Analysis of Potassium (K^+) Efflux Systems (Kef) as an Antibiotic Target in Pathogenic Bacteria



Catherine Anna Haslam

St Hugh's College, University of Oxford

Thesis submitted for the degree

Doctor of Philosophy in Organic Chemistry

Trinity Term 2019

Under the supervision of:
Prof. Stuart Conway (University of Oxford)
Dr Andrew Scott (DSTL)

[dstl] EPSRC

SCHOOL OF CHEMISTRY

DECLARATION OF AUTHORSHIP

Name (in capitals): CATHERINE ANNA HASLAM **Candidate number: 1003778**

College (in capitals): ST HUGH'S **Supervisor: Prof. Stuart Conway**

Title of thesis (in capitals): ANALYSIS OF POTASSIUM (K^+) EFFLUX SYSTEMS (KEF) AS AN ANTIBIOTIC TARGET IN PATHOGENIC BACTERIA

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 8.7; available at https://www.ox.ac.uk/students/academic/student-handbook). ☑

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at https://www.ox.ac.uk/students/academic/guidance/skills?wssl=1 . ☑

The thesis I am submitting is entirely my own work except where otherwise indicated. ☑

It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution. ☑

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. ☑

I have clearly indicated the presence of all paraphrased material with appropriate references. ☑

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor. ☑

I have not copied from the work of any other candidate. ☑

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: http://www.admin.ox.ac.uk/statutes/352-051a.shtml). ☑

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. ☑

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism. ☑

Candidate's signature: …….. Date: ……..

Table of Contents

Abstract ..... **i**

Acknowledgements ..... **v**

Abbreviations ..... **ix**

Chapter 1 : Introduction **1**

1.1 Gram-Negative and Gram-Positive Bacteria 1

1.2 The Global Problem of Antibiotic Resistance 2

1.3 The Discovery of Antibiotics: from Golden Era to Resistance Era 3

1.4 Roles and Regulation of Potassium Ions (K^+) in Bacteria 9

1.5 Potassium (K^+) Efflux Systems (Kef) as Novel Antibiotic Targets 10

1.6 Structure of Kef 13

1.7 Proposed Mechanism of Inhibition/Activation 18

1.8 Structural Requirements for Kef Ligands 21

1.9 Previous Work on Kef as a Novel Antibiotic Target 23

1.10 Summary 27

1.11 Aims 28

Chapter 2 : Elucidation of the Kef Mechanism of Action Using a Fluorinated Kef Protein and ^{19}F NMR Spectroscopy **29**

2.1 Introduction and Aims 29

2.2 Synthesis of a Range of GSX Ligands 32

2.3 Determination of Ligand Function 35

2.4 Chapter 2 Conclusions 48

Chapter 3 : Identification of Kef Systems in Pathogenic Bacteria and Initial Testing of Compounds **49**

3.1 Introduction and Aims 49

3.2 Target Identification Through Bioinformatics 50

3.3 Synthesis of L-Methyl Esters OXFKEF12–14 (31–33) 56
3.4 Synthesis of D-Methyl Esters OXFKEF16 (48) and OXFKEF17 (35) 60
3.5 Molecular Cloning to Synthesise Plasmids for Expression of *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in *E. coli* 62
3.6 Kirby-Bauer Disc Diffusion Assay in *E. coli* Expressing *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2....64
3.6.1 Disc Assay for *E. coli* Expressing *Bp*Kef1 and *Bp*Kef2 64
3.6.2 Disc Assays for *E. coli* Expressing *Yp*KefGBP 66
3.6.2.1 Assays for *E. coli* MJF645 Expressing *Yp*KefGBP....66
3.6.2.2 Assays for *E. coli* Frag 1 Expressing *Yp*KefGBP 70
3.7 Determination of a Minimum Inhibitory Concentration (MIC) for *E. coli* MJF645 Expressing *Yp*KefGBP....74
3.8 Disc Assay in Wild-Type Bacteria....77
3.9 Determination of the Expression Level of *Yp*KefBH$_6$ in *E. coli* MJF645 and Frag 1....80
3.9.1 Molecular Cloning to Produce pTrcYpKefG(BH$_6$)P 81
3.9.2 Disc Assay for *E. coli* MJF645 and Frag 1 Expressing *Yp*KefG(BH$_6$)P....82
3.9.3 SDS-PAGE/Western Blotting of *Yp*KefBH$_6$ in *E. coli* MJF645 and Frag 1....84
3.10 Chapter 3 Conclusions....90

Chapter 4 : Synthesis and Testing of Thiol-ene-Derived Thioethers as Analogues of OXFKEF12–14 91

4.1 Introduction and Aims....91
4.2 Prediction of Compound Solubility Using the Solubility Forecast Index (SFI)....94
4.3 Synthesis of Compounds 80, 81, 82, and 83....96
4.4 Assays 100
4.4.1 Biophysical Assays with Compounds 82 and 83 100
4.4.1.1 WaterLOGSY....100
4.4.1.2 DSF 104
4.4.2 Kirby-Bauer Disc Diffusion Assays with Compounds 80 and 81....108
4.4.2.1 Disc Assay with Compound 80 108

4.4.2.2 Disc Assays with Compound 81 110

4.5 Chapter 4 Conclusions 121

Chapter 5 : Conclusions and Future Work 123

5.1 Conclusions 123

5.2 Future Work 126

Chapter 6 : Experimental 129

6.1 General Chemistry Methods 129

6.2 Chemistry Experimental Procedures 136

6.3 General Biochemistry Methods 168

6.3.1 Molecular Cloning to Synthesise Plasmids for Expression of *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in *E. coli* 168

6.3.1.1 General Strategy 168

6.3.1.2 Gibson Assembly to Produce pTrcBpKef2H_6 and pTrcYpKefGBP 169

6.3.1.3 Traditional Cloning to Produce pTrcBpKef1H_6 171

6.3.2 Molecular Cloning to Synthesise pTrcYpKefG(BH_6)P for Expression of *Yp*KefG(BH_6)P in *E. coli* 174

6.3.3 Chemicals and Cell Culture Reagents 176

6.3.4 Proteins and Plasmids 178

6.3.5 Protein Concentration Measurement 178

6.3.6 pH Measurement 179

6.3.7 Bacterial Cell Culture 179

6.4 Biochemistry Experimental Procedures 180

6.4.1 Constructs 180

6.4.2 Preparation of Competent *E. coli* 180

6.4.3 Molecular Cloning 181

6.4.3.1 Site-Directed Mutagenesis 181

6.4.3.2 Vector Linearisation 181

6.4.3.3 Gibson Assembly 182

6.4.3.4 Ligation 183
6.4.4 Agarose Gel Electrophoresis 183
6.4.5 Extraction of DNA from Agarose Gels 183
6.4.6 Transformation of Plasmid DNA into Bacteria 183
6.4.7 Extraction of Plasmid DNA from *E. coli* 185
6.4.8 Restriction Pattern Analysis of Plasmid DNA 185
6.4.9 Sequencing of Plasmid DNA 186
6.4.10 Expression and Purification of *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) 188
6.4.11 Expression of *Yp*KefBH$_6$ for Western Blotting Analysis 191
6.4.11.1 Growth Experiments 191
6.4.11.2 Membrane Preparation 192
6.4.12 Pierce™ 660 nm Protein Assay 194
6.4.13 SDS-PAGE 194
6.4.13.1 Sample Preparation 195
6.4.14 Western Blot Analysis 196
6.4.15 LC-MS 197
6.4.16 CD 197
6.4.17 AUC 198
6.4.18 SEC-MALS 198
6.4.19 Biophysical Assays 198
6.4.19.1 ITC 198
6.4.19.2 WaterLOGSY 199
6.4.19.3 Protein-Observed ^{19}F NMR Spectroscopy 201
6.4.19.4 DSF 201
6.4.20 Assays in Bacteria 202
6.4.20.1 Kirby-Bauer Disc Diffusion Assays 202
6.4.20.2 MIC and MBC Assays 205

References 207

Appendices 211

ANALYSIS OF POTASSIUM (K^+) EFFLUX SYSTEMS (KEF) AS AN ANTIBIOTIC TARGET IN PATHOGENIC BACTERIA

Catherine Haslam **St Hugh's College** **DPhil in Organic Chemistry** **Trinity Term 2019**

Abstract

Potassium (K^+) efflux systems (Kef) are K^+/H^+ antiporters found in most Gram-negative bacteria. Kef systems play a vital role in the protection of bacteria from toxic electrophiles by regulating cytoplasmic pH and consequently demonstrate potential as an antibiotic target. Kef is inhibited by glutathione (γ-L-Glu-L-Cys-Gly, GSH) and activated by glutathione S-conjugates (GSX). The proposed model of Kef activation in both *Ec*KefC, from *Escherichia coli* (*E. coli*), and *Sd*KefQCTDH$_6$, a construct of Kef from *Shewanella denitrificans* (*S. denitrificans*), is a single amino acid switch at the F441 or F448 capping residue, respectively. The aims of this dissertation were to confirm the proposed mechanism of Kef activation and to investigate novel Kef systems in pathogenic bacteria as an antibiotic target.

In work to probe the mechanism of Kef activation, *p*-trifluoromethyl-L-phenylalanine (*p*tfmF) was incorporated into *Sd*KefQCTDH$_6$ at position 448 to give *Sd*KefQCTDH$_6$(F448*p*tfmF). The conformational change of *Sd*KefQCTDH$_6$(F448*p*tfmF) upon adding inhibitory or activating ligands was analysed using ^{19}F NMR spectroscopy. The data obtained are consistent with the proposed model for Kef activation: activators displace the capping residue, whereas inhibitors, including GSH, do not. This novel ^{19}F NMR-based assay is an efficient method to determine whether a compound is a Kef activator or inhibitor, vital in the design of new Kef ligands.

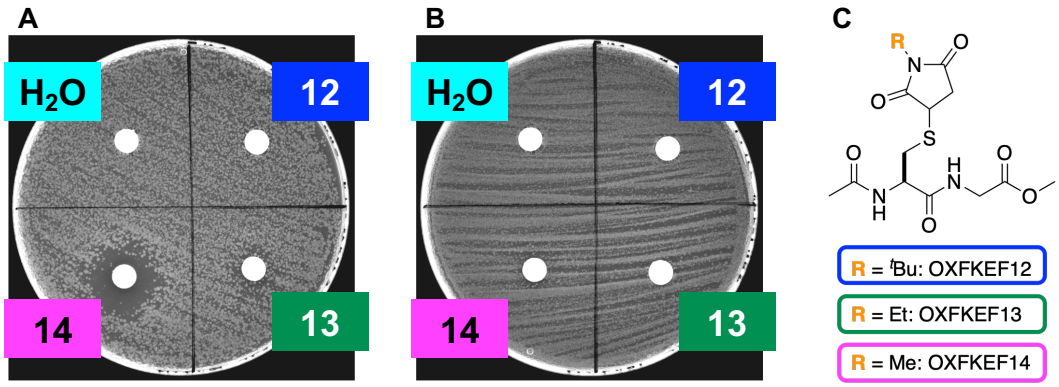


Figure: Kirby-Bauer disc diffusion assay data for *E. coli* MJF645 and Frag 1 expressing *Yp*KefGBP. 12 = OXFKEF12, 13 = OXFKEF13, and 14 = OXFKEF14. OXFKEF12–14 were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcYpKefGBP (0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG)). B) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). C) Structures of OXFKEF12–14.

To investigate whether the role played by Kef in *E. coli* is conserved across Gram-negative bacteria, Kef systems were identified in two pathogenic bacteria: *Yersinia pestis* (*Yp*KefB; with ancillary proteins -G and -P) and *Burkholderia pseudomallei* (*Bp*Kef1 and *Bp*Kef2). Dipeptides OXFKEF12–14 were synthesised and tested in *E. coli* expressing *Yp*KefGBP, *Bp*Kef1, or -2 using the Kirby-Bauer disc diffusion assay (Figures A and B). In *E. coli* MJF645 (mutant strain lacking GSH) and Frag 1 (GSH-containing strain) expressing *Yp*KefGBP, OXFKEF13 and -14 were active only in the MJF645 transformant, as a zone of growth inhibition/death was observed (Figure A). To rationalise these results, the expression level of *Yp*KefBH_6 (a His-tagged variant of *Yp*KefB) was determined using western blotting and was found to be similar in both MJF645 and Frag 1. Due to its sequence similarity to *Yp*KefBH_6, it can be inferred that the expression level of *Yp*KefB is also comparable in both cell strains, and so the opposing disc assay results must be caused by another factor. GSH may be inhibiting *Yp*KefB, and so the dipeptides may be unable to outcompete it in the ligand binding

site. OXFKEF12–14 could also be being degraded by GSH, giving compounds which inhibit *Yp*KefB or are inert.

As OXFKEF12–14 may be unstable in the presence of GSH, we wanted to synthesise dipeptides with improved stability with respect to GSH. Novel thioether analogues with aromatic S-substituents were synthesised using the thiol-ene reaction: two methyl esters for the disc assay and their carboxylic acid analogues for biophysical assays. Both carboxylic acids were shown to bind to *Sd*KefQCTDH$_6$ using two orthogonal assays. Both methyl esters did not inhibit or kill MJF645 expressing *Sd*Kef or *Yp*KefGBP in the disc assay. The inactivity of the compounds may have been caused by several factors, including their poor solubility. The results indicate that aromatic S-substituents are tolerated in terms of Kef binding.

In conclusion, we have shown that *Yp*KefB is similar to *Ec*KefC, both in terms of protein structure and function. In contrast, *Bp*Kef1 and *Bp*Kef2 behave differently from *Ec*KefC. In *Burkholderia pseudomallei*, the endogenous low molecular weight thiol may not be GSH or the role of its Kef systems may be dissimilar from *E. coli*.

Acknowledgements

I would like to thank my supervisor Prof. Stuart Conway for giving me the opportunity to study for a DPhil in his group. Thank you for your continued guidance and support, and for encouraging me, especially when the science wasn't going to plan. You've taught me so much about scientific research, I'm sure it will be invaluable for my future career.

I would also like to thank my supervisors from DSTL: originally Dr Helen Atkins, and later, Dr Andrew Scott. Thank you both for your support and expertise. Our meetings really helped to shape the project and solve biological problems. Thank you also to Dr Phillip Ireland for his advice with cloning at the start of the project. I would also like to thank Dr Christopher Jenkins for supervising me during my short placement at DSTL. It was a great learning experience and I'm really pleased with the data I gained.

Thank you to all the staff in the CRL who are constantly working to keep the department functioning. I would also like to thank the CVR group, managed by Dr Victoria Cullen and Dr Andrew Dolan, who kindly allowed me to use several pieces of equipment in their lab.

Now to thank the Conway Group, who have been a lovely group of people to work with over the last four years. Firstly, thank you to everyone who read chapters of my thesis. Secondly, I would like to thank Team Kef: Anthony, thank you for your help with biochemistry. I really appreciated your patience; Angelika, thanks for making the two compounds which were useful positive controls for the biophysical assays; Holly, my partner in Krime, it's been lovely working with you on Kef. Thanks for being a caring

friend; Mary, thanks for being a great Part II in 2017/18. Best of luck with the rest of your DPhil! To the postdocs: Amélie, thank you for making me feel so welcome in the group. Thanks for teaching me your lab tips and tricks, I hope you publish these as a book very soon! Ewen, thanks for your advice and help over the last 3 years, especially with the HPLC! Thanks for putting up with my lab chat, I hope your new neighbour at the Crick is quieter; James, fountain of knowledge, thank you for all of your chemistry advice. Thanks for being so approachable, it was always great hearing your anecdotes about when you were in Ann Arbor! Maze, thanks for your advice, especially with biochemistry. I was amazed by your scientific knowledge and musical talent; Debs, I'm so glad you joined the Conway Group (finally). Thanks for your advice in the lab and for being a supportive friend; Pietro, thanks for the good science chats. Thanks for delivering the group's Christmas presents as Santa last year, it was great! To my fellow DPhils: Sarah, I really appreciated you taking the time to give me sage advice, especially at the start of my DPhil. Thanks for being a great friend. Larissa and Coco, it's been lovely sharing this DPhil journey with you both. Larissa, thank you for bringing your delicious homemade cakes into the lab, you're a very talented baker. It was lovely hosting you at my home last summer, you're welcome anytime! Coco, I will always remember us dancing around G11, your sense of fun is so infectious! Joe, thank you for always being a listening ear. Thanks also for your very helpful PyMOL lesson; Amy, thanks for always being up for a chat about problematic chemistry. I think we concluded that everything can be solved with ice cream! Mus keep on wearing amazing t-shirts! I enjoyed all of our chats about halloumi, Stormzy and everything in between; Jess, I admire your positive attitude in the lab. Thanks for all of the cute cat videos! Adam, thanks for bringing both quality jokes and music to the lab. I'll keep

listening to your Sunday Morning Breakfast & Chill playlist; Darius and Katrina, unfortunately we haven't had much overlap of time in the lab. Good luck with your DPhils in the group! Many Part IIs have come and gone during my four years in the group, so I'll just mention the 2018/19 crew: Ali, Andrew, James, Michelle, and Oli – it's been a blast, good luck for your future careers.

Thanks to the staff of St Hugh's College for helping me with everything from accommodation to course administration. It's been great being a member of the MCR too. Thanks especially to Josh and Paula for being supportive friends. Thanks also to my undergrad friends for being so understanding about the demands of a DPhil. Finally, I would like to thank my family, especially my parents and brother, for supporting me throughout my DPhil, as well as through my undergraduate degree.

Abbreviations

2D	Two-dimensional
A	Absorbance
Å	Angstroms
A. baumannii, *Ab*	*Acinetobacter baumannii*
Ac	Acetyl
AMP	Adenosine monophosphate
Amp	Ampicillin
AMR	Antimicrobial resistance
aq.	aqueous
Ar	Argon
ATR	Attenuated total reflection
AUC	Analytical ultracentrifugation
BAB	Brucella agar base
B. anthracis, *Ba*	*Bacillus anthracis*
B. pseudomallei, *Bp*	*Burkholderia pseudomallei*
br	broad
BSA	Bovine serum albumin
c	concentration
C. burnetii, *Cb*	*Coxiella burnetii*
CD	Circular dichroism
cm^{-1}	Reciprocal centimetres
COSY	Correlation spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CTD	C-terminal domain
CV	Column volumes
Cys	Cysteine
d	doublet
Da	Daltons
DDM	*n*-Dodecyl β-D-maltoside
dec.	decomposed
DELTA-BLAST	Domain Enhanced Lookup Time Accelerated-Basic Local Alignment Search Tool
DIPEA	*N*,*N*-Diisopropylethylamine

DMAP	*N*,*N*-Dimethyl-4-aminopyridine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorimetry
DSTL	Defence Science Technology Laboratory
ε	Molar extinction coefficient
E. coli, *Ec*	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
e.e.	enantiomeric excess
eq.	equivalents
ESG	*N*-Ethylsuccinimido-*S*-glutathione
ESI	Electrospray ionisation
ESKAPE	***E****nterococcus faecium,* ***S****taphylococcus aureus,* ***K****lebsiella pneumoniae*, ***A****cinetobacter baumannii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* species
Et	Ethyl
et al.	and others
Fmoc	9-Fluorenylmethoxycarbonyl
FPLC	Fast protein liquid chromatography
F. tularensis, *Ft*	*Francisella tularensis*
g	gram(s)
Glu	Glutamic acid
Gly	Glycine
GSH	Glutathione (γ-L-Glu-L-Cys-Gly)
GSX	Glutathione S-conjugate
h	hour(s)
HBTU	*N,N,N′,N′*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HF	High-fidelity
HMBC	Heteronuclear multiple quantum correlation
HPLC	High performance liquid chromatography
HRMS	High-resolution mass spectrometry
HRP	Horseradish peroxidase

HSQC	Heteronuclear single quantum correlation
IMAC	Immobilised metal affinity chromatography
IPA	Isopropyl alcohol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
ITC	Isothermal calorimetry
J	Coupling constant
kb	kilobases
kDa	kilodaltons
Kef	Potassium (K^+) efflux system
K. pneumoniae, *Kp*	*Klebsiella pneumoniae*
KTN	Potassium (K^+) transport and nucleotide-binding
l	Path length
LacI	*lac* repressor
LacO	*lac* operon
LC-MS	Liquid chromatography-mass spectroscopy
LDS	Lithium dodecyl sulfate
lit.	Literature value
LRMS	Low-resolution mass spectrometry
M	Molar (mol dm^{-3})
m	multiplet, medium intensity
MBC	Minimum bactericidal concentration
MDR	Multidrug-resistant
Me	Methyl
MF	Membrane fraction
MG	Methylglyoxal
mg	milligram
MGS	Methylglyoxal synthase
MHz	Megahertz
MIC	Minimum inhibitory concentration
min	minute(s)
mL	millilitre(s)
mp	Melting point
MS	Molecular sieves

MWCO	Molecular weight cut-off
N/A	Not applicable
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
NEM	*N*-Ethylmaleimide
NMR	Nuclear magnetic resonance
NOE	Nuclear overhauser effect
NP	Normal phase
NPM	*N*-Phenylmaleimide
nr	Non-redundant
ns	Not significant
OD, OD_{600}	optical density, OD measured using light with a wavelength (λ) of 600 nm
PBS	Phosphate-buffered saline
PDB	Protein Data Bank
PDR	Pandrug-resistant
PEI	Polyethyleneimine
Ph	Phenyl
ppm	parts per million
*p*tfmF	*para*-Trifluoromethyl-L-phenylalanine
py	Pyridine
Q	Q-linker
q	quartet
ret. time	retention time
R_f	Retention factor
RNA	Ribonucleic acid
rpm	revolutions per minute
rt	Room temperature
s	singlet, strong intensity
SAR	Structure-activity relationship
s.d.	Standard deviation
S. denitrificans, *Sd*	*Shewanella denitrificans*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography

SEC-MALS	Size exclusion chromatography-multi-angle laser scattering
SFI	Solubility forecast index
SLG	*S*-Lactoylglutathione
SN	Supernatant
Sol.	Solvent
Spec	Spectinomycin
SPS	Solvent purification system
Syr	Syringe
T	Temperature
t	triplet
TAE buffer	Tris, acetic acid, and EDTA buffer
TBS	Tris-buffered saline
^{t}Bu	*tert*-Butyl
^{t}BuSG	*N*-*tert*-Butylsuccinimido-*S*-glutathione
TEMED	*N*,*N*,*N*′,*N*′-Tetramethylethane-1,2-diamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting temperature
ΔT_m	Thermal shift
TMS	Tetramethylsilane
Tris (Trizma®)	Tris(hydroxymethyl)aminomethane
Trt	Trityl
TTBS	TBS with Tween® 20
UV	Ultraviolet
v/v	volume per unit volume (volume to volume ratio)
w	weak intensity
waterLOGSY	water-ligand observed *via* gradient spectroscopy
WHO	World Health Organisation
w/v	weight per unit volume (weight to volume ratio)
w/w	weight per unit weight (weight to weight ratio)
XDR	Extensively drug-resistant
Y. pestis, *Yp*	*Yersinia pestis*
Y. pseudotuberculosis, *Ys*	*Yersinia pseudotuberculosis*

Chapter 1: Introduction

1.1 Gram-Negative and Gram-Positive Bacteria

Bacteria are either Gram-negative or Gram-positive, according to the results obtained with the Gram stain.[1] The modern staining procedure uses a crystal violet–iodine complex and a safranin counterstain, which stains Gram-negative bacteria pink and Gram-positive bacteria purple.[2] In the cell walls of Gram-negative bacteria, there is an additional outer membrane and smaller peptidoglycan layer than in the cell walls of Gram-positive bacteria (Figure 1.1).

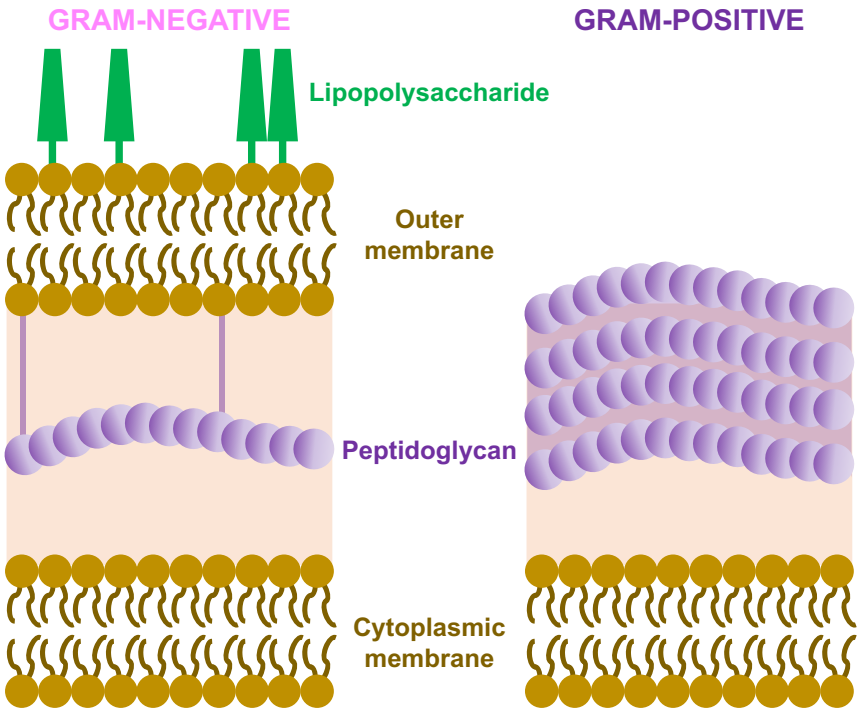


Figure 1.1. Schematic of the cell walls in Gram-negative and Gram-positive bacteria.[3]

1.2 The Global Problem of Antibiotic Resistance

Antibiotic resistance is defined as the continued growth of bacteria in the presence of a lethal concentration of an antibiotic.[4,5] Bacteria can develop resistance by several different mechanisms: transformation, transduction, conjugation, and mutation.[6,7] Overuse of antibiotics in agriculture and medicine has led to widespread resistance,[8] with the emergence of multidrug-resistant (MDR), extensively drug-resistant (XDR), and pandrug-resistant (PDR) strains.[9] MDR bacteria are resistant to at least one antibiotic in three or more different classes.[9] XDR bacteria are susceptible to one or two antibiotic classes only and are non-susceptible to all others.[9] PDR bacteria are resistant to all antibiotics from all classes tested.[9] Multidrug resistance in the ESKAPE pathogens (***E****nterococcus faecium,* ***S****taphylococcus aureus,* ***K****lebsiella pneumoniae*, ***A****cinetobacter baumannii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* species) is highly concerning, as these pathogens cause the majority of hospital-acquired infections in the United States.[10] It is estimated that antimicrobial resistance (AMR) currently causes 700,000 deaths per year (low estimate) but this is expected to rise to 10 million deaths per year by 2050.[11] In this "resistance era", medical advances, including surgery, cancer chemotherapy, and organ transplantation, are at risk.[8] The problem of antibiotic resistance is compounded by the general lack of new antibiotics being discovered, which will be described in the next section.

1.3 The Discovery of Antibiotics: from Golden Era to Resistance Era

Penicillin (**1,** Figure 1.2), discovered in 1929, was one of the first antibiotic drugs. Streptomycin (**2**, Figure 1.2) was discovered subsequently, in 1943. These early discoveries sparked a golden era of antibiotic discovery, in the 1930s–1960s.[12] During this time, most of the currently available antibiotic drugs were discovered (Table 1.1).[8,13] However, a discovery void followed in the 1970s–1990s.[8] This was caused by a lack of new antibiotic targets being identified and a general failure of high-throughput screens to reveal lead compounds.[14,15] Due to these factors, and a lack of financial incentive, many large pharmaceutical companies abandoned their research and development programs for antibiotics.[16] During the period of the discovery void, all new antibiotics were analogues of existing drugs, except for mupirocin (**3**, Figure 1.2), a topical antibiotic for the treatment of infections caused by Gram-positive bacteria, which was introduced in 1985.[13]

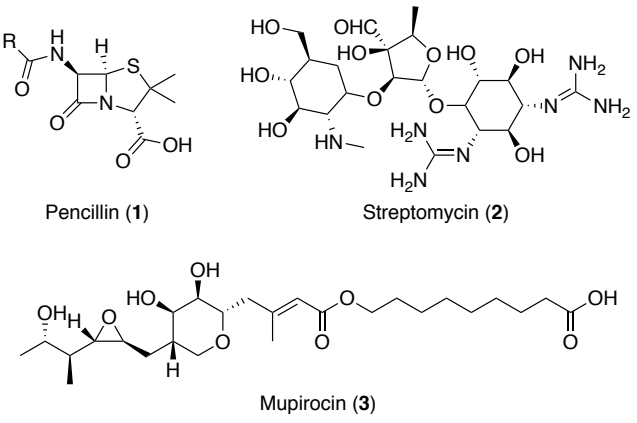
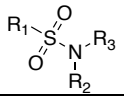
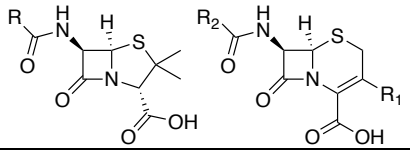
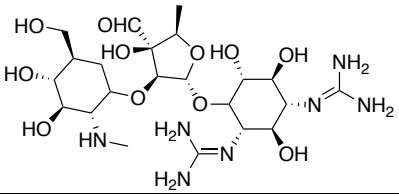
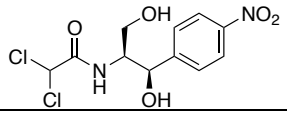
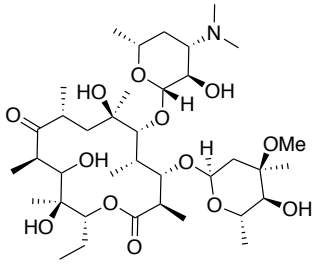
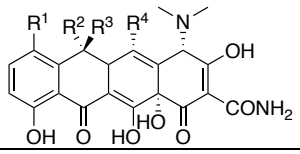
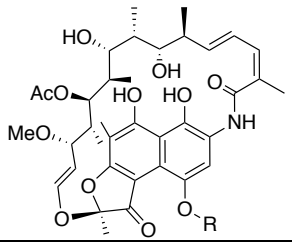
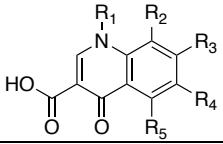
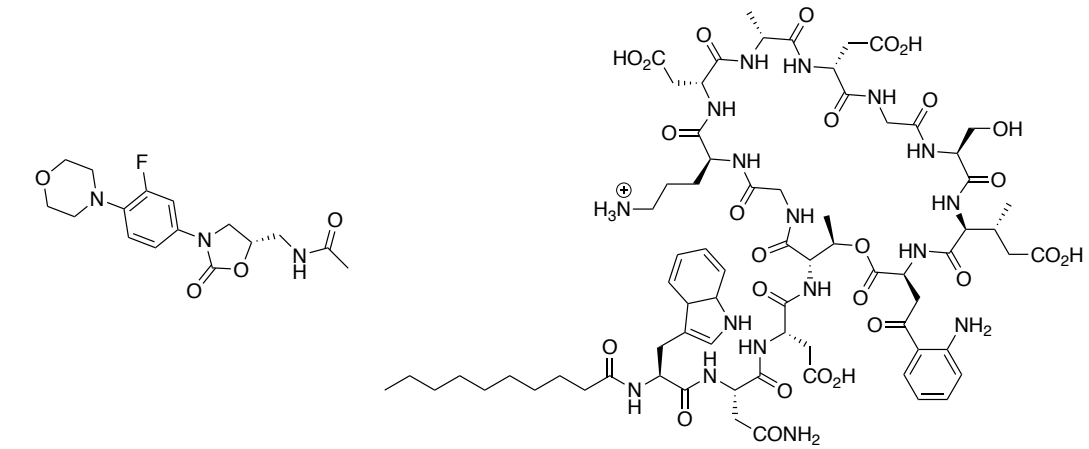


Figure 1.2. Structures of penicillin (1), streptomycin (2), and mupirocin (3).[17]

Table 1.1. Antibiotic classes introduced in the 1930s–1960s.[18]

Antibiotic Class	Introduced	Mechanism of Action
Sulfonamides 	1936	Inhibit folic acid biosynthesis, preventing bacterial growth and reproduction
β-Lactams 	1938	Inhibit bacterial cell wall biosynthesis
Aminoglycosides 	1946	Bind to the 30S ribosomal subunit, preventing translation
Chloramphenicols 	1948	Bind to the 50S ribosomal subunit, preventing translation
Macrolides 	1951	Bind to the 50S ribosomal subunit, preventing translation
Tetracyclines 	1952	Bind to the 30S ribosomal subunit, preventing translation
Rifamycins 	1958	Bind to the RNA polymerase β-subunit, preventing transcription
Glycopeptides	1958	Inhibit bacterial cell wall biosynthesis
Quinolones 	1968	Inhibit deoxyribonucleic acid (DNA) synthesis

Since 2000, five new antibiotic classes have been introduced (Figure 1.3): linezolid (**4**, oxazolidinone), daptomycin (**5**, lipopeptide), retapamulin (**6**, pleuromutilin), fidaxomicin (**7**, tiacumicin), and bedaquiline (**8**, diarylquinoline).[19] However, these are only effective against Gram-positive bacteria.[19] It is therefore critical that new antibacterial classes are developed for Gram-negative bacteria,[19] where the outer membrane represents a major barrier for antibiotic permeability into cells.[7]

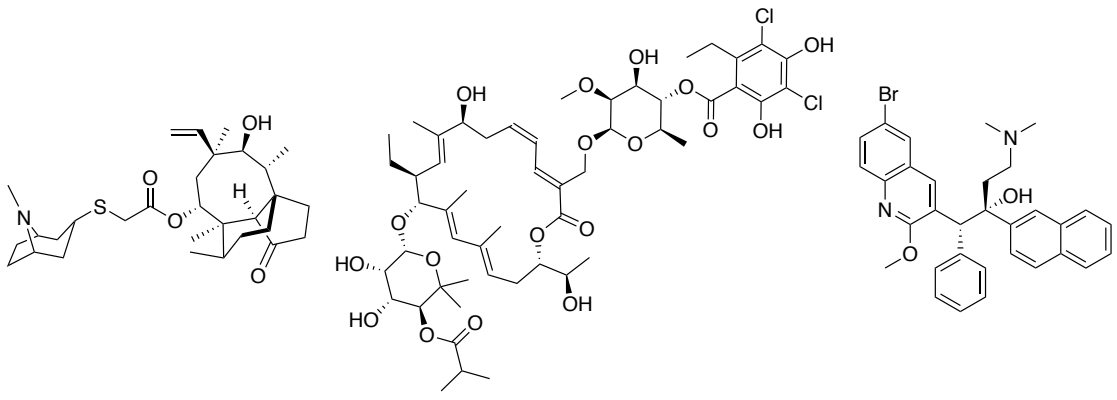


Linezolid (**4**)

(Approved in 2000)

Daptomycin (**5**)

(Approved in 2003)



Retapamulin (**6**)

(Approved in 2007)

Fidaxomicin (**7**)

(Approved in 2010)

Bedaquiline (**8**)

(Approved in 2012)

Figure 1.3. Structures of recently approved antibiotics.[17,20,21]

To assess the extent of worldwide antibiotic resistance, the World Health Organisation (WHO) published its first report on the topic in 2014.[22] Information was amassed about resistance to antibiotics frequently used to treat infections caused by bacterial species of global concern, for nine combinations of antibiotics and bacteria (Table 1.2, Figure 1.4).[22] The information was obtained from multiple international sources and analysed by WHO.[22] The results for *Escherichia coli* (*E. coli*) and *Klebsiella pneumoniae* (*K. pneumoniae*) are particularly concerning. *E. coli* generally showed high levels of resistance to the last generation antibiotics used to treat severe infections, fluoroquinolones (**9**) and 3rd-generation cephalosporins (**10**). In the majority of studies, *K. pneumoniae* showed >30% resistance to 3rd-generation cephalosporins (**10**), rising to >60% in some cases.[22] In some patient groups, *K. pneumoniae* demonstrated >50% resistance to carbapenems (**11**), which are generally considered to be antibiotics of "last resort" for these infections.[22]

Table 1.2. Summary of results from the WHO study on antimicrobial resistance.[22] These data show that there is a high incidence of resistance to the antibiotics commonly used to treat infections caused by these pathogenic bacteria.

Bacterial Species	Classification	Resistance
Escherichia coli (*E. coli, Ec*)	Gram-negative	• Fluoroquinolones (**9**) • 3rd-Generation cephalosporins (**10**)
Klebsiella pneumoniae (*K. pneumoniae, Kp*)	Gram-negative	• Carbapenems (**11**) • 3rd-Generation cephalosporins (**10**)
Neisseria gonorrhoeae	Gram-negative	Reduced susceptibility to 3rd-generation cephalosporins (**10**)
Shigella	Gram-negative	Fluoroquinolones (**9**)
Nontyphoidal *Salmonella*	Gram-negative	Fluoroquinolones (**9**)
Staphylococcus aureus	Gram-positive	β-Lactams, including methicillin (**12**)
Streptococcus pneumoniae	Gram-positive	Resistance and/or non-susceptibility to penicillin (**1**)

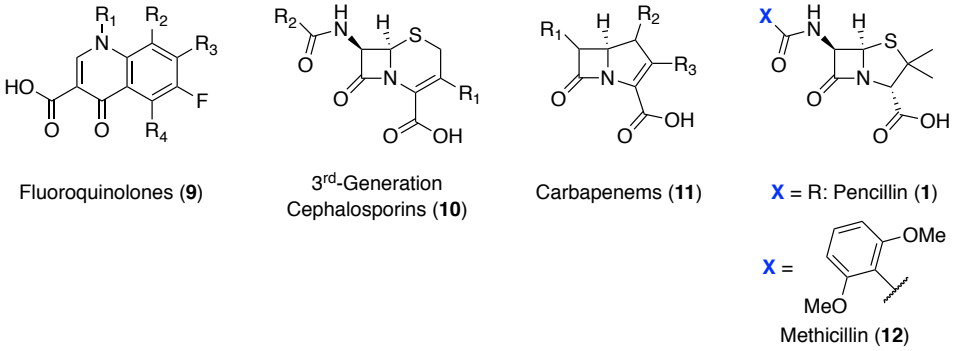


Figure 1.4. Structures of antibiotics analysed in the WHO study on antimicrobial resistance.[22]

The severity of antibiotic resistance is not matched by the number of drugs in the pipeline. Antibacterial drug discovery in this current "resistance era" has been dominated by a target-based approach, which has generally failed to reveal novel antibiotics.[8] In February 2019, the current antibiotic pipeline was analysed in an extensive review.[23] The current pipeline contains 30 novel antibacterials but these are primarily derivatives of existing antibiotics.[23] Of particular urgency are antibiotics for the treatment of XDR and PDR Gram-negative bacteria.[23] These novel antibiotics must also lack cross-resistance, which is when treatment with one antibiotic causes bacteria to acquire resistance to another drug in the same class. Of the 30 new antibiotics, eleven are predicted to show activity in at least one species of XDR Gram-negative bacteria.[23] Out of the 11 compounds, ten are derivatives of existing classes of β-lactams.[23] Although these drugs have been designed to solve some of the class-specific resistance mechanisms, it is not always viable to do so using antibiotics from the same class.[23] Co-resistance, where a bacterial strain is resistant to more than one class of antibiotics, is also possible.[23] These XDR bacterial strains therefore represent a considerable challenge in the design of new antibacterial drugs. An effective strategy to treat these pathogenic bacteria could be to develop novel antibiotics which disrupt bacterial homeostasis.

1.4 Roles and Regulation of Potassium Ions (K^+) in Bacteria

The four main roles of potassium ions (K^+) in bacteria are: acting as an osmotic solute to control turgor pressure; activation of intracellular enzymes; regulation of cytoplasmic pH; and acting as a secondary messenger.[24] Due to the plethora of roles played by K^+ in bacteria, it is vital that the concentration of K^+ is precisely regulated. Uptake of K^+ is controlled by the Trk (**K**$^+$ **tr**ansport), Kdp (**K**$^+$ **d**e**p**endent), and Kup (**K**$^+$ **up**take) systems.[24] In Gram-negative bacteria, K^+ efflux is controlled by potassium (K^+) efflux systems (Kef), which are putative K^+/H^+ antiporters.[24] In Gram-positive bacteria, there is a related protein, the KhtSTU K^+ efflux pump.[25] The mechanosensitive channels in bacteria provide a non-specific path for K^+ efflux, whereby the channels open upon osmotic downshock, to prevent cell rupture.[24]

1.5 Potassium (K^+) Efflux Systems (Kef) as Novel Antibiotic Targets

Kef systems not only regulate K^+ concentration, but also protect bacteria from toxic electrophiles, including methylglyoxal (MG, **13**, Figure 1.5). In bacteria, MG (**13**) is formed *via* glycolysis, using methylglyoxal synthase (MGS), under conditions of carbon excess or phosphate limitation.[26,27] MG (**13**) can react with nucleophilic sites in macromolecules including DNA, RNA, and proteins, leading to cell death.[28]

In *E. coli*, enzymes glyoxalase I and II metabolise toxic MG (**13**) to non-toxic D-lactate (**18**) (Figure 1.5). This detoxification is dependent on cytoplasmic glutathione (γ-L-Glu-L-Cys-Gly, GSH, **14**, Figure 1.6), which is present at ~20 mM in *E. coli*.[29,30] MG (**13**) reacts spontaneously with GSH (**14**) to give the hemithiol acetal **15** (Figure 1.5).[28,31] Compound **15** is isomerised by glyoxalase I to compound **16**, which is further isomerised to the glutathione S-conjugate (GSX, Figure 1.6) *S*-lactoylglutathione (SLG, **17**) (Figure 1.5). Once the concentration of MG (**13**) has been sufficiently reduced, SLG (**17**) is metabolised by glyoxalase II to D-lactate (**18**), which also regenerates the cytoplasmic GSH (**14**).[28]

E. coli contains two Kef systems, *Ec*KefFC and *Ec*KefGB, which provide additional protection from MG (**13**) by acidification of the bacterial cytoplasm.[32] *Ec*KefF and -G are ancillary proteins for *Ec*KefC and -B, respectively, and are required for optimal Kef activity.[33] Both *Ec*KefFC and -GB are inhibited by GSH (**14**) and activated by the GSX SLG (**17**) (Figure 1.5).[28] Activation of *Ec*KefFC and -GB leads to K^+ efflux and concomitant H^+ influx, resulting in a drop in cytoplasmic pH. In *E. coli*, MG (**13**) causes an initial decrease of 0.4–0.6 pH units,[28] which is proposed to partially protonate intracellular nucleophiles, reducing their reactivity with the electrophile and

protecting the cell.[32,34] In work by Booth *et al.*, *E. coli* mutants lacking both *Ec*KefB and -C were more sensitive to MG (**13**) and *N*-ethylmaleimide (NEM, **19**, Figure 1.6) than their wild-type counterparts.[34,35]

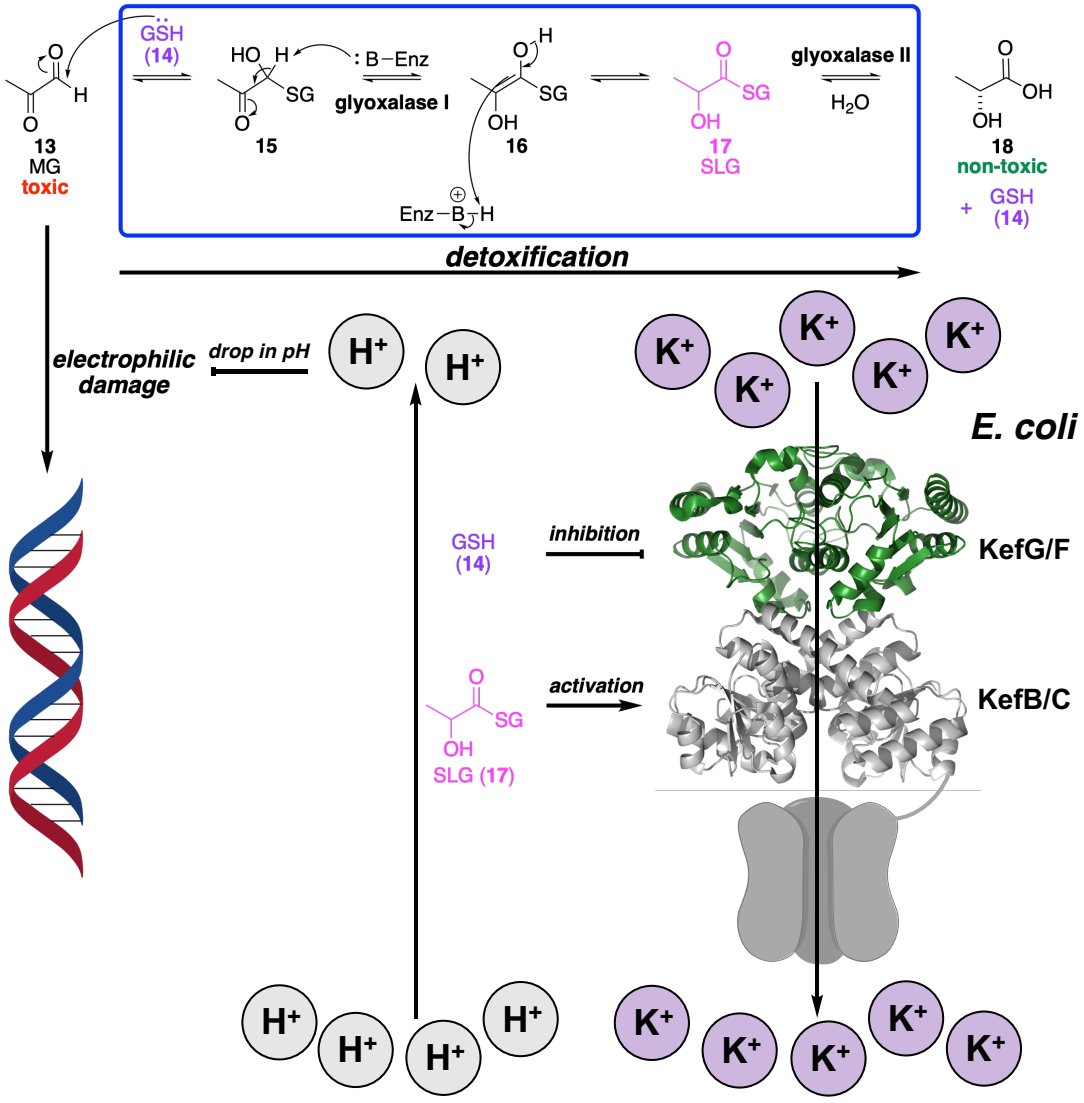


Figure 1.5. Schematic of the synergistic actions in *E. coli* to protect against MG (13).[31]

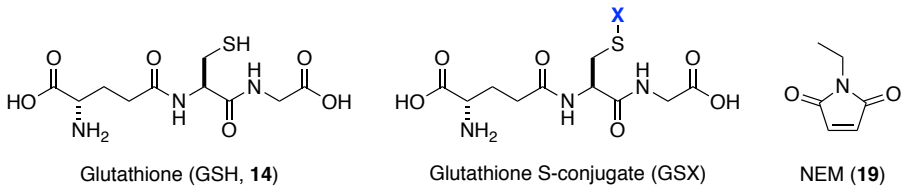


Figure 1.6. Structures of glutathione (GSH, 14), glutathione S-conjugate (GSX), and *N*-ethylmaleimide (NEM, 19).

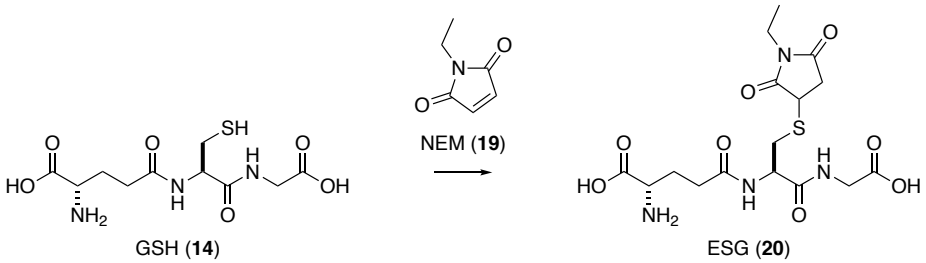
In work by Booth *et al.*, activation of *Ec*KefB was shown to aid cell survival after exposure to MG (**13**) and this effect was more important for survival than detoxification.[35,36] Kef systems have therefore been proposed as potential antibiotic targets.[31,35] A DELTA-BLAST search by Dr Anthony Chan revealed Kef homologues in many pathogenic bacteria, which were subsequently aligned (Table 1.3, unpublished). This analysis was carried out relative to *Sd*Kef (from *Shewanella denitrificans*, a Gram-negative marine bacterium). *Sd*Kef is used as a model for *Ec*KefC, due to its sequence homology and lack of ancillary protein.[31,37] Kef systems are present in two of the ESKAPE pathogens: *Klebsiella pneumoniae* (*Kp*KefC) and *Acinetobacter baumannii* (*Ab*Kef) (Table 1.3). Kef systems do not possess any analogues in humans. These data demonstrate the opportunity to develop antibiotics which target Kef.

Table 1.3. Alignment of Kef homologues. This analysis was carried out by Dr Anthony Chan (unpublished). Each corresponding Kef protein sequence was aligned to the full-length native *Sd*Kef (GenBank ID: ABE53663.1) by using global pairwise alignment (EMBOSS Needle: http://www.ebi.ac.uk/Tools/psa/),[38] with default settings (matrix: BLOSUM62; gap open penalty: 10.0; gap extend penalty: 0.5). EMBOSS Needle creates an optimal global alignment of two sequences using the Needleman-Wunsch algorithm. Identity (%): Extent to which amino acid sequences of the Kef system and *Sd*Kef query have the same residues at identical positions in the alignment, expressed as a percentage. Positives (%): Extent to which amino acid sequences of the Kef system and *Sd*Kef query have residues with similar physicochemical properties at identical positions in the alignment, expressed as a percentage. Gaps (%): Spaces introduced in the alignment of the amino acid sequences of the Kef system and *Sd*Kef query to account for insertions and deletions in each sequence compared to the other, expressed as a percentage.[39]

Species	Kef Protein	GenBank ID	Identity (%)	Positives (%)	Gaps (%)
Escherichia coli	*Ec*KefC	NP_414589.1	42.9	60.5	7.5
	*Ec*KefB	NP_417809.1	42.6	64.2	5.0
Acinetobacter baumannii	*Ab*Kef	KDF03402.1	39.1	58.5	6.8
Citrobacter amalonaticus	*Ca*KefC	WP_046475364.1	44.0	61.1	7.2
Klebsiella pneumoniae	*Kp*KefC	AHI34262.1	42.7	61.7	6.2
Legionella pneumophila	*Lp*Kef	WP_027222075.1	46.8	65.8	4.1
Salmonella enterica	*Se*KefC	NP_454698.1	43.4	60.2	6.9
Shigella flexneri	*Sf*KefC	WP_039055329.1	42.8	60.5	7.5
Shigella sonnei	*Ss*KefC	WP_024249657.1	42.8	60.3	7.5
Vibrio cholerae	*Vc*KefB	AAF95747.1	41.5	59.8	15.7

1.6 Structure of Kef

Three X-ray crystal structures of the C-terminal domain (CTD) of *Ec*KefC (*Ec*KefCCTD) bound to KefF have been published: an *apo* structure (Protein Data Bank (PDB) ID: 3EYW; 2.4 Å resolution);[40] a structure with the inhibitor GSH (**14**) bound (PDB ID: 3L9W; 1.75 Å resolution);[41] and a structure with the Kef activator *N*-ethylsuccinimido-*S*-glutathione (ESG, **20**) bound (PDB ID: 3L9X; 2.10 Å resolution).[41] ESG (**20**) is the GSX of NEM (**19**) (Scheme 1.1). ESG (**20**) is a 1:1 mixture of 2 diastereoisomers, which are formed by attack of the thiol group of GSH (**14**) on both faces of NEM (**19**).[41]



Scheme 1.1. Formation of ESG (20) from GSH (14) and NEM (19).

The three X-ray structures of *Ec*KefCCTD-KefF are broadly similar, and show that *Ec*KefFC is a homodimeric protein, comprising two KefC subunits and two KefF subunits (Figure 1.7, ESG-bound structure). Each *Ec*KefC subunit contains a transmembrane domain at its N-terminus and a potassium (K^+) transport and nucleotide-binding (KTN) domain at its C-terminus (*Ec*KefCCTD).[41] KTN domains are present in many K^+ transporters and channels, and the Rossman fold (G-X-G-X-X-G) is a common motif, to which nucleotides bind.[42] Each *Ec*KefC subunit contains a Rossman fold with one adenosine monophosphate (AMP) molecule bound.[41] Each Rossman fold contains a six-stranded β-sheet at its core, with a set of three α-helices at either side.[42] In each, the C-terminal α-helix (α6) is exchanged with the one from the other motif.[42] The two component Rossman folds are connected *via* a hinge region, which is formed by the loops linking the C-terminal strands (β6) with the exchanged alpha helices (α6).[42] The hinge is thought to be conformationally flexible and binding of ligands is proposed to alter the conformation of the KTN domain.[42] In each *Ec*KefC monomer, the transmembrane and the KTN domains are connected *via* a flexible linker of 19 amino acids, which is hypothesised to allow communication between the two domains.[42] Conformational change in the KTN domain will likely change the conformation of the transmembrane domain.[42]

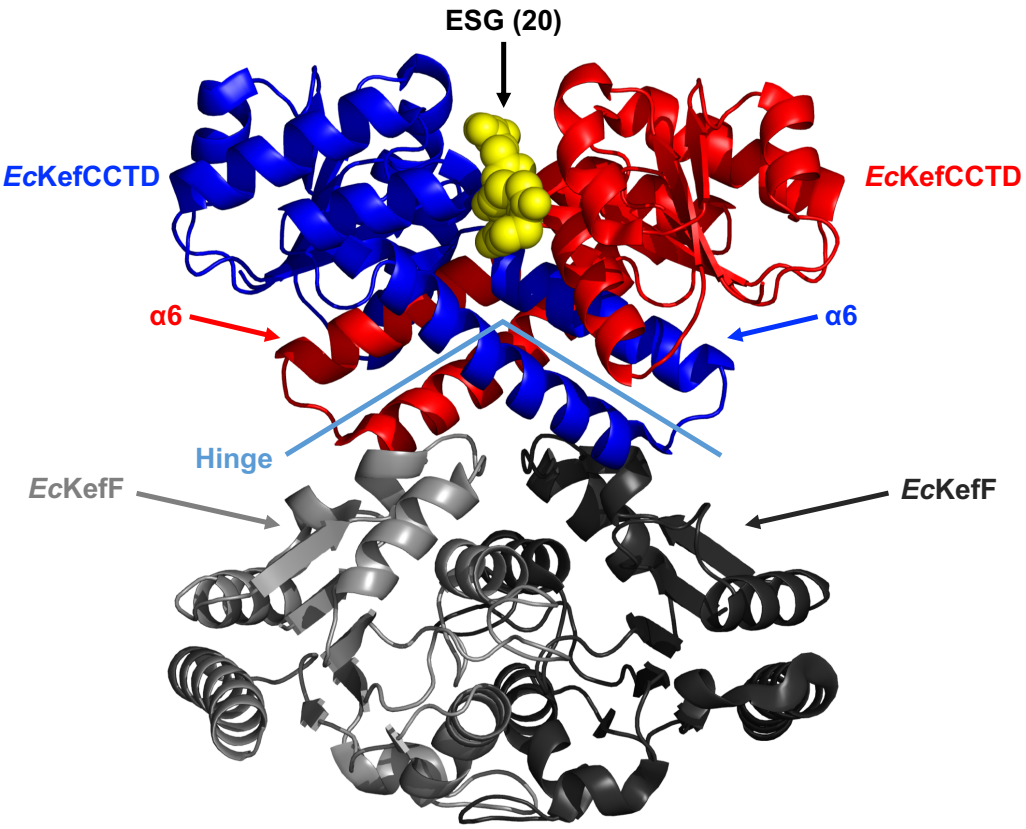


Figure 1.7. The ESG-bound X-ray crystal structure of *Ec*KefCCTD bound to the ancillary subunit KefF (PDB ID: 3L9X; 2.10 Å resolution).[41] Each *Ec*KefCCTD or KefF monomer is shown as a cartoon. The exchanged alpha helices (α6) in each *Ec*KefCCTD monomer and the hinge, formed by the loops linking the C-terminal strands (β6) with α6, are labelled. Both diastereoisomers of ESG (**20**) are shown. Image generated in PyMOL™ version 1.8.x.

The ancillary subunit KefF binds at the hinge of *Ec*KefCCTD, holding it at approximately 120° (Figure 1.7).[40] The crystallographic data suggest that KefF directly affects the hinge angle, potentially stabilising a particular conductive state of the K^+/H^+ antiporter in KefC.[40] KefF and *Ec*KefCCTD each dimerise, and so the overall *Ec*KefCCTD-KefF protein complex is a "dimer of dimers" (Figure 1.7).[40]

Both GSH (**14**) and ESG (**20**) bind at the interface between the two *Ec*KefCCTD domains (Figure 1.7).[41] There are two possible ligand binding sites per dimer but only one is occupied in the X-ray crystal structures.[41] The binding modes of GSH (**14**) and ESG (**20**) are predicted to be similar (Figure 1.8A and B, respectively). The peptide backbones of both ligands are predicted to form interactions with the same amino acid residues: Q412 and R416, from one *Ec*KefCCTD subunit, and R498, D499, V500, R516, and N551 from the other *Ec*KefCCTD subunit (Figure 1.8A and B). The Q412 side chains are predicted to form two interactions with the ligand.[41] R416 and R516 interact directly with the Gly carboxylate.[41] R416 and N551 interact with the Gly carboxylate *via* a structurally-conserved water molecule.[41] R416 interacts with N551, which also interacts with R516.[41] The V500 backbone N–H interacts with the carbonyl of the Cys-Gly amide bond.[41] The side chain of D499 interacts with the Glu-NH_2.[41] R498 forms two interactions with the Glu carboxylate.[41]

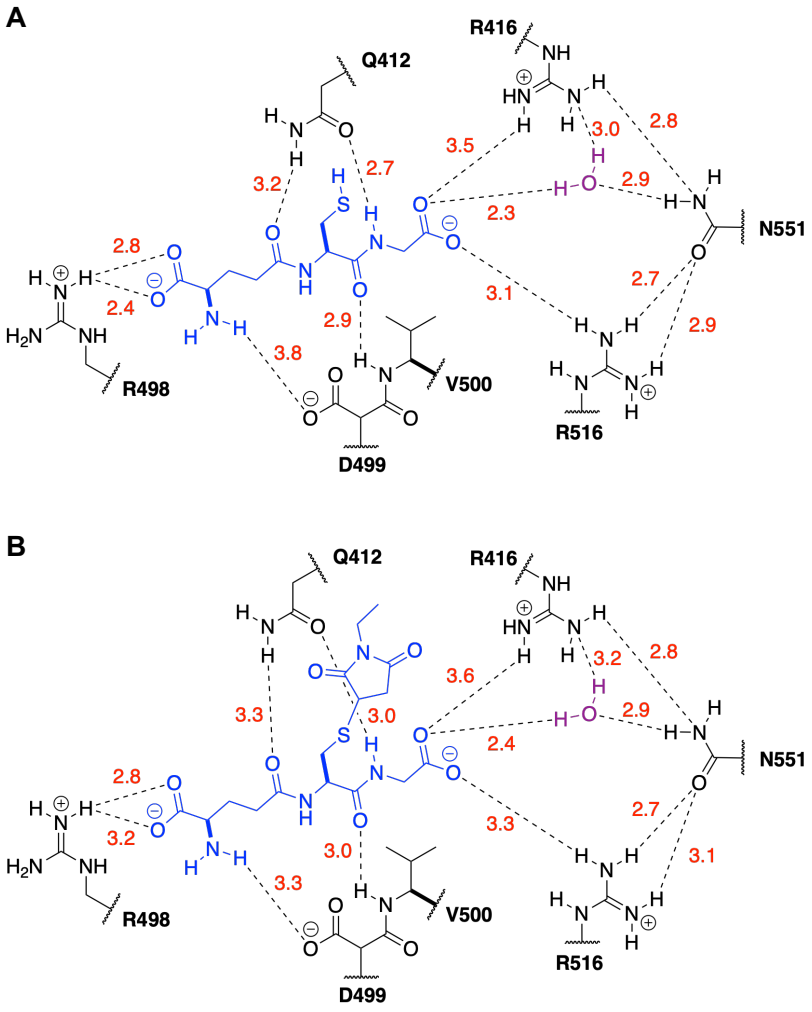


Figure 1.8. Schematic of the residues which bind GSH (14) and ESG (20) in *Ec*KefCCTD-KefF.[41] The residues (black) are as follows: Q412 and R416 (both from one subunit), and R498, D499, V500 (backbone N–H), R516, and N551 (all from the other subunit). The conserved water molecule is shown in purple. Hydrogen bonds are shown as dashed lines (black), with bond lengths (red) in angstroms (Å). A) Schematic for GSH (**14**, shown in blue). B) Equivalent schematic for ESG (**20**, shown in blue).

1.7 Proposed Mechanism of Inhibition/Activation

As the peptide backbones of GSH (**14**, Kef inhibitor) and ESG (**20**, Kef activator) are predicted to form identical interactions with *Ec*KefCCTD-KefF (Figure 1.8A and B, respectively), the difference in activity between these compounds must be caused by their S-substituents.[41] The GSH- and ESG-bound X-ray crystal structures of the protein were compared (Figure 1.9A and B, respectively).[41] The interactions between the different S-substituents and the protein were elucidated, and a mechanism of Kef inhibition and activation was proposed.[41] In the GSH-bound structure, the F441 capping residue, from α2 from one subunit, occupies a defined hydrophobic pocket formed by residues from α7 and α8 from the other subunit (Figure 1.9A). Comparison of this structure with the *apo* X-ray crystal structure shows that the thiol group of GSH (**14**) stabilises the protein by displacing water from the region adjacent to F441.[41] Analysis of the ESG-bound structure reveals a marked conformational change in the region of the F441 capping residue, which is displaced by 12 Å relative to its position in the GSH-bound structure (Figure 1.9B).[41] Superimposing ESG (**20**) into the GSH-bound structure shows that its *N*-ethylsuccinimido head group would sterically clash with the F441 capping residue. F441 and α2 are therefore displaced to a new position, where they interact solely with amino acid residues from their own KTN domain (Figure 1.9B).[41] A stark change also occurs in the other KTN subunit, as α8 is disordered due to the loss of interaction with F441, such that this α-helix could not be observed in the X-ray crystal structure (Figure 1.9B).[41]

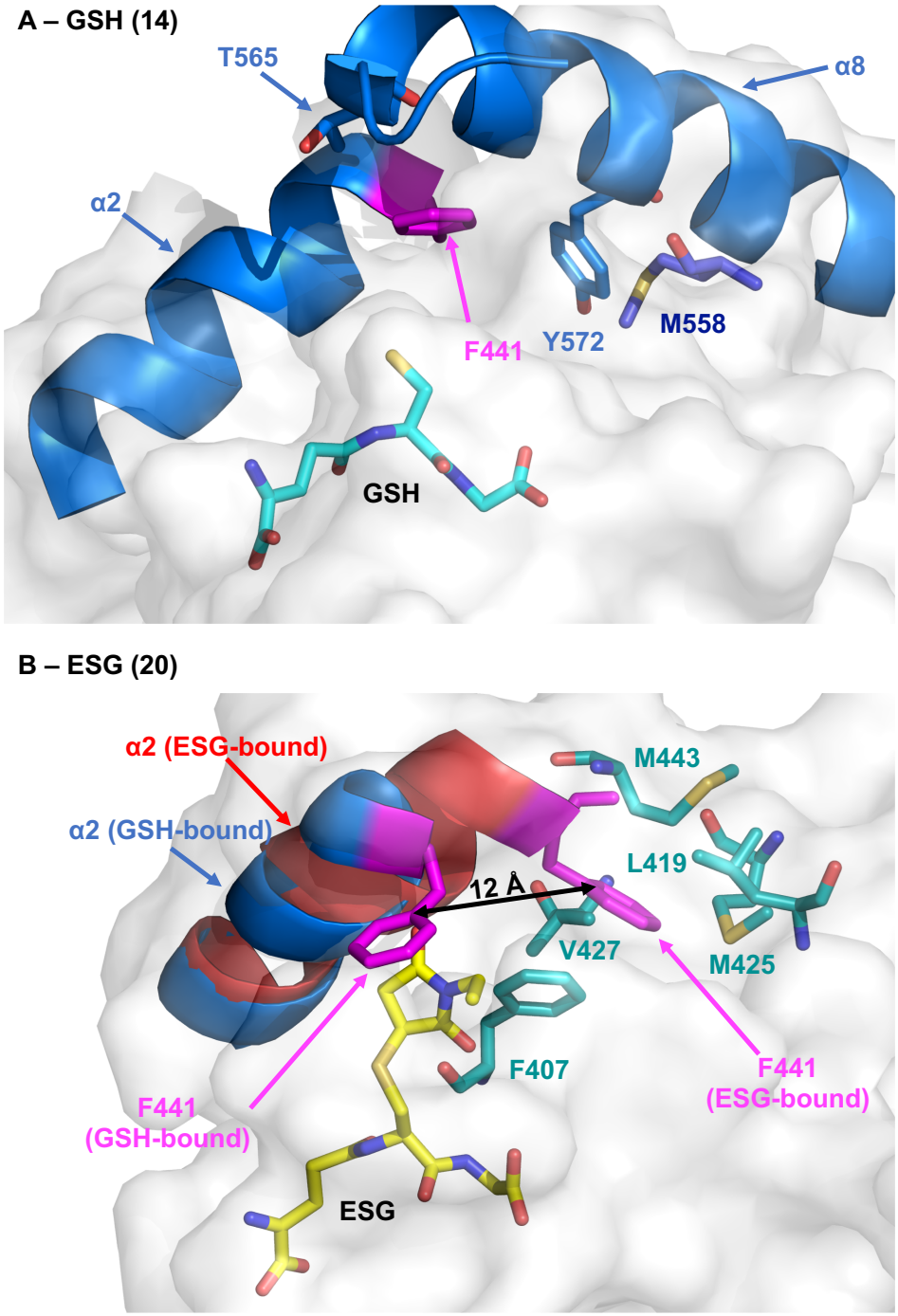


Figure 1.9. Proposed mechanism of Kef activation by GSH/GSX binding.[41] A) Based on PDB ID 3L9W. The thiol group (yellow) of GSH (**14**) stabilises the binding of F441 (pink) from α2 (blue) from one KTN domain to a hydrophobic pocket, formed by four residues from the other subunit: M558 (from α7, dark blue), T565, A569, and Y572 (all from α8, blue). B) Based on PDB ID 3L9X. In ESG (**20**), the *N*-ethylsuccinimido head group (both diastereoisomers of ESG (**20**) are shown) displaces F441 to a new position 12 Å away from its position in A). The displaced F441 now interacts with other hydrophobic amino acid residues from its own KTN domain (blue-green): F407, L419, M425, V427, and M443. The disruption of the previous protein-protein interaction also leads to the displacement of α8, which is no longer resolved crystallographically. Images generated in PyMOL™ version 1.8.x.

Based on these X-ray crystallographic data, the proposed model for Kef activation is that the capping residue F441 acts as a single amino acid switch.[41] GSH (**14**) binds to Kef, but does not displace F441, and stabilises the conformation of the protein.[41] In contrast, the binding of a GSX with both a rigid and bulky S-substituent, such as ESG (**20**), displaces the F441 capping residue, leading to K^+ efflux and concomitant H^+ influx (Kef activation).[41] Displacement of F441 in the ligand binding domain is hypothesised to trigger a conformational change in the transmembrane domain, due to putative communication between the two domains.

It is important to note that the paper published by *Roosild et al.* was subsequently corrected.[43] In 3L9X, the electron density of the *N*-ethylsuccinimido head group in ESG (**20**) is very weak and so its location is unclear.[43] In 3L9W, there is additional electron density adjacent to the sulfur atom, which could correspond to the succinimido ring of ESG (**20**).[43] It is therefore possible that the ligands GSH (**14**) and ESG (**20**) were modelled into the incorrect X-ray crystal structures, which would invalidate our proposed mechanism of Kef activation.

Overall, the mechanism of Kef inhibition and activation is likely to be more complex than the proposed single amino acid switch model at F441.[41] Nevertheless, the X-ray crystallographic data indicate that ligand binding causes large changes in protein conformation, even though the identity of the ligand in each structure can be disputed.[43] The hypothesis that the opposing activities of GSH (**14**) and ESG (**20**) originate from the interaction of the different S-substituents with Kef is still conceivable. However, further work is required to elucidate the full mechanism of activation.

1.8 Structural Requirements for Kef Ligands

Firstly, a major requirement for GSH-derived Kef ligands is the stereochemistry at the Cys-α stereogenic centre. Only L-GSH and L-GSX bind to Kef, while the corresponding D-enantiomers do not bind to the protein.[44]

Secondly, the S-substituent in GSH-derived ligands is also known to affect ligand activity. To explore the structure-activity relationships (SAR) of GSH-derived ligands in the S-substituent region, Healy *et al.* attempted to determine the affinities of GSH (**14**) and a range of GSX for *Sd*KefQCTDH$_6$, the soluble KTN domain of *Sd*Kef.[31] Q is an abbreviation for the Q-linker, a solubilising linker which is located between the transmembrane and KTN domains in full-length *Sd*Kef, comprising amino acids 391–408. The biophysical assays employed by Healy et *al.* were isothermal calorimetry (ITC), competition fluorescence, and differential scanning fluorimetry (DSF, described in Chapter 4; section 4.4.1.2).[31]

The inhibitor GSH (**14**) binds *Sd*KefQCTDH$_6$ with an affinity of K_d = 900 ± 200 μM (fluorescence emission spectra), while the activator ESG (**20**) binds the same protein with affinities of K_d = 12 ± 3 μM (fluorescence emission spectra) and 23 ± 4 μM (ITC).[31] The higher affinity of ESG (**20**) relative to GSH (**14**) was hypothesised to be caused by the *N*-ethylsuccinimido head group, doing one or more of the following: forming hydrophobic interactions with *Sd*KefQCTDH$_6$; causing a conformational change to *Sd*KefQCTDH$_6$, resulting in enthalpic gains in the Gibbs free energy of binding; causing a conformational change to *Sd*KefQCTDH$_6$, displacing structural water molecules and resulting in entropic gains in the Gibbs free energy of binding. To determine the effect of the S-substituent on affinity, Healy *et al.* synthesised a range of GSX and determined

their affinity for *Sd*KefQCTDH$_6$ qualitatively, and in some cases quantitatively (Figure 1.10).[31]

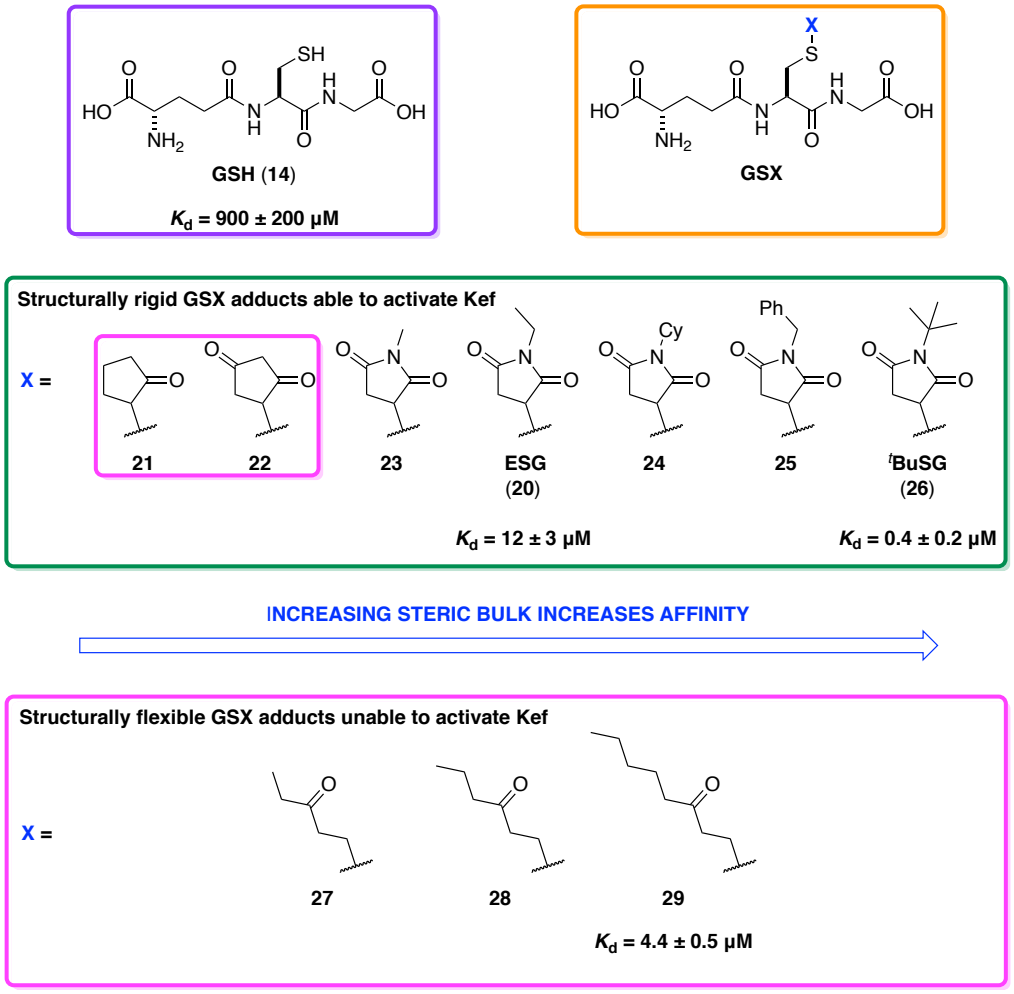


Figure 1.10. Summary of affinities of GSH (14) and a range of GSX for *Sd*KefQCTDH$_6$.[31]

The trends observed in the affinity data are summarised in Figure 1.10. For the ligands with rigid S-substituents (**20** and **23**–**26**), increasing the steric bulk of the S-substituent was shown to increase the affinity for *Sd*KefQCTDH$_6$.[31] *N*-*tert*-Butylsuccinimido-*S*-glutathione (^{t}BuSG, **26**), an analogue of ESG (**20**) with a *tert*-butyl group in place of the ethyl group in ESG (**20**), had the highest affinity of all the ligands tested (K_d = 0.4 ± 0.2 µM, fluorescence emission spectra).[31] A similar trend in affinity was observed for the GSX with flexible S-substituents (**27**–**29**), where increasing the length

of the alkyl chain resulted in an increase in affinity for *Sd*KefQCTDH$_6$, with compound **29** having an affinity of K_d = 4.4 ± 0.5 μM (fluorescence emission spectra).[31]

To determine whether the GSX (Figure 1.10) activated full-length *Sd*Kef, the corresponding enones were tested in *E. coli* supplemented with GSH (**14**) in K^+ efflux assays.[31] The GSX with structurally rigid S-substituents (**20** and **23**–**26**) activate *Sd*Kef, while the analogues with more flexible substituents (**21** and **22**) do not.[31] As compounds **21** and **22** lack the substituted nitrogen atom, this demonstrates that steric bulk at this position is required for activation of Kef.[31] Notably, the GSX with aliphatic S-substituents (**27**–**29**) do not activate Kef, despite possessing similar affinity to the more structurally rigid compounds (**20** and **23**–**26**).[31] These data indicate that the structure of the S-substituent directly affects whether a GSX will activate Kef. S-Substituents must be both sterically bulky and structurally rigid to activate Kef.[31]

1.9 Previous Work on Kef as a Novel Antibiotic Target

Previous work in the Conway Group sought to investigate Kef as a novel antibiotic target in pathogenic bacteria, specifically the ESKAPE pathogens.[44,45] The aim was to develop small molecules, which could bind and activate Kef. Binding to *Sd*KefQCTDH$_6$ was assessed using biophysical assays, including ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) NMR.[44,46] Activation of Kef was determined using the Kirby-Bauer disc diffusion assay.[44,45] In the disc assay, an agar plate is streaked with bacteria, to which paper discs containing the compound of interest are added. After incubation, any zones of exclusion around a disc indicate that the compound kills or inhibits the growth of bacteria (Figure 1.11).

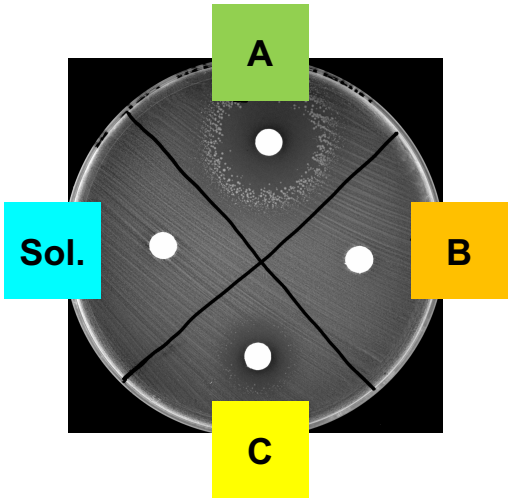


Figure 1.11. Example plate from the Kirby-Bauer disc diffusion assay. Image from Dr Anthony Chan (unpublished). Sol. = solvent, which is tested as a negative control. A–C = Test compounds. In this example, compounds A and C both inhibit/kill the bacteria. The effect seen for A is strong (large zone of exclusion), while the effect for C is weak (small zone of exclusion).

Complementary to the SAR described in 1.8, Dr Samuel Grayer synthesised a range of GSX analogues in which the peptide backbone of ^{t}BuSG (**26**) was truncated.[44] This approach revealed dipeptide OXFKEF04 (**30**, Figure 1.12), which binds *Sd*KefQCTDH$_6$ (K_d = 1.8 ± 0.2 µM, ^{1}H CPMG NMR), but is not cell permeant in the disc assay.[44] Methyl ester analogue OXFKEF12 (**31**, Figure 1.12) was synthesised, which activates *E. coli* MJF335 expressing full-length *Sd*Kef in the disc assay.[44] Alexander Axer synthesised analogues of OXFKEF12 (**31**) with different S-substituents, OXFKEF13 (**32**) and OXFKEF14 (**33**), and D-enantiomers OXFKEF15 (**34**) and OXFKEF17 (**35**), as negative controls (Figure 1.12).[45] These were tested in *E. coli* MJF335 expressing full-length *Sd*Kef, *Ec*KefFC, and *Ab*Kef (from *A. baumannii*).[45] The antibacterial effect of OXFKEF12–14 (**31**–**33**) was dependent on the S-substituent (Table 1.4).[45] These results were confirmed by two negative control experiments: all compounds were non-toxic to the MJF335 *E. coli* strain and D-methyl esters OXFKEF15 (**34**) and OXFKEF17 (**35**) did not cause growth inhibition/death for any of the *E. coli* transformants (Table 1.4).[45]

These data suggest that OXFKEF12–14 (**31**–**33**) cause growth inhibition/death due to activation of Kef and not through off-target effects or compound toxicity.[45]

Figure 1.12. Structures of dipeptide compounds.

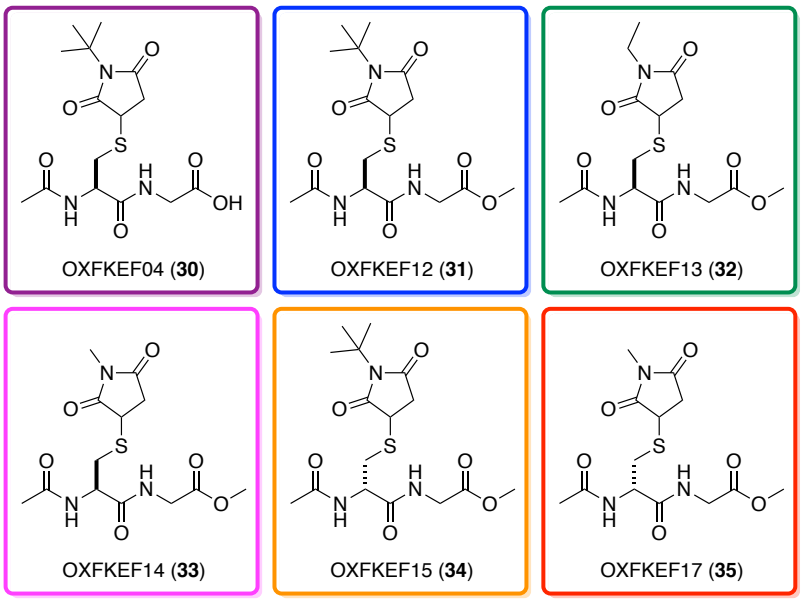
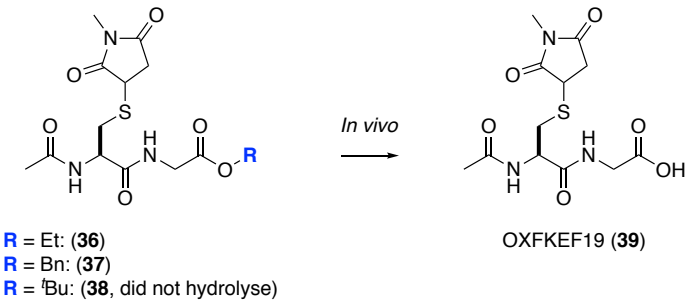


Table 1.4. Summary of Kirby-Bauer disc diffusion assay data for dipeptide compounds.[44,45] Compounds were tested in *E. coli* MJF335 expressing each of the full-length Kef systems, as indicated. Compounds were tested at 20 mg mL^{-1} in H_2O. The solvent (H_2O) and D-methyl esters (OXFKEF15 (**34**) and OXFKEF17 (**35**)) were used as negative controls. ✓ = Compound activates the *E. coli* transformant. ✕ = Compound does not activate the *E. coli* transformant.

Compound	Kef System			
	***Sd*Kef**	***Ec*KefFC**	***Ab*Kef**	**MJF335**
H_2O	✕	✕	✕	✕
OXFKEF12 (**31**)	✓	✕	✕	✕
OXFKEF13 (**32**)	✓	✓	✕	✕
OXFKEF14 (**33**)	✓	✓	✓	✕
OXFKEF15 (**34**)	✕	✕	✕	✕
OXFKEF17 (**35**)	✕	✕	✕	✕

As OXFKEF12 (**31**) does not to bind to *Sd*KefQCTDH$_6$,[44] it was hypothesised that L-methyl esters OXFKEF12–14 (**31**–**33**) act as prodrugs, which are hydrolysed by intracellular esterases to the corresponding carboxylic acids, which bind and activate Kef.[44,45] Jessie Tucker synthesised esters (**36**–**38**) of the *N*-methyl carboxylic acid OXFKEF19 (**39**) (Scheme 1.2).[47] The ethyl and benzyl esters, compounds **36** and **37** respectively, were hydrolysed to OXFKEF19 (**39**) by esterases isolated from *E. coli* MJF335 (Scheme 1.2).[47] However, in the same assay, the *tert*-butyl ester **38** did not hydrolyse (Scheme 1.2).[47] In the disc assay, compounds **36** and **37** activated *E. coli* MJF335 expressing *Sd*Kef, *Ec*KefFC, and *Ab*Kef, whereas compound **38** was inactive.[47] These results suggest that there is at least one esterase in *E. coli* MJF335 which can hydrolyse esters **36** and **37** to the active compound OXFKEF19 (**39**).[47]



Scheme 1.2. Summary of compound hydrolysis experiments using bacterial esterases.[47]

1.10 Summary

Antibiotic resistance is a global problem, and multi- and pan-resistant Gram-negative bacteria are of particular concern. To treat these resistant bacteria, novel antibiotic targets must be identified. Regulation of K^+ concentration is vital for maintaining turgor pressure and cytoplasmic pH. Potassium (K^+) efflux systems (Kef) are responsible for K^+ efflux, which also protects the bacteria from toxic electrophiles. Kef systems are inhibited by glutathione (GSH, **14**) and activated by glutathione S-conjugates (GSX), formed by the reaction of GSH (**14**) with electrophiles. Kef systems are present in many species of Gram-negative bacteria and possess no homologues in humans. Kef proteins are therefore potential novel antibiotic targets: inhibitors could be used to shut down this protective system, or activators could remove K^+, disrupting cell homeostasis.

Kef proteins are homodimeric and each monomer contains a transmembrane domain and a KTN domain, connected *via* a flexible linker. In the X-ray crystal structures of the KTN domain from *Ec*KefC (*Ec*KefCCTD) co-crystallised with either GSH (**14**) or ESG (**20**), both ligands bind at the interface of KTN domains. This binding is hypothesised to induce conformation changes in the KTN domain, leading to either inhibition or activation. Biophysical assays with *Sd*KefQCTDH_6 and K^+ efflux assays in *E. coli* revealed that the S-substituent must be both bulky and rigid to bind and activate Kef. SAR studies on ^{t}BuSG (**26**) were used to develop cell permeant dipeptide probes (L-methyl esters OXFKEF12–14, **31**–**33**), which cause growth inhibition/death of *E. coli* expressing various Kef systems, including *Ab*Kef (from *A. baumannii*). These data demonstrate the potential of Kef as a novel target for antibacterial drugs.

1.11 Aims

There are two aims to this study. Firstly, we seek to elucidate the full mechanism of action of Kef, using a fluorinated protein and ^{19}F NMR spectroscopy. The second aim is to investigate Kef as an antibiotic target in pathogenic bacteria. Novel Kef systems will be revealed (*Yp*KefB, *Bp*Kef1, and *Bp*Kef2), and dipeptide tool compounds will be synthesised and tested in *E. coli* expressing these Kef systems using the disc assay, to identify compounds which activate them. Western blotting will be used to compare the expression level of *Yp*KefBH_6 in two different *E. coli* cell strains, to elucidate the disc assay data. To improve the stability of the existing compounds, novel dipeptides will be designed, synthesised, and tested using biophysical assays and the disc assay.

Chapter 2: Elucidation of the Kef Mechanism of Action Using a Fluorinated Kef Protein and ^{19}F NMR Spectroscopy

2.1 Introduction and Aims

The aim of the work described in this chapter was to further elucidate the mechanism of Kef activation. The proposed model for Kef activation in *Ec*KefCCTD is that the F441 capping residue acts as a single amino acid switch (Figure 2.1).[41,43] The F448 residue in *Sd*KefQCTDH$_6$ is hypothesised to have the same function.[31] This model was proposed based on biophysical studies on *Sd*KefQCTDH$_6$[31] and X-ray crystal structures of *Ec*KefCCTD obtained in the presence and absence of GSH/GSX ligands.[41,43] While these X-ray crystal "snapshots" were essential for the generation of hypotheses, these lack the dynamic aspect required to provide more direct evidence for the proposed mechanism. In addition, neither the biophysical nor the X-ray crystallographic data provide any information about ligand function, whether a ligand will be an activator or inhibitor of Kef.

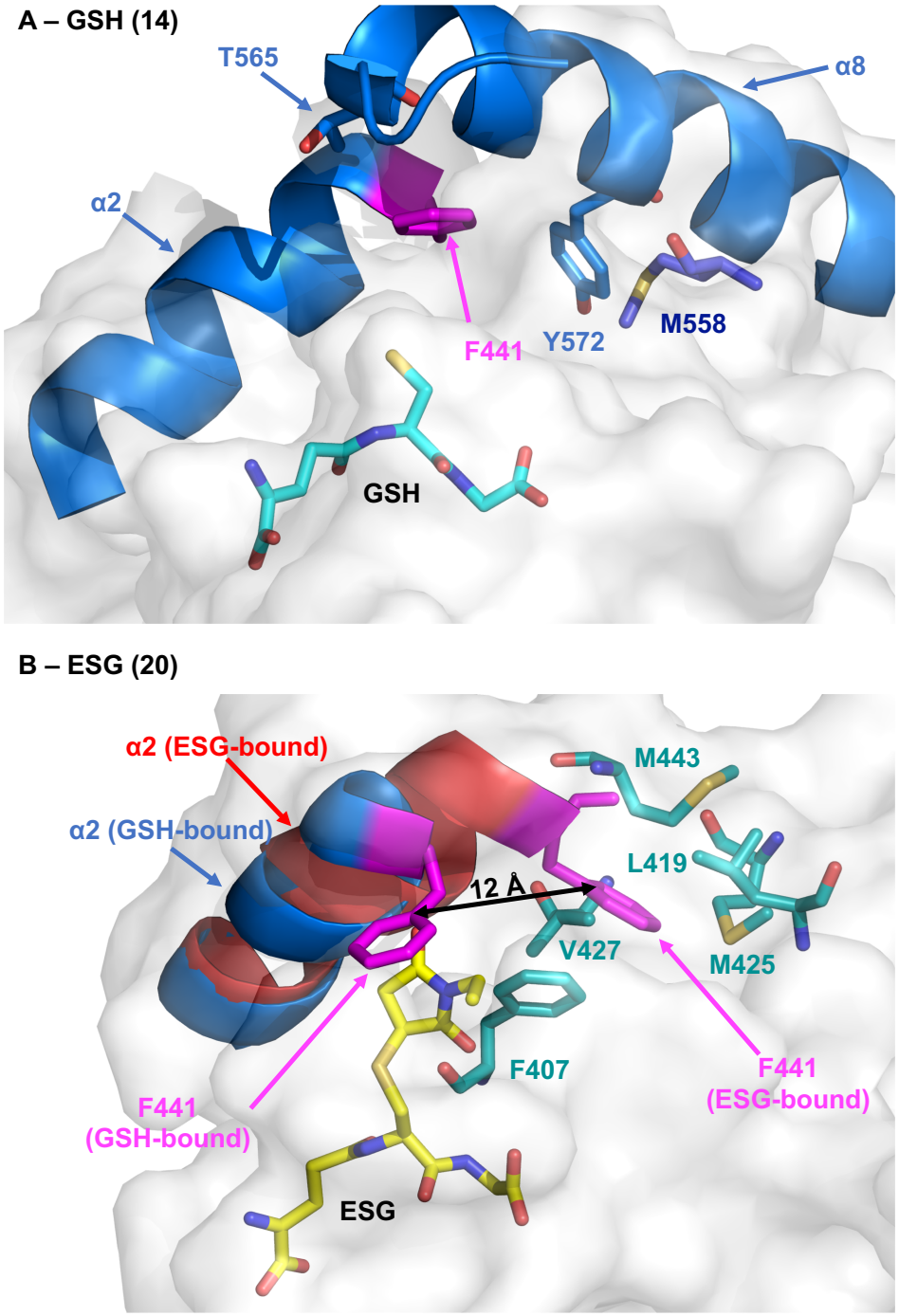


Figure 2.1. Proposed mechanism of Kef activation by GSH/GSX binding.[41] A) Based on PDB ID 3L9W. The thiol group (yellow) of GSH (**14**) stabilises the binding of F441 (pink) from α2 (blue) from one KTN domain to a hydrophobic pocket, formed by four residues from the other subunit: M558 (from α7, dark blue), T565, A569, and Y572 (all from α8, blue). B) Based on PDB ID 3L9X. In ESG (**20**), the *N*-ethylsuccinimido head group (both diastereoisomers of ESG (**20**) are shown) displaces F441 to a new position 12 Å away from its position in A). The displaced F441 now interacts with other hydrophobic amino acid residues from its own KTN domain (blue-green): F407, L419, M425, V427, and M443. The disruption of the previous protein-protein interaction also leads to the displacement of α8, which is no longer resolved crystallographically. Images generated in PyMOL™ version 1.8.x.

To investigate the molecular mechanism of Kef activation, we wanted to incorporate a fluorinated amino acid into the Kef protein and then attempt to analyse the protein conformational change using ^{19}F NMR spectroscopy. The first aim was to replace the F448 capping residue in *Sd*KefQCTDH$_6$ with the non-standard, fluorinated amino acid, *p*-trifluoromethyl-L-phenylalanine (*p*tfmF), to give the fluorinated Kef protein *Sd*KefQCTDH$_6$(F448*p*tfmF). It was hypothesised that the fluorinated protein would behave in a similar manner to the wild-type protein, both in terms of protein folding and ligand binding. The second aim was to study the resulting fluorinated protein using ^{19}F NMR spectroscopy in the presence of known activators and inhibitors, to monitor how ligand function affects the protein dynamics in the region of this F448*p*tfmF capping residue. In the ^{19}F NMR assay, an inhibitor was expected to not affect the chemical shift of the F448*p*tfmF resonance in the protein. In contrast, for an activator, the chemical shift of the F448*p*tfmF resonance should change markedly, indicating displacement of the capping residue upon ligand binding. Using these NMR data, the aim was to validate the proposed model of Kef activation.[41,43]

2.2 Synthesis of a Range of GSX Ligands

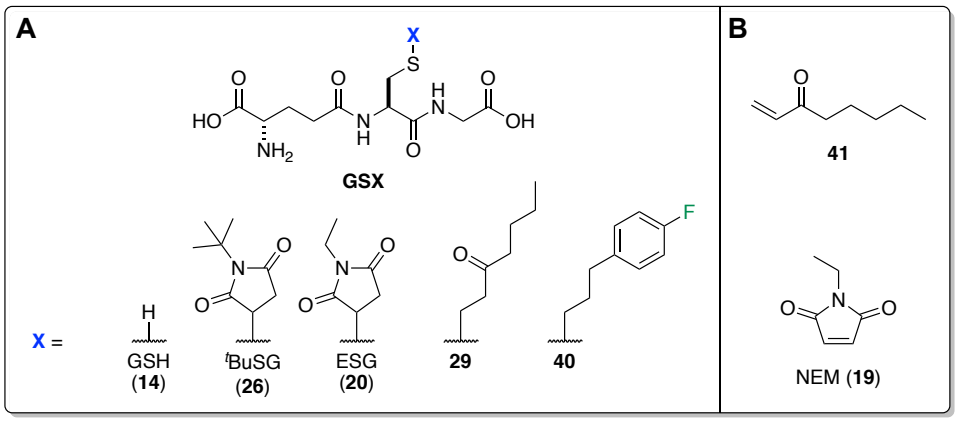
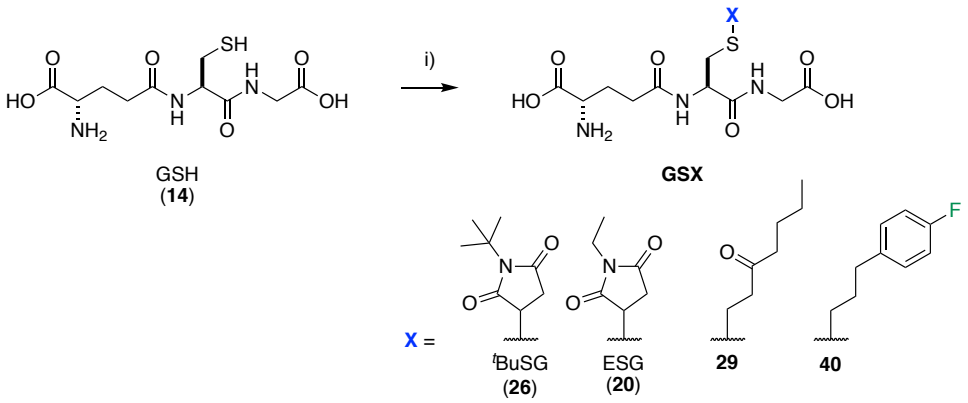


Figure 2.2. A) Structures of GSH (14) and GSX ligands: ^{t}BuSG (26), ESG (20), 29, and 40. B) Structures of electrophiles 1-octen-3-one (41) and NEM (19).

Five ligands were tested in the ^{19}F NMR assay with the fluorinated protein: GSH (**14**), ^{t}BuSG (**26**), ESG (**20**), compound **29**, and compound **40** (Figure 2.2A). GSH (**14**) was tested because it is the endogenous inhibitor of *Ec*KefC and *Ec*KefB.[32,34] GSH (**14**) binds *Sd*KefQCTDH$_6$ with an affinity of K_d = 900 ± 200 µM (fluorescence competition assay).[31] ^{t}BuSG (**26**) and ESG (**20**) were selected because they are known to activate *Sd*Kef.[31] Both ligands bind *Sd*KefQCTDH$_6$ with affinities of K_d = 0.4 ± 0.2 µM and 12 ± 3 µM, respectively.[31] Compound **29** was chosen because it was previously shown to bind *Sd*KefQCTDH$_6$ with an affinity similar to ESG (K_d (**29**) = 4.4 ± 0.5 µM) but its corresponding electrophile, 1-octen-3-one (**41**, Figure 2.2B), elicited minimal K^+ efflux in the *in vivo* assay,[31] suggesting it might act as a high-affinity inhibitor of Kef. In the *in vivo* assay, the rate constant for K^+ efflux from *E. coli* MJF335 expressing *Sd*Kef for 1-octen-3-one (**41**) was far lower than the rate constant for NEM (**19**, Figure 2.2B).[31] The low propensity for K^+ efflux observed for 1-octen-3-one (**41**) could have been caused by its high lipophilicity, resulting in poor cell permeability and therefore minimal formation of the GSX **29** in the cell.[31] It was hoped that the ^{19}F NMR assay

using the fluorinated protein would elucidate whether compound **29** was an activator or inhibitor of Kef. Compound **40** was selected because it binds *Sd*KefQCTDH_6, as shown using the fluorescence competition assay and differential scanning fluorimetry (DSF),[48] and also contains a fluorine atom, potentially enabling study of both the protein and ligand chemical environments at the same time. As neither the fluorescence competition assay nor DSF provide any information about Kef activation or inhibition, it was unknown whether compound **40** was an activator or inhibitor of Kef, and so we planned to investigate this using the ^{19}F NMR assay.

^{t}BuSG (**26**), ESG (**20**), and compound **29** were prepared *via* conjugate addition reactions with GSH (**14**) and their corresponding enones (Scheme 2.1).[31] Compound **40** was synthesised *via* a thiol-ene reaction between GSH (**14**) and 1-allyl-4-fluorobenzene (Scheme 2.1), as reported in a prior study by Healy *et al.*[48] The yield obtained for compound **40** (5%) was much lower than in the literature[48] (90%), as it was necessary to purify the compound several times to obtain it in the required purity (>95%, by reverse phase (RP) analytical high performance liquid chromatography (HPLC)) for biological testing.



Scheme 2.1. Synthesis of GSX. *Reagents and conditions* i), for ^{t}BuSG (**26**): *N*-*tert*-Butylmaleimide (**42**), NaOH, H_2O, rt, 30 min, 33%; For ESG (**20**): NEM (**19**), H_2O, NaOH, rt, 5 h, 51%; For **29**: 1-Octen-3-one, **41**, (50% *w/w* in 1-octen-3-ol), NaOH, H_2O:MeOH 1:1, rt, 42 h, 23%; For **40**: 1-Allyl-4-fluorobenzene, 2,2-dimethoxy-2-phenylacetophenone (DPAP), tetrahydrofuran (THF):H_2O 1:2, rt, ultraviolet (UV) light (λ = 365 nm), 5 h, 5%.

2.3 Determination of Ligand Function

Having synthesised all of the ligands required, it was then possible to begin investigating their effect on *Sd*KefQCTDH$_6$(F448*p*tfmF) using ^{19}F NMR spectroscopy. All experiments described in the following section were carried out by Dr Anthony Chan, in the Conway Group, University of Oxford.

To produce *Sd*KefQCTDH$_6$(F448*p*tfmF), *p*-trifluoromethyl-L-phenylalanine (*p*tfmF), was site-specifically incorporated into *Sd*KefQCTDH$_6$ using amber stop codon suppression technology.[49] In the plasmid for *Sd*KefQCTDH$_6$ (pTrcSdKefQCTDH$_6$), the codon for F448 was mutated to the amber nonsense codon (ATG) using site-directed mutagenesis to give pTrcSdKefQCTDH$_6$(F448Am). An engineered amber suppressor tRNA$_{CUA}$ and aminoacyl-tRNA synthetase pair and pTrcSdKefQCTDH$_6$(F448Am) were transformed into *E. coli*, followed by addition of the *p*tfmF amino acid, to produce *Sd*KefQCTDH$_6$(F448*p*tfmF).

Having obtained *Sd*KefQCTDH$_6$(F448*p*tfmF), it was important to verify that its secondary structure was similar to *Sd*KefQCTDH$_6$. Circular dichroism (CD) spectroscopy of both proteins showed that the characteristic alpha helical structure of *Sd*KefQCTDH$_6$ was preserved in *Sd*KefQCTDH$_6$(F448*p*tfmF) (Figure 2.3).

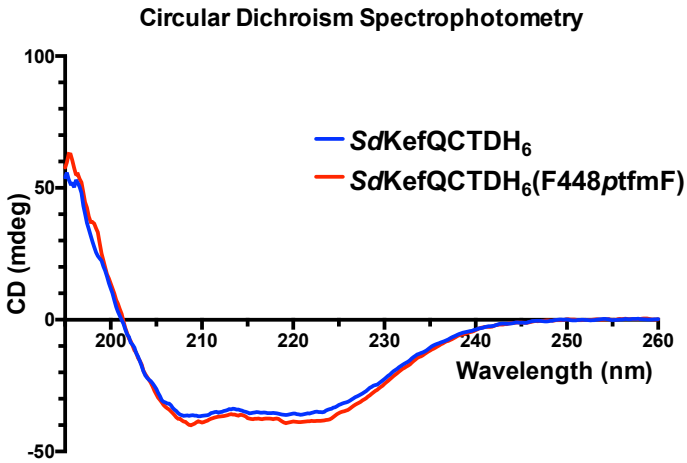


Figure 2.3. Characterisation of *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) using CD spectroscopy. Data from Dr Anthony Chan (unpublished).

Having confirmed that *Sd*KefQCTDH$_6$(F448*p*tfmF) folded in a similar manner to *Sd*KefQCTDH$_6$, biophysical assays were undertaken on both proteins to assess the binding of four Kef ligands: ^{t}BuSG (**26**), ESG (**20**), compound **29**, and GSH (**14**) (Figure 2.4). Given their relatively high affinities for Kef,[31] binding of ^{t}BuSG (**26**), ESG (**20**), and compound **29** to both *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) was investigated using ITC. All three ligands showed reduced affinities for *Sd*KefQCTDH$_6$(F448*p*tfmF) versus *Sd*KefQCTDH$_6$, with ESG (**20**) showing the largest difference, corresponding to approximately a 35-fold reduction in affinity (Table 2.1). These data are consistent with the increased sterics in the ligand binding site in *Sd*KefQCTDH$_6$(F448*p*tfmF), as an H atom in the *para*-position of F448 in *Sd*KefQCTDH$_6$ has been replaced with a sterically bulky -CF_3 group (A-value = 2.1 kcal/mol)[50–53] in the fluorinated protein.

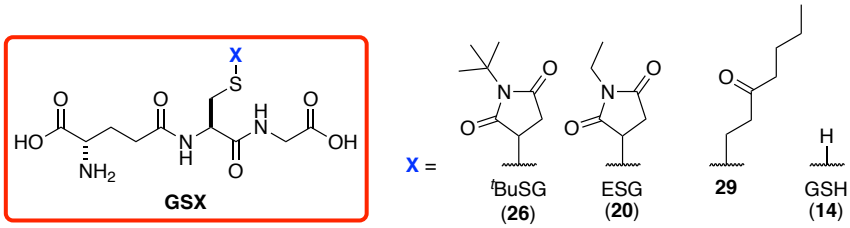
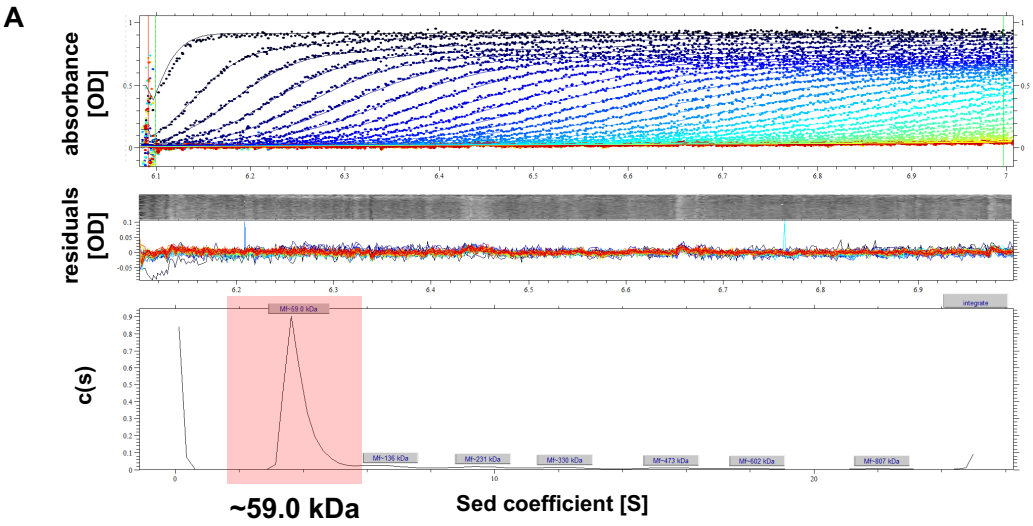


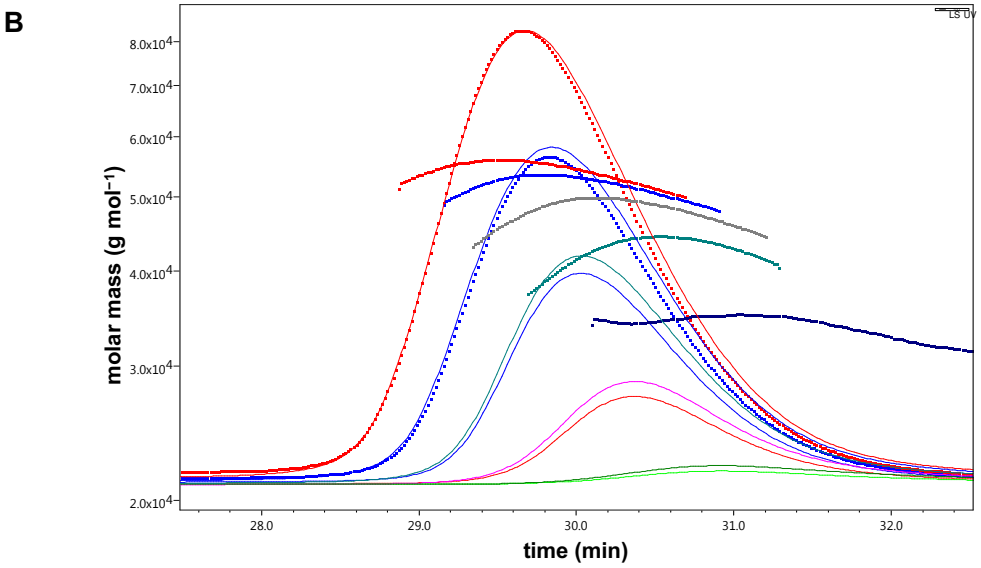
Figure 2.4. Structures of Kef ligands.

Table 2.1. Summary of ITC data for *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF). Data from Dr Anthony Chan (unpublished). Data shown are mean ± standard deviation (s.d.) from three independent ITC experiments (n = 3).

Ligand	**K_d (μM)** *ITC N Number*	
	***Sd*KefQCTDH$_6$**	***Sd*KefQCTDH$_6$(F448*p*tfmF)**
^{t}BuSG (**26**)	**5.9 ± 1.2** *0.72 ± 0.09*	**36.7 ± 12.0** *0.55 ± 0.15*
ESG (**20**)	**22.9 ± 3.5** *0.48 ± 0.09*	**803.7 ± 43.1** *0.12 ± 0.01*
29	**15.9 ± 2.5** *0.49 ± 0.15*	**75.2 ± 32.2** *1.04 ± 0.41*

The ITC N numbers of the ligands for the two different proteins, which demonstrate the protein to ligand ratio, are approximately 0.5 (Table 2.1). This is because *Sd*KefQCTDH$_6$ homodimerises in solution, as shown by previous analytical ultracentrifugation (AUC) experiments,[37] with two ligand binding sites per dimer. These ligand binding sites are at the dimer interface and only the dimerised protein can bind ligands. Both AUC and size exclusion chromatography-multi-angle laser scattering (SEC-MALS) experiments were performed on *Sd*KefQCTDH$_6$(F448*p*tfmF). The AUC data show that the fluorinated protein at the ITC assay concentration (23.4 µM) exists predominately as a homodimer (Figure 2.5A). The SEC-MALS data for *Sd*KefQCTDH$_6$(F448*p*tfmF) demonstrate that the protein tends to homodimerise at >150 µM (Figure 2.5B). The ITC N number of approximately 0.5 for *Sd*KefQCTDH$_6$(F448*p*tfmF), therefore, suggests that only half of the available ligand binding sites are occupied. As the ITC data were analysed using a single-site binding model, this corresponds to one GSX ligand per dimer.





*Sd*KefQCTDH$_6$(F448*p*tfmF)					
Concentration (μM)	25	50	100	150	200
Molecular weight (kDa)	34.1	42.9	48.3	52.0	54.3

Figure 2.5. AUC And SEC-MALS data for *Sd*KefQCTDH$_6$(F448*p*tfmF). Data from Dr Anthony Chan (unpublished). A) AUC Data for *Sd*KefQCTDH$_6$(F448*p*tfmF). Protein concentration = 23.4 μM, same as used for the ITC experiments. The major peak, which corresponds to a protein of approximately 59.0 kDa in size, is highlighted in red. These data suggest that *Sd*KefQCTDH$_6$(F448*p*tfmF) at 23.4 μM forms a homodimer. B) SEC-MALS Data for *Sd*KefQCTDH$_6$(F448*p*tfmF) at increasing concentrations (25–200 μM). These data show that the protein homodimerises at >150 μM.

As GSH (**14**) has lower affinity for *Sd*KefQCTDH$_6$ compared to the GSX (^{t}BuSG (**26**), ESG (**20**), and compound **29**), the binding of GSH (**14**) to *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) was investigated using waterLOGSY (water-ligand observed *via* gradient spectroscopy), which is a qualitative, NMR-based assay.[54] WaterLOGSY relies on the transfer of bulk water magnetisation to the free ligand *via* the protein-ligand complex.[54] Signals for non-binders have opposite sign to the protein and are often weaker than peaks for binders, which have the same sign as the protein.[54]

The water-suppressed ^{1}H NMR and waterLOGSY spectra for GSH (**14**) with both *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) are shown in Figure 2.6A and B, respectively. In each spectrum, a single ligand peak is highlighted in red and a single protein peak (used as a reference) is highlighted in blue. In the ^{1}H NMR spectra of the ligand (GSH, **14**) only, all of the peaks are positive but are inverted in the waterLOGSY spectra of GSH (**14**) alone (Figure 2.6A and B). In the waterLOGSY spectra of GSH (**14**) and each of two proteins, the ligand peaks are positive, as are the protein peaks (Figure 2.6A and B). These data indicate that GSH (**14**) binds to both proteins. To further validate these results, heat-denaturation experiments were also carried out, in which the protein-ligand samples were heated at 100 °C for 15 min to denature the protein in the samples. WaterLOGSY spectra were then acquired on these denatured samples. In the case of both proteins, these spectra resemble the waterLOGSY spectra for the ligand only, indicating that the results seen for ligand plus protein are due to ligand binding to the protein (Figure 2.6A and B). Overall, waterLOGSY confirmed the binding of GSH (**14**) to *Sd*KefQCTDH$_6$, as shown previously by Healy *et al.*,[31] and

allowed us to observe the binding of GSH (**14**) to *Sd*KefQCTDH$_6$(F448*p*tfmF) for the first time.

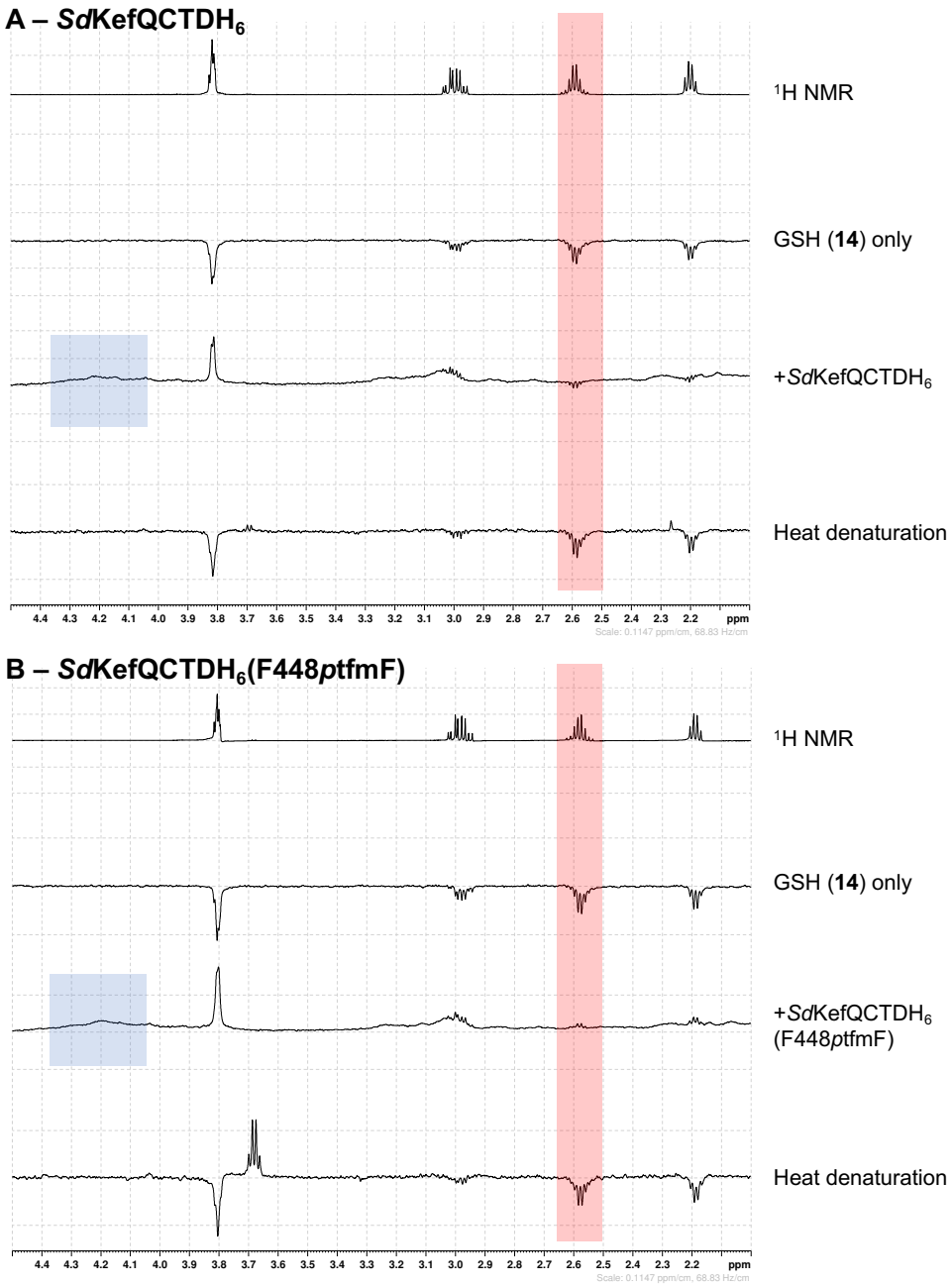


Figure 2.6. Detection of binding of GSH (14) to *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) using waterLOGSY. Data from Dr Anthony Chan (unpublished). A) Water-suppressed ^{1}H NMR and waterLOGSY spectra for GSH (**14**) with *Sd*KefQCTDH$_6$. The order of spectra (from top to bottom) is as follows: Water-suppressed ^{1}H NMR spectrum of GSH (**14**, 5 mM), waterLOGSY spectrum of GSH (**14**, 5 mM), waterLOGSY spectrum of GSH (**14**, 5 mM) with *Sd*KefQCTDH$_6$ (120 µM), and waterLOGSY spectrum of GSH (**14**, 5 mM) with *Sd*KefQCTDH$_6$ (120 µM) denatured at 100 °C for 15 min. In all spectra, a single ligand peak is highlighted in red, and, for the protein-ligand samples, a single protein peak is highlighted in blue. B) Equivalent schematic for *Sd*KefQCTDH$_6$(F448*p*tfmF).

The results for *Sd*KefQCTDH$_6$(F448*p*tfmF) and *Sd*KefQCTDH$_6$ in the biophysical assays are consistent with all four ligands (^{t}BuSG (**26**), ESG (**20**), compound **29**, and GSH (**14**)) binding *Sd*KefQCTDH$_6$(F448*p*tfmF) in an analogous way to *Sd*KefQCTDH$_6$. The fluorinated protein *Sd*KefQCTDH$_6$(F448*p*tfmF) was therefore used in the ^{19}F NMR assay to investigate the activation dynamics of Kef, using ligands GSH (**14**), ^{t}BuSG (**26**), ESG (**20**), **29**, and **40** (Figure 2.7).

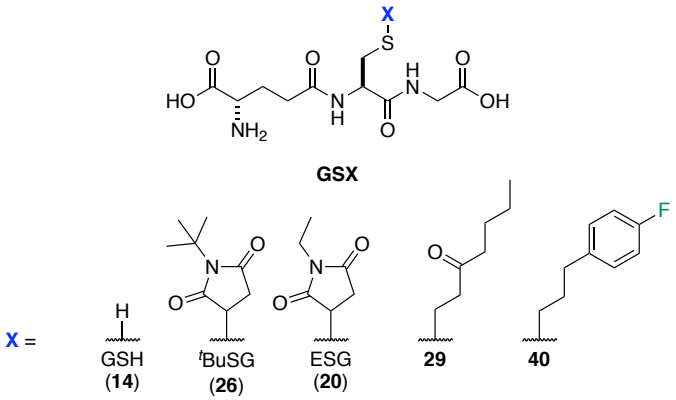


Figure 2.7. Ligands tested in the ^{19}F NMR assay.

In the ^{19}F NMR assay, using the fluorinated protein *Sd*KefQCTDH$_6$(F448*p*tfmF), we aimed to confirm that GSH (**14**) was an inhibitor of Kef, and that ^{t}BuSG (**26**) and ESG (**20**) were activators of Kef. We also wanted to determine whether compounds **29** and **40** were inhibitors or activators of Kef. The results obtained for GSH (**14**) are shown in Figure 2.8A. The results for compound **29** are shown in Figure 2.8B and Figure 2.10. The results for ESG (**20**) are shown in Figure 2.9A and Figure 2.10A. The results for ^{t}BuSG (**26**) are shown in Figure 2.9B and Figure 2.10B. The results for compound **40** are shown in Figure 2.11. In all cases, the spectrum shown in blue is for the protein only (*Sd*KefQCTDH$_6$(F448*p*tfmF), 23.4 μM).

The ^{19}F NMR spectrum for the protein and GSH (**14**) is consistent with Kef inhibition, as the ^{19}F chemical shift of the protein remains approximately constant in the presence of GSH (**14**) compared to the protein alone (Figure 2.8A). The data for compound **29** suggest that this ligand is also an inhibitor of Kef as the ^{19}F chemical shift of the protein is essentially unchanged when the ligand is present (Figure 2.8B). This result is consistent with the hypothesis that compound **29** binds Kef but subsequently its flexible S-substituent cannot displace the capping residue, and so the protein remains in an inactive, closed state.

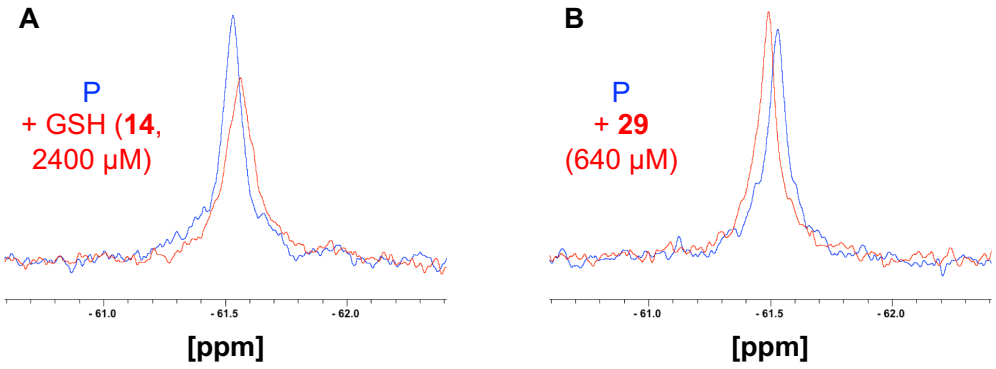


Figure 2.8. Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and either GSH (14) or compound 29. Data from Dr Anthony Chan (unpublished). P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF). A) Blue: P; Red: P + 2400 µM GSH (**14**). These data are consistent with Kef inhibition. B) Blue: P; Red: P + 640 µM **29**. These data are consistent with Kef inhibition.

The results for ESG (**20**) and ^{t}BuSG (**26**) are also consistent with the proposed model of Kef activation (Figure 2.9A and B, respectively). For both ligands, large changes in ^{19}F chemical shift of F448*p*tfmF are obtained when the ligand is present compared to the spectrum for the protein only. These data suggest that the capping residue is being moved to a new location upon ligand binding, consistent with the protein changing conformation from a closed, inactive, state to an open, active, one. The effect is also accompanied by peak broadening, which suggests that the capping residue is moving between several different positions in the open, active, conformation of the protein.

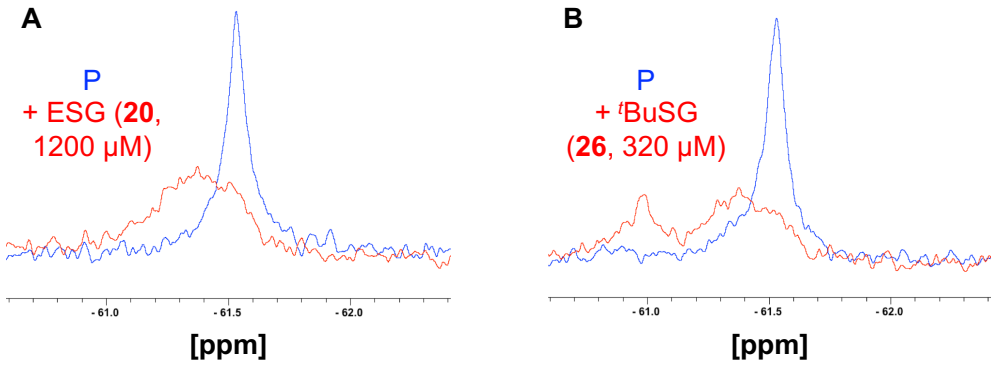


Figure 2.9. Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and either ESG (20) or ^{t}BuSG (26). Data from Dr Anthony Chan (unpublished). P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF). A) Blue: P; Red: P + 1200 µM ESG (**20**). These data are consistent with Kef activation. B) Blue: P; Red: P + 320 µM ^{t}BuSG (**26**). These data are consistent with Kef activation.

^{19}F NMR-Based competition experiments using compound **29** (inhibitor) and either ESG (**20**) or ^{t}BuSG (**26**) (both activators), Figure 2.10A and B, respectively, showed that both ligands are competing for the same ligand binding site. The ^{19}F NMR spectrum acquired when both ligands were present is an average of the two spectra acquired with each of the two ligands. These experiments also demonstrated the opposing activity of compound **29** compared to ESG (**20**) and ^{t}BuSG (**26**), providing further evidence for the inhibitory nature of compound **29**.

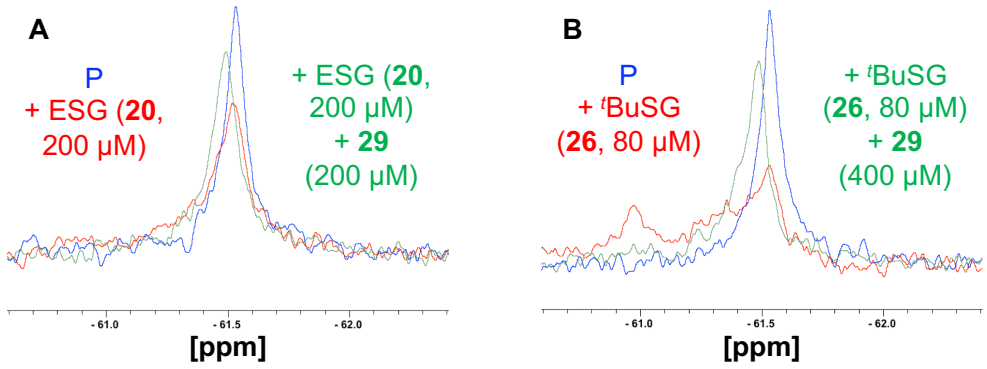


Figure 2.10. ^{19}F NMR Spectroscopy competition experiments with compound 29 and either ESG (20) or ^{t}BuSG (26). Data from Dr Anthony Chan (unpublished). P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF). A) Blue: P; Red: P + 200 µM ESG (**20**); Green: P + 200 µM ESG (**20**) + 200 µM **29**. These data are consistent with Kef inhibition by compound **29**. B) Blue: P; Red: P + 80 µM ^{t}BuSG (**26**); Green: P + 80 µM ^{t}BuSG (**26**) + 400 µM **29**. These data are consistent with Kef inhibition by compound **29**.

To determine whether compound **40** was an inhibitor or an activator of Kef, this ligand was tested in a competition experiment with *Sd*KefQCTDH$_6$(F448*p*tfmF) plus the activator ^{t}BuSG (**26**). The *Sd*KefQCTDH$_6$(F448*p*tfmF) and ^{t}BuSG (**26**) combination was denoted P2 (Figure 2.11). Compound **40** was titrated into P2 until the protein was saturated by **40** (Figure 2.11A and B). As the concentration of ligand **40** is increased, the chemical shift of the peak for its 4-F atom remains constant (Figure 2.11A). This implies that compound **40** is binding to the protein but not displacing the F448*p*tfmF capping residue. Addition of compound **40** to P2 simplifies the ^{19}F NMR spectrum of the protein, as the additional F448*p*tmfF peak, adjacent to the major one, thought to correspond to multiple ^{t}BuSG-bound, open conformations of the fluorinated protein, is gradually removed as the concentration of compound **40** is increased (Figure 2.11B). At saturation by compound **40**, only one major resonance remains, which is thought to correspond to the **40**-bound, closed conformation of the protein. The chemical shift of this major peak was compared to the chemical shifts of the resonances for the protein only and the protein with ligand **40** only, detected in independent experiments (Figure 2.11A and B). The chemical shifts of all three resonances were identical, which supports our hypothesis that the remaining peak in the competition experiment corresponds to the **40**-bound protein. These data suggest that **40** is an inhibitor of *Sd*KefQCTDH$_6$(F448*p*tfmF) as the chemical shift of the protein is unchanged in the presence of **40**, consistent with non-displacement of the capping residue upon ligand binding. The data from the competition experiment also suggest that compound **40** is a Kef inhibitor as it demonstrates opposing activity compared to the activator, ^{t}BuSG (**26**).

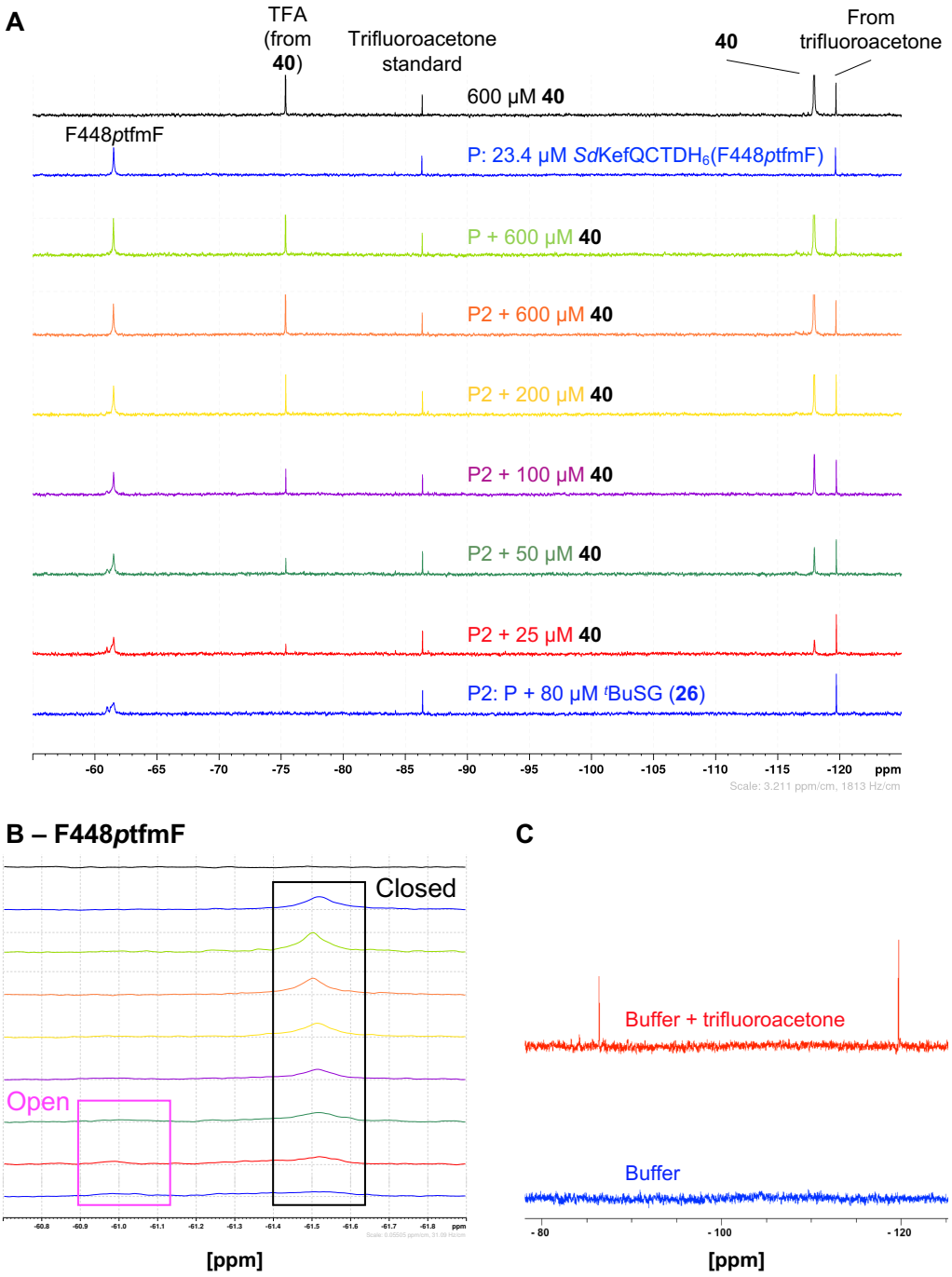


Figure 2.11. Elucidation of the function of compound 40 using *Sd*KefQCTDH$_6$(F448*p*tfmF) and ^{19}F NMR spectroscopy. Data from Dr Anthony Chan (unpublished). A) Titration of compound **40** into P2 (23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF) with 80 µM ^{t}BuSG (**26**)) and monitoring by ^{19}F NMR spectroscopy until saturation by compound **40** (600 µM). Spectra for compound **40** (600 µM), the protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)) only, and P + 600 µM **40** are shown for comparison. Trifluoroacetone is used as an internal ^{19}F standard. B) Overlay view of A), showing the peaks for *Sd*KefQCTDH$_6$(F448*p*tfmF) only. At 0–50 µM compound **40**, two protein peaks are visible, thought to correspond to the open and closed forms of the protein. At >50 µM compound **40**, only one peak remains, thought to correspond to the closed form of the protein. C) ^{19}F NMR Spectrum of NMR buffer + trifluoroacetone standard shows 2 peaks whereas the ^{19}F NMR spectrum of the buffer alone shows none. This explains the presence of the additional peak at ~−120 ppm in A).

2.4 Chapter 2 Conclusions

The data obtained in this project are consistent with the single amino acid "switch" model for Kef activation proposed previously.[31,37,41] The ^{19}F NMR data indicate that Kef activators, such as ESG (**20**) and ^{t}BuSG (**26**), are able to displace the F448*p*tfmF residue of *Sd*KefQCTDH$_6$(F448*p*tfmF), whereas inhibitors, including GSH (**14**), do not displace the capping residue. The inhibitory nature of known ligand **29** has been demonstrated using ^{19}F NMR spectroscopy, as although previous K^+ efflux data suggested that it was an inhibitor, the validity of the K^+ efflux assay was under doubt, due to the high lipophilicity of the corresponding electrophile.[31] The activity of compound **40** has been revealed for the first time, as data obtained using *Sd*KefQCTDH$_6$(F448*p*tfmF) and the ^{19}F NMR assay indicate that compound **40** is an inhibitor of Kef. This result demonstrates the utility of this ^{19}F NMR-based assay in predicting the activity of novel Kef ligands, vital in the design of novel antibacterials which act on this drug target.

Chapter 3: Identification of Kef Systems in Pathogenic Bacteria and Initial Testing of Compounds

3.1 Introduction and Aims

In previous work in the Conway Group, L-methyl esters OXFKEF12–14 (**31**–**33**) (Figure 3.1) were developed as cell permeant probes for Kef. OXFKEF12–14 (**31**–**33**) were shown to cause growth inhibition/death of *E. coli* MJF335 expressing *Sd*Kef, *Ec*KefC, and *Ab*Kef in the Kirby-Bauer disc diffusion assay.[44,45] These results demonstrated the potential of Kef as a novel antibiotic target in these pathogenic bacteria.

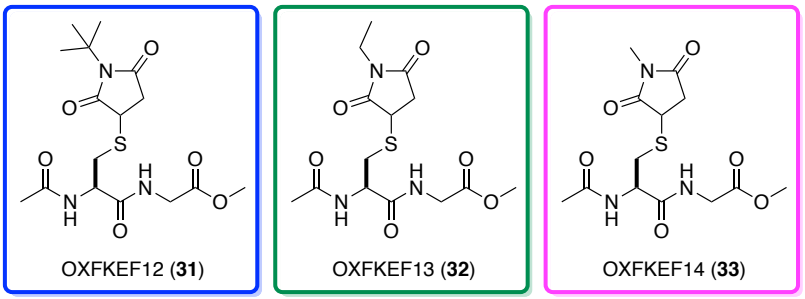


Figure 3.1. Structures of L-methyl esters OXFKEF12–14 (31–33).

To increase the scope of this work, we wanted to investigate Kef systems in other pathogenic bacteria, in collaboration with Defence Science Technology Laboratory (DSTL). These bacteria are resistant to the antibiotics currently available and often suitable vaccines have not been developed. It is therefore essential to identify new antibiotic targets in these pathogenic bacteria.

The overall project aim was to validate Kef as a novel antibiotic target in these pathogenic bacteria. To achieve this aim, the presence of Kef systems in these bacteria was investigated initially. From this analysis, three Kef systems were selected for further study. Having identified suitable targets, the second aim was to test the

L-methyl esters OXFKEF12–14 (**31**–**33**) (Figure 3.1) in *E. coli* expressing the newly-identified Kef systems using the Kirby-Bauer disc diffusion assay. It was envisaged that this approach would reveal some compounds that activate these Kef systems and help to validate them as novel antibiotic targets.

3.2 Target Identification Through Bioinformatics

To identify novel Kef systems in the pathogenic bacteria, we carried out a Domain Enhanced Lookup Time Accelerated-Basic Local Alignment Search Tool (DELTA-BLAST) search of the National Center for Biotechnology Information (NCBI) database, using full-length *Sd*Kef as the query sequence. From each search, only the results with the highest alignment scores (≥200) were analysed to ensure that the homology scores were reliable. *Francisella tularensis* SCHU S4 (*F. tularensis*) and *Coxiella burnetii* (*C. burnetii*) did not contain any proteins with considerable homology to *Sd*Kef (Table 3.1). A homologue of *Sd*Kef was identified in *Bacillus anthracis* Sterne (*B. anthracis*). *B. anthracis* is a Gram-positive pathogen, whereas *S. denitrificans* is Gram-negative, and so the Kef analogue from *B. anthracis* showed poor homology to *Sd*Kef (17% identity, 35% positives) (Table 3.1). Kef homologues were present in *Yersinia pestis* KIM10+ (*Y. pestis*, *Yp*) and *Burkholderia pseudomallei* K96243 (*B. pseudomallei*, *Bp*) (Table 3.1). The Kef system from *Y. pestis* KIM10+ (*Yp*KefB) gave the highest homology score to *Sd*Kef (42% identity, 63% positives) of all the Kef homologues (Table 3.1). The two Kef systems from *B. pseudomallei* K96243, *Bp*Kef1 and *Bp*Kef2, gave lower homology scores to *Sd*Kef (*Bp*Kef1: 31% identity, 50% positives; *Bp*Kef2: 24% identity, 42% positives) (Table 3.1).

Table 3.1. Summary of the search for *Sd*Kef homologues in pathogenic bacteria. The sequence of *Sd*Kef (GenBank ID: ABE53663.1) was used as a query sequence for this NCBI protein BLAST search.[55] The BLAST analysis was performed to search against the non-redundant (nr) protein sequences database, in which repeated protein sequences have been removed, using the DELTA-BLAST algorithm.[56] DELTA-BLAST is sensitive to remote homologues of a query sequence.[56] [a] The genome of *Y. pestis* KIM10+ (taxid:187410) was selected. [b] The genome of *F. tularensis* Biovar A strain SCHU S4 (taxid:177416) was selected. [c] The genome of *B. pseudomallei* K96243 (taxid:272560) was selected. [d] The genome of *C. burnetii* (taxid:777) was selected. [e] The genome of *B. anthracis* Sterne (taxid:260799) was selected. N/A = not applicable. Identity (%): Extent to which amino acid sequences of the Kef system and *Sd*Kef query have the same residues at identical positions in the alignment, expressed as a percentage. Positives (%): Extent to which amino acid sequences of the Kef system and *Sd*Kef query have residues with similar physicochemical properties at identical positions in the alignment, expressed as a percentage. Gaps (%): Spaces introduced in the alignment of the amino acid sequences of the Kef system and *Sd*Kef query to account for insertions and deletions in each sequence compared to the other, expressed as a percentage.[39]

Bacterial Species/Strain	Kef Homologue	Identity (%)	Positives (%)	Gaps (%)
Y. pestis KIM10+ [a]	*Yp*KefB	42	63	2
F. tularensis SCHU S4 [b]	N/A	N/A	N/A	N/A
B. pseudomallei K96243 [c]	*Bp*Kef1	31	50	3
B. pseudomallei K96243 [c]	*Bp*Kef2	24	42	4
C. burnetii [d]	N/A	N/A	N/A	N/A
B. anthracis Sterne [e]	N/A	17	35	9

Table 3.2. Identities of capping residues in the different Kef systems.

Kef System	Capping Residue
*Sd*Kef	F448
*Ec*KefC	F441
*Ec*KefB	Y442
*Yp*KefB	Y442
*Bp*Kef1	A453
*Bp*Kef2	D478

The primary sequences of *Yp*KefB, *Bp*Kef1, and *Bp*Kef2 were overlaid with those of *Sd*Kef, *Ec*KefC, and *Ec*KefB, so that the conservation of residues across the different proteins could be analysed. Due to their proposed role in the Kef mechanism of action, described in 1.7, the capping residues were analysed (Table 3.2). The capping residue in *Yp*KefB was identified as tyrosine (Y442), the same as *Ec*KefB (Y442), and similar to

*Sd*Kef and *Ec*KefC, where the capping residues are phenylalanines (F448 and F441, respectively). In contrast, the capping residues in *Bp*Kef1 and *Bp*Kef2 were alanine (A453) and aspartic acid (D478), respectively.

The GSH/GSX-binding residues of the different proteins were revealed by overlaying their primary sequences (using CLC Main Workbench 7) and the conservation of residues across the sequences was analysed relative to *Ec*KefC (Figure 3.2). This analysis showed that Q412, which interacts with the Glu amide C=O and the Gly amide C=O of ESG (**20**), is conserved in *Yp*KefB (Q413) and *Bp*Kef1 (Q424) but not in *Bp*Kef2 (T449). R416 (*Ec*KefC), which forms one interaction with N551 and two with the Gly carboxylate of ESG (**20**), is conserved in *Yp*KefB (R417) and *Bp*Kef1 (R428), but not in *Bp*Kef2 (H453). R498 in *Ec*KefC, which interacts with the Glu carboxylate of ESG (**20**), is conserved in *Yp*KefB (R499), but not in *Bp*Kef1 (V510) or in *Bp*Kef2 (N535). Notably, the corresponding residue in *Sd*Kef is I505, and so hydrophobic residues at this position are expected to allow the protein to function as in *Ec*KefC and *Sd*Kef. D499, which interacts with the α-NH_2 of ESG (**20**), is conserved in *Bp*Kef1 (D511), but not in *Bp*Kef2 (S536) or *Yp*KefB (G500). The latter is the same as *Ec*KefB (G500). V500, which forms an interaction with the C=O of the Gly amide, is not conserved in any of the proteins. The corresponding residues are R501 in *Yp*KefB, D512 in *Bp*Kef1, and D537 in *Bp*Kef2. However, as it is the amide backbone of V500 which interacts with the ligand, the identity of the side chain is predicted to be less important for ligand binding. R516, which interacts with N551 and the Gly carboxylate of ESG (**20**), is conserved in *Yp*KefB (R517), but not in *Bp*Kef1 (P528) or *Bp*Kef2 (V553). N551 in *Ec*KefC interacts with R416, R516, and the Gly carboxylate of ESG (**20**) *via* the structurally-conserved water

molecule. The residue corresponding to N551 is R563 in *Bp*Kef1, a gap in *Bp*Kef2, and D552 in *Yp*KefB. The corresponding residue is also aspartic acid in *Ec*KefB (D552) and in *Sd*Kef (D558), and so it is known that this residue is tolerated in this position.

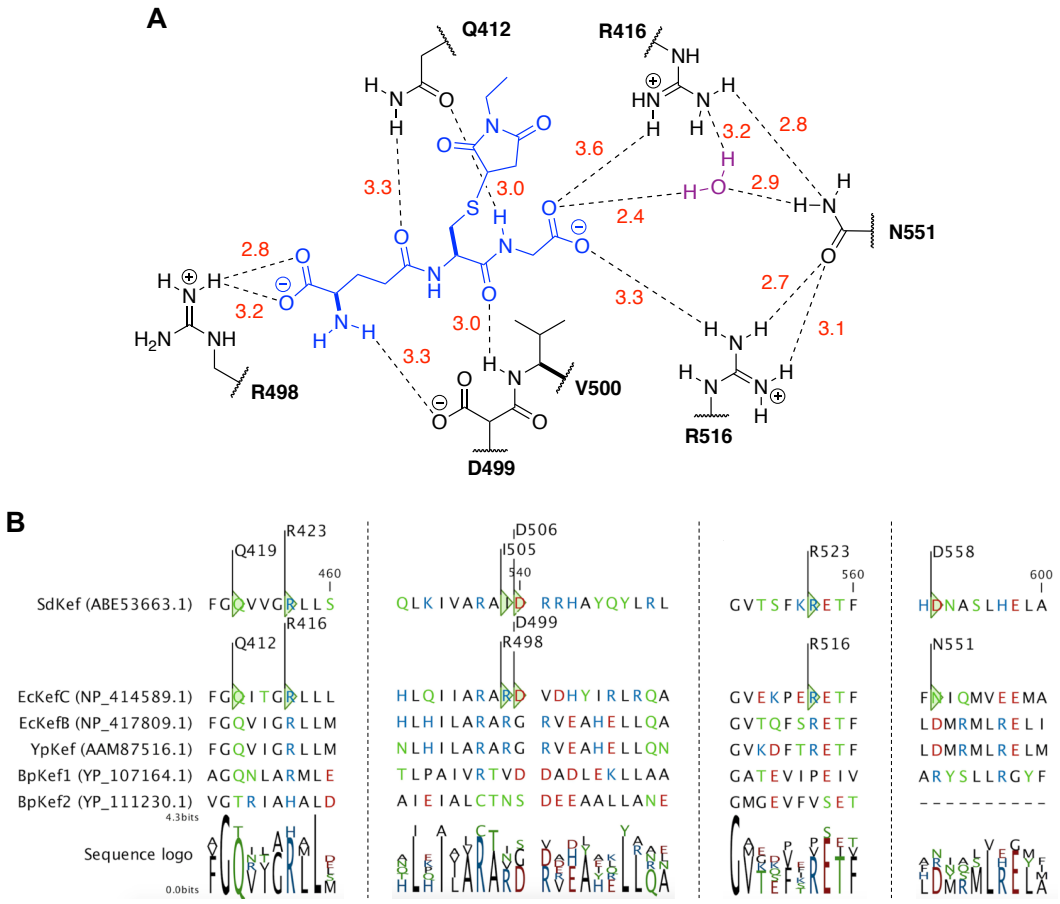


Figure 3.2. Analysis of GSH/GSX-binding residues in Kef homologues. A) Schematic of the residues which bind ESG (**20**, shown in blue) in *Ec*KefCCTD-KefF.[41] The residues (black) are as follows: Q412 and R416 (both from one subunit), and R498, D499, V500 (backbone N–H), R516, and N551 (all from the other subunit). The conserved water molecule is shown in purple. Hydrogen bonds are shown as dashed lines (black), with bond lengths (red) in angstroms. B) Alignment of primary protein sequences. Selected regions of the protein sequence are shown, in which the GSH/GSX-binding residues are located (labelled in figure). The sequence logo shows the conversation of a residue across the different sequences. The colour of each residue indicates its polarity: Black (hydrophobic: A, F, G, I, L, M, P, V, and W), blue (basic: H, K, and R), red (acidic: D and E) and green (hydrophilic, uncharged: C, N, Q, S, T, and Y).

Finally, the requirement of ancillary proteins for *Yp*KefB, *Bp*Kef1, and *Bp*Kef2 was determined (Figure 3.3). *Sd*Kef does not require an ancillary protein for full function, whereas *Ec*KefC requires KefF, and *Ec*KefB requires KefG for full function.[33] *Ec*KefB is predicted to have an additional ancillary protein KefP, which is of unknown function (Dr Anthony Chan, unpublished). The genes in the regions of *YpkefB*, *Bpkef1*, and *Bpkef2* were analysed in the genomes of the different bacterial species (*Yp*, *Bp*). *Yp*KefB is predicted to have two ancillary proteins, *Yp*KefG and *Yp*KefP, the same as *Ec*KefB (Figure 3.3). *Bp*Kef1 and *Bp*Kef2 are not predicted to have ancillary proteins, as is the case for *Sd*Kef (Figure 3.3).

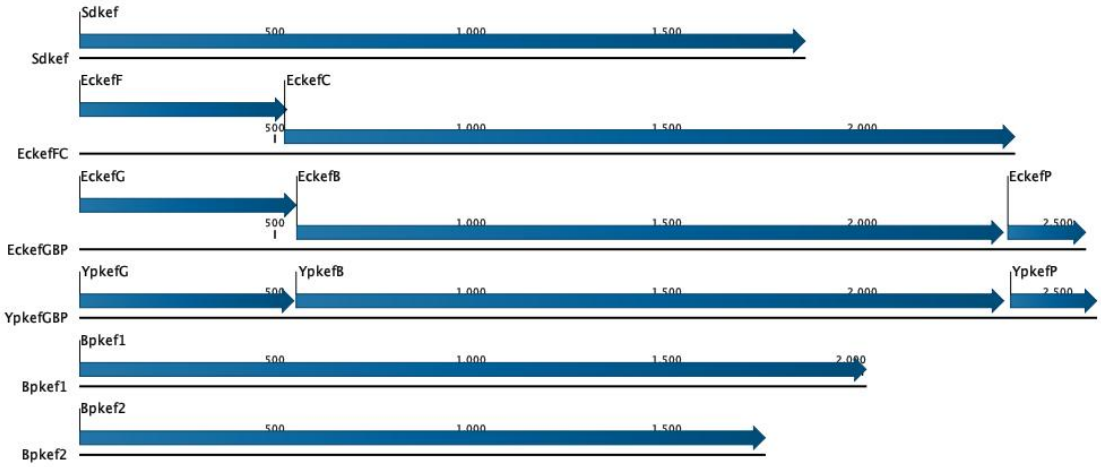
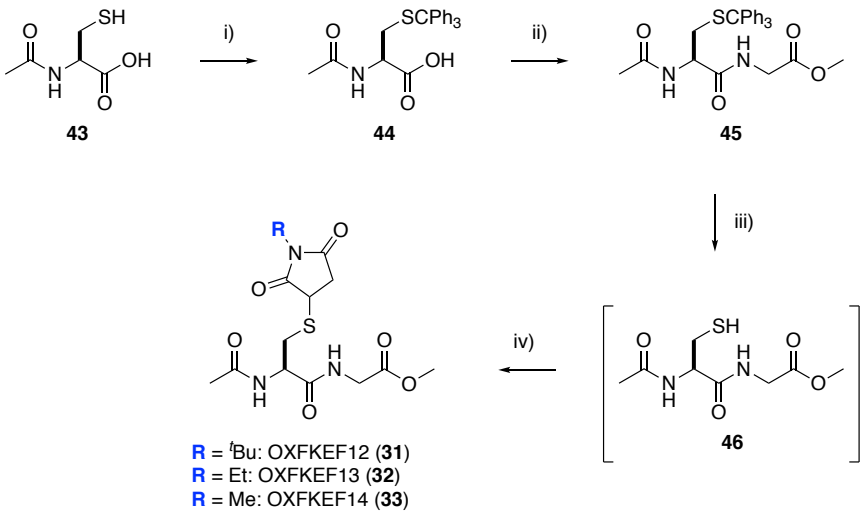


Figure 3.3. Arrangement of *kef* genes in *S. denitrificans*, *E. coli*, *Y. pestis*, and *B. pseudomallei*.

In summary, Kef systems were identified: *Yp*KefB, *Bp*Kef1, and *Bp*Kef2. *Yp*KefB was shown to be similar to *Sd*Kef, *Ec*KefC, and *Ec*KefB, showing moderate homology scores to *Sd*Kef (42% identity, 63% positives) and good conservation of the GSH/GSX-binding residues. *Yp*KefB was predicted to possess two ancillary proteins, *Yp*KefG and *Yp*KefP, similar to *Ec*KefB. In comparison, *Bp*Kef1 and *Bp*Kef2 gave lower homology scores to *Sd*Kef (*Bp*Kef1: 31% identity and 50% positives; *Bp*Kef2: 24% identity and 42% positives) and poorer conservation of the GSH/GSX-binding residues. It was therefore hypothesised that the L-methyl esters OXFKEF12–14 (**31**–**33**) were more likely to activate *Yp*KefGBP than *Bp*Kef1 or *Bp*Kef2. To investigate this hypothesis, OXFKEF12–14 (**31**–**33**) would be tested in *E. coli* expressing *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 using the Kirby-Bauer disc diffusion assay.

3.3 Synthesis of L-Methyl Esters OXFKEF12–14 (31–33)

The L-methyl esters OXFKEF12–14 (**31**–**33**) were synthesised in four steps from *N*-acetyl-L-cysteine **43** using the previously reported route (Scheme 3.1).[44,45] Trityl protection of the free thiol in the starting material **43** occurred in 45–91% yields. The resulting trityl-protected derivative **44** was coupled to glycine methyl ester hydrochloride to give the fully-protected dipeptide intermediate **45**, in yields of 44–93%. Deprotection of the trityl group in compound **45** revealed the thiol **46**, which was submitted to conjugate addition reactions with the three different maleimides (*N*-*tert*-butyl (**42**), *N*-ethyl (**19**), and *N*-methyl (**47**)) to give OXFKEF12, OXFKEF13, and OXFKEF14 (**31, 32**, and **33**, respectively) in 38–52% yields for **31**, in 14–36% yields for **32**, and in a 26% yield for **33**.



Scheme 3.1. Synthesis of L-methyl esters OXFKEF12–14 (31–33). *Reagents and conditions*: i) Ph_3CCl, *N*,*N*-Dimethylformamide (DMF), rt, 20–27 h, 45–91% (n = 5); ii) Glycine methyl ester hydrochloride, *N,N,N′,N′*-tetramethyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate (HBTU), *N*,*N*-diisopropylethylamine (DIPEA), DMF, rt, 16–74 h, 44–93% (n = 10); iii) Et_3SiH, 20% *v*/*v* Trifluoroacetic acid (TFA) in CH_2Cl_2, rt, 20 min–16 h; iv) R = ^{t}Bu: *N*-*tert*-Butylmaleimide (**42**), H_2O, rt, 10–46 h, 38–52% (over 2 steps, n = 2); R = Et: NEM (**19**), H_2O, rt, 15–44 h, 14–38% (over 2 steps, n = 2); R = Me: *N*-Methylmaleimide (**47**), H_2O, rt, 24 h, 26% (over 2 steps).

All final compounds (**31**, **32**, and **33**) were of >95% purity as determined by RP analytical HPLC. The enantiomeric excess (*e.e.*) values of final compounds **32** and **33** were calculated using chiral HPLC. The chiral HPLC traces for compounds **32** and **33** are shown in Figure 3.4B and Figure 3.5B, respectively. The chiral HPLC traces for the corresponding D-enantiomers, OXFKEF16 (**48**) and OXFKEF17 (**35**), are shown in Figure 3.4C and Figure 3.5C, respectively. For each compound, the stereochemistry was referenced to a sample of scalemic compound (Figure 3.4A and Figure 3.5A). Both **32** and **33** contain a mixture of two diastereoisomers, in an approximate 1:1 ratio, as for ESG (**20**).[41] Separation of these diastereoisomers was not observed during chiral HPLC. However, separation of the species with L- and D- stereochemistry at the Cys-α stereogenic centre was observed (Figure 3.4 and Figure 3.5). For **32** and **33**, the quoted *e.e.* is the difference between the total proportion of the two diastereoisomers with L- stereochemistry at the Cys-α stereogenic centre and the total proportion of the two diastereoisomers with D- stereochemistry at the Cys-α stereogenic centre. The *e.e.* of compound **32** was 66% and the *e.e.* of compound **33** was approximately 71%. The *e.e.* for **33** was not accurate because the peaks did not separate completely by chiral HPLC (Figure 3.5B). Epimerisation of the Cys-α stereogenic centre most likely occurred during the amide coupling, as the enantiopurity of the substrate was greatly reduced during this step (98% to 77% *e.e.*). This result was unsurprising, as *N,S*-protected cysteine derivatives have been shown to undergo rapid racemisation upon activation for amide coupling reactions.[57] As the *e.e.* values for the final compounds (**32** and **33**) were lower than for the intermediate **45**, this suggested that further epimerisation of the Cys-α stereogenic centre occurred during the trityl deprotection and/or conjugate addition reactions to the two different maleimides.

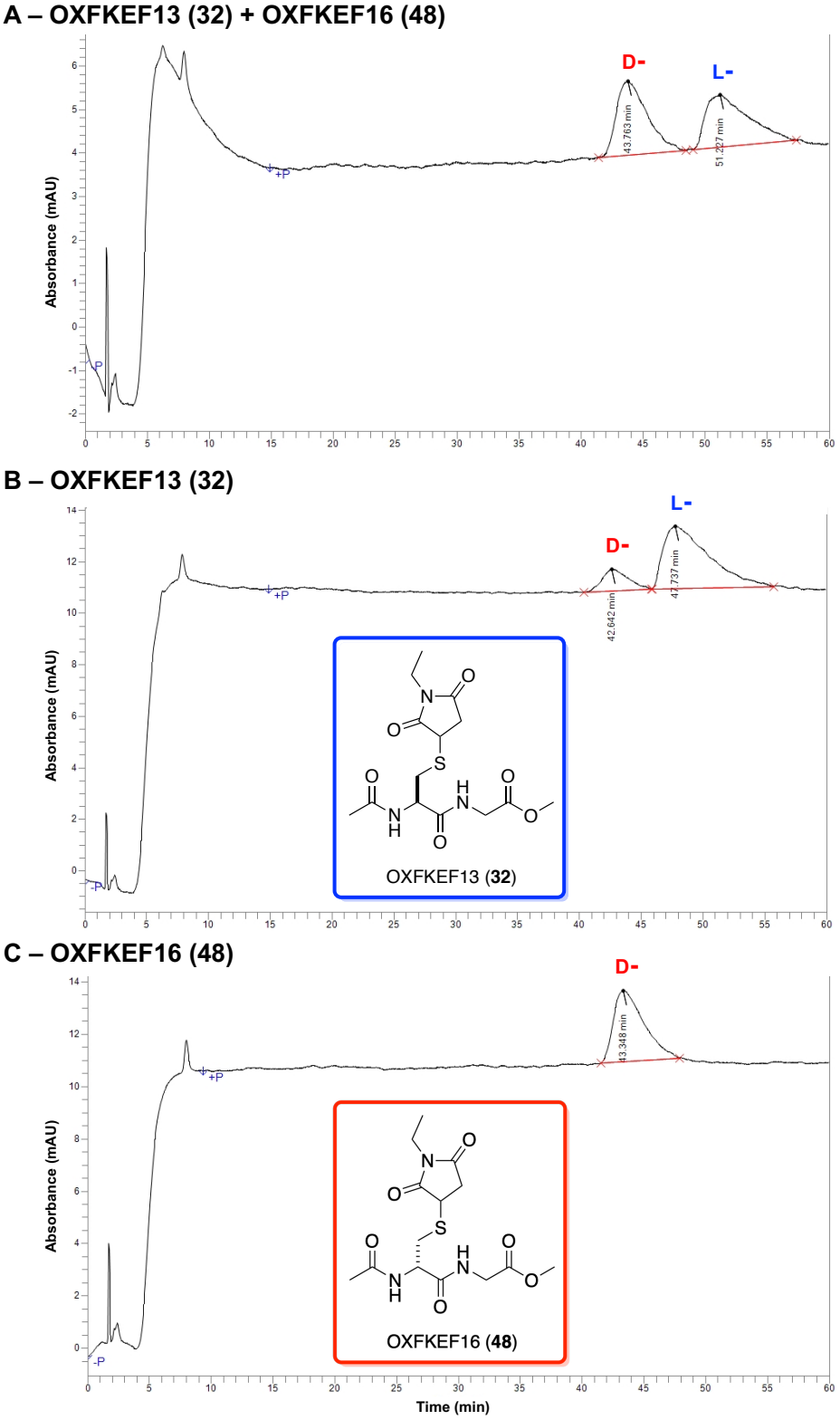


Figure 3.4. Chiral HPLC traces for OXFKEF13 (32) and OXFKEF16 (48). A) Scalemic mixture: **32** + **48**. B) Compound **32** only. C) Compound **48** only. All three compounds were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental.

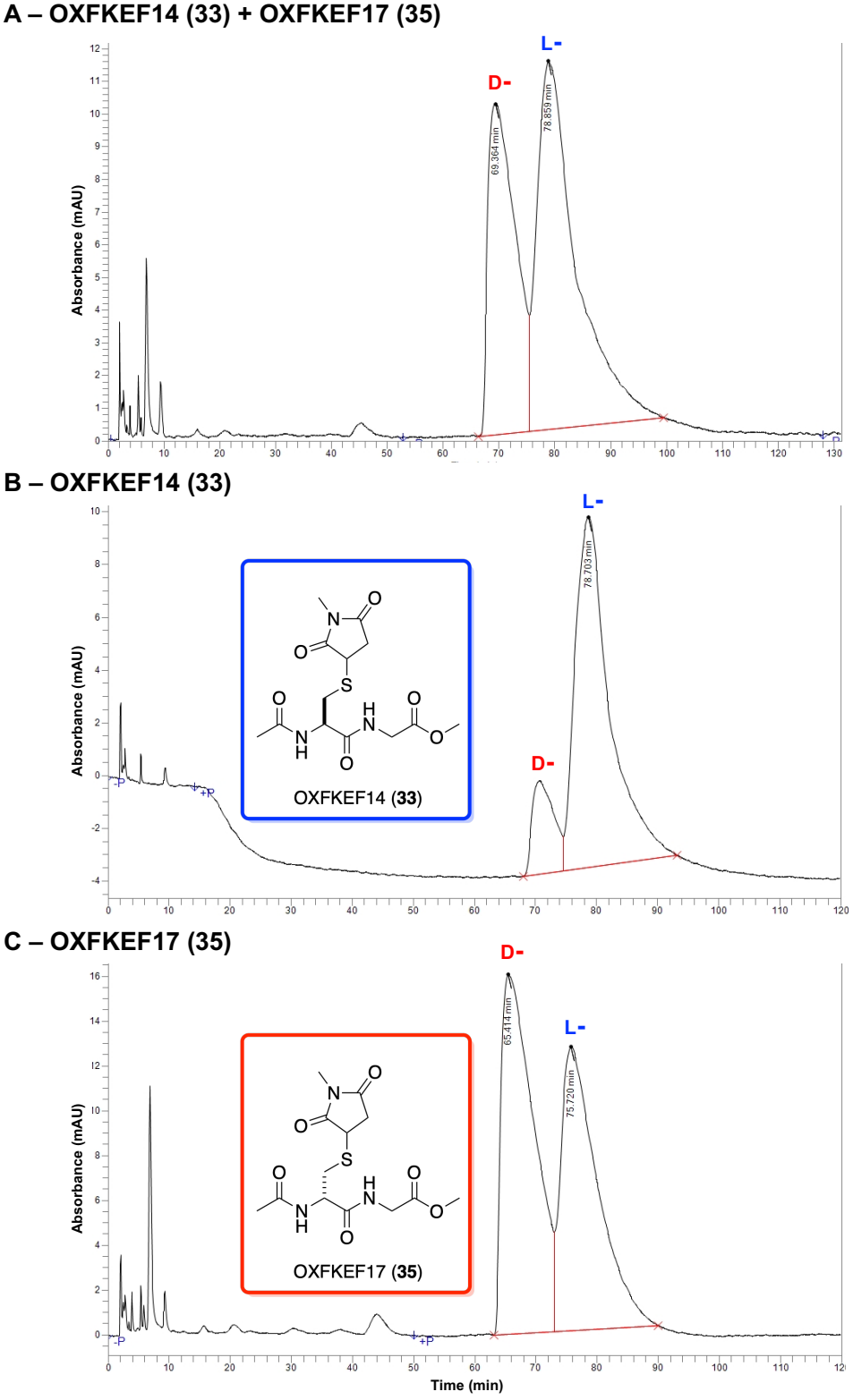
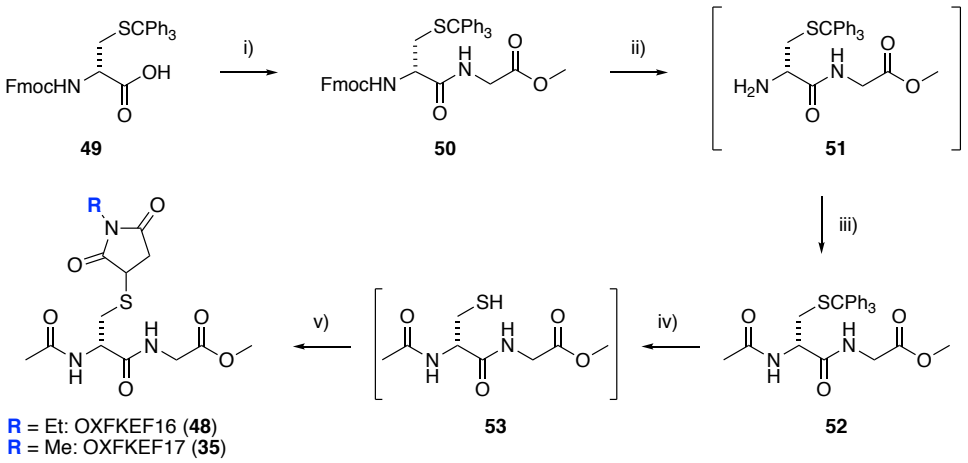


Figure 3.5. Chiral HPLC traces for OXFKEF14 (33) and OXFKEF17 (35). A) Scalemic mixture: **33** + **35**. B) Compound **33** only. C) Compound **35** only. All three compounds were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental.

3.4 Synthesis of D-Methyl Esters OXFKEF16 (48) and OXFKEF17 (35)

The D-methyl esters OXFKEF16 (**48**) and OXFKEF17 (**35**) (Scheme 3.2) were synthesised as they are excellent negative control compounds in the Kirby-Bauer disc diffusion assay. It was predicted that these compounds would not activate *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2, due to the incorrect vector of the S-substituent. The *N*-methyl compound OXFKEF17 (**35**) has previously been shown to not activate *Sd*Kef, *Ec*KefFC, or *Ab*Kef in the disc assay.[44,45]



Scheme 3.2. Synthesis of D-methyl esters OXFKEF16 (48) and OXFKEF17 (35). *Reagents and conditions*: i) Glycine methyl ester hydrochloride, HBTU, DIPEA, DMF, rt, 16–29 h, 61–92% (n = 8); ii) 20% *v/v* Piperidine in DMF, rt, 15–40 min; iii) Method 1: AcCl, NEt_3, CH_2Cl_2, 0 °C → rt, 3.5 h, 14% (over 2 steps); iii) Method 2A: Ac_2O, *N*,*N*-Dimethyl-4-aminopyridine (DMAP), pyridine, rt, 11–19 h, 27–41% (over 2 steps, n = 3); iii) Method 2B: Ac_2O, Pyridine, THF, rt, 11–18 h, 34–48% (over 2 steps, n = 2); (iv) Et_3SiH, 20% *v/v* TFA in CH_2Cl_2, rt, 20–30 min; v) R = Et: NEM (**19**), H_2O, rt, 19 h, 75% (over 2 steps); R = Me: *N*-Methylmaleimide (**47**), H_2O, rt, 3–6 h, 34–46% (over 2 steps, n = 2).

To synthesise OXFKEF16 (**48**) and OXFKEF17 (**35**), Fmoc-D-Cys(Trt)-OH (**49**) was coupled to glycine methyl ester hydrochloride, using the same conditions as for the L-amino acid **44**, to give the fully-protected dipeptide **50** in 61–92% yields (Scheme 3.2). Deprotection of the 9-fluorenylmethoxycarbonyl (Fmoc) group revealed the amine **51**, which was acetylated directly to give intermediate **52**. Compound **52** is the D-enantiomer of **45**, the intermediate from the L-methyl ester synthesis (Scheme 3.1).

Acetylation of the amine **51** using acetyl chloride (Method 1) gave a 14% yield over 2 steps. Acetic anhydride was investigated as an alternative, with either DMAP in pyridine (Method 2A) or pyridine in THF (Method 2B). These conditions gave yields of 27–41% and 34–48%, respectively. As in the synthesis of the L-methyl esters, trityl deprotection of intermediate **52** to the thiol **53** and conjugate addition to the maleimides gave the final compounds, **48** (*N*-ethyl) and **35** (*N*-methyl), in a 75% yield for compound **48** and in 34–46% yields for compound **35**.

The purity of both final compounds (**48** and **35**) was >95% by RP analytical HPLC. The *e.e.* values of compounds **48** and **35** were determined using chiral HPLC, in the same manner as for their corresponding L-enantiomers (**32** and **33**). The chiral HPLC traces for compounds **48** and **35** are shown in Figure 3.4C and Figure 3.5C, respectively. The *e.e.* calculated for compound **48** was >99% but the *e.e.* of compound **35** was only 6%. As for the L-enantiomer **33**, the *e.e.* for compound **35** was not accurate because the L- and D-enantiomers only separated partially by chiral HPLC. In the trace for compound **35** (Figure 3.5C), two peaks were clearly visible, which were in an approximate 1:1 ratio and seemed to correspond to the two peaks in the scalemic sample (**33** + **35**, Figure 3.5A), presumably the L- and D-enantiomers. This suggested that compound **35** had been partially epimerised at the Cys-α stereogenic centre but not in compound **48**.

The two differences between the synthesis of these compounds (**35**, **48**) were the acetylation conditions used to prepare the intermediate **52** and the identity of the maleimide used in the conjugate addition reaction. In the preparation of both batches of the *N*-methyl compound **35**, two batches of the intermediate **52** were combined,

one made using Method 2A (Ac_2O, DMAP, pyridine) and the other using Method 2B (Ac_2O, pyridine, THF). Chiral HPLC of intermediate **52** was analysed using a batch from Method 2A, which showed that this intermediate was enantiopure (>99% *e.e.*). The *e.e.* of the batch of compound **52** made using Method 2B was not analysed. Epimerisation of the Cys-α stereogenic centre in compound **52** could have occurred during the acetylation, which would explain the *e.e.* of the final compound (**35**). Epimerisation of the Cys-α stereogenic centre could also have occurred during the conjugate addition of the thiol **53** to *N*-methylmaleimide (**47**) or after the final compound was isolated. The identity of the maleimide used in the conjugate addition could have affected the *e.e.* of the final compounds (**48**, **35**), but, as the two maleimides only differed by a methylene group, which was spatially removed from the site of reaction, this was deemed less likely to be the cause of the epimerisation. Due to the very low *e.e.* (6%) for the *N*-methyl compound **35**, it was not tested in the disc assay.

3.5 Molecular Cloning to Synthesise Plasmids for Expression of *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in *E. coli*

To test compounds against *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in the Kirby-Bauer disc diffusion assay, it was necessary to design and make plasmids for the heterologous expression of these Kef systems in *E. coli*. In the disc assay, rather than testing compounds against the pathogenic bacteria directly, the Kef systems were overexpressed in *E. coli*, using plasmids. The resulting *E. coli* transformants were used in the assay. Transformation into *E. coli* was assumed to lead to overproduction of the Kef system compared to its production in the relevant wild-type bacterial species.

The designed plasmids were pTrcYpKefGBP, pTrcBpKef1H$_6$, and pTrcBpKef2H$_6$ (Figure 3.6A–C). Preparation of these plasmids using cloning is described in section 6.3.1 of the experimental. In the disc assay, the presence of *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in the *E. coli* transformants was inferred by ampicillin (Amp) resistance, as all three plasmids contain the *bla* gene for β-lactamase, promoted by Pbla (Figure 3.6A–C).

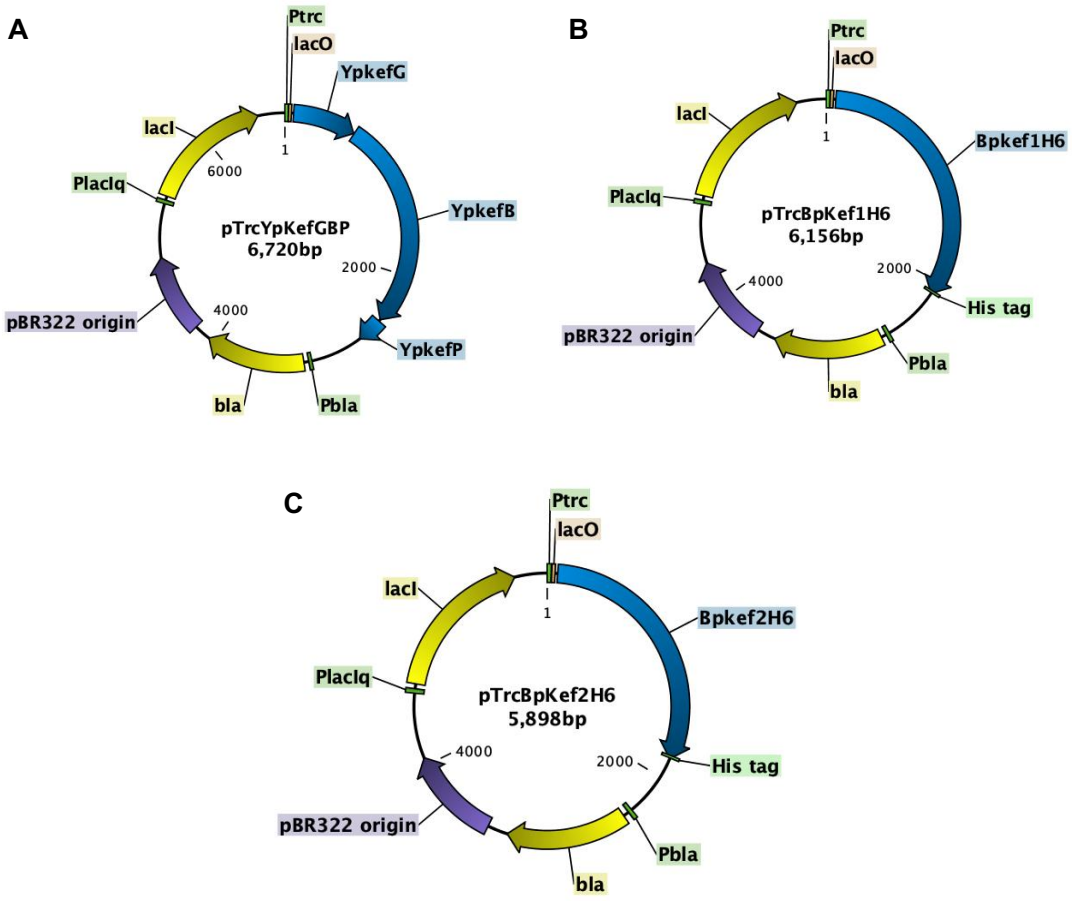


Figure 3.6. Plasmid maps of pTrcYpKefGBP (A), pTrcBpKef1H$_6$ (B), and pTrcBpKef2H$_6$ (C). The *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, and *Bpkef2H$_6$* genes code for the *Yp*KefG, *Yp*KefB, *Yp*KefP, *Bp*Kef1H$_6$, and *Bp*Kef2H$_6$ proteins, respectively. Ptrc is the promoter for the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, and *Bpkef2H$_6$* genes. His tag = H$_6$ tag. *bla* is the gene for β-lactamase, which causes resistance to Amp. Pbla is the promoter for the gene *bla*. pBR322 origin is the origin of replication, which ensures that the plasmid is replicated when it is transformed into *E. coli*. *lacI* is the gene for the *lac* repressor (LacI), which prevents transcription of the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, and *Bpkef2H$_6$* genes by binding to the *lac* operon (LacO). IPTG binds to LacI, preventing LacI from binding to LacO, promoting transcription of the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, and *Bpkef2H$_6$* genes. PlacIq is the promoter for the gene *lacI*.

3.6 Kirby-Bauer Disc Diffusion Assay in *E. coli* Expressing *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2

In the majority of the disc assays described in this section, the MJF645 *E. coli* strain was used. MJF645 is a mutant *E. coli* strain, lacking the genes for the endogenous Kef proteins and glutathione biosynthesis (*kefB*, *kefC*, *kefF*, and *kefG*, and *gshA*, respectively).[31] MJF645 is therefore appropriate for the expression of Kef systems from other bacterial species.

3.6.1 Disc Assay for *E. coli* Expressing *Bp*Kef1 and *Bp*Kef2

To determine whether L-methyl esters OXFKEF12–14 (**31**–**33**) would activate *Bp*Kef1 and *Bp*Kef2, the compounds were tested in *E. coli* MJF645 expressing either *Bp*Kef1 or *Bp*Kef2. The L-methyl esters were also tested in MJF645 expressing *Sd*Kef, using the pTrcSdKefH_6 plasmid, as a positive control. In the positive control of *Sd*Kef, all three compounds OXFKEF12–14 (**31**–**33**) activated MJF645 expressing *Sd*Kef (Figure 3.7A). However, none of the compounds caused zones of growth inhibition/death for MJF645 expressing either *Bp*Kef1 or *Bp*Kef2 (Figure 3.7B and C, respectively). The results obtained for *Bp*Kef1 and *Bp*Kef2 were consistent with our previous hypothesis, as L-methyl esters OXFKEF12–14 (**31**–**33**) did not activate *E. coli* MJF645 expressing either *Bp*Kef1 or *Bp*Kef2. There are several possible reasons for this, which include: the succinimide moiety in these compounds is not large and/or rigid enough to displace the capping residue in either protein (alanine in *Bp*Kef1 and aspartic acid in *Bp*Kef2); the capping residue is being displaced, but this does not lead to activation of Kef; or, the endogenous low molecular weight thiol in *B. pseudomallei* is not glutathione, and therefore these truncated GSX cannot bind or activate its two Kef systems, *Bp*Kef1 and *Bp*Kef2.

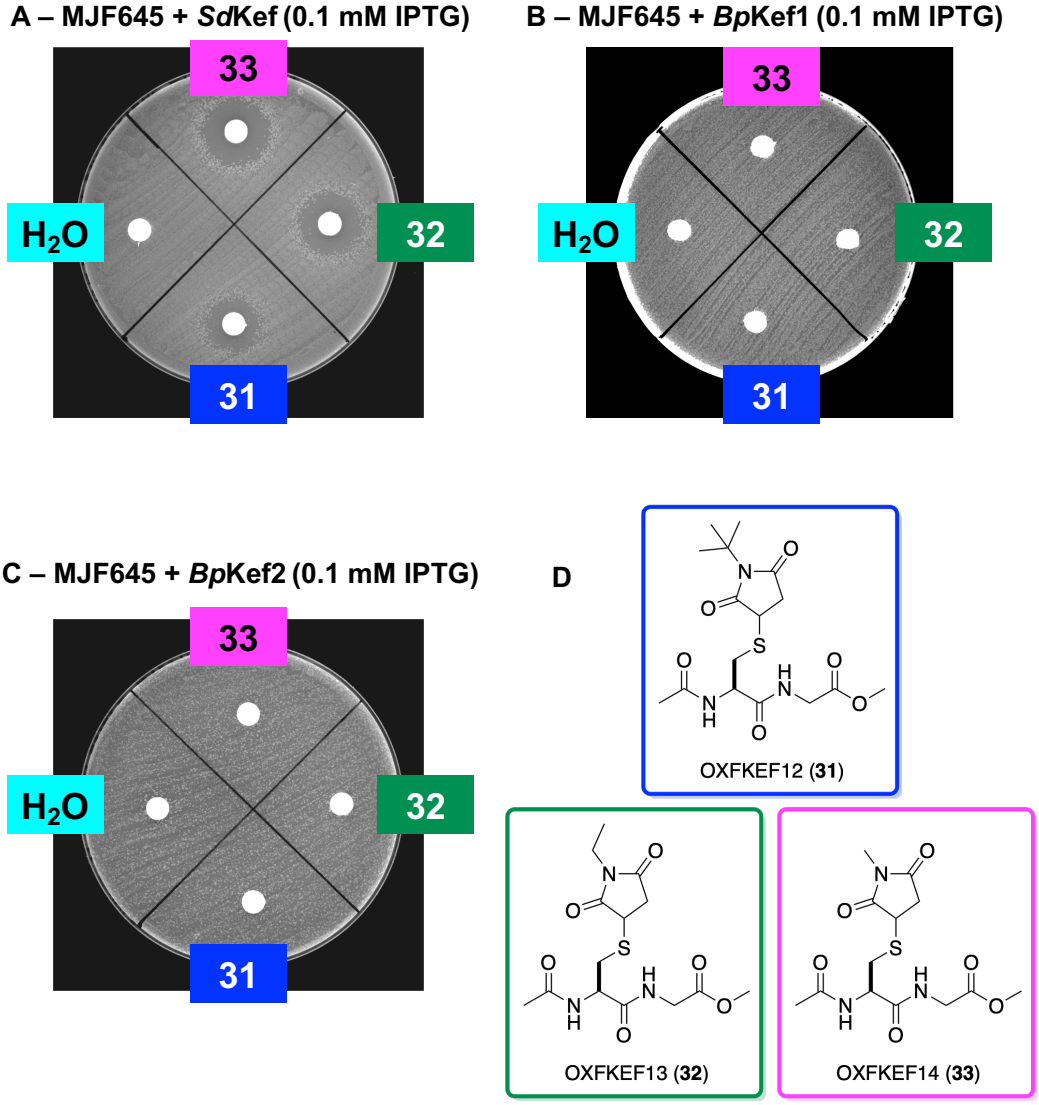


Figure 3.7. Kirby-Bauer disc diffusion assay to test OXFKEF12–14 (31–33) in *E. coli* MJF645 expressing either *Bp*Kef1 or *Bp*Kef2. 31 = OXFKEF12 (**31**), **32** = OXFKEF13 (**32**), and **33** = OXFKEF14 (**33**). All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (positive control, 0.1 mM IPTG). B) MJF645 + pTrcBpKef1H_6 (0.1 mM IPTG). C) MJF645 + pTrcBpKef2H_6 (0.1 mM IPTG). D) Structures of L-methyl esters OXFKEF12–14 (**31**–**33**).

3.6.2 Disc Assays for *E. coli* Expressing *Yp*KefGBP

3.6.2.1 Assays for *E. coli* MJF645 Expressing *Yp*KefGBP

The L-methyl esters OXFKEF12–14 (**31**–**33**) were tested in MJF645 expressing *Yp*KefGBP. As for the previous experiments, the L-methyl esters OXFKEF12–14 (**31**–**33**) were also tested in MJF645 expressing *Sd*Kef, as a positive control. The positive control was successful, as all three compounds OXFKEF12–14 (**31**–**33**) activated *E. coli* MJF645 expressing *Sd*Kef (Figure 3.8A (0.1 mM IPTG) and B (0.05 mM IPTG)). Two different concentrations of IPTG were used for MJF645 expressing *Yp*KefGBP (0.1 mM and 0.05 mM), as too high a concentration of IPTG caused poor cell growth during the disc assay, presumably by affecting the amount of *Yp*KefGBP protein being produced in the cells. The positive control, MJF645 expressing *Sd*Kef, was therefore carried out using both concentrations of IPTG (0.1 mM and 0.05 mM).

When L-methyl esters OXFKEF12–14 (**31**–**33**) were tested in MJF645 expressing *Yp*KefGBP, zones of growth inhibition/death were seen for the *N*-ethyl compound **32** and the *N*-methyl compound **33** (Figure 3.8C (0.1 mM IPTG) and D (0.05 mM IPTG)). These results were consistent with our hypothesis, as compounds **32** and **33** activated MJF645 expressing *Yp*KefGBP. It is important to note the low *e.e.* of these compounds (66% for **32** and ~71% for **33**). If the *e.e.* values of these compounds were >99%, it is predicted that these compounds would be more active in the disc assay. The *N*-methyl compound OXFKEF14 (**33**) produced a larger zone of exclusion than the *N*-ethyl compound OXFKEF13 (**32**) for MJF645 expressing *Yp*KefGBP (Figure 3.8C and D). Compound **33** was therefore inferred to be more active than compound **32**. These

results are consistent with the SAR seen previously for *E. coli* expressing *Sd*Kef and *Ec*KefFC.[45]

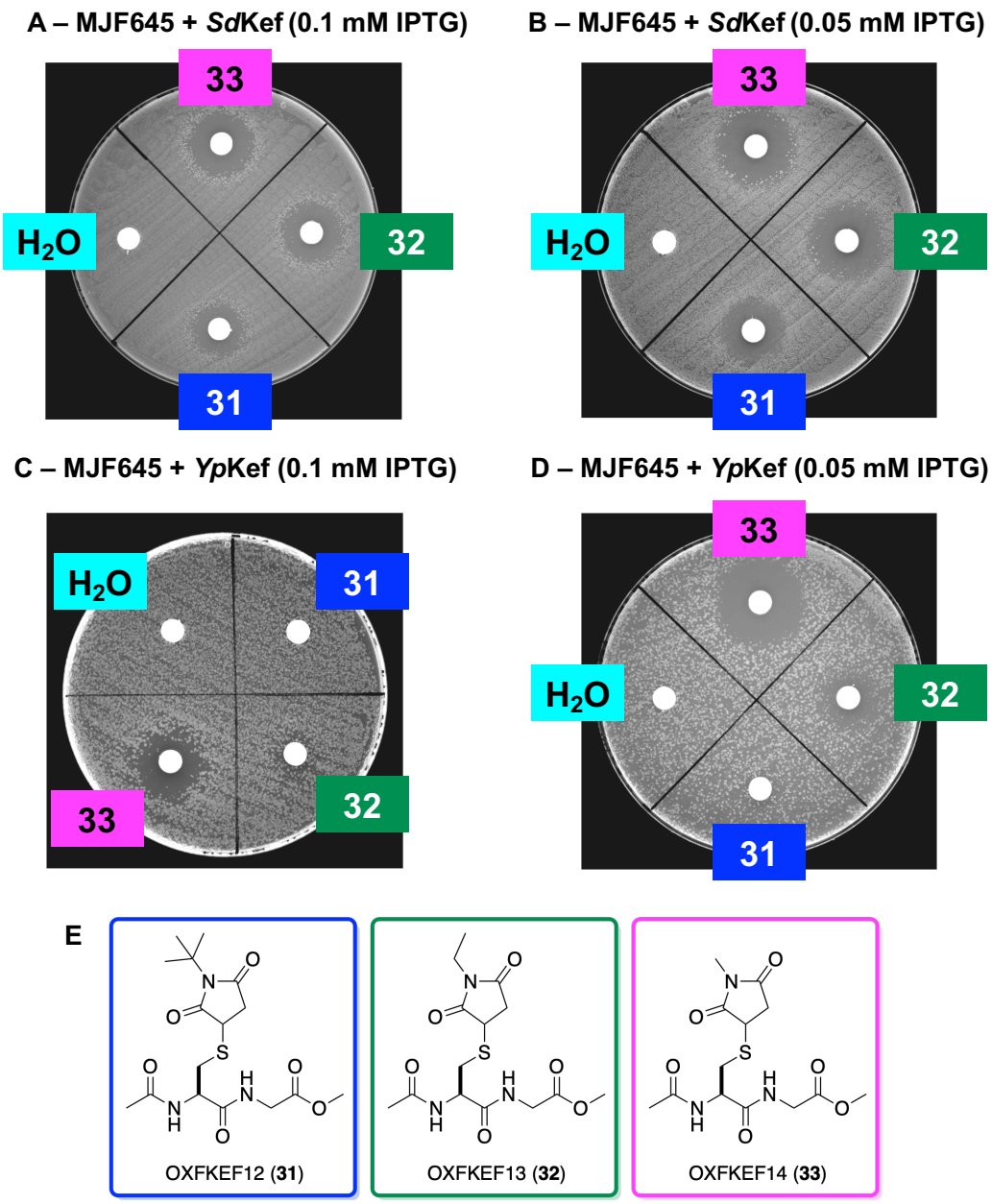


Figure 3.8. Kirby-Bauer disc diffusion assay to test L-methyl esters OXFKEF12–14 (31–33) in *E. coli* MJF645 expressing *Yp*KefGBP. 31 = OXFKEF12 (**31**), **32** = OXFKEF13 (**32**), and **33** = OXFKEF14 (**33**). All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcSdKefH$_6$ (positive control, 0.1 mM IPTG). B) MJF645 + pTrcSdKefH$_6$ (positive control, 0.05 mM IPTG). C) MJF645 + pTrcYpKefGBP (0.1 mM IPTG). D) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). E) Structures of L-methyl esters OXFKEF12–14 (**31**–**33**).

To validate the results obtained with OXFKEF13 (**32**) and OXFKEF14 (**33**) for MJF645 expressing *Yp*KefGBP, several control experiments were carried out. Firstly, the L-methyl esters OXFKEF12–14 (**31**–**33**) were investigated in MJF645 transformed with pTrcYpKefGBP in the absence of IPTG. None of the compounds caused inhibition of growth/death (Figure 3.9C), which suggests that the antibacterial activity seen in the presence of IPTG (Figure 3.9A and B, included as positive controls) is caused by the compounds acting on *Yp*KefGBP.

Secondly, the L-methyl esters OXFKEF12–14 (**31**–**33**) were tested in the MJF645 cell strain in the absence of both IPTG and plasmid, to check for compound toxicity. None of the compounds caused inhibition of growth/death to MJF645 (Figure 3.9D), which provides further evidence that the activity seen for MJF645 expressing *Yp*KefGBP (Figure 3.9A and B) was due to the compounds activating *Yp*KefGBP, as opposed to off-target effects or compound toxicity.

As a third negative control, the *N*-ethyl D-methyl ester OXFKEF16 (**48**) was tested in MJF645 expressing *Yp*KefGBP. While the positive control compounds OXFKEF13 (**32**) and OXFKEF14 (**33**) produced zones of growth inhibition/death on the assay plate, OXFKEF16 (**48**) did not activate *E. coli* expressing *Yp*KefGBP (Figure 3.9E). This result provides further evidence that the effect seen for the L-enantiomer **32** in MJF645 expressing *Yp*KefGBP (Figure 3.9B and E) was due to activation of *Yp*KefGBP, as opposed to off-target effects or compound toxicity.

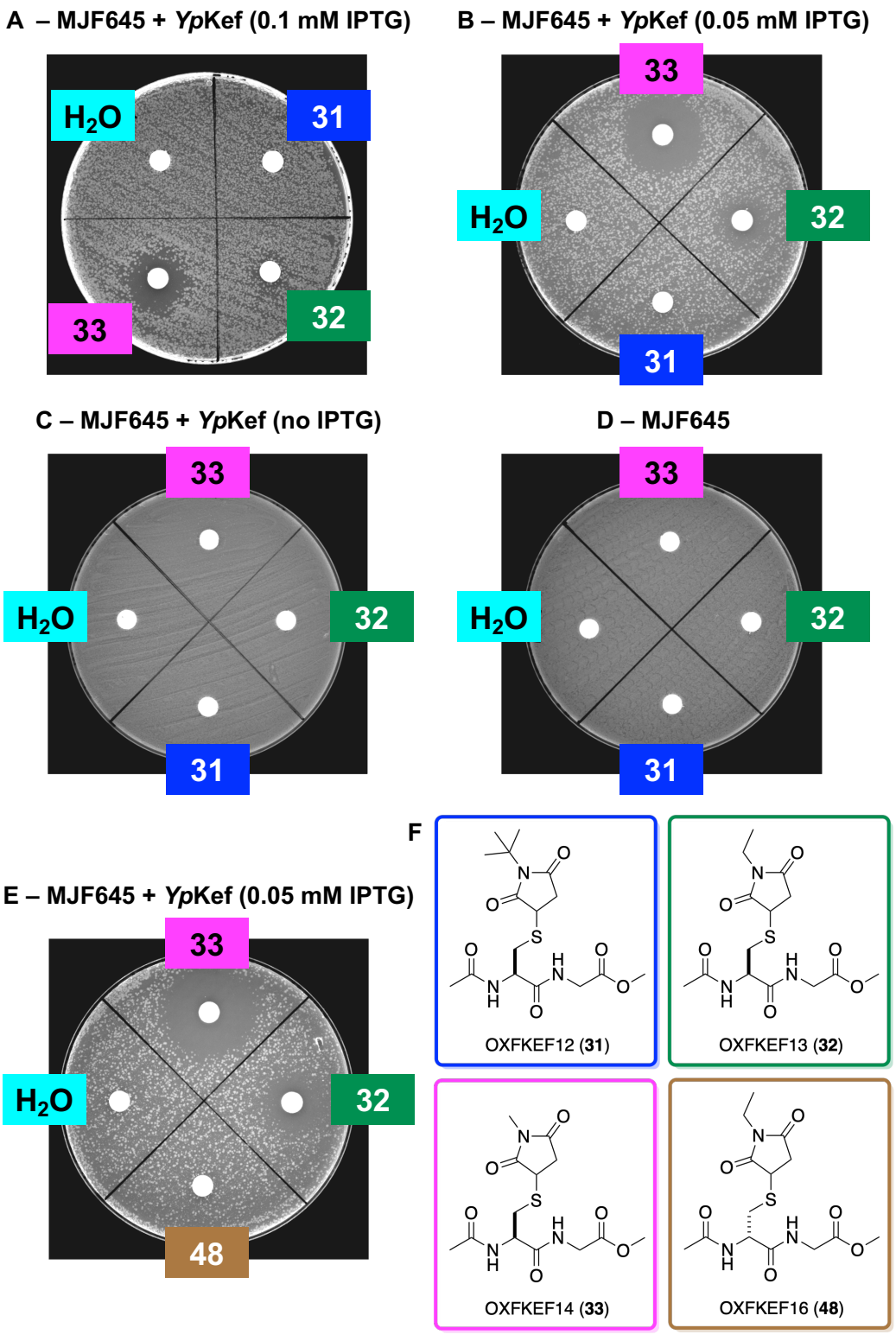


Figure 3.9. Control experiments for MJF645 expressing *Yp*KefGBP. 31 = OXFKEF12 (**31**), **32** = OXFKEF13 (**32**), **33** = OXFKEF14 (**33**), and **48** = OXFKEF16 (**48**). All compounds (OXFKEF12–14 (**31**–**33**) and OXFKEF16 (**48**)) were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcYpKefGBP (positive control, 0.1 mM IPTG). B) MJF645 + pTrcYpKefGBP (positive control, 0.05 mM IPTG). C) MJF645 + pTrcYpKefGBP (no IPTG). D) MJF645. E) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). F) Structures of L-methyl esters OXFKEF12–14 (**31**–**33**) and D-methyl ester OXFKEF16 (**48**).

3.6.2.2 Assays for *E. coli* Frag 1 Expressing *Yp*KefGBP

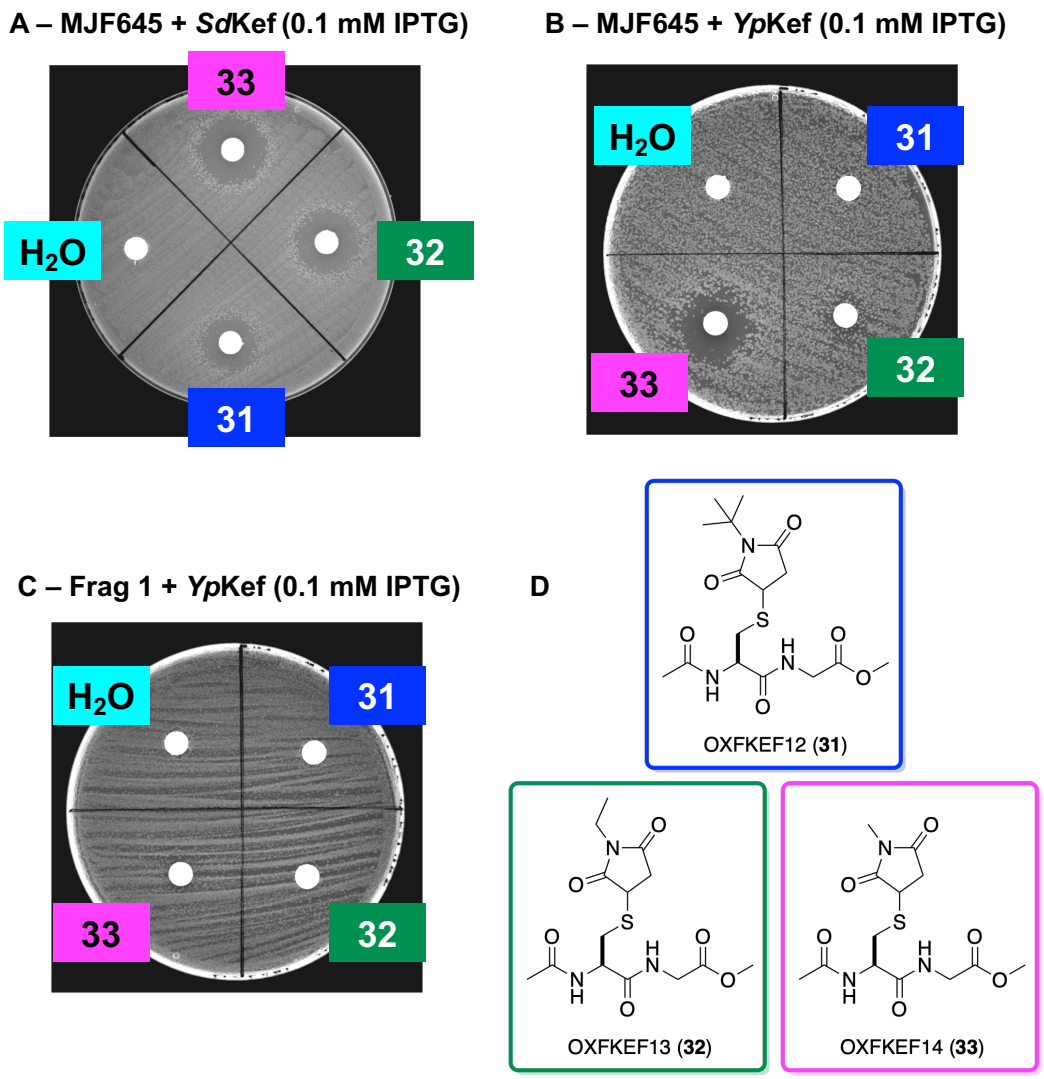


Figure 3.10. Kirby-Bauer disc diffusion assay to test L-methyl esters OXFKEF12–14 (31–33) in *E. coli* Frag 1 expressing *Yp*KefGBP. 31 = OXFKEF12 (**31**), **32** = OXFKEF13 (**32**), and **33** = OXFKEF14 (**33**). All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (positive control, 0.1 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.1 mM IPTG). C) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). D) Structures of L-methyl esters OXFKEF12–14 (**31**–**33**).

The L-methyl esters OXFKEF12–14 (**31**–**33**) were tested in Frag 1 expressing *Yp*KefGBP. Frag 1 is a wild-type *E. coli* strain, containing all four *kef* genes (*kefB*, *kefC*, *kefF*, and *kefG*) and the gene for glutathione biosynthesis (*gshA*). OXFKEF12–14 (**31**–**33**) were also tested in MJF645 expressing *Yp*KefGBP. The two cell strains, Frag 1 and MJF645, were used to compare the function of *Yp*KefGBP in the presence and absence of GSH

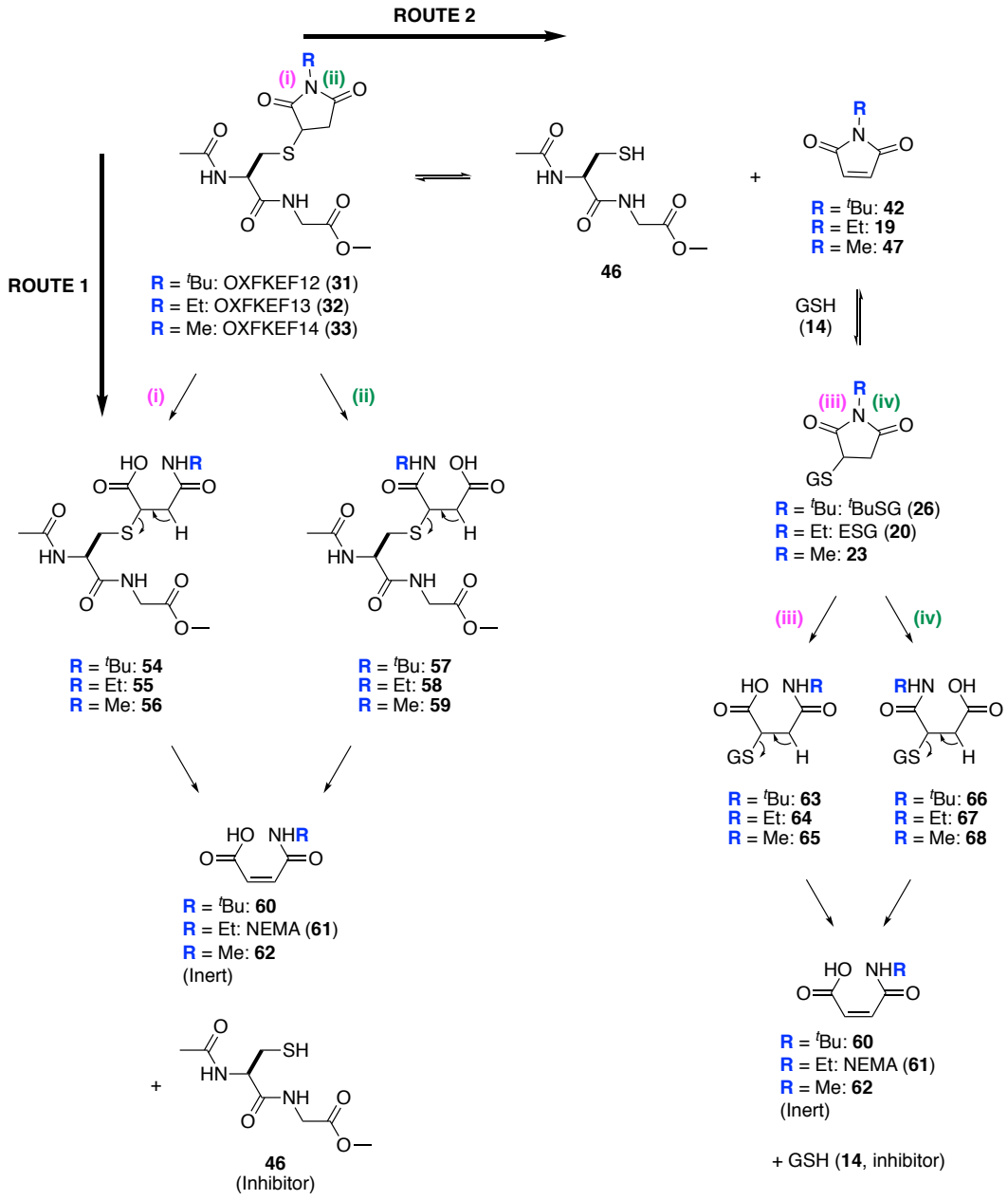
(**14**), respectively. As for the previous experiments, MJF645 expressing *Sd*Kef was included as a positive control. This positive control was successful, as OXFKEF12–14 (**31**–**33**) activated MJF645 expressing *Sd*Kef (Figure 3.10A). Both OXFKEF13 (**32**) and OXFKEF14 (**33**) activated MJF645 expressing *Yp*KefGBP (Figure 3.10B). In contrast, OXFKEF12–14 (**31**–**33**) were inactive in Frag 1 expressing *Yp*KefGBP (Figure 3.10C).

Based on the difference in activity of compounds **32** and **33** in MJF645 and Frag 1 expressing *Yp*KefGBP, several hypotheses were proposed. In the Frag 1 transformants, GSH (**14**), which is not present in MJF645 due to the lack of *gshA*, the gene for GSH (**14**) synthesis (Table 3.3), could be acting as an inhibitor of *Yp*KefB and/or participating in the degradation of the L-methyl esters OXFKEF12–14 (**31**–**33**), to give products which inhibit *Yp*KefB or are inert with respect to this Kef system. It is also possible that *Yp*KefB has lower expression levels in Frag 1 or that the protein is incorrectly folded. The disc assay does not provide direct evidence for the production of *Yp*KefGBP in MJF645 or Frag 1. Production of *Yp*KefGBP in both cell strains was inferred from resistance of the *E. coli* transformants to Amp, as the pTrcYpKefGBP plasmid contains a *bla* gene for β-lactamase (Figure 3.6A).

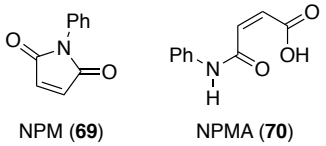
Table 3.3. *E. coli* strains used in the disc assay.

***E. coli* Strain**	**Genotype**
MJF645	MG1655, *kefF*, *kefC*, *kefG*, *kefB*, *gshA*
Frag 1	F^-, *thi*, *rha*, *lacZ*

There are several possible mechanisms by which the L-methyl esters OXFKEF12–14 (**31**–**33**) could be degraded in GSH-containing *E. coli* Frag 1. The first mechanism (Scheme 3.3) is based on a mechanism for the GSH-based detoxification of NEM (**19**) and *N*-phenylmaleimide (NPM, **69**, Figure 3.11) as reported in a study by Booth *et al*.[29] This mechanism is predicted to require an imidase enzyme with broad specificity for the hydrolysis of succinimide rings.[29] The mechanism for OXFKEF12–14 (**31**–**33**) includes two possible routes (Scheme 3.3). In route 1, the succinimide ring in OXFKEF12–14 (**31**–**33**) could hydrolyse, which would give two sets of species, **54**–**56** and **57**–**59**, due to hydrolysis of the succinimide on either side ((i) and (ii), Scheme 3.3). Elimination from both **54**–**56** and **57**–**59** would give the parent thiol **46**, which is predicted to inhibit *Yp*KefB, and maleamic acids **60**, **61**, and **62**, which are predicted to be inert with respect to *Yp*KefGBP-expressing *E. coli*. In studies by Booth *et al*., *N*-ethylmaleamic acid (NEMA, **61**, Scheme 3.3) and *N*-phenylmaleamic acid (NPMA, **70**, Figure 3.11) were shown to be non-toxic to *E. coli*.[29,34] In route 2, L-methyl esters OXFKEF12–14 (**31**, **32**, and **33**) are expected to be in equilibrium with the thiol **46** and the maleimides (**42**, **19**, and **47**) (Scheme 3.3), due to the reversibility of this conjugate addition.[29] The GSH (**14**) in Frag 1 could then react with the maleimides (**42**, **19**, and **47**) to give the GSX (**26**, **20**, and **23**). Finally, the succinimide groups in **26**, **20**, and **23** could then be hydrolysed to give intermediates **63**–**65** and **66**–**68**. Elimination would give the putatively non-toxic maleamic acids (**60**–**62**) and GSH (**14**), which is hypothesised to inhibit *Yp*KefB.



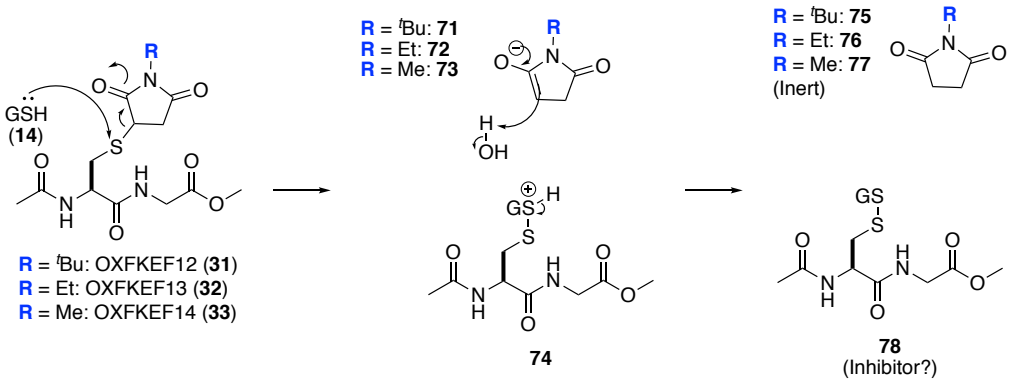
Scheme 3.3. First proposed mechanism for the degradation of L-methyl esters OXFKEF12–14 (31–33) by GSH (14). This mechanism is proposed based on a previous study by Booth *et al.*[58] Mechanism occurs by 2 possible routes: Route 1 or route 2. Route 1 occurs *via* hydrolysis of OXFKEF12–14 (**31**–**33**) to **54–56** and **57–59**. Route 2 occurs *via retro*-conjugate addition of OXFKEF12–14 (**31**–**33**) to the parent thiol (**46**) and the maleimides (**42**, **19**, and **47**).



NPM (**69**) NPMA (**70**)

Figure 3.11. Structures of *N*-phenylmaleimide (NPM, 69) and *N*-phenylmaleamic acid (NPMA, 70).

In the second possible mechanism (Scheme 3.4) GSH (**14**) could attack at the S atom in OXFKEF12–14 (**31**–**33**), which would give species **71**–**73** and **74**. After proton transfer, this would give the succinimides **75**–**77** and the disulfide **78**. The succinimides **75**–**77** are predicted to be inert, due to their lack of conjugate acceptor functionality. The disulfide **78** is predicted to inhibit *Yp*KefB, due to the flexibility of its S-substituent.



Scheme 3.4. Second proposed mechanism for the degradation of L-methyl esters OXFKEF12–14 (31–33) by GSH (14). Mechanism occurs by attack of GSH (**14**) at the S atom of OXFKEF12–14 (**31**–**33**).

Of the three possible factors for reduced compound activity in Frag 1 expressing *Yp*KefGBP versus MJF645, we decided to investigate the expression levels of *Yp*KefGBP in MJF645 and Frag 1 first. This work will be described at the end of the chapter.

3.7 Determination of a Minimum Inhibitory Concentration (MIC) for *E. coli* MJF645 Expressing *Yp*KefGBP

OXFKEF14 (**33**, Figure 3.12A), which was inferred to be the most active compound in MJF645 expressing *Yp*KefGBP in the disc assay, was subsequently tested against the same transformant in a MIC assay. MJF645 expressing *Yp*KefGBP was grown in the presence of OXFKEF14 (**33**) at varying concentrations, starting from 1 mg mL^{-1} (1000 µg mL^{-1}), which was serially diluted twofold to the lowest concentration of 3.91 µg mL^{-1} (Figure 3.12B). The cell culture of MJF645 expressing *Yp*KefGBP was used

as a negative control. After overnight incubation at 37 °C, the OD_{600} (optical density (OD) measured using light with a wavelength (λ) of 600 nm) values of the negative control wells and the compound wells were measured. The MIC was defined as the lowest concentration of compound for which the bacterial growth was inhibited such that the mean OD_{600} was less than 20% of the mean OD_{600} of the negative control wells. For this transformant, the MIC of compound **33** was 15.6 µg mL^{-1}, which is high for a conventional antibiotic.

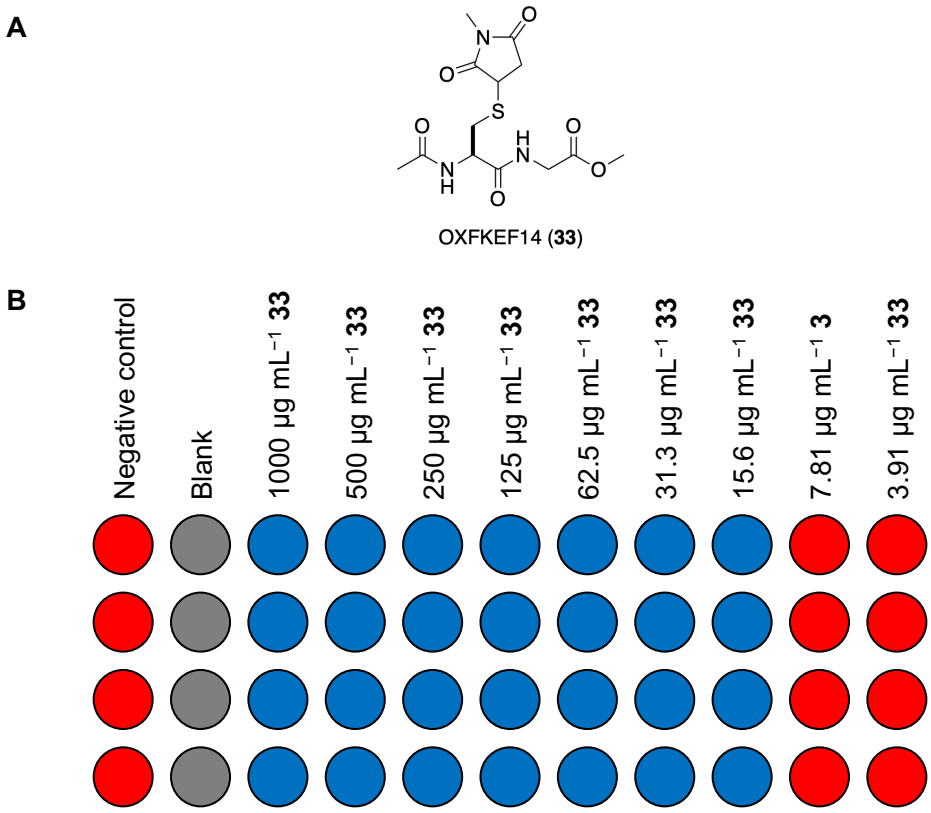


Figure 3.12. Determination of a MIC for OXFKEF14 (33) in *E. coli* MJF645 expressing *Yp*KefGBP. A) Structure of compound OXFKEF14 (**33**). B) Schematic of MIC assay plate. Determination of a MIC for OXFKEF14 (**33**) in *E. coli* MJF645 expressing *Yp*KefGBP. Blank: 2× TY Medium containing 0.2 mg mL^{-1} Amp and 0.2 mM IPTG. Negative control: Cell culture of *E. coli* MJF645 expressing *Yp*KefGBP (using pTrcYpKefGBP). Negative control wells contain (final concentrations) 0.1 mg mL^{-1} Amp and 0.1 mM IPTG. The same culture of *E. coli* MJF645 expressing *Yp*KefGBP was grown in the presence of 0.1 mg mL^{-1} Amp, 0.1 mM IPTG, and OXFKEF14 (**33**) at varying concentrations (as indicated). Compound (OXFKEF14, **33**) wells which showed growth such that mean OD_{600} > 20% (mean OD_{600} of the negative control wells) are shown in red. Wells which showed growth inhibition/death such that mean OD_{600} < 20% (mean OD_{600} of the negative control wells) are shown in blue. Blank wells are shown in grey.

As OXFKEF14 (**33**) caused growth inhibition/death of MJF645 expressing *Yp*KefGBP in the MIC assay, a minimum bactericidal concentration (MBC) assay was carried out subsequently. Aliquots of cell culture from some of the wells showing growth inhibition/death in the MIC assay (15.6, 31.3, 62.5, 125, and 250 μg mL^{-1} compound **33**) were subcultured into fresh medium and incubated. Cell culture from the 7.81 μg mL^{-1} compound well, which did not show growth inhibition/death in the MIC assay, was subcultured as a positive control. 2× TY medium was used as a negative control. All of the cultures which had been subcultured from the MIC assay plate showed visible bacterial growth, indicating that OXFKEF14 (**33**) was bacteriostatic to MJF645 expressing *Yp*KefGBP in the concentration range tested (15.6–250 μg mL^{-1} **33**). This suggested that the K^+ efflux from the cells expressing *Yp*KefGBP inhibited their growth but was non-lethal.

It was important, however, not to overinterpret the results of both the MIC and MBC assays. There are several reasons for this. Firstly, in both assays, there was a lack of a positive control antibiotic. This compound should both inhibit cell growth in the MIC assay and cause cell death in the MBC assay. Secondly, as each assay was carried out once only, more repeats are required to demonstrate reproducibility. Finally, in both assays, the exogenous *Yp*KefGBP protein is being produced in *E. coli* and it is hypothesised that the expression level of *Yp*KefGBP is much higher than in wild-type *Y. pestis*.

3.8 Disc Assay in Wild-Type Bacteria

As OXFKEF14 (**33**, Figure 3.13F) showed activity in *E. coli* MJF645 expressing *Yp*KefGBP in the disc assay (described in 3.6.2.1), the compound was tested in wild-type *Y. pestis*, to determine whether activation of *Yp*KefGBP in the wild-type bacteria would inhibit bacterial growth. OXFKEF14 (**33**) was also tested in wild-type *Yersinia pseudotuberculosis* (*Y. pseudotuberculosis*, *Ys*), which is used as a model organism for *Y. pestis* due to their genetic similarity[59] and its lower biosafety level compared to *Y. pestis*.

A DELTA-BLAST search of the NCBI database, using full-length *Yp*KefB as the query sequence, revealed a single Kef system in *Y. pseudotuberculosis*, which was denoted *Ys*Kef. This Kef system showed 99% sequence homology to *Yp*KefB (Table 3.4), validating *Y. pseudotuberculosis* as an appropriate model for *Y. pestis* in the disc assay.

Table 3.4. Summary of the search for *Yp*KefB homologues in *Y. pseudotuberculosis*. The aa sequence of *Yp*KefB (GenBank ID: AAM87516.1) was used as a query sequence for this NCBI protein BLAST search.[55] The BLAST analysis was performed to search against the nr protein sequences database using the DELTA-BLAST algorithm.[56] The genome of *Y. pseudotuberculosis* (taxid: 633) was selected. Identity (%): Extent to which amino acid sequences of the Kef system and *Yp*KefB query have the same residues at identical positions in the alignment, expressed as a percentage. Positives (%): Extent to which amino acid sequences of the Kef system and *Yp*KefB query have residues with similar physicochemical properties at identical positions in the alignment, expressed as a percentage. Gaps (%): Spaces introduced in the alignment of the amino acid sequences of the Kef system and *Yp*KefB query to account for insertions and deletions in each sequence compared to the other, expressed as a percentage.[39]

Species	Kef	Identity (%)	Positives (%)	Gaps (%)
Y. pseudotuberculosis	*Ys*Kef	99	99	0

OXFKEF14 (**33**) was next tested in wild-type *Y. pseudotuberculosis* using the disc assay. *E. coli* MJF645 expressing *Yp*KefGBP was included as a positive control (Figure 3.13A). *E. coli* MJF645 expressing either *Bp*Kef1 or *Bp*Kef2 were used as negative controls (Figure 3.13B and C). OXFKEF14 (**33**) was active in MJF645 expressing *Yp*KefGBP (Figure 3.13A) and inactive in MJF645 expressing either *Bp*Kef1 or *Bp*Kef2 (Figure 3.13B and C, respectively), consistent with the previous experiments (described in 3.6). OXFKEF14 (**33**) was inactive in wild-type *Y. pseudotuberculosis* (Figure 3.13D).

OXFKEF14 (**33**) was also tested in wild-type *Y. pestis* using the disc assay. Due to the higher biosafety level of *Y. pestis* compared to *Y. pseudotuberculosis*, the disc assay for *Y. pestis* was carried out by Dr Andrew E. Scott (DSTL). Notably, OXFKEF14 (**33**) was inactive in wild-type *Y. pestis* (Figure 3.13E).

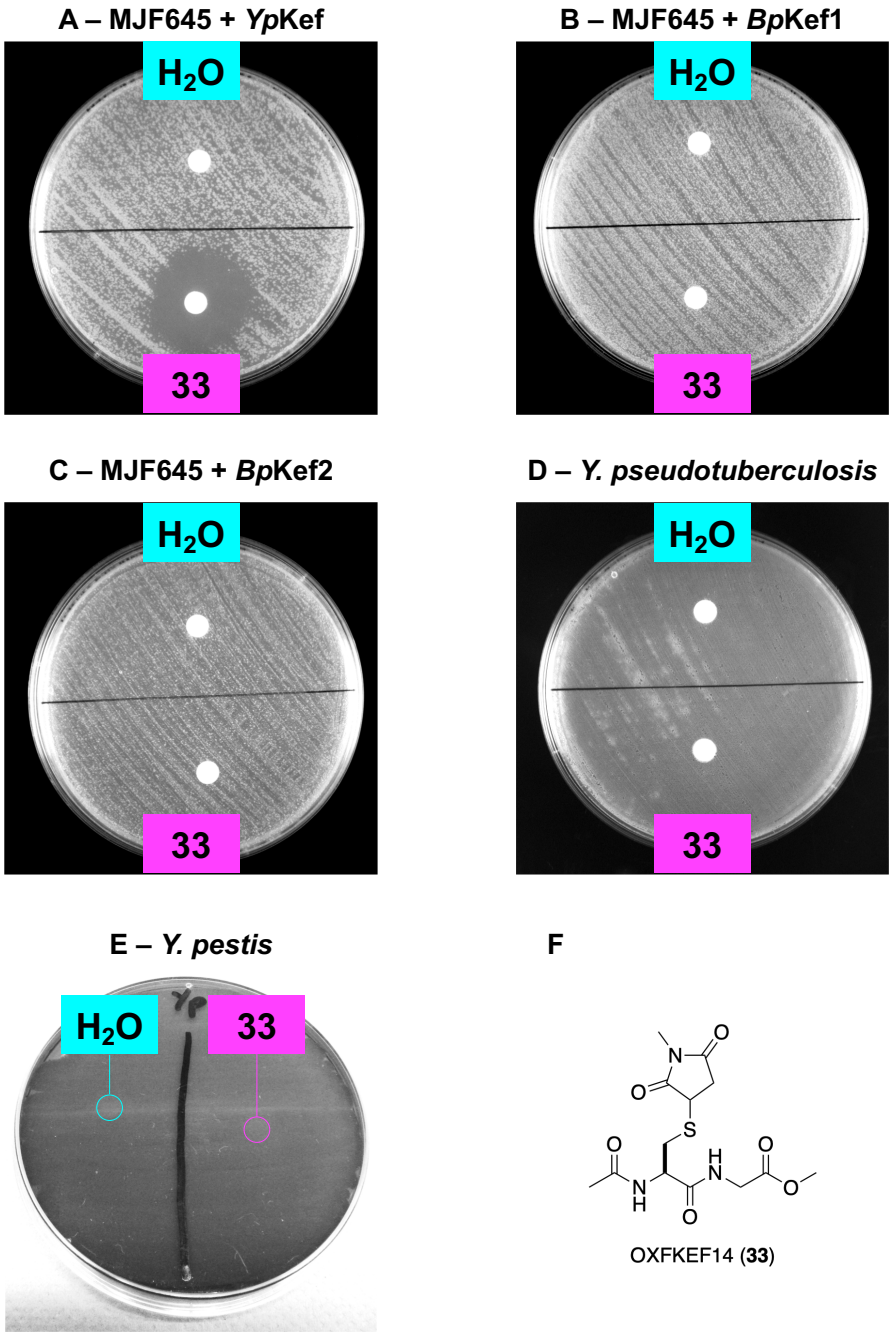


Figure 3.13. Kirby-Bauer disc diffusion assay for OXFKEF14 (33) in wild-type *Y. pseudotuberculosis* and *Y. pestis*. **33** = OXFKEF14 (**33**). OXFKEF14 (**33**) was tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcYpKefGBP (positive control, 0.1 mM IPTG); B) MJF645 + pTrcBpKef1H_6 (negative control, 0.1 mM IPTG); C) MJF645 + pTrcBpKef2H_6 (negative control, 0.1 mM IPTG); D) *Y. pseudotuberculosis*; E) *Y. pestis*. Data from Dr Andrew Scott (unpublished). The positions of the two compound discs on the agar are highlighted. F) Structure of OXFKEF14 (**33**).

The disc assay data acquired in wild-type *Y. pseudotuberculosis* and *Y. pestis* imply that, with the currently available compound (OXFKEF14, **33**), activation of Kef

expressed at wild-type levels is not sufficient to cause growth inhibition. These data could also suggest that OXFKEF14 (**33**) is not cell permeant to *Y. pseudotuberculosis* and *Y. pestis*, or that the compound is being rapidly effluxed by the bacteria. OXFKEF14 (**33**) may not be being hydrolysed to the active carboxylic acid, OXFKEF19 (**39**, Figure 3.14), if the esterase(s) required are not present in these bacterial species. OXFKEF14 (**33**) could also be being metabolised by the bacteria to give compound(s) which inhibit *Yp*KefB or are inert.

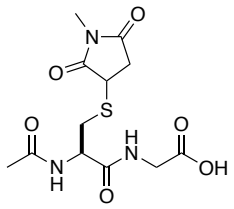


Figure 3.14. Structure of OXFKEF19 (39).

3.9 Determination of the Expression Level of *Yp*KefBH$_6$ in *E. coli* MJF645 and Frag 1

To elucidate the contrasting disc assay results for OXFKEF12–14 (**31**–**33**) in *Yp*KefGBP-expressing MJF645 and Frag 1 (Figure 3.10), we wanted to compare the expression level of *Yp*KefGBP in the two *E. coli* strains. We chose to determine the protein expression levels using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of the protein samples from MJF645 and Frag 1 producing *Yp*KefGBP, followed by western blotting. In the *Yp*KefGBP protein complex, *Yp*KefB was the protein subunit we wanted to study specifically as it is hypothesised to be responsible for K^+ efflux. To detect *Yp*KefB, it was desirable to blot for a His tag as an antibody for this tag is readily available. As the *Yp*KefB construct encoded by the pTrcYpKefGBP plasmid did not contain a His tag, it was necessary to install the codons

for this tag in the existing plasmid. The Kirby-Bauer disc diffusion assay would confirm if *Yp*KefBH$_6$ functions in a similar way to *Yp*KefB. The expression level of *Yp*KefBH$_6$ in MJF645 and Frag 1 would then be analysed by SDS-PAGE and western blotting.

3.9.1 Molecular Cloning to Produce pTrcYpKefG(BH$_6$)P

The His tag was installed at the C-terminus of *Yp*KefB. A modification in this location was hypothesised not to interfere with protein function, as it is 50 amino acid residues away from the last proposed ligand binding residue (D552) in the primary protein sequence. The novel protein construct, denoted *Yp*KefG(BH$_6$)P, was encoded by the pTrcYpKefG(BH$_6$)P plasmid (Figure 3.15A). This plasmid was made efficiently from the pTrcYpKefGBP vector (Figure 3.15B) using molecular cloning techniques, which is described in section 6.3.2 in the experimental.

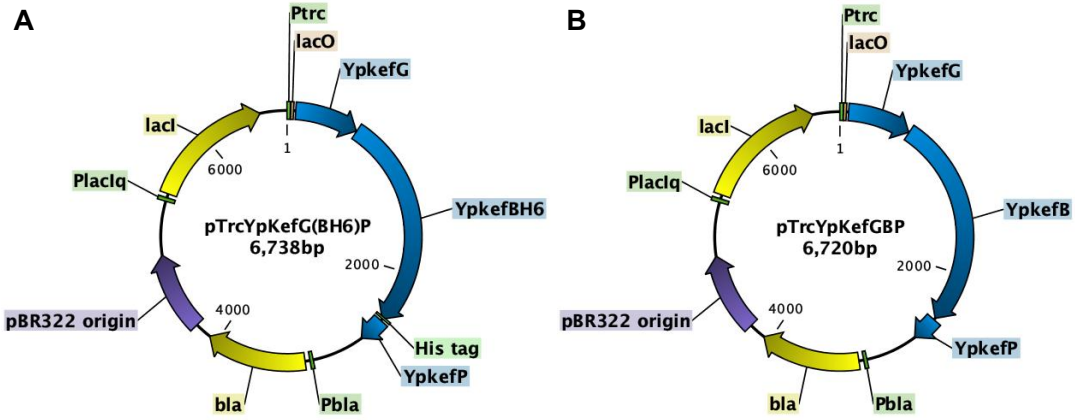


Figure 3.15. Plasmid maps for pTrcYpKefG(BH$_6$)P (A) and pTrcYpKefGBP (B). The *YpkefG*, *YpkefBH$_6$*, *YpkefB*, and *YpkefP* genes code for the *Yp*KefG, *Yp*KefBH$_6$, *Yp*KefB, and *Yp*KefP proteins, respectively. Ptrc is the promoter for the *YpkefG*, *YpkefBH$_6$*, *YpkefB*, and *YpkefP* genes. *bla* is the gene for β-lactamase, which causes resistance to Amp. Pbla is the promoter for the gene *bla*. pBR322 origin is the origin of replication, which ensures that the plasmid is replicated when it is transformed into *E. coli*. *lacI* is the gene for LacI, which prevents transcription of the *YpkefG*, *YpkefBH$_6$*, *YpkefB*, and *YpkefP* genes by binding to LacO. IPTG binds to LacI, preventing LacI from binding to LacO, promoting transcription of the *YpkefG*, *YpkefBH$_6$*, *YpkefB*, and *YpkefP* genes. PlacIq is the promoter for the gene *lacI*.

3.9.2 Disc Assay for *E. coli* MJF645 and Frag 1 Expressing *Yp*KefG(BH_6)P

To determine whether *Yp*KefBH_6 functioned similarly to the wild-type *Yp*KefB protein, L-methyl esters OXFKEF12–14 (**31**–**33**) were tested in MJF645 and Frag 1 transformed with pTrcYpKefGBP and pTrcYpKefG(BH_6)P using the disc assay. MJF645 and Frag 1 transformed with the non-tagged pTrcYpKefGBP were included as positive and negative controls, respectively. OXFKEF13 (**32**) and OXFKEF14 (**33**) were active in *Yp*KefGBP-expressing MJF645 (Figure 3.16A) but were inactive in Frag 1 in the same experiment (Figure 3.16B). OXFKEF13 (**32**) and OXFKEF14 (**33**) were active in MJF645 expressing the tagged *Yp*KefG(BH_6)P protein complex (Figure 3.16C). Conversely, the compounds were inactive in Frag 1 expressing the same protein (Figure 3.16D). Overall, these results showed that *Yp*KefBH_6 functioned similarly to wild-type *Yp*KefB in the disc assay. This was consistent with our hypothesis that the addition of a His tag at the C-terminus of *Yp*KefB would not interfere with its function.

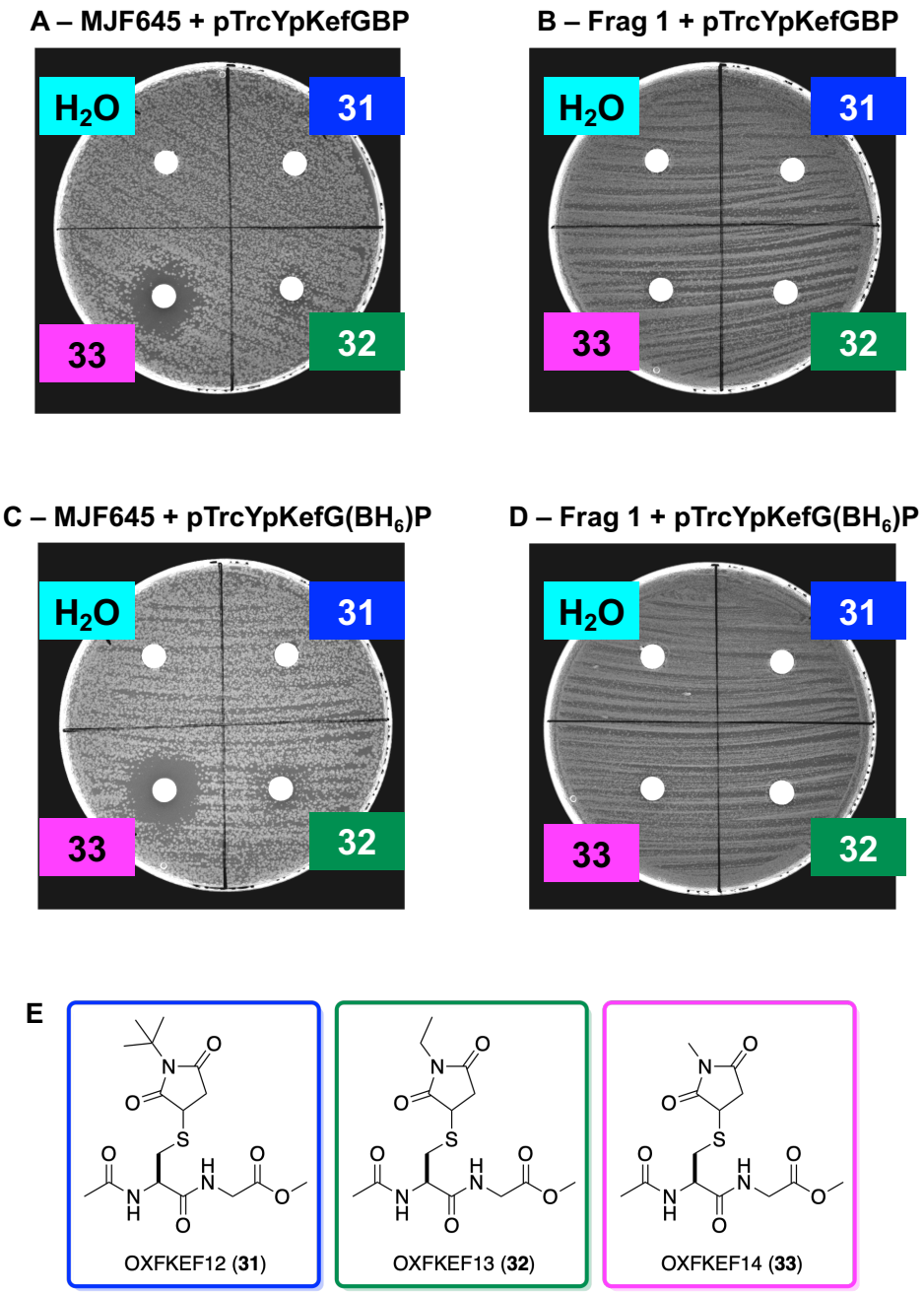


Figure 3.16. Kirby-Bauer disc diffusion assay for OXFKEF12–14 (31–33) in *E. coli* MJF645 and Frag 1 expressing *Yp*KefGBP and *Yp*KefG(BH$_6$)P. 31 = OXFKEF12 (**31**), **32** = OXFKEF13 (**32**), and **33** = OXFKEF14 (**33**). All compounds (OXFKEF12–14, **31–33**) were tested at 15 mM in H_2O. H_2O was tested as a negative control. A) MJF645 + pTrcYpKefGBP (0.1 mM IPTG). B) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). C) MJF645 + pTrcYpKefG(BH$_6$)P (0.1 mM IPTG). D) Frag 1 + pTrcYpKefG(BH$_6$)P (0.1 mM IPTG). E) Structures of L-methyl esters OXFKEF12–14 (**31–33**).

3.9.3 SDS-PAGE/Western Blotting of *Yp*KefBH$_6$ in *E. coli* MJF645 and Frag 1

*Yp*KefG(BH$_6$)P was expressed in both *E. coli* MJF645 and Frag 1 and the resulting protein samples were analysed using SDS-PAGE and western blotting. Initially, a protocol for the production, SDS-PAGE, and western blotting of full-length Kef proteins, including *Sd*Kef, *Ec*KefFC, and *Ec*KefGB, was used.[37,60] The cells were grown to an OD of 0.4, then induced (0.3 mM IPTG, 30 min). After cell lysis (phosphate-buffered saline (PBS) buffer) and centrifugation, ultracentrifugation was used to isolate the membrane protein fractions, which were resuspended in PBS buffer. The supernatants from the ultracentrifugation were also retained for analysis to check if they contained any of the desired *Yp*KefBH$_6$ protein.

The protein concentrations of the membrane fractions were estimated using a Pierce™ 660 nm protein assay. In this assay, a dye is added to the protein samples, which then reacts with the protein(s) in the samples. This causes a change in absorbance at 660 nm, which is measured using spectrophotometric methods. The absorbance change is used to estimate the protein concentration, using a standard curve for a model protein. The standard curve (Figure 3.17) was generated using bovine serum albumin (BSA) samples of concentrations in the range of 50–2000 μg mL^{-1}, which is the working range of the assay. The absorbance values of the samples of *Yp*KefBH$_6$, expressed in both MJF645 and Frag 1, were measured. These absorbances corresponded to very low concentrations: membrane fraction (MF) from MJF645 = 113 μg mL^{-1}, supernatant (SN) from MJF645 = 223 μg mL^{-1}, Frag 1 MF = 85 μg mL^{-1}, and Frag 1 SN = 200 μg mL^{-1} (Table 3.5). As the well volume for the SDS-PAGE gels was

restricted to 20 µL, we were unable to load the recommended 20 µg of protein as per the literature protocol (Table 3.5).[60]

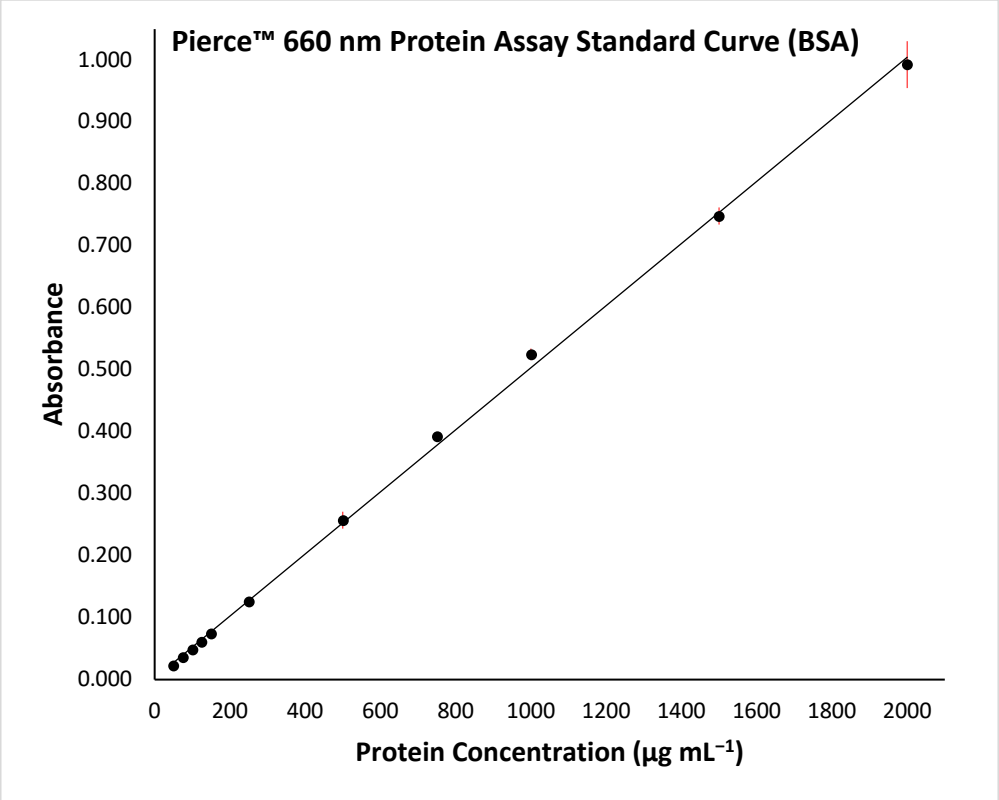


Figure 3.17. Standard curve for Pierce™ 660 nm protein assay, generated using BSA. The standard curve was constructed by measuring the absorbance values of commercial BSA standards of known concentrations. Each data point is the arithmetic mean of the absorbance values of three different assay samples (technical triplicate). Error bars (in red) are ± 1 standard deviation (s.d.).

Table 3.5. Pierce™ 660 nm protein assay 1: Estimation of the concentration of *Yp*KefBH$_6$ protein samples (in PBS buffer). MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from either *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). All of the volumes required are much larger than the well volume (20 µL) on the SDS-PAGE gels.

Sample	Protein Concentration (µg mL^{-1})	Volume for 20 µg Protein (µL)
MJF645 MF	113	178
MJF645 SN	223	90
Frag 1 MF	85	236
Frag 1 SN	200	100

This protocol was therefore adapted to produce protein samples of higher concentration. The cells were grown to a higher OD (0.5) before being induced

overnight (0.3 mM IPTG). Lysis, centrifugation, and ultracentrifugation were carried out according to the previous procedure. The isolated membrane fractions were insoluble in PBS buffer, and so the samples were resuspended in PBS buffer containing 20% *v*/*v* glycerol, as this was predicted to improve the solubility of the membrane protein. The concentrations of these samples were estimated using the Pierce™ assay: MJF645 MF = 3300 μg mL^{-1}, MJF645 SN = 2470 μg mL^{-1}, Frag 1 MF = 1620 μg mL^{-1}, and Frag 1 SN = 1210 μg mL^{-1} (Table 3.6). These concentrations were higher than the ones from the previous experiment (Table 3.5). Solubility was still an issue, however, as precipitation occurred during the Pierce™ assay in the wells for the membrane fractions. The concentrations of these samples were therefore likely to be inaccurate. The concentration value for MJF645 SN (2470 μg mL^{-1}) was also considered to be inaccurate, as it was outside the assay working range (50–2000 μg mL^{-1}).

Table 3.6. Pierce™ 660 nm protein assay 2: Estimation of the concentration of *Yp*KefBH_6 protein samples (in PBS buffer containing 20% *v*/*v* glycerol). MJF645, Frag 1 = *Yp*KefBH_6 Protein produced from either *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH_6)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). The data for the membrane fraction samples (highlighted in red) were considered to be inaccurate due to precipitation in the assay.

Sample	Protein Concentration (μg mL^{-1})	Volume for 20 μg Protein (μL)
MJF645 MF	3300	6
MJF645 SN	2470	8
Frag 1 MF	1620	12
Frag 1 SN	1210	17

All of the protein samples were analysed using SDS-PAGE (see appendix). As the protein concentrations were considered to be inaccurate, equal volumes of each sample were loaded onto the gel. The membrane protein samples were poorly soluble during SDS-PAGE, as some of the sample precipitated in the stacking gel. Western blotting (Figure 3.18) confirmed that *Yp*KefBH_6 was being produced in both MJF645

and Frag 1, as the protein was present in both membrane fraction samples. The western blot also showed that some of the *Yp*KefBH$_6$ protein was soluble on the SDS-PAGE gel, as bands corresponding to the molecular weight of *Yp*KefBH$_6$ (~66 kDa) were visible in the lower region of the blot (Figure 3.18), corresponding to the resolving gel. A large proportion of *Yp*KefBH$_6$ was insoluble, indicated by the presence of bands at the top of the western blot, in the region corresponding to the SDS-PAGE stacking gel. Due to the poor solubility, the expression level of *Yp*KefBH$_6$ in the two *E. coli* strains could not be compared accurately. We therefore investigated a different expression method to improve the solubility of the membrane protein.

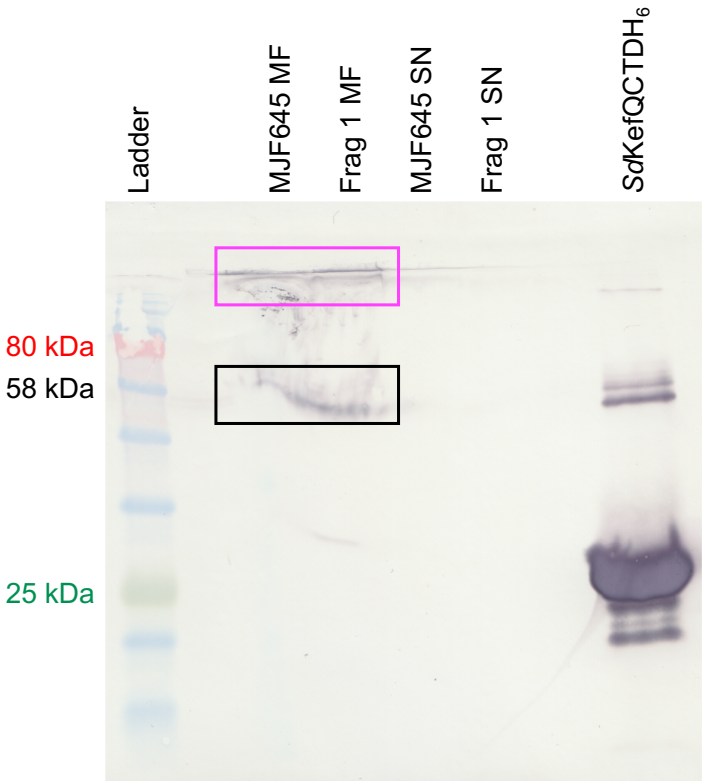


Figure 3.18. Western blot for *Yp*KefBH$_6$ protein samples (suspended in PBS buffer containing 20% *v/v* glycerol) expressed in *E. coli* MJF645 and Frag 1. MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). Bands corresponding to *Yp*KefBH$_6$ (~66 kDa) are highlighted: Soluble (in black) and insoluble (in pink). *Sd*KefQCTDH$_6$ (~26 kDa) was used as both a loading control and a control for the visualisation of the His tag using western blotting.

An expression and purification method for full-length *Sd*KefH$_6$[37] was investigated next for *Yp*KefBH$_6$. The composition of the lysis and solubilisation buffers were different from the previous procedure. Both buffers were Tris-based, whereas the ones from the previous procedure were PBS-based. A non-ionic surfactant (*n*-dodecyl β-D-maltoside (DDM), 1.5% *w/v*) was also included in the solubilisation buffer to improve the solubility of the membrane protein.[37] The *E. coli* cultures were grown, and cell pellets isolated in the same way as described in the previous section. However, this time the cell pellets were lysed into Tris buffer, and the membrane fractions isolated using centrifugation and ultracentrifugation. Samples of each supernatant from ultracentrifugation were retained for analysis. The membrane protein samples were resuspended in solubilisation buffer (Tris buffer containing 1.5% *w*/*v* DDM). The concentrations of the membrane samples produced from MJF645 and Frag 1 were assumed to be similar, as the ratio of solubilisation buffer to the mass of membrane was constant for both samples. SDS-PAGE and western blotting were carried out using the same conditions as described above. Equal volumes of the membrane and supernatant samples were loaded onto the SDS-PAGE gel (see appendix). The samples were all soluble during SDS-PAGE, as no precipitation in the stacking gel was observed. Subsequent western blot analysis (Figure 3.19) confirmed the production of *Yp*KefBH$_6$ (~66 kDa) in both *E. coli* strains. All of the *Yp*KefBH$_6$ was detected in the membrane fractions, and not in the supernatants (Figure 3.19). The *Yp*KefBH$_6$ was fully soluble under these conditions, as intense bands were not observed at the top of the western blot, in the region for the stacking gel. A weak band was observed, however, in the Frag 1 membrane fraction lane, which was hypothesised to correspond to undenatured *Yp*KefBH$_6$ in the sample. In both membrane fraction lanes, the bands

thought to correspond to *Yp*KefBH$_6$ were of similar size and intensity (Figure 3.19). Due to constant sample volume and approximately equal sample concentration, as described above, this western blot qualitatively demonstrated that the expression levels of *Yp*KefBH$_6$ were similar in both *E. coli* strains, MJF645 and Frag 1. As the sequences of *Yp*KefBH$_6$ and the untagged *Yp*KefB are essentially identical, it is assumed that the expression levels of the non-His tagged *Yp*KefB are approximately equal in both *Yp*KefGBP-expressing MJF645 and *Yp*KefGBP-expressing Frag 1.

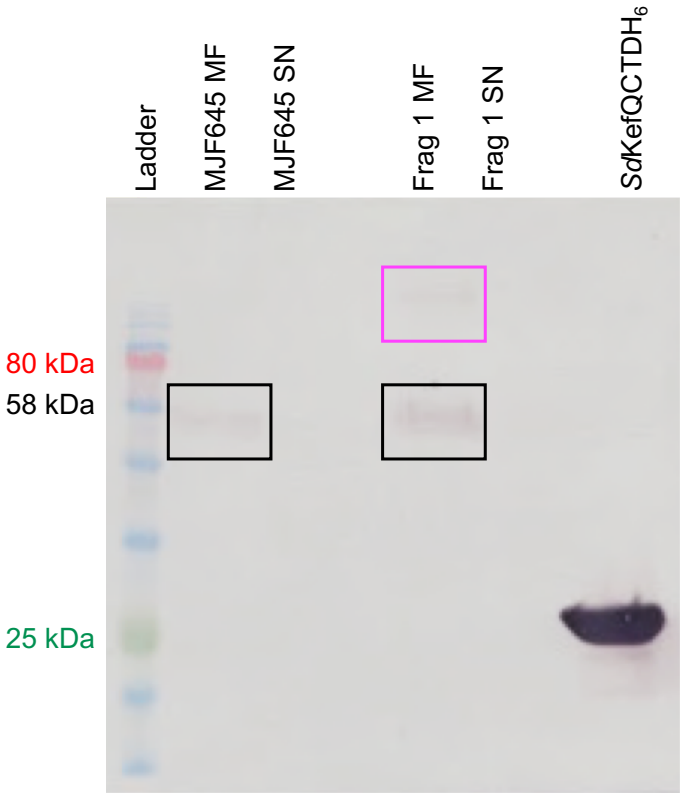


Figure 3.19. Western blot for *Yp*KefBH$_6$ protein samples (suspended in Tris buffer containing 1.5% *w/v* DDM, according to reference[37]) produced in *E. coli* MJF645 and Frag 1. MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). Bands corresponding to *Yp*KefBH$_6$ (~66 kDa) are highlighted: Soluble (in black) and insoluble/undenatured (in pink). *Sd*KefQCTDH$_6$ (~26 kDa) was used as both a loading control and a control for the visualisation of the His tag using western blotting.

3.10 Chapter 3 Conclusions

Homologues of *Sd*Kef were identified in *Y. pestis* (*Yp*KefB) and *B. pseudomallei* (*Bp*Kef1 and *Bp*Kef2) using bioinformatics. L-Methyl esters OXFKEF12–14 (**31**–**33**) were synthesised and tested in *E. coli* expressing *Yp*KefGBP, *Bp*Kef1, or *Bp*Kef2, but only showed activity for *Yp*KefGBP, which has the highest homology score to *Sd*Kef. OXFKEF13 (**32**) and OXFKEF14 (**33**) activated MJF645 expressing *Yp*KefGBP but were inactive in the Frag 1 transformant. MIC and MBC assays showed that OXFKEF14 (**33**) is bacteriostatic in MJF645 expressing *Yp*KefGBP (MIC = 15.6 μg mL^{-1}), meaning that the K^+ efflux elicited from the cells inhibits their growth but is non-lethal. In the disc assay, OXFKEF14 (**33**) is inactive in wild-type *Y. pestis*, presumably due to the lower expression level of *Yp*KefGBP compared to the recombinant system (in *E. coli*). To elucidate the contrasting disc assay results for OXFKEF12–14 (**31**–**33**) in *Yp*KefGBP-expressing MJF645 and Frag 1, western blotting was used to compare the expression level of the *Yp*KefBH_6 model protein in both strains. The expression levels of *Yp*KefBH_6 were shown to be qualitatively similar in MJF645 and Frag 1. Due to the sequence similarity between *Yp*KefBH_6 and *Yp*KefB, it is assumed that the expression level of *Yp*KefB in the two *E. coli* strains is also similar. Based on this assumption, the potential K^+ efflux from MJF645 and Frag 1 expressing *Yp*KefB is approximately equal, and so the contrasting results for OXFKEF12–14 (**31**–**33**) in the disc assay must be caused by another factor. In Frag 1, the endogenous GSH (**14**) could be degrading the compounds or inhibiting *Yp*KefB, or the protein may be incorrectly folded.

Chapter 4: Synthesis and Testing of Thiol-ene-Derived Thioethers as Analogues of OXFKEF12–14

4.1 Introduction and Aims

In Chapter 3, OXFKEF13 (**32**) and OXFKEF14 (**33**) (Figure 4.1) were shown to cause growth inhibition/death in *Yp*KefGBP-expressing MJF645 in the disc assay but were inactive in *Yp*KefGBP-expressing Frag 1. Several parameters were hypothesised to be causing this difference. One hypothesis was that *Yp*KefB was being overexpressed in MJF645 but not in Frag 1. However, we showed that the expression level of *Yp*KefBH_6 was similar in the two *E. coli* strains. Due to the sequence similarity of this protein to *Yp*KefB, it is likely that *Yp*KefB is also expressed to a similar extent in both MJF645 and Frag 1.

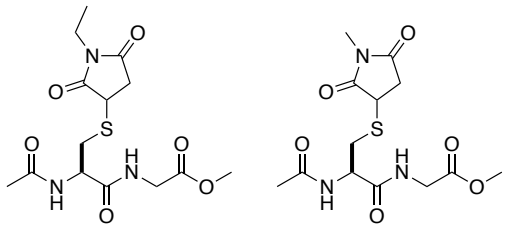


Figure 4.1. L-Methyl esters OXFKEF13 (32) and OXFKEF14 (33).

These data suggest that another factor is causing the difference in compound activity between Frag 1 and MJF645. Frag 1 is a GSH-containing strain whereas MJF645 is a mutant strain lacking GSH (**14**). In Frag 1, GSH (**14**) is likely to be inhibiting *Yp*KefB and so the dipeptides might be unable to outcompete this ligand in the Kef binding site. This could explain the lack of activity observed for L-methyl esters **32** and **33** in Frag 1 expressing *Yp*KefGBP in the disc assay. GSH (**14**) could also be degrading **32** and **33**, giving compounds which either inhibit *Yp*KefB or are inert. We proposed that the S–C

bonds in compounds **32** and **33** could be labile, which would allow the above degradation to occur. We therefore wanted to make analogues of **32** and **33** with inert S–C bonds.

In a previous study by Healy *et al.*, GSX with inert S–C bonds were synthesised using the thiol-ene reaction.[48] Compounds **79** and **40** (Figure 4.2A) are two examples, which are of particular interest as they contain ^{19}F atoms, facilitating the use of ^{19}F NMR spectroscopy to detect ligand binding to Kef. In the study by Healy *et al.*, compounds **79** and **40** were shown to bind *Sd*KefQCTDH$_6$ using both the fluorescence competition assay and DSF.[48] Based on compounds **79** and **40**, we proposed dipeptides **80**, **81**, **82**, and **83** (Figure 4.2B) as analogues of OXFKEF13 (**32**) and OXFKEF14 (**33**) with non-labile S–C bonds.

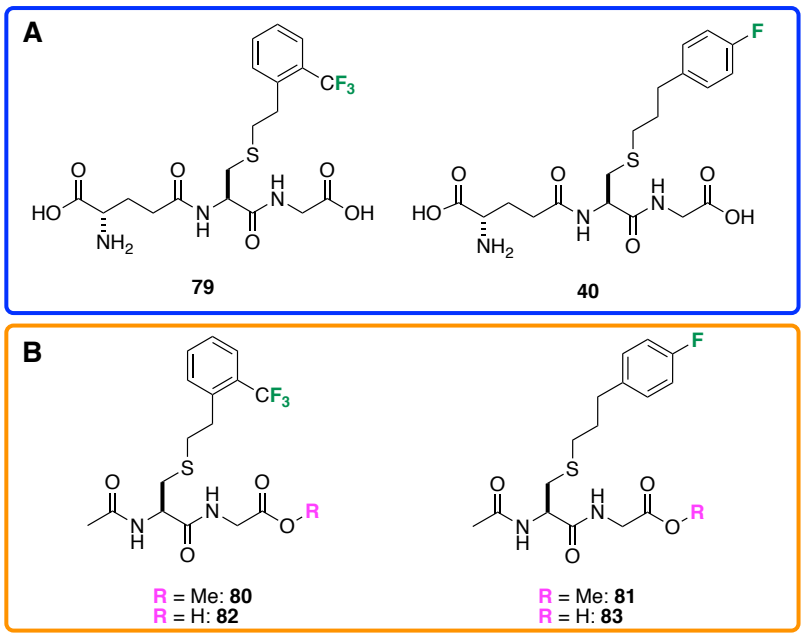
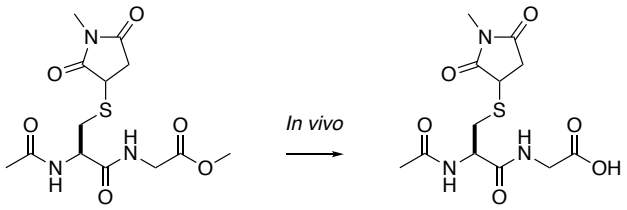


Figure 4.2. Structures of Kef ligands. A) GSX **79** And **40**. B) R = Me: Methyl esters **80** and **81**. R = H: Carboxylic acids **82** and **83**.

The aim of the work described in this chapter was to synthesise compounds **80**, **81**, **82**, and **83** and test them in biological assays to investigate their effect on Kef. The binding

of carboxylic acids **82** and **83** to *Sd*KefQCTDH$_6$ would be detected using two orthogonal biophysical assays: waterLOGSY (described in Chapter 2) and DSF. Methyl esters **80** and **81** (Figure 4.2B) would be tested in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP using the Kirby-Bauer disc diffusion assay, to probe activation of Kef. Compounds **80** and **81** (Figure 4.2B) would be tested in combination with the activator OXFKEF14 (**33**) to test for inhibition. As OXFKEF14 (**33**) is hydrolysed to carboxylic acid OXFKEF19 (**39**) *in vivo* (Scheme 4.1), it was hypothesised that methyl esters **80** and **81** would be hydrolysed to carboxylic acids **82** and **83** (Figure 4.2B) *in vivo*. As compound **40** was predicted to inhibit *Sd*Kef, as shown in Chapter 2, we predicted that its truncated analogue **81** would inhibit *Sd*Kef in the disc assay.



Scheme 4.1. *In vivo* hydrolysis of OXFKEF14 (33) to OXFKEF19 (39).

4.2 Prediction of Compound Solubility Using the Solubility Forecast Index (SFI)

Before synthesising and testing compounds **80**, **81**, **82**, and **83**, their solubility forecast index (SFI) values were calculated to predict solubility. The SFI is calculated using the following formula:

$$\text{SFI} = c\log D_{\text{pH7.4}} + \#\text{Ar}$$

where $c\log D_{\text{pH7.4}}$ is the calculated log*D* at pH 7.4 and #Ar is number of aromatic rings.[61] A SFI < 5 is desirable as this indicates that a compound is likely to be soluble.[61]

Table 4.1. SFI values of compounds 80, 81, 82, and 83. The SFI values of OXFKEF14 (**33**) and OXFKEF19 (**39**) were calculated for comparison. The $c\log D_{\text{pH7.4}}$ values were calculated using the ACD/I-Lab tool (available at http://ilab.cds.rsc.org/?cdsrdr=1).

Compound Class	Compound	$c\log D_{\text{pH7.4}}$	#Ar	SFI
Methyl Esters	OXFKEF14 (**33**)	−0.75	0	−0.75
	80	3.19	1	4.19
	81	2.49	1	3.49
Carboxylic Acids	OXFKEF19 (**39**)	−6.36	0	−6.36
	82	−2.80	1	−1.80
	83	−3.25	1	−2.25

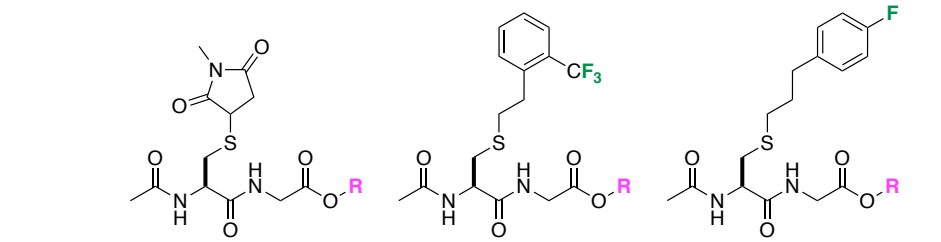
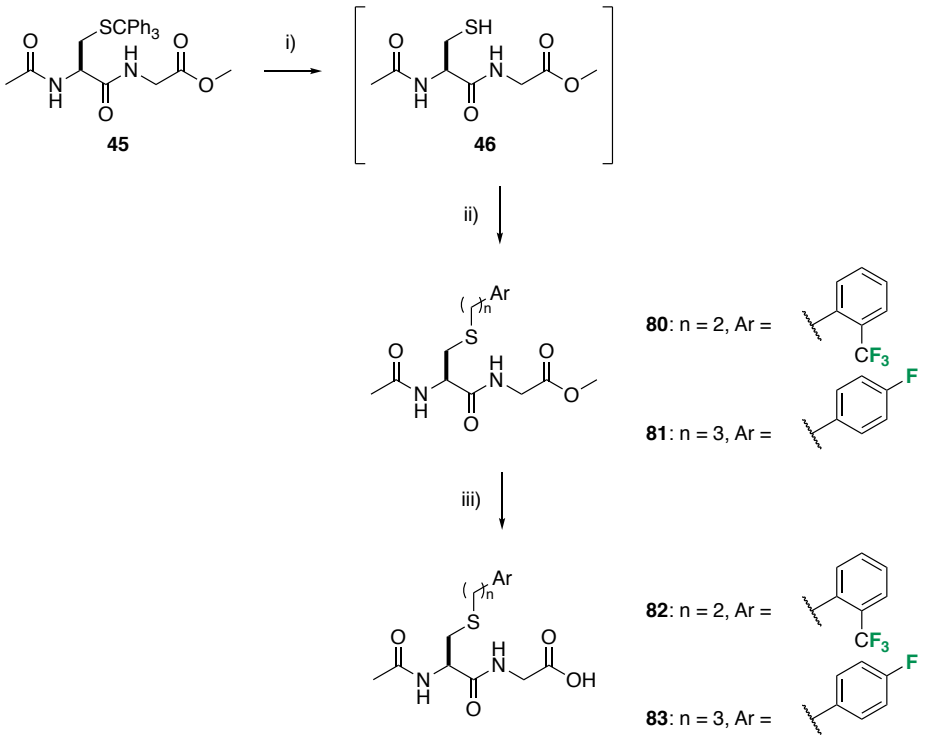


Figure 4.3. Structures of methyl esters OXFKEF14 (33), 80, and 81, and carboxylic acids OXFKEF19 (39), 82, and 83.

The SFI values of compounds **80**, **81**, **82**, and **83** (Figure 4.3) are shown in Table 4.1. The values for OXFKEF14 (**33**) and OXFKEF19 (**39**) (Figure 4.3) were calculated for

comparison. The SFI values of all 6 compounds are < 5, which indicates that they are likely to be soluble. The values for methyl esters **80** and **81** are higher than for their respective acids, **82** and **83**. The values for methyl esters **80** and **81** are also higher than for OXFKEF14 (**33**), as are the values for acids **82** and **83** compared to OXFKEF19 (**39**). These data suggest that replacing the succinimide headgroup with the aromatic thioether decreases the solubility of the compounds. As the SFI values for compounds **80**, **81**, **82**, and **83** were < 5, they were predicted to be soluble. Therefore, we synthesised and tested the compounds.

4.3 Synthesis of Compounds 80, 81, 82, and 83



Scheme 4.2. Synthesis of compounds 80, 81, 82, and 83 from intermediate 45. *Reagents and conditions*: i) Et_3SiH, 20% *v*/*v* TFA in CH_2Cl_2, rt, 20 min–16 h; ii) For **80**: 2-(Trifluoromethyl)styrene, tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl), DPAP, THF:H_2O 1:2, UV light (λ = 365 nm), rt, 5 h, 10–26% (over 2 steps, n = 3); ii) For **81**: 1-Allyl-4-fluorobenzene, TCEP·HCl, DPAP, THF:H_2O 1:2, UV light (λ = 365 nm), rt, 5 h, 12–26% (over 2 steps, n = 2); iii) For **82**: LiOH, THF:MeOH:H_2O 3:2:1, 0 °C, 1.75–2.5 h, 26–77% (n = 4); iii) For **83**: LiOH, THF:MeOH:H_2O 3:2:1, 0 °C, 1.75–2.5 h, 25–79% (n = 3).

Methyl esters **80** and **81** were synthesised from protected dipeptide **45** *via* the thiol **46** (Scheme 4.2). Both intermediates (**45** and **46**) were used in the synthesis of the L-methyl esters OXFKEF12–14 (**31**–**33**) (Scheme 3.1). Thiol-ene reactions between thiol **46** and either 2-(trifluoromethyl)styrene or 1-allyl-4-fluorobenzene gave compounds **80** and **81** in 10–23% and 12–26% yields, respectively (Scheme 4.2). The conditions used for the thiol-ene reactions were reported in the prior study by Healy *et al.*[48] These conditions were the same as used for GSX **40** in Chapter 2 but with the addition of a reducing agent: TCEP·HCl. This reducing agent was predicted to increase the reaction yield by preventing formation of the disulfide of thiol **46**, which was

shown to be the case for the reactions of GSH (**14**) with complex alkenes.[48] Hydrolysis of the methyl esters **80** and **81** under basic conditions (LiOH) gave carboxylic acids **82** and **83** in 26–77% and 25–79% yields, respectively (Scheme 4.2).

All final compounds (**80**, **81**, **82**, and **83**) were of >95% purity, as determined by RP analytical HPLC. The *e.e.* values of these final compounds were calculated using chiral HPLC. In this case, the *e.e.* is the difference between the proportions of the L- and D-enantiomers, calculated using the peak areas from chiral HPLC. The chiral HPLC traces are shown in Figure 4.4 (**80** and **81**) and Figure 4.5 (**82** and **83**). The stereochemistry of **81** (Figure 4.4C) was referenced to a sample of racemic compound (**84**, Figure 4.4B). The trace for the carboxylic acid **83** (Figure 4.5C) was compared to the trace for the racemic compound **85** (Figure 4.5B). The synthesis of both racemic compounds (**84** and **85**) is described in the experimental (section 6.2). The *e.e.* of intermediate **45** (77%) was approximately conserved in the final compounds, which had *e.e.* values as follows: 80% (**80**), 65% (**81**), 76% (**82**), and 72% (**83**). These data suggest that the thiol-ene reactions and ester hydrolyses do not cause substantial epimerisation of the Cys-α stereogenic centre.

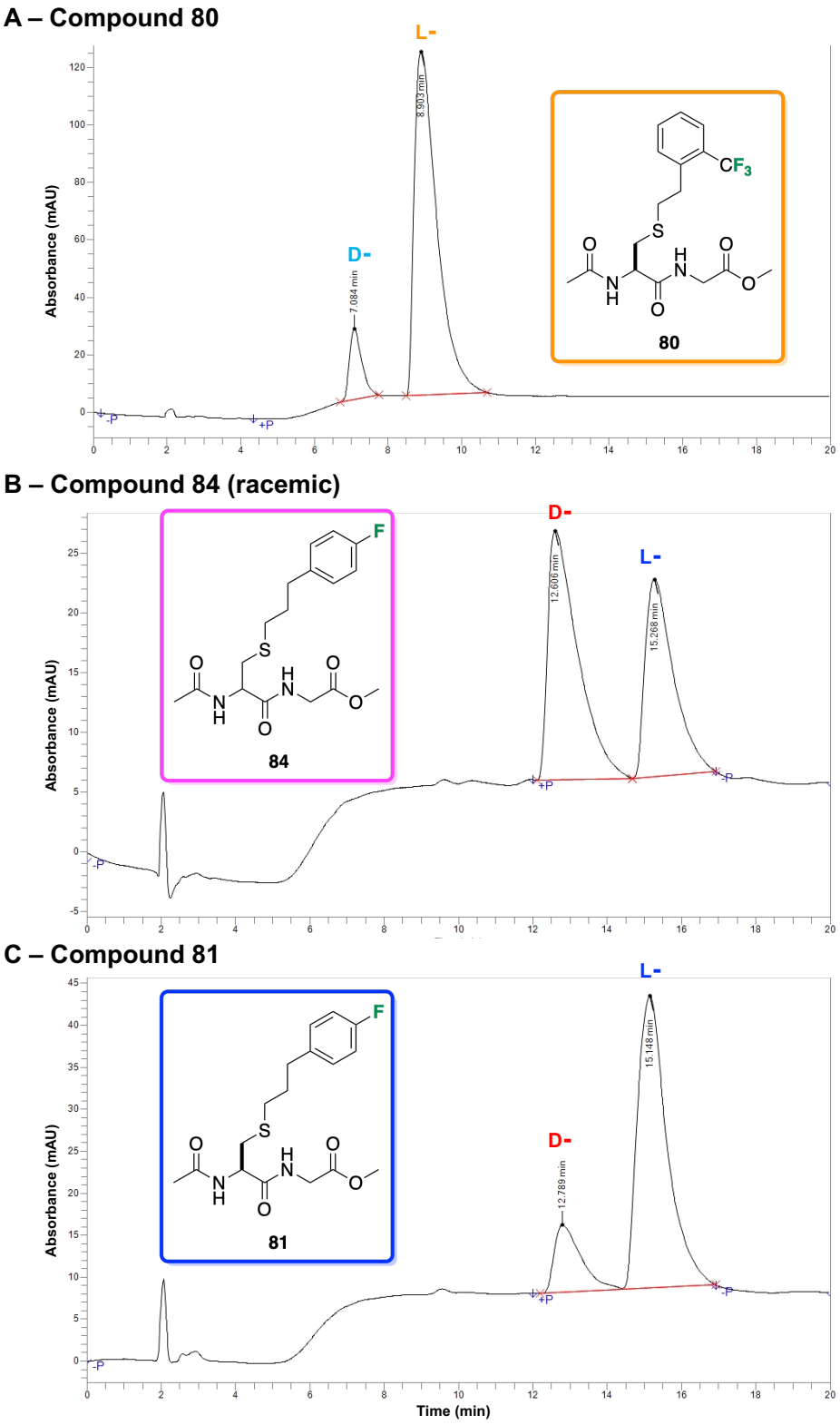


Figure 4.4. Chiral HPLC traces for compounds 80, 84, and 81. A) Compound **80** only. B) Compound **84** (racemic). C) Compound **81** only. All three compounds were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental.

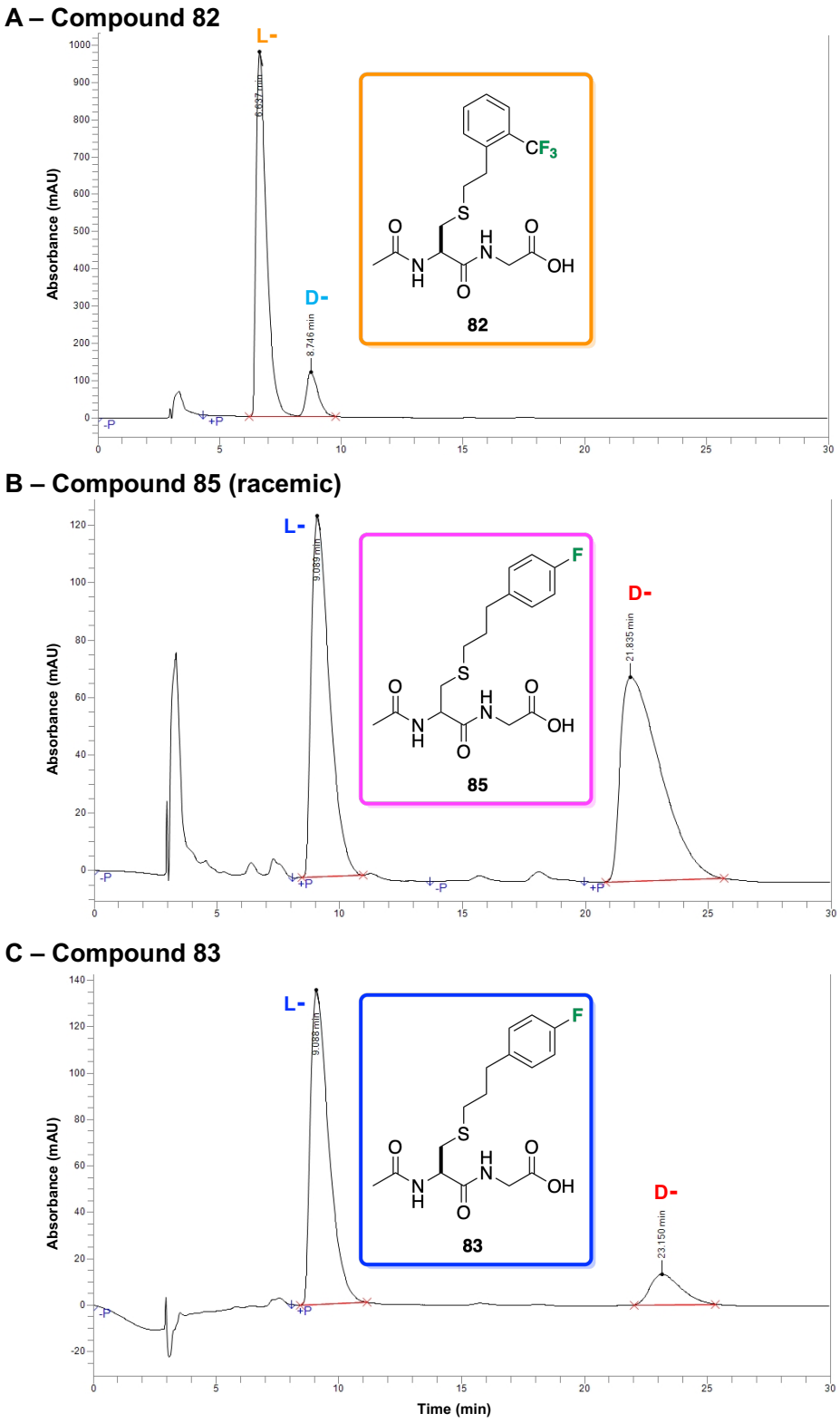


Figure 4.5. Chiral HPLC traces for compounds 82, 85, and 83. A) Compound **82** only. B) Compound **85** (racemic). C) Compound **83** only. All three compounds were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental.

4.4 Assays

Having synthesised the required ligands (**80**, **81**, **82**, and **83**), we began investigating their effects on Kef.

4.4.1 Biophysical Assays with Compounds 82 and 83

4.4.1.1 WaterLOGSY

The binding of carboxylic acids **82** and **83** to *Sd*KefQCTDH$_6$ was detected using waterLOGSY. This assay was described in Chapter 2. In each spectrum, a single ligand peak is highlighted in black, and the dimethyl sulfoxide (DMSO) peak (used as a reference) is highlighted in orange. The *N*-*tert*-butyl compound OXFKEF04 (**30**, synthesised by Angelika Wünsch,[62] Figure 4.6C) was tested as a positive control, as this compound had previously been shown to bind to *Sd*KefQCTDH$_6$ using waterLOGSY.[63] The binding of OXFKEF04 (**30**) to *Sd*KefQCTDH$_6$ was confirmed in this experiment (Figure 4.6A). The *N*-methyl analogue OXFKEF19 (**39**, Figure 4.6C, also synthesised by Angelika Wünsch[62]) was also shown to bind *Sd*KefQCTDH$_6$ using waterLOGSY (Figure 4.6B).

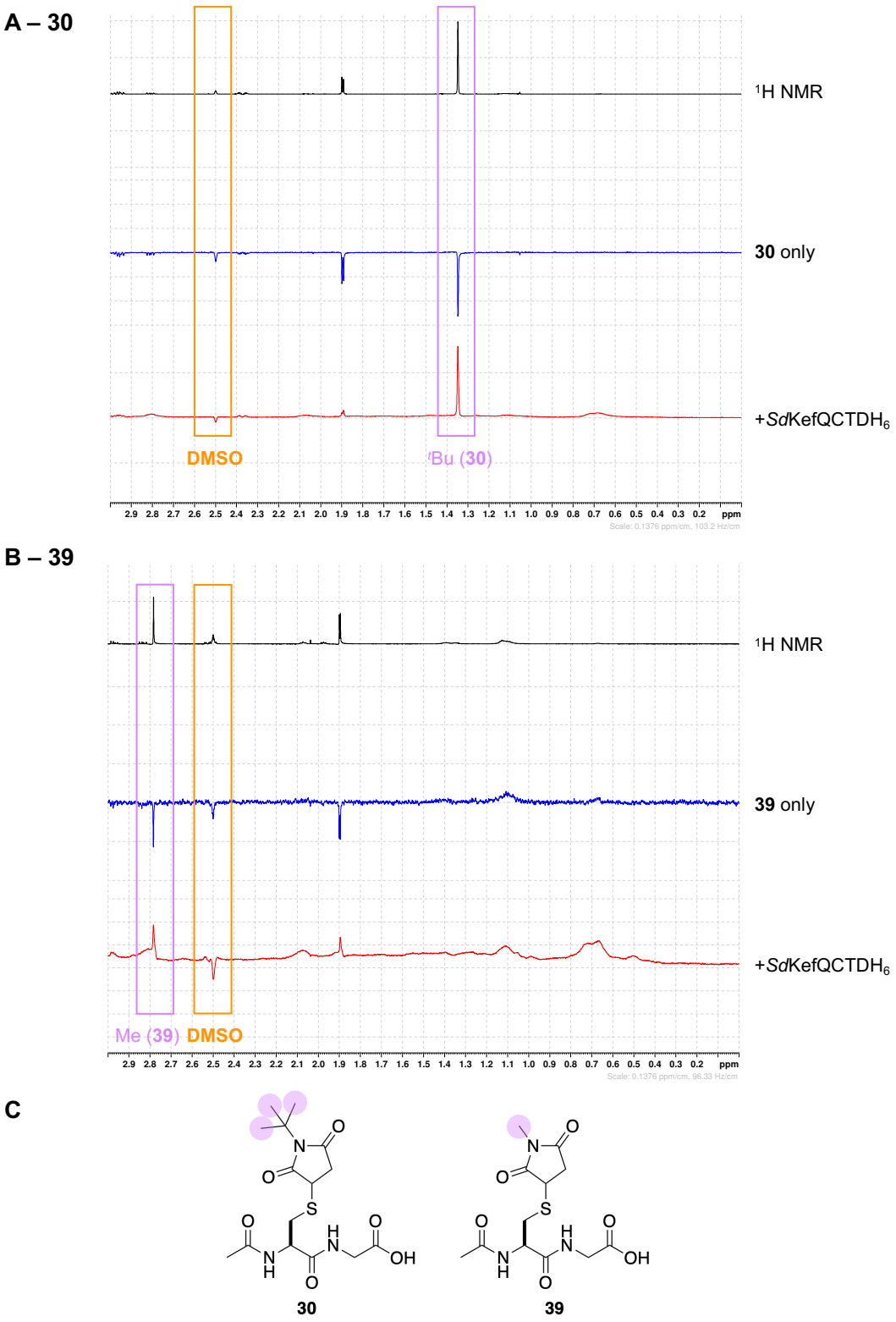


Figure 4.6. Water-suppressed ^{1}H NMR and waterLOGSY spectra for dipeptides OXFKEF04 (30) and OXFKEF19 (39) with *Sd*KefQCTDH$_6$. A) Spectra for OXFKEF04 (**30**, 500 µM) with *Sd*KefQCTDH$_6$ (20 µM). B) Spectra for OXFKEF19 (**39**, 500 µM) with *Sd*KefQCTDH$_6$ (20 µM). In all spectra, the DMSO peak is labelled (orange) and a single ligand peak is assigned (purple). C) Structures of OXFKEF04 (**30**) and OXFKEF19 (**39**). In each structure, the ^{1}H environment(s) corresponding to the assigned peak in the NMR spectra are labelled (purple circles).

The water-suppressed ^{1}H NMR and waterLOGSY spectra for compounds **82** and **83** with *Sd*KefQCTDH$_6$ are shown in Figure 4.7A and B, respectively. WaterLOGSY revealed that compounds **82** and **83** both bind to *Sd*KefQCTDH$_6$. In the waterLOGSY spectrum of **83** and *Sd*KefQCTDH$_6$, inversion of the DMSO peak could not be observed, as it was concealed by one of the ligand peaks, and so it could not be used as a reference (Figure 4.7B). Therefore, the protein peaks were used as a reference. In the waterLOGSY spectrum of compound **83** and *Sd*KefQCTDH$_6$, both the ligand and protein peaks were positive, which indicates that compound **83** binds to *Sd*KefQCTDH$_6$ (Figure 4.7B). Broadening of the ligand peaks in the presence of the protein compared to the ligand only was further evidence for ligand binding (Figure 4.7B).

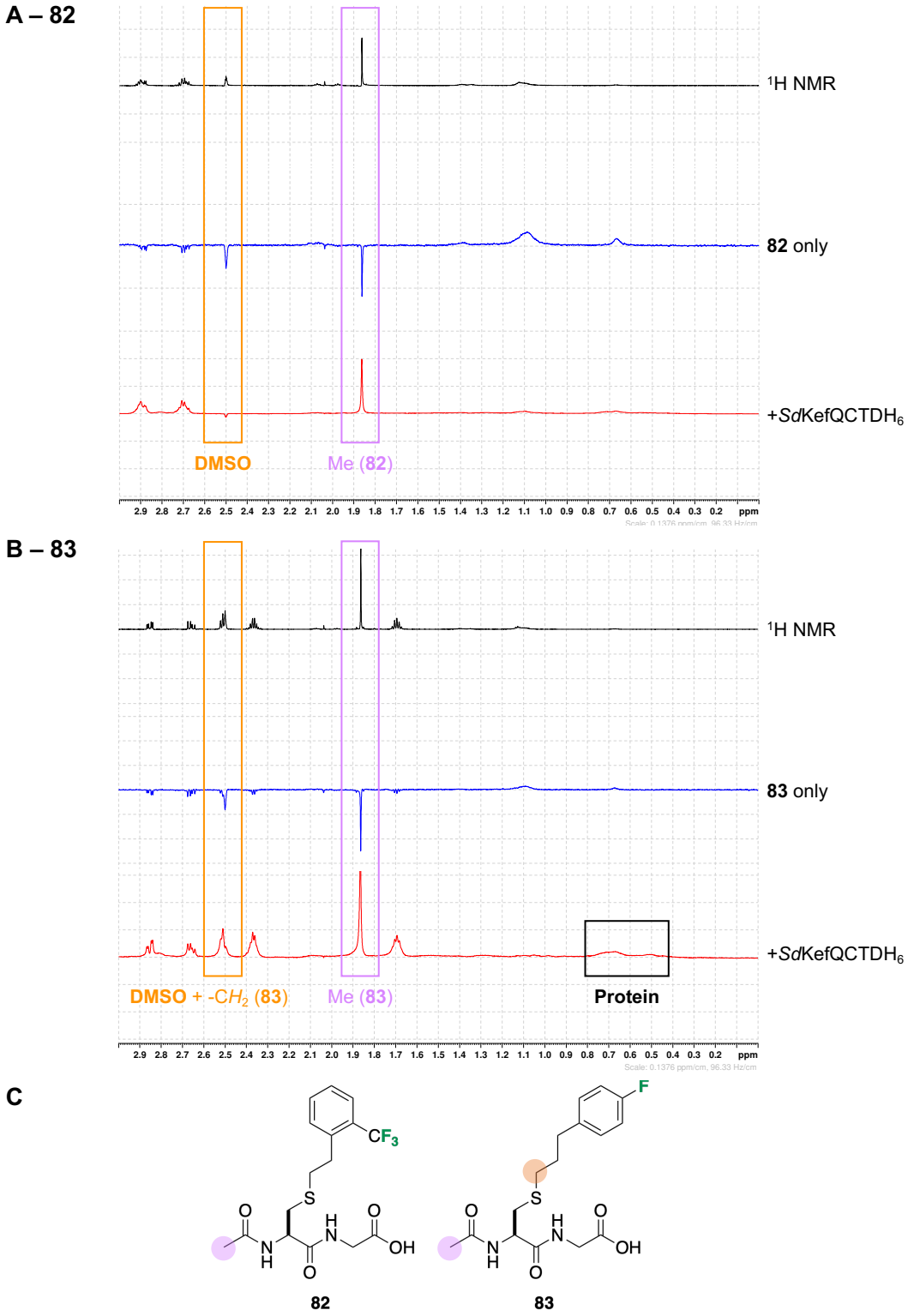


Figure 4.7. Water-suppressed ^{1}H NMR and waterLOGSY spectra for compounds 82 and 83 with *Sd*KefQCTDH$_6$. A) Spectra for **82** (500 µM) with *Sd*KefQCTDH$_6$ (20 µM). B) Spectra for **83** (500 µM) with *Sd*KefQCTDH$_6$ (20 µM). The DMSO peak overlaps with a ligand peak (orange). A second ligand peak is assigned (purple). The ligand and protein peaks (purple and black, respectively) are both positive, which indicates that **83** binds to *Sd*KefQCTDH$_6$. C) Structures of compounds **82** and **83**. The ^{1}H environment(s) corresponding to the assigned ligand peaks in the NMR spectra are labelled (purple or orange circles).

4.4.1.2 DSF

The binding of compounds **82** and **83** to *Sd*KefQCTDH$_6$ was detected using DSF, also known as the thermal shift assay. In DSF, the melting temperature (T_m) of a protein is measured in the presence of a fluorescent dye (SYPRO orange) (Figure 4.8).[64] Over the course of the assay, the temperature is gradually increased and the fluorescence of the dye is measured. The dye binds to hydrophobic regions of the protein, which are revealed as the protein unfolds, resulting in an increase in fluorescence. The T_m is the midpoint of the steepest part of the melting curve (Figure 4.8). When the T_m of the protein is measured in the presence of ligands which bind to it, the T_m increases, assuming that the ligands provide a stabilising interaction. The thermal shift, ΔT_m, which can be used to infer how well a ligand binds to a protein, is calculated by subtracting the T_m of the protein in the absence of ligand from the T_m measured in the presence of ligand. ΔT_m Values obtained for ligands with Kef are frequently large, as these proteins are homodimeric and so ligand binding provides a large stabilisation effect. For example, in a previous study by Healy *et al.* the ΔT_m of ^{t}BuSG (**26**) measured for *Sd*KefQCTDH$_6$ was ~13 °C.[31]

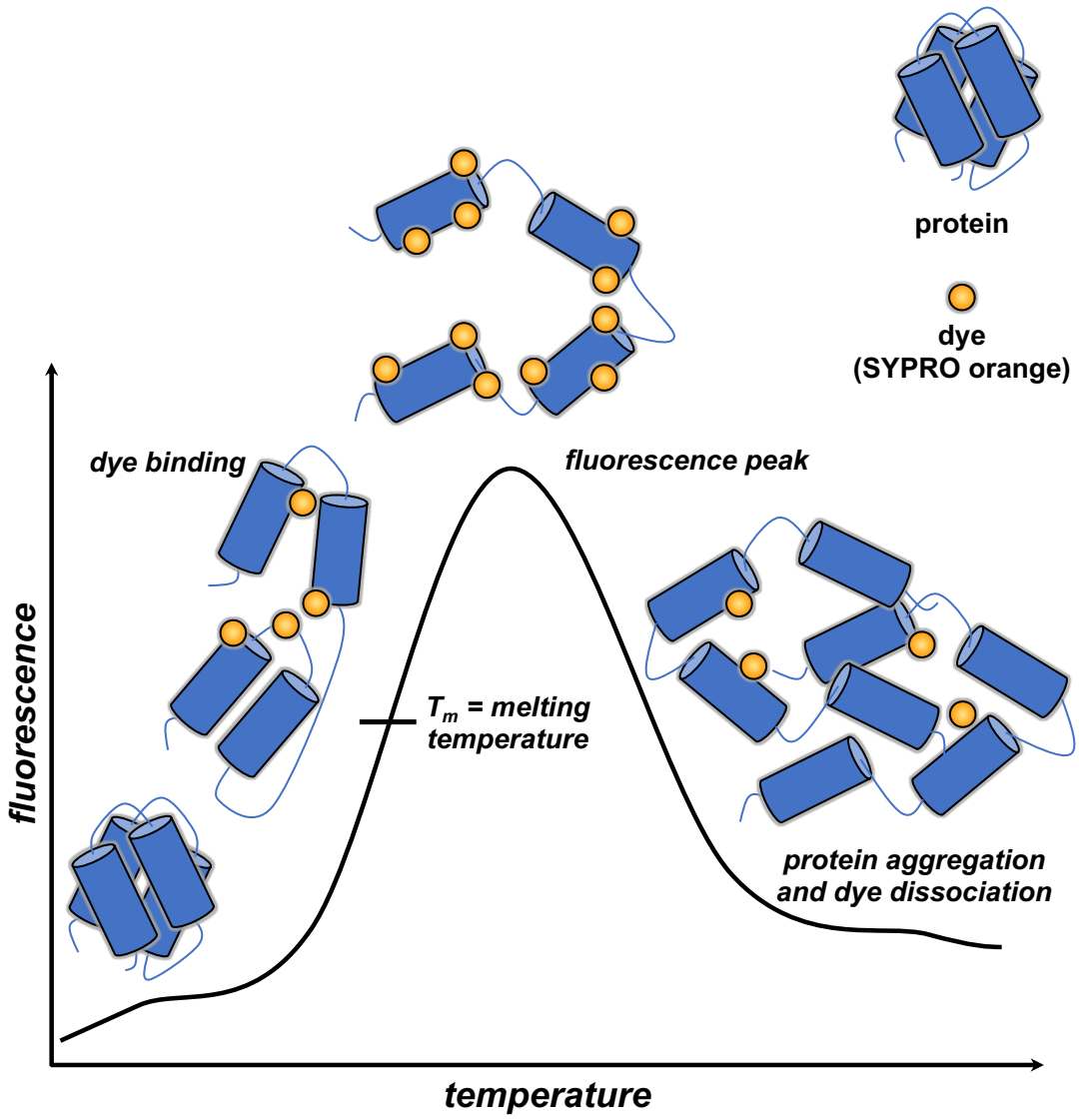


Figure 4.8. Schematic of DSF.[64]

The results of the DSF assay with *Sd*KefQCTDH$_6$ are shown in Figure 4.9. Melting of the protein plus DMSO was measured as a negative control and the corresponding T_m value is shown in Figure 4.9A. ^{t}BuSG (**26**) and ESG (**20**) were included as positive controls, as they had been shown to bind to this protein using DSF in a previous study by Healy *et al.*[31] This assay confirmed their binding to *Sd*KefQCTDH$_6$, as the ΔT_m of ^{t}BuSG (**26**) was 15 °C, and for ESG (**20**) 8 °C (Figure 4.9A), consistent with the values obtained previously (~13 °C and ~7 °C, respectively).[31] Compound **40** was included as a further positive control, as it was previously shown to bind to *Sd*KefQCTDH$_6$ using

DSF.[48] The binding of **40** to *Sd*KefQCTDH$_6$ was verified in this assay, as the ΔT_m was 9 °C (Figure 4.9A), in agreement with the ΔT_m of ~7 °C obtained previously.[48] The binding of dipeptides **30** and **39** to *Sd*KefQCTDH$_6$ was detected using DSF, with **30** giving a ΔT_m of 10 °C and **39** a ΔT_m of 3 °C (Figure 4.9A). These data suggest that the affinity of compound **30** for *Sd*KefQCTDH$_6$ is higher than the affinity of **39** for the protein. It is hypothesised that the difference in affinity is caused by the interaction of the different S-substituents in **30** and **39** with the F448 capping residue in *Sd*KefQCTDH$_6$. Finally, DSF provided further evidence that compounds **82** and **83** bind to *Sd*KefQCTDH$_6$, as **82** gave a ΔT_m of 13 °C, while **83** had a ΔT_m of 7 °C (Figure 4.9A). The larger ΔT_m for compound **82** than for **83** implies that the affinity of **82** for *Sd*KefQCTDH$_6$ is higher than the affinity of **83** for the protein. This could be caused by the increased steric bulk of the S-substituent in **82**, due to its -CF_3 group, or the reduced flexibility of the alkyl linker, due to its shorter length, resulting in a more energetically favourable interaction with F448 in *Sd*KefQCTDH$_6$. The increased lipophilicity of **82** compared to **83** could also contribute to its higher ΔT_m for *Sd*KefQCTDH$_6$.

The difference between the ΔT_m values of ^{t}BuSG (**26**) and its analogue **30** (15 °C and 10 °C, respectively) was analysed using a t test and was very significant (P = 0.0013). In contrast, the difference between the ΔT_m values of **40** and **83** (9 °C and 7 °C, respectively) was not significant (P = 0.25). These data suggest that truncation of **26** and **40** reduces their affinity for *Sd*KefQCTDH$_6$. The smaller ΔT_m difference between **40** and **83** may be caused by the S-aryl group, which is more hydrophobic than glutamate. Truncation of the glutamate in **40** is therefore predicted to cause a smaller reduction in ΔT_m than its truncation in **26**, where the S-substituent is more hydrophilic.

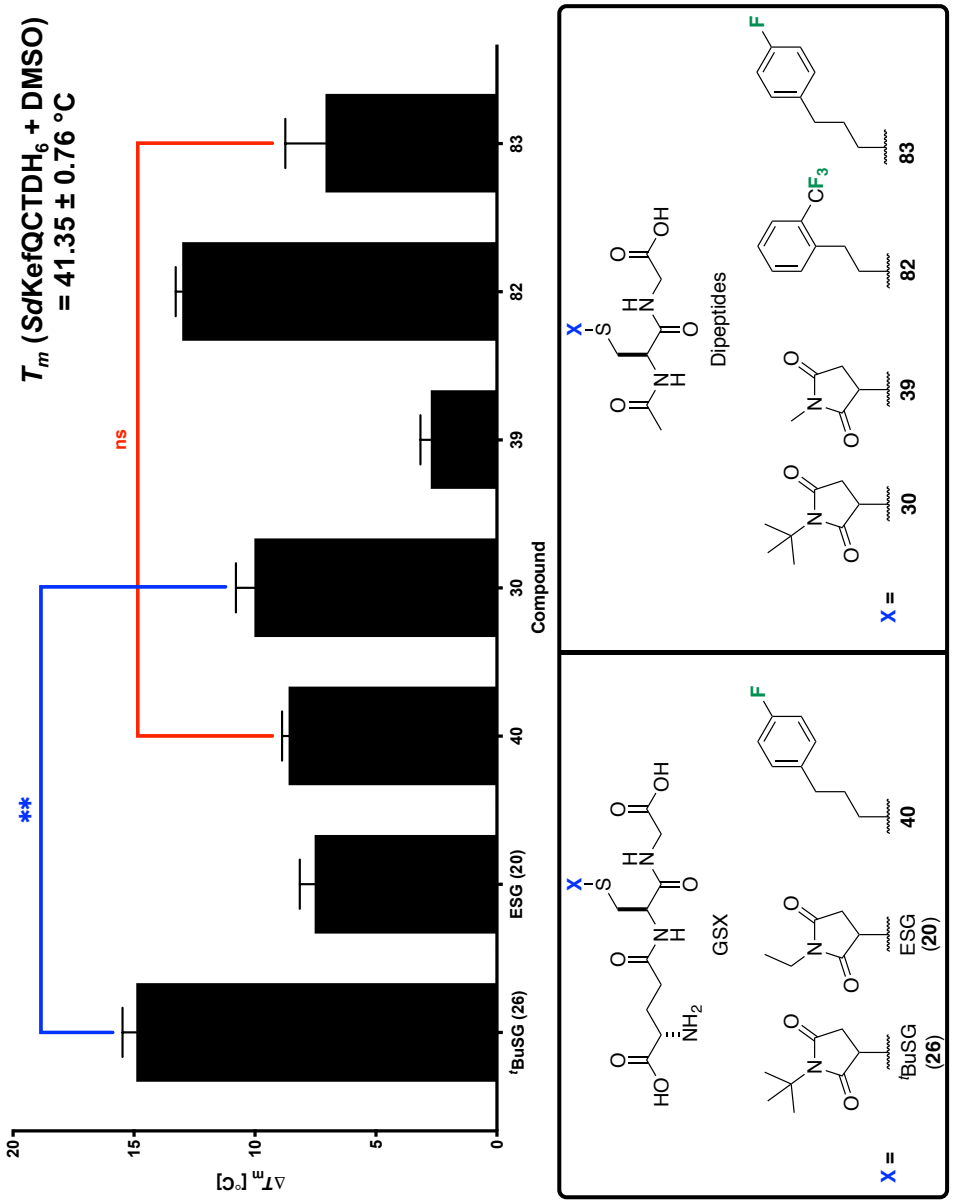


Figure 4.9. A) ΔT_m values in (°C) for *Sd*KefQCTDH$_6$. Data are mean values (n = 3). Error bars indicate 1 s.d. T_m Of *Sd*KefQCTDH$_6$ + DMSO (n = 2) shown for comparison. Significance of changes (**26/30** (very significant, **)) and **40/83** (not significant, ns)) evaluated using an unpaired t test with Welch's correction, which does not assume equal s.d. values. P Values (two-tailed): ****P ≤ 0.0001, ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05, and ns: P > 0.05. B) Structures of ligands.

4.4.2 Kirby-Bauer Disc Diffusion Assays with Compounds 80 and 81

Methyl esters **80** and **81** were tested in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP using the disc assay. Compounds **80** and **81** were tested on their own to probe activation of Kef and in combination with OXFKEF14 (**33**) to test for inhibition. OXFKEF14 (**33**) was included in all assays as a positive control compound, as it was known to be active in both *E. coli* transformants, as shown in Chapter 3. All compounds were tested in the MJF645 *E. coli* strain as a negative control, to test for off-target effects and compound toxicity.

4.4.2.1 Disc Assay with Compound 80

Compound **80** (Figure 4.10D) was tested at 15 mM, the concentration at which OXFKEF12–14 (**31**–**33**) were tested in the same *E. coli* transformants using the disc assay (described in Chapter 3). In the **33**/**80** compound combination, OXFKEF14 (**33**) was tested at 7.5 mM, half of the concentration of compound **80**. The reduced concentration of OXFKEF14 (**33**) compared to **80** was used to favour competition of the different ligands in the Kef binding sites, as their affinities for Kef were unknown. Compound **80** did not cause growth inhibition/death of either *Sd*Kef-expressing or *Yp*KefGBP-expressing MJF645 by Kef activation (Figure 4.10A and B, respectively). These data also suggest that **80** is not inhibiting the Kef systems, as the sizes of the zones for the **33**/**80** combination are the same as for OXFKEF14 (**33**) alone. The MJF645 negative control showed that neither OXFKEF14 (**33**) nor compound **80** is toxic at the concentrations tested.

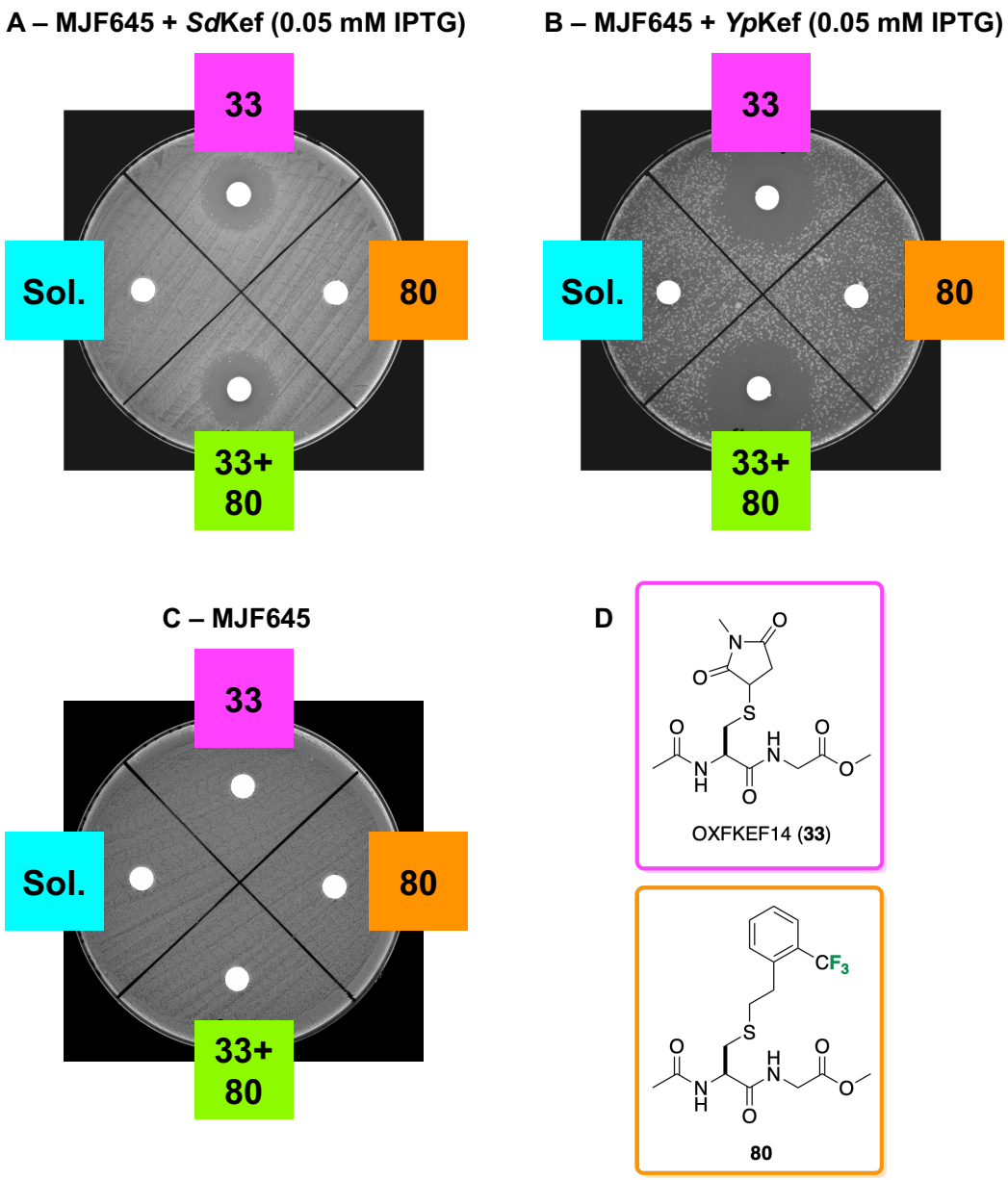


Figure 4.10. Disc assay to test compound 80 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. 33 = OXFKEF14 (**33**) and **80** = compound **80**. Both compounds (**80**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **80** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (0.05 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). C) MJF645 (negative control). D) Structures of OXFKEF14 (**33**) and compound **80**.

4.4.2.2 Disc Assays with Compound 81

Compound **81** was tested using the same conditions as for compound **80** (15 mM **81**, 7.5 mM OXFKEF14 (**33**)). Compound **81** was inactive in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP (Figure 4.11A and B, respectively). Compound **81** did not inhibit either Kef system, as the zones of exclusion for the **33**/**81** combination are the same as for OXFKEF14 (**33**) alone (Figure 4.11A and B). The MJF645 negative control shows that neither compound is toxic at the concentrations tested (Figure 4.11C).

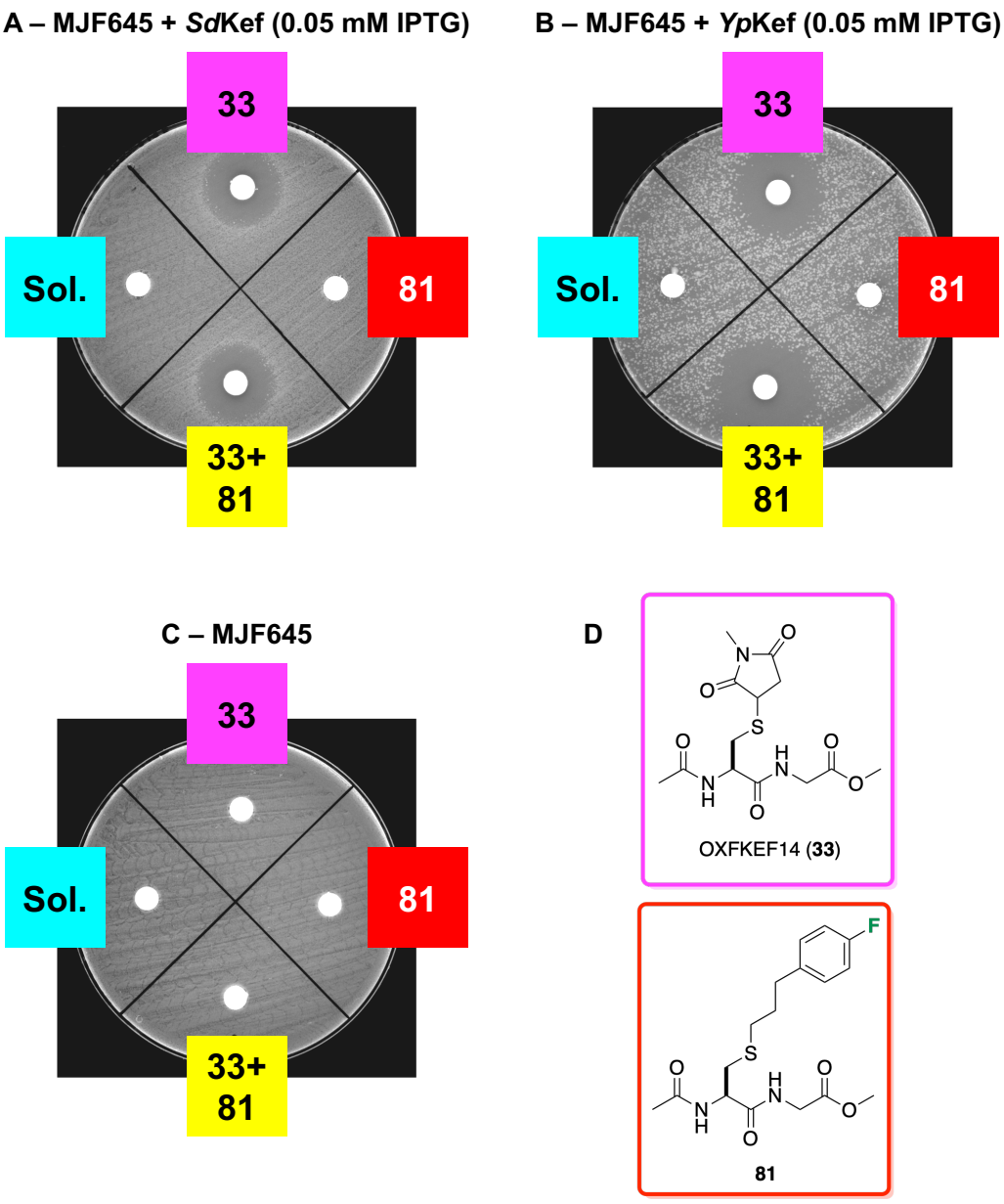


Figure 4.11. Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. 33 = OXFKEF14 (**33**) and **81** = compound **81**. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (0.05 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). C) MJF645 (negative control). D) Structures of OXFKEF14 (**33**) and compound **81**.

These initial disc assays with compounds **80** and **81** showed that both compounds behaved similarly, as neither caused any observable effect. This was hypothesised to be caused by several parameters: the concentrations of compounds **80** and **81** may have been too low; the concentration of the activator OXFKEF14 (**33**) may have been too high; or compounds **80** and **81** may have been insoluble under the conditions of the assay. To elucidate this, further experiments were carried out with compound **81** only, as it was more soluble than compound **80**.

Firstly, compound **81** was tested at 30 mM, double the previous concentration (15 mM). The concentration of OXFKEF14 (**33**) was kept constant at 7.5 mM. The results for compound **81** at 30 mM were identical to those obtained at 15 mM. Compound **81** did not activate or inhibit *Sd*Kef and *Yp*KefGBP expressed in MJF645 (Figure 4.12A and B, respectively). Neither compound **81** nor OXFKEF14 (**33**) was toxic to the MJF645 cell strain (Figure 4.12C).

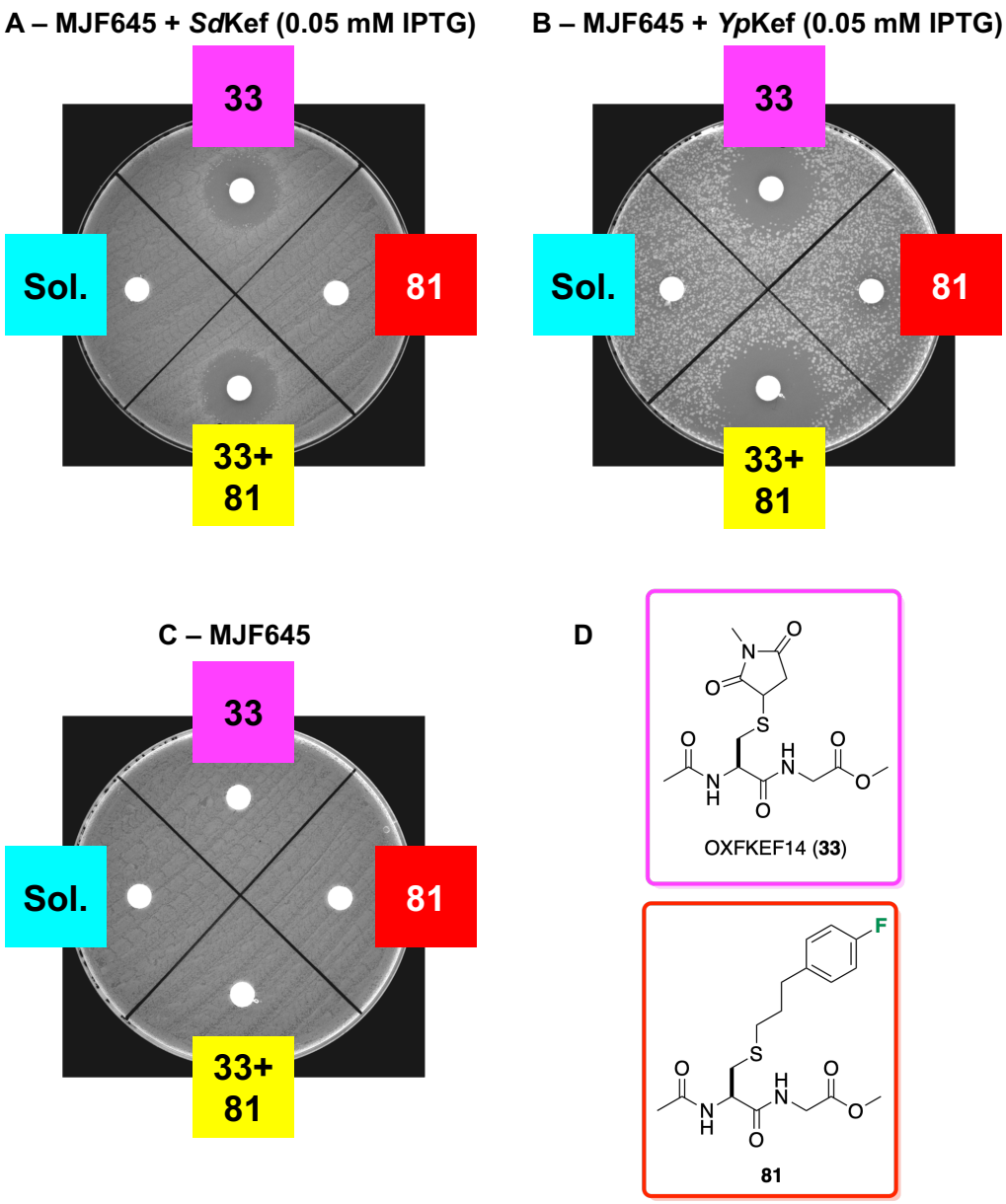


Figure 4.12. Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. 33 = OXFKEF14 (**33**) and **81** = compound **81**. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 30 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (0.05 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). C) MJF645 (negative control). D) Structures of OXFKEF14 (**33**) and compound **81**.

Compound **81** was tested with lower concentrations of the activator OXFKEF14 (**33**) to determine whether **81** would show an inhibitory effect. The concentration of OXFKEF14 (**33**) was reduced from 7.5 mM to 5 mM (Figure 4.13) and 2.5 mM (Figure 4.14), while the concentration of compound **81** was kept constant at 15 mM. The results were similar to those obtained previously, as compound **81** did not inhibit either *Sd*Kef or *Yp*KefGBP, as the sizes of the zones for the **33**/**81** combination were the same as for OXFKEF14 (**33**) alone (Figure 4.13A and B, Figure 4.14A and B). The MJF645 negative controls showed that OXFKEF14 (**33**) is non-toxic at both 5 and 2.5 mM (Figure 4.13C and Figure 4.14C, respectively).

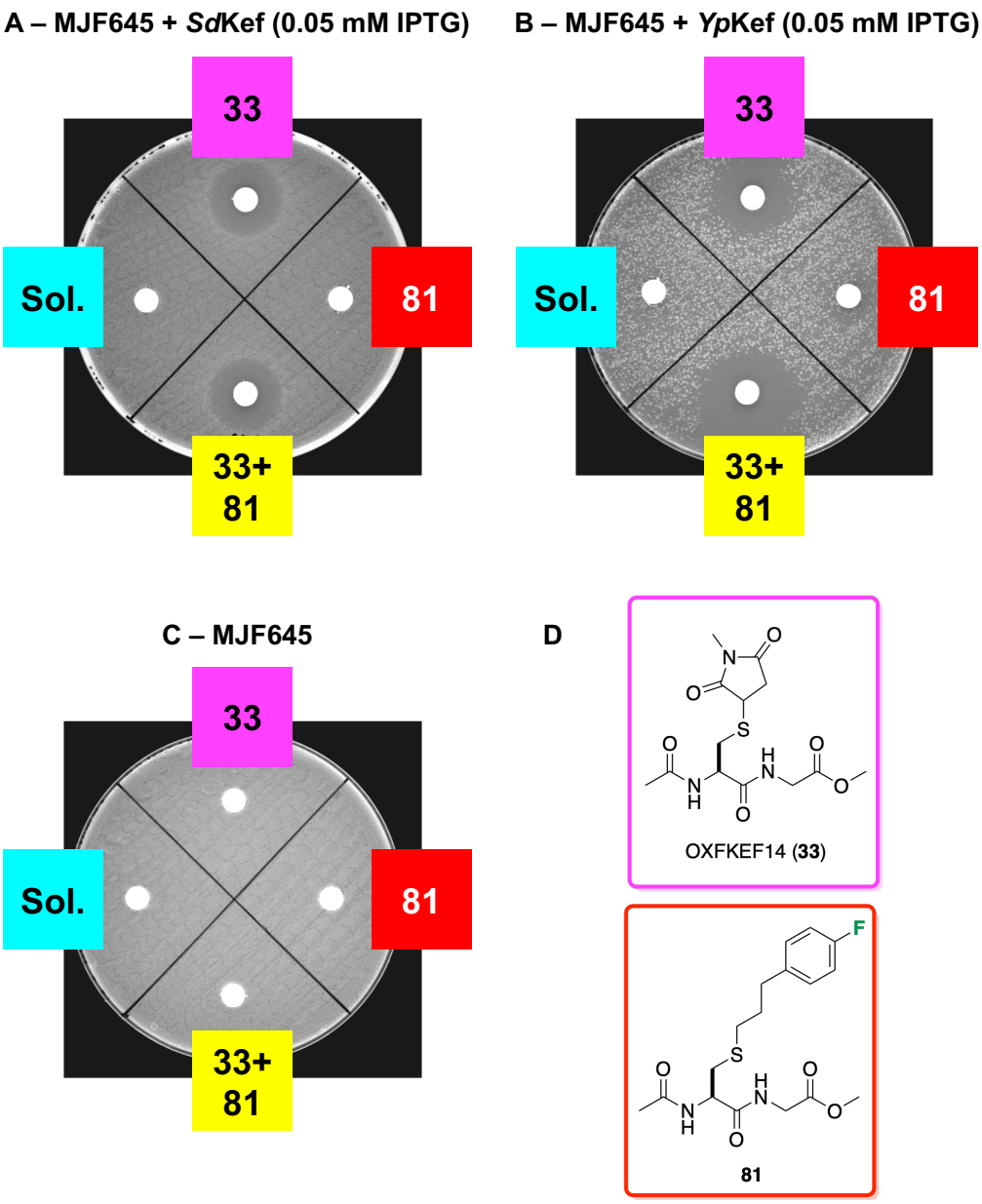


Figure 4.13. Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. 33 = OXFKEF14 (**33**) and **81** = compound **81**. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 5 mM. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) MJF645 + pTrcSdKefH$_6$ (0.05 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). C) MJF645 (negative control). D) Structures of OXFKEF14 (**33**) and compound **81**.

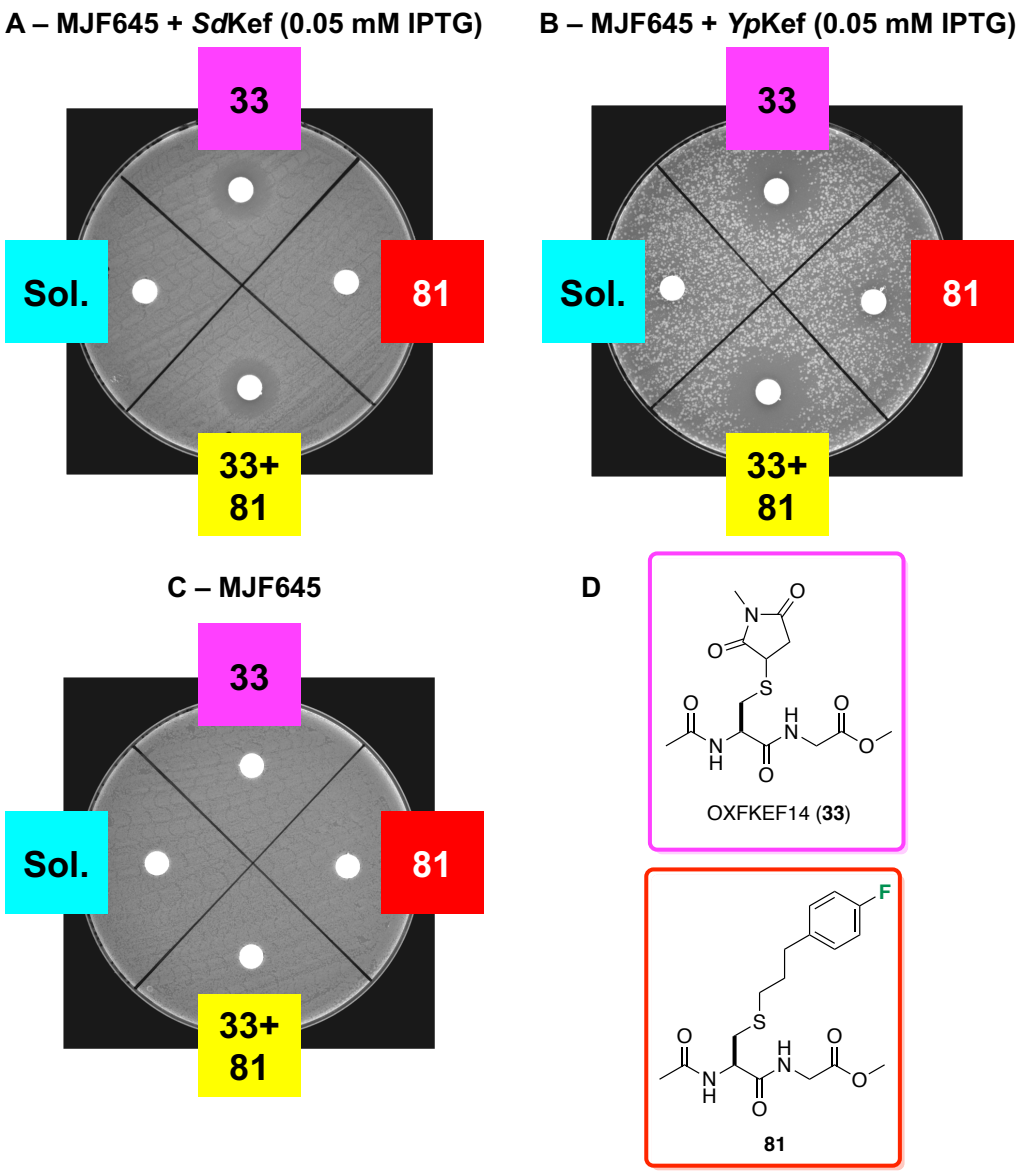
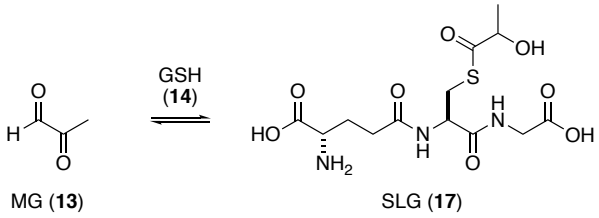


Figure 4.14. Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. 33 = OXFKEF14 (**33**) and **81** = compound **81**. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 2.5 mM. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) MJF645 + pTrcSdKefH_6 (0.05 mM IPTG). B) MJF645 + pTrcYpKefGBP (0.05 mM IPTG). C) MJF645 (negative control). D) Structures of OXFKEF14 (**33**) and compound **81**.

From these experiments with OXFKEF14 (**33**) and **81**, it was hypothesised that the affinity of OXFKEF19 (**39**) for *Sd*Kef and *Yp*KefGBP was high enough to prevent ligand exchange with **83** (formed from **81**). We therefore wanted to repeat the experiments with compound **81** using an activator which had lower affinity for *Sd*Kef and *Yp*KefGBP.

We selected SLG, **17**, (Scheme 4.3) as our Kef ligand of lower affinity. SLG (**17**) is formed from the reaction of GSH (**14**) with the electrophile MG, **13**, (Scheme 4.3). SLG (**17**) binds *Sd*KefQCTDH$_6$ with an affinity of K_d = 900 ± 200 µM (fluorescence emission spectra).[31] SLG (**17**) was predicted to be charged at physiological pH and therefore not cell permeant. However, MG (**13**) was predicted to be neutral and cell permeant.



Scheme 4.3. Formation of SLG (17) from MG (13) and GSH (14).

We proposed using MG (**13**) to generate SLG (**17**) in *E. coli* containing GSH (**14**) in the disc assay. As *E. coli* MJF645 does not contain GSH (**14**), it was necessary to use a different *E. coli* strain, which contained GSH (**14**). We selected the wild-type strain Frag 1. For the disc assay, both *Sd*Kef and *Yp*KefGBP would be expressed in Frag 1. We would first carry out an assay with MG (**13**) only, to test for the optimal concentration of the electrophile. We wanted to produce a zone size large enough to permit clear observation. However, we wanted to produce an activation effect small enough to allow inhibition of Kef by compound **81**. Once the optimal concentration of MG (**13**) had been identified, compound **81** would be tested in the Frag 1 transformants, in combination with MG (**13**), to test for the inhibitory activity of compound **81**.

We first tested MG (**13**) in Frag 1 expressing either *Sd*Kef or *Yp*KefGBP. The Frag 1 cell strain was used as a control, to check for compound toxicity. We carried out a broad screen of concentrations of MG (**13**), starting at ~84 mM and diluting serially twofold

to a lowest concentration of ~2.6 mM. The concentrations of MG (**13**) were approximate as the test solutions were prepared from a commercially available MG solution of approximate concentration (~40% *v/v*). The data for ~2.6, ~5.3, and ~11 mM MG (**13**) are shown in Figure 4.15, and for ~21, ~42, and ~84 mM MG (**13**) in Figure 4.16.

MG (**13**) at ~2.6 and ~5.3 mM did not produce any zones of growth inhibition/death in Frag 1 expressing either *Sd*Kef or *Yp*KefGBP (Figure 4.15A and B, respectively). MG (**13**) at ~11 mM produced small zones of exclusion (Figure 4.15A and B, respectively). For ~21–84 mM MG (**13**), the zone size increased dose-dependently (Figure 4.16A and B). However, the zone sizes for Frag 1 expressing either *Sd*Kef or *Yp*KefGBP were the same as for Frag 1 only (Figure 4.15C (~11 mM MG) and Figure 4.16C (~21–84 mM MG)). These data suggested that the effects seen when Kef was overexpressed were due to general toxicity of MG (**13**), as opposed to activation of Kef by SLG (**17**). Frag 1 was not an entirely valid control, however, as this cell strain contains the endogenous *E. coli* Kef proteins (KefB, C, F, and G).

Further control experiments were therefore carried out in the MJF335 *E. coli* cell strain, which lacks KefB and -C. As MJF335 also lacks *gshA*, the experiments were carried out with and without supplementary GSH (**14**, 1 mM). The results for MJF335 with and without GSH (**14**) are analogous to the results obtained in Frag 1 expressing either *Sd*Kef or *Yp*KefGBP (Figure 4.15D and E (~2.6–11 mM MG) and Figure 4.16D and E (~21–84 mM MG)). These data confirmed that the effects of MG (**13**) seen when Kef was overexpressed were due to general toxicity of the electrophile, as opposed to activation of Kef by SLG (**17**).

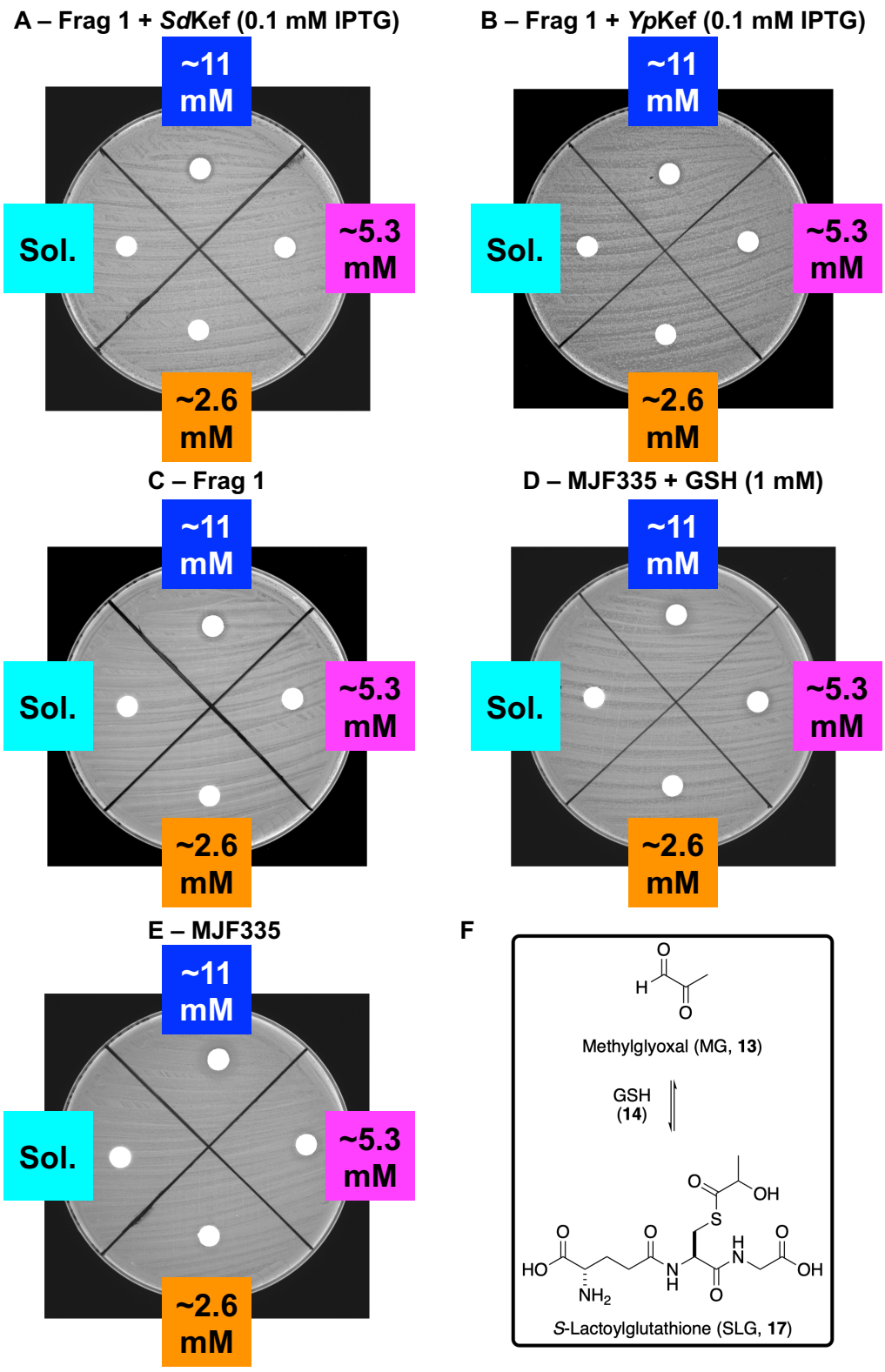


Figure 4.15. Disc assay for ~2.6–11 mM MG (13) in *E. coli* for the formation of activator SLG (17) in the cell. MG (**13**) was tested in 80% *v/v* DMSO at the concentrations indicated. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) Frag 1 + pTrcSdKefH$_6$ (0.1 mM IPTG). B) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). C) Frag 1. D) MJF335 + GSH (**14**, 1 mM). E) MJF335. F) Scheme for the formation of SLG (**17**) from MG (**13**) and GSH (**14**).

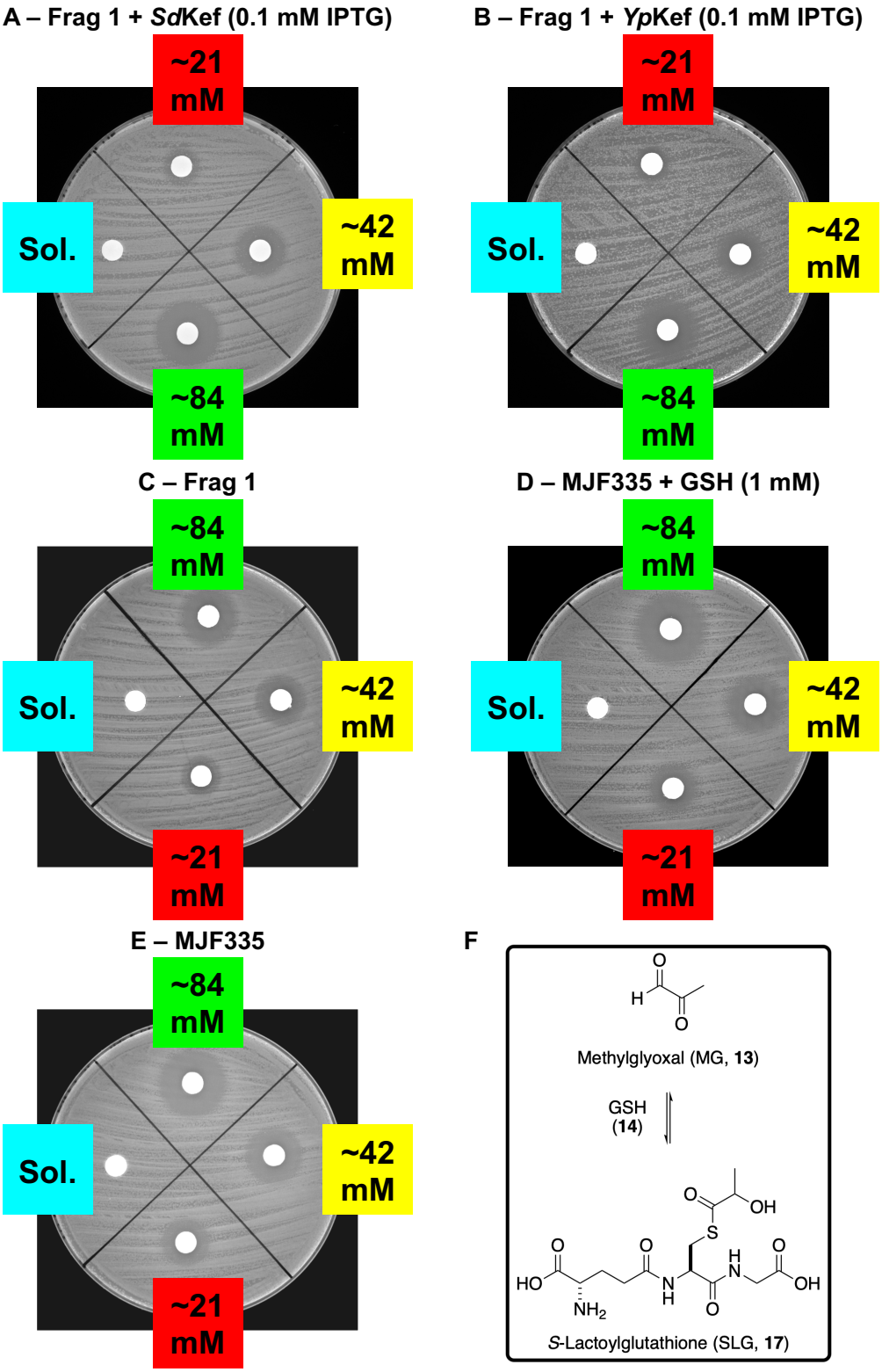


Figure 4.16. Disc assay for ~21–84 mM MG (13) in *E. coli* for the formation of activator SLG (17) in the cell. MG (**13**) was tested in 80% *v/v* DMSO at the concentrations indicated. Sol. = solvent = 80% *v/v* DMSO, which was tested as a negative control. A) Frag 1 + pTrcSdKefH_6 (0.1 mM IPTG). B) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). C) Frag 1. D) MJF335 + GSH (**14**, 1 mM). E) MJF335. F) Scheme for the formation of SLG (**17**) from MG (**13**) and GSH (**14**).

4.5 Chapter 4 Conclusions

Novel dipeptides **80**, **81**, **82**, and **83** were synthesised using the thiol-ene reaction. Carboxylic acids **82** and **83** were shown to bind *Sd*KefQCTDH$_6$ using both waterLOGSY and DSF, which indicates that aromatic S-substituents are tolerated in terms of Kef binding. Methyl esters **80** and **81** did not activate or inhibit either *Sd*Kef or *Yp*KefGBP expressed in MJF645 in the disc assay. We hypothesised that the lack of effect of **80** and **81** may have been caused by several factors: the poor solubility of **80** and **81**; the concentration of **80** and **81** being too low; or the concentration of the activator **33** being too high. Subsequently, **81** was tested at a higher concentration and with a lower concentration of **33** but **81** was still inactive. SLG (**17**) was trialled as an activator of lower affinity for Kef, which would be formed in Frag 1 from its endogenous GSH (**14**) and MG (**13**) in the disc assay. MG (**13**) caused growth inhibition/death of Frag 1 expressing either *Sd*Kef or *Yp*KefGBP but control experiments showed that this was caused by the toxicity of MG (**13**), as opposed to activation of Kef by SLG (**17**). Therefore, using MG (**13**) was not an effective method to generate SLG (**17**) in the cell.

Chapter 5: Conclusions and Future Work

5.1 Conclusions

This dissertation sought to further elucidate the Kef mechanism of action and to investigate novel Kef systems as antibiotic targets in pathogenic bacteria.

The data obtained in Chapter 2 are consistent with the single amino acid switch model for Kef activation. This novel ^{19}F NMR-based assay is an efficient method to determine whether a compound is a Kef activator or inhibitor, vital in the design of new Kef ligands. Activators, including ESG (**20**) and ^{t}BuSG (**26**) (Figure 5.1), displace the F448*pt*fmF capping residue of the protein, whereas inhibitors, including GSH (**14**), do not. The ^{19}F NMR assay also characterised compounds **29** and **40** (Figure 5.1) as Kef inhibitors.

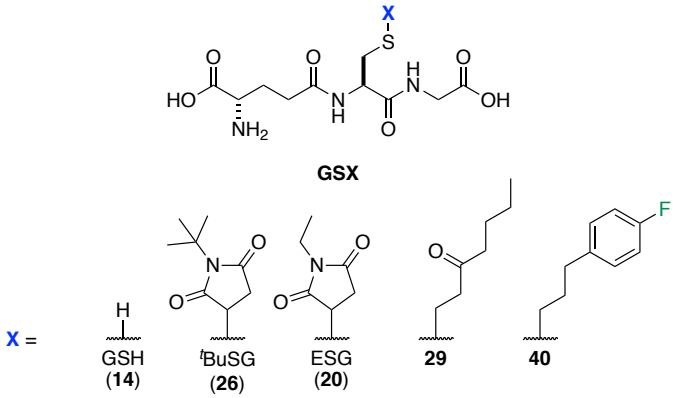


Figure 5.1. Structures of GSH (14) and GSX ligands.

In Chapter 3, novel Kef systems were identified using bioinformatics: *Yp*KefGBP in *Y. pestis* and *Bp*Kef1 and -2 in *B. pseudomallei*. Testing L-methyl esters OXFKEF12–14 (**31**–**33**) in *E. coli* expressing *Yp*KefGBP, *Bp*Kef1, and -2 using the disc assay revealed antibacterial activity against *Yp*KefGBP only. OXFKEF13 (**32**) and -14 (**33**) caused growth inhibition/death of MJF645 expressing *Yp*KefGBP (Figure 5.2A) but were

inactive in the Frag 1 transformant (Figure 5.2B). A MIC assay revealed that OXFKEF14 (**33**) is bacteriostatic in MJF645 expressing *Yp*KefGBP (MIC = 15.6 μg mL^{-1}). The results for *Yp*KefGBP in MJF645 show that activation of *Yp*KefGBP can inhibit bacterial growth, demonstrating its potential as an antibiotic target. In contrast, OXFKEF14 (**33**) did not inhibit the growth of wild-type *Y. pestis* in the disc assay (Figure 5.2C). The expression level of *Yp*KefBH_6 in both MJF645 and Frag 1 was determined using western blotting and was found to be similar, which is also assumed to be the case for *Yp*KefB. Based on this, the potential K^+ efflux from the two *E. coli* transformants above is approximately equal. Therefore, the opposing disc assay results must either be caused by GSH (**14**) in Frag 1 degrading the compounds or inhibiting *Yp*KefB, or incorrect folding of *Yp*KefB in Frag 1.

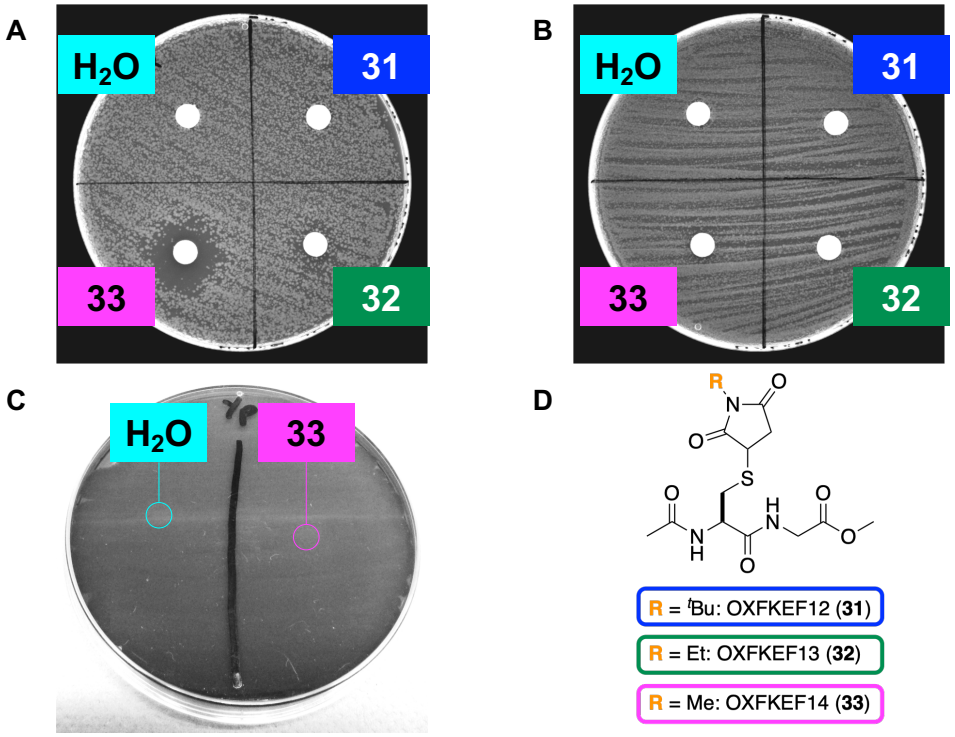


Figure 5.2. Summary of disc assay data for OXFKEF12–14 in *E. coli* expressing *Yp*KefGBP and in wild-type *Y. pestis*. OXFKEF12–14 (**31**–**33**) were tested at 15 mM in H_2O. A) MJF645 + pTrcYpKefGBP (0.1 mM IPTG). B) Frag 1 + pTrcYpKefGBP (0.1 mM IPTG). C) *Y. pestis*. D) Structures of OXFKEF12–14 (**31**–**33**).

Carboxylic acids **82** and **83** (Figure 5.3) were shown to bind to *Sd*KefQCTDH$_6$ using waterLOGSY and DSF. Methyl esters **80** and **81** (Figure 5.3) were inactive in MJF645 expressing *Sd*Kef and *Yp*KefGBP in the disc assay, which may have been caused by several reasons: the poor solubility of **80** and **81**; the concentration of **80** and **81** being too low; or the concentration of the activator **33** being too high. The more soluble compound (**81**) was tested at a higher concentration, and with a lower concentration of activator **33**, but was still inactive. It was hypothesised that the affinity of **39**, formed from prodrug **33** *in vivo*, was too high to see an effect of **81**. SLG (**17**, Figure 5.3) was investigated as an alternative activator with lower affinity for *Sd*Kef and *Yp*KefGBP compared to **39**. In the disc assay, the electrophile MG (**13**) was used to form SLG (**17**) in GSH-containing *E. coli*. The activity for MG (**13**) in *E. coli* was caused by its general toxicity, and so this was not an effective method to generate SLG (**17**) in the cell.

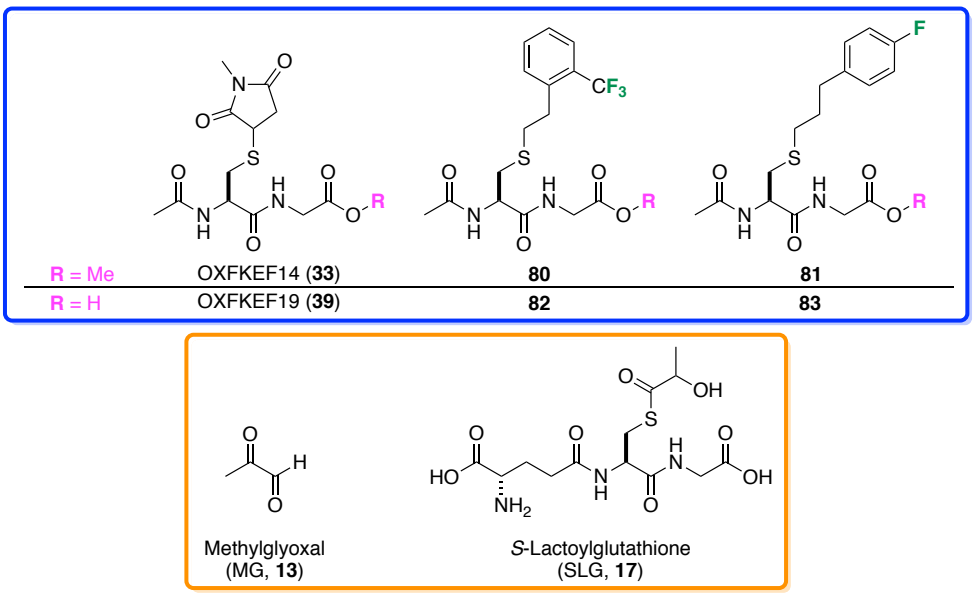


Figure 5.3. Structures of compounds from Chapter 4.

In summary, the ^{19}F NMR data acquired using the fluorinated Kef protein are consistent with the previous model of Kef activation. As antibacterial activity was only identified in *Yp*KefGBP-expressing *E. coli*, and not in wild-type *Y. pestis*, ligand optimisation is required to further assess the potential of *Yp*KefGBP as a novel antibiotic target.

It is unknown whether Kef will be a viable antibiotic target. We have not yet shown that activation of Kef at wild-type expression levels can inhibit or kill bacteria. Further work is needed to explore this and whether Kef inhibition could be effective. If Kef antibiotics are developed in the future, these will likely form part of combination therapies, as this strategy can help to control resistance.[8]

5.2 Future Work

To support the work described in Chapter 2, a control is required to ensure that the effect seen in the ^{19}F NMR spectra with *Sd*KefQCTDH$_6$(F448*p*tfmF) is caused by ligand binding to the protein and displacement of the F448*p*tfmF capping residue. One option would be to use a second mutant protein, in which an amino acid spatially removed from the ligand binding site is mutated to *p*tfmF. Upon ligand binding, the ^{19}F chemical shift is predicted to remain approximately constant, as the mutation is not located in the ligand binding site.

To continue investigating *Yp*KefGBP as a potential antibiotic target, it is necessary to synthesise analogues of the active compound **39** and determine their affinity for *Yp*KefGBP using biophysical assays. These assays require a soluble CTD construct of *Yp*KefB to be produced and purified. To improve the affinity of **39** for *Yp*KefGBP, there

are four main sites for potential modification (Figure 5.4): 1) the acetyl group on the cysteinylamine could be extended to gain affinity. In the parent tripeptides, the carboxylate of the glutamate is predicted to interact with R498 in *Ec*KefCCTD.[41,43] To attempt to form interaction(s) with the corresponding residue (R499) in *Yp*KefB, carboxylate isosteres could be incorporated in the analogous position in **39** (Figure 5.4); 2) the S-substituent could be varied, as this has been shown to affect activity for different Kef systems *in vivo*. Based on **82** and **83** from Chapter 4, aromatic groups with solubilising groups (Figure 5.4) could be incorporated using the thiol-ene reaction. Alternatively, enones with hydrophilic moieties could be added using conjugate addition; 3) the Cys-Gly amide bond could be replaced with isosteres, such as a 1,2,3-triazole (Figure 5.4), which may improve metabolic stability compared to **39**; 4) the glycine carboxylic acid could also be replaced with isosteres, for example an isoxazole (Figure 5.4). Modified compounds must either be cell permeant for *in vivo* assays, or made permeant, for example by forming a prodrug, as for methyl ester **33**. After biophysical assays to determine affinity, compounds would be tested in *E. coli* expressing *Yp*KefGBP, using disc and MIC assays. Compounds of higher affinity than **39** would be tested in wild-type *Yersinia pestis*.

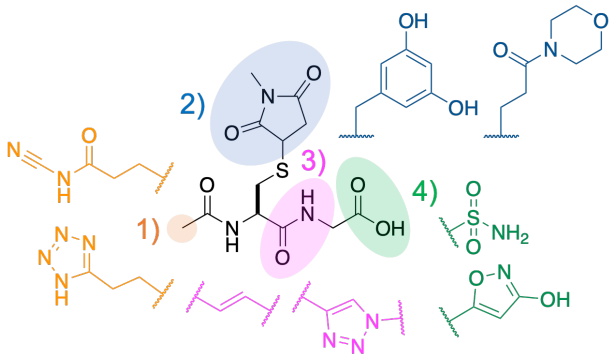
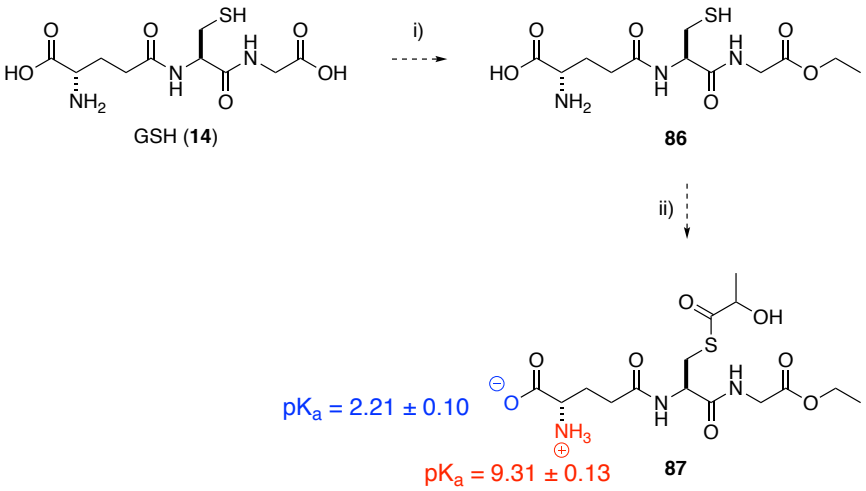


Figure 5.4. Sites for potential modifications in OXFKEF19 (39).

The low sequence identity for *Bp*Kef1 and -2 relative to *Sd*Kef and lack of compound hits identified from OXFKEF12–14 (**31**–**33**) potentially indicate that the low molecular weight thiol in *B. pseudomallei* is not GSH (**14**). The endogenous thiol could be identified through pull-down experiments using maleimide-based probes and detection using mass spectrometry. Compound libraries based on the thiol S-conjugates identified in the above experiments could then be tested against *Bp*Kef1 and -2, to identify compounds which bind and activate the Kef systems.

A potential solution to generating *S*-lactoylglutathione (SLG, **17**) in *E. coli* is to synthesise a cell permeant analogue for the disc assay. Unlike SLG (**17**), its monoethyl ester (**87**) is predicted to be neutral at physiological pH (Scheme 5.1) and therefore cell permeant. The ester **87** could be synthesised in two steps from GSH (**14**), by esterification under acidic conditions to the monoethyl ester (**86**)[65] and subsequent conjugate addition of compound **86** to MG (**13**).



Scheme 5.1. Proposed synthesis of compound 87. *Reagents and conditions*: i) EtOH, H_2SO_4, 35 °C;[65] ii) MG (**13**), H_2O, rt. pK_a Values for compound **87** were calculated using the ACD/I-Lab pK_a tool (available at http://cds.rsc.org).

Chapter 6: Experimental

6.1 General Chemistry Methods

^{1}H NMR Spectra were measured on a Bruker AVIIIHD 400 nanobay (400 MHz) spectrometer or Bruker AVII 500 (500 MHz) spectrometer with dual ^{13}C(^{1}H) cryoprobe, in the stated solvents as a reference for the internal deuterium lock. The spectra are calibrated using the residual solvent peak with the data provided by Fulmer *et al.*[66] The chemical shift data for each signal are given as δ in units of parts per million (ppm), relative to tetramethylsilane (TMS), where δ(TMS) = 0.00, to the nearest 0.01 ppm. The multiplicity of each signal is indicated by: s (singlet); d (doublet); t (triplet); q (quartet); m (multiplet), or combinations thereof. The number of protons (n) for a given resonance signal is indicated by nH. Where appropriate, coupling constants (*J*) are quoted in Hz, recorded to the nearest 0.1 Hz, and were determined by analysis using Bruker TopSpin v3.5pl7 software. Identical coupling constants were averaged. ^{1}H Environments in ABX systems,[67] and the corresponding coupling constants, are labelled (for example, *H*-5$_X$ and J_{AX}). Spectra were assigned using two-dimensional (2D) NMR experiments: ^{1}H-^{1}H correlation spectroscopy (^{1}H-^{1}H COSY), ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC), and ^{1}H-^{13}C heteronuclear multiple quantum correlation (HMBC), as necessary.

When two **diastereoisomers** are present in the sample, A and B denotes each of the 2 diastereoisomers. A Is arbitrarily assigned to the diastereoisomer with the higher ppm shift for the signal corresponding to the Cys-α-*H*. The remaining resonances were assigned to each of the two specific diastereoisomers using COSY, HSQC, and HMBC experiments as necessary. Where resonances could not be assigned to a particular

diastereoisomer, both peaks are reported. The proportions of the 2 diastereoisomers A and B in a mixture were determined from ^{1}H NMR spectra. In each spectrum, 2 distinct resonances, 1 for each diastereoisomer, and both corresponding to the same ^{1}H environment in the molecule, were identified. The value of the larger integral was divided by the smaller one to obtain a ratio of the 2 diastereoisomers A:B, either 1:X or X:1, where X is a number greater than 1, given to 2 significant figures.

^{19}F NMR Spectra were recorded on a Bruker AVIIIHD 400 nanobay (377 MHz) or Bruker AVIIIHD 500 (470 MHz) spectrometer using a broadband proton decoupling pulse sequence and deuterium internal lock. The chemical shift data for each signal are given as δ in units of parts per million (ppm), to the nearest 0.1 ppm. The multiplicity of each signal is singlet.

^{13}C NMR Spectra were measured on a Bruker AVII 500 (126 MHz) spectrometer, with dual ^{13}C(^{1}H) cryoprobe, in the stated solvents as a reference for the internal deuterium lock using the standard ^{13}C experiment. The spectra are calibrated using the solvent peak with the data provided by Fulmer *et al.*[66] The chemical shift data for each signal are given as δ in units of parts per million (ppm) relative to tetramethylsilane (TMS), where δ(TMS) = 0.00, to the nearest whole 0.01 ppm. Signals are quoted to 1 decimal place unless peaks are indistinguishable, in which case 2 decimal places are used. The multiplicity of each signal is indicated by: d (doublet) or q (quartet). Where appropriate, coupling constants (*J*) are quoted in Hz, recorded to the nearest 0.1 Hz, and were determined by analysis using Bruker TopSpin v3.5pl7 software. Spectra were assigned using HSQC and HMBC experiments as necessary, including assignment to

each diastereoisomer (A and B). Where resonances could not be assigned to a particular diastereoisomer, both peaks are reported.

Mass spectra were acquired on either an Agilent 6120 (low resolution) or Bruker microToF spectrometer (high resolution) using electrospray ionisation (ESI) from solutions of MeOH or H_2O. *m/z* Values are reported in Daltons (Da) and followed by their percentage abundance in parentheses. Peak intensities were determined using MestReNova version 12.0.2 software.

Melting points were determined using a Leica Galen III hot stage microscope and are uncorrected. The one or more solvents from which the sample was crystallised are given in parentheses. Samples which decomposed after melting are denoted "dec.".

Infrared (IR) spectra were either obtained from neat samples or from compound solutions (in CH_2Cl_2, $CHCl_3$, or MeOH), using a diamond attenuated total reflection (ATR) module, as indicated. The spectra were recorded on a Bruker Tensor 27 spectrometer. Absorption maxima are reported in wavenumbers (cm^{-1}) and reported as s (strong), m (medium), w (weak), or br (broad). Only the major peaks are reported.

Specific optical rotations were measured using a Schmidt + Haensch Unipol L2000 polarimeter, using a sodium lamp at 589 nm and a path length of 1.0 dm. The concentration (*c*) is expressed in g/100 mL (equivalent to g/0.1 dm^3). Specific rotations are denoted $[\alpha]_D^{\mathrm{T}}$ and are given in implied units of 10^{-1} deg cm^2 g^{-1} at the temperature, T, in °C, stated.

Analytical high-performance liquid chromatography (HPLC) was carried out on a PerkinElmer Flexar system with a Binary LC Pump and UV/Vis LC Detector set at 220 or 254 nm. For determination of compound purity, a Dionex Acclaim® 120 column (RP, C_{18}, 5 μm, 120 Å, 4.6 × 150 mm) was employed using a flow rate of 1.0 or 1.5 mL min^{-1}. Samples were injected in H_2O, MeOH, MeCN, or H_2O:DMSO 9:1. The methods used were as described below:

Method 1

Start Time (min)	End Time (min)	Flow Rate (mL min^{-1})	Start %(A)	Finish %(A)
0	20	1.5	95	5
20	25	1.5	5	5

Mobile phases: A = H_2O + 0.1% TFA; B = MeCN + 0.1% TFA.

Method 2

Start Time (min)	End Time (min)	Flow Rate (mL min^{-1})	Start %(A)	Finish %(A)
0	10	1.5	95	5
10	15	1.5	5	5

Mobile phases: A = H_2O + 0.1% TFA; B = MeCN + 0.1% TFA.

When reporting data for compounds which are a mixture of two diastereoisomers A and B, peaks in the analytical HPLC spectra are assigned in accordance with the A:B ratio obtained from ^{1}H NMR spectra.

Enantiomeric purity was determined by **chiral HPLC** on a PerkinElmer Flexar system with a Binary LC Pump and UV/Vis LC Detector using a ChiralPak® AD-H column (5 μm, 4.6 × 150 mm) with eluents as follows: either IPA (A) and hexane (B), or EtOH + 0.1% TFA (A) and hexane + 0.1% TFA (B). Samples were injected in

IPA or IPA:MeOH 7:1 (when A = IPA), or EtOH (when A = EtOH + 0.1% TFA). Isocratic methods (A:B) were used as indicated, with detection at 220 nm. For each compound, the stereochemistry was referenced to a sample of racemic or scalemic compound. For each of the compounds which form diastereoisomers (OXFKEF13 (**32**), OXFKEF14 (**33**), OXFKEF17 (**35**), and OXFKEF16 (**48**)) the quoted *e.e.* value is the difference between the total proportion of the two diastereoisomers with L- stereochemistry at the Cys-α stereogenic centre and the total proportion of the two diastereoisomers with D- stereochemistry at the Cys-α stereogenic centre. For all other compounds, which do not form diastereoisomers, *e.e.* values were calculated using the following equation:

$$e.e. = \frac{|R - S|}{R + S} \times 100\%$$

where *R* and *S* correspond to the proportions of the individual enantiomers and *R* + *S* = 1.

Semi-preparative HPLC was performed using an Agilent ZORBAX 300SB-C_{18} column (5 µm, 9.4 mm × 250 mm) on a DIONEX P680 HPLC pump at a flow rate of 2.0 mL min^{-1} and detection at 225 and 250 nm. Samples were injected in H_2O. The method used was as described below:

Start Time (min)	End Time (min)	Start %(A)	Finish %(A)
0	3	95	95
3	18	95	5
18	20	5	95
20	25	95	95

Mobile phases: A = H_2O:MeCN 95:5; B = H_2O:MeCN 5:95.

Normal phase silica gel analytical thin layer chromatography (NP TLC) was carried out on normal phase Merck silica gel 60 F_{254} aluminium-supported TLC sheets. Visualisation was by absorption of UV light (λ_{max} = 254 or 365 nm) or thermal development after staining with either an aqueous (aq.) solution of $KMnO_4$ or an ethanolic solution of ninhydrin.

Reverse phase silica gel analytical thin layer chromatography (RP TLC) was carried out on Merck silica gel 60 RP-18 F_{254}s aluminium-supported TLC sheets. Visualisation was by absorption of UV light (λ_{max} = 254 or 365 nm) or thermal development after staining with either an aq. solution of $KMnO_4$ or an ethanolic solution of ninhydrin.

Normal phase (NP) silica gel flash column chromatography was carried out manually on Merck Geduran® silica gel 60 (40–63 μm), eluting with solvents as supplied, under a positive pressure of compressed N_2 (g).

Reverse phase (RP) silica gel flash column chromatography was either carried out manually on C_{18}-reversed phase silica gel, fully end-capped (Sigma-Aldrich, #60756 (15–25 μm, 100 Å) or #60757 (40–63 μm, 90 Å)) eluting with the appropriate solvents as supplied, under a positive pressure of compressed N_2 (g), or performed on a Biotage SP4 System using Biotage® SNAP KP-C_{18}-HS cartridges.

Anhydrous solvents (CH_2Cl_2, DMF, Et_3N, pyridine, and THF) were obtained through an MBRAUN MB5 SPS (solvent purification system), stored in flame-dried round-bottomed flasks over 3 Å molecular sieves under an inert atmosphere of Ar, and used for up to 2 weeks from when they were first obtained. Redistilled piperidine was purchased from Sigma-Aldrich.

Chemicals were obtained from Fluorochem, AnaSpec, Sigma-Aldrich, Acros Organics, Alfa Aesar, MP Biomedicals, Apollo, Fisher Scientific, and BDH.

Reagents and solvents used were of commercially available reagent grade quality and were used without further purification. All non-aqueous reactions were carried out in flame-dried round-bottomed flasks under an inert atmosphere of Ar.

Concentrated *in vacuo* refers to removal of solvent on a Büchi™ rotary evaporator under reduced pressure in a water bath at 40 °C.

Celite® refers to Celite® 545 filter aid, treated with sodium carbonate, flux-calcined (Sigma-Aldrich).

Brine refers to a saturated aq. solution of sodium chloride.

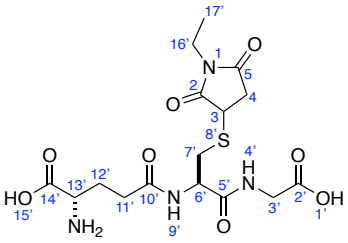
Petroleum ether refers to the fraction in the boiling point range 40–60 °C.

Hexane refers to a mixture of hexane isomers.

Lyophilisation refers to the removal of H_2O from aq. solutions by freeze-drying using a CHRIST-ALPHA 2-4-LD lyophiliser.

6.2 Chemistry Experimental Procedures

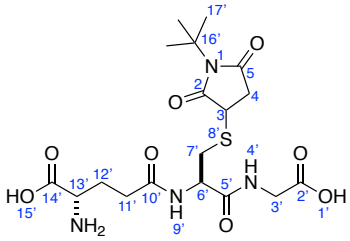
L-γ-Glutamyl-*S*-(1-ethyl-2,5-dioxopyrrolidin-3-yl)-L-cysteinylglycine, *N*-ethylsuccinimido-*S*-glutathione (ESG) (20)[31,68]



Aq. NaOH (1.0 M, 0.32 mL, 0.32 mmol, 1.0 eq.) was added to a solution of L-glutathione, **14**, (0.10 g, 0.32 mmol, 1.0 eq.) in H_2O (2.2 mL), followed by NEM, **19**, (42 mg, 0.33 mmol, 1.0 eq.). The resulting colourless reaction mixture was stirred at rt for 5 h. After this time, TLC analysis (EtOAc:IPA:H_2O 1:3:4, NP) indicated complete consumption of L-glutathione (**14**). The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel flash column chromatography (H_2O:MeOH 70:30 → 50:50). A second purification (H_2O:MeOH 100:0 → 90:10 → 80:20) gave the title compound (**20**) as a colourless, hygroscopic solid in a 1:1.2 ratio (^{1}H NMR) of 2 diastereoisomers A and B (72 mg, 51%): R_f 0.37 (EtOAc:IPA:H_2O 1:1:1, NP); mp 193–195 °C (dec., from H_2O) [lit.[68] 190–191 °C (dec., from EtOH/H_2O)]; $[\alpha]_D^{25}$ = −15.0 (*c* 1.0, H_2O) [lit.[31] $[\alpha]_D^{20}$ = −12.3 (*c* 1.0, H_2O)]; $\bar{\nu}_{max}$ (neat)/cm^{-1} 3254 (N–H, w, br), 3059 (O–H, w, br), 2936 (C–H, w), 1694 (C=O (carboxylic acid), m), 1638 (C=O (amide, succinimide), s), 1594 (N–H (amide), s), 1528 (s), 1401 (s), 1348 (s), 1225 (s), 1127 (m), 790 (m), 685 (s), 607 (s); ^{1}H NMR (400 MHz; D_2O) (diastereoisomer A) δ 4.69 (1H, dd, J_{BX} 8.2, J_{AX} 5.1, *H*-6'$_X$), 4.06 (1H, dd, J_{AX} 9.1, J_{BX} 4.2, *H*-3$_X$), 3.82–3.71 (3H, m, H_aH_b-3', *H*-13'$_X$), 3.53 (2H, q, *J* 7.3, H_2-16'), 3.32–3.22 (2H, m, H_aH_b-4, H_aH_b-7'), 3.16 (1H, dd, J_{AB} 14.1, J_{BX} 8.2, H_aH_b-7'), 2.64 (1H, dd,

J_{AB} 8.0, J_{BX} 4.2, H_aH_b-4), 2.61–2.47 (2H, m, H_aH_b-11'), 2.20–2.12 (2H, m, H_aH_b-12'), 1.12 or 1.11 (3H, t, J 7.3, H_3-17'); ^{1}H NMR (400 MHz; D_2O) (diastereoisomer B) δ 4.67 (1H, dd, J_{BX} 9.2, J_{AX} 5.0, H-6'$_X$), 4.04 (1H, dd, J_{AX} 9.0, J_{BX} 4.2, H-3$_X$), 3.82–3.71 (3H, m, H_aH_b-3', H-13'$_X$), 3.53 (2H, q, J 7.3, H_2-16'), 3.36 (1H, dd, J_{AB} 14.1, J_{AX} 5.0, H_aH_b-7'), 3.32–3.22 (1H, m, H_aH_b-4), 3.01 (1H, dd, J_{AB} 14.1, J_{BX} 9.2, H_aH_b-7'), 2.71 (1H, dd, J_{AB} 8.0, J_{BX} 4.2, H_aH_b-4), 2.61–2.47 (2H, m, H_aH_b-11'), 2.20–2.12 (2H, m, H_aH_b-12'), 1.12 or 1.11 (3H, t, J 7.3, H_3-17'); *m/z* (ESI$^-$) 431 ([M−H]$^-$, 100%), 453 ([M+Na−2H]$^-$, 29%); Analytical HPLC (method 1, 220/254 nm) retention (ret.) time = 5.5 min, area = 44.9% (diastereoisomer A), ret. time = 5.7 min, area = 55.1% (diastereoisomer B). These data are in agreement with the available literature values.[31,68]

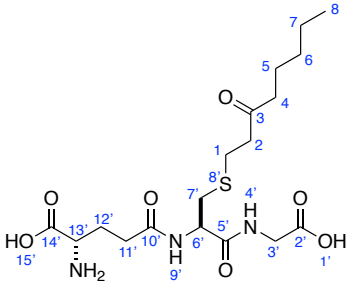
L-γ-Glutamyl-*S*-(1-*tert*-butyl-2,5-dioxopyrrolidin-3-yl)-L-cysteinylglycine, *N*-*tert*-butylsuccinimido-*S*-glutathione (^{t}BuSG) (26)[31,44]



N-*tert*-Butylmaleimide, **42**, (94 µL, 0.65 mmol, 1.0 eq.) was added to a solution of L-glutathione, **14**, (0.20 g, 0.65 mmol, 1.0 eq.) and NaOH (26 mg, 0.65 mmol, 1.0 eq.) in H_2O (5.0 mL). The resulting colourless reaction mixture was stirred at rt for 30 min. After this time, TLC analysis (EtOAc:IPA:H_2O 1:1:1, NP) indicated complete consumption of L-glutathione (**14**). The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel flash column chromatography using a Biotage system (30 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeCN 99:1 → 90:10, 15 column volumes (CV)) to give the title compound (**26**) as a colourless, hygroscopic

solid in a 1:1.2 ratio (^{1}H NMR) of 2 diastereoisomers A and B (0.10 g, 33%): R_f 0.40 (EtOAc:IPA:H_2O 1:1:1, NP); mp 184–189 °C (from H_2O, wide range because compound is hygroscopic); $[\alpha]_D^{25}$ = −20.3 (*c* 1.0, H_2O) [lit.[44] $[\alpha]_D^{20}$ = −13.3 (*c* 1.0, H_2O)]; ^{1}H NMR (400 MHz; D_2O) (diastereoisomer A) δ 4.67 (1H, dd, J_{BX} 8.7, J_{AX} 5.0, *H*-6'$_X$), 3.90 (1H, dd, J_{AX} 9.3, J_{BX} 4.3, *H*-3$_X$), 3.83–3.71 (3H, m, H_aH_b-3', *H*-13'$_X$), 3.24 (1H, dd, J_{AB} 14.1, J_{AX} 5.0, H_aH_b-7'), 3.20–3.09 (2H, m, H_aH_b-4, H_aH_b-7'), 2.57–2.48 (3H, m, H_aH_b-4, H_aH_b-11'), 2.20–2.12 (2H, m, H_aH_b-12'), 1.54 or 1.53 (9H, s, $(H_3)_3$-17'); ^{1}H NMR (400 MHz; D_2O) (diastereoisomer B) δ 4.66 (1H, dd, J_{BX} 9.1, J_{AX} 4.8, *H*-6'$_X$), 3.87 (1H, dd, J_{AX} 9.3, J_{BX} 4.3, *H*-3$_X$), 3.83–3.71 (3H, m, H_aH_b-3', *H*-13'$_X$), 3.34 (1H, dd, J_{AB} 14.1, J_{AX} 4.8, H_aH_b-7'), 3.20–3.09 (1H, m, H_aH_b-4), 2.99 (1H, dd, J_{AB} 14.1, J_{BX} 9.1, H_aH_b-7'), 2.58 (1H, dd, J_{AB} 6.7, J_{BX} 4.3, H_aH_b-4), 2.57–2.48 (2H, m, H_aH_b-11'), 2.20–2.12 (2H, m, H_aH_b-12'), 1.54 or 1.53 (9H, s, $(H_3)_3$-17'); *m/z* (ESI$^+$) 461 ([M+H]$^+$, 100%), 483 ([M+Na]$^+$, 60%); Analytical HPLC (method 1, 220/254 nm) ret. time = 7.6 min, area = 43.1% (diastereoisomer A), ret. time = 7.8 min, area = 56.9% (diastereoisomer B). These data are in agreement with the available literature values.[31,44]

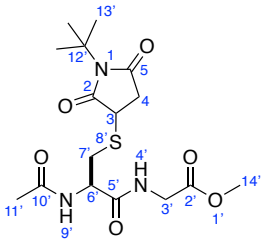
L-γ-Glutamyl-*S*-(3-oxooctyl)-L-cysteinylglycine (29)[31]



NaOH (26 mg, 0.64 mmol, 2.0 eq.) was added to a solution of L-glutathione, **14**, (0.20 g, 0.64 mmol, 2.0 eq.) in H_2O:MeOH 1:1 (5.0 mL). The resulting solution was stirred at rt for 5 min. 1-Octen-3-one, **41**, (50% *w/w* in 1-octen-3-ol, 97 µL, 0.32 mmol,

1.0 eq.) was then added, and the reaction mixture was stirred at rt for 42 h. The MeOH was removed *in vacuo* and the reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel flash column chromatography using a Biotage system (12 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeCN 99:1 → 80:20, 15 CV). A second purification (12 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeCN 99:1 → 90:10, 20 CV) gave the title compound (**29**) as a colourless, hygroscopic solid (32 mg, 23%): R_f 0.59 (EtOAc:IPA:H_2O 1:1:1, NP); $[\alpha]_D^{25}$ = –27.7 (*c* 1.0, H_2O) [lit.[31] $[\alpha]_D^{20}$ = −26.4 (*c* 0.25, H_2O)]; mp 185–190 °C (dec., from H_2O, wide range because compound is hygroscopic); ^{1}H NMR (500 MHz; D_2O) δ 4.58 (1H, dd, J_{BX} 8.8, J_{AX} 4.9, *H*-6'$_X$), 3.82–3.71 (3H, m, H_aH_b-3', *H*-13'$_X$), 3.08 (1H, dd, J_{AB} 14.1, J_{AX} 4.9, H_aH_b-7'), 2.91–2.84 (3H, m, H_aH_b-2, H_aH_b-7'), 2.81–2.74 (2H, m, H_aH_b-1), 2.57–2.49 (4H, m, H_aH_b-4, H_aH_b-11'), 2.20–2.11 (2H, m, H_aH_b-12'), 1.59–1.51 (2H, m, H_aH_b-5), 1.34–1.21 (4H, m, H_aH_b-6, H_aH_b-7), 0.86 (3H, t, *J* 7.0, H_3-8); *m/z* (ESI$^-$) 432 ([M–H]$^-$, 100%); Analytical HPLC (method 2, 220 nm (compound not detectable at 254 nm)) ret. time = 6.6 min, purity = 100%. These data are in agreement with the available literature values.[31]

Methyl *N*-acetyl-*S*-(1-*tert*-butyl-2,5-dioxopyrrolidin-3-yl)-L-cysteinylglycinate, OXFKEF12 (31)[44,45]



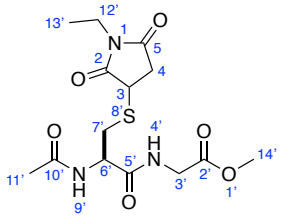
Triethylsilane (0.40 mL, 2.5 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate, **45**, (0.40 g, 0.84 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (30 mL) at rt under Ar. The resulting reaction mixture

was stirred at rt under Ar for 2.5 h. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**). The reaction mixture was concentrated *in vacuo*. The resulting residue was purified by RP silica gel flash column chromatography (H_2O) to give *N*-acetyl-L-cysteinylglycine methyl ester (**46**) as a colourless, hygroscopic foam (0.16 g), which was used without further purification.

N-*tert*-Butylmaleimide, **42**, (33 µL, 0.23 mmol, 1.0 eq.) was added to a solution of *N*-acetyl-L-cysteinylglycine methyl ester, **46**, (80 mg, 0.34 mmol, 1.5 eq.) in H_2O (2.0 mL). The resulting reaction mixture was stirred at rt for 46 h. After this time, ^{1}H NMR analysis (D_2O) indicated complete consumption of *N*-*tert*-butylmaleimide (**42**). The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel flash column chromatography using a Biotage system (12 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeCN 95:5 → 5:95, 19 CV) to give the title compound (**31**) as a yellow, hygroscopic solid, in a 1:1.2 ratio (^{1}H NMR) of 2 diastereoisomers A and B (58 mg, 52% (over 2 steps)): R_f 0.68 (H_2O:MeCN 1:3, RP); $[\alpha]_D^{25}$ = −14.6 (*c* 1.0, H_2O) [lit.[44] $[\alpha]_D^{20}$ = −14.5 (*c* 0.55, H_2O)]; mp 47–50 °C (from H_2O); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer A) δ 4.66 (1H, dd, J_{BX} 7.3, J_{AX} 5.8, *H*-6'$_X$), 4.09–4.00 (2H, m, H_aH_b-3'), 3.91 (1H, dd, J_{AX} 9.2, J_{BX} 4.2, *H*-3$_X$), 3.76 (3H, s, H_3-14'), 3.24–3.11 (3H, m, H_aH_b-4, H_aH_b-7'), 2.55 (1H, dd, J_{AB} 4.2, J_{BX} 4.2, H_aH_b-4), 2.08 (3H, s, H_3-11'), 1.55 (9H, s, $(H_3)_3$-13'); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer B) δ 4.64 (1H, dd, J_{BX} 8.6, J_{AX} 5.4, *H*-6'$_X$), 4.09–4.00 (2H, m, H_aH_b-3'), 3.88 (1H, dd, J_{AX} 9.3, J_{BX} 4.3, *H*-3$_X$), 3.76 (3H, s, H_3-14'), 3.30 (1H, dd, J_{AB} 14.0, J_{AX} 5.4, H_aH_b-7'), 3.24–3.11 (1H, m, H_aH_b-4), 3.01 (1H, dd, J_{AB} 14.0, J_{BX} 8.6, H_aH_b-7'), 2.58 (1H, dd, J_{AB} 4.3, J_{BX} 4.3, H_aH_b-4), 2.08 (3H, s, H_3-11'), 1.55 (9H, s, $(H_3)_3$-13'); *m/z* (ESI$^+$) 410 ([M+Na]$^+$, 100%);

Analytical HPLC (method 1, 220/254 nm) ret. time = 9.0 min, area = 42.8% (diastereoisomer A), ret. time = 9.1 min, area = 52.0% (diastereoisomer B). These data are in agreement with the available literature values.[44,45]

Methyl *N*-acetyl-*S*-(1-ethyl-2,5-dioxopyrrolidin-3-yl)-L-cysteinylglycinate, OXFKEF13 (32)[45]

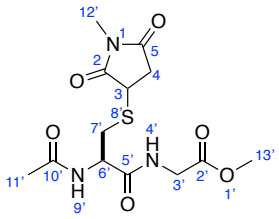


Triethylsilane (1.3 mL, 8.1 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate, **45**, (1.3 g, 2.7 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (97 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 25 min. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (100 mL) and extracted with H_2O (5 × 20 mL). The aq. components were combined and lyophilised to give *N*-acetyl-L-cysteinylglycine methyl ester (**46**) as a colourless, hygroscopic foam (0.61 g), which was used without further purification.

NEM, **19**, (0.14 g, 1.1 mmol, 1.0 eq.) was added to a solution of *N*-acetyl-L-cysteinylglycine methyl ester, **46**, (0.39 g, 1.7 mmol, 1.5 eq.) in H_2O (9.6 mL). The resulting reaction mixture was stirred at rt for 15 h. After this time, TLC analysis (H_2O:MeCN 4:1, RP) indicated complete consumption of NEM (**19**). The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel flash column chromatography using a Biotage system (30 g Biotage® SNAP KP-C_{18}-HS

cartridge, H_2O:MeCN 95:5 → 60:40, 19 CV) to give the title compound (**32**) as a colourless, hygroscopic solid, in a 1:1.2 ratio (^{1}H NMR) of 2 diastereoisomers A and B (0.16 g, 38% (over 2 steps)): R_f 0.38 (H_2O:MeCN 4:1, RP); $[\alpha]_D^{25}$ = −14.8 (*c* 1.0, H_2O) [lit.[45] $[\alpha]_D^{25}$ = −18.1 (*c* 0.38, H_2O)]; mp 133–141 °C (from H_2O, wide range because compound is hygroscopic); ^{1}H NMR (400 MHz; D_2O) (diastereoisomer A) δ 4.65 (1H, dd, J_{BX} 7.3, J_{AX} 5.7, *H*-6'$_X$), 4.07–4.00 (3H, m, *H*-3$_X$, H_aH_b-3'), 3.73 (3H, s, H_3-14'), 3.53 (2H, q, *J* 7.3, H_2-12'), 3.35–3.22 (1H, m, H_aH_b-4), 3.21 (1H, dd, J_{AB} 13.9, J_{AX} 5.7, H_aH_b-7'), 3.16 (1H, dd, J_{AB} 13.9, J_{BX} 7.3, H_aH_b-7'), 2.65 (1H, dd, J_{AB} 5.3, J_{BX} 4.3, H_aH_b-4), 2.06 (3H, s, H_3-11'), 1.11 (3H, t, *J* 7.3, H_3-13'); ^{1}H NMR (400 MHz; D_2O) (diastereoisomer B) δ 4.63 (1H, dd, J_{BX} 8.7, J_{AX} 5.2, *H*-6'$_X$), 4.07–4.00 (3H, m, *H*-3$_X$, H_aH_b-3'), 3.73 (3H, s, H_3-14'), 3.53 (2H, q, *J* 7.3, H_2-12'), 3.35–3.22 (2H, m, H_aH_b-4, H_aH_b-7'), 3.01 (1H, dd, J_{AB} 14.1, J_{BX} 8.7, H_aH_b-7'), 2.69 (1H, dd, J_{AB} 5.3, J_{BX} 4.3, H_aH_b-4), 2.06 (3H, s, H_3-11'), 1.11 (3H, t, *J* 7.3, H_3-13'); *m/z* (ESI$^+$) 382 ([M+Na]$^+$, 100%); Analytical HPLC (method 1, 220/254 nm) ret. time = 6.9 min, area = 58.1% (diastereoisomer B), ret. time = 7.1 min, area = 38.2% (diastereoisomer A); Chiral HPLC (90:10 hexane:IPA, 1.8 mL min^{-1}, 60 min) ret. time = 49.5 min (L-**32**), 66.4% *e.e.* (other enantiomer observed, ret. time = 43.3 min, area = 16.8% (D-**32** = **48**)). These data are in agreement with the available literature values.[45]

Methyl *N*-acetyl-*S*-(1-methyl)-2,5-dioxopyrrolidin-3-yl)-L-cysteinylglycinate, OXFKEF14 (33)[45]

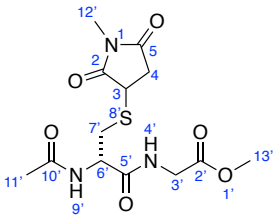


Triethylsilane (0.19 mL, 1.2 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate, **45**, (0.19 g, 0.40 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (15 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 1.5 h. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**). The reaction mixture was concentrated *in vacuo*. The resulting residue was purified by RP silica gel flash column chromatography (H_2O) to give *N*-acetyl-L-cysteinylglycine methyl ester (**46**) as a colourless, hygroscopic foam (44 mg), which was used without further purification.

N-Methylmaleimide, **47**, (7.0 mg, 0.063 mmol, 1.0 eq.) was added to a solution of *N*-acetyl-L-cysteinylglycine methyl ester, **46**, (21 mg, 0.090 mmol, 1.4 eq.) in H_2O (1.6 mL). The resulting reaction mixture was stirred at rt for 24 h. After this time, TLC analysis (H_2O:MeCN 3:2, RP) indicated complete consumption of *N*-methylmaleimide (**47**). The reaction mixture was lyophilised to obtain the crude product, which was purified by semi-preparative HPLC to give the title compound (**33**) as a colourless, hygroscopic solid in a 1:1.1 ratio (^{1}H NMR) of 2 diastereoisomers A and B (12 mg, 26% (over 2 steps)): R_f 0.67 (H_2O:MeCN 3:2, RP); $[\alpha]_D^{25}$ = −12.8 (*c* 1.0, H_2O) [lit.[45] $[\alpha]_D^{25}$ = −14.4 (*c* 0.17, H_2O)]; mp 40–50 °C (from H_2O, wide range because compound is hygroscopic); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer A) δ 4.67 (1H,

dd, J_{BX} 7.2, J_{AX} 5.7, *H*-6'$_X$), 4.09 (1H, dd, J_{AX} 8.5, J_{BX} 4.5, *H*-3$_X$), 4.06–4.02 (2H, m, H_aH_b-3'), 3.76 (3H, s, H_3-13'), 3.37–3.26 (1H, m, H_aH_b-4), 3.23 (1H, dd, J_{AB} 14.2, J_{AX} 5.7, H_aH_b-7'), 3.19 (1H, dd, J_{AB} 14.2, J_{BX} 7.2, H_aH_b-7'), 2.98 (3H, s, H_3-12'), 2.68 (1H, dd, J_{AB} 9.4, J_{BX} 4.5, H_aH_b-4), 2.08 (3H, s, H_3-11'); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer B) δ 4.65 (1H, dd, J_{BX} 8.5, J_{AX} 5.0, *H*-6'$_X$), 4.07 (1H, dd, J_{AX} 8.5, J_{BX} 4.4, *H*-3$_X$), 4.06–4.02 (2H, m, H_aH_b-3'), 3.76 (3H, s, H_3-13'), 3.37–3.26 (2H, m, H_aH_b-4, H_aH_b-7'), 3.04 (1H, dd, J_{AB} 14.1, J_{BX} 8.5, H_aH_b-7'), 2.98 (3H, s, H_3-12'), 2.72 (1H, dd, J_{AB} 9.5, J_{BX} 4.4, H_aH_b-4), 2.08 (3H, s, H_3-11'); *m/z* (ESI$^+$) 368 ([M+Na]$^+$, 100%); Analytical HPLC (method 1, 220/254 nm) ret. time = 5.9 min, area = 45.1% (diastereoisomer A), ret. time = 6.0 min, area = 49.9% (diastereoisomer B); Chiral HPLC (90:10 hexane:IPA, 1.5 mL min^{-1}, 120 min) ret. time = 77.8 min (L-**33**), ~71.2% *e.e.* (peaks not fully separated, other enantiomer observed, ret. time = 68.5 min, area ~14.4% (D-**33** = **35**)). These data are in agreement with the available literature values.[45]

Methyl *N*-acetyl-*S*-(1-methyl)-2,5-dioxopyrrolidin-3-yl)-D-cysteinylglycinate, OXFKEF17 (35)[45]



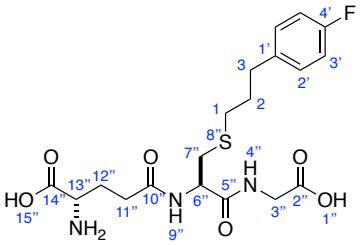
Triethylsilane (0.24 mL, 1.5 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-D-cysteinylglycinate, **52**, (0.24 g, 0.50 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (18 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 30 min. After this time, TLC analysis (EtOAc) indicated

complete consumption of the starting material (**52**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (25 mL) and extracted with H_2O (3 × 20 mL). The combined aq. components were lyophilised to give *N*-acetyl-D-cysteinylglycine methyl ester (**53**) as a colourless, hygroscopic foam (97 mg), which was used without further purification.

N-Methylmaleimide, **47**, (22 mg, 0.20 mmol, 1.0 eq.) was added to a solution of *N*-acetyl-D-cysteinylglycine methyl ester, **53**, (90 mg, 0.38 mmol, 1.9 eq.) in H_2O (7.1 mL). The resulting reaction mixture was stirred at rt for 3 h. After this time, TLC analysis (H_2O:MeCN 3:2, RP) indicated that some *N*-methylmaleimide (**47**) remained. The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel column chromatography (H_2O:MeCN 100:0 → 95:5 → 90:10 → 85:15) to give the title compound (**35**) as a colourless, hygroscopic solid in a 1:1.1 ratio (^{1}H NMR) of 2 diastereoisomers A and B (40 mg, 46% (over 2 steps)): R_f 0.39 (H_2O:MeCN 3:2, RP); $[\alpha]_D^{25}$ = +28.4 (*c* 1.0, H_2O) [lit.[45] $[\alpha]_D^{25}$ = +33.3 (*c* 0.5, MeOH)]; mp 48–57 °C (from H_2O, wide range because compound is hygroscopic); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer A) δ 4.68 (1H, dd, J_{BX} 7.3, J_{AX} 5.7, *H*-6'$_X$), 4.09 (1H, dd, J_{AX} 8.5, J_{BX} 4.1, *H*-3$_X$), 4.05–4.02 (2H, m, H_aH_b-3'), 3.76 (3H, s, H_3-13'), 3.37–3.26 (1H, m, H_aH_b-4), 3.23 (1H, dd, J_{AB} 14.3, J_{AX} 5.7, H_aH_b-7'), 3.18 (1H, dd, J_{AB} 14.3, J_{BX} 7.3, H_aH_b-7'), 2.98 (3H, s, H_3-12'), 2.68 (1H, dd, J_{AB} 9.3, J_{BX} 4.1, H_aH_b-4), 2.08 (3H, s, H_3-11'); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer B) δ 4.65 (1H, dd, J_{BX} 8.4, J_{AX} 5.0, *H*-6'$_X$), 4.07 (1H, dd, J_{AX} 8.3, J_{BX} 4.1, *H*-3$_X$), 4.05–4.02 (2H, m, H_aH_b-3'), 3.76 (3H, s, H_3-13'), 3.37–3.26 (2H, m, H_aH_b-4, H_aH_b-7'), 3.04 (1H, dd, J_{AB} 14.1, J_{BX} 8.4, H_aH_b-7'), 2.98 (3H, s, H_3-12'), 2.72 (1H, dd, J_{AB} 9.3, J_{BX} 4.1, H_aH_b-4), 2.08 (3H, s, H_3-11'); *m/z* (ESI$^+$) 368 ([M+Na]$^+$, 100%);

Analytical HPLC (method 1, 220/254 nm) ret. time = 5.4 min, area = 45.8% (diastereoisomer A), ret. time = 5.6 min, area = 50.7% (diastereoisomer B); Chiral HPLC (90:10 hexane:IPA, 1.5 mL min^{-1}, 120 min) ret. time = 68.5 min (D-**35**), ~6.0% *e.e.* (peaks not fully separated, other enantiomer observed, ret. time = 77.8 min, area ~47.0% (L-**35** = **33**)). These data are in agreement with the available literature values.[45]

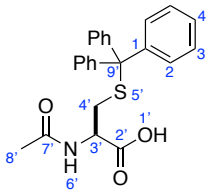
L-γ-Glutamyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycine (40)[48]



Three identical reactions were set up as follows: 1-allyl-4-fluorobenzene (92 µL, 0.68 mmol, 1.0 eq.) was added to a suspension of L-glutathione, **14**, (0.84 g, 2.7 mmol, 4.0 eq.) and DPAP (36 mg, 0.14 mmol, 0.2 eq.) in THF:H_2O 1:2 (6.0 mL). The resulting reaction mixtures were stirred in the presence of UV light (λ = 365 nm) at rt for 5 h. The three mixtures were combined and filtered to give the crude product, which was washed with H_2O and EtOH, and dried *in vacuo*. The crude solid was purified by crystallisation from boiling H_2O. This compound was adsorbed onto Celite® and further purified by RP silica gel flash column chromatography using a Biotage system (12 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeOH:TFA 75:25:0.1 → 0:100:0.1, 20 CV). A second purification (12 g Biotage® SNAP KP-C_{18}-HS cartridge, H_2O:MeCN:TFA 90:10:0.1 → 50:50:0.1) gave the title compound (**40**) as a colourless, hygroscopic solid (43 mg, 5%): R_f 0.30 (H_2O:MeOH 1:1, RP); $[\alpha]_D^{25}$ = −21.2 (*c* 0.5, DMSO)

[lit.[48] $[\alpha]_D^{20}$ = −6.2 (*c* 0.5, 2 M aq. NaOH)]; mp 213–221 °C (dec., from H_2O/DMSO, wide range because compound is hygroscopic) [lit.[48] 170–173 °C (from EtOH/H_2O)]; ^{1}H NMR (400 MHz; D_6-DMSO) δ 8.69 (1H, dd, J_{AX} 5.7, J_{BX} 5.6, *H*-4''$_X$), 8.35 (1H, d, *J* 8.8, *H*-9''), 7.23 (2H, dd, *J* 8.7, 5.7, *H*-2'), 7.09 (2H, dd, J_{HF} 8.9, *J* 8.7, *H*-3'), 4.42 (1H, ddd, J_{BX} 9.8, *J* 8.8, J_{AX} 4.4, *H*-6''$_X$), 3.73 (1H, dd, J_{AB} 17.3, J_{AX} 5.7, H_aH_b-3''), 3.68 (1H, dd, J_{AB} 17.3, J_{BX} 5.6, H_aH_b-3''), 3.38 (1H, dd, *J* 6.5, 6.5, *H*-13''$_X$), 2.89 (1H, dd, J_{AB} 13.6, J_{AX} 4.4, H_aH_b-7''), 2.70–2.58 (3H, m, H_aH_b-3, H_aH_b-7''), 2.53–2.47 (2H, m, H_aH_b-1), 2.39–2.23 (2H, m, H_aH_b-11'), 1.99–1.83 (2H, m, H_aH_b-12''), 1.83–1.72 (2H, m, H_aH_b-2); ^{19}F NMR (470 MHz; D_6-DMSO) δ −117.6 (*F*-4'); *m/z* (ESI$^+$) 444 ([M+H]$^+$, 11%), 466 ([M+Na]$^+$, 5%); Analytical HPLC (method 2, 220/254 nm) ret. time = 6.8 min, purity = 98.3%. These data are in agreement with the available literature values.[48]

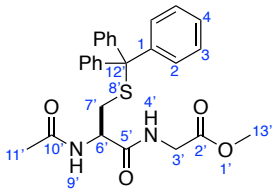
_N_-Acetyl-_S_-(triphenylmethyl)-L-cysteine (44)[44,45,47,62,69]



Triphenylmethyl chloride (6.41 g, 23.0 mmol, 1.50 eq.) was added to a solution of *N*-acetyl-L-cysteine, **43**, (2.50 g, 15.3 mmol, 1.00 eq.) in DMF (40.0 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 21 h. After this time, TLC analysis (EtOAc:IPA:H_2O 2:2:1, NP) indicated consumption of *N*-acetyl-L-cysteine (**43**). The reaction mixture was quenched with H_2O (2 mL) and concentrated to 20 mL *in vacuo*. The mixture was extracted with Et_2O (350 mL) and washed with 10% *w/v* aq. LiCl solution (1 × 250 mL, 3 × 110 mL). The organic component was separated, and the combined aq. components were extracted with Et_2O (4 × 200 mL). The combined

organic components were dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo* to obtain the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (EtOAc:MeOH 100:0 → 95:5) to give the title compound (**44**) as a colourless solid (5.62 g, 91%); R_f 0.50 (EtOAc:IPA:H_2O 2:2:1, NP); $[\alpha]_D^{25}$ = +33.2 (*c* 1.0, MeOH) [lit.[47] $[\alpha]_D^{25}$ = +34.5 (*c* 1.0, MeOH)]; mp 78–81 °C (from CH_2Cl_2) [lit.[45] 78–81 °C (from petroleum ether/EtOAc)]; ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.45–7.38 (6H, m, *H*-2), 7.32–7.25 (6H, m, *H*-3), 7.25–7.19 (3H, m, *H*-4), 5.81 (1H, d, *J* 7.0, *H*-6'), 4.43–4.35 (1H, m, *H*-3'$_X$), 2.78 (1H, dd, J_{AB} 12.9, J_{AX} 6.4, H_aH_b-4'), 2.70 (1H, dd, J_{AB} 12.9, J_{BX} 4.6, H_aH_b-4'), 1.95 (3H, s, H_3-8'); *m/z* (ESI$^-$) 404 ([M−H]$^-$, 100%); Analytical HPLC (method 2, 220/254 nm) ret. time = 7.1 min, purity = 98.9%; Chiral HPLC (90:10 hexane:IPA, 1.9 mL min^{-1}, 30 min) ret. time = 5.4 min (L-**44**), 98.0% *e.e.* (other enantiomer observed, ret. time = 4.6 min, area = 1.0% (D-**44**)). These data are in agreement with the available literature values.[44,45,47,62,69]

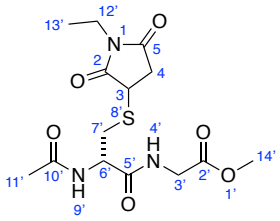
Methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate (45)[44,45]



Glycine methyl ester hydrochloride (730 mg, 5.78 mmol, 1.00 eq.) was added to a solution of *N*-acetyl-*S*-triphenylmethyl-L-cysteine, **44**, (2.46 g, 6.07 mmol, 1.05 eq.) in anhydrous DMF (13.0 mL) at rt under Ar, followed by HBTU (2.30 g, 6.07 mmol, 1.05 eq.). The resulting reaction mixture was stirred at rt under Ar for 5 min. After this time, DIPEA (2.04 mL, 11.7 mmol, 2.00 eq.) was added and the reaction mixture was stirred at rt under Ar for 23 h. After this time, TLC analysis (EtOAc:EtOH 3:2) indicated

complete consumption of glycine methyl ester hydrochloride. The reaction mixture was extracted with EtOAc (350 mL) and washed with 10% *w/v* aq. LiCl solution (5 × 100 mL). The organic component was separated, and the combined aq. components were extracted with EtOAc (3 × 250 mL). The combined organic components were dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo* to obtain the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (petroleum ether:EtOAc 40:60 → 20:80 → 10:90 → 0:100) to give the title compound (**45**) as a colourless solid (2.57 g, 93%): R_f 0.37 (EtOAc); $[\alpha]_D^{25}$ = +1.7 (*c* 1.0, $CHCl_3$) [lit.[44] $[\alpha]_D^{20}$ = +1.9 (*c* 1.0, $CHCl_3$)]; mp 59–62 °C (from CH_2Cl_2) [lit.[44] 59–63 °C (from $CHCl_3$)]; ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.42–7.35 (6H, m, *H*-2), 7.28–7.21 (6H, m, *H*-3), 7.19–7.13 (3H, m, *H*-4), 6.54–6.46 (1H, m, *H*-4'$_X$), 5.71–5.60 (1H, m, *H*-9'), 4.12–4.02 (1H, m, *H*-6'$_X$), 3.93 (1H, dd, J_{AB} 18.3, J_{AX} 5.5, H_aH_b-3'), 3.83 (1H, dd, J_{AB} 18.3, J_{BX} 5.1, H_aH_b-3'), 3.66 (3H, s, H_3-13'), 2.73 (1H, dd, J_{AB} 13.2, J_{AX} 6.4, H_aH_b-7'), 2.51 (1H, dd, J_{AB} 13.2, J_{BX} 5.6, H_aH_b-7'), 1.84 (3H, s, H_3-11'); *m/z* (ESI$^-$) 475 ([M−H]$^-$, 39%), 511 ([M+Cl]$^-$, 100%), 521 ([M+HCO_2H−H]$^-$, 78%); Analytical HPLC (method 2, 220/254 nm) ret. time = 7.3 min, purity = 100%; Chiral HPLC (90:10 hexane:IPA, 1.9 mL min^{-1}, 30 min) ret. time = 14.7 min (L-**45**), 77.0% *e.e.* (other enantiomer observed, ret. time = 12.6 min, area = 11.5% (D-**45** = **52**)). These data are in agreement with the available literature values.[44,45]

Methyl *N*-acetyl-*S*-(1-ethyl-2,5-dioxopyrrolidin-3-yl)-D-cysteinylglycinate, OXFKEF16 (48)

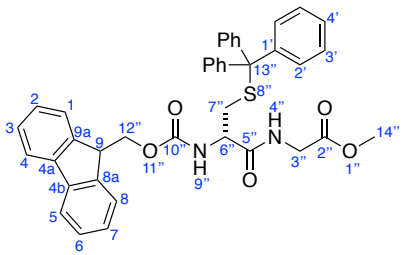


Triethylsilane (0.24 mL, 1.5 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-D-cysteinylglycinate, **52**, (0.24 g, 0.50 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (18 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 30 min. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**52**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (25 mL) and extracted with H_2O (3 × 10 mL). The combined aq. components were lyophilised to give *N*-acetyl-D-cysteinylglycine methyl ester (**53**) as a colourless, hygroscopic foam (0.11 g), which was used without further purification.

NEM, **19**, (35 mg, 0.28 mmol, 1.0 eq.) was added to a solution of *N*-acetyl-D-cysteinylglycine methyl ester, **53**, (98 mg, 0.42 mmol, 1.5 eq.) in H_2O (2.4 mL). The resulting reaction mixture was stirred at rt for 19 h. After this time, TLC analysis (H_2O:MeCN 3:2, RP) indicated that some NEM (**19**) remained. The reaction mixture was lyophilised to obtain the crude product, which was purified by RP silica gel column chromatography (H_2O:MeCN 100:0 → 95:5 → 90:10 → 85:15) to give the title compound (**48**) as a colourless, hygroscopic solid in a 1:1.1 ratio (^{1}H NMR) of 2 diastereoisomers A and B (82 mg, 75% (over 2 steps)): R_f 0.26 (H_2O:MeCN 3:2, RP); mp 137–144 °C (from H_2O, wide range because compound is

hygroscopic); $[\alpha]_D^{25}$ = +17.8 (*c* 1.0, H_2O); $\bar{\nu}_{max}$ (neat)/cm^{-1} 3279 (N–H, m), 3078, 2975, 2939 (C–H, w), 1755 (C=O (ester), m), 1692 (C=O (amide), s), 1648 (C=O (amide, succinimide), s), 1546 (N–H (amide), m), 1403 (w), 1352 (m), 1216 (m), 1119 (m), 980 (w), 790 (w), 731 (w), 685 (m); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer A) δ 4.69 (1H, dd, J_{BX} 7.5, J_{AX} 5.9, *H*-6'$_X$), 4.08 (1H, dd, J_{AX} 8.9, J_{BX} 4.1, *H*-3$_X$), 4.08–4.03 (2H, m, H_aH_b-3'), 3.77 (3H, s, H_3-14'), 3.56 (2H, q, *J* 7.3, H_2-12'), 3.38–3.27 (1H, m, H_aH_b-4), 3.24 (1H, dd, J_{AB} 14.3, J_{AX} 5.9, H_aH_b-7'), 3.20 (1H, dd, J_{AB} 14.3, J_{BX} 7.5, H_aH_b-7'), 2.69 (1H, dd, J_{AB} 6.9, J_{BX} 4.1, H_aH_b-4), 2.10 or 2.09 (3H, s, H_3-11'), 1.14 (3H, t, *J* 7.3, H_3-13'); ^{1}H NMR (500 MHz; D_2O) (diastereoisomer B) δ 4.66 (1H, dd, J_{BX} 8.6, J_{AX} 5.2, *H*-6'$_X$), 4.08–4.03 (3H, m, *H*-3$_X$, H_aH_b-3'), 3.77 (3H, s, H_3-14'), 3.56 (2H, q, *J* 7.3, H_2-12'), 3.38–3.27 (2H, m, H_aH_b-4, H_aH_b-7'), 3.05 (1H, dd, J_{AB} 14.0, J_{BX} 8.6, H_aH_b-7'), 2.72 (1H, dd, J_{AB} 6.8, J_{BX} 4.1, H_aH_b-4), 2.10 or 2.09 (3H, s, H_3-11'), 1.14 (3H, t, *J* 7.3, H_3-13'); ^{13}C NMR (126 MHz; D_2O) (diastereoisomer A) δ 179.5 (*C*-2), 178.2 (*C*-5), 174.4 or 174.3 (*C*-10'), 172.5 (*C*-5'), 171.7 (*C*-2'), 53.1 (*C*-6'), 52.8 (*C*-14'), 41.2 (*C*-3'), 40.7 (*C*-3), 35.8 (*C*-4), 34.3 (*C*-12'), 32.8 (*C*-7'), 21.71 or 21.70 (*C*-11'), 11.7 (*C*-13'); ^{13}C NMR (126 MHz; D_2O) (diastereoisomer B) δ 179.4 (*C*-2), 178.2 (*C*-5), 174.4 or 174.3 (*C*-10'), 172.5 (*C*-5'), 171.7 (*C*-2'), 52.8 (*C*-14'), 52.7 (*C*-6'), 41.2 (*C*-3'), 40.0 (*C*-3), 36.0 (*C*-4), 34.3 (*C*-12'), 32.5 (*C*-7'), 21.71 or 21.70 (*C*-11'), 11.7 (*C*-13'); HRMS *m/z* (ESI^+) [Found $(M+Na)^+$ 382.10436, $C_{14}H_{21}N_3O_6SNa$ requires M^+ 382.10433]; *m/z* (ESI^+) 382 ($[M+Na]^+$, 100%); Analytical HPLC (method 1, 220/254 nm) ret. time = 6.4 min, area = 50.5% (diastereoisomer B), ret. time = 6.5 min, area = 48.1% (diastereoisomer A); Chiral HPLC (90:10 hexane:IPA, 1.8 mL min^{-1}, 60 min) ret. time = 43.3 min (D-**48**), >99.0% *e.e.* (other enantiomer not observed, ret. time = 49.5 min (L-**48** = **32**)).

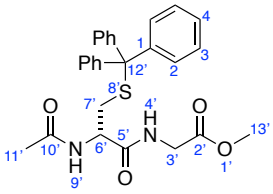
Methyl-*N*-{[(9*H*-fluoren-9-yl)methoxy]carbonyl}-*S*-(triphenylmethyl)-D-cysteinylglycinate (50)



Fmoc-D-Cys(Trt)-OH, **49**, (860 mg, 1.47 mmol, 1.05 eq.), glycine methyl ester hydrochloride (176 mg, 1.40 mmol, 1.00 eq.), and HBTU (557 mg, 1.47 mmol, 1.05 eq.) were dissolved in DMF (3.20 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 5 min. DIPEA (492 µL, 2.80 mmol, 2.00 eq.) was then added, and the reaction mixture was stirred at rt under Ar for 24 h. After this time, TLC analysis (EtOAc:EtOH 3:2) indicated complete consumption of glycine methyl ester hydrochloride. The reaction mixture was extracted with EtOAc (25 mL) and washed with 0.5 M aq. LiCl solution (3 × 25 mL). The organic component was separated, and the combined aq. components were extracted with EtOAc (25 mL). The combined organic components were dried over Na_2SO_4, filtered, and concentrated *in vacuo* to give the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (petroleum ether:EtOAc 1:1) to give the title compound (**50**) as a colourless solid (847 mg, 92%): R_f 0.80 (EtOAc:MeOH 9:1); $[\alpha]_D^{25}$ = −4.9 (*c* 1.0, $CHCl_3$); mp 76–82 °C (from $CHCl_3$/CH_2Cl_2); $\bar{\nu}_{max}$ (thin film)/cm^{-1} 3311 (N–H, m), 3060 (C–H (aromatic), w), 2951, 2919, 2850 (C–H (aliphatic), w), 1747 (C=O (ester), m), 1711 (C=O (carbamate), s), 1669 (C=O (amide), s), 1528 (m), 1371 (m), 1156 (s), 740 (s), 701 (s); ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.74 (2H, dd, *J* 7.3, 7.3, *H*-4, *H*-5), 7.57 (2H, d, *J* 7.2, *H*-1, *H*-8), 7.44–7.39 (6H, m, *H*-2'), 7.39–7.35 (2H, m, *H*-3, *H*-6), 7.31–7.24 (8H, m, *H*-2,

H-7, *H*-3'), 7.24–7.18 (3H, m, *H*-4'), 6.35–6.24 (1H, m, *H*-4''$_X$), 4.94 (1H, d, *J* 7.2, *H*-9''), 4.42–4.37 (2H, m, H_aH_b-12''), 4.20 (1H, dd, *J* 6.8, 6.7, *H*-9$_X$), 3.98 (1H, dd, J_{AB} 18.4, J_{AX} 5.4, H_aH_b-3''), 3.91 (1H, dd, J_{AB} 18.4, J_{BX} 5.1, H_aH_b-3''), 3.84–3.75 (1H, m, *H*-6''$_X$), 3.72 (3H, s, H_3-14''), 2.74–2.60 (2H, m, H_aH_b-7''); ^{13}C NMR (126 MHz; $CDCl_3$) δ 170.4 (*C*-5''), 169.8 (*C*-2''), 156.1 (*C*-10''), 144.4 (*C*-1'), 143.84 (*C*-8a or *C*-9a), 143.79 (*C*-8a or *C*-9a), 141.5 (*C*-4a, *C*-4b), 129.7 (*C*-2'), 128.3 (*C*-3'), 127.91 (*C*-3 or *C*-6), 127.88 (*C*-3 or *C*-6), 127.23 (*C*-2 or *C*-7), 127.21 (*C*-2 or *C*-7), 127.1 (*C*-4'), 125.18 (*C*-1 or *C*-8), 125.17 (*C*-1 or *C*-8), 120.1 (*C*-4, *C*-5), 67.5 (*C*-13''), 67.2 (*C*-12''), 54.0 (*C*-6''), 52.5 (*C*-14''), 47.3 (*C*-9), 41.4 (*C*-3''), 33.8 (*C*-7''); HRMS *m/z* (ESI$^+$) [Found (M+Na)$^+$ 679.22335, $C_{40}H_{36}N_2O_5SNa$ requires M$^+$ 679.22371]; *m/z* (ESI$^+$) 679 ([M+Na]$^+$, 100%); Analytical HPLC (method 2, extended 1 min, 220/254 nm) ret. time = 12.7 min, purity = 100%; Chiral HPLC (75:25 hexane:IPA, 1.0 mL min^{-1}, 30 min) ret. time = 11.5 min (D-**50**), >99.0% *e.e.* (other enantiomer not observed, ret. time = 16.8 min (L-**50**)).

Methyl *N*-acetyl-*S*-(triphenylmethyl)-D-cysteinylglycinate (52)



Method 1

Piperidine (0.50 mL) was added to a solution of methyl-*N*-{[(9*H*-fluoren-9-yl)methoxy]carbonyl}-*S*-(triphenylmethyl)-D-cysteinylglycinate, **50**, (0.20 g, 0.30 mmol, 1.0 eq.) in anhydrous DMF (2.0 mL) at rt under Ar. The resulting yellow reaction mixture was stirred at rt under Ar for 5 min. After this time, TLC analysis (toluene:EtOAc 6:4) indicated complete consumption of the starting material (**50**). The

reaction mixture was diluted with toluene (10 mL) and concentrated *in vacuo* to give the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (toluene:EtOAc:Et_3N 100:0:0 → 90:10:0 → 80:20:0 → 70:30:0 → 60:40:0 → 0:100:1) to give methyl *S*-(triphenylmethyl)-D-cysteinylglycinate (**51**) as a colourless solid (99 mg), which was used without further purification.

A subsequent reaction on the starting material, **50**, (0.57 g, 0.87 mmol, 1.0 eq.) under the same conditions as above gave methyl *S*-(triphenylmethyl)-D-cysteinylglycinate (**51**) as a colourless solid (0.19 g), which was used without further purification. This batch was combined with the previous one for the next reaction (acetylation).

Anhydrous Et_3N (1.92 mL, 13.8 mmol, 25.0 eq.) was added dropwise to a solution of methyl *S*-(triphenylmethyl)-D-cysteinylglycinate, **51**, (240 mg, 0.552 mmol, 1.00 eq.) in anhydrous CH_2Cl_2 (18.2 mL) at 0 °C under Ar. The reaction mixture was stirred at 0 °C under Ar for 20 min. Acetyl chloride (78.5 µL, 1.10 mmol, 2.00 eq.) was added dropwise, the reaction mixture was warmed to rt and stirred under Ar for 2.5 h. After this time, TLC analysis (EtOAc) indicated that some starting material (**51**) remained. Further acetyl chloride (39.2 µL, 0.552 mmol, 1.00 eq.) was added at rt and the reaction was stirred under Ar for 30 min. After this time, TLC analysis (EtOAc) indicated that some starting material (**51**) remained. The reaction was diluted with CH_2Cl_2 (100 mL), washed with 1 M aq. HCl (2 × 60 mL), saturated aq. Na_2CO_3 solution (2 × 70 mL), and H_2O (100 mL). The organic component was separated, dried over $MgSO_4$, filtered, and concentrated *in vacuo* to obtain the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (petroleum ether:EtOAc:Et_3N 70:30:1 → 60:40:1 → 50:50:1 → 40:60:1 → 20:80:1) to

give the title compound (**52**) as a colourless solid (62.0 mg, 14% (over 2 steps)): R_f 0.46 (EtOAc); $[\alpha]_D^{25}$ = −3.7 (*c* 1.0, $CHCl_3$); mp 61–67 °C (from $CHCl_3$/CH_2Cl_2); $\bar{\nu}_{max}$ (neat)/cm^{-1} 3382 (N–H, w), 3056 (C–H (aromatic), w), 2950, 2849 (C–H (aliphatic), w), 1749 (C=O (ester), m), 1650 (C=O (amide), m), 1525 (m), 1489 (m), 1442 (m), 1371 (m), 1209 (m), 1181 (m), 1034 (m), 743 (s), 699 (s), 675 (m), 618 (m); ^{1}H NMR (500 MHz; $CDCl_3$) δ 7.47–7.41 (6H, m, *H*-2), 7.33–7.27 (6H, m, *H*-3), 7.25–7.20 (3H, m, *H*-4), 6.59 (1H, dd, J_{AX} 5.4, J_{BX} 5.1, *H*-4'$_X$), 5.77 (1H, d, *J* 7.6, *H*-9'), 4.18–4.11 (1H, m, *H*-6'$_X$), 3.98 (1H, dd, J_{AB} 18.3, J_{AX} 5.4, H_aH_b-3'), 3.89 (1H, dd, J_{AB} 18.3, J_{BX} 5.1, H_aH_b-3'), 3.72 (3H, s, H_3-13'), 2.77 (1H, dd, J_{AB} 13.2, J_{AX} 7.0, H_aH_b-7'), 2.58 (1H, dd, J_{AB} 13.2, J_{BX} 5.7, H_aH_b-7'), 1.90 (3H, s, H_3-11'); ^{13}C NMR (126 MHz; $CDCl_3$) δ 170.4 (*C*-5', *C*-10'), 169.9 (*C*-2'), 144.5 (*C*-1), 129.7 (*C*-2), 128.2 (*C*-3), 127.1 (*C*-4), 67.4 (*C*-12'), 52.5 (*C*-13'), 52.1 (*C*-6'), 41.3 (*C*-3'), 33.1 (*C*-7'), 23.2 (*C*-11'); HRMS *m/z* (ESI^+) [Found $(M+Na)^+$ 499.16618, $C_{27}H_{28}N_2O_4SNa$ requires M^+ 499.16620]; *m/z* (ESI^-) 475 ($[M-H]^-$, 82%), 511 ($[M+Cl]^-$, 40%), 521 ($[M+HCO_2H-H]^-$, 100%), 589 ($[M+TFA-H]^-$, 10%); Analytical HPLC (method 2, 220/254 nm) ret. time = 10.0 min, purity = 94.5%; Chiral HPLC (90:10 hexane:IPA, 1.9 mL min^{-1}, 30 min) ret. time = 12.5 min (D-**52**), >99.0% *e.e.* (other enantiomer not observed, ret. time = 14.7 min (L-**52**)).

Methods 2A and 2B

Part 1 (Fmoc Deprotection)

Anhydrous piperidine (1.0 mL) was added to a solution of methyl-*N*-{[(9*H*-fluoren-9-yl)methoxy]carbonyl}-*S*-(triphenylmethyl)-D-cysteinylglycinate, **50**, (0.40 g, 0.61 mmol, 1.0 eq.) in anhydrous DMF (4.0 mL) at rt under Ar. The resulting yellow reaction mixture was stirred at rt under Ar for 15 min. After this time, TLC analysis (petroleum

ether:EtOAc 6:4) indicated complete consumption of the starting material (**50**). The reaction mixture was diluted with EtOAc (20 mL), and the solvent was evaporated under a constant stream of N_2 (g) to give methyl *S*-(triphenylmethyl)-D-cysteinylglycinate (**51**) as a yellow solid (0.38 g), which was used without further purification. Half of this material (0.19 g) was acetylated using Method 2A and the other half (0.19 g) using Method 2B. Both acetylation methods are described below.

Method 2A (Acetylation of 51)

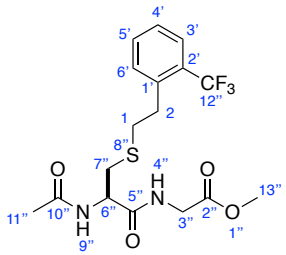
DMAP (7.4 mg, 0.061 mmol, 0.20 eq.) was added to a solution of methyl *S*-(triphenylmethyl)-D-cysteinylglycinate, **51**, (0.19 g, 0.30 mmol, 1.0 eq.) in anhydrous pyridine (0.72 mL) at rt under Ar. Acetic anhydride (58 µL, 0.61 mmol, 2.0 eq.) was added, and the reaction mixture was stirred at rt under Ar for 11 h. After this time, TLC analysis (petroleum ether:EtOAc 1:4) indicated complete consumption of the starting material (**51**). The solvent was evaporated under a constant stream of N_2 (g) to obtain the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (toluene:EtOAc:Et_3N 50:50:1). A second purification (petroleum ether:EtOAc:Et_3N 60:40:1 → 50:50:1 → 40:60:1 → 30:70:1) gave the title compound (**52**) as a colourless solid (62 mg, 41%). Data for compound **52** are shown in Method 1.

Method 2B (Acetylation of 51)

Anhydrous pyridine (0.12 mL, 1.5 mmol, 5.0 eq.) was added to a solution of methyl *S*-(triphenylmethyl)-D-cysteinylglycinate, **51**, (0.19 g, 0.30 mmol, 1.0 eq.) in anhydrous THF (0.60 mL) at rt under Ar. Acetic anhydride (58 µL, 0.61 mmol, 2.0 eq.) was added, and the reaction mixture was stirred at rt under Ar for 11 h. After this time, TLC

analysis (petroleum ether:EtOAc 1:4) indicated complete consumption of the starting material (**51**). The solvent was evaporated under a constant stream of N_2 (g) to obtain the crude product, which was adsorbed onto Celite® and purified by NP silica gel flash column chromatography (toluene:EtOAc:Et_3N 50:50:1). A second purification (petroleum ether:EtOAc:Et_3N 60:40:1 → 50:50:1 → 40:60:1 → 30:70:1) gave the title compound (**52**) as a colourless solid (70 mg, 48%). Data for compound **52** are shown in Method 1 (above).

Methyl *N*-acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycinate (80)



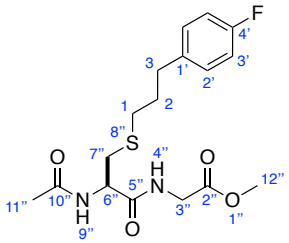
Triethylsilane (2.43 mL, 15.2 mmol, 3.00 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate, **45**, (2.42 g, 5.08 mmol, 1.00 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (180 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 1 h. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (60 mL) and extracted with H_2O (5 × 30 mL). The combined aq. components were lyophilised to give *N*-acetyl-L-cysteinylglycine methyl ester (**46**) as a colourless, hygroscopic foam (1.10 g), which was used without further purification.

To a suspension of methyl *N*-acetyl-L-cysteinylglycinate, **46**, (1.09 g, 4.65 mmol, 4.00 eq.) and DPAP (60.0 mg, 0.233 mmol, 0.200 eq.) in THF:H_2O 1:2 (6.60 mL) was

added 2-(trifluoromethyl)styrene (170 µL, 1.16 mmol, 1.00 eq.), followed by TCEP·HCl (667 mg, 2.33 mmol, 2.00 eq.). The resulting reaction mixture was stirred in the presence of UV light (λ = 365 nm) at rt for 5 h. After this time, TLC analysis (EtOAc) indicated complete consumption of 2-(trifluoromethyl)styrene. The reaction mixture was concentrated *in vacuo* and remaining solvent evaporated under a constant stream of N_2 (g). The reaction mixture was filtered, and the solid was washed with H_2O (40 mL). The crude solid was dried *in vacuo* and purified by RP silica gel flash column chromatography (H_2O:MeCN 80:20 → 70:30 → 65:35 → 60:40 → 50:50) to give the title compound (**80**) as a colourless, hygroscopic solid (133 mg, 26% (over 2 steps)): R_f 0.46 (EtOAc); $[\alpha]_D^{25}$ = −1.9 (*c* 2.0, $CHCl_3$); mp 136–140 °C (from hexane/CH_2Cl_2); $\bar{\nu}_{max}$ (thin film)/cm^{-1} 3283 (N–H, m), 3068 (C–H (aromatic), w), 2928, 2849 (C–H (aliphatic), w), 1753 (C=O (ester), m), 1651 (C=O (amide), s), 1541 (m), 1314 (s), 1117 (s), 1038 (m), 770 (m), 652 (w), 617 (w); ^{1}H NMR (500 MHz; $CDCl_3$) δ 7.62 (1H, d, *J* 7.7, *H*-3'), 7.49 (1H, dd, *J* 7.6, 7.5, *H*-5'), 7.41 (1H, d, *J* 7.6, *H*-6'), 7.32 (1H, dd, *J* 7.7, 7.5, *H*-4'), 6.99 (1H, dd, J_{AX} 5.1, J_{BX} 5.1, *H*-4''$_X$), 6.48 (1H, d, *J* 7.3, *H*-9''), 4.60 (1H, ddd, *J* 7.3, J_{BX} 5.3, J_{AX} 5.3, *H*-6''$_X$), 4.07 (1H, dd, J_{AB} 18.2, J_{AX} 5.1, H_aH_b-3''), 4.02 (1H, dd, J_{AB} 18.2, J_{BX} 5.1, H_aH_b-3''), 3.75 (3H, s, H_3-13''), 3.11–3.02 (3H, m, H_aH_b-2, H_aH_b-7''), 2.90–2.78 (3H, m, H_aH_b-1, H_aH_b-7''), 2.05 (3H, s, H_3-11''); ^{19}F NMR (470 MHz; $CDCl_3$) δ −59.5 (F_3-12''); ^{13}C NMR (126 MHz; $CDCl_3$) δ 170.7 (*C*-5''), 170.4 (*C*-10''), 169.9 (*C*-2''), 138.7 (*C*-1'), 132.1 (*C*-5'), 131.6 (*C*-6'), 128.6 (q, *J* 29.8, *C*-2'), 126.8 (*C*-4'), 126.2 (q, *J* 5.7, *C*-3'), 124.7 (q, *J* 273.6, *C*-12''), 52.6 (*C*-13''), 52.5 (*C*-6''), 41.5 (*C*-3''), 33.9 (*C*-7''), 33.5 (*C*-1), 33.2 (*C*-2), 23.3 (*C*-11''); HRMS *m/z* (ESI$^+$) [Found $(M+Na)^+$ 429.10619, $C_{17}H_{21}N_2O_4F_3SNa$ requires M^+ 429.10663]; *m/z* (ESI$^+$) 429 ($[M+Na]^+$, 100%);

Analytical HPLC (method 2, 220/254 nm) ret. time = 8.6 min, purity = 98.7%; Chiral HPLC (90:10 hexane:IPA, 1.5 mL min^{-1}, 20 min) ret. time = 8.9 min (L-**80**), 79.8% *e.e.* (other enantiomer observed, ret. time = 7.1 min, area = 10.1% (D-**80**)).

Methyl *N*-acetyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycinate (81)



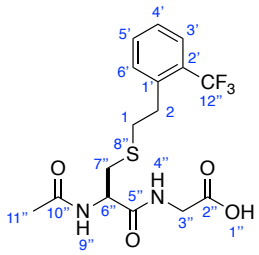
Triethylsilane (2.59 mL, 16.2 mmol, 3.00 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L-cysteinylglycinate, **45**, (2.57 g, 5.39 mmol, 1.00 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (190 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 50 min. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (60 mL) and extracted with H_2O (4 × 30 mL). The combined aq. components were lyophilised to give *N*-acetyl-L-cysteinylglycine methyl ester (**46**) as a colourless, hygroscopic foam (1.45 g), which was used without further purification.

TCEP·HCl (859 mg, 2.99 mmol, 2.00 eq.) was added to a suspension of methyl *N*-acetyl-L-cysteinylglycinate, **46**, (1.40 g, 5.39 mmol, 3.60 eq.) and DPAP (76.6 mg, 0.299 mmol, 0.200 eq.) in THF:H_2O 1:2 (9.00 mL), followed by 1-allyl-4-fluorobenzene (202 µL, 1.50 mmol, 1.00 eq.). The resulting reaction mixture was stirred in the presence of UV light (λ = 365 nm) at rt for 5 h. After this time, TLC analysis (EtOAc) indicated that some 1-allyl-4-fluorobenzene remained. The reaction mixture was

diluted with MeOH (30 mL), and solvent evaporated under a constant stream of N_2 (g). The resulting residue was extracted with EtOAc (80 mL) and washed with 1 M aq. NaOH (3 × 25 mL), H_2O (3 × 20 mL), and brine (25 mL). The organic component was separated, and the combined aq. components were extracted with EtOAc (3 × 60 mL). The combined organic components were dried over Na_2SO_4, filtered, and concentrated *in vacuo* to obtain the crude product, which was purified by RP silica gel flash column chromatography (H_2O:MeCN 70:30 → 65:35) to give the title compound (**81**) as a colourless, hygroscopic solid (145 mg, 26% (over 2 steps)): R_f 0.40 (H_2O:MeOH 1:4, RP); $[\alpha]_D^{25}$ = −2.5 (*c* 2.0, $CHCl_3$); mp 99–104 °C (from MeCN, wide range because compound is hygroscopic); $\bar{\nu}_{max}$ (thin film)/cm^{-1} 3281 (N–H, m), 3067 (C–H (aromatic), w), 2939, 2852 (C–H (aliphatic), w), 1750 (C=O (ester), m), 1651 (C=O (amide), s), 1541 (m), 1509 (s), 1438 (w), 1415 (w), 1373 (w), 1260 (s), 1219 (s), 1093 (s), 1017 (s), 703 (w), 625 (w); ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.17–7.09 (2H, m, *H*-2'), 7.01–6.92 (3H, m, *H*-3', *H*-4''), 6.43 (1H, d, *J* 7.1, *H*-9''), 4.55 (1H, ddd, J_{BX} 7.6, *J* 7.1, J_{AX} 5.1, *H*-6''$_X$), 4.06 (1H, dd, J_{AB} 18.1, J_{AX} 5.2, H_aH_b-3''), 4.01 (1H, dd, J_{AB} 18.1, J_{BX} 5.1, H_aH_b-3''), 3.76 (3H, s, H_3-12''), 2.99 (1H, dd, J_{AB} 13.9, J_{AX} 5.1, H_aH_b-7''), 2.78 (1H, dd, J_{AB} 13.9, J_{BX} 7.6, H_aH_b-7''), 2.72–2.66 (2H, m, H_aH_b-3), 2.66–2.54 (2H, m, H_aH_b-1), 2.04 (3H, s, H_3-11''), 1.97–1.85 (2H, m, H_aH_b-2); ^{19}F NMR (470 MHz; $CDCl_3$) δ −117.5 (*F*-4'); ^{13}C NMR (126 MHz; $CDCl_3$) δ 170.7 (*C*-5''), 170.4 (*C*-10''), 170.0 (*C*-2''), 161.5 (d, *J* 243.6, *C*-4'), 137.0 (d, *J* 3.0, *C*-1'), 129.9 (d, *J* 7.8, *C*-2'), 115.3 (d, *J* 21.0, *C*-3'), 52.6 (*C*-12''), 52.5 (*C*-6''), 41.5 (*C*-3''), 33.94 (*C*-3), 33.89 (*C*-7''), 31.8 (*C*-1), 31.2 (*C*-2), 23.3 (*C*-11''); HRMS *m/z* (ESI^+) [Found $(M+Na)^+$ 393.12561, $C_{17}H_{23}N_2O_4FSNa$ requires M^+ 393.12548]; *m/z* (ESI^+) 393 ($[M+Na]^+$, 100%), 409 ($[M+K]^+$, 6%); Analytical HPLC (method 1, 220/254 nm) ret. time = 12.2 min,

purity = 99.8%; Chiral HPLC (90:10 hexane:IPA, 1.5 mL min^{-1}, 20 min) ret. time = 15.2 min (L-**81**), 65.2% *e.e.* (other enantiomer observed, ret. time = 12.7 min, area = 17.4% (D-**81**)).

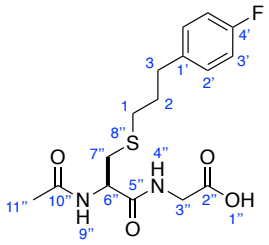
_N_-Acetyl-_S_-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycine (82)



Aq. LiOH (0.37 M, 92 µL, 0.62 eq.) was added to a solution of methyl *N*-acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycinate, **80**, (22 mg, 0.055 mmol, 1.0 eq.) in THF:MeOH 3:2 (0.46 mL) at 0 °C. The resulting reaction mixture was stirred at 0 °C for 50 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated that some starting material (**80**) remained. Further aq. LiOH (0.37 M, 92 µL, 0.62 eq.) was added, and the reaction was stirred at 0 °C for 40 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated that some starting material (**80**) remained. Further aq. LiOH (0.37 M, 92 µL, 0.62 eq.) was added, and the reaction was stirred at 0 °C for 40 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated complete consumption of the starting material (**80**). The reaction was acidified to pH 2 using 1 M aq. HCl. This crude mixture was purified directly by RP silica gel flash column chromatography (H_2O:MeCN 100:0 → 75:25) to give the title compound (**82**) as a colourless, hygroscopic solid (17 mg, 77%): R_f 0.76 (H_2O:MeCN 1:1, RP); $[\alpha]_D^{25}$ = −10.2 (*c* 1.0, MeOH); mp 137–143 °C (from H_2O, wide range because compound is hygroscopic); $\bar{\nu}_{max}$ (thin film)/cm^{-1} 3291 (N–H, m), 3080 (C–H (aromatic),

w), 2928, 2853 (C–H (aliphatic), w), 1639 (C=O (acid), s), 1586 (C=O (amide), s), 1452 (m), 1415 (m), 1314 (s), 1175 (m), 1116 (s), 1039 (m), 768 (m), 652 (w); ^{1}H NMR (400 MHz, D_4-MeOH) δ 7.64 (1H, d, *J* 7.8, *H*-3'), 7.56 (1H, dd, *J* 7.5, 7.3, *H*-5'), 7.51 (1H, d, *J* 7.5, *H*-6'), 7.37 (1H, dd, *J* 7.8, 7.3, *H*-4'), 4.61 (1H, dd, J_{BX} 8.6, J_{AX} 5.4, *H*-6''$_X$), 4.00–3.85 (2H, m, H_aH_b-3''), 3.07 (1H, dd, J_{AB} 13.9, J_{AX} 5.4, H_aH_b-7''), 3.08–3.00 (2H, m, H_aH_b-2), 2.82 (1H, dd, J_{AB} 13.9, J_{BX} 8.6, H_aH_b-7''), 2.82–2.77 (2H, m, H_aH_b-1), 2.03 (3H, s, H_3-11''); ^{19}F NMR (376 MHz; D_4-MeOH) δ −60.8 (F_3-12''); ^{13}C NMR (126 MHz; D_4-MeOH) δ 177.2 (*C*-2''), 173.9 (*C*-10''), 172.7 (*C*-5''), 140.3 (*C*-1'), 133.4 (*C*-5'), 132.8 (*C*-6'), 129.3 (q, *J* 29.6, *C*-12''), 127.8 (*C*-4'), 126.9 (q, *J* 5.9, *C*-3'), 126.1 (q, *J* 272.9, *C*-2'), 55.0 (*C*-6''), 44.5 (*C*-3''), 34.5 (*C*-7''), 34.3 (*C*-1), 34.1 (*C*-2), 22.5 (*C*-11''); HRMS *m/z* (ESI$^+$) [Found (M+H)$^+$ 393.10903, $C_{16}H_{20}N_2O_4F_3S$ requires M$^+$ 393.10904]; *m/z* (ESI$^-$) 391 ([M−H]$^-$, 100%), 783 ([2M−H]$^-$, 44%); Analytical HPLC (method 2, 220/254 nm) ret. time = 7.8 min, purity = 99.3%; Chiral HPLC (50:50:0.1 hexane:EtOH:TFA, 1.0 mL min^{-1}, 30 min) ret. time = 6.6 min (L-**82**), 75.6% *e.e.* (other enantiomer observed, ret. time = 8.7 min, area = 12.2% (D-**82**)).

_N_-Acetyl-_S_-[3-(4-fluorophenyl)propyl]-L-cysteinylglycine (83)

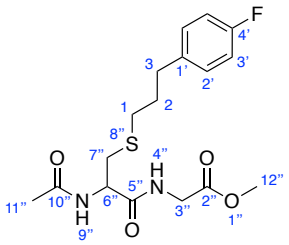


Aq. LiOH·H_2O (0.38 M, 0.11 mL, 0.62 eq.) was added to a solution of methyl *N*-acetyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycinate, **81**, (25 mg, 0.067 mmol, 1.0 eq.) in THF:MeOH 3:2 (0.57 mL) at 0 °C. The resulting reaction mixture was stirred

at 0 °C for 35 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated that some starting material (**81**) remained. Further aq. LiOH·H_2O (0.38 M, 0.11 mL, 0.62 eq.) was added, and the reaction was stirred at 0 °C for 40 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated complete consumption of the starting material (**81**). The reaction was acidified to pH 3 using 1 M aq. HCl. This crude mixture was purified directly by RP silica gel flash column chromatography (H_2O:MeCN 100:0 → 75:25) to give the title compound (**83**) as a colourless, hygroscopic solid (19 mg, 79%); R_f 0.47 (H_2O:MeCN 1:1, RP); $[\alpha]_D^{25}$ = −12.6 (*c* 1.0, MeOH); mp 119–128 °C (from MeOH, wide range because compound is hygroscopic); $\bar{\nu}_{max}$ (thin film)/cm^{-1} 3277 (N–H, m), 3078 (C–H (aromatic), w), 2931, 2857 (C–H (aliphatic), w), 1645 (C=O (acid), m), 1598 (C=O (amide), s), 1509 (s), 1415 (m), 1310 (m), 1221 (m), 1158 (w), 1037 (w), 829 (m), 702 (m); ^{1}H NMR (500 MHz, D_4-MeOH) δ 7.19 (2H, dd, *J* 8.6, 5.5, *H*-2'), 6.97 (2H, dd, J_{HF} 8.9, *J* 8.6, *H*-3'), 4.47 (1H, dd, J_{BX} 8.8, J_{AX} 5.3, *H*-6''$_X$), 3.84 (1H, d, J_{AB} 17.2, H_aH_b-3''), 3.68 (1H, d, J_{AB} 17.2, H_aH_b-3''), 3.00 (1H, dd, J_{AB} 13.8, J_{AX} 5.3, H_aH_b-7''), 2.77 (1H, dd, J_{AB} 13.8, J_{BX} 8.8, H_aH_b-7''), 2.71–2.65 (2H, m, H_aH_b-3), 2.60–2.50 (2H, m, H_aH_b-1), 2.03 (3H, s, H_3-11''), 1.91–1.82 (2H, m, H_aH_b-2); ^{19}F NMR (470 MHz; D_4-MeOH) δ −117.5 (*F*-4'); ^{13}C NMR (126 MHz; D_4-MeOH) δ 177.2 (*C*-2''), 173.9 (*C*-10''), 172.7 (*C*-5''), 162.7 (d, *J* 241.9, *C*-4'), 138.8 (d, *J* 3.0, *C*-1'), 131.1 (d, *J* 7.7, *C*-2'), 115.9 (d, *J* 21.2, *C*-3'), 55.1 (*C*-6''), 44.5 (*C*-3''), 34.7 (*C*-3), 34.4 (*C*-7''), 32.40 (*C*-1 or *C*-2), 32.38 (*C*-1 or *C*-2), 22.6 (*C*-11''); HRMS *m/z* (ESI$^-$) [Found (M−H)$^-$ 355.11329, $C_{16}H_{20}N_2O_4FS$ requires M$^-$ 355.11333]; *m/z* (ESI$^-$) 355 ([M−H]$^-$, 100%), 711 ([2M−H]$^-$, 32%); Analytical HPLC (method 2, 220/254 nm) ret. time = 7.6 min, purity = 97.4%; Chiral HPLC (50:50:0.1 hexane:EtOH:TFA, 1.0 mL min^{-1}, 30 min) ret. time = 9.1 min

(L-**83**), 72.4% *e.e.* (other enantiomer observed, ret. time = 22.5 min, area = 13.8% (D-**83**)).

Methyl *N*-acetyl-*S*-[3-(4-fluorophenyl)propyl]-L/D-cysteinylglycinate (84)

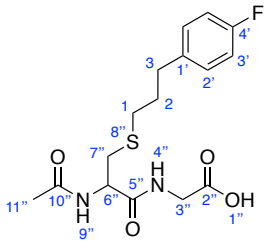


Triethylsilane (0.20 mL, 1.2 mmol, 3.0 eq.) was added to a solution of methyl *N*-acetyl-*S*-triphenylmethyl-L/D-cysteinylglycinate, **45**/**52**, (1:1 mixture by mass, total compound: 0.20 g, 0.40 mmol, 1.0 eq.) in 20% *v*/*v* TFA in anhydrous CH_2Cl_2 (15 mL) at rt under Ar. The resulting reaction mixture was stirred at rt under Ar for 35 min. After this time, TLC analysis (EtOAc) indicated complete consumption of the starting material (**45**/**52**). The reaction mixture was concentrated *in vacuo*. The resulting residue was dissolved in EtOAc (25 mL) and extracted with H_2O (3 × 20 mL). The combined aq. components were lyophilised to give *N*-acetyl-L/D-cysteinylglycine methyl ester (**46**/**53**) as a colourless, hygroscopic foam (99 mg), which was used without further purification.

1-Allyl-4-fluorobenzene (14 µL, 0.11 mmol, 1.0 eq.) was added to a suspension of methyl *N*-acetyl-L/D-cysteinylglycinate (**46**/**53**) (99 mg, 0.42 mmol, 4.0 eq.), DPAP (5.0 mg, 0.021 mmol, 0.20 eq.), and TCEP·HCl (61 mg, 0.21 mmol, 2.0 eq.) in THF:H_2O 1:2 (0.63 mL). The resulting reaction mixture was stirred in the presence of UV light (λ = 365 nm) at rt for 5 h. The solvent was evaporated under a constant stream of N_2 (g) to obtain the crude product, which was purified by RP silica gel flash column

chromatography (H_2O:MeCN 75:25 → 70:30) to give the title compound (**84**) as a colourless, hygroscopic solid (22 mg, 56% (over 2 steps)): R_f 0.38 (H_2O:MeOH 1:4, RP); $[\alpha]_D^{25}$ = +0.4 (*c* 1.5, $CHCl_3$, value of zero is consistent with compound being approximately racemic); mp 110–116 °C (from $CHCl_3$, wide range because compound is hygroscopic); ^{1}H NMR (400 MHz; $CDCl_3$) δ 7.13 (2H, dd, *J* 8.6, 5.5, *H*-2'), 7.05 (1H, dd, J_{AX} 5.0, J_{BX} 4.9, *H*-4''$_X$), 6.96 (2H, dd, J_{HF} 8.7, *J* 8.6, *H*-3'), 6.50 (1H, d, *J* 7.2, *H*-9''), 4.57 (1H, ddd, J_{BX} 7.5, *J* 7.2, J_{AX} 5.3, *H*-6''$_X$), 4.06 (1H, dd, J_{AB} 18.1, J_{AX} 5.0, H_aH_b-3''), 4.00 (1H, dd, J_{AB} 18.1, J_{BX} 4.9, H_aH_b-3''), 3.75 (3H, s, H_3-12''), 2.97 (1H, dd, J_{AB} 13.9, J_{AX} 5.3, H_aH_b-7''), 2.79 (1H, dd, J_{AB} 13.9, J_{BX} 7.5, H_aH_b-7''), 2.71–2.64 (2H, m, H_aH_b-3), 2.64–2.53 (2H, m, H_aH_b-1), 2.04 (3H, s, H_3-11''), 1.96–1.84 (2H, m, H_aH_b-2); ^{19}F NMR (376 MHz; $CDCl_3$) δ −117.5 (*F*-4'); *m/z* (ESI$^+$) 393 ([M+Na]$^+$, 100%); Analytical HPLC (method 1, 220/254 nm) ret. time = 5.1 min, purity = 98.6%; Chiral HPLC (90:10 hexane:IPA, 1.5 mL min^{-1}, 20 min) both enantiomers observed: ret. time = 12.7 min, area = 57.2% (D-**84**), ret. time = 15.2 min, area = 42.8% (L-**84** = **81**), 14.4% *e.e.* All data, except for the $[\alpha]_D^{25}$ and *e.e.* values, are in agreement with the values for compound **81** (reported above), consistent with **84** being an approximately racemic sample of **81**.

_N_-Acetyl-_S_-[3-(4-fluorophenyl)propyl]-L/D-cysteinylglycine (85)



LiOH·H_2O (0.37 M, 68 µL, 0.62 eq.) was added to a solution of methyl _N_-acetyl-_S_-[3-(4-fluorophenyl)propyl]-L/D-cysteinylglycinate, **84**, (15 mg, 0.040 mmol, 1.0 eq.) in THF:MeOH 3:2 (0.34 mL) at 0 °C. The resulting reaction mixture was stirred at 0 °C for 30 min. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated that some starting material (**84**) remained. Further aq. LiOH·H_2O (0.37 M, 68 µL, 0.62 eq.) was added, and the reaction was stirred at 0 °C for 1 h. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated that some starting material (**84**) remained. Further aq. LiOH·H_2O (0.37 M, 68 µL, 0.62 eq.) was added, and the reaction was stirred at 0 °C for 1 h. After this time, TLC analysis (H_2O:MeCN 1:1, RP) indicated complete consumption of the staring material (**84**). The reaction was acidified to pH 2 using 1 M aq. HCl. This crude mixture was purified directly by RP silica gel flash column chromatography (H_2O:MeCN 100:0 → 75:25 → 50:50) to give the title compound (**85**) as a colourless oil (3.6 mg, 26%); R_f 0.45 (H_2O:MeCN 1:1, RP); ^{1}H NMR (400 MHz, D_4-MeOH) δ 7.20 (2H, dd, _J_ 8.7, 5.5, _H_-2'), 6.98 (2H, dd, J_{HF} 8.8, _J_ 8.7, _H_-3'), 4.55 (1H, dd, J_{BX} 8.9, J_{AX} 5.1, _H_-6''$_X$), 3.85 (1H, d, J_{AB} 17.4, H_aH_b-3''), 3.76 (1H, d, J_{AB} 17.4, H_aH_b-3''), 3.03 (1H, dd, J_{AB} 13.9, J_{AX} 5.1, H_aH_b-7''), 2.75 (1H, dd, J_{AB} 13.9, J_{BX} 8.9, H_aH_b-7''), 2.72–2.66 (2H, m, H_aH_b-3), 2.61–2.50 (2H, m, H_aH_b-1), 2.02 (3H, s, H_3-11''), 1.93–1.82 (2H, m, H_aH_b-2); ^{19}F NMR (376 MHz; D_4-MeOH) δ −120.0 (_F_-4'); _m/z_ (ESI$^-$) 355 ([M−H]$^-$, 100%), 711 ([2M−H]$^-$, 32%); Analytical HPLC (method 2, 220/254 nm) ret. time = 4.5 min,

purity = 94.9%; Chiral HPLC (50:50:0.1 hexane:EtOH:TFA, 1.0 mL min^{-1}, 30 min) both enantiomers observed: ret. time = 9.1 min, area = 44.4% (L-**85** = **83**), ret. time = 22.5 min, area = 55.6% (D-**85**), 11.2% *e.e.* All data, except for the *e.e.* value, are in agreement with the values for compound **83** (reported above), consistent with **85** being an approximately racemic sample of **83**.

6.3 General Biochemistry Methods

6.3.1 Molecular Cloning to Synthesise Plasmids for Expression of *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 in *E. coli*

6.3.1.1 General Strategy

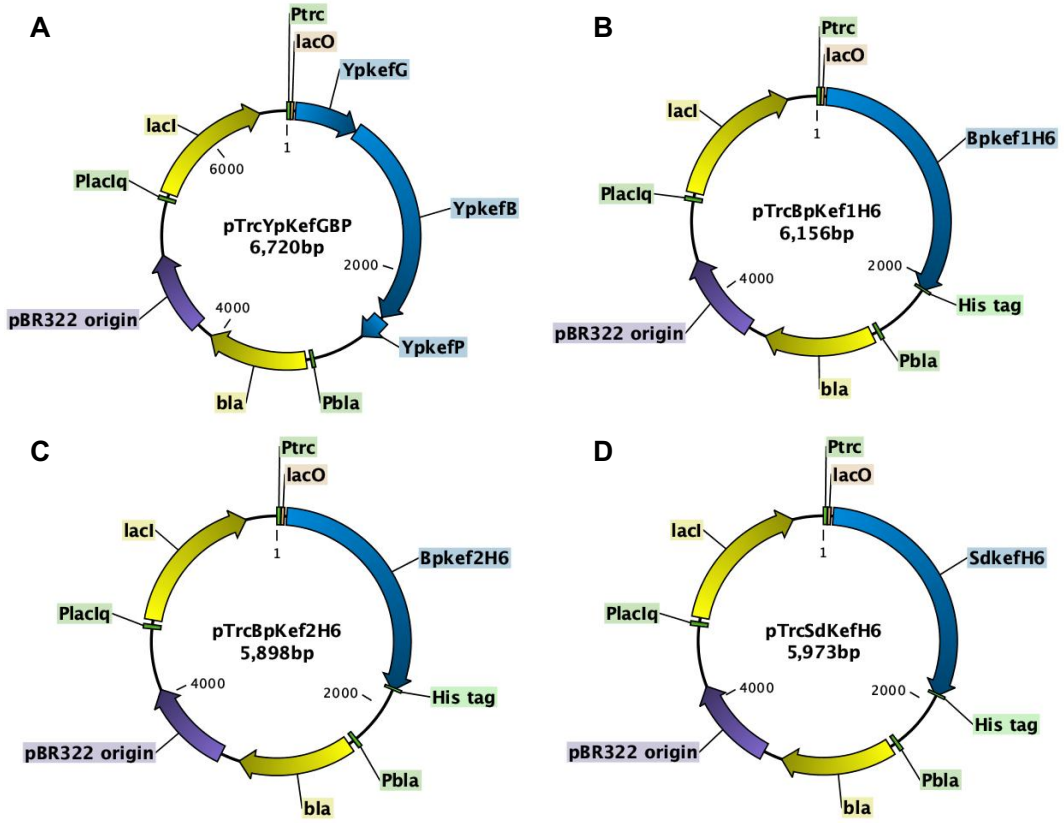


Figure 6.1. Plasmid maps of pTrcYpKefGBP (A), pTrcBpKef1H$_6$ (B), pTrcBpKef2H$_6$ (C), and pTrcSdKefH$_6$ (D). The *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, *Bpkef2H$_6$*, and *SdkefH$_6$* genes code for the *Yp*KefG, *Yp*KefB, *Yp*KefP, *Bp*Kef1H$_6$, *Bp*Kef2H$_6$, and *Sd*KefH$_6$ proteins, respectively. Ptrc is the promoter for the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, *Bpkef2H$_6$*, and *SdkefH$_6$* genes. *bla* is the gene for β-lactamase, which causes resistance to Amp. Pbla is the promoter for the gene *bla*. pBR322 origin is the origin of replication, which ensures that the plasmid is replicated when it is transformed into *E. coli*. *lacI* is the gene for LacI, which prevents transcription of the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, *Bpkef2H$_6$*, and *SdkefH$_6$* genes by binding to LacO. IPTG binds to LacI, preventing LacI from binding to LacO, promoting transcription of the *YpkefG*, *YpkefB*, *YpkefP*, *Bpkef1H$_6$*, *Bpkef2H$_6$*, and *SdkefH$_6$* genes. PlacIq is the promoter for the gene *lacI*.

The designed plasmids were pTrcYpKefGBP, pTrcBpKef1H$_6$, and pTrcBpKef2H$_6$ (Figure 6.1A–C). Genetic fragments containing *YpkefGBP*, *Bpkef1*, or *Bpkef2*, and codon-optimised for *E. coli*, were synthesised commercially. To prepare the pTrcSdKefH$_6$ vector (Figure 6.1D) for cloning, the *Sdkef* gene was removed by

restriction digestion. Gibson assembly[70,71] was then used to ligate the linearised vector and each synthetic gene fragment. This cloning strategy was employed successfully for pTrcYpKefGBP and pTrcBpKef2H$_6$. For pTrcBpKef1H$_6$, the commercial synthesis of the gene fragment for *Bpkef1* was unsuccessful. Consequently, pTrcBpKef1H$_6$ was prepared using traditional cloning.

6.3.1.2 Gibson Assembly to Produce pTrcBpKef2H$_6$ and pTrcYpKefGBP

The pTrcSdKefH$_6$ vector (Figure 6.2B) was prepared for insertion of the synthetic gene fragments, containing either *Bpkef2* or *YpkefGBP*, BpKef2 _frag and YpKefGBP_frag respectively, using restriction digestion. For insertion of BpKef2 _frag, the vector was digested using high-fidelity *Nco*I (*Nco*I-HF) and *Xho*I, and for YpKefGBP_frag, using *Nco*I-HF and *Hind*III-HF. For each digest, the two resulting fragments were separated using agarose gel electrophoresis (Figure 6.2A). In each case, the upper band, corresponding to the larger DNA fragment, the digested vector, was scised from the gel and purified. Gibson assembly[70,71] was then used to insert BpKef2_frag and YpKefGBP_frag into the two different vector fragments. After transformation and minipreparation, the clones were isolated, four per transformant (1-1, 1-2, 1-3, and 1-4 for *Bpkef2*, and 2-1, 2-2, 2-3, and 2-4 for *YpkefGBP*). Restriction pattern analysis (Figure 6.2E), using the restriction enzymes *Eco*RV-HF and *Hind*III-HF, showed that 1-3 (*Bpkef2*) and 2-1, 2-2, and 2-4 (*YpkefGBP*) gave restriction patterns consistent with pTrcBpKef2H$_6$ (Figure 6.2C) and pTrcYpKefGBP (Figure 6.2D), respectively. Subsequent plasmid sequencing of 1-3 and 2-4 confirmed that 1-3 was pTrcBpKef2H$_6$, and that 2-4 was pTrcYpKefGBP.

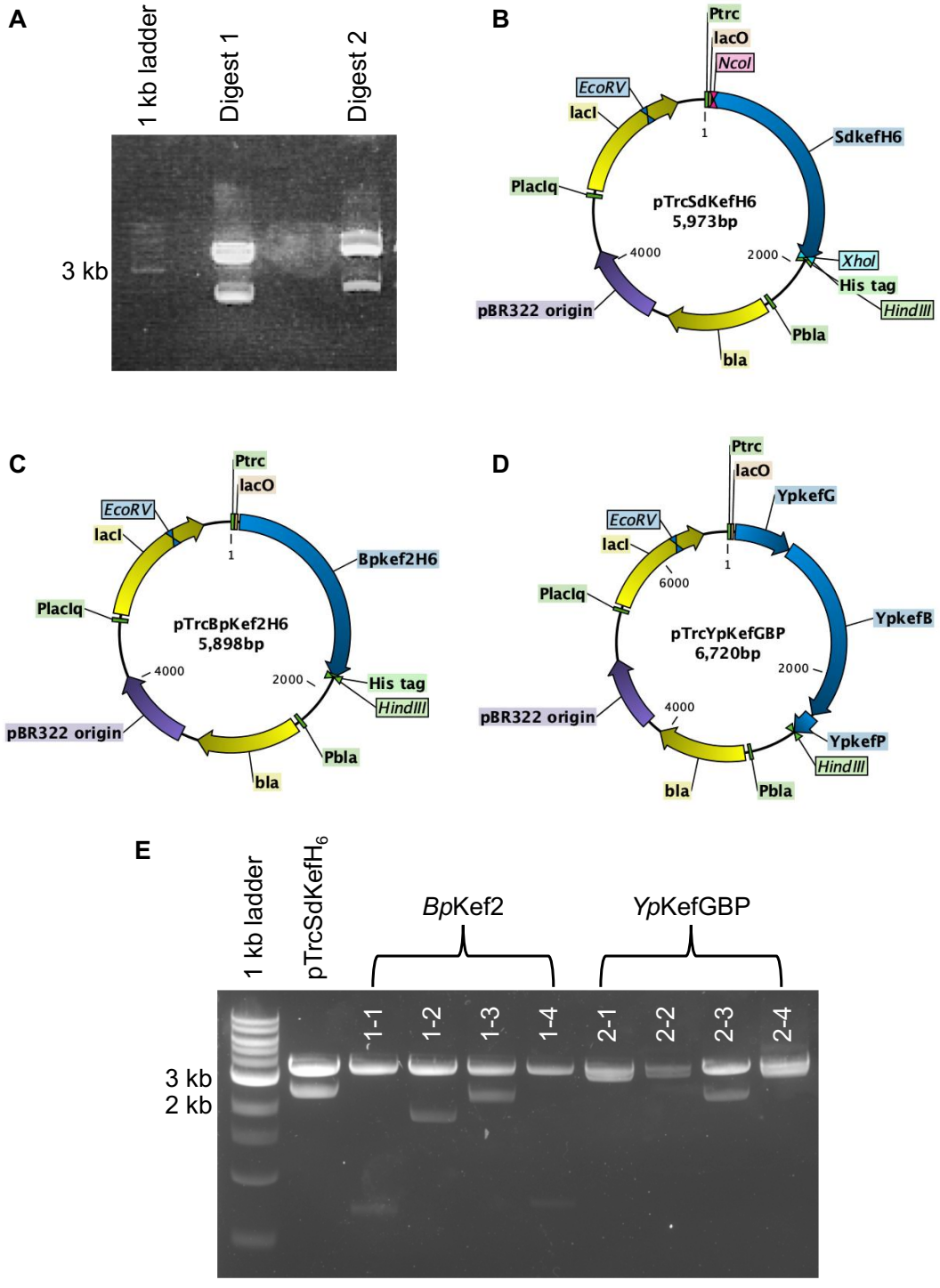


Figure 6.2. Molecular cloning to produce pTrcYpKefGBP and pTrcBpKef2H$_6$ using Gibson assembly. A) Agarose gel electrophoresis of the two different digests of the pTrcSdKefH$_6$ vector. Digest 1: *Nco*I-HF/*Xho*I, For insertion of *Bpkef2*. Digest 2: *Nco*I-HF/*Hind*III-HF, For insertion of *YpkefGBP*. B) Plasmid map of pTrcSdKefH$_6$, showing the *Eco*RV, *Nco*I, *Xho*I, and *Hind*III restriction sites and the His tag. C) Plasmid map of pTrcBpKef2H$_6$, showing the *Eco*RV and *Hind*III restriction sites and the His tag. D) Plasmid map of pTrcYpKefGBP, showing the *Eco*RV and *Hind*III restriction sites. E) Agarose gel electrophoresis of the *Eco*RV-HF/*Hind*III-HF digests of clones 1-1, 1-2, 1-3, and 1-4 (*Bpkef2*) and 2-1, 2-2, 2-3, and 2-4 (*YpkefGBP*). The pTrcSdKefH$_6$ vector was digested as a control.

6.3.1.3 Traditional Cloning to Produce pTrcBpKef1H$_6$

As commercial synthesis of the gene fragment for *Bpkef1* was unsuccessful, *Bpkef1* was made *via* full gene synthesis, where the *Bpkef1* gene was already inserted into a vector. *Bpkef1* was then removed from this shuttle vector and subcloned into the pTrc vector, using restriction enzyme-mediated cloning. This approach necessitates identical restriction sites at either end of the vector and insert, such that the vector and insert can be digested with the same restriction enzymes to generate complementary sticky ends, which are then ligated to generate the desired plasmid. The *Nco*I and *Xho*I sites in the pTrcSdKefH$_6$ vector were selected for the cloning, as these are located at either end of the *Sdkef* gene (Figure 6.3C). The *Bpkef1* insert, however, did not possess a *Nco*I site at the start of the *Bpkef1* gene and so this had to be added. There was also an undesired *Nco*I site within the *Bpkef1* construct, which needed to be removed. The required *Nco*I site was added by changing the codon for Ile2 to GTT (for Val). This changed the *N*-terminal sequence of the *Bp*Kef1 protein encoded by this plasmid from MISP to MVSP, but, as these two residues are very similar, it was assumed that the structure and function of the protein would not be drastically affected. The undesired *Nco*I site was removed from the construct by a CAT to CAC switch for the His621 codon.

This modified *Bpkef1* sequence was commercially synthesised inserted into a pMA-T shuttle vector (Bpkef1_pMA-T). Bpkef1_pMA-T was doubly digested using *Nco*I-HF and *Xho*I to liberate the *Bpkef1* insert from the vector. At the same time, the pTrcSdKefH$_6$ vector was doubly digested with the same restriction enzymes to liberate the unwanted *Sdkef* gene from the desired pTrc vector. The products of these digests

were separated using agarose gel electrophoresis (Figure 6.3A), and the bands corresponding to the pTrc vector and *Bpkef1* were separately scised and the DNA extracted. Ligation of the resulting pTrc vector and *Bpkef1* insert was then performed using T4 DNA ligase. Both 1:1.2 and 1:3.4 ratios of vector:insert were investigated. Each reaction mixture was separately used to transform *E. coli*. Three colonies were screened: one from the 1:1.2 ligation reaction, and two from the 1:3.4 ligation reaction. After minipreparation, these three clones (3-1, from the 1:1.2 reaction, and 3-2 and 3-3, from the 1:3.4 reaction) were screened through a digest using *Nco*I-HF and *Xho*I (Figure 6.3D). Only 3-3 showed the correct restriction pattern for pTrcBpKef1H_6 (Figure 6.3B), and so it was further characterised by a digest with *BamH*I-HF (Figure 6.3E). 3-3 Gave a restriction pattern consistent with pTrcBpKef1H_6 (Figure 6.3E), and sequencing data subsequently confirmed that it was the desired plasmid.

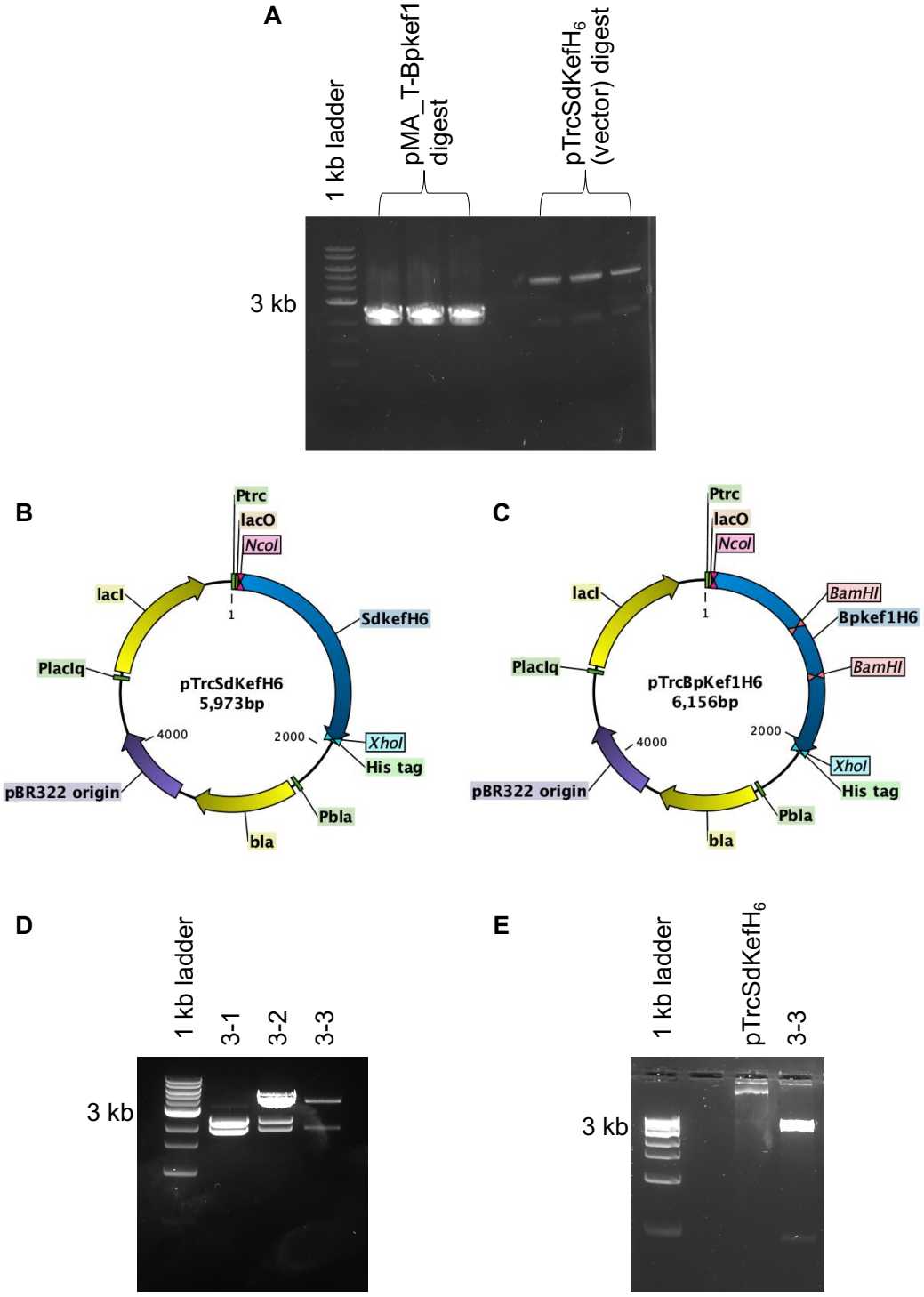


Figure 6.3. Molecular cloning to produce pTrcBpKef1H$_6$ using traditional cloning (*Nco*I-HF/*Xho*I). A) Agarose gel electrophoresis of the *Nco*I-HF/*Xho*I digests of both pMA_T-Bpkef1 and the pTrcSdKefH$_6$ vector. Due to their large volumes, both digests were loaded across three wells on the gel. B) Plasmid map of pTrcSdKefH$_6$, showing the *Nco*I and *Xho*I restriction sites and the His tag. C) Plasmid map of pTrcBpKef1H$_6$, showing the *Nco*I, *Xho*I, and *BamH*I restriction sites and the His tag. D) Agarose gel electrophoresis of the *Nco*I-HF/*Xho*I digests of *Bpkef1* clones 3-1, 3-2, and 3-3. E) Agarose gel electrophoresis of the *BamH*I-HF digest of 3-3. The pTrcSdKefH$_6$ vector was reacted with *BamH*I-HF as a control.

6.3.2 Molecular Cloning to Synthesise pTrcYpKefG(BH_6)P for Expression of *Yp*KefG(BH_6)P in *E. coli*

The pTrcYpKefGBP vector (Figure 6.4B) was first digested with restriction enzymes *Nco*I-HF and *Hind*III-HF to remove part of the *YpkefB* gene, from its the 3' end, and the entire *YpkefP* gene. The resulting pTrcYpKefGB' DNA fragment, corresponding to the upper band on the agarose gel (Figure 6.4C), was purified by gel extraction. This was then ligated to a *B"H_6P* synthetic gene fragment using Gibson assembly. The *B"H_6P* gene fragment contained part of the *YpkefB* gene, from its 3' end, corresponding to the section missing from the pTrcYpKefGB' DNA fragment. Following this section of *YpkefB*, the $(CAT)_6$ sequence was inserted, which codes for a His tag. Of the two codons for His, the CAT codon was selected over CAC, as it is more frequently used in *E. coli*, according to the *E. coli* Gene Expression Database.[72] After the $(CAT)_6$ sequence in the *B"H_6P* gene fragment was the entire *YpkefP* gene. The *B"H_6P* gene fragment also contained the overhangs required for ligation to the pTrcYpKefGB' vector fragment using Gibson assembly. After transformation of this Gibson assembly reaction mixture into *E. coli*, the plasmid DNA was extracted using minipreparation. The resulting 5 clones (sample names: 4-1, 4-2, 4-3, 4-4, and 4-5) were screened by a digest with restriction enzymes *Eco*RV-HF and *Hind*III-HF. Clones 4-1, 4-3, 4-4, and 4-5 gave the restriction pattern consistent with pTrcYpKefG(BH_6)P (Figure 6.4D). Sequencing of clone 4-3 subsequently confirmed that it was pTrcYpKefG(BH_6)P.

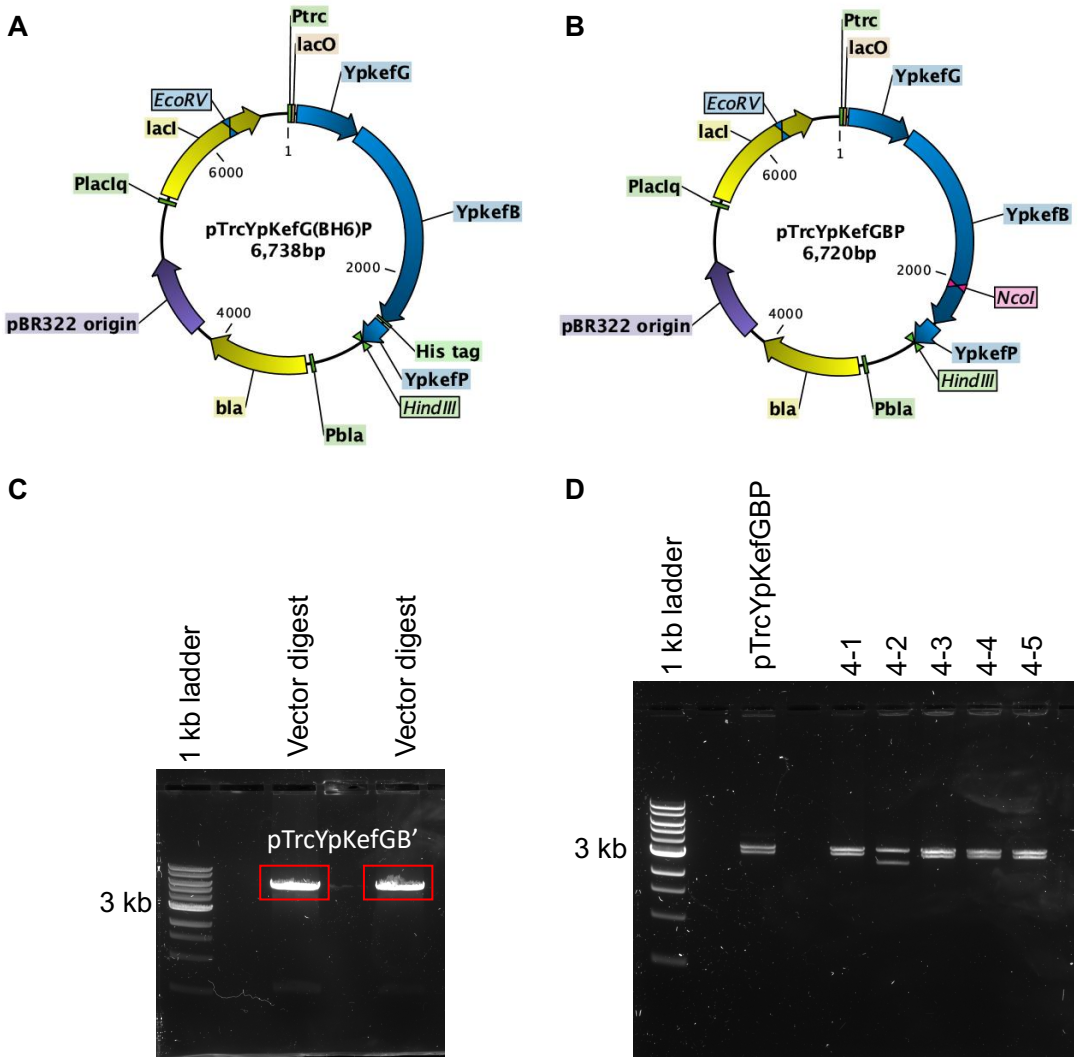


Figure 6.4. Molecular cloning to produce pTrcYpKefG(BH$_6$)P. A) Plasmid map of pTrcYpKefG(BH$_6$)P, showing the *Eco*RV and *Hind*III restriction sites and the His tag. B) Plasmid map of pTrcYpKefGBP, showing the *Eco*RV, *Nco*I, and *Hind*III restriction sites. C) Agarose gel electrophoresis of the digest of the pTrcYpKefGBP vector using *Nco*I-HF/*Hind*III-HF. Due to its large volume, the digested vector was loaded across two wells on the gel. The upper band in each digest lane, corresponding to the pTrcYpKefGB' DNA fragment, is labelled and highlighted in red. D) Agarose gel electrophoresis of the digest of *YpkefBH$_6$* clones 4-1, 4-2, 4-3, 4-4, and 4-5 using *Eco*RV-HF and *Hind*III-HF. The pTrcYpKefGBP vector was digested as a control.

6.3.3 Chemicals and Cell Culture Reagents

Name	Supplier
Acetic acid	Fisher Scientific
30% *w/w* Acrylamide/bis-acrylamide solution	Sigma-Aldrich
Agarose	Acros Organics, Fisher Scientific
Ammonium persulfate (APS)	Sigma-Aldrich
Ampicillin (Amp)	Apollo, Fisher Scientific
β-Mercaptoethanol	Sigma-Aldrich
Bacteriological (bacto) agar	Oxoid™
Benzamidine hydrochloride hydrate	Sigma-Aldrich
Benzonase nuclease®	Novagen®
Brilliant Blue G	Sigma-Aldrich
Calcium chloride	Merck
Deuterium oxide (D_2O)	Sigma-Aldrich
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
D_6-Dimethyl sulfoxide (D_6-DMSO)	Sigma-Aldrich
Disodium hydrogen phosphate (Na_2HPO_4)	Sigma-Aldrich
n-Dodecyl β-D-maltoside (DDM)	Apollo
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Ethanol	Fisher Scientific, Merck
L-Glutathione (GSH, **14**)	Fluorochem
Glycerol	Fisher Scientific

Name	**Supplier**
Glycine	ChemCruz, Fisher Scientific
4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Alfa Aesar
Imidazole	Alfa Aesar, Fluorochem
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Fisher Scientific
Methanol	Fisher Scientific
Methylglyoxal solution (~40% *v/v* in H_2O)	Sigma-Aldrich
Nickel (II) Sulfate ($NiSO_4$)	Acros Organics
Polyethyleneimine (PEI)	Sigma-Aldrich
Potassium chloride (KCl)	Fisher Scientific
Potassium dihydrogen phosphate (KH_2PO_4)	Sigma-Aldrich
Super optimal broth with catabolite repression (SOC) medium	NEB
Sodium azide (NaN_3)	MP Biomedicals
Sodium chloride (NaCl)	Fisher Scientific, Merck
Sodium dodecyl sulfate (SDS, dust-free pellets)	Sigma-Aldrich
Sodium dihydrogen phosphate dihydrate ($NaH_2PO_4 \cdot 2H_2O$)	Sigma-Aldrich
Sodium hydroxide (NaOH)	Fisher Scientific
Spectinomycin (Spec)	Apollo
SYBR Safe™	Invitrogen™
SYPRO® Orange	Invitrogen™
N,*N*,*N*',*N*'-Tetramethylethane-1,2-diamine (TEMED)	Sigma-Aldrich
1,1,1-Trifluoroacetone	Sigma-Aldrich

Name	Supplier
4-(Trifluoromethyl)-L-phenylalanine	Fluorochem
Tris-buffered saline (TBS)	Alfa Aesar
TBS with Tween® 20 (TTBS)	Alfa Aesar
Tris(hydroxymethyl)aminomethane (Tris, Trizma®)	Sigma-Aldrich
Tryptone	Oxoid™
Yeast Extract	Oxoid™

6.3.4 Proteins and Plasmids

*Sd*KefQCTDH$_6$ protein was provided by Holly Bunce, Dr Anthony Chan, and Mary Wilson (Conway Group, University of Oxford). *Sd*KefQCTDH$_6$(F448*p*tfmF), pTrcSdKefQCTDH$_6$, and pTrcSdKefH$_6$ were obtained from Dr Anthony Chan. The pULTRA-CNF plasmid[73] was a gift from Prof. Peter G. Schultz at the Scripps Research Institute (Addgene plasmid #48215).

6.3.5 Protein Concentration Measurement

Absorbance values of protein samples were measured using a NanoDrop Lite spectrophotometer (Thermo Scientific™). The protein concentrations were calculated using Beer-Lambert law: $A = \varepsilon lc$ (where A is absorbance, ε is molar extinction coefficient in $M^{-1}\,cm^{-1}$, l is the normalised path length of 1 cm, and c is concentration in M). The ε value for *Sd*KefQCTDH$_6$ is 34840 $M^{-1}\,cm^{-1}$, with respect to the protein homodimer. This coefficient was calculated using the ProtParam tool, using a duplicate primary amino acid sequence of *Sd*KefQCTDH$_6$ (minus the first N-terminal Met residue

in each monomer) as the query input. The ε value for *Sd*KefQCTDH$_6$(F448*p*tfmF) was assumed to be 34840 M^{-1} cm^{-1}.

6.3.6 pH Measurement

The pH values of buffers were determined using a VWR pH 100 pH meter. Electrodes were stored in a phthalate buffer solution (pH 4.0). Calibration of the pH meter was carried out before use, between pH 7.0–10.0, using buffer solutions of phosphate (pH 7.0) and borate (pH 10.0) from Fisher Scientific.

6.3.7 Bacterial Cell Culture

Standard sterile practices were followed throughout. Media and culture flasks were autoclaved (*Thermo Live Sciences* MAT90 LEI) at 121 °C for 15 min before use.

Optical density at 590, 600, 625, or 650 nm was measured using a Novaspec® II spectrophotometer, Thermo Fisher BioMate 3 UV spectrophotometer, or Biochrom WPA CO8000 cell density meter.

Cells were grown with shaking (180 rpm) at 37 °C in 2× TY medium [16 g L^{-1} tryptone, 10 g L^{-1} yeast extract, 5 g L^{-1} NaCl], unless otherwise specified.

When solid 2× TY medium was required, bacto agar was added to a final concentration of 1.5% *w/v*.

Ampicillin (Amp) and spectinomycin (Spec) were prepared as stock solutions at 100 mg mL^{-1} in Milli-Q® H_2O and sterilised by filtering through 0.22 µm filters (Millex® (33 mm) sterile filter unit, Merck). The final concentrations of Amp and Spec were both 100 µg mL^{-1}, unless otherwise specified.

6.4 Biochemistry Experimental Procedures

6.4.1 Constructs

Table 6.1. Plasmids used in this study.

Plasmid	Reference	Antibiotic Resistance
pTrcSdKefQCTDH$_6$	[31]	Amp
pTrcSdKefQCTDH$_6$(F448Am)	This dissertation	Amp
pULTRA-CNF	[73]	Spec
pTrcSdKefH$_6$	[74]	Amp
pTrcYpKefGBP	This dissertation	Amp
pTrcBpKef1H$_6$	This dissertation	Amp
Bpkef1_pMA-T	N/A	Amp
pTrcBpKef2H$_6$	This dissertation	Amp
pTrcYpKefG(BH$_6$)P	This dissertation	Amp

Bpkef1_pMA-T was synthesised by Invitrogen™ and was diluted to ~100 ng µL^{-1} in 10 mM Tris-HCl (pH 8.0, Thermo Fisher Scientific™).

6.4.2 Preparation of Competent *E. coli*

Competent *E. coli* were prepared using the $CaCl_2$ method.[75] A single colony was inoculated into 10 mL of 2× TY medium and incubated (~16–18 h). 400 µL Of this starter culture was subcultured into 100 mL of 2× TY medium and grown until OD_{600} ~0.4. Alternatively, a single colony was inoculated into 100 mL of 2× TY medium and grown until OD_{600} ~0.4. The culture flask(s) were placed on ice for 10 min. The culture was divided into 3 tubes and cells harvested by centrifugation (6000 × *g*, 5 min, 4 °C). Cell pellets were resuspended in freshly autoclaved $CaCl_2$ solution (15 mL, 60 mM $CaCl_2$, 15% glycerol, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.0, 4 °C). The cells were placed on ice for 0.5–8 h, harvested by centrifugation (6000 × *g*, 5 min, 4 °C), and resuspended in $CaCl_2$ solution (1.5 mL). Cells

were aliquoted (50 or 100 µL) in 2 mL Eppendorf tubes, and either used immediately or frozen in liquid N_2 (cold shock) and stored at −80 °C prior to use.

6.4.3 Molecular Cloning

6.4.3.1 Site-Directed Mutagenesis

Novel plasmid pTrcSdKefQCTDH$_6$(F448Am) was prepared by Dr Anthony Chan. The mutagenic primer pair, SdKef_TTCtoTAG-f and SdKef_TTCtoTAG-r (Table 6.2), were synthesised by Sigma-Aldrich Life Science, resuspended in Milli-Q® H_2O to a final primer concentration of 100 µM, and stored at −20 °C. The pTrcSdKefQCTDH$_6$ template[31] was amplified with SdKef_TTCtoTAG-f and SdKef_TTCtoTAG-r using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies). Products were transformed into either NEB 5-alpha or One Shot™ TOP10 competent *E. coli* cells. As the mutagenesis protocol involved amplification of the whole DNA template, mutation(s) may have occurred in location(s) other than where the mutation was desired. pTrcSdKefQCTDH$_6$(F448Am) was fully sequenced to ensure authenticity of the entire plasmid.

Table 6.2. Primers used in the site-directed mutagenesis of pTrcSdKefQCTDH$_6$ to pTrcSdKefQCTDH$_6$(F448Am). The intended mutations are highlighted in grey.

Primer	Orientation	Sequence (5' to 3')
SdKef_TTCtoTAG-f	Forward	CTCAAATTGACATGCTGCGCCG TTAGGGTAATAAGGTTTTCTA
SdKef_TTCtoTAG-r	Reverse	TAGAAAACCTTATTACCCTAAC GGCGCAGCATGTCAATTTGAG

6.4.3.2 Vector Linearisation

Restriction enzymes were purchased from New England Biolabs (NEB): *Nco*I-HF® (#R3193), *Hind*III-HF® (#R3104), and *Xho*I (#R0146). Vectors were digested using the

required restriction enzymes (Table 6.3) according to the manufacturer's (NEB) instructions. The *Bpkef1* gene insert was prepared by digestion of Bpkef1_pMA-T. In each reaction, 2–5 µg of vector was digested in a 20–80 µL reaction at 37 °C for 2–7 h, in a dry heat block or water bath. After digestion, the linearised vectors were purified using agarose gel electrophoresis.

Table 6.3. Restriction enzymes used for vector linearisation.

Initial Vector	Plasmid Product Target	Restriction Enzymes
pTrcSdKefH_6	pTrcYpKefGBP	*Nco*I-HF® *Hind*III-HF®
pTrcSdKefH_6 (vector) Bpkef1_pMA-T (insert)	pTrcBpKef1H_6	*Nco*I-HF® *Xho*I
pTrcSdKefH_6	pTrcBpKef2H_6	*Nco*I-HF® *Xho*I
pTrcYpKefGBP	pTrcYpKefG(BH_6)P	*Nco*I-HF® *Hind*III-HF®

6.4.3.3 Gibson Assembly

Ligation to produce pTrcYpKefGBP, pTrcBpKef2H_6, and pTrcYpKefG(BH_6)P was carried out using Gibson assembly.[70] The inserts used were DNA fragments YpKefGBP_frag, BpKef2_frag, and B'H_6_frag, which were synthesised by Thermo Fisher Scientific™ (GeneArt® Strings synthesis) and diluted to 20 ng μL^{-1} in Milli-Q® H_2O. Cloning was achieved using NEBuilder® HiFi DNA Assembly Master Mix (2×, NEB, #E2621), using 0.02–0.05 pmol of vector in a 1:2 ratio of vector:insert. Reactions (14–20 µL total volume) were run at 50 °C for 1 h. The reaction mixtures were used directly to transform NEB 5-alpha competent *E. coli* cells.

6.4.3.4 Ligation

Ligation to produce pTrcBpKef1H_6 was performed by a one-step isothermal reaction using T4 DNA ligase (NEB). Vector (39.5 fmol) and *Bpkef1* gene insert (46.4 fmol or 133 fmol), equivalent to a vector:insert ratio of 1:1.2 or 1:3.4, were reacted (20 µL total volume) at rt for 2 h. The reaction mixtures were used directly to transform NEB 5-alpha competent *E. coli* cells.

6.4.4 Agarose Gel Electrophoresis

DNA samples were analysed and purified by gel electrophoresis on 1% *w/v* agarose gels prepared with SYBR Safe™ DNA gel stain. Samples were mixed with 6× loading dye (blue (#B7021, NEB) or purple (#B7024, NEB)) before loading onto the gels. Gels were run in 1× Tris, acetic acid, and EDTA (TAE) buffer (diluted from 50× (Bio-Rad)) on a Bio-Rad system at a constant voltage of 100 V for 30 min. Gels were visualised using a UV transilluminator (Gel Logic 200) and the size of the fragments was referenced to a 1 kb DNA ladder (NEB, #N3232).

6.4.5 Extraction of DNA from Agarose Gels

After purification by agarose gel electrophoresis, digested DNA was purified using a GeneJET gel extraction kit (Thermo Scientific™, #K0691) according to the manufacturer's instructions. DNA Concentration was estimated by measuring the absorbance at 260 nm using a NanoDrop Lite spectrophotometer (Thermo Scientific™).

6.4.6 Transformation of Plasmid DNA into Bacteria

Competent *E. coli* cells (Table 6.4) were either used on ice immediately after preparation or thawed on ice for ~20–30 min before transformation. Once the cells had thawed, purified plasmid DNA (0.5–1.5 µL, ~60–290 ng μL^{-1} plasmid), a Gibson

assembly product (2 µL of mixture or the entire mixture), or an entire ligation reaction mixture, was added to 50 or 100 µL of competent cells. After the addition of plasmid DNA, each tube was tapped 10 times to mix the cells and plasmid and were not mixed again until plating. The cells were left on ice for 10–30 min before placement into a water (or dry) bath at 42 °C (heat shock) for exactly 30 s. The tubes were returned to ice for 5 min. If necessary, a recovery step was included as follows: SOC medium (950 µL) was added to the cells, which were incubated (37 °C, 1 h, water bath). For plasmids containing the *bla* gene, which confers Amp resistance to the resulting *E. coli* transformants, the recovery step can be omitted. The cells were mixed thoroughly, diluted with SOC medium (if necessary), and 5–350 µL of the mixture was plated onto 2× TY agar plates containing the appropriate antibiotic(s) and incubated (37 °C, ~12–24 h) before selection of well-isolated colonies.

Table 6.4. *E. coli* strains used in this study. NEB 5-Alpha cells were purchased from NEB (#C2987), and One Shot™ TOP10 from Invitrogen™ (#C404010).

***E. coli* strain**	**Genotype**	**Reference**
NEB 5-alpha	*fhu*A2 Δ(*arg*F-*lac*Z)U169 *pho*A *gln*V44 Φ80*lac*ZΔM15 *gyr*A96 *rec*A1 *rel*A1 *end*A1 *thi*-1 *hsd*R17	N/A
One Shot™ TOP10	F⁻ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	N/A
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	[76]
MJF645	MG1655, *kefF*, *kefC*, *kefG*, *kefB*, *gshA*	[31]
Frag 1	F⁻, *thi*, *rha*, *lacZ*	[77]
MJF335	Frag5, *lacI*, *kefB*, *kefC*::Tn*10* (*kan*), *gshA*::Tn*10* (*kan*)	[74]

6.4.7 Extraction of Plasmid DNA from *E. coli*

A single *E. coli* colony was inoculated into 10 or 100 mL of 2× TY medium, containing the appropriate antibiotic, and the culture was incubated (37 °C, 180 rpm, ~12–18 h). Plasmid DNA was purified using a GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific™, #K0503), according to the manufacturer's instructions. 4–6 mL Of cell culture was processed per miniprep column.

6.4.8 Restriction Pattern Analysis of Plasmid DNA

Clones of each plasmid made in this study were screened by digestion using restriction enzymes, and DNA fragments separated using agarose gel electrophoresis. Clones showing the restriction pattern consistent with the target plasmid were sequenced. The enzymes used for each digest are shown in Table 6.5. Restriction enzymes were purchased from NEB: *Eco*RV-HF® (#R3195) and *Bam*HI-HF (#R3136).

Table 6.5. Restriction enzymes used for digests of plasmid DNA.

Plasmid	Restriction Enzymes
pTrcYpKefGBP	*Eco*RV-HF®, *Hind*III-HF®
pTrcBpKef1H_6	Digest 1: *Nco*I-HF®, *Xho*I; Digest 2: *Bam*HI-HF®
pTrcBpKef2H_6	*Eco*RV-HF®, *Hind*III-HF®
pTrcYpKefG(BH_6)P	*Eco*RV-HF®, *Hind*III-HF®

6.4.9 Sequencing of Plasmid DNA

Sanger sequencing was performed by Source Bioscience. The primers used for each plasmid are shown in Table 6.6, and the primer sequences in Table 6.7. The upPtrc_f, dnPBADR, and PBADR sequencing primers are available from Source Bioscience. The other sequencing primers were synthesised by Sigma-Aldrich Life Science, resuspended in Milli-Q® H_2O to a final primer concentration of 100 µM, and stored at −20 °C. Melting temperatures (T_m) were predicted using the NEB T_m Calculator (http://tmcalculator.neb.com/#!/) using the settings for Q5 High-Fidelity DNA Polymerase.

Table 6.6. Sequencing primers used for each plasmid.

Plasmid	Sequencing Primers
pTrcSdKefQCTDH_6(F448Am)	uptrcp-f, PBADR, 30nt_M13R, Sdkef-f2, bla-f1, bla-f2, pBR322-f, PGEX 3’
pTrcYpKefGBP	upPtrc_f, YpKefGBP_seqf1, YpKefGBP_seqf2, dnPBADR
pTrcBpKef1H_6	upPtrc_f, BpKef1_Nseqf1, BpKef1_Nseqf2, BpKef1_seqf1, dnPBADR, PBDAR
pTrcBpKef2H_6	upPtrc_f, BpKef2_seqf1, BpKef2_seqf2, BpKef2_seqfA, BpKef2_seqrB, BpKef2_seqfC, BpKef2_seqrD, dnPBADR
pTrcYpKefG(BH_6)P	upPtrc_f, YpKefGBP_seqf1, YpKefGBP_seqf2, YpKefB_seqf3, YpKefB_seqf4, YpKefB_seqf5, YpKefB_seqr1, YpKefB_seqr2, dnPBADR

Table 6.7. Sequencing primers used in this study.

Sequencing Primer	Orientation	Sequence (5' to 3')
uptrcp-f	Forward	CTCGCAATCAAATTCAGCCGATA
PBADR	Reverse	GATTTAATCTGTATCAGG
30nt_M13R	Forward	AGCGGATAACAATTTCACACAGGA
Sdkef-f2	Forward	GAATACGTGTAAAGGTGGTGACTGT
bla-f1	Forward	GGTGCACGAGTGGGTTACATC
bla-f2	Forward	CGACGGGGAGTCAGGCAAC
pBR322-f	Forward	CGAAGGGAGAAAGGCGGACA
PGEX 3'	Forward	CCGGGAGCTGCATGTGTCAGAGG
upPtrc_f	Forward	TCCCGTTCTGGATAATG
YpKefGBP_seqf1	Forward	CTGGTTTTTCTGTTTGCA
YpKefGBP_seqf2	Forward	TATTGCAGGTATTCTGCT
YpKefB_seqf3	Forward	TGGCGATGTCGTCTA
YpKefB_seqf4	Forward	ATGGCATTGCAACTGAT
YpKefB_seqf5	Forward	TGGGGCTATCGATGG
YpKefB_seqr1	Reverse	CCATGCCGACAGAAATAA
YpKefB_seqr2	Reverse	GCCCTTTAAACGGCT
BpKef1_Nseqf1	Forward	GGTTCGCGGCGGTGA
BpKef1_Nseqf2	Forward	CCTCGCGCGGATGCT
BpKef1_seqf1	Forward	GCTTTGCCGCAGTTA
BpKef2_seqfA	Forward	CTGGTGCAATTGTTCAG
BpKef2_seqfC	Forward	TACTGCTGCCTCCTC
BpKef2_seqf1	Forward	TTCATGCCGGTGAAG
BpKef2_seqrD	Reverse	TAAACAGTTCACGGCTA
BpKef2_seqrB	Reverse	ATGCAACACCAACTGC
BpKef2_seqf2	Forward	AGCCGTGATGATCCA
dnPBADR	Reverse	CCAGGCAAATTCTGTTTTA

6.4.10 Expression and Purification of *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF)

Expression and Purification of *Sd*KefQCTDH$_6$

*Sd*KefQCTDH$_6$ was expressed and purified by Holly Bunce, Dr Anthony Chan, and Mary Wilson. *E. coli* BL21(DE3) was transformed with pTrcSdKefQCTDH$_6$. The transformant was aerobically cultured in 2 L of 2× TY medium (1 L medium per 2.5 L Tunair™ shake flask (Sigma-Aldrich)) containing Amp at 30 °C with an agitation speed of 180 rpm. When the bacterial culture reached OD_{600} = 1.991 ± 1.290 (mean ± s.d.; n = 3), IPTG was added (final concentration of 0.5 mM). After induction for 4.5 h, the bacterial cell pellet was collected by centrifugation (11325 × *g*, 4 °C, 5 min) and stored at –80 °C until protein purification. The *Sd*KefQCTDH$_6$ protein was purified at 4 °C by immobilised metal affinity chromatography (IMAC) and then size-exclusion chromatography (SEC), operated through a computerized ÄKTA fast protein liquid chromatography (ÄKTA-FPLC) system (GE Healthcare™).

Before purification, the cell pellet containing the target protein was completely suspended in extraction buffer (5 mL of extraction buffer/g of cell pellet). Extraction buffer was prepared by dissolving a Sigma*FAST*™ protease inhibitor cocktail tablet (EDTA-free, Sigma-Aldrich) in 100 mL binding buffer (Table 6.8). Benzonase® nuclease (250 U) was added to the cell suspension to degrade genomic DNA. The cells were lysed on ice by sonication (50% amplitude, 2 s pulse, 5 s rest, 60 cycles; VCX 500, Sonics & Materials) to release proteins. Polyethyleneimine, PEI, (0.15% *v*/*v*) was also added from a 5% *v*/*v* stock solution (pH 7.4) in Milli-Q® H_2O, and the mixture was incubated on ice for 15 min to precipitate DNA. Insoluble cell debris and precipitated DNA were removed by centrifugation (25000 × *g*, 4 °C, 15 min). The supernatant was

collected and filtered through 0.45 µm pore size syringe filters (Merck-Millipore). The clarified cell lysate was diluted to 100 mL with binding buffer supplemented with 20 mM imidazole and 10 mM β-mercaptoethanol.

Table 6.8. Composition of buffers for purification of *Sd*KefQCTDH$_6$ using IMAC and SEC. All buffers were filtered under vacuum through a 0.22 µm filter paper (Sartorius UK), and stored at 4 °C before use.

Buffer	Reagent	Concentration
Binding buffer (pH 7.4)	$NaH_2PO_4 \cdot 2H_2O$ Na_2HPO_4	50 mM (total)
	NaCl	500 mM
	Glycerol	10%
Strip buffer (pH 7.6)	HEPES	50 mM
	NaCl	500 mM
	EDTA	50 mM
Elution buffer (pH 7.4)	$NaH_2PO_4 \cdot 2H_2O$ Na_2HPO_4	50 mM (total)
	NaCl	500 mM
	Imidazole	500 mM
	Glycerol	10%
SEC buffer (pH 7.4)	$NaH_2PO_4 \cdot 2H_2O$ Na_2HPO_4	50 mM (total)
	NaCl	150 mM

For protein purification by IMAC, a HisTrap HP 5 mL column (GE Healthcare™) was used. The reusable HisTrap HP column was previously stripped with and stored in strip buffer (Table 6.8). To replenish Ni^{2+} ions to the affinity column, the column was rinsed with 10 CV of Milli-Q® H_2O, followed by recharging with 2 CV of 0.1 M $NiSO_4$. The column was equilibrated with 10 CV of binding buffer supplemented with 10 mM β-mercaptoethanol. The pre-treated cell lysate (100 mL mixture from the above sample preparation procedure) was then loaded into the equilibrated HisTrap HP column. Mixtures of binding buffer and elution buffer (Table 6.8) were used to remove non-specific binders and elute the protein of interest. The column was washed with

10 CV of 9% elution buffer (equivalent to 45 mM imidazole), then 10 CV of 18% elution buffer (equivalent to 90 mM imidazole). The protein was eluted in 2 mL fractions using a linear gradient (8 CV, 18–100% elution buffer, equivalent to ≤ 500 mM imidazole). The column was washed with 10 CV of Milli-Q® H_2O, then stripped with 10 CV of strip buffer.

The IMAC fractions were selected and further purified by SEC. Fractions were loaded *via* a 2 mL injection loop into a gel filtration column XK 16/70 (GE Healthcare™) packed with 120 mL of Superdex 75 resin (GE Healthcare™). This SEC column was pre-equilibrated with 150 mL degassed SEC buffer (Table 6.8). The SEC procedure was operated at a constant flow rate (1 mL min^{-1}) over a total volume of 150 mL, and fractions (5 mL) were collected in the 20–120 mL section of the run. The SEC-purified protein was concentrated using a Vivaspin™ sample concentrator (30 kDa molecular weight cut-off (MWCO), GE Healthcare™). The protein concentration was measured, and samples were retained for SDS-PAGE analysis. The protein was rapidly frozen in liquid N_2 and stored at –80 °C. The protein yield was 1.6 ± 0.3 mg protein/g pellet (mean ± s.d.; n = 3). The purity of *Sd*KefQCTDH$_6$ was verified using SDS-PAGE and liquid chromatography-mass spectroscopy (LC-MS).

Expression and Purification of *Sd*KefQCTDH$_6$(F448*p*tfmF)

*Sd*KefQCTDH$_6$(F448*p*tfmF) was expressed and purified by Dr Anthony Chan. *E. coli* BL21(DE3) was co-transformed with pTrc*Sd*KefQCTDH$_6$(F448Am) and pULTRA-CNF. The transformant was aerobically cultured in 2 L of 2× TY medium (1 L medium per 2.5 L Tunair™ shake flask (Sigma-Aldrich)) containing Amp and Spec. 4-(Trifluoromethyl)-L-phenylalanine (100 mM stock solution in 1 M NaOH, 1 mM final

concentration) was supplemented to the culture medium. When the bacterial culture reached OD_{600} = 1.953 ± 0.680 (mean ± s.d.; n = 3), IPTG was added (0.5 mM). After induction for 4.5 h, the bacterial cell pellet was collected by centrifugation (11325 × *g*, 4 °C, 5 min) and stored at –80 °C until protein purification. The *Sd*KefQCTDH$_6$(F448*p*tfmF) protein was purified by IMAC and then SEC, using the same procedures as for *Sd*KefQCTDH$_6$. The protein yield was 2.8 ± 0.3 mg protein/g pellet (mean ± s.d.; n = 3). The purity of *Sd*KefQCTDH$_6$(F448*p*tfmF) was verified using SDS-PAGE and LC-MS.

6.4.11 Expression of *Yp*KefBH$_6$ for Western Blotting Analysis

6.4.11.1 Growth Experiments

For all methods (1–3), two different *E. coli* strains, MJF645 and Frag 1, were each transformed with pTrcYpKefG(BH$_6$)P. For each of the two transformants, a single colony was inoculated into 10 mL of 2× TY medium (in a sterile Falcon™ 50 mL conical centrifuge tube, Thermo Fisher Scientific™) containing Amp and incubated (~13–16 h), as a starter culture.

Method 1

For each of the two transformants, 300 µL of starter culture was inoculated into 10 mL of 2× TY medium containing Amp as a pre-culture (to give OD_{650} ~0.05) and incubated until OD_{650} ~0.4. Subsequently, each pre-culture was diluted (~1:10) in 200 mL of 2× TY medium containing Amp (to give OD_{650} ~0.05) and incubated until OD_{650} ~0.4. The cultures were induced (0.3 mM IPTG) and incubated for another 30 min. Cell pellets were isolated by centrifugation (9600 × *g*, 4 °C, 10 min) and stored at −80 °C until membrane preparation.

Method 2

For each of the two transformants, 4.5 mL of starter culture was inoculated into 200 mL of 2× TY medium containing Amp (to give OD_{600} ~0.05) and incubated until OD_{600} ~0.5. The cultures were induced (0.3 mM IPTG) and incubated (~23 h). Cell pellets were isolated by centrifugation (9600 × *g*, 4 °C, 5 min) and stored at −80 °C until membrane preparation.

Method 3

For each of the two transformants, 8.5 mL of starter culture was inoculated into 900 mL of 2× TY medium containing Amp (in a 2.5 L Tunair™ shake flask (Sigma-Aldrich)) to give OD_{600} ~0.02 and incubated until OD_{600} ~1. The cultures were induced (0.3 mM IPTG) and incubated (~16 h). Cell pellets were isolated by centrifugation (11000 × *g*, 4 °C, 10 min) and stored at −80 °C until membrane preparation.

6.4.11.2 Membrane Preparation

Method 1

The two different cell pellets were thawed on ice. PBSi buffer was prepared by dissolving a Sigma*FAST*™ protease inhibitor cocktail tablet (EDTA-free, Sigma-Aldrich) in 100 mL PBS buffer (pH 7.6, Table 6.9). Each cell pellet was resuspended in 10 mL PBSi buffer. The cells were lysed on ice by sonication (50% amplitude, 2 s pulse, 5 s rest, 60 cycles; VCX 500, Sonics & Materials) to release proteins. Bulk cell debris was removed by centrifugation (4500 × *g*, 4 °C, 10 min). Membrane fractions were isolated by ultracentrifugation (90,000 × *g*, 4 °C, 1 h). For each sample, 4 mL of the supernatant

was collected, and the membrane fraction was resuspended in 250 µL PBSi buffer. All samples were stored at −80 °C until the Pierce™ 660 nm protein assay.

Table 6.9. Composition of 10 mM Na/K PBS buffer (pH 7.6). The buffer was filtered under vacuum through a 0.22 µm filter paper (Sartorius UK), and stored at 4 °C before use.

Reagent	Concentration (mM)
KH_2PO_4 Na_2HPO_4	10 (total)
NaCl	150

Method 2

Lysis, centrifugation, and ultracentrifugation were carried out as in Method 1. For each sample, 4.5 mL of the supernatant was collected, and the membrane fraction was resuspended in 250 µL PBSi buffer supplemented with 20% *v*/*v* glycerol. All samples were stored at −80 °C until the Pierce™ 660 nm protein assay.

Method 3

The two different cell pellets were thawed on ice. Each pellet was resuspended in 30 mL PBS buffer (Table 6.9). For each set of cells, 10 mL were transferred to a new tube and repelleted using centrifugation (7700 × *g*, 4 °C, 10 min). Each resulting cell pellet (~1.3 g) was resuspended in 13 mL lysis buffer (Table 6.10). The cells were lysed on ice by sonication (50% amplitude, 2 s pulse, 5 s rest, 60 cycles; VCX 500, Sonics & Materials) to release proteins. Bulk cell debris was removed by centrifugation (4000 × *g*, 4 °C, 20 min). Membrane fractions were isolated by ultracentrifugation (90,000 × *g*, 4 °C, 1 h). For each sample, the entire supernatant was collected. Each membrane fraction (0.18 g) was resuspended in 4.5 mL solubilisation buffer (Table 6.10) using a homogeniser and left to shake at 4 °C for 1 h. Both the supernatants and

membrane fractions were stored as 100 µL aliquots at −80 °C before analysis using SDS-PAGE and western blotting.

Table 6.10. Composition of lysis and solubilisation buffers.[37] Both buffers were filtered under vacuum through a 0.22 µm filter paper (Sartorius UK) before use, and stored at 4 °C.

Buffer	Reagent	Concentration
Lysis buffer (pH 7.8)	Trizma®	50 mM
	KCl	300 mM
	Benzamidine hydrochloride hydrate	1 mM
Solubilisation buffer (pH 7.8)	Trizma®	50 mM
	KCl	300 mM
	Benzamidine hydrochloride hydrate	1 mM
	n-Dodecyl β-D-maltoside (DDM)	1.5% *w/v*
	Glycerol	10% *v/v*
	Imidazole	25 mM

6.4.12 Pierce™ 660 nm Protein Assay

The concentration of *Yp*KefBH_6 protein samples (prepared in 6.4.11, Methods 1 and 2) was estimated using the Pierce™ 660 nm Protein Assay Kit (Thermo Scientific™, #23208), according to the manufacturer's instructions. Samples were prepared in Corning® 96-well tissue culture-treated microplates (Sigma-Aldrich). Absorbances were measured using a PHERAstar® *FS* microplate reader (BMG Labtech). Data were analysed and plotted using Microsoft® Excel version 16.21.1.

6.4.13 SDS-PAGE

SDS-PAGE was performed using a mini-PROTEAN® 3 electrophoresis module (Bio-Rad). The Tris-glycine discontinuous buffer system was used (Table 6.11). Gels (Table 6.12) were cast in mini-PROTEAN® 3 cells (Bio-Rad), using 70 × 100 mm glass plates, with 0.75 mm spacers. TEMED And APS were added immediately before casting.

Table 6.11. Composition of buffers for SDS-PAGE.

Buffer	Reagent	Composition	Amount required per litre
SDS-PAGE Running Buffer (10×)	Tris (Trizma®)	N/A	30 g
	Glycine	N/A	144 g
	SDS	N/A	10 g
SDS-PAGE Staining Solution	Brilliant Blue G	0.5% *w/v*	5 g
	Acetic acid	10% *v/v*	100 mL
	Methanol	50% *v/v*	500 mL
SDS-PAGE Destaining Solution	Acetic acid	10% *v/v*	100 mL
	Methanol	40% *v/v*	400 mL

Table 6.12. Composition of SDS-PAGE gels (quantities for 1 gel).

Reagent	Resolving gel (16%)	Stacking gel (4%)
Tris-HCl (1.5 M, pH 8.8)	1.25 mL	N/A
Tris-HCl (0.5 M, pH 6.8)	N/A	0.65 mL
Milli-Q® H_2O	1.05 mL	1.45 mL
Acrylamide/bis acrylamide (30% *w*/*w*)	2.65 mL	0.35 mL
SDS (10% *w*/*v*)	50 µL	25 µL
APS (10% *w*/*v*)	50 µL	25 µL
TEMED	4 µL	2.5 µL

6.4.13.1 Sample Preparation

Samples of *Yp*KefBH_6 (membrane fractions and supernatants from ultracentrifugation, 20 µL) and *Sd*KefQCTDH_6 (loading control, 10 µL) for SDS-PAGE were prepared in lithium dodecyl sulfate (LDS) sample buffer (diluted from 4× NuPAGE™ LDS sample buffer, #NP0007, Invitrogen™) containing 2.5% *v/v* β-mercaptoethanol and denatured at 70 °C for 10 min (dry heat block).

Samples (10, 15, or 20 µL) were loaded onto the cast gel (4–16% Bis-Tris gel) and run against colour prestained protein standard (11–245 kDa, #P7712, NEB). Gels were run at a constant voltage (100 or 200 V). After SDS-PAGE, protein bands in the gels were either visualised using Coomassie® blue stain (Table 6.11) (15 min stain, then

destained using destaining solution (Table 6.11) and Milli-Q® H_2O) or transferred to polyvinylidene difluoride (PVDF) membranes and analysed using western blotting.

6.4.14 Western Blot Analysis

Proteins in the electrophoresed gels were transferred onto PVDF membranes (Immun-Blot® PVDF membrane sandwiches, Bio-Rad) using a MiniTrans Blot® transfer system and transfer buffer (25 mM Tris, 192 mM glycine, 20% *v/v* MeOH, pH 8.3). Transfers were run for 1 h at 100 V (constant) and 350 mA. The membranes were washed with Tris-buffered saline (1× TBS, prepared from 10× TBS, pH 7.4) for 5–10 min, then blocked for 1 h using blocking solution (5% *w/v* Blotting grade blocker (Bio-Rad) in TBS). The membranes were washed with TBS containing Tween® 20 (1× TTBS, prepared from 10× TTBS (contains 0.5% Tween® 20)) twice (5–10 min per wash). The membranes were then incubated in blocking solution with 6× His epitope tag antibody (horseradish peroxidase (HRP) conjugate HIS.H8, 1:2000 dilution, Invitrogen™) for 3 h at rt with gentle shaking. The membranes were washed with 1× TTBS (4 × 5–10 min), then 1× TBS (5 min) prior to visualisation using colorimetric detection (Opti-4CN Substrate Kit, Bio-Rad).

6.4.15 LC-MS

LC-MS Experiments on *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF), expressed and purified as described in 6.4.10, were carried out by Holly Bunce, Dr Anthony Chan, and Mary Wilson. Experiments were performed at ambient temperature using a Meck chromolith C_{18} 2 × 5 mm guard column and a Waters LCT premier XE mass spectrometer equipped with electrospray interface. The method used was as follows:

Time (min)	Flow (mL min^{-1})	Solvent A (%)	Solvent B (%)
0	0.4	95	5
1	0.4	95	5
5	0.4	0	100
8	1	0	100
8.1	1	95	5

Mobile phases: A = 99.9% H_2O + 0.1% formic acid; B = MeCN.

Inspection and deconvolution of spectra were performed using MassLynx 4.1 software (Waters). Mass spectra are shown in the appendix.

6.4.16 CD

CD Experiments were carried out by Dr Anthony Chan. Purified *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) were exchanged into CD buffer (10 mM NaH_2PO_4/Na_2HPO_4, 20 mM NaCl, pH 7.4) using Slide-A-Lyzer™ MINI Dialysis Devices (2 kDa MWCO, Thermo Scientific™) at 4 °C for 12 h. Subsequently, both proteins were diluted to 7.3 µM in CD buffer. CD Measurements were made using a Jasco J-815 CD spectropolarimeter at 20 °C. Raw data were imported into Prism 6.0h (GraphPad software) for graph plotting.

6.4.17 AUC

Samples of *Sd*KefQCTDH$_6$(F448*p*tfmF) for analysis using AUC were prepared by Dr Anthony Chan. The concentration of *Sd*KefQCTDH$_6$(F448*p*tfmF) was 23.4 µM, with respect to its homodimer, in SEC buffer (6.4.10, Table 6.8). Sedimentation velocity experiments were performed by Dr David Staunton (Department of Biochemistry, University of Oxford) using a Beckman Optima XL-I analytical ultracentrifuge.

6.4.18 SEC-MALS

Samples of *Sd*KefQCTDH$_6$(F448*p*tfmF) for analysis using SEC-MALS were prepared by Dr Anthony Chan. *Sd*KefQCTDH$_6$(F448*p*tfmF) (before SEC purification) was tested at five different concentrations (25–200 µM) in SEC buffer (6.4.10, Table 6.8). and subjected to five SEC-MALS runs under the same conditions. Experiments were performed by Dr David Staunton (Department of Biochemistry, University of Oxford) using a Wyatt Dawn HELEOS-II 8-angle light scattering detector and Wyatt Optilab rEX refractive index monitor linked to a Shimadzu HPLC system comprising LC-20AD pump, SIL-20A autosampler, and SPD20A UV/Vis detector.

6.4.19 Biophysical Assays

6.4.19.1 ITC

All ITC experiments were carried out by Dr Anthony Chan.

Experiments were carried out on a MicroCal PEAQ-ITC Automated system (Malvern) and data were analysed using MicroCal PEAQ-ITC Analysis software (version 1.1.0.1262, Malvern), using a single site binding model. Nunc™ 96-well polypropylene DeepWell™ storage plates (#260251, Thermo Scientific™) were used. Proteins were centrifuged at 13000 rpm (Hettich MIKRO 20 Hematocrit Centrifuge, MedWOW Ltd.)

at room temperature for 1 min to remove insoluble protein precipitates. If necessary, the protein samples were buffer exchanged into SEC buffer (6.4.10, Table 6.8) using PD SpinTrap G-25 columns (GE Healthcare™). Ligand stock solutions were prepared in SEC buffer. Protein and ligand solutions were prepared at the concentrations indicated (Table 6.13). Protein solutions in the sample cell were titrated with the ligand solutions in the syringe. The ligands were titrated into the protein solution (19 injections, 1 × 0.4 and 18 × 2.0 µL), with 150 s spacing between injections. Control experiments (heats of ligand dilution) were performed by titrating the ligand solutions into SEC buffer. The integrated heat data were corrected for the heats of ligand dilution. Each protein/ligand combination was tested in three independent ITC experiments (n = 3).

Table 6.13. Protein and ligand concentrations used in ITC experiments.

Ligand	**Protein Concentration (µM)** Ligand Concentration (µM)	
	***Sd*KefQCTDH$_6$**	***Sd*KefQCTDH$_6$(F448*p*tfmF)**
^{t}BuSG (**26**)	**25** 300	**25** (n = 1), **50** (n = 2) 800
ESG (**20**)	**25** 500, 600, 800 (all n = 1)	**25** 4000 (n = 2), 5000 (n = 1)
29	**25** 300	**25** 1000

6.4.19.2 WaterLOGSY

WaterLOGSY Assays for L-Glutathione (GSH, 14) with *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF)

The waterLOGSY assays for GSH (**14**) with *Sd*KefQCTDH$_6$ and *Sd*KefQCTDH$_6$(F448*p*tfmF) were carried out by Dr Anthony Chan.

A water-suppressed ^{1}H NMR spectrum for GSH (**14**) was recorded as a reference, followed by a waterLOGSY spectrum of GSH (**14**) in the absence of the protein, and

finally a waterLOGSY spectrum in the presence of the specified protein. Samples were prepared to a total volume of 550 µL in SEC buffer (6.4.10, Table 6.8), with 5 mM GSH (**14**) from a 80 mM stock solution (in SEC buffer), 10% *v*/*v* D_2O (for signal locking), and 120 µM protein (in SEC buffer, for samples with protein). Samples were loaded into Norell® Select Series™ 5 mm NMR tubes (Sigma-Aldrich), and briefly centrifuged in a manual benchtop rotor. Spectra were recorded on a Bruker AVIIIHD 600 MHz spectrometer with a 5 mm BB-F/^{1}H Prodigy N_2 broadband cryoprobe, using a standard waterLOGSY pulse sequence.[78]

WaterLOGSY Assays for Dipeptides with *Sd*KefQCTDH$_6$

The waterLOGSY assays for dipeptide ligands (**30**, **39**, **82**, and **83**) with *Sd*KefQCTDH$_6$ were carried out with Holly Bunce.

If necessary, the samples of *Sd*KefQCTDH$_6$ protein were buffer exchanged into PBS buffer (Table 6.9) using PD SpinTrap G-25 columns (GE Healthcare™). For each compound, a water-suppressed ^{1}H NMR spectrum was recorded as a reference, followed by a waterLOGSY spectrum of the compound in the absence of the protein, and finally a waterLOGSY spectrum in the presence of the specified protein. Samples were prepared to a total volume of 160 µL in PBS buffer (Table 6.9), with 500 µM compound from 50 mM or 100 mM stock solutions (in D_6-DMSO), 10% *v*/*v* D_2O (for signal locking), and 20 µM protein (for samples with protein). Samples were loaded into 3 mm Hilgenberg NMR tubes for Bruker Match™ system, and briefly centrifuged in a manual benchtop rotor. Spectra were acquired at 298 K using either a 750 MHz Bruker Avance III HD spectrometer with 5 mm TCI cryoprobe (for **30** only), or a Bruker

AVIII 700 MHz spectrometer with inverse TCI cryoprobe (for all other ligands), using a standard waterLOGSY pulse sequence.[78]

6.4.19.3 Protein-Observed ^{19}F NMR Spectroscopy

All ^{19}F NMR spectroscopy experiments were carried out by Dr Anthony Chan.

A fixed concentration of *Sd*KefQCTDH$_6$(F448*p*tfmF) (23.4 μM) was titrated with various concentrations of ligand to the point of saturation. Each combination of protein and ligand was freshly prepared in an individual NMR tube (Norell® Select Series™ 5 mm NMR tube, Sigma-Aldrich) to avoid protein aggregation and dilution effects caused by adding extra ligand to the same sample. Samples were equilibrated at rt for at least 10 min before acquisition. 1,1,1-Trifluoroacetone was used as an internal ^{19}F standard. ^{19}F NMR Spectra were acquired at 298 K using a Bruker AVIIIHD 600 MHz spectrometer with a 5 mm BB-F/^{1}H Prodigy N_2 broadband cryoprobe. TopSpin software (version 3.5pl6, Bruker) was used to operate data acquisition, using the zgflqn pulse program (^{1}H-coupled ^{19}F NMR spectroscopy).

6.4.19.4 DSF

The DSF assay was carried out with Holly Bunce.

Experiments were performed using a Stratagene Mx3005P real-time polymerase chain reaction (qPCR) system, with filters for the SYPRO® Orange dye (excitation at 465 nm, emission at 590 nm). The set initial temperature was 25 °C, increasing in increments of 1 °C per 30 s to 85 °C, with fluorescence measurements at each interval. The melting temperature (T_m) was identified by fitting to the Boltzmann equation,[64] using Prism 8 software. The change in melting temperature (ΔT_m) was calculated as the shift in T_m

relative to the T_m of the protein in the absence of ligand. For selected ligands (**26**/**30** and **40**/**83**), the significance of changes was evaluated using an unpaired t test with Welch's correction. P Values (two-tailed) were defined as follows: ****P ≤ 0.0001 (extremely significant), ***P ≤ 0.001 (extremely significant), **P ≤ 0.01 (very significant), *P ≤ 0.05 (significant), and P > 0.05 (not significant, ns). Stock solutions of ligands were prepared at a final concentration of 100 mM in DMSO. These stock solutions were diluted to a concentration of 10 mM in PBS buffer (Table 6.9). A protein master mix was prepared, which contained *Sd*KefQCTDH_6 (12 µM) and SYPRO® Orange dye (1:200 dilution). 96-Well PCR plates (Axygen) were prepared using the diluted ligand solutions (2.5 µL) and the protein master mix (22.5 µL). The 96-well PCR plates were sealed and centrifuged (1000 rpm, 2 min, 25 °C) before measurement.

6.4.20 Assays in Bacteria

6.4.20.1 Kirby-Bauer Disc Diffusion Assays

Disc Assays in *E. coli*

The full-length *Sd*Kef (GenBank ID: ABE53663.1), *Yp*KefGBP, *Bp*Kef1, *Bp*Kef2, and *Yp*KefG(BH_6)P proteins, encoded by the pTrcSdKefH_6,[74] pTrcYpKefGBP, pTrcBpKef1H_6, pTrcBpKef2H_6, and pTrcYpKefG(BH_6)P plasmids, respectively, were transformed into the *E. coli* hosts, as described in 6.4.6. For each isolate, 3–5 morphologically similar colonies were inoculated into 5 mL of 2× TY medium (in a sterile Falcon™ 50 mL conical centrifuge tube, Thermo Fisher Scientific™) containing Amp for *E. coli* transformants, and 1 mM GSH (**14**, from a 250 mM stock solution in Milli-Q® H_2O) for MJF335 + GSH cells, using an inoculation loop. Subsequently, the bacterial cultures were incubated.

During this incubation period, solutions of test compounds (at the concentrations indicated, in either Milli-Q® H_2O or 80% *v/v* DMSO) were prepared. A stock solution of MG, **13**, (~84 mM, in 80% *v/v* DMSO) was prepared from MG solution (~40% *v*/*v* in H_2O). The stock solution was serially diluted twofold in 80% *v/v* DMSO to give the solutions of lower concentration (~2.6–42 mM). Compound discs were prepared by impregnating 20 µL of solutions of test compounds onto Whatman® antibiotic assay discs of 6 mm diameter (Sigma-Aldrich). Negative control discs containing the appropriate solvent (Milli-Q® H_2O or 80% *v/v* DMSO) were also prepared.

If necessary, the *E. coli* cultures were diluted to the required cell density of 0.5 McFarland turbidity standard[79] using 2× TY medium, and inoculated onto 2× TY agar plates (containing Amp for all *E. coli* transformants, 0.05 mM or 0.1 mM IPTG (where indicated), and 1 mM GSH (**14**) for MJF335 + GSH cells) using a sterile cotton swab. Subsequently, the compound discs were aseptically placed onto the bacterial plates. After 24 h of incubation at 37 °C, inhibition zone diameters associated with the inhibition effect of the compounds were recorded by averaging both the vertical and horizontal diameters of each zone and are given to the nearest whole millimetre. All experimental procedures were performed according to the standards set by the Clinical and Laboratory Standards Institute.[80,81]

Disc Assays in Wild-Type *Y. Pseudotuberculosis*

The Kirby-Bauer disc diffusion assays in *Y. pseudotuberculosis* IP32953, and *E. coli* MJF645 controls, described in 3.8, were carried out in the microbiology laboratories at DSTL, Porton Down.

The full-length *Yp*KefGBP, *Bp*Kef1, and *Bp*Kef2 proteins, encoded by the pTrcYpKefGBP, pTrcBpKef1H_6, and pTrcBpKef2H_6 plasmids, respectively, were transformed into *E. coli* MJF645. At the same time, an overnight culture of *Y. pseudotuberculosis* IP32953 was set up, using colonies obtained from Dr Andrew E. Scott (DSTL). These colonies were prepared by inoculation of *Y. pseudotuberculosis* IP32953 onto 2× TY agar plates and incubation at 37 °C for 24 h. For the overnight culture, one colony was inoculated into 5 mL of 2× TY medium in a sterile Universal 30 mL centrifuge tube (DD Biolab) and incubated (30 °C, 180 rpm, ~18 h).

On the following day, all bacterial cultures for the disc assay were set up and grown simultaneously. For *Y. pseudotuberculosis* IP32953, 5 µL of the overnight culture was subcultured into 5 mL of 2× TY medium in a sterile Universal 30 mL centrifuge tube (DD Biolab). For each *E. coli* transformant, 3 morphologically similar colonies were inoculated into 5 mL of 2× TY medium (in a sterile Universal 30 mL centrifuge tube, DD Biolab) containing Amp. All bacterial cultures were incubated simultaneously, at either 37 °C (for the *E. coli* transformants) or 30 °C (for *Y. pseudotuberculosis* IP32953), with agitation (180 rpm).

During this incubation period, the OXFKEF14 (**33**) compound discs were prepared, in the same manner as described in the previous section. The cell cultures were diluted to 0.5 McFarland, and the plates for the assay were set up, both in the same way as described in the previous section. The plates were incubated at either 37 °C (for the *E. coli* transformants) or 30 °C (for *Y. pseudotuberculosis* IP32953). Inhibition zone diameters were recorded as described in the previous section. All experimental

procedures were performed according to the standards set by the Clinical and Laboratory Standards Institute.[80,81]

Disc Assay in Wild-Type *Y. Pestis*

The Kirby-Bauer disc diffusion assay in *Y. pestis* CO92, described in 3.8, was performed by Dr Andrew E. Scott (DSTL).

Y. pestis CO92 was streaked from a glycerol stock (stored at −80 °C) onto Brucella agar base (BAB)-haemin agar and incubated at 28 °C for 48 h. 5 To 10 colonies were inoculated into 10 mL BAB-haemin broth and incubated (28 °C, 180 rpm, 48 h). 1 mL Of this culture was subcultured into 9 mL of BAB-haemin broth and incubated (28 °C, 180 rpm, 48 h) until OD_{590} ~0.6. This culture was used directly to swab BAB-haemin agar plates for the assay. The compound discs and assay plates were prepared in the same manner as described previously in this dissertation. Plates were incubated at 28 °C for 48 h to allow sufficient bacterial growth.

6.4.20.2 MIC and MBC Assays

MIC Assay

OXFKEF14 (**33**) (2 mg mL^{-1} in 2× TY medium) was serially diluted twofold into 2× TY medium containing Amp (200 µg mL^{-1}) and IPTG (0.2 mM) across four lanes in the interior wells of a 96-well flat-bottom polystyrene culture plate (Corning Costar). The final concentration of OXFKEF14 (**33**) was 3.91–1000 µg mL^{-1} in 2× TY medium, with final assay volumes of 100 µL per well. 200 µL Of sterile 2× TY medium containing Amp (200 µg mL^{-1}) and IPTG (0.2 mM) was used as the blank, and 100 µL of the same medium was used as a negative control. An aliquot of the overnight culture of *E. coli* MJF645 transformed with pTrcYpKefGBP (prepared as a control for the Kirby-Bauer

disc diffusion assay in *Y. pseudotuberculosis*, described in 6.4.20.1) was subcultured into 2× TY medium to give an OD_{600} = 0.37. Amp stock solution was subsequently added to a final concentration of 100 μg mL^{-1}. This cell culture was serially diluted tenfold into two sterile Universal 30 mL centrifuge tubes (DD Biolab), each containing 9 mL of sterile 2× TY medium supplemented with Amp. 100 μL Of the resulting bacterial culture was subsequently added to each negative control and compound well. The plate was covered and incubated (37 °C, ~18 h). After incubation, OD_{600} measurements were made using a plate reader (Thermo Multiskan FC, Thermo Fisher Scientific™). The MIC was the lowest concentration of OXFKEF14 (**33**) at which mean OD_{600} (compound wells) < 20% mean OD_{600} (negative control wells).

MBC Assay

After the MIC assay, 100 μL from the compound wells indicated, which showed growth inhibition/death such that the mean OD_{600} (compound wells) < 20% mean OD_{600} (negative control wells), were each subcultured into 10 mL of 2× TY medium in a sterile Universal 30 mL centrifuge tube (DD Biolab) containing Amp and 0.1 mM IPTG. Cell culture from a compound well which did not show growth inhibition/death, according to criteria above, was subcultured as a positive control. 2× TY Medium containing Amp and 0.1 mM IPTG was subcultured as a negative control. The resulting cultures were incubated (~68 h). The MBC was the lowest concentration showing no visible bacterial growth.

References

1 J. W. Bartholomew and T. Mittwer, *Bacteriol. Rev.*, 1952, **16**, 1–29.
2 R. Coico, *Curr. Protoc. Immunol.*, 1997, **23**, A.3O.1-A.3O.2.
3 Gram Staining, https://ib.bioninja.com.au/options/untitled/b1-microbiology-organisms/gram-staining.html, (accessed May 2019).
4 G. D. Wright, *Nat. Rev. Microbiol.*, 2007, **5**, 175–186.
5 D. Schillaci, V. Spanò, B. Parrino, A. Carbone, A. Montalbano, P. Barraja, P. Diana, G. Cirrincione and S. Cascioferro, *J. Med. Chem.*, 2017, **60**, 8268–8297.
6 H. C. Neu, *Science*, 1992, **257**, 1064–1073.
7 M. F. Chellat, L. Raguž and R. Riedl, *Angew. Chemie Int. Ed.*, 2016, **55**, 6600–6626.
8 E. D. Brown and G. D. Wright, *Nature*, 2016, **529**, 336–343.
9 A. P. Magiorakos, A. Srinivasan, R. B. Carey, Y. Carmeli, M. E. Falagas, C. G. Giske, S. Harbarth, J. F. Hindler, G. Kahlmeter, B. Olsson-Liljequist, D. L. Paterson, L. B. Rice, J. Stelling, M. J. Struelens, A. Vatopoulos, J. T. Weber and D. L. Monnet, *Clin. Microbiol. Infect.*, 2011, **18**, 268–281.
10 L. B. Rice, *J. Infect. Dis.*, 2008, **197**, 1079–1081.
11 J. O. Neill, *Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations. The Review on Antimicrobial Resistance.*, 2014.
12 J. Davies, *Can. J. Infect. Dis. Med. Microbiol.*, 2006, **17**, 287–290.
13 M. S. Butler, M. A. Blaskovich and M. A. Cooper, *J. Antibiot. (Tokyo).*, 2013, **66**, 571–591.
14 D. Friedman and J. Alper, *Technological Challenges in Antibiotic Discovery and Development: A Workshop Summary*, The National Academies Press, Washington, DC, 2014.
15 S. K. Urahn, A. Coukell, E. Jungman, K. Talkington, C. Shore and P. Kothari, *A Scientific Roadmap for Antibiotic Discovery*, 2016.
16 M. A. Cooper and D. Shlaes, *Nature*, 2011, **472**, 32.
17 G. L. Patrick, in *An Introduction to Medicinal Chemistry*, Oxford University Press, Oxford, 5th edn., 2013, pp. 413–513.
18 K. Lewis, *Nat. Rev. Drug Discov.*, 2013, **12**, 371–387.
19 M. S. Butler, M. A. Blaskovich and M. A. Cooper, *J. Antibiot. (Tokyo).*, 2017, **70**, 3–24.
20 D. J. Newman and G. M. Cragg, in *Comprehensive Natural Products II: Chemistry and Biology*, Elsevier Ltd, Oxford, 1st edn., 2010, pp. 623–650.
21 P. J. Choi, H. S. Sutherland, A. S. T. Tong, A. Blaser, S. G. Franzblau, C. B. Cooper, M. U. Lotlikar, A. M. Upton, J. Guillemont, M. Motte, L. Queguiner, K. Andries, W. Van den Broeck, W. A. Denny and B. D. Palmer, *Bioorganic Med. Chem. Lett.*, 2017, **27**, 5190–5196.
22 WHO, *Antimicrobial Resistance: Global Report on Surveillance*, 2014.
23 U. Theuretzbacher, S. Gottwalt, P. Beyer, M. Butler, L. Czaplewski, C. Lienhardt, L. Moja, M. Paul, S. Paulin, J. H. Rex, L. L. Silver, M. Spigelman, G. E. Thwaites, J. P. Paccaud and S. Harbarth, *Lancet Infect. Dis.*, 2019, **19**, e40–e50.
24 W. Epstein, *Prog. Nucleic Acid Res. Mol. Biol.*, 2003, **75**, 293–320.
25 P. Chandrangsu, R. Dusi, C. J. Hamilton and J. D. Helmann, *Mol. Microbiol.*, 2014,

91, 706–715.
26 R. S. Ackerman, N. R. Cozzarelli and W. Epstein, *J. Bacteriol.*, 1974, **119**, 357–362.
27 I. R. Booth, G. P. Ferguson, S. Miller, C. Li, B. Gunasekera and S. Kinghorn, *Biochem. Soc. Trans.*, 2003, **31**, 1406–1408.
28 G. P. Ferguson, S. Tötemeyer, M. J. MacLean and I. R. Booth, *Arch. Microbiol.*, 1998, **170**, 209–219.
29 D. McLaggan, H. Rufino, M. Jaspars and I. R. Booth, *Appl. Environ. Microbiol.*, 2000, **66**, 1393–1399.
30 B. D. Bennett, E. H. Kimball, M. Gao, R. Osterhout, S. J. Van Dien and J. D. Rabinowitz, *Nat. Chem. Biol.*, 2009, **5**, 593–599.
31 J. Healy, S. Ekkerman, C. Pliotas, M. Richard, W. Bartlett, S. C. Grayer, G. M. Morris, S. Miller, I. R. Booth, S. J. Conway and T. Rasmussen, *Biochemistry*, 2014, **53**, 1982–1992.
32 G. P. Ferguson, D. McLaggan and I. R. Booth, *Mol. Microbiol.*, 1995, **17**, 1025–1033.
33 S. Miller, L. S. Ness, C. M. Wood, B. C. Fox and I. R. Booth, *J. Bacteriol.*, 2000, **182**, 6536–6540.
34 G. P. Ferguson, Y. Nikolaev, D. McLaggan, M. Maclean and I. R. Booth, *J. Bacteriol.*, 1997, **179**, 1007–1012.
35 G. P. Ferguson, A. W. Munro, R. M. Douglas, D. McLaggan and I. R. Booth, *Mol. Microbiol.*, 1993, **9**, 1297–1303.
36 E. Ozyamak, S. S. Black, C. A. Walker, M. J. MacLean, W. Bartlett, S. Miller and I. R. Booth, *Mol. Microbiol.*, 2010, **78**, 1577–1590.
37 C. Pliotas, S. C. Grayer, S. Ekkerman, A. K. N. Chan, J. Healy, P. Marius, W. Bartlett, A. Khan, W. A. Cortopassi, S. A. Chandler, T. Rasmussen, J. L. P. Benesch, R. S. Paton, T. D. W. Claridge, S. Miller, I. R. Booth, J. H. Naismith and S. J. Conway, *Biochemistry*, 2017, **56**, 4219–4234.
38 W. Li, A. Cowley, M. Uludag, T. Gur, H. McWilliam, S. Squizzato, Y. M. Park, N. Buso and R. Lopez, *Nucleic Acids Res.*, 2015, **43**, W580–W584.
39 J. Fassler and P. Cooper, in *BLAST® Help*, National Center for Biotechnology Information, Bethesda, 2008, pp. 55–63.
40 T. P. Roosild, S. Castronovo, S. Miller, C. Li, T. Rasmussen, W. Bartlett, B. Gunasekera, S. Choe and I. R. Booth, *Structure*, 2009, **17**, 893–903.
41 T. P. Roosild, S. Castronovo, J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway and I. R. Booth, *Proc. Natl. Acad. Sci.*, 2010, **107**, 19784–19789.
42 T. P. Roosild, S. Miller, I. R. Booth and S. Choe, *Cell*, 2002, **109**, 781–791.
43 J. Healy, S. Miller, C. Pliotas, T. Rasmussen, W. Bartlett, S. J. Conway and I. R. Booth, *Proc. Natl. Acad. Sci.*, 2014, **111**, 3895.
44 S. C. Grayer, University of Oxford, 2015.
45 A. Axer, Münster University, 2015.
46 H. Y. Carr and E. M. Purcell, *Phys. Rev.*, 1954, **94**, 630–638.
47 J. Tucker, University of Oxford, 2016.
48 J. Healy, T. Rasmussen, S. Miller, I. R. Booth and S. J. Conway, *Org. Chem. Front.*, 2016, **3**, 439–446.
49 C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989,

244, 182–188.
50 J. A. Hirsch, in *Topics in Stereochemistry*, 1967, vol. 1, pp. 199–222.
51 F. R. Jensen and C. H. Bushweller, in *Advances in Alicyclic Chemistry*, 1971, pp. 139–194.
52 E. L. Eliel and S. H. Wilen, in *Stereochemistry of Organic Compounds*, 1993, pp. 690–699.
53 E. L. Eliel, S. H. Wilen and M. P. Doyle, in *Basic Organic Stereochemistry*, 2001, pp. 439–446.
54 C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi and B. Stockman, *J. Biomol. NMR*, 2001, **21**, 349–359.
55 M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis and T. L. Madden, *Nucleic Acids Res.*, 2008, **36**, 5–9.
56 G. M. Boratyn, A. A. Schäffer, R. Agarwala, S. F. Altschul, D. J. Lipman and T. L. Madden, *Biol. Direct*, 2012, **7**, 12.
57 Y. Han, F. Albericio and G. Barany, *J. Org. Chem.*, 1997, **62**, 4307–4312.
58 A. D. Baldwin and K. L. Kiick, *Bioconjug. Chem.*, 2011, **22**, 1946–1953.
59 A. McNally, N. R. Thomson, S. Reuter and B. W. Wren, *Nat. Rev. Microbiol.*, 2016, **14**, 177–190.
60 S. Ekkerman, University of Aberdeen, 2016.
61 A. P. Hill and R. J. Young, *Drug Discov. Today*, 2010, **15**, 648–655.
62 A. Wünsch, University of Oxford, 2015.
63 M. Wilson, University of Oxford, 2018.
64 F. H. Niesen, H. Berglund and M. Vedadi, *Nat. Protoc.*, 2007, **2**, 2212–2221.
65 M. E. Anderson and A. Meister, *Anal. Biochem.*, 1989, **183**, 16–20.
66 G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176–2179.
67 D. H. Williams and I. Fleming, *Spectroscopic Methods in Organic Chemistry*, McGraw-Hill Education, 6th edn., 2007.
68 T. Sakaue, M. Ohhira, K. Ogata and S. Ohmori, *Arzneimittel-Forschung/Drug Res.*, 1992, **42**, 1482–1486.
69 C. Loukou, P. Changenet-Barret, M.-N. Rager, P. Plaza, M. M. Martin and J.-M. Mallet, *Org. Biomol. Chem.*, 2011, **9**, 2209–2218.
70 D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison III and H. O. Smith, *Nat. Methods*, 2009, **6**, 343–345.
71 D. G. Gibson, H. O. Smith, C. A. Hutchison III, J. C. Venter and C. Merryman, *Nat. Methods*, 2010, **7**, 901–903.
72 The *E. coli* Gene Expression Database (GenExpDB), https://genexpdb.okstate.edu/databases/genexpdb/, (accessed March 2016).
73 K. C. Schultz, L. Supekova, Y. Ryu, J. Xie, R. Perera and P. G. Schultz, *J. Am. Chem. Soc.*, 2006, **128**, 13984–13985.
74 S. Miller, R. M. Douglas, P. Carter and I. R. Booth, *J. Biol. Chem.*, 1997, **272**, 24942–24947.
75 M. R. Green and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 3rd edn., 2001.
76 H. Jeong, V. Barbe, C. H. Lee, D. Vallenet, D. S. Yu, S.-H. Choi, A. Couloux, S.-W. Lee, S. H. Yoon, L. Cattolico, C.-G. Hur, H.-S. Park, B. Ségurens, S. C. Kim, T. K. Oh, R. E. Lenski, F. W. Studier, P. Daegelen and J. F. Kim, *J. Mol. Biol.*, 2009, **394**,

644–652.
77 G. P. Ferguson and I. R. Booth, *J. Bacteriol.*, 1998, **180**, 4314–4318.
78 C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti and M. Sundström, *J. Biomol. NMR*, 2000, **18**, 65–68.
79 J. McFarland, *J. Am. Med. Assoc.*, 1907, **XLIX**, 1176–1178.
80 F. R. Cockerill III, M. A. Wikler, J. Alder, M. N. Dudley, G. M. Eliopoulos, M.-J. Ferraro, D. J. Hardy, D. W. Hecht, J. A. Hindler, J. B. Patel, M. Powell, J. M. Swenson, R. B. Thomson Jr., M. M. Traczewski, J. D. Turnidge, M. P. Weinstein and B. L. Zimmer, *Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Second Informational Supplement*, 2013, vol. 32.
81 J. B. Patel, F. R. Cockerill III, J. Alder, P. A. Bradford, G. M. Eliopoulos, D. J. Hardy, J. A. Hindler, S. G. Jenkins, J. S. Lewis II, L. A. Miller, M. Powell, J. M. Swenson, M. M. Traczewski, J. D. Turnidge, M. P. Weinstein and B. L. Zimmer, *Performance Standards for Antimicrobial Susceptibility Testing; Twenty-Fourth Informational Supplement*, 2014, vol. 34.

Appendices

A. **Protein Characterisation**

B. **ITC Curves**

C. **Protein-Observed ^{19}F NMR Spectra**

D. **Kirby-Bauer Disc Diffusion Assays**

E. **MIC Assay**

F. **Pierce™ 660 nm Protein Assays**

G. **SDS-PAGE Gels**

H. **Chiral HPLC Traces for Intermediates**

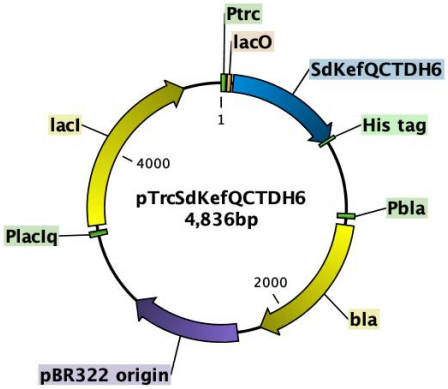
I. **NMR Spectra of Novel Compounds**

A. Protein Characterisation

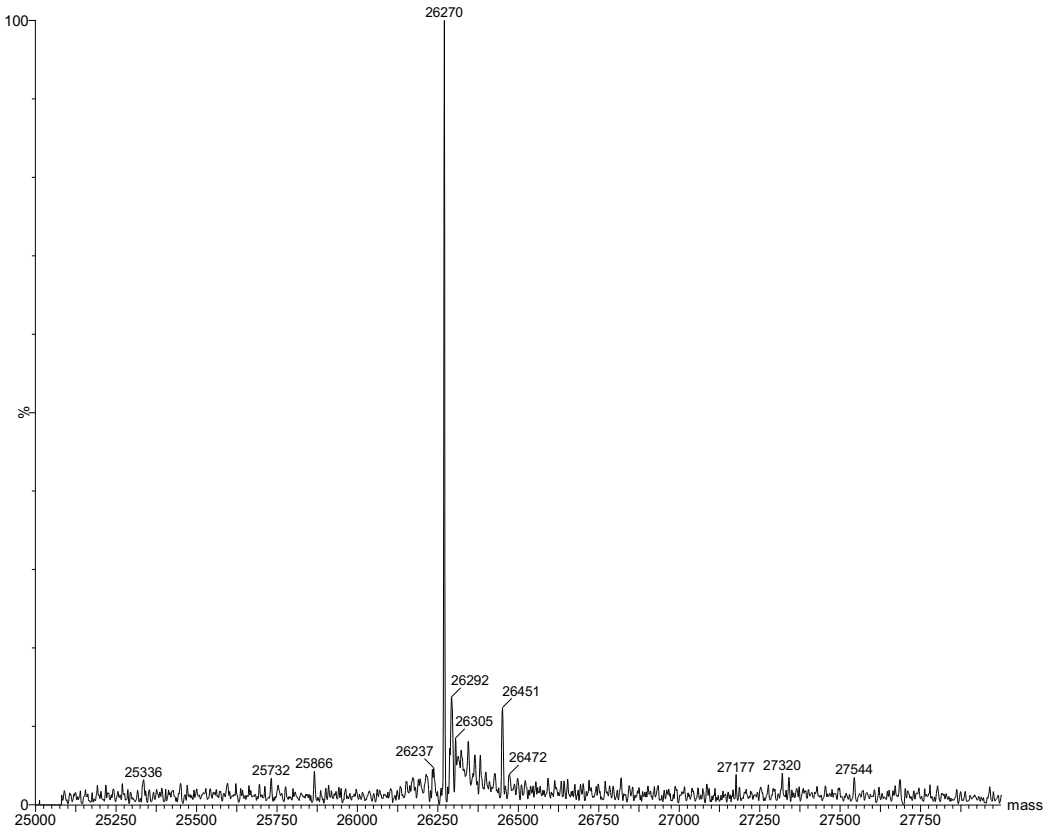
Molecular weights of proteins in Daltons (Da) were predicted using the ProtParam tool (available at: http://web.expasy.org/protparam/).

i. *Sd*KefQCTDH$_6$

Plasmid	pTrcSdKefQCTDH$_6$
Kef gene	*SdkefQCTDH$_6$*
DNA sequence (5' to 3')	ATGGGCCACGAGTTAGAAGTGGATATCGAACCGAAGAAAACCCCAGAAA ATCCGGAATACGATAATGAAATTGAGGCAACTAATAGCGTTATCATTGCC GGTTATGGTCGCTTTGGCCAAGTTGTTGGTCGTCTGCTGTCTGCGCAGGGC TACCACCTGAGCATCTTGGACCATAGCCCGTCTCAAATTGACATGCTGCGC CGTTTCGGTAATAAGGTTTTCTATGGCGATGCGGCGCGTAAAGATCTGCTG GAGGCGGCAGGCGCGAAGGACGCCCAACTGCTGGTGATTGCGATCGACG CCCCTGACAAGGCGCTGGAGATCGTTGAGCTGGCTCATAAACACTACCCG CAATTGAAAATCGTGGCCCGTGCGATTGATCGCCGCCACGCGTACCAGTA CCTGCGTCTGGGCGTGACTTCTTTCAAACGTGAAACCTTCGACTCCGCGGT GAATCTGGGCATTGAAGCGCTGACCCTGCTGGGTAACAGCTCCACCGTGG CAGAGCGCGCTGGCGACCTGTTCTCCCAGCACGATAATGCAAGCCTGCAC GAGTTGGCAGCCCTGTGGGGTGATGACCACACGTATGGCGTCGCGGTCCG CCAGCGTATGGAAGATCTGAAGCAAGTTCTGGCAAAAGATAAGGCAGCAC AGGAACGCCTGAATACGTGTAAAGGTGGTGACTGTCTCGAGCACCACCAC CACCACCACTGATAA
Protein sequence (N- to C-terminus)	MGHELEVDIEPKKTPENPEYDNEIEATNSVIIAGYGRFGQVVGRLLSAQGYHLS ILDHSPSQIDMLRRFGNKVFYGDAARKDLLEAAGAKDAQLLVIAIDAPDKALEI VELAHKHYPQLKIVARAIDRRHAYQYLRLGVTSFKRETFDSAVNLGIEALTLLG NSSTVAERAGDLFSQHDNASLHELAALWGDDHTYGVAVRQRMEDLKQVLA KDKAAQERLNTCKGGDCLEHHHHHH
Molecular weight (Da)	26402.7



Plasmid map of pTrcSdKefQCTDH$_6$. The *SdkefQCTDH$_6$* gene codes for the *Sd*KefQCTDH$_6$ protein. Ptrc is the promoter for the *SdkefQCTDH$_6$* gene. *bla* is the gene for β-lactamase, which causes resistance to Amp. Pbla is the promoter for the gene *bla*. pBR322 origin is the origin of replication, which ensures that the plasmid is replicated when it is transformed into *E. coli*. *lacI* is the gene for LacI, which prevents transcription of the *SdkefQCTDH$_6$* gene by binding to LacO. IPTG binds to LacI, preventing LacI from binding to LacO, promoting transcription of the *SdkefQCTDH$_6$* gene. PlacIq is the promoter for the gene *lacI*.



Mass spectrum of *Sd*KefQCTDH$_6$. Data from Dr Anthony Chan (unpublished). These data are consistent with the expected mass of the *Sd*KefQCTDH$_6$ monomer. Predicted molecular weight of *Sd*KefQCTDH$_6$ monomer without first N-terminal Met = 26271.5 Da.

ii. *Sd*KefQCTDH$_6$(F448*p*tfmF)

Plasmid pTrcSdKefQCTDH$_6$(F448Am)

Kef gene *SdkefQCTDH$_6$(F448Am)*

DNA sequence (5' to 3')

ATGGGCCACGAGTTAGAAGTGGATATCGAACCGAAGAAAACCCCAGAAA
ATCCGGAATACGATAATGAAATTGAGGCAACTAATAGCGTTATCATTGCC
GGTTATGGTCGCTTTGGCCAAGTTGTTGGTCGTCTGCTGTCTGCGCAGGGC
TACCACCTGAGCATCTTGGACCATAGCCCGTCTCAAATTGACATGCTGCGC
CGT**TAG**GGTAATAAGGTTTTCTATGGCGATGCGGCGCGTAAAGATCTGCT
GGAGGCGGCAGGCGCGAAGGACGCCCAACTGCTGGTGATTGCGATCGAC
GCCCCTGACAAGGCGCTGGAGATCGTTGAGCTGGCTCATAAACACTACCC
GCAATTGAAAATCGTGGCCCGTGCGATTGATCGCCGCCACGCGTACCAGT
ACCTGCGTCTGGGCGTGACTTCTTTCAAACGTGAAACCTTCGACTCCGCGG
TGAATCTGGGCATTGAAGCGCTGACCCTGCTGGGTAACAGCTCCACCGTG
GCAGAGCGCGCTGGCGACCTGTTCTCCCAGCACGATAATGCAAGCCTGCA
CGAGTTGGCAGCCCTGTGGGGTGATGACCACACGTATGGCGTCGCGGTCC
GCCAGCGTATGGAAGATCTGAAGCAAGTTCTGGCAAAAGATAAGGCAGC
ACAGGAACGCCTGAATACGTGTAAAGGTGGTGACTGTCTCGAGCACCACC
ACCACCACCACTGATAA

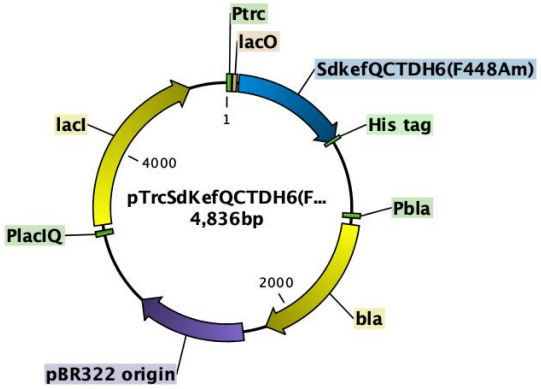
TAG = Amber stop codon

Protein sequence (N- to C-terminus)

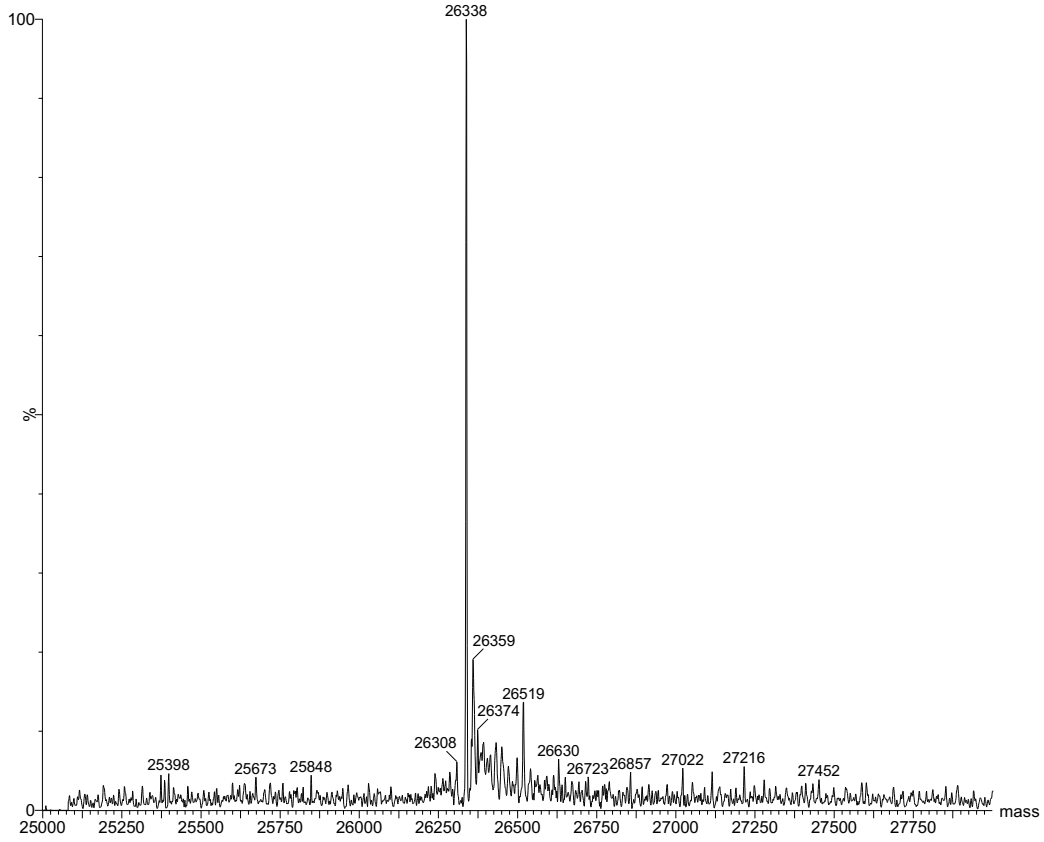
MGHELEVDIEPKKTPENPEYDNEIEATNSVIIAGYGRFGQVVGRLLSAQGYHLS
ILDHSPSQIDMLRR**F**GNKVFYGDAARKDLLEAAGAKDAQLLVIAIDAPDKALEI
VELAHKHYPQLKIVARAIDRRHAYQYLRLGVTSFKRETFDSAVNLGIEALTLLG
NSSTVAERAGDLFSQHDNASLHELAALWGDDHTYGVAVRQRMEDLKQVLA
KDKAAQERLNTCKGGDCLEHHHHHH

F = *p*tfmF

Molecular weight (Da) 26470.7



Plasmid map of pTrcSdKefQCTDH$_6$(F448Am). The *SdkefQCTDH$_6$(F448Am)* gene allows *p*tfmF to be incorporated at position 448 in *Sd*KefQCTDH$_6$. Ptrc is the promoter for the *SdkefQCTDH$_6$(F448Am)* gene. *bla* is the gene for β-lactamase, which causes resistance to Amp. Pbla is the promoter for the gene *bla*. pBR322 origin is the origin of replication, which ensures that the plasmid is replicated when it is transformed into *E. coli*. *lacI* is the gene for LacI, which prevents transcription of the *SdkefQCTDH$_6$(F448Am)* gene by binding to LacO. IPTG binds to LacI, preventing LacI from binding to LacO, promoting transcription of the *SdkefQCTDH$_6$(F448Am)* gene. PlacIq is the promoter for the gene *lacI*.



Mass spectrum of *Sd*KefQCTDH$_6$(F448*p*tfmF). Data from Dr Anthony Chan (unpublished). These data are consistent with the expected mass of the *Sd*KefQCTDH$_6$(F448*p*tfmF) monomer. Predicted molecular weight of *Sd*KefQCTDH$_6$(F448*p*tfmF) monomer without first N-terminal Met = 26339.5 Da.

iii. *Sd*KefH$_6$

Plasmid pTrcSdKefH$_6$

Kef gene *SdkefH$_6$*

DNA sequence (5' to 3')

ATGGTGCTGGAAATGGCTGTTATCTATCTGGCGGCTGCAATCGTGGCAGT
TCCGGTGGCGAAACGTCTGGGTCTGGGTAGCGTCCTGGGCTACTTGCTGG
CAGGTATTCTGATCGGCCCGTTTTTGCTGGGCCTGGTCGGCGAGCAAACC
GACGTGATGCACTTTGCGGAGTTCGGTGTCGTGATGATGTTGTTTCTGGTG
GGCCTGGAACTGCAACCGTCCCGTCTGTGGAAACTGCGTCATAGCATCATC
GGTCTGGGTGGTTTGCAGGTTCTGCTGAGCGCAGGCCTGATTTTCGCTATC
GGCTTCTACCTGTTTGAACTGCAATGGCAAACGGCTCTGGCCATCGGTTTG
ATGCTGTCGCTGAGCTCTACCGCAATTGTGCTGCAAACTCTGACGGAAAAA
GGCTGGATTAAGCAGGACGCTGGTCAGAACGCGTTTTCGGTCCTGCTGTT
CCAGGACATCGCTGTCATTCCGATCCTGGCGCTGATCCCGTTGCTGGCTTT
TACGGACATCAATGCGGCCCAAGCAGACCACGGTAATATGATCGCGCATC
TGCCTAGCTATTTGCAGGTCATGATCAGCCTGGGTGTGATTGGTACCATTA
TCCTGGCAGGTCGTTACGTGTCTGCACCGCTGTTTCGTTATATCGCGGAAA
CCCATCTGCGTGAAATTTTCACCATTTCGGCGCTGTTCCTGGTCGTCGCAAT
TGCAGTCATCATGCAGAAGATCGGCCTGTCGCCGGCCTTGGGCACGTTTCT
GGCAGGTGTCGTGCTGGCGGAGAGCGAATTTCGTCATGAACTGGAGGTT
GATATTGAGCCGTTTAAGGGTCTGCTGCTGGGTCTGTTCTTCATCACCGTG
GGTGCGTCGATTGACTTTCCGCTGCTGTTGGAGCAGTCTCTGTTGGTTCTG
GGTTTGGTCCTGGGTTTGCTGTTGGTCAAAGCCCTGGTTCTGTTTGTCCTG
GCGCAAGGCTTTAAGATGCAAGGTGAGCAAAAGCTGCTGTTCACTCTGGC
CCTGGCACAGGGCGGTGAATTTGCGTTCGTCCTGCTGAGCTTGACCCAAA
ATCTGGAAATCCTGACGCTGGAGCAAACGAAGGTCGTCACCCTGGTCGTG
ACCATCTCTATGCTGTGTGCGCCACTGCTGCTGATGTTCTACGAGCGTCTG
CTGGTTAAGAAAACCCCAGAAAATCCGGAATACGATAATGAAATTGAGGC
AACTAATAGCGTTATCATTGCCGGTTATGGTCGCTTTGGCCAAGTTGTTGG
TCGTCTGCTGTCTGCGCAGGGCTACCACCTGAGCATCTTGGACCATAGCCC
GTCTCAAATTGACATGCTGCGCCGTTTCGGTAATAAGGTTTTCTATGGCGA
TGCGGCGCGTAAAGATCTGCTGGAGGCGGCAGGCGCGAAGGACGCCCAA
CTGCTGGTGATTGCGATCGACGCCCCTGACAAGGCGCTGGAGATCGTTGA
GCTGGCTCATAAACACTACCCGCAATTGAAAATCGTGGCCCGTGCGATTG
ATCGCCGCCACGCGTACCAGTACCTGCGTCTGGGCGTGACTTCTTTCAAAC
GTGAAACCTTCGACTCCGCGGTGAATCTGGGCATTGAAGCGCTGACCCTG
CTGGGTAACAGCTCCACCGTGGCAGAGCGCGCTGGCGACCTGTTCTCCCA
GCACGATAATGCAAGCCTGCACGAGTTGGCAGCCCTGTGGGGTGATGACC
ACACGTATGGCGTCGCGGTCCGCCAGCGTATGGAAGATCTGAAGCAAGTT
CTGGCAAAAGATAAGGCAGCACAGGAACGCCTGAATACGTGTAAAGGTG
GTGACTGTCTCGAGCACCACCACCACCACCACTGATAA

Protein sequence

MVLEMAVIYLAAAIVAVPVAKRLGLGSVLGYLLAGILIGPFLLGLVGEQTDVM
HFAEFGVVMMLFLVGLELQPSRLWKLRHSIIGLGGLQVLLSAGLIFAIGFYLFEL
QWQTALAIGLMLSLSSTAIVLQTLTEKGWIKQDAGQNAFSVLLFQDIAVIPILA

(N- to C-terminus)	LIPLLAFTDINAAQADHGNMIAHLPSYLQVMISLGVIGTIILAGRYVSAPLFRYIA ETHLREIFTISALFLVVAIAVIMQKIGLSPALGTFLAGVVLAESEFRHELEVDIEPF KGLLLGLFFITVGASIDFPLLLEQSLLVLGLVLGLLLVKALVLFVLAQGFKMQGE QKLLFTLALAQGGEFAFVLLSLTQNLEILTLEQTKVVTLVVTISMLCAPLLLMFYE RLLVKKTPENPEYDNEIEATNSVIIAGYGRFGQVVGRLLSAQGYHLSILDHSPS QIDMLRRFGNKVFYGDAARKDLLEAAGAKDAQLLVIAIDAPDKALEIVELAHK HYPQLKIVARAIDRRHAYQYLRLGVTSFKRETFDSAVNLGIEALTLLGNSSTVAE RAGDLFSQHDNASLHELAALWGDDHTYGVAVRQRMEDLKQVLAKDKAAQE RLNTCKGGDCLEHHHHHH
Molecular weight (Da)	67347.3
Plasmid map	See Figure 6.1D, Figure 6.2B, and Figure 6.3B.

iv. *Yp*KefGBP

Plasmid	pTrcYpKefGBP
Kef gene	*YpkefB*
DNA sequence (5’ to 3’)	ATGGAGGGCACAGGGCTACTTACCGCTGTTCTAGTGTTTTTATTCGCTGCT GTCGTTGCGGTCCCTATCGCCCAACGCTTAGGTATTGGTGCGGTACTGGGT TACCTGATCGCGGGTATTGCCATTGGCCCGTGGGGGCTAGGATTTATTCGC GATGTTGACGAAATTCTACATTTTTCTGAGCTAGGCGTCGTCTTCCTGATGT TTATCATTGGGTTGGAGTTGAATCCCGCAAAACTTTGGCAATTAAGGCGAT CTATTTTTGGTGTCGGTGCTGGTCAAGTCGTGATAACGGCCGCGGTACTG GGCGCATTGCTTTATTTTACCCAGTTCGCCTGGCAGGCAGCGGTTATCGGT GGCGTAGGTTTGGCGATGTCGTCTACCGCGATGGCATTGCAACTGATGCG GGAAAAGGGGATGAACCGCAATGAAGGCGGCCAGCTCGGTTTTTCCGTTC TACTGTTCCAGGATATGGCGGTGATCCCCGCGTTAGCCTTGATTCCTATTCT GGCGGGTAACGAAGGGGGAGCCAATGACTGGGTCAAAATTGGCCTCAAA ATTGCGGCTTTTGCCGGTATGTTGATTGGGGGCCGTTATTTGTTGCGCCCG TTGTTCCGCTACATTGTCGCCTCTGGGGTACGTGAGGTTTTTACCGCTGCG GCATTGCTGGTGGTGCTGGGGTCCGCGCTGTTTATGGATGCCCTGGGGCT ATCGATGGCATTGGGGACATTTATCGCCGGGATATTACTGGCTGAAAGTG AGTTTCAGCACGAATTGGAAATCGCTATCGAGCCGTTTAAAGGGCTCTTGC TCGGTTTGTTTTTTATTTCTGTCGGCATGGCATTGGACCTCGGCGTGCTGTT TACCCATTTACTGGATGTACTGCTGGGCGTCTTGGCACTGGTCTTTATCAA AAGTGCCATCTTGTATGGTTTGGCGCGGGTGTTTGGCTTGCGGCGTTCGG TCCGTTTGCAATTTGCCGGTGTGCTCAGCCAAGGCGGCGAATTTGCTTTCG TCCTGTTTTCCGCGGCGTTCTCTCAGCGTGTGTTGAATGCAGAACAACTGG CATTGCTATTGGTGGTCGTCACCTTATCGATGATGACTACGCCGTTGCTGA TGCAGGTTATTGACCGTATTTTGGTTCGTCGCTACAACGCGCAGGAAGAG AGTGATGAGAAACCCTTTGTTGAAGATAATGACCCACAGGTGATTATTGTC GGCTTCGGGCGATTTGGTCAGGTGATTGGCCGTTTATTGATGGCGAACAA

	AATGCGCATTACCGTACTGGAACGTGATGTCAGCGCGGTCAGTATGATGC GCAAATATGGCTACAAAGTTTACTATGGTGACGCGACCGAGTTGGAACTG CTACGAGCCGCTGGCGCAGAAAAAGCCAAAGCCATTGTTATCACCTGCAA TGAGCCGGAAGACACCATGGCGCTGGTGCATTTATGCCAGCAACACTTCC CTAATCTGCACATTCTTGCTCGCGCCAGAGGGCGTGTCGAAGCCCATGAAC TGCTCCAAAACGGGGTCAAAGACTTCACGCGTGAAACCTTCTCCAGTGCG CTGGAACTGGGGCGTAAGACATTATTGGAACTGGGCATGCATCCACATCA GGCCTACCGGGCGCAACAACATTTTCGTCGCTTGGACATGCGAATGCTAC GTGAGCTGATGCCACAACATCATGGCGATGTGGCACAAATTTCCAGAATT AAAGAAGCTCGCCGTGAGTTGGAAGATATTTTTCAACGAGAAATGCTACA TGAAAGCCGGCAGTTAGATGGGTGGGATGAATACGAATAA
Protein sequence (N- to C-terminus)	MEGTGLLTAVLVFLFAAVVAVPIAQRLGIGAVLGYLIAGIAIGPWGLGFIRDVD EILHFSELGVVFLMFIIGLELNPAKLWQLRRSIFGVGAGQVVITAAVLGALLYFT QFAWQAAVIGGVGLAMSSTAMALQLMREKGMNRNEGGQLGFSVLLFQD MAVIPALALIPILAGNEGGANDWVKIGLKIAAFAGMLIGGRYLLRPLFRYIVAS GVREVFTAAALLVVLGSALFMDALGLSMALGTFIAGILLAESEFQHELEIAIEPF KGLLLGLFFISVGMALDLGVLFTHLLDVLLGVLALVFIKSAILYGLARVFGLRRSV RLQFAGVLSQGGEFAFVLFSAAFSQRVLNAEQLALLLVVVTLSMMTTPLLMQ VIDRILVRRYNAQEESDEKPFVEDNDPQVIIVGFGRFGQVIGRLLMANKMRIT VLERDVSAVSMMRKYGYKVYYGDATELELLRAAGAEKAKAIVITCNEPEDTM ALVHLCQQHFPNLHILARARGRVEAHELLQNGVKDFTRETFSSALELGRKTLLE LGMHPHQAYRAQQHFRRLDMRMLRELMPQHHGDVAQISRIKEARRELEDIF QREMLHESRQLDGWDEYE
Molecular weight (Da)	66328.2
Plasmid map	See Figure 3.6A, Figure 3.15B, Figure 6.1A, Figure 6.2D, and Figure 6.4B.

v. *Bp*Kef1H$_6$

Plasmid	pTrcBpKef1H$_6$
Kef gene	*Bpkef1H$_6$*
DNA sequence (5' to 3')	ATGGTTAGTCCGCTGGAAATGACCTTACTGCTGCTGCTGGCAAGCGTTGTT GGTGTTGTTGTTTTTCGTAGCCTGAATCTGCCTCCGATGCTGGGTTATCTG ACCGTTGGTATTCTGGTTGGTCCGCATGCACTGGGTTTTGCAGCAGATCTG GAACGTGCAGAACATCTGGCAGAATTTGGTGTGGTTTTTCTGATGTTTAGC ATCGGTCTGGAATTTAGCCTGAGCAAACTGCGTGCAATGCGTCGTCTGGTT TTTGGTCTGGGTCTGGCACAGGTTCTGGGCACCATTGCAATTGCCCTGGTT CTGGGTGCACTGTTTGAACGTTGGATGCATATTACCTGGGAAGGTAGCGT TGCCCTGGGTGGTGCACTGGCAATGAGCAGCACCGCAATTGTTAGCAAAA TGCTGGCAGAACGTCTGGAAATTGAAACCGAACATGGTCGTAACATTTTT

GGTGTTCTGCTGTTTCAGGATCTGGCAGTTGTTCCGCTGCTGATTGTTATT
GCAGCATTTGGTGGTGAAAGCAGCAAAGATCTGGCCATGAGCCTGGGCTT
TGCCGCAGTTAAAATTGTTGTTGCACTGGCCCTGCTGCTGATCGTTGGTCA
GCGTTTTATGACCCGTTGGCTGAATGTTGTTGCCCGTCGTCGTAGCCAAGA
ACTGTTTGTTCTGAATCTGCTGCTGGTTACCCTGGGTGCAGCATTTATTACC
GATCGTTTTGGTCTGAGCCTGGCGCTGGGTGCATTTATTGCAGGTATGCTG
ATTAGCGAAACCCCGTATCGTCATCAGGTTGAAGAAGATATTAAACCGTTT
CGTGATGTGCTGCTGGGTCTGTTTTTTGTTACCACCGGTATGCTGCTGGAT
CCGAGCGTTATTTGGCGTCATCCTCTGCTGGTACTGGCATTTTTTGTTGGTC
AGGTTCTGCTGAAAAGCGTTATGATTACAGGCCTGGCACGTGCGTTTGGT
GCAACACCGGGTGTGGCAATGCGTACCGGACTGGGCCTGGCGCAGGCAG
GCGAATTTGGTTTTGTACTGCTGAATCTGATTCTGGATCGTCATCTGGTGG
ATGCAACCCTGCTGCAGAGCATTCTGGCAAGCATGCTGCTGAGTATGCTG
GCTGCACCGTTTCTGATTCAGAATGCAGATCGTATTGTTATGCGTCTGAGC
GCAGCAGAATGGATGCAGCAGAGCCTGCAGATGACCCGTATTGCAACCCA
GAGCATTAAACAGAAAGGCCATGTGATTATTTGCGGTTATGGTCGTGCAG
GTCAGAATCTGGCACGTATGCTGGAACAAGAGGGTATTGCATATGTTGCG
CTGGATCTGGATCCTGATCGTGTTACCGCAGCAGCCGCAGCCGGTGAAAG
CGTTGTGTTTGGTGATGCAGCACGTCGTGAAAGCCTGCTGGCAGCCGGTA
TTCATCGTGCCGCAGCAGTTGCAGTTACCTATGCAAATACCCCGAGCGCAC
TGCGTGTTCTGCATAATGTGCATGAACTGGAACCGACCCTGCCTGCAATTG
TTCGTACCGTTGATGATGCCGATCTGGAAAAACTGCTGGCTGCGGGTGCA
ACCGAAGTTATTCCGGAAATTGTTGAAGGTAGCCTGATGCTGGCGAGCCA
TACCCTGGTGCTGATGGGTGTTCCGATGCGTCGTGTTGTTCGTCGTGTGGA
AGAAATGCGTGATGCACGTTATAGCCTGCTGCGTGGCTATTTTCATGGTGC
CGATGATGCAGATGATGATGATCATGAACAGGTTCGTCTGCAGAGCGTTC
CGGTTGATGCCCGTGCAGATGCAGTTGGTCGTAGTCTGAGCGAACTGGGA
CTGTTTGCGCTGGGACTGGAAGTTACCGCAATTCGTCGCCACGGTATTCGT
GGTGTTGAACCTGATCCGGAAACAAAACTGCGTGCCAGCGATATTGTTGT
TCTGCGTGGTCTGCCGGAAGCCCTGGCCCAGGCTGAAGAACGCCTGAGCC
GTCATCGTCGTGAAGGTGGTAGCGCAGCCGTTGCACTCGAGCACCACCAC
CACCACCACTGATAA

Protein sequence (N- to C-terminus)

MVSPLEMTLLLLLASVVGVVVFRSLNLPPMLGYLTVGILVGPHALGFAADLER
AEHLAEFGVVFLMFSIGLEFSLSKLRAMRRLVFGLGLAQVLGTIAIALVLGALFE
RWMHITWEGSVALGGALAMSSTAIVSKMLAERLEIETEHGRNIFGVLLFQDL
AVVPLLIVIAAFGGESSKDLAMSLGFAAVKIVVALALLLIVGQRFMTRWLNVV
ARRRSQELFVLNLLLVTLGAAFITDRFGLSLALGAFIAGMLISETPYRHQVEEDIK
PFRDVLLGLFFVTTGMLLDPSVIWRHPLLVLAFFVGQVLLKSVMITGLARAFG
ATPGVAMRTGLGLAQAGEFGFVLLNLILDRHLVDATLLQSILASMLLSMLAAP
FLIQNADRIVMRLSAAEWMQQSLQMTRIATQSIKQKGHVIICGYGRAGQNL
ARMLEQEGIAYVALDLDPDRVTAAAAAGESVVFGDAARRESLLAAGIHRAAA
VAVTYANTPSALRVLHNVHELEPTLPAIVRTVDDADLEKLLAAGATEVIPEIVE
GSLMLASHTLVLMGVPMRRVVRRVEEMRDARYSLLRGYFHGADDADDDDH
EQVRLQSVPVDARADAVGRSLSELGLFALGLEVTAIRRHGIRGVEPDPETKLRA
SDIVVLRGLPEALAQAEERLSRHRREGGSAAVALEHHHHHH

Molecular weight (Da)	73376.1
Plasmid map	See Figure 3.6B, Figure 6.1B, and Figure 6.3C.

vi. *Bp*Kef2H$_6$

Plasmid	pTrcBpKef2H$_6$
Kef gene	*Bpkef2H$_6$*
DNA sequence (5' to 3')	ATGCCGCATGATGTTAGCCTGATTGCACTGCTGGCAGCAGGTTTTGGTCTG GCAATGATTTTTGGTTATCTGGCAAGCCTGCTGAAAATGCCTCCGCTGGTT GGCTATCTGCTGGCTGGTATTGTTATGGGTCCGGGTACACCGGGTTTTGTT GGTGATCTGGCCCTGGCACAGCAGCTGGCAGAAATTGGTGTTATGCTGCT GATGTTTGGTGTTGGTCTGCATTTTAGCCTGGGTGATCTGCTGAGCGTTCG TAAAATTGCCCTGCCTGGTGCAATTGTTCAGATTGCAGCAGCAACCGCACT GGGTGCAGGTCTGGCACTGTGGTGGGGTTGGAGCGTTGGTGGTGCACTG GTTTTTGGCCTGAGCCTGAGCGTTGCAAGCACCGTTGTTCTGCTGCGTGCA CTGGAAGGTCGTGGTCTGATTGAAAGCGTTAATGGTCGTATTGCAGTTGG TTGGCTGGTTGTTGAAGATCTGGTTATGGTTCTGGTGCTGGTACTGCTGCC TCCTCTGGCAGCCCTGCTGGGTGGTGGTGGCGCACCGGCACATGGTGATG CAGCAGTTCATGCCGGTGAAGCCGGTGGTAGCCTGTGGGCAGCGCTGGG TGTTACACTGCTGAAAGTTGCAGCATTTGTTGCACTGATGCTGGTTGTGGG TAAACGTGTTTTTCCGCGTATTCTGTGGCTGGTCGCACGTACCGGTAGCCG TGAACTGTTTACCCTGTGTATGATTGCAGCCGCAGTTGGTGTTGCATTTGG TGCAGCAAAACTGTTTGATGTGAGCTTTGCCCTGGGTGCATTTTTTGCAGG TATGATGATGCGTGAAAGCGAATTTAGCCGTCGTGCAGCAGATGAAACCC TGCCGCTGCGTGATGCATTTAGCGTTCTGTTTTTTATCAGCGTGGGTATGC TGTTTGATCCGCGTGTGCTGATTGATGAACCTCTGCATGTTCTGGAAGTTG CCGCAATTGTGGTGCTGGGTAAAACCCTGGCAGCCGTTGCCCTGGTTCTG GCATTTCGTTATCCGCTGAATACCGCACTGACCGTTGGTGTTTCACTGGCA CAGATTGGTGAATTTAGCTTTATTCTGGCCAGCCTGGGTCGCGGTCTGGGT CTGCTGAGTGCAGAAGGTCAGAGCCTGATTCTGGCAGTTGCGCTGCTGAG CATTGCACTGAATACCCTGCTGTTTGCAGCGGTTGATCCGGTTCTGGCCTG GATTCGTAAACGTAGCGCATTTGCACGTCGTCTGGAAAGCCGTGATGATC CACTGGCAGCACTGCCGATGAGCACACCGCAGGCACATCTGACAGGTCAG GTTGTTATTGTGGGTTATGGTCGTGTTGGCACCCGTATTGCACATGCACTG GATGCACGTGGTATTGCATATGTTGTTGTTGAACAGAATCGTGAAACCGT GGAAAAACTGCGTGCCGATGGTGTTGCCGCAGTTAGCGGTGATGCCGTTG AACCGATTGTTCTGGTTCAGGCACATATTGCACGTGCCGGTATGCTGGTAG TTACCCTGCCGGATGTTTTTGATGTTCGTCAGATTGTTGATATTAGCCGTAC CCTGAATCCGGCAATTGAAATTGCACTGTGTACCAACAGTGATGAAGAAG CTGCTCTGCTGGCAAATGAAGGTATGGGTGAAGTTTTTGTTAGCGAAACC GAACTGGCCCATGGTATGACCGAACATGTGCTGGCACGTATGGGTGCGG

	GTGAAGCACATGCCCGTCGTAGCGCAGCACATCTCGAGCACCACCACCAC CACCACTGATAA
Protein sequence (N- to C-terminus)	MPHDVSLIALLAAGFGLAMIFGYLASLLKMPPLVGYLLAGIVMGPGTPGFVG DLALAQQLAEIGVMLLMFGVGLHFSLGDLLSVRKIALPGAIVQIAAATALGAG LALWWGWSVGGALVFGLSLSVASTVVLLRALEGRGLIESVNGRIAVGWLVVE DLVMVLVLVLLPPLAALLGGGGAPAHGDAAVHAGEAGGSLWAALGVTLLKV AAFVALMLVVGKRVFPRILWLVARTGSRELFTLCMIAAAVGVAFGAAKLFDVS FALGAFFAGMMMRESEFSRRAADETLPLRDAFSVLFFISVGMLFDPRVLIDEP LHVLEVAAIVVLGKTLAAVALVLAFRYPLNTALTVGVSLAQIGEFSFILASLGRGL GLLSAEGQSLILAVALLSIALNTLLFAAVDPVLAWIRKRSAFARRLESRDDPLAA LPMSTPQAHLTGQVVIVGYGRVGTRIAHALDARGIAYVVVEQNRETVEKLRA DGVAAVSGDAVEPIVLVQAHIARAGMLVVTLPDVFDVRQIVDISRTLNPAIEIA LCTNSDEEAALLANEGMGEVFVSETELAHGMTEHVLARMGAGEAHARRSAA HLEHHHHHH
Molecular weight (Da)	61877.0
Plasmid map	See Figure 3.6C, Figure 6.1C, and Figure 6.2C.

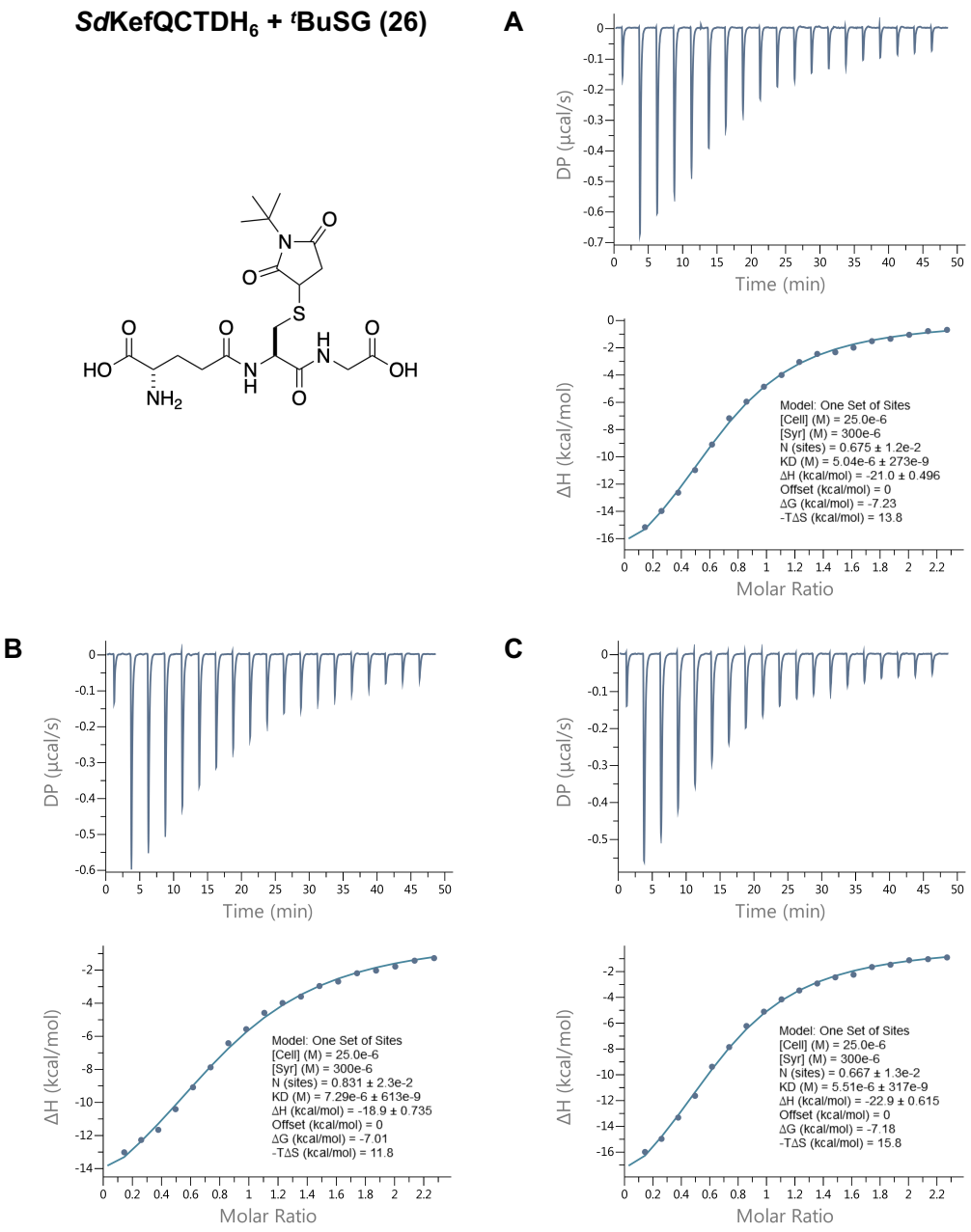
vii. *Yp*KefG(BH_6)P

Plasmid	pTrcYpKefG(BH_6)P
Kef gene	*YpkefBH*$_6$
DNA sequence (5' to 3')	ATGGAGGGCACAGGGCTACTTACCGCTGTTCTAGTGTTTTTATTCGCTGCT GTCGTTGCGGTCCCTATCGCCCAACGCTTAGGTATTGGTGCGGTACTGGGT TACCTGATCGCGGGTATTGCCATTGGCCCGTGGGGGCTAGGATTTATTCGC GATGTTGACGAAATTCTACATTTTTCTGAGCTAGGCGTCGTCTTCCTGATGT TTATCATTGGGTTGGAGTTGAATCCCGCAAAACTTTGGCAATTAAGGCGAT CTATTTTTGGTGTCGGTGCTGGTCAAGTCGTGATAACGGCCGCGGTACTG GGCGCATTGCTTTATTTTACCCAGTTCGCCTGGCAGGCAGCGGTTATCGGT GGCGTAGGTTTGGCGATGTCGTCTACCGCGATGGCATTGCAACTGATGCG GGAAAAGGGGATGAACCGCAATGAAGGCGGCCAGCTCGGTTTTTCCGTTC TACTGTTCCAGGATATGGCGGTGATCCCCGCGTTAGCCTTGATTCCTATTCT GGCGGGTAACGAAGGGGGAGCCAATGACTGGGTCAAAATTGGCCTCAAA ATTGCGGCTTTTGCCGGTATGTTGATTGGGGGCCGTTATTTGTTGCGCCCG TTGTTCCGCTACATTGTCGCCTCTGGGGTACGTGAGGTTTTTACCGCTGCG GCATTGCTGGTGGTGCTGGGGTCCGCGCTGTTTATGGATGCCCTGGGGCT ATCGATGGCATTGGGGACATTTATCGCCGGGATATTACTGGCTGAAAGTG AGTTTCAGCACGAATTGGAAATCGCTATCGAGCCGTTTAAAGGGCTCTTGC TCGGTTTGTTTTTTATTTCTGTCGGCATGGCATTGGACCTCGGCGTGCTGTT TACCCATTTACTGGATGTACTGCTGGGCGTCTTGGCACTGGTCTTTATCAA AAGTGCCATCTTGTATGGTTTGGCGCGGGTGTTTGGCTTGCGGCGTTCGG TCCGTTTGCAATTTGCCGGTGTGCTCAGCCAAGGCGGCGAATTTGCTTTCG

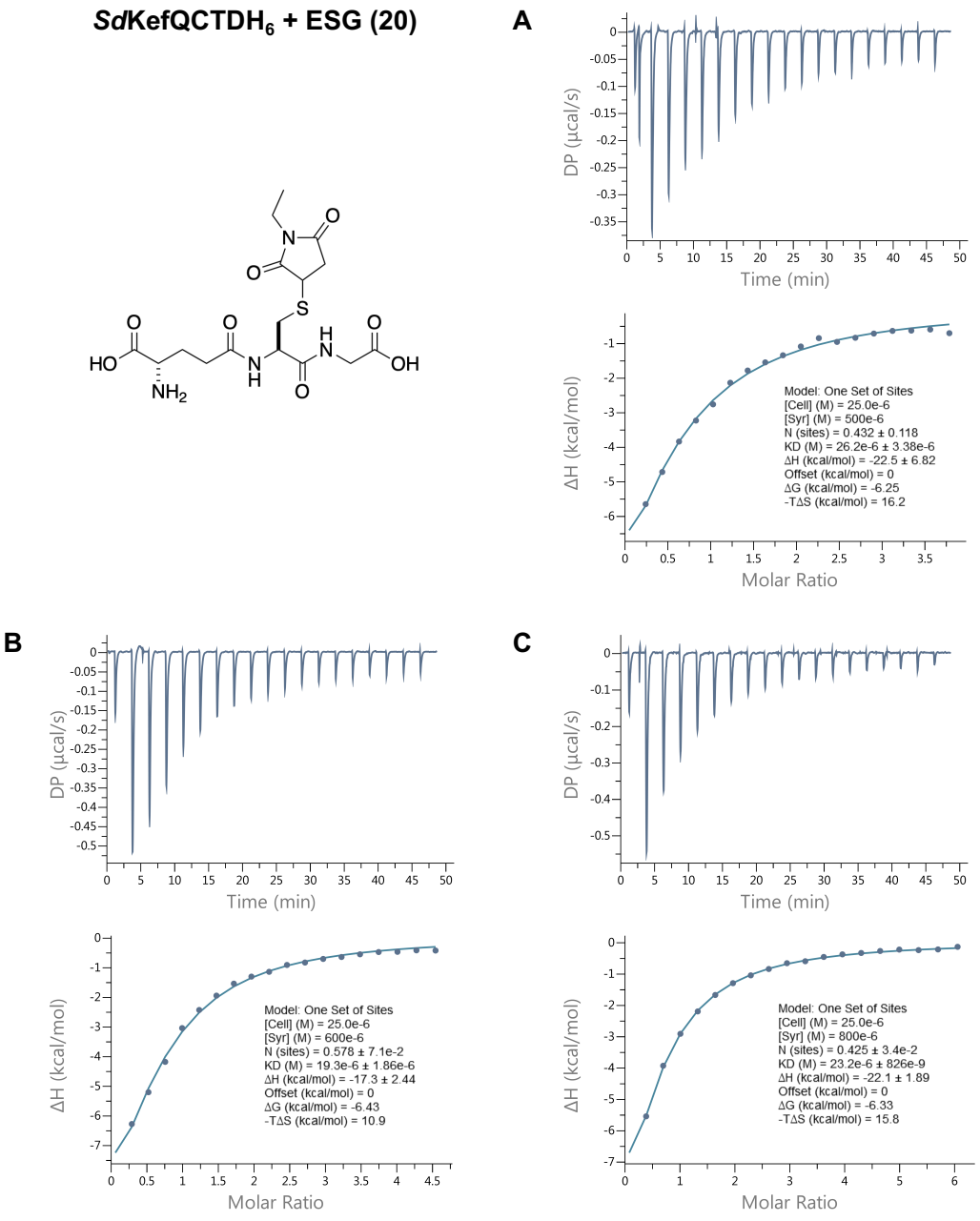
	TCCTGTTTTCCGCGGCGTTCTCTCAGCGTGTGTTGAATGCAGAACAACTGG CATTGCTATTGGTGGTCGTCACCTTATCGATGATGACTACGCCGTTGCTGA TGCAGGTTATTGACCGTATTTTGGTTCGTCGCTACAACGCGCAGGAAGAG AGTGATGAGAAACCCTTTGTTGAAGATAATGACCCACAGGTGATTATTGTC GGCTTCGGGCGATTTGGTCAGGTGATTGGCCGTTTATTGATGGCGAACAA AATGCGCATTACCGTACTGGAACGTGATGTCAGCGCGGTCAGTATGATGC GCAAATATGGCTACAAAGTTTACTATGGTGACGCGACCGAGTTGGAACTG CTACGAGCCGCTGGCGCAGAAAAAGCCAAAGCCATTGTTATCACCTGCAA TGAGCCGGAAGACACCATGGCGCTGGTGCATTTATGCCAGCAACACTTCC CTAATCTGCACATTCTTGCTCGCGCCAGAGGGCGTGTCGAAGCCCATGAAC TGCTCCAAAACGGGGTCAAAGACTTCACGCGTGAAACCTTCTCCAGTGCG CTGGAACTGGGGCGTAAGACATTATTGGAACTGGGCATGCATCCACATCA GGCCTACCGGGCGCAACAACATTTTCGTCGCTTGGACATGCGAATGCTAC GTGAGCTGATGCCACAACATCATGGCGATGTGGCACAAATTTCCAGAATT AAAGAAGCTCGCCGTGAGTTGGAAGATATTTTTCAACGAGAAATGCTACA TGAAAGCCGGCAGTTAGATGGGTGGGATGAATACGAACATCATCATCATC ATCATTAA
Protein sequence (N- to C-terminus)	MEGTGLLTAVLVFLFAAVVAVPIAQRLGIGAVLGYLIAGIAIGPWGLGFIRDVD EILHFSELGVVFLMFIIGLELNPAKLWQLRRSIFGVGAGQVVITAAVLGALLYFT QFAWQAAVIGGVGLAMSSTAMALQLMREKGMNRNEGGQLGFSVLLFQD MAVIPALALIPILAGNEGGANDWVKIGLKIAAFAGMLIGGRYLLRPLFRYIVAS GVREVFTAAALLVVLGSALFMDALGLSMALGTFIAGILLAESEFQHELEIAIEPF KGLLLGLFFISVGMALDLGVLFTHLLDVLLGVLALVFIKSAILYGLARVFGLRRSV RLQFAGVLSQGGEFAFVLFSAAFSQRVLNAEQLALLLVVVTLSMMTTPLLMQ VIDRILVRRYNAQEESDEKPFVEDNDPQVIIVGFGRFGQVIGRLLMANKMRIT VLERDVSAVSMMRKYGYKVYYGDATELELLRAAGAEKAKAIVITCNEPEDTM ALVHLCQQHFPNLHILARARGRVEAHELLQNGVKDFTRETFSSALELGRKTLLE LGMHPHQAYRAQQHFRRLDMRMLRELMPQHHGDVAQISRIKEARRELEDIF QREMLHESRQLDGWDEYEHHHHHH
Molecular weight (Da)	67151.0
Plasmid map	See Figure 3.15A and Figure 6.4A.

B. ITC Curves

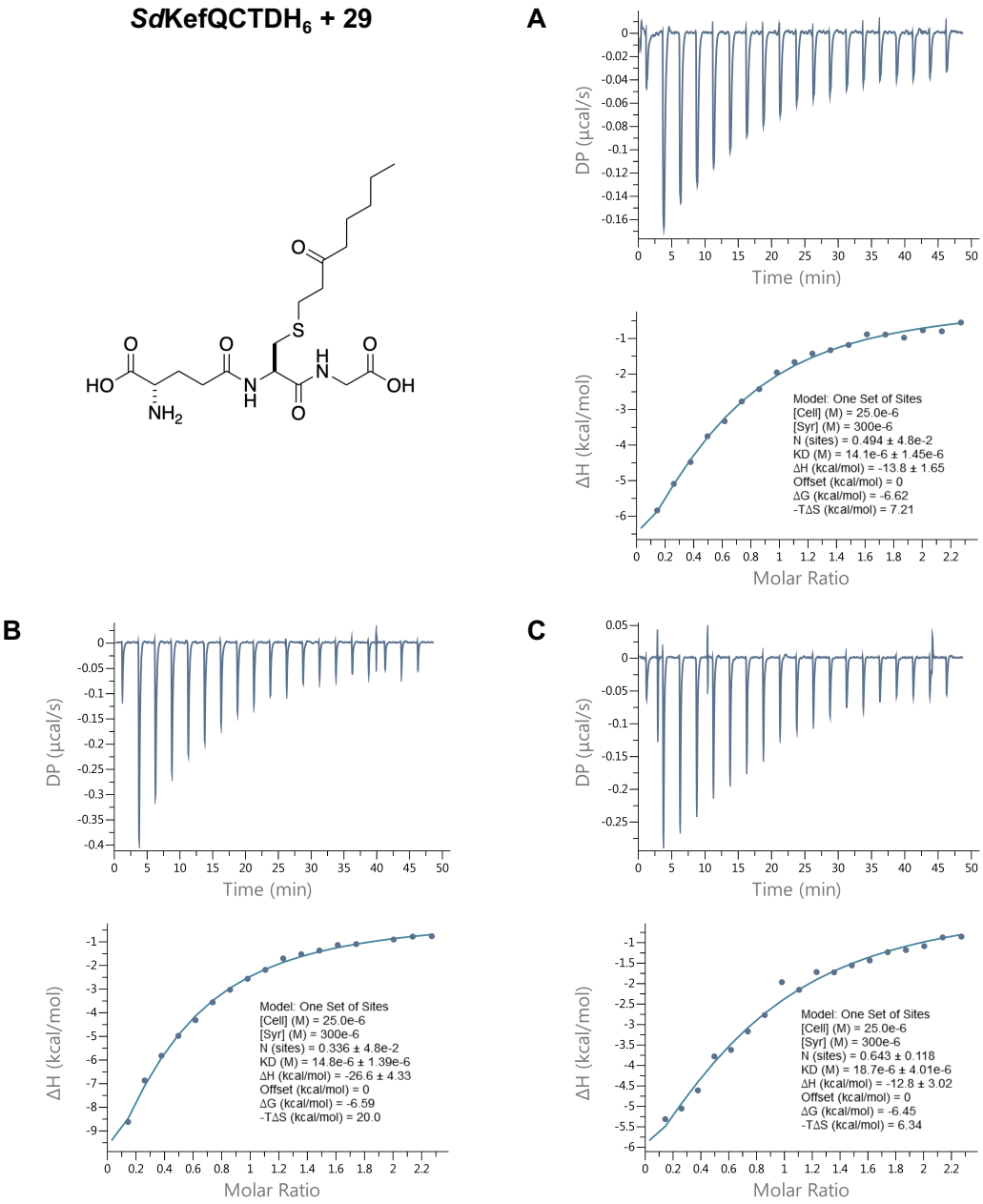
Supplementary data for Table 2.1: Summary of ITC data for *Sd*KefQCTDH$_6$ and SdKefQCTDH$_6$(F448*p*tfmF).



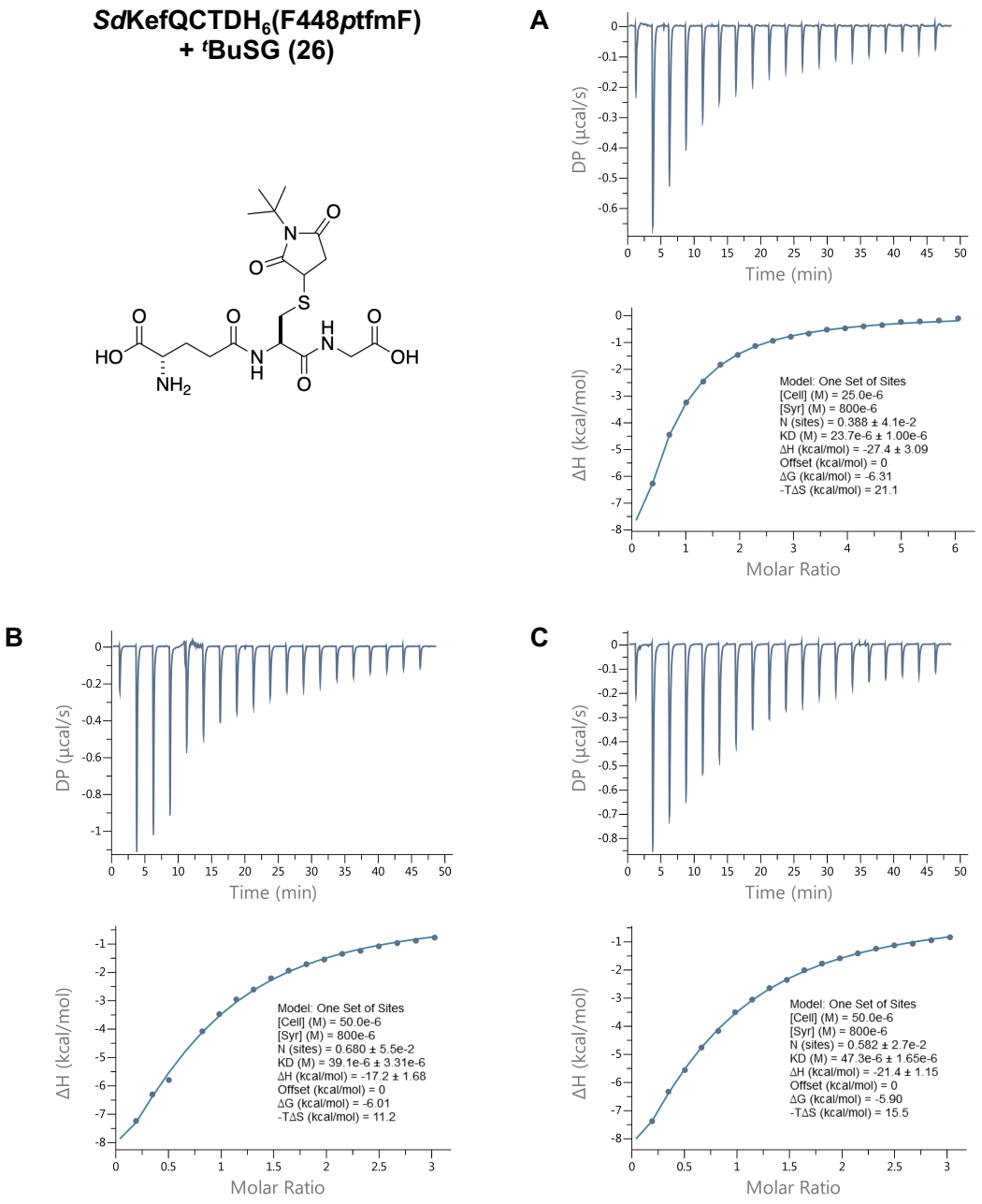
ITC Curves for *Sd*KefQCTDH$_6$ + ^{t}BuSG (26). Cell = *Sd*KefQCTDH$_6$. Syringe (syr) = ^{t}BuSG (**26**). A)–C) Cell concentration = 25 µM, syr concentration = 300 µM.



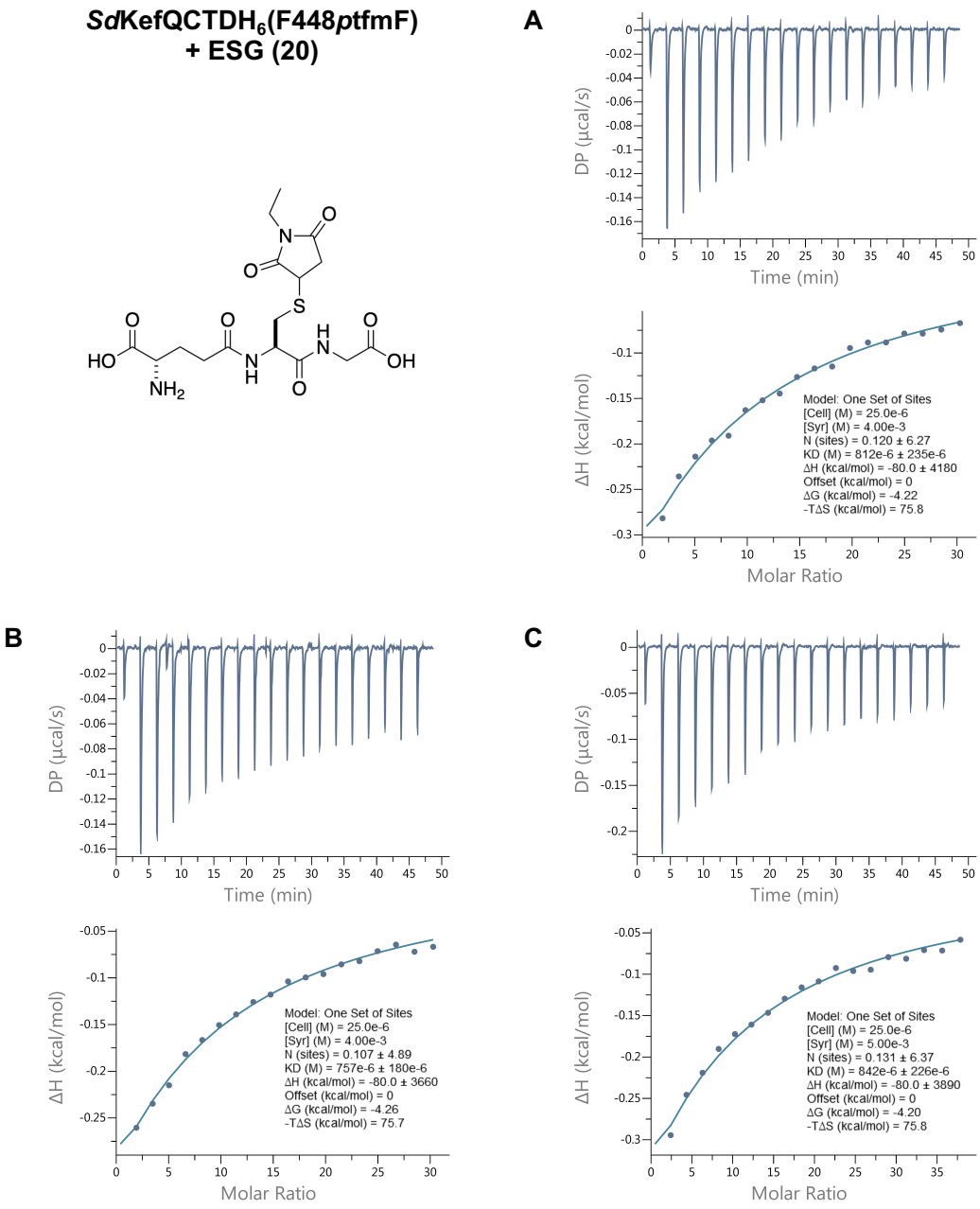
ITC Curves for *Sd*KefQCTDH$_6$ + ESG (20). Cell = *Sd*KefQCTDH$_6$. Syr = ESG (**20**). A) Cell concentration = 25 µM, syr concentration = 500 µM. B) Cell concentration = 25 µM, syr concentration = 600 µM. C) Cell concentration = 25 µM, syr concentration = 800 µM.



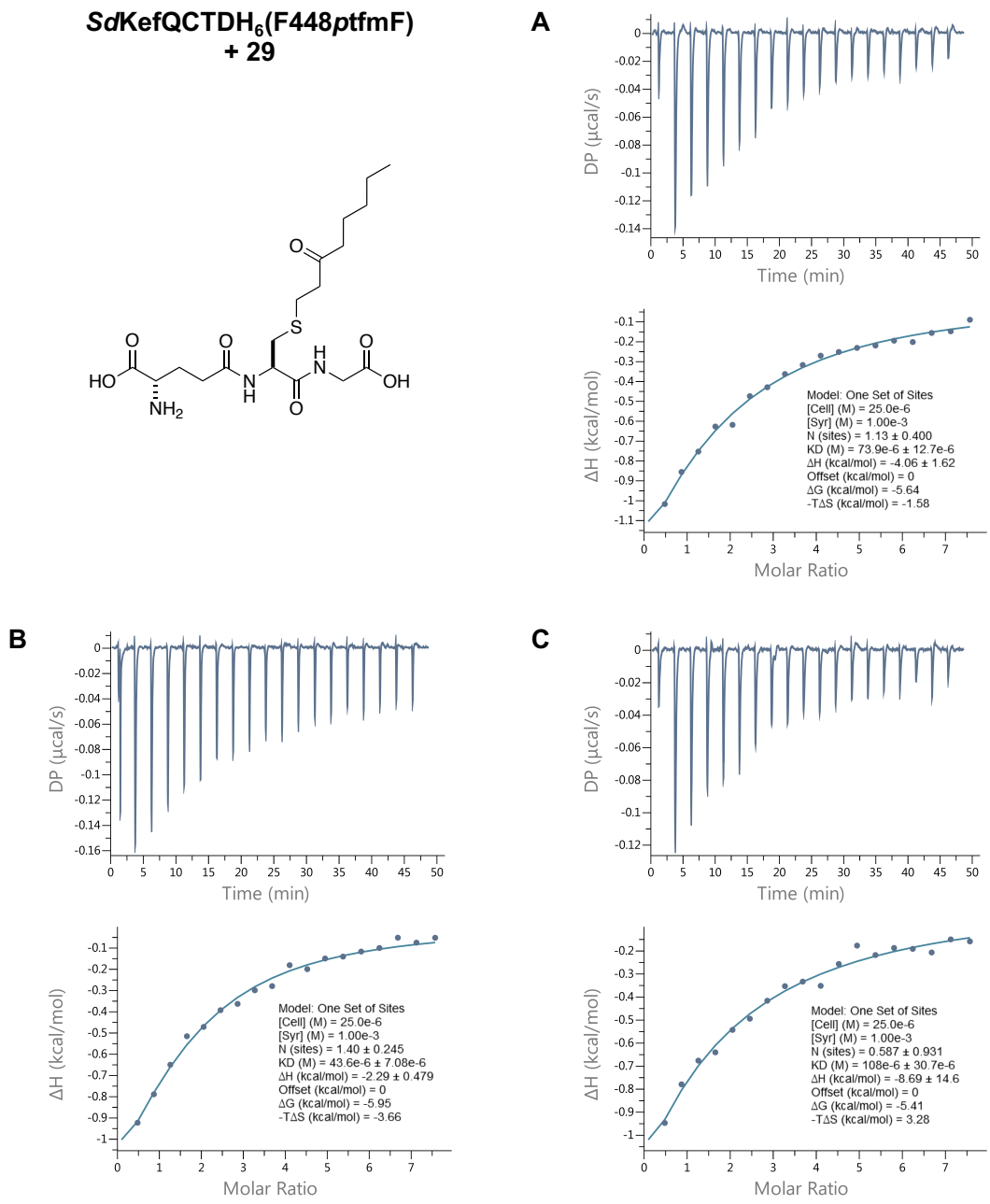
ITC Curves for *Sd*KefQCTDH$_6$ + 29. Cell = *Sd*KefQCTDH$_6$. Syr = **29**. A)–C) Cell concentration = 25 µM, syr concentration = 300 µM.



ITC Curves for *Sd*KefQCTDH$_6$(F448*p*tfmF) + ^{t}BuSG (26). Cell = *Sd*KefQCTDH$_6$(F448*p*tfmF). Syr = ^{t}BuSG (**26**). A) Cell concentration = 25 µM, syr concentration = 800 µM. B) And C) Cell concentration = 50 µM, syr concentration = 800 µM.

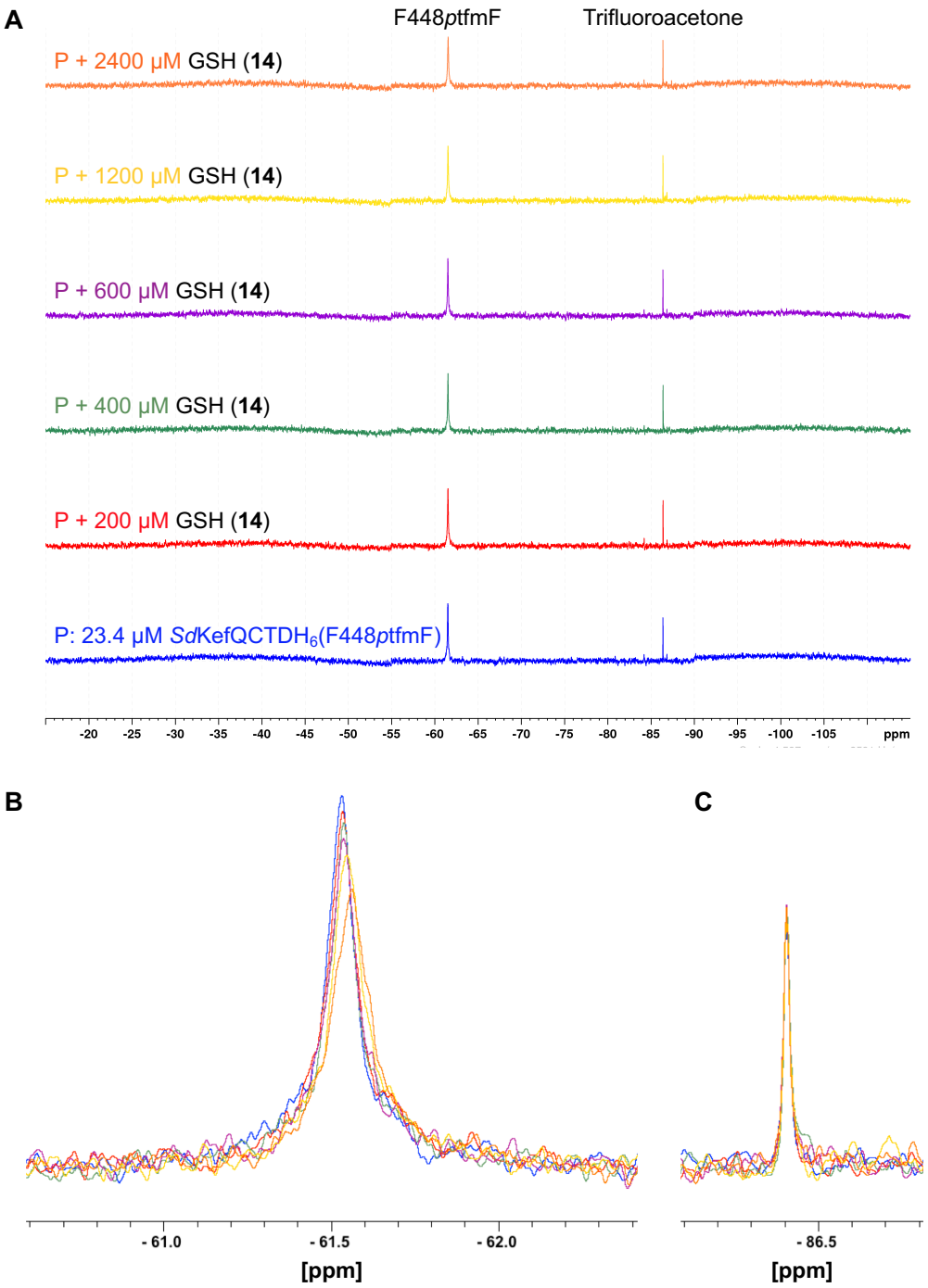


ITC Curves for *Sd*KefQCTDH$_6$(F448*p*tfmF) + ESG (20). Cell = *Sd*KefQCTDH$_6$(F448*p*tfmF). Syr = ESG (**20**). A) And B) Cell concentration = 25 µM, syr concentration = 4000 µM. C) Cell concentration = 25 µM, syr concentration = 5000 µM.

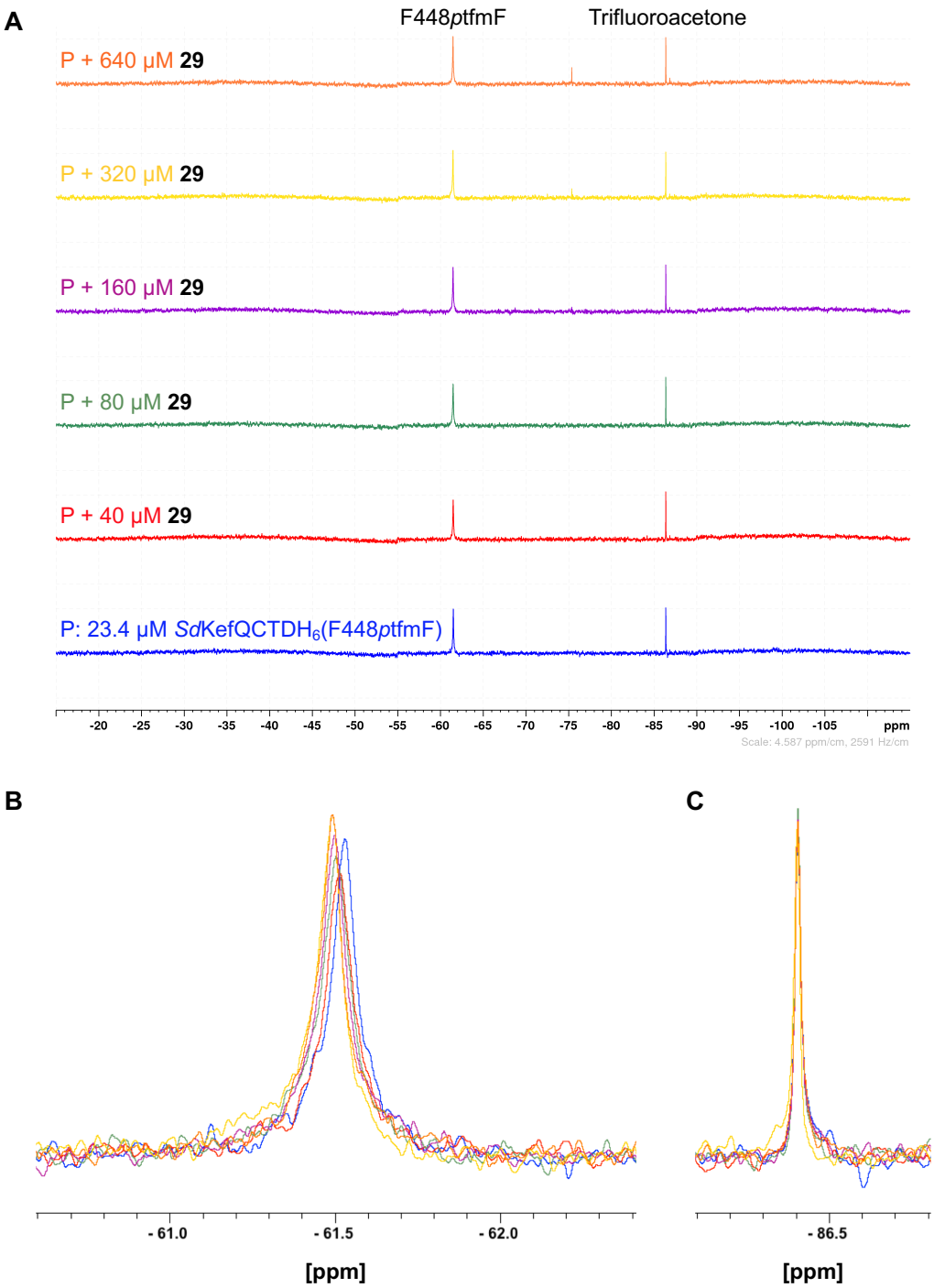


ITC Curves for *Sd*KefQCTDH$_6$(F448*p*tfmF) + 29. Cell = *Sd*KefQCTDH$_6$(F448*p*tfmF). Syr = **29**. A)–C) Cell concentration = 25 µM, syr concentration = 1000 µM.

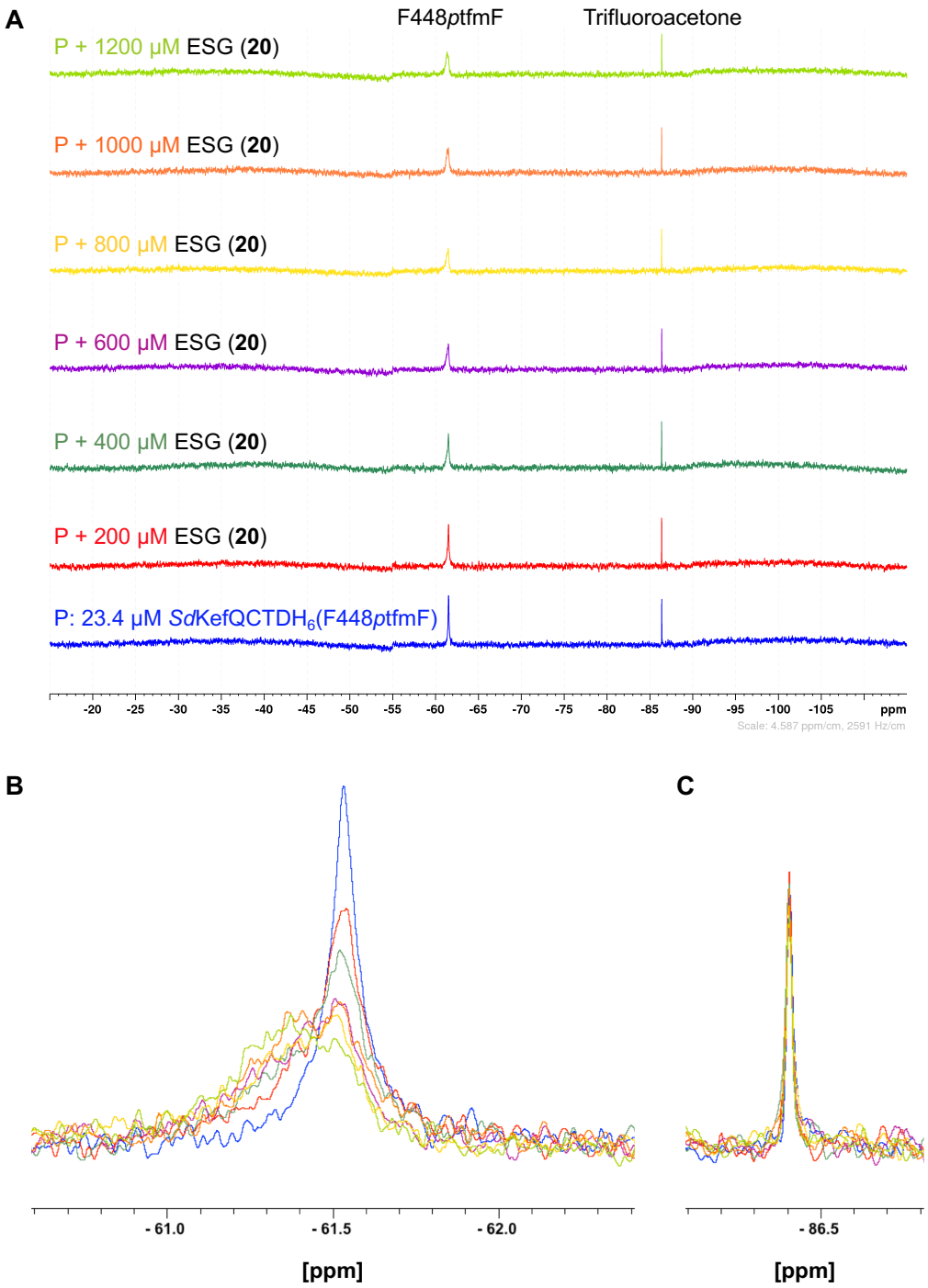
C. Protein-Observed ^{19}F NMR Spectra



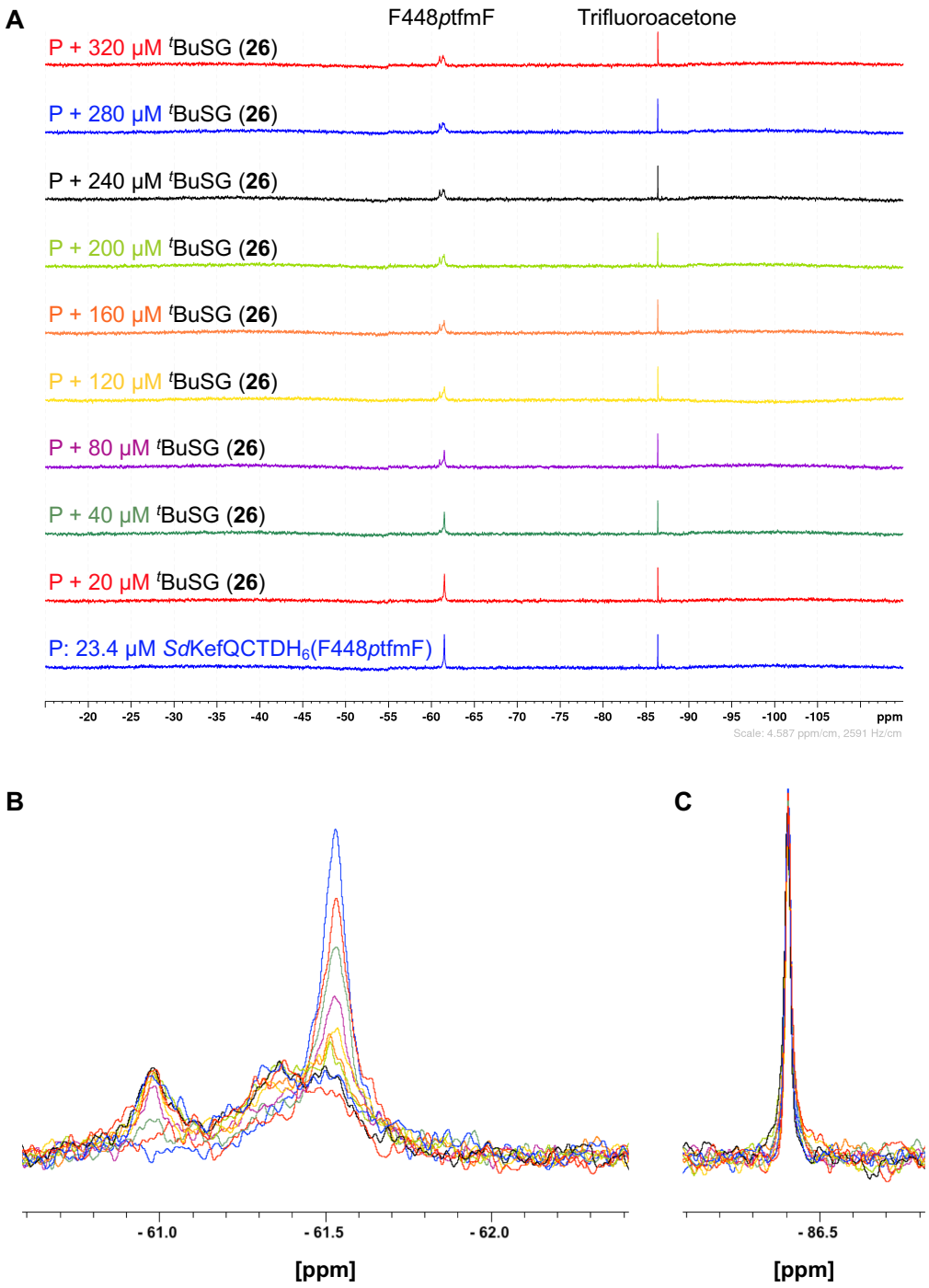
Supplementary ^{19}F NMR data for Figure 2.8A: Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and GSH (14). Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of GSH (**14**) into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef inhibition by GSH (**14**). B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



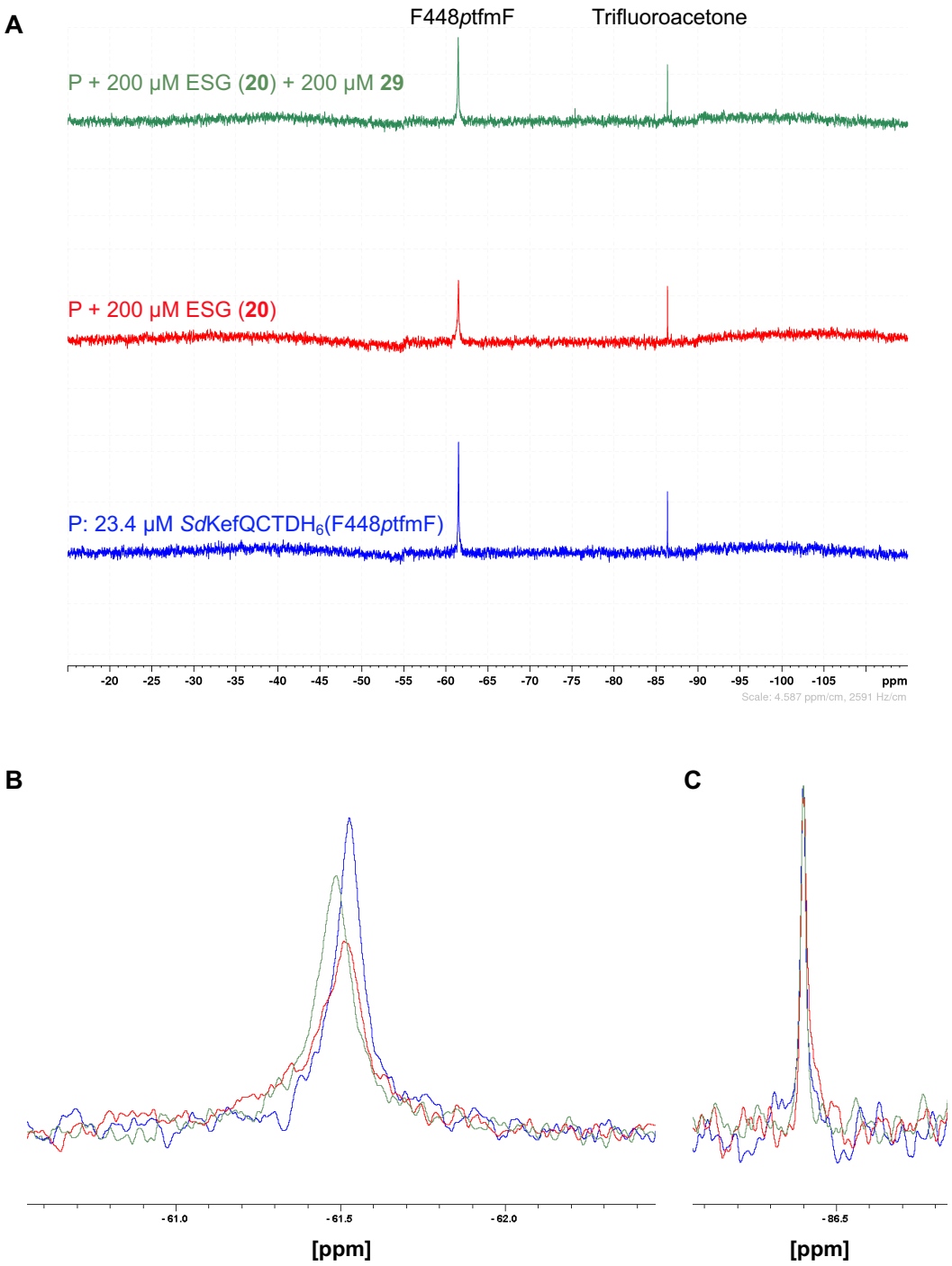
Supplementary ^{19}F NMR data for Figure 2.8B: Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and compound 29. Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of compound **29** into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef inhibition by compound **29**. B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



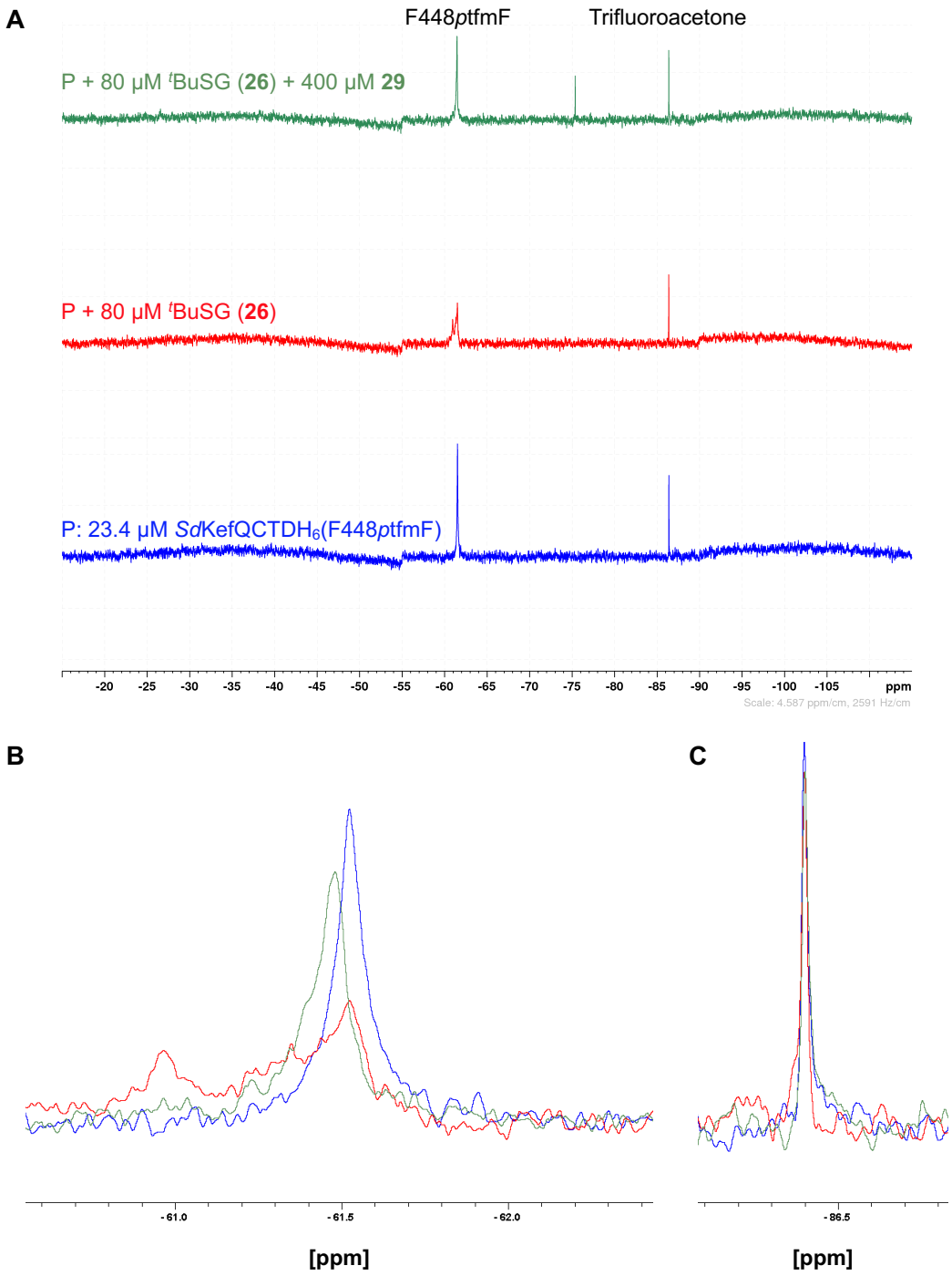
Supplementary ^{19}F NMR data for Figure 2.9A: Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and ESG (20). Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of ESG (**20**) into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef activation by ESG (**20**). B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



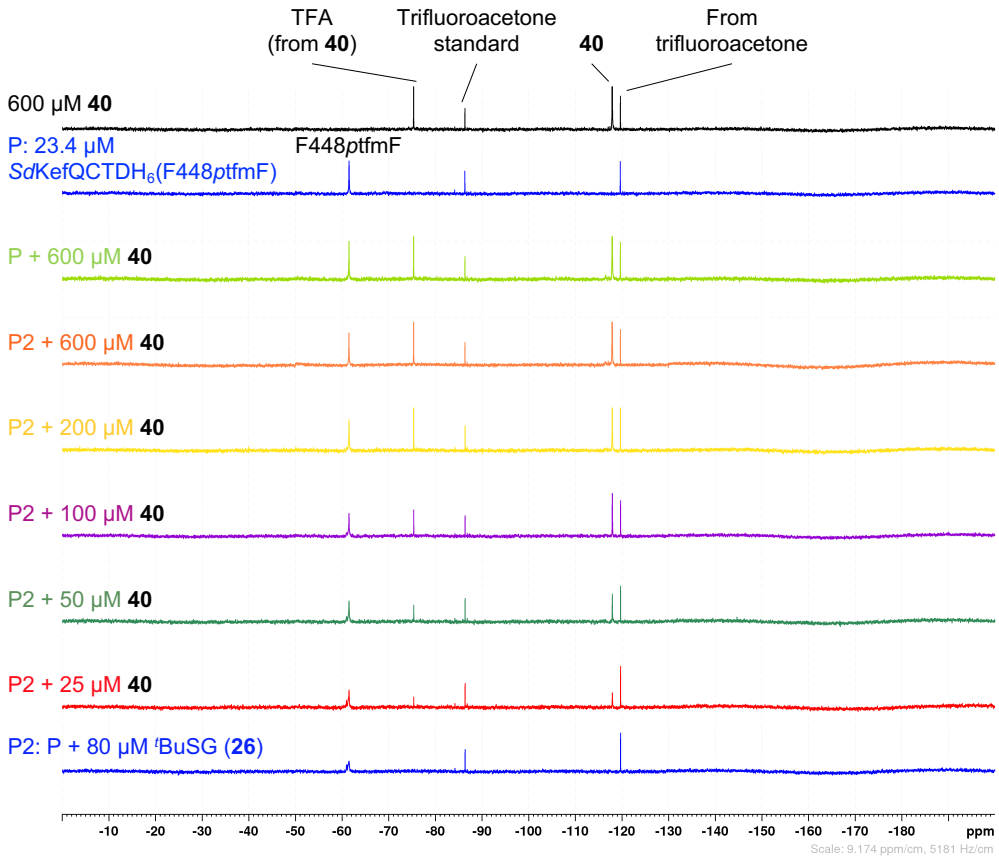
Supplementary ^{19}F NMR data for Figure 2.9B: Probing Kef dynamics *via* ^{19}F NMR spectroscopy using *Sd*KefQCTDH$_6$(F448*p*tfmF) and ^{t}BuSG (26). Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of ^{t}BuSG (**26**) into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef activation by ^{t}BuSG (**26**). B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



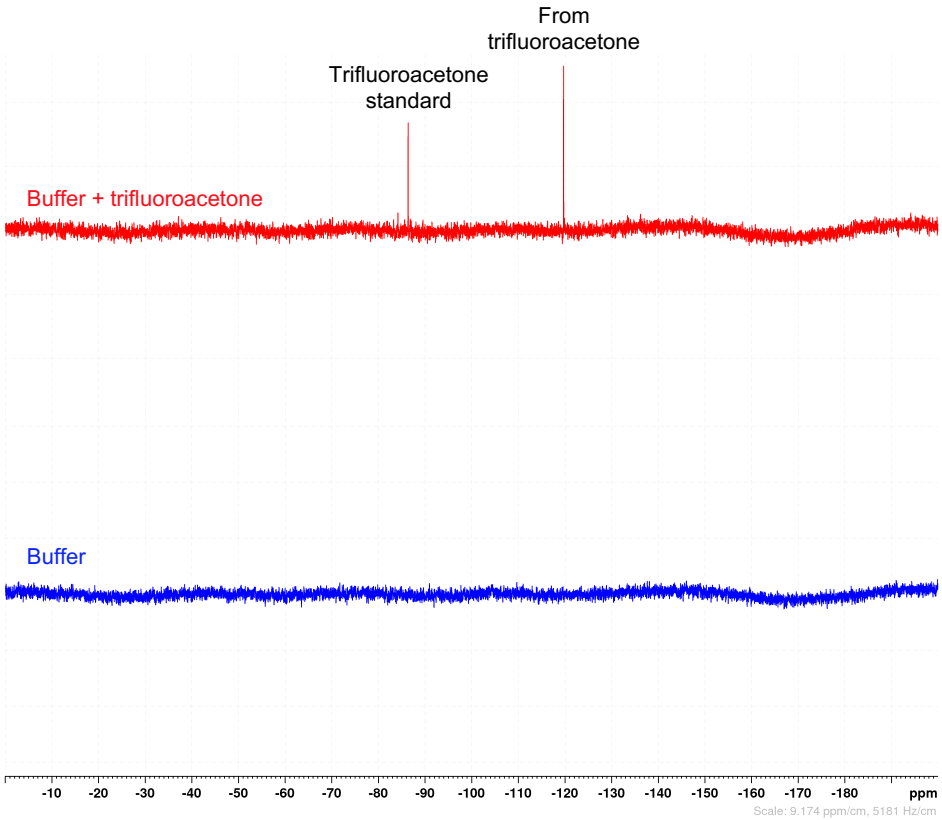
Supplementary ^{19}F NMR data for Figure 2.10A: ^{19}F NMR Spectroscopy competition experiments with compound 29 and ESG (20). Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of ESG (**20**) and compound **29** into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef inhibition by compound **29**. B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



Supplementary ^{19}F NMR data for Figure 2.10B: ^{19}F NMR Spectroscopy competition experiments with compound 29 and ^{t}BuSG (26). Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). A) Titration of ^{t}BuSG (**26**) and compound **29** into a fixed concentration of protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)). The peaks for the F448*p*tfmF capping residue and the trifluoroacetone standard are labelled. These data are consistent with Kef inhibition by compound **29**. B) Side view of the peak for the F448*p*tfmF capping residue. C) Side view of the peak for the trifluoroacetone standard.



Supplementary ^{19}F NMR data for Figure 2.11: Elucidation of the function of compound 40 using *Sd*KefQCTDH$_6$(F448*p*tfmF) and ^{19}F NMR spectroscopy. Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). Titration of compound **40** into P2 (23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF) with 80 µM ^{t}BuSG (**26**)) and monitoring by ^{19}F NMR until saturation by compound **40** (600 µM). Spectra for 600 µM compound **40**, the protein (P: 23.4 µM *Sd*KefQCTDH$_6$(F448*p*tfmF)) only, and P + 600 µM compound **40** are shown for comparison. Trifluoroacetone is used as an internal ^{19}F standard.



Supplementary ^{19}F NMR data for Figure 2.11: Elucidation of the function of compound 40 using *Sd*KefQCTDH$_6$(F448*p*tfmF) and ^{19}F NMR spectroscopy. Data from Dr Anthony Chan (unpublished). The spectra show the full ^{19}F NMR chemical shift range measured (0 to −200 ppm, in buffer). ^{19}F NMR Spectrum of NMR buffer + trifluoroacetone standard shows 2 peaks, whereas the ^{19}F NMR spectrum of the buffer alone shows none. This explains the presence of the additional peak at approximately −120 ppm in Figure 2.11A.

D. Kirby-Bauer Disc Diffusion Assays

Each table shows the data for each disc assay figure from Chapter 3 or 4. In each row, the plate letter (A–E) corresponds to the letter for each plate in the disc assay figure. N/A = Not applicable. The presence of Amp and IPTG in each plate is indicated by + or −, where + = Amp or IPTG included in plate, and − = Amp or IPTG excluded from plate. When IPTG was included, the concentration is indicated in parentheses. Inhibition zone diameters associated with the inhibition effect of the compounds were recorded by averaging both the vertical and horizontal diameters of each zone and are given to the nearest whole millimetre (mm). As each compound disc measures 6 mm in diameter, an inhibition zone diameter of 6 mm represents that a compound does not inhibit bacterial growth. The number of repeats, n, represents the number of independent trials of each plate. For plates with >1 repeat, the inhibition zone diameters were averaged over all repeats.

Chapter 3

Supplementary data for Figure 3.7: Kirby-Bauer disc diffusion assay to test OXFKEF12–14 (31–33) in *E. coli* MJF645 expressing either *Bp*Kef1 or *Bp*Kef2. All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF12 (31)**	**OXFKEF13 (32)**	**OXFKEF14 (33)**	**Number of repeats, n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.1 mM)	6	16	19	20	3	Positive control for plates B and C
B	*E. coli* MJF645	pTrcBpKef1H_6	+	+ (0.1 mM)	6	6	6	6	2	
C	*E. coli* MJF645	pTrcBpKef2H_6	+	+ (0.1 mM)	6	6	6	6	1	

Supplementary data for Figure 3.8: Kirby-Bauer disc diffusion assay to test L-methyl esters OXFKEF12–14 (31–33) in *E. coli* MJF645 expressing *Yp*KefGBP. All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF12 (31)**	**OXFKEF13 (32)**	**OXFKEF14 (33)**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.1 mM)	6	16	19	20	3	Positive control for plate C
B	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.05 mM)	6	18	20	21	1	Positive control for plate D
C	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	17	25	3	
D	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	6	14	24	1	

Supplementary data for Figure 3.9: Control experiments for MJF645 expressing *Yp*KefGBP. All compounds (OXFKEF12–14 (**31**–**33**) and OXFKEF16 (**48**)) were tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)						
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF12 (31)**	**OXFKEF13 (32)**	**OXFKEF14 (33)**	**OXFKEF16 (48)**	**n**	**Remark**
A	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	17	25	N/A	3	Positive control for plates C–E
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	6	14	24	N/A	1	Positive control for plates C–E
C	*E. coli* MJF645	pTrcYpKefGBP	+	−	6	6	6	6	N/A	1	
D	*E. coli* MJF645	N/A	−	−	6	6	6	6	N/A	1	
E	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	N/A	20	30	6	1	

Supplementary data for Figure 3.10: Kirby-Bauer disc diffusion assay to test L-methyl esters OXFKEF12–14 (31–33) in *E. coli* Frag 1 expressing *Yp*KefGBP. All compounds (OXFKEF12–14, **31**–**33**) were tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF12 (31)**	**OXFKEF13 (32)**	**OXFKEF14 (33)**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.1 mM)	6	16	19	20	3	Positive control for plate C
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	17	25	3	Positive control for plate C
C	*E. coli* Frag 1	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	6	6	1	

Supplementary data for Figure 3.13: Kirby-Bauer disc diffusion assay for OXFKEF14 (33) in wild-type *Y. pseudotuberculosis* and *Y. pestis*. OXFKEF14 (**33**) was tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)			
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF14 (33)**	**n**	**Remark**
A	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.1 mM)	6	29	3	Positive control for plates D and E
B	*E. coli* MJF645	pTrcBpKef1H_6	+	+ (0.1 mM)	6	6	3	Negative control for plates D and E
C	*E. coli* MJF645	pTrcBpKef2H_6	+	+ (0.1 mM)	6	6	3	Negative control for plates D and E
D	*Y. pseudotuberculosis* IP32953	N/A	−	−	6	6	3	
E	*Y. pestis* CO92	N/A	−	−	6	6	2	Experiment performed by Dr Andrew Scott

Supplementary data for Figure 3.16: Kirby-Bauer disc diffusion assay for OXFKEF12–14 (31–33) in *E. coli* MJF645 and Frag 1 expressing *Yp*KefGBP and *Yp*KefG(BH$_6$)P. All compounds were tested at 15 mM in Milli-Q® H_2O. Milli-Q® H_2O was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**Milli-Q® H_2O**	**OXFKEF12 (31)**	**OXFKEF13 (32)**	**OXFKEF14 (33)**	**n**	**Remark**
A	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	17	25	3	Positive control for plates C and D
B	*E. coli* Frag 1	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	6	6	1	Negative control for plates C and D
C	*E. coli* MJF645	pTrcYpKefG(BH$_6$)P	+	+ (0.1 mM)	6	6	14	20	1	
D	*E. coli* Frag 1	pTrcYpKefG(BH$_6$)P	+	+ (0.1 mM)	6	6	6	6	1	

Chapter 4

Supplementary data for Figure 4.10: Disc assay to test compound 80 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. Both compounds (**80**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **80** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**OXFKEF14 (33, 7.5 mM)**	**80 (15 mM)**	**OXFKEF14 (33) + 80**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.05 mM)	6	19	6	20	1	OXFKEF14 (**33**) is a positive control
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	26	6	28	1	OXFKEF14 (**33**) is a positive control
C	*E. coli* MJF645	N/A	−	−	6	6	6	6	2	

Supplementary data for Figure 4.11: Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**OXFKEF14 (33, 7.5 mM)**	**81 (15 mM)**	**OXFKEF14 (33) + 81**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.05 mM)	6	21	6	20	1	OXFKEF14 (**33**) is a positive control
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	24	6	27	1	OXFKEF14 (**33**) is a positive control
C	*E. coli* MJF645	N/A	−	−	6	6	6	6	2	

Supplementary data for Figure 4.12: Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 30 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 7.5 mM. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**OXFKEF14 (33, 7.5 mM)**	**81 (30 mM)**	**OXFKEF14 (33) + 81**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.05 mM)	6	20	6	19	1	OXFKEF14 (**33**) is a positive control
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	24	6	24	1	OXFKEF14 (**33**) is a positive control
C	*E. coli* MJF645	N/A	−	−	6	6	6	6	2	

Supplementary data for Figure 4.13: Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 5 mM. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**OXFKEF14 (33, 5 mM)**	**81 (15 mM)**	**OXFKEF14 (33) + 81**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH_6	+	+ (0.05 mM)	6	19	6	18	1	OXFKEF14 (**33**) is a positive control
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	21	6	26	1	OXFKEF14 (**33**) is a positive control
C	*E. coli* MJF645	N/A	−	−	6	6	6	6	1	

Supplementary data for Figure 4.14: Disc assay to test compound 81 in *E. coli* MJF645 expressing either *Sd*Kef or *Yp*KefGBP. Both compounds (**81**, OXFKEF14 (**33**)) were tested in 80% *v/v* DMSO. Compound **81** was tested at 15 mM and OXFKEF14 (**33**, tested as a positive control) was tested at 2.5 mM. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**OXFKEF14 (33, 2.5 mM)**	**81 (15 mM)**	**OXFKEF14 (33) + 81**	**n**	**Remark**
A	*E. coli* MJF645	pTrcSdKefH$_6$	+	+ (0.05 mM)	6	14	6	15	1	OXFKEF14 (**33**) is a positive control
B	*E. coli* MJF645	pTrcYpKefGBP	+	+ (0.05 mM)	6	19	6	19	1	OXFKEF14 (**33**) is a positive control
C	*E. coli* MJF645	N/A	−	−	6	6	6	6	1	

Supplementary data for Figure 4.15: Disc assay for ~2.6–11 mM MG (13) in *E. coli* for the formation of activator SLG (17) in the cell. MG (**13**) was tested in 80% *v/v* DMSO at the concentrations indicated. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**~2.6 mM MG (13)**	**~5.3 mM MG (13)**	**~11 mM MG (13)**	**n**	**Remark**
A	*E. coli* Frag 1	pTrcSdKefH$_6$	+	+ (0.1 mM)	6	6	6	9	3	
B	*E. coli* Frag 1	pTrcYpKefGBP	+	+ (0.1 mM)	6	6	6	9	3	
C	*E. coli* Frag 1	N/A	−	−	6	6	6	8	2	Control for plates A and B
D	*E. coli* MJF335 + GSH (**14**, 1 mM)	N/A	−	−	6	6	8	11	1	Control for plates A and B
E	*E. coli* MJF335	N/A	−	−	6	6	9	12	1	Control for plates A and B

Supplementary data for Figure 4.16: Disc assay for ~21–84 mM MG (13) in *E. coli* for the formation of activator SLG (17) in the cell. MG (**13**) was tested in 80% *v/v* DMSO at the concentrations indicated. 80% *v/v* DMSO was tested as a negative control.

					Mean inhibition zone diameter (mm)					
Plate	**Host/ Species**	**Plasmid**	**Amp**	**IPTG**	**80% *v/v* DMSO**	**~21 mM MG (13)**	**~42 mM MG (13)**	**~84 mM MG (13)**	**n**	**Remark**
A	*E. coli* Frag 1	pTrcSdKefH$_6$	+	+ (0.1 mM)	6	11	14	19	3	
B	*E. coli* Frag 1	pTrcYpKefGBP	+	+ (0.1 mM)	6	10	15	18	3	
C	*E. coli* Frag 1	N/A	−	−	6	10	13	15	2	Control for plates A and B
D	*E. coli* MJF335 + GSH (**14**, 1 mM)	N/A	−	−	6	10	15	19	1	Control for plates A and B
E	*E. coli* MJF335	N/A	−	−	6	10	16	18	1	Control for plates A and B

E. MIC Assay

Blank-Corrected Data

Supplementary data for Figure 3.12: Determination of a MIC for OXFKEF14 (33) in *E. coli* MJF645 expressing *Yp*KefGBP. Blank: 2× TY Medium containing 0.2 mg mL^{-1} Amp and 0.2 mM IPTG. Negative control: Cell culture of *E. coli* MJF645 expressing *Yp*KefGBP (using pTrcYpKefGBP). Wells contain (final concentrations) 0.1 mg mL^{-1} Amp and 0.1 mM IPTG. The same culture of *E. coli* MJF645 expressing *Yp*KefGBP was grown in the presence of 0.1 mg mL^{-1} Amp, 0.1 mM IPTG, and OXFKEF14 (**33**) at varying concentrations (as indicated). Compound (OXFKEF14, **33**) wells which showed cell growth such that mean OD_{600} > 20% (mean OD_{600} of the negative control wells) are shown in red. Wells which showed growth inhibition/death such that mean OD_{600} < 20% (mean OD_{600} of the negative control wells) are shown in blue. OD_{600} Values for compound wells which were excluded from the mean are shown in light blue. Blank wells are shown in grey.

	Negative Control	Blank	OXFKEF14 (33) Concentration (μg mL^{-1})								
			1000	500	250	125	62.5	31.3	15.6	7.81	3.91
OD_{600}	0.437	0.001	0.001	0.001	0.001	0.001	0.002	0.008	0.054	0.245	0.381
	0.348	0.000	0.002	0.001	0.000	0.002	0.002	0.004	0.063	0.237	0.316
	0.338	0.000	0.001	0.001	0.000	0.001	0.000	0.003	0.024	0.248	0.345
	0.365	0.000	0.003	0.001	0.001	0.001	0.002	0.075	0.008	0.290	0.329
Mean	0.372	0.000	0.002	0.001	0.001	0.001	0.002	0.005	0.047	0.255	0.343
20%	0.074		**MIC = 15.6 μg mL^{-1}**								

F. Pierce™ 660 nm Protein Assays

BSA = Bovine serum albumin, which was used as a model protein to generate a standard curve. MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from either *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation).

Supplementary data for Table 3.5. Pierce™ 660 nm protein assay 1: Estimation of the concentration of *Yp*KefBH$_6$ protein samples (in PBS buffer).

Assay 1 Plate Layout

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		Blank for BSA	Blank for BSA	Blank for BSA	250	250	250		Blank for *Yp*KefBH$_6$	Blank for *Yp*KefBH$_6$	Blank for *Yp*KefBH$_6$	
C		2000	2000	2000	150	150	150					
D		1500	1500	1500	125	125	125		MJF645 MF	MJF645 MF	MJF645 MF	
E		1000	1000	1000	100	100	100		MJF645 SN	MJF645 SN	MJF645 SN	
F		750	750	750	75	75	75		Frag 1 MF	Frag 1 MF	Frag 1 MF	
G		500	500	500	50	50	50		Frag 1 SN	Frag 1 SN	Frag 1 SN	
H												

Assay 1: Absorbance at 660 nm; Raw Data

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		0.070	0.069	0.072	0.192	0.196	0.200		0.067	0.070	0.070	
C		1.032	1.053	1.104	0.144	0.140	0.149					
D		0.806	0.818	0.832	0.130	0.131	0.131		0.128	0.103	0.152	
E		0.589	0.605	0.592	0.120	0.116	0.120		0.182	0.175	0.192	
F		0.462	0.466	0.461	0.106	0.105	0.106		0.110	0.122	0.109	
G		0.313	0.337	0.333	0.093	0.094	0.092		0.165	0.186	0.164	
H												

Assay 1: Absorbance at 660 nm; Blank-Corrected Data

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		Blank (for BSA) mean = 0.070			0.122	0.126	0.130		Blank (for *Yp*KefBH$_6$) mean = 0.069			
C		0.962	0.983	1.034	0.074	0.070	0.079					
D		0.736	0.748	0.762	0.060	0.061	0.061		0.059	0.034	0.083	
E		0.519	0.535	0.522	0.050	0.046	0.050		0.113	0.106	0.123	
F		0.392	0.396	0.391	0.036	0.035	0.036		0.041	0.053	0.040	
G		0.243	0.267	0.263	0.023	0.024	0.022		0.096	0.117	0.095	
H												

BSA Standard Curve Data

BSA Concentration (µg mL^{-1})	Absorbance at 660 nm				
	Blank-Corrected			Mean	Standard Deviation
50	0.023	0.024	0.022	0.023	0.0010
75	0.036	0.035	0.036	0.035	0.0006
100	0.050	0.046	0.050	0.048	0.0023
125	0.060	0.061	0.061	0.060	0.0006
150	0.074	0.070	0.079	0.074	0.0045
250	0.122	0.126	0.130	0.126	0.0040
500	0.243	0.267	0.263	0.257	0.0129
750	0.392	0.396	0.391	0.393	0.0026
1000	0.519	0.535	0.522	0.525	0.0085
1500	0.736	0.748	0.762	0.748	0.0130
2000	0.962	0.983	1.034	0.993	0.0370

Assay 1 Protein Concentration Data: *Yp*KefBH_6 Samples in PBS Buffer

	Mean Absorbance at 660 nm	**Protein Concentration (µg mL^{-1})**	**Volume for 20 µg Protein (µL)**
MJF645 MF	0.059	113	178
MJF645 SN	0.114	223	90
Frag 1 MF	0.045	85	236
Frag 1 SN	0.103	200	100

Supplementary data for Table 3.6. Pierce™ 660 nm protein assay 2: Estimation of the concentration of *Yp*KefBH$_6$ protein samples (in PBS buffer containing 20% *v*/*v* glycerol).

Assay 2 Plate Layout

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		Blank for SN samples	Blank for SN samples	Blank for SN samples					Blank for MF samples	Blank for MF samples	Blank for MF samples	
C												
D												
E		MJF645 SN	MJF645 SN	MJF645 SN					MJF645 MF	MJF645 MF	MJF645 MF	
F												
G		Frag 1 SN	Frag 1 SN	Frag 1 SN					Frag 1 MF	Frag 1 MF	Frag 1 MF	
H												

Assay 2: Absorbance at 660 nm; Raw Data

The data for the membrane fraction samples (highlighted in red) were considered to be inaccurate due to precipitation in the assay.

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		0.076	0.073	0.082					0.065	0.080	0.064	
C												
D												
E		1.322	1.345	1.278					1.676	1.749	1.751	
F												
G		0.659	0.712	0.685					1.008	0.811	0.834	
H												

Assay 2: Absorbance at 660 nm; Blank-Corrected Data

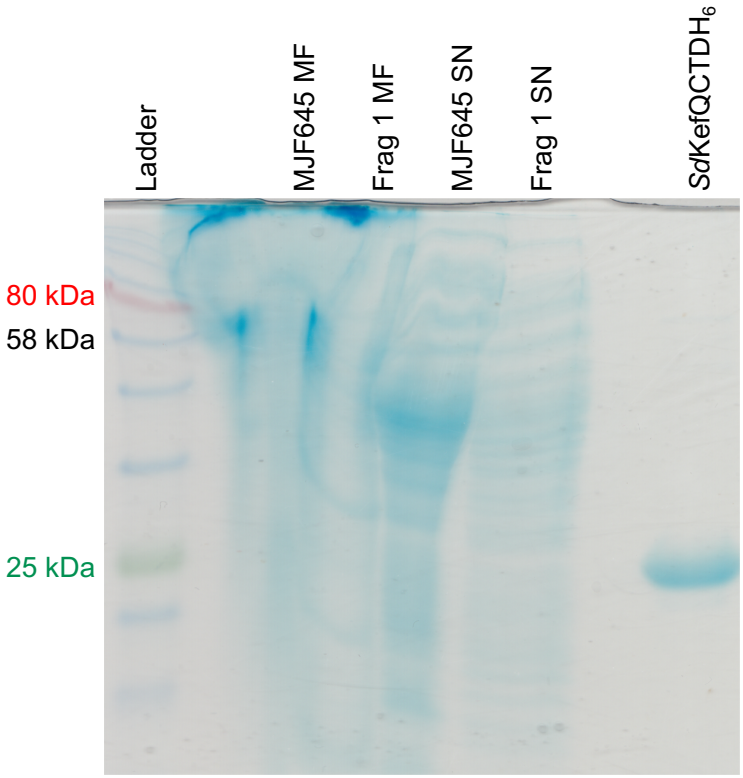
	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		Blank (for SN samples) mean = 0.077							Blank (for MF samples) mean = 0.070			
C												
D												
E		1.245	1.268	1.201					1.606	1.679	1.681	
F												
G		0.582	0.635	0.608					0.938	0.741	0.764	
H												

Assay 2 Protein Concentration Data: *Yp*KefBH$_6$ Samples in PBS buffer Containing 20% *v/v* Glycerol

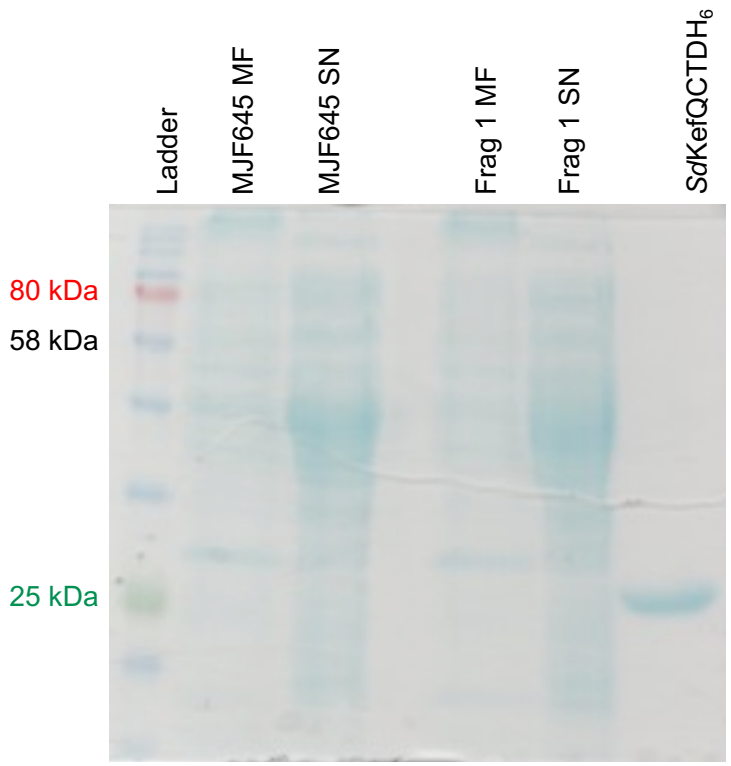
The data for the membrane fraction samples (highlighted in red) were considered to be inaccurate due to precipitation in the assay.

	Mean Absorbance at 660 nm	Protein Concentration (µg mL^{-1})	Volume for 20 µg Protein (µL)
MJF645 MF	1.656	3300	6
MJF645 SN	1.238	2470	8
Frag 1 MF	0.815	1620	12
Frag 1 SN	0.608	1210	17

G. SDS-PAGE Gels



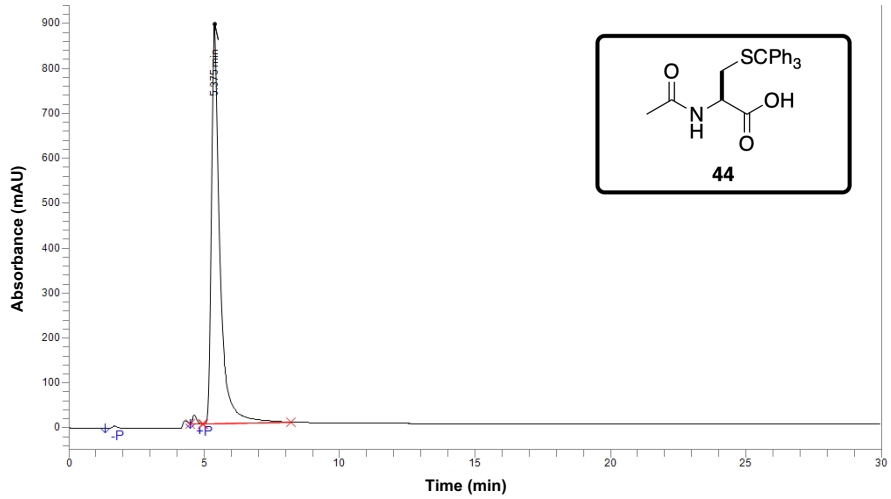
Supplementary Figure for Figure 3.18: Western blot for *Yp*KefBH$_6$ protein samples (suspended in PBS buffer containing 20% *v/v* glycerol) expressed in *E. coli* MJF645 and Frag 1. SDS-PAGE Gel for *Yp*KefBH$_6$ protein samples (suspended in PBS buffer containing 20% *v/v* glycerol) expressed in *E. coli* MJF645 and Frag 1, stained with Coomassie blue. Unstained SDS-PAGE gel was transferred to a membrane and blotted to give the western blot shown in Figure 3.18. MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). *Sd*KefQCTDH$_6$ (~26 kDa) was used as both a loading control and a control for the visualisation of the His tag using western blotting.



Supplementary Figure for Figure 3.19: Western blot for *Yp*KefBH$_6$ protein samples (suspended in Tris buffer containing 1.5% *w/v* DDM) produced in *E. coli* MJF645 and Frag 1. SDS-PAGE Gel for *Yp*KefBH$_6$ protein samples (suspended in buffer containing 1.5% *w/v* DDM) expressed in *E. coli* MJF645 and Frag 1, stained with Coomassie blue. Unstained SDS-PAGE gel was transferred to a membrane and blotted to give the western blot shown in Figure 3.19. MJF645, Frag 1 = *Yp*KefBH$_6$ Protein produced from *E. coli* MJF645 or Frag 1 transformed with pTrcYpKefG(BH$_6$)P and induced with IPTG (0.3 mM). MF = Membrane fraction; SN = Supernatant (from ultracentrifugation). *Sd*KefQCTDH$_6$ (~26 kDa) was used as both a loading control and a control for the visualisation of the His tag using western blotting.

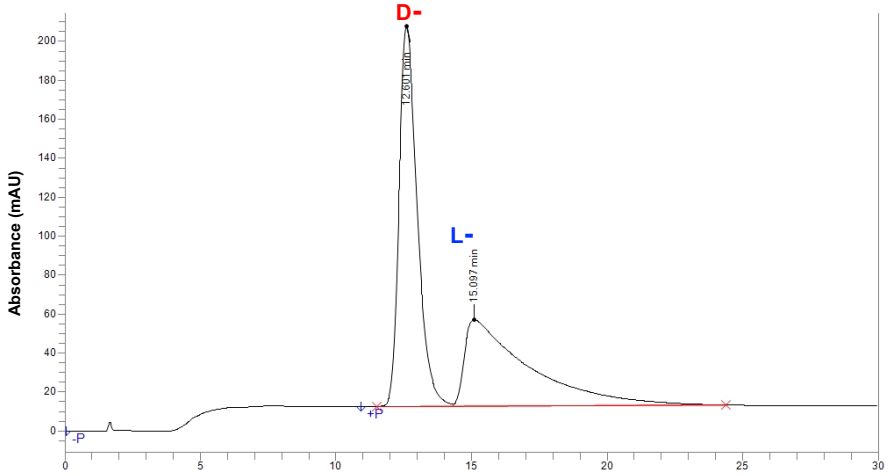
H. Chiral HPLC Traces for Intermediates

Chiral HPLC spectra are provided for intermediates. For chiral HPLC spectra of final compounds, see Chapters 3 and 4.

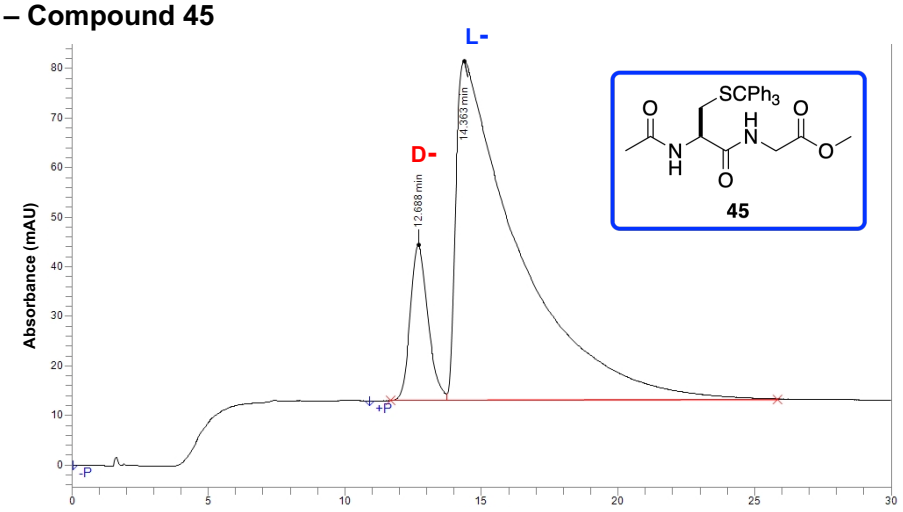


Chiral HPLC trace for compound 44. For details about the HPLC method and column used, see sections 6.1 and 6.2 in the experimental.

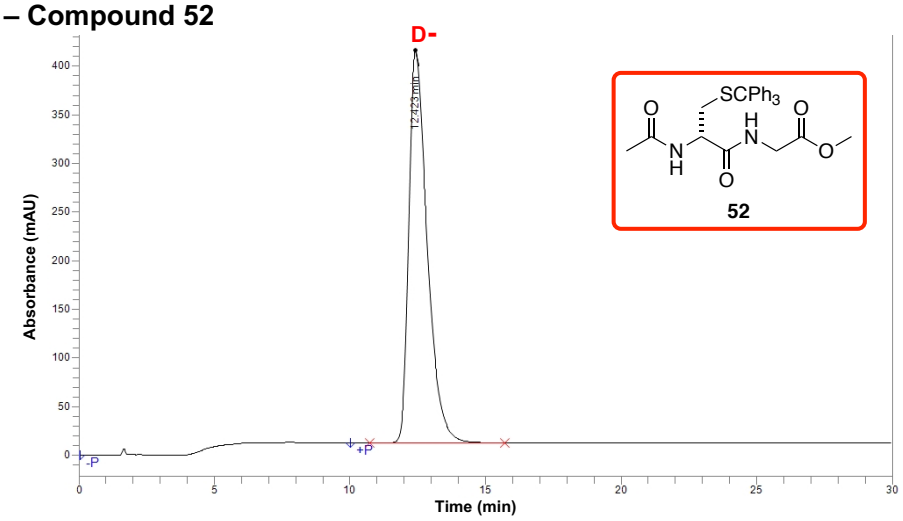
A – Scalemic mixture: 45 + 52



B – Compound 45

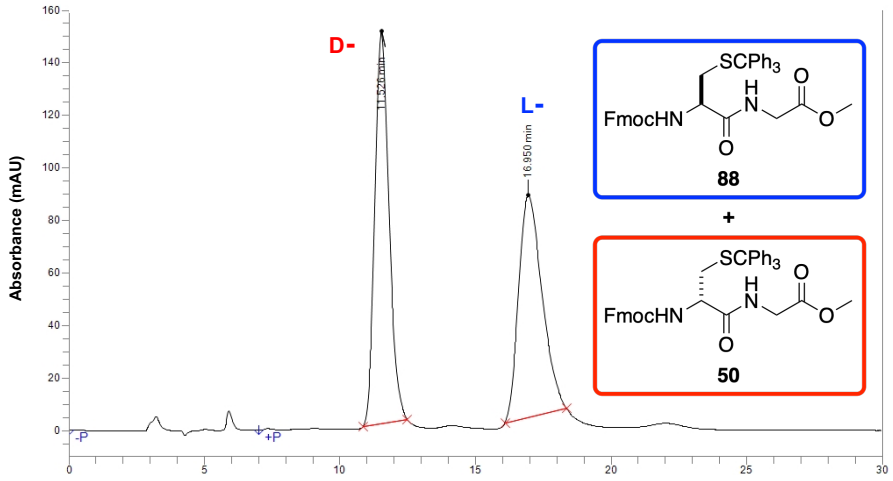


C – Compound 52

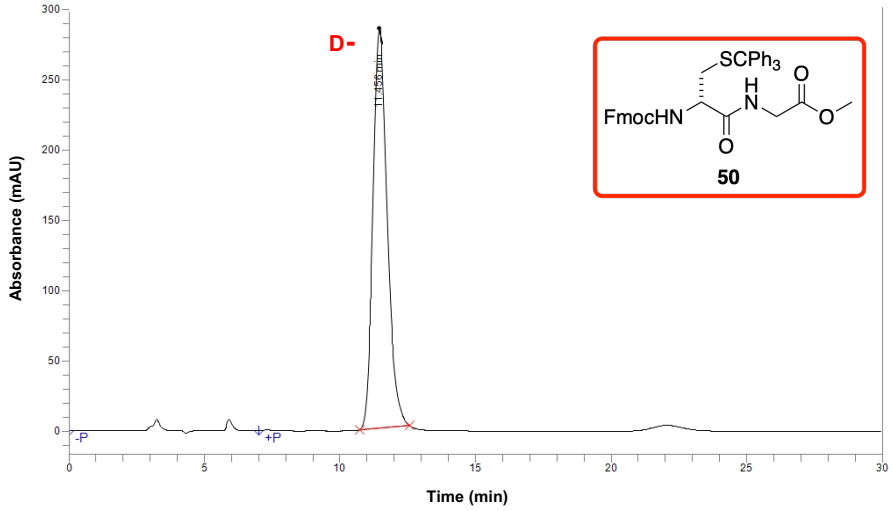


Chiral HPLC traces for compounds 45 and 52. A) Scalemic mixture: **45** + **52**. B) Compound **45** only. C) Compound **52** only. All three compounds were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental.

A – Scalemic mixture: 88 + 50



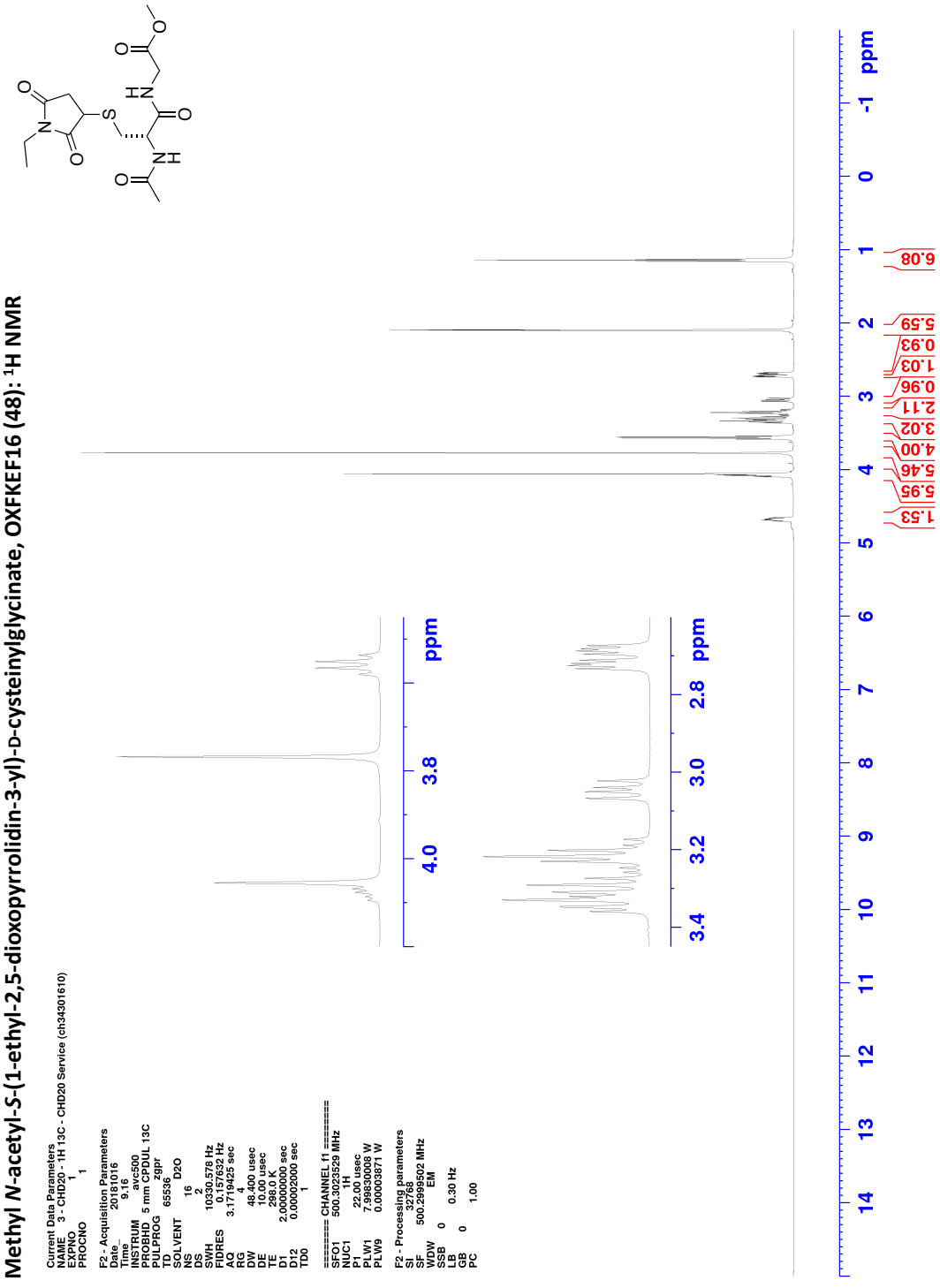
B – Compound 50

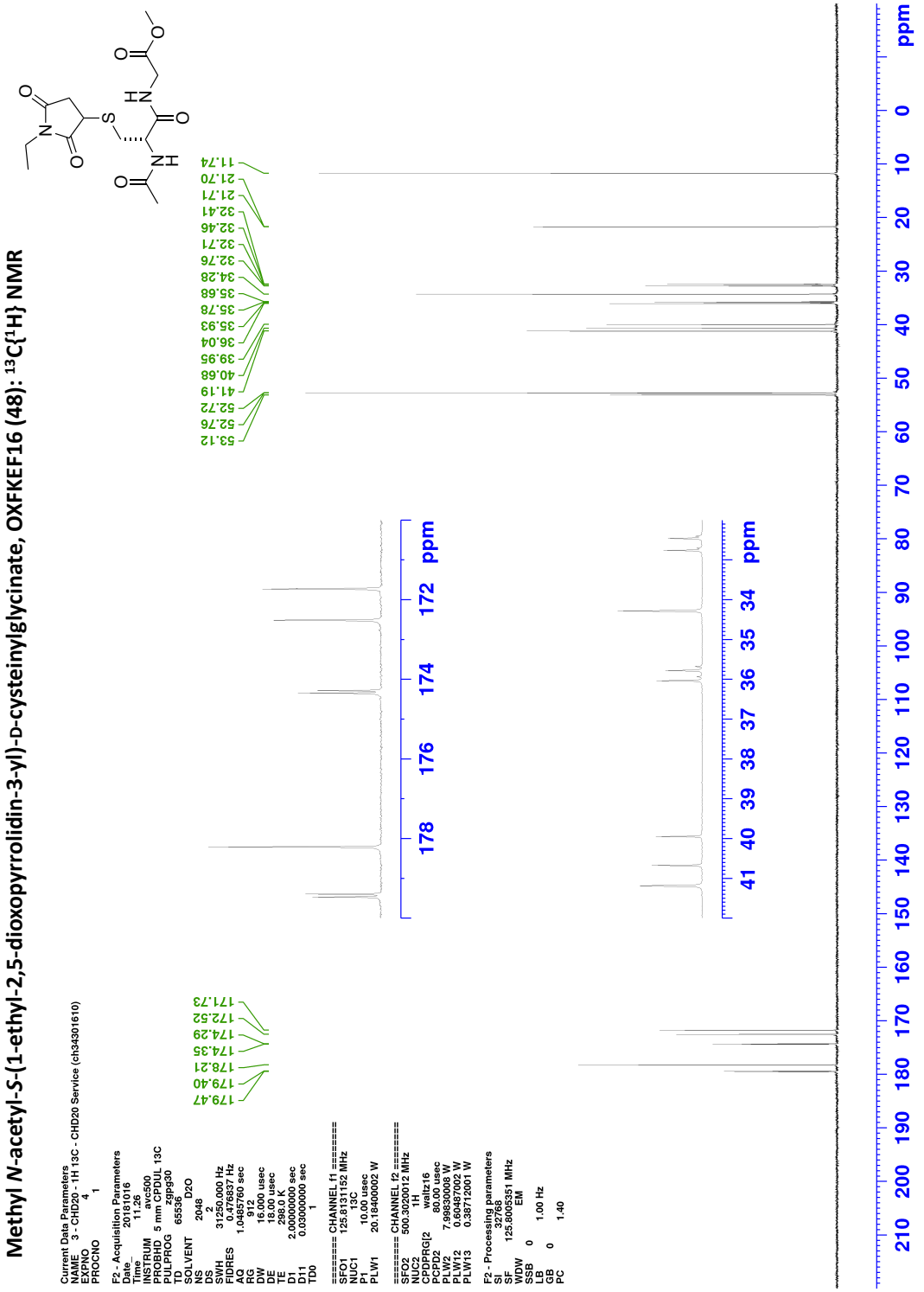


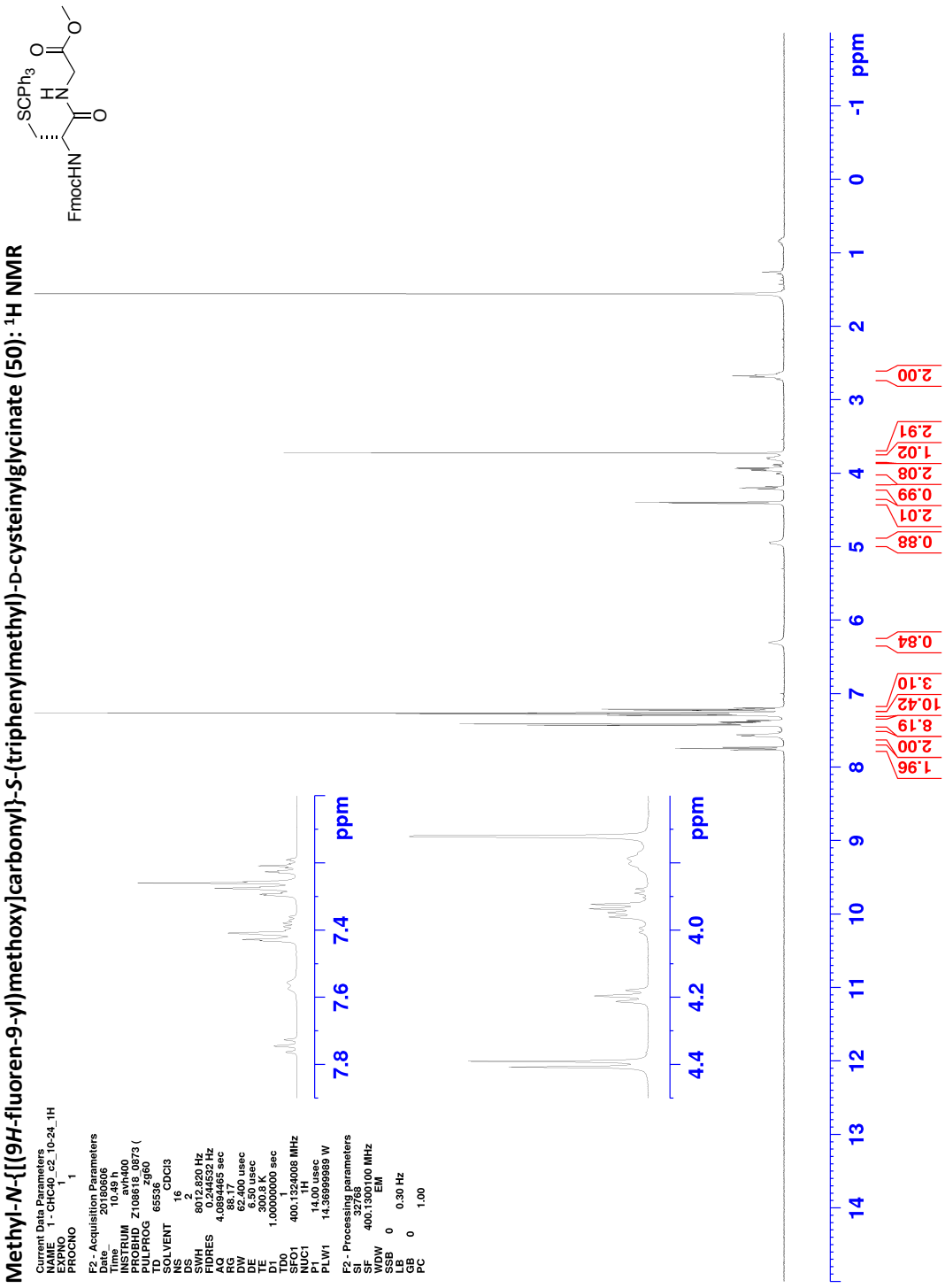
Chiral HPLC trace for compound 50. A) Scalemic mixture: **88** + **50**. B) Compound **50** only. Both samples were run using the same HPLC column and method. For further details, see sections 6.1 and 6.2 in the experimental. Compound **88** was synthesised by Holly Bunce (Conway Group, University of Oxford).

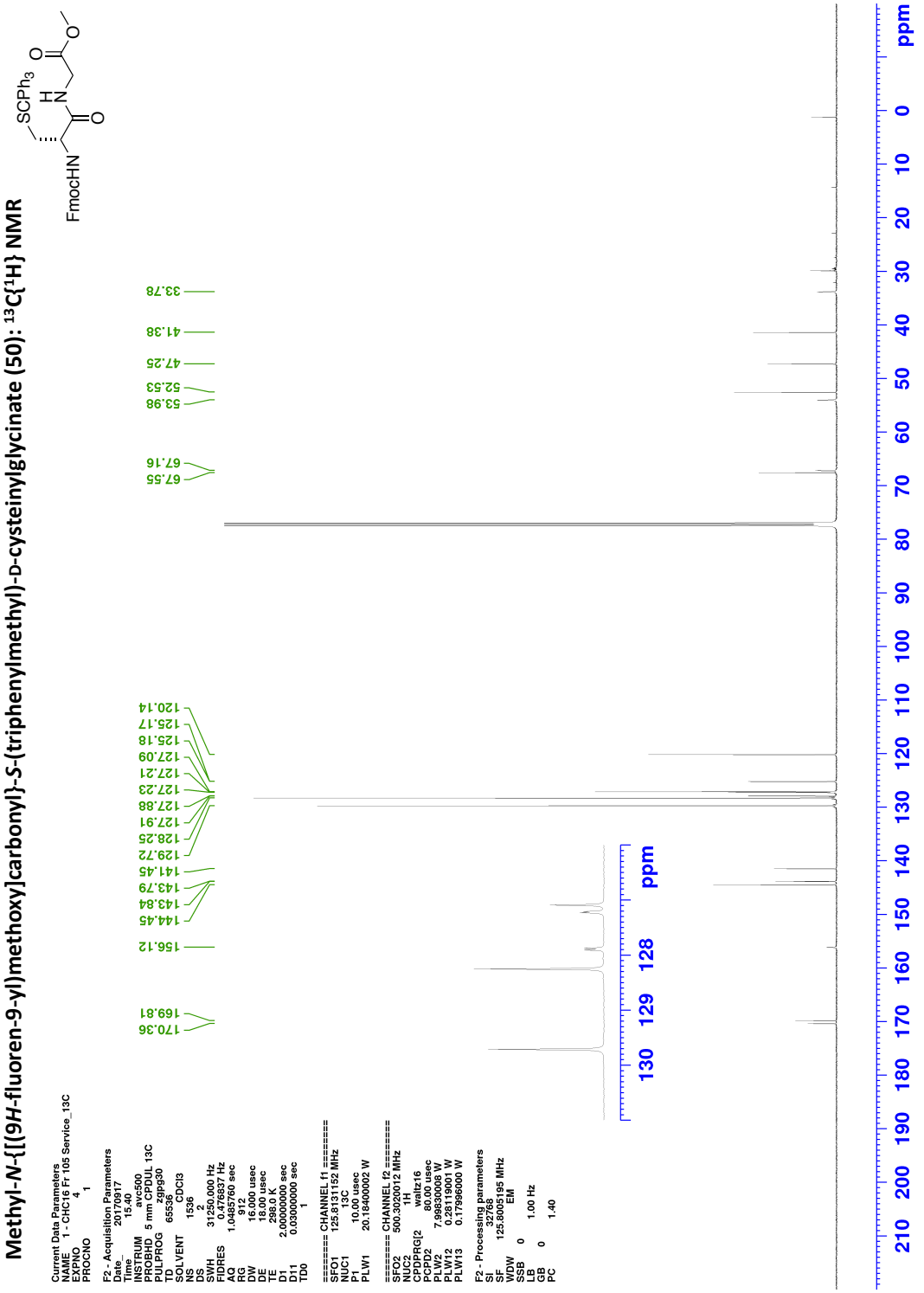
I. NMR Spectra of Novel Compounds

^{1}H, ^{13}C, And ^{19}F NMR (where appropriate) spectra for all novel compounds. Spectra are presented in numerical order according to compound number.

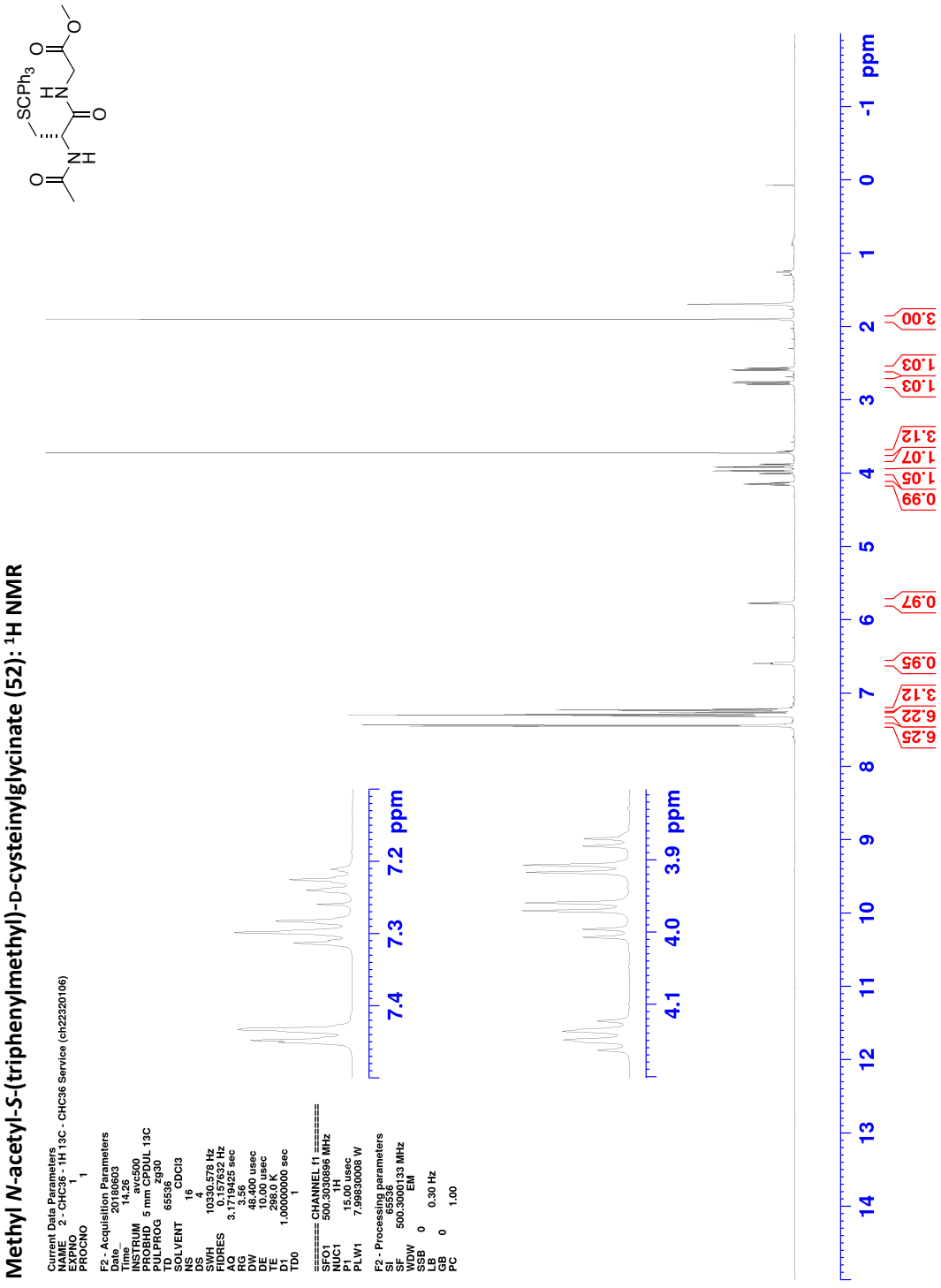

Methyl N-acetyl-S-(1-ethyl-2,5-dioxopyrrolidin-3-yl)-D-cysteinylglycinate, OXFKEF16 (48): 1H NMR
Current Data Parameters
NAME 3 - CHD20 - 1H 13C - CHD20 Service (ch34301610)
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20181016
Time 9.16
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zgpr
TD 65536
SOLVENT D2O
NS 16
DS 2
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 4
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 2.00000000 sec
D12 0.00002000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3023529 MHz
NUC1 1H
P1 22.00 usec
PLW1 7.99830008 W
PLW9 0.00003871 W
F2 - Processing parameters
SI 32768
SF 500.2999502 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
1.53
5.95
5.46
4.00
3.02
2.11
0.96
1.03
0.93
5.59
6.08
ppm


Methyl N-acetyl-S-(1-ethyl-2,5-dioxopyrrolidin-3-yl)-D-cysteinylglycinate, OXFKEF16 (48): 13C{1H} NMR
Current Data Parameters
NAME 3 - CHD20 - 1H 13C - CHD20 Service (ch34301610)
EXPNO 4
PROCNO 1
F2 - Acquisition Parameters
Date_ 20181016
Time 11.26
INSTRUM av500
PROBHD 5 mm CPDUL 13C
PULPROG zgpg30
TD 65536
SOLVENT D2O
NS 2048
DS 16
SWH 31250.000 Hz
FIDRES 0.476837 Hz
AQ 1.0485760 sec
RG 912
DW 16.000 usec
DE 18.00 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 125.8131152 MHz
NUC1 13C
P1 10.00 usec
PLW1 20.18400002 W
======== CHANNEL f2 ========
SFO2 500.3020012 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 7.99830008 W
PLW12 0.60487002 W
PLW13 0.38712001 W
F2 - Processing parameters
SI 32768
SF 125.8005351 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40
179.47
179.40
178.21
174.35
174.29
172.52
171.73
53.12
52.76
52.72
41.19
40.68
39.95
36.04
35.93
35.78
35.68
34.28
32.76
32.71
32.46
32.41
21.71
21.70
11.74
178 176 174 172 ppm
41 40 39 38 37 36 35 34 ppm
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 ppm

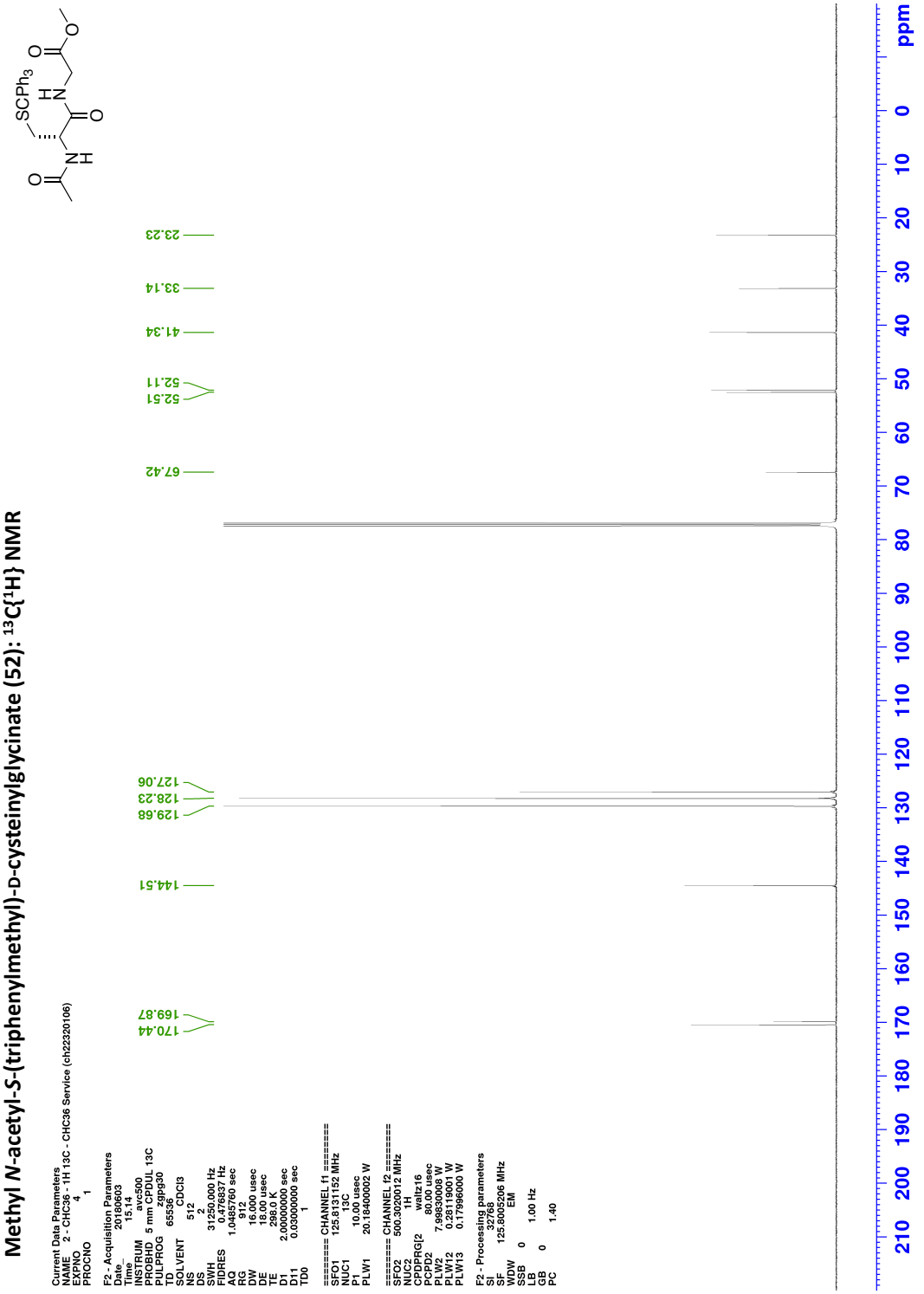

Methyl-*N*-{[(9*H*-fluoren-9-yl)methoxy]carbonyl}-*S*-(triphenylmethyl)-D-cysteinylglycinate (50): ^{1}H NMR
FmocHN
SCPh3
Current Data Parameters
NAME 1 - CHC40_c2_10-24_1H
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180606
Time_ 10.49 h
INSTRUM avh400
PROBHD Z108618_0873 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 88.17
DW 62.400 usec
DE 6.50 usec
TE 300.8 K
D1 1.00000000 sec
TD0 1
SFO1 400.1324008 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.36999989 W
F2 - Processing parameters
SI 32768
SF 400.1300100 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.8
7.6
7.4
ppm
4.4
4.2
4.0
ppm
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
ppm
1.96
2.00
8.19
10.42
3.10
0.84
0.88
2.01
0.99
2.08
1.02
2.91
2.00


Methyl-N-{[(9H-fluoren-9-yl)methoxy]carbonyl}-S-(triphenylmethyl)-D-cysteinylglycinate (50): 13C{1H} NMR
FmocHN
SCPh3
Current Data Parameters
NAME 1 - CHC16 Fr 105 Service_13C
EXPNO 4
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170917
Time 15.40
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 1536
DS 2
SWH 31250.000 Hz
FIDRES 0.476837 Hz
AQ 1.0485760 sec
RG 912
DW 16.000 usec
DE 18.00 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 125.8131152 MHz
NUC1 13C
P1 10.00 usec
PLW1 20.16400002 W
======== CHANNEL f2 ========
SFO2 500.3020012 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 7.99800014 W
PLW12 0.28119001 W
PLW13 0.17996000 W
F2 - Processing parameters
SI 32768
SF 125.8005195 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40
170.36
169.81
156.12
144.45
143.84
143.79
141.45
129.72
128.25
127.91
127.88
127.23
127.21
127.09
125.18
125.17
120.14
67.55
67.16
53.98
52.53
47.25
41.38
33.78
130
129
128
ppm
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 ppm

Methyl *N*-acetyl-*S*-(triphenylmethyl)-D-cysteinylglycinate (52): ^{1}H NMR



Methyl *N*-acetyl-*S*-(triphenylmethyl)-D-cysteinylglycinate (52): $^{13}C\{^{1}H\}$ NMR



Current Data Parameters
NAME 2 - CHC36 - 1H 13C - CHC36 Service (ch22320106)
EXPNO 4
PROCNO 1

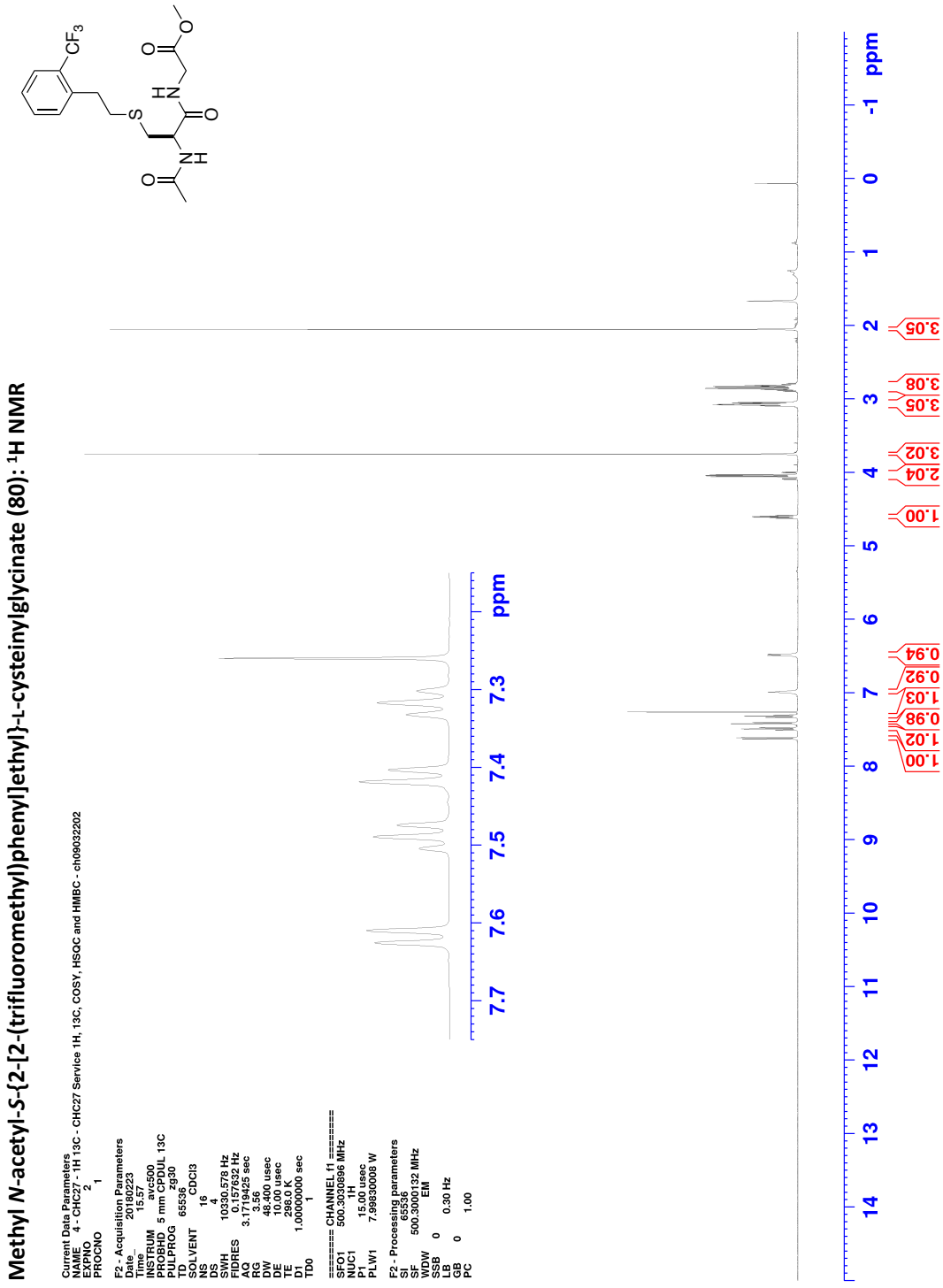
F2 - Acquisition Parameters
Date_ 20180603
Time 15.14
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 512
DS 2
SWH 31250.000 Hz
FIDRES 0.476837 Hz
AQ 1.0485760 sec
RG 912
DW 16.000 usec
DE 18.00 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1

======== CHANNEL f1 ========
SFO1 125.8131152 MHz
NUC1 13C
P1 10.00 usec
PLW1 20.18400002 W

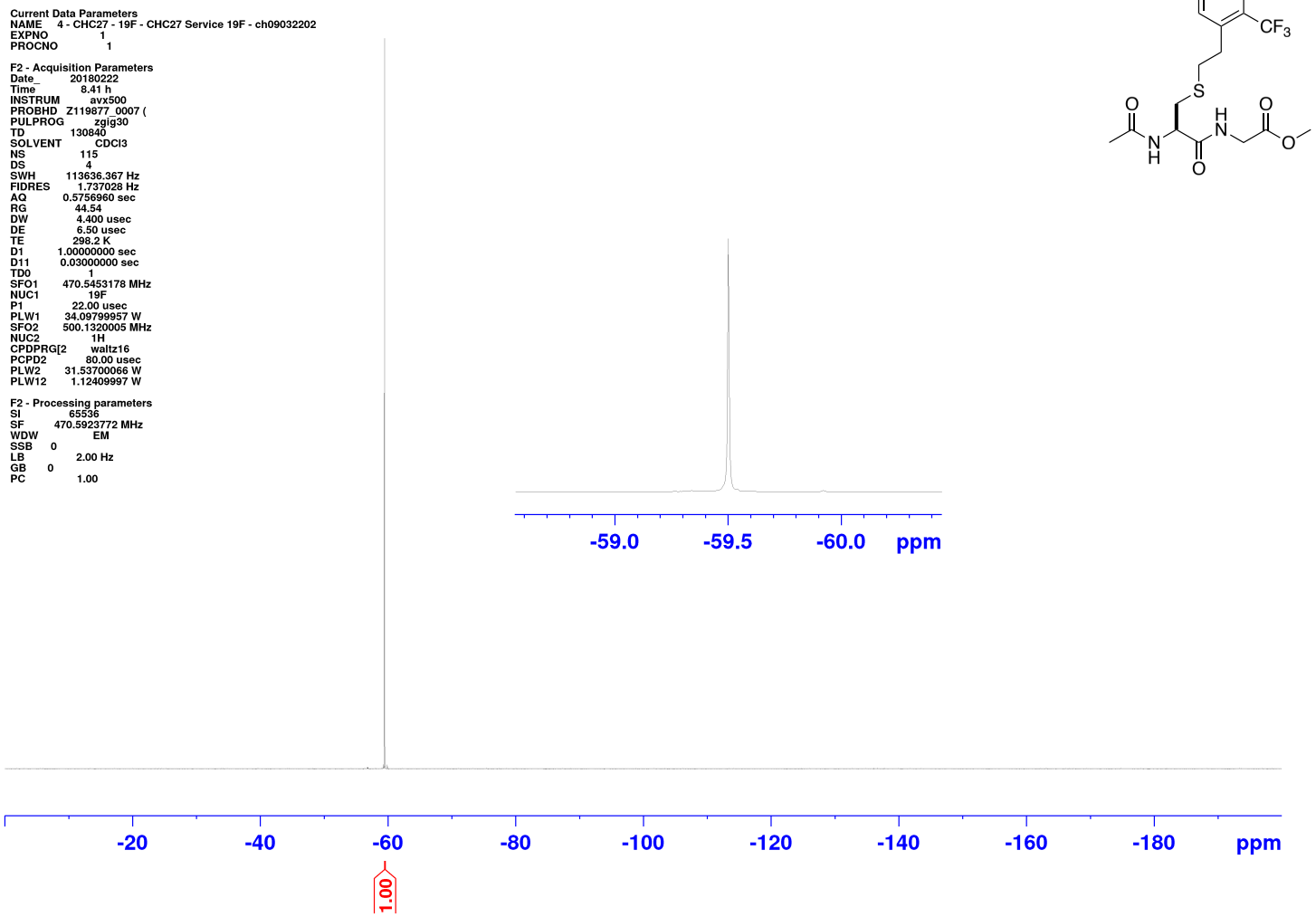
======== CHANNEL f2 ========
SFO2 500.3020012 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 7.98830000 W
PLW12 0.23513000 W
PLW13 0.17996000 W

F2 - Processing parameters
SI 32768
SF 125.8005206 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

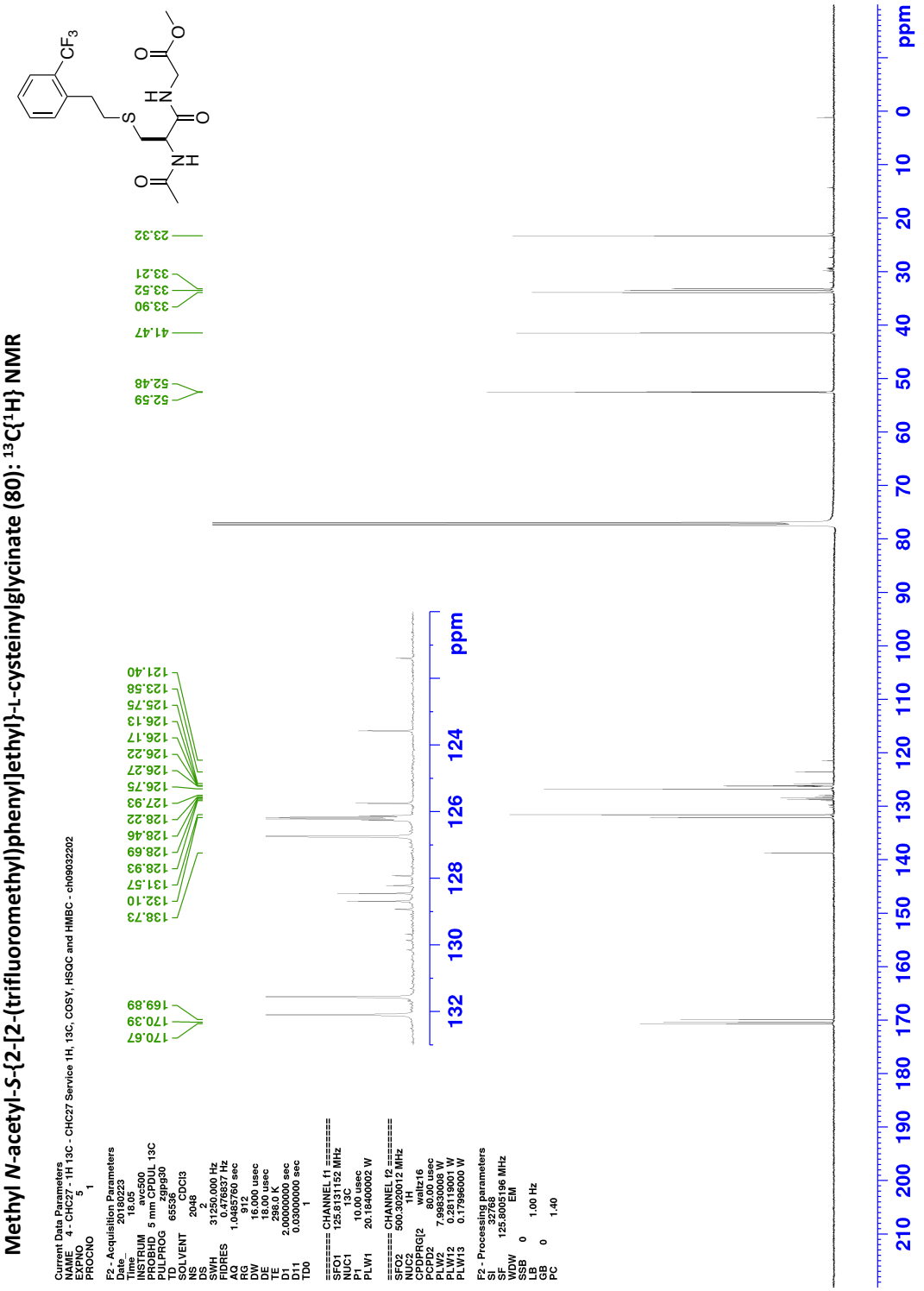
Methyl *N*-acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycinate (80): ^{1}H NMR



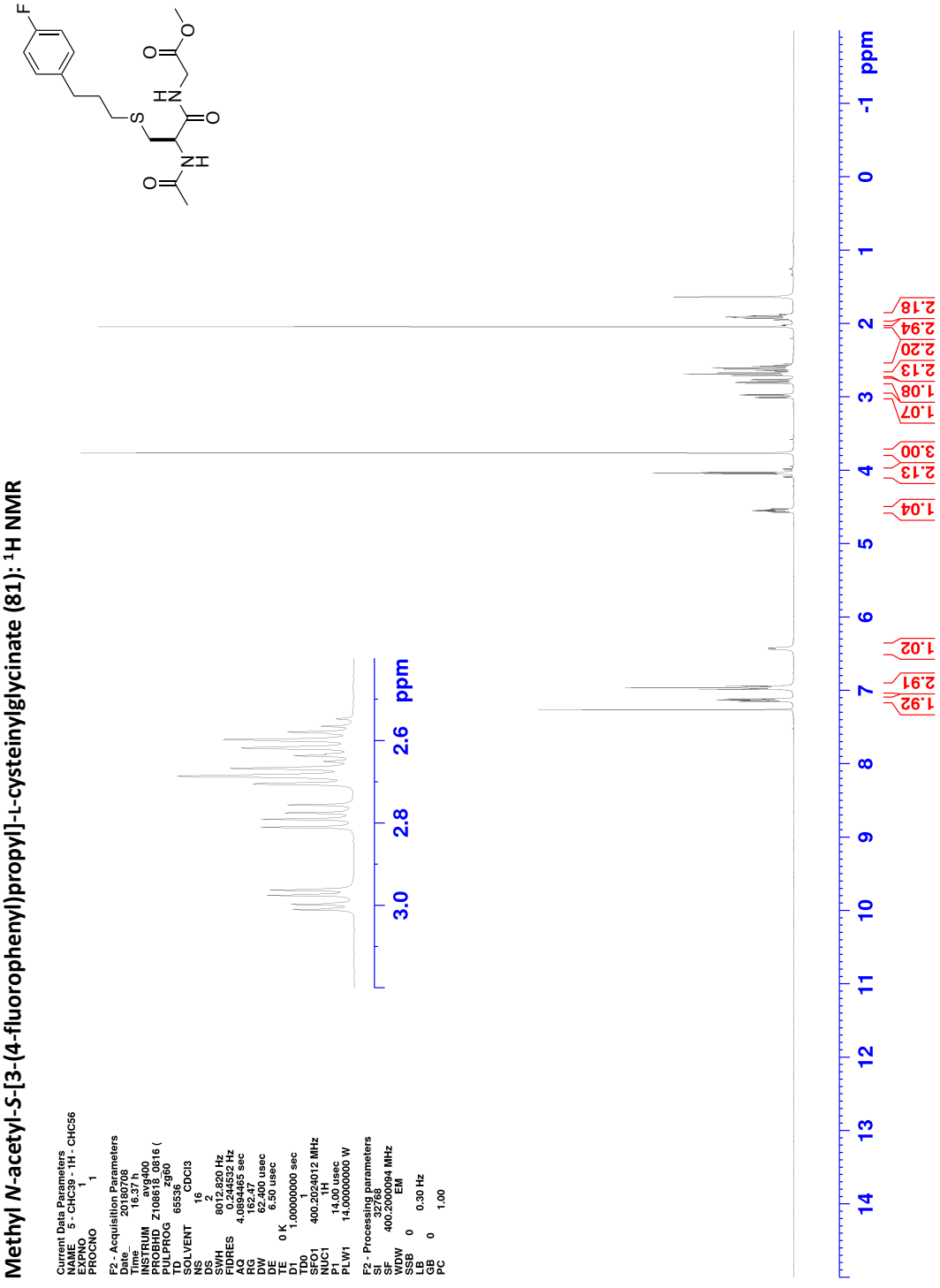
Methyl *N*-acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycinate (80): $^{19}F\{^{1}H\}$ NMR



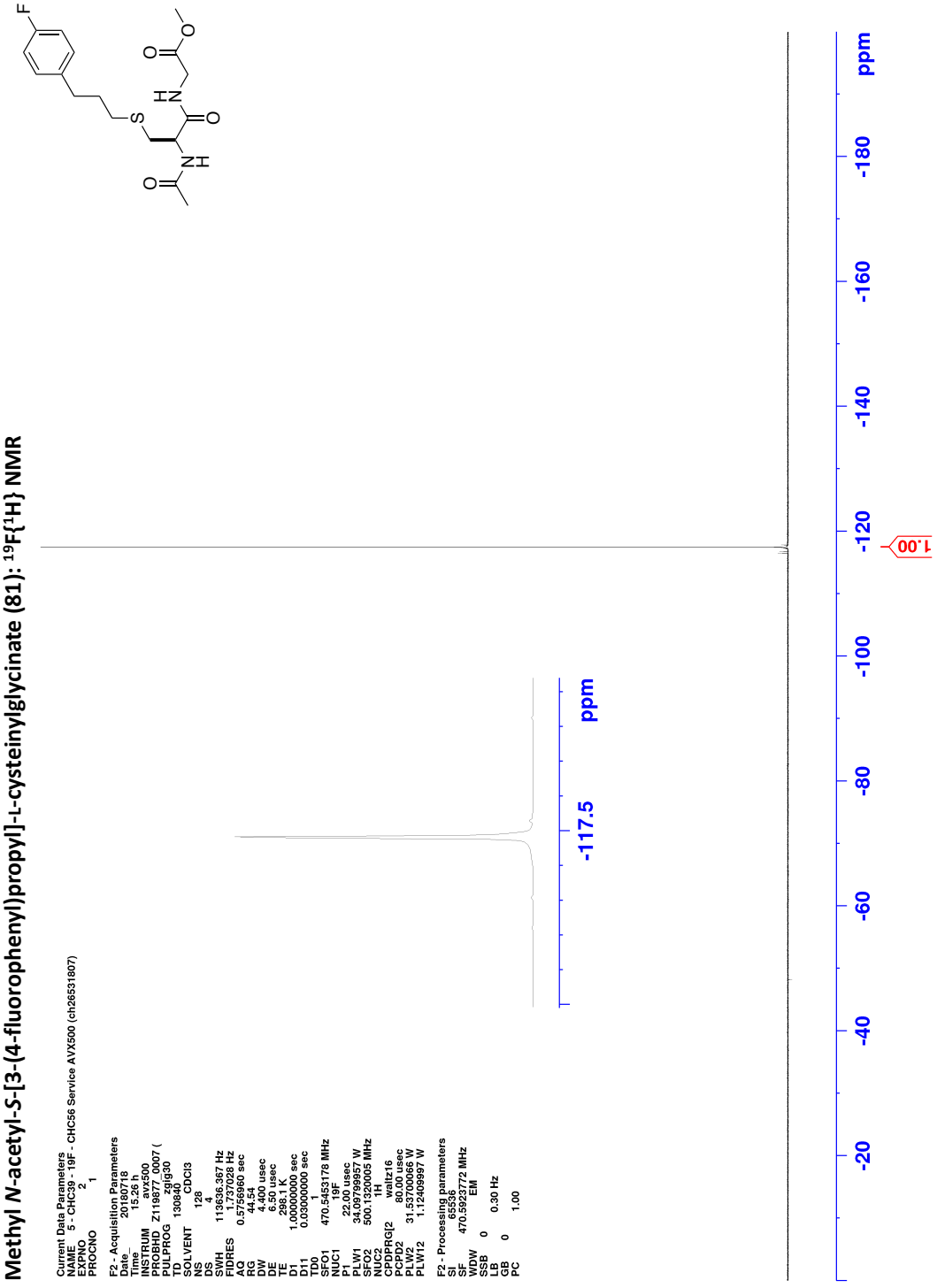
Methyl *N*-acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycinate (80): $^{13}C\{^{1}H\}$ NMR

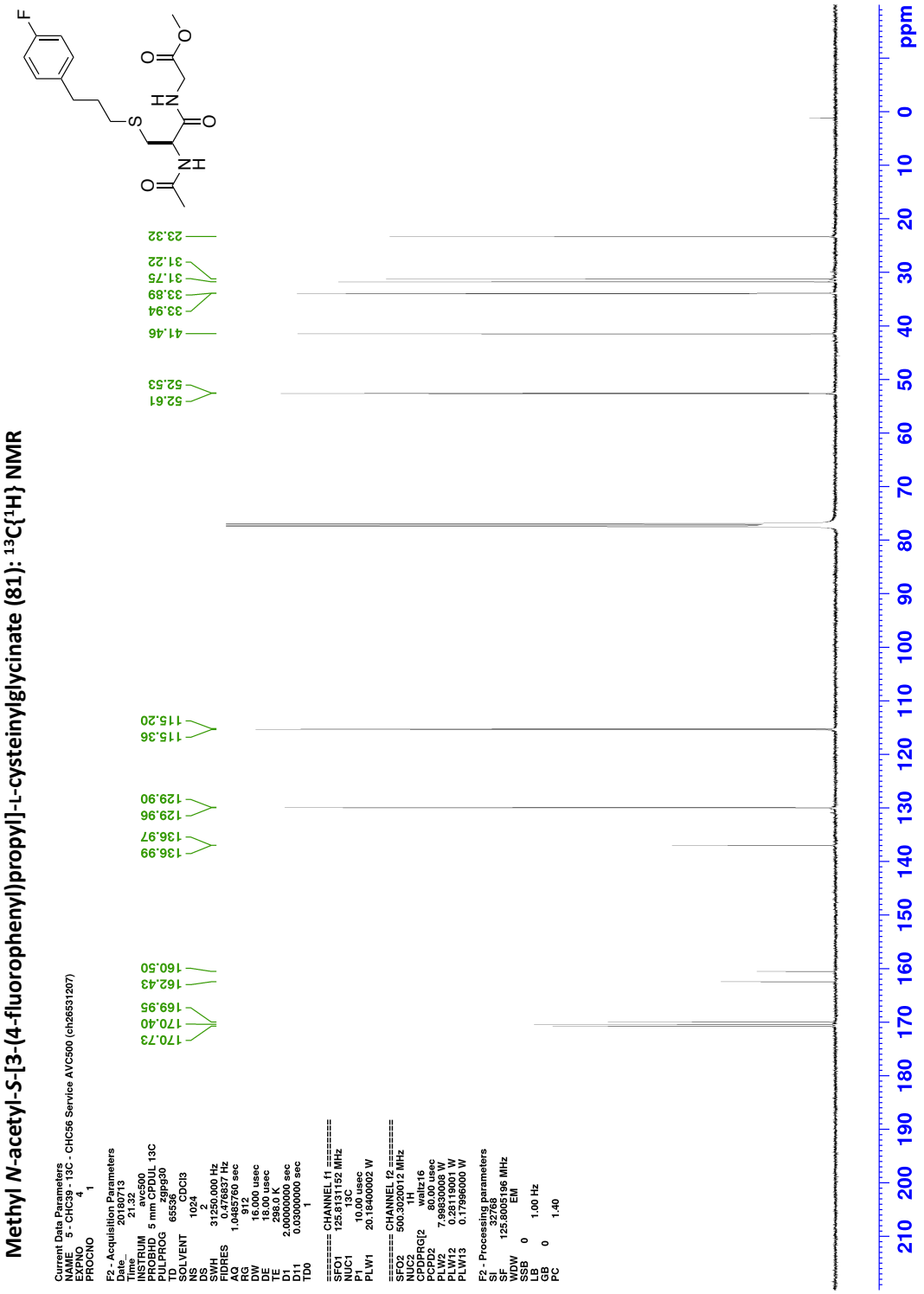


Methyl *N*-acetyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycinate (81): ^{1}H NMR

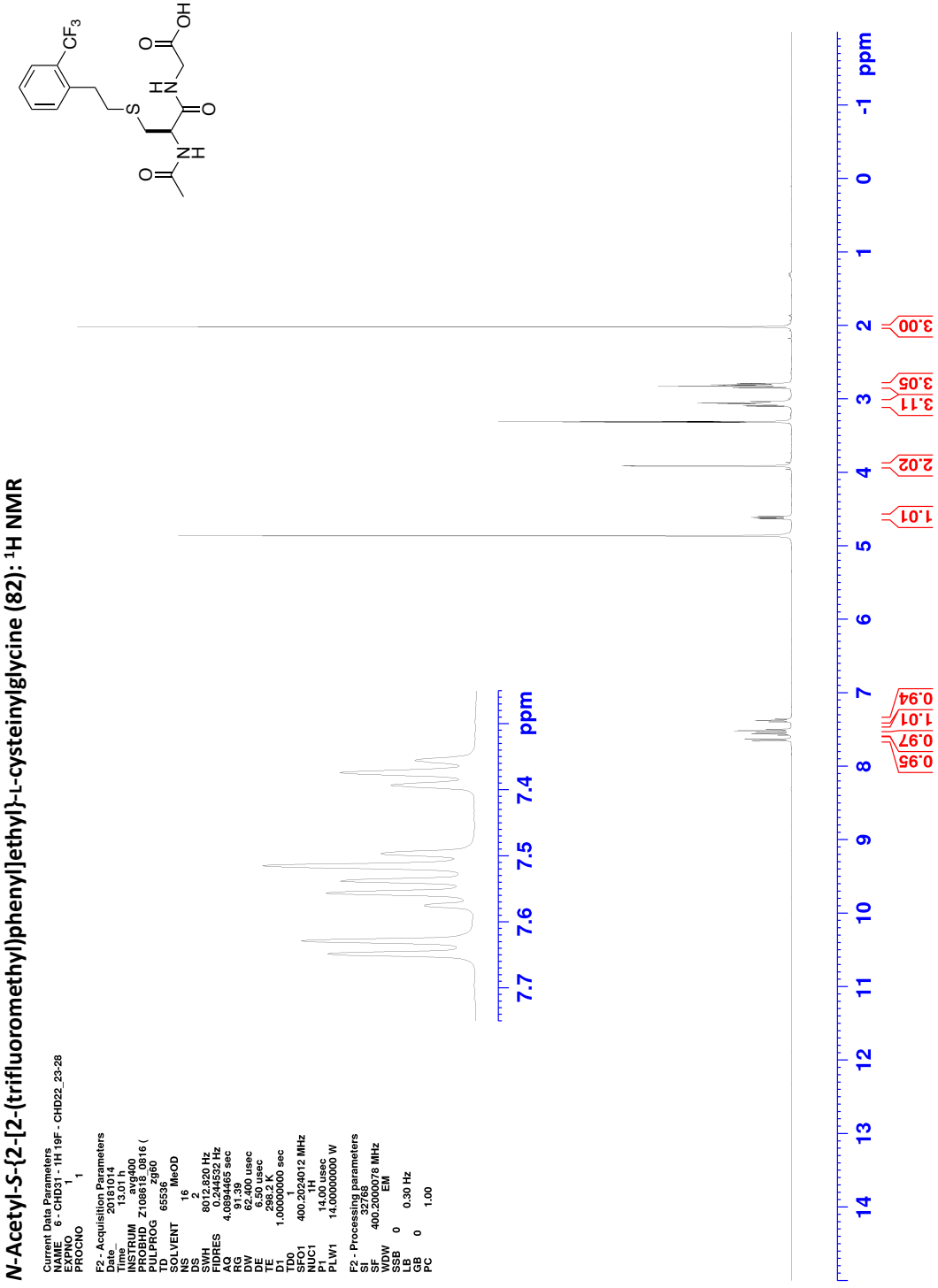


Methyl *N*-acetyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycinate (81): $^{19}F\{^{1}H\}$ NMR

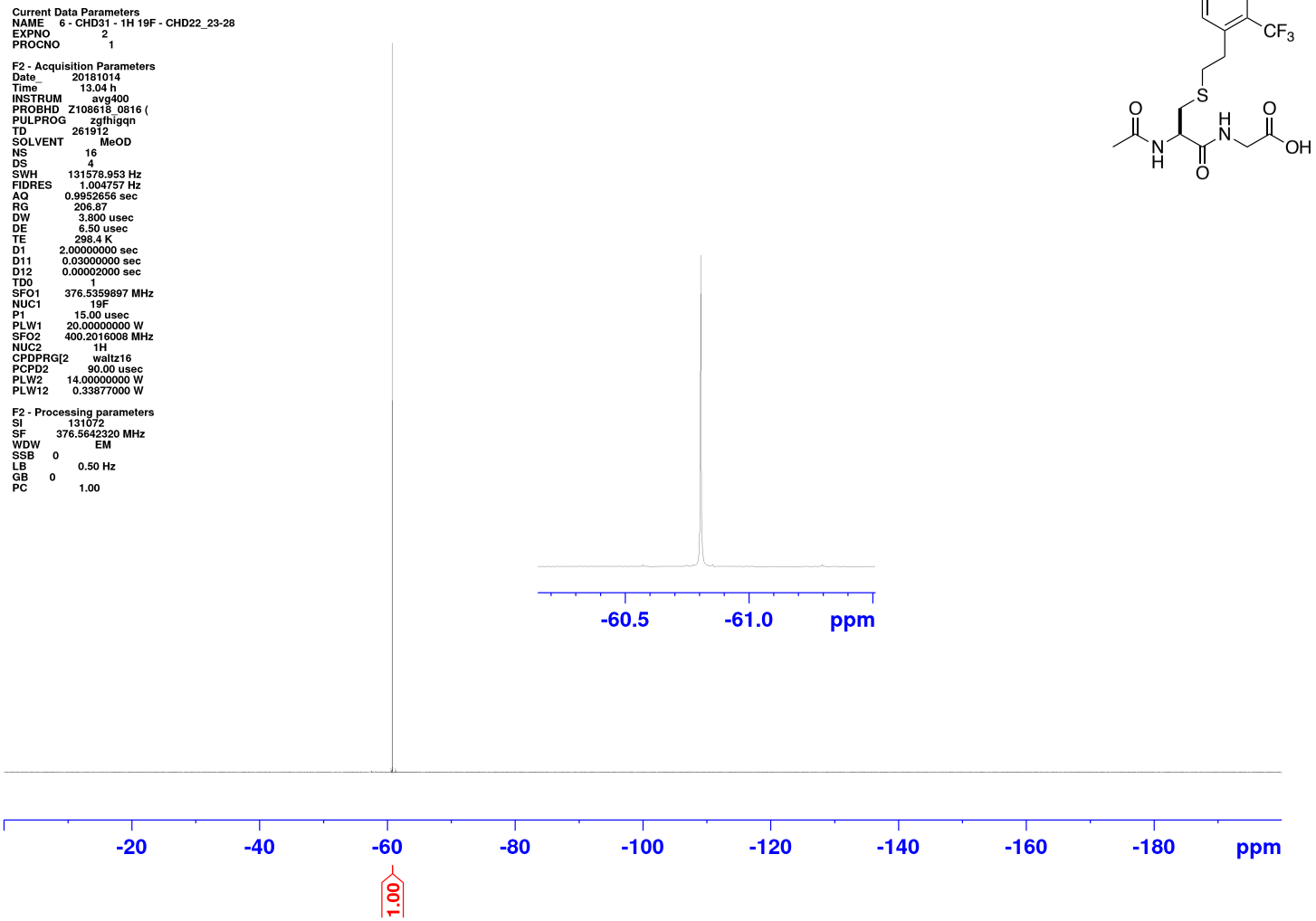



Methyl N-acetyl-S-[3-(4-fluorophenyl)propyl]-L-cysteinylglycinate (81): 13C{1H} NMR
Current Data Parameters
NAME 5 - CHC39 - 13C - CHC56 Service AVC500 (ch26531207)
EXPNO 4
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180713
Time 21.32
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 1024
DS 2
SWH 31250.000 Hz
FIDRES 0.476837 Hz
AQ 1.0485760 sec
RG 161
DW 16.000 usec
DE 18.00 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 125.81311152 MHz
NUC1 13C
P1 10.00 usec
PLW1 20.18400002 W
======== CHANNEL f2 ========
SFO2 500.3020012 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 7.99830008 W
PLW12 0.28119001 W
PLW13 0.17998000 W
F2 - Processing parameters
SI 32768
SF 125.8005196 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40
23.32
31.22
31.75
33.89
33.94
41.46
52.53
52.61
115.20
115.36
129.90
129.96
136.97
136.99
160.50
162.43
169.95
170.40
170.73
ppm

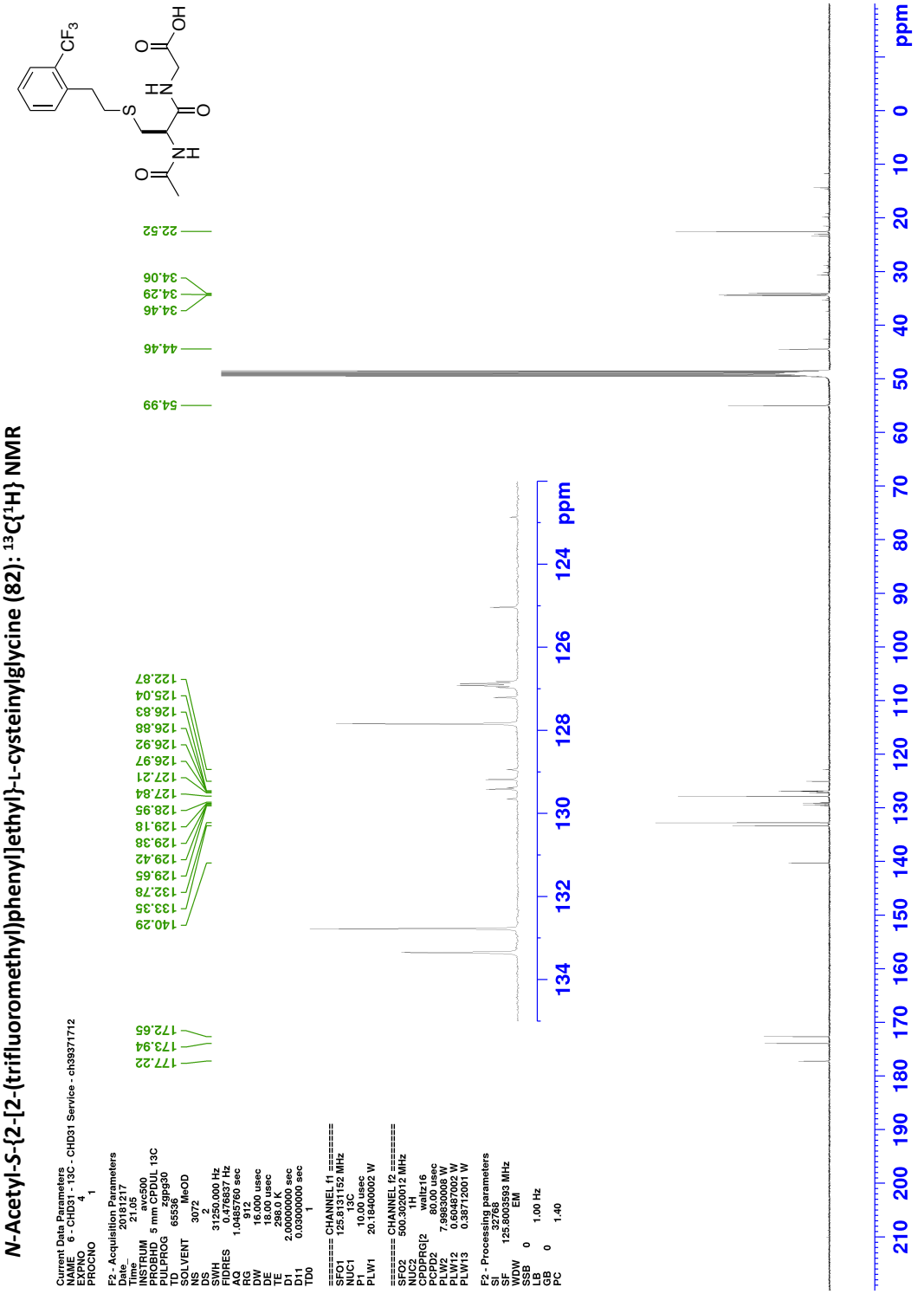
***N*-Acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycine (82): ^{1}H NMR**

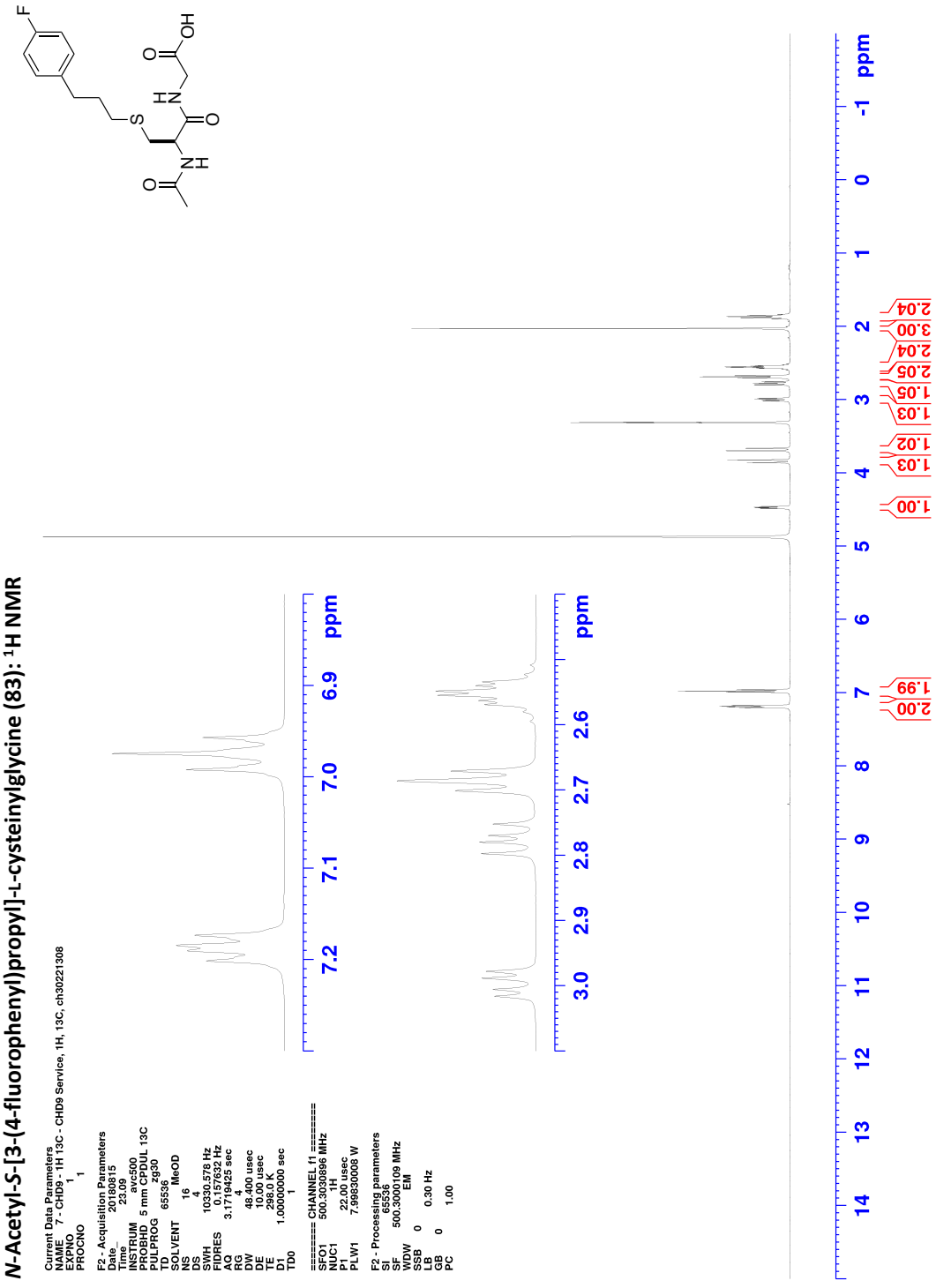


N-Acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycine (82): $^{19}F\{^{1}H\}$ NMR

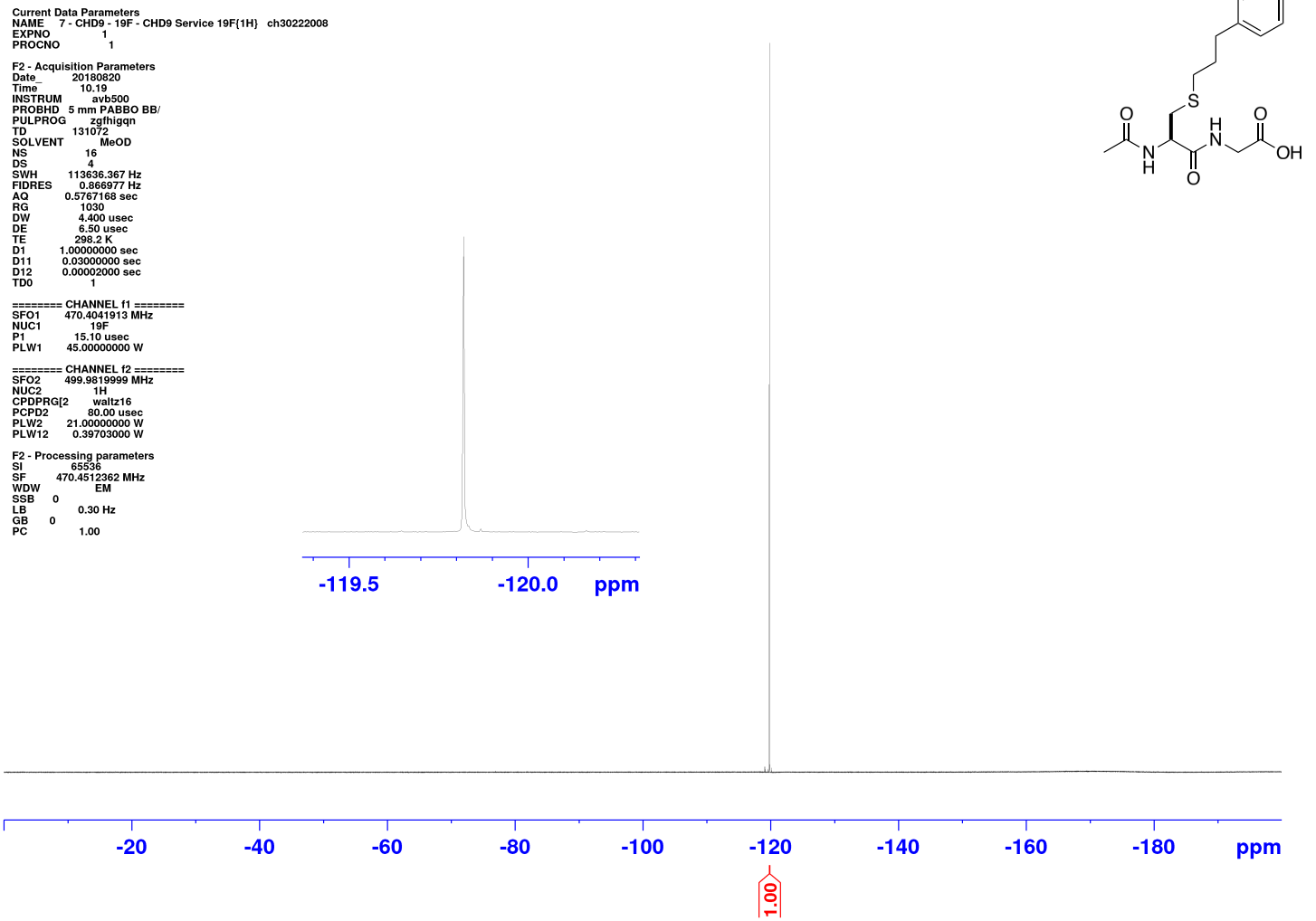


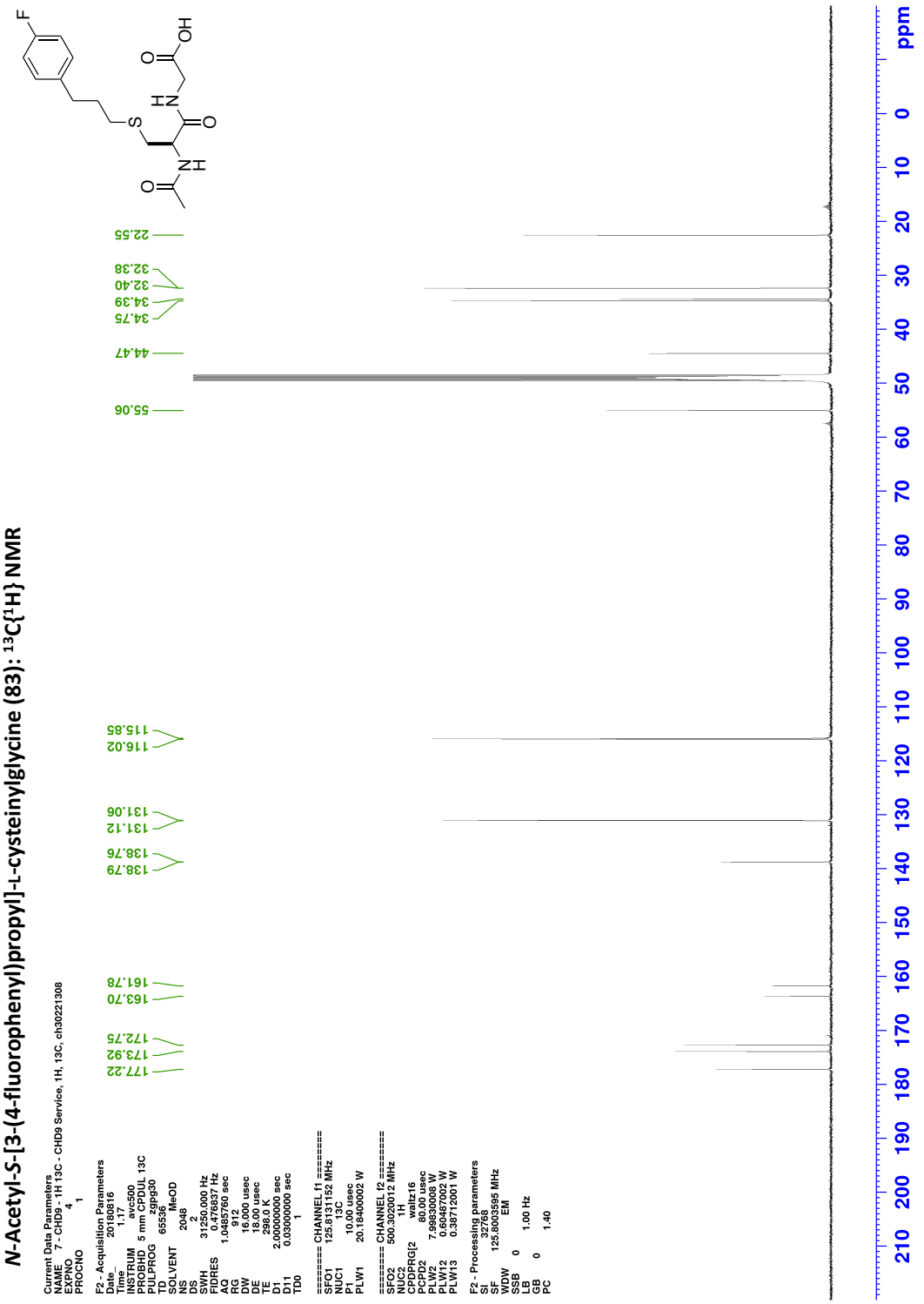
***N*-Acetyl-*S*-{2-[2-(trifluoromethyl)phenyl]ethyl}-L-cysteinylglycine (82): $^{13}C\{^{1}H\}$ NMR**




N-Acetyl-S-[3-(4-fluorophenyl)propyl]-L-cysteinylglycine (83): 1H NMR
Current Data Parameters
NAME 7 - CHD9 - 1H 13C - CHD9 Service, 1H, 13C, ch30221308
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180815
Time_ 23.09
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 4
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3000996 MHz
NUC1 1H
P1 22.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000109 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
F
O
OH
N
H
O
S
N
H
O
O
7.2
7.1
7.0
6.9
ppm
3.0
2.9
2.8
2.7
2.6
ppm
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
-1
ppm
2.00
1.99
1.00
1.03
1.02
1.03
1.05
2.05
2.04
3.00
2.04

N-Acetyl-*S*-[3-(4-fluorophenyl)propyl]-L-cysteinylglycine (83): $^{19}F\{^{1}H\}$ NMR




N-Acetyl-S-[3-(4-fluorophenyl)propyl]-L-cysteinylglycine (83): 13C{1H} NMR
Current Data Parameters
NAME 7 - CHD9 - 1H 13C - CHD9 Service, 1H, 13C, ch30221308
EXPNO 4
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180816
Time_ 1.17
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zgpg30
TD 65536
SOLVENT MeOD
NS 2048
DS 2
SWH 31250.000 Hz
FIDRES 0.476837 Hz
AQ 1.0485760 sec
RG 912
DW 16.000 usec
DE 16.00 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 125.8131152 MHz
NUC1 13C
P1 10.00 usec
PLW1 20.18400002 W
======== CHANNEL f2 ========
SFO2 500.0000012 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 7.99830008 W
PLW12 0.60487002 W
PLW13 0.38712001 W
F2 - Processing parameters
SI 32768
SF 125.8003595 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40
22.55
32.38
32.40
34.39
34.75
44.47
55.06
115.85
116.02
131.06
131.12
138.76
138.79
161.78
163.70
172.75
173.92
177.22
ppm