48.9 ± 25.1
9		>50, 17.0 ± 11.3	**27**		>50, 48.3 ± 9.6
10		>50, >50	**28**		>50, 35.3 ± 2.2

Continued on next page

Table 3.1 – continued from previous page

Cmpd	Structure	IC_{50}[a,b] (μM)	Cmpd	Structure	IC_{50}[a,b] (μM)
11		>50, 45.7 ± 19.7	**29**		>50, >50
12		>50, >50	**30**		>50, 13.3 ± 5.3
13		5.7 ± 1.4, 5.8 ± 2.0	**31**		>50, 22.9 ± 13.3
14		>50, 49.3 ± 9.9	**32**		>50, >50
15		16.0 ± 2.2, 10.2 ± 2.2	**33**		>50, >50
16		>50, 43.3 ± 19.7	**34**		>50, 70.6 ± 2.7
17		24.4 ± 2.3, 9.2 ± 3.7	**35**		>50, 41.8 ± 5.4
18		>50, 34.0 ± 8.8	**36**		>50, >50

Continued on next page

Table 3.1 – continued from previous page

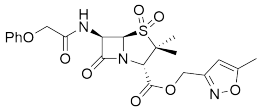
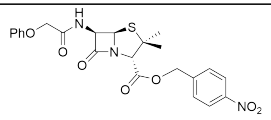
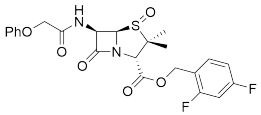
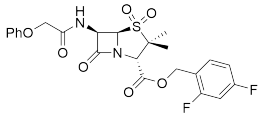
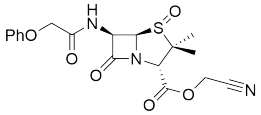
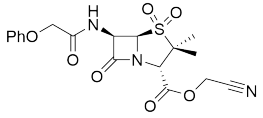
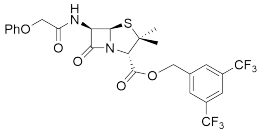
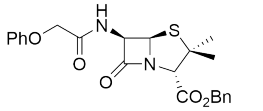
Cmpd	Structure	IC_{50}[a,b] (μM)	Cmpd	Structure	IC_{50}[a,b] (μM)
19		14.1 ± 2.3 , 9.0 ± 2.3	**37**		>50, 19.7 ± 1.9
20		>50 , 19.4 ± 3.7	**38**		7.9 ± 1.6, 3.6 ± 0.98
21		>50, 32.1 ± 13.9	**39**		>50, 9.7 ± 3.2
22		>50, 73.5 ± 30.3	**40**		>50, >50

Table 3.2: IC_{50}s of penicillin sulfone/sulfoxide C3 ester and C6 derivatives. [a] = IC_{50}s determined using the 11-mer substrate and [b] = IC_{50}s using the nsp4/5 cleavage site derived 37-mer substrate. Assays were performed as described [253,303,314]. Conditions: 0.15 µM M^{pro}, 2.0 µM 11-mer (TSAVLQ/SGFRK-NH_2) or 37-mer (ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2) in the reaction buffer. Mean of three independent replicates (n = 3; mean ± SD).

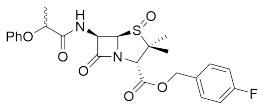
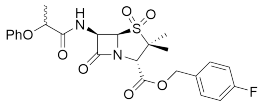
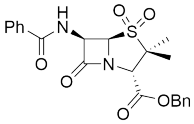
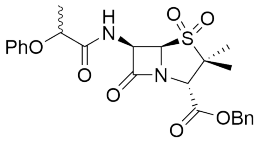
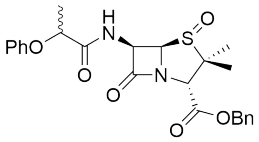
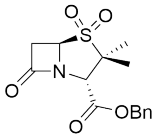
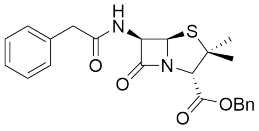
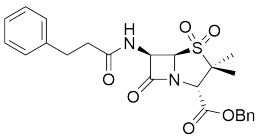
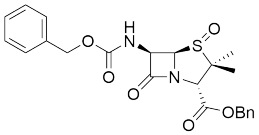
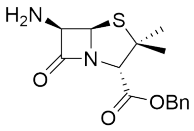
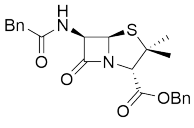
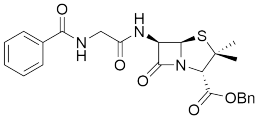
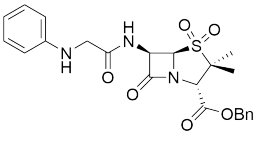
C3 ester and C6 derivatives	Structure	IC_{50}[a,b] (µM)
41		>50, >50
42		>50, 41.8 ± 5.3

Table 3.3: IC_{50}s of C6 derivatives of penicillin sulfone/sulfoxide esters. [a] and [b] represent IC_{50}s obtained using the nsp4/5 cleavage site derived 11-mer and 37-mer substrate peptides, respectively. Assays were performed as described [253, 303, 314]. Mean of three independent replicates (n = 3; mean ± SD).

C6 derivatives	Structure	IC_{50}[a,b] (μM)
43		>50, >50
44		>50, >50
45		>50, >50
46		>50, >50
47		>50, >50
48		>50, 45.7 ± 0.9
49		>50, >50
50		>50, >50
51		>50, >50
52		>50, >50
53		>50, 41.8 ± 5.3

The SPE-MS assays employing the the 11-mer substrate peptide were effective in identifying good-to-moderately potent M^{pro} inhibitors. However, a flat structure-activity relationship was observed for less efficient M^{pro} inhibitors, such as the C3 ester and C6 derivatives of the penicillin sulfone esters (IC_{50} values >15 μM), making it difficult to rank the potencies of these compounds.

The linear relationship of product formation over time and excellent assay parameters afforded confidence in the 11-mer assays. However, MS assays can be impacted by ion suppression or enhancement which may result in false positives (product ion suppression) or false negatives (product ion enhancement) [289]. At high inhibitor concentrations, ion suppression of product/substrate peptides may occur. This may explain the discrepancies observed in the IC_{50} values determined using the shorter 11-mer substrate compared to the longer 37-mer substrates, as the longer peptide likely ionises better than the 11-mer substrate [315–317]. In contrast to the general belief that longer peptides ionise better, Liigand *et al.* [318] have reported that ionisation efficiency can begin to level off once peptides reach a certain length. However, their studies were only conducted with oligopeptides up to 14 amino acid residues in length. Thus, the use of the 37-mer substrate to address ionisation issues was a viable solution as this peptide may ionise more readily. Longer M^{pro} substrates such as the 37-mer peptide may also possess better catalytic parameters than the 11-mer peptide, as it should more closely resemble the natural M^{pro} substrate. The kinetic parameters determined for the 11-mer substrate (**TSAVLQ/SGFRK**-NH_2, k_{cat} = 23.5 min^{-1} and K_m = 14.3 μM, Figure 2.14) and the 37-mer substrate (ALNDFSNSGSDVLYQPPQTSI**TSAVLQ/SGFRK**MAFPS-NH_2, k_{cat} = 19.5 min^{-1} and K_m = 5.42 μM, Figure 3.4) using SPE-MS support this assumption.

S-Oxidised C3-penicillin esters were rescreened with assays using the 37-mer substrate. The results indicate that the compounds which manifest weak inhibition would not have been identified by the SPE-MS assays using the 11-mer substrate. Although weak M^{pro} inhibitors are of little medicinal interest, the ability to accurately determine their inhibitory properties may be useful in future SAR studies to investigate the importance of the different functional groups.

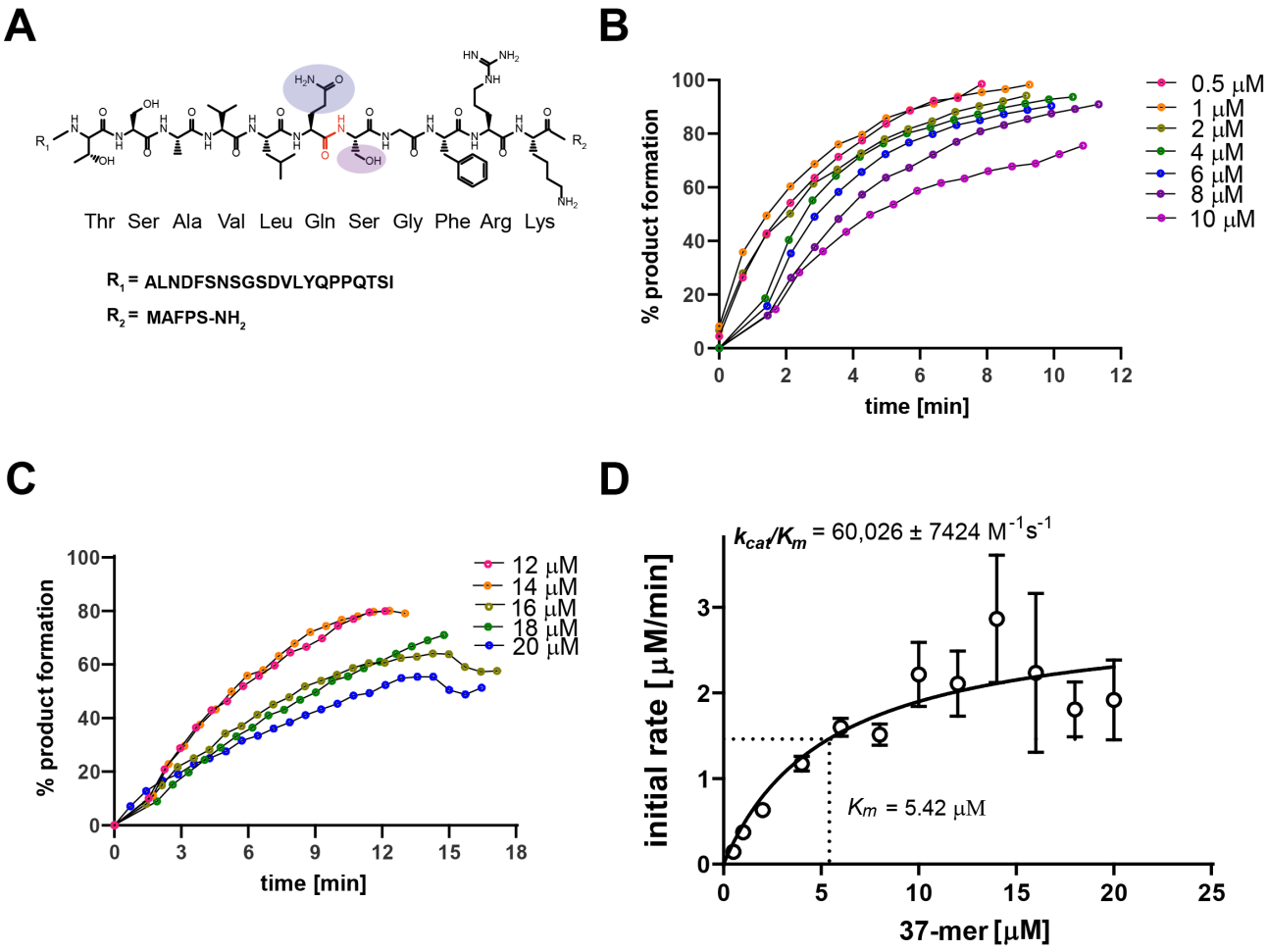


Figure 3.4 | Steady state kinetic analyses for M^{pro} with the 37-mer M^{pro} nsp4/5 substrate. A. The 37-mer substrate shares 11 residues with the 11-mer substrate. The substrate residues in the P1 position (Gln; blue) and P1′ position (Ser; purple) are shaded. M^{pro} (0.15 µM) was incubated with **B.** 0.5-10 µM and **C.** 12-20 µM of the 37-mer peptide, and conversion was monitored with SPE-MS. **D.** Initial rates and substrate concentrations were fitted to the Michaelis-Menten equation using non-linear regression. Note that interconversion between monomeric and dimeric forms of M^{pro} increases the complexity of its kinetics.

3.3.2 Comparison of SPE-MS assay results obtained with 11-mer and 37-mer substrate peptides

Scatter plot analysis shows that the pIC_{50} values obtained with the 37-mer substrate are more distributed compared with those obtained with the 11-mer substrate (Figure 3.5A). The correlation plot ($R^2 = 0.62$) indicates good agreement between the pIC_{50} values. The most difference arises from the prediction of weak inhibitors. A log value of no inhibition i.e., -log0 equals an undefined value, thus, analyses based on pIC_{50} values cannot investigate compounds that show no inhibition with 11-mer assay but

weak/moderate inhibition with the 37-mer assay. A correlation plot based on IC_{50} values clearly demonstrates good detection of weak inhibitory activity using the 37-mer peptide, whereas most weak inhibitors (highlighted in red in Figure 3.9B) have IC_{50} values of 0 (labelled as IC_{50} >100 μM) when screened using the 11-mer substrate. This is not apparent in the correlation plot derived from pIC_{50} values as it omits the compounds that showed no inhibition with the 11-mer assay (Figure 3.5B). A quantile-quantile (QQ) plot can be used to probe the normal distribution of the given data in Prism, which predicts the Y values for each of the data points, assuming that the Y values are obtained from a Gaussian distribution using the maximum-likelihood method [319]. The predicted Y values are plotted on the Y axis against the actual Y values on the horizontal X axis. If the data were truly sampled from a normal distribution, it would be expected to follow the line of identity at 45 degrees in the X-axis, matching the predicted Y values. QQ plot obtained with pIC_{50} values from 11-mer and 37-mer substrate demonstrate that they follow a normal distribution (Figure 3.5C). In the case of the IC_{50} values derived analysis, IC_{50}s obtained from the assay using the 37-mer substrate follow the line of identity more closely than the IC_{50}s obtained using the 11-mer substrate (Figure 3.9C). A non-parametric two-tailed paired Wilcoxon Signed-Rank test, performed in GraphPad Prism, indicated that the median score of the IC_{50} values obtained with the 11-mer peptide was significantly different from the median score of IC_{50} values obtained with the 37-mer peptide (median of differences = 22.4, $p < 0.0001$). IC_{50} values were most different for compounds that were shown to be inactive when using the 11-mer peptide; the statistical significance explains the superiority of the 37-mer assay in detecting the activity of weak inhibitors, and therefore this assay may be superior for SAR studies. Bland-Altman analysis that compares the differences between the data obtained from two different methods by plotting the difference between the two values obtained from different methods against the average value obtained by taking both values shows that the pIC_{50} values obtained for active compounds are highly corroborative between 11-mer and 37-mer substrate based assay (bias=-0.07 ± 0.33, Figure 3.5D). Therefore, despite the superiority of 37-mer substrate based assay, both 11-mer and 37-mer SPE-MS assays are robust and predict identical values for active compounds which are generally of pharmaceutical interest.

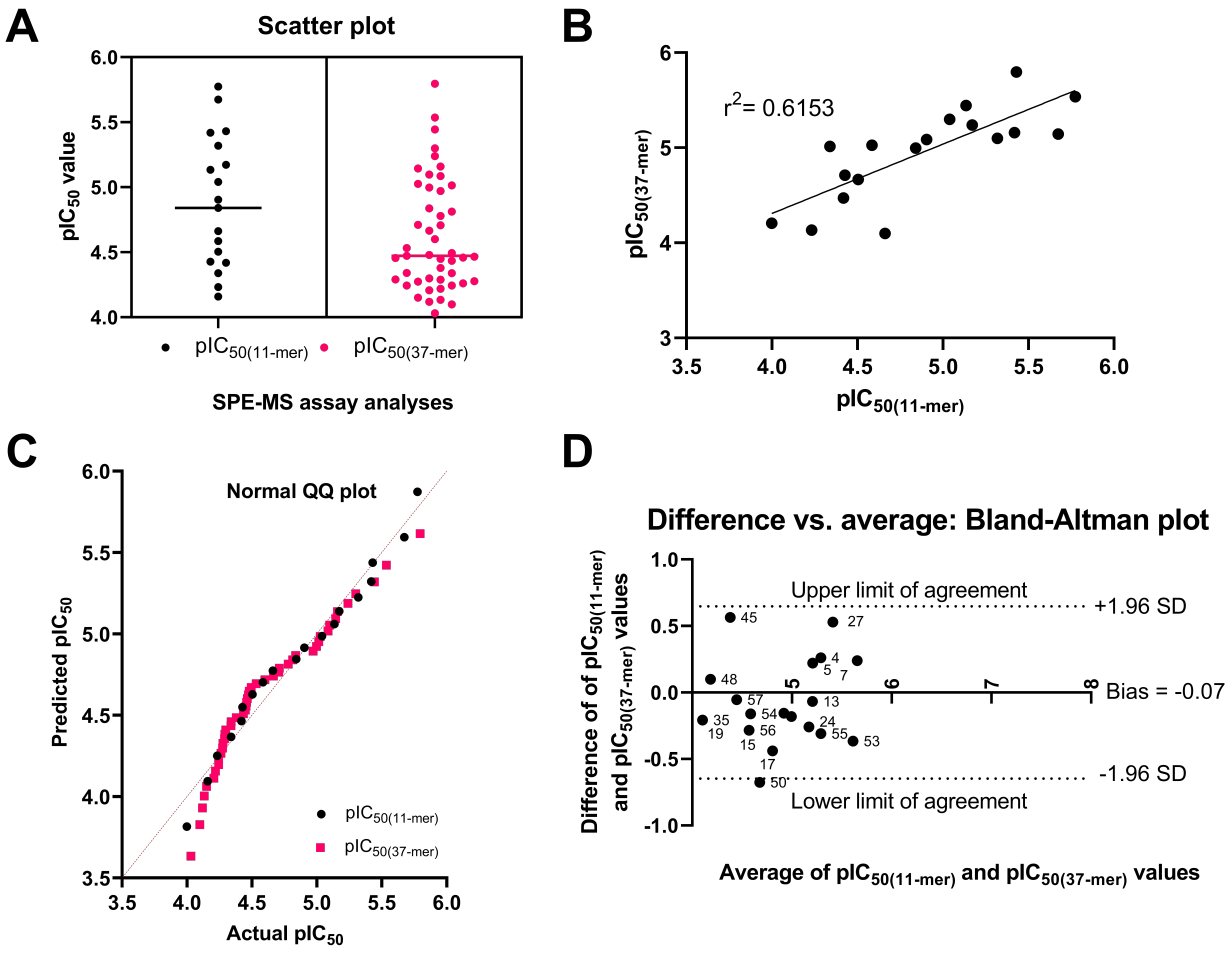


Figure 3.5 | Scatter plot, correlation plot, normal QQ plot and Bland-Altman analyses of the pIC_{50}s obtained with the 11-mer and the 37-mer M^{pro} peptide substrates. A. The scatter plot demonstrates that pIC_{50}s obtained with the 37-mer peptide are more dispersed than the IC_{50}s obtained with the 11-mer substrate. **B.** Correlation plot of pIC_{50}s obtained using the 37-mer substrate plotted against the IC_{50}s obtained with the 11-mer substrate. **C.** The normal QQ plot shows that only normal distribution of pIC_{50}s obtained with the 11-mer or 37-mer substrate. **D.** Bland-Altman plot analysis reveals low bias between the pIC_{50} values revealing both assays predict identical pIC_{50} values. Note that inhibitors that displayed no activity with 11-mer substrate could not be analysed here as a log value of no inhibition i.e., -log0 equals to an undefined value. See Appendix Figure 3.9 for scatter correlation and normal plot analyses derived using IC_{values}.

3.3.3 Investigating the mode of interaction of bromopenicillins with M^{pro} using a denaturing MS assay

Protein-observed MS studies were carried out to investigate the mechanism by which C6-halo penicillin sulfones inhibit M^{pro}. Following incubation of M^{pro} with selected C6-bromo compounds, only partial modification of M^{pro} was observed. Some evidence of protein: compound adduct formation was

observed for the dibromopenicillin sulfone [observed Δm/z = 478 Da; MW(compound) = 481 Da] (Figure 3.6A). A 42 Da modification of the protein peak was observed with **27** (dibromo penicillin sulfone with C3 para-nitrobenzyl ester), **28** (dibromopenicillin sulfone benzyl ester) and **29** (mono-bromopenicillin sulfone benzyl ester). Fragmentation of the bound adduct was observed with **31** (dibromopenicillin sulfoxide ester). Non-denaturing MS analyses were carried out with Dr. Victor Mikhailov. However, incubation of M^{pro} with the bromopenicillins resulted in a poor signal-to-noise ratio (data not shown).
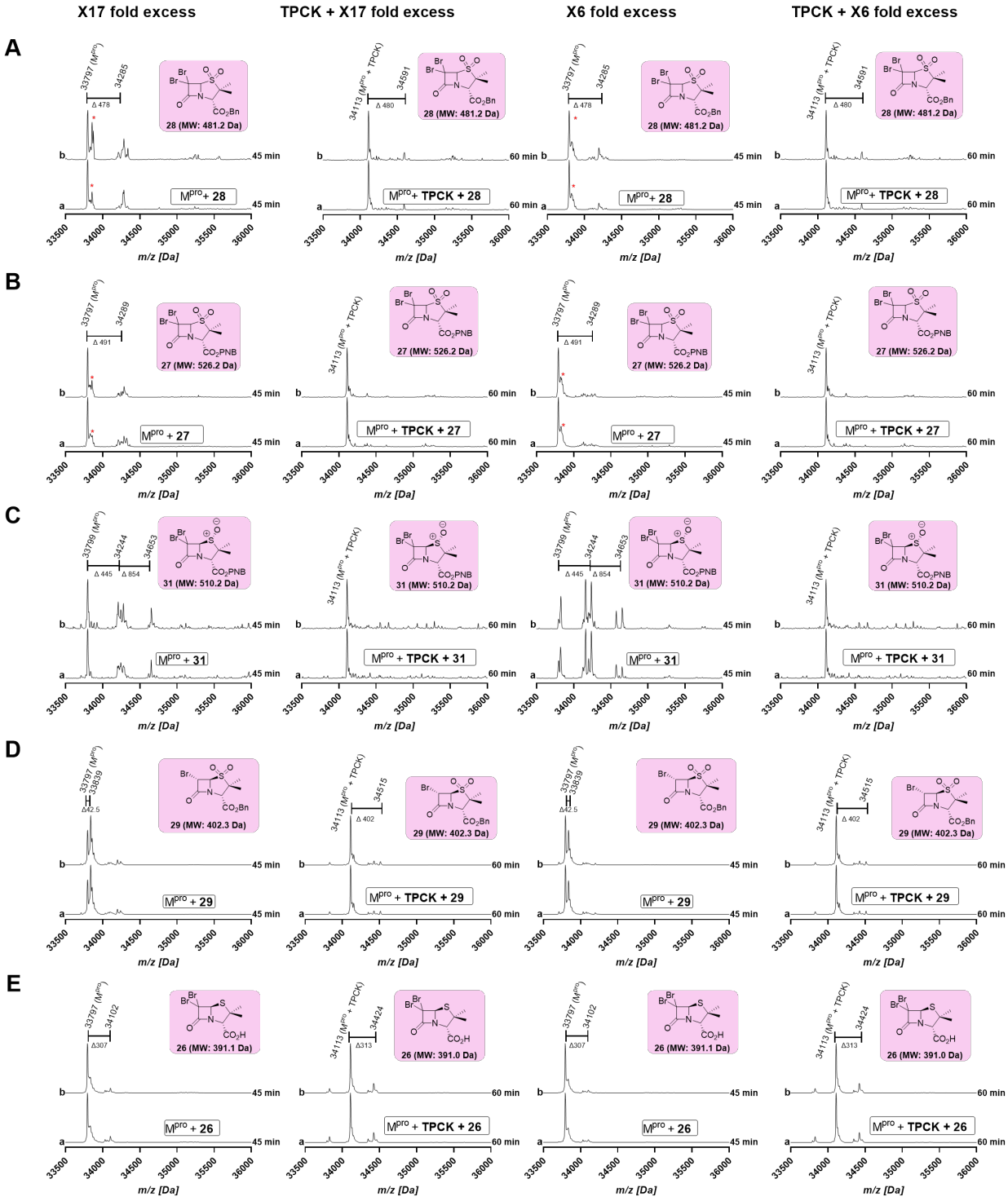


Figure 3.6 | C6-bromopenicillin sulfone ester derivatives inhibit M^{pro} by complex mechanisms. M^{pro} assays were performed using SPE-MS as described in the Experimental Section. Samples contained SARS-CoV-2 M^{pro} (2 μM), the indicated **(A-E)** bromopenicillin sulfone ester derivatives (33 μM or 11 μM), and TPCK (10 μM). A 42 Da mass difference from the protein peak was indicated with an asterisk (red). Note that compound numbering is the same as that in Manuscript 2 [303]. a and b represent the technical duplicates.

3.3.4 Investigating the effect of γ-lactams and representative sulfoximines on M^{pro} activity

Initially, γ-lactams were screened as potential M^{pro} inhibitors at a concentration of 50 µM. However, potent M^{pro} inhibition was not observed under these conditions. Dose response studies were performed to accurately determine the inhibitory activity and to inform on the SAR studies with bicyclic γ-lactams (Table 3.4, Figure 3.7C).

Table 3.4: Estimated IC_{50}s of γ-lactam derivatives. Assays were performed as described [253,303,314]. Conditions: 0.15 µM M^{pro} and 2.0 µM TSAVLQ/SGFRK-NH_2 in reaction buffer (20 mM HEPES pH 7.5, 50 mM NaCl). Structures of the γ-lactams are shown in Figure 3.2. Note that **65** is a control compound with an oxo-tetrahydrofuranyl ring. **90** and **91** are penicillin derivatives.

Compound	SPE-MS ($\sim IC_{50}$ / µM)	FRET ($\sim IC_{50}$ / µM)	Compound	SPE-MS ($\sim IC_{50}$ / µM)	FRET ($\sim IC_{50}$ / µM)
54	>100	>100	**59**	90.7	>100
55	>100	>100	**60**	20.0	>100
56	52.8	>100	**84**	>100	>100
57	72.1	>100	**54**	>100	>100
61	>100	>100	**66**	14.8	>100
62	>100	84.5	**67**	>100	>100
63	93.6	>100	**89**	77.3	>100
64	>100	>100	**71**	99.8	>100
68	>100	>100	**72**	>100	>100
69	60.3	15.1	**78**	21.5	37.5
70	>100	>100	**79**	25.8	>100
73	>100	>100	**77**	>100	>100
74	39.9	>100	**75**	50.0	>100
79	>100	>100	**83**	>100	>100
80	16.0	>100	**88**	8.2	>100
81	15.6	>100	**58**	25.6	>100
82	>100	>100	**87**	21.7	>100
85	>100	>100	**86**	20.6	>100
65	>100	>100	**91**	62.6	>100
90	1.7	18.8			

82, a methyl ester γ-lactam sulfoxide, was the most potent γ-lactam inhibitor observed in the SPE-MS assays (IC_{50} ∼8.2 μM). This observation contrasts with the results obtained for the penicillin derivatives, for which the sulfoxides were mostly inactive, apart from the C6-amido phenoxyacetyl ether penicillin ester sulfoxide (**14** in manuscript 2 [303], IC_{50} ∼22.9 μM) and the C6-dibromo penicillin ester sulfoxide (**32**) in manuscript 2 [303], IC_{50} ∼0.5 μM). Moderate inhibition (IC_{50} ∼8.2-25.8 μM) was observed with compounds **60**, **66**, **80**, **81**, **86**, **87**, **88**, and **90** using the SPE-MS assay. However, no inhibition was observed for these compounds using an alternative FRET assay (Table 3.7). Likewise, γ-lactam inhibitors shown to have weak activity in the SPE-MS assay manifested little-to-no activity when investigated using FRET assays (e.g. **56**, **57**, **59**, **63**, **71**, **74**, **75**, **81**, **86**, **87**, and **89**). **78**, a bicyclic benzyl ester γ-lactam thioether derivative which resembles the penicillin sulfone M^{pro} inhibitor **4** (manuscript 1) displayed moderate activity [IC_{50} ∼21.5 μM (SPE-MS), ∼37.5 μM (FRET)] in both assays.

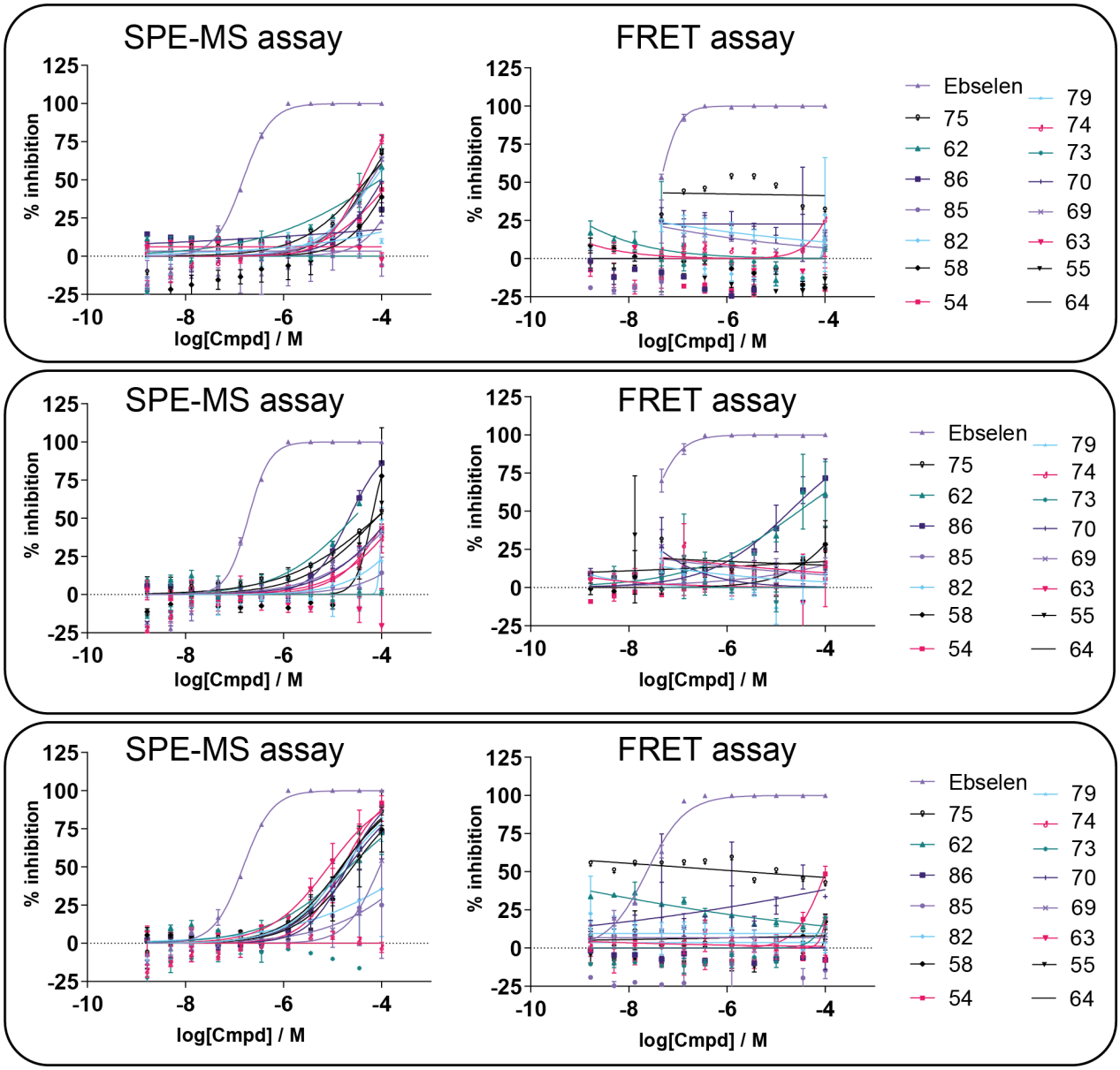


Figure 3.7 | M^{pro} dose response curves for γ-lactam derivatives obtained using SPE-MS and FRET assays. Structures of γ-lactams are shown in Figure 3.2.

Both the *N*-Boc protected sulfoximine penicillin **90** (IC_{50} ∼18.8 μM) and ebselen, the latter used as a positive control, were observed to inhibit M^{pro} in both the SPE-MS and FRET assays (Figure 3.6A, 3.6B). As the dose response curves of the γ-lactams were incomplete, the poor agreement between the estimated IC_{50}s of weak inhibitors from the two different assays is expected.

The C6-bromo compounds **28**, **31** and **32** were the most potent β-lactam penicillin inhibitors of M^{pro} [Figure 3.6, Table 4 (manuscript 2 [303])]. In contrast with the C6-amido penicillin sulfoxides (Table 3.1), the C6-monobromo penicillin sulfoxides (Figure 3.8A) inhibited M^{pro}. Notably, the inhibition by the C6-monobromo penicillin sulfoxides was more potent than the inhibition by the corresponding sulfone derivatives (Figure 3.8B, 3.8C). However, the inhibitory activity of the monobromo penicillins is conserved upon substitution of the sulfoxide with a sulfoximine (Figure 3.8D-3.8F). Thus, sulfoximine penicillin derivatives, which are likely to possess better solubility than sulfoxide and sulfone derivatives [313], could be explored for cellular inhibition studies. The sulfoximine group could also be exploited as a chemical handle to couple with other functional groups for further SAR optimisation and/or improvement of physicochemical properties.

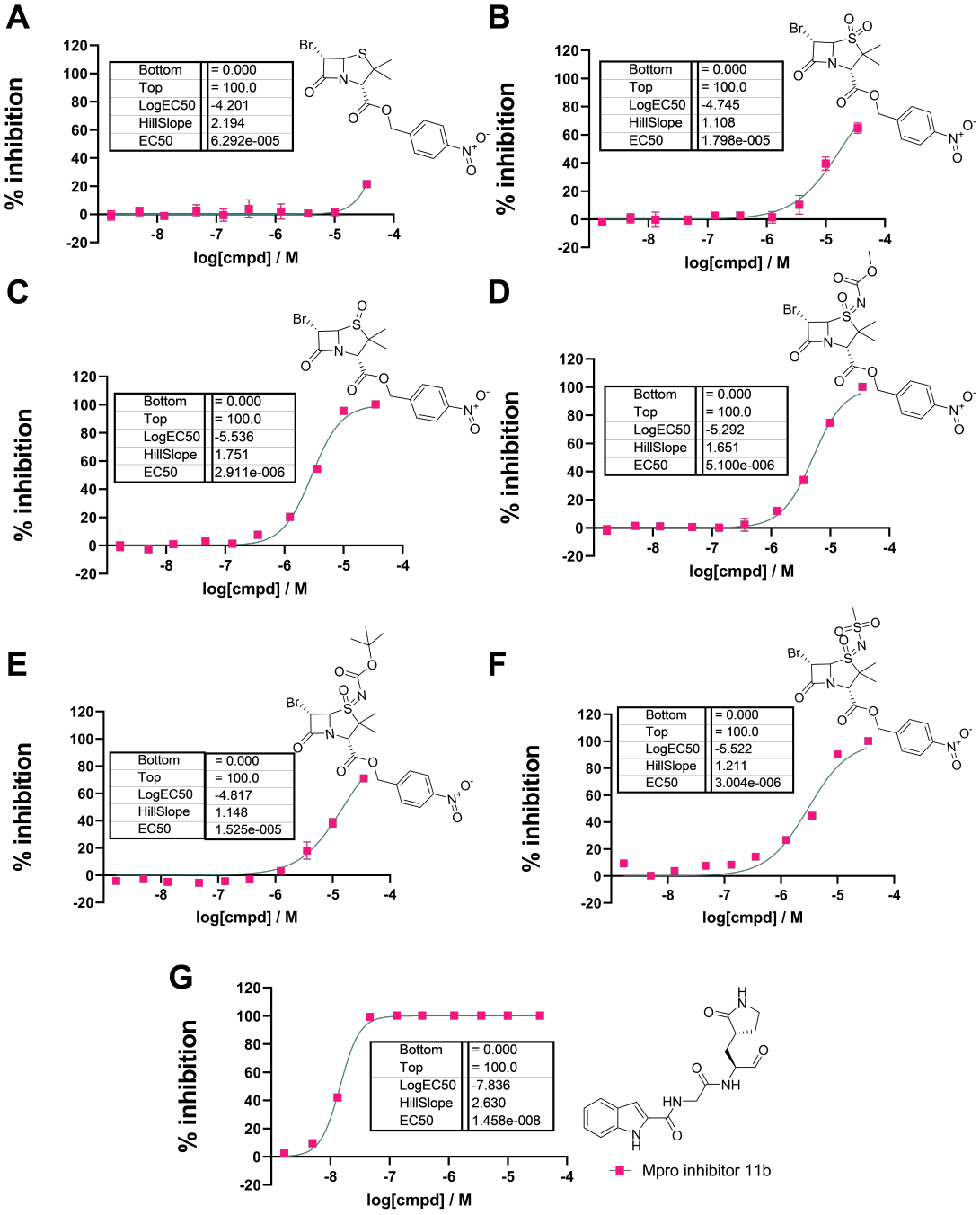


Figure 3.8 | M^{pro} dose response curve analyses with C6-bromo penicillin sulfoximines. **A** C6-monobromo compound is active upon oxidation of thioether to **B.** sulfoxide. **D.-F.** Sulfoximine derivatives retain M^{pro} inhibitory activity. **G.** An α-ketoamide inhibitor was used as a control. Conditions: 0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2 in the reaction buffer. Compounds were synthesized by Dr. Lennart Brewitz (Department of Chemistry, University of Oxford).

3.4 Discussion

The 11-mer M^{pro} substrate suffered from ion suppression in the presence of some compounds, therefore dose response studies were carried out with the 37-mer peptide. Excellent Z' [256] values (a statistical measurement of assay quality) were obtained with both the 11-mer and the 37-mer peptides. The 37-mer peptide could be used to identify weak inhibitors, as shown in the distribution plot (Figure 3.5A). This property was important, as it helped guide the SAR studies carried out on the penicillin sulfone derivatives.

Penicillin sulfone derivatives are excellent inhibitors of M^{pro}, as demonstrated using the SPE-MS assays. The SAR studies indicate that both aromatic and aliphatic C3-esters of the penicillin scaffold can be tolerated. However, of the tested compounds, the presence of a C6-amido phenoxyacetyl ether substituent was found to be crucial for activity. The C6-bromo penicillin derivatives **28, 31**, and **32** were notable exceptions, manifesting the most potent β-lactam-derived M^{pro} inhibition of all the screened β-lactams, despite the absence of the C6-phenoxyacetyl ether oxygen. The results obtained from non-denaturing protein-observed MS studies were inconclusive with the C6-bromo compounds, but suggest their mechanism(s) of reaction with M^{pro} may be complex.

Monocyclic [309] and bicyclic [311] γ-lactams have been investigated as inhibitors of bacterial and viral (hepatitis C virus) serine proteases, respectively. However, the tested γ-lactams were, in general, less potent than the analogous β-lactams in terms of M^{pro} inhibition. This can, in part, be explained by the lower chemical reactivity of the γ-lactam ring compared to the β-lactam ring [310]. The bicyclic nature of the β-lactam penicillin scaffold imposes a non-planar conformation on the β-lactam nitrogen, making it more reactive as an electrophile.

The sulfoximine γ-lactam derivatives were devoid of M^{pro} inhibition activity. By contrast, the β-lactam sulfoximines showed activity, and hence can be exploited to improve the potency or the solubility of penicillin sulfone derivatives.

3.5 Future directions

The mode of inhibition of the C6-bromo penicillins, which manifest potent M^{pro} inhibition, is yet to be determined and may involve fragmentation [320–323]. When incubated with M^{pro}, some of the dibromopenicillin sulfone esters (i.e. **28, 27**, and **29**) appear to shift the M^{pro} mass by +42 Da (in addition to other adducts) as observed by protein-observed MS. Crystallographic studies may help to characterise these M^{pro}-bound fragments. Similarly, further optimisation of the non-denaturing MS assays may improve the low signal-to-noise ratio; notably, this may enable investigations into their mode of inhibition. Inhibition studies using FRET assays could be performed in the presence of low levels of detergent to rule out the possibility that the dibromo compounds inhibit M^{pro} via compound aggregation, while the inclusion of low levels of EDTA [281] could help rule out the presence of metal ion contaminants in the compound/sample preparations.

β-Lactam penicillin sulfones are promising alternatives to peptidomimetic M^{pro} inhibitors. The excellent safety window of β-lactam antibiotics, even at high concentrations [324, 325], is a reason to pursue them as M^{pro} coronavirus inhibitors.

3.6 Methods

3.6.1 FRET assays

Compounds were dispensed into Greiner 384 low volume non-binding plates using an Echo 555 acoustic dispenser. DMSO was back-filled for a uniform final concentration of 0.5%($_{v/v}$) in the assay. Reagents for M^{pro} assays were dispensed into the assay plate (12.5 µL/well of M^{pro}, followed by 12.5 µL/well of peptide substrate) using a Multidrop Combi liquid dispenser. Reaction conditions were: 20 mM HEPES pH 7.3, 50 mM NaCl, 0.01% Triton-X100, 10% glycerol, 50 nM M^{pro} and 375 nM fluorogenic peptide substrate ([5-FAM]-AVLQSGFR-[Lys(Dabcyl)]-K-amide). M^{pro} was pre-incubated for 15 min at room temperature with the compound before the addition of the substrate and centrifugation for 15 s. Reactions were continuously monitored without quenching in a BMG Pherastar FS plate reader with a 480/520 ex/em filter set. Data analyses were performed using GraphPad Prism 8.

All other methods are listed in the experimental section of manuscript 2 [303].

3.7 Appendix

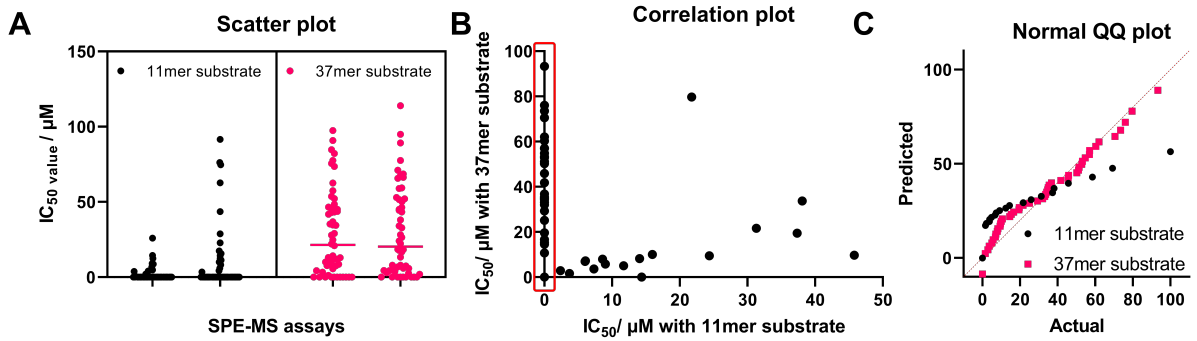


Figure 3.9 | Scatter plot, correlation plot, normal QQ plot analyses of the pIC_{50}s obtained with the 11-mer and the 37-mer M^{pro} peptide substrates. A. The scatter plot demonstrates that IC_{50}s obtained with the 11-mer peptide are more clustered than the IC_{50}s obtained with the 37-mer substrate. **B.** Correlation plot of pIC_{50}s obtained using the 37-mer substrate plotted against the IC_{50}s obtained with the 11-mer substrate. **C.** The normal QQ plot shows that only IC_{50}s obtained with the 37-mer substrate are normally distributed.

4

Studies on M^{pro} substrates and peptide inhibitors of M^{pro}

Contents

4.1 Abstract **155**
4.2 Introduction **156**
4.3 Results **159**
4.4 Future directions **166**

4.1 Abstract

The work described in manuscript 4 [314], which is associated with the work described in this chapter, involved inhibition and biochemical studies with linear peptide inhibitors of M^{pro}. The linear peptides were designed following *in silico* studies aimed at predicting the most tightly binding residues spanning the P4-P5′ positions which occupy the S4-S5′ subsites in the protease. The experimental evidence supports the proposal that the *in silico* designed peptides manifest inhibitory activity in a substrate-competitive manner with the 11-mer M^{pro} substrate. In addition to the work presented in the manuscript, inhibition studies suggest that N-terminal product peptides of varying lengths manifest product mediated M^{pro} inhibition. The tetramer substrate-based peptide VLQS moderately inhibited M^{pro}, whereas the tripeptide LQS did not. This implies that the P3 residue plays a role in M^{pro} binding, as was observed with N3 [28] where the backbone NH and carbonyl groups of the P3 Val form hydrogen bonds with backbone carbonyl and NH groups of Gln166, respectively. Amino acid substitutions at the P1 position

show that a P1 Gln is important for optimal M^{pro} activity. However, substrate turnover was observed when His was present at the P1 position, which may reflect the ability of M^{pro} to cleave a pool of host substrates. Finally, assays with the 37-mer substrate revealed that the inhibitory activity of the peptide inhibitors is substrate-dependent. The inhibition of M^{pro} by the designed 11-mer peptide inhibitors was ablated when assays were performed with the 37-mer substrate instead of the 11-mer substrate, reflecting the higher affinity of the 37-mer substrate for M^{pro}. Thus, linear peptide inhibitors of M^{pro} require further optimisation to be developed as effective M^{pro} inhibitors.

4.2 Introduction

Oligopeptides (500-5000 Da) are important drugs, with an estimated global revenue of 70 billion dollars in 2019 [52, 326, 327]. Peptide drugs can bind directly to allosteric or active sites of pathologically relevant enzymes, and disrupt protein-protein interactions [328, 329]. Although the development of peptide-based inhibitors which can bind at the dimer interface of the catalytically active M^{pro} homodimer is plausible, the active site of M^{pro} is a more obvious target for peptidomimetic drug development. M^{pro} cleaves its polyprotein substrate at 11 distinct sites, but the current M^{pro} peptidomimetic drugs and drug leads are solely based on nsp4/5 cleavage sites (Table 1.1). Therefore, investigations into the nature of the substrate recognition and binding affinity of M^{pro} may enable the design of novel inhibitors based on other cleavage sites.

At the beginning of this work in March 2020, the interactions of M^{pro} with its substrates were not fully defined. A consortium was thus established to carry out molecular dynamics simulations to model the interaction of M^{pro} with 11-mer peptide substrates (s01-s11, Table 1.1) representing all 11 M^{pro} cleavage sites in SARS-Cov-2 pp1ab, in order to identify conserved hydrogen bonding interactions across all substrates. Turnover studies with these 11-mer substrates were performed to complement the modelling studies; the *in silico* and experimental studies afforded comparable M^{pro} substrate reactivity profiles. Subsequently, the results of alanine scanning experiments were used to design linear peptides which bind tightly at the M^{pro} active site. The resulting peptides manifested a substrate-competitive mode of M^{pro} inhibition.

Structures of some substrate peptides (s01, s02, s03, s05, s06, s10, s11) have recently been solved in complex with catalytically inactive variants, i.e., M^{pro} H41A/C145A [330,331]. The P4-P1′ residues of these peptides occupy the S4-S1′ subsites. The P1 Gln fits tightly into the S1 subsite, as seen with substrates s01 (Figure 4.1A), s02 (Figure 4.1B), s03 (Figure 4.1C), s05 (Figure 4.1D), s06 (Figure 4.1E), s10 (Figure 4.1F) and s11 (Figure 4.1G). The Leu at P2 in s01, s05, s06, s10, and s11 forms hydrophobic contacts at the S2 subsite, similar to what was observed with the P2 Leu of the N3 inhibitor [28], whereas Phe and Val occupy the S2 subsite in s02 and s03, respectively (Figure 4.1B, 4.1C). The backbone carbonyl and NH groups of the P3 Val in s01 are poised to form hydrogen bonds with the backbone NH and carbonyl groups of Glu166, respectively. By contrast, s02 which has Val at the P4 position and s3 which possesses Val at P1 position form H-bonds with Gln189 [330]. All P3 residues such as Thr in s02 and s03, Lys in s05 and s10, and Arg in s06 and s11 interact with Glu166. Notably, additional ′cross subsite′ interactions via the P3 side chains are formed with Asn142 in the S1′ subsite in s06 (Arg) and s11 (Lys) (Figure 4.1H); this is an interesting observation as the interaction of prime site residues with non-prime subsites is yet to be exploited for inhibitor design. In the P1′ position, Ser and Ala residues are most prevalent, although an Asn residue is uniquely present at the P1′ position of s05 (Figure 4.1D), and is positioned to form a hydrogen bond with the Nϵ atom of His41. Thus, M^{pro} substrate structures can be informative for inhibitor design and optimisation.

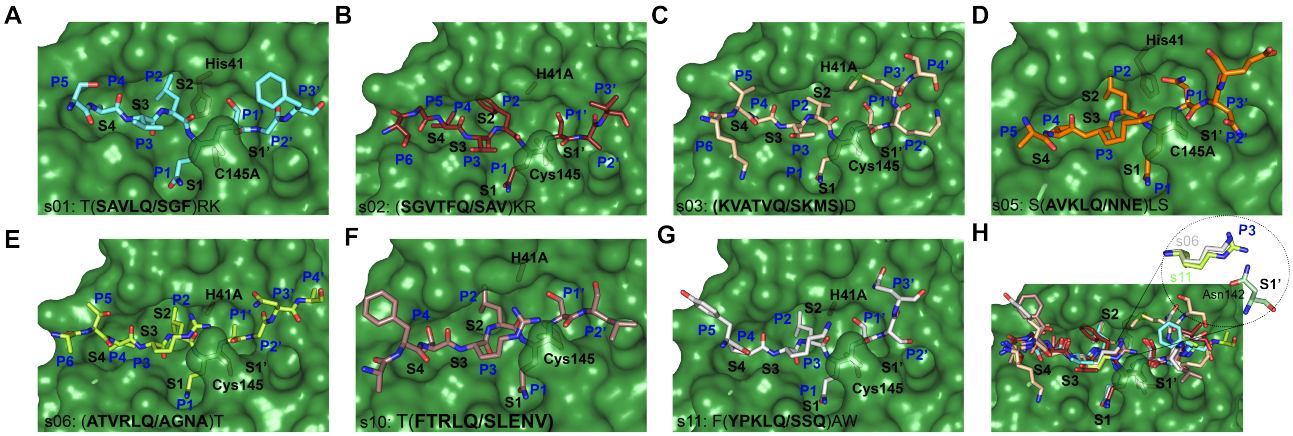


Figure 4.1 | Views of the substrate peptides at the active site subsites of the M^{pro} H41A/C145A variant. Substrates **A.** s01 (PDB: 7DVP), **B.** s02 (PDB: 7DVW), **C.** s03 (PDB: 7DVX), **D.** s05 (PDB: 7MGR), **E.** s06 (PDB: 7DVY), **F.** s10 (PDB: 7DW6) and **G.** s11 (PDB: 7DW0) are represented in stick models, and the surface of M^{pro} is represented in green. **H.** Only the non-prime site binding residue in s06 and s11 interact with a prime site subsite residue (i.e. Arg and Lys at the P3 position of substrates s06 and s11, respectively interact with the side chain of Asn142). The electron densities were observed for the residues highlighted in bold. The positions of the substrate residues and the M^{pro} subsites are indicated. See Table 1.1 M^{pro} substrates section in chapter 1 for substrate sequences. Electron densities at the active site was only observed for the residues of the substrates enclosed in parenthesis.

4.2.1 Contributions to the manuscript: Discovery of SARS-CoV-2 M^{pro} peptide inhibitors based on substrate modelling and ligand binding experiments

Manuscript 3 [314] describes *in silico* and biochemical work with peptides based on the natural substrate sequences of M^{pro}, as well as studies with acyclic peptide inhibitors. These peptide inhibitors were designed by: (i) computational alanine scanning to identify the residues that contribute the most to the binding energy, followed by (ii) predictive saturation variation scanning to identify amino acid substitutions that may give rise to higher peptide binding energies.

My work for manuscript 3 included investigating the natural substrate preferences of M^{pro} through *in vitro* inhibition studies with linear peptides. In detail, my contributions were:

- Developing and conducting SPE-MS kinetic studies with the 11 M^{pro} substrate peptides.
- LC-MS studies to monitor the turnover of the natural substrate peptides.

- Analysis of the mode of inhibition of the acyclic peptides.
- Solid-phase peptide synthesis of peptides variants of the M^{pro} natural substrate s01 (i.e. s01_LP2W and s01_QP1W) and the peptide inhibitor p13_WP2L.

In addition to the work described in manuscript 3, I also carried out the following unpublished experiments:

- Determining kinetic parameters for M^{pro} with natural substrates.
- Investigating N- and C-terminal product peptides of the s01 M^{pro} substrate as potential inhibitors.
- Screening truncated peptides based on the nsp4/5 cleavage site for M^{pro} catalysed turnover and inhibition.
- Determining the inhibitory effect of the acyclic peptides on M^{pro} catalysis using SPE-MS assays using the 37-mer substrate peptide.

4.3 Results

4.3.1 M^{pro} catalysis with natural peptide substrates

In manuscript 3, SPE-MS turnover assays revealed that only the s01, s02, s06, and s11 M^{pro} peptide substrates (Table 1.1 for substrate sequence) were turned over during the standard incubation time (Figure S2.23, manuscript 4). Subsequently, the kinetic parameters of these substrates were determined to identify the substrate for which M^{pro} demonstrates the highest catalytic efficiency. The catalytic efficiencies (k_{cat}/K_m) of M^{pro} for both s06 (4,300 $M^{-1}s^{-1}$) and s011 (6,800 $M^{-1}s^{-1}$) (Figure 4.2) were substantially lower than that for s01 (27,000 $M^{-1}s^{-1}$, Figure 2.14). However, curve-fitting of the initial rates determined at different substrate concentrations using Michaelis-Menten kinetics was sub-optimal for s06 and s11, and the V_{max} was thus not accurately determined, as product-mediated inhibition of M^{pro} was observed at substrate concentrations above 30 µM. Similarly, product-mediated inhibition was observed with s02; as the substrate concentration (red) was increased,

product formation (green) decreased (Figure 4.2). The catalytic efficiency of M^{pro} for s02 could also therefore not be accurately determined.

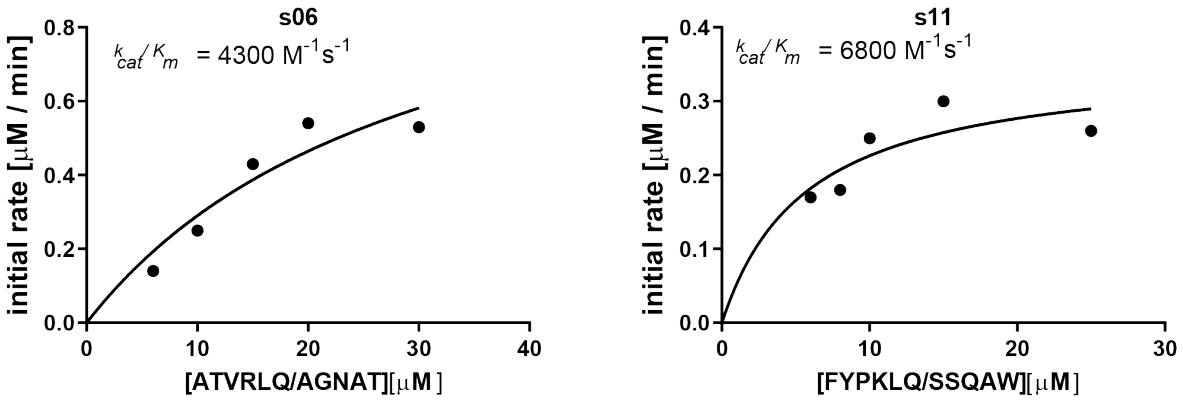


Figure 4.2 | Steady state kinetic analyses for M^{pro} with substrate peptides s06 ATVRLQ/AGNAT-NH_2 and s11 FYPKLQ/SSQAW-NH_2. M^{pro} (0.15 μM) was incubated with a range of substrate concentrations. The initial rates and substrate concentrations were fitted to the Michaelis-Menten equation using non-linear regression. M^{pro} displayed a higher catalytic efficiency with s11 (6,800 $M^{-1}s^{-1}$) than s06 (4,300 $M^{-1}s^{-1}$). Note that product-mediated inhibition was observed for s11 peptide concentrations greater than 30 μM.

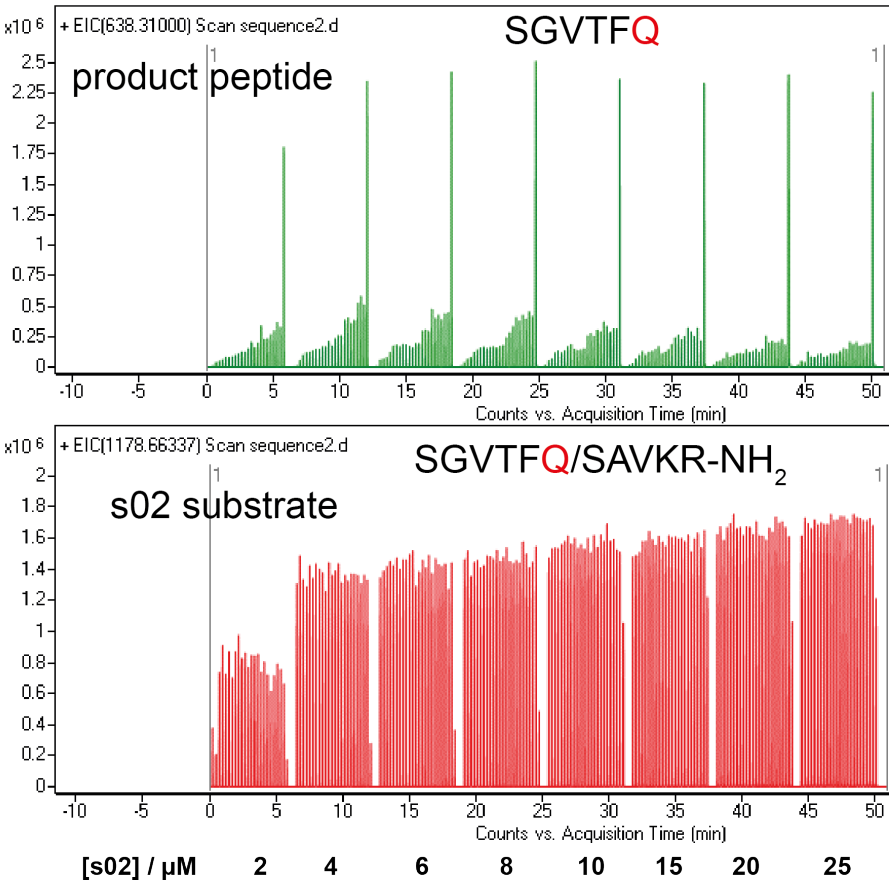


Figure 4.3 | Decreased s02 turnover by M^{pro} at increasing s02 concentrations. M^{pro} (0.15 µM) was incubated with 2-25 µM of the s02 peptide substrate. Integration of the +1 charge state of the extracted SGVTFQ N-terminal product peptide (green peaks) and of the SGVTFQ/SAVKR-NH_2 substrate (red peaks) suggest reduced product formation at higher substrate concentrations.

4.3.2 Product inhibition of M^{pro}

To further investigate the observed product mediated inhibition of M^{pro} with s02, product peptides from the 11-mer and 37-mer substrates were synthesized and evaluated as M^{pro} inhibitors. Dose response studies suggest that M^{pro} activity is inhibited with the 11-mer substrate upon preincubation with the product fragments derived from both the 11-mer and 37-mer substrate peptides corresponding to the nsp4/5 cleavage site. The N-terminal peptide derived from M^{pro}-catalysed hydrolysis of the 11-mer s01 substrate peptide (Figure 2.10) was capped at the N-terminal amino group with an acetyl group. This peptide, i.e. Ac-TSAVLQ, does not potently inhibit M^{pro} (IC_{50} ~57.9 µM, Figure 4.4A), whereas the N-terminal product derived from the 37-mer substrate, i.e. ALNDFSNSGSDVLYQP PQTSITSAVLQ-NH_2, displays higher inhibitory activity (IC_{50} 12.0 ± 3.0 µM, Figure 4.4 C). The

C-terminal product of the 11-mer s01 substrate peptide has a comparable inhibitory activity (IC_{50} 11.4 ± 0.2 μM, Figure 4.4B) to the N-terminal product peptide formed from the 37-mer substrate. Surprisingly, the 10-mer C-terminal product peptide derived from the 37-mer substrate with the extended C-terminal FS residues (compared to the 8-mer C-terminal product peptide derived from the 11-mer substrate) shows weaker activity (IC_{50} ~28.8 μM, Figure 4.4D) suggesting that increased peptide length may improve catalysis (11-mer vs 37-mer substrates), but not inhibition (10-mer vs 8-mer C-terminal products). Note that the 8-mer C-terminal product was used as an internal standard for the optimised assay with the 37-mer substrate, as described in Chapter 5.

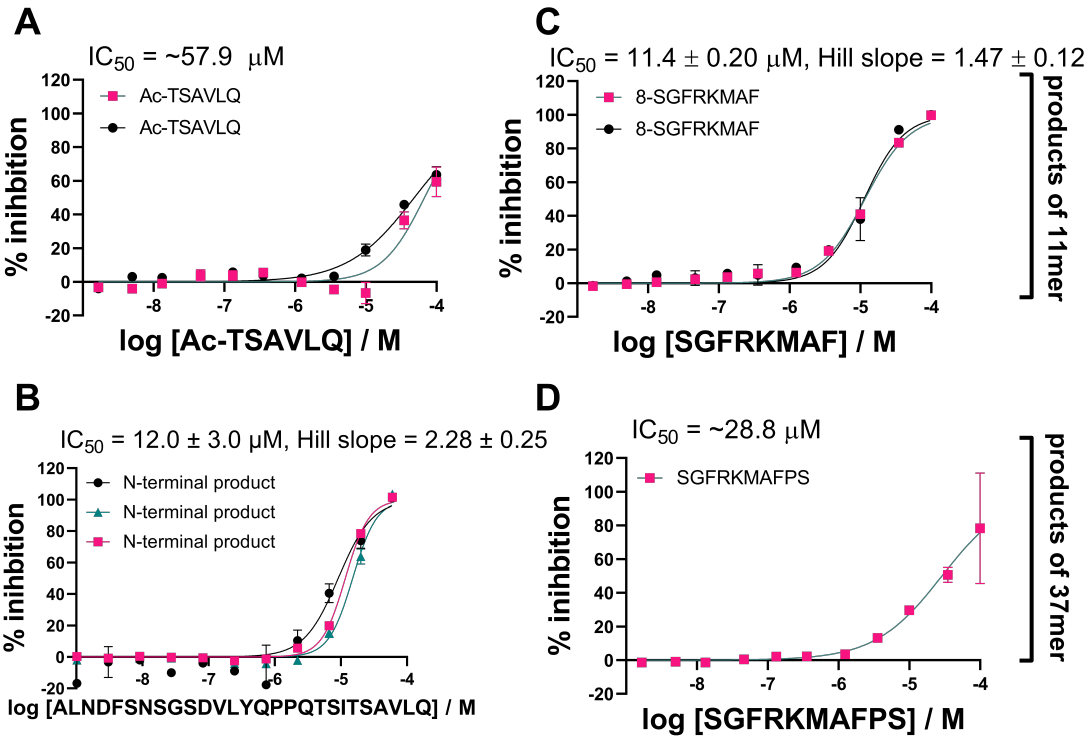


Figure 4.4 | Representative dose response curves indicate that peptide products inhibit M^{pro}. **A.** and **B.** N-terminal and (**C.** and **D.** C-terminal product fragments of the 11-mer and 37-mer substrate peptides, respectively, inhibit M^{pro} activity. Assay were performed as described [253,303]. Conditions: 0.15 μM M^{pro}, 2.0 μM TSAVLQ/SGFRK-NH_2 in the reaction buffer. IC_{50} values are the mean of at least two independent repeats composed of technical duplicates. Note that the IC_{50} value in panel D is composed of the mean of technical duplicates. Red square and black circle in panels A-C represent independent repeats composed of technical duplicates.

4.3.3 Investigation of truncated s01 substrate peptides as M^{pro} substrates and inhibitors

A collaboration with Prof. Allen Orville (Diamond Light Source) on time-resolved M^{pro} X-ray crystallography required water soluble peptide substrates in order to monitor M^{pro} catalysis. Thus, short tripeptides and tetrapeptides with and without capped N-terminal amino groups were designed. However, M^{pro}-catalysed turnover of these truncated substrates was not observed by NMR spectroscopy or LC-MS. Notably, LC-MS was used instead of SPE-MS as the latter is not suitable for detecting short hydrophilic peptides. M^{pro} inhibition studies were performed to investigate potentially substrate-competitive inhibition from these truncated substrate peptides. The tetrapeptide VLQS was found to be a weak M^{pro} inhibitor (IC_{50} $\sim$39.8 µM, Figure 4.5A), whereas other tripeptides and tetrapeptides do not inhibit M^{pro} (Figures 4.5B- 4.5D).

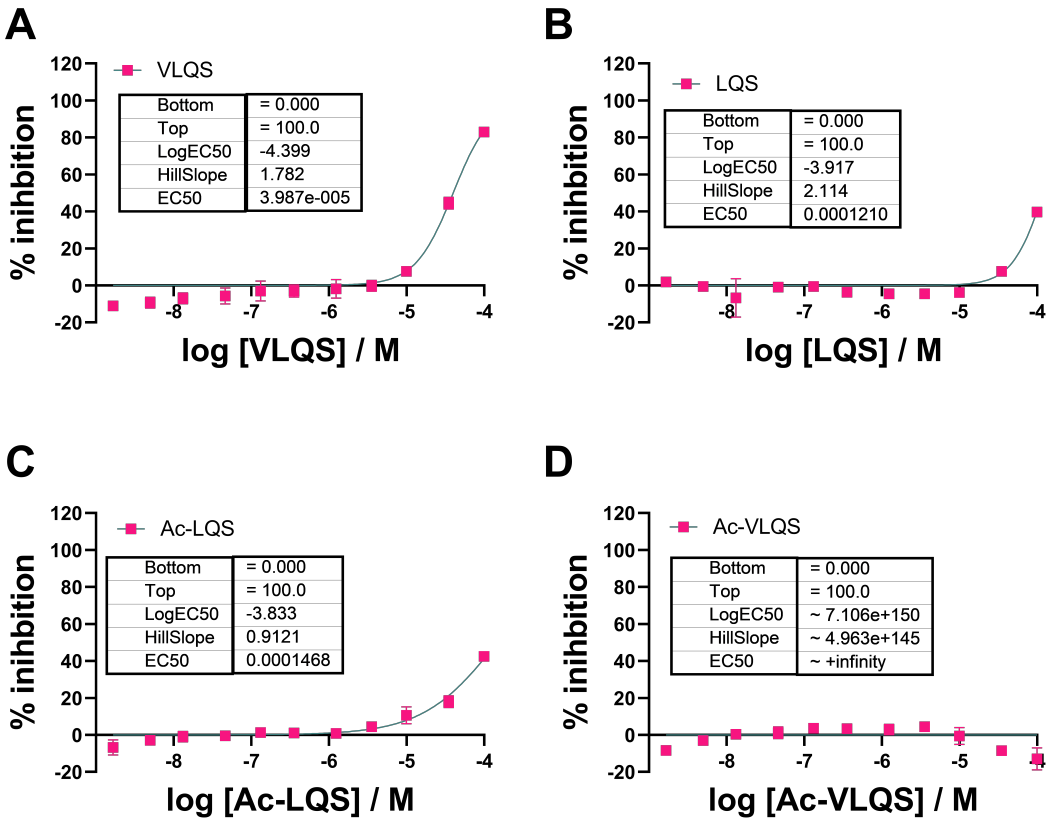


Figure 4.5 | Representative dose response curves to investigate M^{pro} inhibition by truncated s01-derived substrate peptides. A. Tetrapeptide VLQS display moderate M^{pro} inhibition, whereas **B.** and **C.** tripeptides are weak inhibitors. **D.** The M^{pro} inhibition from tetrapeptide VLQS is lost upon *N*-acetylation of the tetrapeptide VLQS. Assays were performed as described [253,303]. Conditions: 0.15 µM M^{pro}, 2.0 µM TSAVLQ/SGFRK-NH_2 in the reaction buffer.

4.3.4 Investigation of a 20-mer substrate variant at the P1 position

Kinetic analyses were performed using several variants of a 20-mer substrate designed based on the nsp4/5 N-terminal *cis* M^{pro} autocleavage site (Chapter 2, Figure 2.7). The amino acid residue at the P1 position of the substrate, which is Gln across all 11 cleavage sites of M^{pro}, was substituted with other residues. SARS-CoV-1 M^{pro} accepts substrates with His or Met residues at the P1 position, therefore the tolerance of SARS-CoV-2 M^{pro} for His or the positively charged residues Arg and Lys at the P1 position was investigated [332]. Only the peptide variant with His at P1 (i.e. Q10H) is cleaved by M^{pro} (Figure 4.6A). Substitution of L-Gln with D-Gln in the 20-mer peptide (Q10q) does not result in turnover by M^{pro}. Next, potential competitive inhibition that could result from the binding of the uncleaved peptide

variants (Q10R, Q10K, Q10q) was assessed. No inhibitory activity was detected (Figure 4.6B-4.6D), suggesting that the P1 residue is critical for binding of the substrate to the active site of M^{pro}.

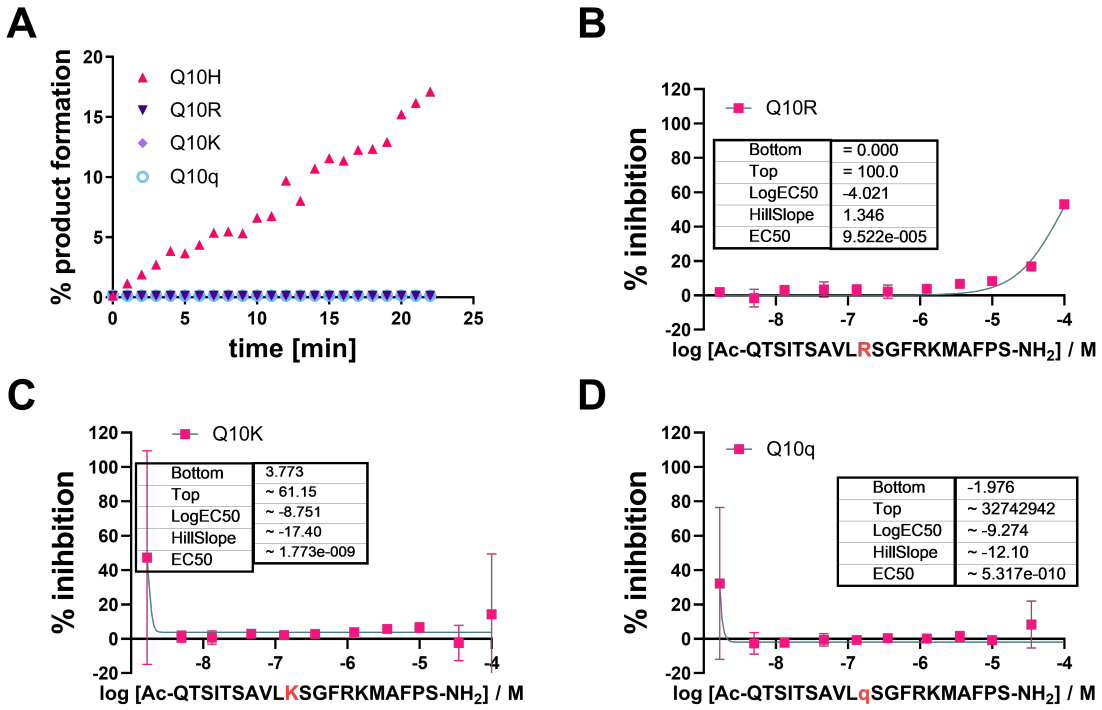


Figure 4.6 | The effect of substitutions at the P1 position of the M^{pro} substrate on catalysis. A. A M^{pro} substrate variant with His at position P1 is hydrolysed, whereas substrate variants with the positively charged residues Arg and Lys or the non-canonical D-Gln at the P1 position are not M^{pro} substrates or (**B.-D.**) inhibitors. Assay were performed as described [253, 303]. Conditions: 0.15 µM M^{pro} and 2 µM substrate peptides in 20 mM HEPES pH 7.5, 50 mM NaCl

4.3.5 Validating the inhibitory effect of the peptide inhibitors on M^{pro} catalysis using assays with a 37-mer substrate

Manuscript 3 describes the determination of the mode of M^{pro} inhibition by the designed peptide inhibitors (p12, p13, p15, and p16) with varying concentrations of the 11-mer substrate. A linear increase in IC_{50} values was observed with increasing substrate concentrations, suggesting a substrate-competitive mode of inhibition. In parallel, ratiometric analysis of IC_{50}s, at substrate concentrations below and above the K_m of the 11-mer substrate, as proposed by Wei *et al.* [333], was used to validate the substrate-competitive mode of inhibition.

As the 37-mer substrate has a lower K_m than the 11-mer substrate, the peptide inhibitors were

re-tested using the 37-mer substrate. However, no inhibition was observed in the dose response studies conducted with the peptides, suggesting that they can only compete with the 11-mer substrate for binding to the M^{pro} active site because this shorter peptide has a ∼5-fold higher K_m than the 37-mer substrate, and thus it has a reduced affinity to M^{pro}.

4.4 Future directions

Hydrolysis of the computationally designed peptide inhibitors was not observed under the standard assay conditions; however, structural studies with M^{pro} may provide insights into the lack of the turnover observed for the peptide inhibitors with Gln at P1. The truncated tripeptide LQS, derived from the s01 substrate, displays M^{pro} inhibition. This inhibition was not observed when the amino group of the N-terminal Leu was capped with an acetyl group. Structural and/or modelling studies would be useful to investigate the effect of the acetyl group on the interaction of the tripeptide with M^{pro}.

Only variations at the P1 position of the nsp4/5-derived s01 substrate were investigated in this work; the s01 variants with positively charged residues at the P1 position were neither substrates nor inhibitors of M^{pro}. Studies on substitutions at other positions, e.g. P2 or prime sites, may be useful to draw conclusions on intrapeptide interactions of prime and non-prime sites residues of the substrate. For instance, evidence for sub-site co-operativity has been reported with the C-terminal auto cleavage site of SARS-CoV-1 M^{pro} [279]. In sub-site co-operativity, the binding of a residue in one active site pocket can change the plasticity of the nearby pocket in the active site. As the active sites of SARS-CoV-1 and SARS-CoV-2 M^{pro} are identical [184], subsite co-operativity would be expected to be conserved between the two proteins. By contrast with the SARS-CoV-1 cleavage site, the C-terminal autocleavage site of SARS-CoV-2 M^{pro} possesses a Phe in the P2 but not in the P3′ position; a Val residue is found in the P3′ position in the SARS-CoV-2 M^{pro} cleavage site. Catalytic studies with variants of C-terminal peptides with P3′ Phe, in combination with crystallographic studies, could be undertaken to investigate if sub-site co-operativity can be deliberately introduced in SARS-CoV-2 M^{pro}. These studies could contribute towards inhibitor design, as, other from the P1′ site position, current inhibitors do not target the prime subsites.

5

Macrocyclic inhibitors of M^{pro}

Contents

5.1 Abstract **167**
5.2 Introduction **168**
5.3 Contributions to the manuscript **168**
5.4 Results **169**
5.5 Discussion **185**
5.6 Future directions **187**

5.1 Abstract

The RaPID selection screen was applied with the aim of identifying novel cyclic peptide inhibitors of M^{pro}, including those with unnatural γ-amino acid residues. This screen led to the discovery of the potent M^{pro} inhibitor TMGM4-05, which has a His residue at the P1 equivalent position, and a K_D of 10 nM. Other charged residues were not tolerated in the P1 position of this cyclic peptide, as was also observed with acyclic peptide substrates. Tryptophan was tolerated at the P1 position in both a substrate peptide (as shown with the s01Q→W peptide in manuscript 3) and a variant of the TMGM4-05 cyclic peptide inhibitor (manuscript 4). Furthermore, the incorporation of features of the M^{pro} substrates into TMGM4-05 by introducing Gln at the P1 position improved potency. Thus, results obtained from the M^{pro} substrate profiling can be applied to improve the potency of cyclic peptide inhibitors (and inhibitors of other classes) that hold the potential to be optimised as antiviral COVID-19 therapeutics.

5.2 Introduction

Macrocyclic peptides often display superior chemical stability and permeability compared to linear peptides [334]. Their clinical efficacy is exemplified by the existing antiviral therapies for HCV and HIV [13,60]. As the inhibitors that were reported during the earlier phases of the COVID-19 pandemic were based on inhibitors of SARS-CoV-1 M^{pro} [28] or other viral proteases [77,78], application of the RaPID technology was considered in order to identify novel tight-binding M^{pro} cyclic peptide inhibitors.

5.3 Contributions to the manuscript

A novel macrocyclic M^{pro} inhibitor with non-canonical amino acids is reported in Manuscript 4. This work involved collaborative research between Oxford University, Diamond Light Source, the University of Tokyo, and the University of Newcastle. At the beginning of the pandemic, a novel method for the incorporation of unnatural amino acids into peptides had just been devised by the group of Prof. Hiroaki Suga. This methodology was then applied to identify peptidic inhibitors for M^{pro} containing unnatural amino acids.

5.3.1 Manuscript 4 (published prior to thesis deposition)

My contributions to manuscript 4 included investigations into the inhibitory activity and mode of inhibition of cyclic peptide M^{pro} inhibitors provided by the Suga group (University of Tokyo). Macrocyclic peptides e.g. TMGM4-05 were identified using RaPID method and subsequently synthesized using SPPS by Takashi Miura (University of Tokyo). Macrocyclic peptides used for alanine scanning the TMGM4-05 cyclic peptide were supplied by Siddhique Amin (University of Newcastle). My specific contributions to the manuscript were as follows:

- Modifying the M^{pro} assay, using the 37-mer substrate in combination with an internal peptide standard.
- Carrying out LC-MS studies to monitor M^{pro}-catalysed turnover of the macrocyclic peptides.

- Designing cyclic peptides (CPs) variants based on insights obtained from work with natural M^{pro} substrates and variants thereof.

Further to the work described in manuscript 4, I also performed studies with:

- A library of variants of the lead cyclic peptide TMGM4-05.
- A library of truncated cyclic and linear peptide variants of TMGM4-05.
- A library of TMGM4-05 peptide variants for alanine scanning without the γ amino acid (TMGM4-05A).
- Protein-observed mass spectrometry M^{pro} assays with CPs
- A library of disulfide-rich cyclic peptides (provided by Dr. Wenyu Liu, University of Tokyo) as a new class of M^{pro} inhibitors.

5.4 Results

5.4.1 Optimisation of the M^{pro} assay with a 37-mer substrate and an internal peptide standard

Initial screening of CP inhibitors using the 11-mer M^{pro} substrate by SPE-MS indicated that the CPs tested inhibit M^{pro} with IC_{50} values in the nanomolar range below the assay limit (i.e. IC_{50} < (enzyme concentration used/2), data not shown). It, therefore, seemed reasonable to lower the enzyme concentration in the assay. As the CPs are likely to inhibit via a substrate-competitive mode of inhibition, the use of a substrate concentration close to the K_m (5.74 µM) was considered during the assay optimisation. As MS assays can be susceptible to reduced substrate/product fragment ionisation due to the presence of inhibitors (as discussed in chapter 3), an internal standard derived from the product peptide was incorporated into the assay. Considering that the CP samples were not available in high quantities, optimisation of the assay volume was undertaken to reduce the volume of samples and reagents consumed.

Under the previous M^{pro} assay conditions with the 37-mer substrate, the conversion was linear for up to 10 min after initiating the reaction (Figure 5.1A). Reducing the M^{pro} concentration and increasing the 37-mer substrate concentration under the optimised conditions prolonged the linear conversion range to 20 min after initiation of the reaction (Figure 5.1B, right). An internal standard (SGFRMAF-NH_2, 942 Da) was designed based on the C-terminal product (SGFRMAFPS-NH_2, 1125 Da) obtained from M^{pro}-catalysed hydrolysis of the 37-mer substrate. The internal standard and the C-terminal product were observed to ionise to a similar degree (Figure 5.1B, left). Chapter 4 describes studies on product-mediated inhibition of M^{pro}. To prevent any alteration of M^{pro} activity by the product-derived internal standard, the internal standard was only added into the reaction mixture once the reaction had been quenched with aqueous formic acid (Figure 5.1C).

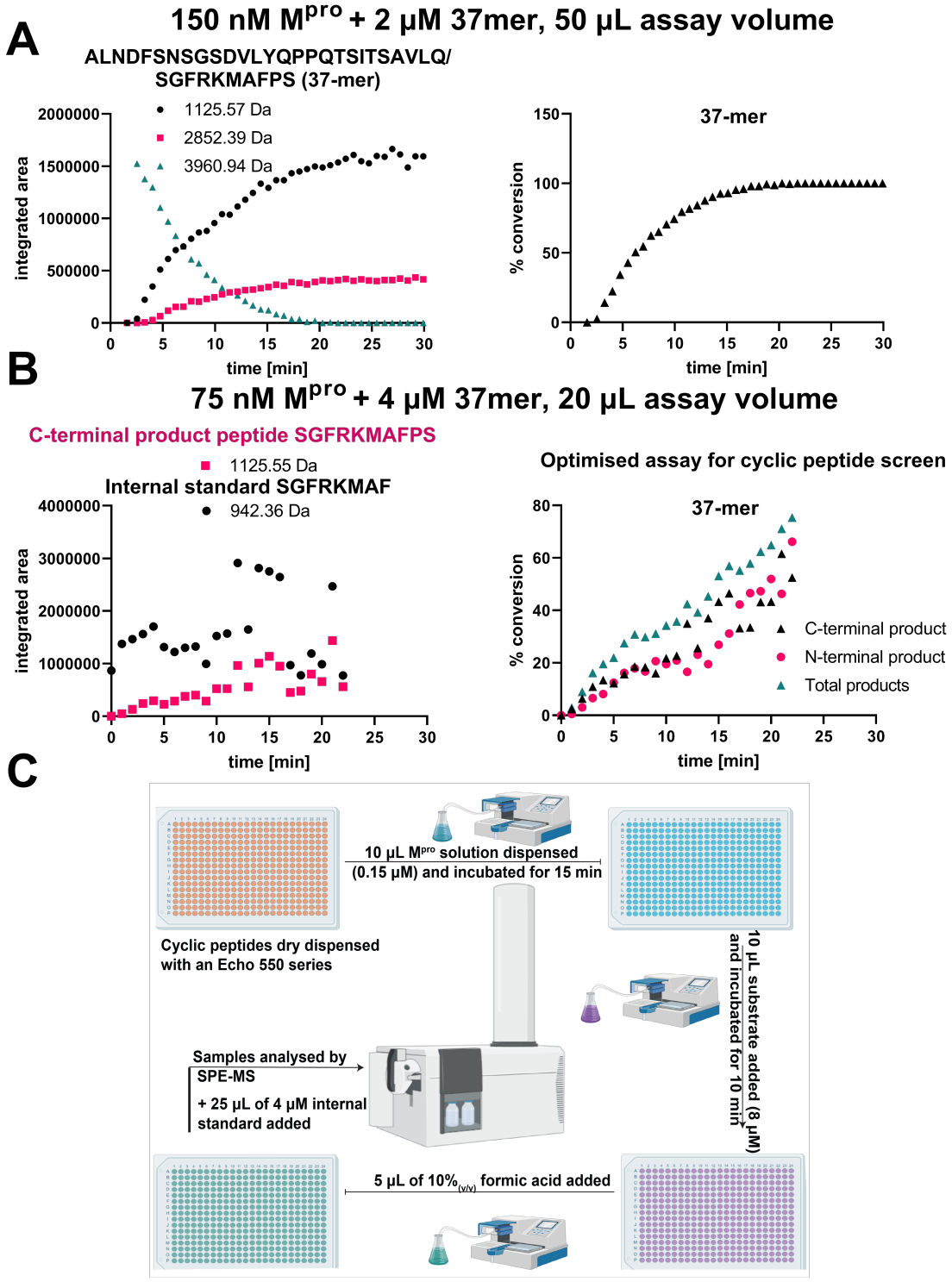


Figure 5.1 | Optimisation of the M^{pro} assay with the 37-mer substrate for screening cyclic peptides. **A.** M^{pro}-catalysed hydrolysis of the 37-mer substrate, as described in chapter 3. 2852 Da and 1125 Da correspond to the N- and C-terminal products. The abundance of the C-terminal product was higher than the N-terminal product, hence, product formation (right) was calculated based on the abundance of the C-terminal product only. **B.** The optimised M^{pro} assay with the 37-mer substrate for cyclic peptide screening. **C.** Set up for the 37-mer assay for screening cyclic peptides. Conditions: 75 nM M^{pro}, 2 µM 37-mer substrate (ALNDFSNSGSDVLYQPPQTSITSAVLQ/SGFRKMAFPS-NH_2) in the reaction buffer. The internal standard peptide (2 µM, SGFRKMAF-NH_2), was added after the reaction was quenched.

5.4.2 Comparison of pIC_{50} values obtained for cyclic peptide M^{pro} inhibitors with or without normalisation to an internal standard

IC_{50} values were determined for 26 cyclic peptides using ebselen as a positive control using both: A) the standard method [% product formation = (integration of C-terminal product peak) / (sum of integration for C-terminal product peak and substrate peak) * 100%] (Chapter 2), and B) quantification of observed C-terminal product based on the integration of the 2 μM internal standard [normalised concentration of product peptide = (peak integration of C-terminal product peptide * 2)/ peak integration of C-terminal product peptide based internal standard]. Note that only one set of dose response studies was performed with the 37-mer substrate and the internal standard, but the results were analysed by the above stated methods using Excel and subsequently in GraphPad Prism. IC_{50} values were then transformed using a logarithmic function to obtain pIC_{50} values. The pIC_{50} values obtained via methods A and B were further assessed to determine if they were statistically different. This analysis enabled probing of the robustness of the 37-mer substrate assay. Unlike the method comparison between the 11-mer substrate and the 37-mer substrate described in Chapter 3, where a poor correlation was observed for weak inhibitors, the IC_{50} values of CPs that did not display relatively potent M^{pro} inhibition afforded similar results using both methods A and B. A good distribution of pIC_{50} values was observed upon inspection of the resultant scatter plot (Figure 5.2A). A positive correlation was also observed between the two data sets (Pearson correlation coefficient = 0.93, r^2 = 0.82, Figure 5.2B). The log of the IC_{50} for inactive compounds (IC_{50} 0 μM) cannot be calculated, therefore, inactive compounds were omitted from the pIC_{50} analyses.

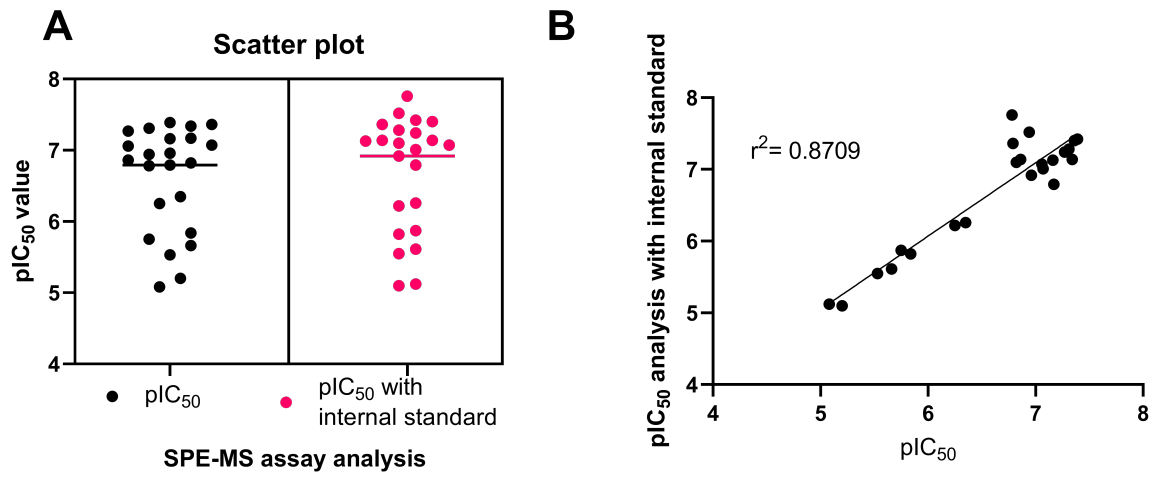


Figure 5.2 | Distribution and correlation analyses of pIC_{50} values obtained with or without normalisation to an internal standard. A. Scatter plot of the two data sets suggests a good distribution of pIC_{50} values. **B.** Correlation analysis of the data set from two different analytical methods to obtain pIC_{50} values.

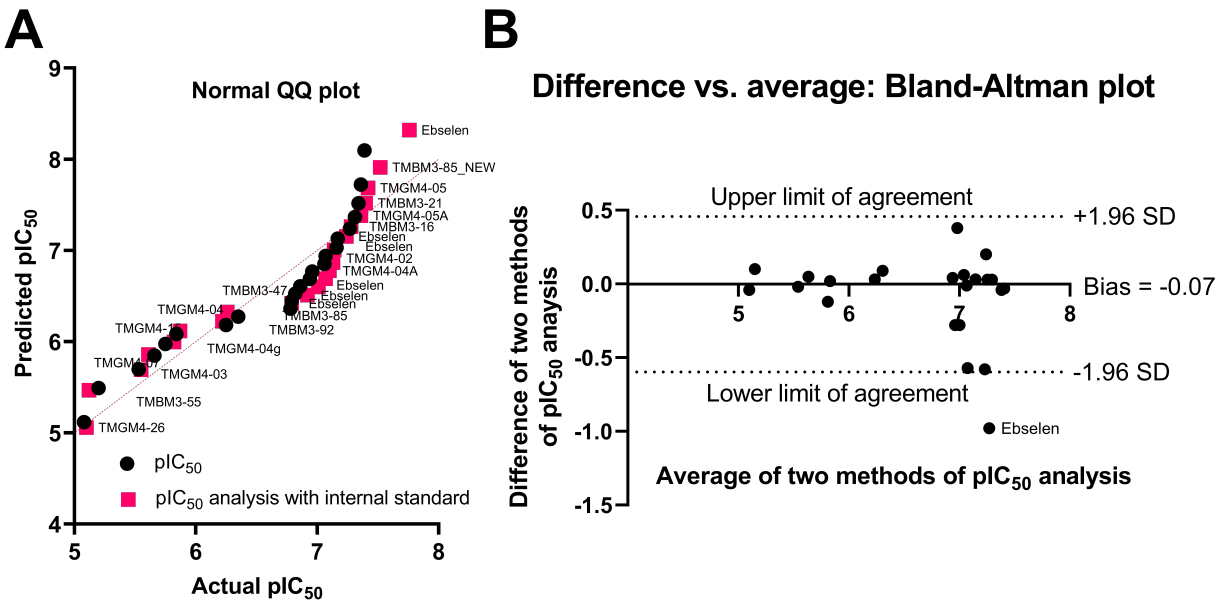


Figure 5.3 | Bland-Altman plot to analyse the method of pIC_{50} calculation via the standard method (A) and the normalised method using an internal standard (B). A. QQ plot to analyse the normality of the data sets obtained from the two different methods of evaluating pIC_{50} values. **B.** The Bland-Altman plot reveals the data are near identical with a bias value of only -0.07 ± 0.27 μM.

As both the pIC_{50} data sets displayed a normal distribution (Figure 5.3), a condition that is required for Bland-Altman analysis to compare differences in methods, the Bland-Altman method was used

to calculate whether the pIC_{50} values obtained from methods A and B generate pIC_{50} values in agreement [335]. The difference between the pIC_{50} values obtained from the two methods (i.e., A - B) was plotted against the mean of the two values (i.e., A + B / 2) (Figure 5.3B). If the two sets of pIC_{50} values are identical, there will be no difference between them and they will lie on the horizontal x-axis. In cases where the values are different, they will appear above or below the x-axis. The difference between the two methods of analysis can be represented as a bias value which represents the average of the differences between all of the values. A very small value for bias (-0.07 ± 0.27) was obtained, showing that methods A and B generate very similar results (Figure 5.3B). Furthermore, all data were within limits of agreement (mean ± 1.96 SD), with the exception of one of the measurements for ebselen which lies below the lower limit. These results suggest that methods A and B are equally suited for dose response curve estimations. The internal standard was included in the subsequent assay setup for screening cyclic peptides, while IC_{50} values were determined using method A.

5.4.3 Inhibition studies of peptide variants derived from TMGM4-05

A crystal structure of the TMGM4-05 cyclic peptide in complex with M^{pro} has been obtained [experiments performed by Dr. David Owen (Diamond Light Source, Oxford)], as described in manuscript 4. These data reveal that TMGM4-05 binds at the active site of M^{pro}. Residues of the CPs are represented by a single letter code in manuscript 4, apart from the D-Tyr residue which is involved in backbone thioether cyclization. This nomenclature has been conserved in this chapter (Figure 5.4). The γ-amino acid in the TMGM4-05 CP occupies the S1′ subsite, the H3 residue. occupies the S1 subsite, F2 occupies the S2 subsite, and the tyrosinyl side chain of ^{Ac}Tyr (D-Tyr) sits in the S4 subsite while its main chain lies in the S3 subsite. The side chains of other residues (e.g. the side chains of L5, N6, and R10) of this CP are solvent exposed.

Substitutions were carried out at various positions of the CP to assess the significance of 1) the γ-amino acid, and 2) the solvent exposed residues on inhibitory activity. Interestingly, an intrapeptide interaction between Y9 and H3 was observed in the M^{pro}: CP complex (manuscript 4). Therefore,

variants of TMGM4-05 were screened to investigate whether the crystallographically observed Y9-H3 intrapeptide interaction is critical for the inhibitory activity of TMGM4-05.

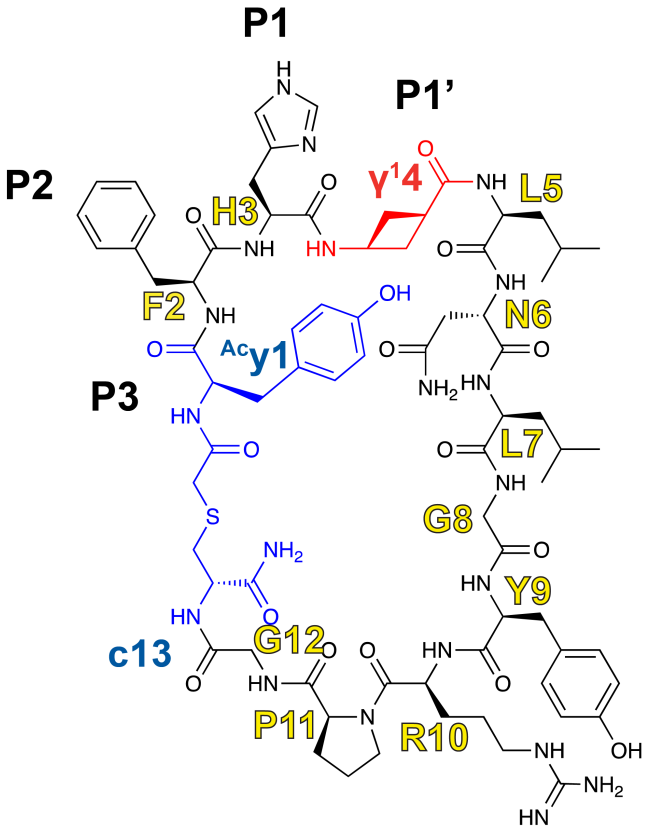


Figure 5.4 | Nomenclature of the residues in macrocyclic peptide TMGM4-05. The γ-amino acid at the P1′ position is shown in red. F2 and H3 (in the P2 and P1′ positions of TMGM4-05) occupy the S2 and S1′ subsites, respectively, in M^{pro}.

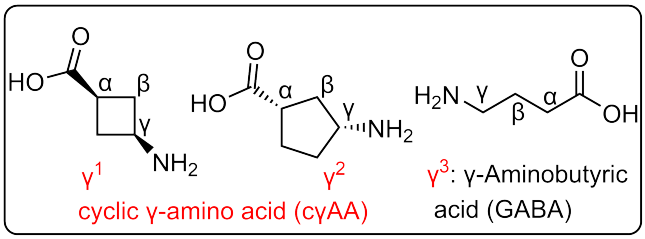
5.4.3.1 Significance of the unnatural amino acid at the P1′ position towards the potency of TMGM4-05 and related macrocyclic peptides

The incorporation of unnatural amino acids into CPs by the RaPID selection method has been applied for the development of diverse CP inhibitors. The crystallographic studies described in manuscript 4 reveal that the γ-amino acid in TMGM4-05 is positioned at the P1′ equivalent position, where various substitutions (e.g. small amino acids, hydrophobic amino acids, benzyl esters, other ester derivatives) are known to be tolerated [28,330,331]. Based on these observations, the contributions of the γ-amino acid on the M^{pro} inhibitory activity were investigated.

Dose response studies with the TMGM4-05A variant derived from TMGM4-05 by substituting the γ-amino acid with an Ala residue suggested that the γ-amino acid (γ^1) can be substituted with a small

hydrophobic residue (i.e. Ala) without compromising the potency of inhibition (IC_{50} 0.01 ± 0.00 μM, close to the limit of assay detection). A polar residue (i.e. Asn) at the P1′ (TMGM4-05N) equivalent position can also be tolerated, although a reduction in potency was observed (IC_{50} 1.00 ± 0.10 μM). Substitution with a positively charged residue (i.e. Arg) (TMGM4-05R) resulted in the complete ablation of its inhibitory activity. Thus, the active site of M^{pro} favours polar overcharged residues at the S1′ subsite, at least in the cases of TMGM4-05R and TMGM4-05N. Additionally, other variants were prepared in which the γ-amino acid was substituted with a natural signalling molecule that has an acyclic γ-amino acid core [γ^3, γ-aminobutyric acid (GABA)]. This led to a minor reduction in potency (IC_{50} 0.20 ± 0.10 μM). Thus, the synthetic cyclic γ^1 residue, though not essential, is the superior γ-amino acid at the P1′ position to achieve potent M^{pro} inhibition. γ-Amino acids were also incorporated at the P1′ position of TMGM4-02 and TMGM4-10, as these macrocyclic peptides shared the same FH motif present in the TMGM4-05 at the P2 and P3 positions which are adjacent to the site of the backbone thioether cyclisation. Inhibitory studies with these peptides indicate that the presence of the γ-amino acid was more important for the M^{pro} inhibition by TMGM4-02, but not by TMGM4-10 (Table 5.1).

Table 5.1: Dose response curve studies with the γ-amino substituted cyclic peptide variants of TMGM4-05, TMGM4-02 and TMGM4-10. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). The lowercase represents the amino acids with a D-configuration. The amino acid residues in blue are part of the thioether backbone of the CP, whereas the letter in red represents the P1′ equivalent position.



Cyclic peptides	Sequence	IC_{50} (µM)
TMGM4-05	^{Ac}yFHγ^{1}LNLGYRPGc-NH_2	0.05 ± 0.01
TMGM4-05g	^{Ac}yFHγ^{3}LNLGYRPGc-NH_2	0.20 ± 0.10
TMGM4-05A	^{Ac}yFHALNLGYRPGc-NH_2	0.01 ± 0.00
TMGM4-05N	^{Ac}yFHNLNLGYRPGc-NH_2	1.00 ± 0.10
TMGM4-05R	^{Ac}yFHRLNLGYRPGc-NH_2	>25
TMGM4-02	^{Ac}yFHγ^{1}FPPHTYTLc-NH_2	0.46 ± 0.32
TMGM4-02A	^{Ac}yFHAFPPHTYTLc-NH_2	>25
TMGM4-10	^{Ac}yFHγ^{1}FRDLYILGc-NH_2	0.04 ± 0.02
TMGM4-10A	^{Ac}yFHAFRDLYILGc-NH_2	0.05 ± 0.03

5.4.3.2 Significance of the solvent exposed residues on the potency of TMGM4-05

The potential of a macrocyclic peptide as therapeutic agents is often limited due to low cell permeability [336] (Chapter 1); synthetic modifications, including capping polar residues or incorporating cell penetrating peptide tails, can be performed to increase their cell permeability [334]. The R10 residue of TMGM4-05 was selected as a model solvent-exposed residue to investigate substitutions at this position for improved cell permeability, as a substitution at this position was not expected to impact inhibitory potency. The substitution of R10 to Ala and Phe is tolerated with a minor reduction in inhibitory potency, suggesting modifications of solvent exposed residue at this position (to improve the peptide permeability) may compromise the potency. Hence, the systematic substitution of the residues at these positions is crucial to determine the ideal position for chemical modifications without tremendously affecting the potency of these cyclic peptides.

Table 5.2: Dose response curve studies with the γ-amino substituted cyclic peptide variants. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). Lower case represents amino acid residues with a D-configuration. The amino acid residues in blue are part of the thioether backbone of the CP, whereas the letter in red represents the P1′ equivalent position.

Cyclic peptides	Sequence	IC_{50} (µM)
TMGM4-05	^{Ac}yFHγ^{1}LNLGYRPGc-NH_2	0.05 ± 0.01
TMGM4-05 R10A	^{Ac}yFHγ^{1}LNLGYAPGc-NH_2	0.44 ± 0.10
TMGM4-05 R10F	^{Ac}yFHγ^{1}LNLGYFPGc-NH_2	0.27 ± 0.09

5.4.3.3 Investigating the significance of the crystallographically observed H3-Y9 intrapeptide interaction for TMGM4-05 inhibitory potency

Crystallographic studies reveal a potential interaction between the tyrosinyl and imidazole rings of Y9 and H3, respectively, in the complex of TMGM4-05 bound to M^{pro}. Initially, cyclic peptides with substitutions to either H3 or Y9 were designed, followed by the double variant TMGM4-05 H3Y and Y9W. Substitutions of H3 by small non-polar Ala (TMGM4-05 H3A), negatively charged Glu (TMGM4-05 H3E), serine (TMGM4-05 H3S), and tyrosine (TMGM4-05 H3Y) all resulted in a loss in inhibitory activity (Table 5.3). Interestingly, the substitution of H3 by a bulky Trp (TMGM4-05 H3W) does not impact the activity of the macrocycle. As expected, the substitution of H3 by Gln (TMGM4-05 H3Q) at the P1 position does not negatively alter the activity of TMGM4-05 (Table 5.3). Substitution of Y9 with an Ala residue (TMGM4-05 Y9A) renders the CP inactive; however, as observed with TMGM4-05 H3W, a peptide with a Trp residue at Y9 (TMGM4-05 Y9W) inhibited M^{pro}, though 10-fold less potently than TMGM4-05. These results suggest that the H3-Y9 interaction plays an important role in the activity of TMGM4-05.

To further assess the H3-Y9 intrapeptide interaction, truncated cyclic (i.e. TMGM4-05 trun1, TMGM4-05 trun2) and linear variants (i.e. TMGM4-05 lin0, TMGM4-05 lin1, TMGM4-05 lin2, TMGM4-05 lin3, TMGM4-05 lin5) of the TMGM4-05 peptide were screened for M^{pro} inhibition (Table 5.3). No inhibitory activity was detected from the truncated peptides, suggesting that the conformation of the 13-mer macrocyclic peptide is crucial for its interactions with M^{pro}.

Table 5.3: Dose response curve studies to assess the contributions of the H3-Y9 intrapeptide interaction on M^{pro} inhibition. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). The lowercase represents amino acid residues with a D-configuration. The amino acid residues in blue are part of the thioether backbone of the CP, whereas the letter in red represents the P1′ equivalent position.

Cyclic peptides	Sequence	IC_{50} (µM)
TMGM4-05	^{Ac}yFHγ^1LNLGYRPGc-NH_2	0.05 ± 0.01
TMGM4-05 H3A	^{Ac}yFAγ^1LNLGYRPGc-NH_2	>25
TMGM4-05 H3E	^{Ac}yFEγ^1LNLGYRPGc-NH_2	>25
TMGM4-05 H3S	^{Ac}yFSγ^1LNLGYRPGc-NH_2	>25
TMGM4-05 H3Y	^{Ac}yFYγ^1LNLGYRPGc-NH_2	>25
TMGM4-05 H3W	^{Ac}yFWγ^1LNLGYRPGc-NH_2	0.06 ± 0.02
TMGM4-05 H3Q	^{Ac}yFQγ^1LNLGYRPGc-NH_2	0.01 ± 0.00
TMGM4-05 Y9A	^{Ac}yFHγ^1LNLGAPGc-NH_2	>25
TMGM4-05 Y9W	^{Ac}yFHγ^1LNLGARPGc-NH_2	0.57 ± 0.33
TMGM4-05 H3Y and Y9W	^{Ac}yFYγ^1LNLGARPGc-NH_2	>25
TMGM4-05 trun1	^{Ac}yFHγ^1 LNLGc-NH_2	>25
TMGM4-05 trun2	^{Ac}yHγ^1 LNLGc-NH_2	>25
TMGM4-05 lin0	yFHγ^1LNLGYRPG	>25
TMGM4-05 lin1	YFHγ^1LNL	>25
TMGM4-05 lin2	yFQγ^1LNL	>25
TMGM4-05 lin3	yFHγ^1LNL	>25
TMGM4-05 lin5	AyFHγ^1LNL	>25

5.4.3.4 Alanine scan of the TMGM4-05A peptide

As the TMGM4-05A variant displayed similar activity as TMGM4-05, alanine scanning of the TMGM4-05A peptide was performed. Alanine scanning of the TMGM4-05A did not ablate the activity of the peptide with alanine at any position. However, the TMGM4-05A F2A, TMGM4-05A H3A, TMGM4-05A L7A, TMGM4-05A G8A, and TMGM4-05A Y9A variants manifested 10-fold lower inhibitory activity, suggesting that the side chains of these residues may be important in interactions with M^{pro}. For instance, the leucine of TMGM4-05A L7 is observed to fit snugly into the S3′ pocket, and its substitution with Ala would be expected to reduce the Van der Waals interactions at that site. Interactions at the S1 and S2 subsites are crucial for substrate recognition [330,331], consistent with

the reduced activity of the F2A and H3A variants. Similarly, as discussed in the above section, the H3-Y9 interaction is crucial for M^{pro} inhibition; consistent with this observation, the TMGM4-05A Y9A manifests reduced inhibitory activity compared to TMGM4-05A. However, in comparison to the original TMGM4-05 CP, the inhibitory activity of TMGM4-05A Y9A is reduced but not ablated (Tables 5.3, 5.4). In conclusion, these results indicate that the inhibitory activity of TMGM4-05 is a result of multiple additive factors rather than arising from interactions of a single or small set of residues (e.g. the FH motif that occupies the S1 and S2 M^{pro} subsites) with M^{pro}.

Table 5.4: Dose response curve studies with the TMGM4-05A variants. [a] represents peptide purity <90%, [b] represents peptide purity >90%. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). The lowercase represents amino acid residues with a D-configuration. The amino acid residues in blue are part of the thioether backbone of the CP, whereas the letter in red represents the P1′ equivalent position.

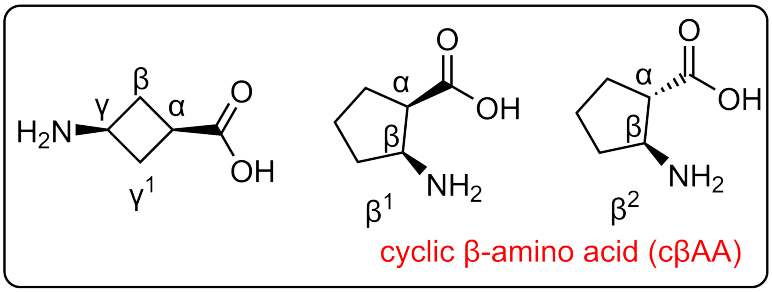
Cyclic peptides	Sequence	IC_{50} (µM)[a]	IC_{50} (µM)[b]
TMGM4-05	^{Ac}yFHγ^1LNLGYRPGc-NH_2	0.05 ± 0.01	
TMGM4-05A	^{Ac}yFHALNLGYRPGc-NH_2	0.79 ± 0.21	0.01 ± 0.01
TMGM4-05A F2A	^{Ac}yAHALNLGYRPGc-NH_2	4.3 ± 0.2	2.2 ± 1.3
TMGM4-05A H3A	^{Ac}yFAALNLGYRPGc-NH_2	3.4 ± 0.1	0.28 ± 0.24
TMGM4-05A L5A	^{Ac}yFHAANLGYRPGc-NH_2	0.58 ± 0.01	0.04 ± 0.03
TMGM4-05A N6A	^{Ac}yFHALALGYRPGc-NH_2	0.56 ± 0.02	0.23 ± 0.15
TMGM4-05A L7A	^{Ac}yFHALNAGYRPGc-NH_2	4.06 ± 0.35	∼2.72
TMGM4-05A G8A	^{Ac}yFHALNLAYRPGc-NH_2	3.1 ± 0.5	1.19 ± 1.08
TMGM4-05A Y9A	^{Ac}yFHALNLGARPGc-NH_2	4.93 ± 0.08	∼7
TMGM4-05A R10A	^{Ac}yFHALNLGYAPGc-NH_2	0.53 ± 0.03	0.06 ± 0.03
TMGM4-05A P11A	^{Ac}yFHALNLGYRAGc-NH_2	0.84 ± 0.12	0.05 ± 0.03
TMGM4-05A G12A	^{Ac}yFHALNLGYRPAc-NH_2	0.95 ± 0.33	0.40 ± 0.30
TMGM4-05A L-Tyr	YFHALNLGYRPAc-NH_2	4.98 ± 1.28	4.27 ± 1.54

5.4.4 M^{pro} inhibition studies with shorter cyclic peptides

In addition to the truncated TMGM4-05 peptides, additional shorter CPs (TMGMS001, TMGMS004, TMGMS005, TMGMS006, TMGMS007, TMGMS008, TMGMS014, TMGMS017, TMGMS505, TMGMS506, and TMGMS901) were identified using the RaPID selection method, which contained a

non-canonical β-amino acid or the previously incorporated γ-amino acid (Table 5.5). However, most of the subsequently selected peptides generated using the RaPID selection display poor inhibitory activity compared to TMGM4-05. The most potent CP of these is TMGMS014 (^{Ac}yFHNHγ^{1}YRLG c-NH_2, IC_{50} 2.05 ± 0.91 μM), which also contained the FH motif present in TMGM4-05. Other macrocyclic peptides with the FH motif, TMGMS004 and TMGMS001, also display micromolar IC_{50}s (Table 5.5). Interestingly, TMGMS505, a short CP with an IH motif, manifests activity (IC_{50} 5.68 ± 1.39 μM) comparable to TMGMS014. Hence, isoleucine can apparently be substituted at the P2 position, based on the activity observed for TMGMS505. Note that, although shorter CPs (7-11-mer) are less effective M^{pro} inhibitors than TMGM4-05, they are more potent than many identified small molecule M^{pro} inhibitors, and are potentially more selective [337].

Table 5.5: Dose response curve studies with cyclic peptides containing unnatural β and γ-amino acids. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). The lowercase represents amino acid residues with a D-configuration. The amino acid residues in blue are part of the thioether backbone of the CP, whereas the letter in red represents the non-canonical amino acid residues.



Cyclic peptides	Sequence	IC_{50} (µM)
TMGM4-05	^{Ac}yFHγ^{1}LNLGYRPGc-NH_2	0.05 ± 0.01
TMGMS001	^{Ac}yHNYFHRHγ^{1}H c-NH_2	17.5 ± 1.8
TMGMS004	^{Ac}yFHNQγ^{1}YRFGc-NH_2	10.9 ± 4.9
TMGMS005	^{Ac}yHNYFHβ^{1}Hγ^{1}Hc-NH_2	>25
TMGMS006	^{Ac}yFHNDYYQγ^{1}Gc-NH_2	>25
TMGMS007	^{Ac}yFHβ^{2}FQγ^{1}LF c-NH_2	>25
TMGMS008	^{Ac}yFHNSLYQγ^{1}Gc-NH_2	>25
TMGMS014	^{Ac}yFHNHγ^{1}YRLG c-NH_2	2.1 ± 0.9
TMGMS017	^{Ac}yFHβ^{2}Hγ^{1} LFTNc-NH_2	>25
TMGMS505	^{Ac}yTQγ^{1}IT c-NH_2	>25
TMGMS506	^{Ac}yIHγ^{1}SYC c-NH_2	5.7 ± 1.4
TMGMS901	^{Ac}yFIHTFHγ^{1}LNc-NH_2	∼9

5.4.5 Screening of tricyclic disulfide rich cyclic peptides as M^{pro} inhibitors

Cell based assays were performed by the groups of Prof. Nadia Naffakh (Institut Pasteur) and Prof. Miles Caroll (University of Oxford) to screen CP inhibitors of M^{pro}. Prof. Naffakh used a luciferase reporter-based M^{pro} assay [338]. However, no notable inhibitory activity of the CPs (i.e. TMGM4-05, TMGM4-05A, TMGM4-04A, TMGM4-02) was observed (data not shown), which likely arises from the poor cell permeability of the screened CPs.

Some cyclic disulfide rich peptides (DRPs) not only possess superior protease stability due to disulfide crosslinking but can be more cell permeable, e.g. Kalata B1 and MCOTI-II cyclotides

(Chapter 1, Section 1.3.4.1) [127]. An attempt was therefore made to screen DRPs that bind tightly to M^{pro}. Wenyu Liu in the lab of Prof. Hiroki Suga applied the RaPID platform to identify DRPs, using controlled enzymatic and Pd(II) catalysed deprotection of thiols to control the formation of disulfide bonds. Disulfide reshuffling was observed during the Pd(II)-catalysed deprotection step. Different fractions were collected during HPLC purification, and the purified peptides were analysed by SPE-MS for M^{pro} inhibition. Each fraction (labelled as peak a, b, or c) likely represents a different disulfide connectivity in the DRP; only the DRP MP3062, based on the scaffold of the natural DRP MCOTI-II, was purified with a single peak.

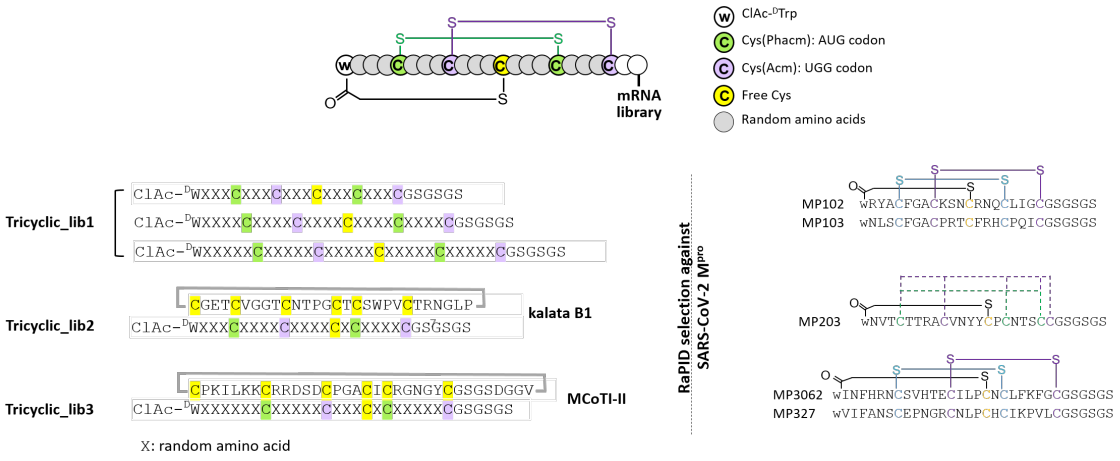


Figure 5.5 | Selection of disulfide rich peptides identified from RaPID screening. Three tricyclic libraries, including the peptide frameworks of Kalata B1 and MCOTI-II, were used to perform RaPID selection against M^{pro}. This led to the identification of the DRPs MP102, MP103, MP203, MP3062 and MP327. The experiment was performed by Dr. Wenyu Liu (University of Tokyo).

5.4.5.1 Results

Different levels of M^{pro} inhibition were observed from the different HPLC purified peptides that were obtained from the same parent DRP. For instance, peak b of MP102 DRP shows the highest M^{pro} inhibition (IC_{50} 0.94 ± 0.04 μM) compared to peak a (IC_{50} 4.4 ± 0.4 μM), a mixture of peaks b and c (IC_{50} 1.49 ± 0.23 μM), and peak c (IC_{50} 2.07 ± 0.08 μM). By contrast, the three peaks from the MP103 DRP display comparable levels of M^{pro} inhibition: peak a (IC_{50} 2.63 ± 0.25 μM), mixture of peaks a and b (IC_{50} 3.72 ± 0.39 μM), and a mixture of peaks b and c (IC_{50} 3.26 ± 0.26 μM), while peak c displays nanomolar inhibition (IC_{50} 0.71 ± 0.00 μM). MP3062 displays good M^{pro} inhibition (IC_{50} 1.53 ± 0.14

µM); the activity of MP3062 is comparable to the M^{pro} inhibition observed for the two peaks collected for MP327 (peak a IC_{50} 1.10 ± 0.47 µM, peak b IC_{50} 2.37 ± 0.29 µM), a DRP which is based on the MCOTI-II scaffold. However, no inhibitory activity was observed for the DRP MP203, which was based on the Kalata B1 scaffold. Despite the undetermined disulfide connectivity patterns of these DRP M^{pro} inhibitors, the results manifest the potential of DRPs to be developed as potent M^{pro} inhibitors.

Table 5.6: Dose response studies with the DRP M^{pro} inhibitors. M^{pro} inhibition assays were performed using SPE-MS, as described in manuscript 4. Conditions: 0.075 µM M^{pro}, 4.0 µM 37-mer substrate in the reaction buffer, and 2.0 µM 8-mer peptide (internal standard). Lower case letters represents amino acids with a D-configuration. The disulfide pairs are colour coded in green and red. Cysteines indicated in yellow forms the thioether backbone with D-Tyr in the macrocyclic DRPs.

Cyclic peptides	Sequence	IC_{50} (µM)
MP102 peak a	wRYACFGACKSNCRNQCLIGGSGSGS	4.4 ± 0.4
MP102 peak b	wRYACFGACKSNCRNQCLIGGSGSGS	0.94 ± 0.04
MP102 peak b,c	wRYACFGACKSNCRNQCLIGGSGSGS	1.49 ± 0.23
MP102 peak c	wRYACFGACKSNCRNQCLIGCGSGSGS	2.07 ± 0.08
MP103 peak a	wNLSCFGACPRTCFRHCPQICGSGSGS	2.63 ± 0.25
MP103 peak a,b	wNLSCFGACPRTCFRHCPQICGSGSGS	3.72 ± 0.39
MP103 peak b,c	wNLSCFGACPRTCFRHCPQICGSGSGS	3.26 ± 0.26
MP103 peak c	wNLSCFGACPRTCFRHCPQICGSGSGS	0.71 ± 0.00
MP327 peak a	wVIFANSCEPNGRCNLPCHCIKPVLCGSGSGS	1.10 ± 0.47
MP327 peak b	wVIFANSCEPNGRCNLPCHCIKPVLCGSGSGS	2.37 ± 0.29
MP203 peak a	wNVTCTTRACVNYYCPCNTSCGSGSGS	>12.5
MP203 peak b	wNVTCTTRACVNYYCPCNTSCGSGSGS	>12.5
MP3062 single peak	wINFHRNCSVHTECILPCNCLFKFGCGSGSGS	1.53 ± 0.14

LC-MS experiments monitoring interactions between DRPs and M^{pro}

M^{pro} has 11 cysteine residues, in addition to Cys145 which forms the catalytic dyad together with His41. As disulfide reshuffling was observed to occur in DRPs, LC-MS experiments were performed to investigate if the DRPs can covalently interact with M^{pro} by forming disulfide bonds with any of the M^{pro} cysteine residues. Interestingly, no evidence for the covalent reaction of the DRPs that inhibited M^{pro} substrate turnover was obtained. Conversely, inactive DRP MP203 slowly reacted with M^{pro}, as observed by LC-MS, suggesting that DRPs can covalently react with M^{pro} (Figure 5.6). However, as

this DRP did not inhibit M^{pro}, DRP MP203 likely modified non-active site M^{pro} cysteine residues.

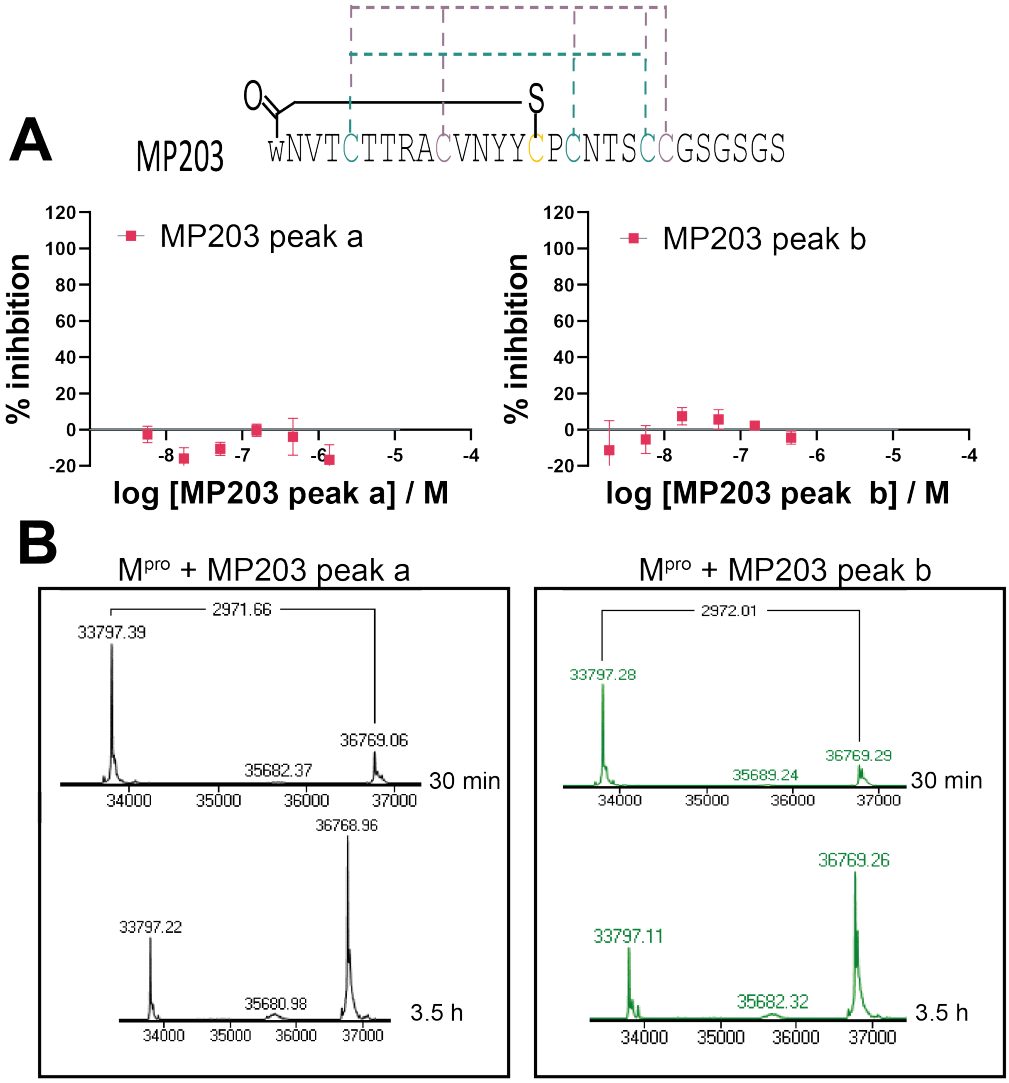


Figure 5.6 | LC-MS analysis reveals that the inactive DRP MP203 covalently binds to M^{pro}. A. MP203 peak a and peak b do not inhibit M^{pro}. **B.** LC-MS studies suggest that the MP203 peaks can covalently react with M^{pro} in a time-dependent manner. Conditions: 1.5 µM M^{pro}, 5 µM DRP in 20 mM HEPES pH 7.5, 50 mM NaCl.

5.5 Discussion

Backbone thioether cyclised CPs display potent inhibition against M^{pro}; in particular, TMGM4-05 was observed to inhibit with an IC_{50} value of 0.05 ± 0.01 µM. The FH motif in these macrocycles was found to be important for activity. Substitution of the His in the FH motif resulted in the loss of activity. However, in TMGM4-05A, a variant of TMGM4-05, the substitution of the H in the FH motif with Ala resulted in a reduction but not ablation of activity. The H3-Y9 intrapeptide interaction was

found to be important for the activity of TMGM4-05, as evidenced by the loss of inhibitory activity for the TMGM4-05 H3A and TMGM4-05 Y9A variants, but not for the TMGM4-05A variant. As observed by crystallography, H3 of TMGM4-05 is positioned to interact with Glu166 and Ser1, while Y9 interacts with Ser1. The absence of the γ-amino acid in TMGM4-05A may allow the macrocyclic peptides TMGM4-05A H3A and TMGM4-05A Y9A to adopt alternative conformations where the H3-Y9 interaction is not crucial for inhibition. The observation that the potencies of TMGM4-05A H3A and TMGM4-05A Y9A are lower than those of TMGM4-05A and TMGM4-05, indicates that the H3-Y9 interaction is favourable for potent M^{pro} inhibition.

The investigated CPs are more potent than reported isoxazole-containing macrocycles [148] or CPs based on the C-terminal auto-cleavage sequence of M^{pro} [147] (Chapter 1, Section 1.3.4.3). A RaPID selection screen was applied to identify shorter (7-11)-mer CPs yielded moderately potent inhibitors (Table 5.5) which manifest higher inhibition than reported isoxazole or substrate-based macrocyclic peptides [147,148]. A RaPID selection screen has also been performed for M^{pro} by a different group, who identified CPs that were similarly potent to TMGM4-05 [146]. Johansen-Leete *et al.* identified CP Se-1 (sequence: ^{Ac}yLQYAVLRHKRREC-NH_2, IC_{50} 0.07 ± 0.01) which has an identical 13-mer length as TMGM4-05, but which uses a terminal non-canonical residue, D-Tyr, for thioether backbone cyclisation. The Leu and Gln at the P2 and P1 positions of Se-1 are identical to many M^{pro} cleavage sites and occupy the S2 and S1 subsites as observed in a crystal structure of the Se-1:M^{pro} complex (PDB: 7RNW). By contrast, TMGM4-05 consists of Phe and His at P2 and P1 equivalent positions, respectively. Substitution of the P1 His to the conserved Gln residue in TMGM4-05 reduces its dissociation constant by ∼10 fold (manuscript 4), suggesting Gln is preferred at P1 as expected. The D-Tyr, part of the backbone thioether in both TMGM4-05 and Se-1, occupies the S4 subsite.

Despite the potent inhibition of M^{pro} by Se-1 and TMGM4-05, to date these CPs have been observed to be inactive in cells. Addition of Arg residues or a cell penetrating peptide may improve the cell based activity of the identified macrocyclic peptides. Indeed, incorporation of the cell penetrating peptide penetratin to Se-1 led to moderate activity in cells [146]. This might be further improved by capping peptide amides with *N*-methyl group and by substituting the P1 Gln with a pyrrolidone

moiety [311]. However, the most potent CP from Johansen-Leete *et al*., Se-1, was observed to display weak activity against the human protease Cathepsin L, leaving room for improving the selectivity of the reported CPs against host cysteine proteases [146].

Another potential means of improving cell permeability is via the use of disulfide rich CPs. Many natural disulfide rich peptides have superior stability compared to linear and synthetic macrocyclic peptides [115]. However, with a few exceptions (including the cell penetrating cyclotides [127, 128]; see Chapter 1), cell permeability is still a limitation for many DRPs. This challenge can be exacerbated by the size of the DRP CPs, which can hamper cell permeability in part due to the increased number of hydrogen bond acceptors and donors in their peptide backbones [87, 127]. The disulfide rich peptides derived from plant and animal sources (e.g. venoms), may also trigger adverse immunogenic responses [87]. However, the shuffling of disulfide bonds in DRP scaffolds likely led to the observation of multiple species during HPLC purification. Peaks isolated from MP102 (peak b) and MP103 (peak c) display nanomolar M^{pro} potency; however, their disulfide connection has not yet been determined. *In silico* approaches for molecular grafting of cell-penetrating peptides and identified target sequences onto small disulfide rich scaffolds, in combination with synthetic methods, may improve the potency and potentially cell permeability of the DRP M^{pro} inhibitors.

5.6 Future directions

The mechanisms of cysteine proteases remain incompletely understood, in particular from the perspective of conformational changes during catalysis. The CP inhibitors of M^{pro} present an opportunity to design prosubstrates with a pH/light sensitive linker that can be cleaved to provide a linear substrate to trigger M^{pro} catalysis, which might be used in biophysical studies to inform on its catalytic mechanism. Time-resolved (TR) crystallography using X-ray free electron laser (XFEL) methods provides intense femtosecond long pulses of X-rays that can enable diffraction from micron to micro sized crystals enabling both TR crystallography and limit X-ray induced radiation damage observed in synchrotron studies [339]. Macrocyclic peptides with a light/pH-sensitive linker could be used as a prosubstrate to

co-crystallise with macrocycles and the reaction initiated by pH change/light to crystallographically probe time-resolved catalytic steps of M^{pro} substrate catalysis (Figure 5.7).

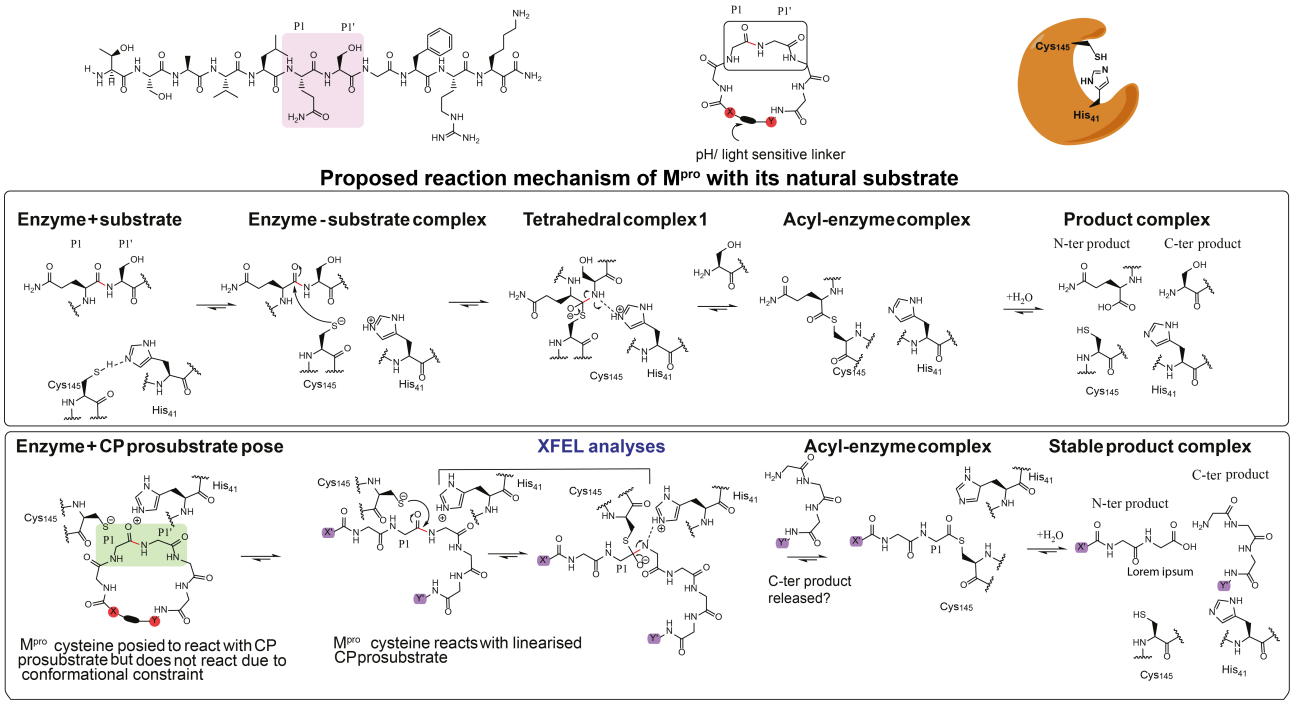


Figure 5.7 | Potential application of CPs as prosubstrates for time-resolved crystallographic studies with M^{pro}.

Improving the cell permeability of the CPs by incorporating disulfide links or cell penetrating peptides, along with selectivity screens against human proteases will be crucial in pursuing the development of the identified CPs as a potential antiviral for the treatment of COVID-19. A library of macrocyclic peptides incorporating cell permeability enhancing peptides or chemical modifications could be designed and assessed for optimising the CPs to work in cells. In addition, the tagged CPs could be used as biological tools for patient-sample imaging in SARS-CoV-2 clinical samples in studies of disease progression and infection, as exemplified by work from Rut *et al.* [340].

6

Non-peptidomimetic inhibitors of M^{pro}

Contents

6.1 Abstract **189**
6.2 Introduction **190**
6.3 Contributions to the manuscripts 5-7 . . . . . **191**
6.4 Future directions **193**

6.1 Abstract

The solid phase extraction coupled to mass spectrometry-based M^{pro} assay (Chapter 2) was employed to screen non-peptidomimetic inhibitors. The investigated compounds included (i) ebsulfur derivatives, (ii) heterocyclic compounds specifically designed for M^{pro} inhibition, and (iii) pentafluorosulfanyl-group containing compounds. The latter two compound classes were provided by the COVID Moonshot Consortium. The results reveal that the substitution patterns of ebsulfur derivatives modulate the stoichiometry of their covalent reaction with M^{pro}. Some of the ebsulfur derivatives display enhanced selectivity while maintaining similar potency. Inhibition studies with >1000 heterocyclic compounds demonstrated that M^{pro} inhibition can be improved by modification of the S1 and S2 subsites binding substituents. Replacement of the 4-methylpyridine group of TRY-UNI-714a760b-6 with a benzotriazole group (ADA-UCB-6c2cb422-1, Figure 6.1) enabled potent inhibition (IC_{50} 0.73 μM). Potency was further optimised by bicyclization of the chlorophenol group by the introduction of an *N*-sulfonylated heterocycle (Figure 6.1). Potent non-covalent inhibitors identified are promising hits to translate into

clinical development. Pentafluorosulfanyl-substituted scaffolds are of increasing interest in medicinal chemistry and were shown to be of use as M^{pro} inhibitions.

6.2 Introduction

Fragment based screening using X-ray crystallography (XChem) led to the identification of lead compounds for M^{pro} inhibition [248]. Subsequent efforts were made to combine or expand (to increase interactions with the neighbouring residues) the fragments to design potent *de novo* inhibitors of M^{pro}. Synthesis of the optimised fragments and molecules submitted from an open science discovery program (`https://covid.postera.ai/submit`), which included a collection of varied scaffold e.g. pentafluorosulfanyl-substituted scaffolds, was performed by Enamine for the COVID Moonshot Consortium. I screened the compounds via the SPE-MS coupled MS assays. The data were then released on a web based platform (`https://covid.postera.ai/covid/activity_data`) for facilitating the SAR of the submitted compounds and openly informing on the design of improved M^{pro} inhibitors in a strictly IP-free method.

Whilst the compounds provided by the Moonshot consortium were primarily non-covalent non-peptidomimetic molecules, potent covalent non-peptidomimetic inhibitors including S-containing compounds PX-12, disulfiram and Se-containing ebselen were reported [28]. In Chapter 2, protein observed mass spectrometric experiments revealed that ebselen covalently reacts with all cysteines in M^{pro}. Work was carried out on ebsulfur compounds with the aim of improving selectivity for the nucleophilic cysteine of M^{pro}.

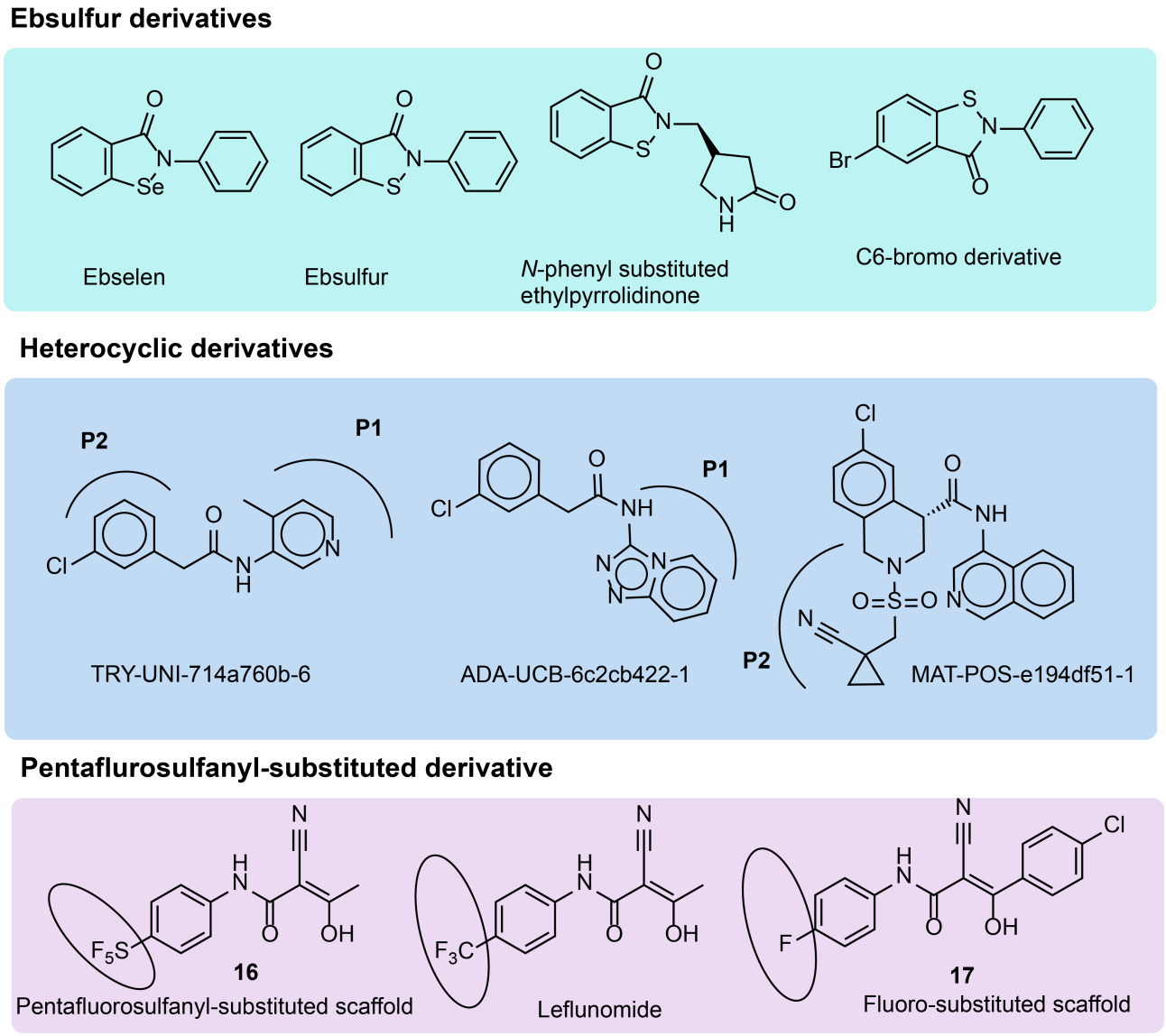


Figure 6.1 | Structures of the ebsulfur derivatives, heterocyclic compounds, and a pentafluorosulfanyl-substituted scaffold relevant to the text of this chapter. See corresponding manuscripts (5, 6, and 7) for full details. In the case of the heterocyclic compounds, P1 and P2 elements of the compounds interacting with the S1 and S2 subsites are highlighted. Note also the pentafluorosulfanyl group, which is highlighted, and which was pursued in one set of M^{pro} inhibitors (manuscript 5).

6.3 Contributions to the manuscripts 5-7

Much of the work outlined in this chapter involved work performed in collaboration with the synthetic chemists Timothy Suits and Siegfried Thun-Hohenstein (Department of Chemistry, University of Oxford) and the COVID Moonshot consortium. Following on the collaboration on ebselen/ebsulfur like compounds with Ldt_{Mt2} with Timothy Suits, I with Timothy Suits and Siegfried Thun-Hohenstein

(who carried out the synthesis) worked on the inhibition of M^{pro} by ebsulfur derivatives as described in Manuscript 5. The COVID Moonshot consortium consisted of a large cohort of domestic and global scientists in an open science-based effort to develop *de novo* antivirals against COVID-19 (`https://dndi.org/research-development/portfolio/covid-moonshot/`). A part of this work involving the identification of pentafluorosulfanyl compounds as M^{pro} inhibitors was performed in collaboration with Prof. John Spencer (University Of Sussex).

6.3.1 Manuscript 5: Structure-Activity Studies Reveal Scope for Optimisation of Ebselen-Type Inhibitors of SARS-CoV-2 Main Protease

For the described ebsulfur work, I carried out the following:

- Screening of ebsulfur compounds, via the SPE-MS M^{pro} assay.
- Optimisation of a FRET based assay as an orthogonal M^{pro} inhibition assay.
- Protein observed MS assays to investigate the mode of inhibition of ebsulfur compounds.
- Studies on the selectivity of reactions of inhibitors with M^{pro} were also conducted by blocking the active site cysteine with the selective covalent inhibitor TPCK, followed by incubation with inhibitors and analyses by protein observed MS assays.

6.3.2 Manuscripts 6 and 7 describing the work carried out as part of the COVID Moonshot Consortium

As a contributing member of the COVID Moonshot work, I performed the following work for the listed manuscripts 6 and 7:

- Screening of >1000 *de novo* compounds, including pentafluorosulfanyl-substituted compounds, via the SPE-MS M^{pro} assay.
- Protein observed MS assay to determine the mode of inhibition of selected compounds.

- Optimisation of the 37mer substrate based peptide assay with low resolution MS at the Target Drug Discovery Institute, to enable the application of the SPE-Mass Spectrometry Assay by Professor Frank Von Delft's group at Nuffield Department of Medicine, University of Oxford.

6.4 Future directions

The SPE-MS based assays were useful for the identification of new classes of M^{pro} inhibitors providing multiple lines of future investigation, both with isolated M^{pro} and in cellular and *in vivo* studies. Compounds from the Moonshot consortium are being pursued toward clinical development. In the case of the ebsulfur compounds, further work could be done to improve their selectivity with respect to reaction with Cys145 of M^{pro}. As with other compounds covalently reacting with M^{pro}, proteomic type-MS could be used to validate the modification of Cys145. Given that ebselen/ebsulfur compounds are unlikely to be completely selective (though ebselen is of clinical interest [341]), cellular studies will also be important to determine if the improved selectivity is reflected by reduced cytotoxicity. The work also highlighted the potential of pentafluorosulfanyl group in medicinal chemistry. It should be noted that some of the pentafluorosulfanyl-substituted compounds possess a nitrile group (manuscript 7), which may react covalently with M^{pro}. This is seen for the nitrile moiety in nirmatrelvir [38]. This possibility could be investigated by biophysical analyses including MS and crystallography.

7

Mechanistic insights into interactions of Ldts with β-lactam antibiotics

Contents

7.1 Abstract **194**
7.2 Introduction **195**
7.3 Results and Discussion **198**
7.4 Conclusions **237**
7.5 Future directions **239**
7.6 Methods **240**
7.7 Appendix **247**

7.1 Abstract

The L,D-transpeptidases from *Mycobacterium tuberculosis* are nucleophilic cysteine enzymes that catalyse the formation of 3-3 crosslinks between tetrapeptides in the peptidoglycan layer. This chapter reports evidence that Ldt_{Mt2} reacts with ampicillin to form a diketopiperazine product via 5-exo trig cyclisation of the amino group of the ampicillin side chain with the thioester formed by the reaction of the ampicillin β-lactam with the nucleophilic cysteine. This observation apparently correlates with the differences in reactivities of ester/thioester complexes of nucleophilic enzymes. With the carbapenem class of antibiotics, the corresponding acyl-enzyme complexes with Ldt_{Mt2} were observed to react reversibly or to deacylate through the formation of lactone products. Formation of lactone products and the reversible reaction of carbapenems was more evident with the L, D-transpeptidase

from *Enterococcus faecium* than with Ldt_{Mt2}. Similarly to the carbapenems, evidence is presented that acyl-enzyme complexes derived from cephalosporins can undergo recyclisation, as monitored by ^1H NMR spectroscopic studies. Faropenem was observed to undergo rapid C5-C6 fragmentation following acylation of the Ldt_{Mt2} nucleophilic cysteine. This complex can be deacylated by a reaction with D-amino acids; the rate of reaction remains dependent on the amino acid. The deacylation of Ldt_{Mt2} complexes formed by penicillins that undergo C5-C6 fragmentation is also promoted by D-amino acids. Deacylation of Ldt_{Mt2} and related enzymes mediated by D-amino acids may be of biological interest.

7.2 Introduction

Enzymes that employ the reactions of nucleophilic serine, cysteine, and aspartate residues with ester and amide substrates are of immense importance in biology (Chapter 1). The acyl-enzyme complexes (AECs) formed through these reactions are commonly deacylated by nucleophilic substrates, i.e. water or amines to give hydrolysed or transamidated products, respectively. The covalent reaction between the nucleophilic residues of these enzymes with inhibitors is medicinally important and contributes to the activity of many important pharmaceuticals, including the penicillins, carbapenems, penems, and related β-lactam antibiotics.

β-Lactams react with the nucleophilic serine of their D,D-transpeptidase (i.e., PBP) targets to form a relatively stable AEC. The serine β-lactamases, which likely evolved from transpeptidases [342,343], efficiently hydrolyse the acyl-enzyme complex to (normally) give β-amino acids. In some cases, non-hydrolytic fragmentation of the intact acyl-enzyme complex has been reported, e.g. for some penicillins and cephalosporins. The AECs derived from certain cephalosporins, such as cefotaxime, sometimes undergo the elimination of a C3′ leaving group (e.g. an acetoxy group), which can increase the stability of the AEC, particularly when C3′ elimination is faster than hydrolysis of the intact acyl-enzyme complex [344]. C5-C6 fragmentation of the AECs derived from PBPs or β-lactamases with some penicillins has been reported, although this has been observed to be a slow process [345]. In the case of β-lactam-containing β-lactamase scaffold inhibitors, such as clavulanic acid, the resulting AECs can undergo further reactions (including fragmentations) which compete with hydrolysis [346,347].

Some β-lactamases can react with carbapenems to give both β-lactone and hydrolysis products, with the β-lactones undergoing reversible reaction with the β-lactamase. With the penem subclass of β-lactams, β-lactamases can react to form multiple hydrolysed products, as shown by studies with faropenem (Chapter 8, manuscript 7 [348]). However, to date, C5-C6 fragmentation of faropenem has only been reported with the Ldts. The so-formed 3-hydroxy propionyl AEC apparently allows the β-hairpin loop at the active of Ldt_{Mt2} to close tightly (as compared to intact carbapenems), possibly contributing to the slow hydrolysis of this AEC [178].

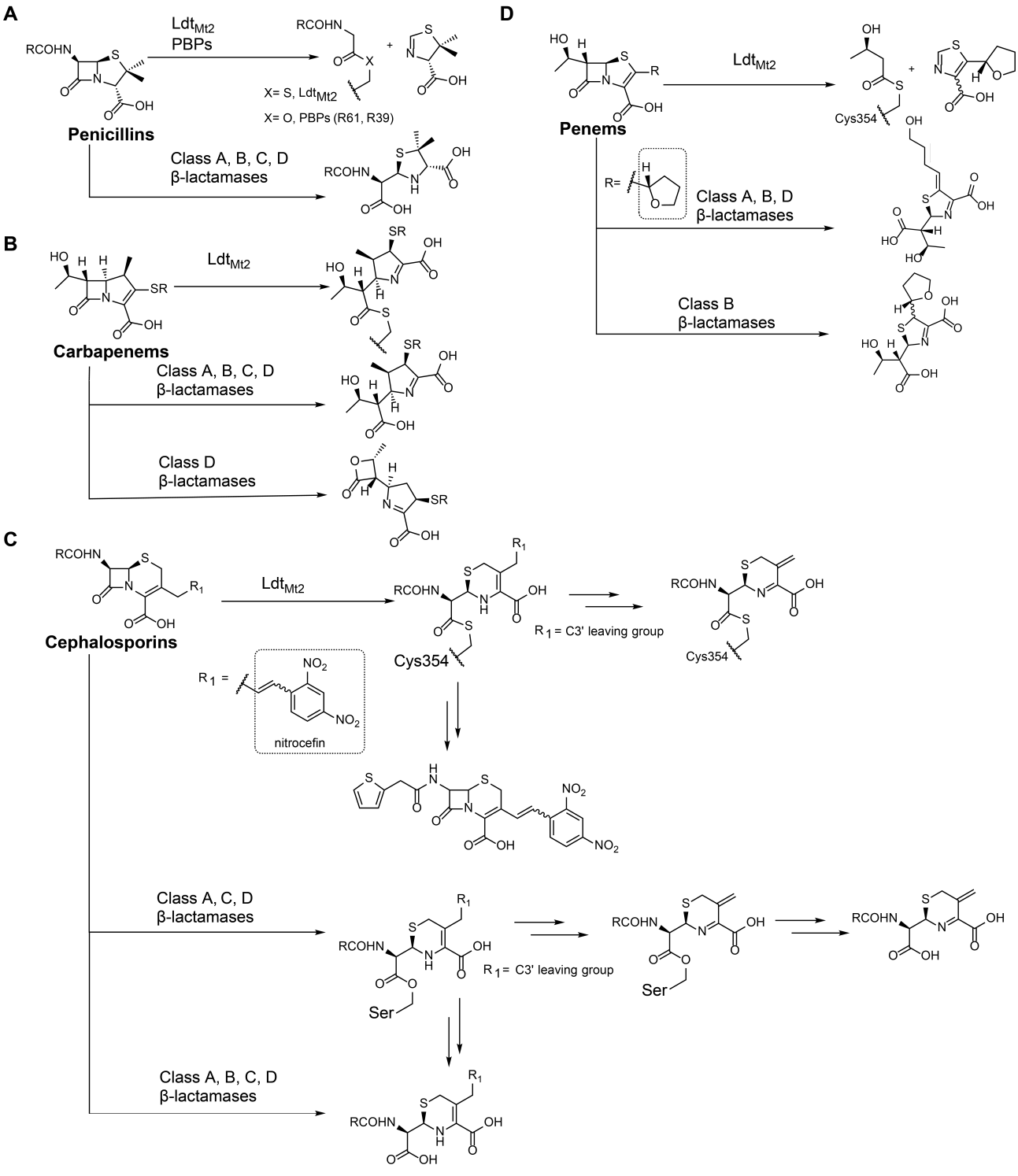


Figure 7.1 | Proposed reactions of major subclasses of β-lactams with L,D-transpeptidases and β-lactamases. Hydrolysis products or acyl-enzyme complexes are observed upon reaction with **A.** penicillins, **B.** penems, **C.** carbapenems, and **D.** cephalosporins. Fragmentation of AECs has been observed with penicillins, penems, and cephalosporins, and deacylation of some carbapenems occurs via a rearrangement.

The inhibition of nucleophilic cysteine enzymes by acylation has been less well-explored than the analogous serine enzymes, with recent work on the cysteine protease M^{pro} being a rare example

(Chapter 2-6). The L, D-transpeptidases are targets for the treatment of tuberculosis. Of the five Ldt_{Mt} paralogs in *Mycobacterium tuberculosis*, studies with mouse models have shown that only Ldt_{Mt2} is essential for virulence, while gene expression and protein profiling reveals that Ldt_{Mt1} is overexpressed during non-replicative stages of *M. tuberculosis* persistence [175]. The *in vivo* functions of Ldt_{Mt3}-Ldt_{Mt5} have not been well characterised, but they may have roles in (i) anchoring cell-wall tetrapeptides by covalently crosslinking the carboxyl group of *meso*-diaminopimelic acid (DAP) to the amino group of the C-terminal lysine in Braun's lipoprotein, e.g. as seen with LdtB in *E. coli* [349], (ii) catalysing the hydrolysis of the lipid anchor crosslink (LdtF, which has been renamed as *meso*-diaminopimelic acid protein amidase A (DpaA)) [350], or (iii) incorporation of non-canonical D-amino acids, as shown by work in *Vibrio cholera* (LdtB) and *Clostridium difficile* (LdtC) [351].

Unlike the 4→3 crosslinking D,D-transpeptidases, also known as PBPs, the L,D-transpeptidases use a Cys-His catalytic dyad, the nucleophilic cysteine of which can be acylated by β-lactam antibiotics. In a mouse model of tuberculosis [203], certain β-lactams, in particular, carbapenems have shown potential as anti-tuberculosis therapeutics. This is encouraging as tuberculosis is becoming resistant to existing therapies, in part, due to the lack of compliance to the current 6 months of antibiotic regimen and improper use of antibiotics. Here I report mechanistic studies on the interactions of the L,D-transpeptidase active site cysteine residues with penicillin, cephalosporin, carbapenem, and penem antibiotics, and identify novel products formed by the Ldt_{Mt}-catalysed degradation of β-lactams.

7.3 Results and Discussion

7.3.1 Protein production and purification

To study the interaction of Ldts with different β-lactam antibiotics, genes encoding for recombinant $Ldt_{Mt1\text{-}5}$ and Ldt_{fm} were cloned into the pET28 vector and expressed in *Escherichia coli* BL21(DE3), by modifying reported methods [228]. The N-terminally hexaHis-tagged Ldts were purified by IMAC, followed by size-exclusion chromatography. SDS-PAGE analyses of the purified material indicated that purified $Ldt_{Mt1\text{-}3}$ co-purify with low levels of endogenous proteins. Several adducts on all Ldts

were observed by MS, including a +54 Da adduct (consistent with modification of arginine with methylglyoxal forming a hydroimidazolone [352,353], a +97 Da adduct (unidentified), and a +178 Da adduct (consistent with phosphogluconoylation of the hexaHis-tag [354]). Purified $LdtMt_{1\text{-}3}$ and $LdtMt_5$ appear stable in solution, as indicated by a single peak observed by SDS-PAGE analysis. However, multiple poorly separated bands were observed for Ldt_{Mt4} suggesting possible degradation during size exclusion chromatography. Purification of Ldt_{Mt4} was optimised in sodium phosphate buffer, which was found to improve stability (see methods section). As the yields of Ldt_{Mt4} and Ldt_{Mt5} were low, studies requiring large quantities of proteins, focused on $Ldt_{Mt1\text{-}3}$ and Ldt_{fm}.

Contrary to reports describing the loss of the N-terminal methionine for hexaHis-tagged Ldts, the observed masses [304] for all purified proteins implied the presence of an N-terminal methionine [174]. The molecular weights of Δ1-28Ldt_{Mt1}, Δ1-55Ldt_{Mt2}, Δ32-271Ldt_{Mt3}, Ldt_{Mt4} (full length) and Δ91-451Ldt_{Mt5}, including the N-terminal hexaHis-tags, as observed by MS were 26536 Da (calc'd: 26538 Da), 40440 Da (calc'd: 40444 Da), 28438 Da (calc: 28439 Da), 41457 Da (calc'd: 41456 Da), 45477 Da (calc'd: 45418 Da), respectively (Figure 7.2). The observed masses of the purified Ldts correspond to their calculated masses, apart from Ldt_{Mt5} which is 60 Da greater than expected. MS experiments with the reducing agent DTT and reaction with two ebselen species showed that this post-translational modification likely does not occur at the active site Cys360 residue (Appendix Figure 7.28). Note that expression of the Δ140-406Ldt_{Mt2}, Δ18-366Ldt_{Mt4} constructs was not successful. An L,D-transpeptidase from *Enterococcus faecium* (Δ215-466Ldt_{fm}) was purified using the same methodology, its calculated mass was consistent with the mass observed by SPE-MS, that is 30555 Da ± 6 Da (calc'd: 30556 Da).

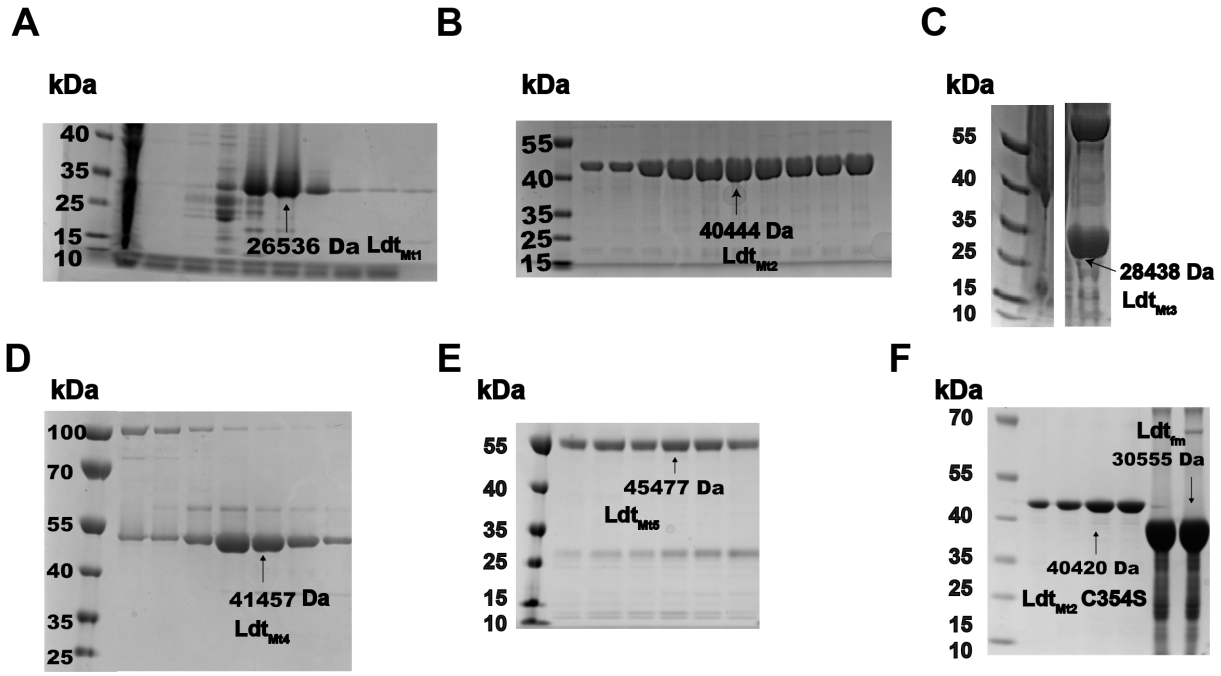


Figure 7.2 | SDS-PAGE gel analysis of L,D-transpeptidases produced by over-expression in *E. coli* BL21(DE3) cells. Protein bands correspond to HisTrap and size exclusion column purified fractions of **B.** Ldt_{Mt2}, **D.** Ldt_{Mt4}, **E.** Ldt_{Mt5}, **F.** Ldt_{fm} and Ldt_{Mt2} C354S. Protein bands of **A.** Ldt_{Mt1} and **C.** Ldt_{Mt3} correspond to HisTrap column only purified fractions. Conditions: 10-12% Bis-Tris Nu-PAGE Gradient gel, 160 mV, 50 minutes.

7.3.2 Interaction of mycobacterial and enterococcal L,D-transpeptidases with representative β-lactams

The interactions of all six purified Ldts with four major subclasses of β-lactams were investigated by incubating them with an excess of representative β-lactam antibiotics: ertapenem and meropenem (carbapenems), cefditoren (a cephalosporin), faropenem (a penem), or oxacillin (a penicillin). SPE-MS analyses indicated the formation of covalent adducts with most of the tested β-lactams. In some cases, stable acyl-enzyme complexes and/or C5-C6 fragmentation of the acyl-enzyme complexes were observed.

7.3.2.1 Interaction of Ldts with faropenem

Previous work with faropenem and Ldt_{Mt}s has shown evidence for C5-C6 fragmentation to give a +86 Da mass shift corresponding to 3-hydroxybutyryl-bound fragment of faropenem [178,228]. In my work, an analogous +86 Da mass shift was observed with $Ldt_{Mt1\text{-}4}$ (Figure7.3) and Ldt_{fm} (Figure 7.7). However, in the case of Ldt_{Mt5}, there was no evidence for a +86 Da mass shift upon reaction with

faropenem. By contrast, there was evidence for a partial reaction of faropenem without fragmentation. However, it cannot be certain that this is due to the reaction of the active site cysteine. Overall, the results imply that Ldt_{Mt5} is less reactive with faropenem than other tested Ldt_{Mt}s and Ldt_{fm}.

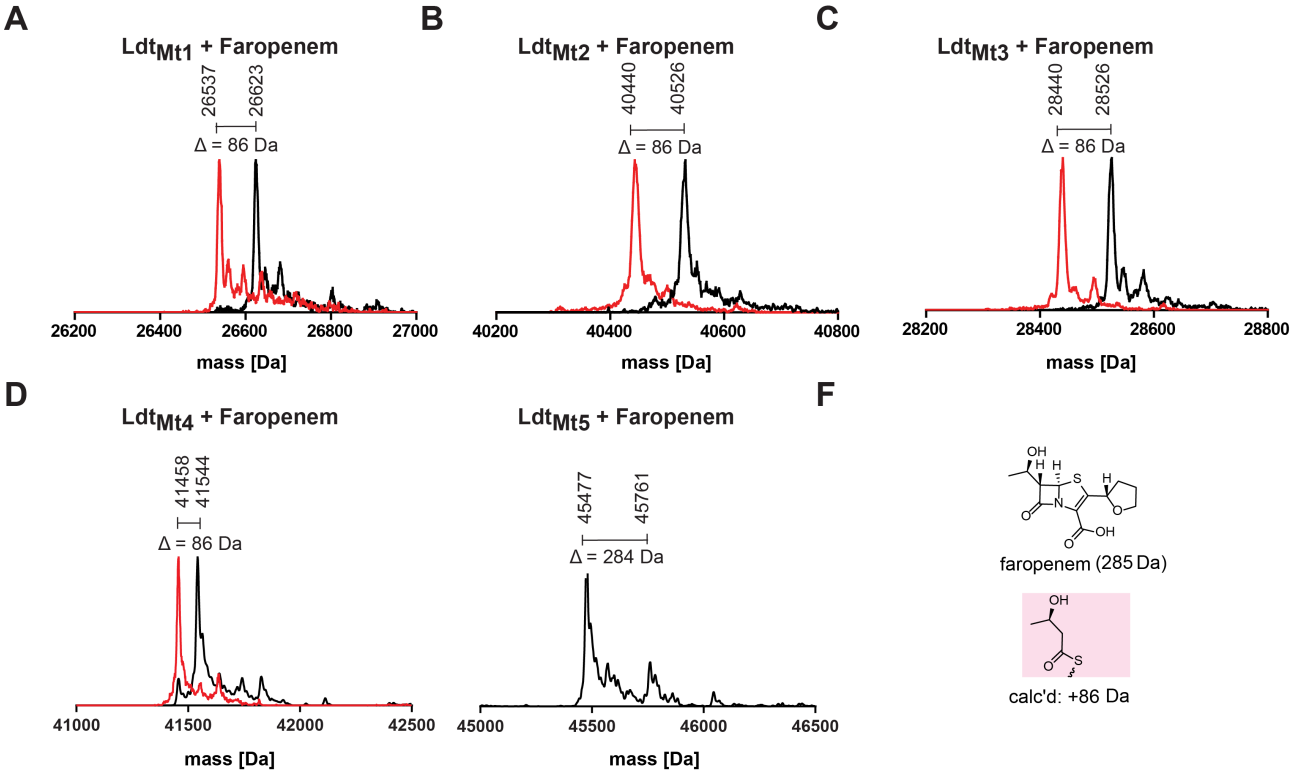


Figure 7.3 | Mass spectrometric analyses of the interactions of penems and Ldt_{Mt}s. Purified **A.** Ldt_{Mt1} (1 µM), **B.** Ldt_{Mt2} (2 µM), **C.** Ldt_{Mt3} (2 µM), **D.** Ldt_{Mt4} (0.5 µM) and **E.** Ldt_{Mt5} (1 µM) were incubated with an excess of faropenem (100 µM) for 10-14 min. C5-C6 fragmentation of the acyl-enzyme complex was observed with faropenem (+86 Da adduct) and Ldt_{Mt1}, Ldt_{Mt2}, Ldt_{Mt3} and Ldt_{Mt4} but not Ldt_{Mt5}. Unexpectedly, a minor peak corresponding to an intact acyl-enzyme complex was observed with Ldt_{Mt5}. All samples were prepared in 50 mM sodium phosphate, pH 7.5.

7.3.2.2 Interaction of Ldts with representative carbapenems

Apparently complete acylation was observed upon the reaction of ertapenem or meropenem and all of the tested Ldts (Figures 7.4 and 7.7A), with the exception of Ldt_{Mt5}. Although no *in vitro* MS evidence has been reported for Ldt_{Mt5} acylation by carbapenems, crystallographic studies have shown acylation of Ldt_{Mt5} with meropenem [182]. Weak peaks corresponding to a +32 Da mass shift compared to the protein only peak were observed with Ldt_{Mt2} (Figure 7.4B), which may correspond to oxidation of the active site Cys to a sulfinic acid [355]; the +32 Da modified enzyme was apparently not

modified further by reaction with ertapenem, suggesting that the +32 Da modification may occur at the active site. A 44 Da mass loss was observed with some of the Ldt-carbapenem complexes, possibly representing an artefact resulting from mass spectrometric conditions, or indicating decarboxylation for Ldt_{Mt2} using MS studies or the loss of the hydroxyethyl side chain as acetaldehyde for BlaC on the basis of MS and crystallographic studies [178, 233].

The covalent carbapenem complexes with Ldt_{Mt}s were apparently stable, and there was no evidence for the C5-C6 fragmentation observed for faropenem (Figure 7.4) Fragmentation of the carbapenems may be unfavourable due to the lack of a heteroatom at position 1, or may in part reflect the lack of aromaticity of the potential product formed upon C5-C6 fragmentation.

The protein adducts observed following purification (e.g., resulting from phosphogluconoylation) were also acylated by the β-lactams, confirming that these modifications occur away from the active site (Figures 7.4, 7.7A). In the case of Ldt_{fm}, a +475 Da adduct was observed, confirming the reaction with carbapenem to form a stable AEC (Figure 7.7).

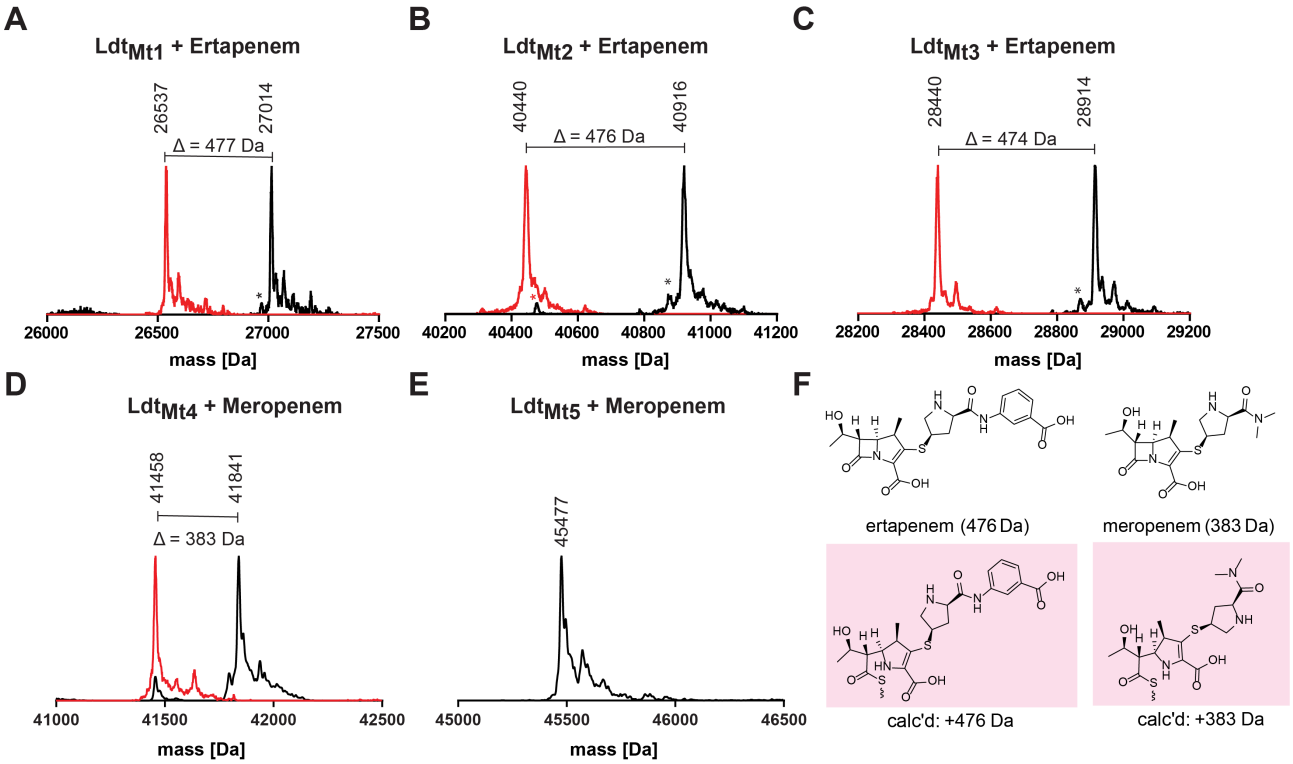


Figure 7.4 | Mass spectrometric analyses of the interactions of carbapenem and Ldt$_{Mt}$s. Purified **A.** Ldt$_{Mt1}$ (1 µM), **B.** Ldt$_{Mt2}$ (2 µM), **C.** Ldt$_{Mt3}$ (2 µM), **D.** Ldt$_{Mt4}$ (0.5 µM) and **E.** Ldt$_{Mt5}$ (1 µM) were incubated with an excess of meropenem or ertapenem (100 µM) for 10-14 min. Full acylation was observed with Ldt$_{Mt1}$, Ldt$_{Mt2}$, Ldt$_{Mt3}$ and Ldt$_{Mt4}$; no evidence for C5-C6 fragmentation of carbapenem was observed upon reaction with the Ldt$_{Mt}$s. Acylation of Ldt$_{Mt5}$ was not observed upon the addition of meropenem. Low level of a +32 Da adduct (red *) was observed with Ldt$_{Mt2}$ treated with ertapenem, potentially representing oxidation of Cys to sulfinic acid, or a methanol derived artifact associated with the RapidFire mass spectrometer. The asterisk represents the low level of a 44 Da mass loss from the Ldt$_{Mt}$-carbapenem complex. All samples were prepared in 50 mM sodium phosphate, pH 7.5.

7.3.2.3 Interaction of Ldts with a representative cephalosporin

By contrast with the complete modification of all the Ldts (except Ldt$_{Mt5}$) with carbapenems, the level of modification by cefditoren was not uniform across the Ldts. Full acylation of Ldt$_{Mt1}$, Ldt$_{Mt2}$ and Ldt$_{fm}$ (Figure 7.7C) was observed with a mass shift of ∼506 Da, but only partial modification of Ldt$_{Mt3}$ and Ldt$_{Mt4}$ (Figure 7.5) was observed, under the standard conditions. In addition to the varying active site structures of the Ldts, the differences in reaction with cefditoren may be explained, in part, by different binding affinities and hydrolytic efficiencies of the Ldts [230,356,357].

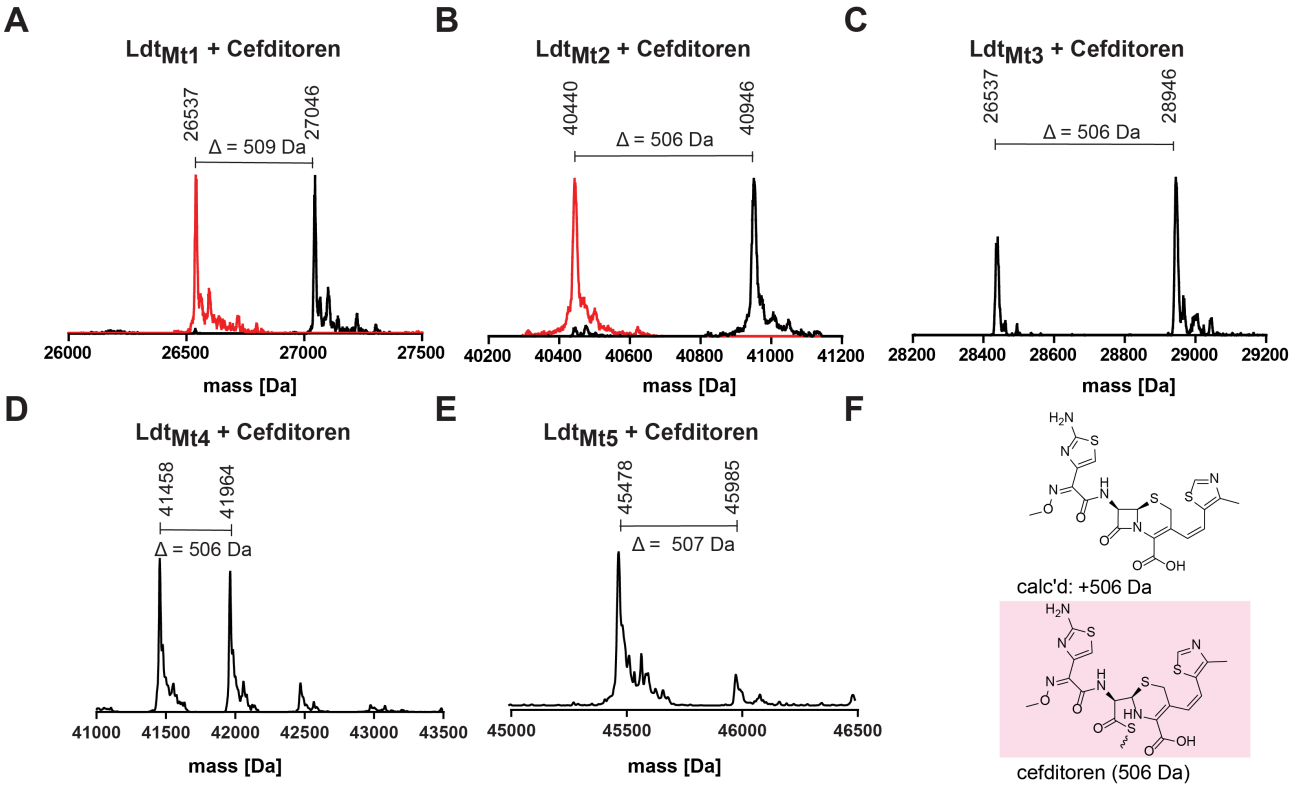


Figure 7.5 | Mass spectrometric analyses of the interactions of cefditoren and Ldt$_{Mt}$s. Purified **A.** Ldt$_{Mt1}$ (1 μM), **B.** Ldt$_{Mt2}$ (2 μM), **C.** Ldt$_{Mt3}$ (2 μM), **D.** Ldt$_{Mt4}$ (0.5 μM) and **E.** Ldt$_{Mt5}$ (1 μM) were incubated with an excess of cefditoren (100 μM) for 10-14 min. Complete acylation was observed with Ldt$_{Mt1}$, Ldt$_{Mt2}$, but not Ldt$_{Mt3}$ and Ldt$_{Mt4}$; no evidence for C5-C6 fragmentation of cefditoren was observed upon reaction with Ldt$_{Mt}$s. Acylation of Ldt$_{Mt5}$ was not observed upon the addition of cefditoren. All samples were prepared in 50 mM sodium phosphate, pH 7.5.

7.3.2.4 Interaction of Ldts with representative penicillin

The acylation profile of oxacillin was also different across the Ldts. Complete modification occurred with Ldt$_{Mt1\text{-}3}$, but only partial acylation of Ldt$_{Mt4}$ and Ldt$_{fm}$ was observed, under the standard conditions (Figure 7.6). The extent of apparent C5-C6 fragmentation of oxacillin varied, with extensive fragmentation only occurring with Ldt$_{Mt2}$. No evidence for C5-C6 fragmentation of the oxacillin-derived complexes was observed with Ldt$_{Mt1}$, Ldt$_{Mt3}$, or Ldt$_{Mt5}$. Note that only low levels of partial reaction were observed upon the addition of oxacillin to Ldt$_{Mt5}$.

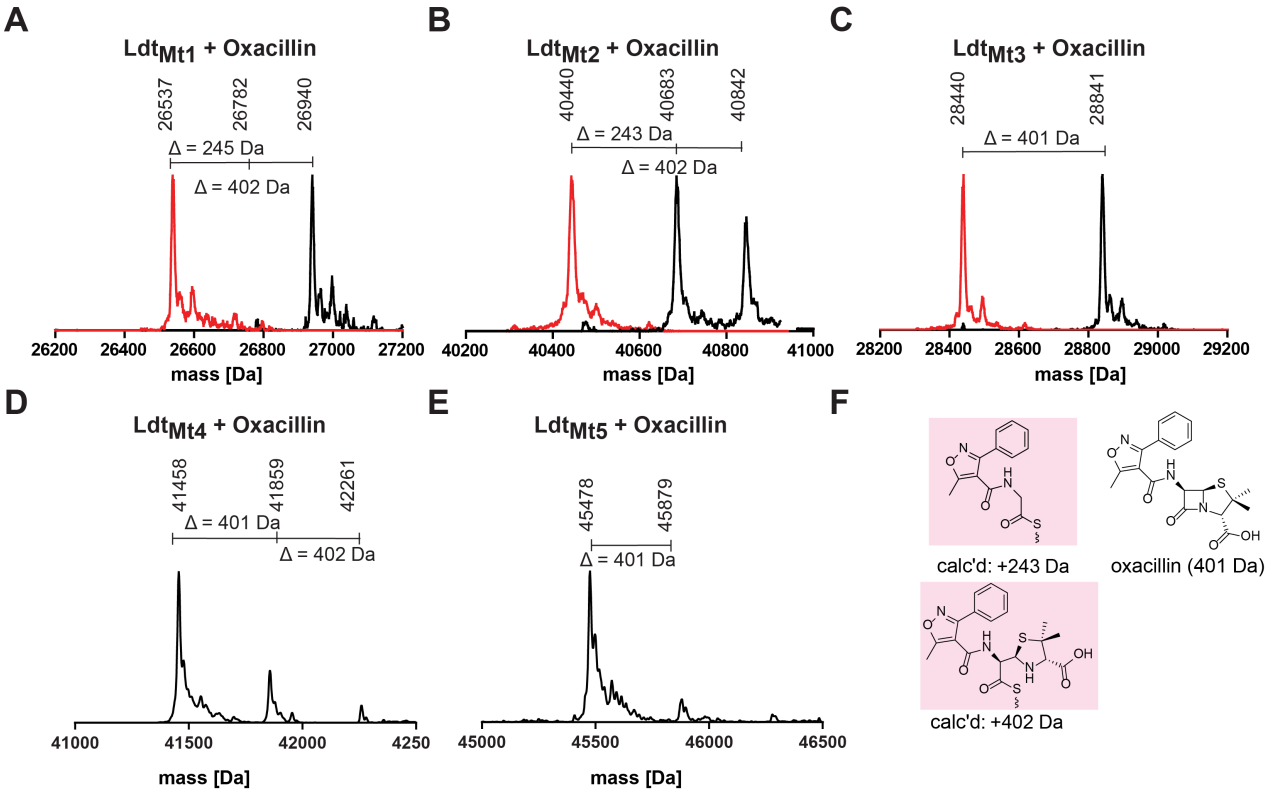


Figure 7.6 | Mass spectrometric analyses of the interactions of penicillins and Ldt$_{Mt}$s. Purified **A.** Ldt$_{Mt1}$ (1 μM), **B.** Ldt$_{Mt2}$ (2 μM), **C.** Ldt$_{Mt3}$ (2 μM), **D.** Ldt$_{Mt4}$ (0.5 μM) and **E.** Ldt$_{Mt5}$ (1 μM) were incubated with an excess of oxacillin (100 μM) for 2 h. Full acylation was observed with Ldt$_{Mt1}$, Ldt$_{Mt2}$ (+245 ± 2 Da adduct) and Ldt$_{Mt3}$, but not Ldt$_{Mt4}$; C5-C6 fragmentation of the acyl-enzyme complex (i.e., a +243 Da adduct) was observed with oxacillin and Ldt$_{Mt2}$, but not the other Ldt$_{Mt}$s. No major peaks corresponding to an acyl-enzyme complex were observed with oxacillin upon incubation with Ldt$_{Mt5}$. All samples were prepared in 50 mM sodium phosphate, pH 7.5. The proposed adduct structures highlighted in pink are consistent with the mass adducts observed with Ldts with an error of ±3 Da. Red spectra show Ldt$_{Mt}$ alone and are superimposed with black spectra showing Ldt$_{Mt}$s incubated with the β-lactams.

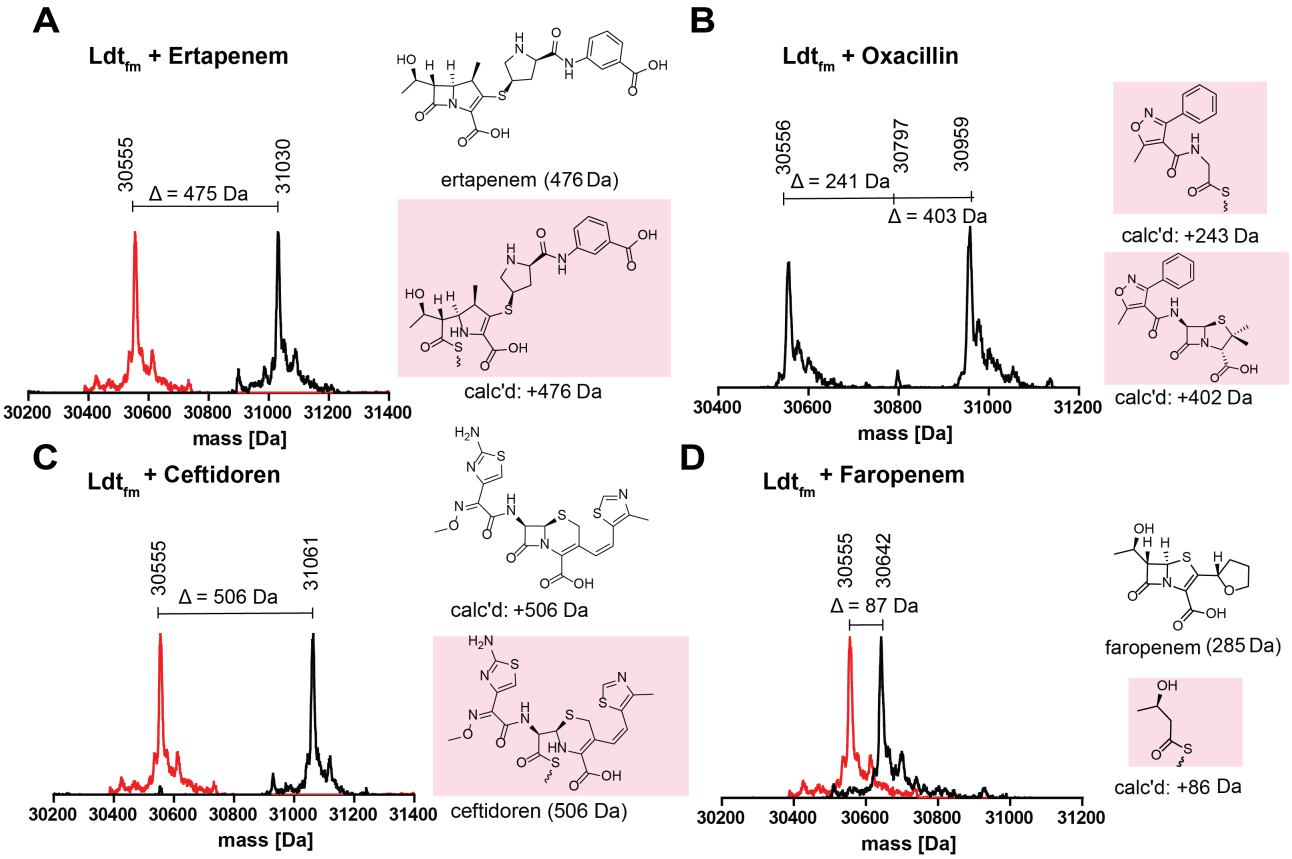


Figure 7.7 | Mass spectrometric analyses of the interactions between representative β-lactams and Ldt_{fm}. Purified Ldt_{fm} (2 µM) was incubated with a 50-fold excess of **A.** ertapenem for 10 min, B) oxacillin for 2 h (using 1 µM Ldt_{fm} and a 100-fold excess of oxacillin) **C.** cefditoren for 10 min, and D) faropenem for 10 min. C5-C6 fragmentation of the acyl-enzyme complexes was observed with oxacillin (+243 ± 2 Da adduct) and faropenem (+86 Da adduct). The ertapenem and cefditoren AEC apparently remained intact following acylation. Along with the evidence for C5-C6 fragmentation of oxacillin, evidence for intact oxacillin bound Ldt_{fm} was also observed with a mass shift of 401 Da. The proposed structures highlighted in pink are consistent with the mass adducts observed with Ldt_{fm} with an error of ±2 Da. Minimal unmodified Ldt_{fm} was observed when treated with excess cefditoren. All samples were prepared in 50 mM sodium phosphate, pH 7.5. Red spectra represent Ldt_{fm} (2 µM) alone and are superimposed on black spectra representing Ldt_{fm} incubated with β-lactam antibiotics.

7.3.2.5 The nucleophilic serine variant of Ldt_{Mt2} does not react with β-lactams

PBPs employ a nucleophilic serine to catalyse transpeptidation reactions, and this process can be inhibited following reaction with β-lactams in a similar manner as with the Ldts, except that an ester intermediate is involved. However, C5-C6 fragmentation of acyl-enzyme complexes derived from PBPs and β-lactams have not been widely studied [345, 358, 359]. The nucleophilic Cys in Ldt_{Mt2} was therefore substituted with a serine to probe its interaction and role in the fragmentation of

acyl-enzyme complexes with β-lactams. Site directed mutagenesis was performed on the Δ1-55Ldt_{Mt2} construct to generate the Δ1-55Ldt_{Mt2} C354S variant. A little surprisingly, no evidence of covalent modification of the Ldt_{Mt2} C354S variant was obtained, suggesting the active site Cys nucleophile is critical for the reaction of Ldt_{Mt2} with β-lactams (Figure 7.8).

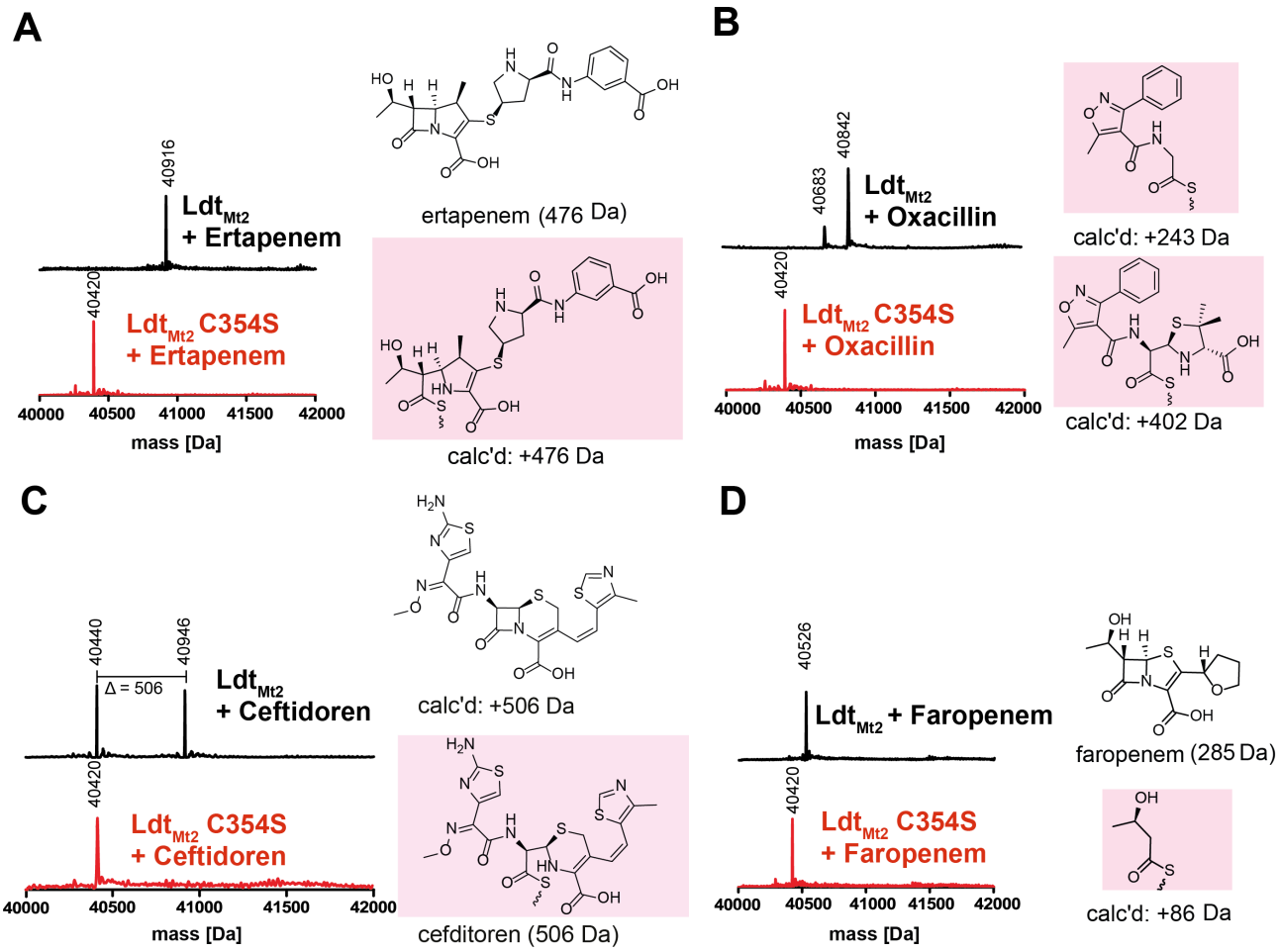


Figure 7.8 | Mass spectrometric analyses of interactions of representative β-lactams with Ldt_{Mt2} C354S. Mass spectra showing the absence of acylation of Ldt_{Mt2} C354S (2.2 μM), in comparison to wild-type Ldt_{Mt2} (1.64 μM), upon treatment with **A.** ertapenem (100 μM) for 10 min (wild-type) or 50 min (C354S variant), **B.** oxacillin (100 μM) for 50 min (wild-type) or 70 min (C354S variant), **C.** cefditoren (100 μM) for 5 min (both enzymes), and **D.** faropenem (100 μM) for 20 min (wild-type) or 50 min (C354S variant). Samples were prepared in 50 mM sodium phosphate, pH 7.5. The proposed structures of the acyl-enzyme complexes are shown with a pink background. Wild-type Ldt_{Mt2} in the absence of β-lactam was observed at 40440 Da, while Ldt_{Mt2} C354S in the absence of β-lactam was observed at 40420 Da. Black spectra: wild-type Ldt_{Mt2} with the indicated β-lactam; red spectra: Ldt_{Mt2} C354S with the indicated β-lactam. This figure was prepared for and adapted from published work [228].

7.3.3 Mechanistic insights into the interaction of penicillins with Ldts

To further investigate the C5-C6 fragmentation of Ldt_{Mt2}-oxacillin complexes, time course MS investigations were carried out with oxacillin and ampicillin and three Ldts (Ldt_{Mt1-3}). Evidence for C5-C6 fragmentation of the acyl-enzyme complex was accrued with both penicillins with Ldt_{Mt1} and Ldt_{Mt2}, but not for Ldt_{Mt3}. Strong evidence for C5-C6 fragmentation of oxacillin adduct was observed in the case of the Ldt_{Mt2}-oxacillin complex. In order to characterise the released C5-C6 fragmentation product and identify the potential hydrolysis products, NMR spectroscopic studies were carried out.

Figure 7.9 | C5-C6 fragmentation of oxacillin- and ampicillin-derived Ldt_{Mt1-3} acyl-enzyme complexes over time. A 50-fold excess of oxacillin was incubated with (2 μM) **A.** Ldt_{Mt1}, **B.** Ldt_{Mt2}, and **C.** Ldt_{Mt3}, and a 50-fold excess of ampicillin was incubated with (2 μM) **D.** Ldt_{Mt1}, **E.** Ldt_{Mt2}, and **F.** Ldt_{Mt3}. C5-C6 fragmentation of both penicillins was clearly observed with Ldt_{Mt2}, but not with Ldt_{Mt3}. A weak peak corresponding to the C5-C6 fragmentation product formed by Ldt_{Mt1} was observed with oxacillin but not with ampicillin. All samples were prepared in 50 mM sodium phosphate, pH 7.5. Red spectra represent control Ldt_{Mt2} (2 μM) alone, and black spectra represent the mixture of Ldt_{Mt2} with β-lactam penicillins. An unassigned + 33 Da adduct was observed with Ldt_{Mt2} when treated with oxacillin (mass: 40476 Da).

7.3.3.1 Application of NMR spectroscopy to monitor C5-C6 fragmentation of penicillins upon reaction with Ldts

NMR spectroscopy has been used to investigate the reaction of ampicillin and oxacillin with Ldt_{Mt2} [228]. This work reported that, following the addition of these penicillins to N-terminally hexaHis-tagged Ldt_{Mt2}, in addition to the expected hydrolysis products **5** and low level of C5-C6 fragmentation product **2**, weak signals corresponding to an uncharacterised product **3** were observed [228].

Ampicillin turnover by Ldt_{Mt2} using enzyme prepared via a rigorous purification protocol

After the publication of this work [228], subsequent batches of Ldt_{Mt2} protein were prepared using NaOH washed gel filtration columns; these batches manifested altered product profiles upon ampicillin turnover. Peaks corresponding to a new product (subsequently shown to be **3**) were observed to predominate, with the C5-C6 fragmentation product **2** being the next most abundant product (Appendix Figure 7.29). No significant peaks corresponding to the hydrolysis product **5/5*** were observed

with the Ldt_{Mt2} purified using the NaOH column treatment for up to 55 h (Appendix Figure 7.30), suggesting that formation of **5** may result from a contaminant.

To test this hypothesis, experiments were performed with Ldt_{Mt2} which was purified according to the previous protocol without the use NaOH wash step for the gel filtration column. This batch of Ldt_{Mt2} was treated with the TEV protease to cleave the N-terminal hexaHis-tag and subjected to an additional purification step by IMAC compared to the N-terminal hexaHis-tagged Ldt_{Mt2}. ^{1}H NMR studies with the hexaHis-tag cleaved Ldt_{Mt2} revealed the presence of relatively lower level of hydrolysis products **5/5*** and increased abundance of **2** and **3** compared to the N-terminal hexaHis-tagged Ldt_{Mt2} purified without using the NaOH washed gel filtration columns (Appendix Figure 7.29). These results suggest that the additional step of purification reduced the contaminant present in the hexaHis-tag cleaved Ldt_{Mt2} compared to the N-terminal hexaHis-tagged Ldt_{Mt2}, supporting the hypothesis for the presence of a β-lactamase contaminant.

Care was taken to prepare subsequent batches of purified protein using HisTrap and size exclusion columns that were treated with NaOH prior to use. All Ldts used for ^{1}H NMR studies were purified using chromatographic columns with the rigorous NaOH wash assisted method.

7.3.3.2 Degradation of aminopenicillins via recyclisation to form diketopiperazines

Penicillins with a sidechain amino group can undergo polymerisation via the intermolecular nucleophilic attack by the nitrogen of the sidechain amino group onto the β-lactam carbonyl of another penicillin molecule [360]. The potential of the sidechain amino group in penicillins to act as a nucleophile and to attack the thioester bond in the AECs was investigated by monitoring the products formed upon the reaction of sidechain amino penicillins with Ldts.

Mechanistic studies of the degradation of penicillins by Ldt_{Mt2}

Upon incubation of Ldt_{Mt2} with ampicillin, in addition to the ampicillin-derived C5-C6 fragmentation product i.e., thiazolidine (**2**), unassigned peaks at 1.32 ppm and 1.55 ppm were observed (Figure 7.10). The hydrolysis products (**5** and **5***) were not observed. The uncharacterised product was proposed to be a diketopiperazine **3** following NMR spectroscopic characterisation of the crude mixture.

The ampicillin-derived diketopiperazine product (**3**) was isolated by HPLC and characterised by NMR spectroscopy (Appendix Table 7.8, Appendix Figures 7.31,7.32,7.33,7.34,7.35,7.36). It is proposed that once the acyl-enzyme complex is formed with ampicillin and Ldt_{Mt2}, the side chain amino group reacts with the thioester group in the acyl-enzyme complex to form a diketopiperazine (**3**) and an unmodified Ldt_{Mt2}.

The non-enzymatic degradation of ampicillin to yield a diketopiperazine following prolonged incubation or under alkaline conditions has been reported [361,362]. A mixture of ampicillin diketopiperazine diastereomers (Cas No. 94659-47-9) was purchased and characterised by NMR spectroscopy. The chemical shift assignments for the commercial ampicillin diketopiperazine resembled the chemical shift assignments for the HPLC-purified ampicillin-derived product generated by Ldt_{Mt2} (Appendix Tables 7.8,7.9, Appendix Figures 7.31,7.37). **3** is likely to be epimeric at C5 and C6 and, maybe, the C13 positions; the stereochemistry of these carbons could not be assigned based on the data collected (Appendix Figure 7.38). However, the configurations at C6 and C13 are expected to be (*R*), based on the structure of intact ampicillin. In addition to the ampicillin-derived diketopiperazine (**3**), chemical shifts were also assigned for the degraded thiazolidine product **4**, i.e. *N*-formyl-D-penicillamine and its formylthio derivative (**4***), which have previously been reported to be formed by PBPs from *Streptomyces* R61 and R39 (Appendix Table 7.10, Appendix Figures 7.39, 7.40) [358].

Upon reaction of Ldt_{Mt2} with amoxicillin, the products formed were assigned as thiazolidine (**2**) and diketopiperazine (**3**), formed via C5-C6 fragmentation and cyclisation via rearrangement, respectively. However, lower levels of diketopiperazine (**3**) were formed from amoxicillin compared to ampicillin, suggesting that the hydroxyl group may change the interaction of penicillin in the active site making deacylation via cyclisation less favourable at least in the case of Ldt_{Mt2}. These results suggest that aminopenicillin-derived complexes preferentially deacylate via cyclisation to give diketopiperazine (**3**) or C5-C6 fragmentation to give thiazolidine (**2**) rather than hydrolysis. Penicillin G, which lacks an amino group in the acylamino side chain was also tested. After mixing penicillin G with Ldt_{Mt2}, only low levels of C5-C6 fragmentation and hydrolysis products were observed,

suggesting that the C6-amino group and not the nitrogen in the C6-amido group can participate in this deacylation reaction to give diketopiperazine (**3**).

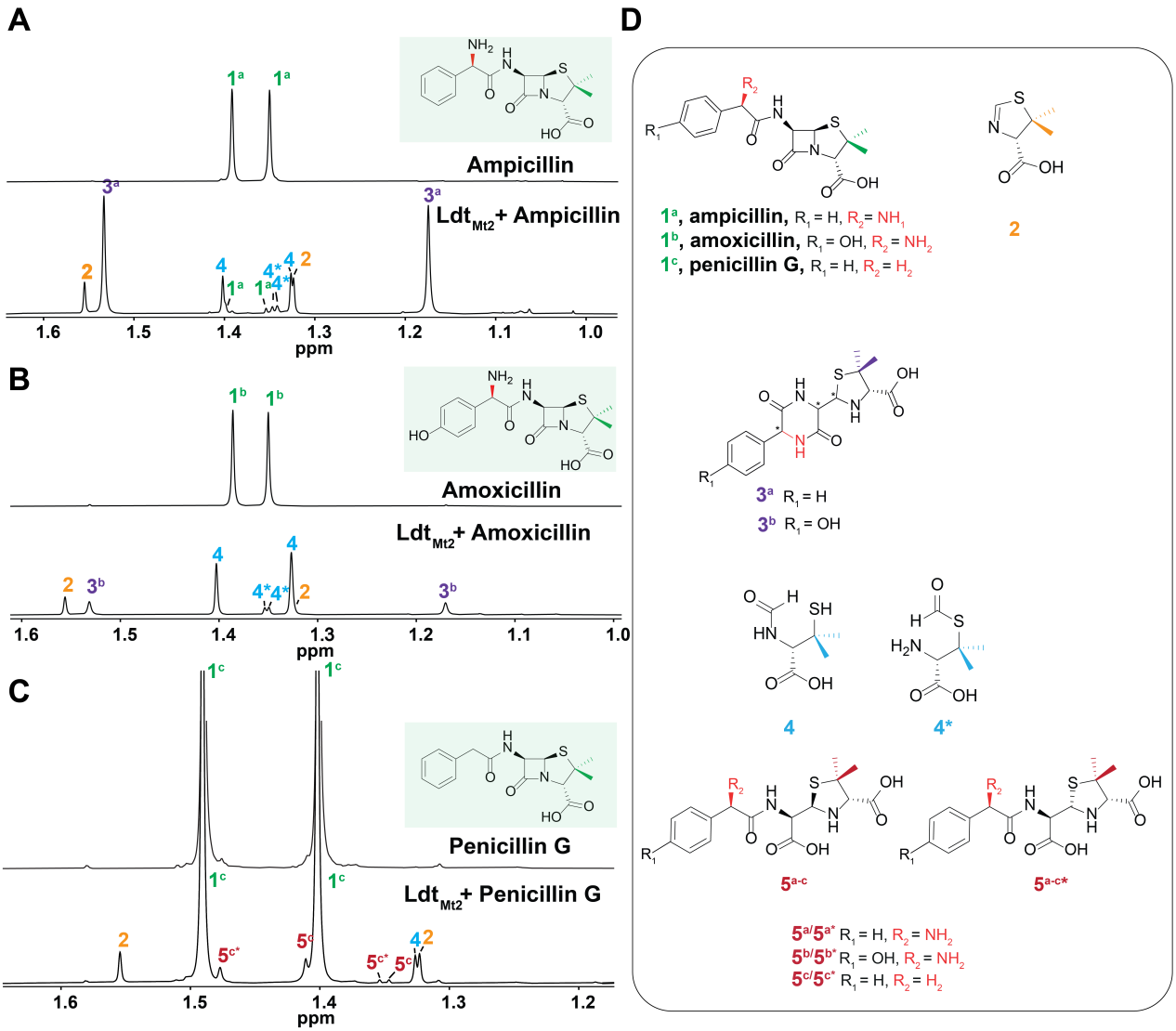


Figure 7.10 | Reaction of penicillins with Ldt$_{Mt2}$. ^{1}H NMR spectra (600 MHz) showing evidence for different product profiles of **A.** ampicillin, **B.** amoxicillin and **C.** penicillin G on reaction with Ldt$_{Mt2}$. C5-C6 fragmentation of the acyl-enzyme complex and cyclisation of the acyl-enzyme complex is observed in penicillins with amino groups in the C6 acylamino sidechain, releasing diketopiperazine product **3**. Hydrolysis products **5** and its epimer **5*** were only observed with penicillin G, which lacks an amino group in the C6 acylamino sidechain. The thiazolidine **2** was the major product observed as a result of C5-C6 fragmentation of the acyl-enzyme complex. The likely epimeric carbons of **3** are indicated with *.

Mechanistic studies with penicillins and Ldt$_{Mt2}$ paralogs

The reaction of penicillins with the four Ldt$_{Mt2}$ paralogs produced by *Mycobacterium tuberculosis*

(i.e., Ldt_{Mt1}, Ldt_{Mt3}, Ldt_{Mt4}, and Ldt_{Mt5}) were next investigated by NMR spectroscopy. With Ldt_{Mt1} and ampicillin (where R_1 = H, Figure 7.11)), formation of hydrolysis products (**5**, its C5-epimer **5***), and low levels of diketopiperazine (**3**) were observed (Figure 7.11). With Ldt_{Mt3}, diketopiperazine (**3**) and hydrolysis products (**5/5***) were observed. No C5-C6 fragmentation or hydrolysis was observed with Ldt_{Mt3} within detection limits, suggesting possible interaction of the side chain amino group with the active site, restricting the movement of the side chain amino group. With Ldt_{Mt4} and ampicillin, formation of **5/5***, **3** and **2** was observed. With Ldt_{Mt5}, **5/5*** and **3** were observed (Figure 7.11).

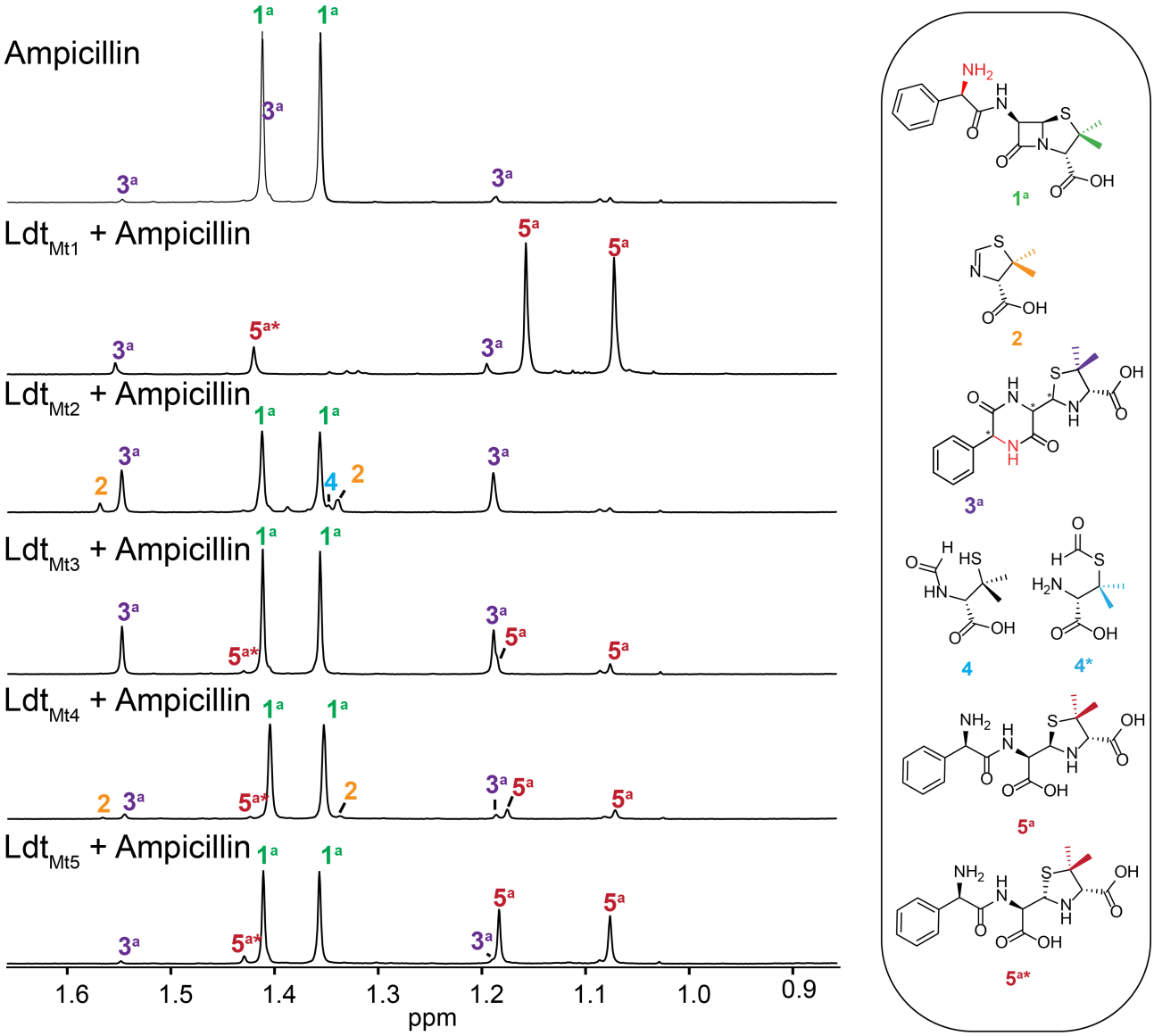


Figure 7.11 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of ampicillin with Ldt$_{Mt2}$ paralogs. Note that **5*** represents the C5 epimer of **5**. The peaks shown correspond to the methyl resonances of the thiazolidine ring or related products. The likely epimeric carbons of **3** are indicated with *. **5** is a C5 epimer of **5***.

In addition to ampicillin, the nature of products formed upon reaction of the Ldt$_{Mt2}$ paralogs with amoxicillin and penicillin G was investigated. With Ldt$_{Mt1}$ and amoxicillin (where R_1 = OH and R_2 = NH_2, Figure 7.12) formation of **5/5*** and low levels of **3** were observed. With Ldt$_{Mt3}$, **3** and **5/5*** were observed. With Ldt$_{Mt4}$ and amoxicillin, formation of **5/5*** and low levels of **2** and **3** were observed. In the case of Ldt$_{Mt5}$ and amoxicillin, formation of **5/5*** and low levels of **3** were observed. With Ldt$_{Mt1}$ and penicillin G (where R_1 = H and R_2 = H, Figure 7.12), similarly to the observations with ampicillin

and amoxicillin and Ldt_{Mt1}, formation of **5/5*** was observed. With Ldt_{Mt2} and penicillin G, formation of low levels of **2**, a *N*-formyl-D-penicillamine (**4**) and its formylthio derivative (**4***) and **5/5*** were observed. With Ldt_{Mt3} and penicillin G, evidence for the formation of low levels of **5/5*** and **2** was observed. In the case of Ldt_{Mt4} and Ldt_{Mt5} and penicillin G, evidence for formation of **5/5*** and **2** were observed. With Ldt_{Mt5} and penicillin G, evidence for degradation of **2** to **4/4*** was also observed.

Overall, Ldt_{Mt1} primarily manifested β-lactamase-like activity with relatively fast degradation of all three tested penicillins, when compared to all other Ldt_{Mt}s. Hydrolysis of ampicillin and amoxicillin was apparently faster than that of penicillin G with $Ldt_{2\text{-}3}$, but not Ldt_{Mt5}, where a similar trend was observed for hydrolysis of all three tested penicillins. The indiscriminate hydrolysis of all tested penicillins was also observed with Ldt_{Mt1}. In the case of the Ldt_{Mt4}, an apparent slow rate of reaction was observed with all three tested penicillins, as compared to all other Ldt_{Mt}s.

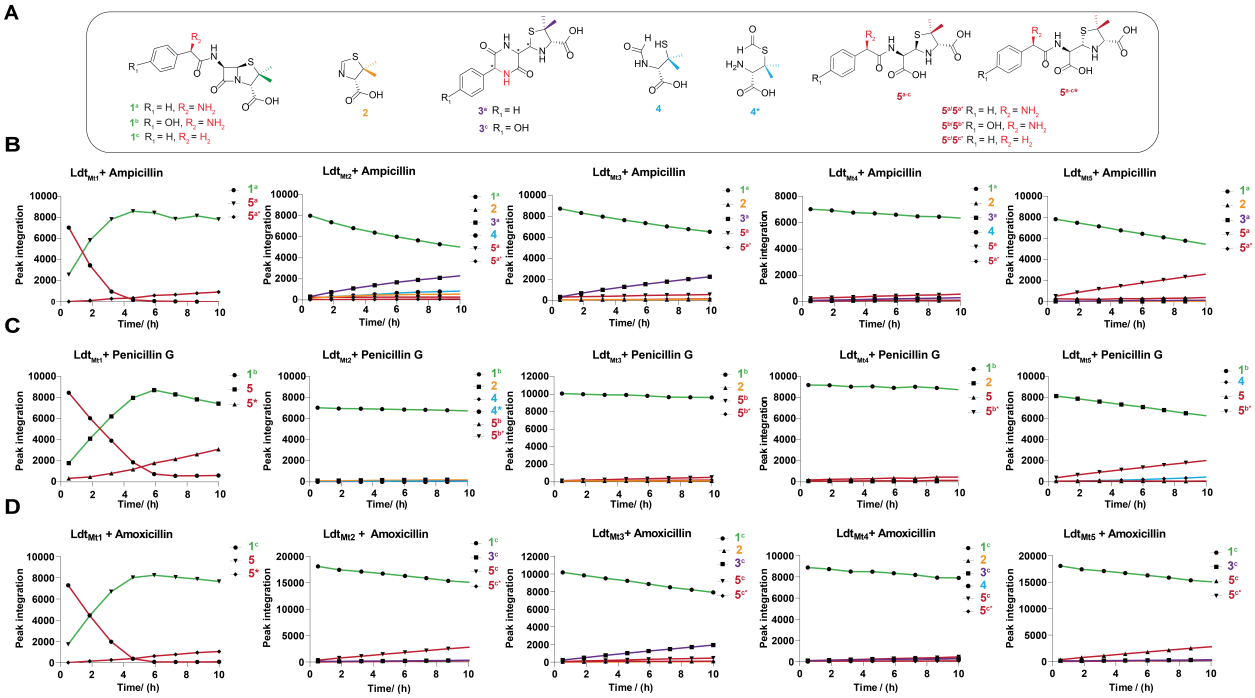


Figure 7.12 | ^{1}H NMR spectra (600 MHz) showing the reaction of ampicillin by Ldt_{Mt}s.A. Structures of cmpd **1-5*** are shown. Peak integration of methyl resonances of cmpd **1-5*** formed upon reaction of $Ldt_{Mt1\text{-}5}$ with **A.** ampicillin, **B.** penicillin G, and **C.** amoxicillin.Ldt_{Mt2} and Ldt_{Mt3} non-hydrolytically deacylate penicillins with C6 side chain amino groups via 6-exo trig cyclisation, releasing diketopiperazine product **3**. While hydrolase-like activity was observed with Ldt_{Mt1} and Ldt_{Mt5}, a poor reaction of penicillins was observed with Ldt_{Mt4}. Note that **4*** corresponds to the formylthio derivative of *N*-formyl-D-penicillamine (**4**), and **5*** represents the C5 epimer of **5**.

To examine if the rate of penicillin turnover by Ldt_{Mts} corresponds to the rate of acylation, time-course MS assays were performed. Evidence for acylation of Ldts was observed with all penicillins, although at varying levels. A dominant peak corresponding to a mass shift of 333 Da was observed with $Ldt_{Mt1\text{-}3}$ upon reaction with penicillin G, suggesting that the apparent poor turnover of penicillin G observed by NMR is not due to a slow rate of acylation (Figure 7.13). The partial reaction of Ldt_{Mt4} and Ldt_{Mt5} to form an apparent AEC was observed with penicillin G, as indicated by mass shifts of 334 Da and 330 Da, respectively. In the case of the reaction of Ldt_{Mt4} and Ldt_{Mt5} with amoxicillin, no evidence for the covalent reaction was observed. With Ldt_{Mt4} and ampicillin, a mass shift corresponding to covalent acylation with ampicillin was also not observed; in the case of the reaction of ampicillin with Ldt_{Mt5}, a mass shift of 595 Da was observed, suggesting reaction of two molecules with ampicillin with Ldt_{Mt5}. This likely indicates the modification of both of the cysteines present in the Ldt_{Mt5} construct (Cys288 and Cys360).

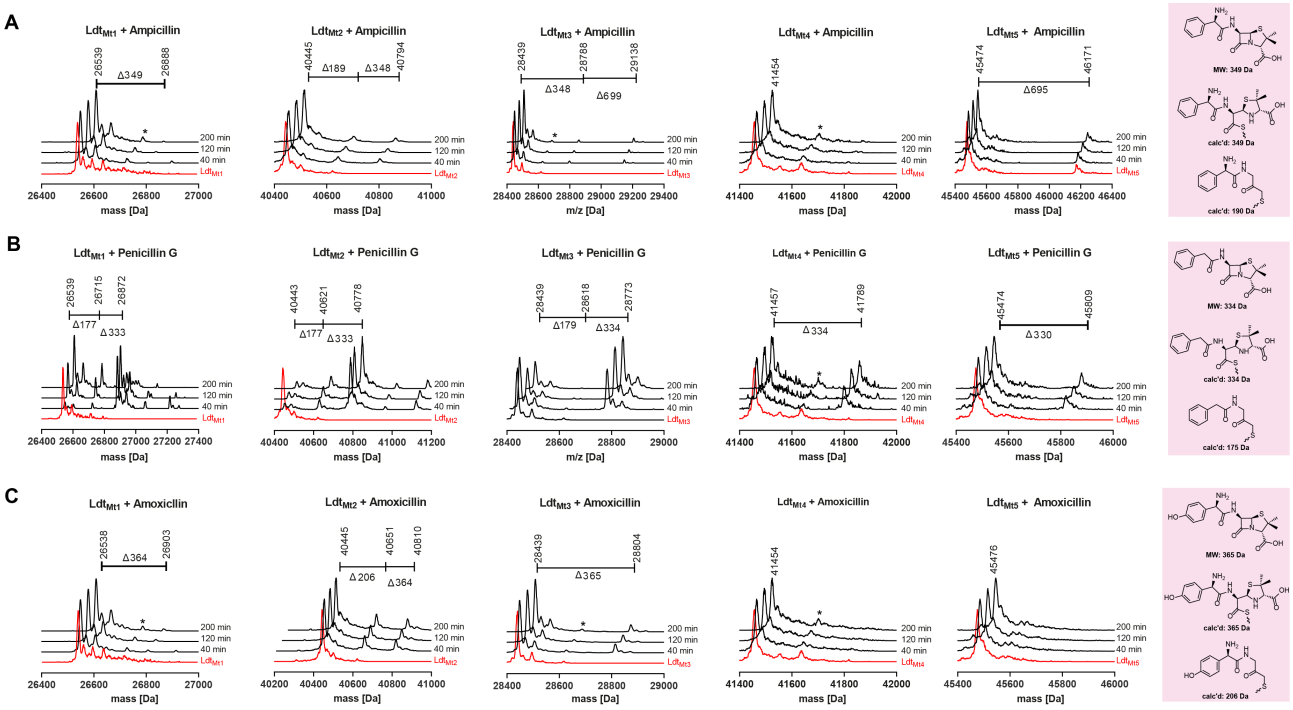


Figure 7.13 | MS spectra for evidence of reaction of Ldt_{Mt}s with ampicillin, amoxicillin and penicillin G. The reaction of **A.** ampicillin, **B.** penicllin G, and **C.** amoxicillin with $Ldt_{Mt1\text{-}5}$ was monitored overtime (∼200 minutes). Evidence for C5-C6 fragmentation of ampicillin (189 Da adduct) was only observed with Ldt_{Mt2} whereas evidence for penicillin G was observed with Ldt_{Mt2} and Ldt_{Mt1}. The peak corresponding to 178 Da modification due to phosphogluconoylation is shown with an *. Intact adducts were observed for all three tested penicillins with $Ldt_{Mt1\text{-}5}$ apart from Ldt_{Mt4} and amoxicillin and Ldt_{Mt5} and amoxicillin.

Mechanistic studies with penicillins and Ldt_{fm}, OXA-48, BlaC and PBP5

To investigate the generality of penicillin deacylation by diketopiperazine (**3**) formation or C5-C6 fragmentation (**2**), ampicillin degradation by related enzymes, e.g., PBPs and β-lactamases, was monitored over time via ^{1}H NMR spectroscopy. With Ldt_{fm}, evidence for formation of **3**, **2**, **4/4*** was observed (Figure 7.14). C5-C6 fragmentation of the Ldt_{fm}-penicillin complex has been previously reported [229,363]. By contrast, with penicillin-resistant PBP5 and serine-β-lactamases BlaC and OXA-48, evidence was only obtained for the formation of **5** and **5***. Therefore, these non-hydrolytic mechanisms of penicillin deacylation, i.e., via the formation of **3** and **2** appear to be more efficient with the Ldts which employ a cysteine nucleophile. However, other features in the active sites of Ldts may also promote the alternative mechanisms, as the Ldt_{Mts} display a varied profile for penicillin degradation. In this regard, it is of interest that a variant of OXA-48 in which the serine nucleophile was substituted with cysteine was observed to form **5/5***, **2**, **4/4***, suggesting the extent of penicillin degradation via C5-C6 fragmentation, at least in the case of OXA-48, arises from the nature of active site nucleophile. The composition of the active site may also influence the rate with which the penicillin side chain amino group attacks the thioester intermediate as indicated by the varying level of formation of **3** with the tested Ldt_{Mts}.

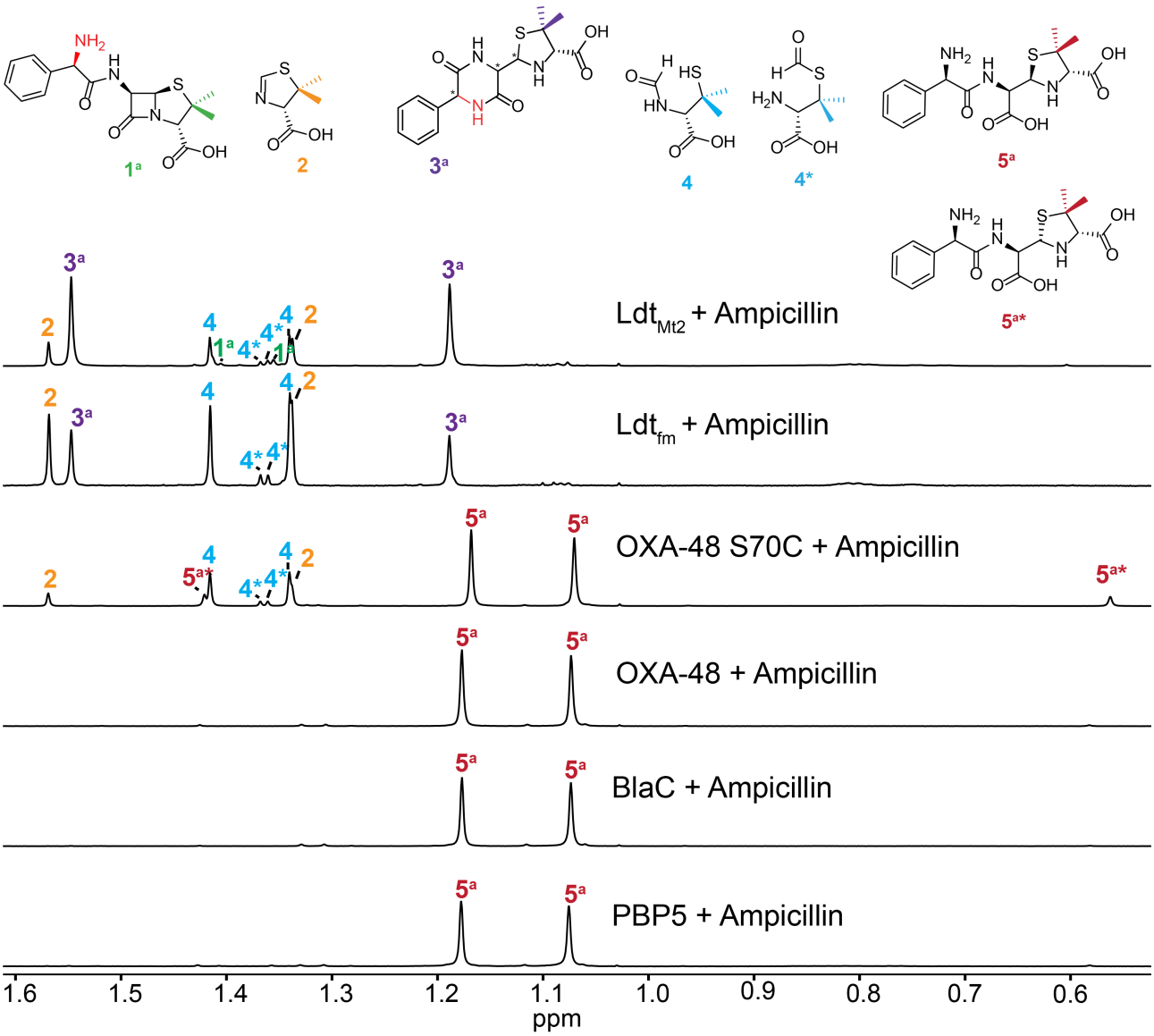


Figure 7.14 | ^{1}H NMR spectra (600 MHz) showing the reaction of ampicillin by Ldt$_{fm}$, PBP5 and BlaC. Note that **4*** corresponds to formylthio derivative of *N*-formyl-D-penicillamine (**4**) and **5*** represents the C5 epimer of **5**.

7.3.4 Mechanistic insights into interaction and degradation of carbapenems

Carbapenems react covalently with Ldt$_{Mts}$ by modifying their active site cysteine residue. Loss of a 44 Da mass from the Ldt$_{Mt2}$-meropenem acyl-enzyme complex (likely a result of MS-related decarboxylation) or a loss of the hydroxyethyl side chain as acetaldehyde has been reported [178, 233]. Evidence for this modification was observed at low level in the MS studies with carbapenem and Ldt$_{Mts}$ (Figure 7.4). However, solution studies for the identification of the fragmented or slow hydrolytic products of carbapenem degradation catalysed by Ldts have, thus far, not been conducted.

7.3.4.1 Lactone formation

To investigate the products formed by the reaction of a carbapenem with the Ldts, the reaction of Ldt_{Mt2} with meropenem was first monitored by NMR. Slow turnover of meropenem was observed, as indicated by the presence of low levels of the *(S)*-Δ^1 hydrolysis product along with a new set of peaks around 1.64 ppm (Appendix Figure 7.42,7.43). To investigate if the formation of the new product is observed with a non-mycobacterial L,D-transpeptidase, reactions were performed with Ldt_{fm}. The observations with Ldt_{Mt2} and meropenem provided evidence for the formation of low levels of a product giving a peak at 1.64 ppm as well as the *(S)*-Δ^1 hydrolysis product was accrued. 1D-TOCSY spin transfer experiments revealed that the peak at 1.64 ppm is coupled to protons which have chemical shifts consistent with a meropenem-derived lactone product. The *(S)*-Δ^1 lactone is reported to be formed on a reaction of meropenem with OXA-48 (Appendix Figure 7.44) [364].

Following addition of OXA-48 to the crude reaction mixture of meropenem and Ldt_{Mt2}, peaks corresponding to the *(R)*/*(S)*-Δ^1 hydrolysis and *(R)*/*(S)*-Δ^1 lactone products were observed (Figure 7.15). The intensity of the peak at 1.64 ppm increased (overlapping with the signals corresponding to the *(S)*-Δ^1 lactone product observed with OXA-48), supporting the proposal that the 1.64 ppm peak corresponds to the *(S)*-Δ^1 lactone. In the case of OXA-48, enzymatically released *(R)*-Δ^1 lactone from meropenem-enzyme complex equilibrates to give the *(S)*-Δ^1 lactone [364]. It may be that the *(R)*-Δ^1 lactone product is present in the Ldt_{Mt2}/Ldt_{fm}-meropenem mixture, but at a level below the detection limit.

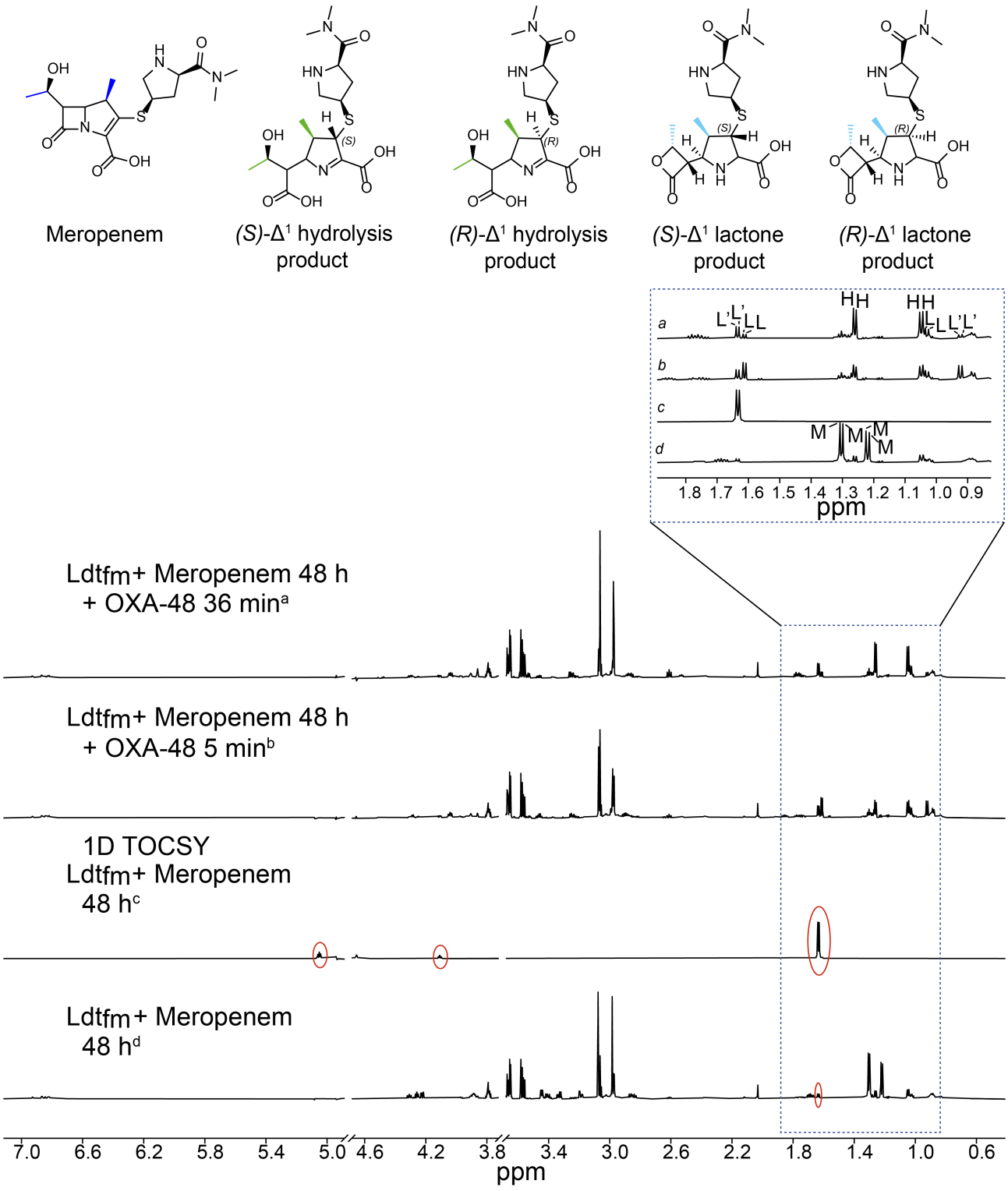


Figure 7.15 | ^{1}H NMR spectra (600 MHz) showing evidence for lactone formation from reaction of Ldt_{fm} and meropenem. The chemical shift for the peak at 1.64 ppm in Ldt_{fm}-meropenem mixture is consistent with the formation of meropenem-lactone product, as observed with OXA-48 [364]. Expansion of methyl region is shown on the inset on the top right where LL, L'L', MM and HH correspond to the methyl resonances from *(R)*-Δ^1 lactone product, *(S)*-Δ^1 lactone product, meropenem and *(R)*-Δ^1 hydrolysis product, respectively.

7.3.4.2 Recyclisation of AECs to give β-lactams

Recently, evidence for deacylation of an acyl-enzyme complex via recyclisation to form a β-lactam ring has been reported with Ldt_{fm} and nitrocefin [356]. To investigate if this mechanism of deacylation occurs with a carbapenem, NMR studies were performed. The penem faropenem forms a stable AEC with Ldts by acylating the active site cysteine nucleophile [178]. An excess of faropenem was thus added to a Ldt_{fm} reacted with a carbapenem to investigate the potential released carbapenem-derived product.

In the experiment, Ldt_{fm} was present in excess compared to the meropenem (2:1) with the aim of ensuring a complete carbapenem reaction to give (an) AEC(s); faropenem was then added to the reaction mixture. The methyl peaks corresponding to meropenem were observed to increase in intensity over time after the treatment of the Ldt_{fm}-meropenem complex with faropenem (Figure 7.16). To validate that the newly formed methyl resonances correspond to the methyl peaks of meropenem, the Ldt_{fm}-meropenem mixture was spiked with an excess of meropenem; the newly formed peaks following the addition of faropenem were consistent with the methyl resonances of meropenem (Appendix Figure 7.45).

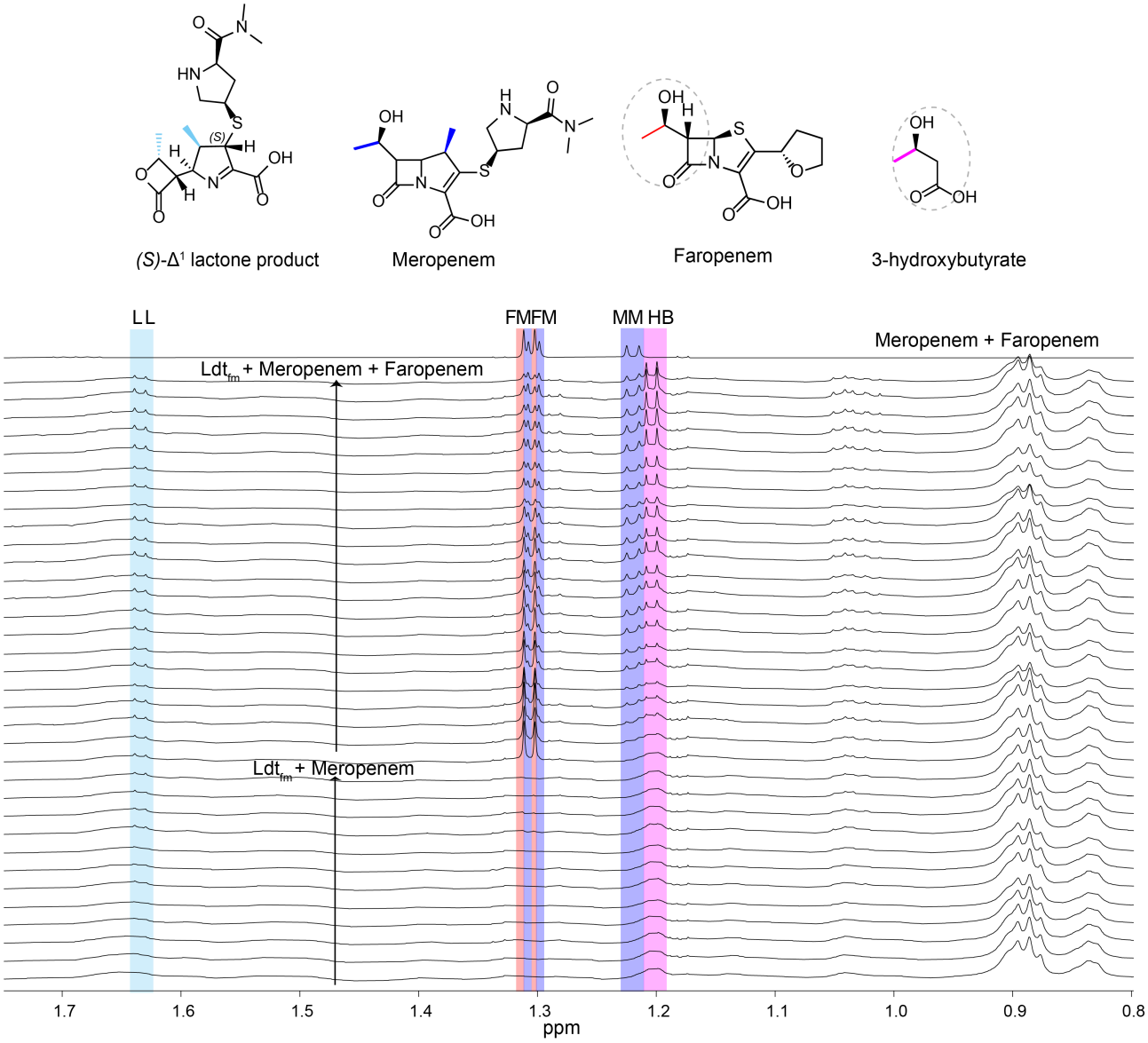


Figure 7.16 | ^{1}H NMR spectra (700 MHz) showing evidence for recyclisation of the meropenem derived from an AEC and lactone formation upon treatment of the Ldt_{fm}-meropenem complex with faropenem. The chemical shifts of the methyl peaks at 1.22 ppm and 1.30 ppm in the Ldt_{fm}-meropenem mixture are consistent with the peak assignments of meropenem observed in the enzyme free meropenem and faropenem mixture (top spectrum). The Ldt_{fm}-meropenem mixture was incubated for 1 h before the addition of faropenem. L'L', FF, MM, and HB correspond to the methyl resonances from *(S)*-Δ^1 lactone product, faropenem, meropenem, and faropenem derived 3-hydroxybutyrate, respectively.

The nature of the C2-substituents and the presence or absence of a 1-β-methyl group in carbapenems can affect their hydrolysis rates and interactions with Ldts [195, 236]. To investigate if the C2-substituents and β-methyl group can affect the recyclisation of carbapenems, the reaction of other carbapenems i.e., doripenem, biapenem, and imipenem, were investigated with Ldt_{fm}. Evidence for lactone formation

as well as reversible recyclisation was observed in all cases (Figures 7.17, 7.18, 7.19).

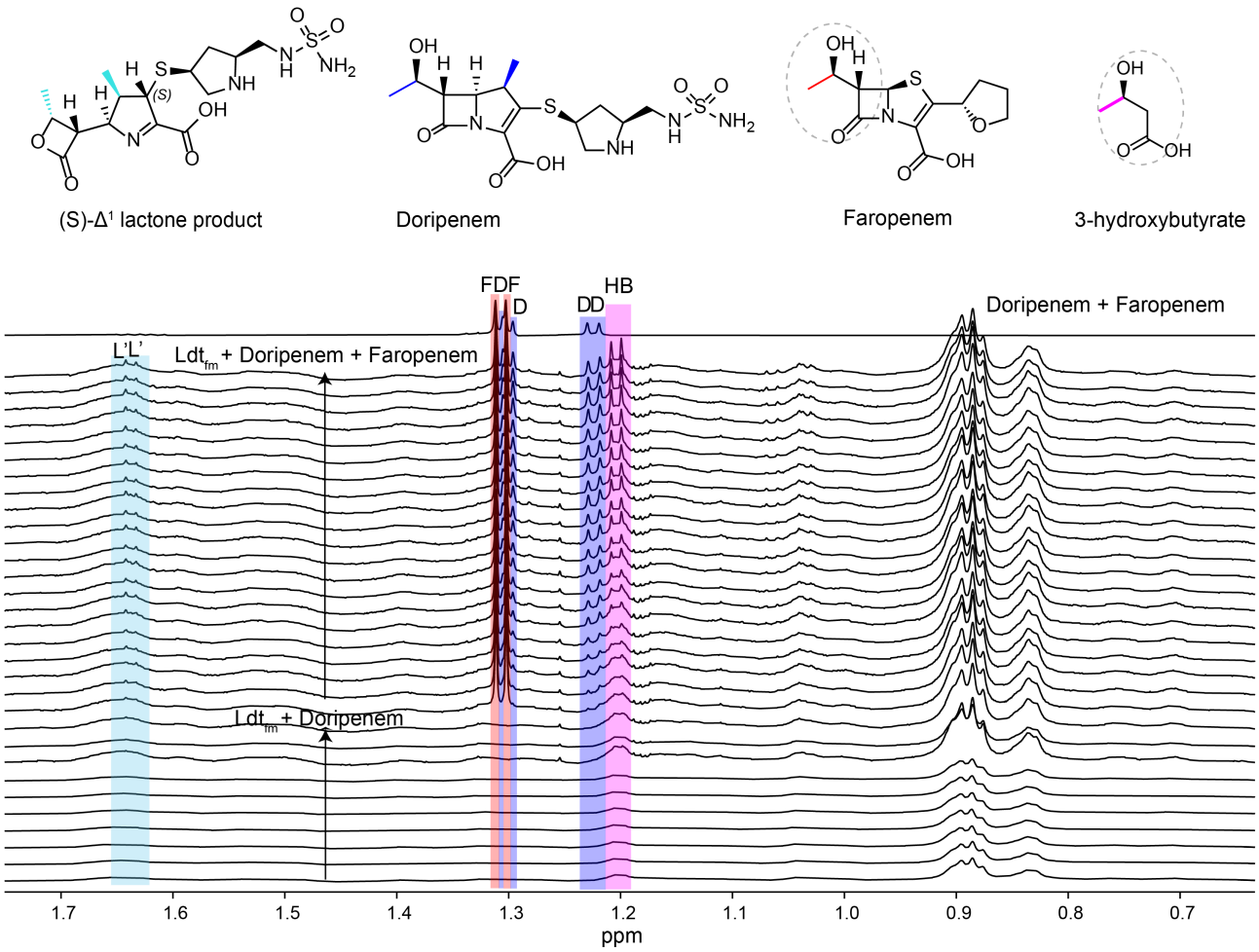


Figure 7.17 | ^{1}H NMR spectra (700 MHz) showing evidence for recyclisation of the doripenem derived from an AEC and lactone formation upon treatment of the Ldt$_{fm}$-doripenem complex with faropenem. The chemical shifts of the methyl peaks at 1.22 ppm and 1.30 ppm in the Ldt$_{fm}$-doripenem mixture are consistent with the peak assignments of doripenem observed in the enzyme free doripenem and faropenem mixture (top spectrum). The Ldt$_{fm}$-doripenem mixture was incubated for 2.5 h before the addition of faropenem. L'L', FF, DD, and HB correspond to the methyl resonances from *(S)*-Δ^{1} lactone, faropenem, doripenem, and faropenem-derived 3-hydroxybutyrate, respectively.

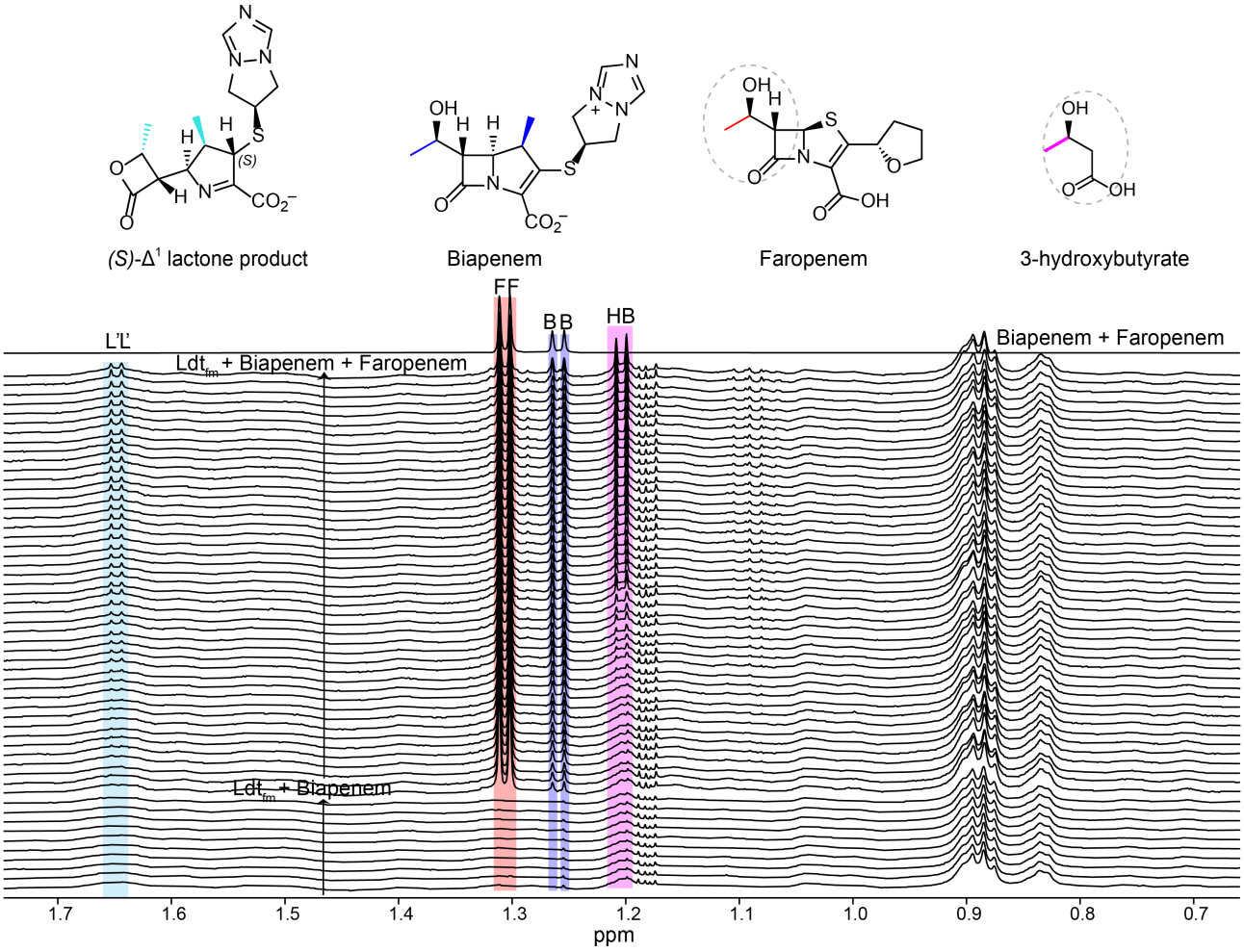


Figure 7.18 | ^{1}H NMR spectra (700 MHz) showing evidence for recyclisation of the biapenem derived from an AEC and lactone formation upon treatment of the Ldt_{fm}-biapenem complex with faropenem. The chemical shifts of the methyl peaks at 1.26 ppm and 1.30 ppm in the Ldt_{fm}-biapenem mixture are consistent with the peak assignments of biapenem observed in the enzyme free biapenem and faropenem mixture (top spectrum). The Ldt_{fm}-biapenem mixture was incubated for 1.3 h before the addition of faropenem. L'L', FF, BB, and HB correspond to the methyl resonances from *(S)*-Δ^1 lactone, faropenem, biapenem and faropenem-derived 3-hydroxybutyrate, respectively.

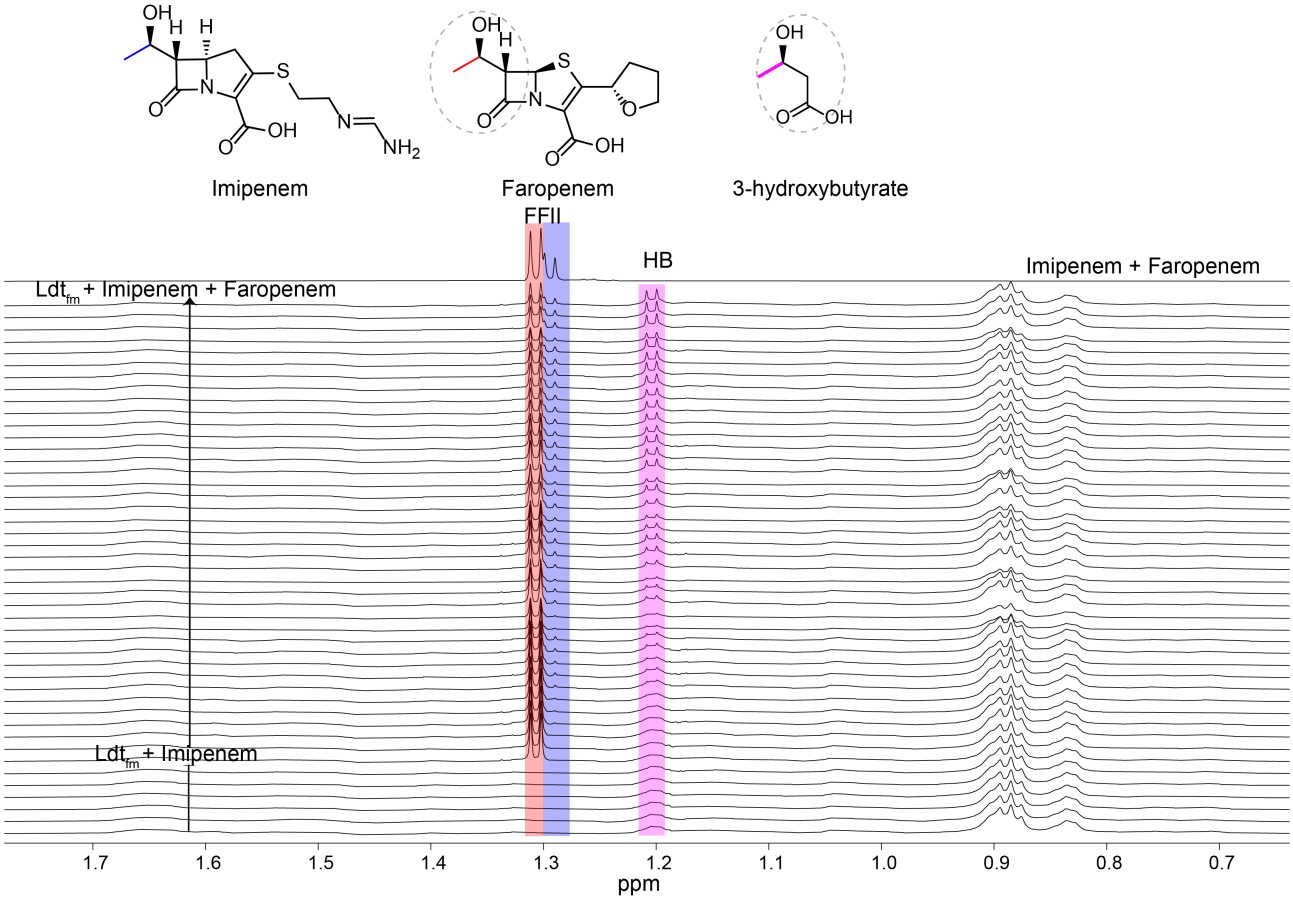


Figure 7.19 | ^{1}H NMR spectra (700 MHz) showing evidence for the recyclisation of the imipenem derived from an AEC but not the formation of the lactone product upon treatment of Ldt_{fm}-imipenem complex with faropenem. Chemical shifts of the methyl peaks at 1.29 ppm in Ldt_{fm}-imipenem mixture are consistent with the peak assignments of imipenem observed in the enzyme free imipenem and faropenem mixture (top spectrum). The Ldt_{fm}-imipenem mixture was incubated for 2.5 h before the addition of faropenem. FF, II, and HB correspond to the methyl resonances from faropenem, imipenem, and faropenem-derived 3-hydroxybutyrate, respectively.

MS studies were used to investigate the reversible reaction of meropenem by monitoring the displacement of the Ldt_{fm}-meropenem acyl-enzyme complex with faropenem. Upon treatment of the Ldt_{fm}-meropenem complex with faropenem, a new peak corresponding to the C5-C6 fragmented Ldt_{fm}-faropenem adduct was observed (Figure 7.20). Similar observations were made upon treatment of Ldt_{fm}-meropenem complex with ebselen (Appendix Figure 7.49). Additional evidence for the reversible reaction of meropenem was provided by the treatment of the Ldt_{fm}-meropenem complex with NDM-1 (24318 Da), which rapidly hydrolyses meropenem, thus apparently depleting the available pool of intact meropenem to acylate Ldt_{fm}. In agreement with this proposal, only minor

evidence for the covalent reaction of meropenem with Ldt_{fm} was observed in the presence of NDM-1 (Figure 7.20C). In the case of both the Ldt_{Mt1} and Ldt_{Mt2}, displacement of meropenem adduct was observed by faropenem (Appendix Figure 7.50).

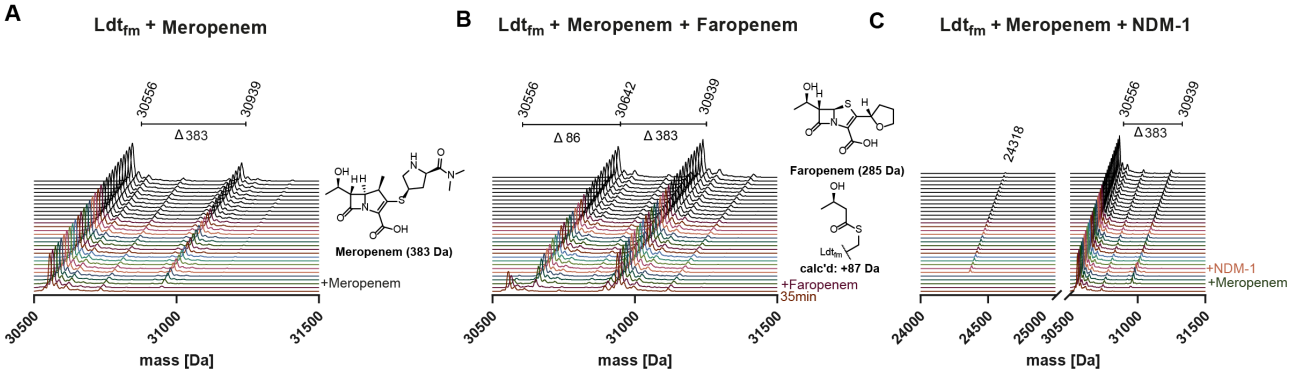


Figure 7.20 | Faropenem competitively binds to meropenem-treated Ldt_{fm}. **A.** Acylation of Ldt_{fm} (1.5 µM) by meropenem (0.75 µM). **B.** An Ldt_{fm}-faropenem adduct at 30642 Da with a mass shift of +86 Da is observed compared to +383 Da adduct with meropenem over time following treatment of the Ldt_{fm}-meropenem complex with faropenem (1.5 µM). **C.** Low level of peak corresponding to the Ldt_{fm}-meropenem complex is observed over time upon treatment with NDM-1 (0.1 µM). NDM-1 hydrolyses meropenem, therefore, reducing the amount of intact meropenem available to react with Ldt_{fm}. Spectra are stacked from bottom to top, and a new spectrum was acquired every 1.35 min.

7.3.5 Mechanistic insights into interaction and degradation of cephalopsorins

Edoo *et al.* [356] have reported a reversible reaction of nitrocefin with Ldt_{fm} based on spectrophotometric assay and MS measurements monitoring the transfer of nitrocefin reacted to Ldt_{fm} to isotopically labelled Ldt_{fm}. However, evidence for the direct observation of the recyclised nitrocefin was not reported in their work.

7.3.5.1 Recyclisation

In an attempt to directly study the recyclised nitrocefin molecule and help validate the reversible binding of nitrocefin with Ldt_{fm}, the Ldt_{fm}-nitrocefin mixture with an excess of Ldt_{fm} (2:1, to ensure complete nitrocefin reaction) was treated with an excess of faropenem then analysed by NMR (Figure 7.21A). New peaks at 5.24 ppm and 5.70 ppm, which were consistent with the C5 and C6 protons of nitrocefin were observed (Figure 7.21B, Appendix Figure 7.51).

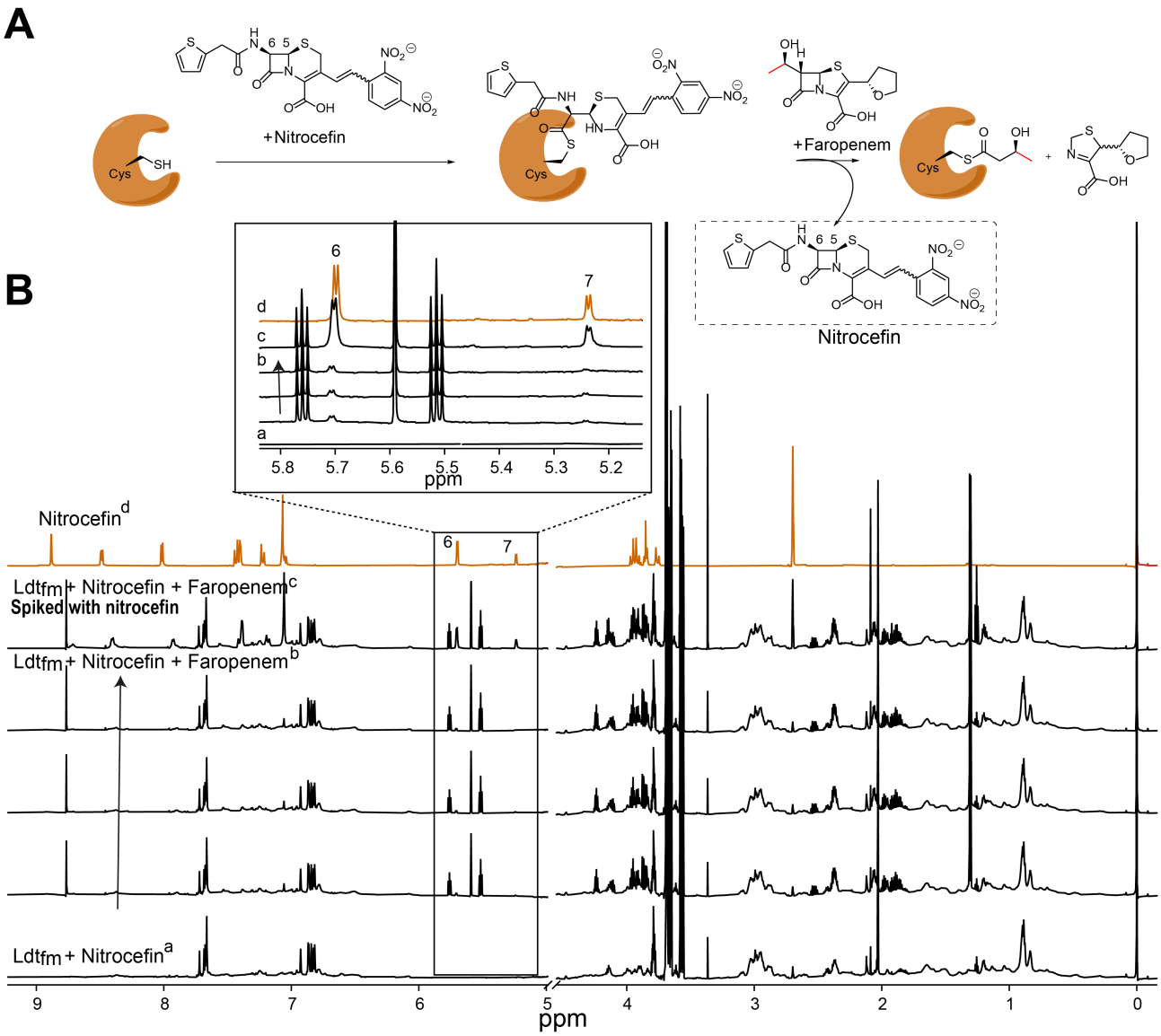


Figure 7.21 | ^{1}H NMR spectra (700 MHz) showing evidence for the recyclisation of nitrocefin upon treatment of the Ldt$_{fm}$-nitrocefin complex with excess faropenem. **A.** A representation of the NMR experiment setup. **B.** Ldt$_{fm}$ (200 μM) was incubated with nitrocefin (100 μM) for 10 min (spectrum a). The Ldt$_{fm}$-nitrocefin complex was incubated for 0.75 h, then treated with an excess of faropenem (600 μM). The three spectra (spectra b) were subsequently acquired after further incubation of 5 min, 20 min, and 53 min. The chemical shifts of the peaks at 5.24 ppm and 5.74 ppm in the Ldt$_{fm}$-nitrocefin mixture are consistent with the assignments of protons attached to C6 and C7, respectively, of nitrocefin in the faropenem treated Ldt$_{fm}$-nitrocefin mixture (spectrum c) and the nitrocefin observed in the enzyme-free sample (spectrum d).

MS studies were performed to monitor the competitive reaction of nitrocefin reacted Ldt$_{fm}$ with imipenem. The Ldt$_{fm}$-nitrocefin complex (28607 Da) was observed to be formed on the treatment of the Ldt$_{fm}$-nitrocefin complex with an excess of imipenem (Figure 7.22). This observation suggests nitrocefin reacts reversibly with Ldt$_{fm}$ and can be displaced by treatment with a carbapenem.

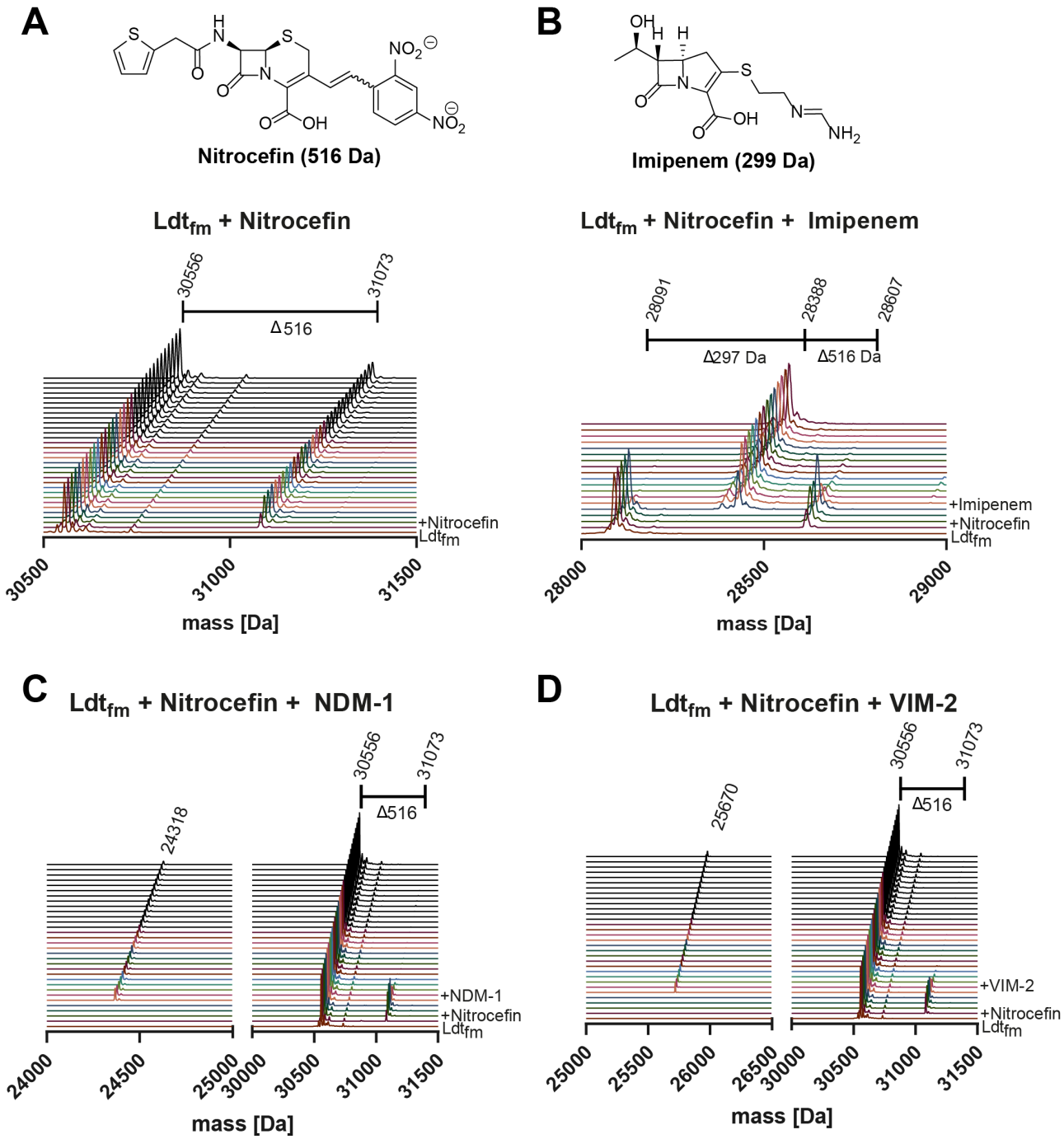


Figure 7.22 | Evidence that nitrocefin reacts reversibly with Ldt_{fm}. **A.** Incubation of Ldt_{fm} (2.5 μM) with nitrocefin (1.25 μM) gives an AEC at 31073 Da. **B.** Treatment of the Ldt_{fm}-nitrocefin mixture with imipenem (1.25 μM) gives a peak corresponding to the imipenem AEC with a mass shift of 297 Da compared to a mass shift of +517 Da with nitrocefin. **C.** Treatment of Ldt_{fm}-nitrocefin mixture with NDM-1 (0.01 μM) leads to a loss of a peak corresponding to the nitrocefin AEC. **D.** Treatment of Ldt_{fm} (2.5 μM) incubated with nitrocefin (1.25 μM) with VIM-2 (0.01 μM).

In addition to monitoring the recyclisation of nitrocefin using NMR and MS, additional MS studies were performed to monitor transfer of nitrocefin from the Ldt_{fm}-nitrocefin complex to the hexaHis-tag

cleaved Ldt_{fm} and vice versa. It was hypothesised that the reversible binding of nitrocefin to Ldt_{fm} should release some of the recyclised nitrocefin into the solution which can react with Ldt_{fm} again. An excess of Ldt_{fm} was added to nitrocefin (2:1) fold and incubated for 5 minutes before the addition of hexaHis-tag cleaved Ldt_{fm} (Figure 7.23). The acylation of hexaHis-tag cleaved Ldt_{fm} was observed suggesting the β-lactam ring in nitrocefin is reformed to react with the added excess protein. Identical experiments were also performed with hexaHis-tag cleaved and hexaHis-tagged Ldt_{Mt2}, the results of which provided corroborative evidence for recyclisation of nitrocefin (Appendix Figure 7.52).

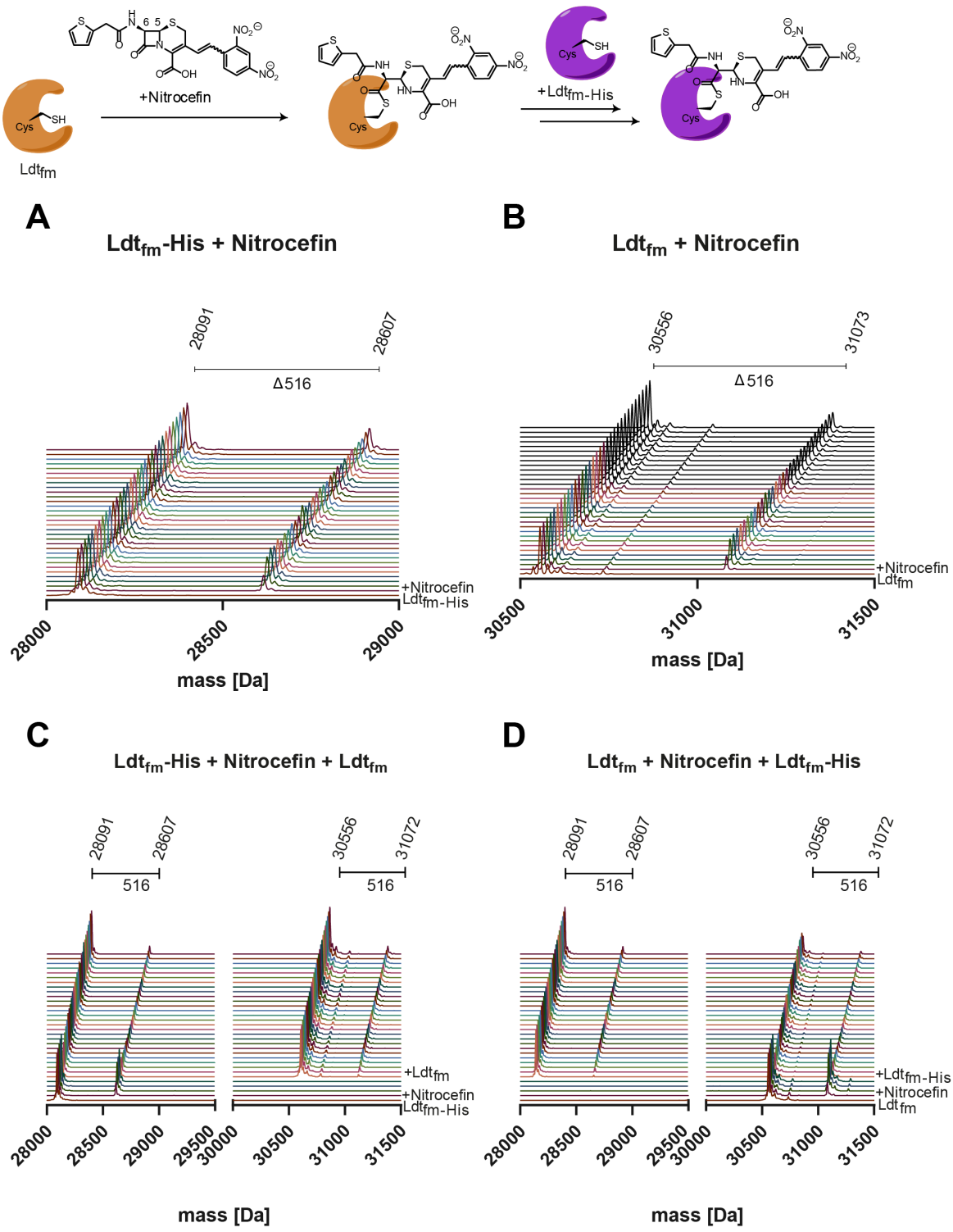


Figure 7.23 | Monitoring recyclisation of nitrocefin using protein transfer MS assays. **A.** Covalent reaction of hexaHis-tag cleaved Ldt_{fm} (1.5 µM) with nitrocefin (0.75 µM). **B.** Covalent reaction of Ldt_{fm} (1.5 µM) with nitrocefin (0.75 µM). **C.** hexaHis-tag cleaved Ldt_{fm} (1.5 µM) was treated with nitrocefin (0.75 µM) and incubated for 5 min before the addition of Ldt_{fm} (1.5 µM). Following the addition of Ldt_{fm}, the relative acylation of hexaHis-tag cleaved Ldt_{fm} with nitrocefin decreased and evidence for acylation of Ldt_{fm} was observed. **D.** Ldt_{fm} (1.5 µM) was treated with nitrocefin (0.75 µM) and incubated for 5 min before addition of the hexaHis-tag cleaved Ldt_{fm} (1.5 µM). Following the addition of hexaHis-tag cleaved Ldt_{fm}, the relative acylation of Ldt_{fm} with nitrocefin decreased, and evidence for acylation of hexaHis-tag cleaved Ldt_{fm} was observed.

7.3.6 Mechanistic insights into interaction and degradation of faropenem, a penem subclass of β-lactam antibiotics

Earlier work with the Ldt_{Mt2} C354S variant and the OXA-48 S70C variant with faropenem revealed that the nature of the active site nucleophile determines the fate of C5-C6 fragmentation of acyl-enzyme complex formed with faropenem [228]. To further investigate the role of cysteine and serine nucleophiles on the nature of C5-C6 fragmentation of faropenem, reactions of faropenem with free L- and D- cysteine and serine were studied. Evidence for the formation of thiazole product was only observed with L-Cys (that is not with D-Cys, L-Ser and D-Ser), suggesting that both the configuration and the nature of nucleophile is important in the reaction with faropenem (Figure 7.24, Appendix Figures 7.54,7.55 and 7.53). These results likely reflect the lower pKa and increased nucleophilicity of thiols versus alcohols. It is of interest that the Ldt_{Mt2} H336A variant with faropenem leads to Cys354 acylation after 30 minutes of incubation [178]. In addition, some *in silico* and crystallographic studies with M^{pro} report that the cysteine of the Cys-His catalytic dyad, is protonated, at least at the start of the reaction [173,314,365]. By contrast, combinatorial scanning mutagenesis studies with SBLs support the notion that the charge relay system is important for the catalysis mediated by the serine nucleophile [366].

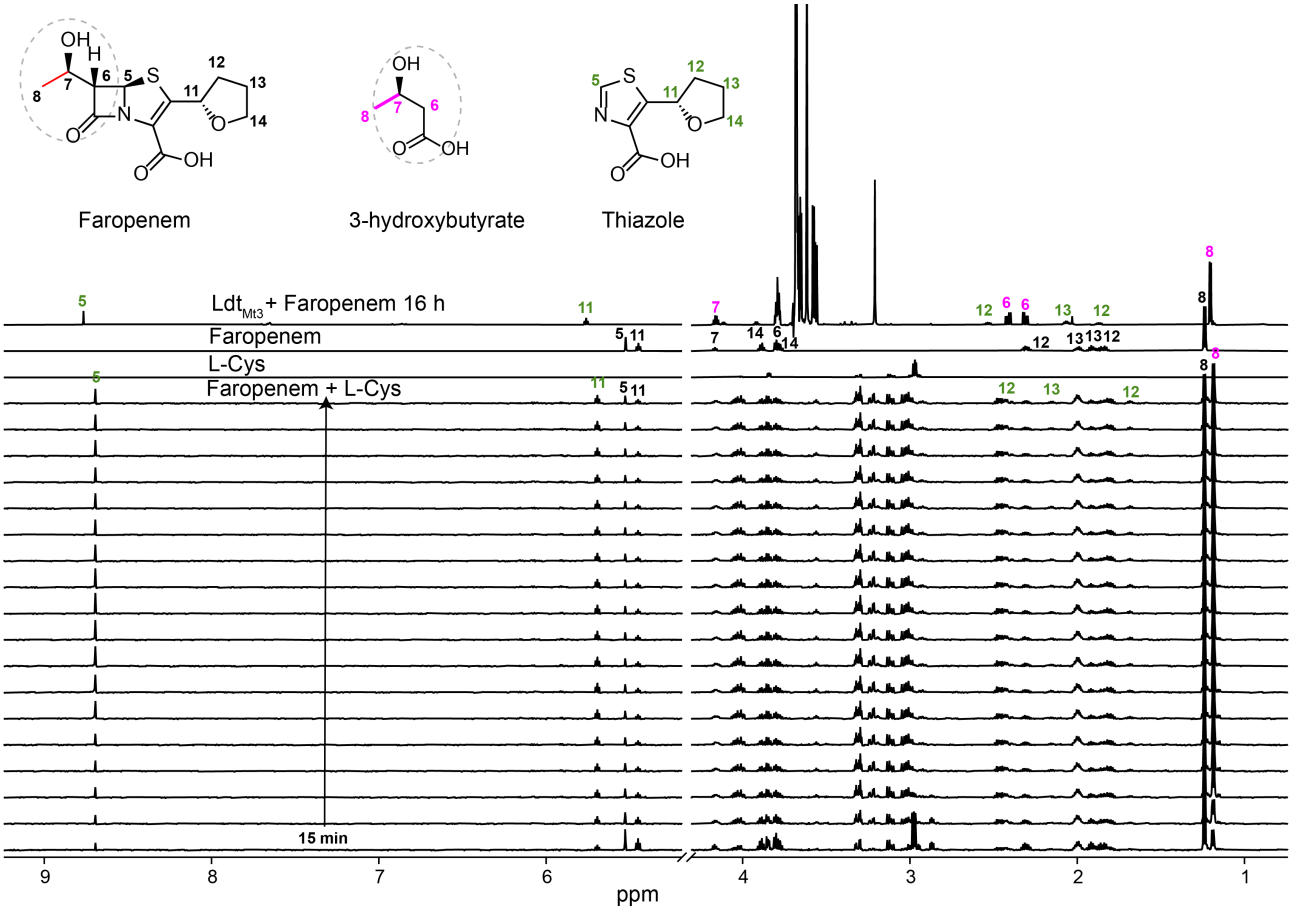


Figure 7.24 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of faropenem with L-cysteine amino acid. The reaction of faropenem with L-Cys gives a faropenem derived thiazole product resulting from C5-C6 fragmentation of faropenem. The first spectrum at the bottom was acquired 15 minutes after sample preparation. Each spectrum was then acquired every 1.3 h (32 scans).

7.3.6.1 D-amino acid mediated deacylation

The acyl-enzyme complex formed upon reaction of faropenem with Ldt$_{Mt2}$ resembles the thioester intermediate formed upon reaction of the Ldt$_{Mt2}$ with donor tetrapeptide substrate stem, which releases the terminal D-Ala residue (Chapter 1, Figure 1.28). To investigate the potential reversibility of this reaction by the reformation of the tetrapeptide stem by a transpeptidation reaction between the thioester intermediate with D-Ala, the faropenem-Ldt$_{Mt2}$ complex was used as a model of the AEC formed with a substrate. ^{1}H NMR studies revealed that D-Ala reacts with the Ldt$_{Mt2}$-faropenem complex to release a product (Figure 7.25). The product was provisionally assigned as an amide product based on NMR studies. An HMBC experiment provided evidence that the chiral protons attached to C2 (faropenem-derived) and C5 (D-Ala derived) of the amide product, each possesses a vicinal carbonyl

group at chemical shifts of 173.2 ppm and 180.3 ppm, respectively (Appendix Figure 7.57). Although evidence for spin-coupling between the C6 proton and the C4 carbonyl or between the C2 proton and the C4 carbonyl was not observed, the lower chemical shift for the C2 carbon compared to the C5 carbon supports the assumption that C2 is likely an amide carbonyl. Thus, an isopeptide bond between the faropenem-derived fragment from the Ldt_{Mt2}-faropenem complex and D-Ala through a S-to-N acyl transfer may give an amide product. Such S-to-N acyl transfers are common in biological processes such as enzymatic ubiquitination, sortase ligation, protein splicing and transglutamination [367].

The observation that D-Ala reacts with the Ldt_{Mt2}-faropenem complex prompted the study of the reaction of *meso*-DAP (as amidated *meso*-DAP is the third residue in the tetrapeptide substrate) with the Ldt_{Mt2}-faropenem complex. The observation of new methyl resonances at 1.16 ppm as well as increased formation of the thiazole product upon the reaction of excess faropenem with regenerated Ldt_{Mt2} in the presence of *meso*-DAP provided evidence for *meso*-DAP-mediated deacylation of the Ldt_{Mt2}-faropenem complex. When *meso*-DAP was reacted with the Ldt_{Mt2}-faropenem complex, new peaks at 1.16 ppm and 1.17 ppm were observed, providing evidence for slow deacylation of the Ldt_{Mt2}-faropenem complex in the presence of *meso*-DAP (Appendix figure 7.56). This observation is consistent with studies which show that the amidation of the *meso*-DAP residue in the tetrapeptide acceptor stem is critical for the 3→3 transpeptidation reaction catalysed by Ldt_{Mt2}, consistent with the lack of evidence for deacylation of the Ldt_{Mt2}-faropenem complex by *meso*-DAP residue using MS assays in the timescale of 40 min [171, 178]. As the amidated *meso*-DAP residue in the substrate stem does not possess a free carboxylic acid group, experiments were also performed with amino acid methyl esters, i.e., glycine methyl ester and D-Ala methyl ester. In both cases, no evidence for the deacylation of Ldt_{Mt2}-faropenem complex was observed upon treatment with these methyl esters (Appendix Figures 7.58, 7.59). These results suggest that the free D-amino acids likely mimic the D-Ala amino acid released after the reaction of Ldt_{Mt2} with donor substrate to form the thioester complex, and not the amidated *meso*-DAP nucleophile that is present in the acceptor stem.

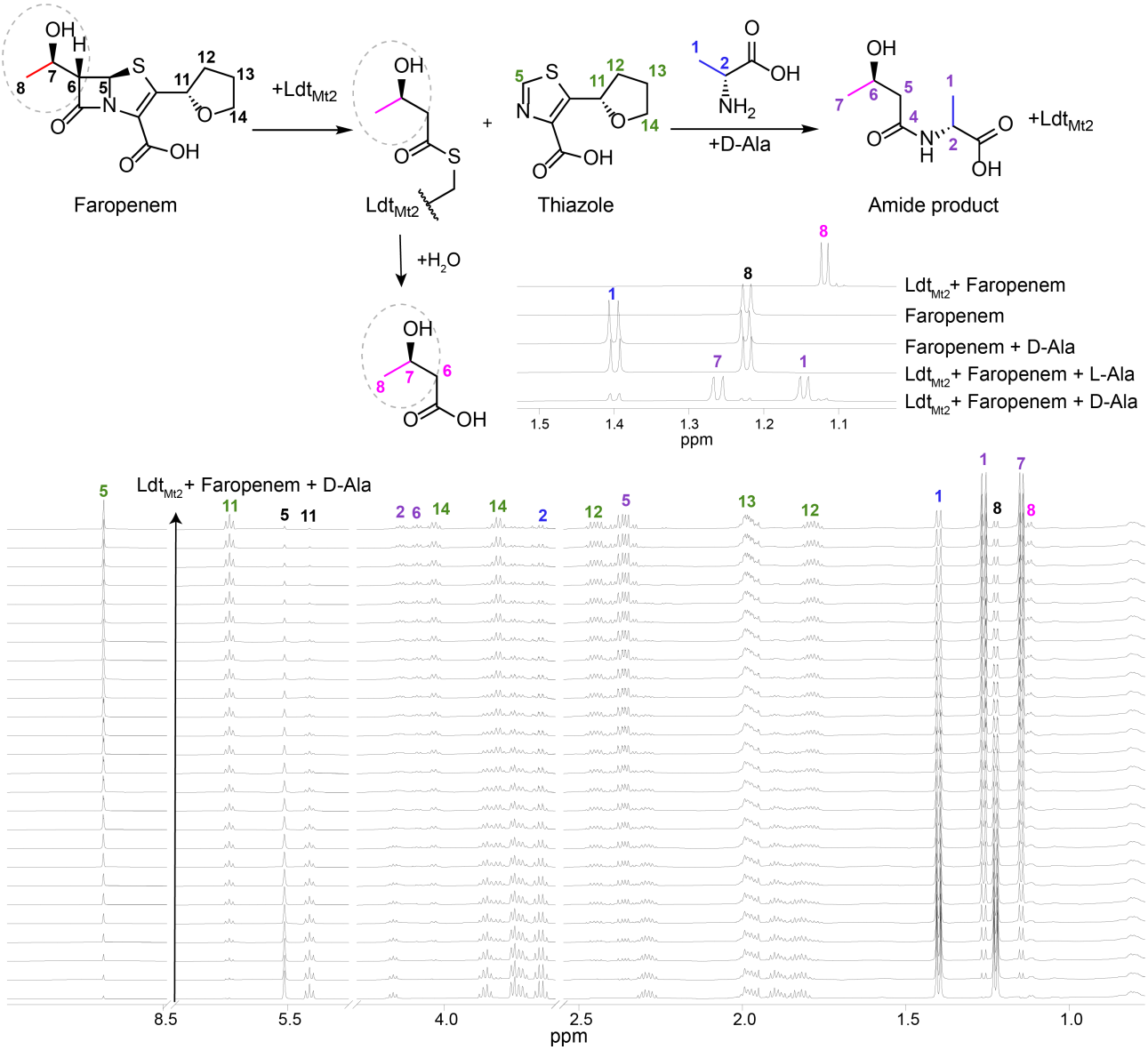


Figure 7.25 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of D-Ala with the Ldt$_{Mt2}$-faropenem complex. The reaction of D-Ala (2 mM) with Ldt$_{Mt2}$ (100 µM)-faropenem (1 mM) complex leads to the formation of an amide product. The first spectrum (at the bottom) was acquired after 5 minutes of reaction initiation. Subsequent spectra were acquired every 45 min (32 scans). The methyl resonances in the top right stacked spectra show that the formation of the amide product was not observed upon the reaction of the Ldt$_{Mt2}$ (100 µM)-faropenem (2 mM) complex with L-Ala (2 mM) for 2.5 h.

Next, the generality of the reaction of the Ldt$_{Mt2}$-faropenem complex with other D-amino acids was assessed. In all cases, apart from L-Cys, only D-amino acids reacted with the Ldt$_{Mt2}$-faropenem complex. Under the non-reducing conditions employed for these NMR assays, low level of peaks corresponding to cystines was observed when investigating the reaction of L- and D-Cys with faropenem. The thiols of L- and D-cysteines likely act as nucleophiles to deacylate the Ldt$_{Mt2}$-faropenem complex

and form a thioester product, compared to the amide product observed with other amino acids where their α-amino group act as nucleophiles. Note that degradation of the initially formed product was observed with L-Cys (Appendix Figure 7.60), potentially via spontaneous intramolecular nucleophilic rearrangement involving the α-amino group of L-Cys and the newly formed thioester carbonyl in the thioester product [368]. The results suggest the following approximate order of reactivity for the reaction of amino acids with the Ldt_{Mt2}-faropenem complex: D-Val>D-Phe=D-Leu>D-Cys=L-Cys>D-Met=D-Ile>D-Trp=D-His=D-Tyr=D-Gln=D-Lys>D-Ala=D-Thr. The characterisation of the amide product derived from D-Phe and the Ldt_{Mt2}-faropenem complex was performed using NMR spectroscopy and MS (Appendix Table 7.11, Appendix Figures 7.61, 7.62 and 7.63). As the β-hairpin lid that is present in a closed conformation in the active site has hydrophobic composition with aromatic residues, the side chains of D-Phe, D-Val and, maybe, D-His might interact with the lid more readily than other amino acids to expose the thioester carbonyl of AEC, leading to efficient deacylation of Ldt_{Mt2}-faropenem complex via S-to-N acyl transfer step [173].

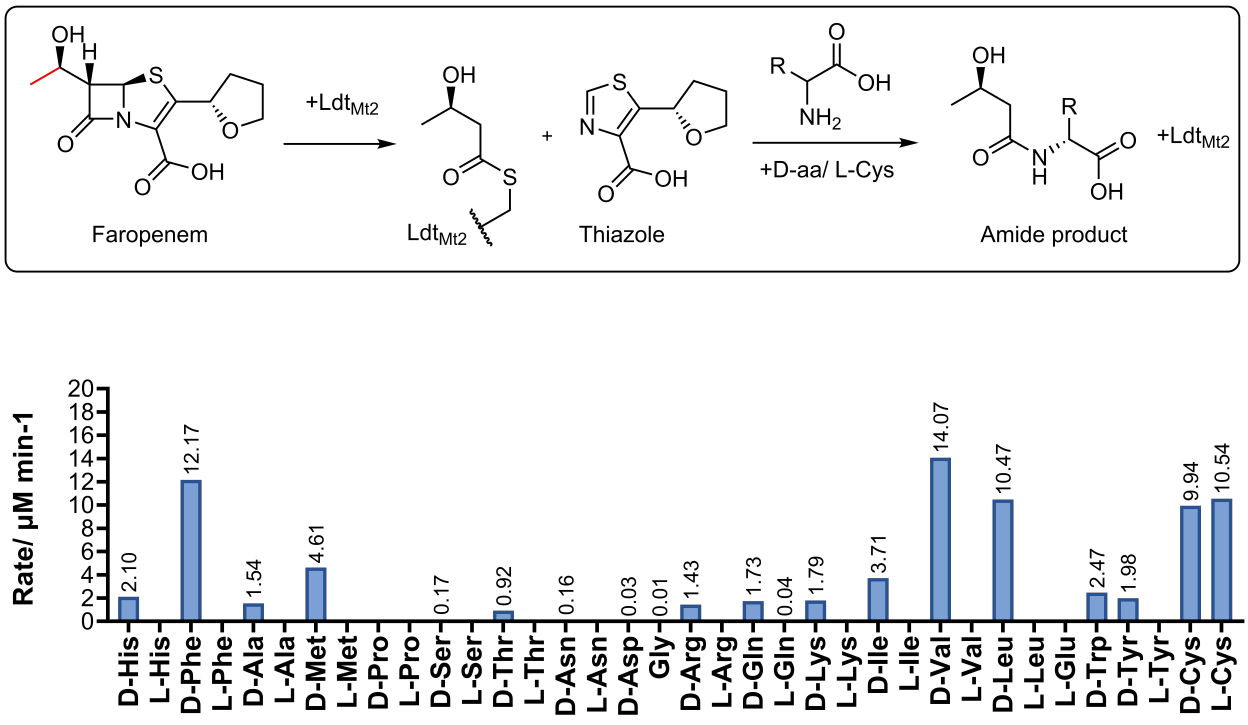
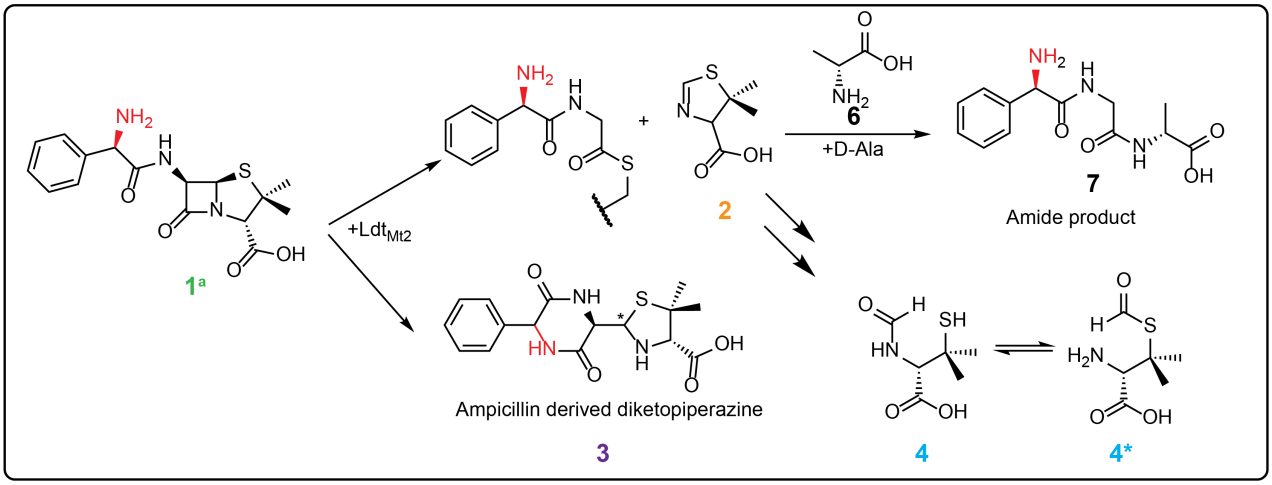


Figure 7.26 | Determination of the rate of deacylation of Ldt$_{Mt2}$-faropenem complex by D-amino acids. The reaction of D-amino acids (1 or 2 mM) with Ldt$_{Mt2}$ (100 µM)-faropenem (1 or 2 mM) thioester intermediate leads to the formation of amide product, which was monitored by NMR. The rate was calculated by normalising the integration of the chemical shifts corresponding to the methyl resonances of the faropenem and amide product formed after D-amino acid mediated deacylation of the Ldt$_{Mt2}$-faropenem complex.

7.3.7 D-amino acid mediated deacylation of penicillin derived AECs

Following the reaction of penicillins such as ampicillin and oxacillin with Ldt$_{Mt2}$, C5-C6 fragmentation may occur to release a thiazolidine product (Section 7.3.3.1). Similar to faropenem, which undergoes C5-C6 fragmentation and can then accept D-amino acids as acyl acceptors, D-amino acid mediated deacylation of Ldt$_{Mt2}$ complexes derived from penicillins were investigated using ampicillin, oxacillin, and piperacillin. In all cases, the formation of new products was observed (Figures 7.27, Appendix Figures 7.65, 7.66). MS evidence for the formation of ampicillin derived amide product was obtained (Appendix Figure 7.64).



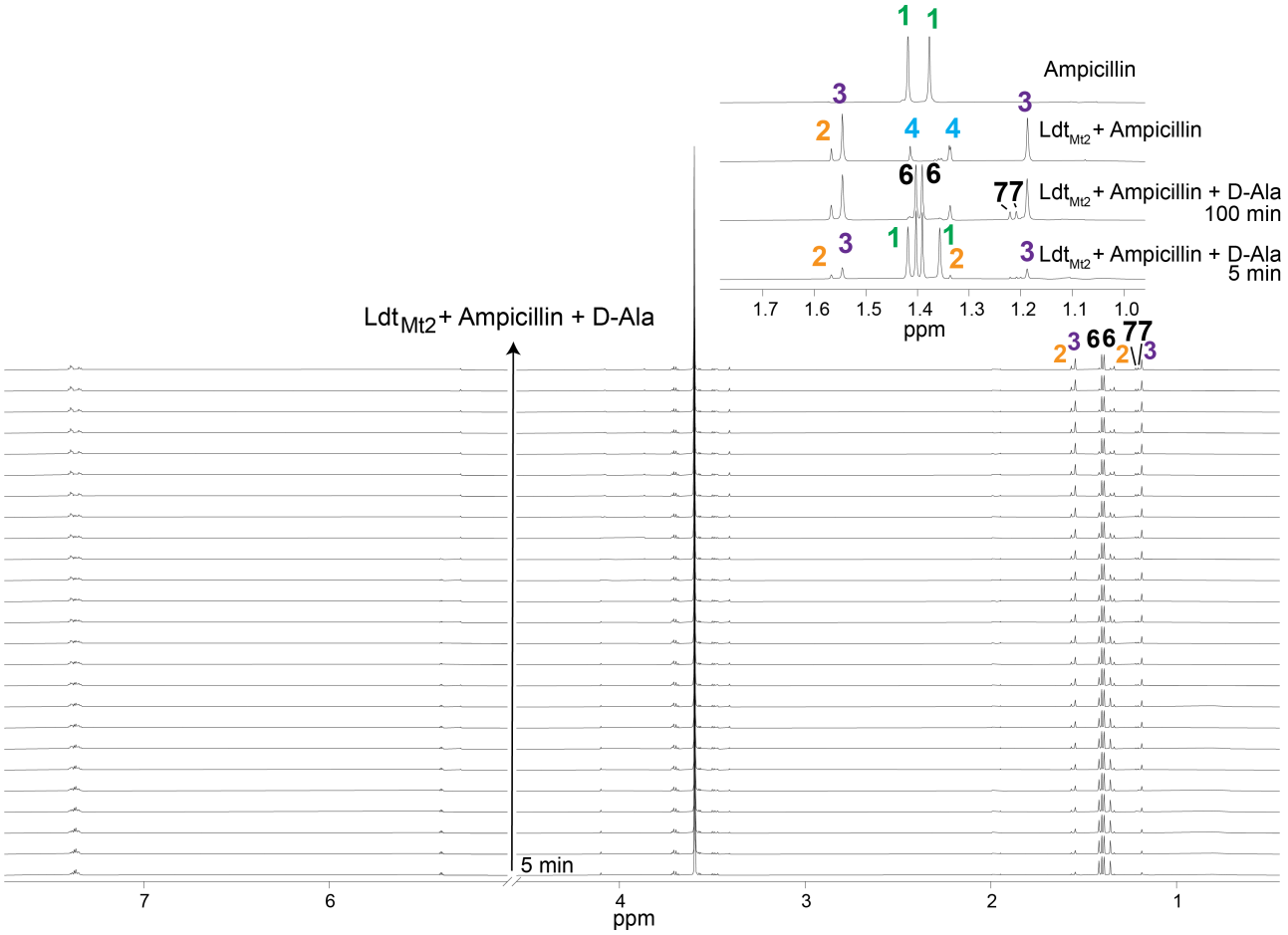


Figure 7.27 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of D-Ala with the ampicillin-derived Ldt$_{Mt2}$ thioester complex. The reaction of D-Ala (2 mM) with Ldt$_{Mt2}$ (100 µM)-ampicillin (1 mM) thioester intermediate leads to the formation of an amide product. The first spectrum at the bottom was acquired 5 minutes after sample preparation. Each subsequent spectrum was acquired every 2.5 min (32 scans). The methyl resonances in the top right stacked spectra show that the formation of the amide product was not observed upon the reaction of Ldt$_{Mt2}$ (100 µM) with ampicillin (2 mM).

7.4 Conclusions

The NMR studies with both carbapenems (meropenem, biapenem, doripenem, and imipenem) and cephalosporins (nitrocefin) support the proposed reversal of the reaction of certain β-lactams with Ldts. Thus, the complexes formed between β-lactams and Ldt_{fm}/Ldt_{Mt2}, which were formed under conditions where there was an excess of the enzyme to β-lactam, were disrupted by treatment with an (apparently) irreversibly binding inhibitor. Importantly, such treatment resulted in NMR signals corresponding to the intact β-lactams. No evidence for other products was accrued (except in the case of the carbapenems, where the formation of lactone was observed). The results are therefore supportive of the reversible reaction of β-lactams with Ldt_{fm}/Ldt_{Mt2}. Evidence for reversibility was also observed by MS that mixing one Ldt-β-lactam derived AEC complex with another Ldt with a different mass, leads to 'transfer' of the inhibitor.

Previously the observation of the reversible reaction of a β-lactam has been interpreted as indicating the reaction of an AEC to reform the β-lactam. However, there are other options. The binding of the β-lactam to the Ldt could be non-covalent, though there is extensive evidence, including by MS presented here, that this is not the case, at least for some β-lactams. Assuming a covalent reaction occurs, the reversibly forming complex could be a tetrahedral intermediate, an AEC complex in which β-lactam derived nitrogen is protonated (as observed in crystallographic studies) or it could be a deprotonated AEC with an amine anion [369]. Further work is required to differentiate between these possibilities.

The results show that the Ldt_{Mt2} AECs can be deacylated by reaction with D-amino acids. In this regard, it is of interest that studies have shown that dipeptides such as Ac_2-L-Lys-D-Ala react with Ldt_{fm} to form a thioester [187]. Amino acid-mediated deacylation of this thioester complex was reported with D-Met, D-Asn, D-Gln, D-Ser, D-Glu and D-Asp [187]. The incorporation of non-canonical D-amino acids into the PG layer has also been reported with *Vibrio cholerae* (D-Met and D-Leu) and *Bacillus subtilis* (D-Phe and D-Tyr) [370]. In the case of *Clostridium difficile*, D-Met and D-Leu can deacylate the AEC formed with its Ldts and their substrates [371]. Incorporation of D-amino acids has been proposed to have biological significance by causing a change in cell shape from a rod to a spherical in

V. cholerae, in a manner linked to induction of stationary phase of growth and reduced PG content [370]. Interestingly, bacteria such as *E. coli* that do not have mechanisms to produce non-canonical amino acids can still incorporate D-amino acids in their PG sacculus upon exogenous treatment with these non-canonical D-amino acids. This D-amino acid exchange is also catalysed by PBPs which may enable resistance to vancomycin, an antibiotic that interacts with the terminal D-Ala-D-Ala motif present in the pentapeptide substrates of PBPs [372–375]. It is possible that the D-amino acid exchange activity of PBPs and the transpeptidase activity and carboxypeptidase activity of Ldts may enable resistance in *Mycobacterium* against both β-lactams and glycopeptides in a synergistic manner.

In general, of the Ldts investigated for interaction with β-lactams, $Ldt_{Mt1\text{-}3}$ and Ldt_{fm} were observed to react efficiently with all subclasses of β-lactams as demonstrated by MS studies. Exceptions were observed, including the reactions of Ldt_{Mt3} and Ldt_{Mt4} with cefditoren, where only partial acylation was observed; the reactions of Ldt_{Mt4} and Ldt_{fm} with oxacillin also resulted in incomplete modification. Ldt_{Mt5} was observed to react poorly with all the tested β-lactams. Given the reaction of the other Ldt_{Mt}s and β-lactams, the lack of reaction of Ldt_{Mt5} with β-lactams is unexpected, as it is analogous to the first step in the transpeptidation reaction [174, 182]. However, consistent with the observations described here, Ldt_{CD3}, one of the three L,D-transpeptidases in *Clostridium difficile*, which also catalyses 3→3 transpeptidation reaction in vitro, was not acylated by the tested β-lactams [371].

In the reaction of Ldt_{Mt5} and faropenem, a peak corresponding to the mass adduct of intact faropenem was observed; this is in contrast to the +86 Da mass adduct fragment bound to the other Ldts. Spectrophotometric studies have been reported for Ldt_{Mt5}-mediated nitrocefin hydrolysis along with (a) crystallographic evidence for covalent acylation with meropenem [182]. Substituting the active site Cys360 of Ldt_{Mt5} with alanine has been reported to reduce the rate of nitrocefin hydrolysis [182]. The low levels of hydrolysis observed with the Ldt_{Mt5} C360A variant suggest the potential for hydrolysis of nitrocefin via non-active site cysteines, or at the active site in a manner independent of cysteine. Thus, there is a possibility that the low level modifications of Ldt_{Mt5} observed with the tested β-lactams may be a result of modification of a non-active site cysteine. Tandem MS

following selective proteolysis could be applied to investigate the identity of the cysteine residue that was observed to undergo low level modification with β-lactams in this work.

Overall, the studies with β-lactams and Ldts provide evidence for several different mechanisms for the different classes of β-lactams following reaction with Ldts. It is hoped that these insights on mechanistic diversity will help in the design of improved inhibitors of the Ldts.

7.5 Future directions

Ldt_{Mt1} and its homolog Ldt_{fm} display hydrolase-like activity against the tested penicillins. The hydrolase-like activity of Ldt_{Mt1} is of interest as the expression of Ldt_{Mt1} is upregulated during the stationary phase of *Mycobacterium* [172]. In the case of Ldt_{Mt2}, meropenem hydrolysis is increased upon substitution of Tyr318 with Ala [176] in the β-hairpin lid. Crystal structures show that Tyr308 and Tyr318 in the β-hairpin lid interact with the hydroxy group and carboxylic acid group of meropenem (Chapter 1, Figure 1.32) [176,224]. The Tyr308 equivalent residue in Ldt_{Mt1} corresponds to an isoleucine residue which may rationalise the hydrolase activity of Ldt_{Mt1}. Site-directed mutagenesis studies with Ldt_{Mt1} I180F and Ldt_{Mt2} Y308I variants, in combination with NMR and crystallographic analyses, may enable swapping of the hydrolytic profiles of Ldt_{Mt1} and Ldt_{Mt2}.

C5-C6 fragmentation was observed with some penicillins and faropenem, but no evidence for C5-C6 fragmentation was observed with meropenem and Ldts. By contrast, a recent report suggests that modification of the C5 position of meropenem with methyl or ethyl groups leads to C5-C6 fragmentation of the meropenem derivatives [376]. MS and crystallographic studies with meropenem derivatives at C5-C6 position and/or reversed stereochemistry of the hydroxy group may lead to the identification of novel mechanisms of interaction of carbapenems with Ldts or reduction in reversible reactions of carbapenems with Ldts. In the context of modifications at the C5-C6 position of β-lactams, it is of interest to perform studies with synthetic derivatives of C6 sidechain amino penicillins with inverted stereochemistry of the amino group and investigate how these impact diketopiperazine formation or C5-C6 fragmentation.

In the case of the reversible reaction of nitrocefin with Ldt_{fm} and Ldt_{Mt2}, NMR studies and two MS based assays were employed. In the protein transfer MS based assays, though His-tagged Ldt_{fm}

or His-tagged Ldt_{Mt2} was in excess of nitrocefin prior to the addition of the second His-tag cleaved Ldt_{fm} or His-tag cleaved Ldt_{Mt2}, residual nitrocefin might have been present in solution. It may be that this nitrocefin, as opposed to the recyclised nitrocefin deacylated from hexaHis-tagged Ldt_{fm} and hexaHis-tagged Ldt_{Mt2} reacted with hexaHis-tag cleaved Ldt_{fm} and hexaHis-tag cleaved Ldt_{Mt2}, respectively. Further experiments should be performed to eliminate this possibility.

MS based evidence has been presented for the formation of amide products upon the reaction of D-amino acids with Ldt_{Mt2}-faropenem and Ldt_{Mt2}-penicillin complexes. Further studies to perform detailed characterisation of all the amide products using MS and HPLC techniques will provide additional supporting evidence.

7.6 Methods

7.6.1 Cloning, protein production and purification

The plasmid pNIC28-Bsa4-Ldt_{Mt1} (Δ1-28) was prepared by ligation independent cloning [228], with a codon-optimised synthetic gene (GeneArt, Thermo Fisher Scientific), and was used to transform *E. coli* BL21(DE3) (Tables 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 and 7.7). Oligos were purchased from Sigma-Aldrich and reagents were sourced from New England Biolabs unless stated otherwise.

Table 7.1: Primers used for PCR.

Genes	Forward and Reverse Primers
Ldt_{Mt1}	5′-TACTTCCAATCCATGGCAGCACCGCCTCTG-3′ and 5′-TATCCACCTTTACTGCTATTAACCCACAACTTCAATCGGATCACC-3′
Ldt_{Mt2}	5′-TACTTCCAATCCATGGATCTGCTGGTTCCGAAACTG-3′ and 5′-TATCCACCTTTACTGCTATTATGCTTTTGCATTACCTGCACGC-3′
Ldt_{Mt3}	5′-TACTTCCAATCCATGCAGAGCTATGGTTTTGCAGTTGC-3′ and 5′-TATCCACCTTTACTGCTATTCTTGCACAATCACCGGATCACC-3′
Ldt_{Mt4}	5′-TACTTCCAATCCATGCCGCATTGGGCTGAAGAAC-3′ and 5′-TATCCACCTTTACTGCTAGATCTGCCAATCCTGTGCAC-3′
Ldt_{Mt5}	5′-TACTTCCAATCCATGGCAGGTAAAGTTACCAAACTG-3′ and 5′-TATCCACCTTTACTGCTAACCACCAGGACCATTGG-3′
Ldt_{fm}	5′-TACTTCCAATCCATGAAAGAACAGCTGGCAAGCATG-3′ and 5′-TATCCACCTTTACTGCTAAAAAACCAGAACCGGTGTACC-3′

Table 7.2: PCR protocol for plasmid DNA.

Components	Volume [µL]	Final Concentration
DNA template	50 ng	
5x Phusion HF Buffer	10	
Cloned Pfu DNA Polymerase	1	
BSA	1 µL (10 mg/mL)	0.2 mg/mL
dNTPs	1 µL (10 mM)	200 µM
Forward primer	1 µL (10 mM)	200 nM
Reverse primer	1 µL (10 mM)	200 nM
H_2O	Add to 50 µL	
Total Volume	50	

Table 7.3: Plasmid DNA PCR Conditions.

	Temperature	**Time**	**Cycles**
Initial denaturation	95 °C	2 min	
Denaturation	95 °C	30 s	
Annealing	Tm calculated using NEBaseChanger	30 s	
Elongation	72 °C	1 min per kb	30
Final elongation	72 °C	10 min	

Table 7.4: Restriction Digestion of LIC plasmid.The sample was incubated at 50 ºC, 300 rpm, 160 min. The digested DNA was subsequently cleaned using a PCR purification kit.

Components	**Volume [µL]**
DNA	5 µg
10x CutSmart Buffer	5
BSA (10 mg/mL)	1
Bsa I (3 µg)	3
H_2O	Add to 100 µL
Total Volume	100

Table 7.5: Preparing cohesive ends for insert. Sample was incubated at r.t. for 30 min, then heated at 75 ºC for 25 min, and stored at -20 ºC.

Components	**Volume [µL]**
Gel extracted insert	5
Lucigen Buffer A	1
dCTP (25 mM)	1
BSA (10 mg/mL)	0.2
DTT (100 mM)	0.5
T4 DNA Ligase	1
Lucigen T4 DNA Polymerase	0.5
H_2O	2.05
Total Volume	10

Table 7.6: Preparing cohesive ends for vector. Sample was incubated at r.t. for 30 min, then heated at 75 ºC for 25 min, and stored at -20 ºC.

Components	Volume [µL]
Plasmid (digested with appropriate enzyme)	50
Lucigen Buffer A	10
dGTP (25 mM)	1
BSA (20 mg/mL)	0.5
DTT (100 mM)	5
T4 DNA Ligase	1
Lucigen T4 DNA Polymerase	1
H_2O	32.4
Total Volume	100

Table 7.7: Transformation of cloned vector in *E. coli* XL10-Gold competent cells. A heat-shock transformation protocol was performed as described [228]. Cells were plated on 2TY agar plates (containing antibiotic and 5% sucrose) and incubated overnight at 37 ºC.

Components	Volume [µl]
LIC-ready vector	1
LIC-ready insert	2
Incubate for 30 min at r.t. and cool on ice for 5 min	
Chemically competent XL10-Gold	33-50

An overnight culture of *E. coli* BL21 (DE3) pNIC28-Bsa4-Ldt_{Mt2} (Δ1-55) was grown at 37 °C with shaking (180 rpm) in 2TY media supplemented with 50 µg/mL kanamycin. This culture (1%) was used to inoculate a larger 2TY culture (with 50 µg/mL kanamycin), which was grown at 37°C with shaking (180 rpm) to an OD600 of 0.6-1. Isopropyl β-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.5 mM, and the culture was incubated at 18 °C, 180 rpm overnight. The cells were harvested by centrifugation (11,000 g, 10 min), and stored at -80 °C.

Cells were resuspended in His A buffer (50 mM Tris-HCl, pH 8, 20 mM imidazole, 500 mM NaCl) and supplemented with DNAse I. Cells were lysed by sonication for 10 min (3s on, 3 s off) at 60% amplitude followed by centrifugation at 4°C for 30 min at 24 krpm. The supernatant was filtered (0.45

μm) and the filtrate loaded into 5 mL HiTrap His column. 20 column volume (CV) washes with His A was performed prior to elution using a linear gradient of 0-100% of His B buffer (50 mM Tris-HCl, pH 8, 500 mM imidazole, 500 mM NaCl) over 8 CV. The eluted fractions were pooled and concentrated, then loaded into S75 column (Superdex®, Sigma-Aldrich) for size exclusion gel filtration purification. The column was equilibrated with 50 mM Tris-HCl, pH 8, 100 mM NaCl after 0.5 M NaOH wash. The purified protein was determined to be >95% pure by SDS PAGE gel and the SPE-MS. The mass of Ldt_{Mt1} was obtained as 26537.9 Da (cal'd 26538.83 Da). Ldt_{Mt2} (Δ1-55, MW 28439 Da), Ldt_{Mt3} (Δ32-271, MW 28439 Da), Ldt_{Mt4} (MW 41454 Da) and Ldt_{Mt5} (Δ1-50, MW 45475 Da) were purified as per the method described for Ldt_{Mt1} with a few differences. 0.2 mM IPTG was used for induction of Ldt_{Mt3} protein in BL21-DE3 cells. Ldt_{Mt5} (0.1 mM IPTG, 24h growth) had 1 mM TCEP present in all buffers. Ldt_{Mt4} was purified with a different buffer composition (His A: 50 mM sodium phosphate, pH 6.4, 20 mM imidazole, 500 mM NaCl, 1 mM TCEP, 5% glycerol), His B (50 mM sodium phosphate, pH 6.4, 500 mM NaCl, 1 mM TCEP, 5% glycerol) and size exclusion gel filtration buffer (50 mM sodium phosphate, pH 6.4, 100 mM NaCl, 1 mM TCEP, 5% glycerol). A step gradient (0-20% 2 CV, 20-40% 2 CV, 40-60% 2 CV, 60-805 2 CV, and 80-100% 2 CV) was used for elution of Ldt_{Mt4} from the His column and concentration determined using the standard Pierce BCA Protein Assay Kit (Thermo Scientific). Δ215-466Ldt_{fm} (30555 Da) was purified as described for Ldt_{Mt2} [228].

The sequences of the Ldts were:

Ldt_{Mt1}: MHHHHHHSSGVDLGTENLYFQSMAAPPLQPIPGVASVSPANGAVVGVAHPVV VTFTTPVTDRRAVERSIRISTPHNTTGHFEWVASNVVRWVPHRYWPPHTRVSVGVQELT EGFETGDALIGVASISAHTFTVSRNGEVLRTMPASLGKPSRPTPIGSFHAMSKERTVVMD SRTIGIPLNSSDGYLLTAHYAVRVTWSGVYVHSAPWSVNSQGYANVSHGCINLSPDNA AWYFDAVTVGDPIEVVG

Ldt_{Mt2}: MHHHHHHSSGVDLGTENLYFQSMDLLVPKLTASVTDGAVGVTVDAPVSVT AADGVLAAVTMVNDNGRPVAGRLSPDGLRWSTTEQLGYNRRYTLNATALGLGGAATR QLTFQTSSPAHLTMPYVMPGDGEVVGVGEPVAIRFDENIADRGAAEKAIKITTNPPVEGA FYWLNNREVRWRPEHFWKPGTAVDVAVNTYGVDLGEGMFGEDNVQTHFTIGDEVIATA

DDNTKILTVRVNGEVVKSMPTSMGKDSTPTANGIYIVGSRYKHIIMDSSTYGVPVNSPNG
YRTDVDWATQISYSGVFVHSAPWSVGAQGHTNTSHGCLNVSPSNAQWFYDHVKRGDIV
EVVNTVGGTLPGIDGLGDWNIPWDQWRAGNAKA

Ldt_{Mt3}: MHHHHHHSSGVDLGTENLYFQSMQSYGFAVASVLPTRGQVVGVAHPVVVT
FSAPITNPANRHAAERAVEVKSTPAMTGKFEWLDNDVVQWVPDRFWPAHSTVELSVGSL
SSDFKTGPAVVGVASISQHTFTVSIDGVEEGPPPPLPAPHHRVHFGEDGVMPASMGRPEYP
TPVGSYTVLSKERSVIMDSSSVGIPVDDPDGYRLSVDYAVRITSRGLYVHSAPWALPALG
LENVSHGCISLSREDAEWYYNAVDIGDPVIVQE

Ldt_{Mt4}: MHHHHHHSSGVDLGTENLYFQSMPHWAEERHRRESNYVALEAGLDEGESIR
RSEHSRSGCGADAGCWRCRGGPGRGSRRSRRSRGPGGTAGPVDPPAVDLLAPPPDPLALP
PALDPLAPPPPDPLAPPPPDPLAVPVAAGPVAGQDPTSFVGPPPFRPPTFNPVDGAMVGVA
KPIVINFAVPIADRAMAESAIHISSIPPVPGKFYWMSPTQVRWRPFEFWPANTAVNIDAAGT
KSSFRTGDSLVATADDATHQMTITRNGVVQKTFPMSMGMVSGGHQTPNGTYYVLEKFA
TVVMDSSTYGVPVNSAQGYKLTVSDAVRIDNSGNFVHSAPWSVADQGKRNVTHGCINLS
PANAKWFYDNFGSGDPVVVKNSVGTYNKNDGAQDWQI

Ldt_{Mt5}: MHHHHHHSSGVDLGTENLYFQSMAGKVTKLAEKRPPPAPRLTFRPADSAAD
VVPIAPISVEVGDGWFQRVALTNSAGKVVAGAYSRDRTIYTITEPLGYDTTYTWSGSAVG
HDGKAVPVAGKFTTVAPVKTINAGFQLADGQTVGIAAPVIIQFDSPISDKAAVERALTVT
TDPPVEGGWAWLPDEAQGARVHWRPREYYPAGTTVDVDAKLYGLPFGDGAYGAQDM
SLHFQIGRRQVVKAEVSSHRIQVVTDAGVIMDFPCSYGEADLARNVTRNGIHVVTEKYS
DFYMSNPAAGYSHIHERWAVRISNNGEFIHANPMSAGAQGNSNVTNGCINLSTENAEQY
YRSAVYGDPVEVTGSSIQLSYADGDIWDWAVDWDTWVSMSALPPPAAKPAATQIPVTA
PVTPSDAPTPSGTPTTTNGPGG

Ldt_{fm}: MHHHHHHSSGVDLGTENLYFQSMKEQLASMNAIANVKATYSINGETFQIPSSD
IMSWLTYNDGKVDLDTEQVRQYVTDLGTKYNTSTNDTKFKSTKRGEVTVPVGTYSWTIQ
TDSETEALKKAILAGQDFTRSPIVQGGTTADHPLIEDTYIEVDLENQHMWYYKDGKVALE

TDIVSGKPTTPTPAGVFYVWNKEEDATLKGTNDDGTPYESPVNYWMPIDWTGVGIHDSD
WQPEYGGDLWKTRGSHGCINTPPSVMKELFGMVEKGTPVLVF

7.6.2 HPLC purification of diketopiperatizine 3

10 µM Ldt_{fm} NMR was incubated with 500 fold excess ampicillin (15 mL reaction volume) and incubated at r.t. for 7 days. The protein was removed using 4 mL Amicon 10 KDa filter and the flow through was filtered using a 0.25 µM filter. Filtrate was manually injected into an HPLC machine (Shimadzu) equipped with a reverse phase semi-preparative column (Gemini 00G-4454-U0-AX; phase: NX-C18). Sample was eluted using a gradient of ($3\%_{(v/v)}$- $45\%_{(v/v)}$) MeCN + $0.1\%_{(v/v)}$ CF_3CO_2H over 55 min (t_R= 2.4 min).

7.6.3 NMR Spectroscopy

NMR Spectroscopy NMR experiments were carried out in Bruker AVIIIHD 600 MHz spectrometer equipped with a Prodigy N2 broadband cryoprobe or a Bruker AVIII 700 MHz spectrometer equipped with an inverse TCI cryoprobe.

7.6.3.1 Monitoring enzymatic hydrolytic products of penicillins

Samples were prepared with $10\%_{(v/v)}$ D_2O in 50 mM sodium phosphate, pH 7.9 buffer. For monitoring the degradation of intact penicillins, mycobacterial Ldts (10 µM Ldt_{Mt1}, Ldt_{Mt2}, Ldt_{Mt3}, Ldt_{Mt4} or Ldt_{Mt5}), Ldt_{fm} (100 uM), PBP5 (10 µM), OXA-48 (1.7 µM), OXA-48 S70C (100 µM) and BlaC (5 µM) were investigated with 2 mM penicillins (ampicillin, amoxicillin and penicillin G) were treated with 2 mM penicillins. 2D NMR methods ^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY, ^{1}H-^{13}C HMBC, ^{1}H-^{13}C-HSQC were employed for characterisation of the products from the crude mixtures.

7.6.3.2 D-amino acid mediated deacylation of C5-C6 fragment bound Ldt_{Mt2} acyl-enzyme.

Samples were prepared with deuterated in 50 mM sodium phosphate, pH 7.9 buffer or deuterated 50 mM sodium phosphate, pH 7.9. Purified Ldt_{Mt2} (100 µM) was incubated with faropenem (2 mM)

on ice for 5 min prior to the addition of L- or D-amino acids (2 mM). ^{1}H-^{13}C HMBC was employed for characterisation of the new amide products from the crude mixtures.

7.6.3.3 Recyclisation of carbapenems and nitrocefin.

NMR experiments were carried out on Bruker AVIII 700 MHz spectrometer due to the higher sensitivity afforded for protons. Samples consisted of 100 µM Ldt_{Mt2} treated with 50 µM carbapenem prior to the addition of 100 µM Faropenem in 50 mM sodium phosphate, pH 7.5 buffer. Two fold higher enzyme and the corresponding compounds were used for experiments with Ldt_{fm} [pulse program: zgesgppe (water suppression using excitation sculpting with gradients), ns=32, o1p: 4.70].

7.7 Appendix

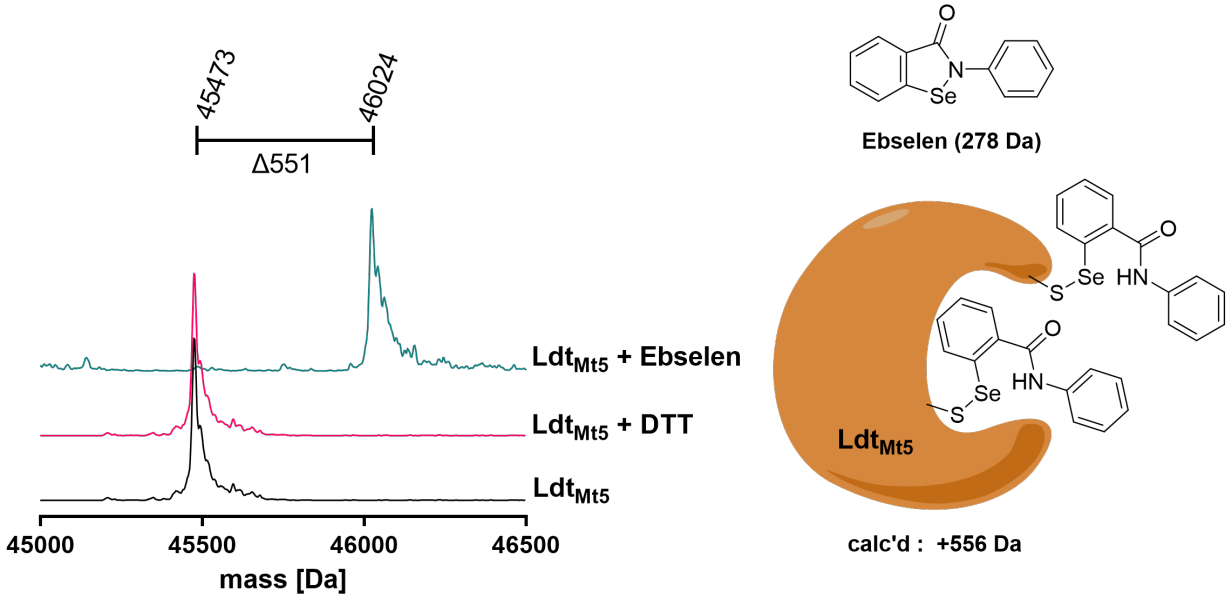


Figure 7.28 | Evidence that the +60 Da adduct observed for Ldt_{Mt5} does not involve modification of cysteine residues, including active site nucleophile Cys360. Treatment with a reducing agent, DTT did not impact on the peak corresponding to Ldt_{Mt5}, suggesting that the cysteine residues are not reversibly oxidised. The observed reaction of Ldt_{Mt5} with two ebselen molecules supports the proposal that the two cysteine residues present in the Ldt_{Mt5} construct tested in this work are not post-translationally modified. Conditions: 2 µM Ldt_{Mt5} with 2 mM DTT or 30 µM ebselen, in 50 mM sodium phosphate, pH 7.5 (15 min).

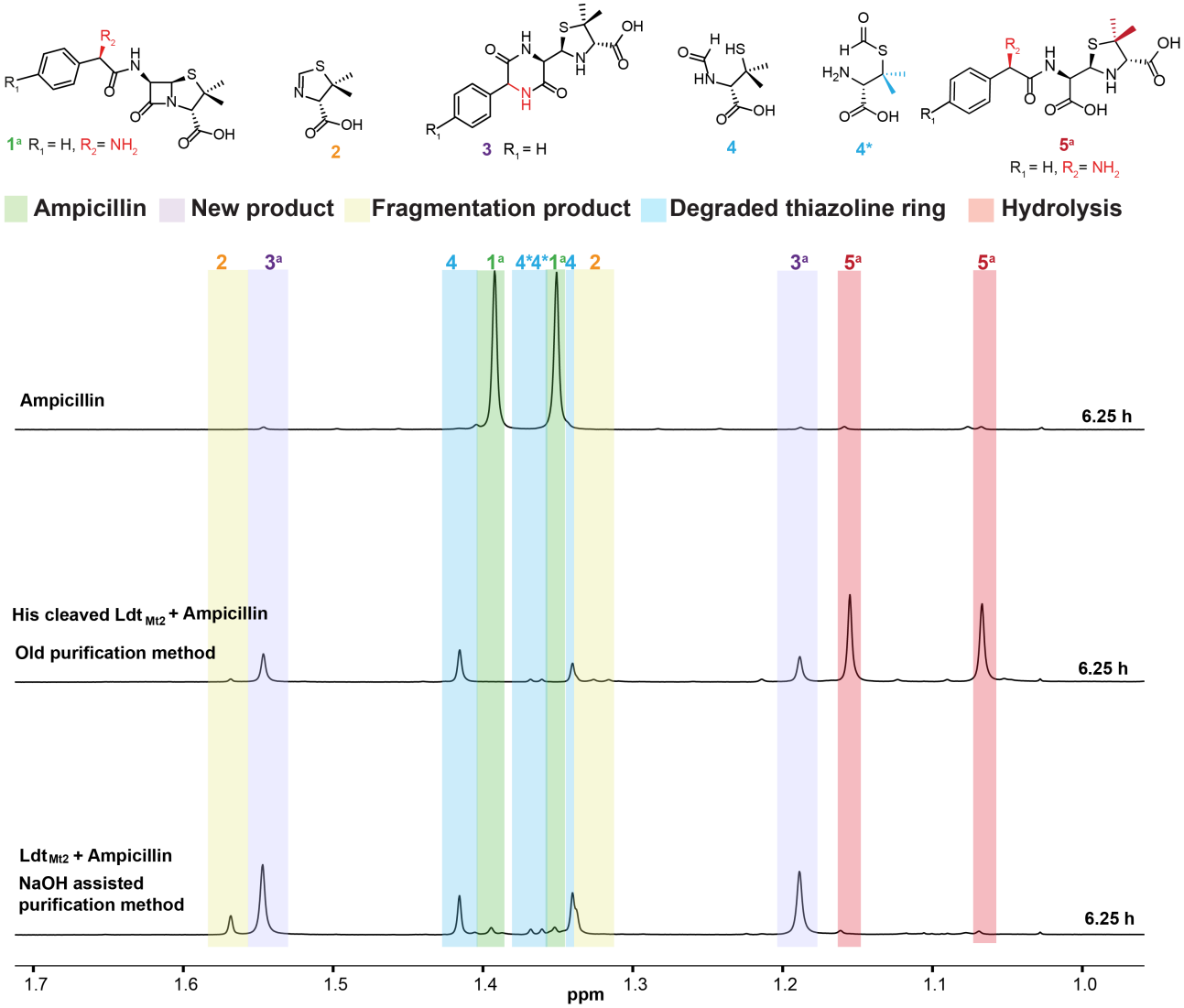


Figure 7.29 | Comparison of ampicillin hydrolysed products from Ldt_{Mt2} purified via two different methods after 6.5 h incubation. The His-tag cleaved Ldt_{Mt2} preparation displayed weak penicillinase activity, as evidenced by the presence of **5** and**5***. No hydrolysis products on incubation of Ldt_{Mt2} purified using the NaOH washed column was observed suggesting the absence of contaminating β-lactamase. Note that with Ldt_{Mt2} purified via the NaOH wash free method, hydrolysed ampicillin product was observed as a major product [228].Conditions: 100 µM Ldt_{Mt2}, 5 mM Ampicillin, 50 mM sodium phosphate, pH 7.5, 6.5 h incubation at r.t. Note that **5*** is the *(5S)* epimer of 5.

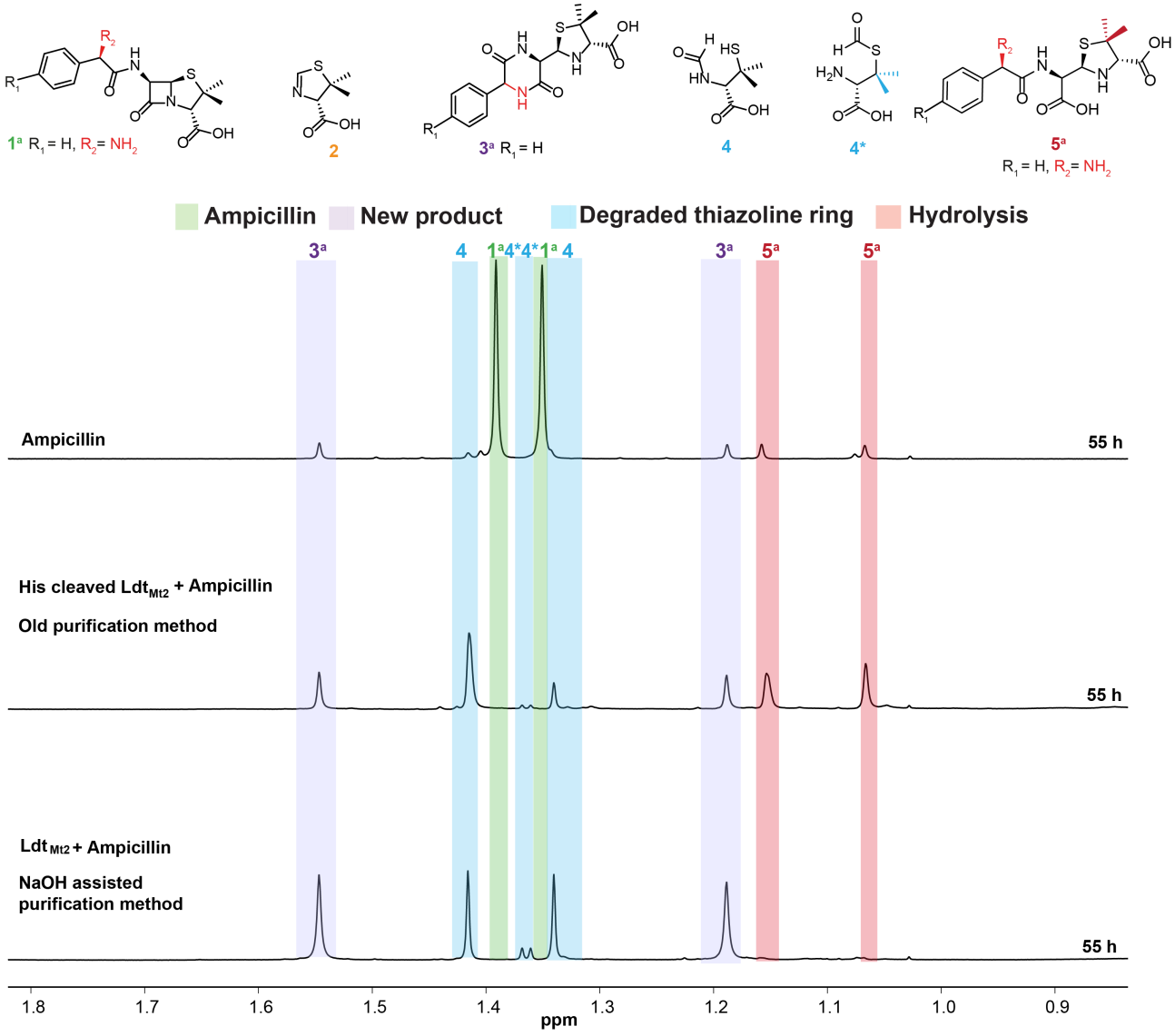


Figure 7.30 | Comparison of ampicillin hydrolysed products from Ldt_{Mt2} purified via two different methods after 55 h incubation. The hexaHis-tag cleaved Ldt_{Mt2} preparation displayed penicillinase activity, as evidenced by presence of **5** and**5***. No hydrolysis products were observed on incubation of Ldt_{Mt2} purified using the NaOH treated column suggesting the absence of contaminating β-lactamase. Note the absence of the fragmentation product, **3** which has fully degraded to **4** or **4***. Conditions: 100 μM Ldt_{Mt2}, 5 mM Ampicillin, 50 mM sodium phosphate, pH 7.5, 55 h incubation at r.t. Note that **5*** is the *(5S)* epimer of 5.

Table 7.8: The chemical shift assignments for the HPLC purified ampicillin derived diketopiperaine product in 10%$_{(v/v)}$ D_2O. Compound is likely a mixture of C5 epimers. Based on the intact structure of ampicillin the *R*-configuration is expected at C13 and C6 positions. Refer to Appendix Figures 7.31, 7.33, 7.34, 7.35 and 7.36.

Position	^{1}H (ppm)	^{13}C (ppm)
2		60.24
3	3.41	76.09
4		
5	5.21	67.75
6	4.08	60.56
7		170.33
8	1.55	25.75
9	1.19	26.31
10		175.19
11		
12		168.65
13	5.29	59.06
14	5.29	136.81
15-19	7.32-7.39	127.34-129.66

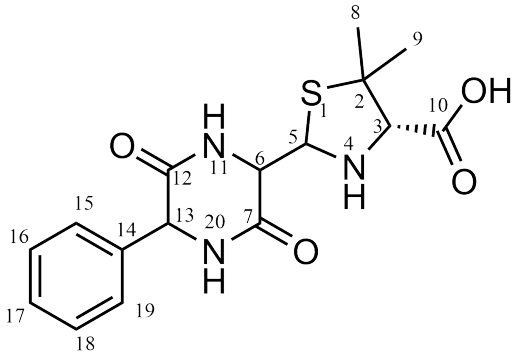


Table 7.9: The chemical shift assignments for ampicillin diketopiperazine sodium salt (mixture of diastereomers) in 50 mM sodium phosphate, pH 7.5, 10%$_{(v/v)}$ D_2O. Refer to Appendix Figures 7.31 and 7.37.

Position	^{1}H (ppm)	^{13}C (ppm)
2		59.71
3	3.39	76.08
4		
5	5.08	67.46
6	4.06	60.07
7		170.33
8	1.52	25.18
9	1.10	25.66
10		175.20
11		
12		169.17
13	5.27	59.10
14	5.29	136.85
15-19	7.33-7.39	127.36-128.66

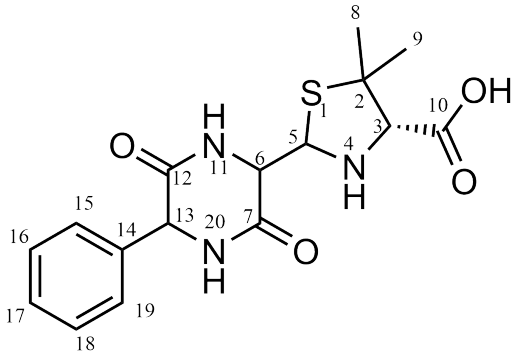
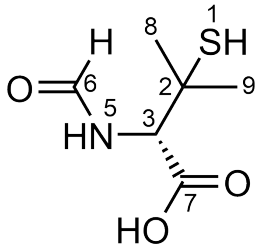


Table 7.10: Chemical shift assignments for the *N*-formyl-D-penicillamine (4) in 90%$_{(v/v)}$ H_2O, 10%$_{(v/v)}$ D_2O. Refer to Appendix Figures 7.39 and 7.40. Note that low levels of formylthiol derivative (**4***) could not be assigned in the crude mixture.

Position	**^{1}H (ppm)**	**^{13}C (ppm)**
2		45.36
3	4.24	63.04
4		
5		
6		175.43
7		
8	1.34	29.59
9	1.42	29.77



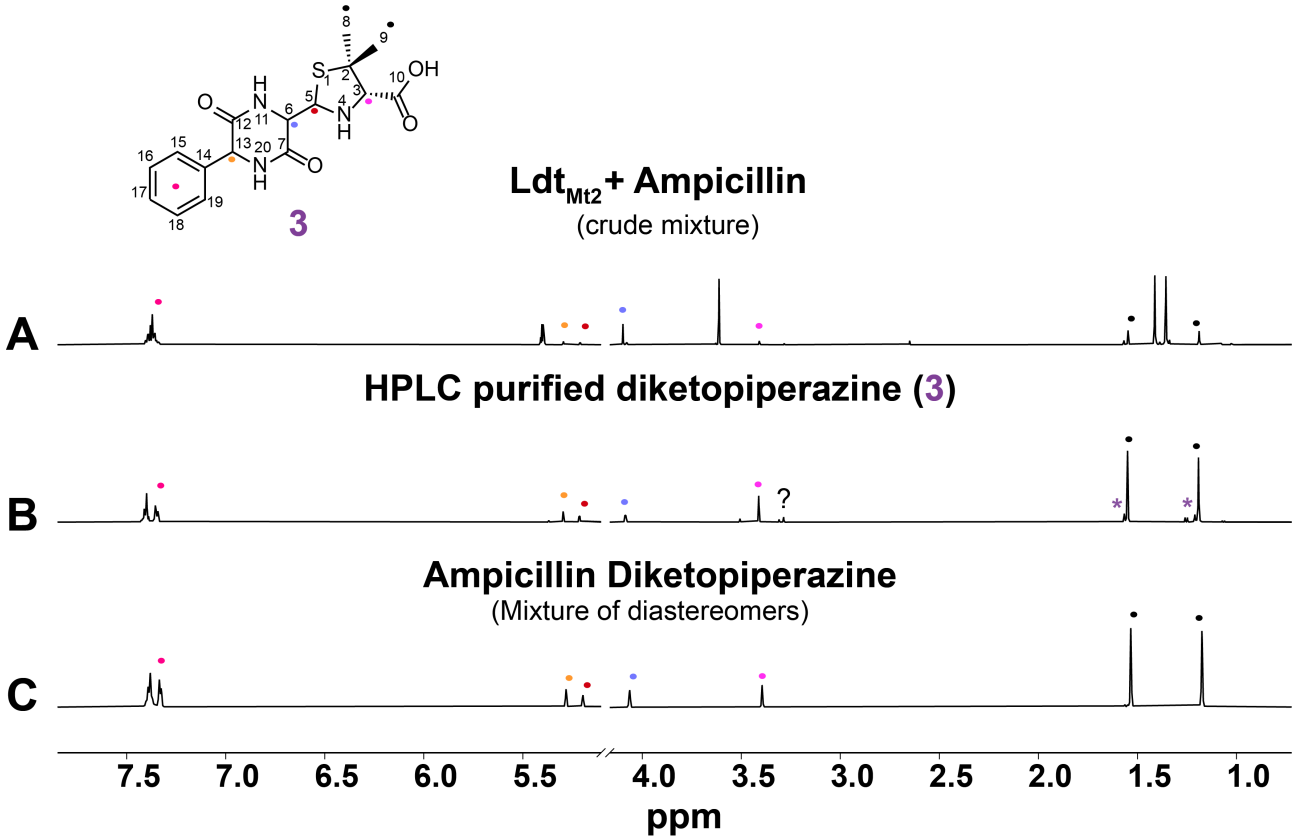


Figure 7.31 | ^{1}H NMR spectra of the ampicillin derived diketopiperazine (3). A. The crude mixture of Ldt_{Mt2} (10 µM) was incubated with ampicillin (2 mM). The major product corresponds to the chemical shifts of methyl peaks of *N*-formyl-D-penicillamine (**4**). **B.** HPLC purified sample consisted of 50 mM ampicillin derived diketopiperazine product (following treatment of 50 mM ampicillin with 20 µM Ldt_{fm}) in 50 mM sodium phosphate, pH 7.5, 10%$_{(v/v)}$ D_2O, and was measured on a 600 MHz NMR spectrometer. Note that the diastereomers (represented by asterisk) are present in the HPLC purified sample in an approximate ratio of 7:1. **C.** The commercial ampicillin diketopiperazine is a mixture of diastereomers- according to the suppliers compound data sheet. The spectrum of the commercial ampicillin diketopiperazine mixture is very similar to that obtained for the major product in the HPLC purified diketopiperazine.

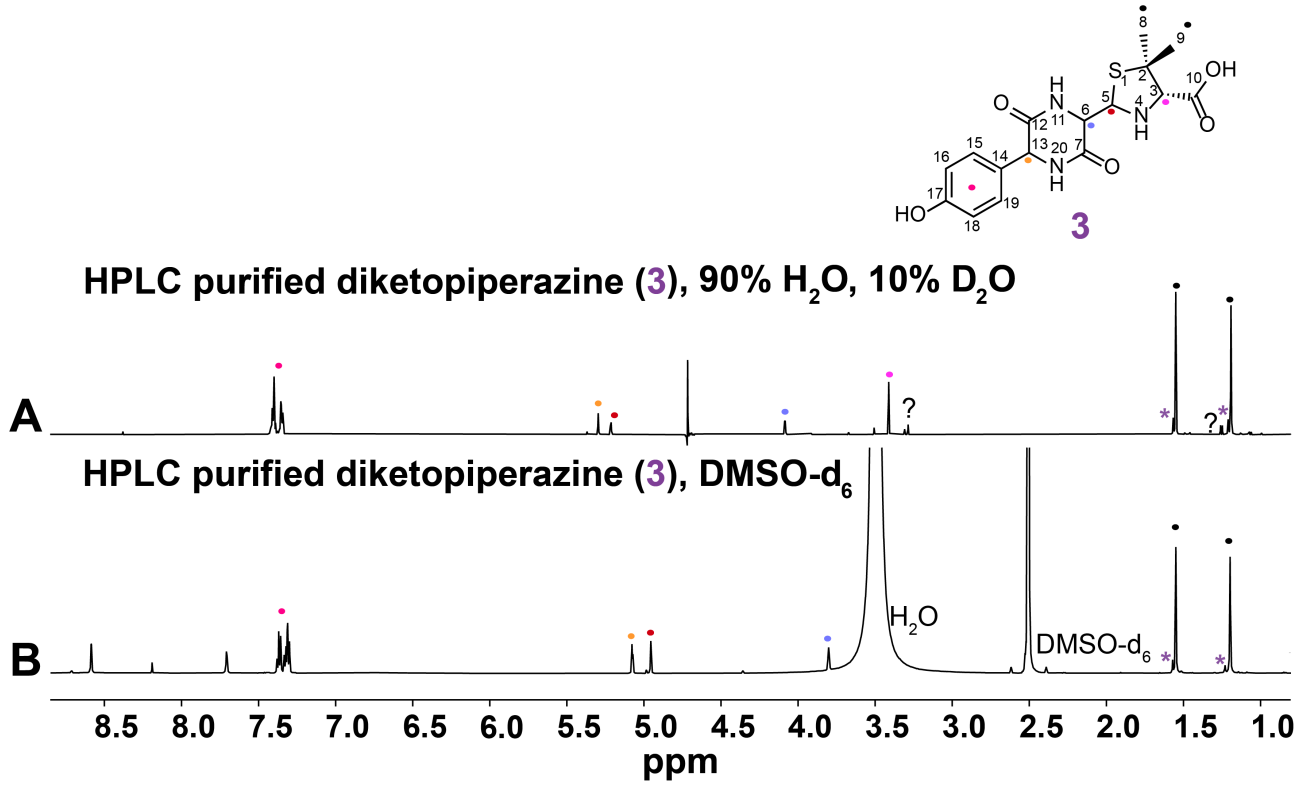


Figure 7.32 | ^{1}H NMR spectra of the HPLC purified ampicillin derived diketopiperazine product. Proton spectrum was acquired in **A.** 90%$_{(v/v)}$ H_2O, 10%$_{(v/v)}$ D_2O or **B.** DMSO-d_6. The peaks only visible around 7.5 ppm - 9 ppm in the bottom spectrum (DMSO-d_6 solvent) correspond to the NH protons which will undergo deuterium exchange in the 90%$_{(v/v)}$ H_2O, 10%$_{(v/v)}$ D_2O solvent.

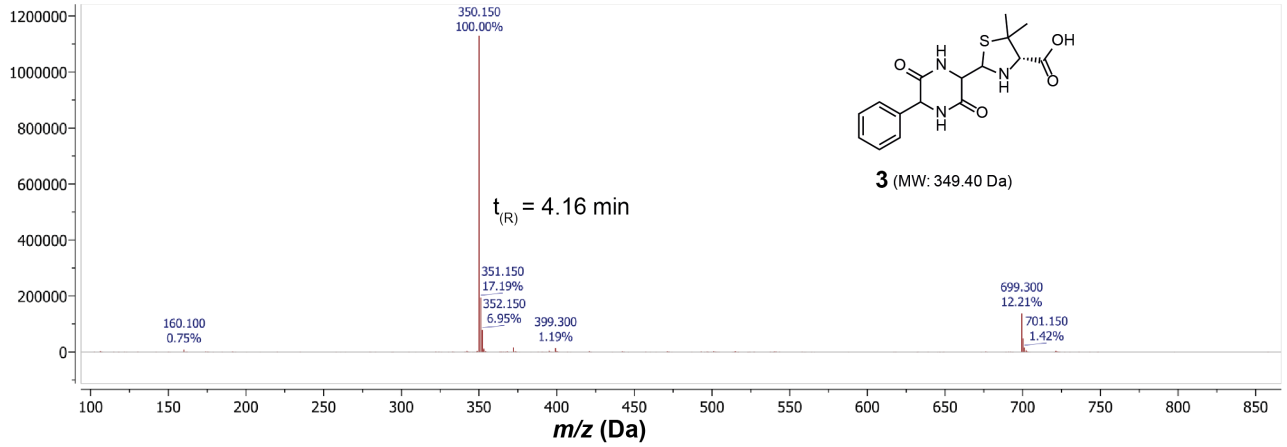


Figure 7.33 | Mass spectrum of the HPLC purified ampicillin derived diketopiperazine (3). Retention time is indicated (4.16 min); the observed mass of 350 Da corresponds to the protonated mass of the expected product.

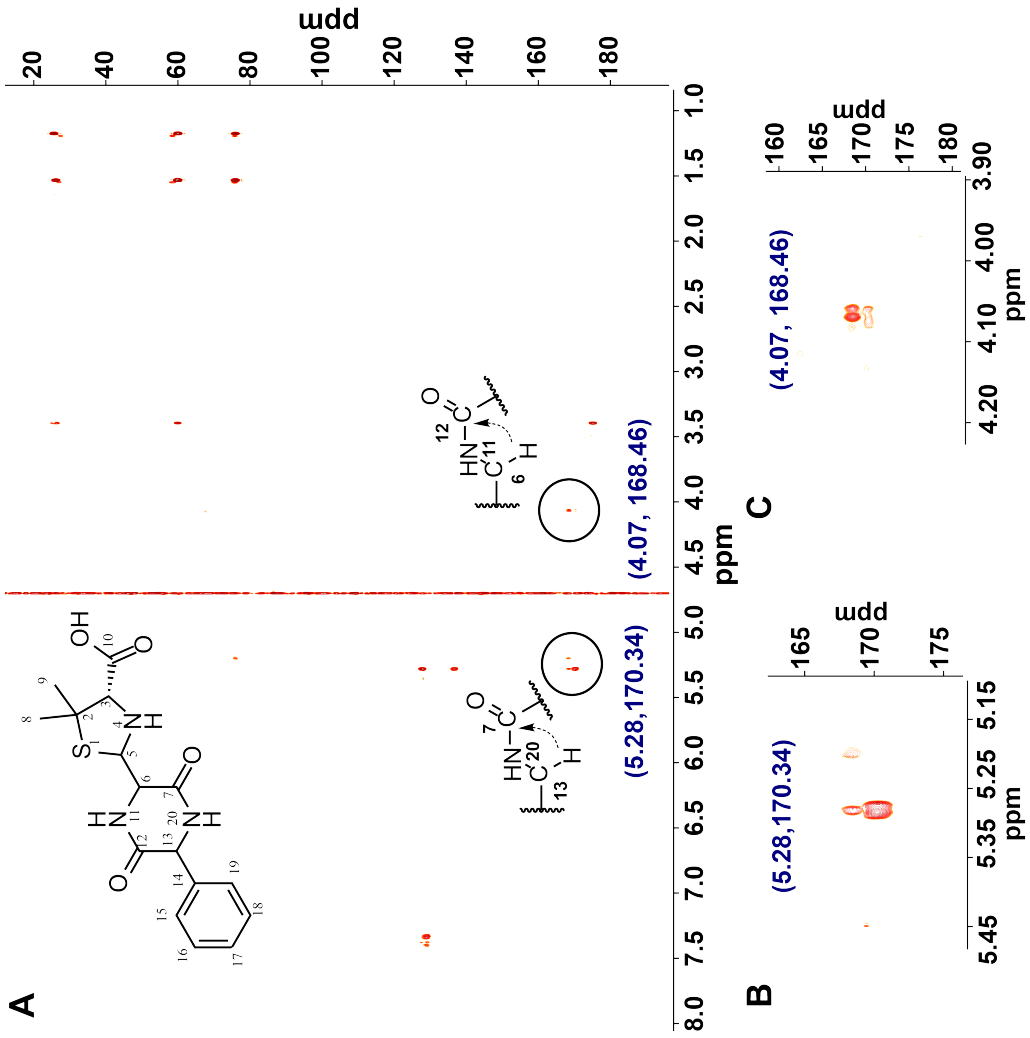


Figure 7.34 | ^{1}H-^{13}C-HMBC spectra of the HPLC purified ampicillin derived diketopiperazine product (6 mM). Evidence for cyclisation to form diketopiperazine ring is observed through the spin transfer of **B.** proton attached to C13 to C7 and **C.** proton attached to C6 to C11. Note that other diastereomers potentially at C5 and C13 were present in an approximate ratio of 7:1.

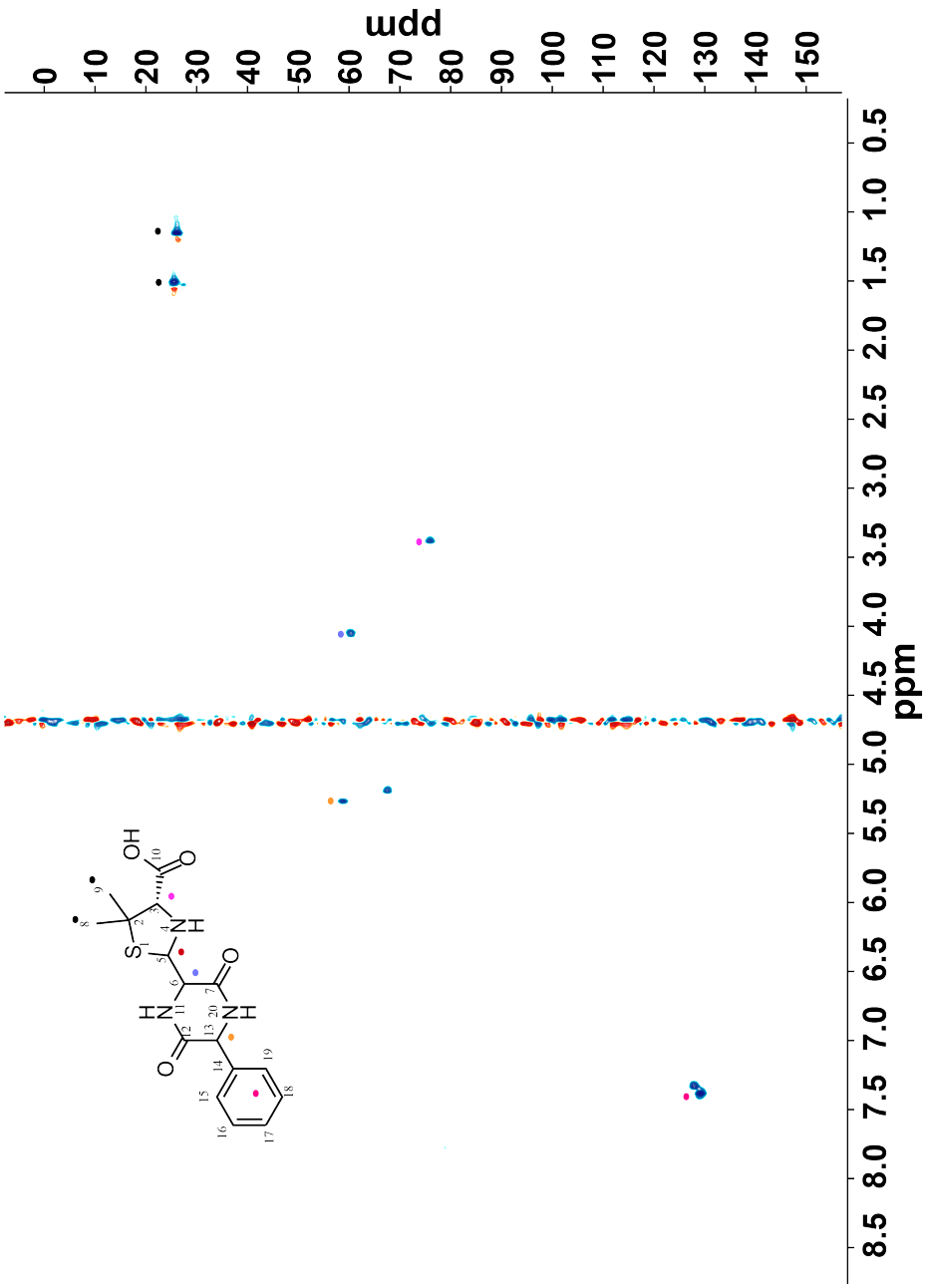


Figure 7.35 | ^{1}H-^{13}C-HSQC spectrum of the HPLC purified ampicillin derived diketopiperazine product (6 mM).

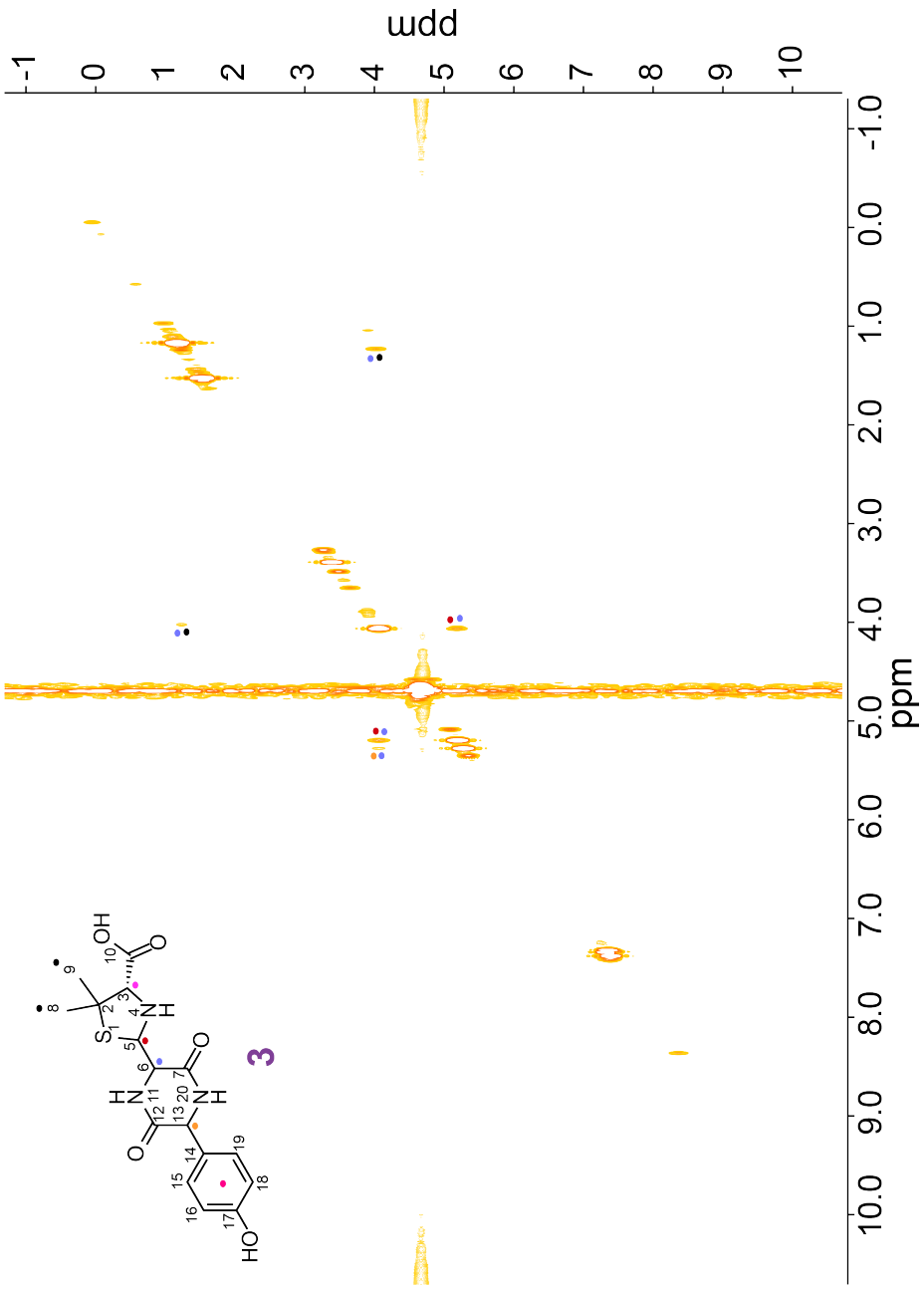


Figure 7.36 | ^{1}H-^{1}H-COSY spectrum of the HPLC purified ampicillin derived diketopiperazine product(6 mM). Evidence for spin coupling of protons attached to C5-C6 and protons attached to C6-C13 were observed.

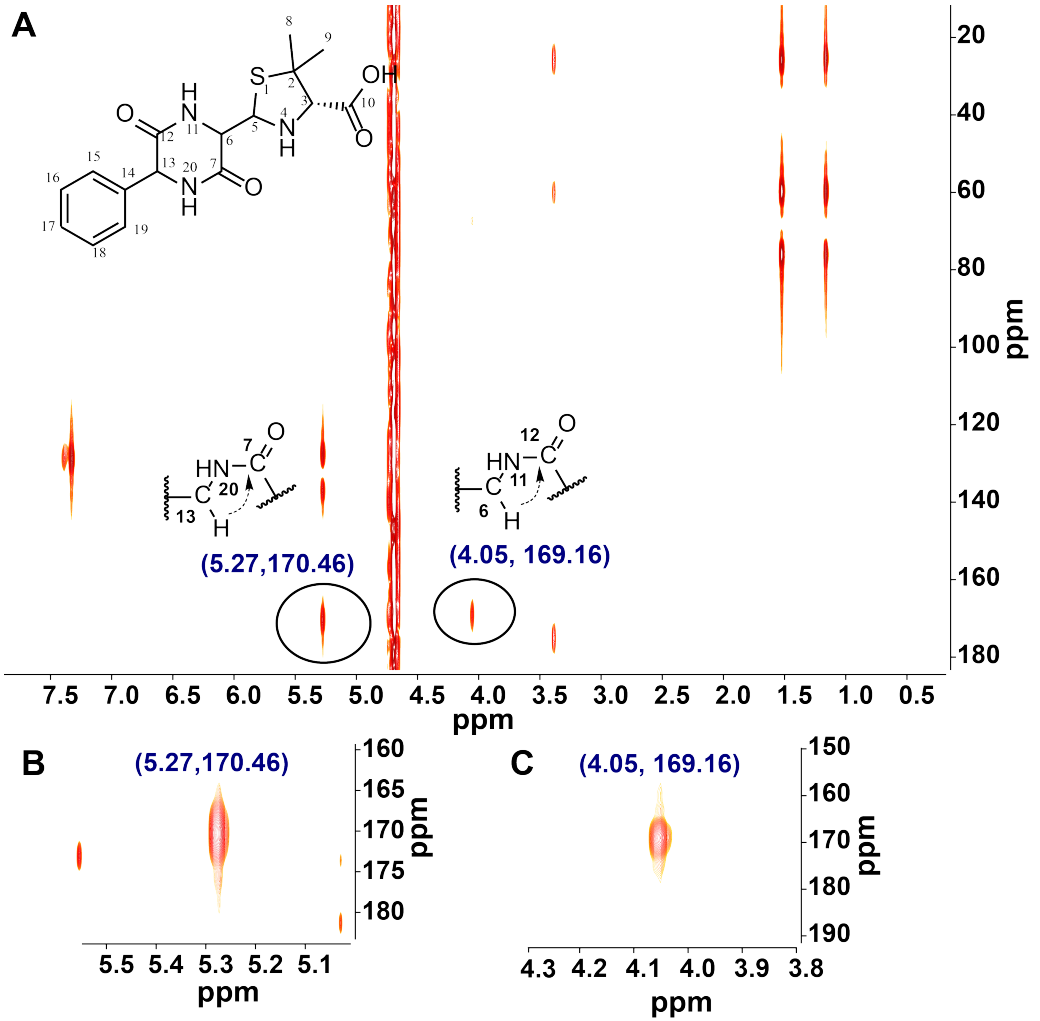


Figure 7.37 | ^{1}H-^{13}C-HMBC spectra of the commercial ampicillin derived diketopiperazine product (6 mM). Evidence for cyclisation to form diketopiperazine ring is observed through the spin transfer of **B.** the proton attached to C13 to C7 and **C.** the proton attached to C6 to C11 through chemical bonds. Note that the commercial stock of ampicillin diketopiperazine was available as a mixture of diastereomers.

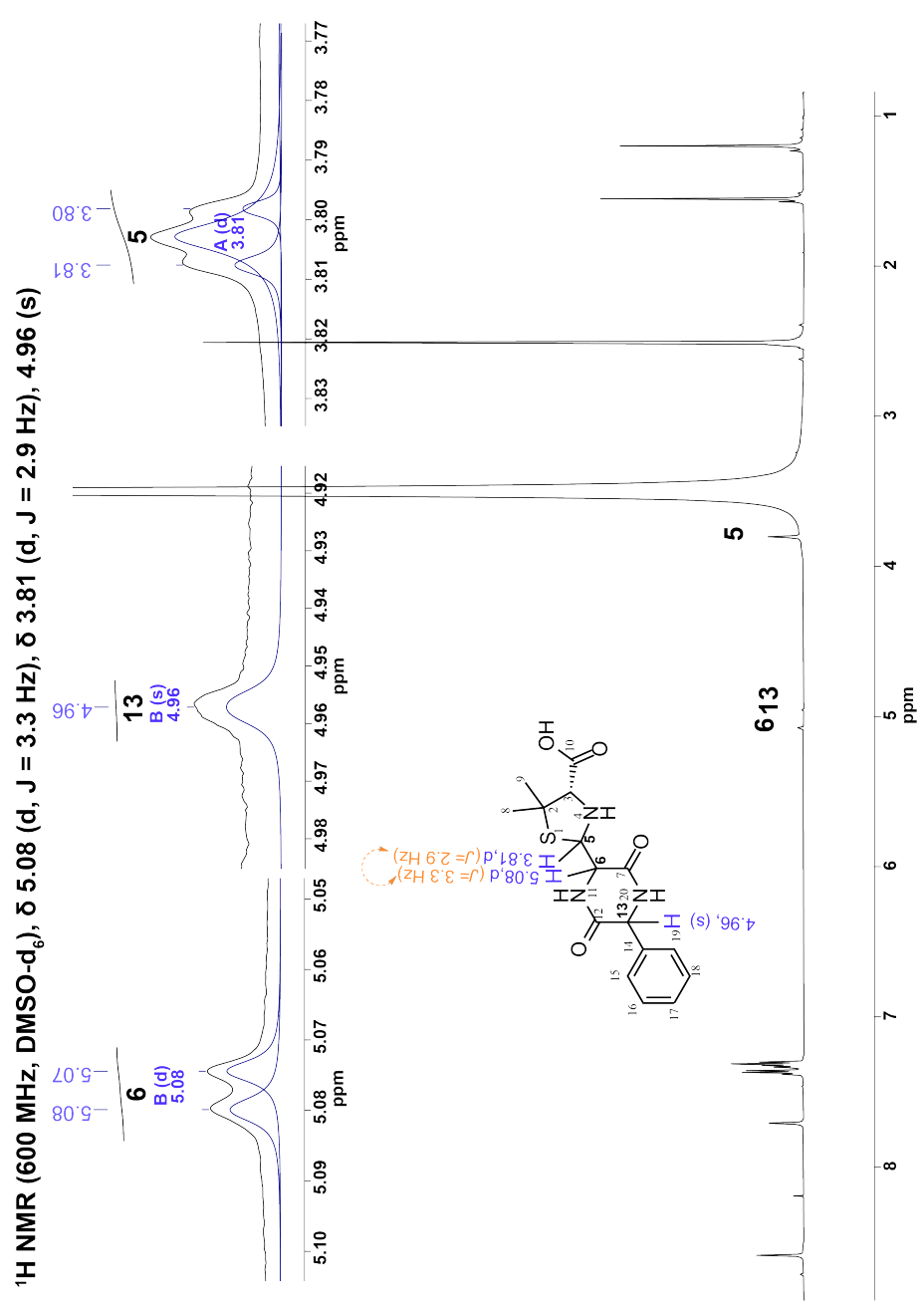


Figure 7.38 | ^{1}H NMR spectrum (600 MHz) of HPLC purified ampicillin derived diketopiperazine product.

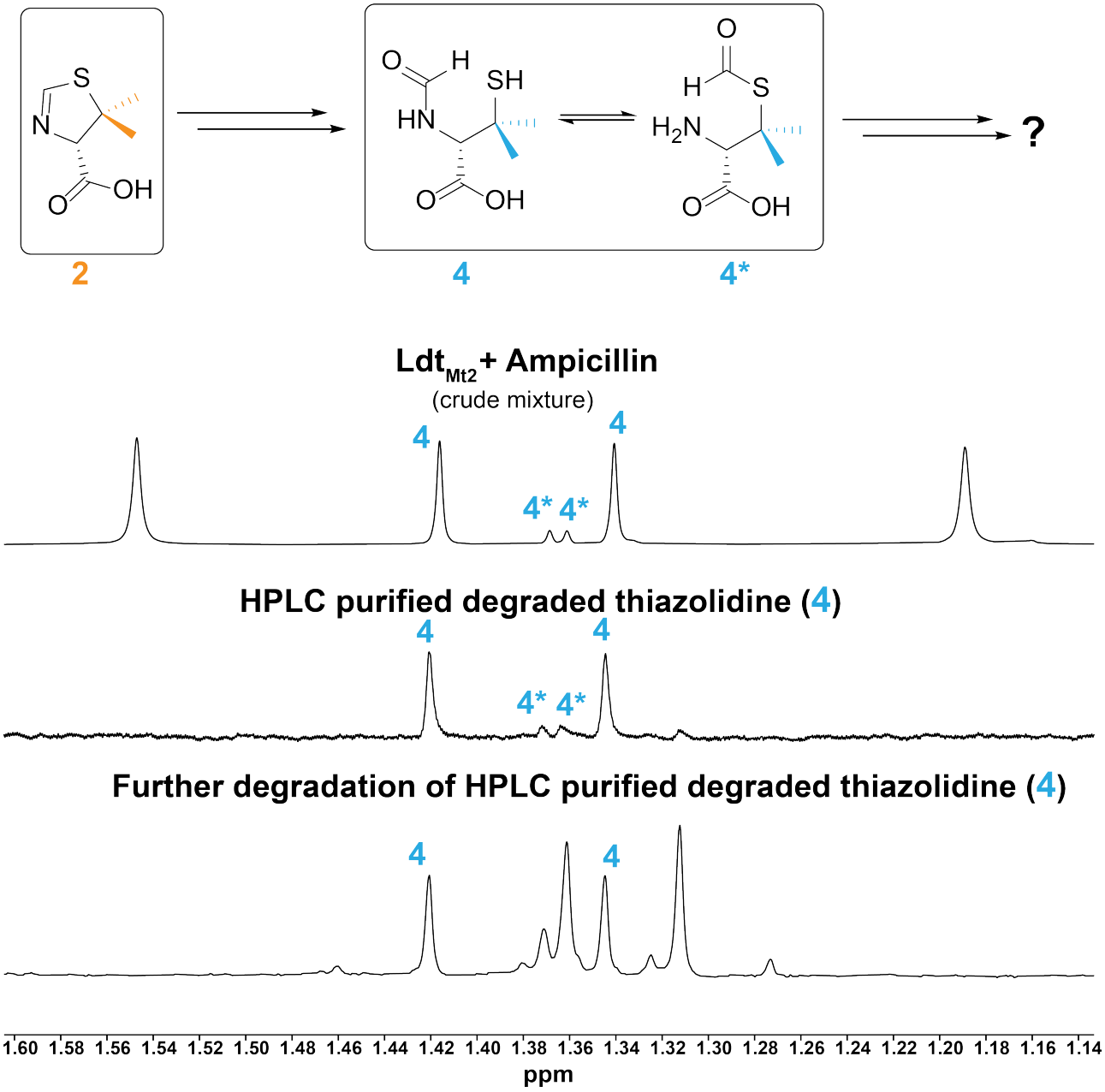


Figure 7.39 | ^{1}H NMR spectra of the HPLC purified ampicillin derived degraded thiazolidine product in 10%$_{(v/v)}$ D_2O. Note that the HPLC purified mixture of **4** and **4*** degraded further whilst performing the characterisation experiments (bottom spectrum), so subsequent characterisation was performed of **4**. See Table 7.10 for assignments for **4**. Note that the low levels of **4*** could not be assigned in the crude mixture. The conversion of **4**→**4*** has been reported [345, 359].

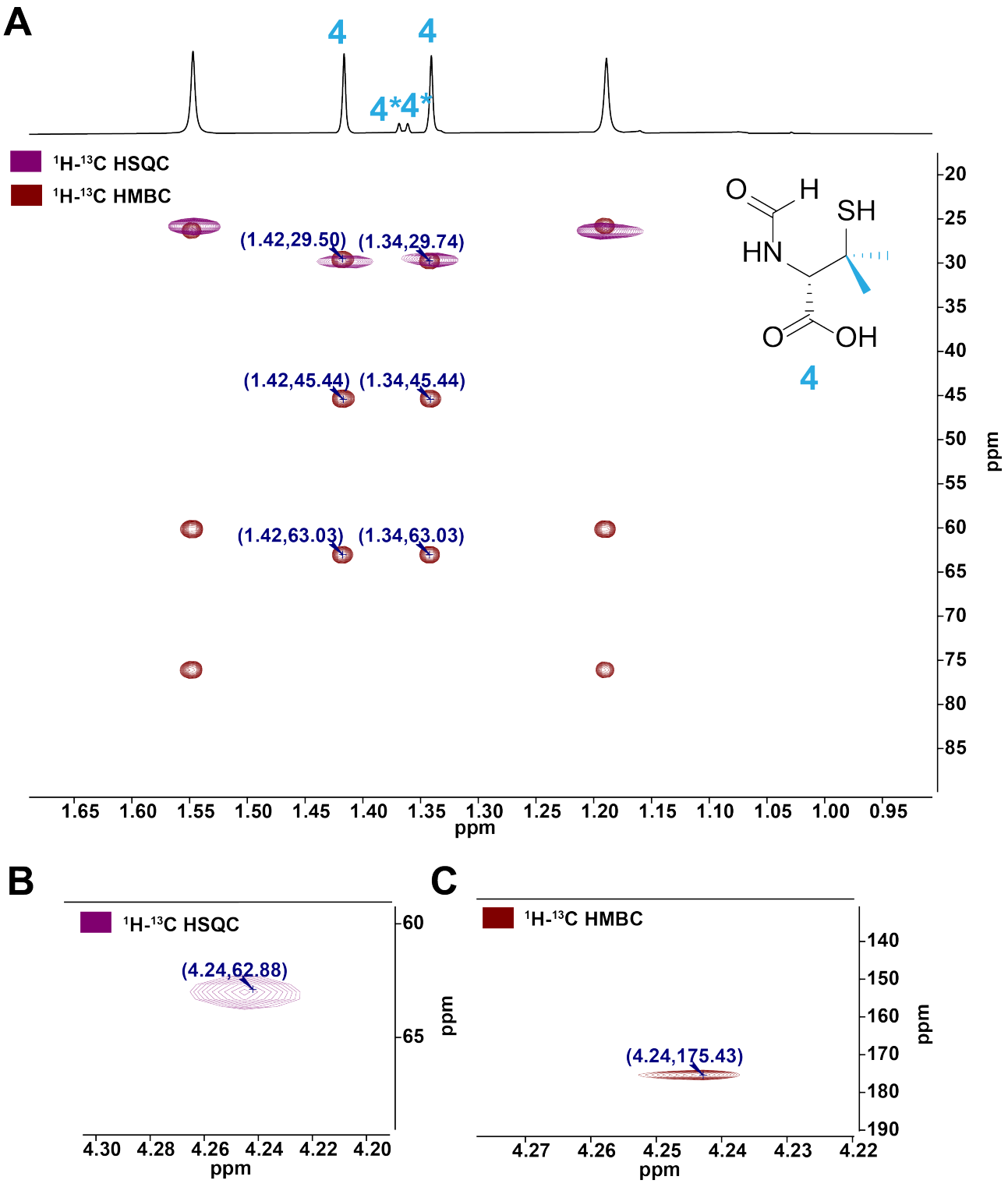


Figure 7.40 | 2D-NMR spectra of the crude ampicillin derived degraded thiazolidine product in 90%$_{(v/v)}$ H_2O, 10%$_{(v/v)}$ D_2O. A. ^{1}H-^{13}C HMBC spectrum (red) superimposed with the ^{1}H-^{13}C HSQC spectrum (purple) shows spin-transfer of the methyl protons to each other and to C2 and C3. **B.** ^{1}H-^{13}C HSQC crosspeak representing spin transfer between C3 and its attached proton. **C.** ^{1}H-^{13}C HMBC crosspeak corresponding to the spin transfer between the C3 proton and the *N*-formyl carbonyl carbon.

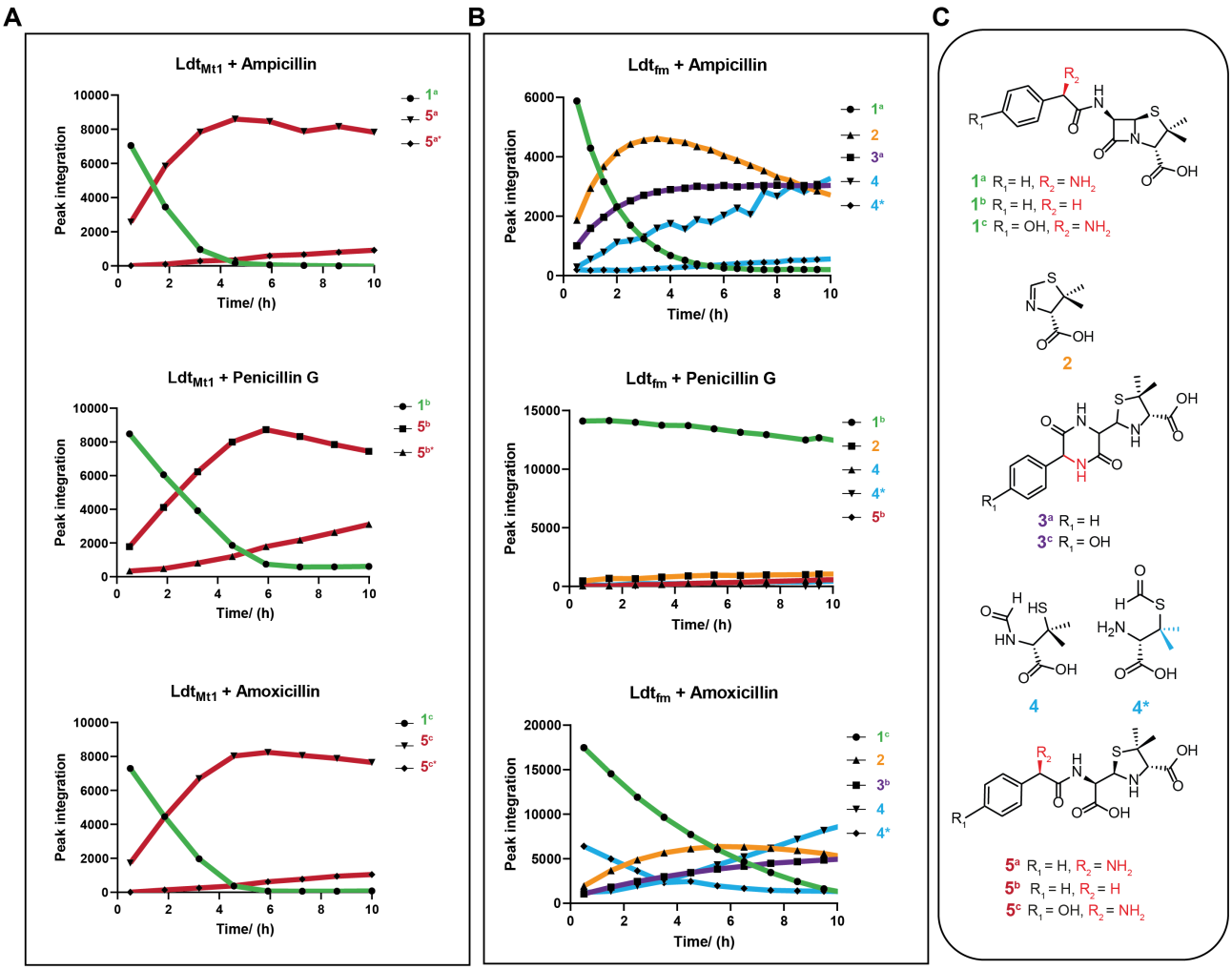


Figure 7.41 | Peak integration of ^{1}H NMR spectra (600 MHz) for the reaction of Ldt$_{Mt1}$ and Ldt$_{fm}$ with ampicillin, penicillin G and amoxicillin. Note that **4*** corresponds to formylthio derivative of *N*-formyl-D-penicillamine (**4**) and **5*** represents the C5 epimer of **5**. Spectra (32 scans), stacked from bottom to top, were acquired every 2.5 min.

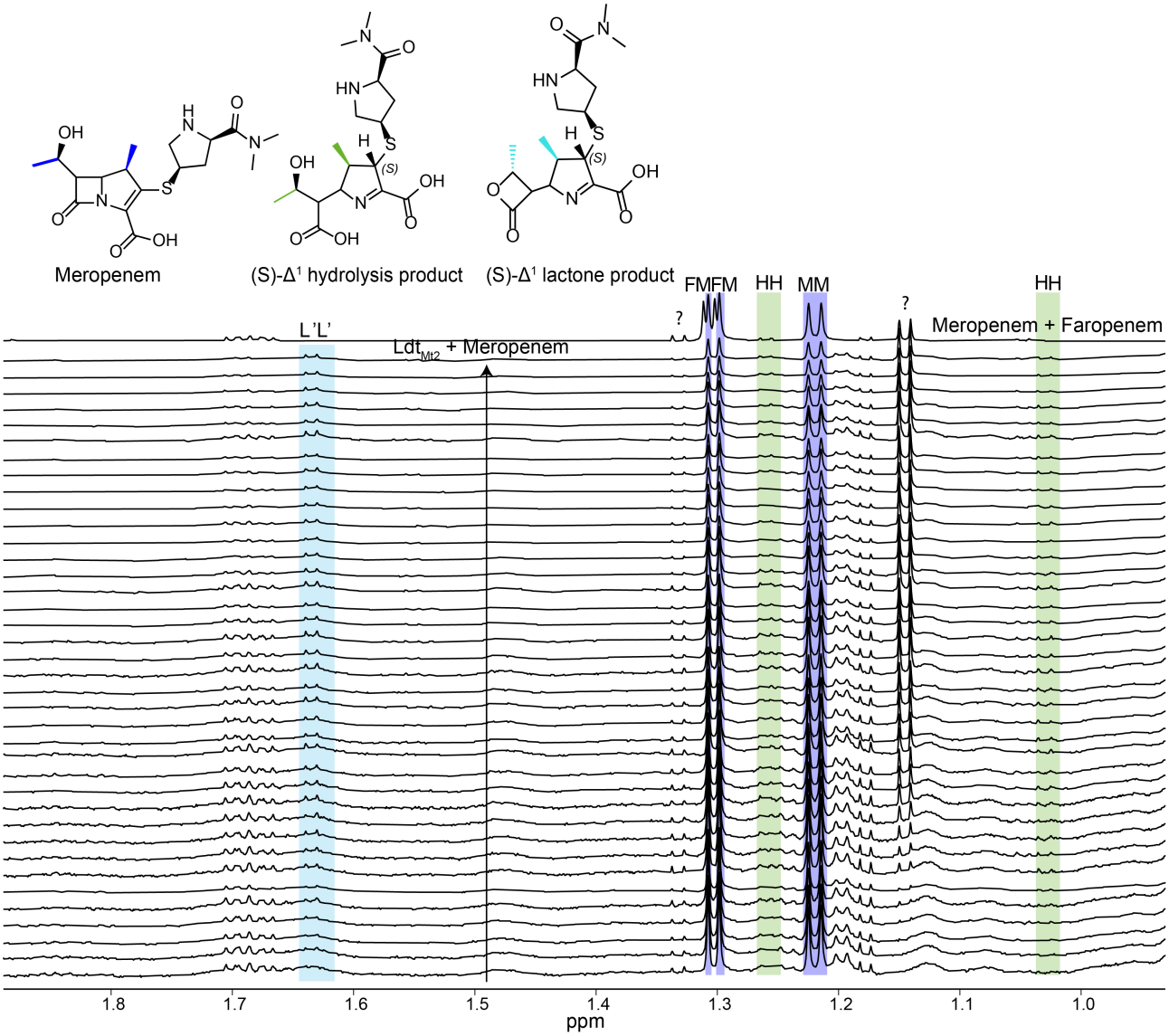


Figure 7.42 | ^{1}H NMR spectra (700 MHz) for the reaction of meropenem (50 µM) with Ldt$_{Mt2}$ (50 µM). A question mark indicates an unassigned peak, which was also observed with enzyme free meropenem and faropenem mixture (refer to Figure 7.43). Spectra (32 scans), stacked from bottom to top, were acquired every 26 min. L'L', FF, MM and HH correspond to the methyl resonances from *(S)*-Δ^1 lactone, faropenem, meropenem, and meropenem derived *(S)*-Δ^1 hydrolysis product, respectively.

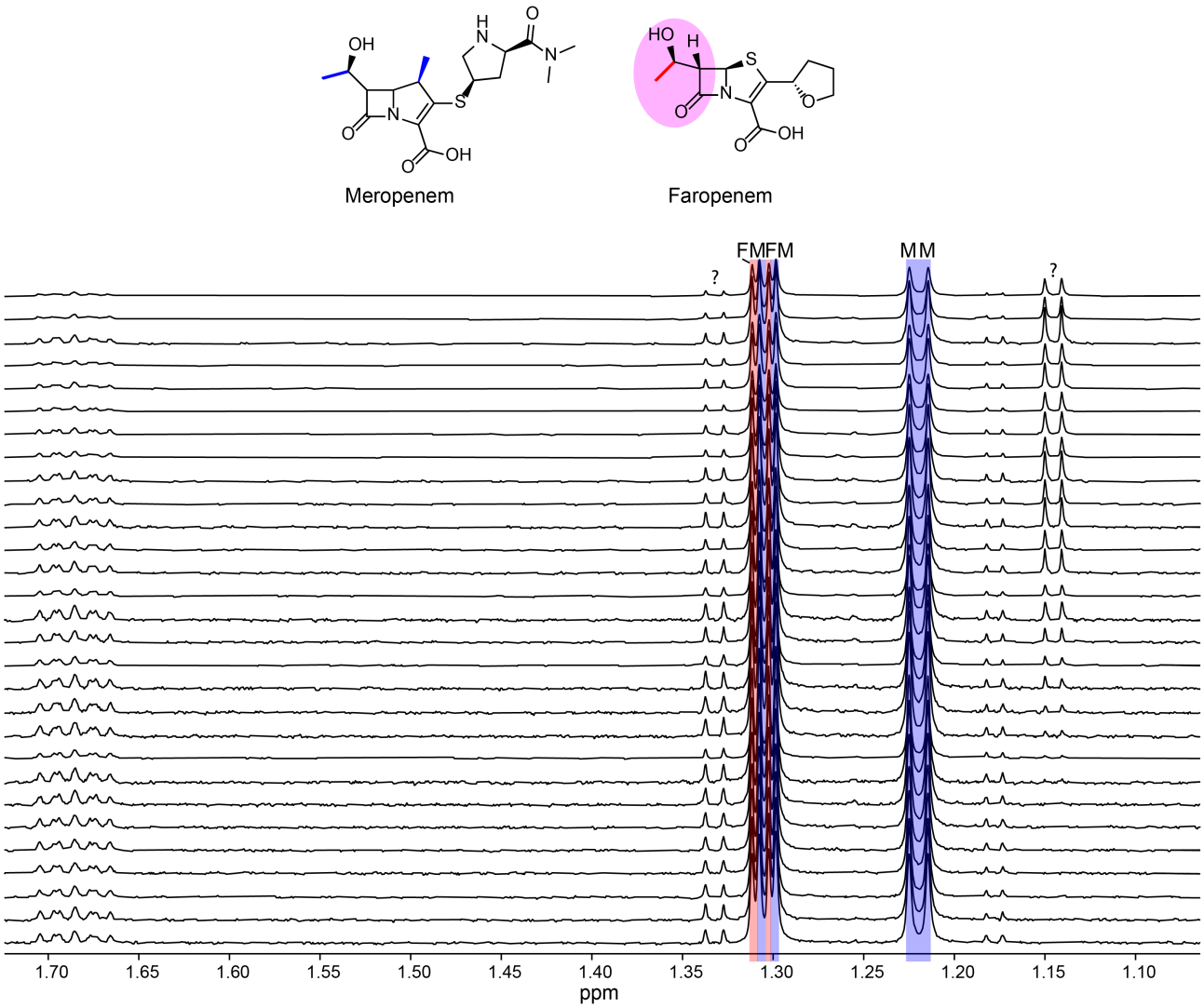


Figure 7.43 | ^{1}H NMR spectra (700 MHz) on for the time dependent reaction of meropenem (50 µM) and faropenem (50 µM). A question mark indicates the unassigned peak at ∼1.32 ppm and ∼1.15 ppm. Spectra (32 scans), stacked from bottom to top, were acquired every 26 min. FF and MM correspond to the methyl resonances from the faropenem and meropenem, respectively.

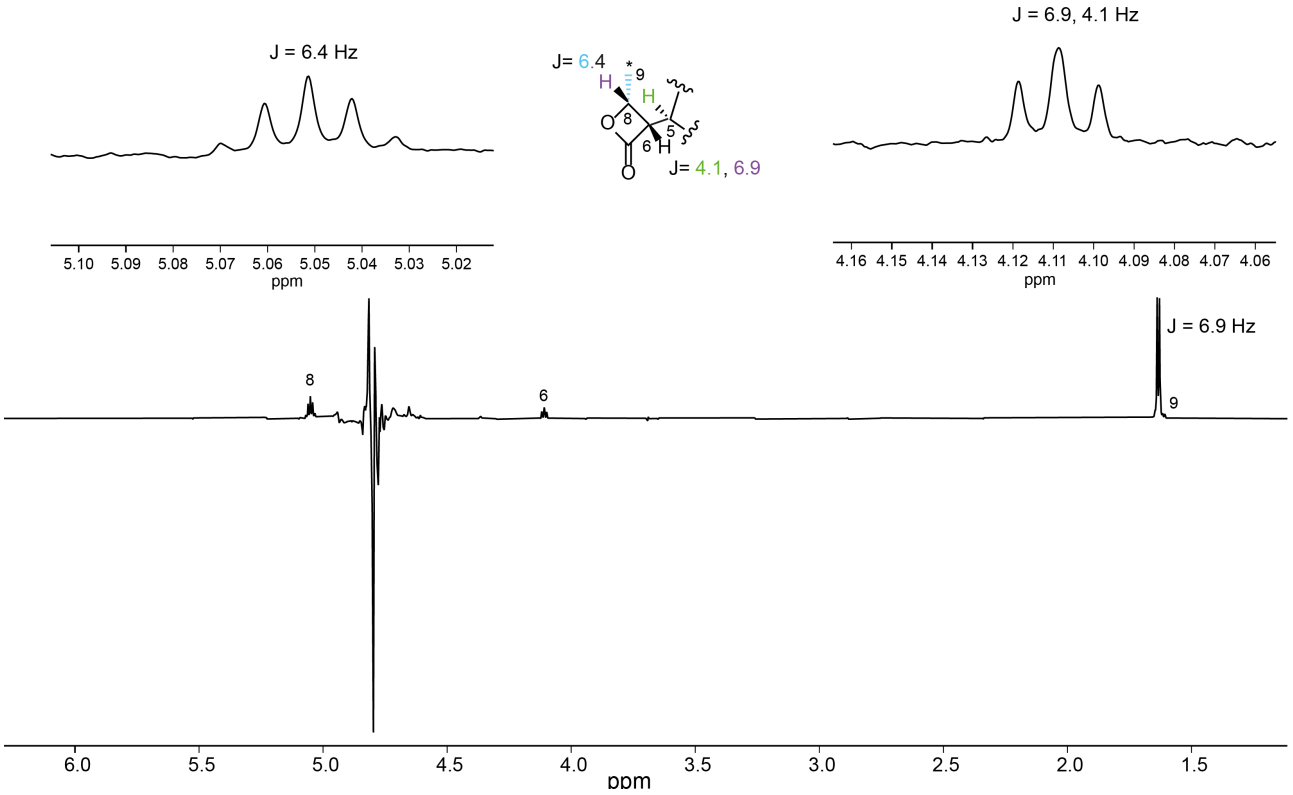


Figure 7.44 | 1D TOCSY NMR spectra (700 MHz) of the new product observed upon incubation of Ldt_{fm} (200 µM) with meropenem (600 µM). Long range spin-coupling of the methyl proton was observed with protons attached to C6 and C8 supporting the proposal that this product is likely to be a meropenem derived *(S)*-Δ^1 lactone. As observed with some SBLs including OXA-48, meropenem reacts to give the *(S)*-Δ^1 lactone (amongst other products) [364].

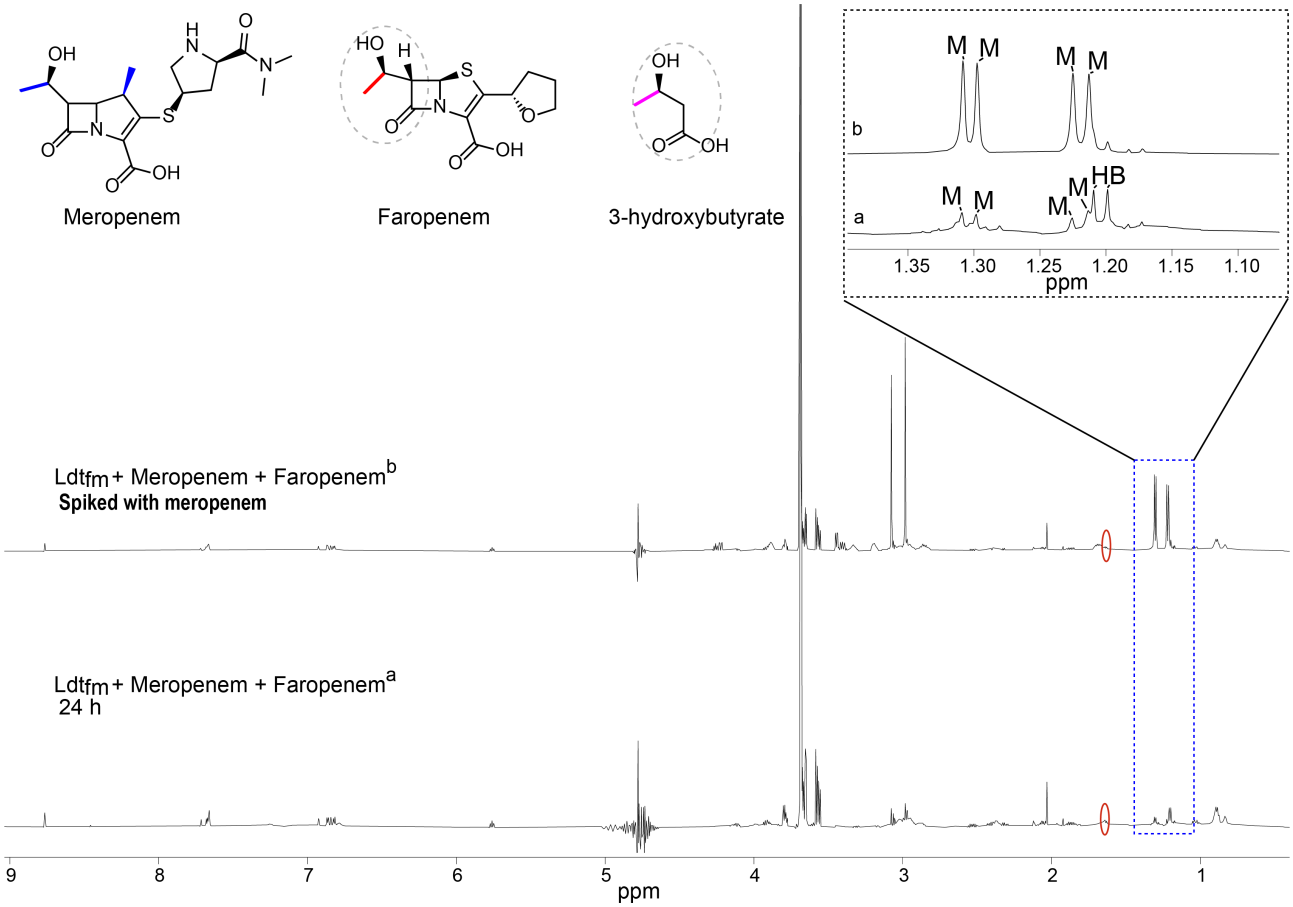


Figure 7.45 | ^{1}H NMR spectra (700 MHz) show evidence that the newly formed methyl peaks in the faropenem treated Ldt_{fm}-meropenem mixture correspond to the methyl peaks of meropenem. Low level peak which may correspond to the lactone product is circled in red. Spectra (32 scans), stacked from bottom to top, were acquired every 26 minutes. Conditions: 200 µM Ldt_{fm}, 100 µM meropenem, 200 µM faropenem (24 h incubation) followed by the addition of excess meropenem (1 mM) in 50 mM sodium phosphate, pH 7.5, 10%$_{(v/v)}$ D_2O.

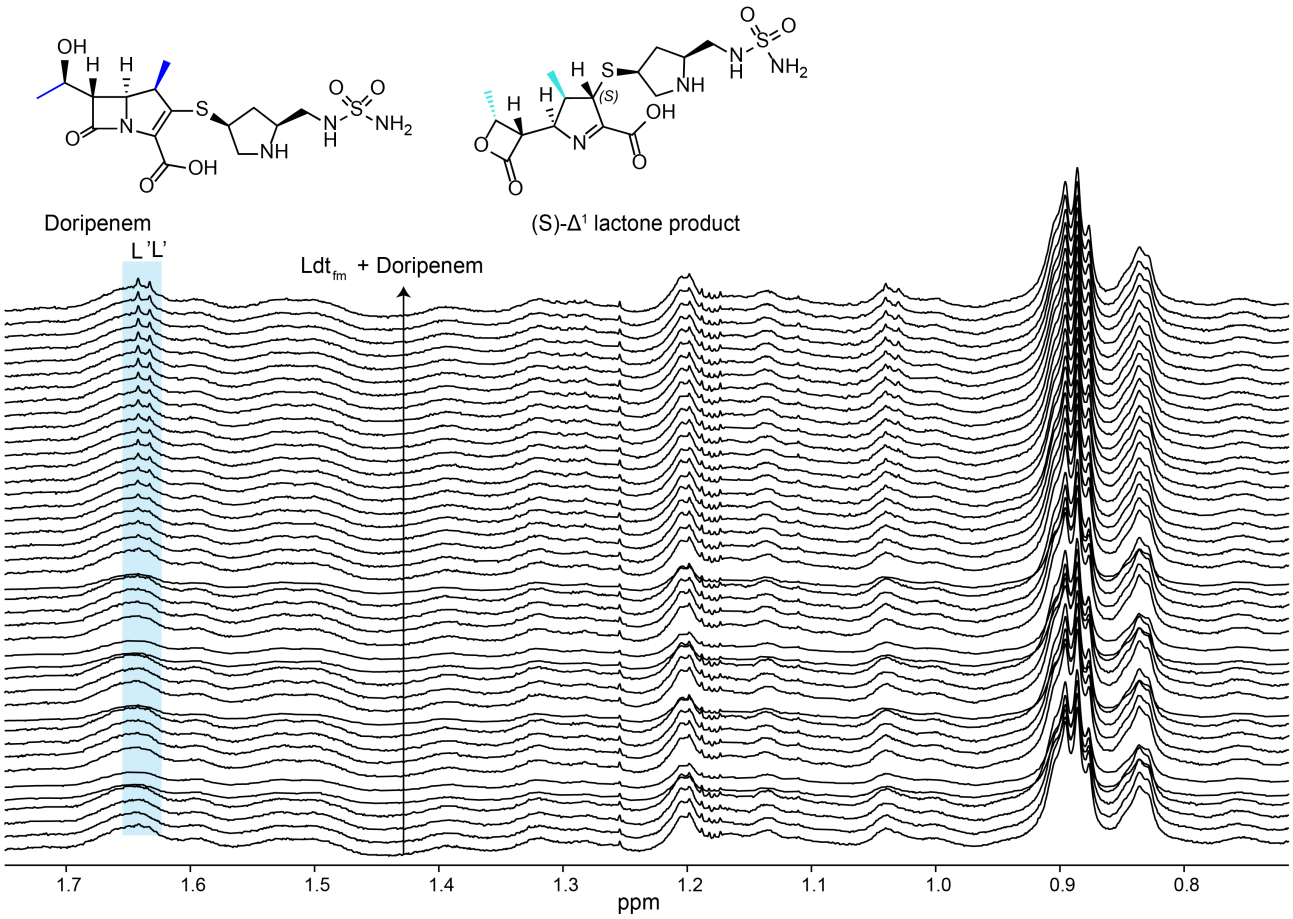


Figure 7.46 | ^{1}H NMR spectra (700 MHz) for the reaction of doripenem (100 µM) with Ldt$_{fm}$ (200 µM). Spectra (32 scans), stacked from bottom to top, were acquired every 26 minutes. L'L' corresponds to the methyl resonances from *(S)*-Δ^1 lactone.

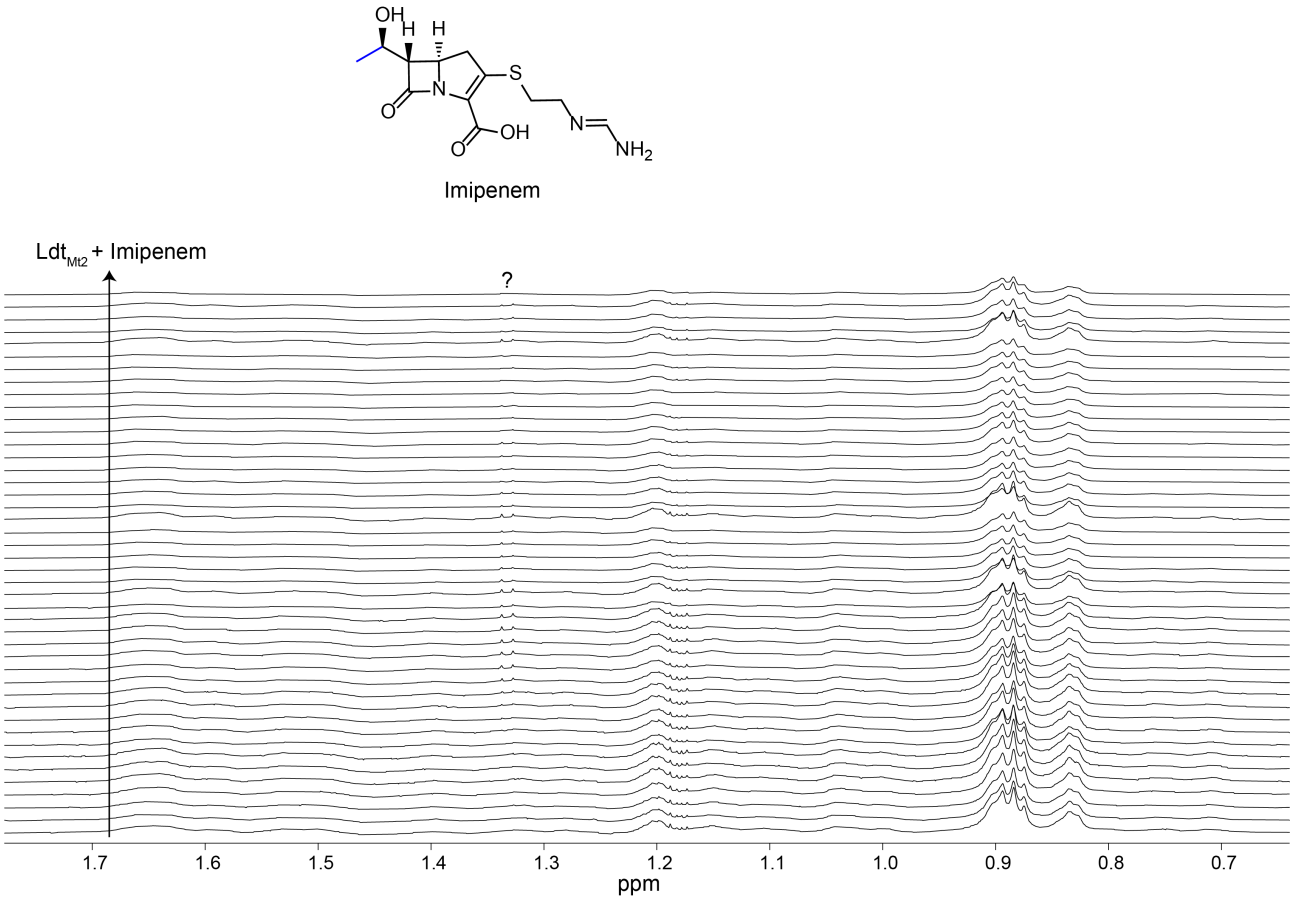


Figure 7.47 | ^{1}H NMR (700 MHz) evidence shows that the reaction of imipenem (100 µM) with Ldt_{fm} (200 µM) does not yield a lactone product. No peak corresponding to *(S)*-Δ^1 lactone product was observed (in the expected region of 1.64 ppm). Note the presence of an unassigned peak around 1.32 ppm. Spectra (32 scans), stacked from bottom to top, were acquired every 26 minutes.

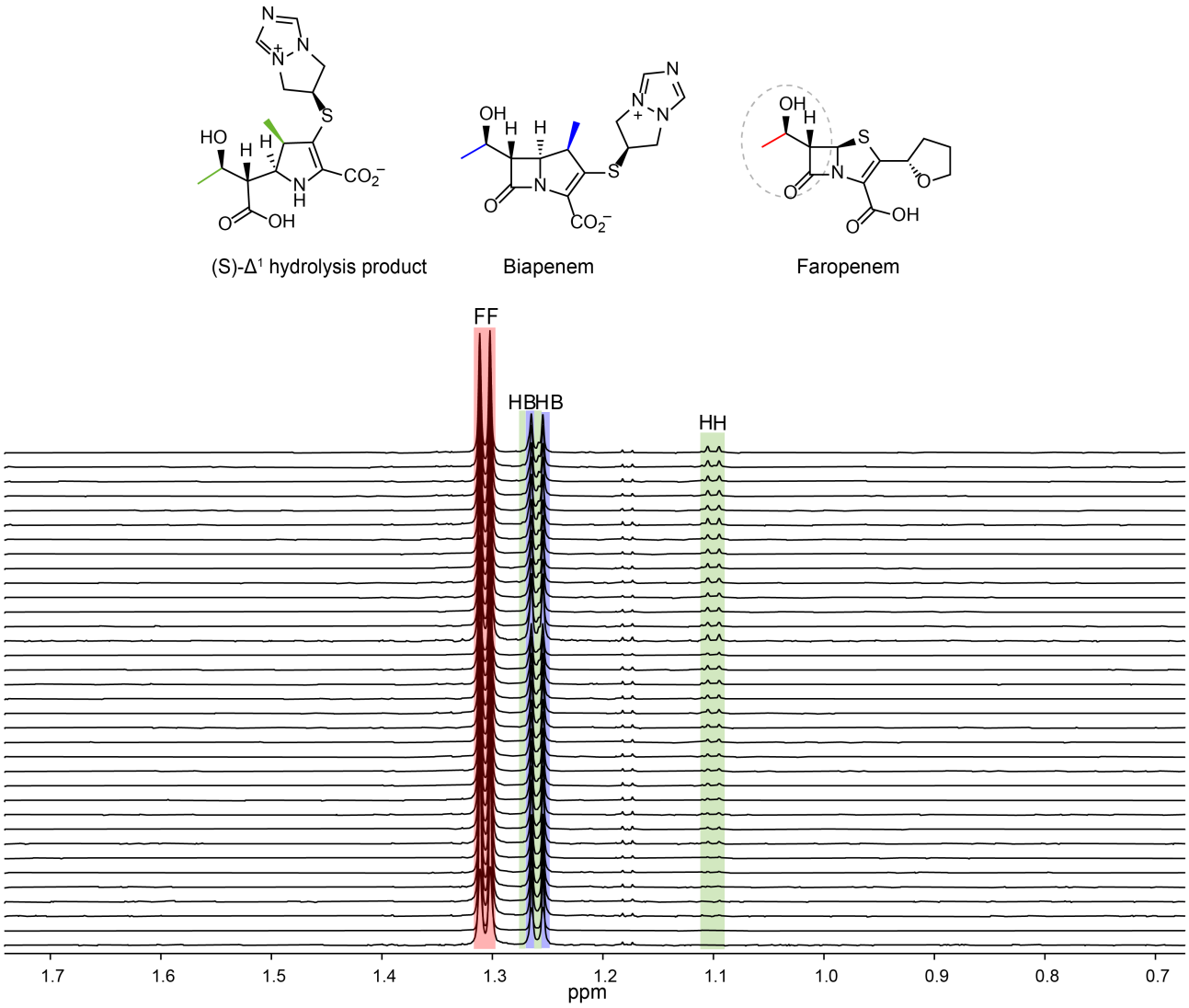


Figure 7.48 | ^{1}H NMR spectra (700 MHz) on the hydrolysis of the biapenem (100 µM)-faropenem (200 µM) mixture. Note the presence of an unassigned peak at 1.32 ppm. Spectra (32 scans), stacked from bottom to top, were acquired every 26 minutes. FF, HH, and BB correspond to the methyl resonances of intact faropenem, *(S)*-Δ^1 hydrolysis product and biapenem, respectively. Note that β-methyl resonance of biapenem overlaps with the methyl resonance of faropenem.

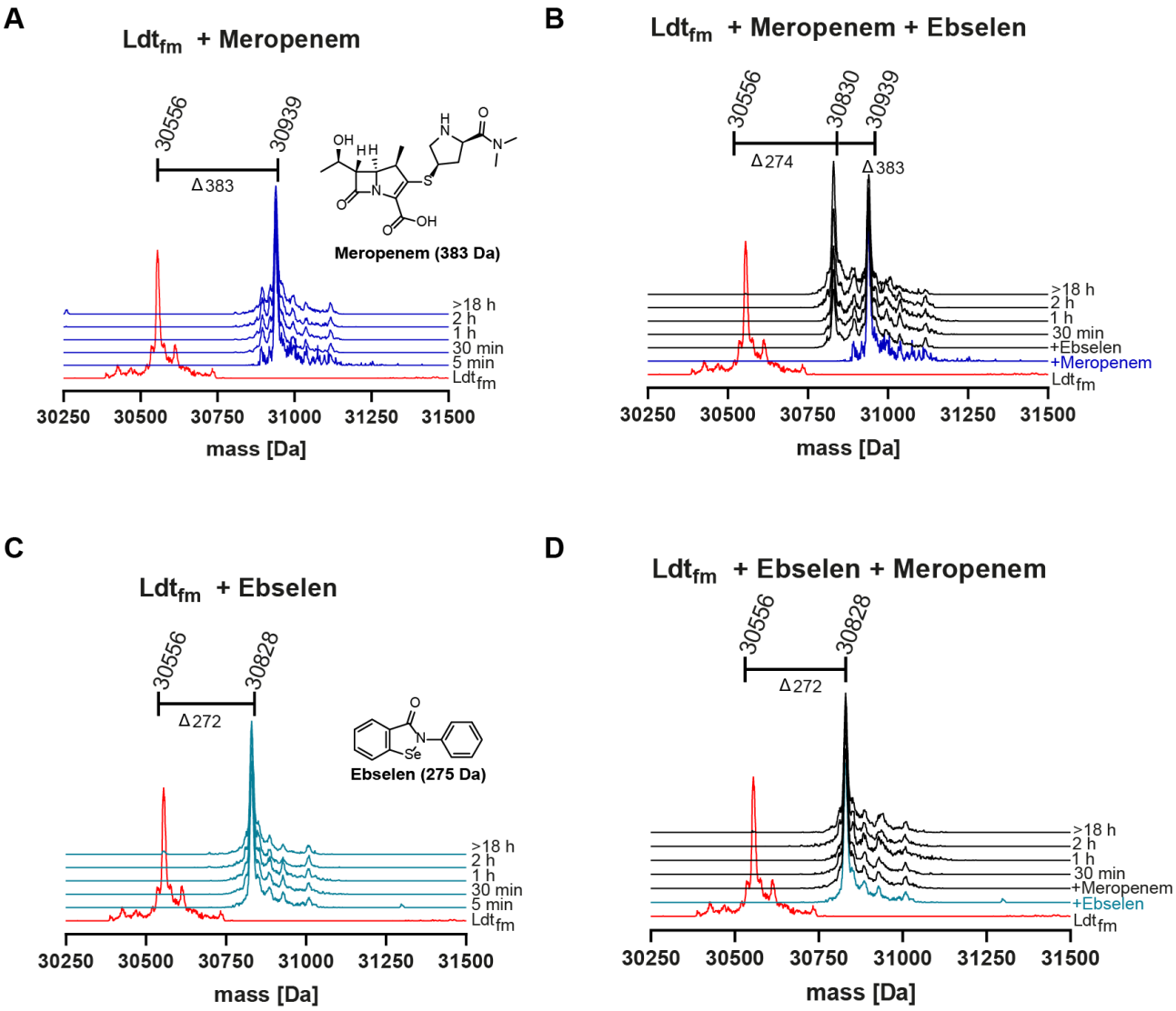


Figure 7.49 | Ebselen irreversibly reacts with the Ldt_{fm}-meropenem complex. **A.** Reaction of Ldt_{fm} (5 µM) with meropenem (50 µM). **B.** An adduct corresponding to the mass of the intact ebselen was observed over time upon treatment of Ldt_{fm}-meropenem mixture with ebselen (100 µM). **C.** Reaction of Ldt_{fm} (5 µM) with ebselen (100 µM). **C.** Ldt_{fm}-ebselen complex is stable over time. **D.** Addition of meropenem to Ldt_{fm}-ebselen does not result in displacement of ebselen from the Ldt_{fm}-ebselen complex. The spectra are stacked from bottom to top and were acquired every 1.35 min.

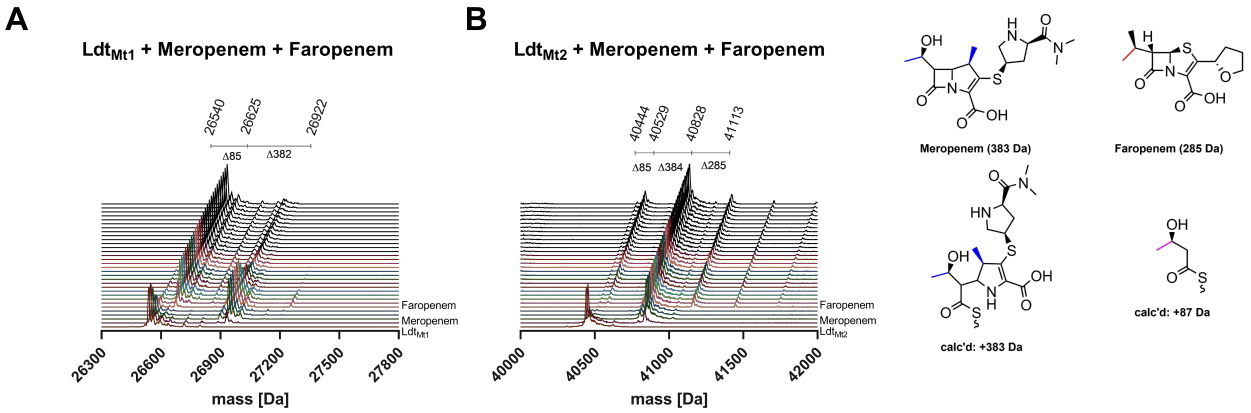


Figure 7.50 | SPE-MS evidence for the reversible reaction for meropenem with Ldt_{Mt1} and Ldt_{Mt2}. **A.** The level of the Ldt_{Mt1}-meropenem complex decreased by reaction with faropenem, accompanied by the increase in peak corresponding to the Ldt_{Mt1}-faropenem complex. **B.** The level of the Ldt_{Mt2}-meropenem complex (slowly) decreased by reaction with faropenem, accompanied by the increase in peak corresponding to the Ldt_{Mt2}-faropenem complex. The spectra are stacked from bottom to top and subsequent spectrum was acquired every 1.4 min.

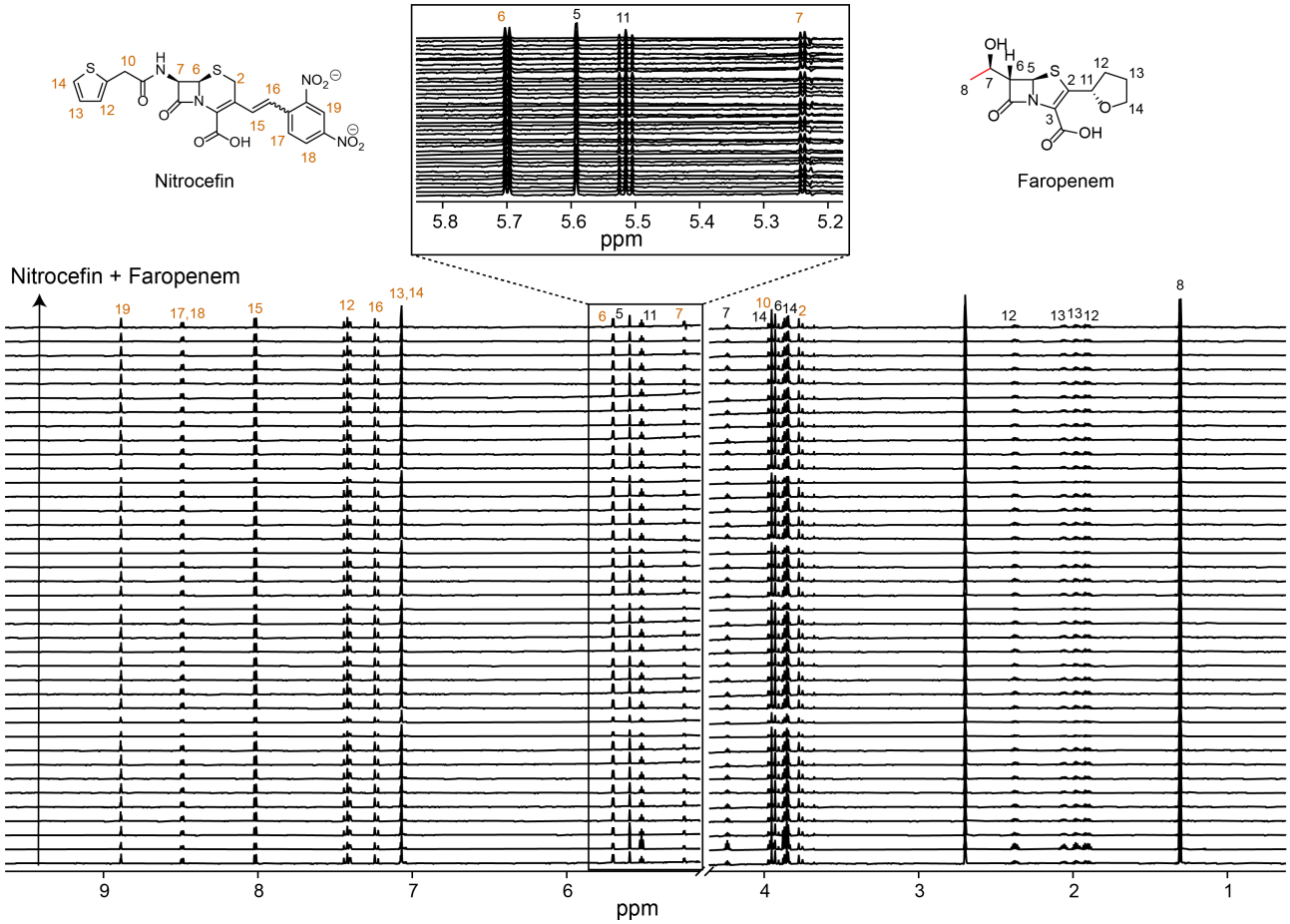


Figure 7.51 | ^{1}H NMR spectra (700 MHz) show evidence for the lack of reaction of nitrocefin with faropenem. The spectra are stacked from bottom to top and subsequent spectrum was acquired every 26 min.

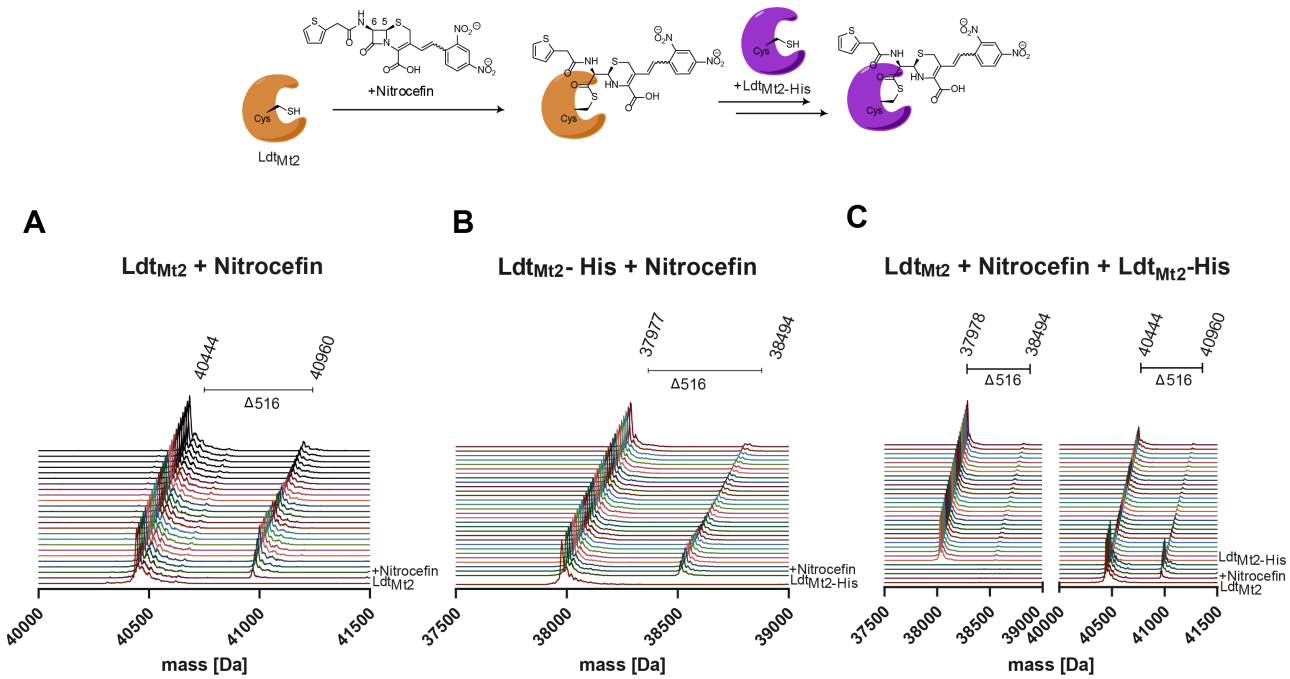


Figure 7.52 | Monitoring recyclisation of nitrocefin using protein transfer MS assays. **A.** Covalent reaction of hexaHis-tag cleaved Ldt_{Mt2} (1.5 µM) with nitrocefin (0.75 µM). **B.** Covalent reaction of hexaHis-tag cleaved Ldt_{Mt2} (1.5 µM) with nitrocefin (0.75 µM). **C.** hexaHis-tag cleaved Ldt_{Mt2} (1.5 µM) was treated with nitrocefin (0.75 µM) and incubated for 5 min before the addition of Ldt_{Mt2} (1.5 µM). Following the addition of Ldt_{Mt2}, the relative acylation of hexaHis-tag cleaved Ldt_{Mt2} with nitrocefin decreased and evidence for acylation of Ldt_{Mt2} was observed. **D.** Ldt_{Mt2} (1.5 µM) was treated with nitrocefin (0.75 µM) and incubated for 5 min before the addition of the hexaHis-tag cleaved Ldt_{Mt2} (1.5 µM). Following the addition of hexaHis-tag cleaved Ldt_{Mt2}, the level of the Ldt_{Mt2}-nitrocefin complex decreased and evidence for acylation of hexaHis-tag cleaved Ldt_{Mt2} was observed.

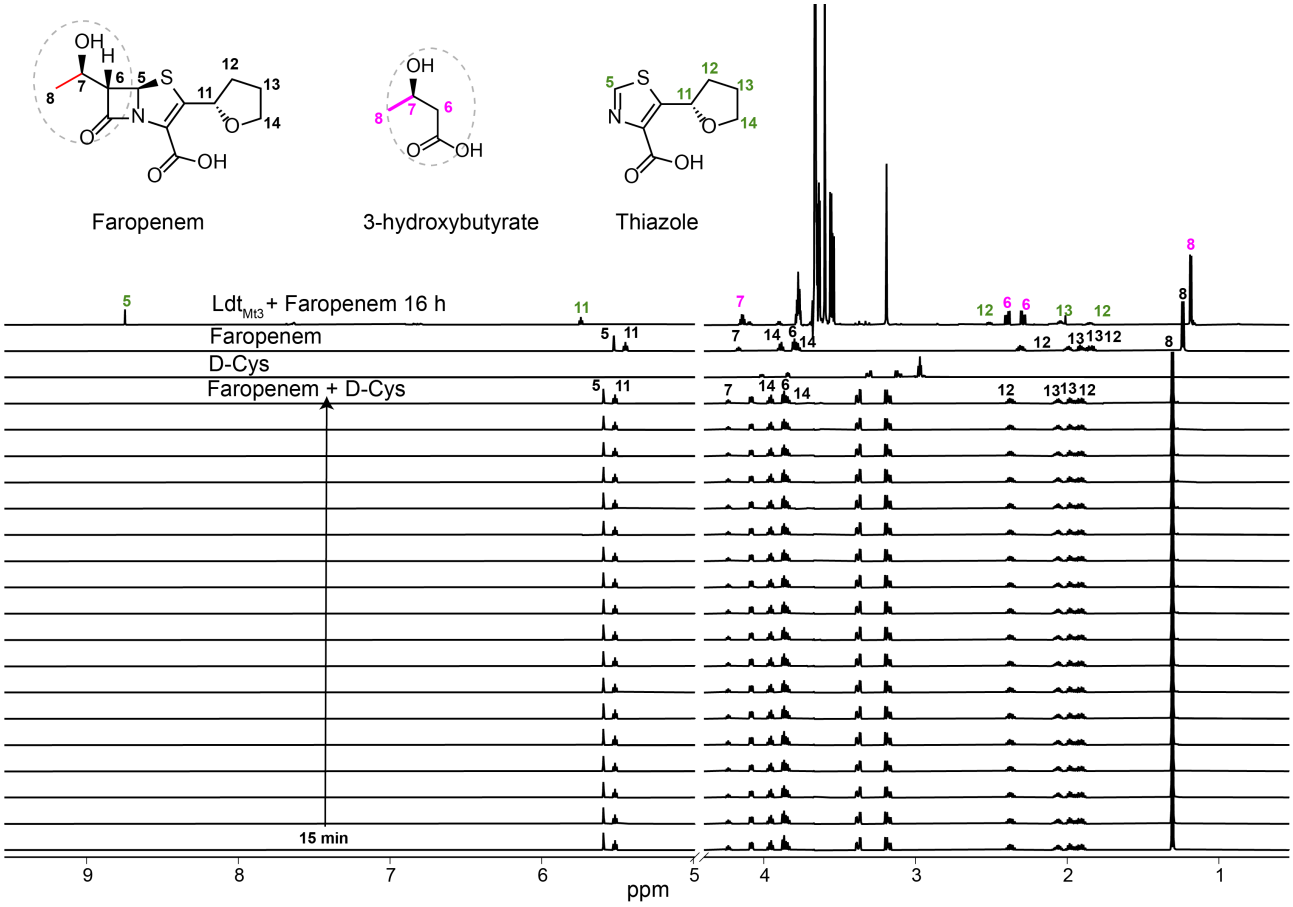


Figure 7.53 | ^{1}H NMR spectra (600 MHz) showing evidence for lack of reaction of faropenem with D-Cys amino acid. The first spectrum at the bottom was acquired after 15 minutes of sample preparation. Each spectrum was then acquired every 1.3 h (32 scans). Note that the ratio of free D-Cys and cystine may affect the observations.

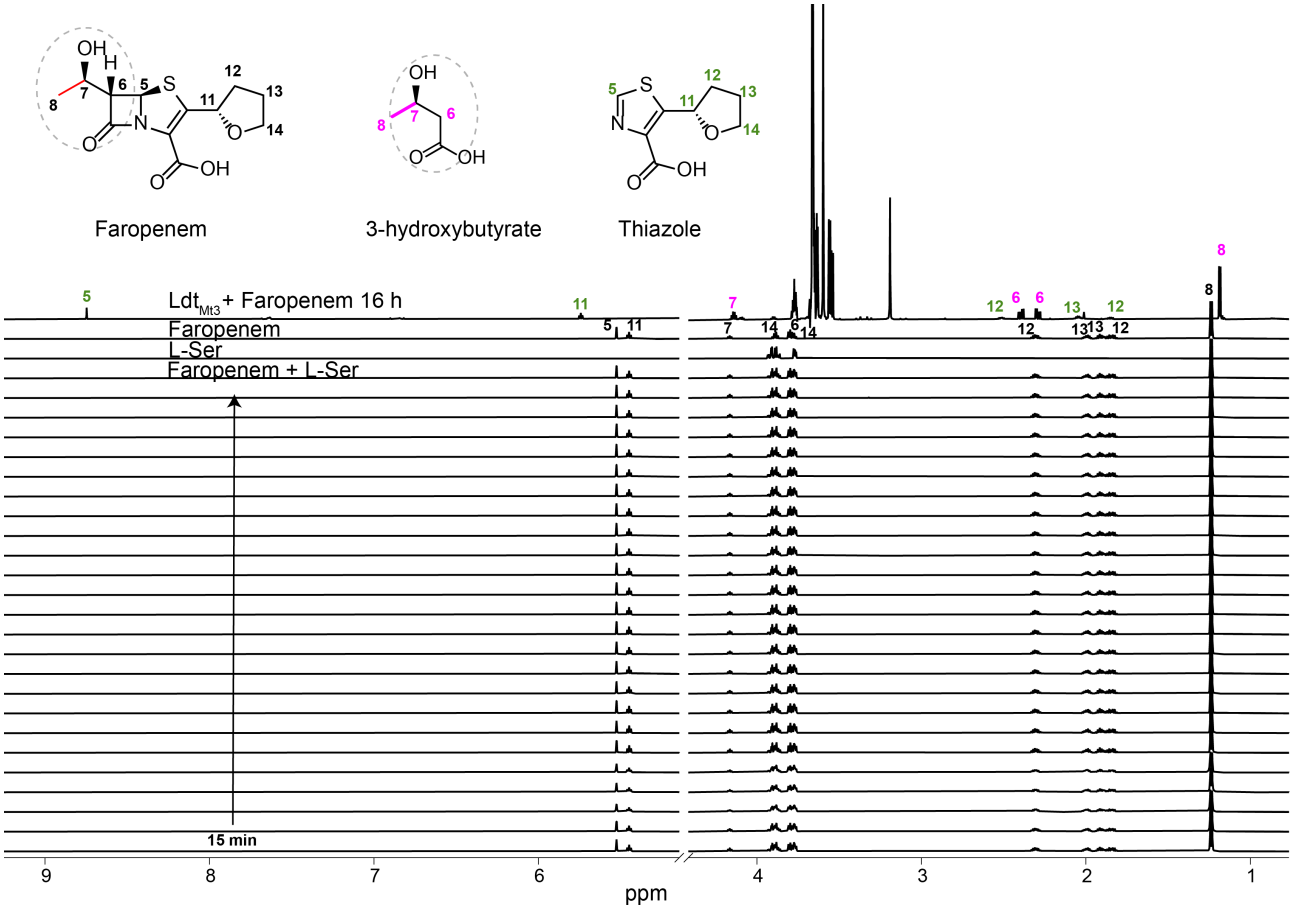


Figure 7.54 | ^{1}H NMR spectra (600 MHz) showing evidence for lack of reaction of faropenem with L-Ser amino acid. The first spectrum at the bottom was acquired after 15 minutes of sample preparation. Each spectrum was then acquired every 1.3 h (32 scans).

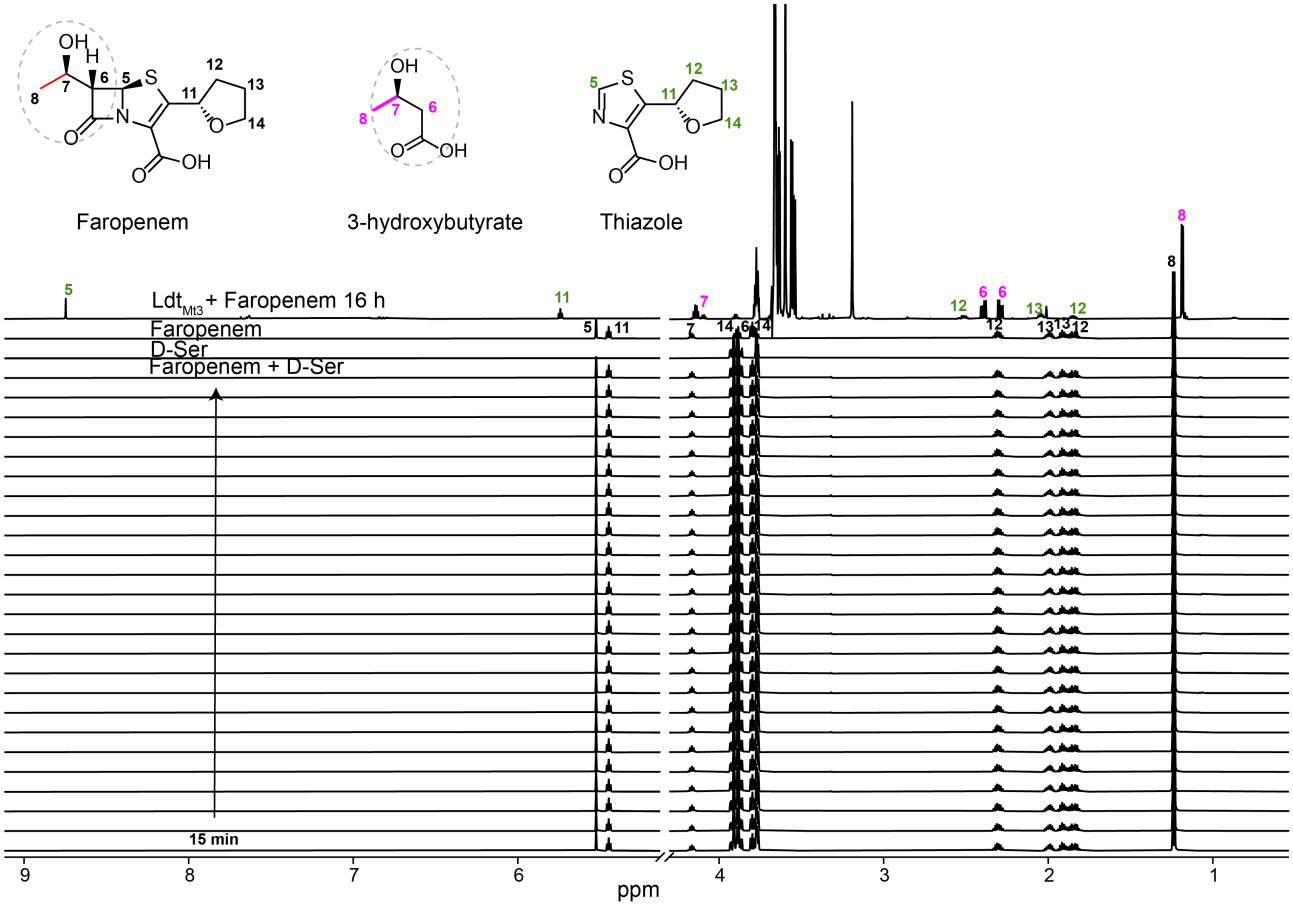


Figure 7.55 | ^{1}H NMR (600 MHz) evidence for lack of reaction of faropenem with D-Ser amino acid. The first spectrum at the bottom was acquired after 15 minutes of sample preparation. Each spectrum was then acquired every 1.3 h (32 scans).

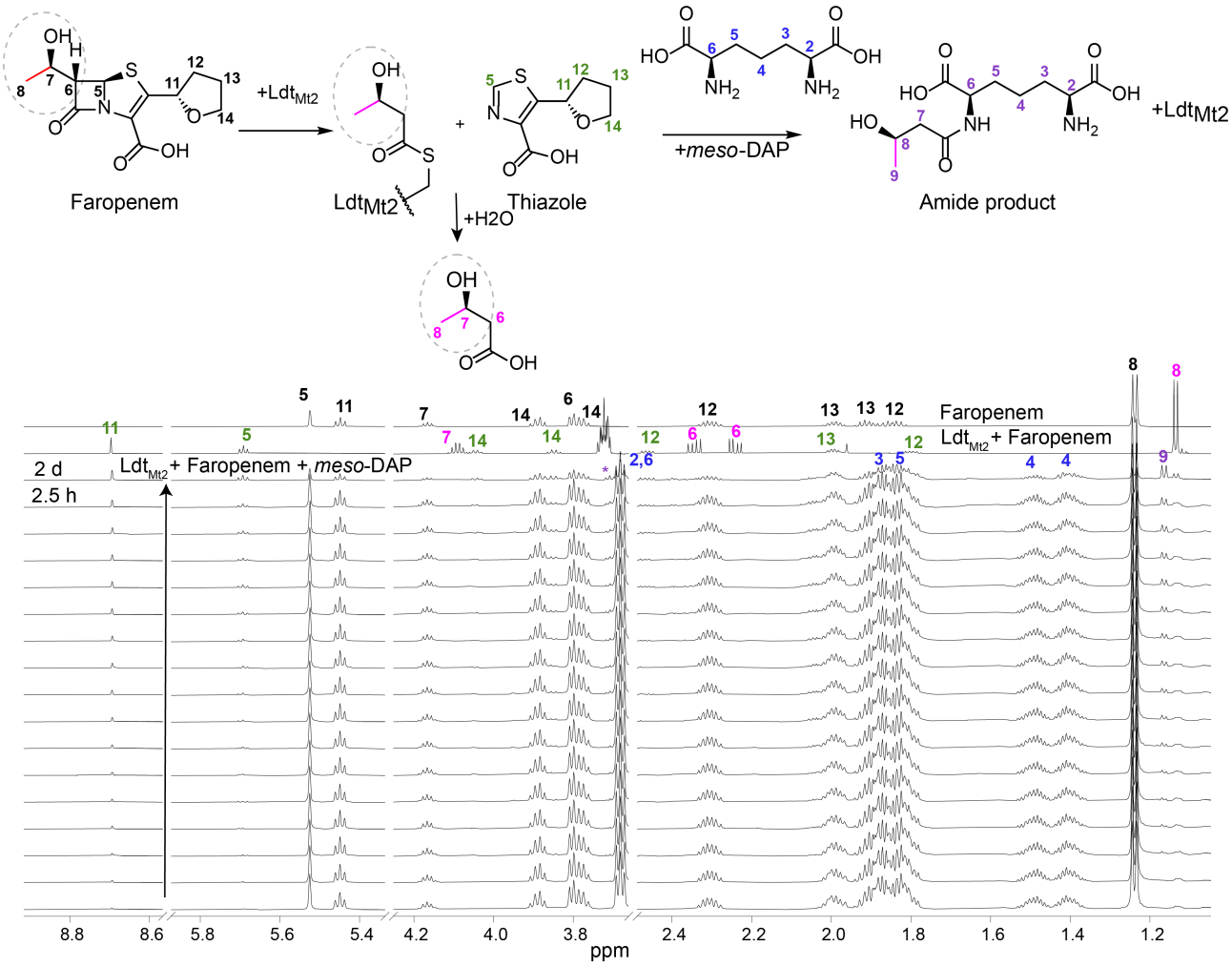


Figure 7.56 | ^{1}H NMR spectra (600 MHz) showing evidence for deacylation of Ldt$_{Mt2}$-faropenem complex by *meso*-DAP. In the presence of *meso*-DAP (2 mM), new methyl resonances (1.16 ppm) and a triplet at 3.71 ppm (indicated by an asterisk) corresponding to the chemical shift of C2/C6 proton of the expected amide product was observed in the Ldt$_{Mt2}$ (100 μM)-faropenem (2 mM) mixture. The first spectrum at the bottom was acquired after 5 minutes of sample preparation. Each spectrum was then acquired every 60 min (32 scans).

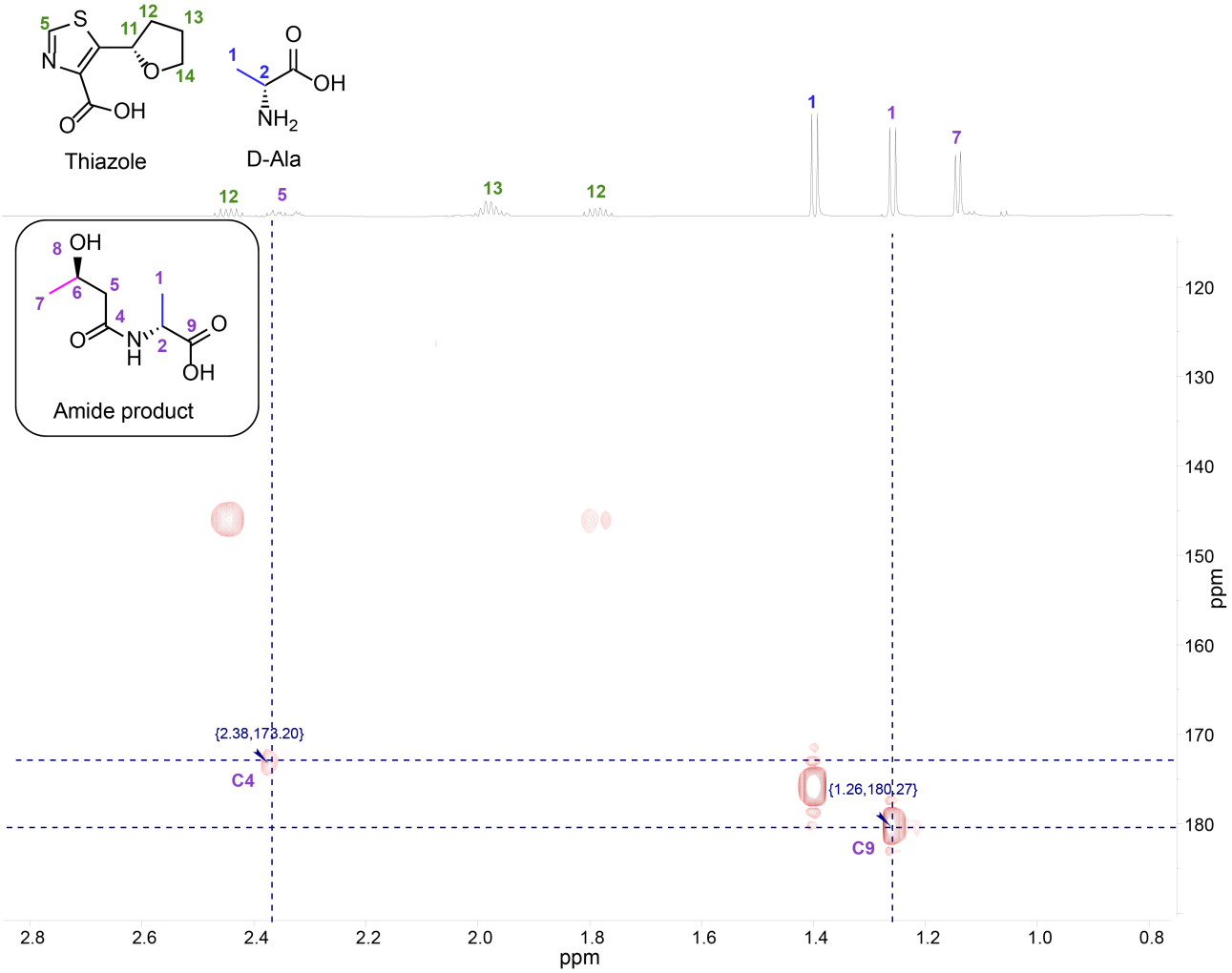


Figure 7.57 | ^{1}H-^{13}C HMBC spectrum (600 MHz) suggests that the C4 carbonyl of amide product has a lower chemical shift than the C9 carbonyl in the amide product formed from the deacylation of Ldt_{Mt2}-faropenem complex by D-Ala. The analysis was performed using a crude mixture of the D-Ala treated Ldt_{Mt2}-faropenem complex.

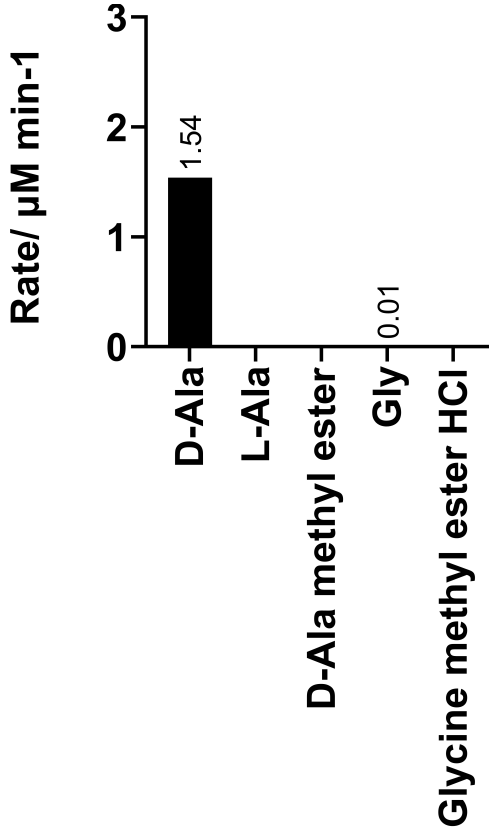


Figure 7.58 | Amino acid esters are poor deacylators of Ldt_{Mt2}-faropenem complex compared to the free amino acids.

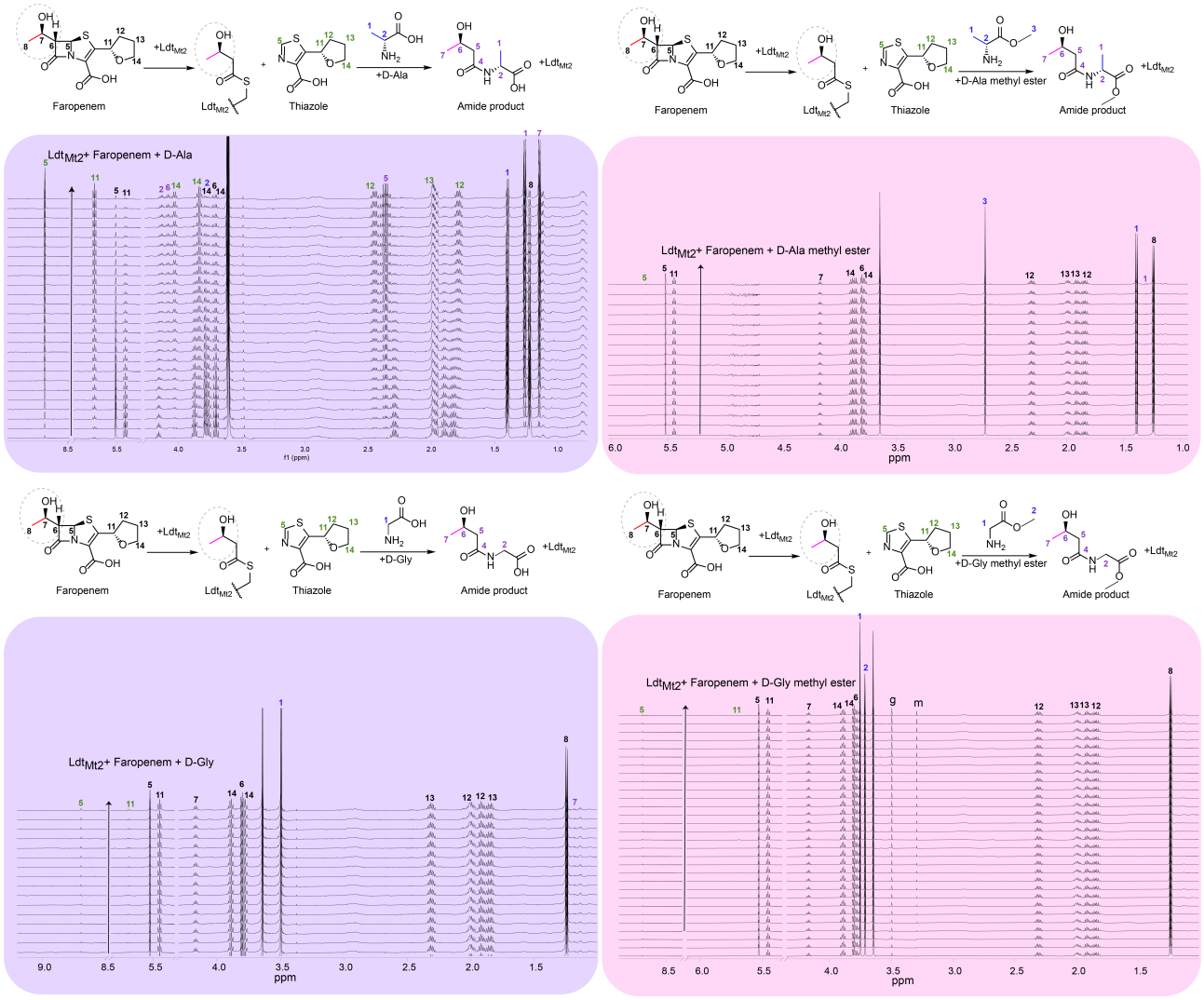


Figure 7.59 | ^{1}H NMR spectra (600 MHz) showing evidence for deacylation of Ldt$_{Mt2}$-faropenem complex only by a free carboxylic acid group containing amino acids compared to their respective ester derivatised forms. "g" and "m" correspond to spontaneous hydrolysis products of glycine methyl ester i.e., glycine and methanol, respectively. Each spectrum was acquired every 1.3 h, 60 min, 60 min, and 3 min for Ldt$_{Mt2}$-faropenem mixture treated with D-Ala, D-Ala methyl ester, Gly and Gly methyl ester, respectively.

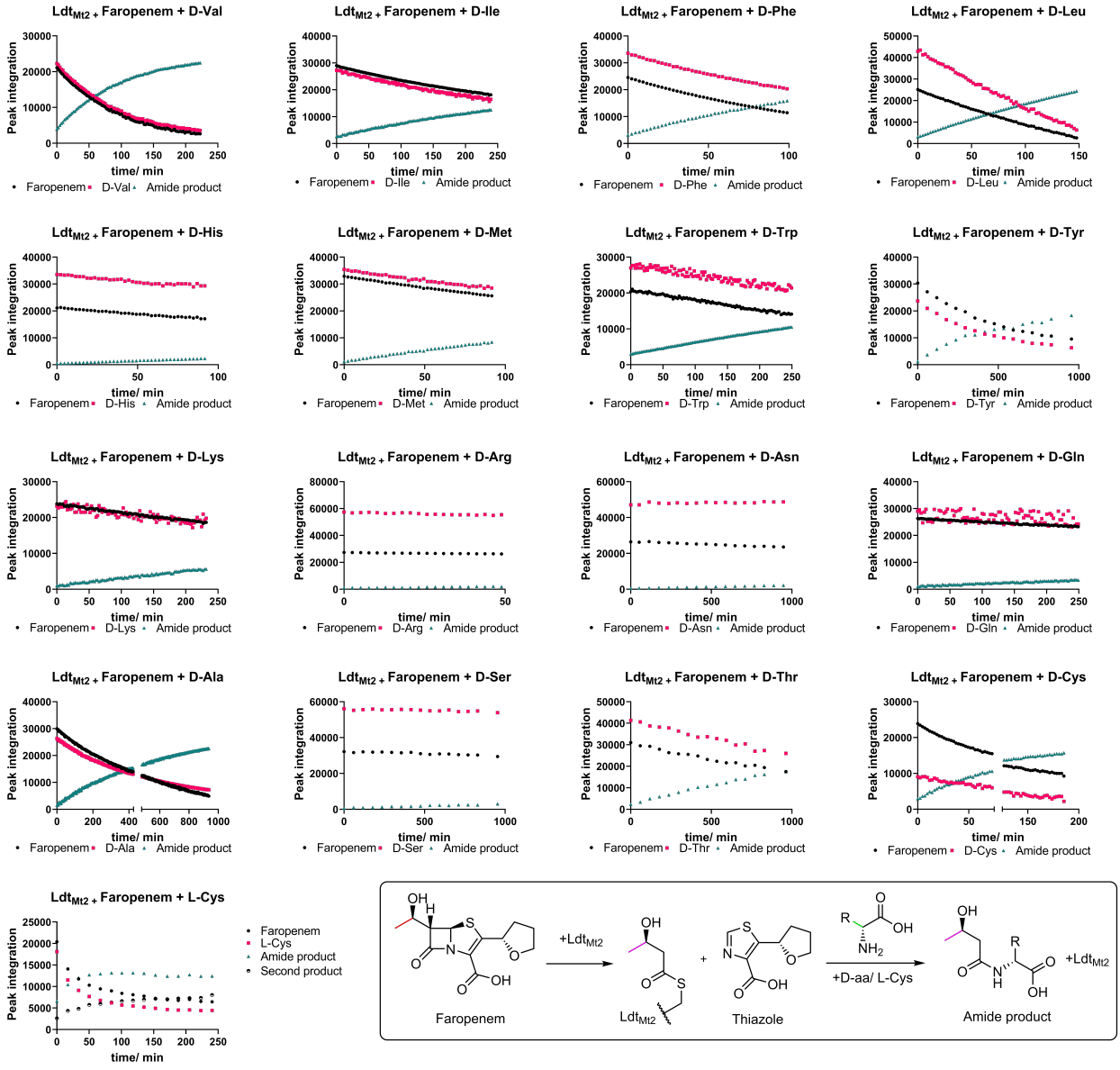


Figure 7.60 | Integration of the peaks corresponding to methyl resonances of faropenem, D-amino acids or L-Cys and the new peptide like amide product from the time course ^{1}H NMR spectra (AV600). The amount of faropenem and D-amino acid present in the reaction decreases over time, whereas the product formation occurs at varying rates with different D-amino acids. Only in the case of L-Cys was new product formation via deacylation of Ldt$_{Mt2}$-faropenem complex observed of all the tested L-amino acids. Conversion of the initial product to a more stable second product was observed with L-Cys. Note that the reaction with D- or L-Cys may be complex due to the presence of mixtures of thiols and/or cystines in the solution. A scheme of the proposed reaction and the expected amide product is shown.

Table 7.11: The chemical shift assignments for the D-Phe and Ldt_{Mt2}-faropenem complex derived amide product. The crude reaction product mixture of amide product and thiazole fragment formed upon reaction of faropenem with Ldt_{Mt2}. The sample was prepared in 50 mM sodium phosphate pH 7.5, 10%$_{(v/v)}$ D_2O. Chemical shift assignments were made based on HMBC and HSQC spectra. Refer to Appendix Figures 7.61, 7.62 and 7.63.

Position	^{1}H (ppm)	^{13}C (ppm)
1	2.84, 3.14	37.88
2	3.98	65.25
3		
4		172.99
5	2.31, 2.23	44.64
6	4.42	56.23
7	0.96	21.79
8		
9		137.82
10-14	7.21, 7.23, 7.30	126.80, 128.67, 129.14
15		178.14

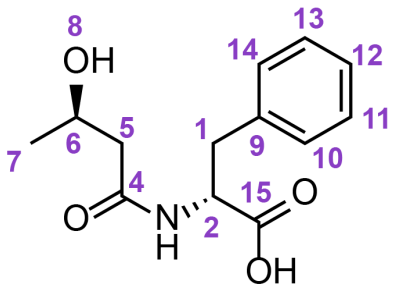
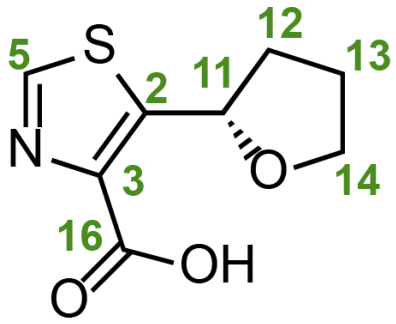


Table 7.12: The chemical shift assignments for the Ldt_{Mt2}-faropenem complex derived thiazole product. Thiazole was assigned in the crude reaction product mixture of amide product and thiazole fragment formed upon reaction of faropenem with Ldt_{Mt2}. The sample was prepared in 50 mM sodium phosphate pH 7.5, 10%$_{(v/v)}$ D_2O. Chemical shift assignments were made based on HMBC and HSQC spectra. Refer to Appendix Figures 7.61, 7.62.

Position	**^{1}H (ppm)**	**^{13}C (ppm)**
2		146.3
3		
5	8.68	74.98
11	5.68	75.23
12	1.79, 2.45	34.66
13	1.99	25.33
14	3.84, 4.03	69.11
16		



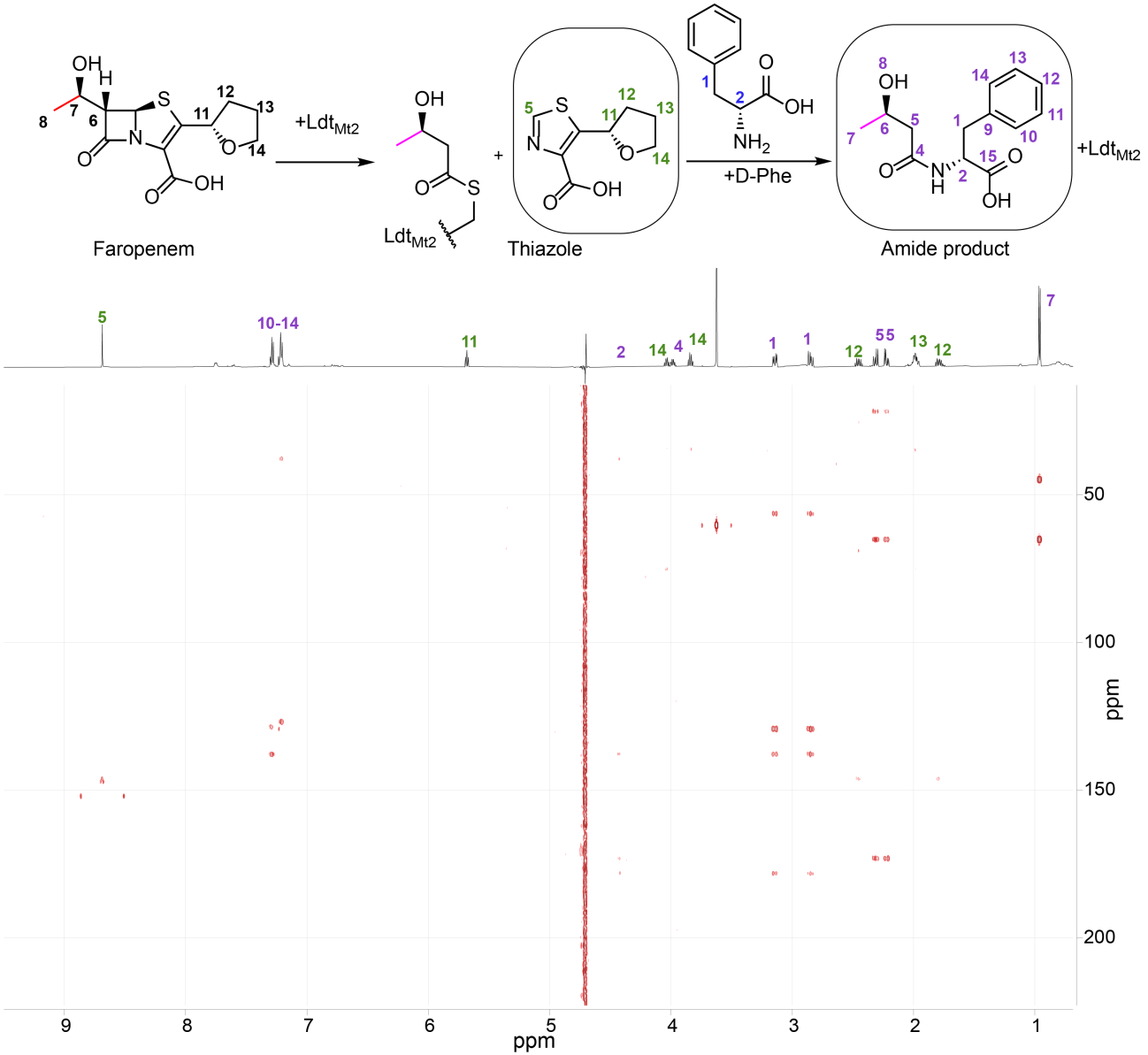


Figure 7.61 | ^{1}H-^{13}C-HMBC spectrum of the amide product derived from deacylation of Ldt$_{Mt2}$-faropenem complex by D-Phe.

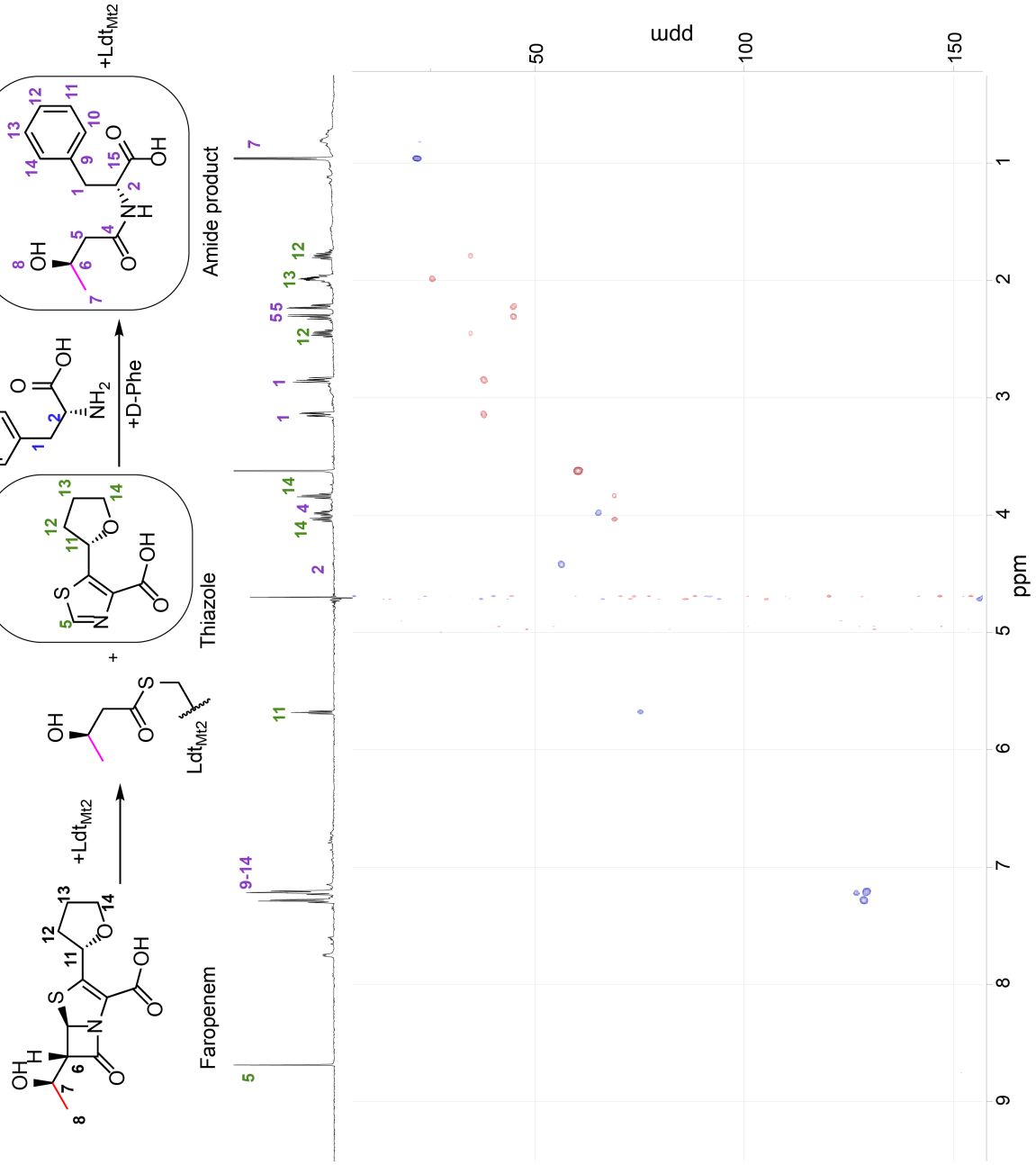


Figure 7.62 | ^{1}H-^{13}C-HSQC spectrum of the amide product derived from deacylation of Ldt$_{Mt2}$-faropenem complex by D-Phe.

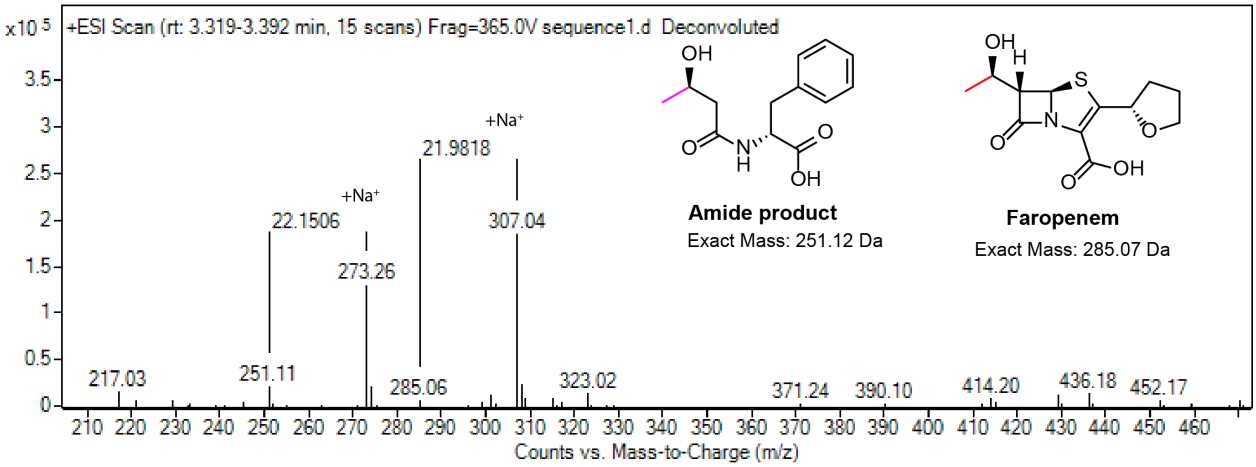


Figure 7.63 | Mass spectrum of crude mixture of Ldt_{Mt2}-faropenem complex treated with D-Phe. Ldt_{Mt2} (0.6 µM) was incubated with faropenem (12 µM) for 5 min before treatment with excess D-Phe (25 µM) in 50 mM sodium phosphate, pH 7.5 for >12 h prior to analysis by SPE-MS. The mass at 251 Da and 285 Da corresponds to the amide product and faropenem, respectively. The sodiated adducts of both the amide product and faropenem were observed with a mass shift of ∼22 Da.

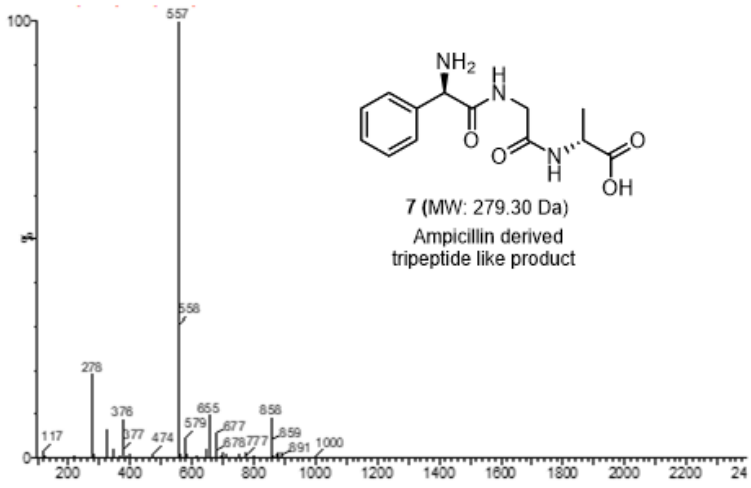


Figure 7.64 | Mass spectrum of crude mixture of Ldt_{Mt2}-ampicillin complex treated with D-Ala. Ldt_{Mt2} was incubated with faropenem before treatment with excess D-Ala in 50 mM sodium phosphate, pH 7.5, and incubated >18 h. Protein was removed from the sample by using Amicon Ultra-0.5, Ultracel-3 Membrane, 3 kDa and filtrate analysed using LC-MS. The mass at 278 Da corresponds to the ampicillin and D-Ala derived amide product.

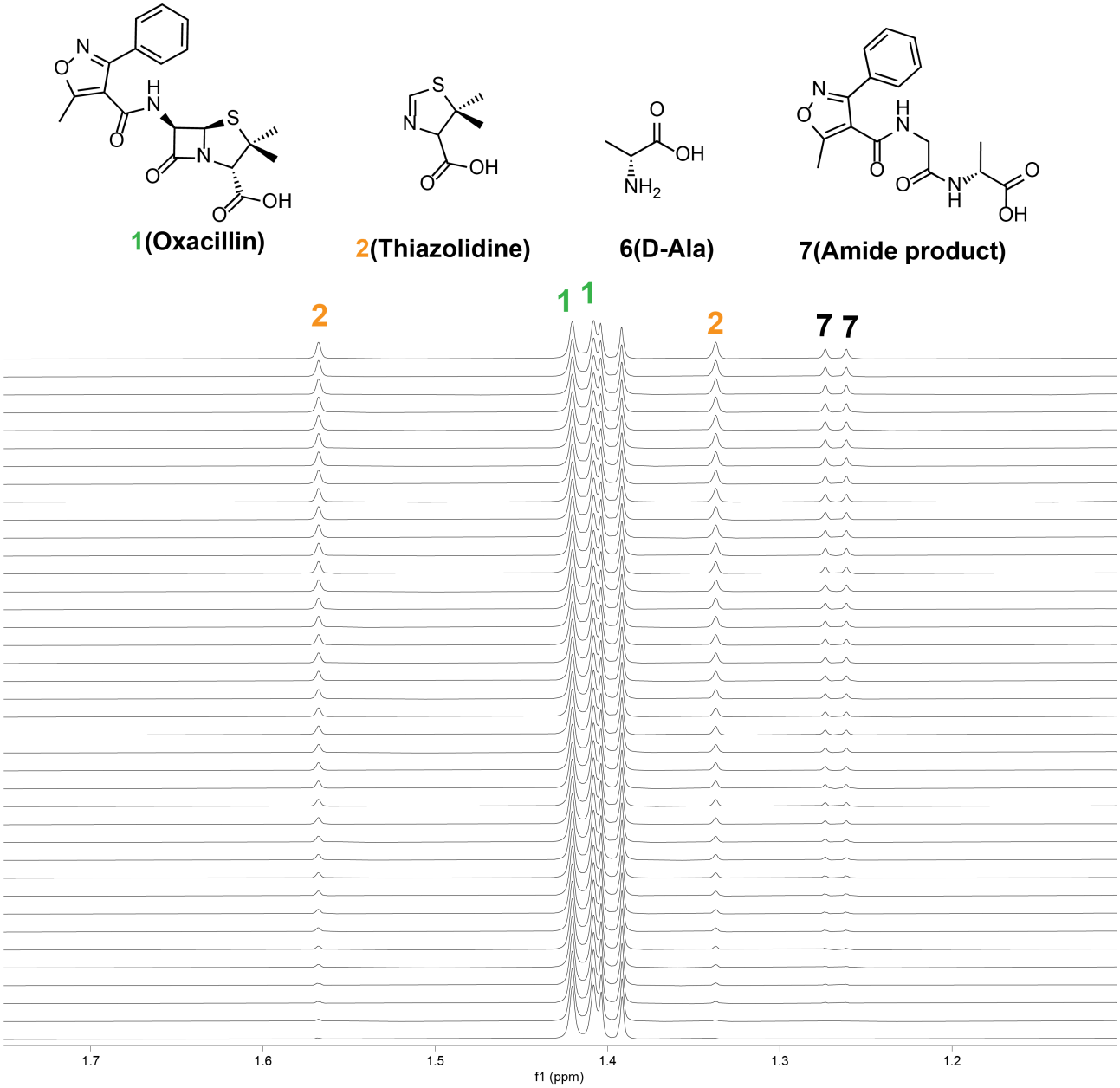


Figure 7.65 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of D-Ala with the Ldt_{Mt2}-oxacillin thioester complex. The reaction of D-Ala (2 mM) with the Ldt_{Mt2} (100 µM)-oxacillin (1 mM) thioester complex leads to the formation of amide product. The first spectrum at the bottom was acquired after 5 minutes of sample preparation. Each spectrum was acquired every 2.5 min (32 scans).

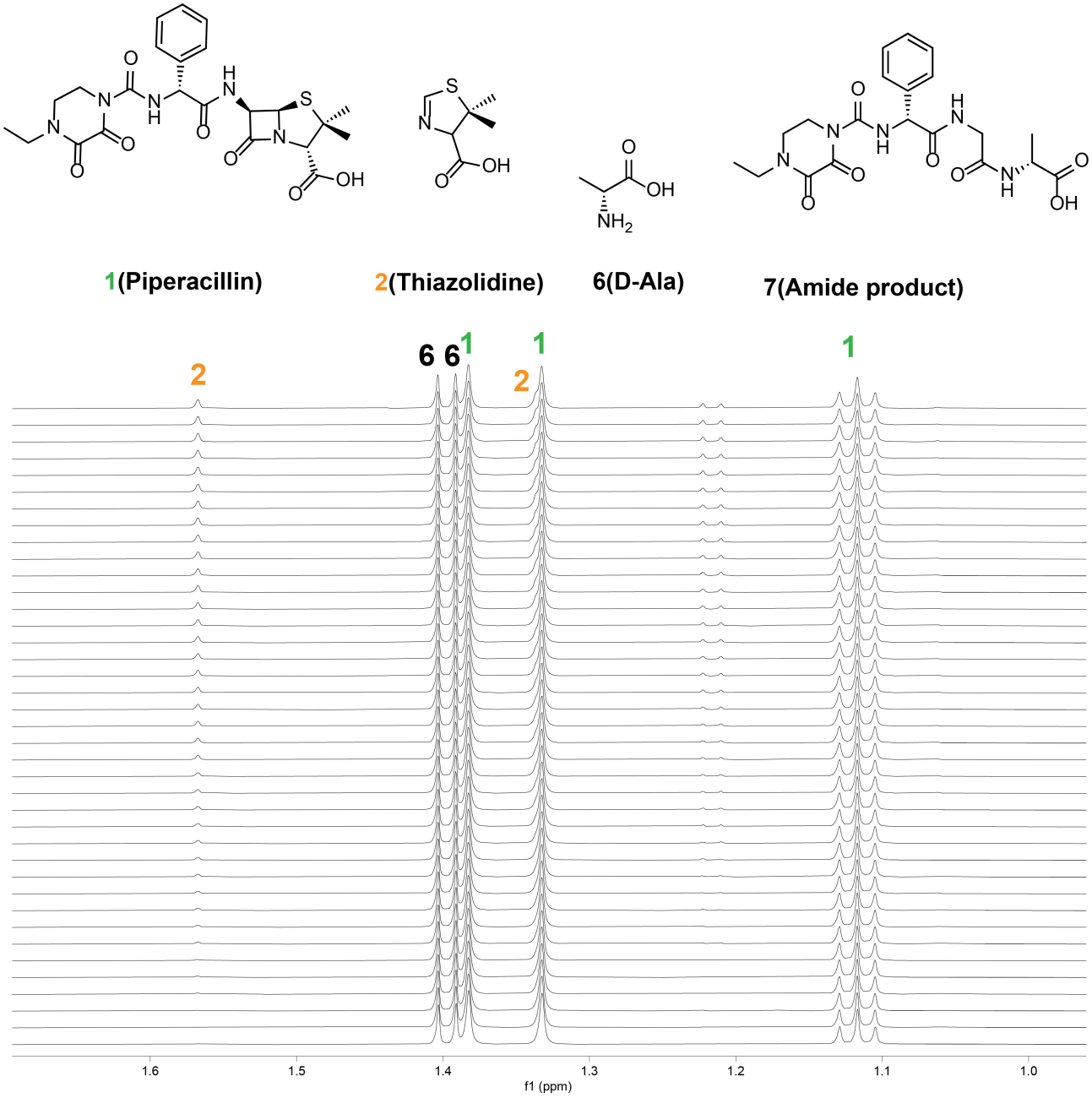


Figure 7.66 | ^{1}H NMR spectra (600 MHz) showing evidence for the reaction of D-Ala with piperacillin thioester intermediate. The reaction of D-Ala (2 mM) with Ldt_{Mt2} (100 µM)-piperacillin (1 mM) thioester intermediate leads to the formation of tripeptide like amide product. The first spectrum at the bottom was acquired after 5 minutes of sample preparation. Each spectrum was acquired every 2.5 min (32 scans)

8

Mechanistic studies employing NMR on β-lactamase catalysed hydrolysis of carbapenems and a penem

Contents

8.1 Abstract **291**
8.2 Introduction **291**
8.3 Contributions to the manuscripts **292**
8.4 Additional work to that described in manuscript 9 . . . . . **293**
8.5 Future directions **298**

8.1 Abstract

The work described in this chapter involved NMR studies conducted by me on the hydrolysis of biapenem (a carbapenem) and faropenem (a penem). The NMR studies in solution complement collaborative crystallographic studies by others (refer to the section on Contributions to the manuscripts). The results inform on the products observed in solution with the serine-β-lactamase (KPC-2) and the metallo-β-lactamases (VIM-1, VIM-2, L1). The faropenem derived THF-ring closed-(imine)-thiazole was the major product observed with VIM-1 and VIM-2. The THF-ring opened-(imine)-thiazole alkene was the major product observed with L1, KPC-2, OXA-10, OXA-23 and OXA-48 (by NMR studies). In the case of faropenem, the major product in solution was not reflected in the crystallographic studies, at least, with metallo-β-lactamases L1 in complex with THF-ring closed-(imine)-thiazole and VIM-2 with The THF-ring opened-(imine)-thiazole alkene. The results show that the hydrolytic products in the solution may not be represented in the crystal structures. By contrast, in the case of biapenem, the *((2S)*-imine)-pyrroline, the major product observed in solution was also observed in complex with VIM-2 by crystallographic studies. Therefore, mechanistic insights for downstream applications, i.e. drug design should be drawn from a combination of solution and crystallographic studies.

8.2 Introduction

β-Lactamases catalyse the hydrolysis of the β-lactam antibiotics, and can be divided into four classes [377] [378]. Classes A, C and D employ a nucleophilic serine in catalysis; class B MBLs employ one or two zinc ions in catalysis [378].

Analysis of β-lactamase products can provide insight into their mechanisms, and may inform on the design of more efficient inhibitors. My work involved investigation of the hydrolytic profiles of (i) faropenem with serine-β-lactamases, KPC-2 (Class A), OXA-10, OXA-23, OXA-48 (Class D), and metallo-β-lactamases NDM-1, VIM-1, VIM-2 and L1 (Class B) and (ii) biapenem with β-lactamases (VIM-1, VIM-2, and KPC-2).

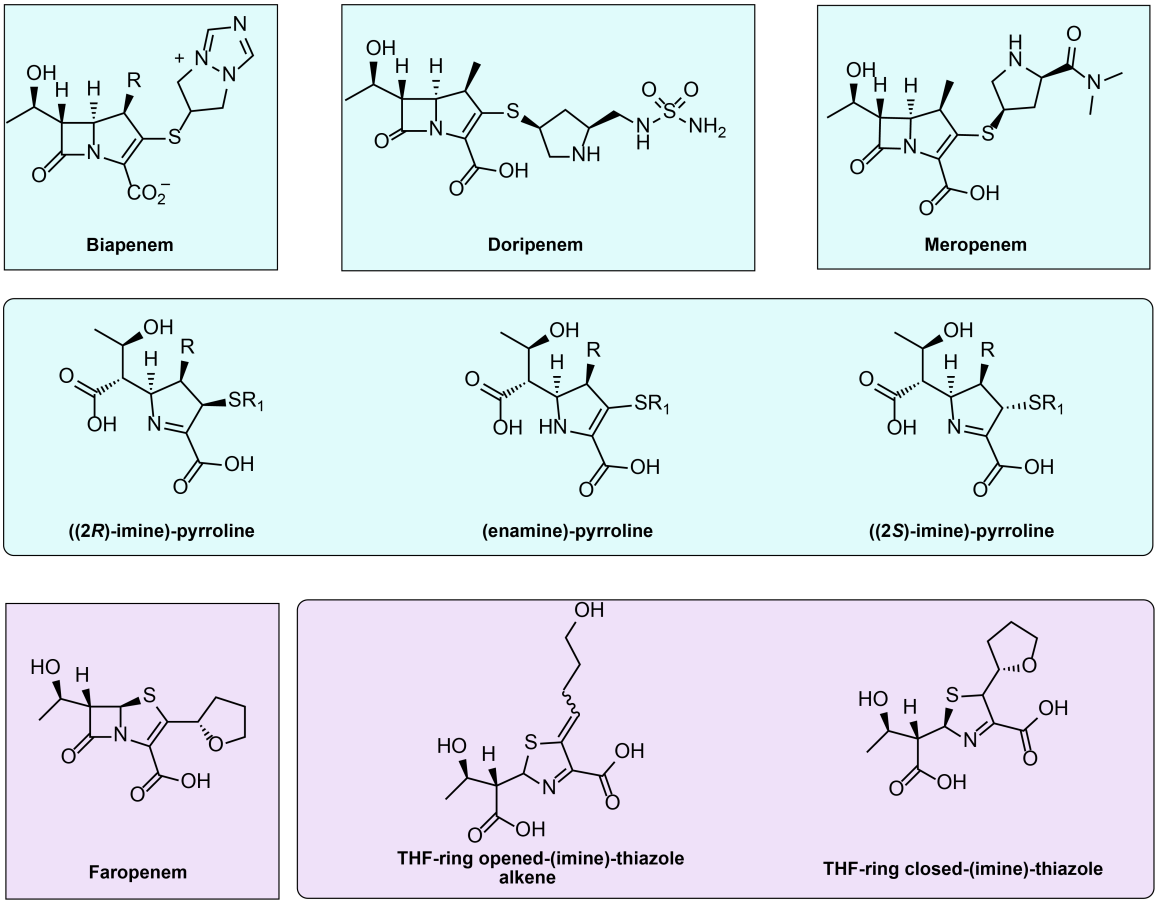


Figure 8.1 | Chemical structures of carbapenems (biapenem, doripenem, meropenem) and penem (faropenem) and their hydrolytic products. Note that the structures shown only include those that are relevant in the additional results section of this chapter (refer to manuscripts 8 and 9 for other products).

8.3 Contributions to the manuscripts

I collaborated with Dr.Anka Lucik (Supervisor: Prof. Christopher Schofield) and Dr. Philip Hinchliffe (Supervisor: Prof. Jim Spencer, University of Bristol), who carried out structural studies on the reactions of β-lactamases with faropenem. This collaboration was later extended to perform studies with biapenem. All NMR experiments in the papers were carried out by me. My work involved identifying and characterising the hydrolysis product of faropenem (manuscript 8 [348] and biapenem (manuscript 9 [379]) using NMR time-course studies.

8.3.1 Manuscript 8: Faropenem reacts with serine and metallo-β-lactamases to give multiple products

In manuscript 8 [348], my work entailed:

- Employing the methods described in detail in manuscript 8 [348]. In particular, I characterised the THF-ring closed product of faropenem for the first time in solution.
- Unstable kinetic products are usually difficult to capture by crystallographic studies. My experiments provided evidence for the formation of the enzyme catalysed in solution products and their subsequent conversion to hydrolytic products.

8.3.2 Manuscript 9: Studies on the Reactions of Biapenem with VIM metallo-β-lactamases and the serine-β-lactamase KPC-2

Manuscript 9 [379] describes crystallographic and in-solution studies with biapenem, where the biapenem derived *(2S)*-imine was observed as the major product in solution. My work in Manuscript 9 involved the following:

- Monitoring and assigning the hydrolytic products of biapenem upon incubation with MBLs (VIM-1 and VIM-2) and SBL(KPC-2).
- NMR characterisation of the apparent nascent enamine product.

8.4 Additional work to that described in manuscript 9

In addition to the work described in manuscript 9, experiments were also performed with the N-terminally His-tagged and His-tag cleaved variants of KPC-2 to determine if the His-tag affects catalysis. Time-course analysis shows that the His-tag does not affect the substrate turnover profile of KPC-2 (Figure 8.2). With the existing work, it was difficult to determine the stereochemistry of the THF-ring closed product of faropenem hydrolysis by MBLs or to fully characterise its nascent

products. Experiments performed with a reduced enzyme concentration did not result in the variation in the observed hydrolytic product profile under the tested conditions (Figure 8.3).

Further studies with VIM-2 were carried out with other carbapenems, i.e. doripenem and meropenem (Figure 8.4), where the (enamine)-pyrroline, and (*(2R)*,*(2S)*-imine)-pyrroline products were observed. Crystallographic evidence for all three hydrolytic products has been reported in crystal structures with SBLs [380,381]. The carbapenem derived *(2S)*-imine was the major product observed with both doripenem and meropenem (at least on the NMR timescale); this observation is in accord with those for biapenem and VIM-2 as reported in manuscript 9 [379]. Studies with imipenem, a carbapenem without a β-methyl group, report the slow tautomerisation of the (*(2R)*-imine)-pyrroline to the (*(2S)*-imine)-pyrroline [364]. The work with imipenem and other work [364] suggests that the (*(2R)*-imine)-pyrroline is likely the kinetic imine product from carbapenem hydrolysis, despite the apparent observation of the biapenem derived (*(2S)*-imine)-pyrroline in both NMR and crystallographic studies with VIM-2. The combined results are not conclusive with regard to the identity of the nascent product(s) of β-lactamase hydrolysed carbapenem and penem catalysis. In some cases, observations are consistent with the (enamine)-pyrroline/(enamine)-thiazole being the nascent product, at least under conditions of efficient catalysis. However, in particular, when stable acyl-enzyme complexes are formed, the (*(2S)*- and/or the *(2R)*-imine)-pyrroline may (also) be nascent products. Evidence for the (enamine)-pyrroline form of carbapenem isomerising to the (imine)-pyrroline products has been observed in acyl-enzyme complex with BlaC, an SBL present in *Mycobacterium tuberculosis* [382].

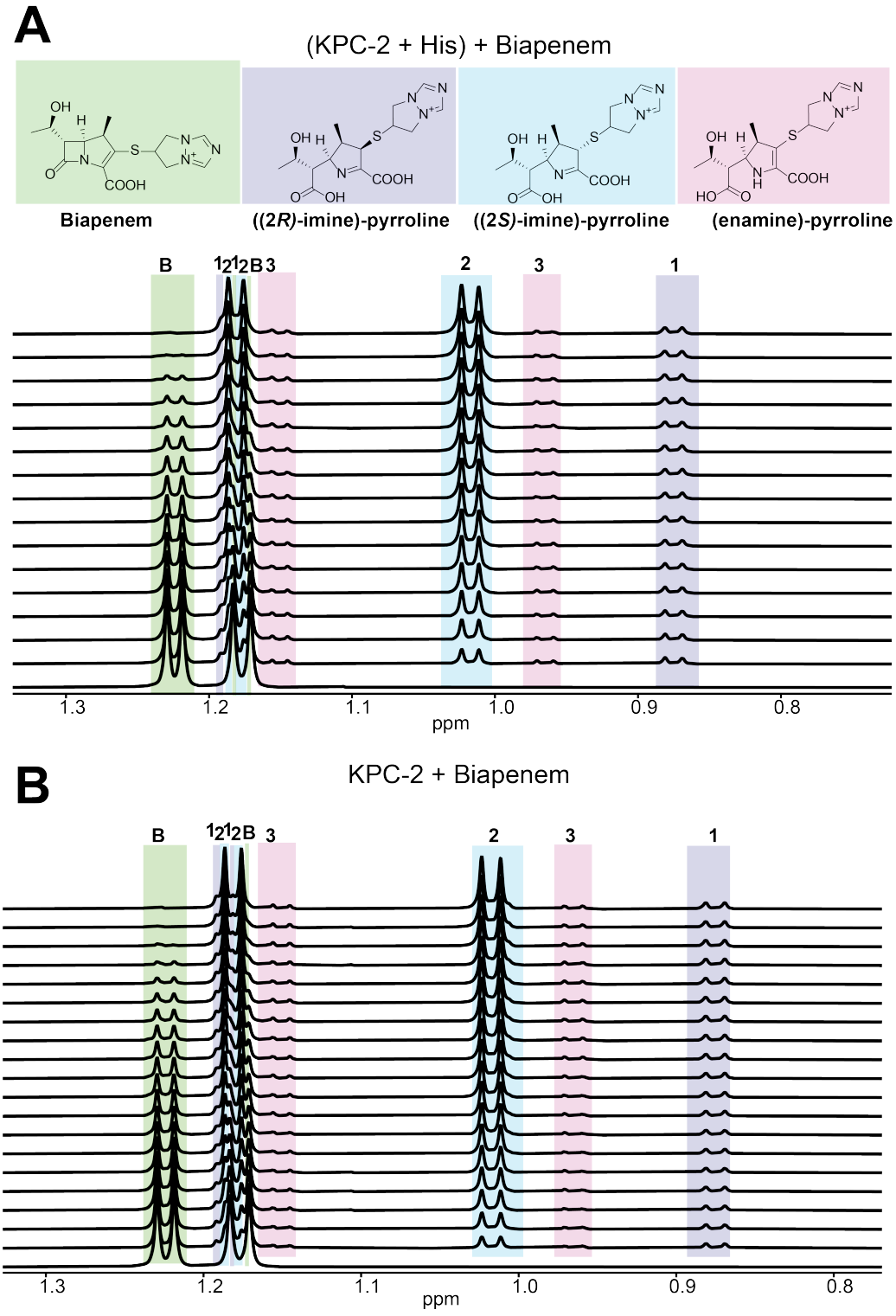


Figure 8.2 | The presence of an N-terminal His-tag does not affect the substrate turnover of KPC-2. A. ^{1}H NMR reaction monitoring of His-tagged KPC-2 treated with biapenem (5 mM) **B.** ^{1}H NMR reaction monitoring of His-tag cleaved KPC-2 treated with biapenem (5 mM). The result suggests that the presence of an N-terminal His-tag does not affect the catalysis by KPC-2. Spectra, stacked from bottom to top, were acquired every 2.5 minutes.

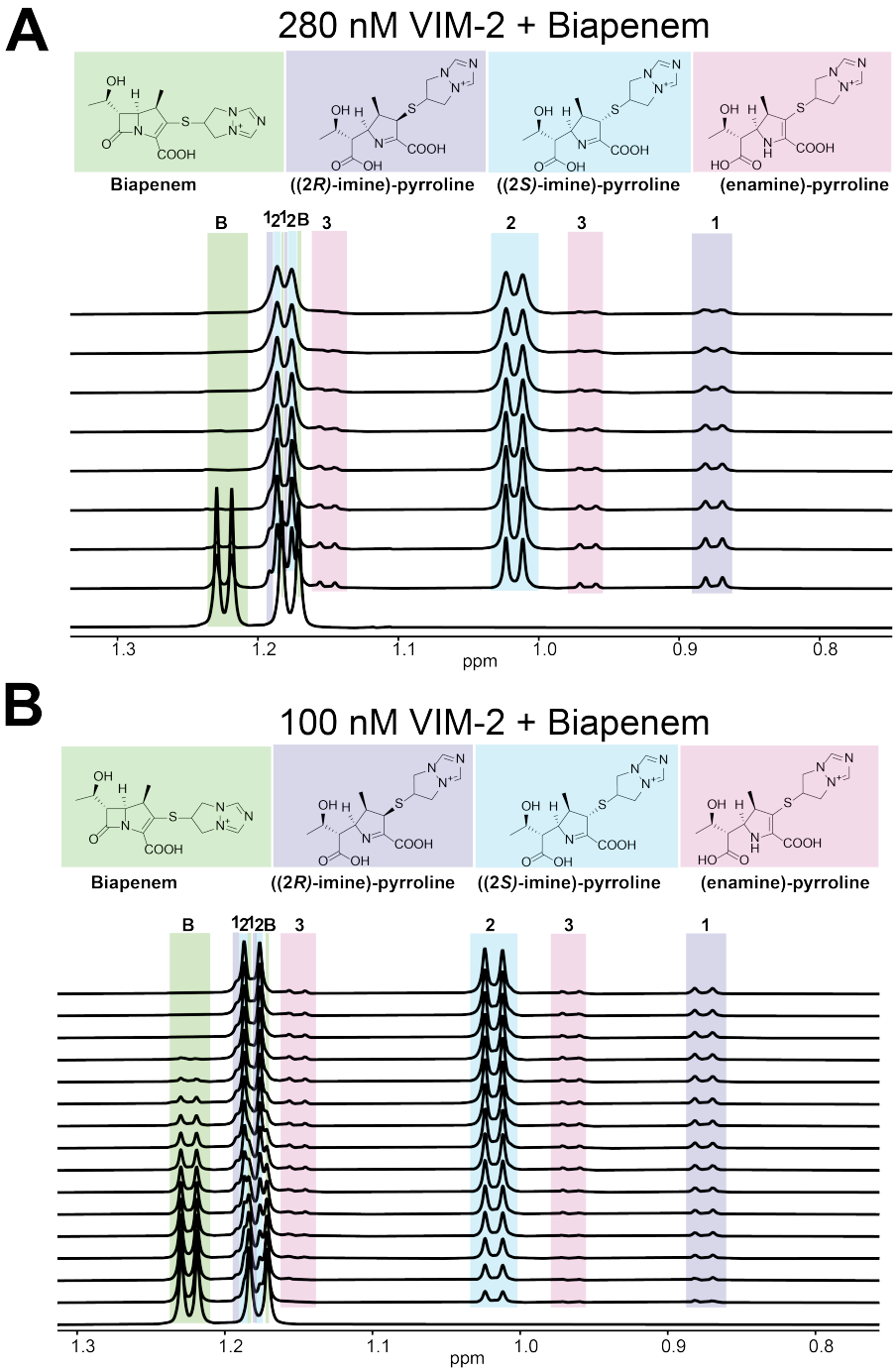


Figure 8.3 | ^{1}H NMR time course of the methyl resonances of the species formed upon reaction of biapenem with varying concentrations of VIM-2. **A.** 280 nM or **B.** 100 nM VIM-2 (5 mM) was treated with biapenem (5 mM) in 50 mM sodium phosphate at pH 7.5, 10%($_{v/v}$) D_2O. The hydrolytic product profile of biapenem did not vary with altered enzyme concentrations. Spectra, stacked from bottom to top, were acquired every 2.5 minutes.

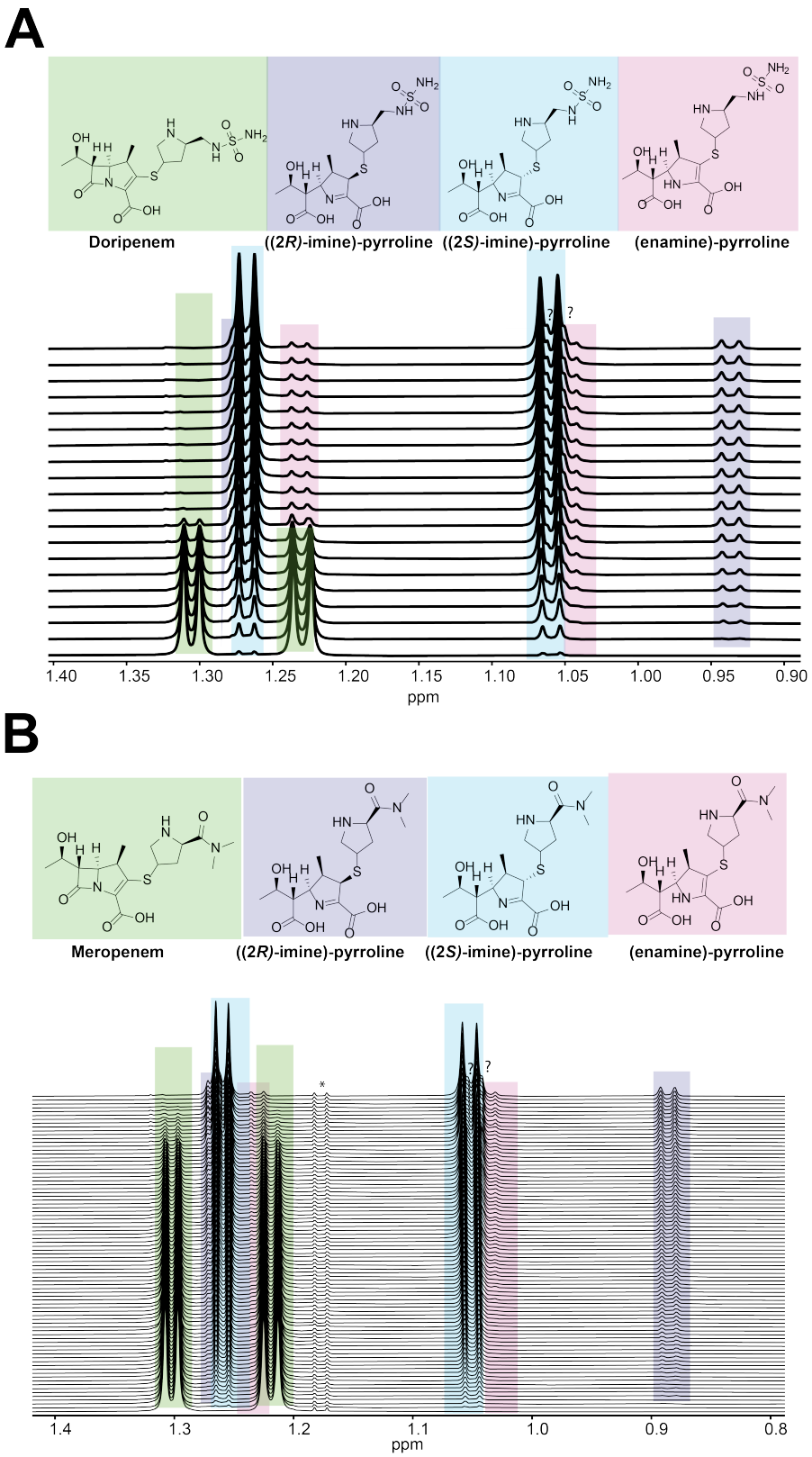


Figure 8.4 | ^{1}H NMR time course of the methyl resonances of species formed upon reaction of doripenem or meropenem with VIM-2. **A.** Doripenem (5 mM) or **B.** meropenem (5 mM) was treated with VIM-2 (280 nM) in 50 mM sodium phosphate at pH 7.5, 10%($_{v/v}$) D_2O. Uncharacterised product peaks are labeled with a question mark or an asterisk. Evidence was accrued for the formation of both *(2R)*-imine, *(2S)*-imine and enamine products (blue, purple, and pink, respectively). Spectra, stacked from bottom to top, were acquired every 2.5 minutes.

8.5 Future directions

Stopped flow kinetics experiments coupled with UV-spectroscopy could be performed with MBLs and SBLs to characterise the nascent hydrolytic products of both faropenem (manuscript 8) and biapenem (manuscript 9). In the work described here, HMBC experiments were used to characterise the nascent enamine products observed with both faropenem and carbapenem. HPLC experiments could be performed following large scale enzymatic crude mixtures to isolate the nascent products, likely the enamines, and perform more detailed NMR characterisation. These will be challenging due to the likely tautomerisation of the enamines to imines.

There is scope for designing improved inhibitors by expanding the understanding of the covalent acyl-enzyme complexes. Stabilisation of the acyl-enzyme complex by preventing the tautomerisation from the (enamine)-pyrroline to (imine)-pyrroline form could be exploited to obtain more efficient inhibitors. Thus, the C2 side chain of carbapenem may be modified to vary the electronic properties of the carbapenem and subsequently formed acyl-enzyme complexes. In addition to inhibition of β-lactamases, analogous methods could be applied to design improved inhibitors of L, D-transpeptidases, or PBPs by carbapenems and penems.

9
Discussion

This thesis describes biochemical and inhibition studies on the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) main protease (M^{pro}), which is a therapeutic target to treat coronavirus disease 2019 (COVID-19) (Chapters 2-6), and the mycobacterial L,D-transpeptidases, which are investigational targets for tuberculosis treatment (Chapter 7). Both M^{pro} and L,D-transpeptidases are nucleophilic cysteine enzymes, and the same techniques were employed to investigate their reaction with small molecules, i.e. MS, crystallography, and NMR spectroscopy. In addition, NMR studies were performed with antibacterial targets, β-lactamases (Chapter 8) to investigate their hydrolytic products upon reaction with β-lactams, which can inhibit both M^{pro} and L,D-transpeptidases (Figure 9.1).

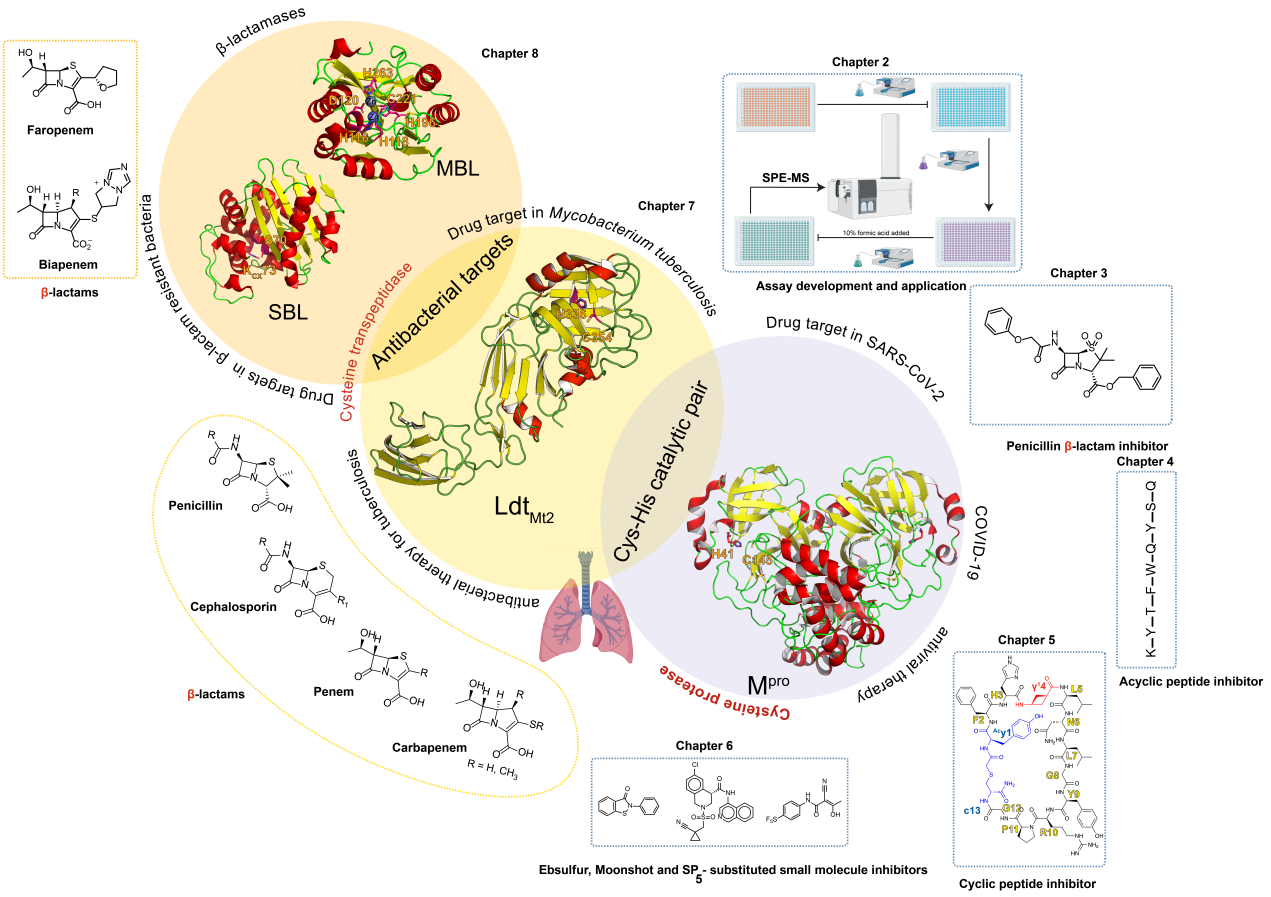


Figure 9.1 | Interconnectivity of the chapters described in this thesis.

Initial work presented in the thesis focuses on the development of a high throughput mass spectrometry-based assay for SARS-CoV-2 M^{pro} and its application to identify different classes of M^{pro} inhibitors; this work was performed at the onset of the COVID-19 pandemic in early 2020, as traditionally used fluorescence-based M^{pro} assays may suffer from fluorescence quenching of the substrate and, when screening small-molecule inhibitors, from compound-induced fluorescence interference. At that time, no direct MS-based M^{pro} assays had been reported. The novel assay employed solid phase extraction coupled to MS, a semi-automated technique that enables high-throughput applications (Chapter 2). The assay was validated by investigating commercial sets of small molecules for M^{pro} inhibition, i.e. the LOPAC^{1280} library and a set of FDA-approved active pharmaceutical ingredients, as well as reported M^{pro} inhibitors including N3, ebselen, disulfiram and boceprevir. Some of the identified inhibitors from commercial libraries include TPCK, TLCK, 5-thioguanine, auranofin, NSC9593, cisplatin, and

IPA-3, and these compounds belonged to different classes of inhibitors i.e. protease inhibitors, kinase inhibitors, metal-based inhibitors, and disulfides. The mode of inhibition of the identified M^{pro} inhibitors was investigated using MS which revealed that mostly covalent M^{pro} inhibitors have been identified, except for the metal-based inhibitors and non-covalently reacting 5-thioguanine.

After the validation of the MS M^{pro} inhibition assay, the SPE-MS assay was applied to screen commercial protease inhibitors and a set of compounds available in the Schofield lab (Department of Chemistry, University of Oxford), including synthetic β-lactams and γ-lactams, for M^{pro} inhibition to identify novel classes of M^{pro} inhibitors that could be of potential clinical relevance. Initially, the screens suffered from low reproducibility due to the reduced solubility of some of the tested compounds in the buffer system used. A CyBio-Liquid Handler robot was initially used to dispense the compounds; subsequently, an acoustic dry dispenser was used to reinvestigate the initial hits for M^{pro} inhibition. While PBIT and ebselen were identified as potent M^{pro} inhibitors irrespective of the dispensing method employed, certain β-lactams, but not cephalosporins which were mostly inactive, displayed lower levels of M^{pro} inhibition when using an acoustic dispenser, which likely reflects their higher solubility in DMSO over the buffer. Consequently, inhibition assays were then only performed for dry dispensed compounds.

From the in-house screen, a penicillin-derived β-lactam inhibitor that covalently reacts with M^{pro}, was identified (Chapter 3). Structure activity relationship (SAR) studies were initiated to investigate the effect of the configuration of the C6 stereocenter, the nature of the C6 amido side chain, the nature of the C2 carboxyester, and the β-lactam group on inhibitor potency. Initially, the results of the MS M^{pro} inhibition assays were ambiguous, potentially due to compound-induced suppression of product peptide ionization; therefore, the MS inhibition assays were performed with a 37-mer peptide substrate rather than the previously used 11-mer substrate, which improved assay robustness while not detrimentally affecting assay results as revealed by the comparison of IC_{50} of reported inhibitors with the two different substrates. The thus optimised assay helped the SAR studies on the penicillin derivatives and enabled the identification of C6-bromo substituted penicillin sulfone derivatives as inhibitors with nanomolar potency. A crystal structure of one of the synthetic penicillin sulfone derivatives in complex with M^{pro}

confirmed its covalent reaction with the active site cysteine to form an acyl-enzyme complex.

Apart from small molecules, linear peptides were computationally designed and screened as potential substrate/inhibitors using the MS assays (Chapter 4). No turnover of the linear peptides was observed. However, substrate-competitive inhibition was observed in dose response studies (manuscript 3). Product mediated inhibition was also investigated and inhibition from both *N*-and C-terminal products was observed, suggesting product peptides can affect activity under the tested assay conditions. This was informative on designing the assay incorporating a C-terminal product peptide-derived internal peptide standard in Chapter 5. The internal peptide standard was added to the assay mixture after the reaction was quenched to prevent M^{pro} inhibition arising from the internal peptide. Short linear substrates (LQS, VLQS, and their derivatives) were screened as potential substrates to be used for X-ray time-resolved studies, however, only inhibitory activity was detected from tetrapeptide VLQS, suggesting the P3 residue is critical in binding. The inhibition from both the short linear peptide VLQS, and the computationally designed linear peptide is moderate. Thus, further optimisation is required to apply linear substrate based peptides as potent inhibitors.

In collaboration with Prof. Suga's group at the University of Tokyo, macrocyclic peptides were identified as inhibitors using the RaPID screening method; they were investigated for their inhibitory activity in MS assays (Chapter 5). The SPE-MS assay utilising a 37-mer substrate was optimised to reduce assay volume and total enzyme concentration to account for the high potency of the cyclic peptides and to include an internal peptide standard (Figure 9.2). Based on the crystal structure of TMGM4-05 (Ac-yFHγ^1LNLGYRPGc-NH_2) bound and SAR studies performed employing derivatives of TMGM4-05 and MS based inhibition assays, the FH motif was identified to be important for inhibition as it likely occupies the S2 and S1 subsites. Interestingly, an intrapeptide H3-Y9 interaction of the TMGM4-05 macrocyclic peptide was observed in the same crystal structure. TMGM4-05 with substitutions of H3 and Y9 residue showed that this H3-Y9 interaction is critical for potent TMGM4-05 activity.

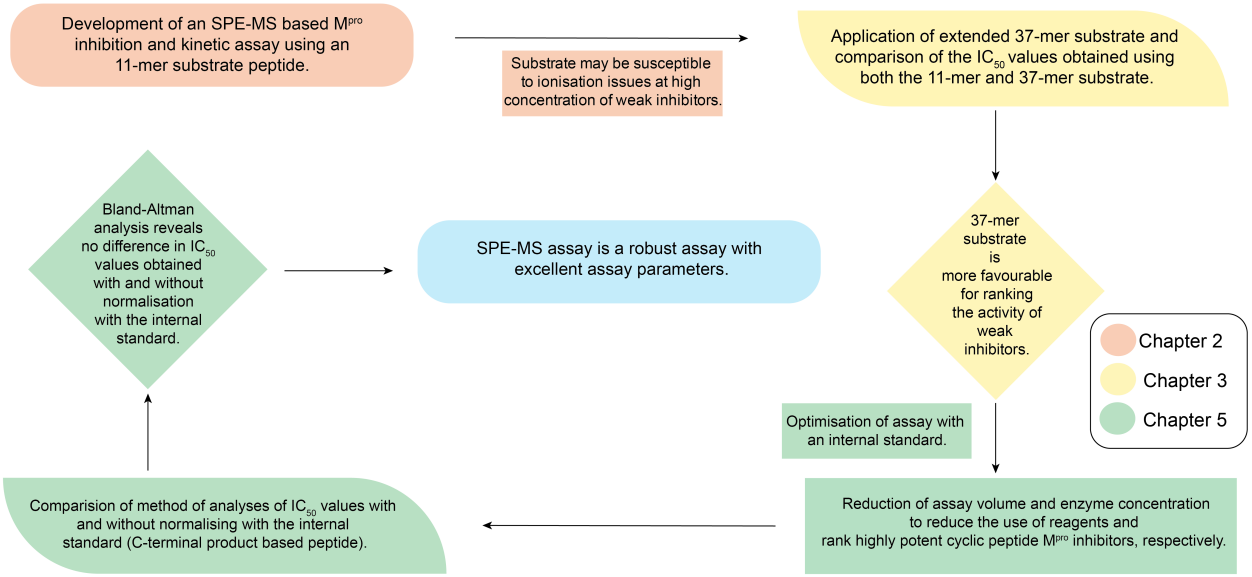


Figure 9.2 | Flowchart summarising the SPE-MS assay optimisation performed in chapters 2, 3 and 5.

Studies with TMGM4-05 variants at the P1′ position revealed that the γ-amino acid incorporated in the initially identified cyclic peptide inhibitors is not crucial for efficient inhibition; cyclic peptide derivatives with an acyclic γ-amino acid (TMGM4-05g) or natural amino acid (TMGM4-05) at the P1′ position manifested slightly lower or comparableIC_{50}s, respectively than TMGM4-05. It may be that the cyclic γ-amino acid at the P1′position improves serum stability of the cyclic peptide compared to variants with the linear γ-amino acid or natural peptide at the P1′ position.

Chapter 6 presents SAR studies on small molecule inhibitors other than penicillins. Ebsulfur compounds derived from ebselen were synthesized to improve the potency and selectivity of ebselen for covalent active site cysteine modification (Manuscript 5). The ebsulfur analogue of ebselen displayed similar potency to ebselen, suggesting an S-N unit is similarly as efficient as an Se-N unit in ebselen for potent inhibition. Protein-observed MS studies on the synthetic ebsulfur derivatives showed proof-of-concept evidence that the ebsulfur substitution pattern affects the reactivity of the compound with cysteine residues. While the covalent reaction of ebsulfur with all 11 cysteine residues was observed by MS, ebsulfur γ-lactam derivatives reacted only with two cysteines. In addition to ebsulfur derivatives, heterocyclic compounds (as part of the COVID Moonshot project) (Manuscript 6) and pentafluorosulfanyl-group containing compounds (Manuscript 7) were also investigated for

inhibition. Screening of >1000 compounds with P1 and P2 substituents (which bind to S1 and S2 subsites, respectively) led to the identification of a potent non-covalent inhibitor *N*-sulfonylated heterocycle (MAT-POS-e194df51-1). Pentafluorosulfanyl-substituted scaffolds in amide linked oxindole compounds also inhibited, providing evidence for the utility of the pentafluorosulfanyl group, which is more stable and lipophilic than the trifluoro group, for inhibitors.

Chapter 7 describes mechanistic studies with L,D-transpeptidases that use a Cys-His catalytic dyad identical to M^{pro}, and their reactions with β-lactams. Five Ldt_{Mt} paralogs present in *Mycobacterium tuberculosis* were investigated for their interactions with the penicillin, carbapenem, penem, and cephalosporin β-lactam subclasses using MS and NMR spectroscopy. The most clinically relevant Ldt_{Mt}, Ldt_{Mt2} catalyses C5-C6 fragmentation following reaction with penicillins and penems, but not with cephalosporins and carbapenems. All Ldts that form covalent complexes with penems, i.e. Ldt_{Mt1}, Ldt_{Mt3}, and Ldt_{Mt4}, were observed to catalyse C5-C6 fragmentation. In the case of the C6-side chain of the aminopenicillins, e.g. ampicillin, C5-C6 fragmentation is followed by cyclisation which leads to the formation of diketopiperazine products. Though hydrolysis of ampicillin was not observed with Ldt_{Mt2}, evidence for ampicillin hydrolysis is observed predominantly with Ldt_{Mt1}, and to a lower extent with Ldt_{Mt5}, Ldt_{Mt4} and Ldt_{Mt3}. Notably, evidence for C5-C6 fragmentation of ampicillin was not observed with Ldt_{Mt3}, supporting the notion that Ldt_{Mt}s may have non-redundant roles in *M. tuberculosis*.

Carbapenems form a stable acyl-enzyme complex upon reaction with Ldt_{Mt2}. Slow deacylation of the carbapenems from Ldt_{Mt2} complex was observed to occur via lactone formation and recyclisation. Evidence for recyclisation was observed more readily with Ldt_{fm}, an L,D-transpeptidase from *Enterococcus faecium*.

In Chapter 8, solution NMR studies were conducted with MBLs and SBLs with biapenem and faropenem. Multiple hydrolytic products including enamine, *2(R)*-imine and *2(S)*-imine hydrolytic products were observed. In the case of biapenem, the *2(S)*-imine hydrolytic product was the major hydrolytic product observed. Further experiments with doripenem and meropenem also suggested that the *2(S)*-imine is the major hydrolytic product. However, the *2(S)*-imine form of the hydrolytic product

may not be the nascent kinetic product, as is the case with imipenem where slow conversion of *2(R)*-imine to *2(S)*-imine was reported [364]. Additionally, the enamine product could only be characterised via the HMBC experiments, and future work should include detailed characterisation of this product.

In the case of faropenem, the tautomerisation of the enamine hydrolytic product to the imine product was also observed. However, in faropenem, the formation of the imine product was followed by C2-THF ring opening to give stable ring-opened imine alkene products with KPC-2, L1, NDM-1, OXA-10, OXA-23, and OXA-48. Crystallographic evidence corroborated the in-solution studies in the case of KPC-2 (class A β-lactamase), where the C2-tetrahydrofuran ring-open imine alkene product was observed. By contrast, C2-tetrahydrofuran ring-closed imine products were the major hydrolytic products observed with VIM-1 and VIM-2 in these solution studies. Surprisingly, a C2-tetrahydrofuran ring-closed imine product was captured with L1, and a C2-tetrahydrofuran ring-open imine alkene product was observed with VIM-2 by crystallography. The C2-tetrahydrofuran ring-closed imine product with VIM-1 and VIM-2 was observed to hydrolyse over time to give the C2-tetrahydrofuran ring-opened imine alkene product. However, no evidence for L1 with C2-tetrahydrofuran ring-closed product was observed in solution studies; the data obtained with L1 and VIM-2 suggests that the crystallographic complexes may not reflect the in-solution dynamics of a reaction. In the case of carbapenems, (*2R*) to (*2S*) epimerisation of the imine hydrolysis product is rapid. Only (*2S*) imine product was identified as the thermodynamic product, which was in accord with the crystallographic studies with VIM-2.

Overall, the work presented in this thesis led to the identification of a structurally diverse profile of M^{pro} inhibitors with potential for adaptation for antiviral drug development, i.e. β-lactams and macrocyclic peptide inhibitors. In addition, biochemical studies with β-lactams and L,D-transpeptidases, SBLs, and MBLs reveal that the various mechanisms employed by these enzymes upon reaction with β-lactams can be exploited to design improved β-lactam based antibiotics to tackle β-lactam resistance.

10 Conclusions

Vaccines for the treatment of COVID-19 have been developed and have been widely implemented. The microRNA, non-coding RNA-based vaccines are the first of their kind that have been approved for clinical use (as developed by Pfizer-BioNTech and Moderna [383]). The Oxford-AstraZeneca vaccine which is delivered using a modified adenovirus has been generously donated around the world [384]. Together with the development of cutting-edge vaccines, approval of nirmatrelvir (co-administered with ritonavir) [38] as an antiviral treatment is revolutionising the clinical care of COVID-19 patients.

Considerations on M^{pro} resistance

Vaccines cannot currently be administered to people on immunosuppressants for the treatment of chronic conditions which comprises about 3 million individuals in Europe [385]. Moreover, vaccines may not generate an adequate response in some people or may trigger allergic responses to reagents in vaccines or drugs in clinical therapy [42]. There thus remains a need for alternative COVID-19 therapies.

Much of the work described in this thesis focusing on SARS-CoV-2 aimed to help identify M^{pro} inhibitors for use in COVID-19 treatments. This work developed new M^{pro} assays and identified multiple new leads for M^{pro} inhibition, notably including acylating agents, e.g., β-lactams. These developments remain of interest, despite the clinical approval of a nirmatrelvir for COVID-19 therapy; the development of alternatives to nirmatrelvir that do not require co-administration with ritonavir is of interest. β-Lactams are of particular interest given their widespread use as antibacterials and excellent safety profile. History would also suggest that second generation M^{pro} inhibiting drugs

will also be required to combat potential resistance of SARS-CoV-2 to antivirals, as evidenced in the case of HIV/HCV treatments involving protease inhibitors [59, 386–388].

To date, several variants of SARS-CoV-2 have emerged, resulting in the onset of multiple waves of COVID infections. Analysis of the SARS-CoV-2 viral genome has suggested that some of the mutations occur in the nsp5 (M^{pro}) site. P132H, P132L, P132S, L89F, K90R, and K88R are some of the common substitutions that have been observed in the gene encoding for M^{pro} [389]. Notably, mutations are frequently observed away from the active site or dimerisation interface and hence variants of M^{pro} are observed to display similar catalytic efficiency against their substrates, and inhibitory constants against M^{pro} inhibitors, that is nirmatrelvir, GC-376, and PF-00835231, respectively [389]. In the case of the HCV protease, single-site mutations at Arg155, Ala156, and Asp168 apparently were the major cause of antiviral protease inhibitor drug resistance [386, 387]. These mutations usually occur at the active site but are found outside of the consensus volume (the region of the active site where the substrates interact [47]). Based on the observations with the proteases in HCV and other viruses, the SARS-CoV-2 virus may develop resistance against nirmatrelvir via mutation at residues that interact with nirmatrelvir and lie outside of the consensus volume of M^{pro}, e.g. the S2 subsite of M^{pro}. In this regard, it is of interest that nirmatrelvir has a bicyclic proline derivative at the P2 equivalent position which makes additional interactions with the S2 subsite of M^{pro} compared to the P2 leucine substrate residue, as observed in M^{pro}: N3 complex structure [38]. The larger bicyclic proline derivative in nirmatrelvir is likely to make interactions outside of the consensus volume, hence providing a potential hotspot for developing resistance via mutation of the S2 subsite residues which interact with the P2 substituent, outside of the P2 consensus volume.

Part of the work described in this thesis focused on the identification of different warheads with the potential to inhibit M^{pro} activity by applying a robust SPE-MS based assay (Chapter 2). Small covalent and heterocyclic non-covalent small molecules including β-lactam derivatives, acyclic peptides, and macrocyclic peptides were identified as M^{pro} inhibitors, as described in Chapters 2-6.

Different classes of drugs offer different advantages and suitability to be developed as drugs. Small molecule (molecular weight < 900 Da) inhibitors comprise 90% of pharmacologically used drugs. In

2021, of the 50 drugs approved by the FDA, half were small molecules [390]. Broadly classified as peptidomimetic or non-peptidomimetic, small molecules can act as covalent or non-covalent inhibitors, which have potentially different advantages and disadvantages. It can be argued that certain types of irreversibly covalently reacting drugs may be less prone to resistance, which is an immense challenge in infectious disease and anticancer therapy [391]. However, irreversibly covalently reacting inhibitors are not suitable when partial inhibition or short residency times are desired in a drug – though this is not normally the case in anti-infective targets. The electrophilic warheads present in covalent drugs, such as Michael acceptors and epoxide electrophiles, can be associated with toxicity [391]. In the cases of the ebsulfur derived and chloromethyl ketone based covalent M^{pro} inhibitors e.g. TPCK/TLCK, though some selectivity was observed for reaction with the active site cysteine in M^{pro}, toxicity may still be likely due to cross-reactions with cysteine nucleophile containing enzymes. Thus, non-covalent small molecules, such as heterocyclic compounds (MAT-POS-e194df51-1, Figure 6.1), which are potent and selective M^{pro} inhibitors may be preferable in terms of toxicity. However, both covalent (e.g., ebsulfur and chloromethyl ketone inhibitors) and non-covalent compounds (e.g., small heterocyclic molecules) remain susceptible to idiosyncratic toxicity (e.g. where drug metabolites may trigger the immune response [391]), thus, ongoing *in vivo* studies to investigate the potential of different types of M^{pro} inhibitors will be crucial in informing their therapeutic potential.

Peptide based inhibitors are useful to mimic catalytic units or substrate based selective inhibitors, however, compared to small molecules, peptides often have poor pharmacokinetic properties and cell permeability [334]. The linear M^{pro} peptide inhibitors (i.e. p12:KYTFWQYSQFY, p13:KYLTWQNSQIN, p15:LTINWQKYFNT, and p16:WFTLKQYWQTN, described in manuscript 3), which manifest moderate M^{pro} inhibition, are very unlikely to be useful as drugs, however, the SAR generated in the work may be useful in inhibitor development in understanding the substrate selectivity of M^{pro}, both with respect to viral and human substrates.

As described in Chapter 5, macrocyclic peptides e.g. TMGM4-05 and TMGM4-10, can be highly potent M^{pro} inhibitors and are more stable and possess superior pharmacokinetics compared to the linear peptides [95, 96, 334]. Optimisation of cell permeability may lead to the improvement

of cyclic peptides to the point where they are suitable for in vivo studies and ultimately clinical use. In addition, disulfide rich macrocyclic peptides e.g. MP103 which have different properties to the initially developed CP M^{pro} inhibitors can also be exploited further for therapeutic applications. It is reported that relatively low toxicity is observed with peptidomimetic inhibitors compared to non-peptidomimetic inhibitors in the context of HIV/HCV protease inhibition suggesting that peptidomimetic inhibitors of M^{pro} (e.g. CP derivatives) may provide superior toxicological profile compared to non-peptidomimetic inhibitors (e.g., ebsulfur derivatives) [47].

Though there is varying applicability of different classes of M^{pro} inhibitors as potential chemotherapeutic agents, the combined kinetic and inhibition studies presented here and elsewhere suggest that diverse substituents are tolerated in inhibitors binding at the S1 and S2 subsites in M^{pro} [28, 29, 37, 83, 337]. Information from the crystal structures of M^{pro} with its substrates and available inhibitors can be used to improve or design better M^{pro} inhibitors. The work in this thesis shows Gln (TMGM4-05 H3Q), His (TMGM4-05), or Trp (TMGM4-05 H3W) residues can be accommodated at the S1 subsite. Studies elsewhere with a calpain inhibitor and fragment screening have shown that the methionine side chain and other groups (e.g. a pyrimidine moiety), can also be accommodated at the S1 subsite [86, 248]. The results show that in addition to the preferred Leu residue at the S2 subsite, Phe (TMGM4-05) and Trp (s01_LP2W, manuscript 3) residues can be accommodated. The Phe at the P2 equivalent position of TMGM4-05 is orientated in an identical manner as the Phe in s02, and its binding mode partially overlaps with that of the bicyclic proline groups of nirmatrelvir (Figure 10.1). The tolerance of various substituents at the S1 and S2 subsites provides evidence for the plasticity of S1 and S2 subsites and the possibility that M^{pro} may accept a wider substrate pool during host infection than is presently perceived [25]. However, the accommodation of a side chain at a subsite does not necessarily correspond to the catalysis of the substrate counterpart. For example, TMGM405 H3W is active with a Trp at the P1 equivalent position, but no evidence of substrate turnover by MS was observed with the s01 variant with a Trp at the P1 position (s01_QP1W). Therefore, findings from inhibitor profiling for P1 and P2 equivalent substituents have to be translated cautiously in terms of interpreting the suitability of analogous groups/ binding modes in terms of predicting M^{pro} substrates [314]. This is

further evidenced by calpain inhibitors with Met at the P1 equivalent position, which is a potent M^{pro} inhibitor, while only poor turnover of substrate variants with P1 Met has been reported [392].

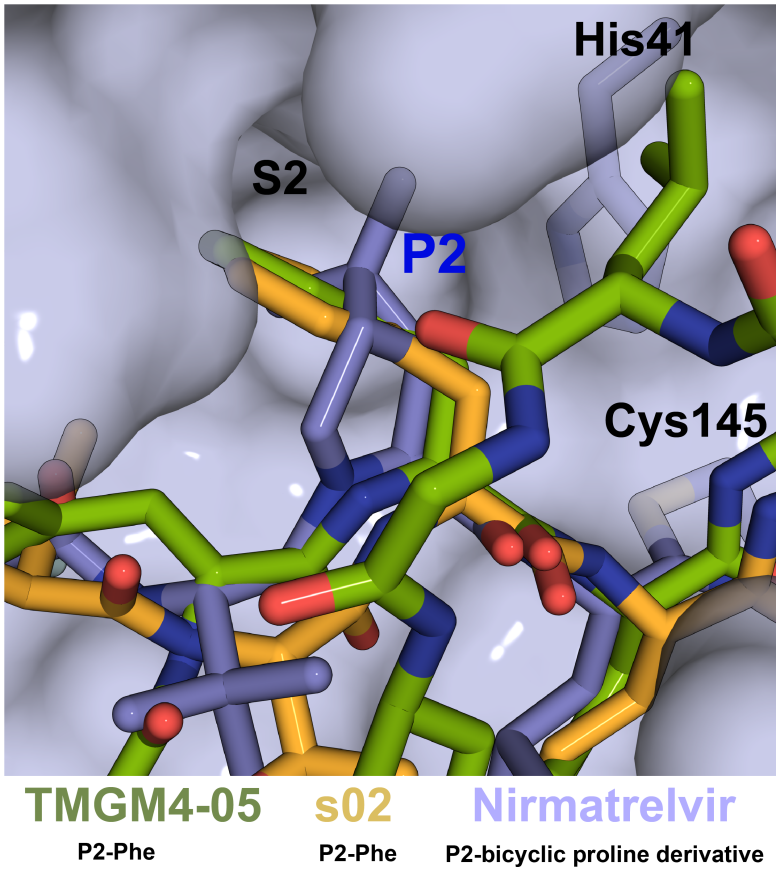


Figure 10.1 | Superimposition views from crystal structures of M^{pro} reacted to TMGM4-05, nirmatrelvir and the nsp4/5 cleavage site derived s02 peptide. The P2 equivalent substituent (Phe) in TMGM4-05 and Phe at the P2 of s02 partially overlap with the bicyclic proline moiety at the P2 equivalent position of nirmatrelvir.

The results show that some β-lactam inhibitors covalently react with M^{pro}, though a C6-methoxy penicillin sulfone (**5**, Chapter 2, Section, 2.3.5.9) represents an exception for which covalent modification was not observed. There is considerable scope for optimising β-lactams as covalently or non-covalently reacting M^{pro} inhibitors.

Evidence for β-lactam mediated covalent inhibition is widely reported with serine and cysteine proteases and transpeptidases [169, 173, 181, 183, 393–395]. Evidence for inhibition by non-covalent interactions is observed in *Staphylococcus aureus*, where ceftaroline, a cephalosporin subclass of β-lactam inhibits PBP2a by binding at an allosteric site, in addition to the covalent reaction with Ser70 residue, increasing the sensitivity of *Staphylococcus aureus* to ceftaroline [396, 397]. Monocyclic

β-lactam containing inhibitors, e.g., x10789, have been reported as M^{pro} inhibitors though their mechanism of action is not clear [248,314].

In a crystal structure of penicillin sulfone-reacted M^{pro} (Manuscript 3, PDB ID:7Z59), the sulfinic acid is positioned away from the S1 subsite; no interaction of the penicillin β-lactam was observed with the sidechain of Glu166, backbone carbonyl of Phe140 and imidazole of His163 in S1 subsite [303]. Similar interactions at S1 occur with the γ-lactam Gln side chain analogue at P1 in nirmatrelvir or the His at P1 in TMGM4-05 (Figure 10.2), and likely contribute to the high potency of these compounds. The C6 amido carbonyl of penicillin sulfone M^{pro} inhibitors interacts with the backbone NH group of Glu166; a similar interaction with the backbone carbonyl and NH groups of Glu166 has been observed with the backbone NH and carbonyl groups of P3 substituents in reported M^{pro} inhibitors [28,29,37,84]. The C6-phenyl group of β-lactam is positioned to form π-π stacking interactions with the imidazole ring of catalytic His41. Inhibition studies have shown that the phenoxyacetyl oxygen atom is important in enabling this π-π stacking interaction and hence the activity of penicillin sulfone compounds. By contrast, the phenylalanine at P2 of TMGM4-05 is pointed away from the His41 and thus, is devoid of π-stacking observed in an offset manner with phenyl ring of penicillin β-lactam. The crystal structure of the reacted penicillin sulfone in complex with M^{pro} provides insight that will be useful to enhance the potency of improved C6-penicillin penicillin derivatives, potentially by substituting the C6 proton of phenoxyacetyl penicillin sulfone with a substituent which can form interactions at the S1 subsite.

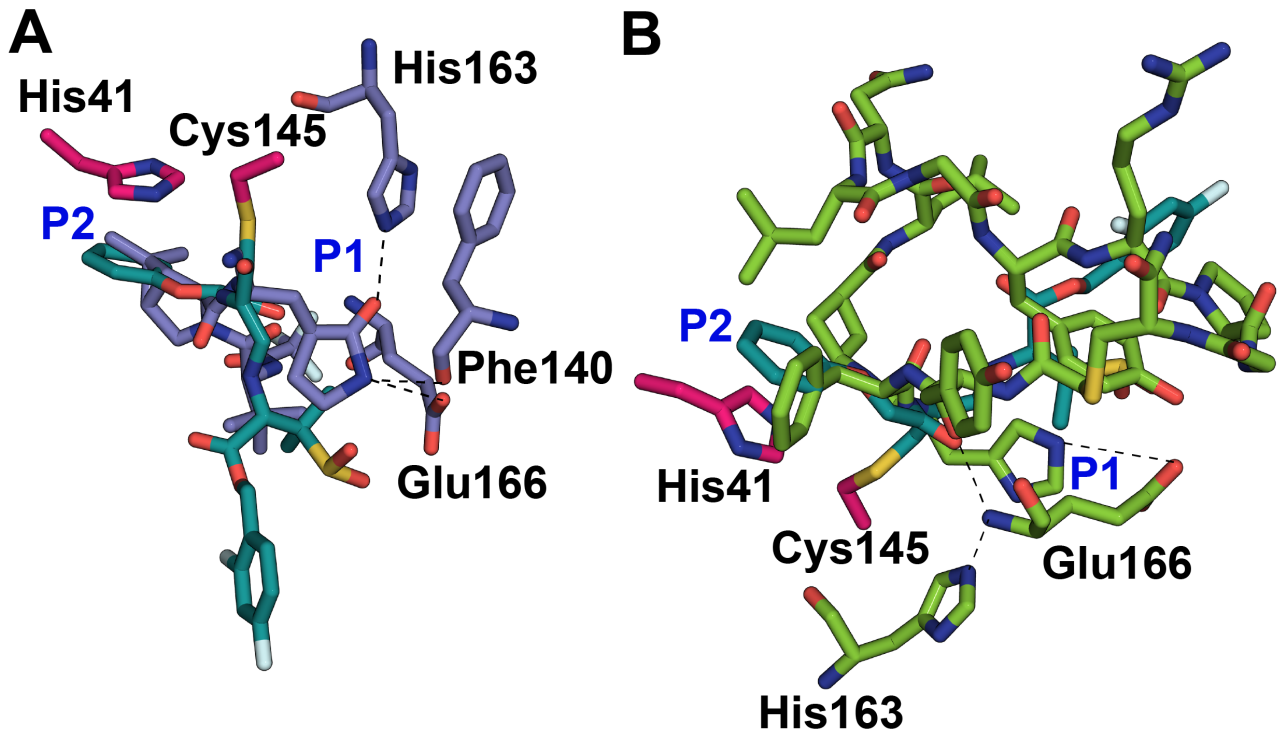


Figure 10.2 | Comparisons of the interaction of the P1 and P2 substituents of TMGM4-05 and nirmatrelvir with penicillin sulfone (teal blue) reacted tp M^{pro}. A. Superimposition views of M^{pro} reacted to nirmatrelvir (violet blue) and penicillin sulfone reacted M^{pro}. **B.** Superimposition views of TMGM4-05 (green) reacted with M^{pro} with penicillin sulfone reacted M^{pro} (PDB:7Z59) [303].

As reported in this thesis, β-lactams can potently inhibit M^{pro}; they are also a safe and efficacious class of antibiotics that target serine transpeptidases (i.e., PBPs) in bacteria. The clinical use of β-lactams has enabled treatment against many pathogenic bacteria e.g., staphylococci, enteric bacteria, anaerobes, *Pseudomonas aeruginosa*, and *Salmonella* spp., and they remain the most widely used class of antibiotics for treatment of infectious diseases [398]. Alterations in PBPs via spontaneous mutations, limited access to PBPs via restricted entry or increased activity of efflux pumps, and inactivation of β-lactams by the acquisition of β-lactamases have resulted in β-lactam resistance [399]. The latter is the most prevalent mechanism of β-lactam resistance in Gram negative bacteria [400].

Following the initial reaction with PBPs, SBLs, and MBLs, subsequent non-hydrolytic reactions of the enzyme-inhibitor complexes can occur. Understanding the nature of these reactions is important in deducing how the inhibitors work and in optimising their potency. In addition to understanding the mechanism of inhibition of antibacterial targets (PBPs, SBLs, MBLs), there is evidence that such

information is relevant to the inhibition of proteases by β-lactams as evidenced in the inhibition of M^{pro} by penicillin sulfones as described above and in prior work, including on elastase inhibition, where conformational changes following initial reaction were shown to be important [401].

As described in this thesis, in the case of faropenem, reaction with L1, VIM-1, and VIM-2 (all MBLs), gives an initial enamine product which tautomerises to give an imine product. Crystallographic work (described in manuscript 8) shows evidence for the formation of an imine product with L1 and an SBL, KPC-2. However, a crystal structure with a class A SBL SED-G238C (PDB:3BFF) shows evidence of an enamine at the active site. Tautomerisation of enamine to imine 'intermediates' has also been extensively reported with carbapenems, in particular with respect to their reactions with SBLs and MBLs [364]. In the case of SBLs (at least in some cases) it is proposed that an enamine is initially formed and can undergo hydrolysis to give an enamine product in solution which can undergo tautomerisation to give imine products. Tautomerisation from enamine to imine forms can also occur on the protein [224, 233]. There have been fewer studies of the reactivity of inhibitor-derived acyl-enzyme complexes of the bacterial nucleophilic cysteine transpeptidases Ldts which are nucleophilic cysteine enzymes and which were the subject of the studies described in Chapter 7.

Studies with the Ldts have provided insights into the mechanism of their inhibition by β-lactams. Evidence was presented that penicillins can undergo C5-C6 fragmentation, deacylation of aminopenicillins by diketopiperazine formation can occur, carbapenems can undergo lactone formation, carbapenems and cephalosporins can undergo deacylation by reversible recyclization, and that D-amino acids can enable deacylation of acyl-enzyme complexes. In corroboration with the observations with carbapenem and Ldts reported here, recently reported MS studies by others have provided evidence for lactone formation and reversible reaction with carbapenems and Ldt_{Mt2} [402]. Collectively, these and other studies reveal that inhibition of nucleophilic cysteine and serine enzymes by β-lactams involves mechanisms beyond simple acylation. The rational optimisation of these processes beyond initial acylation with respect to inhibition of different classes of a target for β-lactams is of interest.

Thus, the different mechanisms that are prevalent in infections caused by bacteria and other viral diseases set the precedent to monitor the resistance mechanisms in SARS-CoV-2 drug tar-

gets, M^{pro} and PL^{pro} to develop a diverse portfolio of alternative potent antivirals in anticipation of future drug resistance.

Bibliography

[1] Barrett, A. J. Bioinformatics of proteases in the MEROPS database. *Curr. Opin. Drug. Disc. Dev.* **7**, 334–341 (2004).

[2] Quesada, V., Ordonez, G. R., Sanchez, L. M., Puente, X. S. & López-Otín, C. The degradome database: mammalian proteases and diseases of proteolysis. *Nu. Acid. Res.* **37**, D239–D243 (2009).

[3] López-Otín, C. & Bond, J. S. Proteases: multifunctional enzymes in life and disease. *J. Biol. Chem.* **283**, 30433–30437 (2008).

[4] Riedl, S. J. & Salvesen, G. S. The apoptosome: signalling platform of cell death. *Nat. Rev. Mol. Cell. Bio.* **8**, 405–413 (2007).

[5] Sela-Passwell, N., Rosenblum, G., Shoham, T. & Sagi, I. Structural and functional bases for allosteric control of mmp activities: can it pave the path for selective inhibition? *Biochim. et Biophys. Acta. (BBA)-Molecular Cell Res.* **1803**, 29–38 (2010).

[6] Richter, C., Tanaka, T. & Yada, R. Y. Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. *Biochem. J.* **335**, 481–490 (1998).

[7] Organization, W. H. *et al.* Latest HIV estimates and updates on HIV policies uptake (2020).

[8] Salam, K. A. & Akimitsu, N. Hepatitis C virus NS3 inhibitors: current and future perspectives. *BioMed Res. in.* **2013** (2013).

[9] Rawlings, N. D. *et al.* The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the panther database. *Nu. Acid. Res.* **46**, D624–D632 (2018).

[10] Buller, A. R. *et al.* Autoproteolytic activation of thnt results in structural reorganization necessary for substrate binding and catalysis. *J. Mol. Bio.* **422**, 508–518 (2012).

[11] Guan, C. *et al.* Activation of glycosylasparaginase: Formation of active N-terminal threonine by intramolecular autoproteolysis. *J. Biol. Chem.* **271**, 1732–1737 (1996).

[12] Love, R. A. *et al.* The crystal structure of hepatitis c virus ns3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **87**, 331–342 (1996).

[13] Rodríguez, A., Oliva, C. & González, M. A comparative QM/MM study of the reaction mechanism of the Hepatitis C virus ns3/ns4a protease with the three main natural substrates NS5A/5B, NS4B/5A and NS4A/4B. *Phys. Chem. Chem. Phys.* **12**, 8001–8015 (2010).

[14] Hofmann, T., Hodges, R. S. & James, M. N. Effect of ph on the activities of penicillopepsin and Rhizopus pepsin and a proposal for the productive substrate binding mode in penicillopepsin. *Biochemistry* **23**, 635–643 (1984).

[15] Xu, R.-H. *et al.* Epidemiologic clues to SARS origin in China. *Emerg. Infect. Dis.* **10**, 1030 (2004).

[16] Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet Infect. Dis.* **20**, 533–534 (2020).

[17] Mackenzie, J. *et al.* The WHO response to SARS and preparations for the future. In *Learning from SARS: preparing for the next disease outbreak: workshop summary*, 42 (National Academies Press, 2004).

[18] Wu, A. *et al.* Genome composition and divergence of the novel coronavirus (2019-ncov) originating in china. *Cell Host. & Microbe.* **27**, 325–328 (2020).

[19] Thiel, V. *et al.* Mechanisms and enzymes involved in sars coronavirus genome expression. *J. Gen. Virol.* **84**, 2305–2315 (2003).

[20] Ziebuhr, J. Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **7**, 412–419 (2004).

[21] Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. Virus-encoded proteinases and proteolytic processing in the nidovirales. *J. Gen. Virol.* **81**, 853–879 (2000).

[22] Ziebuhr, J., Thiel, V. & Gorbalenya, A. E. The autocatalytic release of a putative rna virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *J. Biol. Chem.* **276**, 33220–33232 (2001).

[23] Berger, A. & Schechter, I. On the size of active site in proteases. I. Papain. *Biochem. Biophys Res Commun* **27**, 157–162 (1967).

[24] Hegyi, A. & Ziebuhr, J. Conservation of substrate specificities among coronavirus main proteases. *J. Gen. Virol.* **83**, 595–599 (2002).

[25] Hameedi, M. A. *et al.* Structural and functional characterization of NEMO cleavage by SARS-CoV-2 3CLpro. *Biorxiv* (2021).

[26] Koudelka, T. *et al. N*-terminomics for the identification of in vitro substrates and cleavage site specificity of the SARS-CoV-2 main protease. *Proteomics* **21**, 2000246 (2021).

[27] O'Loughlin, T. *et al.* OPTN recruitment to a golgi-proximal compartment regulates immune signalling and cytokine secretion. *J. Cell. Sci.* **133**, jcs239822 (2020).

[28] Jin, Z. *et al.* Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).

[29] Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).

[30] Grottesi, A. *et al.* Computational studies of SARS-CoV-2 3CLpro: Insights from MD simulations. *Int. J. Mol. Sci.* **21**, 5346 (2020).

[31] Fan, K. *et al.* Biosynthesis, purification, and substrate specificity of severe acute respiratory syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **279**, 1637–1642 (2004).

[32] Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).

[33] Mengist, H. M., Dilnessa, T. & Jin, T. Structural basis of potential inhibitors targeting SARS-CoV-2 main protease. *Front. Chem.* 7 (2021).

[34] Osipiuk, J. *et al.* Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nat. Commun.* **12**, 1–9 (2021).

[35] Hermanns, T., Pichlo, C., Baumann, U. & Hofmann, K. A structural basis for the diverse linkage specificities within the zufsp deubiquitinase family. *Nat. Commun.* **13**, 1–14 (2022).

[36] Hilgenfeld, R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS J.* **281**, 4085–4096 (2014).

[37] Zhang, L. *et al.* α-Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: structure-based design, synthesis, and activity assessment. *J. Med. Chem.* **63**, 4562–4578 (2020).

[38] Owen, D. R. *et al.* An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of covid-19. *Science* **374**, 1586–1593 (2021).

[39] Lamb, Y. N. Nirmatrelvir Plus Ritonavir: First Approval. *Drugs* 1–7 (2022).

[40] Medicines & products Regulatory Agency, H. First oral antiviral for COVID-19, lagevrio (molnupiravir), approved by MHRA (2021).

[41] Kabinger, F. *et al.* Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat. structural Mol. Bio.* **28**, 740–746 (2021).

[42] Wise, J. Covid-19: Evusheld is approved in UK for prophylaxis in immunocompromised people (2022).

[43] Medicines & products Regulatory Agency, H. Evusheld approved to prevent COVID-19 in people whose immune response is poor (2022).

[44] Thakur, S. *et al.* Exploring the magic bullets to identify achilles' heel in SARS-CoV-2: Delving deeper into the sea of possible therapeutic options in covid-19 disease: An update. *Food Chem. Toxicol.* **147**, 111887 (2021).

[45] Hoofnagle, J. H. Livertox: a website on drug-induced liver injury. In *Drug-Induced Liver Disease*, 725–732 (Elsevier, 2013).

[46] Zhou, S. *et al.* β-dn 4-hydroxycytidine (nhc) inhibits SARS-CoV-2 through lethal mutagenesis but is also mutagenic to mammalian cells. *The J. Infect. Dis.* (2021).

[47] Zephyr, J., Yilmaz, N. K. & Schiffer, C. A. Viral proteases: Structure, mechanism and inhibition. In *The Enzymes*, vol. 50, 301–333 (Elsevier, 2021).

[48] Tözsér, J. Comparative studies on retroviral proteases: substrate specificity. *Viruses* **2**, 147–165 (2010).

[49] Dunn, B. M., Goodenow, M. M., Gustchina, A. & Wlodawer, A. Retroviral proteases. *Genome Bio.* **3**, 1–7 (2002).

[50] Kausar, S. *et al.* A review: Mechanism of action of antiviral drugs. *Int. J. Immunopathol. Pharmacol.* **35**, 20587384211002621 (2021).

[51] Vardanyan, R. & Hruby, V. Antiviral drugs. *Synth. best-seller Drug.* 687 (2016).

[52] Njardarson, J. *et al.* Top 200 pharmaceuticals by retail sales in 2019. *Univ. Ariz.* (2019).

[53] Organization, W. H. *et al.* Global health sector strategy on viral hepatitis 2016-2021. towards ending viral hepatitis. Tech. Rep., World Health Organization (2016).

[54] Lv, Z., Chu, Y. & Wang, Y. HIV protease inhibitors: a review of molecular selectivity and toxicity. *HIV/AIDS (Auckland, NZ)* **7**, 95 (2015).

[55] Ghosh, A. K., Osswald, H. L. & Prato, G. Recent progress in the development of HIV-1 protease inhibitors for the treatment of hiv/aids. *J. Med. Chem.* **59**, 5172–5208 (2016).

[56] Park, K., Kwon, I. C. & Park, K. Oral protein delivery: Current status and future prospect. *React. Funct. Polym.* **71**, 280–287 (2011).

[57] James, J. S. Saquinavir (invirase): first protease inhibitor approved–reimbursement, information hotline numbers. *AIDS treatment news* 1–2 (1995).

[58] Pokorná, J., Machala, L., Řezáčová, P. & Konvalinka, J. Current and novel inhibitors of HIV protease. *Viruses* **1**, 1209–1239 (2009).

[59] Agnello, S., Brand, M., Chellat, M. F., Gazzola, S. & Riedl, R. A structural view on medicinal chemistry strategies against drug resistance. *Angew. Chemie Int. Ed.* **58**, 3300–3345 (2019).

[60] Pokorná, J., Machala, L., Řezáčová, P. & Konvalinka, J. Current and novel inhibitors of HIV protease. *Viruses* **1**, 1209–1239 (2009).

[61] Tomasselli, A. G. & Heinrikson, R. L. Targeting the HIV-protease in aids therapy: a current clinical perspective. *Biochim. et Biophys. Acta. (BBA)-Protein Struct. Mol. Enzymol.* **1477**, 189–214 (2000).

[62] McCoy, C. Darunavir: a nonpeptidic antiretroviral protease inhibitor. *Clin. Ther.* **29**, 1559–1576 (2007).

[63] Hoofnagle, J. H. Livertox: a website on drug-induced liver injury. In *Drug-Induced Liver Disease*, 725–732 (Elsevier, 2013).

[64] Viekirax, E. Summary of product characteristics (2018).

[65] Chen, K. X. & Njoroge, F. G. The journey to the discovery of boceprevir: an NS3−NS4 HCV protease inhibitor for the treatment of chronic Hepatitis C. *Prog. Med. Chem.* **49**, 1–36 (2010).

[66] Kwong, A. D., Kauffman, R. S., Hurter, P. & Mueller, P. Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nat. biotechnology* **29**, 993–1003 (2011).

[67] Clark, V. C., Peter, J. A. & Nelson, D. R. New therapeutic strategies in hcv: second-generation protease inhibitors. *Liver Int.* **33**, 80–84 (2013).

[68] Scola, P. M. *et al.* Discovery and early clinical evaluation of bms-605339, a potent and orally efficacious tripeptidic acylsulfonamide NS3 protease inhibitor for the treatment of hepatitis C virus infection. *J. Med. Chem.* **57**, 1708–1729 (2014).

[69] Imperiali, B. & Abeles, R. H. Extended binding inhibitors of chymotrypsin that interact with leaving group subsites S1'-S3'. *Biochemistry* **26**, 4474–4477 (1987).

[70] Echeverría, N., Moratorio, G., Cristina, J. & Moreno, P. Hepatitis c virus genetic variability and evolution. *World J. Hepatol.* **7**, 831 (2015).

[71] Izquierdo, L. *et al.* Simeprevir for the treatment of hepatitis C virus infection. *Pharmacogenomics Pers. Med.* **7**, 241 (2014).

[72] Moreno, C. *et al.* P1319 once-daily simeprevir (tmc435) with peginterferon/ribavirin in treatment-naive or treatment-experienced chronic hcv genotype 4-infected patients: Final results of a phase III trial. *J. Hepatol.* **1**, S535 (2014).

[73] Ljungman, P. Herpes virus infections in immunocompromised patients: problems and therapeutic interventions. *Annals Med.* **25**, 329–333 (1993).

[74] Tandon, R., Mocarski, E. S. & Conway, J. F. The a, b, cs of herpesvirus capsids. *Viruses* **7**, 899–914 (2015).

[75] V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).

[76] Food, U., Administration, D. *et al.* Fact sheet for health care providers emergency use authorization (EUA) of remdesivir (GS-5734™). *Silver Spring, MD, US Food Drug Adm.* (2020).

[77] Fu, L. *et al.* Both boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **11**, 1–8 (2020).

[78] Vuong, W. *et al.* Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **11**, 1–8 (2020).

[79] Redhead, M. A. *et al.* Bispecific repurposed medicines targeting the viral and immunological arms of covid-19. *Sci. Rep.* **11**, 1–14 (2021).

[80] Hoffman, R. L. *et al.* Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of covid-19. *J. Med. Chem.* **63**, 12725–12747 (2020).

[81] Huang, C., Wei, P., Fan, K., Liu, Y. & Lai, L. 3c-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry* **43**, 4568–4574 (2004).

[82] Solowiej, J. *et al.* Steady-state and pre-steady-state kinetic evaluation of severe acute respiratory syndrome coronavirus (SARS-CoV) 3CLpro cysteine protease: development of an ion-pair model for catalysis. *Biochemistry* **47**, 2617–2630 (2008).

[83] Qiao, J. *et al.* Sars-cov-2 Mpro inhibitors with antiviral activity in a transgenic mouse model. *Science* **371**, 1374–1378 (2021).

[84] Li, L. & Huang, S. Newly synthesized Mpro inhibitors as potential oral anti-SARS-CoV-2 agents. *Signal Transduct. Target. Ther.* **6**, 1–2 (2021).

[85] Chamakuri, S. *et al.* Dna-encoded chemistry technology yields expedient access to SARS-CoV-2 Mpro inhibitors. *Proc. Natl. Acad. Sci.* **118** (2021).

[86] Sacco, M. D. *et al.* Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci. Adv.* **6**, eabe0751 (2020).

[87] Dougherty, P. G., Sahni, A. & Pei, D. Understanding cell penetration of cyclic peptides. *Chem. Rev.* **119**, 10241–10287 (2019).

[88] Agbowuro, A. A., Huston, W. M., Gamble, A. B. & Tyndall, J. D. Proteases and protease inhibitors in infectious diseases. *Med. Res. Rev.* **38**, 1295–1331 (2018).

[89] Tan, J. *et al.* 3C protease of enterovirus 68: structure-based design of Michael acceptor inhibitors and their broad-spectrum antiviral effects against picornaviruses. *J. Virol.* **87**, 4339–4351 (2013).

[90] Dragovich, P. S. *et al.* Solid-phase synthesis of irreversible human rhinovirus 3C protease inhibitors. part 1: Optimization of tripeptides incorporating n-terminal amides. *Bioorganic & Med. Chem.* **7**, 589–598 (1999).

[91] Cunha, F. M. *et al.* Intracellular peptides as natural regulators of cell signaling. *J. biological Chem.* **283**, 24448–24459 (2008).

[92] Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A. & Ghasemi, Y. A comprehensive review of signal peptides: Structure, roles, and applications. *Eur. J. Cell. Bio.* **97**, 422–441 (2018).

[93] Joo, S. H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol. Ther.* **20**, 19 (2012).

[94] Huan, Y., Kong, Q., Mou, H. & Yi, H. Antimicrobial peptides: classification, design, application and research progress in multiple fields. *Front. Microbiol.* 2559 (2020).

[95] Sohrabi, C., Foster, A. & Tavassoli, A. Methods for generating and screening libraries of genetically encoded cyclic peptides in drug discovery. *Nat. Rev. Chem.* **4**, 90–101 (2020).

[96] Zhang, H. & Chen, S. Cyclic peptide drugs approved in the last two decades (2001-2021). *RSC Chem. Biol.* (2021).

[97] Abdalla, M. A. & McGaw, L. J. Natural cyclic peptides as an attractive modality for therapeutics: a mini review. *Molecules* **23**, 2080 (2018).

[98] Zhao, S. M. *et al.* Antitumor cyclic hexapeptides from rubia plants: history, chemistry, and mechanism (2005-2011). *CHIMIA Int. J. for Chem.* **65**, 952–956 (2011).

[99] Tan, N.-H. & Zhou, J. Plant cyclopeptides. *Chem. Rev.* **106**, 840–895 (2006).

[100] Abdalla, M. A. Three new cyclotetrapeptides isolated from streptomyces sp. 447. *Nat. product Res.* **31**, 1014–1021 (2017).

[101] Falanga, A. *et al.* Cyclic peptides as novel therapeutic microbicides: engineering of human defensin mimetics. *Molecules* **22**, 1217 (2017).

[102] Bucy, R., Xu, X. Y., Li, J. & Huang, G. Cyclosporin a-induced autoimmune disease in mice. *The J. Immunol.* **151**, 1039–1050 (1993).

[103] Miach, P. J. Cyclosporin a in organ transplantation. *Med. J. Aust.* **145**, 146–150 (1986).

[104] Katsu, T., Imamura, T., Komagoe, K., Masuda, K. & Mizushima, T. Simultaneous measurements of k+ and calcein release from liposomes and the determination of pore size formed in a membrane. *Anal. Sci.* **23**, 517–522 (2007).

[105] Holland, T. L., Arnold, C. & Fowler, V. G. Clinical management of staphylococcus aureus bacteremia: a review. *Jama* **312**, 1330–1341 (2014).

[106] Terlain, B. & Thomas, J. Structure of griselimycin, polypeptide antibiotic extracted streptomyces cultures. i. identification of the products liberated by hydrolysis. *Bull. de la Soc. chimique de France* **6**, 2349–2356 (1971).

[107] Kling, A. *et al.* Targeting dnan for tuberculosis therapy using novel griselimycins. *Science* **348**, 1106–1112 (2015).

[108] Bao, J. *et al.* New cyclic tetrapeptides and asteltoxins from gorgonian-derived fungus aspergillus sp. scsgaf 0076. *Tetrahedron* **69**, 2113–2117 (2013).

[109] Liang, X., Nong, X.-H., Huang, Z.-H. & Qi, S.-H. Antifungal and antiviral cyclic peptides from the deep-sea-derived fungus simplicillium obclavatum eiodsf 020. *J. Agric. Food Chem.* **65**, 5114–5121 (2017).

[110] Jing, X. & Jin, K. A gold mine for drug discovery: Strategies to develop cyclic peptides into therapies. *Med. Res. Rev.* **40**, 753–810 (2020).

[111] Ványolós, A. *et al.* Gymnopeptides a and b, cyclic octadecapeptides from the mushroom gymnopus fusipes. *Org. Lett.* **18**, 2688–2691 (2016).

[112] Li, J., Hu, S., Jian, W., Xie, C. & Yang, X. Plant antimicrobial peptides: structures, functions, and applications. *Bot. Stud.* **62**, 1–15 (2021).

[113] Smith, J. J., Cummins, T. R., Alphy, S. & Blumenthal, K. M. Molecular interactions of the gating modifier toxin protx-ii with nav1. 5: implied existence of a novel toxin binding site coupled to activation. *J. Biol. Chem.* **282**, 12687–12697 (2007).

[114] Colgrave, M. L., Korsinczky, M. J., Clark, R. J., Foley, F. & Craik, D. J. Sunflower trypsin inhibitor-1, proteolytic studies on a trypsin inhibitor peptide and its analogs. *Pept. Sci.* **94**, 665–672 (2010).

[115] Wang, C. K. & Craik, D. J. Designing macrocyclic disulfide-rich peptides for biotechnological applications. *Nat. Chem. Bio.* **14**, 417–427 (2018).

[116] Luckett, S. *et al.* High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Bio.* **290**, 525–533 (1999).

[117] Olivera, B. M. *et al.* Peptide neurotoxins from fish-hunting cone snails. *Science* **230**, 1338–1343 (1985).

[118] Olivera, B. M. *et al.* Diversity of conus neuropeptides. *Science* **249**, 257–263 (1990).

[119] Azam, L. & McIntosh, J. M. Alpha-conotoxins as pharmacological probes of nicotinic acetylcholine receptors. *Acta. Pharmacol. Sinica* **30**, 771–783 (2009).

[120] Sandall, D. *et al.* A novel α-conotoxin identified by gene sequencing is active in suppressing the vascular response to selective stimulation of sensory nerves in vivo. *Biochemistry* **42**, 6904–6911 (2003).

[121] Clark, R. J. *et al.* The engineering of an orally active conotoxin for the treatment of neuropathic pain. *Angew. Chemie Int. Ed.* **49**, 6545–6548 (2010).

[122] Satkunanathan, N. *et al.* Alpha-conotoxin vc1. 1 alleviates neuropathic pain and accelerates functional recovery of injured neurones. *Brain Res.* **1059**, 149–158 (2005).

[123] Yu, R. *et al.* Less is more: design of a highly stable disulfide-deleted mutant of analgesic cyclic α-conotoxin vc1. 1. *Sci. Rep.* **5**, 1–11 (2015).

[124] R Stroud, M., J Hansen, S. & M Olson, J. In vivo bio-imaging using chlorotoxin-based conjugates. *Curr. pharmaceutical Des.* **17**, 4362–4371 (2011).

[125] Dardevet, L. *et al.* Chlorotoxin: a helpful natural scorpion peptide to diagnose glioma and fight tumor invasion. *Toxins* **7**, 1079–1101 (2015).

[126] Craik, D. J., Daly, N. L., Bond, T. & Waine, C. Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Bio.* **294**, 1327–1336 (1999).

[127] Colgrave, M. L. & Craik, D. J. Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. *Biochemistry* **43**, 5965–5975 (2004).

[128] Henriques, S. T. *et al.* The prototypic cyclotide kalata b1 has a unique mechanism of entering cells. *Chem. Bio.* **22**, 1087–1097 (2015).

[129] Thongyoo, P., Roqué-Rosell, N., Leatherbarrow, R. J. & Tate, E. W. Chemical and biomimetic total syntheses of natural and engineered mcoti cyclotides. *Org. Biomol. Chem.* **6**, 1462–1470 (2008).

[130] Sommerhoff, C. P. *et al.* Engineered cystine knot miniproteins as potent inhibitors of human mast cell tryptase β. *J. Mol. Bio.* **395**, 167–175 (2010).

[131] Glotzbach, B. *et al.* Combinatorial optimization of cystine-knot peptides towards high-affinity inhibitors of human matriptase-1. *PloS one* **8**, e76956 (2013).

[132] Chan, L. Y. *et al.* Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds. *Blood, The J. Am. Soc. Hematol.* **118**, 6709–6717 (2011).

[133] Wong, C. T. *et al.* Orally active peptidic bradykinin b1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. *Angew. Chemie Int. Ed.* **51**, 5620–5624 (2012).

[134] Eliasen, R. *et al.* Design, synthesis, structural and functional characterization of novel melanocortin agonists based on the cyclotide kalata b1. *J. Biol. Chem.* **287**, 40493–40501 (2012).

[135] Thongyoo, P., Bonomelli, C., Leatherbarrow, R. J. & Tate, E. W. Potent inhibitors of β-tryptase and human leukocyte elastase based on the mcoti-ii scaffold. *J. Med. Chem.* **52**, 6197–6200 (2009).

[136] Aboye, T. L. *et al.* Design of a novel cyclotide-based cxcr4 antagonist with anti-human immunodeficiency virus (HIV)-1 activity. *J. Med. Chem.* **55**, 10729–10734 (2012).

[137] Wang, C. K. *et al.* Molecular grafting onto a stable framework yields novel cyclic peptides for the treatment of multiple sclerosis. *ACS Chem. Bio.* **9**, 156–163 (2014).

[138] Bozovičar, K. & Bratkovič, T. Evolving a peptide: library platforms and diversification strategies. *Int. J. Mol. Sci.* **21**, 215 (2019).

[139] Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu. Rev. Biochem.* **83**, 727–752 (2014).

[140] Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).

[141] Bazan, J., Całkosiński, I. & Gamian, A. Phage display—a powerful technique for immunotherapy: 1. Introduction and potential of therapeutic applications. *Hum. vaccines & immunotherapeutics* **8**, 1817–1828 (2012).

[142] Kawakami, T., Murakami, H. & Suga, H. Ribosomal synthesis of polypeptoids and peptoid-peptide hybrids. *J. Am. Chem. Soc.* **130**, 16861–16863 (2008).

[143] Passioura, T. & Suga, H. A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem. Commun.* **53**, 1931–1940 (2017).

[144] Oba, M., Kunitake, M., Kato, T., Ueda, A. & Tanaka, M. Enhanced and prolonged cell-penetrating abilities of arginine-rich peptides by introducing cyclic α, α-disubstituted α-amino acids with stapling. *Bioconjug. Chem.* **28**, 1801–1806 (2017).

[145] Liu, W. *et al.* An ultrapotent and selective cyclic peptide inhibitor of human β-factor xiia in a cyclotide scaffold. *J. Am. Chem. Soc.* **143**, 18481–18489 (2021).

[146] Johansen-Leete, J. *et al.* Antiviral cyclic peptides targeting the main protease of SARS-CoV-2. *Chem. Sci.* **13**, 3826–3836 (2022).

[147] Kreutzer, A. G. *et al.* A cyclic peptide inhibitor of the SARS-CoV-2 main protease. *Eur. J. Med. Chem.* **221**, 113530 (2021).

[148] Al-Wahaibi, L. H. *et al.* Discovery of novel oxazole-based macrocycles as anti-coronaviral agents targeting SARS-CoV-2 main protease. *Bioorganic Chem.* **116**, 105363 (2021).

[149] Seung, K. J., Keshavjee, S. & Rich, M. L. Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harb. Perspect. Med.* **5**, a017863 (2015).

[150] Talbot, E. & Pai, M. Tackling drug-resistant tuberculosis: we need a critical synergy of product and process innovations. *The Int. J. Tuberc. Lung Dis.* **23**, 774–782 (2019).

[151] Chakaya, J. *et al.* Global tuberculosis report 2020–reflections on the global tb burden, treatment and prevention efforts. *Int. J. Infect. Dis.* **113**, S7–S12 (2021).

[152] Buchmann, J. P. & Holmes, E. C. Cell walls and the convergent evolution of the viral envelope. *Microbiol. Mol. Biol. Rev.* **79**, 403–418 (2015).

[153] Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167 (2008).

[154] Vigouroux, A. *et al.* Class-a penicillin binding proteins do not contribute to cell shape but repair cell-wall defects. *Elife* **9**, e51998 (2020).

[155] Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Bio.* **2**, a000414 (2010).

[156] Schleifer, K. H. & Kandler, O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* **36**, 407–477 (1972).

[157] Heijenoort, J. v. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology* **11**, 25R–36R (2001).

[158] Alderwick, L. J., Harrison, J., Lloyd, G. S. & Birch, H. L. The mycobacterial cell wall—peptidoglycan and arabinogalactan. *Cold Spring Harb. Perspect. Med.* **5**, a021113 (2015).

[159] Shockman, G. D. & Barren, J. F. Structure, function, and assembly of cell walls of gram-positive bacteria. *Annu. Rev. Microbiol.* **37**, 501–527 (1983).

[160] Costerton, J., Ingram, J. & Cheng, K. Structure and function of the cell envelope of gram-negative bacteria. *Bacteriol. Rev.* **38**, 87–110 (1974).

[161] Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochim. et Biophys. Acta. (BBA)-Proteins Proteomics* **1794**, 808–816 (2009).

[162] Matsuhashi, M. Biosynthesis in the bacterial cell wall. *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme* **11**, 875–886 (1966).

[163] Lederer, E., Adam, A., Ciorbaru, R., Petit, J.-F. & Wietzerbin, J. Cell walls of mycobacteria and related organisms; chemistry and immunostimulant properties. *Mol. Cell. Biochem.* **7**, 87–104 (1975).

[164] Raymond, J. B., Mahapatra, S., Crick, D. C. & Pavelka, M. S. Identification of the namh gene, encoding the hydroxylase responsible for the n-glycolylation of the mycobacterial peptidoglycan. *J. Biol. Chem.* **280**, 326–333 (2005).

[165] Nikaido, H. & Brennan, P. The envelope of mycobacteria. *Annu. Rev. Biochem.* **64**, 29–63 (1995).

[166] Petit, J., Adam, A., Wietzerbin-Falszpan, J., Lederer, E. & Ghuysen, J.-M. Chemical structure of the cell wall of *Mycobacterium smegmatis*. I—Isolation and partial characterization of the peptidoglycan. *Biochem. biophysical Res. Commun.* **35**, 478–485 (1969).

[167] Aliashkevich, A. & Cava, F. L,D-transpeptidases: the great unknown among the peptidoglycan cross-linkers. *The FEBS J.* (2021).

[168] Wietzerbin-Falszpan, J. *et al.* Isolation and mass spectrometric identification of the peptide subunits of mycobacterial cell walls. *Biochem. Biophys. Res. Commun.* **40**, 57–63 (1970).

[169] Tipper, D. J. & Strominger, J. L. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-d-alanyl-d-alanine. *Proc. Natl. Acad. Sci. United States Am.* **54**, 1133 (1965).

[170] Lavollay, M. *et al.* The peptidoglycan of stationary-phase *Mycobacterium tuberculosis* predominantly contains cross-links generated by L, D-transpeptidation. *J. Bacteriol.* **190**, 4360–4366 (2008).

[171] Ngadjeua, F. *et al.* Critical impact of peptidoglycan precursor amidation on the activity of l, d-transpeptidases from *enterococcus faecium* and *Mycobacterium tuberculosis*. *Chem. Eur. J.* **24**, 5743–5747 (2018).

[172] Gupta, R. *et al.* The mycobacterium tuberculosis protein ldt mt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. *Nat. Med.* **16**, 466–469 (2010).

[173] Böth, D. *et al.* Structure of LdtMt2, an L, D-transpeptidase from *Mycobacterium tuberculosis*. *Acta. Crystallogr. Sect. D: Biol. Crystallogr.* **69**, 432–441 (2013).

[174] Cordillot, M. *et al.* In vitro cross-linking of *Mycobacterium tuberculosis* peptidoglycan by L, D-transpeptidases and inactivation of these enzymes by carbapenems. *Antimicrob. Agents. Chemother.* **57**, 5940–5945 (2013).

[175] Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. & Duncan, K. Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* **43**, 717–731 (2002).

[176] Li, W.-J. *et al.* Crystal structure of L, D-transpeptidase LdtMt2 in complex with meropenem reveals the mechanism of carbapenem against *Mycobacterium tuberculosis*. *Cell Res.* **23**, 728–731 (2013).

[177] de Munnik, M. *et al.* Targeting the *Mycobacterium tuberculosis* transpeptidase LdtMt2 with cysteine-reactive inhibitors including ebselen. *Chem. Commun.* **55**, 10214–10217 (2019).

[178] Steiner, E. M., Schneider, G. & Schnell, R. Binding and processing of β-lactam antibiotics by the transpeptidase LdtMt2 from *Mycobacterium tuberculosis*. *The FEBS J.* **284**, 725–741 (2017).

[179] Caveney, N. A. *et al.* Structural insight into YcbB-mediated beta-lactam resistance in *Escherichia coli*. *Nat. Commun.* **10**, 1–11 (2019).

[180] Gokulan, K., Khare, S., Cerniglia, C. E., Foley, S. L. & Varughese, K. I. Structure and inhibitor specificity of L, D-transpeptidase (LdtMt2) from *Mycobacterium tuberculosis* and antibiotic resistance: calcium binding promotes dimer formation. *The AAPS J.* **20**, 1–14 (2018).

[181] Libreros-Zuniga, G. A., dos Santos Silva, C., Salgado Ferreira, R. & Dias, M. V. B. Structural basis for the interaction and processing of β-lactam antibiotics by L, D-transpeptidase 3 (LdtMt3) from *Mycobacterium tuberculosis*. *ACS Infect. Dis.* **5**, 260–271 (2018).

[182] Basta, L. A. B. *et al.* Loss of a functionally and structurally distinct L, D-transpeptidase, LdtMt5, compromises cell wall integrity in Mycobacterium tuberculosis. *J. Biol. Chem.* **290**, 25670–25685 (2015).

[183] Correale, S., Ruggiero, A., Capparelli, R., Pedone, E. & Berisio, R. Structures of free and inhibited forms of the L, D-transpeptidase LdtMt1 from *Mycobacterium tuberculosis*. *Acta. Crystallogr. Sect. D: Biol. Crystallogr.* **69**, 1697–1706 (2013).

[184] Sanders, A. N., Wright, L. F. & Pavelka Jr, M. S. Genetic characterization of mycobacterial L,D-transpeptidases. *Microbiology* **160**, 1795 (2014).

[185] Zandi, T. A., Marshburn, R. L., Stateler, P. K. & Brammer Basta, L. A. Phylogenetic and biochemical analyses of mycobacterial l, d-transpeptidases reveal a distinct enzyme class that is preferentially acylated by meropenem. *ACS Infect. Dis.* **5**, 2047–2054 (2019).

[186] Magnet, S., Dubost, L., Marie, A., Arthur, M. & Gutmann, L. Identification of the L, D-transpeptidases for peptidoglycan cross-linking in *Escherichia coli*. *J. Bacteriol.* **190**, 4782–4785 (2008).

[187] Mainardi, J.-L. *et al.* A novel peptidoglycan cross-linking enzyme for a β-lactam-resistant transpeptidation pathway. *J. Biol. Chem.* **280**, 38146–38152 (2005).

[188] Patru, M.-M. & Pavelka Jr, M. S. A role for the class a penicillin-binding protein pona2 in the survival of *Mycobacterium smegmatis* under conditions of nonreplication. *J. Bacteriol.* **192**, 3043–3054 (2010).

[189] Goffin, C. & Ghuysen, J.-M. Biochemistry and comparative genomics of sxxk superfamily acyltransferases offer a clue to the mycobacterial paradox: presence of penicillin-susceptible target proteins versus lack of efficiency of penicillin as therapeutic agent. *Microbiol. Mol. Bio. Rev.* **66**, 702–738 (2002).

[190] Wolfe, S., Yang, K. & Khalil, M. Conformation–activity relationships and the mechanism of action of penicillin. *Can. J. Chem.* **66**, 2733–2750 (1988).

[191] Lobanovska, M. & Pilla, G. Focus: drug development: Penicillin's discovery and antibiotic resistance: lessons for the future? *The Yale J. Bio. Med.* **90**, 135 (2017).

[192] Zhang, X., Cui, Y., Liu, C., Zuo, K. & Tang, Y. Antibiotic sales in primary care in Hubei province, china: an analysis of 2012–2017 procurement records. *Int. J. Environ. Res. Public Health.* **16**, 3376 (2019).

[193] Munita, J. M. & Arias, C. A. Mechanisms of antibiotic resistance. *Microbiol. spectrum* **4**, 4–2 (2016).

[194] Alcock, B. P. *et al.* Card 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nu. Acid. Res.* **48**, D517–D525 (2020).

[195] Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems: past, present, and future. *Antimicrob. Agents. Chemother.* **55**, 4943–4960 (2011).

[196] Oliver, A., Levin, B. R., Juan, C., Baquero, F. & Blázquez, J. Hypermutation and the preexistence of antibiotic-resistant pseudomonas aeruginosa mutants: implications for susceptibility testing and treatment of chronic infections. *Antimicrob. Agents. Chemother.* **48**, 4226–4233 (2004).

[197] Meletis, G. Carbapenem resistance: overview of the problem and future perspectives. *Ther. advances Infect. Dis.* **3**, 15–21 (2016).

[198] Kurz, S. G. & Bonomo, R. A. Reappraising the use of β-lactams to treat tuberculosis. *Expert. Rev. Anti. Ther.* **10**, 999–1006 (2012).

[199] Dartois, V. A. & Rubin, E. J. Anti-tuberculosis treatment strategies and drug development: challenges and priorities. *Nat. Rev. Microbiol.* 1–17 (2022).

[200] Hackbarth, C. J., Unsal, I. & Chambers, H. F. Cloning and sequence analysis of a class A beta-lactamase from *Mycobacterium tuberculosis* H37Ra. *Antimicrob. Agents. Chemother.* **41**, 1182–1185 (1997).

[201] Voladri, R. K. R. *et al.* Recombinant expression and characterization of the major β-lactamase of Mycobacterium tuberculosis. *Antimicrob. Agents. Chemother.* **42**, 1375–1381 (1998).

[202] Saxena, A. K. & Singh, A. Mycobacterial tuberculosis enzyme targets and their inhibitors. *Curr. Top. Med. Chem.* **19**, 337–355 (2019).

[203] England, K. *et al.* Meropenem-clavulanic acid shows activity against *Mycobacterium tuberculosis in vivo*. *Antimicrob. Agents. Chemother.* **56**, 3384–3387 (2012).

[204] Wang, F., Cassidy, C. & Sacchettini, J. C. Crystal structure and activity studies of the *Mycobacterium tuberculosis* β-lactamase reveal its critical role in resistance to β-lactam antibiotics. *Antimicrob. Agents. Chemother.* **50**, 2762–2771 (2006).

[205] Tremblay, L. W., Fan, F. & Blanchard, J. S. Biochemical and structural characterization of *Mycobacterium tuberculosis* β-lactamase with the carbapenems ertapenem and doripenem. *Biochemistry* **49**, 3766–3773 (2010).

[206] Kumar, G., Galanis, C., Batchelder, H. R., Townsend, C. A. & Lamichhane, G. Penicillin binding proteins and β-lactamases of *Mycobacterium tuberculosis*: Reexamination of the historical paradigm. *Msphere* **7**, e00039–22 (2022).

[207] Reading, C. & Farmer, T. The inhibition of periplasmic β-lactamase in *Escherichia coli* by clavulanic acid and other β-lactamase inhibitors. *Postgrad. Med.* 163–8 (1984).

[208] Tally, F. P., Jacobus, N. V. & Gorbach, S. L. In vitro activity of thienamycin. *Antimicrob. Agents. Chemother.* **14**, 436–438 (1978).

[209] Weaver, S. S., Bodey, G. P. & LeBlanc, B. M. Thienamycin: new beta-lactam antibiotic with potent broad-spectrum activity. *Antimicrob. Agents. Chemother.* **15**, 518–521 (1979).

[210] Hashizume, T., Ishino, F., Nakagawa, J.-I., Tamaki, S. & Matsuhashi, M. Studies on the mechanism of action of imipenem (*N*-formimidoylthienamycin) in vitro: binding to the penicillin-binding proteins (PBPs) in *Escherichia coli* and *Pseudomonas aeruginosa*, and inhibition of enzyme activities due to the pbps in *e. coli*. *The J. Antibiot.* **37**, 394–400 (1984).

[211] Basker, M. J., Boon, R. & Hunter, P. Comparative antibacterial properties in vitro of seven olivanic acid derivatives: Mm 4550, mm 13902, mm 17880, mm 22380, mm 22381, mm 22382 and mm 22383. *The J. Antibiot.* **33**, 878–884 (1980).

[212] Fukasawa, M. *et al.* Stability of meropenem and effect of 1 beta-methyl substitution on its stability in the presence of renal dehydropeptidase i. *Antimicrob. Agents. Chemother.* **36**, 1577–1579 (1992).

[213] Norrby, S. *et al.* Urinary recovery of n-formimidoyl thienamycin (mk0787) as affected by coadministration of n-formimidoyl thienamycin dehydropeptidase inhibitors. *Antimicrob. Agents. Chemother.* **23**, 300–307 (1983).

[214] Bassetti, M., Nicolini, L., Esposito, S., Righi, E. & Viscoli, C. Current status of newer carbapenems. *Curr. Med. Chem.* **16**, 564–575 (2009).

[215] Nix, D. E., Majumdar, A. K. & DiNubile, M. J. Pharmacokinetics and pharmacodynamics of ertapenem: an overview for clinicians. *J. Antimicrob. Chemother.* **53**, ii23–ii28 (2004).

[216] Kazmierczak, K. M., Karlowsky, J. A., de Jonge, B. L., Stone, G. G. & Sahm, D. F. Epidemiology of carbapenem resistance determinants identified in meropenem-nonsusceptible enterobacterales collected as part of a global surveillance program, 2012 to 2017. *Antimicrob. Agents. Chemother.* **65**, e02000–20 (2021).

[217] Babiker, A. *et al.* Changing epidemiology and decreased mortality associated with carbapenem-resistant gram-negative bacteria, 2000–2017. *Clin. Infect. Dis.* **73**, e4521–e4530 (2021).

[218] Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile genetic elements associated with antimicrobial resistance. *Clin. Microbiol. Rev.* **31**, e00088–17 (2018).

[219] Kumar, P. *et al.* Meropenem inhibits d, d-carboxypeptidase activity in m ycobacterium tuberculosis. *Mol. Microbiol.* **86**, 367–381 (2012).

[220] Dauby, N., Muylle, I., Mouchet, F., Sergysels, R. & Payen, M.-C. Meropenem/clavulanate and linezolid treatment for extensively drug-resistant tuberculosis. *The Pediatr. Infect. Dis. J.* **30**, 812–813 (2011).

[221] De Lorenzo, S. *et al.* Efficacy and safety of meropenem–clavulanate added to linezolid-containing regimens in the treatment of mdr-/xdr-tb. *Eur. Respir. J.* **41**, 1386–1392 (2013).

[222] Dhar, N. *et al.* Rapid cytolysis of *Mycobacterium tuberculosis* by faropenem, an orally bioavailable β-lactam antibiotic. *Antimicrob. Agents. Chemother.* **59**, 1308–1319 (2015).

[223] Kumar, P., Singh, U. B., Lamichhane, G. & Story-Roller, E. Repurposing of carbapenems for the treatment of drug-resistant tuberculosis. In *Mycobacterium tuberculosis: Molecular Infection Biology, Pathogenesis, Diagnostics and New Interventions*, 497–514 (Springer, 2019).

[224] Kim, H. S. *et al.* Structural basis for the inhibition of *Mycobacterium tuberculosis* L, D-transpeptidase by meropenem, a drug effective against extensively drug-resistant strains. *Acta. Crystallogr. Sect. D: Biol. Crystallogr.* **69**, 420–431 (2013).

[225] Bianchet, M. A. *et al.* Structural insight into the inactivation of *Mycobacterium tuberculosis* non-classical transpeptidase LdtMt2 by biapenem and tebipenem. *BMC Biochem.* **18**, 1–14 (2017).

[226] de Munnik, M. *et al.* A fluorescence-based assay for screening β-lactams targeting the *Mycobacterium tuberculosis* transpeptidase LdtMt2. *ChemBioChem* **21**, 368–372 (2020).

[227] Moon, T. M. *et al.* The structures of penicillin-binding protein 4 (pbp4) and pbp5 from enterococci provide structural insights into β-lactam resistance. *J. Biol. Chem.* **293**, 18574–18584 (2018).

[228] Lohans, C. T. *et al.* Non-hydrolytic β-lactam antibiotic fragmentation by l, d-transpeptidases and serine β-lactamase cysteine variants. *Angew. Chemie Int. Ed.* **58**, 1990–1994 (2019).

[229] Triboulet, S. *et al.* Kinetic features of l, d-transpeptidase inactivation critical for β-lactam antibacterial activity. *PloS one* **8**, e67831 (2013).

[230] Dubée, V. *et al.* Inactivation of *Mycobacterium tuberculosis* L, D-transpeptidase LdtMt1 by carbapenems and cephalosporins. *Antimicrob. Agents. Chemother.* **56**, 4189–4195 (2012).

[231] Lopez Quezada, L. *et al.* Activity-based protein profiling reveals that cephalosporins selectively active on non-replicating *Mycobacterium tuberculosis* bind multiple protein families and spare peptidoglycan transpeptidases. *Front. Microbiol.* 1248 (2020).

[232] Sütterlin, L., Edoo, Z., Hugonnet, J.-E., Mainardi, J.-L. & Arthur, M. Peptidoglycan cross-linking activity of l, d-transpeptidases from clostridium difficile and inactivation of these enzymes by β-lactams. *Antimicrob. Agents. Chemother.* **62**, e01607–17 (2017).

[233] Hugonnet, J.-E., Tremblay, L. W., Boshoff, H. I., Barry 3rd, C. E. & Blanchard, J. S. Meropenem-clavulanate is effective against extensively drug-resistant *Mycobacterium tuberculosis*. *Science* **323**, 1215–1218 (2009).

[234] Kaushik, A. *et al.* Carbapenems and rifampin exhibit synergy against *Mycobacterium tuberculosis* and *Mycobacterium abscessus*. *Antimicrob. Agents. Chemother.* **59**, 6561–6567 (2015).

[235] Kumar, P. *et al.* Non-classical transpeptidases yield insight into new antibacterials. *Nat. Chem. Bio.* **13**, 54–61 (2017).

[236] Zhao, F., Hou, Y.-J., Zhang, Y., Wang, D.-C. & Li, D.-F. The 1-β-methyl group confers a lower affinity of l, d-transpeptidase ldtmt2 for ertapenem than for imipenem. *Biochem. biophysical Res. Commun.* **510**, 254–260 (2019).

[237] Silva, J. R. A. *et al.* Targeting the cell wall of *Mycobacterium tuberculosis*: a molecular modeling investigation of the interaction of imipenem and meropenem with L, D-transpeptidase 2. *J. Biomol. Struct. Dyn.* **34**, 304–317 (2016).

[238] Yan, G. *et al.* Development of a simple and miniaturized sandwich-like fluorescence polarization assay for rapid screening of SARS-CoV-2 main protease inhibitors. *Cell & Biosci.* **11**, 1–14 (2021).

[239] Graziano, V., McGrath, W. J., DeGruccio, A. M., Dunn, J. J. & Mangel, W. F. Enzymatic activity of the SARS coronavirus main proteinase dimer. *FEBS letters* **580**, 2577–2583 (2006).

[240] Xue, X. *et al.* Production of authentic SARS-CoV Mpro with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *J. Mol. Bio.* **366**, 965–975 (2007).

[241] Blanchard, J. E. *et al.* High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem. & Bio.* **11**, 1445–1453 (2004).

[242] Grum-Tokars, V., Ratia, K., Begaye, A., Baker, S. C. & Mesecar, A. D. Evaluating the 3c-like protease activity of sars-coronavirus: recommendations for standardized assays for drug discovery. *Virus Res.* **133**, 63–73 (2008).

[243] Sittampalam, G. S. *et al.* Assay guidance manual [internet] (2004).

[244] Kao, R. Y. *et al.* Characterization of SARS-CoV main protease and identification of biologically active small molecule inhibitors using a continuous fluorescence-based assay. *Febs Lett.* **576**, 325–330 (2004).

[245] Bacha, U., Barrila, J., Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. *Biochemistry* **43**, 4906–4912 (2004).

[246] Dos Remedios, C. G. & Moens, P. D. Fluorescence resonance energy transfer spectroscopy is a reliable" ruler" for measuring structural changes in proteins: dispelling the problem of the unknown orientation factor. *J. structural Bio.* **115**, 175–185 (1995).

[247] Kantsadi, A. L., Cattermole, E., Matsoukas, M.-T., Spyroulias, G. A. & Vakonakis, I. A covid moonshot: assessment of ligand binding to the SARS-CoV-2 main protease by saturation transfer difference nmr spectroscopy. *J. Biomol. NMR* **75**, 167–178 (2021).

[248] Resnick, E. *et al.* Rapid covalent-probe discovery by electrophile-fragment screening. *J. Am. Chem. Soc.* **141**, 8951–8968 (2019).

[249] Yang, H. *et al.* The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci.* **100**, 13190–13195 (2003).

[250] Tokmina-Roszyk, M., Tokmina-Roszyk, D. & Fields, G. B. The synthesis and application of fmoc-lys (5-fam) building blocks. *Pept. Sci.* **100**, 347–355 (2013).

[251] Shitrit, A. *et al.* Conserved interactions required for inhibition of the main protease of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Sci. Rep.* **10**, 1–11 (2020).

[252] Jannetto, P. J. & Langman, L. J. High-throughput online solid-phase extraction tandem mass spectrometry: Is it right for your clinical laboratory? *Clin. Biochem.* **49**, 1032–1034 (2016).

[253] Malla, T. R. *et al.* Mass spectrometry reveals potential of β-lactams as SARS-CoV-2 Mpro inhibitors. *Chem. Commun.* **57**, 1430–1433 (2021).

[254] Product description fda approved compound library. https://cdn.caymanchem.com/cdn/insert/23538.pdf. Accessed: 2022-03-02.

[255] Product description sigma lopac^{1280} Library of Pharmacologically Active Compounds. https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/109/211/lo1280dat.pdf. Accessed: 2022-06-02.

[256] Zhang, J.-H., Chung, T. D. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**, 67–73 (1999).

[257] Schoellmann, G. & Shaw, E. Direct evidence for the presence of histidine in the active center of chymotrypsin. *Biochemistry* **2**, 252–255 (1963).

[258] Ha, K.-H. *et al.* *N*-tosyl-L-phenylalanine chloromethyl ketone inhibits NF-κB activation by blocking specific cysteine residues of IκB kinase β and p65/RelA. *Biochemistry* **48**, 7271–7278 (2009).

[259] Stöppler, H., Stöppler, M. C., Adduci, A., Koval, D. & Schlegel, R. The serine protease inhibitors TLCK and TPCK react with the rb-binding core of HPV-18 E7 protein and abolish its rb-binding capability. *Virology* **217**, 542–553 (1996).

[260] Krishnan, N., Bencze, G., Cohen, P. & Tonks, N. K. The anti-inflammatory compound bay-11-7082 is a potent inhibitor of protein tyrosine phosphatases. *The FEBS J.* **280**, 2830–2841 (2013).

[261] Atkins, S. L. *et al.* Small molecule screening identifies inhibitors of the Epstein-Barr virus deubiquitinating enzyme, BPLF1. *Antivir. Res.* **173**, 104649 (2020).

[262] Canal, B. *et al.* Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of nsp15 endoribonuclease. *Biochem. J.* **478**, 2465–2479 (2021).

[263] Pohjala, L. & Tammela, P. Aggregating behavior of phenolic compounds—a source of false bioassay results? *Molecules* **17**, 10774–10790 (2012).

[264] He, Z. *et al.* Molecules inhibit the enzyme activity of 3-chymotrypsin-like cysteine protease of SARS-CoV-2 virus: the experimental and theory studies. *bioRxiv* (2020).

[265] Baker, J. D., Uhrich, R. L., Kraemer, G. C., Love, J. E. & Kraemer, B. C. A drug repurposing screen identifies hepatitis C antivirals as inhibitors of the SARS-CoV-2 main protease. *PloS one* **16**, e0245962 (2021).

[266] Zhu, W. *et al.* Identification of SARS-CoV-2 3CL protease inhibitors by a quantitative high-throughput screening. *ACS Pharmacol. & Transl. Sci.* **3**, 1008–1016 (2020).

[267] El-Baba, T. J. *et al.* Allosteric inhibition of the SARS-CoV-2 main protease: Insights from mass spectrometry based assays. *Angew. Chemie Int. Ed.* **59**, 23544–23548 (2020).

[268] Waløen, K. *et al.* Cysteine modification by ebselen reduces the stability and cellular levels of 14-3-3 proteins. *Mol. Pharmacol.* **100**, 155–169 (2021).

[269] Capper, M. J. *et al.* The cysteine-reactive small molecule ebselen facilitates effective sod1 maturation. *Nat. Commun.* **9**, 1–9 (2018).

[270] Marthandan, S. Effects of ebselen and n-acetyl cysteine on replicative aging of primary human fibroblast strains. *Immun. Ageing* **12**, 1–3 (2015).

[271] Suda, H., Aoyagi, T., Hamada, M., Takeuchi, T. & Umezawa, H. Antipain, a new protease inhibitor isolated from agtinomycetes. *The J. Antibiot.* **25**, 263–265 (1972).

[272] Kolli, M., Özen, A., Kurt-Yilmaz, N. & Schiffer, C. A. HIV-1 protease-substrate coevolution in nelfinavir resistance. *J. Virol.* **88**, 7145–7154 (2014).

[273] Zhang, R. *et al.* Probing the substrate specificity of hepatitis c virus ns3 serine protease by using synthetic peptides. *J. Virol.* **71**, 6208–6213 (1997).

[274] Ma, W., Tang, C. & Lai, L. Specificity of trypsin and chymotrypsin: loop-motion-controlled dynamic correlation as a determinant. *Biophys. J.* **89**, 1183–1193 (2005).

[275] Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Bio.* **267**, 727–748 (1997).

[276] Eid, R. A. *et al.* Efficacy of Ceftazidime and Cefepime in the Management of COVID-19 Patients: Single Center Report from Egypt. *Antibiotics* **10**, 1278 (2021).

[277] De Meester, F. *et al.* Automated analysis of enzyme inactivation phenomena: application to β-lactamases and dd-peptidases. *Biochem. Pharmacol.* **36**, 2393–2403 (1987).

[278] Michaux, C., Charlier, P., Frère, J.-M. & Wouters, J. Crystal structure of brl 42715, c6-(n 1-methyl-1, 2, 3-triazolylmethylene) penem, in complex with enterobacter c loacae 908r β-lactamase: Evidence for a stereoselective mechanism from docking studies. *J. Am. Chem. Soc.* **127**, 3262–3263 (2005).

[279] Muramatsu, T. *et al.* Sars-cov 3cl protease cleaves its c-terminal autoprocessing site by novel subsite cooperativity. *Proc. Natl. Acad. Sci.* **113**, 12997–13002 (2016).

[280] Shaqra, A. M. *et al.* Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nat. Commun.* **13**, 1–11 (2022).

[281] Gurard-Levin, Z. A. *et al.* Evaluation of SARS-CoV-2 3C-like protease inhibitors using self-assembled monolayer desorption ionization mass spectrometry. *Antivir. Res.* **182**, 104924 (2020).

[282] Lakowicz, J. R. *Principles of fluorescence spectroscopy* (Springer, 2006).

[283] Liu, Y. *et al.* Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* **267**, 331–335 (1999).

[284] Holskin, B. *et al.* A continuous fluorescence-based assay of human cytomegalovirus protease using a peptide substrate. *Anal. Biochem.* **227**, 148–155 (1995).

[285] Palmier, M. O. & Van Doren, S. R. Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal. Biochem.* **371**, 43–51 (2007).

[286] Blanchard, J. E. *et al.* High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem. & Bio.* **11**, 1445–1453 (2004).

[287] Gurard-Levin, Z. A., Scholle, M. D., Eisenberg, A. H. & Mrksich, M. High-throughput screening of small molecule libraries using samdi mass spectrometry. *ACS Comb. Sci.* **13**, 347–350 (2011).

[288] Feng, B. Y. & Shoichet, B. K. A detergent-based assay for the detection of promiscuous inhibitors. *Nat. Protoc.* **1**, 550–553 (2006).

[289] Volmer, D. & Jessome, L. L. Ion suppression: a major concern in mass spectrometry (2006).

[290] Viaud, J. & Peterson, J. R. An allosteric kinase inhibitor binds the p21-activated kinase autoregulatory domain covalentlyallosteric kinase inhibition. *Mol. cancer Ther.* **8**, 2559–2565 (2009).

[291] Wilcox, P. E., Segal, D. M., Powers, J. C., Cohen, G. H. & Davies, D. R. Substrate binding site in bovine chymotrypsin aγ. crystallographic study using peptide chloromethyl ketones as site-specific inhibitors. *Biochemistry* **10**, 3728–3738 (1971).

[292] Shaw, E., Angliker, H., Rauber, P., Walker, B. & Wikstrom, P. Peptidyl fluoromethyl ketones as thiol protease inhibitors. *Biomed. Biochim. Acta.* **45**, 1397–1403 (1986).

[293] Ghahremanpour, M. M. *et al.* Identification of 14 known drugs as inhibitors of the main protease of SARS-CoV-2. *ACS Med. Chem. letters* **11**, 2526–2533 (2020).

[294] Temesi, D., Law, B. & Howe, N. Synthesis and evaluation of peg414, a novel formulating agent that avoids analytical problems associated with polydisperse vehicles such as peg400. *J. pharmaceutical Sci.* **92**, 2512–2518 (2003).

[295] Miller, V. A., Baghalabadi, V., Nickerson, J. & Doucette, A. A. Removal of PEG Contaminants From Peptides Through An Off-Line Ion-Exchange Spin Column.

[296] Gil-Moles, M. *et al.* Metallodrug profiling against SARS-CoV-2 target proteins identifies highly potent inhibitors of the S/ACE2 interaction and the Papain-like Protease PLpro. *Chem. Eur. J.* (2021).

[297] Pace, N. J. & Weerapana, E. Zinc-binding cysteines: diverse functions and structural motifs. *Biomolecules* **4**, 419–434 (2014).

[298] Seamon, K. J. & Stivers, J. T. A high-throughput enzyme-coupled assay for SAMHD1 dNTPase. *J. Biomol. Screen.* **20**, 801–809 (2015).

[299] Brasnett, A. *et al.* Human oxygenase variants employing a single protein feii ligand are catalytically active. *Angew. Chemie* **133**, 14778–14784 (2021).

[300] Grifagni, D. *et al.* SARS-CoV-2 Mpro inhibition by a zinc ion: structural features and hints for drug design. *Chem. Commun.* **57**, 7910–7913 (2021).

[301] Kozak, J. J., Gray, H. B. & Garza-López, R. A. Structural stability of the SARS-CoV-2 main protease: Can metal ions affect function? *J. Inorg. Biochem.* **211**, 111179 (2020).

[302] Roberto, A., Garza-Lopez, J. J., Kozak, H. B. *et al.* Copper (II) inhibition of the SARS-CoV-2 main protease. *chemRxiv* (2020).

[303] Malla, T. R. *et al.* Penicillin derivatives inhibit the SARS-CoV-2 main protease by reaction with its nucleophilic cysteine. *J. Med. Chem.* (2022).

[304] Gasteiger, E. *et al.* Protein identification and analysis tools on the expasy server. *The proteomics Protoc. handbook* 571–607 (2005).

[305] Akova, M. Sulbactam-containing β-lactamase inhibitor combinations. *Clin. Microbiol. Infect.* **14**, 185–188 (2008).

[306] Drawz, S. M. *et al.* Penicillin sulfone inhibitors of class D β-lactamases. *Antimicrob. Agents. Chemother.* **54**, 1414–1424 (2010).

[307] Buynak, J. D. The discovery and development of modified penicillin-and cephalosporin-derived β-lactamase inhibitors. *Curr. Med. Chem.* **11**, 1951–1964 (2004).

[308] Lecaille, F., Kaleta, J. & Brömme, D. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. *Chem. Rev.* **102**, 4459–4488 (2002).

[309] Wilmouth, R. C. *et al.* Mechanistic insights into the inhibition of serine proteases by monocyclic lactams. *Biochemistry* **38**, 7989–7998 (1999).

[310] Scheidt, K. A. *et al.* Structure-based design, synthesis and evaluation of conformationally constrained cysteine protease inhibitors. *Bioorganic Med. Chem.* **6**, 2477–2494 (1998).

[311] Andrews, D. M. *et al.* Pyrrolidine-5, 5-trans-lactams. 5. pharmacokinetic optimization of inhibitors of hepatitis C virus NS3/4A protease. *Org. Lett.* **5**, 4631–4634 (2003).

[312] Mäder, P. & Kattner, L. Sulfoximines as rising stars in modern drug discovery? current status and perspective on an emerging functional group in medicinal chemistry. *J. Med. Chem.* **63**, 14243–14275 (2020).

[313] Lücking, U. Sulfoximines: a neglected opportunity in medicinal chemistry. *Angew. Chemie Int. Ed.* **52**, 9399–9408 (2013).

[314] Chan, H. H. *et al.* Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding. *Chem. Sci.* **12**, 13686–13703 (2021).

[315] Schnier, P. D., Gross, D. S. & Williams, E. R. On the maximum charge state and proton transfer reactivity of peptide and protein ions formed by electrospray ionization. *J. Am. Soc. Mass Spectrom.* **6**, 1086–1097 (1995).

[316] Smith, R. D., Loo, J. A., Loo, R. R. O., Busman, M. & Udseth, H. R. Principles and practice of electrospray ionization—mass spectrometry for large polypeptides and proteins. *Mass Spectrom. Rev.* **10**, 359–452 (1991).

[317] Felitsyn, N., Peschke, M. & Kebarle, P. Origin and number of charges observed on multiply-protonated native proteins produced by esi. *Int. J. Mass Spectrom.* **219**, 39–62 (2002).

[318] Grandori, R. Origin of the conformation dependence of protein charge-state distributions in electrospray ionization mass spectrometry. *J. mass spectrometry* **38**, 11–15 (2003).

[319] Swift, M. L. Graphpad prism, data analysis, and scientific graphing. *J. Chem. in. computer Sci.* **37**, 411–412 (1997).

[320] Cohen, S. & Pratt, R. Inactivation of bacillus cereus. beta.-lactamase i by 6. beta.-bromopenicillanic acid: mechanism. *Biochemistry* **19**, 3996–4003 (1980).

[321] Orlek, B. S., Sammes, P. G., Knott-Hunziker, V. & Waley, S. G. On the chemical binding of 6β-bromopenicillanic acid to β-lactamase i. *J. Chem. Soc. Chem. Commun.* 962–963 (1979).

[322] Kemp, J. E., Closier, M. D., Narayanaswami, S. & Stefaniak, M. H. Nucleophilic sn2 displacements on penicillin-6-and cephalosporin-7-triflates; 6β-iodopenicillanic acid, a new β-lactamase inhibitor. *Tetrahedron Lett.* **21**, 2991–2994 (1980).

[323] Sauvage, E. *et al.* Structural basis of the inhibition of class a β-lactamases and penicillin-binding proteins by 6-β-iodopenicillanate. *J. Am. Chem. Soc.* **131**, 15262–15269 (2009).

[324] McKinnon, P. S., Paladino, J. A. & Schentag, J. J. Evaluation of area under the inhibitory curve (auic) and time above the minimum inhibitory concentration (t> mic) as predictors of outcome for cefepime and ceftazidime in serious bacterial infections. *Int. J. Antimicrob. Agents.* **31**, 345–351 (2008).

[325] Veiga, R. P. & Paiva, J.-A. Pharmacokinetics–pharmacodynamics issues relevant for the clinical use of beta-lactam antibiotics in critically ill patients. *Critical Care* **22**, 1–34 (2018).

[326] Henninot, A., Collins, J. C. & Nuss, J. M. The current state of peptide drug discovery: back to the future? *J. Med. Chem.* **61**, 1382–1414 (2018).

[327] Wang, L. *et al.* Therapeutic peptides: current applications and future directions. *Signal Transduct. Target. Ther.* **7**, 1–27 (2022).

[328] Birk, A. V. *et al.* The mitochondrial-targeted compound ss-31 re-energizes ischemic mitochondria by interacting with cardiolipin. *J. Am. Soc. Nephrol.* **24**, 1250–1261 (2013).

[329] Carvajal, L. A. *et al.* Dual inhibition of mdmx and mdm2 as a therapeutic strategy in leukemia. *Sci. Transl. Med.* **10**, eaao3003 (2018).

[330] Zhao, Y. *et al.* Structural basis for replicase polyprotein cleavage and substrate specificity of main protease from SARS-CoV-2. *Proc. Natl. Acad. Sci.* **119**, e2117142119 (2022).

[331] MacDonald, E. A. *et al.* Recognition of divergent viral substrates by the SARS-CoV-2 main protease. *ACS Infect. Dis.* **7**, 2591–2595 (2021).

[332] Chuck, C.-P. *et al.* Profiling of substrate specificity of SARS-CoV 3CLpro. *PloS one* **5**, e13197 (2010).

[333] Wei, M. *et al.* High-throughput determination of mode of inhibition in lead identification and optimization. *J. Biomol. Screen.* **12**, 220–228 (2007).

[334] Buckton, L. K. & McAlpine, S. R. Improving the cell permeability of polar cyclic peptides by replacing residues with alkylated amino acids, asparagines, and d-amino acids. *Org. letters* **20**, 506–509 (2018).

[335] Altman, D. G. & Bland, J. M. Measurement in medicine: the analysis of method comparison studies. *J. R. Stat. Soc. Ser. D (The Stat.* **32**, 307–317 (1983).

[336] Craik, D. J., Fairlie, D. P., Liras, S. & Price, D. The future of peptide-based drugs. *Chem. Bio. Drug. Des.* **81**, 136–147 (2013).

[337] Yang, H. & Yang, J. A review of the latest research on Mpro targeting SARS-COV inhibitors. *RSC Med. Chem.* **12**, 1026 (2021).

[338] O'Brien, A. *et al.* Detecting SARS-CoV-2 3clpro expression and activity using a polyclonal antiserum and a luciferase-based biosensor. *Virology* **556**, 73–78 (2021).

[339] Orville, A. M. Recent results in time resolved serial femtosecond crystallography at xfels. *Curr. Opin. Struct. Biol.* **65**, 193–208 (2020).

[340] Rut, W. *et al.* SARS-CoV-2 Mpro inhibitors and activity-based probes for patient-sample imaging. *Nat. Chem. Bio.* **17**, 222–228 (2021).

[341] Wang, J. *et al.* Mechanisms of ebselen as a therapeutic and its pharmacology applications. *Futur. Med. Chem.* **12**, 2141–2160 (2020).

[342] Hall, B. G. & Barlow, M. Evolution of the serine β-lactamases: past, present and future. *Drug Resist. Updat.* **7**, 111–123 (2004).

[343] Kelly, J. A. *et al.* On the origin of bacterial resistance to penicillin: comparison of a β-lactamase and a penicillin target. *Science* **231**, 1429–1431 (1986).

[344] Mazzella, L. J. & Pratt, R. Effect of the 3-leaving group on turnover of cephem antibiotics by a class C beta-lactamase. *Biochem. J.* **259**, 255–260 (1989).

[345] Frère, J., Ghuysen, J., Degelaen, J., Loffet, A. & Perkins, H. R. Fragmentation of benzylpenicillin after interaction with the exocellular DD-carboxypeptidase-transpeptidases of Streptomyces R61 and R39. *Nature* **258**, 168–170 (1975).

[346] van den Akker, F. & Bonomo, R. A. Exploring additional dimensions of complexity in inhibitor design for serine β-lactamases: Mechanistic and intra- and inter-molecular chemistry approaches. *Front. Microbiol.* **9**, 622 (2018).

[347] Kalp, M., Totir, M. A., Buynak, J. D. & Carey, P. R. Different intermediate populations formed by tazobactam, sulbactam, and clavulanate reacting with SHV-1 β-lactamases: Raman crystallographic evidence. *J. Am. Chem. Soc.* **131**, 2338–2347 (2009).

[348] Lucic, A. *et al.* Faropenem reacts with serine and metallo-β-lactamases to give multiple products. *Eur. J. Med. Chem.* **215**, 113257 (2021).

[349] Magnet, S. *et al.* Identification of the L, D-transpeptidases responsible for attachment of the braun lipoprotein to *Escherichia coli* peptidoglycan. *J. Bacteriol.* **189**, 3927–3931 (2007).

[350] Winkle, M. *et al.* Dpaa detaches Braun's lipoprotein from peptidoglycan. *Mbio* **12**, e00836–21 (2021).

[351] Cava, F., De Pedro, M. A., Lam, H., Davis, B. M. & Waldor, M. K. Distinct pathways for modification of the bacterial cell wall by non-canonical d-amino acids. *The EMBO J.* **30**, 3442–3453 (2011).

[352] Ferguson, G. P., Tötemeyer, S., MacLean, M. & Booth, I. R. Methylglyoxal production in bacteria: suicide or survival? *Arch. Microbiol.* **170**, 209–218 (1998).

[353] Booth, I. R. *et al.* Bacterial production of methylglyoxal: a survival strategy or death by misadventure? *Biochem. Soc. Trans.* **31**, 1406–1408 (2003).

[354] Geoghegan, K. F. *et al.* Spontaneous α-n-6-phosphogluconoylation of a "his tag" in *Escherichia coli*: The cause of extra mass of 258 or 178 da in fusion proteins. *Anal. Biochem.* **267**, 169–184 (1999).

[355] Baez, N. O. D., Reisz, J. A. & Furdui, C. M. Mass spectrometry in studies of protein thiol chemistry and signaling: opportunities and caveats. *Free. Radic. Biol. Med.* **80**, 191–211 (2015).

[356] Edoo, Z., Arthur, M. & Hugonnet, J.-E. Reversible inactivation of a peptidoglycan transpeptidase by a β-lactam antibiotic mediated by β-lactam-ring recyclization in the enzyme active site. *Sci. Rep.* **7**, 1–8 (2017).

[357] Levine, S. R. & Beatty, K. E. Investigating β-lactam drug targets in *Mycobacterium tuberculosis* using chemical probes. *ACS Infect. Dis.* **7**, 461–470 (2021).

[358] Ghuysen, J.-M., Frere, J.-M., Leyh-Bouille, M., Perkins, H. & Nieto, M. The active centres in penicillin-sensitive enzymes. *Philos. Trans. R. Soc. London. B, Biol. Sci.* **289**, 285–301 (1980).

[359] Marquet, A., Frère, J.-M., Ghuysen, J.-M. & Loffet, A. Effects of nucleophiles on the breakdown of the benzylpenicilloyl–enzyme complex ei formed between benzylpenicillin and the exocellular dd-carboxypeptidase–transpeptiase of streptomyces strain r61. *Biochem. J.* **177**, 909–916 (1979).

[360] Wu, Q., Zhang, X., Du, J. & Hu, C. Discussion on the dimerization reaction of penicillin antibiotics. *J. Pharm. Anal.* **12**, 481–488 (2022).

[361] Robinson-Fuentes, V., Jefferies, T. & Branch, S. Degradation pathways of ampicillin in alkaline solutions. *J. Pharm. Pharmacol.* **49**, 843–851 (1997).

[362] Paternotte, I. *et al.* Syntheses and hydrolysis of basic and dibasic ampicillin esters tailored for intracellular accumulation. *Bioorganic & Med. Chem.* **9**, 493–502 (2001).

[363] Mainardi, J.-L. *et al.* Unexpected inhibition of peptidoglycan LD-transpeptidase from enterococcus faecium by the β-lactam imipenem. *J. Biol. Chem.* **282**, 30414–30422 (2007).

[364] Lohans, C. T. *et al.* Mechanistic insights into β-lactamase-catalysed carbapenem degradation through product characterisation. *Sci. Rep.* **9**, 1–9 (2019).

[365] Kneller, D. W. *et al.* Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* **11**, 1–6 (2020).

[366] Goldberg, S. D., Iannuccilli, W., Nguyen, T., Ju, J. & Cornish, V. W. Identification of residues critical for catalysis in a class c β-lactamase by combinatorial scanning mutagenesis. *Protein Sci.* **12**, 1633–1645 (2003).

[367] Burke, H. M., McSweeney, L. & Scanlan, E. M. Exploring chemoselective s-to-n acyl transfer reactions in synthesis and chemical biology. *Nat. Commun.* **8**, 1–16 (2017).

[368] Johnson, E. C. & Kent, S. B. Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* **128**, 6640–6646 (2006).

[369] Bhattacharjee, N. *et al.* Negative impact of carbapenem methylation on the reactivity of β-lactams for cysteine acylation as revealed by quantum calculations and kinetic analyses. *Antimicrob. Agents. Chemother.* **63**, e02039–18 (2019).

[370] Cava, F., Lam, H., De Pedro, M. A. & Waldor, M. K. Emerging knowledge of regulatory roles of d-amino acids in bacteria. *Cell. Mol. Life Sci.* **68**, 817–831 (2011).

[371] Sütterlin, L., Edoo, Z., Hugonnet, J.-E., Mainardi, J.-L. & Arthur, M. Peptidoglycan cross-linking activity of l, d-transpeptidases from clostridium difficile and inactivation of these enzymes by β-lactams. *Antimicrob. Agents. Chemother.* **62**, e01607–17 (2018).

[372] Fura, J. M., Kearns, D. & Pires, M. M. D-amino acid probes for penicillin binding protein-based bacterial surface labeling. *J. Biol. Chem.* **290**, 30540–30550 (2015).

[373] Lupoli, T. J. *et al.* Transpeptidase-mediated incorporation of d-amino acids into bacterial peptidoglycan. *J. Am. Chem. Soc.* **133**, 10748–10751 (2011).

[374] Kim, S. J., Matsuoka, S., Patti, G. J. & Schaefer, J. Vancomycin derivative with damaged d-ala-d-ala binding cleft binds to cross-linked peptidoglycan in the cell wall of staphylococcus aureus. *Biochemistry* **47**, 3822–3831 (2008).

[375] Cetinkaya, Y., Falk, P. & Mayhall, C. G. Vancomycin-resistant enterococci. *Clin. Microbiol. Rev.* **13**, 686–707 (2000).

[376] Gupta, R. *et al.* Atypically modified carbapenem antibiotics display improved antimycobacterial activity in the absence of β-lactamase inhibitors. *ACS Infect. Dis.* **7**, 2425–2436 (2021).

[377] Ambler, R. P. The structure of β-lactamases. *Philos. Trans. R. Soc. London. B, Biol. Sci.* **289**, 321–331 (1980).

[378] Majiduddin, F. K., Materon, I. C. & Palzkill, T. G. Molecular analysis of beta-lactamase structure and function. *Int. J. medical Microbiol.* **292**, 127–137 (2002).

[379] Lucic, A. *et al.* Studies on the reactions of biapenem with VIM Metallo β-lactamases and the serine β-lactamase KPC-2. *Antibiotics* **11**, 396 (2022).

[380] Charnas, R. L. & Knowles, J. R. Inhibition of the RTEM β-lactamase from *Escherichia coli*. Interaction of the enzyme with derivatives of olivanic acid. *Biochemistry* **20**, 2732–2737 (1981).

[381] Kalp, M. & Carey, P. R. Carbapenems and shv-1 β-lactamase form different acyl-enzyme populations in crystals and solution. *Biochemistry* **47**, 11830–11837 (2008).

[382] Tremblay, L. W., Fan, F. & Blanchard, J. S. Biochemical and structural characterization of *Mycobacterium tuberculosis* β-lactamase with the carbapenems ertapenem and doripenem. *Biochemistry* **49**, 3766–3773 (2010).

[383] Ying, H. *et al.* mirnas; a novel strategy for the treatment of covid-19. *Cell Bio. in.* **45**, 2045–2053 (2021).

[384] Burki, T. K. Challenges in the rollout of covid-19 vaccines worldwide. *The Lancet Respir. Med.* **9**, e42–e43 (2021).

[385] Mode, D. & Stockholm, L. C. Astrazeneca: Evusheld long-acting antibody combination recommended for approval in the eu for the pre-exposure prophylaxis (prevention) of COVID-19.

[386] Sarrazin, C. & Zeuzem, S. Resistance to direct antiviral agents in patients with hepatitis C virus infection. *Gastroenterology* **138**, 447–462 (2010).

[387] Kieffer, T. L., Kwong, A. D. & Picchio, G. R. Viral resistance to specifically targeted antiviral therapies for hepatitis C (STAT-Cs). *J. Antimicrob. Chemother.* **65**, 202–212 (2010).

[388] Organization, W. H. *et al.* Global action plan on HIV drug resistance 2017–2021 (2017).

[389] Sacco, M. D. *et al.* The P132H mutation in the main protease of Omicron SARS-CoV-2 decreases thermal stability without compromising catalysis or small-molecule drug inhibition. *Cell Res.* **32**, 498–500 (2022).

[390] Mullard, A. 2020 fda drug approvals. *Nat. Rev. Drug Discov.* **20**, 85–91 (2021).

[391] Bauer, R. A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug Disc.* **20**, 1061–1073 (2015).

[392] Wong, K., Wan, D. & Chow, H. Substrate specificity and rational design of peptidomimetic inhibitors for SARS coronavirus main protease. *Hong Kong Med. J* **20**, 18–21 (2014).

[393] Drazˇic, T. *et al.* Peptide-β-lactam inhibitors of dengue and west nile virus NS2B-NS3 protease display two distinct binding modes. *J. Med. Chem.* **63**, 140–156 (2019).

[394] Schirmeister, T. & Klockow, A. Cysteine protease inhibitors containing small rings. *Mini Rev. Med. Chem.* **3**, 585–596 (2003).

[395] Siklos, M., BenAissa, M. & Thatcher, G. R. Cysteine proteases as therapeutic targets: does selectivity matter? A systematic review of calpain and cathepsin inhibitors. *Acta. Pharm. Sinica B* **5**, 506–519 (2015).

[396] Otero, L. H. *et al.* How allosteric control of staphylococcus aureus penicillin binding protein 2a enables methicillin resistance and physiological function. *Proc. Natl. Acad. Sci.* **110**, 16808–16813 (2013).

[397] Gonzales, P. R. *et al.* Synergistic, collaterally sensitive β-lactam combinations suppress resistance in MRSA. *Nat. Chem. Biol.* **11**, 855–861 (2015).

[398] Bush, K. & Bradford, P. A. β-lactams and β-lactamase inhibitors: an overview. *Cold Spring Harb. Perspect. Med.* **6**, a025247 (2016).

[399] Tang, S. S., Apisarnthanarak, A. & Hsu, L. Y. Mechanisms of β-lactam antimicrobial resistance and epidemiology of major community-and healthcare-associated multidrug-resistant bacteria. *Adv. Drug. Deliv. Rev.* **78**, 3–13 (2014).

[400] Pitout, J. D., Sanders, C. C. & Sanders Jr, W. E. Antimicrobial resistance with focus on β-lactam resistance in gram-negative bacilli. *The Am. J. Med.* **103**, 51–59 (1997).

[401] Wilmouth, R. C. *et al.* Inhibition of elastase by n-sulfonylaryl β-lactams: Anatomy of a stable acyl- enzyme complex. *Biochemistry* **37**, 17506–17513 (1998).

[402] Zandi, T. A. & Townsend, C. A. Competing off-loading mechanisms of meropenem from an L,D-transpeptidase reduce antibiotic effectiveness. *Proc. Natl. Acad. Sci.* **118**, e2008610118 (2021).