

Design of artificial metalloenzymes for C-C cross-coupling reactions



**A thesis submitted for partial fulfilment of degree of
Doctor of Philosophy in Chemical Biology
at the University of Oxford**

Mahima Sharma

Wolfson College

Trinity 2014

Dedicated to my grandparents

Author's Declaration

I declare the work presented in this thesis was carried out at the Chemistry Research Laboratory and Inorganic Chemistry Laboratory, University of Oxford, UK under the guidance of Professor Benjamin G. Davis. All the work is my own, except where otherwise stated, and has not been submitted for any other degree at this or any other university.

Mahima Sharma

Mahima Sharma

October 2014

Short Abstract for Bodleian Library Indexing

Design of artificial metalloenzymes for C-C cross-coupling reactions

Mahima Sharma

Wolfson College

A thesis submitted for degree of Doctor of Philosophy in Chemical Biology

Trinity 2014

Artificial metalloenzymes enable several non-natural chemical transformations by hosting a metal catalyst and providing activating microenvironments to the metal. New strategies to combine the complementary features of enzyme and transition metal catalysis are crucial for exploiting the pre-evolved selectivities of enzymes in chemical synthesis. In order to exploit the inherent selectivity of a serine protease Subtilisin (SBL) for C-C cross-coupling reactions, the interplay of natural substrate binding pockets and the nucleophilic machinery of the enzyme with the metal catalyst was central to our approach. A catalytically potent metal catalyst (palladium) was appropriately positioned to carry out oxidative Heck reaction between an alkene and arylboronic acid substrates in the active site of SBL. Further characterization of SBL-based artificial metalloenzymes and step-wise elucidation of the biotransformation pathway was accomplished through X-ray crystallography and mass-spectrometry studies. Such interlacing of the metal and the enzyme steps for cross-coupling reactions presents an attractive route for performing important organic reactions enzymatically.

Thesis Abstract

Design of artificial metalloenzymes for C-C cross-coupling reactions

Mahima Sharma, Wolfson College

A thesis submitted for degree of Doctor of Philosophy in Chemical Biology

Trinity 2014

Artificial metalloenzymes are developed by exploiting known ligand affinities or covalent/dative interactions to embed non-native metal catalysts into protein (or DNA) scaffolds. Through protein engineering and activity-optimization, these artificial metalloenzymes have been developed to demonstrate a wide range of chemical reactions. Complementary and attractive features of enzyme and transition metal catalysis can be combined by taking advantage of the evolved selectivities of natural enzymes and the versatility of metal catalysts. The work in this thesis was focused on the design and construction of artificial metalloenzymes based on a serine protease Subtilisin (SBL) and metal catalyst (palladium) by employing covalent chemical modification strategies. Precise positioning of metal in the enzyme active site enabled a natural serine protease to foster synergy with palladium - a biologically absent metal that is otherwise a potent reaction centre for chemical cross-coupling reactions.

The site-selective reaction at dehydroalanine (Dha) or cysteine (Cys) residues enabled chemical installation of bidentate metal binding amino acids into SBL by using two different strategies. Palladium coordination with these unnatural amino acids present at the base of S1' pocket of SBL was demonstrated. Selective and stoichiometric metal interaction of the SBL-based artificial metalloenzymes was verified by several techniques including LC-MS, ICP-MS and microPIXE analyses. The inherent selectivity of the substrate binding hydrophobic pockets and nucleophilic interactions of SBL was used to localise the substrates. Oxidative Heck cross-coupling reaction between a tethered alkene substrate and several arylboronic acids was demonstrated in the enzyme active site. The step-wise elucidation of this biotransformation pathway was accomplished by mass spectrometry studies of the representative intermediates. X-ray crystal structures of the substrate-bound enzyme and apoenzyme species highlighted the key enzyme-substrate and enzyme-ligand interactions. X-ray crystallography was used to carry out a systematic study of the intermediates of the cross-coupling reaction. The precise location of metal in the active site was verified by X-ray crystal structure of SBL-metalloenzyme, using nickel as a metal surrogate.

The rational design of β -lactamase-based artificial metalloenzymes for Suzuki-Miyaura cross-coupling reaction was undertaken. Active site engineering of β -lactamase (AmpC) using chemical and genetic means enabled the generation of a library of AmpC-variants that were modified to corresponding metal-free apoenzymes.

Overall, this work presents a proof-of-concept for utilising naturally evolved selectivities of an enzyme and chemical reactivities of a metal to facilitate localisation of substrates in the enzyme active site. Such bio-cooperative metal catalysis of important organic reactions can be used to device new synthetic routes, especially for C-C cross-coupling reactions.

Table of Contents

Acknowledgements	ix
Publication list	xi
Selected Abbreviations	xii

Chapter 1. An Introduction to artificial metalloenzymes	1
--	----------

1.1	Background	1
1.2	Design of artificial metalloenzymes	2
1.3	Tools for design and optimization of artificial metalloenzymes	7
1.4	Existing metallo-bioconjugates: scope of artificial metalloenzymes	14
1.5	Palladium catalysed C-C cross-coupling reactions	24
1.6	Subtilisin: model serine protease	28
1.7	Aim of the thesis	31
1.8	References	33

Chapter 2. Design and construction of the SBL-based artificial metalloenzymes	40
--	-----------

2.1	Introduction	40
2.2	Results and discussion	42
2.2.1	Design of SBL-based metalloenzymes	42
2.2.2	Construction of SBL-based metalloenzymes	47
2.2.3	Characterization of SBL-based metalloenzymes	63
2.3	Conclusions	69

2.4	References	72
2.5	Experimental procedures	74
2.6	NMR Spectra	121

Chapter 3. Oxidative Heck cross-coupling reaction using SBL-based artificial metalloenzymes

3.1	Introduction	135
3.2	Results and discussion	138
3.2.1	Design of substrates and modelling studies	138
3.2.2	Enzyme-substrate interactions	142
3.2.3	Engineered cycle	145
3.2.4	Control reactions	158
3.2.5	Reloading	159
3.2.6	Substrate scope of oxidative Heck cross-coupling reaction	162
3.2.7	Investigation of pH dependence of cross-coupling reaction	165
3.2.8	Compulsory order of palladation and on-protein cross-coupling reaction	166
3.3	Conclusions	168
3.4	References	170
3.5	Experimental procedures	173
3.6	NMR Spectra	267

Chapter 4. Structural characterization of SBL-based artificial metalloenzymes

4.1	Introduction	279
-----	--------------	-----

4.2	Results and discussion	282
4.2.1	Target structures	282
4.2.2	Protein modifications and characterization	283
4.2.3	Initial Screening	286
4.2.4	Optimization of crystallization conditions	289
4.2.5	Structure descriptions	293
4.2.6	Microbeam Proton Induced X-Ray Emission	305
4.3	Conclusions	307
4.4	References	309
4.5	Experimental procedures	311
 Chapter 5. Design of β-lactamase-based artificial metalloenzymes		336
5.1	Introduction	336
5.2	Results and discussion	337
5.2.1	Design of system	337
5.2.2	Subcloning and site-directed mutagenesis	343
5.2.3	Protein expression and purification: optimization	346
5.2.4	Chemogenetic design: ligand incorporation	351
5.2.5	Metal complexation	354
5.1	Conclusions	355
5.2	References	356
5.3	Experimental procedures	358
 Chapter 6. Summary		387

Acknowledgements

By the way of writing this note, I thank all the people who directly or indirectly formed my support system during my DPhil study. First and foremost, I would like to express my sincere gratitude to my parents, Vijay Sharma and Alka Sharma, the JN TATA Endowment trust, Wolfson College and the University of Oxford for the financial support offered to me, without which this DPhil wouldn't be possible.

I sincerely thank Prof. Benjamin G. Davis for giving me the opportunity to work on this challenging yet exciting project and I appreciate his help and guidance. I would like to acknowledge the Department of Chemistry for providing a great infrastructural set-up and an excellent support staff, both of which contributed hugely in expediting the work presented in this thesis. I appreciate the prior work done by Dr. Alba-Díaz Rodríguez and Dr. Anca Pordea on this research project and I am grateful for their guidance and support throughout my DPhil. I thank my collaborators in Oxford and York for all the important scientific discussions that have contributed in shaping up of this thesis. I would like to express my gratitude to Prof. Gideon J. Davies and Wendy Offen for their contributions in X-Ray crystallography studies reported in this thesis. I also thank my college advisor, Prof. Christina Redfield for her help and encouragement throughout my degree.

I greatly appreciate the time and effort that was put in for proof-editing this thesis – a special thanks to Adhaar Gour, Maria Espling, Balakumar Vijayakrishnan, Mitul Patel, Wendy Offen, Alba-Díaz Rodríguez, Matt Schombs, Chris Spicer, Sebastian Galan, Sylvain Royer, Mercedes, Chris Serpell, Prerna Sharma, Macarena Sanchez-Navarro, Yiqun Geng, Mark Wormald, Max Crispin and Tom Parsons. I would also like to extend my sincere gratitude to my DPhil examiners, Prof. F.A. Armstrong and Prof. Dr. Donald

Hilvert for their insightful comments and suggestions about this work, as well as for a memorable *viva voce*.

Someone recently said to me, ‘..you may not realise this, but life is bigger than your DPhil even while you are doing it.’ I would like to acknowledge all the members of the Davis group, past and present, for their help throughout my study and for contributing in making my time at Oxford most memorable. I also thank my sisters Prerna Sharma, Nayency Sharma, Sonia Sharma, and my beloved Adhaar Gour for their continuous support and encouragement throughout these years. I lovingly thank Prerna Sharma, Milly Emmanuel, Adhaar Gour and Hua Wang for the helpful, stress-busting trips and for pushing me take a break from the DPhil work when most required. Finally, I can never thank enough my parents and my grandfather J.N. Sharma for always keeping me in their prayers and waiting most eagerly for the day I would complete my DPhil degree.

Mahima Sharma

Mahima Sharma

27th March 2015

Department of Chemistry

University of Oxford

Publication list

Cooperative bio-metallic selectivity in a tailored protease enables creation of a C–C cross-coupling “Heck-ase”

M. Sharma, A. Díaz-Rodríguez, W. A. Offen, M. E. Palm-Espling, A. Pordea, M. R. Wormald, E. F. Garman, M. A. McDonough, G. J. Davies, B. G. Davis; *manuscript in preparation*

Selected Abbreviations

°C	degree Celsius
aq.	aqueous
ACN	acetonitrile
BCA	bicinchoninic protein assay
BSA	bovine serum albumin
calc.	calculated
CD	circular dichroism
CDCl ₃	deuterated chloroform
D ₂ O	deuterated water
Da	Dalton
DCM	dichloromethane
DNA	deoxyribose nucleic acid
ES	electrospray
<i>et al.</i>	<i>et alia</i> (Latin: and others)
<i>etc.</i>	<i>et cetera</i> (Latin: and the rest)
EtOAc	ethylacetate
Equiv	equivalents
FA	formic acid
H ₂ O	water
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl β-D-1-thiogalactopyranoside
<i>J</i>	NMR coupling constant (Hz)

k	kilo
LB	Luria Bertani
LC-MS	liquid chromatography-mass spectrometry
LRMS	low resolution mass spectrometry
MeOH	methanol
MHz	megahertz
min	minute
mL	milliliter
MS	mass spectrometry
NMR	nuclear magnetic resonance
PDB	protein data bank
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
ppm	part per million
quant.	quantitative
R_f	retention factor
rpm	revolution per minute
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
THF	tetrahydrofuran

Chapter 1 Introduction to artificial metalloenzymes

1.1 Background

The use of natural catalysts, such as enzymes, to perform useful chemical transformations underpins the idea of biocatalysis in the context of organic synthesis.¹ Natural enzymes provide biologically active microenvironments for efficient catalysis. High catalytic power and substrate specificity are the two innate properties that make natural enzymes powerful catalysts. The chiral make-up of enzymes can confer chemo-, regio- and/or enantioselectivity on the reaction, while providing environmentally benign alternative synthetic methods. Because of these attractive features, enzymes are now increasingly used in the synthesis of organic compounds.²⁻⁵ In addition to their natural function, enzymes sometimes also exhibit promiscuous activities which can be further optimized for chemical synthesis.⁶ Thus, natural enzymes can be engineered to adapt new activities by employing protein engineering techniques.⁷⁻⁹

The use of enzymes for organic synthesis can often be limited due to their high substrate and reaction specificities. In contrast to this, the synthesis of compounds using chemical catalysis is typically wider in substrate breadth and often relies on transformations with no equivalent in nature. In this regard, homogeneous and enzyme catalysis can be combined to carry out enantioselective transformations. Alternative chemical functions can be introduced in enzymes by the use of non-native metals,¹⁰ thereby offering promising opportunities to transfer important organic transformations to the biological world.^{11,12} In an effort to provide alternative synthetic routes and add biomolecular scaffolds to

chemists' toolbox, several enantioselective transformations have been reported using artificial metalloenzymes in the past decade.¹¹⁻¹⁵ Owing to the advances in protein engineering and computational strategies combined with artificial evolutionary fine-tuning, several selective artificial metalloenzymes have been created.¹⁶ Key concepts in the field of artificial metalloenzymes have been reviewed in this chapter under three major headings to discuss the design of artificial metalloenzymes, the different performance optimization strategies and finally the current scope of the reported artificial metalloenzymes. A brief introduction to cross-coupling reactions and the model protein scaffold used for the work described in this thesis is given at the end of this chapter.

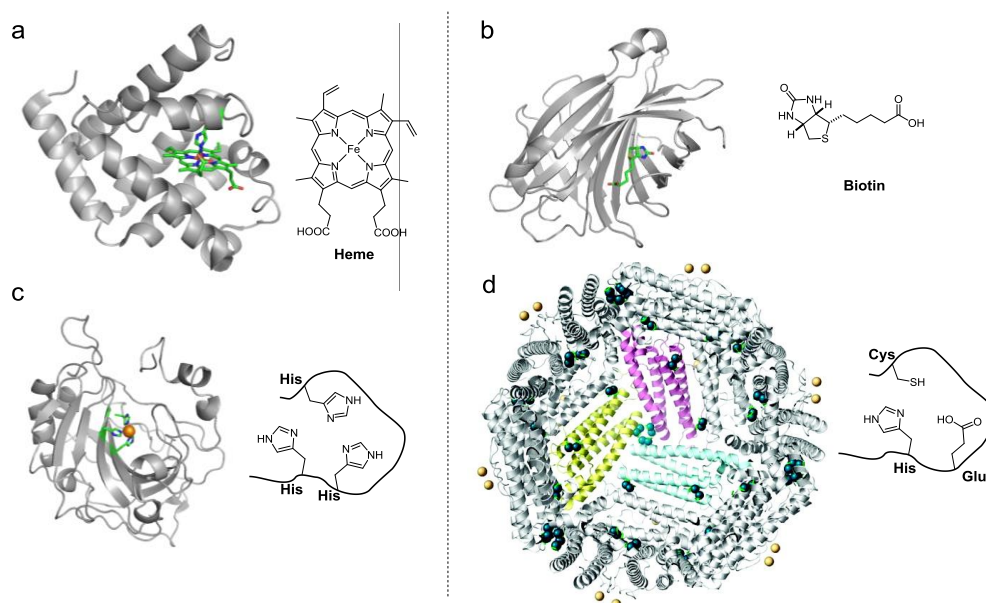
1.2 Design of artificial metalloenzymes

The basic design of artificial metalloenzymes depends on the choice of model biomolecular scaffold and the strategy to anchor the metal complex at a suitable pocket or site on the biomolecule, as discussed in this section.

1.2.1 Choice of scaffold

Late transition metal complexes (Cu, Rh, Ru, Ir, Pd) that form the basis of homogeneous catalysis have been widely incorporated into proteins mainly because of their catalytic versatility. Often, some protein engineering is required to install non-natural metal centres into proteins using organometallic ligands/cofactors. For this reason, a robust biomolecular scaffold which can be easily manipulated by employing different molecular biology techniques is the primary requirement while designing artificial metalloenzymes. It is important to consider the stability of the biomolecular scaffold in the presence of a non-native metal. The aqueous and air stability of these organometallic complexes is also a desirable feature. From a practical standpoint, a ready access to good amounts of both the

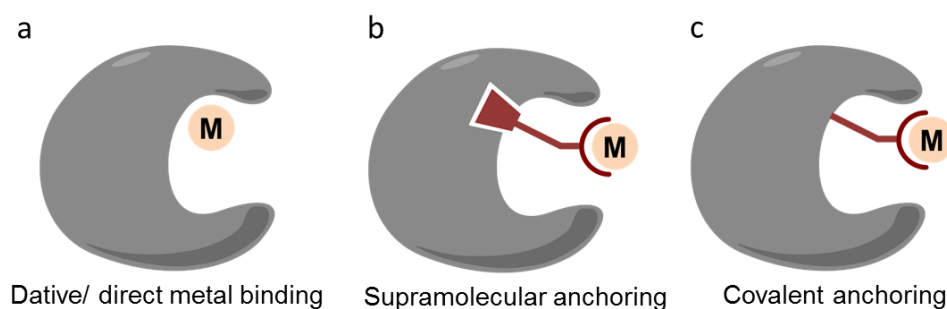
biomolecular scaffolds and the metal complexes is essential for future optimization and screening purposes. The commonly used protein scaffolds for designing artificial metalloenzymes comprise bovine serum albumin, carbonic anhydrase, myoglobin, avidin, streptavidin and proteases like papain, chymotrypsin etc. The structures of some of these common protein scaffolds and their natural ligands/cofactors are shown in Scheme 1-1. The affinity of natural binding pockets for these ligands can be effectively employed to install metal centres in proteins. Large vacant spaces of protein cages like ferritin have also been engineered to deposit non-native metals in the interior of these macromolecules. In addition to this, organometallic catalysts can be bound to polynucleotides like DNA, which can offer opportunities to explore asymmetric catalysis.¹⁷ The classification of the resulting artificial metalloenzymes based on the mode of anchoring of the metal catalyst is discussed in Section 1.2.2.



Scheme 1-1. Ribbon diagram of commonly used scaffolds for the design of artificial metalloenzymes (a) Myoglobin (b) Streptavidin (c) Carbonic anhydrase (d) Ferritin; (a-c) taken from PDB ID: 1MBN, 1MK5, 1CA2, respectively; (d) Reprinted (adapted) with permission from Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, *130*, 10512. Copyright (2008) American Chemical Society.

1.2.2 Anchoring strategies

The existing artificial metalloenzymes can be classified based on the mode of installation of metal complexes. Three different anchoring strategies have been used previously, namely non-covalent direct metal binding, supramolecular anchoring and covalent anchoring (Scheme 1-2).



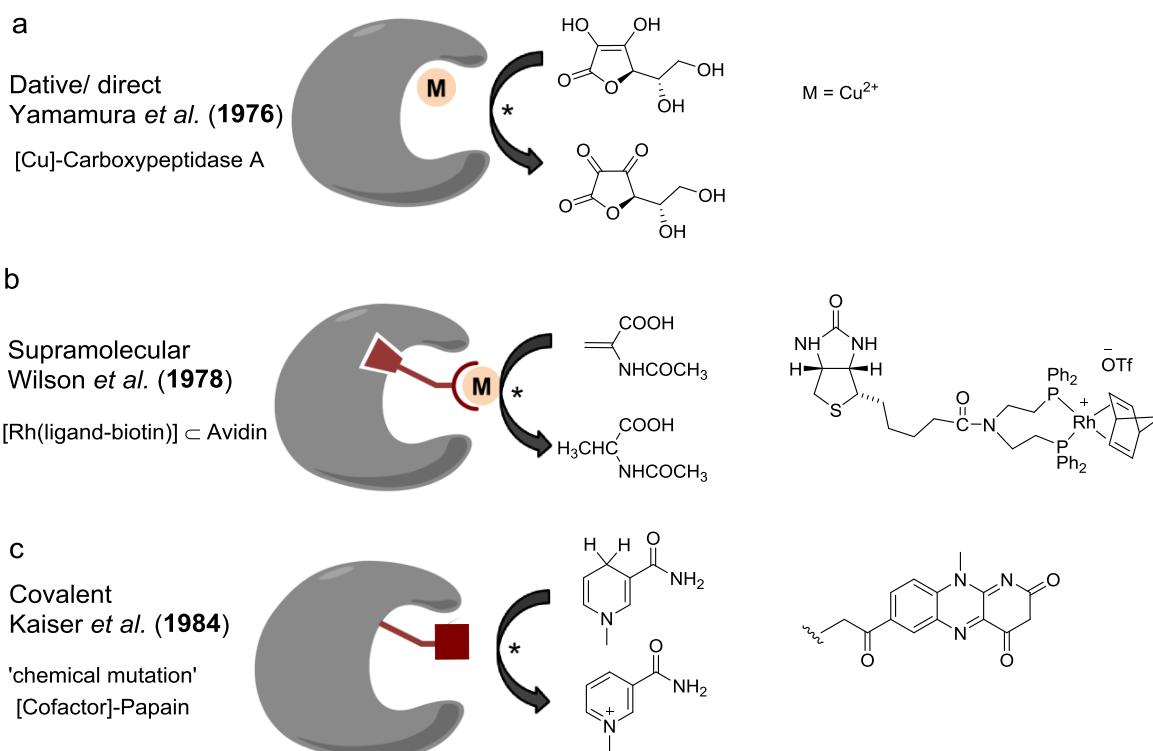
Scheme 1-2. Different modes of anchoring of metal complexes to proteins

Dative/direct metal binding refers to the direct coordination of metal ions to coordinating amino acid residues in the protein scaffolds. Replacement of the native metal ion with another metal catalyst can introduce novel activities in natural enzymes. In 1976, Kaiser and co-workers demonstrated this approach for the first time by replacing zinc in a natural carboxypeptidase A (CPA) with copper (Scheme 1-3a).¹⁸ This completely eliminated the hydrolytic activity of the enzyme (only some thiolesterase activity was retained). Notably, the coordination sphere of installed copper (tetrahedral) differed from the Zn-containing enzyme (with square planar geometry), while the overall binding site of the native enzyme was mainly preserved. This new CPA•Cu(II) enzyme functioned as an oxidase and catalysed the oxidation of ascorbic acid to dehydroascorbic acid.

Dative metal binding can be further differentiated into three approaches that have been employed, namely, (i) metal replacement at pre-evolved binding sites, (ii) *de novo* design

of metal binding site, (iii) non-specific metal binding by chelating thiol, carboxylate and imidazole groups on proteins (other than naturally evolved metal binding sites).¹⁵

Supramolecular anchoring relies on non-covalent interactions such as hydrogen-bonding, van der Waals forces, hydrophobic and electrostatic interactions, which can together lead to specific protein-cofactor affinities. Such strong interaction of biotin with Strept(avidin) ($K_d = 10^{-12} - 10^{-15}$ M) has been successfully employed in the design of artificial metalloenzymes. Whitesides and co-workers demonstrated enantioselective hydrogenation of α -acetamidoacrylic acid using Rh-containing biotin-avidin protein complex.¹⁹ Although only 33-44% ee was achieved, this provided a proof-of-concept study for asymmetric induction using well-defined three-dimensional environment of proteins (Scheme 1-3b). Our knowledge about the biotin-avidin system since Whitesides seminal report has deepened and numerous chemical transformations have been reported using artificial metalloenzymes based on biotin-Strept(avidin) interactions.²⁰ The extensive optimization of this system has been accomplished by Ward and co-workers (discussed in section 1.4), extending this technology to catalyse enantioselective reactions with no equivalent in synthetic chemistry.²¹



Scheme 1-3. Early reports on different modes of anchoring of unnatural cofactors or metals to protein scaffolds

Covalent anchoring of metal cofactors to proteins is a powerful technique for site-selective installation of metal centres. This approach relies on installation of a uniquely reactive amino acid (often cysteine) which can subsequently be modified to install a metal complex. By employing a combination of site-directed mutagenesis and chemical modification techniques, this approach circumvents the restriction of using specific protein-cofactor interactions for installing metal complexes. In other words, theoretically any protein scaffold can be used as a model scaffold for designing artificial metalloenzymes. In 1984, Kaiser and co-workers demonstrated the use of covalent approaches to generate hybrid catalysts by installing a redox active co-factor in a cysteine protease, papain, thereby converting it into an oxido-reductase (Scheme 1-3c).^{22,23} The single cysteine in the active site of papain was alkylated with brominated flavins.

Importantly, the substrate binding properties were preserved to retain natural enzyme-like properties in the hybrid catalysts. It was speculated that the substrates interacting closely with enzyme binding pockets may exhibit larger rate accelerations and stereospecificity may be achieved.²³ Although no metal was involved, this study proved important to demonstrate the huge potential of covalent modification strategies for tailoring existing proteins to perform new functions. Indeed, several artificial metalloenzymes have now been reported by employing different covalent modification strategies.¹⁵

1.3 Tools for design and optimization of artificial metalloenzymes

Major emphasis to-date has been to confer stereo-selectivity on reactions by the way of fine-tuning the proteinaceous (or biomolecular) environment around the metal to complement or enhance the metal activity. To achieve this, several design and performance optimization tools can be employed.

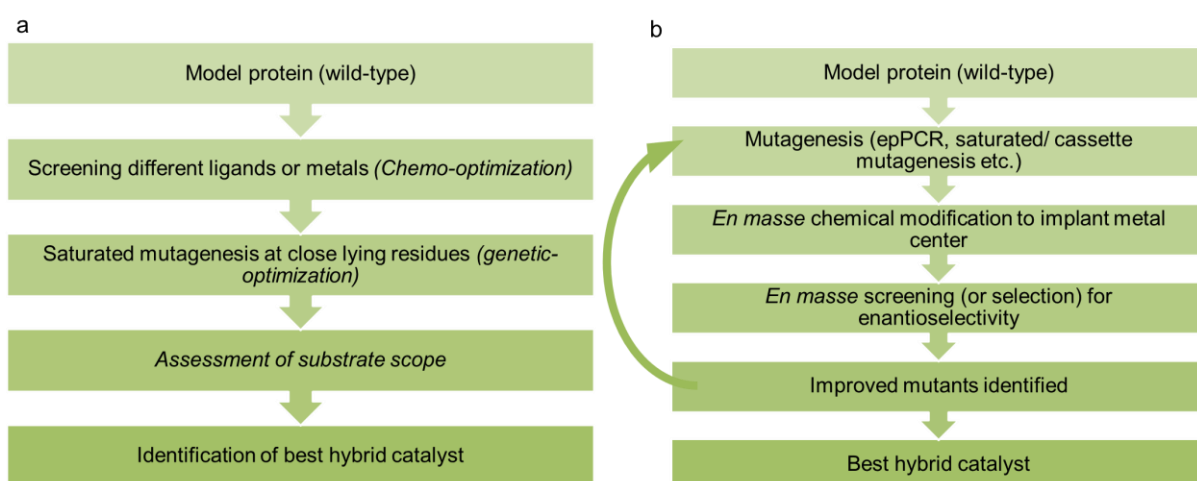
1.3.1 Chemogenetic optimization

The activity and selectivity of the artificial metalloenzymes can be improved by using a combination of chemical and genetic means. In particular, mainly a two-step optimization procedure is used where the first coordination sphere interactions of the metal (that greatly influence the activity) can be altered by chemical methods and the secondary coordination sphere interactions of the enzyme (comprising weak, non-bonding contacts that influence selectivity) can be tuned by genetic means (Scheme 1-4a). This has been successfully demonstrated using the biotin-Strept(avidin) system.²⁴⁻²⁸ As the strong binding of the biotin with Strept(avidin) is mainly due to the bicyclic moiety of biotin,²⁹ the valeric acid side-chain can be modified with different metal complexes that can be screened for their activity towards various chemical reactions (chemo optimization). In general, it was

observed that the alterations in organometallic moiety result in dramatic change in reactivity. The fine-tuning of the enantioselectivity may then be achieved by screening different isoforms of the host protein generated by point mutations at neighbouring amino acid residues. Finally, a third optimization step can be included to evaluate the substrate scope by employing the best protein host-metal complex combination for catalysis.²⁴

1.3.2 Directed evolution

In 2002, Reetz and co-workers proposed the use of directed evolution of enzymes to enhance the enantioselectivity of a given reaction.³⁰ In general, directed evolution involves repeated cycles of random gene mutagenesis, followed by high-throughput screening/selection of desirable mutant proteins to exert evolutionary pressure towards a particular catalytic activity in a Darwinistic manner (Scheme 1-4b).^{30,31} Directed evolution of enantioselective hybrid catalysts based on Rh-containing biotin-strept(avidin) system was first demonstrated for hydrogenation of α -acetamidoacrylic acid ester. 65% ee of R enantiomer and potential to invert the selectivity in order to access both enantiomers was demonstrated by using existing mutagenesis techniques.³²



Scheme 1-4. General Scheme for (a) chemogenetic optimization (b) combination of directed evolution and chemical modification for fine-tuning of activity of artificial metalloenzymes

1.3.3 Covalent chemical modification strategies

Chemical modification is a powerful method to alter structure and function of natural enzymes.³³⁻³⁶ Covalent modification can not only be used to introduce new functions by anchoring a metal catalyst (as described above), but can also be used to alter substrate specificity³⁷⁻³⁹ or enhance the stability⁴⁰ of the protein-based catalysts. In order to expand the scope of natural biocatalysts, Reetz proposed an approach based on the combination of directed evolution and chemical modification of protein-based catalysts (Scheme 1-4b).^{31,41} This involved employing the existing mutagenesis methods to first generate and express mutant libraries of a suitable model protein, which can be modified in a selective manner by reaction at a reactive amino acid side-chain (like cysteine). *En masse* screening of the mutant hybrid catalysts in a given transition metal catalysed reaction and repetitive cycles would lead to the identification of the best catalyst. In this way, the powerful chemical modification and artificial evolutionary methodologies can be combined.

1.3.4 Computational design

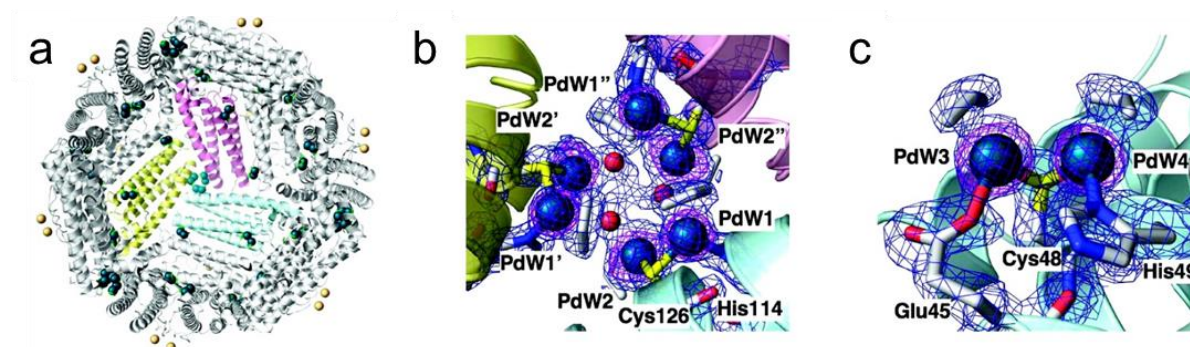
In general, the computational design of *de novo* enzymes comprises of common steps.⁴²⁻⁴⁴ Once an appropriate reaction and the associated mechanism are chosen, the protein functional groups are positioned in a way so as to allow for maximum transition state stabilization, i.e. the model of an ‘idealized active site’ is constructed. This is followed by a search through the available crystal structures for appropriate protein backbone positions that can support this idealized active site. After considering other factors like stability of the protein, affinity for transition state etc., the top ranking designs are filtered and experimentally verified. Several computationally designed enzymes have been experimentally expressed and their activity towards the target reaction was assessed. Notably, some have been shown to catalyse reactions with 10^2 - 10^3 increase in

activity.^{43,45,46} Impressively, the stereoselective enzymes for Diels-Alder reaction created by Baker and co-workers through computational design displayed high substrate specificity.⁴⁷ This design was based on an ‘ideal’ active site containing hydrogen bond-acceptor/donor residues capable of ensuring proper orientation of the substrates for such bimolecular reaction.⁴⁷ More recently, computational design of a metalloprotein based genetic incorporation of unnatural amino acid (bipyridylalanine) was accomplished.⁴⁸ Excellent atomic-level accuracy of the design with the X-ray crystal structure of the metalloprotein shows the huge potential of exploiting computational methods for designing precise metal binding sites in proteins.

1.3.5 X-ray crystallography

Available protein crystal structures can be an important resource in the design of artificial metalloenzymes.⁴⁹ The X-ray crystal structures of metalloproteins can provide information about the primary coordination sphere of the metal⁵⁰ as well as the secondary interactions due to the chiral environment around the metal complex. X-ray crystallography can also highlight the key protein-substrate interactions. This has been used to elucidate mechanistic details by systematically trapping the various intermediates of the catalytic cycle.⁵¹ It may also be possible to elucidate the role of the close lying amino acids in catalysis based on the structural data. Information derived from X-ray crystal structures of artificial metalloenzymes has been previously used for prediction of enantioselection mechanism⁵² and for the future optimization to develop second generation catalysts through saturated mutagenesis at pre-selected amino acid residues (designed evolution).^{24,53} Structural characterization of artificial metalloenzymes based on model proteins, such as ferritin (Fr), Myoglobin (Mb), streptavidin (Sav) and carbonic anhydrase (CA) has been most successful.

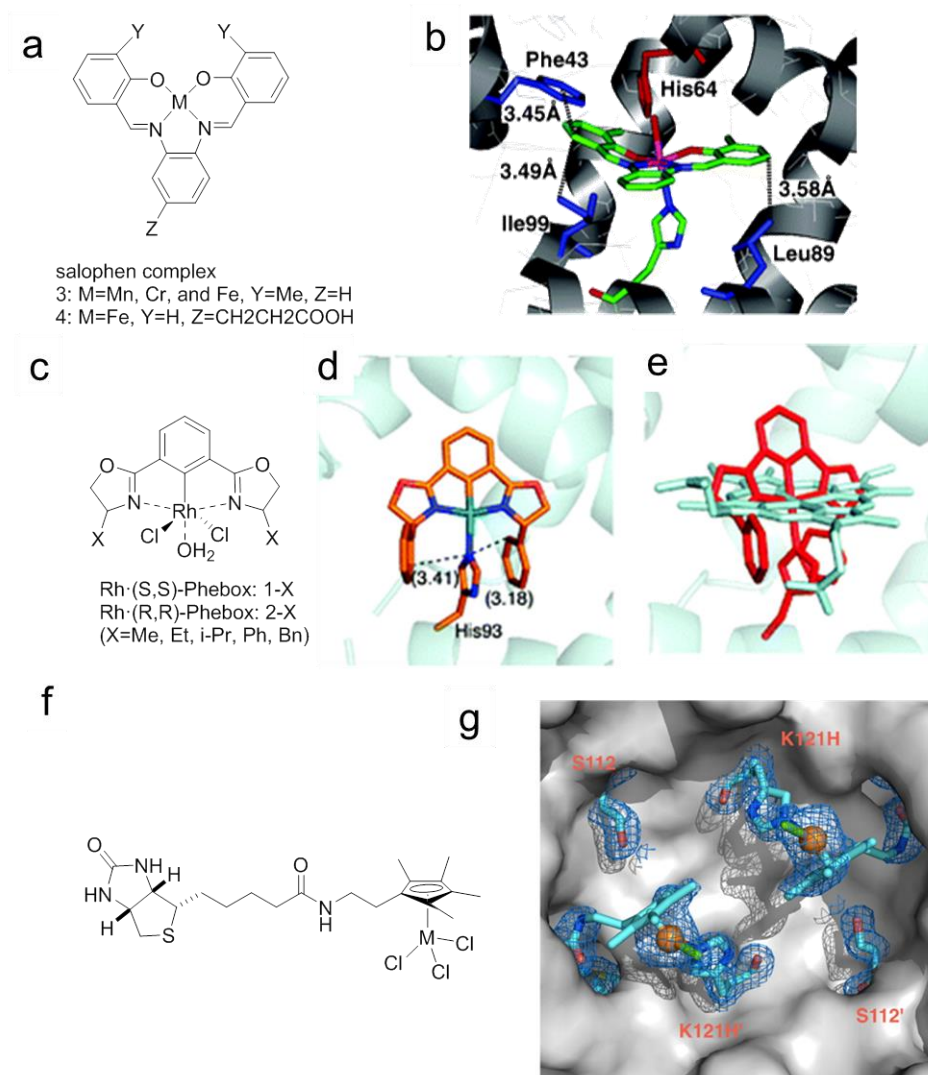
Ferritin, a spherical protein made of 24 subunits has natural metal binding sites capable of accumulating as many as 4,500 iron atoms (Scheme 1-5a). Thiol, carboxylate and imidazole side-chains of natural amino acids in the interior of the ferritin cage can bind different metal ions.⁵⁴⁻⁵⁶ In particular, the crystal structure of **Pd(allyl)•apo-Fr** revealed Pd(allyl) coordinated with key His, Cys and Glu residues in the interior of the ferritin cage (Scheme 1-5b and c).⁵⁶ Further investigation of the accumulation process of Pd(II) ions in the interior of apo-ferritin was undertaken by X-ray crystallographic studies. A series of crystal structures of apo-ferritin with low (144), intermediate (192) and high (216) number of Pd(II) binding sites in the protein cage were elucidated.⁵⁴ This step-wise depiction of the metal binding pattern in ferritin on the 3-fold axis, accumulation centres and another third site on the monomers provided insights into the metal deposition and metal core formation processes.



Scheme 1-5. (a) Crystal structures of Pd(allyl)•apo-Frs. (b-c) The interior structures of Pd(allyl)•apo-Fr, Close views of (b) 3-fold channels and (c) accumulation centers of Pd(allyl)•apo-Fr; the Pd and Cd atoms are indicated as greenish blue and beige spheres, respectively. The anomalous difference Fourier maps at $4.0\ \sigma$ indicate the positions of palladium atoms which are shown in magenta. The selected $2|F_o| - |F_c|$ electron density maps at $1.0\ \sigma$ are shown in blue; (a-c) Reprinted (adapted) with permission from (Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, *130*, 10512.). Copyright (2008) American Chemical Society.

In another approach, the known affinities of natural prosthetic groups or their synthetic analogues can be used to localise the metal catalyst in a binding pocket. The heme binding cavity of Myoglobin (Mb) forms an attractive target to bind unnatural metal cofactors. In particular Mb can bind metal-containing Salophen ligands due to the strong hydrophobic interactions and the coordination of metal with proximal histidine (His) residue⁵⁷ (Scheme 1-6a and b). Watanabe and co-workers reported the X-ray crystal structures of Mb-based metalloenzymes with such non-natural metal cofactors.⁵⁷⁻⁵⁹ Interestingly, the arrangement of two unnatural cofactors crystallised in the Mb cavity was found to be perpendicular to each other showing the importance of hydrophobic interactions to ensure the appropriate positioning of the metal cofactors (Scheme 1-6c, d and e).⁵⁹ Crystal structures of other artificial metalloenzymes based on supramolecular interactions of biotinylated metal complexes with Sav variants are also known.^{52,53,60} However, low occupancy of ligand and metal was an issue.^{52,53} Hence, a dual metal anchoring approach was applied based on supramolecular and dative binding from a neighbouring His residue.⁶⁰ The X-ray crystal structure of $[(\eta^5\text{-Biot-2})\text{Rh}]^{2+}$ showed the dative bonding of Rh with His residues installed at position 112 or 121.

Comparatively limited examples of covalently anchored metal bound unnatural cofactors are known.^{61,62} By employing the covalent anchoring strategy, the crystal structure of **ALBP-Phen117** containing the metal-free phenanthroline (Phen) ligand covalently attached to Cys117 residue in the protein cavity was obtained.⁶¹ No deviation from the native ALBP structure was seen indicating that the ligand was successfully installed without perturbation of the native conformation. However, close inspection of the structure revealed that the phen ligand was quite constrained in the cavity and could not bind Cu^{2+} in the same configuration owing to steric constraints.⁶¹



Scheme 1-6. (a) (c) and (f) chemical structures of Salophen, Phebox and biotin-based ligands respectively, (b) Crystal structures of **Mn(Salophen)•apo-A71GMb**; (d) **Rh(Phebox)•apo-A71GMb** showing hydrophobic interaction between the amino acid residues and Phebox ligand. (e) Comparison of arrangement between Phebox (red) and heme (white), (g) Close-up view of the X-ray crystal structure of **[(η⁵-Biot-2)RhCl]⁺ c K121H Sav**; Rhodium is colored in orange and chlorine in green. Residues H121' and K121' are provided by the adjacent Sav monomer. (b) Reprinted (adapted) with permission from (Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**, 127, 6556). Copyright (2005) American Chemical Society. (d-e) Reprinted (adapted) with permission from (Satake, Y.; Abe, S.; Okazaki, S.; Ban, N.; Hikage, T.; Ueno, T.; Nakajima, H.; Suzuki, A.; Yamane, T.; Nishiyama, H.; Watanabe, Y. *Organometallics* **2007**, 26, 4904). Copyright (2007) American Chemical Society. (g) Reprinted (adapted) with permission from (Zimbron, J. M.; Heinisch, T.; Schmid, M.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *J. Am. Chem. Soc.* **2013**). Copyright (2013) American Chemical Society.

1.4 Existing metallo-bioconjugates: scope of artificial metalloenzymes

An overview of the reactions catalysed by artificial metalloenzymes is given. Several reviews describe the full range of reactions covered under the ambit of artificial metalloenzymes.¹¹⁻¹⁵

1.4.1 Ester Hydrolysis

Davies *et al.* demonstrated amide and ester hydrolysis inside a hydrophobic cavity of Adipocyte lipid binding protein (ALBP) based metalloenzyme.⁶³ Chemical installation of a phenanthroline (Phen) ligand by direct alkylation gave the desired metalloenzyme, ALBP-Phen117•Cu(II) (Scheme 1-7a). Using this, a range of ester and amide substrates were hydrolysed with selectivity up to 86% ee and observed turnover number (TON) of 7.6. Notably, significant rate acceleration was achieved over the background reaction, albeit 2-3 orders of magnitude less than that of free Cu(II) in the solution. Further studies showed the chiral environment around the metal catalyst could influence the reaction rates and selectivity.⁶⁴ By installing Phen ligand at different locations within the cavity of a related protein Intestinal Fatty Acid Binding Protein (IFABP), ester hydrolysis with TON of 1.3 and selectivity up to 93% ee was achieved.

1.4.2 Hydrogenation

Inspired by the Whitesides' seminal report on enantioselective hydrogenation using biotin-avidin based system,¹⁹ the catalytic scope for asymmetric hydrogenation has been explored mainly using Strept(avidin) as host protein.^{20,25,26,32,65-67} It was shown that quantitative conversions and high levels of enantioselectivity up to 95% are possible by employing artificial metalloenzymes (Scheme 1-7b).¹³ Crucially, through the chemogenetic probing and optimization of the system it was concluded that the enantioselectivity and

conversions were mainly determined by the organometallic counterpart. The involvement of the secondary sphere interactions from the protein scaffold in fine-tuning of the selectivity was also assessed. This was evidenced by the improvement in ee values by using genetic optimization methods, whether by switching avidin→streptavidin²⁵ or by saturated mutagenesis of close lying amino acid residues.²⁶ Using chiral amino acid spacer units between the biotin anchor and the organometallic catalyst, higher and tunable selectivity with 3-fold rate acceleration compared to protein-free catalyst was demonstrated.⁶⁶ Using an alternative performance optimization approach, enantioselective hydrogenation of *N*-protected dehydroamino acids was optimized by directed evolution.³² Olefin hydrogenation has also been demonstrated using Pd-nanoclusters enclosed in ferritin cage.⁶⁸ The substrate selectivity due to size discrimination during diffusion through the three-fold axis is, of course, an attractive feature of using ferritin cage for catalysis.

1.4.3 Transfer Hydrogenation

Ward and co-workers have explored artificial transfer hydrogenases^{13,69} based on biotin-Strept(avidin) system for enantioselective reduction of ketones^{27,53,70,71} and imines^{52,60} using aqueous sodium formate as hydrogen source (Scheme 1-7c). Different piano-stool complexes (attached with a biotinylated sulphonamide ligand) and metal catalysts (Ru, Rh, Ir) were screened to exploit the chemo diversity of the system. Out of these, ruthenium complexes were generally more active and the capping arene unit greatly influenced the enantioselectivity of the reaction. Using saturated mutagenesis at close lying amino acids to the metal (S112 and S122), it was shown that both R (92%) and S (62%) enantiomers could be obtained.²⁷ Interestingly, X-ray crystallography gave valuable insights into the absolute configuration at Ru center (S) and other close contacts of the metal with the surrounding amino acids.⁵³ Based on this crystal structure, additional residues were chosen

for saturated mutagenesis which led to the generation of (S)-selective catalysts²⁴ as well as active transfer hydrogenases for more challenging dialkylketone substrates.⁵³

1.4.4 Oxidation

Non-covalent insertion of Cr(III) and Mn(III) Salophen and Salen-based complexes resulted in artificial metalloenzymes active in catalysing sulfoxidations of thioanisoles (Scheme 1-7d).^{57,72,73} In these enzymatic oxidations, rate acceleration (up to 31-fold) and asymmetric induction (32% S or 13% R) was ensured by the localisation of the metal and substrates in the myoglobin cavity.⁵⁷ It was proposed that a two-point attachment of the Mn(Salen) complex by covalent chemical modification strategies could limit the number of conformations of the metal complex in the protein cavity, thereby ensuring precise positioning of the metal for catalysis.⁷⁴ Indeed, this dual anchoring approach led to higher enantioselectivities of up to 51% and the effect of altering the anchor positions both on the reactivity and selectivity was systematically evaluated.^{75,76} Higher enantioselectivities up to 93% ee are reported using vanadium-based protein catalysts.^{77,78} In this regard, the excellent selectivity of the biotin-Strept(avidin) based artificial metalloenzyme perhaps could be attributed to its ability to provide a specific binding site for pentahydrated vanadium ion $[\text{VO}(\text{H}_2\text{O})_5]^{2+}$.

Enantioselective epoxidation of styrene has also been reported in the recent years, however the conversions of these reactions using artificial metalloenzymes remain quite low.⁷⁹⁻⁸¹ Recently, a metalloporphyrin cofactor was inserted into Xylanase and was found to catalyse stereoselective oxidation of a *p*-methoxystyrene substrate with high enantioselectivity (80% ee), albeit in 16% epoxide yield.⁸² Recombinant expression of

ferritin enabled the use of a highly robust and temperature-stable host for generating Pd-nanoparticles for alcohol oxidation.⁸³

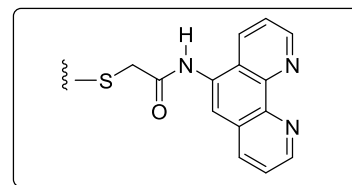
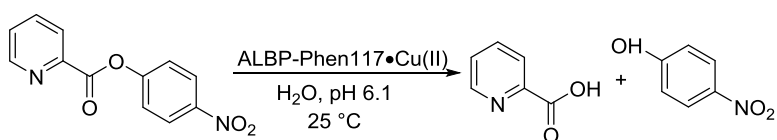
1.4.5 Hydroxylation

Dihydroxylation of model alkene substrates was demonstrated by incorporation of OsO₄ catalytic centre in BSA⁸⁴ and strept(avidin).⁸⁵ Impressively, genetic optimization of the Osmium loaded streptavidin catalyst by point mutations of the neighbouring amino acid residues in the biotin binding site afforded improved enantioselectivities of up to 97%.⁸⁵

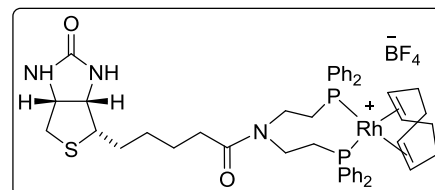
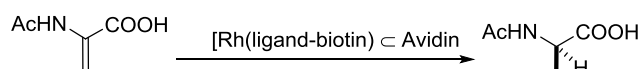
1.4.6 Hydration

By employing a DNA-based catalyst that can be generated by self-assembly of an achiral ligand, Boersma *et al.* succeeded in demonstrating the first enantioselective syn-hydration of α,β -unsaturated ketones.⁸⁶ The chirality transfer was mainly due to the presence of DNA and enantioselectivities of up to 82% were achieved in this reaction, with no equivalent in homogeneous catalysis. Bos *et al.* recently reported an enantioselective metallo-hydratase by covalently anchoring a [Cu(II)-phenanthroline] metal complex at the interface of a homodimeric LmrR protein (Lactococcal multidrug resistance Regulator).⁸⁷ This artificial metalloenzyme catalysed the conjugate addition of water to alkenes with up to 84% ee which could be modulated by altering the second coordination sphere of the metal installed in the hydrophobic pocket of the protein dimer.

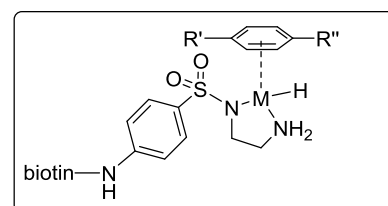
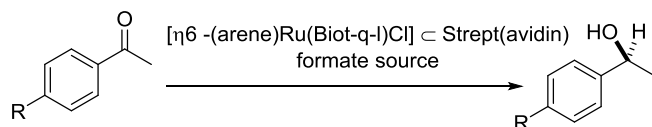
(a) **Davies et al. (1997)**



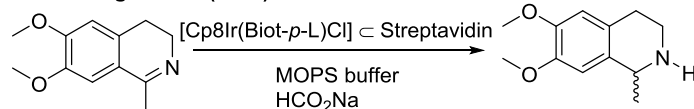
(b) **Collot et al. (2003)**



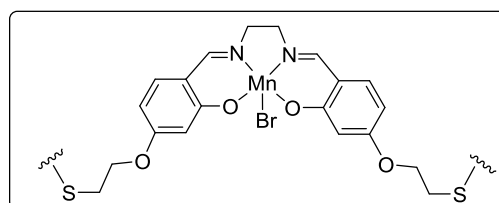
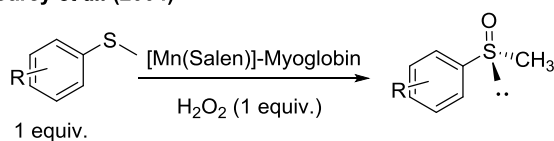
(c) **Letondor et al. (2005)**



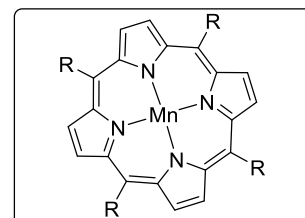
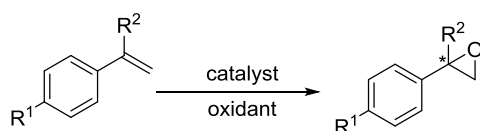
Durrenberger et al. (2011)



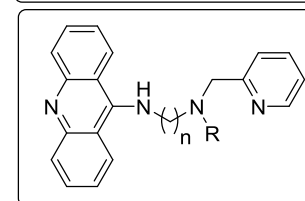
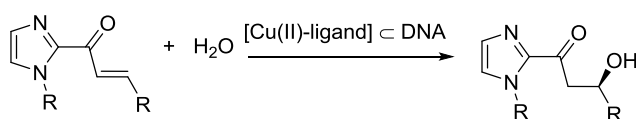
(d) **Carey et al. (2004)**



Allard et al. (2012)



(e) **Boersma et al. (2010)**



Scheme 1-7. Reactions catalysed by artificial metalloenzymes (a) Ester/amide hydrolysis (b) Asymmetric hydrogenation (c) Asymmetric transfer hydrogenation of ketone and imines (d) Sulfoxidation and epoxidation reactions (e) Syn-hydration of alkenes

1.4.7 Hydroformylation

Native zinc ion in carbonic anhydrase (CA) was replaced by rhodium using $[\text{Rh}(\text{CO})_2(\text{acac})]$ as the metal source and the resultant artificial metalloenzymes adduct were used in regioselective hydroformylation reaction with styrene (Scheme 1-8a).^{88,89} Previously, non-specific metal coordination by surface exposed residues of CA was also observed.⁹⁰ However, protein engineering of the scaffold could facilitate specific Rh binding at the active site.⁸⁹ Interestingly, it was shown that the linear hydroformylation may be favoured over branched product when the metal was selectively bound at the active site. This activity was different from the activity of Rh catalyst in solution or bound non-specifically on the surface that mainly gave the branched product.⁸⁹

1.4.8 Allylic Amination

Covalent anchoring of phosphine ligands on photoactive yellow protein (PYP) via imidazole formation of corresponding carboxylic acids has been reported (Scheme 1-8b).⁹¹ The resultant conjugates were used for asymmetric allylic amination reaction with good conversions in 1:1 water/DMF solvent conditions. However, no asymmetric induction could be achieved, presumably because of denaturation of protein due to high percentage of organic solvents in the reaction mixture.

1.4.9 Allylic Alkylation

The design of Pd-based hybrid catalysts for an asymmetric allylic alkylation reaction has been reported by using supramolecular binding of a biotinylated phosphine ligand with Streptavidin (Scheme 1-8c).⁹² Chemogenetic optimization of these protein hybrid catalysts afforded both R/S enantiomers (in 93% ee and 82% ee respectively) in this chemical transformation, for which no natural counterpart exists. In this case, two equiv. of

didodecyldimethylammonium bromide (DMB) detergent was included in the reaction mixture to prevent hydrolysis of the starting 1,3-diphenylallyl acetate to corresponding alcohol.

1.4.10 Benzannulation

Incorporation of a biotinylated Rh(III) complex in streptavidin complemented with a carboxylate residue in the vicinity led to design of an artificial metalloenzyme for C-H activation reactions (Scheme 1-8d).²¹ The first coordination sphere interactions around the metal facilitated enantioselective benzannulation reaction between pivaloyl-protected benzhydroxamic acids and alkenes, with 92-fold rate acceleration as compared to free rhodium complex. This example showed the huge potential of using the chiral environment of the protein to carry out asymmetric reactions as there was no prior precedence of enantioselective version of this benzannulation reaction.

1.4.11 Olefin metathesis

In a quest to add more challenging reactions to artificial metalloenzyme repertoire, olefin metathesis (OM) has attracted much attention recently.⁹³⁻⁹⁵ Interestingly, artificial metalloenzymes for ring closing metathesis (RCM) reaction were generated by using different modes of attachment of the organometallic moiety, namely covalent anchoring using direct alkylation of a cysteine residue⁹³ (Scheme 1-8e) or by employing a covalent protease inhibitor⁹⁴ or by supramolecular anchoring based on biotin-streptavidin interaction.⁹⁵ Linkers were incorporated in the organometallic counterpart to expose the catalytically active metal to the surface for multi-turnover activity.^{94,95} Chemogenetic optimization and lowering the pH of the reaction led to higher conversions, albeit still lower or at best comparable to the free metal complex in the solution. Recently, a hybrid

catalyst for ring-opening metathesis polymerization (ROMP) of 7-oxanorbornene substrates has also been developed using covalent anchoring of Grubbs-Hoveyda type catalyst to a β -barrel transmembrane protein.⁹⁶

1.4.12 Diels-Alder reaction

Diels-Alderases for this synthetically valuable reaction have also been created by using catalytic antibodies,⁹⁷⁻⁹⁹ RNAzymes¹⁰⁰ and *de novo* computational design.⁴⁷ The utility of artificial metalloenzymes as biological Diels-Alder catalysts has been demonstrated by the use of both protein¹⁰¹⁻¹⁰³ and DNA^{17,104} scaffolds (Scheme 1-8f). Artificial metalloenzymes have been reported for lewis acid catalysed Diels-Alder reactions by employing supramolecular and dative anchoring (enabling 98% ee)¹⁰⁵ as well as by using covalent attachment of *N,N*-ligands.¹⁰⁶⁻¹⁰⁸ The influence of protein secondary sphere on the selectivity of a reaction has also been recently demonstrated by creation of an active site on the dimer interface of a transcription factor LmrR that enabled Cu(II) catalysed Diels-Alder reaction with excellent enantioselectivities of up to 97%.¹⁰⁸

1.4.13 Polymerization

Polymerization of a phenylacetylene substrate was catalysed by Rh(nbd) complex bound at discrete sites in the interior of ferritin cage.⁵⁵ Three Rh binding sites per monomer were structurally characterized and the reaction was believed to be regulated by the size and environment in the interior of the protein cage. More recently, Hayashi and co-workers have demonstrated that a heme-free vacant nitrobindin (NB) protein scaffold can be used to covalently attach mononuclear Rh-catalysts in the interior of the protein (Scheme 1-8g).⁶² Interestingly, several trans-selective NBRh-based catalysts were generated yielding polymers with trans content up to 82% when compared to a cis-selective reference

metalloprotein (MbA125C-Rh) with rhodium complex on the surface of myoglobin. Stereopreference was further rationalised based on X-ray crystallography and computational methods.⁶²

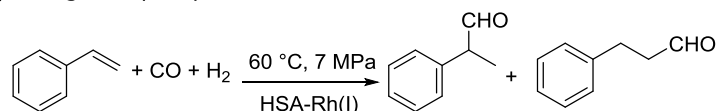
1.4.14 Suzuki-Miyaura cross-coupling reaction

In order to expand the scope of artificial protein conjugates, protein engineering and metal deposition on the interior surface of apo-ferritin (a natural iron storage protein composed of 24 sub-units) was used to form palladium-nanoparticles containing 96 palladium atoms each.⁵⁶ These nanoparticles were used to demonstrate a model Suzuki-Miyaura cross-coupling reaction between 4-iodoaniline and phenylboronic acid, however the catalytic activity was lower than that of free palladium catalyst in solution.⁵⁶

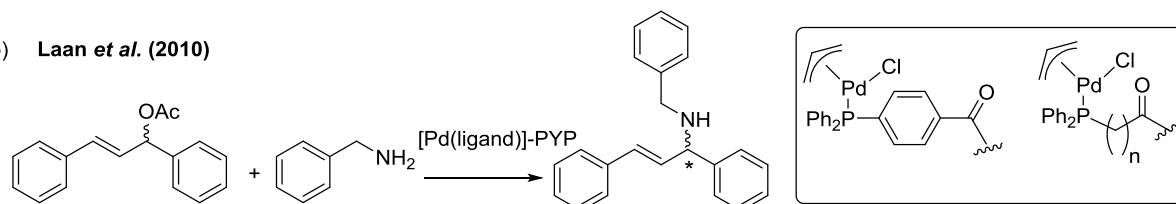
1.4.15 DNAzymes

Supramolecular binding of intercalating achiral ligands to DNA has been used to carry out several enantioselective C-C bond formation reactions, where DNA is the only chiral inducer (Scheme 1-8f).¹⁷ Such chirality transfer from DNA has been exploited to carry out Diels-Alder¹⁰⁴ and Michael addition¹⁰⁹ and Friedel-Crafts¹¹⁰ reactions with up to 99% ee. The reactions were shown to be DNA-accelerated¹¹¹ and chemo-optimization of the system was possible by screening different achiral ligands. It was also observed that both enantiomers were accessible by tuning the denticity of the ligand.¹¹² Covalent anchoring of organometallic complexes has also been a successful approach in yielding enantioselective DNAzymes.^{113,114} A full account of the current scope of DNAzymes has been discussed in several reviews.^{12,115,116}

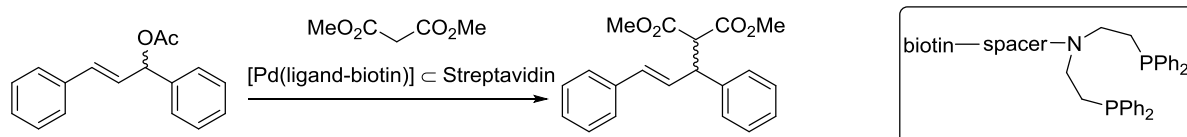
(a) **Jing et al. (2010)**



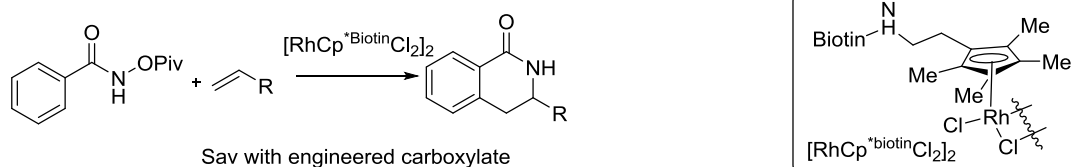
(b) **Laan et al. (2010)**



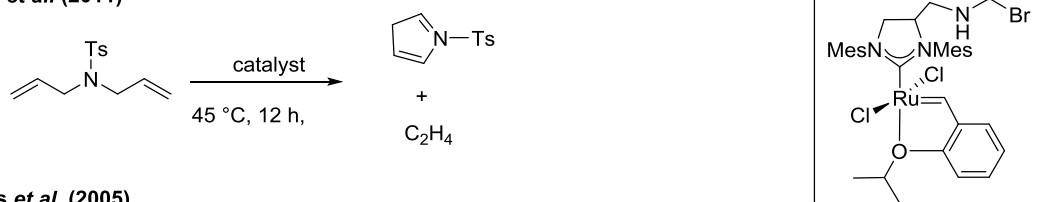
(c) **Pierron et al. (2008)**



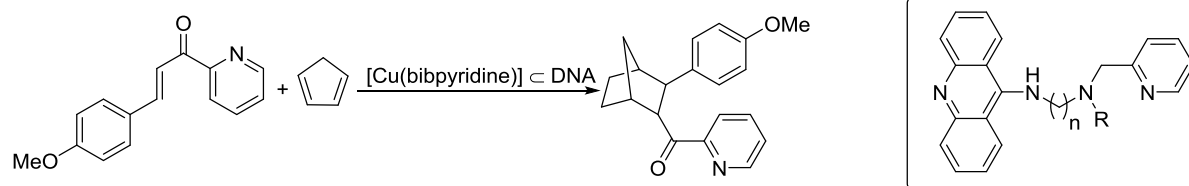
(d) **Hyster et al. (2012)**



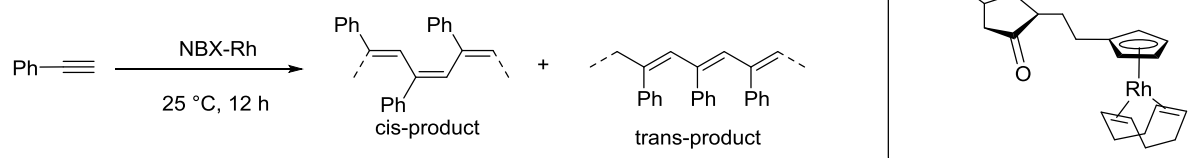
(e) **Mayer et al. (2011)**



(f) **Roelfes et al. (2005)**



(g) **Fukumoto et al. (2014)**



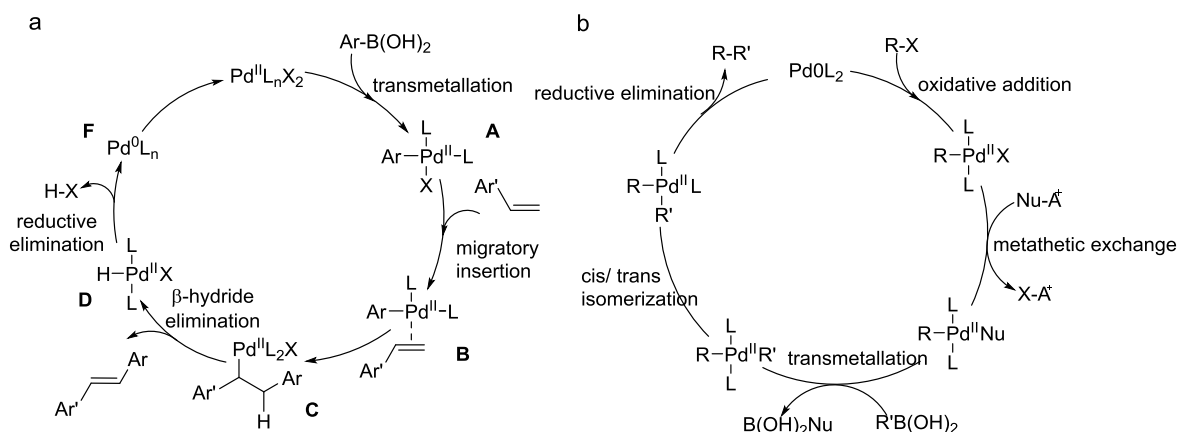
Scheme 1-8. C-C or C-X bond formation reactions catalysed by artificial metalloenzymes (a) Hydroformylation (b) Allylic amination (c) Asymmetric allylic alkylation (d) Benzannulation via C-H activation (e) Ring closing metathesis (f) Diels-alder reaction (g) Polymerization

The following section comprises a description of the chemical catalyst (palladium) and the model enzyme (subtilisin) used for the work described in this thesis.

1.5 Palladium catalysed C-C cross-coupling reactions

Transition metal catalysed cross-coupling reactions offer powerful and diverse chemical routes for the construction of C-C bonds. Amongst these, Pd-catalysed Mizoroki-Heck^{117,118} and Suzuki-Miyaura¹¹⁹ cross-coupling reactions comprise some of the most important reactions in organic synthesis. Because of the widespread applicability and industrial impact of these Pd-catalysed cross-coupling reactions, Richard Heck, Akira Suzuki and Ei-ichi Negishi were awarded the Nobel Prize in Chemistry 2010 for their pioneering work.

An alternative alkenylation reaction, the oxidative Heck cross-coupling has gained importance in recent years.¹²⁰⁻¹²⁴ Importantly, commercially available and less toxic arylboronic acid coupling partners have replaced aryl halides in this variant of the Heck reaction. The various intermediates in the proposed cycle for oxidative Heck cross-coupling reaction between an arylboronic acid and an olefin are shown (Scheme 1-9a).^{120,125} The reaction begins with transmetalation of arylboronic acid to form the arylpalladium(II) species **A**. This is followed by coordination of the alkene with the metal centre to give species **B**, which undergoes migratory insertion to give **C**. β -hydride elimination forms the palladium hydride species **D**, which ultimately dissociates from the olefinic cross-coupled product. With the use of *N,N* ligands the regeneration of the active Pd(II) catalyst is possible by directly employing molecular oxygen.^{126,127}



Scheme 1-9. (a) Proposed catalytic cycle of oxidative Heck cross-coupling reaction between olefin and arylboronic acid (b) Catalytic cycle for Suzuki-Miyaura cross-coupling reaction

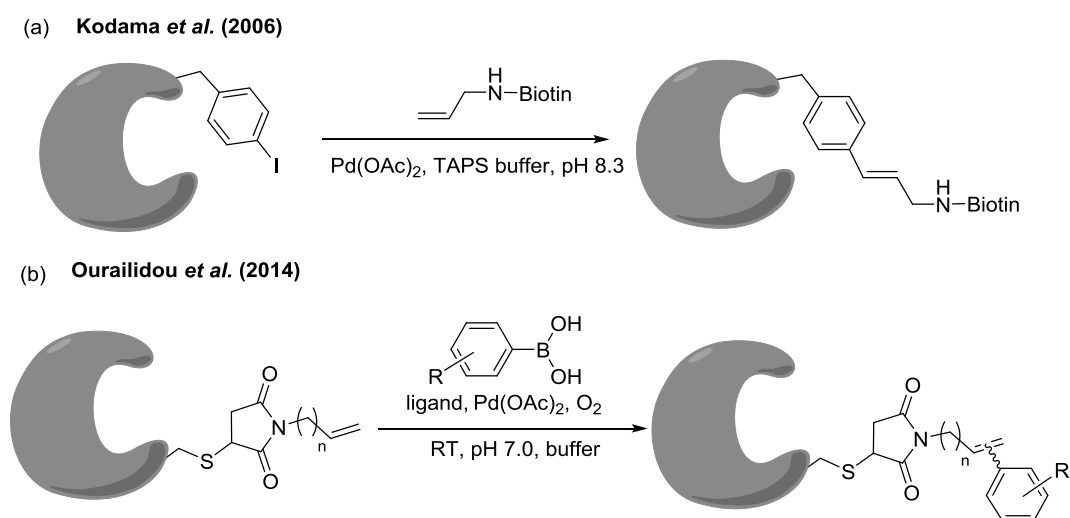
Three basic steps involved in the Suzuki-Miyaura cross-coupling between arylboronic acids and arylhalides involve oxidative addition, transmetalation and reductive elimination sequences (Scheme 1-9b). The presence of base is crucial for this reaction to facilitate the transmetalation step.¹²⁸ Furthermore, Amatore *et al.* have highlighted the critical role of tri- or penta coordinated anionic intermediates in the catalytic cycle providing mechanistic details.¹²⁹ While our knowledge about this important reaction mechanism continues to grow, the Suzuki-Miyaura reaction has also been applied to protein substrates demonstrating the compatibility of these transition metal complexes with biological systems.¹³⁰⁻¹³⁴ In fact, several other transition metal-catalysed reactions have been demonstrated on protein substrates as discussed in the next section.

1.5.1 Metal-catalysed cross-coupling reactions on protein substrates

Transition metal-catalysed reactions have previously been used for site-specific protein modification.¹³⁵ However, several M-catalysed C-C cross-coupling reactions have only been recently demonstrated as effective tools for generating a variety of protein-

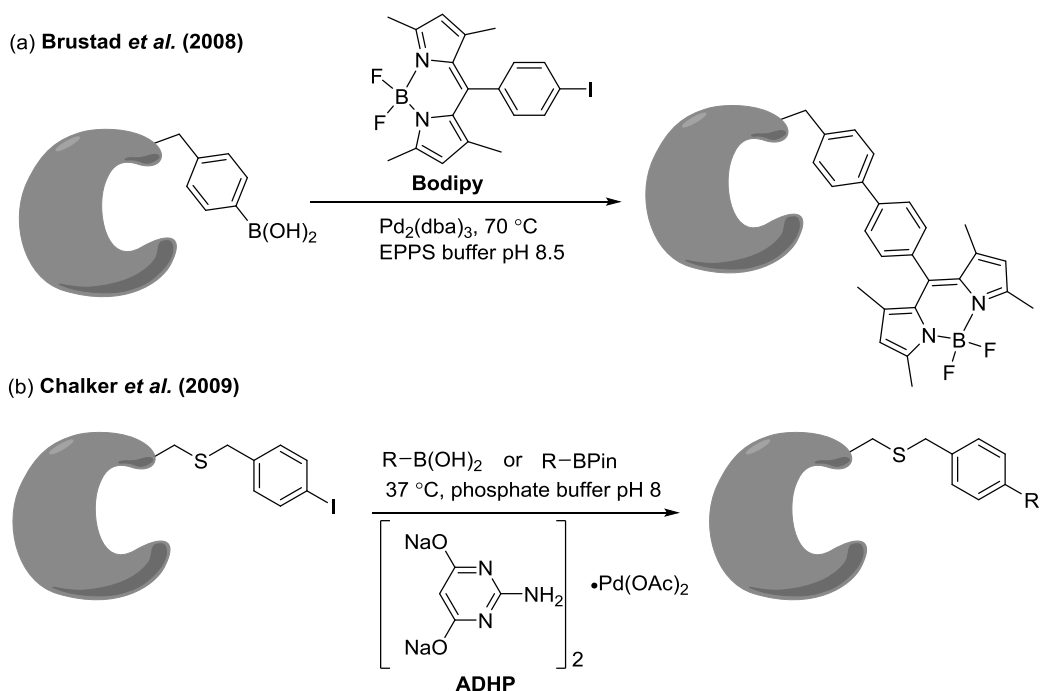
conjugates.¹³⁶ To this end, protein modification using Mizoroki-Heck, oxidative Heck, and Suzuki-Miyaura are discussed in this section.

Proteins were used as substrates for a Pd-catalysed cross-coupling reaction for the first time by Yokoyama and co-workers to demonstrate Mizoroki-Heck cross-coupling reaction at a genetically encoded *p*-iodophenylalanine (*p*IPhe) residue (Scheme 1-10a).^{137,138} The *p*IPhe residue was first incorporated at position 32 into a His6-fused Ras protein (iF32-Ras-His), which was subsequently modified using a biotinylated alkene in 2% yield. A significant amount of dehalogenation side product was also formed (28%), perhaps due to the steric hindrance at position 32.¹³⁸ More recently, a bio-orthogonal protein labelling strategy was demonstrated using protein-bound alkene as coupling partner in the oxidative Heck cross-coupling reaction (Scheme 1-10b).¹³⁹ A single cysteine 4-oxalocrotonate tautomerase variant (4-OT-R61C) was modified with alkene using a maleimide linker. Subsequent, site-selective labelling was achieved by coupling fluorescent phenylboronic acid to the alkene bound to 4-OT-R61C using 20 equiv. of Pd(OAc)₂/BIAN as catalyst and 17% DMF as co-solvent.



Scheme 1-10. (a) Mizoroki-Heck and (b) Oxidative Heck cross-coupling reactions on protein substrates

Genetic incorporation of a *p*-boronophenylalanine residue (*p*BPhe) in a model protein substrate lead to a tagged protein substrate (Scheme 1-10a).¹⁴⁰ Suzuki-Miyaura cross-coupling at this unnatural *p*BPhe residue with fluorescent aryl halide (**Bodipy**) using palladium dibenzylidene acetone (Pd₂(dba)₃) as catalyst enabled partial labelling of the protein (30%), albeit at high temperatures. Hence, water soluble Pd-catalysts were developed that enabled efficient Suzuki cross-coupling reactions on proteins. 1,1-dimethylguanidine¹³³ and the sodium salt of **ADHP** ligand¹³⁰ (a 2-amino-4,6-dihydroxypyrimidine ligand) are both water-soluble, active Pd-catalysts for Suzuki-Miyaura cross-coupling reactions (Scheme 1-11b). Cross-coupling of protein substrates containing chemically installed *p*-iodobenzyl residue (Pic) and arylboronic acids was successfully demonstrated using these ligands.^{130,133} The Suzuki-Miyaura cross-coupling reaction catalysed by ligand **ADHP** has been reported at genetically incorporated *p*IPhe residues on a variety of protein substrates^{131,134,141,142} enabling cell surface labelling.^{134,141} Moreover, site-selective PEGylation of protein substrates was also realised in a ‘self-liganded’ fashion where the PEG boronic acid also served to coordinate the active metal catalyst.¹³²



Scheme 1-11. Suzuki-Miyaura cross-coupling reaction on (a) genetically encoded *p*-boronophenylalanine residue (b) chemically installed Pic residue

Other examples of transition metal chemistry on protein substrates have also been reported including site-selective Sonogashira cross-coupling,^{138,143,144} and olefin-metathesis reaction.¹⁴⁵⁻¹⁴⁸

1.6 Subtilisin: model serine protease

Subtilisin is an attractive model protein scaffold with huge academic and industrial relevance. The specificities of substrate binding pockets combined with the catalytic Asp32-His64-Ser221 machinery are responsible for the remarkable hydrolytic activity of subtilisins (See Chapter 2 Section 2.2.1.2).¹⁴⁹ After binding the native peptide substrates, nucleophilic attack by Ser221 is facilitated by proton transfer to His64. This leads to the formation of the tetrahedral enzyme-substrate intermediate (ES) stabilized by the oxyanion hole interactions (formed by the backbone NH of Ser221 and Asn155).^{150,151} On collapse of the tetrahedral intermediate, a proton is transferred from His64 to the amine group of

the leaving group. Finally, hydrolysis of the ester intermediate by a water molecule regenerates the Ser221 hydroxyl.

The active sites of subtilisins have been traditionally used for engineering new activities.^{9,152-157} Koshland and Bender independently demonstrated the concept of chemical mutagenesis by converting the active site serine of subtilisin to a cysteine residue (Ser221→Cys221).^{152,153} Active site mutation in this new thiolsubtilisin enzyme resulted in considerable loss of protease activity. However, Kaiser and co-workers used this thiolsubtilisin to catalyse a peptide ligation reaction, thereby demonstrating that new functions can be introduced in enzymes with minimum active site mutation.¹⁵⁴ Inspired by Kaiser's peptide ligase, a double mutant of subtilisin (Subtiligase) created by Wells and co-workers showed improved ligase activity while the substrate specificity essentially remained similar to that of wild-type enzyme.^{155,156} This led to the realisation that although the remarkable activities of natural enzymes arise due to high substrate specificity and catalytic activity, it is possible to engineer these two parameters separately and in a decoupled fashion.¹⁵⁶

The cysteine-free wild-type subtilisin also provides exciting opportunities to combine the site-directed mutagenesis and chemical modification techniques. Jones and co-workers have used these techniques to alter the specificity and activity of subtilisin.^{37,158,159} Their approach relied on using single cysteine variants of subtilisin which were modified by reacting them with methanethiosulfonate reagent to form chemically modified mutant proteins (CMMs). Our own group has used Subtilisin *Bacillus lentus* (SBL) as a model system for extensive protein chemical modification studies. The idea of chemical mutation was extended further by establishing chemical methods to convert cysteine (Cys) to

electrophilic dehydroalanine (Dha).^{160,161} This was subsequently used to install post-translational modification to form thio-ether based protein conjugates.^{162,163} Furthermore, Dr. Anca Pordea and Dr. Alba Díaz-Rodríguez in our lab used SBL as a model protein to create Pd-based artificial metalloenzymes. Chemically modified subtilisin mutants (at modification sites 156 and 217) were generated by direct alkylation of the single cysteine with bidentate ligands (2,2'-bipyridine or (2,2'-dipyridyl)methyl) as well as of monodentate pyridyl ligands. Corresponding palladium complexes were obtained by treating with Na₂PdCl₄ and these artificial metalloenzymes were tested for their activity towards the Suzuki-Miyaura cross-coupling of aryl halides with boronic acids. No activity could be observed under the employed conditions in preliminary catalysis experiments (Dr. Anca Pordea).

Further investigation of catalysis of the non-chiral and asymmetric version of this reaction was done with SBL-L217C modified with 2,2'-dipyridyl)methyl-based ligand (Dr. Alba Díaz-Rodríguez). This SBL-based artificial metalloenzyme was found to be active in catalysing non-chiral Suzuki-Miyaura coupling reactions. However, the control reactions using SBL-WT showed similar or higher activity. Additionally, monitoring of the reaction by mass spectrometry (relying on a PMSF based aryl halide and alkene) and crystallization of SBL-based metalloenzyme was attempted. Unfortunately, no cross-coupling could be observed by LC-MS and relevant structural data could not be obtained. The work in this thesis was built on using SBL modified with these bidentate ligands as the model system for designing artificial metalloenzymes for cross-coupling reactions and the specific research objectives are described in the next section.

1.7 Aim of the thesis

The main goal of this thesis was to design artificial metalloenzymes for chemical cross-coupling reactions. As discussed in section 1.6, it is possible to alter the function of the enzyme without affecting the substrate binding properties of a naturally evolved pocket. Inspired by this view and by taking advantage of the existing knowledge about the substrate specificity of Subtilisin *Bacillus lentus* (SBL), we aimed at engineering cross-coupling activity in SBL by localising a chemical catalyst (palladium) in its active site. Our efforts to this end were mainly focused on introducing a metal binding site and demonstrating a cross-coupling reaction in the active site of SBL-based artificial metalloenzymes.

A rational design approach was adopted for creation of SBL-based artificial metalloenzymes that included three essential benchmarks, namely introduction of a single, reactive ‘chemical tag’ in a suitable enzyme pocket, site-selective installation of a metal binding amino acid at the pre-tagged site and metal coordination in the active site. In order to cause minimum alteration of the enzyme’s catalytic residues and binding pockets, we decided to chemically install a metal binding amino acid at the base of the active site by using two alternative covalent chemical modification techniques, namely direct alkylation of cysteine and conjugate addition at dehydroalanine (Dha). While direct alkylation strategies have been used traditionally to design artificial metalloenzymes, to the best of our knowledge, the selective chemistry of the Dha route has not yet been exploited for this purpose. In Chapter 2, we describe our logical design and construction of SBL-based metalloenzymes.

Section 1.4 outlined the catalytic scope of the existing artificial metalloenzymes. While several challenging transformations have been realised, the huge potential of transition metal catalysts have not been exploited thus far. For example, metal catalysed cross-coupling reactions can expand the scope of artificial metalloenzymes dramatically. We chose Pd(II)-catalysed oxidative Heck cross-coupling between olefin and arylboronic acids as the model reaction. Our approach is based on utilising the nucleophilicity and specificity of SBL to localise the two cross coupling partners. In chapter 3, we describe the systematic investigation of various intermediates of an engineered cycle for oxidative Heck-cross coupling between a tethered alkene substrate and various phenylboronic acids.

Several strategies can be employed for design and optimization of artificial metalloenzymes, including X-ray crystallography which can be used to unambiguously confirm the precise location and properties of installed metal centres. In chapter 4, we undertake the structural characterization of the SBL-based metalloenzymes. Moreover, our approach is based on taking advantage of the key enzyme-substrate intermediate for carrying out non-native reactions. Hence, our efforts to characterise the substrate bound enzyme species and other reaction intermediates from the engineered oxidative Heck cycle have also been described.

In chapter 5, we describe the design of artificial metalloenzymes for Suzuki-Miyaura cross-coupling reaction based on another serine hydrolase, AmpC β -lactamase. Taking advantage of the known covalent yet reversible interactions of arylboronic acids with AmpC, we propose that the active site of these enzymes would provide an activating microenvironment for localising the substrates. A rational design approach will be described to combine chemical and genetic methods to generate a library of tailored

metalloenzymes. The catalytic activity of these metalloenzymes in cross-coupling reactions remains to be determined.

1.8 References

- (1) Faber, K. *Springer Berlin* **2011**.
- (2) Koeller, K. M.; Wong, C.-H. *Nature* **2001**, *409*, 232.
- (3) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185.
- (4) Savile, C. K.; Janey, J. M.; Mundorff, E. C.; Moore, J. C.; Tam, S.; Jarvis, W. R.; Colbeck, J. C.; Krebber, A.; Fleitz, F. J.; Brands, J.; Devine, P. N.; Huisman, G. W.; Hughes, G. J. *Science* **2010**, *329*, 305.
- (5) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. *Science* **2003**, *299*, 1694.
- (6) Bornscheuer, U. T.; Kazlauskas, R. J. *Angew. Chem. Int. Ed.* **2004**, *43*, 6032.
- (7) Park, H.-S.; Nam, S.-H.; Lee, J. K.; Yoon, C. N.; Mannervik, B.; Benkovic, S. J.; Kim, H.-S. *Science* **2006**, *311*, 535.
- (8) Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. *Science* **2013**, *339*, 307.
- (9) Toscano, M. D.; Woycechowsky, K. J.; Hilvert, D. *Angew. Chem. Int. Ed.* **2007**, *46*, 3212.
- (10) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855.
- (11) Deuss, P. J.; den Heeten, R.; Laan, W.; Kamer, P. C. J. *Chemistry – A European Journal* **2011**, *17*, 4680.
- (12) Rosati, F.; Roelfes, G. *ChemCatChem* **2010**, *2*, 916.
- (13) Ward, T. R. *Acc. Chem. Res.* **2011**, *44*, 47.
- (14) Ueno, T.; Tabe, H.; Tanaka, Y. *Chemistry – An Asian Journal* **2013**, *8*, 1646.
- (15) Lewis, J. C. *ACS Catalysis* **2013**, *3*, 2954.
- (16) Heinisch, T.; Ward, T. R. *Curr. Opin. Chem. Biol.* **2010**, *14*, 184.
- (17) Roelfes, G.; Feringa, B. L. *Angew. Chem.* **2005**, *117*, 3294.
- (18) Yamamura, K.; Kaiser, E. T. *J. Chem. Soc., Chem. Commun.* **1976**, 830.
- (19) Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**, *100*, 306.
- (20) Lin, C.-C.; Lin, C.-W.; Chan, A. S. C. *Tetrahedron: Asymmetry* **1999**, *10*, 1887.
- (21) Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. *Science* **2012**, *338*, 500.
- (22) Kaiser, E.; Lawrence, D. *Science* **1984**, *226*, 505.
- (23) Levine, H. L.; Nakagawa, Y.; Kaiser, E. T. *Biochem. Biophys. Res. Commun.* **1977**, *76*, 64.
- (24) Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**, *0*, 4239.
- (25) Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *J. Am. Chem. Soc.* **2004**, *126*, 14411.

- (26) Klein, G.; Humbert, N.; Gradinaru, J.; Ivanova, A.; Gilardoni, F.; Rusbandi, U. E.; Ward, T. R. *Angew. Chem. Int. Ed.* **2005**, *44*, 7764.
- (27) Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2006**, *128*, 8320.
- (28) Creus, M.; Ward, T. R. *Organic & Biomolecular Chemistry* **2007**, *5*, 1835.
- (29) Colowick, S. P.; Kaplan, N. O.; Wilchek, M.; Bayer, E. A. *Methods in Enzymology: Avidin-biotin technology / edited by Meir Wilchek, Edward A. Bayer*; Academic Press, 1990.
- (30) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *CHIMIA International Journal for Chemistry* **2002**, *56*, 721.
- (31) Reetz, M. T. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 5716.
- (32) Reetz, M. T.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun.* **2006**, 4318.
- (33) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**, *101*, 3081.
- (34) Tann, C.-M.; Qi, D.; Distefano, M. D. *Curr. Opin. Chem. Biol.* **2001**, *5*, 696.
- (35) DeSantis, G.; Jones, J. B. *Curr. Opin. Biotechnol.* **1999**, *10*, 324.
- (36) Davis, B. G. *Curr. Opin. Biotechnol.* **2003**, *14*, 379.
- (37) Berglund, P.; DeSantis, G.; Stabile, M. R.; Shang, X.; Gold, M.; Bott, R. R.; Graycar, T. P.; Lau, T. H.; Mitchinson, C.; Jones, J. B. *J. Am. Chem. Soc.* **1997**, *119*, 5265.
- (38) DeSantis, G.; Shang, X.; Jones, J. B. *Biochemistry* **1999**, *38*, 13391.
- (39) Plettner, E.; Khumtaveeporn, K.; Xiao, S.; Jones, J. B. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2291.
- (40) Díaz-Rodríguez, A.; Davis, B. G. *Curr. Opin. Chem. Biol.* **2011**, *15*, 211.
- (41) Reetz, M. T. *Tetrahedron* **2002**, *58*, 6595.
- (42) Kries, H.; Blomberg, R.; Hilvert, D. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221.
- (43) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387.
- (44) Nanda, V.; Koder, R. L. *Nat Chem* **2010**, *2*, 15.
- (45) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190.
- (46) Khare, S. D.; Kipnis, Y.; Greisen, P., Jr.; Takeuchi, R.; Ashani, Y.; Goldsmith, M.; Song, Y.; Gallaher, J. L.; Silman, I.; Leader, H.; Sussman, J. L.; Stoddard, B. L.; Tawfik, D. S.; Baker, D. *Nat Chem Biol* **2012**, *8*, 294.

- (47) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St.Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, 329, 309.
- (48) Mills, J. H.; Khare, S. D.; Bolduc, J. M.; Forouhar, F.; Mulligan, V. K.; Lew, S.; Seetharaman, J.; Tong, L.; Stoddard, B. L.; Baker, D. *J. Am. Chem. Soc.* **2013**, 135, 13393.
- (49) Cherrier, M.; Girgenti, E.; Amara, P.; Iannello, M.; Marchi-Delapierre, C.; Fontecilla-Camps, J.; Ménage, S.; Cavazza, C. *JBIC Journal of Biological Inorganic Chemistry* **2012**, 17, 817.
- (50) Hakansson, K.; Wehnert, A.; Liljas, A. *Acta Crystallographica Section D* **1994**, 50, 93.
- (51) Cavazza, C.; Bochot, C.; Rousselot-Pailley, P.; Carpentier, P.; Cherrier, M. V.; Martin, L.; Marchi-Delapierre, C.; Fontecilla-Camps, J. C.; Ménage, S. *Nat Chem* **2010**, 2, 1069.
- (52) Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angew. Chem. Int. Ed.* **2011**, 50, 3026.
- (53) Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; LeTrong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem.* **2008**, 120, 1422.
- (54) Ueno, T.; Abe, M.; Hirata, K.; Abe, S.; Suzuki, M.; Shimizu, N.; Yamamoto, M.; Takata, M.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, 131, 5094.
- (55) Abe, S.; Hirata, K.; Ueno, T.; Morino, K.; Shimizu, N.; Yamamoto, M.; Takata, M.; Yashima, E.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, 131, 6958.
- (56) Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, 130, 10512.
- (57) Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**, 127, 6556.
- (58) Ueno, T.; Ohashi, M.; Kono, M.; Kondo, K.; Suzuki, A.; Yamane, T.; Watanabe, Y. *Inorg. Chem.* **2004**, 43, 2852.
- (59) Satake, Y.; Abe, S.; Okazaki, S.; Ban, N.; Hikage, T.; Ueno, T.; Nakajima, H.; Suzuki, A.; Yamane, T.; Nishiyama, H.; Watanabe, Y. *Organometallics* **2007**, 26, 4904.
- (60) Zimbron, J. M.; Heinisch, T.; Schmid, M.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *J. Am. Chem. Soc.* **2013**.
- (61) Ory, J. J.; Mazhary, A.; Kuang, H.; Davies, R. R.; Distefano, M. D.; Banaszak, L. J. *Protein Eng.* **1998**, 11, 253.
- (62) Fukumoto, K.; Onoda, A.; Mizohata, E.; Bocola, M.; Inoue, T.; Schwaneberg, U.; Hayashi, T. *ChemCatChem* **2014**, 6, 1229.
- (63) Davies, R. R.; Distefano, M. D. *J. Am. Chem. Soc.* **1997**, 119, 11643.
- (64) Davies, R. R.; Kuang, H.; Qi, D.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**, 9, 79.

- (65) Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *J. Am. Chem. Soc.* **2003**, *125*, 9030.
- (66) Rusbandi, U. E.; Lo, C.; Skander, M.; Ivanova, A.; Creus, M.; Humbert, N.; Ward, T. R. *Adv. Synth. Catal.* **2007**, *349*, 1923.
- (67) Collot, J.; Humbert, N.; Skander, M.; Klein, G.; Ward, T. R. *J. Organomet. Chem.* **2004**, *689*, 4868.
- (68) Ueno, T.; Suzuki, M.; Goto, T.; Matsumoto, T.; Nagayama, K.; Watanabe, Y. *Angew. Chem. Int. Ed.* **2004**, *43*, 2527.
- (69) Köhler, V.; Wilson, Y. M.; Dürrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knörr, L.; Häussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R. *Nat Chem* **2013**, *5*, 93.
- (70) Letondor, C.; Humbert, N.; Ward, T. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 4683.
- (71) Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**, 4239.
- (72) Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem.* **2003**, *115*, 1035.
- (73) Ueno, T.; Koshiyama, T.; Abe, S.; Yokoi, N.; Ohashi, M.; Nakajima, H.; Watanabe, Y. *J. Organomet. Chem.* **2007**, *692*, 142.
- (74) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. *J. Am. Chem. Soc.* **2004**, *126*, 10812.
- (75) Zhang, J.-L.; Garner, D. K.; Liang, L.; Chen, Q.; Lu, Y. *Chem. Commun.* **2008**, 1665.
- (76) Garner, D. K.; Liang, L.; Barrios, D. A.; Zhang, J.-L.; Lu, Y. *ACS Catalysis* **2011**, *1*, 1083.
- (77) van de Velde, F.; Konemann, L. *Chem. Commun.* **1998**, 1891.
- (78) Pordea, A.; Creus, M.; Panek, J.; Duboc, C.; Mathis, D.; Novic, M.; Ward, T. R. *J. Am. Chem. Soc.* **2008**, *130*, 8085.
- (79) Okrasa, K.; Kazlauskas, R. *Chemistry – A European Journal* **2006**, *12*, 1587.
- (80) Fernández-Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumilion, P. *ChemBioChem* **2006**, *7*, 1013.
- (81) Ricoux, R.; Dubuc, R.; Dupont, C.; Marechal, J.-D.; Martin, A.; Sellier, M.; Mahy, J.-P. *Bioconjugate Chem.* **2008**, *19*, 899.
- (82) Allard, M.; Dupont, C.; Muñoz Robles, V.; Doucet, N.; Lledós, A.; Maréchal, J.-D.; Urvoas, A.; Mahy, J.-P.; Ricoux, R. *ChemBioChem* **2012**, *13*, 240.
- (83) Kanbak-Aksu, S.; Nahid Hasan, M.; Hagen, W. R.; Hollmann, F.; Sordi, D.; Sheldon, R. A.; Arends, I. W. C. E. *Chem. Commun.* **2012**, *48*, 5745.
- (84) Kokubo, T.; Sugimoto, T.; Uchida, T.; Tanimoto, S.; Okano, M. *J. Chem. Soc., Chem. Commun.* **1983**, 769.
- (85) Köhler, V.; Mao, J.; Heinisch, T.; Pordea, A.; Sardo, A.; Wilson, Y. M.; Knörr, L.; Creus, M.; Prost, J.-C.; Schirmer, T.; Ward, T. R. *Angew. Chem. Int. Ed.* **2011**, *50*, 10863.

- (86) Boersma, A. J.; Coquière, D.; Geerdink, D.; Rosati, F.; Feringa, B. L.; Roelfes, G. *Nat Chem* **2010**, 2, 991.
- (87) Bos, J.; Garcia-Herraz, A.; Roelfes, G. *Chemical Science* **2013**, 4, 3578.
- (88) Bertucci, C.; Botteghi, C.; Giunta, D.; Marchetti, M.; Paganelli, S. *Adv. Synth. Catal.* **2002**, 344, 556.
- (89) Jing, Q.; Kazlauskas, R. J. *ChemCatChem* **2010**, 2, 953.
- (90) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. *Chemistry – A European Journal* **2009**, 15, 1370.
- (91) Laan, W.; Muñoz, B. K.; den Heeten, R.; Kamer, P. C. J. *ChemBioChem* **2010**, 11, 1236.
- (92) Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angew. Chem. Int. Ed.* **2008**, 47, 701.
- (93) Mayer, C.; Gillingham, D. G.; Ward, T. R.; Hilvert, D. *Chem. Commun.* **2011**, 47, 12068.
- (94) Lo, C.; Ringenberg, M. R.; Gnanndt, D.; Wilson, Y.; Ward, T. R. *Chem. Commun.* **2011**, 47, 12065.
- (95) Matsuo, T.; Imai, C.; Yoshida, T.; Saito, T.; Hayashi, T.; Hirota, S. *Chem. Commun.* **2012**, 48, 1662.
- (96) Philippart, F.; Arlt, M.; Gotzen, S.; Tenne, S.-J.; Bocola, M.; Chen, H.-H.; Zhu, L.; Schwaneberg, U.; Okuda, J. *Chemistry – A European Journal* **2013**, 19, 13865.
- (97) Hilvert, D.; Hill, K. W.; Nared, K. D.; Auditor, M. T. M. *J. Am. Chem. Soc.* **1989**, 111, 9261.
- (98) Braisted, A. C.; Schultz, P. G. *J. Am. Chem. Soc.* **1990**, 112, 7430.
- (99) Chen, J.; Deng, Q.; Wang, R.; Houk, K. N.; Hilvert, D. *ChemBioChem* **2000**, 1, 255.
- (100) Seelig, B.; Keiper, S.; Stuhlmann, F.; Jäschke, A. *Angew. Chem. Int. Ed.* **2000**, 39, 4576.
- (101) Coquière, D.; Bos, J.; Beld, J.; Roelfes, G. *Angew. Chem. Int. Ed.* **2009**, 48, 5159.
- (102) Reetz, M. T. *The Chemical Record* **2012**, 12, 391.
- (103) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. *Angew. Chem. Int. Ed.* **2012**, 51, 7472.
- (104) Roelfes, G.; Boersma, A. J.; Feringa, B. L. *Chem. Commun.* **2006**, 635.
- (105) Reetz, M. T.; Jiao, N. *Angew. Chem. Int. Ed.* **2006**, 45, 2416.
- (106) Talbi, B.; Haquette, P.; Martel, A.; de Montigny, F.; Fosse, C.; Cordier, S.; Roisnel, T.; Jaouen, G.; Salmain, M. *Dalton Transactions* **2010**, 39, 5605.
- (107) Deuss, P. J.; Popa, G.; Slawin, A. M. Z.; Laan, W.; Kamer, P. C. J. *ChemCatChem* **2013**, n/a.
- (108) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. *Angew. Chem.* **2012**, 124, 7590.
- (109) Coquière, D.; Feringa, B. L.; Roelfes, G. *Angew. Chem.* **2007**, 119, 9468.
- (110) Boersma, A. J.; Feringa, B. L.; Roelfes, G. *Angew. Chem.* **2009**, 121, 3396.
- (111) Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *J. Am. Chem. Soc.* **2008**, 130, 11783.
- (112) Boersma, A. J.; de Bruin, B.; Feringa, B. L.; Roelfes, G. *Chem. Commun.* **2012**, 48, 2394.

- (113) Oltra, N. S.; Roelfes, G. *Chem. Commun.* **2008**, 6039.
- (114) Fournier, P.; Fiammengio, R.; Jäschke, A. *Angew. Chem. Int. Ed.* **2009**, 48, 4426.
- (115) Park, S.; Sugiyama, H. *Angew. Chem. Int. Ed.* **2010**, 49, 3870.
- (116) Boersma, A. J.; Megens, R. P.; Feringa, B. L.; Roelfes, G. *Chem. Soc. Rev.* **2010**, 39, 2083.
- (117) Mizoroki, T.; Mori, K.; Ozaki, A. *Bull. Chem. Soc. Jpn.* **1971**, 44, 581.
- (118) Heck, R. F.; Nolley, J. P. *The Journal of Organic Chemistry* **1972**, 37, 2320.
- (119) Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, 95, 2457.
- (120) Jung, Y. C.; Mishra, R. K.; Yoon, C. H.; Jung, K. W. *Org. Lett.* **2003**, 5, 2231.
- (121) Andappan, M. M. S.; Nilsson, P.; Larhed, M. *Chem. Commun.* **2004**, 218.
- (122) Andappan, M. S.; Nilsson, P.; Larhed, M. *Mol Divers* **2003**, 7, 97.
- (123) Yoo, K. S.; Yoon, C. H.; Jung, K. W. *J. Am. Chem. Soc.* **2006**, 128, 16384.
- (124) Yoo, K. S.; Park, C. P.; Yoon, C. H.; Sakaguchi, S.; O'Neil, J.; Jung, K. W. *Org. Lett.* **2007**, 9, 3933.
- (125) Enquist, P.-A.; Nilsson, P.; Sjöberg, P.; Larhed, M. *The Journal of Organic Chemistry* **2006**, 71, 8779.
- (126) Stahl, S. S.; Thorman, J. L.; Nelson, R. C.; Kozee, M. A. *J. Am. Chem. Soc.* **2001**, 123, 7188.
- (127) Thiel, W. R. *Angew. Chem. Int. Ed.* **1999**, 38, 3157.
- (128) Amatore, C.; Jutand, A.; Le Duc, G. *Chemistry – A European Journal* **2011**, 17, 2492.
- (129) Amatore, C.; Jutand, A. *Acc. Chem. Res.* **2000**, 33, 314.
- (130) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, 131, 16346.
- (131) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, 47, 1698.
- (132) Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. *Angew. Chem. Int. Ed.* **2013**, 52, 3916.
- (133) Gao, Z.; Gouverneur, V.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, 135, 13612.
- (134) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2013**, 49, 2747.
- (135) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.* **2006**, 10, 253.
- (136) Spicer, C. D.; Davis, B. G. *Nat Commun* **2014**, 5.
- (137) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. *ChemBioChem* **2006**, 7, 134.
- (138) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. *ChemBioChem* **2007**, 8, 232.
- (139) Ourailidou, M. E.; van der Meer, J.-Y.; Baas, B.-J.; Jeronimus-Stratingh, M.; Gottumukkala, A. L.; Poelarends, G. J.; Minnaard, A. J.; Dekker, F. J. *ChemBioChem* **2014**, 15, 209.

- (140) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220.
- (141) Spicer, C. D.; Triemer, T.; Davis, B. G. *J. Am. Chem. Soc.* **2011**, *134*, 800.
- (142) Wang, Y.-S.; Russell, W. K.; Wang, Z.; Wan, W.; Dodd, L. E.; Pai, P.-J.; Russell, D. H.; Liu, W. R. *Molecular BioSystems* **2011**, *7*, 714.
- (143) Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. *J. Am. Chem. Soc.* **2011**, *133*, 15316.
- (144) Li, J.; Lin, S.; Wang, J.; Jia, S.; Yang, M.; Hao, Z.; Zhang, X.; Chen, P. R. *J. Am. Chem. Soc.* **2013**, *135*, 7330.
- (145) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642.
- (146) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714.
- (147) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *J. Am. Chem. Soc.* **2010**, *132*, 16805.
- (148) Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 12156.
- (149) Wells, J. A.; Estell, D. A. *Trends Biochem. Sci* **1988**, *13*, 291.
- (150) Bode, W.; Papamokos, E.; Musil, D. *Eur. J. Biochem.* **1987**, *166*, 673.
- (151) McPhalen, C. A.; James, M. N. G. *Biochemistry* **1988**, *27*, 6582.
- (152) Neet, K. E.; Koshland, D. E. *Proceedings of the National Academy of Sciences of the United States of America* **1966**, *56*, 1606.
- (153) Polgar, L.; Bender, M. L. *J. Am. Chem. Soc.* **1966**, *88*, 3153.
- (154) Nakatsuka, T.; Sasaki, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1987**, *109*, 3808.
- (155) Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151.
- (156) Chang, T. K.; Jackson, D. Y.; Burnier, J. P.; Wells, J. A. *Proceedings of the National Academy of Sciences* **1994**, *91*, 12544.
- (157) Wu, Z. P.; Hilvert, D. *J. Am. Chem. Soc.* **1989**, *111*, 4513.
- (158) DeSantis, G.; Jones, J. B. *Biorg. Med. Chem.* **1999**, *7*, 1381.
- (159) Plettner, E.; DeSantis, G.; Stabile, M. R.; Jones, J. B. *J. Am. Chem. Soc.* **1999**, *121*, 4977.
- (160) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052.
- (161) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chemical Science* **2011**, *2*, 1666.
- (162) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835.
- (163) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. *Acc. Chem. Res.* **2011**, *44*, 730.

Chapter 2

Design and construction of the SBL-based artificial metalloenzymes

2.1 Introduction

Different modes of anchoring are available for design of artificial metalloenzymes, namely non-covalent supramolecular, dative/direct metal binding and covalent anchoring of organometallic complexes (as described in Chapter 1 Section 1.2.2). The ligand/cofactor used to bind metal is an important component in the design of artificial metalloenzymes as it influences both the overall reactivity and the selectivity of the artificial metalloenzyme. Moreover, the ligand plays an important role in localising the metal in the desired pocket of the enzyme.

Although recombinant DNA technology has enabled high fidelity manipulation of protein sequence space, the nature of alteration in this case is typically limited to 20 common amino acids. On the other hand, the concept of site-specific chemical modification of amino acids has facilitated the incorporation of complex structures within native protein scaffolds. Because of its unique reactivity, cysteine is often considered the ‘amino acid of choice’ for site-selective chemical modification of proteins. The high nucleophilicity and natural low abundance of cysteine residues have previously been exploited to generate protein conjugates.^{1,2} In this regard, combination of site-directed mutagenesis and chemical modification at cysteine is often a straightforward and useful methodology for site-selective installation of chemical structures into proteins. Direct alkylation and

conjugate addition to Michael acceptors are conventional methods to selectively modify cysteine residues and have also been employed previously for generating artificial metalloenzymes.³⁻⁵

In general, the inherent high metal affinity of bidentate *N,N*-chelating moieties has been widely used to install metal centres using supramolecular or covalent anchoring,^{3,4,6-12} even through genetic incorporation of an unnatural (2,2'-bipyridin-5-yl)alanine (Bpy-Ala) residue.^{13,14} Lee *et al.* subsequently incorporated Bpy-Ala into a catabolite activator protein (CAP) to synthesize copper-dependent site-specific DNA cleavage enzymes.¹⁵ More recently, Baker and co-workers demonstrated that computational methods can be used to precisely design metalloproteins and control the coordination environment around a metal bound by a genetically encoded bidentate Bpy-Ala residue.¹⁶

Hence, site-specific installation of these metal binding ligands is a pre-requisite. We decided to explore covalent chemical modification strategies to install a metal complex into SBL. Although, the nucleophilic interactions of proteases (due to active site Cys or Ser residues) have been widely used to covalently install a metal complex, this renders the enzyme inactive towards their natural substrates. In order to cause minimum alteration of the enzyme's catalytic residues and binding pockets, we decided to chemically install a metal binding amino acid in the active site by using two alternative covalent chemical modification techniques, namely direct alkylation of cysteine and conjugate addition at dehydroalanine (Dha) as discussed in the next section.

2.2 Results and discussion[†]

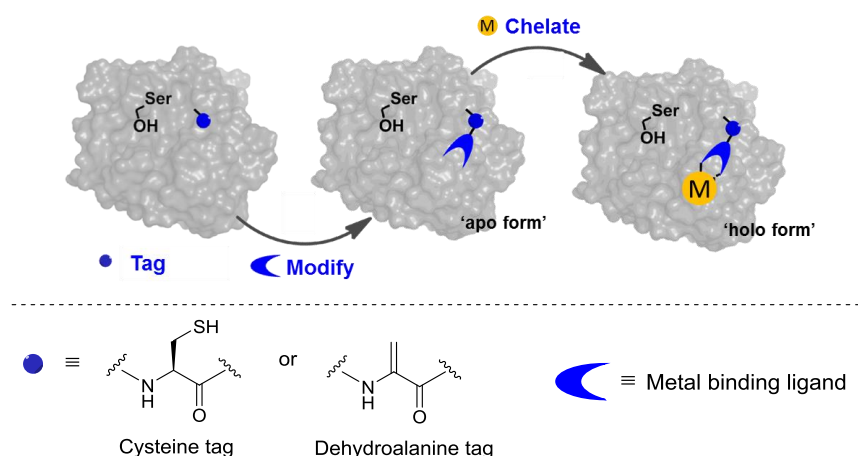
2.2.1 Design of SBL-based metalloenzymes

The design of the proposed biocatalyst relies on the combination of site-directed mutagenesis and chemical modification methodologies. Chemical installation of metal binding amino acids in the enzyme active site was envisaged without the use of long linker moieties to keep the metal centre as close as possible to the natural substrate binding pockets of the enzyme. We envisioned that the metal binding site could be introduced in the enzyme active site by installing a uniquely reactive cysteine residue for bio-orthogonal reaction with a metal chelating ligand in a ‘tag-and-modify’ fashion (Scheme 2-1).^{17,18} This approach enables chemical inter-conversion to unnatural amino acids, and affords variability at different levels of building-up of the artificial metalloenzymes, as described below.

- (i) *Model protein scaffold and mode of anchoring:* A robust protein scaffold that can be used to install metal complexes is an important pre-requisite while designing artificial metalloenzymes. Quite often, the choice of the scaffold depends on the mode of anchoring to be used for incorporation of metal complexes (Chapter 1 Section 1.2). For example, supramolecular anchoring approaches using biotin-(Strept)avidin technology often limits the choice to biotin binding proteins only. On the other hand, covalent protein modification approaches could theoretically be used to modify any protein scaffold.

[†]Dr. Anca Pordea and Dr. Alba Díaz-Rodríguez conducted exploratory work in our laboratory to create artificial SBL-based metalloenzymes. Experimental results that were obtained by them are indicated in the body of this chapter and their contributions are stated explicitly in the experimental sections of this chapter.

- (ii) *Tag*: A single reactive amino acid residue can serve as a ‘chemical tag’ for site-selective installation of an organometallic complex. Site-directed mutagenesis could be used to install a single reactive cysteine residue at any site on a given protein. Alternatively, chemical mutagenesis¹⁹⁻²¹ to inter-convert cysteine to dehydroalanine (Dha) could be used to install a reactive ‘Dha tag’²² (exploiting protein sequence space).
- (iii) *Modify*: Modification with different ligand moieties at the reactive amino acid ‘tag’ can be used to install different metal binding unnatural amino acids (exploiting ligand space), thereby forming the metal-free apo form of the desired metalloenzyme.
- (iv) *Chelate*: Selective binding of different metals on the installed unnatural amino acid would yield the metal-containing holo form, which can be employed for screening various cross-coupling reactions.



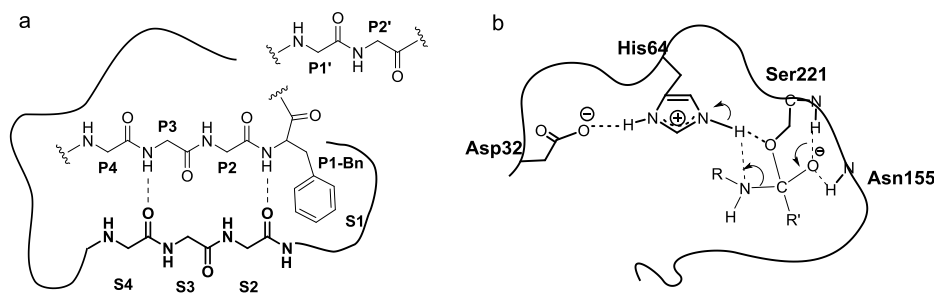
Scheme 2-1. Tag-and-modify strategy showing bio-orthogonal reaction of metal chelating ligand with a reactive Cys or Dha tag, followed by coordination of metal to form the corresponding holoenzyme

2.2.1.1 Choice of model protein scaffold

Subtilisin from *Bacillus lentus* (SBL) was chosen as the model protein scaffold mainly because of its robustness, operational stability and availability of rich structural information.²³⁻²⁶ The innate specificity/selectivity of SBL towards a wide range of substrates was an important consideration for the rational design of the proposed artificial metalloenzyme. Moreover, the active site of this serine protease is pre-equipped with the catalytic machinery (Asp32-His64-Ser221) and the stabilizing oxyanion hole site that offers an activating microenvironment for exploring catalysis of non-native reactions. The cysteine-free SBL wild-type (WT) would enable utilization of site-selective chemical modification strategies. In fact, the active site and substrate specificity of subtilisin have been extensively probed and tailored heretofore by chemical modification strategies.²⁷⁻³³ From a practical standpoint, convenient over-expression of many SBL variants²⁹ provides access to good amounts of proteins, required for extensive screening and further optimization of the system.

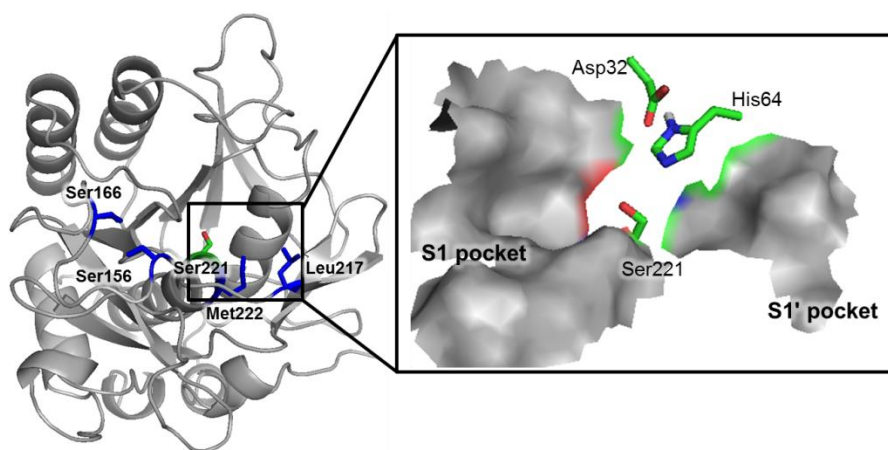
2.2.1.2 Design of metal binding site

A preliminary survey of the enzyme binding pockets was carried out to identify an appropriate installation site for a non-native metal. The goal of this endeavour was to combine the chemoselectivity and substrate specificity of a naturally evolved enzyme pocket with activity of a metal. Both sides of the catalytic centre of proteolytic enzymes (such as papain, subtilisin etc.) are lined with substrate binding subsites. These subsites are numbered as S2, S1, S1', Sn' according to the substrate amino acids P2, P1, P1', Pn' based on Schechter and Berger nomenclature³⁴ (Scheme 2-2). Subsite S1 forms the primary substrate specificity site on the amino terminal side, whereas S1' region hosts the leaving group of the native amide substrates for such enzymes.



Scheme 2-2. Substrate binding pockets and active site of SBL. (a) S1 binding pocket accommodating P1-benzyl (P1-Bn) group of the native peptide substrate; substrate amino acids are denoted as P2, P1, P1', Pn' and the corresponding enzyme binding sites as S2, S1, S1', Sn'. P1-P1' is the hydrolysed bond. (b) Acyl intermediate during hydrolysis of a peptide substrate by SBL; amino acid residues in the active site and Asn155 and main chain amide of Ser221 in the oxyanion hole site are labelled

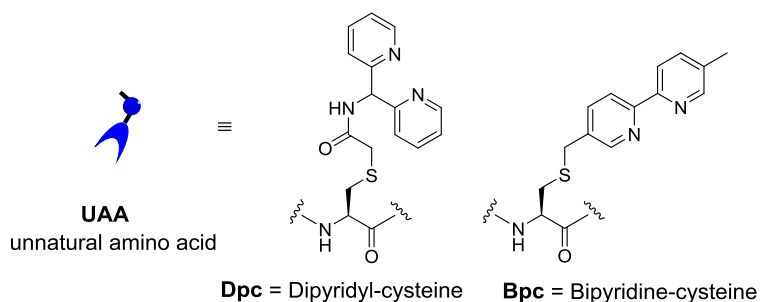
In the case of SBL, natural specificities of these binding pockets are dictated by several electrostatic, hydrophobic and H-bonding interactions of residues with the substrates. These interactions can be altered chemically, ultimately broadening the substrate scope and making natural enzymes synthetically more useful.^{27,32,35,36} The hydrophobic S1 pocket of SBL is lined by residues 125-127 complemented by a loop consisting of residues 155-166 on the other side^{37,38} (Scheme 2-3). S1' subsite is a surface depression consisting of Tyr217, Met222 and His64.^{37,38} Importantly, mapping of SBL substrate binding pockets revealed the hydrophobic crevice in S1 pocket gives rise to an innate selectivity for phenyl/benzyl substrates, whereas S1' has relatively broader specificity.³⁷ For our purposes, we decided to utilize the interplay between the S1 and S1' pockets to create a microenvironment amenable for hosting aryl substrates.



Scheme 2-3. Close-up view of S1 and S1' specificity pockets and the catalytic machinery (Asp32-His64-Ser221) of SBL (PDB ID: 1GCI)²⁵; surface view of the residues 125-127 and 151-156, Ser166 in the S1 pocket and Met222 and Leu217 in S1' pocket are shown

The rationale was that binding affinities of the S1 site could be exploited for localisation of substrates, while the metal catalyst could be accordingly positioned at the base of the S1' subsite (Scheme 2-3). Precise positioning of the metal catalyst in the active site was guided by the crystal structure of SBL-WT (PDB ID: 1GCI).²⁵ Positions 166 and 156 (in S1 pocket), 222 and 217 (in S1' subsite) were considered. The side-chain of Ser156 in the S1 pocket is surface exposed and well-solvated, whereas Ser166 is at the bottom of the S1 pocket. Prior chemical modifications at position 166 have led to altered binding specificities and dimensions of the S1 pocket.³¹ We decided to exploit the natural specificity of S1 pocket for hydrophobic substrates, so Ser166 was not considered suitable for installation of metal complex. Met222 in the S1' site is adjacent to the nucleophilic Ser221, therefore we envisioned metal installation at this position might lead to a constrained active site. Finally, based on the distances and orientations of the substrates and the metal in the active site, position 217 in the S1' pocket and 156 on a flexible loop in S1 pocket were chosen for the installation of metal binding amino acids.

Dipyridyl-based (pyridine rings with a methylene bridge) and bipyridine-based compounds are known to form five- and six- membered metal chelates with different transition metal ions, such as Pd^{2+} , Ni^{2+} , Cu^{2+} etc. Metal binding using these *N,N*-chelating moieties was considered as being ideal due to their oxidative stability, bi-dentate mode of metal coordination and catalytic versatility.^{39,40} Therefore, unnatural amino acids (UAAs) bipyridine-cysteine (Bpc) and dipyridyl-cysteine (Dpc) were designed (Scheme 2-4). Site-directed mutagenesis used in conjunction with chemical modification could enable site-selective installation of such moieties into SBL. Using these, the chelation and localisation of a potent metal catalyst, such as palladium, was envisioned.



Scheme 2-4. Designed metal binding unnatural amino acid residues (UAAs) Dpc and Bpc using dipyridyl- and bipyridine-based core structures respectively

2.2.2 Construction of SBL-based metalloenzymes

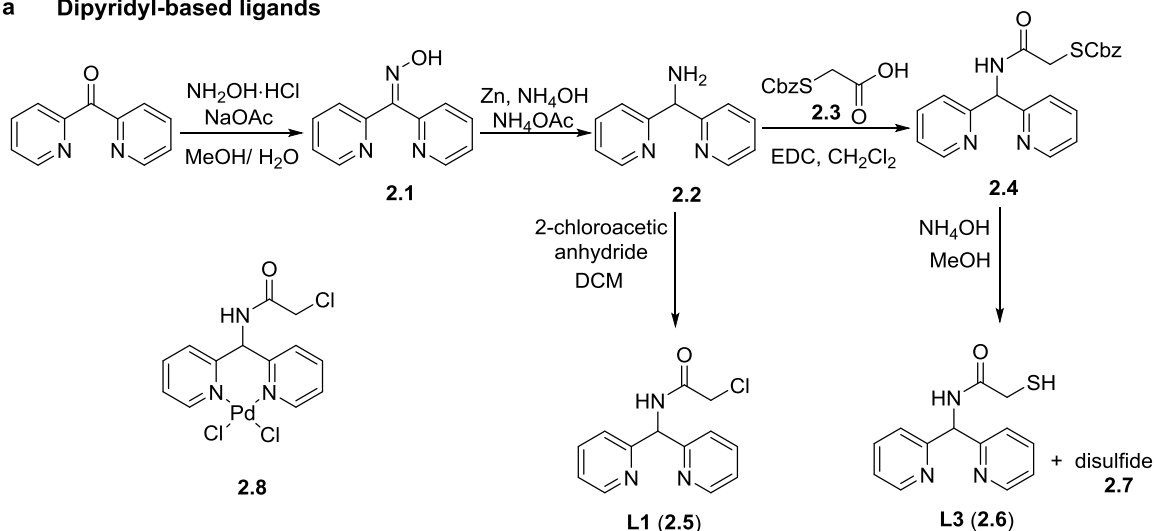
2.2.2.1 Synthesis of metal binding ligands

Two alternative ‘tag-and-modify’ strategies (using dehydroalanine (Dha) or cysteine (Cys) as chemical tags) were employed for the installation of Bpc and Dpc into SBL. The synthesis of the halogenated or thiol ligands **L1-L4** was carried out (Scheme 2-5). Synthesis of dipyridyl-based ligand **L1** (**2.5**) was accomplished in three steps, using 2,2'-dipyridyl ketone as the starting material⁴⁰ (Scheme 2-5a). The corresponding (2,2'-dipyridyl)methylamine **2.2** was obtained by Zn-mediated reduction of the oxime intermediate **2.1**.⁴⁰ This was followed by coupling of amine **2.2** with chloroacetic

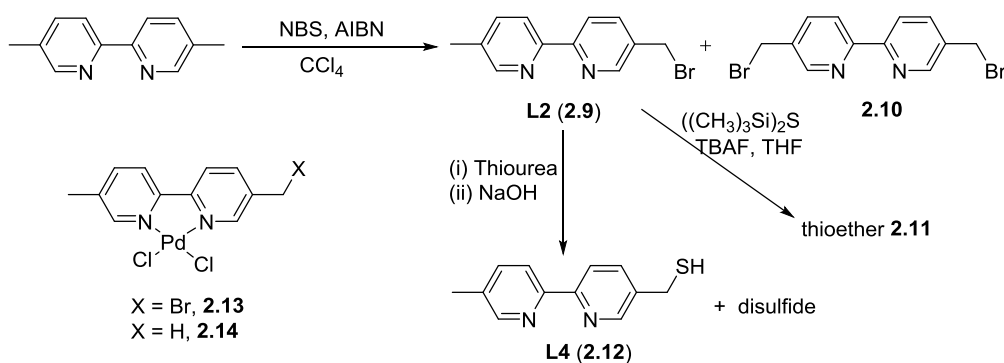
anhydride giving the final product **L1** in good yield. The dipyridyl-based thiol ligand **L3** (**2.6**) was prepared by coupling (2,2'-dipyridyl)methylamine **2.2** with the S-protected mercaptoacetic acid **2.3**, followed by deprotection of carbobenzyloxy (Cbz) group.

The synthesis of **L2** (**2.9**) was accomplished by a radical reaction where 5,5'-dimethyl-2,2'-bipyridine was treated with *N*-bromosuccinimide initiated using azobisisobutyronitrile (AIBN)⁴¹ (Scheme 2-5b). The reaction was quenched after 3 h to minimise the formation of bisfunctionalised derivative **2.10**. Given the difficult separation of the mono- and bisfunctionalised derivatives, the purification of **L2** by column chromatography gave the desired product in relatively moderate yields (44%) with appreciable amounts of dibromo by-product **2.10** (27%). The synthesis of ligand **L4** (**2.12**) using **L2**, by trimethylsilylthioxy dehalogenation, followed by *in situ* desilylation using a mixture of tetra-*n*-butylammonium fluoride (TBAF) and hexamethyldisilathiane in THF was attempted (Scheme 2-5b).^{42,43} However, this approach mainly yielded the corresponding thioether product **2.11** instead (See Experimental Section 2.5.3). Hence, a different approach was followed. The reaction of **L2** with thiourea, followed by hydrolysis of the isothiuronium intermediate gave the desired thiol ligand **L4**.^{44,45} Additionally, palladium complexes, **2.8**[†], **2.13** and **2.14**, were synthesized by reacting the corresponding ligands with a Pd(II) source such as PdCl₂.

a Dipyrrolyl-based ligands



b Bipyridine-based ligands

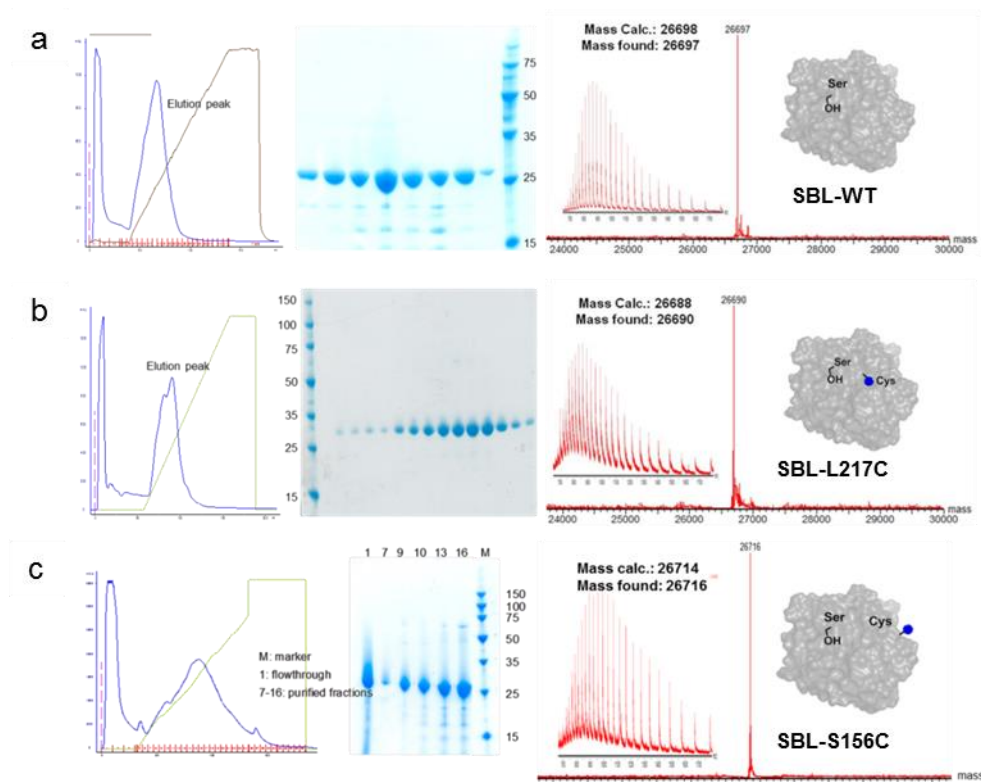


Scheme 2-5. Chemical routes for synthesis of (a) dipyrrolyl-based (b) bipyridine-based metal binding ligands **L1-L4**

2.2.2.2 Protein purification and characterization

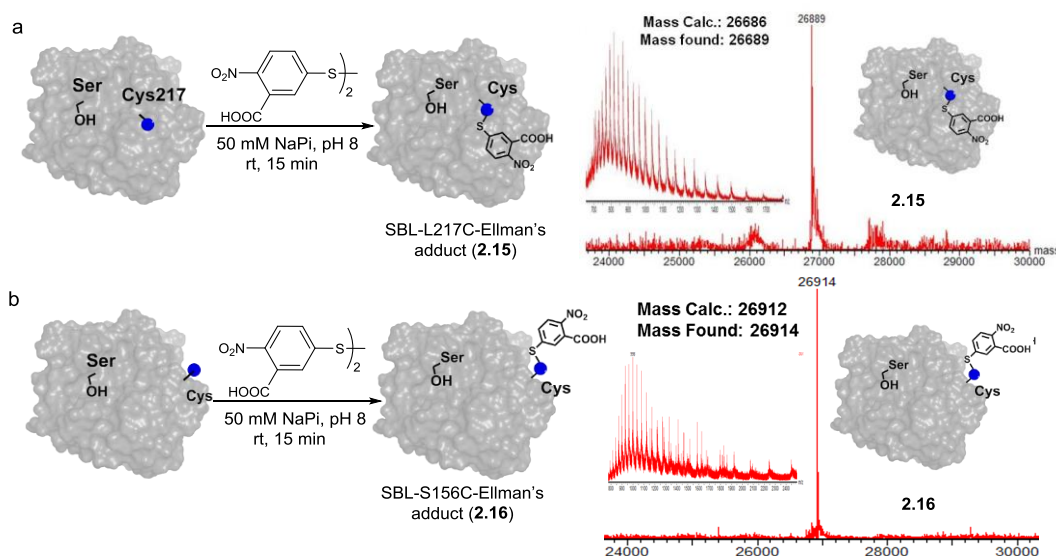
Advantageously, SBL-WT contains no natural cysteine, thereby allowing ready positioning of both tags *via* site-directed mutagenesis. SBL-L217C and SBL-S156C variants were used as the model enzymes. Previously, the gene for subtilisin from *B. lentus* was cloned into the bacteriophage M13mpl9 vector for mutagenesis and subsequently cloned into expression plasmid GGT274.³⁶ *B. subtilis* was used as host for expression of SBL-WT and SBL variants (crude proteins were stored with stabilizers).³⁶ The purification of SBL-WT, SBL-L217C and SBL-S156C[†] was accomplished by fast protein liquid chromatography (FPLC) using a strong cation exchange column (See Experimental

Section 2.5.5). Subsequently, the purity and homogeneity of SBL enzymes was confirmed by non-reducing PAGE gel and LC-MS analysis (Scheme 2-6).



Scheme 2-6. FPLC trace for the purification (left), non-reducing PAGE gel analysis of eluted fractions (middle) and ESI-MS spectra (right) of (a) SBL-WT (b) SBL-L217C (c) SBL-S156C[†]

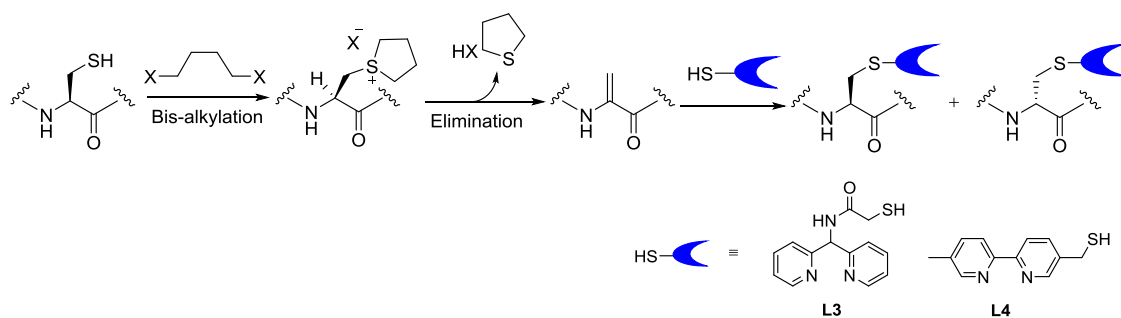
Ellman's test⁴⁶ using 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) is a rapid assay to quantify the number of cysteine residues in a protein sample. The reaction of SBL-L217C and SBL-S156C with Ellman's reagent confirmed the presence of a single cysteine on each by the formation of corresponding disulfides **2.15** and **2.16**, as detected by LC-MS analysis (Scheme 2-7). No modification was observed in the absence of the reactive thiol group when SBL-WT was treated with DTNB.



Scheme 2-7. Ellman's test for the variant proteins (a) SBL-L217C (b) SBL-S156C

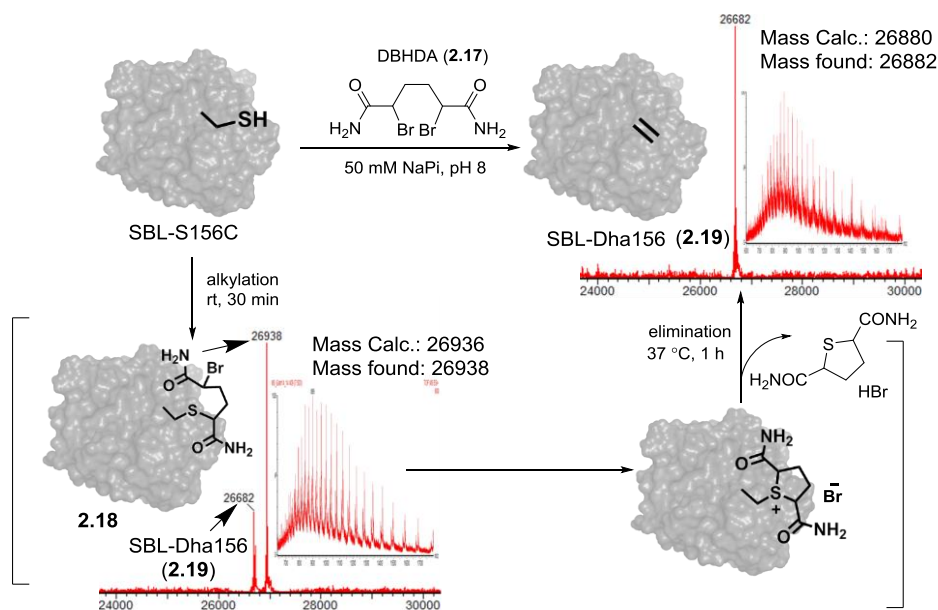
2.2.2.3 Artificial metalloenzymes using dehydroalanine (Dha) route

Chemical installation of Dpc and Bpc was envisaged *via* the Dha route using thiol ligands **L3** and **L4** respectively (Scheme 2-8). The 'Dha tag' can be installed through the direct chemical conversion of cysteine (Cys) to dehydroalanine (Dha).²² Once present, Dha serves as a Michael acceptor for a variety of thiol nucleophiles.^{14,18} One commonly employed method for the chemical introduction of Dha at Cys is through bis-alkylation to form the corresponding sulfonium salt which subsequently undergoes elimination.⁴⁷ While the bis-alkylation-elimination has been previously established as an effective synthetic route for installation of the Dha tag in peptide and protein substrates,⁴⁷ this process (in theory) results in a mixture of two possible diastereomers.^{48,49} Rapid and chemoselective reaction at the Dha tag has previously been used for installing diverse complex structures on the model single cysteine system SBL-S156C.^{22,47} Significantly, this selective method of employing an electrophilic Dha residue to install chemical structures into proteins has not yet been employed in the design of artificial metalloenzymes.



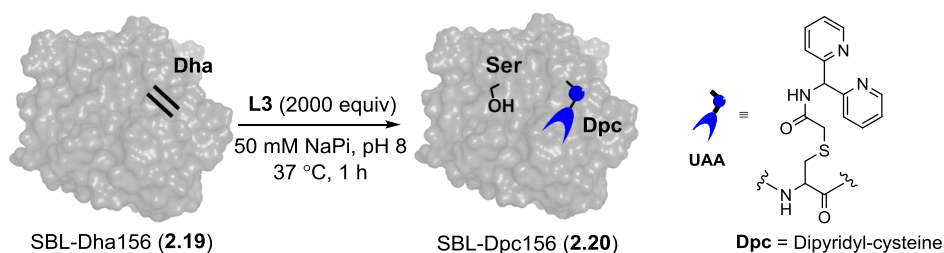
Scheme 2-8. Chemical route to create artificial metalloenzymes using Dha strategy

The bis-alkylation-elimination of Cys156 in SBL-S156C was carried out by treating with an excess of 2,5-dibromohexanediamide reagent (DBHDA, **2.17**)⁴⁷ to give SBL-Dha156 (Scheme 2-9). The reaction proceeded heterogeneously as the DBHDA (**2.17**) was added as a solid and was found to be quite insoluble under aqueous conditions. Even so, rapid and complete consumption of starting material to form the mono-alkylated intermediate **2.18** and 24% SBL-Dha156 was observed after incubation at room temperature for 30 min. By raising the reaction temperature to 37 °C for the elimination step, full conversion to SBL-Dha156 **2.19** was observed. No reaction was observed by treating SBL-Dha156 with 200 equiv. DTNB (Ellman's reagent) by LC-MS analysis. This confirmed the full consumption of cysteine thiol. In addition to this, the presence of Dha was successfully confirmed by conjugate addition of a simple thiol compound to SBL-Dha156. On treatment of SBL-Dha156 with 200 equiv. of β -mercapto ethanol (BME), a mass corresponding to the corresponding thio-ether conjugate (SBL-Dha156-BME) was observed (See Experimental Section 2.5.6).



Scheme 2-9. Synthesis of SBL-Dha156 by bis-alkylation-elimination route using DBHDA

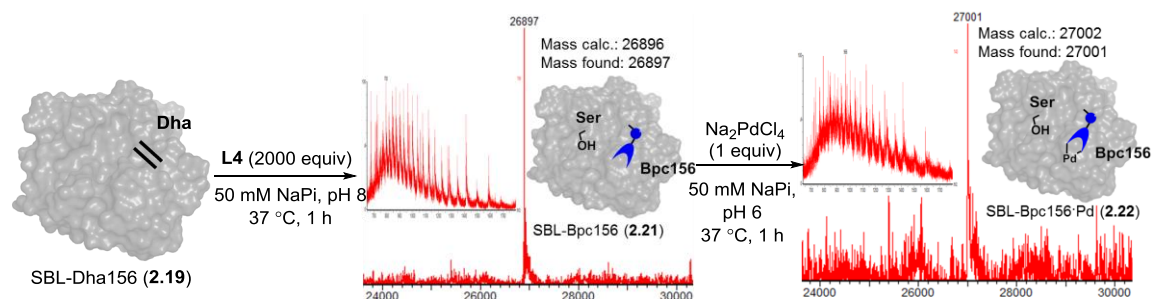
With SBL-Dha156 in hand, the conjugate addition of dipyridyl-based thiol ligand to the electrophilic Dha156 was attempted using **L3** (Scheme 2-10). We were pleased to observe the incorporation of this ligand to form the metal free apoenzyme SBL-Dpc156 (**2.20**). The reaction was found to be fast and chemoselective leading to successful installation of metal binding amino acid Dpc into SBL-S156C (See Experimental Section 2.5.6).



Scheme 2-10. Synthesis SBL-Dpc156 using Dha strategy

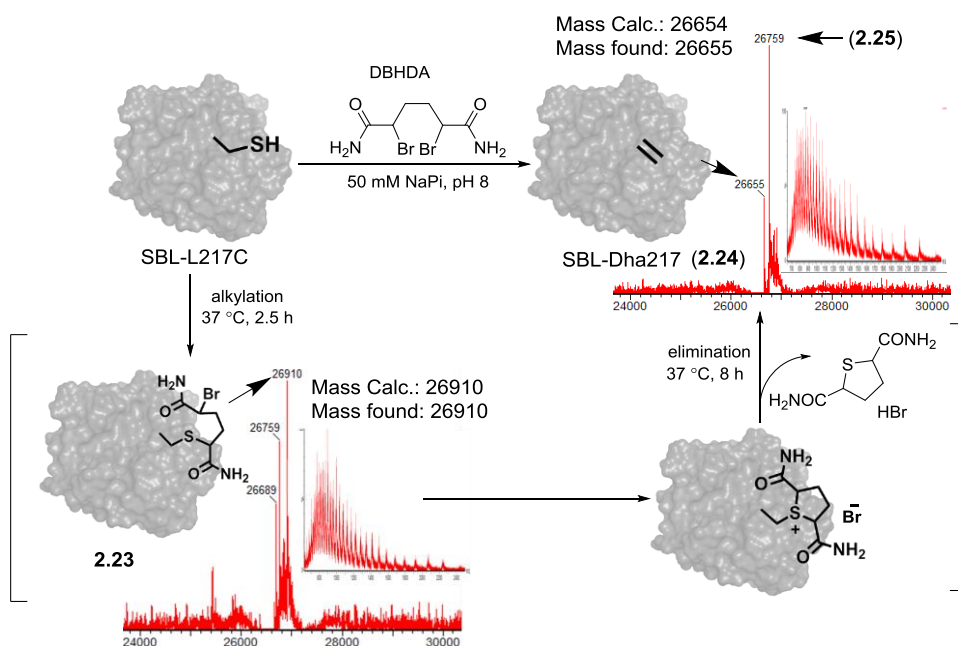
Similar behaviour was observed when ligand **L4** was reacted with SBL-Dha156, when complete conversion to SBL-Bpc156 (**2.21**) was achieved after 1 h of reaction at 37 °C (Scheme 2-11). Crucially, subsequent metal complexation was also successfully carried out to obtain the SBL-based artificial metalloenzyme. On incubation of SBL-Bpc156 with

Pd(II) source Na_2PdCl_4 at 37 °C for 45 min, a mass corresponding to SBL-BpC156•Pd (2.22) was observed showing >95% conversion to the metal containing holoenzyme (See Experimental Section 2.5.6).



Scheme 2-11. Chemical route to create SBL-Bpc156 and SBL-Bpc156•Pd using Dha strategy

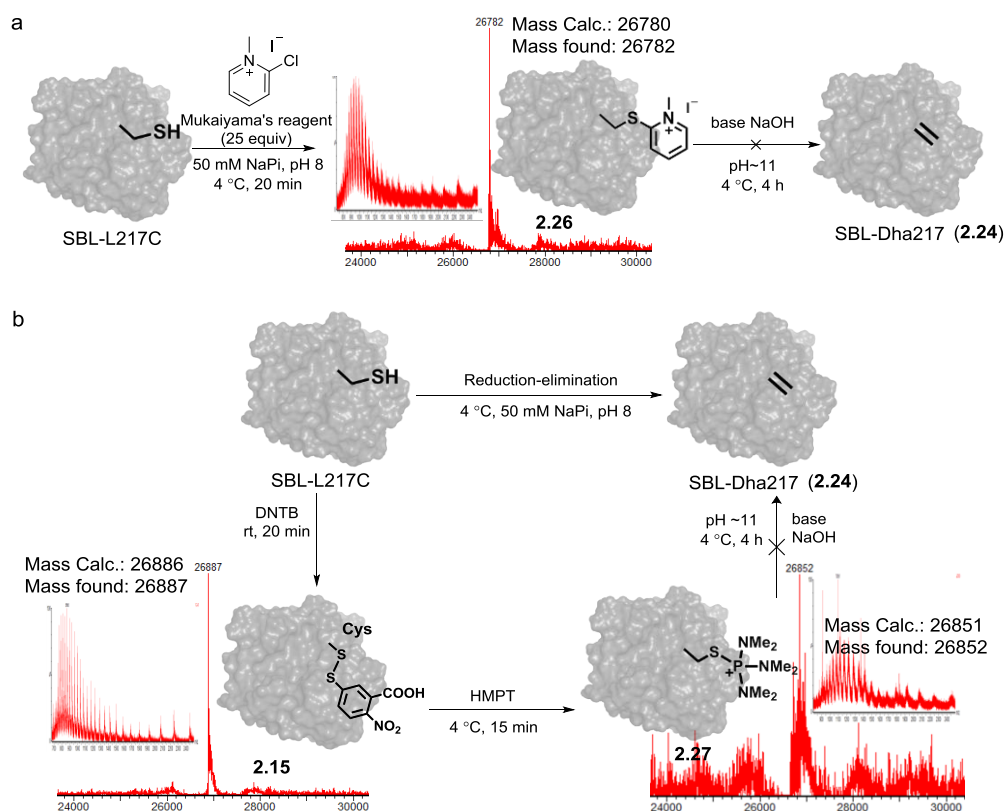
The bis-alkylation-elimination route was then used for another SBL variant SBL-L217C. Attempts using an excess of DBHDA (2.17) resulted in the formation of the corresponding mono-alkylated product 2.23 (Scheme 2-12). When compared to reaction of DBHDA with SBL-S156C, the reaction at Cys217 (present at the base of the active site) was much slower. Mono-alkylation of SBL-L217C was not complete even after 2.5 h and required higher temperature). However, complete consumption of the starting Cys217 thiol was observed after 8 h of incubation at 37 °C. LC-MS analysis of the resulting mixture revealed only 26% conversion to SBL-Dha217 (2.24). Interestingly, a major peak with mass corresponding to 26759 Da (2.25) was observed in this case. Complete conversion to SBL-Dha217 was not achieved even after raising the reaction temperature to 42 °C or increasing the pH to ~11 (See Experimental Section 2.5.6, Table 2-2).



Scheme 2-12. Conditions used to convert Cys217 to Dha using DBHDA

Other methods to convert cysteine to dehydroalanine at position 217 were also explored⁴⁷ (Scheme 2-13). First, base-mediated elimination of aryl-cysteine adducts was considered (Scheme 2-13a). The highly selective reaction of Cys217 with 25 equiv. of 2-chloromethylpyridinium iodide (Mukaiyama's reagent, **MR**) gave rapid conversion to the SBL-L217C-MR adduct (**2.26**). However, addition of base to facilitate the subsequent elimination reaction did not lead to the desired Dha product. As an alternative strategy, the tandem reduction-elimination of mixed disulfides using phosphines was employed (Scheme 2-13b). Previously, this route lead to Dha formation using SBL-S156C-Ellman's disulfide as model substrate and hexamethylphosphorous triamide (HMPT) as the electron rich phosphine.⁴⁷ It was shown that elimination proceeds *via* the attack of phosphine on cysteine sulfur to give SBL-Dha156.⁴⁷ The reduction of SBL-L217C-Ellman's disulfide **2.15** with 150 equiv. of HMPT gave intermediate **2.27**. However, the subsequent elimination step by the addition of base (NaOH) resulted in the loss of protein signal perhaps due to degradation of protein. Such highly basic conditions are detrimental for

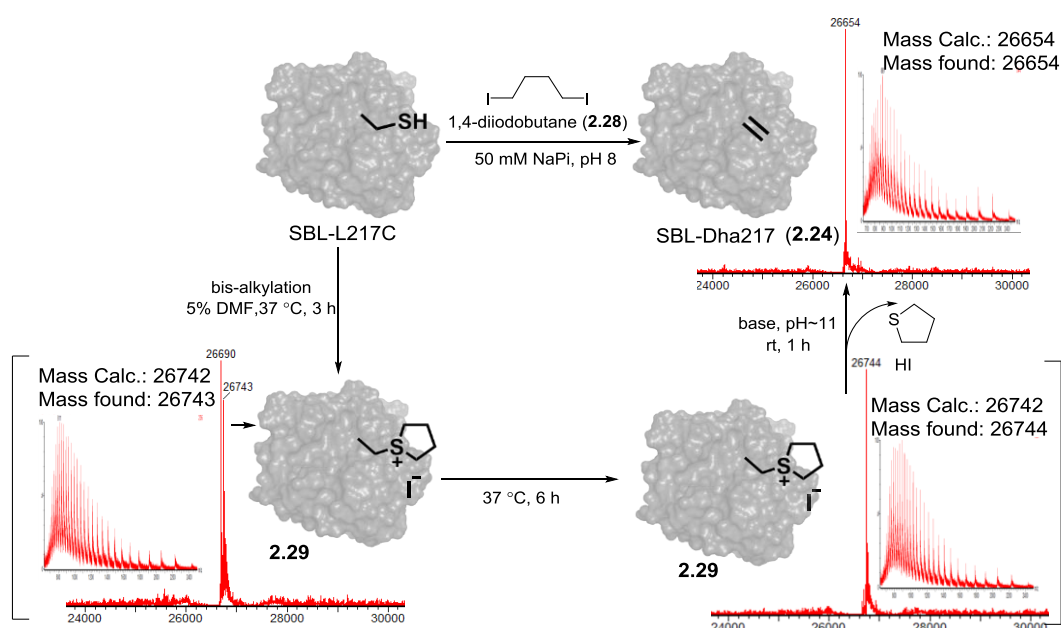
protein structure and hence these routes were found unfavourable to obtain the desired metalloenzymes.



Scheme 2-13. Conditions used to convert Cys217 to Dha (a) using Mukaiyama's reagent (b) using reduction-elimination route using HMPT

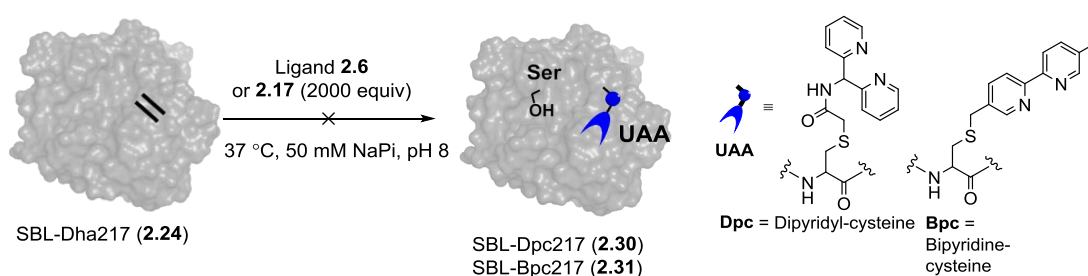
Finally, bis-alkylation-elimination of SBL-L217C using a different diiodo-based alkylating reagent **2.28** was employed (Scheme 2-14). At first, the alkylation of Cys217 with 1,4-diiodobutane reagent **2.28** was found very slow (13% even after 4 h) perhaps due to immiscibility of the diiodo reagent **2.28**. On including 5% DMF as co-solvent in the reaction mixture, the bis-alkylation of the Cys217 with **2.28** at 37 °C proceeded *via* the observance of a mass corresponding to the sulfonium intermediate **2.29**. Complete conversion to the sulfonium **2.29** was observed after 6 h incubation at 37 °C. Increasing the temperature to 42 °C afforded only 13% Dha product **2.24**. However, rapid elimination to SBL-Dha217 (**2.24**) was observed at pH ~10 at room temperature by LC-MS analysis. The buffer was quickly exchanged back to pH 6 for subsequent addition reaction.

Importantly, no reaction was observed upon treatment of SBL-Dha217 with Ellman's reagent (1000 equiv.), confirming the full consumption of Cys217.



Scheme 2-14. Synthesis of SBL-Dha217 by using 1,4-diiodobutane reagent **2.28**

Conjugate addition of **L3** and **L4** to SBL-Dha217 was attempted (Scheme 2-14). However, no shift in mass was observed after 1 h incubation at 37 °C (under same conditions as used for SBL-Dha156). Our attempts for conjugate addition of thiol ligands **L3** or **L4** to SBL-Dha217 (with increased equiv. of the ligands to up to 2000 and prolonged reaction times) were unsuccessful in yielding the desired enzyme species SBL-Dpc217 (**2.30**) or SBL-Bpc217 (**2.31**). The analysis was difficult due to noisy spectra and the thiol ligands were also unstable in the reaction mixture due to formation of disulfides or degradation.



Scheme 2-15. Conditions used for conjugate addition of ligands **L3** and **L4** to SBL-Dha217

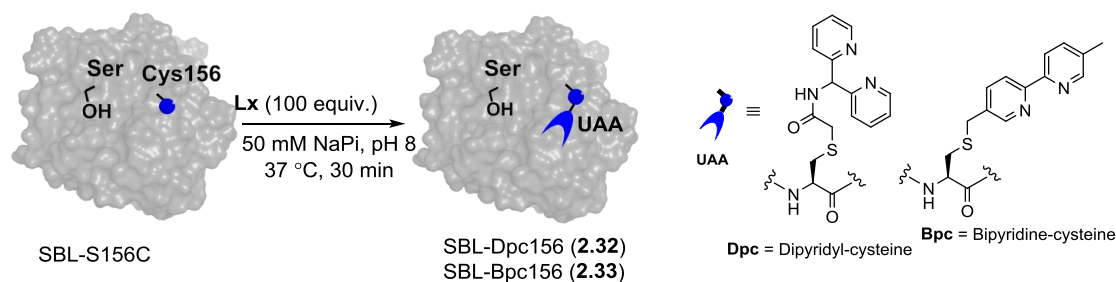
While the Dha route was successfully employed for the chemical installation of metal binding amino acids in model protein SBL-S156C where the reactive Cys156 is present on a flexible loop in the S1 pocket, the same strategy was not successful for SBL-L217C. Moreover, the bis-alkylation-elimination route required highly basic denaturing conditions to facilitate the elimination step, which were found undesirable for our purposes to generate active artificial metalloenzymes.

2.2.2.4 Artificial metalloenzymes using direct alkylation route

The S_N2 substitution reaction of haloacetamides for direct alkylation of cysteine is a well-documented reaction.⁵⁰ For example, iodoacetamides are extensively used as alkylating reagent in mass spectrometry, but such iodo- electrophiles can also modify histidine or lysine residues.⁵¹ However, selective alkylation at cysteine residues may be achieved by working at higher pH (>7.5), limiting the time of contact or using less reactive electrophiles.^{5,51,52} Previously, successful direct alkylation of SBL-S156C using allyl chloride was reported without any non-specific reactions at any surface accessible lysines and histidines.⁵³ Hence, we decided to explore direct alkylation of single cysteine SBL variants using other halogenated ligands (such as chloroacetamide **L1** and bromoalkylating reagent **L2**).

The direct alkylation of SBL-S156C with halogenated ligands **Lx** gave SBL-Dpc156 and SBL-Bpc156 within 30 min incubation at 37 °C (Scheme 2-16).[†] The Cys156 in this SBL-variant is present on a surface exposed loop, thus enabling higher access to the ligands for chemical modification and faster reaction. It was noted that the LC-MS analysis of SBL-Dpc156 revealed the presence of a significant amount of a mass corresponding to the double-alkylated protein species[†] (See Experimental Section 2.5.7). In contrast to this,

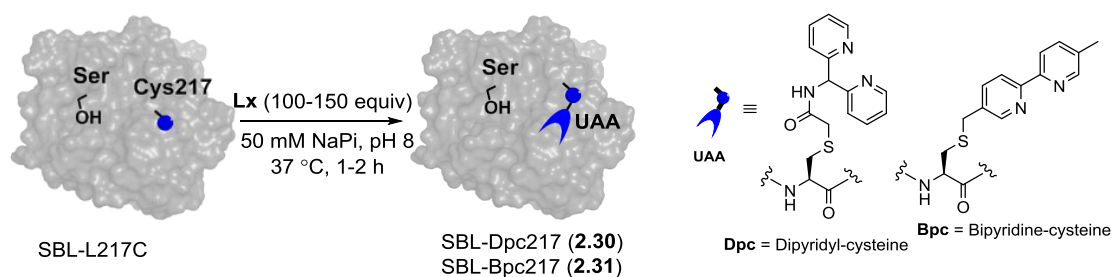
only mono-alkylated SBL-Dpc156 was observed when the Dha route was employed (Section 2.2.2.3).



Scheme 2-16. Chemical modification of the SBL-S156C variant with ligands **L1** and **L2**[†]

As an alternative to the Dha strategy, the direct alkylation of Cys tag proved most successful at the key site 217 to obtain the metal-free ‘apo form’ of the desired metalloenzymes (Scheme 2-17). The alkylation of the reactive single cysteine tag (Cys217) was carried out with a moderate excess of the halogenated ligand **Lx** (100-150 equiv.) to ensure rapid and complete modification within 2 h. While **L1** is a water-soluble acetamide, the reaction of **L2** proceeded heterogeneously, so 2.5-5% of DMF was used as a co-solvent to ensure partial solubility of the ligands. The covalent bio-conjugation of SBL-L217C with the chelating ligands **Lx** was monitored by LC-MS, verifying complete conversion on alkylation with **L1** and **L2** giving homogeneous modified enzymes SBL-Dpc217 (**2.30**) and SBL-Bpc217 (**2.31**), respectively. Successful installation of UAAs and the identities of these synthetic proteins SBL-Dpc217 and SBL-Bpc217 were confirmed through full characterization by LC-MS analysis, peptide mapping and MS/MS[†] analysis (See Experimental Section 2.5.8 for details).

[†]Peptide mapping and MS/MS analysis was done for SBL-Dpc217 using MASCOT⁵⁸ database search. Sample was prepared by Dr. Alba Díaz-Rodríguez and analysis was done by the author.



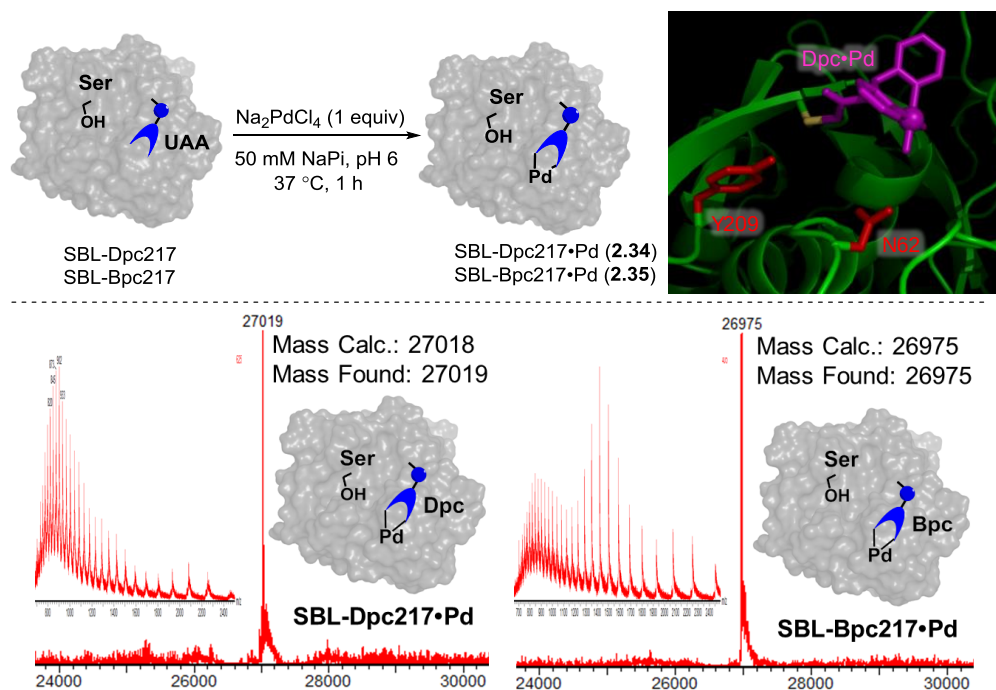
Scheme 2-17. Chemical modification of the SBL-L217C variant with ligands **L1** and **L2**

Negative control reactions with SBL-WT (that lacks the reactive sulfhydryl group) confirmed the chemoselectivity of direct alkylation with both **L1** and **L2** ligands. Moreover, Ellman's test of the modified enzymes SBL-Dpc217 and SBL-Bpc217 confirmed complete alkylation of the installed cysteine residue at position 217, as no change in the mass was observed on treatment with DTNB reagent under standard conditions (See Experimental Section 2.5.8 for details).

Dpc and Bpc can chelate palladium in a bidentate fashion. The metal-free 'apo form' was converted to metal-containing 'holo form' of the Pd-enzymes by coordination of 1 equiv. of Pd(II) at 37 °C. The chelation of palladium to give SBL-Dpc217•Pd (**2.34**) and SBL-Bpc217•Pd (**2.35**) was confirmed by LC-MS; full conversions were obtained in both cases. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd (mass calculated 27018, mass found 27019) and SBL-Bpc217•Pd (mass calculated 26975, mass found 26975), after the loss of two chloride ligands using Na_2PdCl_4 as the palladium source (Scheme 2-18). Preliminary modelling studies[†] were carried out to assess plausible binding interactions of the installed metal with any other neighbouring active site residues using the SBL-WT structure (PDB ID:1GCI).²⁵ Modelling studies to explore the range of

[†] Modelling studies were done by Dr. Mark Wormald

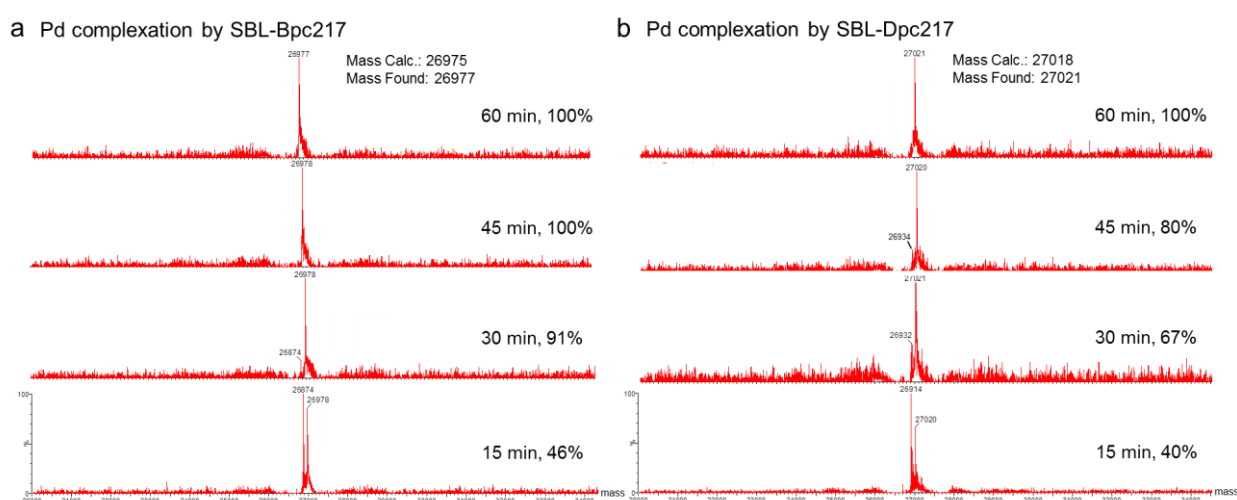
possible conformations of the ligands suggest possible coordination of palladium with residues Asn62 or Tyr209, when **Dpc•Pd** complex was modelled at position 217 (See Experimental Section 2.5.10.4).



Scheme 2-18. Metal coordination to form ‘holo-enzymes’ SBL-Dpc217•Pd and SBL-Bpc217•Pd; model containing Dpc•Pd complex installed at position 217 was generated using SBL-WT structure (PDB ID: 1GCI)²⁵

Interestingly, different palladium Pd(II) sources [such as $\text{Pd}(\text{OAc})_2$ or $\text{trans-Pd}(\text{NH}_3)_2\text{Cl}_2$] gave the same results under the analysis conditions employed, regardless of the difference in labile ligand groups on Pd, thereby highlighting the good affinity of Bpc and Dpc for the metal (See Experimental Section 2.5.9). The mass shift corresponding to complexation of only one Pd atom (+106) was observed in all cases, thus suggesting that the labile Pd ligands (acetate, NH_3 or chloride) were either replaced by coordinating functionalities inside the enzyme, or dissociated during the analysis conditions.

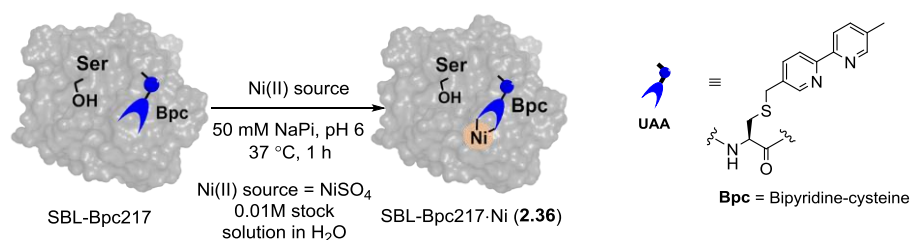
Time course experiments were carried out for the palladium complexation by SBL-Dpc217 and SBL-Bpc217 using Na_2PdCl_4 as the Pd(II) source. While the uptake of palladium was complete within 45 min for SBL-Bpc217 to give the metal containing holoform SBL-Bpc217•Pd (Mass calculated: 26975, Mass found: 26977); the palladium loading on SBL-Dpc217 took slightly longer (~1 h) to give the corresponding holoform SBL-Dpc217•Pd (Mass calculated: 27018, Mass found: 27021). This comparison is reflective of the binding affinities of the Dpc and Bpc ligands installed on the protein.



Scheme 2-19. Time course assay for the palladium complexation by SBL-Dpc217 and SBL-Bpc217

Control reactions with wild-type SBL-WT or SBL-L217C (precursors which both lack designed metal binding UAAs) showed complete absence of any non-specific palladium coordination (to any surface residue or to the installed cysteine residue in the active site) under the same conditions, consistent with their intended role and the proposed design (See Experimental Section 2.5.10.3). Palladium loading at increased equivalents (1-10 equiv.) showed negative results, thus verifying that the installed amino acids Dpc and Bpc serve as selective metal binding sites.

Chelation of other metal ions with Dpc and Bpc was also investigated using LC-MS to detect metal:protein adducts. In particular, nickel was considered as an attractive metal target given its utility in cross-coupling reactions.^{54,55} Complexation of nickel was successful with SBL-Bpc217 to yield the corresponding metalloenzyme SBL-Bpc217•Ni (**2.36**), however no metal adduct could be detected for the same modification for SBL-Dpc217. Metal complexation of SBL-Bpc217 with a Cu(II) source such as Cu(OAc)₂ was unsuccessful in yielding any metalloenzyme species by LC-MS analysis (See Experimental Section 2.5.11).



Scheme 2-20. Complexation of nickel by SBL-Bpc217

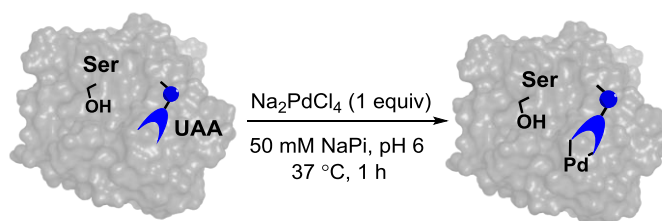
2.2.3 Characterization of SBL-based metalloenzymes

2.2.3.1 Trace metal analysis by ICP-MS

Typically, the levels of palladium in our samples are between 1-10 $\mu\text{g g}^{-1}$. Hence, trace metal analysis was undertaken by employing the ICP-MS technique (which has a detection limit of $\sim 0.09 \text{ ng g}^{-1}$). To confirm the successful incorporation of the metal catalyst in engineered SBL variants SBL-Dpc217 and SBL-Bpc217, metal complexation was carried out using one equiv. of Na₂PdCl₄ as Pd(II) source and their identities were confirmed by LC-MS analysis. Quantitative ICP-MS and protein content analyses (based on bicinchoninic acid assay, BCA) unambiguously confirmed the palladium loading on SBL-Bpc217•Pd and SBL-Dpc217•Pd[†], giving values corresponding to the presence of just one palladium atom per molecule of enzyme (Table 2-1).

As a control, when palladium coordination with SBL-WT was carried out (followed by dialysis to remove the unbound metal), ESI-MS spectra revealed no metal adduct. The ICP-MS analysis revealed metal:protein content of 0.3, indicative of some soluble Pd(II) present in the solution or due to non-specific palladium coordination to SBL-WT (not seen in LC-MS analysis). The analysis of negative control samples, SBL-Dpc217, SBL-WT or blank (without the addition any Pd(II) source to the sample) showed no detection of Pd105 signal in the trace metal analysis, confirming all observed palladium signal was originating from the synthesized metalloenzyme species.

Table 2-1. Results from the trace metal analysis for m/z Pd105



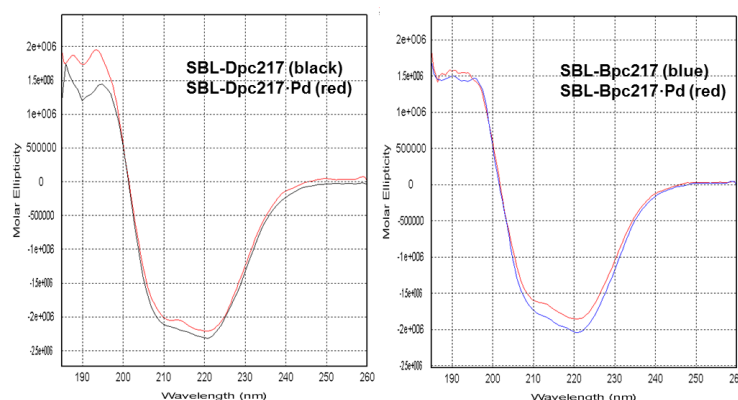
Entry	Sample ID	Sample volume (μL)	[Protein] (mg/mL)	Pd105 ^a (ng/g)	Protein (nmol/g)	Pd105 (nmol/g)
1	SBL-Bpc217•Pd	100	0.2288	884.67156	8.4819	8.4254
2	SBL-Dpc217•Pd ^{b†}	480	-	-	164.58	144.8
3	SBL-WT + Na_2PdCl_4 ^c	100	0.6095	739.40513	22.830	7.04
4	SBL-Dpc217	100	0.7808	27.25618	29.012	0.2595 ^d
5	SBL-WT	100	0.9496	29.59128	35.568	0.2818 ^d
6	Blank (MilliQ)	100	-	21.59986	-	0.2057 ^d

^aMinimum Detection Limit (MDL) for Pd105 was 93.83 ng/g, ^bSample was prepared by Dr. Alba Díaz-Rodríguez and analysis performed by Butterworth Laboratories, ^c Na_2PdCl_4 used as palladium precursor to check for any non-specific binding of palladium with SBL-WT; ^dPd105 levels below minimum detection limit

2.2.3.2 Metal compatibility with SBL

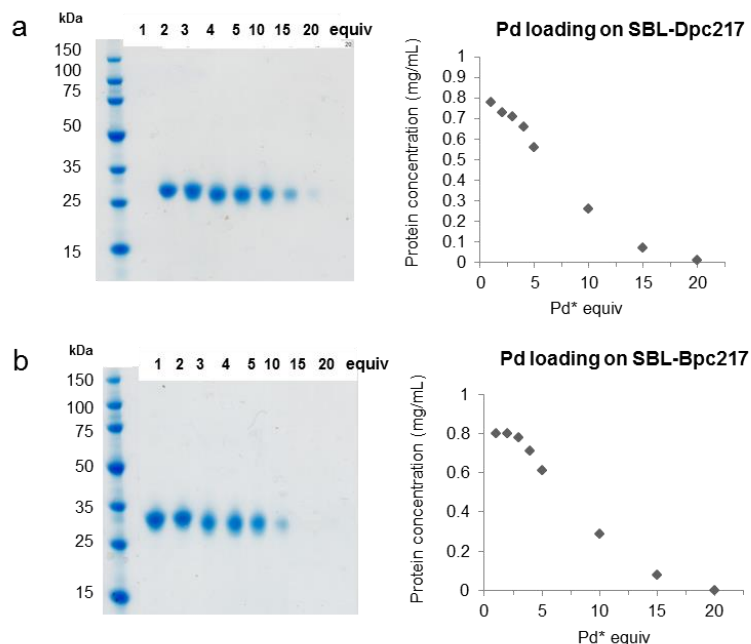
In order to assess any secondary structure changes that may be introduced by incorporation of a non-native metal such as palladium, far-UV circular dichroism (CD) spectra of engineered SBL-Dpc217 and SBL-Bpc217 were recorded before and after Pd complexation (See Experimental Section 2.5.13). LC-MS analysis confirmed the identity

of these chemically modified protein species. More importantly, no overall secondary changes could be detected, as can be judged from the CD spectra of the holo-proteins (Scheme 2-21).



Scheme 2-21. Overlays of far UV CD spectra of (a) SBL-Dpc217 and SBL-Dpc217•Pd (b) SBL-Bpc217 and SBL-Bpc217•Pd

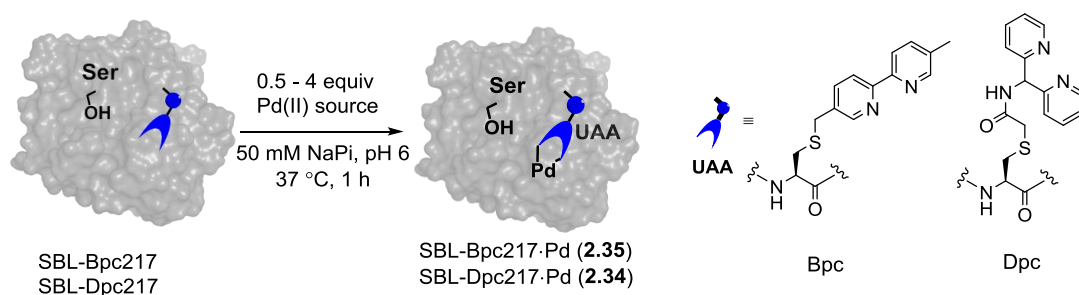
While the compatibility of one equiv. of non-native palladium metal with protein was ascertained by CD analysis and was found to have little effect on protein stability, the limit of metal tolerance and protein stability was speculated. To verify this, we performed a metal tolerance assay with SBL-Dpc217 and SBL-Bpc217 based on monitoring the loss of protein by non-reducing PAGE gel analysis and BCA analysis. This allows measurement of the concentration of soluble protein after incubation with the non-native metal ion. On incubating protein with increasing equiv. of palladium for 1 h, the gel analysis showed clear loss of protein, perhaps due to protein aggregation induced by divalent metal ions acting as bridges (Scheme 2-22). Similar behaviour was observed for both engineered proteins with complete loss of soluble protein content when >15 equiv of Pd(II) source was used, mainly due to protein precipitation (See Experimental Section 2.5.14). With these tests, a limit for future metal titration studies was set to 4 molar excess of metal with respect to the protein to avoid protein aggregation and precipitation.



Scheme 2-22. Non-reducing PAGE analysis of metal loading studies with SBL-Dpc217 and SBL-Bpc217 after 1 h incubation with Pd(II) source

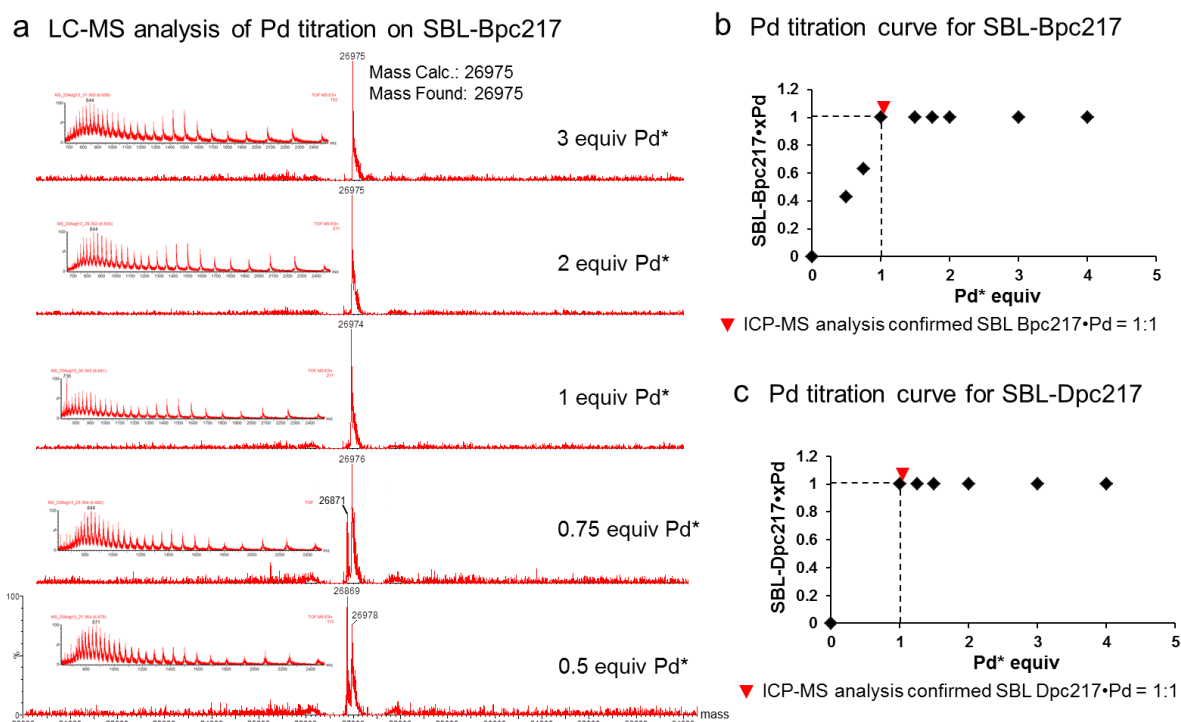
2.2.3.3 Palladium titration

The metal binding selectivity and stoichiometry of the incorporated UAAs was tested by metal titration experiments (See Experimental Section 2.5.14). Prior to titration of palladium, the protein samples were dialysed extensively to ensure no excess unbound ligand molecules were present in the solution. To the conditioned protein samples, SBL-Bpc217 and SBL-Dpc217, 0.5-4 equiv. of Pd(II) source at 37 °C was added and the complexation was monitored by LC-MS after 1 h incubation (Scheme 2-23).



Scheme 2-23. Palladium titration studies with apo-proteins SBL-Bpc217 and SBL-Dpc217

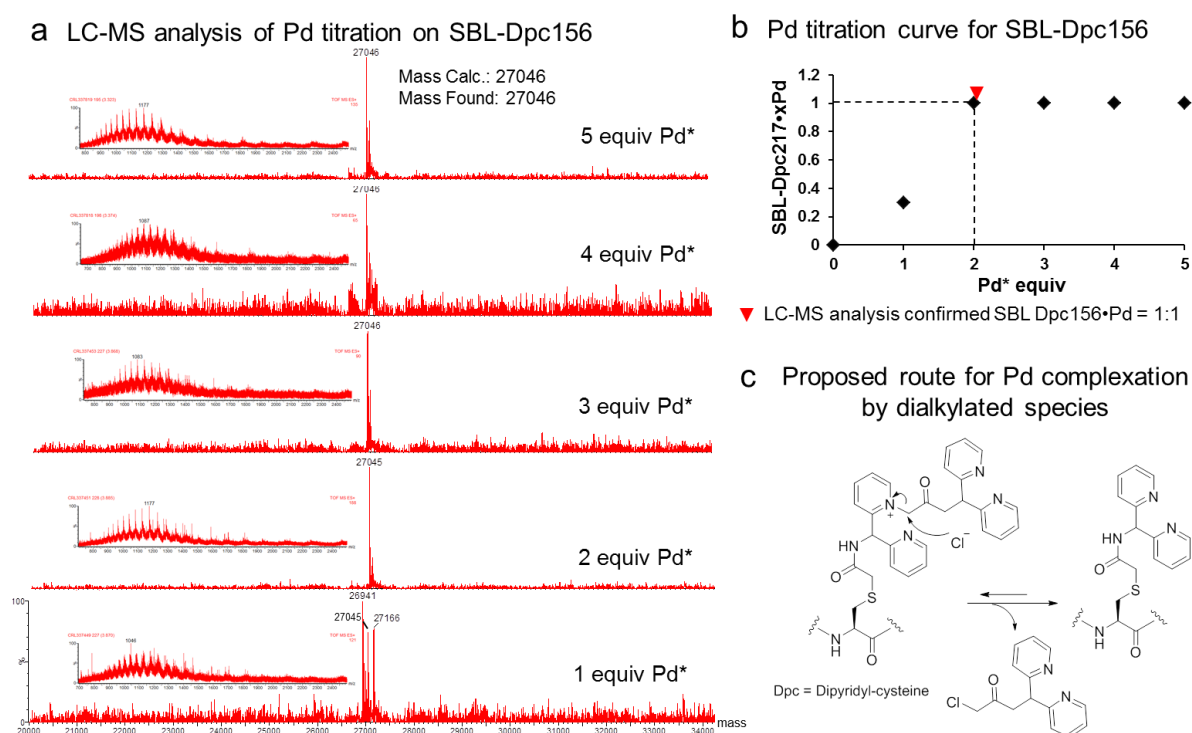
The exquisite selectivity of both Bpc and Dpc was further confirmed by palladium titration of SBL-Bpc217 and SBL-Dpc217, when only one equiv. of Pd(II) was loaded in both cases and clean titration of palladium was confirmed. Also, in accordance with our metal tolerance observations, protein precipitation was observed beyond 4 equiv. of Pd(II) source (Scheme 2-24). Taken together with the results from the ICP-MS analysis, clean titration of metal onto the installed Bpc and Dpc amino acid residues confirms that a metal binding site has been successfully installed that could serve as a reaction centre for the target chemical reactions.



Scheme 2-24. Palladium titration studies on SBL-Bpc217 and SBL-Dpc217 (a) LC-MS analysis of titration of 0.5-4 equiv. of palladium on SBL-Bpc217 (b) Palladium titration curve for SBL-Bpc217 (b) Palladium titration curve for SBL-Dpc217

Additionally using this established procedure, palladium titration was also attempted with SBL-Dpc156. Chemical installation of Dpc on SBL-S156C *via* direct alkylation route yields a mixture of mono- and di-alkylated SBL-Dpc156 (Section 2.2.2.4). It was previously shown that this double-alkylation adduct appears perhaps due to the Dpc itself.

On palladium complexation, the double alkylated SBL-Dpc156 was found to shift to single protein species with the mass corresponding to SBL-Dpc156•Pd species.[†] Presumably, this can be explained by the displacement of the second alkylating group by a nucleophile present in the solution (such as Cl⁻) to give the mono-alkylated species which can then chelate Pd(II) ion in a bidentate fashion (Scheme 2-25). In agreement with this observation, titration curve of Pd(II) ion on SBL-Dpc156[†] indicates the requirement of slight excess of Pd(II) source would be required to obtain the corresponding metalloenzyme species SBL-Dpc156•Pd (2.37). Importantly, no non-specific palladium coordination was observed when higher equiv. of Pd(II) source was titrated onto the protein, as can be judged by the LC-MS assay below.



Scheme 2-25. Palladium titration studies on SBL-Dpc156 (a) LC-MS analysis of titration of 1-5 equiv. of palladium on SBL-Dpc156 (b) Palladium titration curve for SBL-Dpc156 (b) Proposed route for palladium complexation by di-alkylated SBL-Dpc156 species[†]

[†]SBL-Dpc156 was borrowed from Dr. Alba Díaz-Rodríguez and the experiment was done with Dr. Maria Espling

2.2.3.4 Palladium leaching

Both SBL-Bpc217 and SBL-Dpc217 containing bi-dentate metal chelating moieties, have been shown to have excellent palladium complexation capacity.³⁹ In order to determine if there was any metal leaching possible from these metalloenzyme species, SBL-Bpc217•Pd enzyme was monitored for signs of any metal leaching based on a LC-MS assay. Gratifyingly, the enzyme was found to retain the metal at 4 °C for over 60 days, providing a reliable system for future screening.

2.3 Conclusions

Creation of artificial metalloenzymes opens exciting avenues of combining enzyme activity with non-biological metals such as palladium, rhodium, ruthenium etc. that are well-established catalysts in synthetic chemistry. The first step in the logical design of such an artificial metalloenzyme is the choice of enzyme scaffold and the metal catalyst. With an aim to exploit the substrate binding characteristics and enhanced nucleophilicity of an enzyme active site, Subtilisin (SBL) was chosen as the model enzyme. In order to keep the catalytic metal centre close to the SBL active site, two metal binding unnatural amino acids Bpc (Bipyridyl-cysteine) and Dpc (Dipyridyl-cysteine) lacking in any spacer/linker group were designed. A site-selective methodology was required to ensure chemical installation of Bpc and Dpc. Such tailoring of the active site was achieved by employing our tag-and-modify strategy based on well-established covalent chemical modification methods.

Two different methods could be successfully employed to form the apo-SBL structures (namely, SBL-Dpc217, SBL-Bpc217, SBL-Dpc156 and SBL-Bpc156) capable of coordinating metal ions at the active site. The first method involved inter-conversion of an

installed cysteine residue to an electrophilic dehydroalanine (Dha) residue that enables conjugate addition of thio-functionalised ligands in a selective manner. Prior work by Davis and co-workers has successfully demonstrated the incorporation of several other thiol compounds using SBL-S156C scaffold as the model system,^{22,47} showing Dha route can provide chemical access to post-translational modifications. We used these established procedures to introduce the Dha strategy for site-selective installation of metal centres in order to create artificial metalloenzymes. The use of bis-alkylation-elimination reaction of a reactive di-bromide reagent (DBHDA) proved most successful for positioning of Dpc and Bpc chelating amino acid residues in the active site at position 156. This is the first example of use of Dha strategy for generation of artificial metalloenzymes. Further, we also investigated the synthesis of Dha at position 217. We found that bis-alkylation-elimination to Dha residue in SBL-L217C using 1,4-diiodobutane as alkylating reagent proceeded *via* the observance of a mass corresponding to a stable sulfonium intermediate. This difference in reactivity of cysteine residues at different sites on the same scaffold is notable and indicates that other steric and electronic factors are at play. In this context, the formation of such a stable cyclic sulfonium intermediate has been reported recently, which when reacted with β -mercaptoethanol led to a bithioether (presumably through ring opening of the sulfonium).^{56,57} This cysteine-sulfonium intermediate was further used for addition of a range of nucleophiles.^{56,57}

Alternatively, site-selective modification of SBL *via* direct alkylation yielded four distinct synthetic proteins namely, SBL-Bpc217, SBL-Dpc217, SBL-Bpc156 and SBL-Dpc156. Importantly, the site of installation of UAAs could be confirmed by the necessary control reactions and MS/MS analysis. The main advantage of establishing precise chemical installation procedures for these designed amino acids is the general nature of these

covalent methodologies that can be applied to virtually any protein scaffold when compared with other available supramolecular and dative anchoring strategies (see Chapter 1 Section 1.2.2). Chemical installation of unnatural amino acids and the step-wise assimilation of various components would ensure variability at each level (i.e. enzyme, ligand, metal) for future design and screening, which is the essence of our ‘tag-and-modify’ approach.

Crucially, selective metal binding sites were created with confirmed metal:protein stoichiometry of 1:1 giving holoenzymes SBL-Dpc217•Pd, SBL-Bpc217•Pd, SBL-Dpc156•Pd and SBL-Bpc156•Pd. On the basis of LC-MS, ICP-MS techniques and negative control reactions it was concluded that chelating amino acids Bpc and Dpc installed in the active site enable metal coordination. Furthermore metal titrations using 1-5 equiv. of Pd²⁺ precursor to detect any non-specific metal adducts by LC-MS confirmed that other surface coordinating residues were not involved in metal binding. Even though in our design stoichiometric amount of M(II) precursors leads to full conversion to metal-containing holoenzymes, we evaluated the stability and compatibility of the protein in the presence of foreign metal ions (like palladium) by using circular dichroism and evaluating soluble protein content in the presence of increased metal concentration. Conversely, once palladium is bound to SBL any metal inactivation by aggregation was prevented, as evidenced by metal leaching experiments which showed that the metalloenzyme species remained intact over an extended period of time at 4 °C.

In this way, not only does protein provides water solubility to the organometallic complex (such as water insoluble ligand **L2**), it was envisioned that the enzyme would serve to play an active role in bringing the substrates close to the metal complex in an activating

microenvironment. Having successfully incorporated these metal centres in the enzyme active site, we were motivated to test the activity of these enzymes for Pd-catalysed reactions, in particular oxidative Heck cross-coupling reaction as discussed in Chapter 3. Furthermore, structural characterization to unambiguously confirm the location of metal in these artificial metalloenzymes was also undertaken (Chapter 4).

2.4 References

- (1) Crankshaw, M. W.; Grant, G. A. In *Current Protocols in Protein Science*; John Wiley & Sons, Inc.: 2001.
- (2) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chemistry – An Asian Journal* **2009**, 4, 630.
- (3) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *CHIMIA International Journal for Chemistry* **2002**, 56, 721.
- (4) Haquette, P.; Dumat, B.; Talbi, B.; Arbabi, S.; Renaud, J.-L.; Jaouen, G.; Salmain, M. *J. Organomet. Chem.* **2009**, 694, 937.
- (5) Hermanson, G. T. *Bioconjugate Techniques*; Academic Press, 1996.
- (6) Davies, R. R.; Kuang, H.; Qi, D.; Mazhary, A.; Mayaan, E.; Distefano, M. D. *Bioorg. Med. Chem. Lett.* **1999**, 9, 79.
- (7) Haquette, P.; Talbi, B.; Canaguier, S.; Dagorne, S.; Fosse, C.; Martel, A.; Jaouen, G.; Salmain, M. *Tetrahedron Lett.* **2008**, 49, 4670.
- (8) Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. *J. Am. Chem. Soc.* **2008**, 130, 11783.
- (9) Boersma, A. J.; Coquière, D.; Geerdink, D.; Rosati, F.; Feringa, B. L.; Roelfes, G. *Nat Chem* **2010**, 2, 991.
- (10) Boersma, A. J.; de Bruin, B.; Feringa, B. L.; Roelfes, G. *Chem. Commun.* **2012**, 48, 2394.
- (11) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. *Angew. Chem. Int. Ed.* **2012**, 51, 7472.
- (12) Bos, J.; Garcia-Herraiz, A.; Roelfes, G. *Chemical Science* **2013**, 4, 3578.
- (13) Lee, H. S.; Spraggon, G.; Schultz, P. G.; Wang, F. *J. Am. Chem. Soc.* **2009**, 131, 2481.
- (14) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, 51, 1835.
- (15) Lee, H. S.; Schultz, P. G. *J. Am. Chem. Soc.* **2008**, 130, 13194.
- (16) Mills, J. H.; Khare, S. D.; Bolduc, J. M.; Forouhar, F.; Mulligan, V. K.; Lew, S.; Seetharaman, J.; Tong, L.; Stoddard, B. L.; Baker, D. *J. Am. Chem. Soc.* **2013**, 135, 13393.

- (17) Davis Benjamin, G. In *Pure Appl. Chem.* 2009; Vol. 81, p 285.
- (18) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. *Acc. Chem. Res.* **2011**, *44*, 730.
- (19) Neet, K. E.; Koshland, D. E. *Proceedings of the National Academy of Sciences of the United States of America* **1966**, *56*, 1606.
- (20) Polgar, L.; Bender, M. L. *J. Am. Chem. Soc.* **1966**, *88*, 3153.
- (21) Chalker, J. M.; Davis, B. G. *Curr. Opin. Chem. Biol.* **2010**, *14*, 781.
- (22) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052.
- (23) Wright, C. S.; Alden, R. A.; Kraut, J. *Nature* **1969**, *221*, 235.
- (24) Alden, R. A.; Birktoft, J. J.; Kraut, J.; Robertus, J. D.; Wright, C. S. *Biochem. Biophys. Res. Commun.* **1971**, *45*, 337.
- (25) Kuhn, P.; Knapp, M.; Soltis, S. M.; Ganshaw, G.; Thoene, M.; Bott, R. *Biochemistry* **1998**, *37*, 13446.
- (26) Gallagher, T.; Oliver, J.; Bott, R.; Betzel, C.; Gilliland, G. L. *Acta Crystallographica Section D* **1996**, *52*, 1125.
- (27) Bonneau, P. R.; Graycar, T. P.; Estell, D. A.; Jones, J. B. *J. Am. Chem. Soc.* **1991**, *113*, 1026.
- (28) Berglund, P.; DeSantis, G.; Stabile, M. R.; Shang, X.; Gold, M.; Bott, R. R.; Graycar, T. P.; Lau, T. H.; Mitchinson, C.; Jones, J. B. *J. Am. Chem. Soc.* **1997**, *119*, 5265.
- (29) Plettner, E.; Khumtaveeporn, K.; Xiao, S.; Jones, J. B. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2291.
- (30) DeSantis, G.; Jones, J. B. *Bioorg. Med. Chem.* **1999**, *7*, 1381.
- (31) DeSantis, G.; Shang, X.; Jones, J. B. *Biochemistry* **1999**, *38*, 13391.
- (32) Davis, B. G.; Khumtaveeporn, K.; Bott, R. R.; Jones, J. B. *Bioorg. Med. Chem.* **1999**, *7*, 2303.
- (33) Davis, B. G.; Shang, X.; DeSantis, G.; Bott, R. R.; Jones, J. B. *Bioorg. Med. Chem.* **1999**, *7*, 2293.
- (34) Berger, A.; Schechter, I. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **1970**, *257*, 249.
- (35) Jones, J. B.; DeSantis, G. *Acc. Chem. Res.* **1998**, *32*, 99.
- (36) Stabile, M. R.; George Lai, W.; DeSantis, G.; Gold, M.; Bryan Jones, J.; Mitchinson, C.; Bott, R. R.; Graycar, T. P.; Liu, C.-C. *Bioorganic & Medicinal Chemistry Letters* **1996**, *6*, 2501.
- (37) Robertus, J. D.; Kraut, J.; Alden, R. A.; Birktoft, J. J. *Biochemistry* **1972**, *11*, 4293.
- (38) Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337.
- (39) Nájera, C.; Gil-Moltó, J.; Karlström, S.; Falvello, L. R. *Org. Lett.* **2003**, *5*, 1451.
- (40) Nájera, C.; Gil-Moltó, J.; Karlström, S. *Adv. Synth. Catal.* **2004**, *346*, 1798.

- (41) Heck, R.; Dumarcay, F.; Marsura, A. *Chemistry – A European Journal* **2002**, 8, 2438.
- (42) Ipe, B. I.; Yoosaf, K.; Thomas, K. G. *J. Am. Chem. Soc.* **2006**, 128, 1907.
- (43) Hu, J.; Fox, M. A. *The Journal of Organic Chemistry* **1999**, 64, 4959.
- (44) Kumar, J.; Thomas, K. G. *The Journal of Physical Chemistry Letters* **2011**, 2, 610.
- (45) Bertoncello, P.; Kefalas, E. T.; Pikramenou, Z.; Unwin, P. R.; Forster, R. J. *The Journal of Physical Chemistry B* **2006**, 110, 10063.
- (46) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, 82, 70.
- (47) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chemical Science* **2011**, 2, 1666.
- (48) Zhu, Y.; van der Donk, W. A. *Org. Lett.* **2001**, 3, 1189.
- (49) Galonić, D. P.; van der Donk, W. A.; Gin, D. Y. *Chemistry – A European Journal* **2003**, 9, 5997.
- (50) Anson, M. L. *The Journal of General Physiology* **1940**, 23, 321.
- (51) Boja, E. S.; Fales, H. M. *Anal. Chem.* **2001**, 73, 3576.
- (52) Nielsen, M. L.; Vermeulen, M.; Bonaldi, T.; Cox, J.; Moroder, L.; Mann, M. *Nat Meth* **2008**, 5, 459.
- (53) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714.
- (54) Zhou, J.; Fu, G. C. *J. Am. Chem. Soc.* **2004**, 126, 1340.
- (55) Tasker, S. Z.; Standley, E. A.; Jamison, T. F. *Nature* **2014**, 509, 299.
- (56) Nathani, R.; Moody, P.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. *ChemBioChem* **2012**, 13, 1283.
- (57) Nathani, R. I.; Moody, P.; Chudasama, V.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. *Chemical Science* **2013**, 4, 3455.
- (58) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *ELECTROPHORESIS* **1999**, 20, 3551.
- (59) Roepstorff, P.; Fohlman, J. *Biological Mass Spectrometry* **1984**, 11, 601.
- (60) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, 47, 1698.

2.5 Experimental procedures

2.5.1 General Considerations

Chemical Syntheses and characterization.

Proton nuclear magnetic resonance (^1H NMR) spectra were recorded on a Bruker DPX200 (200 MHz), a Bruker AV400 (400 MHz), or a Bruker AVII500 (500 MHz) spectrometer,

as indicated. Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Bruker DPX200 (50 MHz), a Bruker AV400 (100 MHz) or on a Bruker AVII500 (125 MHz) spectrometer, as indicated. Fluorine nuclear magnetic resonance (^{19}F NMR) spectra were recorded on a Bruker AVII500 (470.4 MHz) or on a Bruker DQX400 (376.5 MHz) spectrometer with CFCl_3 as external standard, as indicated. NMR Spectra were assigned using COSY, HMQC, and DEPT135. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (^1H NMR: CDCl_3 = 7.26; CD_3OD = 3.31; DMSO- d_6 = 2.50; CD_3CN = 1.94; D_2O = 4.80 and ^{13}C NMR: CDCl_3 = 77.0; CD_3OD = 49.0; DMSO- d_6 = 39.5; CD_3CN = 1.94). Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet.

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer. Absorption maxima (ν_{max}) are reported in wavenumbers (cm^{-1}).

Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer Waters using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ESI mass spectrometer by the Mass Spectrometry Service in the Department of Chemistry, University of Oxford. Nominal and exact m/z values are reported in Daltons (Da).

Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with 60F254 silica gel. Visualization of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or ammonium molybdate (5% in 2 M H_2SO_4), and/or potassium permanganate (5% KMnO_4 in 1 M NaOH with 5% potassium carbonate). Flash column

chromatography was carried out using BDH PROLAB® 40-63 mm silica gel (VWR). Mobile phases are reported in % volume of more polar solvent in less polar solvent for binary systems (e.g. 20% EtOAc in petrol = 1:4 ethyl acetate:petrol) and in relative composition for ternary systems (e.g. 1:2:4 H₂O: iPrOH: EtOAc). Anhydrous solvents were purchased from Fluka or Acros with the exception of dichloromethane and THF, which were dried in an alumina column under nitrogen. All other solvents were used as supplied (Analytical or HPLC grade), without prior purification. Deionized water was used for chemical reactions and MilliQ® purified water for protein manipulations. All other reagents were purchased from Fisher Scientific or Sigma Aldrich, unless otherwise indicated.

‘Petrol’ refers to the fraction of light petroleum ether boiling in the range 40-60 °C. All reactions using anhydrous conditions were performed using flame-dried apparatus under an atmosphere of argon or nitrogen. Brine refers to a saturated solution of sodium chloride. Anhydrous magnesium sulfate (MgSO₄) or sodium sulphate (Na₂SO₄) were used as drying agents after reaction workup, as indicated.

Protein modifications and characterization. Proteins were purified using ÄKTA FPLC system (GE Healthcare) monitored by UV absorption at 254 nm or temperature compensated conductivity. Protein concentrations were determined by BCA (bicinchoninic acid) Assay (Pierce). Bovine Serum Albumin (BSA) and Horse Heart Myoglobin (HHM) were used as standards and purchased from Sigma Aldrich.

Liquid chromatography-mass spectrometry (LC-MS) was performed on a Micromass LCT (ESI-TOF-MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C4

column (250 x 4.6 mm x 5 μ m). Water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase at a flow rate of 1.0 mL min⁻¹. The gradient was programmed as follows: 95% A (5 min isocratic) to 95% B over 15 min then isocratic for 5 min. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V for Subtilisin *Bacillus lentus* (SBL). Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L/hr. Spectra were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25 V. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v.4.0 from Waters) according to manufacturer's instructions.

Modified and unmodified proteins were not resolved during chromatography and typically eluted between 13 and 16 min. MS data therefore reflects the total protein content of the reaction mixture. Conversions were calculated from the relative intensities of mass peaks after MaxEnt processing. We assume that both modified and unmodified proteins are ionized with similar efficiency (NB: This was also confirmed by a LC-MS assay in Section 3.2.3 of this Thesis).

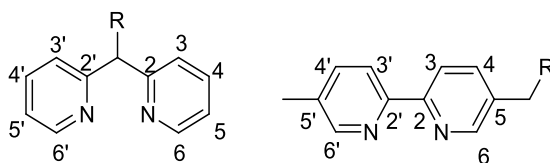
Design of the system

2.5.2 Modelling studies

Preliminary modelling studies to visualize the substrate binding pockets and catalytic triad were carried out using Open-source PyMOLTM 1.7.x software, starting with a published SBL-WT structure (PDB ID: 1GCI).²⁵

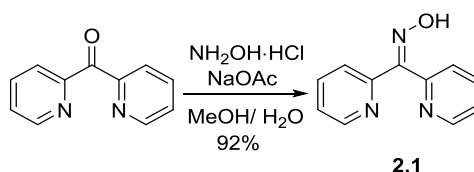
2.5.3 Ligand synthesis

Explanatory note. Bipyridine and dipyridyl structures have been numbered as below.



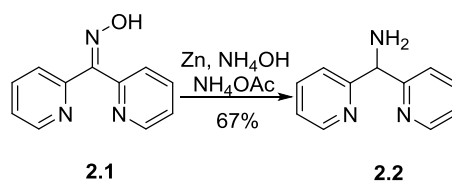
Synthesis of L1

Di(pyridin-2-yl)methanone oxime 2.1



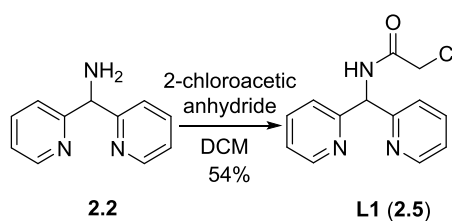
Synthesis of **2.1** was accomplished by following a published procedure.⁴⁰ Hydroxylamine hydrochloride (375 mg, 5.4 mmol) and NaOAc (443 mg, 5.4 mmol) were dissolved in 5 mL H₂O and heated at 60 °C for 1 h. 2,2'-dipyridylketone (500 mg, 2.7 mmol) in MeOH (1 mL) was added and the resulting mixture was stirred overnight at 60 °C. The reaction mixture was cooled to room temperature and then on ice. The corresponding oxime precipitated and was collected by filtration. The product was washed with cold H₂O and dried under vacuum, to obtain a light pink powder (496 mg, 92%), which was used in the subsequent step without further purification. $R_f = 0.29$ (SiO₂, 10% EtOH/ CHCl₃); IR (neat): 2804, 1594, 1568, 1472, 1428, 1337, 1283, 1155, 1095, 1052, 1001, 945, 789 cm⁻¹, ¹H NMR (400 MHz, CDCl₃): δ = 7.36 (ddd, J =6.9, 4.8, 2.1, 1H, H₅), 7.46 (ddd, J =7.6, 5.0, 1.2, 1H, H₅), 7.67 (dt, J =8.2, 1.1, 1H, H₃), 7.80 – 7.91 (m, 3H, H₄ + H₃), 8.63 (dt, J =5.0, 1.3, 1H, H₆), 8.65 (dt, J =4.8, 1.4, 1H, H₆), 16.31 (s, 1H, OH); ¹³C NMR (50.3 MHz, CDCl₃): δ = 123.6 (C₅), 124.3 (C₅), 124.9 (C₄), 125.1 (C₃), 137.2 (C₄), 137.8 (C₃), 145.9 (C₆), 148.5 (C₆), 150.4 (C₂), 151.3 (C₂), 154.5 (C=NOH); LRMS m/z (ESI⁺): 222.07 [M+Na]⁺.

(2, 2'-Dipyridyl)methylamine **2.2**



Synthesis of **2.2** was carried out by following a published procedure.⁴⁰ 2,2'-dipyridylketone oxime **2.1** (400 mg, 2 mmol) was dissolved in EtOH (8 mL). NH₄OAc (280 mg, 3.4 mmol), NH₄OH 25% (6 mL) and H₂O (4 mL) were added to the solution and the mixture was heated at 80 °C. Zinc powder was added in small portions over 30 min, after which the mixture was refluxed for 5 h and then stirred at room temperature overnight. The resulting suspension was filtered to remove the salts and the filtrate was concentrated under vacuum to yield an aqueous solution, which was made alkaline with 5 M NaOH and extracted with CH₂Cl₂. The organic phase was washed with brine, dried over MgSO₄ and evaporated under vacuum, to yield a yellow oil (252 mg, 67%). No further purification was found necessary. IR (neat): 3367, 2922, 2852, 1681, 1587, 1569, 1466, 1433, 1317, 994.29, 947 cm⁻¹; ¹H NMR (400 MHz, CDCl₃): δ = 2.35 (br. s, 2H, NH₂), 5.33 (s, 1H, CHNH₂), 7.14 (ddd, *J*=7.5, 4.8, 1.2, 2H, H₅), 7.39 (dt, *J*=7.8, 1.1, 2H, H₃), 7.63 (td, *J*=7.7, 1.8, 2H, H₄), 8.56 (ddd, *J*=4.8, 1.8, 0.9, 2H, H₆); ¹³C NMR (101 MHz, CDCl₃): δ = 62.3 (CHNH₂), 121.8 (2C₅), 122.1 (2C₃), 136.6 (2C₄), 149.2 (2C₆), 162.7 (2C₂); LRMS *m/z* (ESI⁺): 186.11 [M+H]⁺.

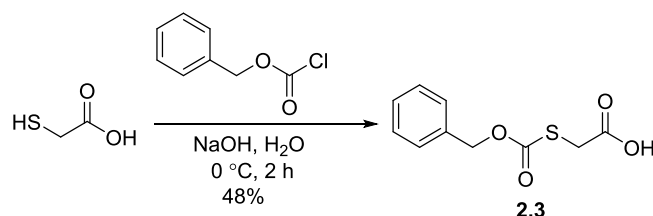
2-Chloro-*N*-(di(pyridin-2-yl)methyl)acetamide **L1**



Synthesis of **L1** was accomplished by following a published procedure.⁴⁰ To a solution of (2,2'-dipyridyl) methylamine (280 mg, 1.4 mmol) in dry CH₂Cl₂ (30 mL) was added chloroacetic anhydride (462 mg, 2.72 mmol) as a solid, under inert atmosphere. After 2 h stirring at room temperature, water was added and the organic phase was separated. The aqueous phase was adjusted to pH 8.5 with sat. NaHCO₃, extracted with CH₂Cl₂ and dried over MgSO₄. The solvent was evaporated and crude product was dried in vacuo. Purification by flash chromatography on silica gel using EtOAc/ petroleum ether gradient (70% to 100% EtOAc) to yield a light yellow solid **L1** (**2.5**) (200 mg, 54%). R_f = 0.35 (SiO₂, 90% EtOAc/ Hexane); IR (neat): 3367, 3056, 3007, 2957, 2922, 2852, 1717, 1681, 1651, 1587, 1466, 1281, 1243, 1048, 994, 784 cm⁻¹; ¹H NMR (400 MHz, DMSO-d₆): δ = 4.30 (s, 2H, CH₂Cl), 6.19 (d, 1H, *J* = 7.8 Hz, CHNH), 7.29 (ddd, 2H, *J* = 7.4, 4.9, 0.9 Hz, H₅), 7.51 (d, 2H, *J* = 7.8 Hz, H₃), 7.79 (td, *J* = 7.6, 1.8, 2H, H₄), 8.51 (d, 2H, *J* = 4.5 Hz, H₆), 9.24 (d, 1H, *J* = 7.8 Hz, CHNH) ppm; ¹³C NMR (50.3 MHz, DMSO-d₆): δ = 42.9 (CH₂Cl), 59.8 (-CHNH), 122.3, (2C₅), 123.2 (2C₃), 137.6 (2C₄), 149.2 (2C₆), 159.2 (2C₂), 166.2 (C=O); LRMS m/z (ESI⁺): 262.08 [M+H]⁺. HRMS m/z (ESI⁺): Found 262.07397 [M+H]⁺; C₁₃H₁₃ClN₃O requires 262.07417.

Synthesis of L3

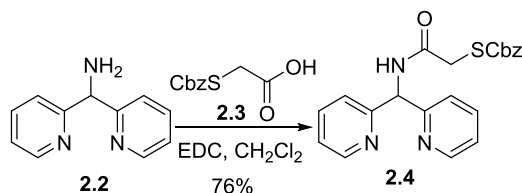
2-(((benzyloxy)carbonyl)thio)acetic acid **2.3**



To a solution of mercaptoacetic acid (5.1 g, 50.8 mmol) at 0 °C in NaOH 2M (48 mL) were added in two portions benzyl chloroformate (4.75 g, 10.8 mmol) and NaOH 2M (6 mL) with vigorous stirring. The cloudy mixture was stirred for 1 h at 0 °C, after which the

suspension was poured in a separatory funnel and washed thrice with Et₂O. The water layer was separated and acidified to pH 2 with HCl 1M to precipitate the product, which was collected by filtration. The white precipitate was washed twice with cyclohexane / toluene mixture and dried under vacuum to yield a white powder **2.3** (5 g, 48%), which was used without further purification. IR (neat): 3007.68, 1697.03, 1419.71, 1379.90, 1317.24, 1220.09, 1141.94, 907.93, 795.14, 742.89, 694.98, 670.36, 626.71; ¹H NMR (400 MHz, CDCl₃): δ = 3.73 (s, 1H, CH₂SCbz), 5.28 (s, 1H, PhCH₂O), 7.38 (s, 2H, CHar), 8.28 (s, 1H, COOH); ¹³C NMR (101 MHz, CDCl₃): δ = 33.1(CH₂SCbz), 69.9 (PhCH₂O), 128.5, 128.7, 128.8 (CHar), 134.6 (Cqar), 169.6 (O(C=O)S), 174.2 (COOH). LRMS m/z (ESI): 225.04 [M-H]⁻.

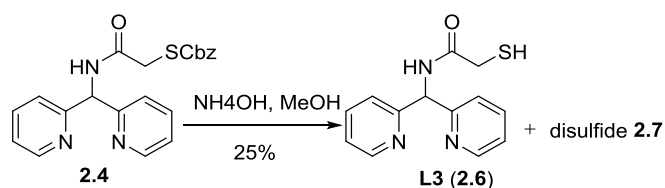
O-benzyl *S*-(2-((di(pyridin-2-yl)methyl)amino)-2-oxoethyl) carbonothioate **2.4**



To a solution of *S*-Carbobenzyloxy-mercaptoacetic acid **2.3** (305 mg, 1.35 mmol) in CH₂Cl₂ (5 mL) at 0 °C was added (2,2'-dipyridyl)methylamine (250 mg, 1.35 mmol) in CH₂Cl₂ (8.5 mL) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 310 mg, 1.62 mmol) as a solid. Additional CH₂Cl₂ (10 mL) was added and the solution was stirred at room temperature 24 h. The solvent was removed under reduced pressure and EtOAc was added to the crude product, followed by H₂O. The organic phase was extracted and washed with saturated NaHCO₃, then H₂O, dried over Na₂SO₄ and evaporated. The crude product was purified by flash chromatography on silica gel using CH₂Cl₂ / MeOH 95 / 5 solvent mixture, to yield 400 mg of a white solid **2.4** (76%). R_f = 0.34 (SiO₂, 5% MeOH/ DCM); IR (neat): 1711.77, 1675.26, 1588.45, 1568.21, 1531.81,

1469.29, 1426.87, 1365.47, 1127.50, 1046.87, 1000.21, 909.60, 803.94, 782.04, 755.18, 700.47, 675.27, 660.39, 614.83; ^1H NMR (400 MHz, CDCl_3) ^1H NMR (400 MHz, CDCl_3): δ = 3.74 (s, 2H, CH_2S), 5.31 (s, 2H, CH_2O), 6.23 (d, $J=6.6$, 1H, CHNH), 7.19 (ddd, $J=7.5$, 4.9, 1.2, 2H, CHar), 7.34 – 7.42 (m, 5H, CHar), 7.45 (dt, $J=8.0$, 1.1, 2H, CHar), 7.65 (td, $J=7.7$, 1.8, 2H, CHar), 8.55 (ddd, $J=4.9$, 1.9, 0.9, 2H, CHar), 8.64 (d, $J=6.4$, 1H, CHNH); ^{13}C NMR (126 MHz, CDCl_3) δ = 34.8 (SCH_2CO), 59.2 (CHNH), 69.7 (CH_2OCOS), 122.2 (2C_3), 122.7 (2C_5), 128.4 (3CHar), 128.6 (2CHar), 134.8 (C_{qar}), 137.3 (2C_4), 148.8 (2C_6), 158.2 (2C_2), 167.1 (CH_2CO), 170.2 ($\text{O}(\text{C}=\text{O})\text{S}$). HRMS m/z (ESI^+): Found 416.1034 $[\text{M}+\text{Na}]^+$; $\text{C}_{21}\text{H}_{19}\text{N}_3\text{NaO}_3\text{S}$ requires 416.1039.

N-(di(pyridin-2-yl)methyl)-2-mercaptoacetamide **L3**



To a solution of *S*-carbobenzyloxy-mercaptoacetic acid, (2,2'-dipyridyl)methylamide (250 mg, 0.63 mmol) in degassed MeOH (2 mL) was added a degassed solution of NH_4OH 25% (2 mL). Additional MeOH (4 mL) and NH_4OH 25% (2 mL) were added to provide a yellow suspension, which was stirred for 4 h at room temperature, after which time ESI-MS analysis revealed complete conversion of the starting material into **L3** and its corresponding disulfide. After evaporation of MeOH, the aqueous phase was adjusted to pH 7 and extracted with CH_2Cl_2 . The organic phase was dried over Na_2SO_4 and evaporated. Flash chromatography on silica gel (CH_2Cl_2 / MeOH 95 / 5 solvent mixture) yielded 42 mg (25%) of **L3** as yellow solid and 100 mg of a mixture of **L3** and the corresponding disulphide. R_f = 0.36 (SiO_2 , 5% MeOH/ DCM); ^1H NMR (400 MHz, CDCl_3): δ = 2.04 (t, $J=8.9$, 1H, SH), 3.36 (d, $J=8.9$, 2H, CH_2SH), 6.21 (d, $J=6.7$, 1H, CHNH), 7.18 (ddd, $J=7.5$, 4.9, 1.2, 3H, C_5), 7.44 (dt, $J=7.9$, 1.1, 2H, C_3), 7.65 (td, $J=7.7$,

1.8, 2H, C₄), 8.57 (ddd, *J*=4.9, 1.8, 0.9, 2H, C₆), 8.81 (d, *J*=6.7, 1H, NH). HRMS *m/z* (ESI⁺): Found 282.0667 [M+Na]⁺; C₁₃H₁₃N₃NaOS requires 282.0672.

NB This compound was found to oxidise spontaneously in air to form the corresponding disulphide compound. For protein modification reactions, the **L3** was used directly after quick purification on silica gel.

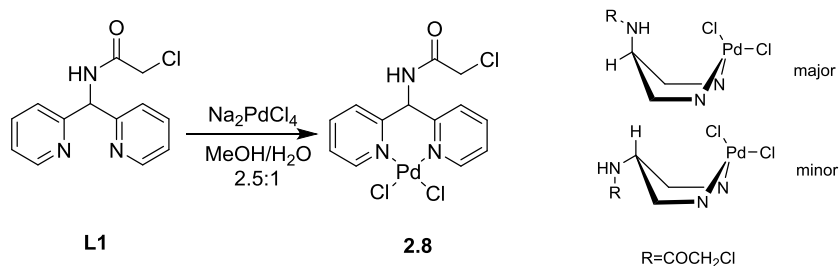
2,2'-disulfanediylbis(N-(di(pyridin-2-yl)methyl)acetamide) disulphide 2.7

R_f = 0.22 (SiO₂, 5% MeOH/ DCM); IR (neat): 3269.60, 1686.05, 1589.04, 1496.16, 1470.13, 1433.91, 1403.22, 1342.23, 1120.94, 1084.53, 1069.88, 1027.39, 995.89, 910.92, 750.35, 731.40, 696.47, 612.90; ¹H NMR (400 MHz, CDCl₃): δ = 3.60 (s, 2H, CH₂S-SCH₂), 6.28 (d, *J*=6.9, 2H, 2xCHNH), 7.12 – 7.18 (m, 4H, 2xC₅), 7.45 (d, *J*=7.8, 4H, 2xC₃), 7.62 (td, *J*=7.7, 1.9, 4H, 2xC₄), 8.49 – 8.54 (m, 4H, 2xC₆), 8.69 (d, *J*=6.9, 4H, 2xNH). ¹³C NMR (101 MHz, CDCl₃) δ = 42.8 (CH₂S), 59.5 (CHNH), 122.3 (C₅), 122.6 (C₃), 137.0 (C₄), 149.1 (C₆), 158.5 (C₂), 167.8 (C=O). HRMS *m/z* (ESI⁺): Found 539.1296 [M_{disulfide}+Na]⁺; C₂₆H₂₄N₆NaO₂S₂ requires 539.1294.

Synthesis of **L1-PdCl₂** complex **2.8**

This synthesis was done by Dr. Anca Pordea in our lab. Characterization was completed to confirm identity before use.

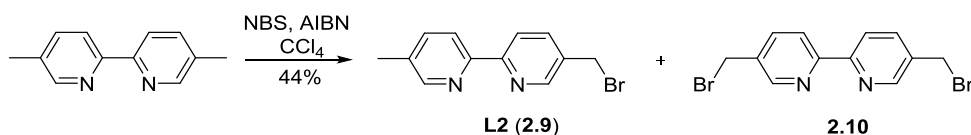
Chloroacetic acid, (2,2'-dipyridyl)methylamide, palladium complex 2.8



Palladium chloride (56.7 mg, 0.32 mmol) and NaCl (37.4 mg, 0.64 mmol) were mixed in a test tube with and 1.28 mL H₂O and the suspension was stirred at 45 °C until complete dissolution of the insoluble PdCl₂, with formation of a dark brown solution of Na₂PdCl₄ (20 min). A solution of **L1** (100 mg, 0.38 mmol) in MeOH (3.2 mL) was added and immediate precipitation of the yellow palladium complex occurred. The mixture was stirred for 2 h at 37 °C, then at room temperature for 1 h. The palladium complex was filtered, thoroughly washed with H₂O and MeOH and dried in vacuo to yield 91 mg (54%) of a yellow powder **2.8**, which was insoluble in most organic solvents and in water. Spectral data are consistent with those reported in literature. IR (neat): 3234.33, 1676.44, 1607.43, 1513.38, 1479.49, 1447.01, 1367.34, 1314.88, 1286.43, 1224.28, 1188.07, 1152.66, 1061.14, 969.47, 854.35, 793.47, 782.68, 758.74, 694.37, 625.97; ¹H NMR (400 MHz, DMSO-d₆): major isomer (~77%) δ = 4.55 (2H, s, CH₂Cl), 7.32 (1H, d, J= 9.7 Hz, CHNH), 7.60 (2H, m, C₅), 7.81 (2H, d, 8.08 Hz, C₃), 8.20 (2H, m, C₄), 8.88 (2H, d, J= 5.81 Hz, C₆), 10.19 (1H, d, J= 9.7 Hz, CHNH) ppm; minor isomer (~23%): 4.33 (2H, s, CH₂Cl), 6.44 (1H, s, CHNH), 7.56 (2H, m, C₅), 7.93 (2H, d, 7.58 Hz, C₃), 8.10 (2H, m, C₄), 8.89 (2H, d, 5.81 Hz, C₆), 10.11 (1H, s, CHNH) ppm; LRMS m/z (ESI⁺): 437.90 [M-H]⁺.

Synthesis of L2

5-(Bromomethyl)-5'-methyl-2,2'-bipyridine **L2**



Synthesis of **L2** was accomplished by following a published procedure.⁴¹ To a solution of 5,5'-dimethyl-2,2'-bipyridine (517 mg, 2.80 mmol) and *N*-bromosuccinimide (500 mg, 2.80 mmol) in dry carbon tetrachloride (25 mL), AIBN was added (10 mg, catalytic

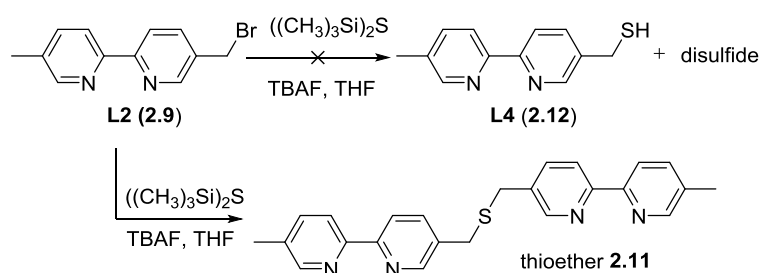
amount) and the reaction mixture was refluxed at 80 °C for 2.5 h under argon. The solution was filtered while hot. The solvent was removed under vacuum and the residue was dissolved in CH₂Cl₂. The solution was left at room temperature for five days to precipitate out most of the less soluble dibromo derivative formed. The solution was filtered and the solvent was evaporated. The residue thus obtained was purified by flash chromatography (SiO₂, 10-20% EtOAc/ DCM) which gave compound **L2** as a white powder (320 mg, 44%). R_f = 0.3 (SiO₂, 40% EtOAc in n-hexane); M.p. = 80 °C. IR (neat): 2917, 1596, 1470, 1253, 1221, 1054, 773 cm⁻¹; ¹H NMR (400 MHz, CDCl₃): δ = 2.42 (s, 3H, CH₃), 4.54 (s, 2H, CH₂Br), 7.69 (dd, 1H, J = 8.0, 1.5 Hz, H_{4'}), 7.86 (dd, 1H, J = 8.0, 2.3 Hz, H₄), 8.33 (d, 1H, J = 8.0 Hz, H_{3'}), 8.43 (d, 1H, J = 8.0 Hz, H₃), 8.54 (br. s., 1H, H_{6'}), 8.68 (s, 1H, H₆) ppm; ¹³C NMR (101 MHz, CDCl₃): δ = 18.4 (CH₃), 29.6 (CH₂Br), 121.0 (C_{3'}), 121.1 (C₃), 133.6 (C_{5'}), 134.1 (C₅), 137.7 (C_{4'}), 138.3 (C₄), 149.0 (C_{6'}), 149.3 (C₆), 152.3 (C_{2'}), 155.2 (C₂); LRMS m/z (ESI⁺): 285.02 [M+Na]⁺.

5,5'-bis(bromomethyl)-2,2'-bipyridine 2.10

Yield = 27%, R_f = 0.44 (SiO₂, 40% EtOAc in n-hexane); IR (neat): 3365, 3015, 2923, 2853, 1963, 1721, 1643, 1594, 1469, 1280, 1254, 1053, 943, 777 cm⁻¹; ¹H NMR (400 MHz, CDCl₃): δ = 4.55 (s, 4H, CH₂Br), 7.87 (dd, J = 8.2, 2.3, 2H, H₄), 8.42 (d, J = 8.2, 2H, H₆), 8.69 (d, J = 2.3, 2H, H₃); ¹³C NMR (101 MHz, CDCl₃): δ = 29.5 (CH₂Br), 121.2 (2C₃), 133.9 (2C₅), 137.8 (2C₄), 149.3 (2C₆), 155.3 (2C₂); LRMS m/z (ESI⁺): 364.93 [M+Na]⁺.

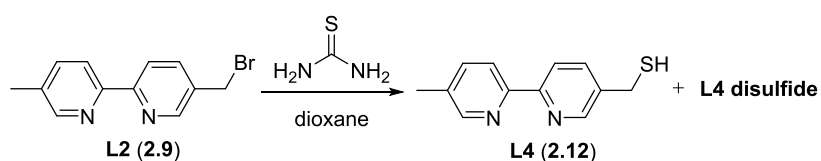
Synthesis of **L4**

bis((5'-methyl-[2,2'-bipyridin]-5-yl)methyl)sulfane thioether 2.11



To an ice-cooled solution of compound **L2** (100 mg, 0.38 mmol) in THF (3 mL) was added a mixture of tetrabutylammonium fluoride (110 mg, 0.42 mmol) and hexamethyldisilathiane (84 mg, 0.46 mmol) in THF (2 mL) under stirring. The reaction mixture was stirred for a further period of 12 h. It was then diluted with dichloromethane and washed with saturated ammonium chloride (10 mL) and then with distilled water (20 mL x 2). The organic layer was separated and dried over anhydrous sodium sulfate, and the solvent was removed under reduced pressure to give a residue that was purified by column chromatography that only afforded **thioether 2.11** product (70 mg, 46%) and a trace amount of **L4 disulfide** (only detected in LRMS). $R_f = 0.37$ (SiO₂, 4% MeOH/DCM); ¹H NMR (400 MHz, CDCl₃): $\delta = 2.41$ (s, 6H, CH₃ x 2), 3.67 (s, 4H, CH₂Br x 2), 7.62 – 7.67 (m, 2H, H_{4'} x 2), 7.78 (dd, J=8.2, 2.2, 2H, H₄ x 2), 8.28 (d, J=8.0, 2H, H_{3'} x 2), 8.35 (d, J=8.2, 2H, H₃ x 2), 8.51 (d, J=1.9, 2H, H_{6'} x 2), 8.56 (d, J=1.9, 2H, H₆ x 2); ¹³C NMR (101 MHz, CDCl₃): $\delta = 18.4$ (CH₃ x 2), 32.6 (CH₂Br x 2), 120.8 (C_{3'} x 2), 120.8 (C₃ x 2), 133.3 (C_{5'} x 2), 133.6 (C₅ x 2), 137.5 (C_{4'} x 2), 137.8 (C₄ x 2), 149.3 (C_{6'} x 2), 149.4 (C₆ x 2), 153.0 (C_{2'} x 2), 155.0 (C₂ x 2); LRMS m/z (ESI⁺): 391.1 [M+H]⁺.

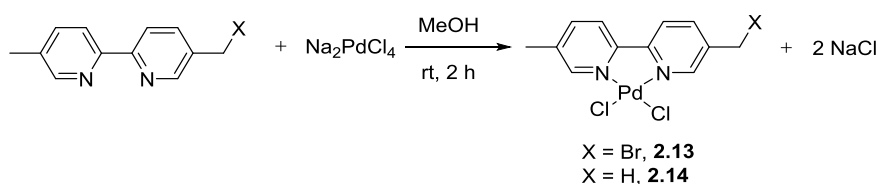
(5'-methyl-[2,2'-bipyridin]-5-yl)methanethiol L4



Synthesis of **L4** was carried out by following a published procedure. A solution of 5-bromomethyl-5'-methyl-2,2'-dipyridyl **L2** (60 mg, 0.23 mmol) and thiourea (55 mg, 0.7 mmol) in 15 mL dry dioxane was heated to reflux (102 °C) under N₂ for 3 h. The reaction mixture steadily changed from pale to cloudy white within first 10 min. After 3 h, NaOH (40 mg, 1 mmol) dissolved in 5 mL of degassed H₂O was added (colour changed to yellow) and the mixture was heated to reflux for another 8 h. The solvent was evaporated, pH adjusted to ~4 using HCl, and extracted with CH₂Cl₂. The organic phase was collected, and the solvent was removed. The residue was column chromatographed over silica using (2% MeOH/ DCM) and the product was obtained as yellow oil which contained a mixture of **L4** and **L4 disulfide** (10 mg). R_f = 0.24 (SiO₂, 2% MeOH/ DCM). IR (neat): 2918.71, 1594.70, 1552.96, 1465.55, 1375.06, 1242.13, 1117.41, 1056.44, 1027.48, 832.72, 778.29, 740.16, 650.56. HRMS m/z (ESI⁺): Found 453.1188 [M+Na]⁺; C₂₄H₂₂N₄NaS₂ requires 453.1178.

NB This compound was found sensitive to auto-oxidation to form the corresponding disulphide compound. Hence, for protein modification reactions, the **L4** was used directly after quick purification on silica gel.

Synthesis of L2-PdCl₂ complex **2.13**



To a solution of **Lx** (50 mg, 0.252 mmol) in CH₃OH (10 mL) were added Na₂PdCl₄ (0.25 mmol, 0.5 M in H₂O). In a few min the product began to precipitate. The mixture was stirred at room temperature for 2 h. The precipitate formed was filtered off and washed with water, ethanol and diethyl ether to obtain the complex **2.13** as a yellow solid (50 mg,

46%). IR (neat): 3044.80, 1602.86, 1473.55, 1403.61, 1309.43, 1248.57, 1208.10, 1159.97, 1141.73, 1052.64, 1036.93, 939.84, 839.67, 761.41, 720.75, 632.69; ^1H NMR (400 MHz, DMSO) δ 9.44 (s, 1H), 9.17 (d, J = 18.3 Hz, 1H), 8.89 (d, J = 8.8 Hz, 1H), 8.52 (t, J = 8.5 Hz, 1H), 8.45 (dd, J = 8.1, 3.5 Hz, 1H), 8.39 (d, J = 8.0 Hz, 1H), 8.19 (d, J = 7.8 Hz, 1H), 4.95 (d, J = 37.4 Hz, 2H), 2.46 (s, 3H). ^{13}C NMR (126 MHz, DMSO) δ = 18.4 (CH_3), 42.1 (CH_2Br), 123.4 (C_3), 123.7 (C_3), 136.9 (C_5), 138.1 (C_5), 141.1 (C_4), 141.5 (C_4), 149.3 (C_6), 149.4 (C_6), 153.4 (C_2), 156.1 (C_2).

Bipyridine-PdCl₂ complex 2.14

Yellow-brown solid (55 mg, 61%), IR (neat): 3041.98, 1607.34, 1468.48, 1390.12, 1326.42, 1309.19, 1250.35, 1228.18, 1157.63, 1139.79, 1052.42, 994.54, 833.13, 717.22, 665.90; ^1H NMR (400 MHz, DMSO): δ = 2.44 (s, 6H, CH_3), 8.17 (dd, J =8.2, 1.2, 2H, H_4), 8.42 (d, J =8.2, 2H, H_6), 8.90 (s, 2H, H_3); ^{13}C NMR (126 MHz, DMSO): δ = 18.3 (CH_3), 123.0 (C_3), 137.4 (C_5), 141.4 (C_4), 149.3 (C_6), 153.9 (C_2).

Construction of the system

2.5.4 Protein Sequences

Average molecular weight was computed by using online Expasy tool:

http://web.expasy.org/compute_pi/

Subtilisin *Bacillus Lentus* wild-type (SBL-WT)

(Active site Ser221 is shown in red; based on Subtilisin BPN' numbering)

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPT
QDGNGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEW
AGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYA
NAMAVGATDQNNNRASFQYAGLDIVAPGVNVQSTYPGSTYASLNGT^SMATPH
VAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Theoretical pI/Mw: 9.30 / 26698.50

SBL-L217C (*Cys217 is shown in red*)

AQSVPWGISRVQAPAAHNRGLTGSQVKVAVLDTGISTHPDLNIRGGASFVPGEPT
QDGNGHGHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEW
AGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYA
NAMAVGATDQNNNRASFQYAGGLDIVAPGVNVQSTYPGSTYAS^CNGTSMATPH
VAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Theoretical pI/Mw: 9.14 / 26688.48

SBL-S156C (*Cys156 is shown in red*)

AQSVPWGISRVQAPAAHNRGLTGSQVKVAVLDTGISTHPDLNIRGGASFVPGEPT
QDGNGHGHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEW
AGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAAS^CGNSGAGSISYPARY
ANAMAVGATDQNNNRASFQYAGGLDIVAPGVNVQSTYPGSTYASLNGTSMATP
HVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Theoretical pI/Mw: 9.14 / 26714.56

2.5.5 Protein purification & characterization

Purification of SBL-S156C was done by Dr. Anca and Dr. Alba Díaz-Rodríguez. The purification of SBL-WT and variant proteins (SBL-L217C and SBL-S156C) was accomplished by following a published procedure.^{28,36} The crude protein extract containing stabilizers (5 mL) was treated with 0.08 M dithiothreitol (DTT). The solution was vortexed and shaken for 30 min at 4 °C to reduce possible disulfide bonds. The small molecular weight contaminants were removed by desalting the protein through a PD-10 desalting column containing Sephadex G-25 medium (GE Healthcare Life Sciences). The protein was eluted with acetate buffer (20 mM sodium acetate, 5 mM CaCl₂, pH 5.2). The protein was subsequently purified on a strong cation exchange HiTrap SP Sepharose HP column (GE Healthcare Life Sciences) and eluted with a gradient of 0-200 mM NaCl in sodium acetate buffer, pH 5.2. The purity and homogeneity of the enzymes was confirmed by non-

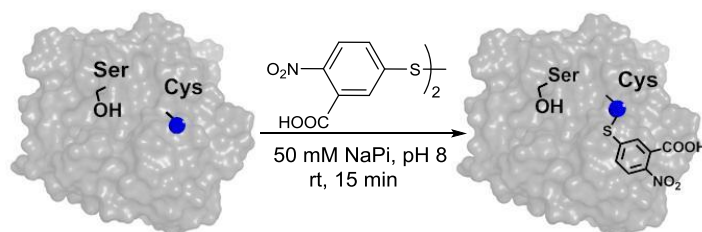
reducing PAGE gel analysis, after denaturation with 0.1 M HCl at 0 °C for 30 min. Prior to charging on the gel, the sample was diluted twice with Laemmli buffer (NB no β -mercaptoethanol was added). The eluted protein fractions were dialysed directly against 50 mM NaPi buffer pH 8, and used for subsequent reactions. Alternatively, the eluted fractions were pooled and dialysed against MilliQ water and salt-free enzyme powder was obtained after subsequent lyophilisation. ESI-MS analysis of the protein samples confirmed the molecular weight of the SBL WT and cysteine variants SBL-L217C and SBL-S156C.

Non-reducing PAGE gel analysis: NuPAGE® Bis-Tris gels (4-12%, Invitrogen) were used. Gels were placed in the gel tank, and the gel tank filled with 1x NuPAGE® MES SDS Running Buffer (800 mL, prepared from 20x Invitrogen concentrate). Samples for the gel were prepared by adding sample (9 μ L) to 1M HCl (1 μ L) and denaturing at 4 °C for 30 min. 10 μ L of this denatured sample was loaded onto the gel. The gel was run in MES running buffer at 200 V for 45 min. Staining of the gels was done with Instant Blue™ Coomassie Brilliant Stain (20 mL). It was allowed to develop for few hours on an orbital rocker, after which it destained overnight with MilliQ water.

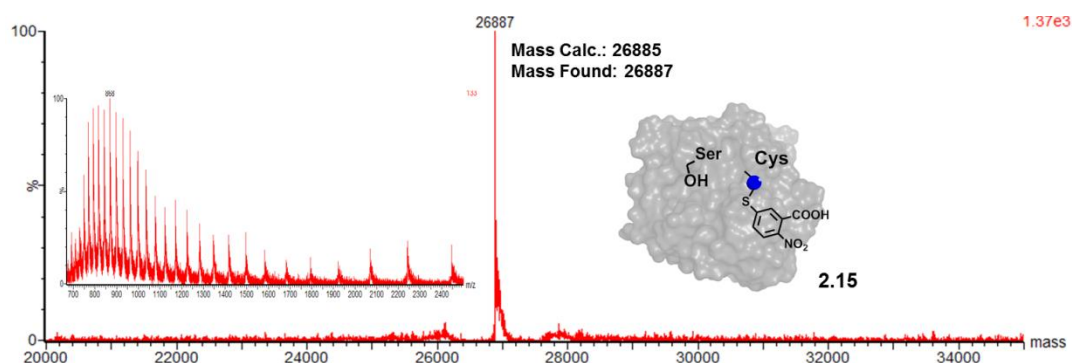
BCA Assay: Protein concentration was measured by BCA assay, prior to all protein modifications. BCA assay was performed in a 96-well format by employing BSA (Bovine Serum Albumin) standards and assay reagents A and B (ThermoScientific) mixed in a ratio 50:1. 200 μ L of the reagent solution was taken in each well and 10 μ L of BSA standards (0-2 mg/mL stock solutions) and the unknown protein samples were added successively. The plate was incubated at 37 °C for 30 min. Absorbance was measured at 562 nm using a

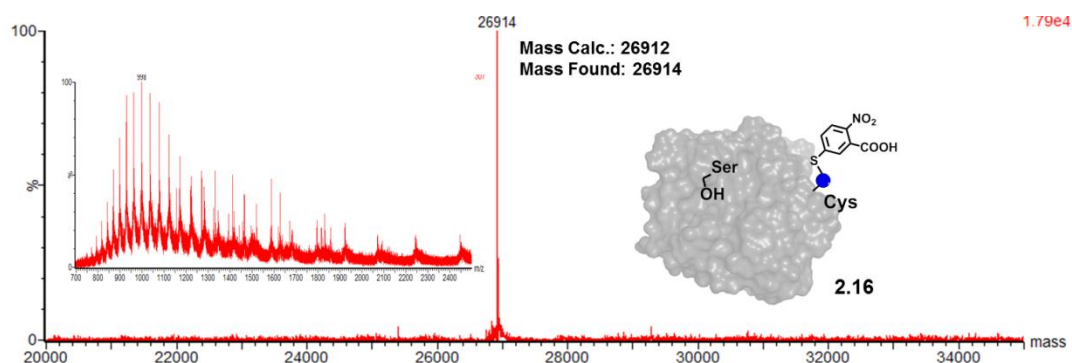
UV/ Visible spectrophotometer and a calibration curve was generated. Linearity was observed for the BSA standards from which unknown concentration was calculated.

Ellman's test for single cysteine mutants



Ellman's test: Presence of a single cysteine was confirmed by treatment of the SBL-L217C and SBL-S156C with 50 equiv. of Ellman's reagent. A 50 μ L aliquot of SBL variant was prepared at 0.8 mg/mL in sodium phosphate buffer (50 mM, pH 8). A 25 mM stock solution of Ellman's reagent was prepared in 50 mM sodium phosphate buffer, pH 8. To the 50 μ L sample of protein (1.1 nmol) was added 2.2 μ L of the Ellman's solution (55 nmol). The reaction was shaken for 30 min at room temperature and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the Ellman's disulfides (For SBL-L217C-Ellman's adduct (**2.15**) mass calculated: 26885, mass found: 26889; For SBL-S156C-Ellman's adduct (**2.16**) mass calculated: 26912, mass found: 26914). ESI-MS corresponding to the modifications are shown below. This experiment is also used as the negative control for the direct alkylation reactions to demonstrate that the alkylation occurs at the reactive Cys217 residue.

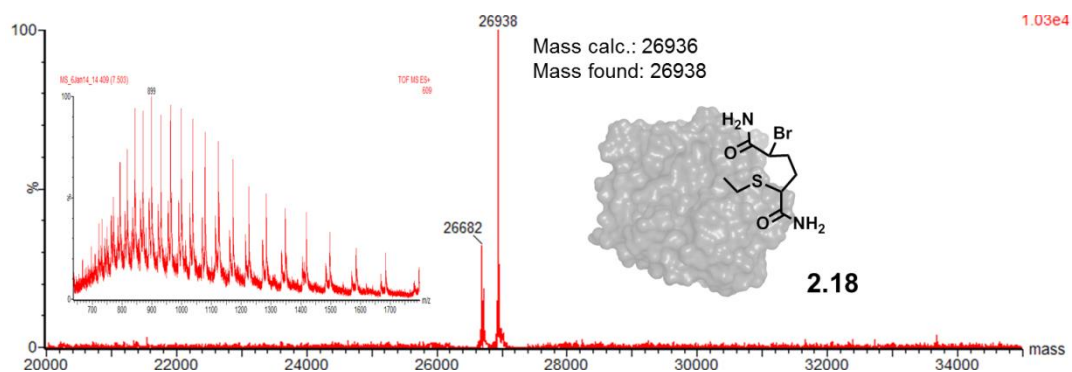




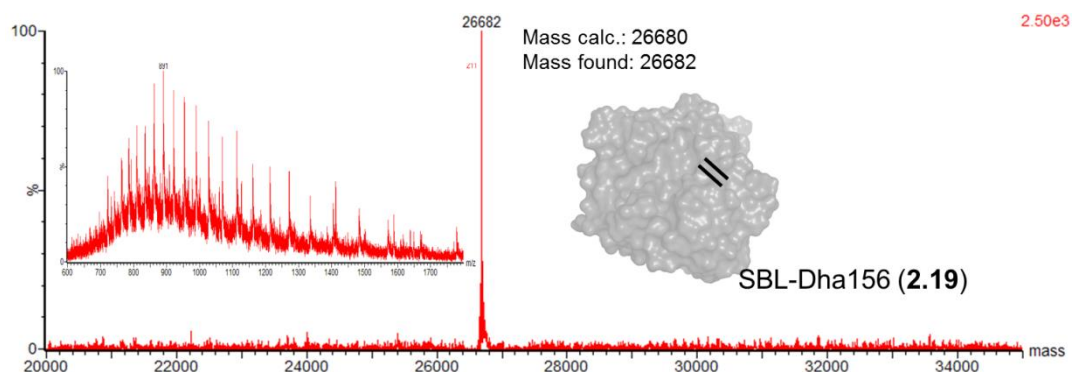
2.5.6 Artificial metalloenzymes using Dha route

Synthesis of SBL-Dha156

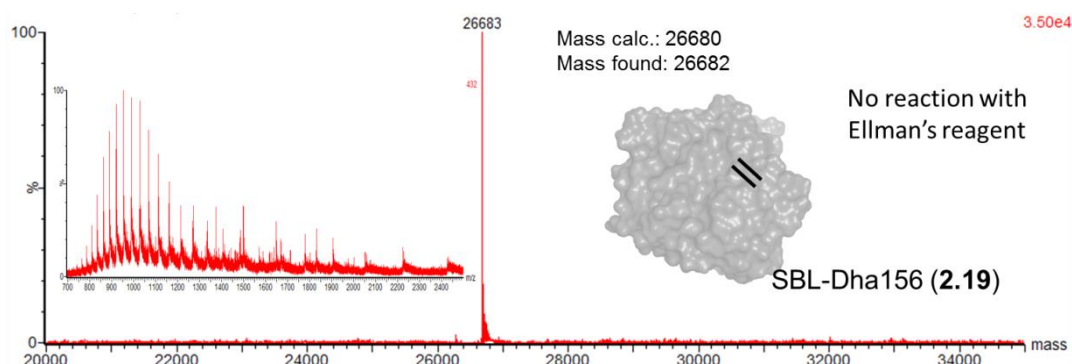
SBL-S156C was prepared at 2.0 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM) and a 1 M stock solution of DTT was prepared. To a 1 mL aliquot of the protein solution, 0.75 μ L of DTT solution (10 equiv.) was added to reduce any unwanted disulfides. After 20 minutes of incubation at room temperature, the DTT was removed by passing the protein solution through a PD MidiTrap™ column (GE Healthcare Life Sciences), eluting with 1.5 mL of sodium phosphate buffer (pH 8.0, 50 mM). To this protein solution (75 nmol) was added 1000 equiv. of 2,5-dibromohexanediamide reagent (DBHDA) **2.17** (75 μ mol) as solid and the reaction was shaken at room temperature for 30 min, after which an aliquot was analysed by LC-MS. Complete consumption of the starting material was observed which was converted into the uncyclized monoalkylated adduct **2.18** (mass calculated: 26936, mass found: 26938).



The reaction was further heated to 37 °C for 1 h. LC-MS analysis of the protein solution revealed the formation of SBL-Dha516 **2.19** (mass calculated: 26880, mass found: 26882). The unreacted DBHDA solid was separated by centrifuging the protein solution at 13,000 g and the protein was subsequently dialysed against 50 mM NaPi, pH 8 for 2 h and stored at -20 °C until further use. The ESI-MS for the modification is shown below.

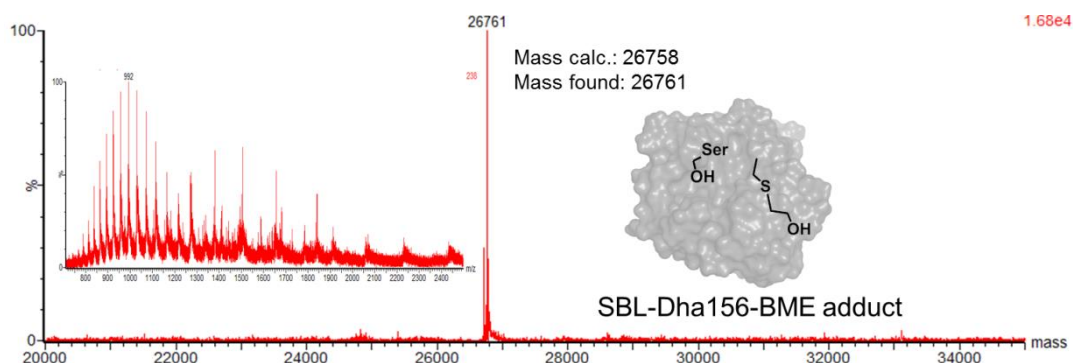


To the 20 μ L sample of protein (0.75 nmol) was added 0.65 μ L of the 0.25 M Ellman's solution (.15 μ mol). The reaction was shaken for 30 min at room temperature and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dha156 was seen (mass calculated: 26680, mass found: 26683). ESI-MS corresponding to the modification is shown below.

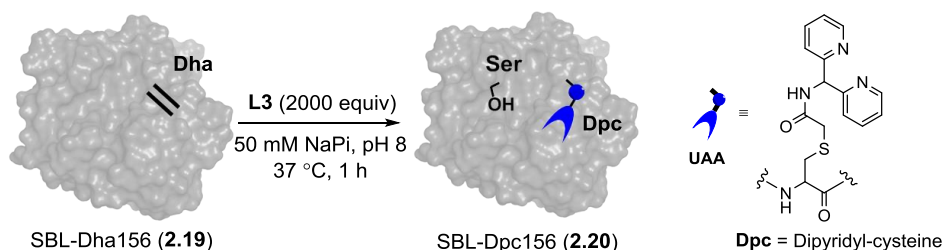


To the 20 μ L sample of protein (0.75 nmol) was added 0.65 μ L of the 0.25 M β -mercaptoethanol solution (.15 μ mol). The reaction was shaken for 30 min at room temperature and then analyzed directly by LC-MS. A single protein species with a mass

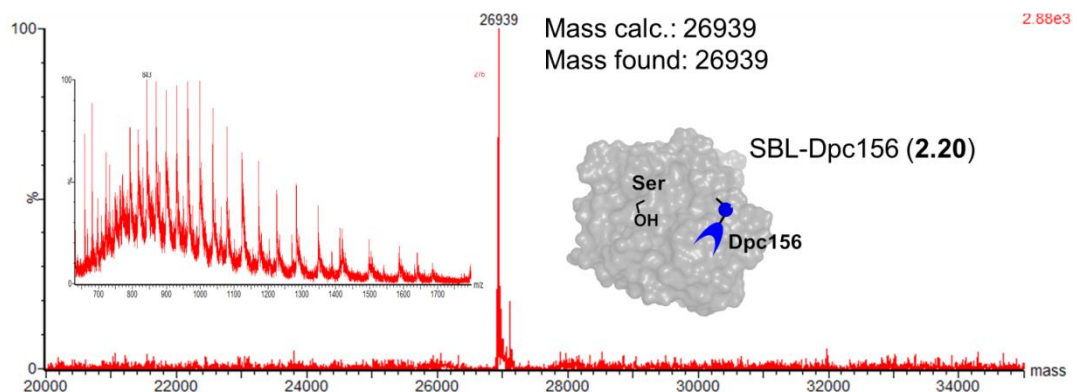
corresponding to the SBL-Dha156-BME was seen (mass calculated: 26758, mass found: 26761). ESI-MS of the modification is shown below.



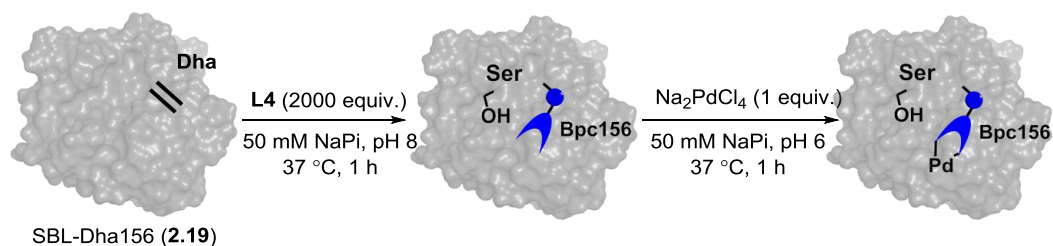
Synthesis of SBL-Dpc156



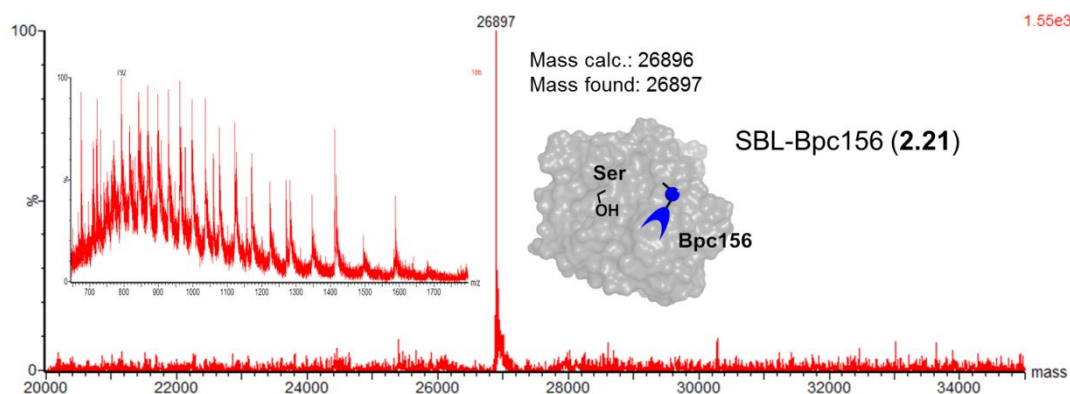
A solution of SBL-Dha156 (**2.19**) was prepared at a concentration of 0.8 mg/mL in 50 mM NaPi pH 8. A 0.72 M stock solution of thiol ligand **L3** was prepared in DMF. To a 180 μ L solution of SBL-Dha156 **2.19** (5.4 nmol) in 50 mM NaPi pH 8, was added 1.5 μ L thiol **L3** (1.07 μ mol). The protein reaction was shaken at 37 °C for 1 h, after which an aliquot was analysed by LC-MS. Complete conversion to SBL-Dpc156 (**2.20**) was observed (mass calculated: 26939, mass found: 26939). ESI-MS of the modification is shown below.



Synthesis of SBL-Bpc156 and SBL-Bpc156•Pd

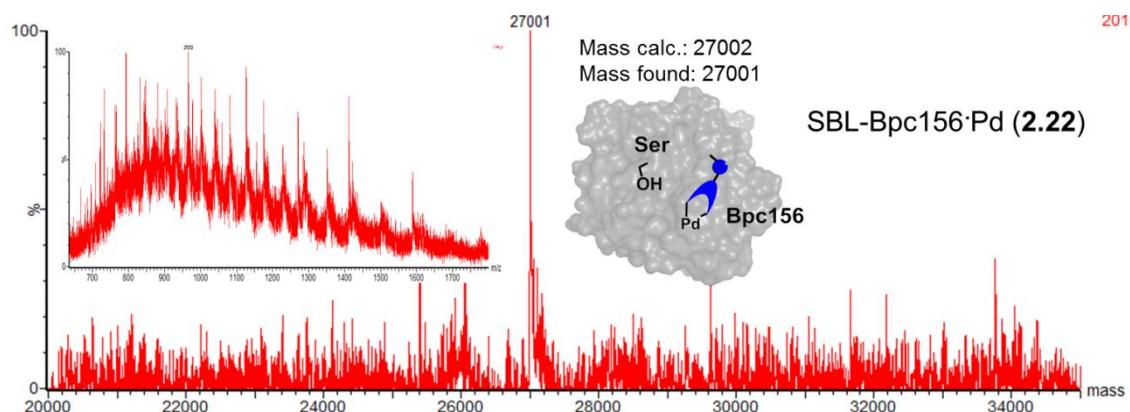


A solution of SBL-Dha156 (**2.19**) was prepared at a concentration of 0.8 mg/mL in 50 mM NaPi pH 8. A 0.72 M stock solution of thiol ligand **2.12** was prepared in DMF. To a 180 μL solution of SBL-Dha156 **2.19** (5.4 nmol) in 50 mM NaPi pH 8, was added 1.5 μL thiol **2.12** (1.07 μmol). The protein reaction was shaken at 37 $^\circ\text{C}$ for 1 h, after which an aliquot was analysed by LC-MS. Complete conversion to SBL-Dpc156 (**2.21**) was observed (mass calculated: 26896, mass found: 26897). The protein was desalted using PD SpinTrap™ G-25 column (GE Healthcare Life Sciences) by eluting with 50 mM NaPi pH 6. The ESI-MS for the modification is shown below.



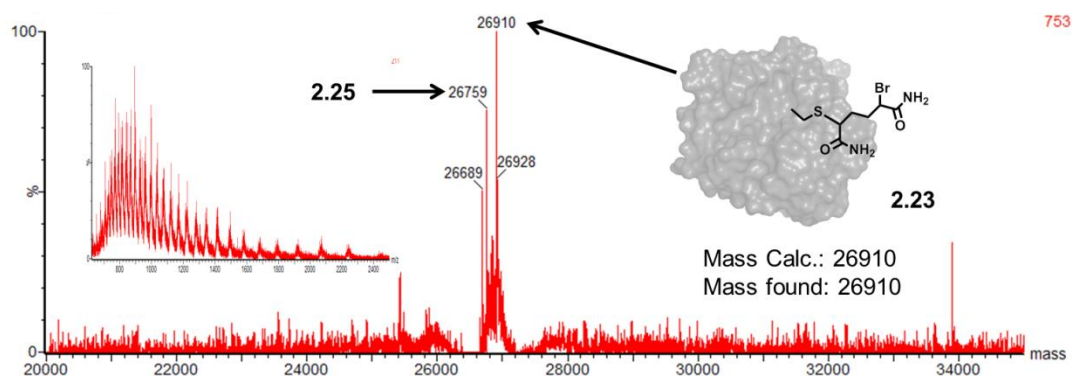
The protein solution containing SBL-Dpc156 was subsequently dialysed against 50 mM NaPi pH 6 to remove any excess thiol ligand in the solution. A solution of SBL-Bpc156 (**2.21**) was prepared at a concentration of 0.6 mg/mL in 50 mM NaPi pH 6. A 0.01 M stock solution of Na_2PdCl_4 was prepared in MilliQ water. To 180 μL of SBL-Bpc156 (4.08 nmol) in 50 mM NaPi pH 6 was added 0.4 μL of Na_2PdCl_4 (4.08 nmol) and the reaction was shaken at 37 $^\circ\text{C}$ for 1 h. LC-MS analysis of the protein solution revealed complete

uptake of metal into the protein to give SBL-Bpc156•Pd (**2.22**). The ESI-MS of the modification is shown below.

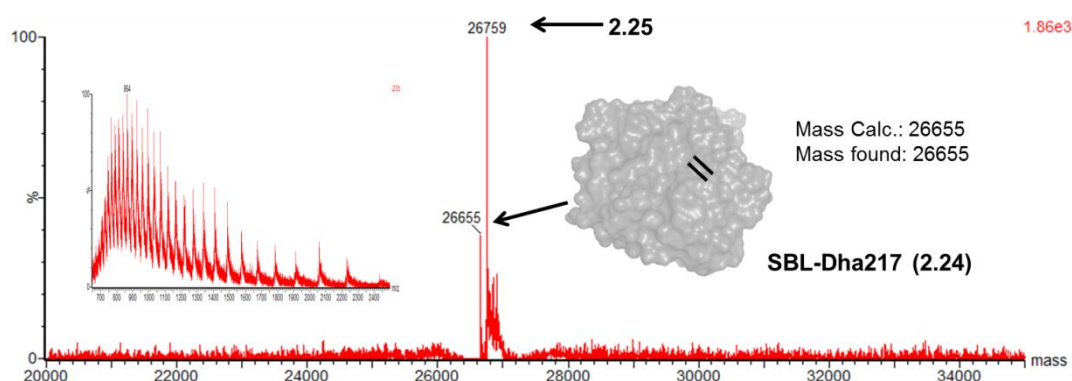


Synthesis of Dha on SBL-L217C using DBHDA reagent **2.17**

SBL-L217C was prepared at 1.0 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). DTT (1.4 mg) was added as a solid to a 500 μ L aliquot of the protein solution to reduce any contaminant disulfide. After 15 minutes of incubation at room temperature, the DTT was removed by passing the protein solution through a PD minitrap (GE Healthcare), eluting with 1 mL of sodium phosphate buffer (pH 8.0, 50 mM). A 100 μ L aliquot of the reduced protein (0.5 mg/mL, 0.002 μ mol) was used for the elimination reaction. Accordingly, a stock solution of DBHDA reagent was prepared by dissolving 35.5 mg in 418 μ L DMF. A 10 μ L aliquot of the DBHDA solution (3 μ mol) was added to the 100 μ L protein solution at room temperature. The reaction was vortexed to homogenize and shaken at room temperature for 30 min. After this time, solids were removed by centrifugation (1 min, 14K g). (Note: some DBHDA crystallizes/ precipitates from the reaction mixture). The supernatant was analyzed by LC-MS, revealing revealing mixture of protein species: non-alkylated starting material (mass calculated = 26688; observed mass = 26689), uncyclized intermediate **2.23** (mass calculated = 26910; observed mass = 26911) and **2.25** (mass calculated = 26654; observed mass = 26655).



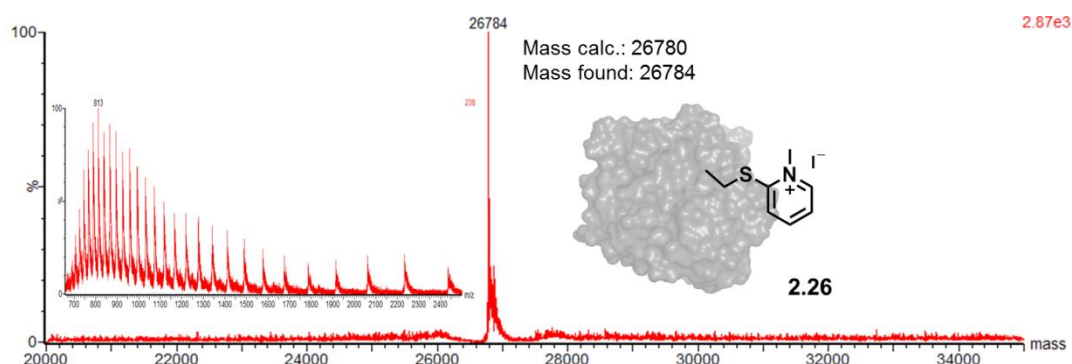
The reaction was then shaken at 37 °C for 8 h to until complete consumption of starting material. After this time, solids were removed by centrifugation and the supernatant analyzed by LC-MS revealing mixture of SBL-Dha217 **2.24** and protein adduct **2.25**. ESI-MS for the modification is shown below.



Dha synthesis on SBL-L217C using Mukaiyama's Reagent

All manipulations were carried out in a cold room at 4 °C according to a published procedure. A freshly prepared solution of 0.16 M 2-chloro-1-methyl-pyridinium iodide (Mukaiyama's Reagent) in 50 mM sodium phosphate buffer (pH 8.0) was used. To a solution of SBL-L217C (500 µL of 0.7 mg/mL, 13.0 nmol) in sodium phosphate buffer (50 mM, pH 8.0), 2 µL of 2-chloro-1-methyl-pyridinium iodide (328.0 nmol) was added and the resulting mixture was vortexed for 30 s. After 15 minutes of additional shaking, an aliquot was analyzed directly by LC-MS and complete conversion to SBL-L217C-mukaiyama's adduct was observed (mass calculated = 26782; observed mass = 26782). To

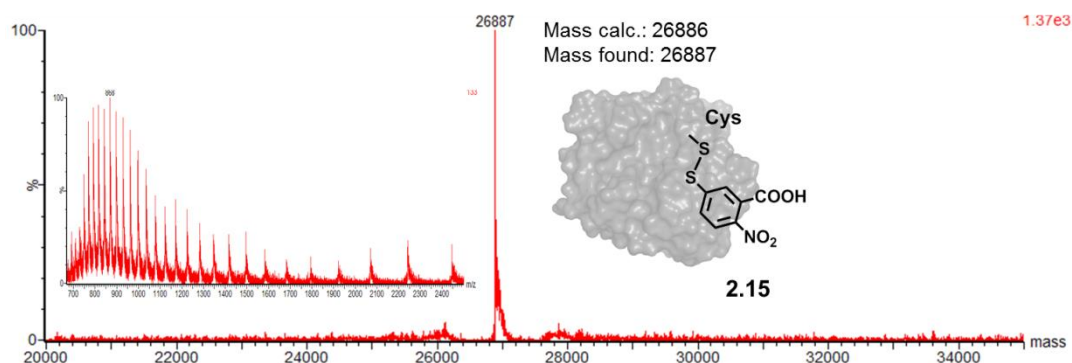
the reaction mixture, 1 M NaOH aq. (40 μ L) was added and the resulting mixture was vortexed for 30 s. After 1-3 h of shaking at 4 $^{\circ}$ C, the reaction was analyzed directly by LC-MS and complete conversion to SBL-Dha was observed. The protein sample was passed through a PD minitrapp to remove excess reagents and the sample was analyzed again by LC-MS. No conversion to SBL-Dha217 was observed even after 3 h of reaction at 4 $^{\circ}$ C. The ESI-MS of the modification is shown below.



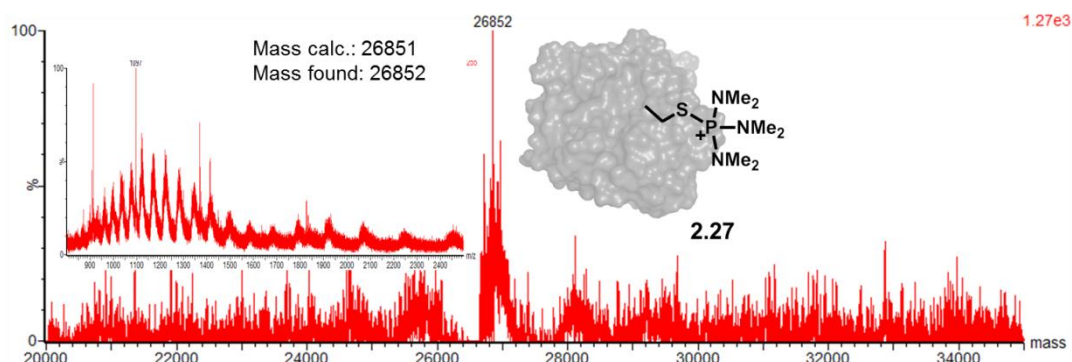
Dha synthesis on SBL-L217C by reduction-elimination strategy

All manipulations were carried out in a cold room at 4 $^{\circ}$ C according to a published procedure. A freshly prepared solution of 0.06 mM solution of hexamethylphosphorous triamide (HMPT) in DMF and 25 mM stock of 5,5'-Dithiobis(2-nitrobenzoic acid) (Ellman's reagent) in 50 mM sodium phosphate buffer (pH 8.0) was used. To a 100 μ L solution of 0.8 mg/mL SBL-L217C (3 nmol) in 50 mM sodium phosphate buffer pH 8 (pre-treated with 10 equiv. TCEP at room temperature for 5 min) was added 25 equiv. of Ellman's reagent (75 nmol). After shaking the reaction at room temperature for 30 min, an aliquot of the resulting protein solution was analysed by LC-MS. LC-MS analysis revealed full conversion to the Ellman's

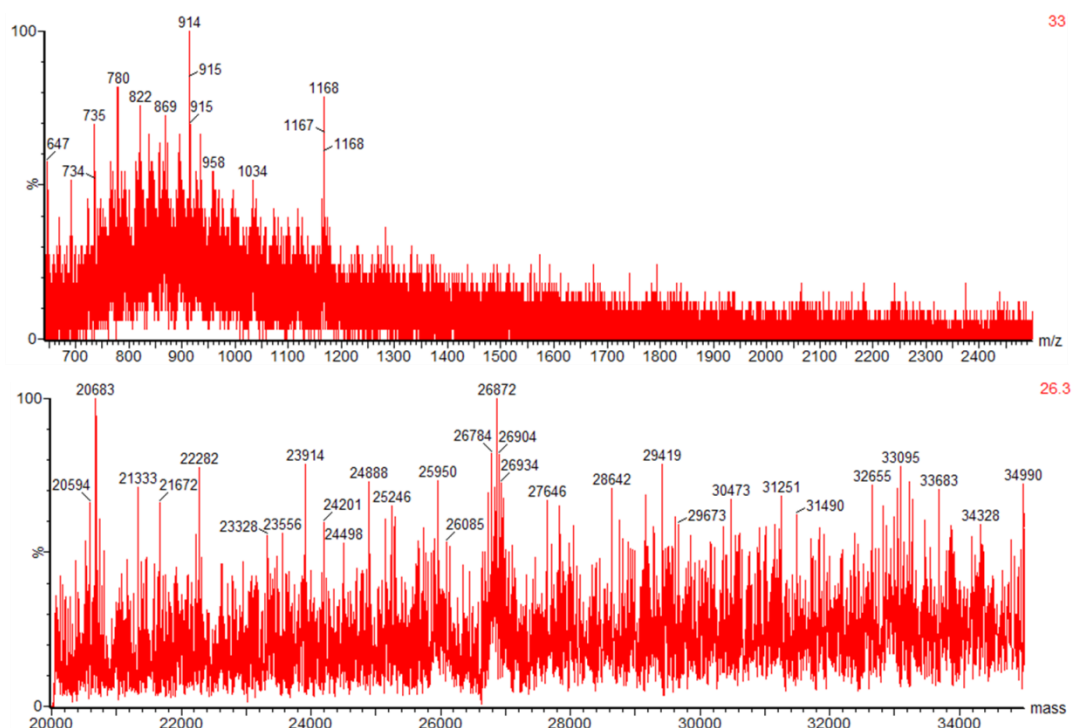
Adduct **2.15** (mass calculated = 26886; observed mass = 26887).



The resulting protein solution was desalted using PD SpinTrap G-25 column (GE Healthcare) and dialysed against 50 mM NaPi, pH 8. To a 150 μ L aliquot of the protein solution (0.3 mg/mL, 0.002 μ mol), HMPT (5 μ L of a 0.06 mM solution in DMF, 0.3 μ mol) was added. The reaction was shaken for 15 min at 4 $^{\circ}$ C and then analyzed by LC-MS. The phosphonium salt **2.27** was observed as the major product (mass calculated = 26851; observed mass = 26852).

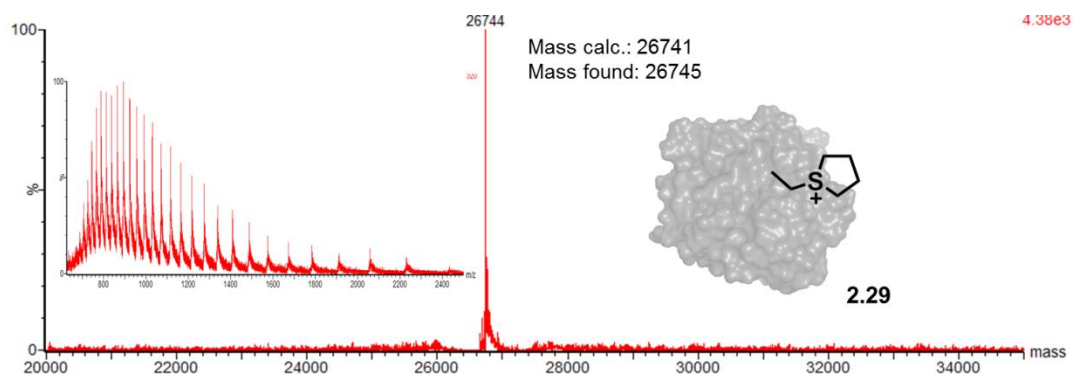


To this solution, 50 μ L of NaOH (1M solution) was added and then reaction shaken for another 20 min at 4 $^{\circ}$ C. LC-MS analysis of protein showed considerable degradation of protein, yet mass corresponding to the desired Dha product was not detected. ESI-MS of the protein at pH 11 is shown below.



Synthesis of Dha on SBL-L217C using 1,4-diiodobutane

To 200 μ L of SBL-L217C in 50 mM phosphate buffer pH 8 (1 mg/mL) was added a solution of diiodobutane (67 μ L of a 0.113M stock in DMF, 1000 equiv., total DMF content 25%). The emulsion was vigorously stirred in an oil bath at 37 $^{\circ}$ C for 6 h, after which an aliquot was withdrawn, the diiodobutane was separated by centrifugation and the supernatant was analyzed by ESI-MS.



The pH was adjusted to 10-11 using dilute NaOH and the reaction was shaken for 1 h at room temperature. The protein solution was then analysed by LC-MS and dialysed back into 50 mM NaPi, pH 8 and stored at -20 °C.

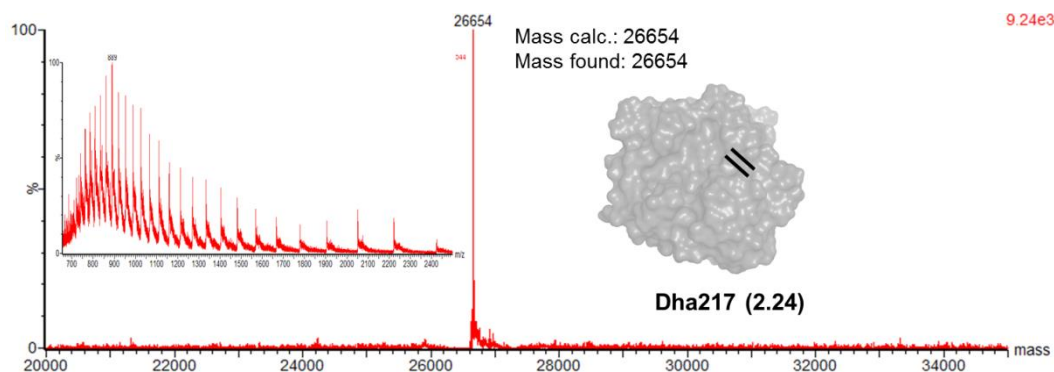


Table 2-2. Summary of conditions used for the synthesis of SBL-Dha217

Entry	Starting Protein	Reagent (equiv.)	Conditions	Mass found	remarks
1.	SBL-S156C	DBHDA , 1000 equiv.	1000 equiv., 30 min rt then 1 h at 37 °C	26682	>95% Dha156
2.	SBL-L217C	MR , 25 equiv. then base	1 h, 4 °C, pH 11-12	26779	No elimination
3.	SBL-L217C	2.28 , 1000 equiv.	37 °C, 2.5 h	26749	20% bisalkylated adduct
4.	SBL-L217C	2.28 , 1000 equiv.	37 °C, 6 h, 5% DMF	26747	43% bisalkylated adduct
5.	SBL-L217C	DBHDA , 1000 equiv.	37 °C, 6 h, 5% DMF	26763	66% - 2.25
6.	SBL-L217C-Ellman's disulphide	HMPT	pH 11, 4 °C, 5 h	-	degradation
7.	SBL-L217C	MR , 25 equiv. then base	3 h, 4 °C, pH 11-12	26784	No elimination
8.	SBL-L217C	DBHDA , 1000 equiv.	2 h, 37 °C, 5% DMF	26915	44%, noisy spectra
9.	SBL-L217C	DBHDA , 1000 equiv.	4 h 37 °C, 5% DMF	26912	>70%, very noisy spectra
10.	SBL-L217C	DBHDA , 660	2.5 h, 37 °C, no	26910	47% mono-

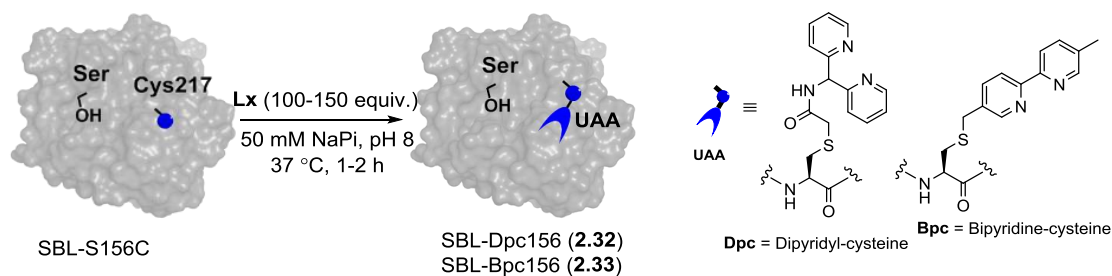
		equiv.	cosolvent		alkylation, 33% 2.25 29% Dha217, 71% 2.25
11.	SBL-L217C	DBHDA 1000 equiv.	8 h, 37 °C, no cosolvent	26655, 26759	
12.	SBL-L217C	DBHDA 660 equiv.	8 h, 37 °C, no cosolvent then 2 h 42 °C	26655, 26759	29% Dha217, 71% 2.25
13.	SBL-L217C	2.28 , 2000 equiv.	4 h, 37 °C, no cosolvent	26745	13% bisalkylated adduct
14.	SBL-L217C	2.28 , 2000 equiv.	6 h, 37 °C, 5% DMF	26744	100% bisalkylated adduct
15.	SBL-L217C	2.28 , 2000 equiv.	6 h, 37 °C, 5% DMF Then 2 h 42 °C	26744	87% bisalkylated adduct, 13% Dha adduct
16.	SBL-L217C	DBHDA , 1000 equiv.	6 h, 37 °C, no cosolvent	26756	87% 2.25 , 13% Dha217
17.	SBL-L217C	2.28 , 2000 equiv.	37 °C 10 h	26741	67% bisalkylation adduct
18.	2.29		pH adjusted to 11, rt 20 min		>95% Dha217
19.	SBL-L217C	2.28 , 2000 equiv.	16 h, 37 °C, 5% DMF	26745 (2.29), 26657 (Dha217)	77% 2.29 , 23% Dha217
20.	2.29		pH adjusted to 11, rt 2 h	26655	>95% Dha217
21.	2.25		pH adjusted to 11, rt 5 h	26759, (2.25), 26657 (Dha217)	87% 2.25 , 23% Dha217

Attempts for conjugate addition of thiol ligands **L3** and **L4** on SBL-Dha217

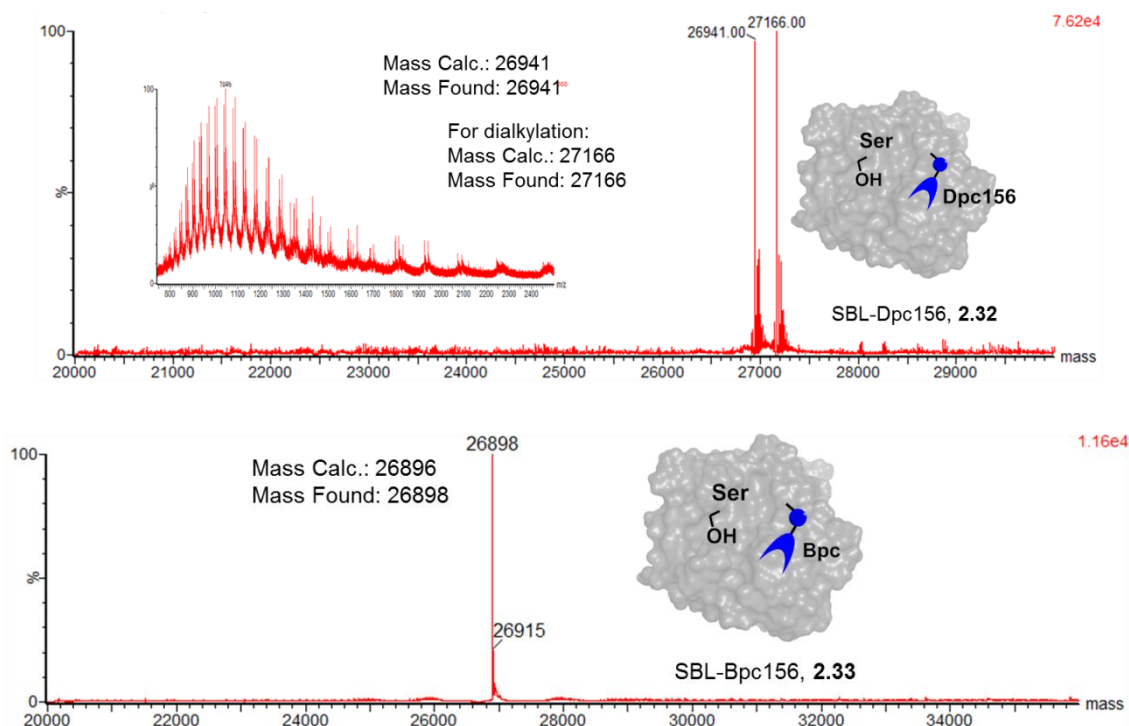
A solution of SBL-Dha217 (**2.24**) was prepared at a concentration of 0.4 mg/mL in 50 mM NaPi pH 8. To a 180 µL solution of SBL-Dha217 **2.24** (2.7 nmol) in 50 mM NaPi pH 8, was added thiol **L3** or **L4** (5.4 µmol). The protein reaction was shaken at 37 °C for 1 h, after which an aliquot was analysed by LC-MS, showing no conversion even up to 5 h. The protein was desalted using PD SpinTrap™ G-25 column (GE Healthcare Life Sciences) by eluting with 50 mM NaPi pH 6.

2.5.7 Direct alkylation of SBL-S156C with ligand Lx

Chemical modification of SBL-S156C was accomplished by direct alkylation using halogenated ligands **L1** and **L2**. The synthesis of SBL-Dpc156 and SBL-Bpc156 was done by Dr. Anca Pordea.

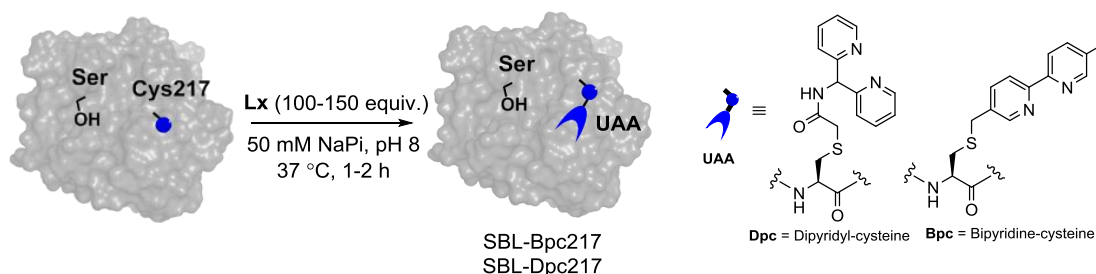


To the SBL-S156C (2.5 mL, 2 mg/mL=75 μM concentration established by BCA assay) in 50 mM sodium phosphate buffer at pH 8 was added in a test tube a solution of DTT (4.7 μL of a 1M solution, 4.7 μmol , 25 equiv.), followed by **Lx** (0.1 M or 1 M in DMF, 100 equivalents). The resulting solution was shaken at 37 $^{\circ}\text{C}$ until complete consumption of the starting material within 30 min (verified by ESI-MS), then passed through a PD-10 column to obtain SBL-Dpc156 and SBL-Bpc156. ESI-MS for the modification is shown below.

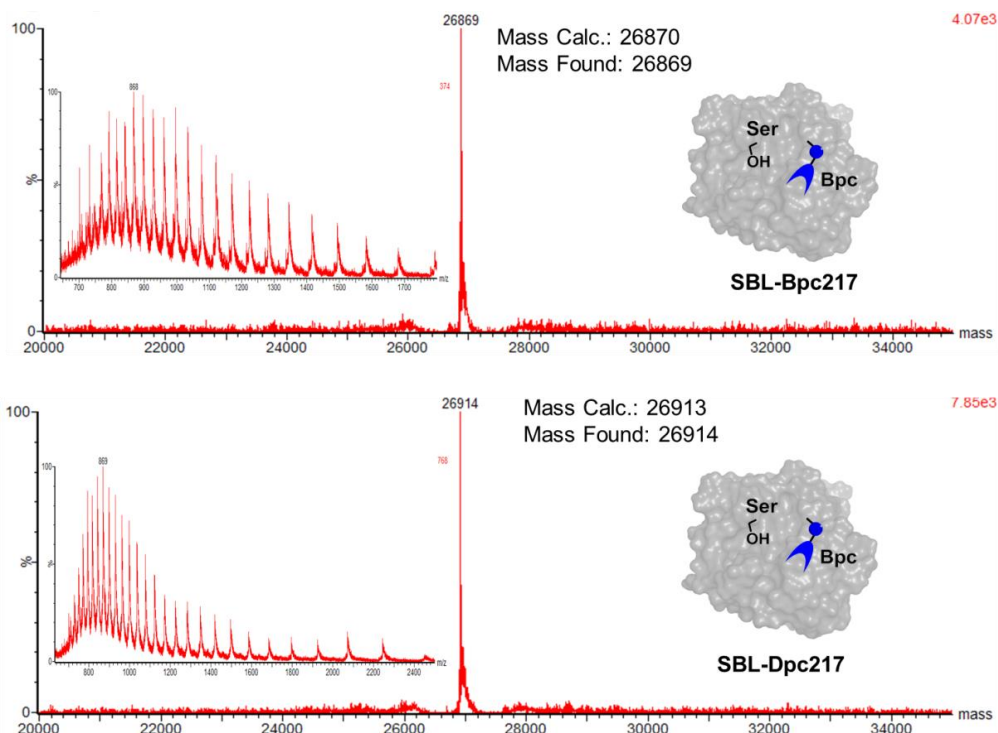


2.5.8 Direct alkylation of SBL-L217C with ligand Lx

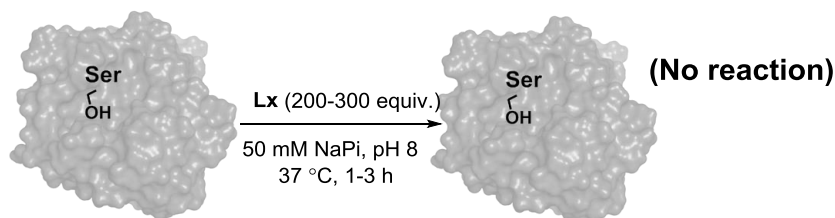
Chemical modification of SBL-L217C was accomplished by direct alkylation using halogenated ligands **L1** and **L2**.



5 mg of the lyophilised protein was dissolved in sodium phosphate buffer (50 mM, pH 8). The concentration was determined by BCA assay as described above. A 1 M stock solution of DTT was prepared in MilliQ water and a 1 M solution of **Lx** in DMF. Typically, to 1 mL of SBL-L217C (56 nmol) in 50 mM phosphate buffer, pH 8 at 4 °C, was added 1.4 μ L of DTT (1.4 μ mol) to reduce the possible disulfide bonds. The reaction was shaken for 15 min at room temperature. The protein solution was reacted with **Lx** (100 equiv. for **L1** and 150 equiv. for **L2**). The reaction mixture was incubated at 37 °C for upto 2 h (1 h for **L1** and 3 h for **L2**). The reaction was monitored at an interval of 1 h by LC-MS. A single protein species was detected showing a mass corresponding to the apo-enzyme species (For SBL-Dpc217 **2.30** modification: mass calculated 26913, mass found 26914; For SBL-Bpc217 **2.31** modification: mass calculated 26870, mass found 26869). PD-10 G-25 column were pre-equilibrated with 50 mM sodium phosphate buffer, pH 6 at 4 °C. The reaction mixture was de-salted using PD-10 spin trap G-25 column on completion of the reaction, followed by dialysis in sodium phosphate buffer (50 mM, pH 6) for 2 h. ESI-MS for the modifications is shown below.

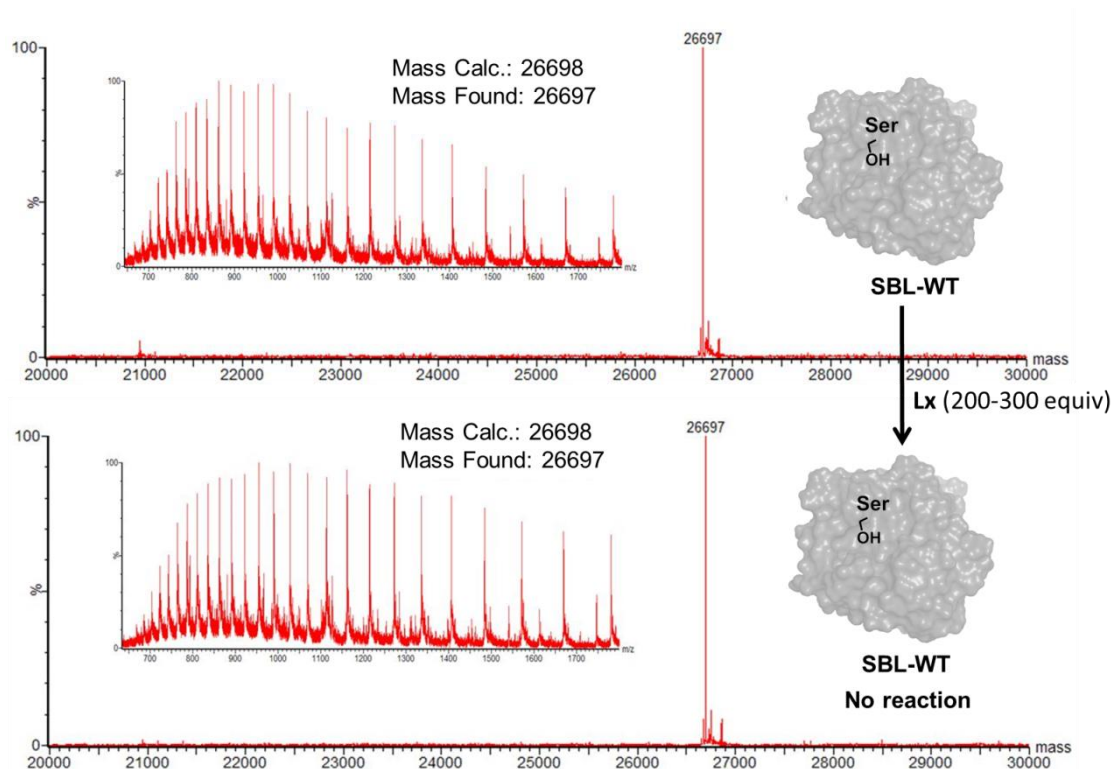


Control reaction: SBL WT treated with Lx. As a negative control, SBL-WT was treated with **Lx** under similar conditions as above.

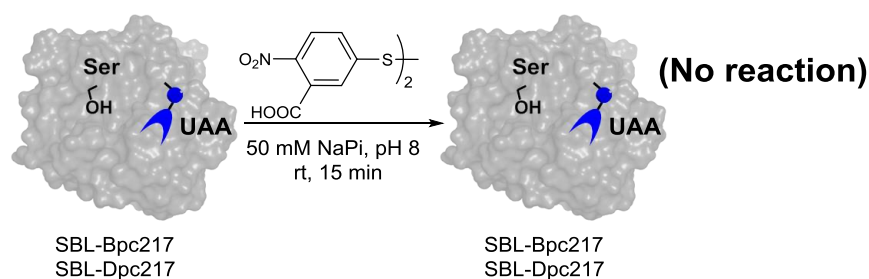


A 1 mL sample of SBL-WT was prepared at 1.5 mg/mL in sodium phosphate buffer (50 mM, pH 8). A 1 M stock solution of DTT was prepared in MilliQ water and 1 M solution of **Lx** in DMF. To 1 mL of SBL-WT (55 nmol) in 50 mM phosphate buffer, pH 8 at 4 °C, was added 1.4 µl of DTT (1.4 µmol). The reaction was shaken for 15 min at room temperature. The protein solution was reacted with **Lx** (200 equiv. for **L1** and 300 equiv. for **L2**). The reaction mixture was incubated at 37 °C for upto 3 h (1 h for **L1** and 3 h for **L2**). The reaction was monitored at an interval of 1 h by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT (mass calculated 26698, mass

found 26697) showing that no modification took place in the absence of the cysteine residue. The ESI-MS for the negative controls for **L1** and **L2** modifications after 3 h of reaction are shown below.

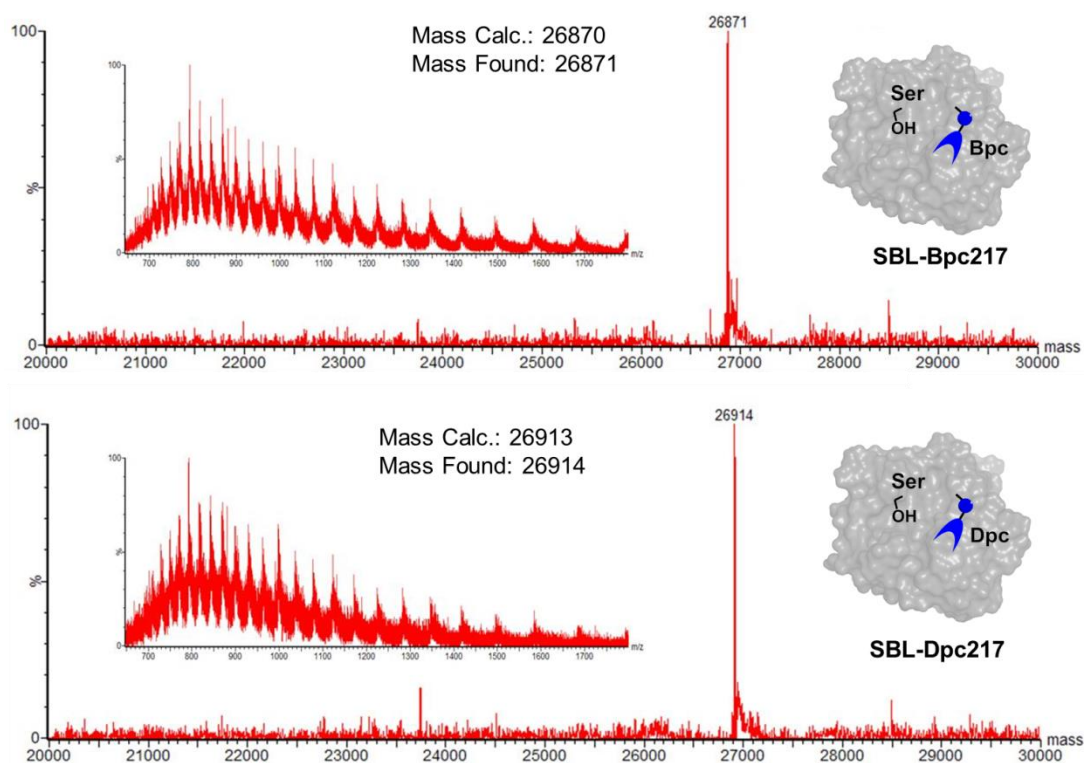


Ellman's test. Alkylation of protein with ligand **Lx** at the reactive Cys217 residue was confirmed by a negative Ellman's test.



A 0.5 M stock solution of Ellman's reagent was prepared in 50 mM sodium phosphate buffer, pH 8. To a 50 μ L aliquot of modified SBL-Bpc217 and SBL-Dpc217 protein samples (2.8 nmol) was added 0.22 μ L of the Ellman's solution (280 nmol). The reaction was shaken for 15 min at room temperature and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to apo-enzyme species. No

shift in mass was observed, showing complete alkylation of the Cys217. (For SBL-Dpc217: mass calculated 26913, mass found 26914; For SBL-Bpc217: mass calculated 26870, mass found 26871). Under identical conditions, SBL-L217C with free cysteine residue gave the Ellman disulfide (See Previous Experiment).



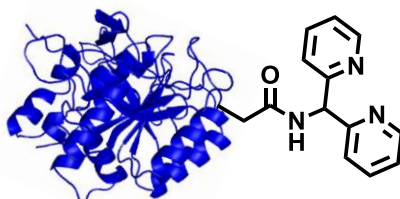
2.5.9 Peptide mapping and MS/MS analysis

Peptide Mapping: The peptide mapping of SBL-Dpc217 was accomplished by carrying out an in-solution protein digest using elastase (Promega) and analysing the peptide fragments by MASCOT⁵⁸ database search.

In-solution digestion procedure. This sample was prepared by Dr. Alba Díaz-Rodríguez. To a 10 μ L of SBL-Dpc217 sample (20 μ g, protein concentration 2 mg/mL) in sodium phosphate buffer (50 mM, pH 8) was added urea (5 μ L, 8 M stock solution). The reaction

mixture was treated with 1 µg of elastase. Total reaction volume was 40 µL. The reaction was incubated at 37 °C overnight.

The modification ‘**Cys modified with L1**’ was uploaded on Unimod server (www.unimod.org/) as shown below.



Scheme 2-26. Structure of modification as introduced in Unimod Server

Accession Number #	PSI-MS Name	Interim Name	Description	Monoisotopic Mass	Average Mass	Composition
#1277		Dipyridyl	Cys modified with dpy ligand	225.090212	225.2459	H(11) C(13) N(3) O

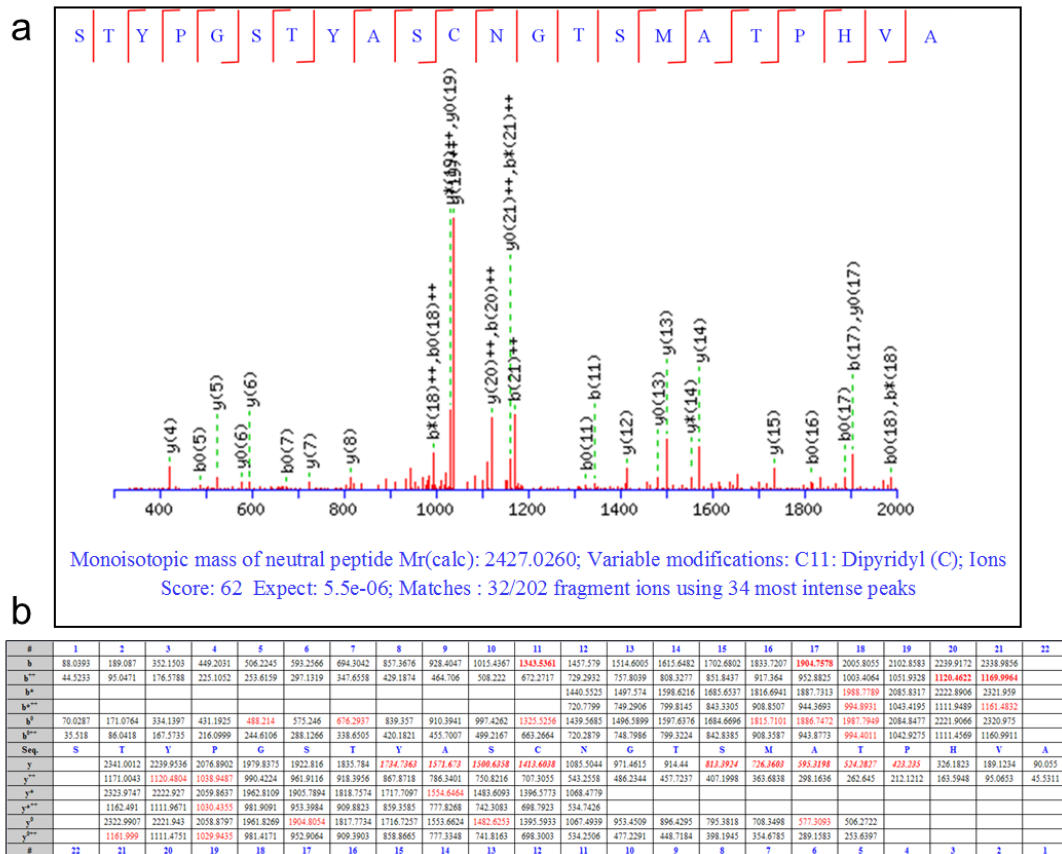
Peptide Sequence Identification. Peptide fingerprinting was done by using MASCOT. 100% sequence coverage was obtained for SBL-Dpc217. The annotated fragment spectra were subjected to database search using the Mascot search engine (version 2.2, Matrix Science, London, UK). The following parameters were used for the searches: mass tolerance (MS: ± 20 ppm, MS/MS: ± 0.5 Da), enzyme (trypsin or none), fixed modifications (none), variable modifications (See figure and table for modification) [Oxidation \(M\)](#), Dipyridyl (C), maximum precursor charge (± 3) and instrument (ESI-TRAP). Peptide match with a score above the confidence threshold ($p \pm 0.05$) was considered to be a significant hit.

Match to: SBL Score: 5043; Nominal mass (M_r): 26672; Calculated pI value: 9.14
 Variable modifications: Oxidation (M), Dipyridyl (C)
 No enzyme cleavage specificity; Sequence Coverage: 100%; Matched peptides shown in **Bold Red**

1 AQSVPWGISR VQAPAAHNRG LTGSGVKVAV LDTGISTHPD LNIRGGASFV
51 PGEPTQDGN GHGTHVAGTIAALNNSIGVL GVAPSAELYA VKVLGASGSG
101 SVSSIAQGLE WAGNNGMHVA NLSLGSPSPS ATLEQAVNSA TSGVLVVA
151 SGNAGAGSIS YPARYANAMA VGATDQNNNR ASFSQYGAGL DIVAPGVNVQ
201 STYPGSTYAS CNGTSMATPH VAGAAALVKQ KNPSWSNVQI RNHLKNTATS
251 LGSTNLYGSG LVNAEAATR

Scheme 2-27. Peptide mapping results obtained using MASCOT; 100% sequence coverage was obtained identifying the peptide carrying the **L1** modification

MS/ MS analysis of SBL-Dpc217. The results for the peptide finger printing are as follows. The dipyridyl modification corresponding to ligand **L1** could be found on Cys217 as shown below.

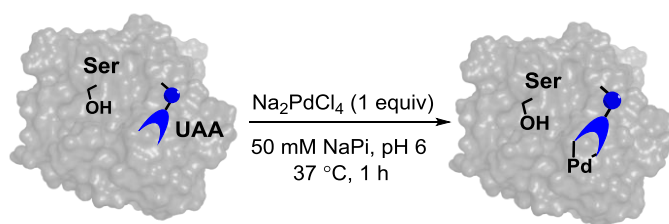


Scheme 2-28. (a) MS/MS Fragmentation of **STYPGSTYASCNGTSMATPHVA**, cysteine is shown to be modified by the dipyridyl ligand **L1**; peptide fragmentation notation was built (according to the Roepstroff-Fohlman scheme⁵⁹) (b) Table showing 32/202 fragment ions using 34 most intense peaks, the found fragment ions are highlighted in red

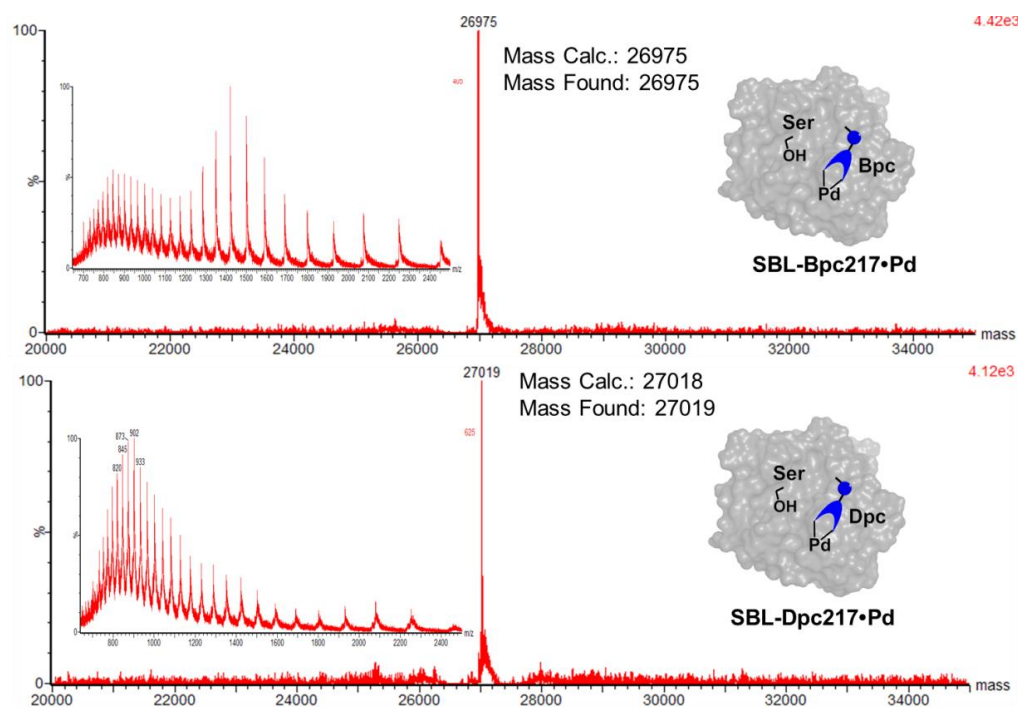
2.5.10 Metal complexation

2.5.10.1 Palladium complexation reaction

Preparation of stock solution of Na_2PdCl_4 . 0.5 M stock solution was made by suspending NaCl (58.5 mg, 1 mmol) and PdCl_2 (89 mg, 0.5 mmol) in 1 mL distilled H_2O . The mixture was stirred in a sealed flask at 45 °C until PdCl_2 dissolved completely. This stock solution was diluted to 10 mM using MilliQ water for enzyme reactions.



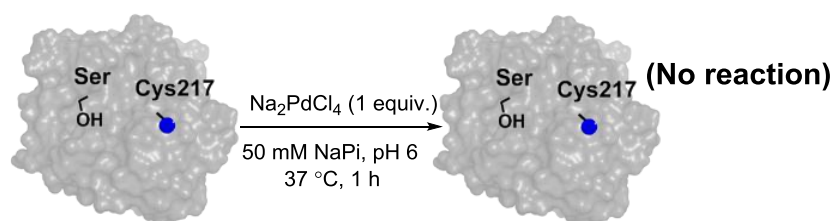
A 100 μL sample of SBL-Dpc217 (or SBL-Bpc217) was prepared at 1.0 mg/mL in sodium phosphate buffer (50 mM, pH 6) and the enzyme was stored on ice until needed. A 10 mM stock solution of Na_2PdCl_4 was prepared in MilliQ water as described above. To 100 μL sample of SBL-Dpc217 or SBL-Bpc217 (3.7 nmol) in 50 mM phosphate buffer, pH 6 at 4 °C, was added 0.37 μL of Na_2PdCl_4 (3.7 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. The reaction was monitored by LC-MS. A single protein species was detected with a mass corresponding to the holo-form of the artificial metalloenzyme (For SBL-Dpc217•Pd **2.32**: mass calculated 27019, mass found 27022; For SBL-Bpc217•Pd **2.33**: mass calculated 26976, mass found 26975) showing uptake of palladium. The ESI-MS on completion of reaction after 1 h are shown below.



2.5.10.2 Control reaction: SBL-L217C and Na₂PdCl₄ only (No UAA control)

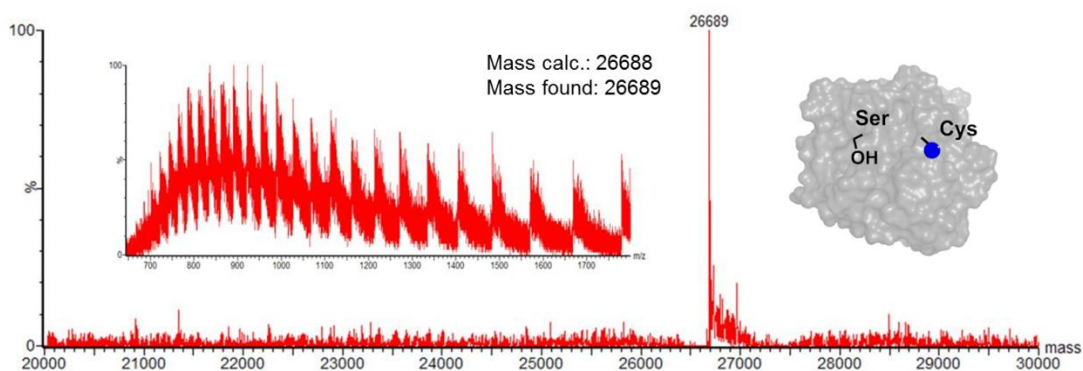
Palladium loading was attempted with the SBL-L217C with upto 10 equiv. of Na₂PdCl₄ under the same conditions as described above. This control was considered necessary to confirm if the cysteine residue was chelating palladium in the absence of a ligand.

Using 1 equiv. of Na₂PdCl₄:

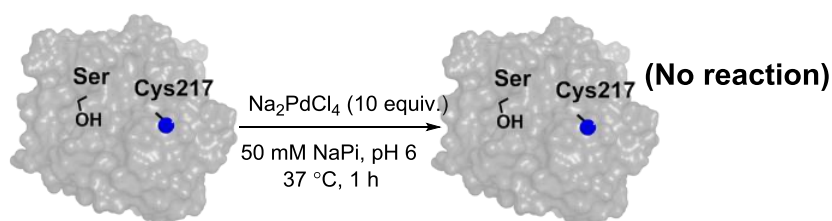


A 100 μ L sample of SBL-L217C was prepared at 1.0 mg/mL in sodium phosphate buffer (50 mM, pH 8). A 0.5 M stock solution of DTT was prepared in MilliQ water. To 100 μ L of SBL-L217C (4 nmol) in 50 mM phosphate buffer, pH 8 at 4 °C, was added 0.2 μ L of DTT (100 nmol) to reduce the possible disulfide bonds. The excess DTT was removed and the buffer was exchanged by passing the protein through a Sephadex G-25 medium

(Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was stored on ice until needed and concentration was measured using BCA Assay. A 10 mM stock solution of Na_2PdCl_4 was prepared in MilliQ water. To 100 μL sample of SBL-L217C (3.7 nmol) in 50 mM phosphate buffer, pH 6 at 4 $^\circ\text{C}$, was added 0.37 μL of Na_2PdCl_4 (3.7 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 $^\circ\text{C}$ for 1 h. A single protein species was detected with a mass corresponding to the SBL-L217C (mass calculated 26688, mass found 26689) showing no uptake of palladium. The ESI-MS are shown below. Under the same reaction conditions, the chemically modified SBL-Dpc217 (or SBL-Bpc217) is converted to corresponding metalloenzyme species (see previous experiment).

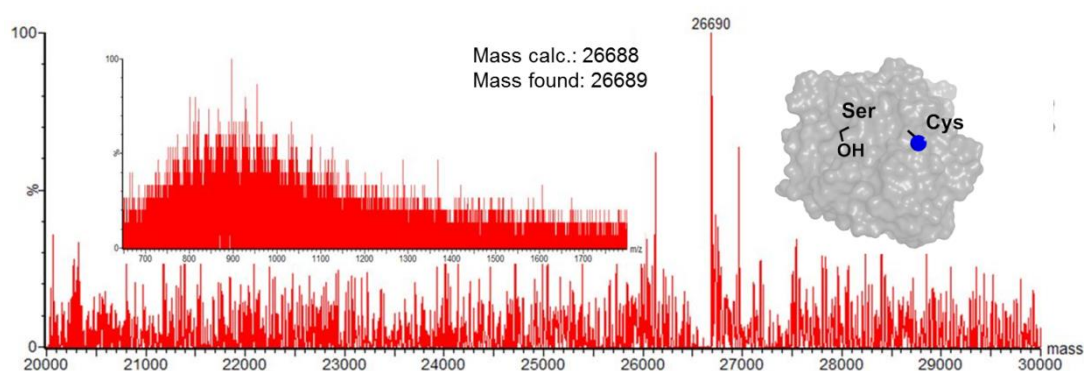


Using 10 equiv. of Na_2PdCl_4 :



A 100 μL sample of SBL-L217C was prepared at 1.0 mg/mL in sodium phosphate buffer (50 mM, pH 8). A 0.5 M stock solution of DTT was prepared in MilliQ water. To 100 μL of SBL-L217C (4 nmol) in 50 mM phosphate buffer, pH 8 at 4 $^\circ\text{C}$, was added 0.2 μL of DTT (100 nmol) to reduce the possible disulfide bonds. The excess DTT was removed and

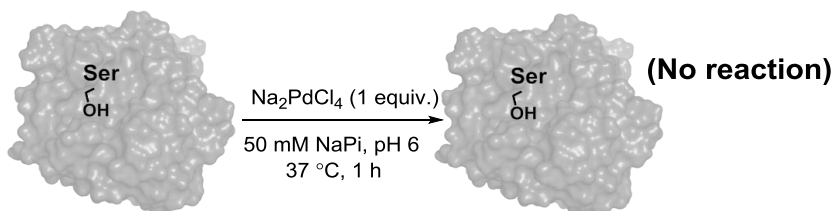
the buffer was exchanged by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was stored on ice until needed and concentration was measured using BCA Assay. A 10 mM stock solution of Na_2PdCl_4 was prepared in MilliQ water. To 100 μL sample of SBL-L217C (3.7 nmol) in 50 mM phosphate buffer, pH 6 at 4 $^\circ\text{C}$, was added 3.7 μL of Na_2PdCl_4 (37 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 $^\circ\text{C}$ for 1 h. A single protein species was detected with a mass corresponding to the SBL-L217C (mass calculated 26688, mass found 26690) showing no uptake of palladium. The ESI-MS are shown below.



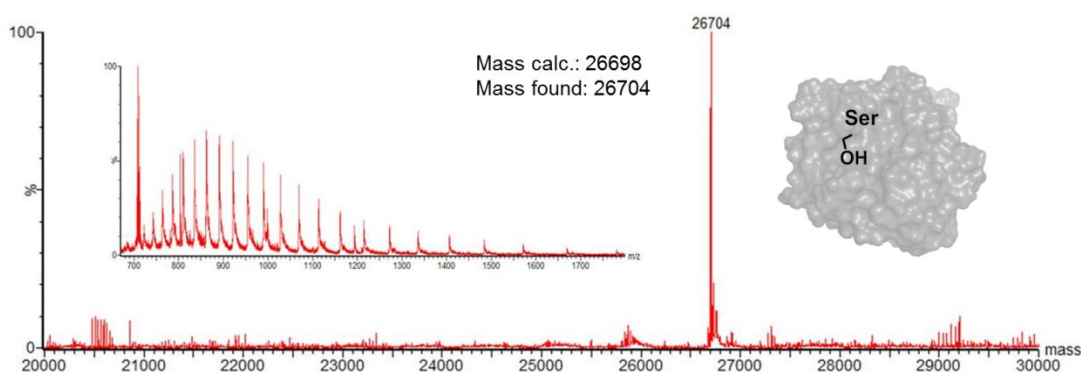
Note regarding the palladium loading on cysteine residue of SBL-L217C. These results should be considered keeping in mind that the reactive cysteine residue in the variant is susceptible to oxidation and is difficult to be managed under the reaction conditions in the absence of a reducing agent like DTT (which was desalted before addition of palladium to avoid its coordination with DTT). Also addition of higher palladium may give non-specific coordination by the free cysteine residue or other protein residues on the surface giving several isoforms of the enzyme or precipitation of protein, leading to poor intensity in ESI-MS.⁶⁰

2.5.10.3 Control reaction: SBL-WT and Na₂PdCl₄

Palladium loading was attempted with the SBL-WT with upto 5 equiv. of Na₂PdCl₄ under the same conditions as described above. This control was performed to confirm no other surface residue was chelating palladium.



A 100 μ L sample of SBL-WT was prepared at 1.0 mg/mL in sodium phosphate buffer (pH 6, 50 mM). To 100 μ L sample of SBL-WT (3.7 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.37 μ L of Na₂PdCl₄ (3.7 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. A single protein species was detected with a mass corresponding to the SBL-WT (mass calculated 26698, mass found 26704) showing no unspecific palladium coordination took place in the absence of the alkylated cysteine residue for upto 5 equiv. of palladium. The ESI-MS for the negative control is shown below.



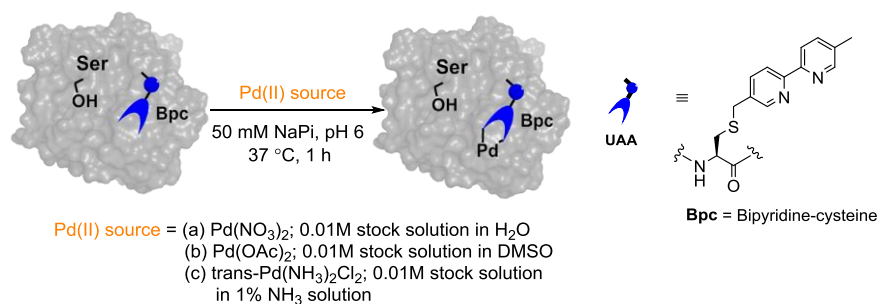
2.5.10.4 Modelling studies

Modelling experiments were done by Dr. Mark Wormald. A preliminary modelling study was performed using Insight II and Discover software (Accelrys) on a Silicon Graphics

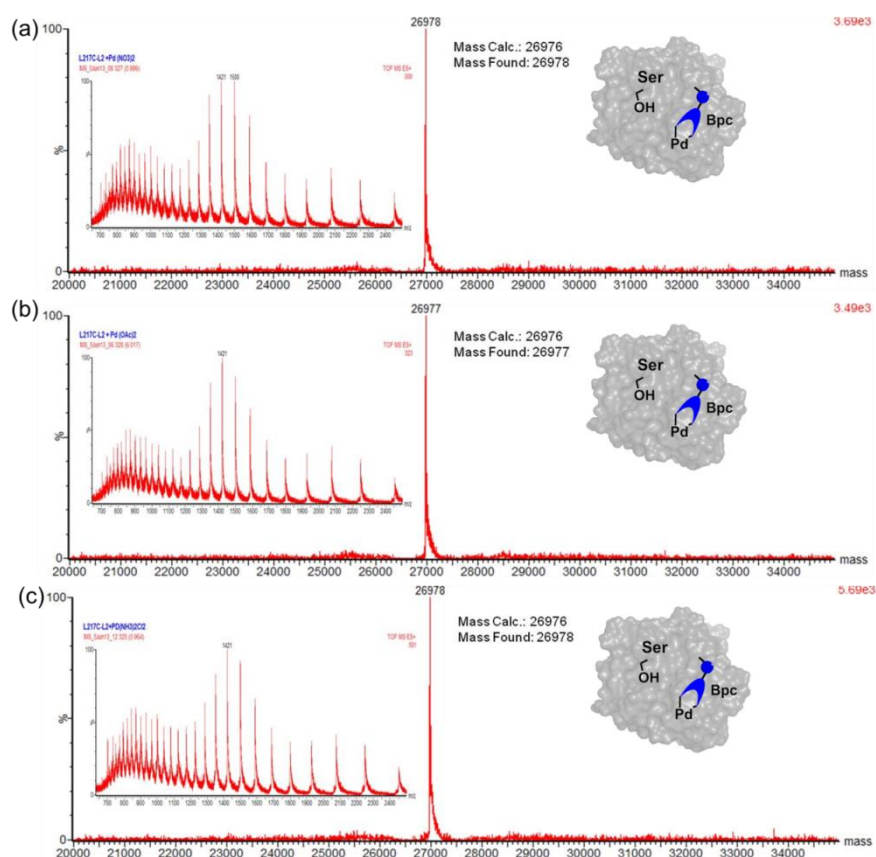
Indigo 2 workstation. The SBL-WT crystal structure (PDB ID: IGCI) was obtained from the Protein Data Bank.²⁵ Residue 217 was converted to a cysteine. The Dpc·Pd complex was installed at position 217 by adding corresponding ligand **L1** to Cys217 and phenylboronic acid at Ser221. The coordination geometry of the palladium was fixed as square planar. The resulting model was energy minimised using the CVFF force field in Discover, keeping the peptide backbone completely rigid. Only relaxation of the ligands and the amino acid side-chains in the immediate vicinity of the ligands was allowed.

2.5.10.5 Palladium complexation using different precursors

A 50 μ L sample of SBL-Bpc217 was prepared at 0.55 mg/mL in sodium acetate buffer (50 mM, pH 5.9) and the enzyme was stored on ice until needed. A 10 mM stock solution of different Pd(II) sources was prepared in respective solvent systems as described in the scheme below.

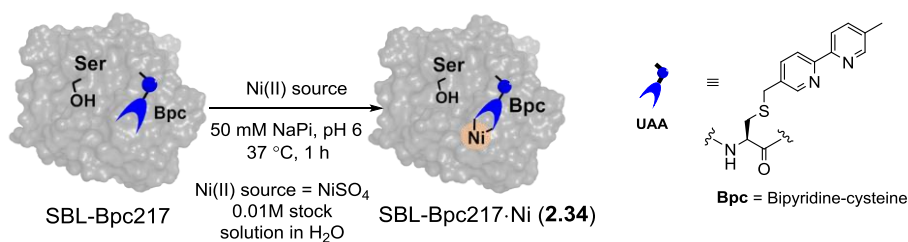


To 50 μ L sample of SBL-Bpc217 (1.1 nmol) in 50 mM sodium acetate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 0.2 μ L of Pd(II) source (1.9 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 $^{\circ}\text{C}$ for 1 h. The reaction was monitored LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217·Pd (mass calculated 26976, mass found 26978) showing uptake of palladium. The ESI-MS on completion of reaction after 1 h are shown below.



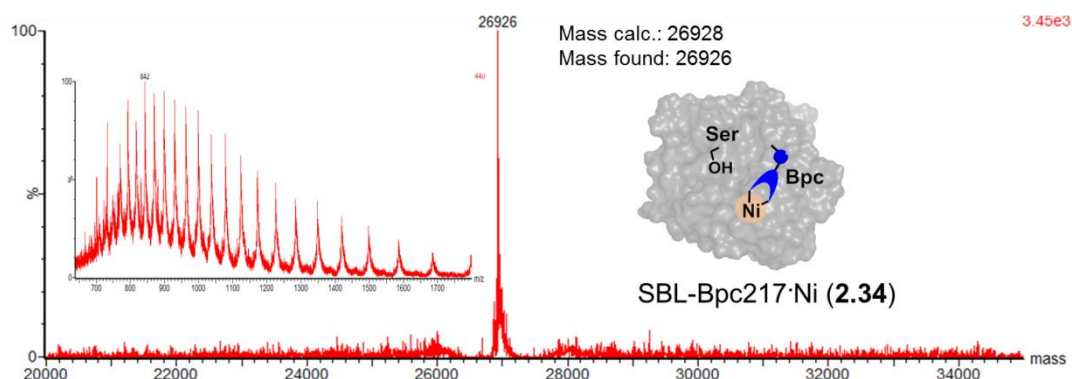
2.5.11 Loading of different metal ions

A 50 μL sample of SBL-Bpc217 was prepared at 0.55 mg/mL in sodium phosphate buffer (50 mM, pH 5.9) and the enzyme was stored on ice until needed. A 10 mM stock solution of different M(II) sources was prepared in MilliQ water.

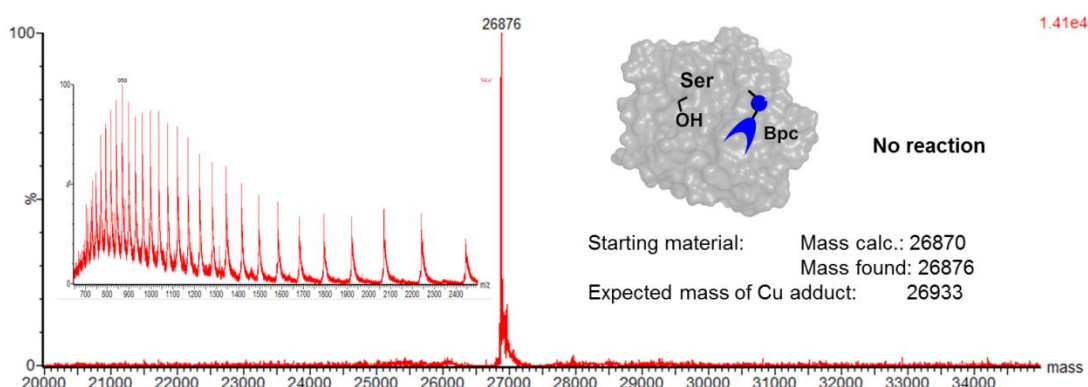


To 50 μL sample of SBL-Bpc217 (1.1 nmol) in 50 mM sodium acetate buffer, pH 6 at 4 °C, was added 0.2 μL of Ni(II) source (1.9 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. The reaction was monitored LC-MS. A single protein species was detected with a mass corresponding to the SBL-

Bpc217•Ni (mass calculated 26976, mass found 26978) showing uptake of palladium. The ESI-MS on completion of reaction after 1 h are shown below.



Metal complexation with other M(II) sources (such as $\text{Cu}(\text{OAc})_2$) was carried out using standard procedure described above. No complexation of copper or any other metal was detectable by LC-MS analysis. ESI-MS of the reaction of $\text{Cu}(\text{OAc})_2$ with SBL-Bpc217 is shown below.



Characterization of Pd-enzymes

2.5.12 ICP-MS sample preparation

All protein samples (except SBL-Dpc217•Pd; see below) were prepared by following this standard procedure. ICP-MS measurements were taken on Perkin Elmer Elan 6000DRC ICP-MS from m/z 105 at Department of Earth Sciences at University of Oxford, unless otherwise stated. All blanks, standards and samples were also spiked with 1 ng/g Rh, so

that any general instrument drift could be normalised. Dilutions were made using a 2% HNO₃ solution, prepared using in-house distilled nitric acid and 18.2 Mohm de-ionised water. Typically, the identity of each protein sample was confirmed by LC-MS, prior to preparation of ICP-MS samples for trace metal analysis. To a 100 µL protein sample was added 200 µL concentrated HNO₃. Each sample was diluted 50 times to bring the Pd levels to an estimated value of 40 ppb before analysis.

SBL-Dpc217•Pd: This sample was prepared by Dr. Alba Díaz-Rodríguez. Trace metal analysis to confirm the palladium loading on SBL-Dpc217•Pd was accomplished by ICP-MS analysis performed at the Butterworth Laboratories. To 960 µL sample of SBL-Dpc217 (79 nmol) in 50 mM phosphate buffer, pH 6 at 4 °C, was added 15.8 µl of Na₂PdCl₄ (79 nmol). The reaction solution was incubated at 37 °C for 1 h. The reaction was monitored LC-MS showing uptake of palladium by chemically modified SBL variant as described above. Sample was split into two parts- **A** and **B**, for its analysis before and after dialysis, to check for any unspecific palladium coordination. Part **B** was dialysed against 50 mM phosphate buffer (2 x 1 h). ICP-MS values obtained are given below.

Part **A**: 8.2 µg Pd/ mL for L217CL1Pd before dialysis

Part **B**: 7.3 µg Pd/ mL for L217CL1Pd after dialysis

2.5.13 Circular Dichroism

CD measurements were made using a Chirascan spectropolarimeter fitted with a Peltier temperature controller. SBL-WT and modified SBL-Bpc217 and SBL-Dpc217 were each made at 0.5 mg/mL solutions each in 50 mM sodium phosphate, pH 6. Prior to recording CD spectra, the identity of each protein sample was confirmed by LC-MS. CD measurements were taken in a 1 mm quartz cuvette at room temperature using a scan rate

of 50 nm/ min, 1 nm interval, 1 nm bandwidth and a response time of 0.5 s. After baseline correction, ellipticities in deg were converted to molar ellipticities ($\text{deg cm}^2 \text{dmol-res}^{-1}$) by normalizing for the concentration of peptide bonds and path length.

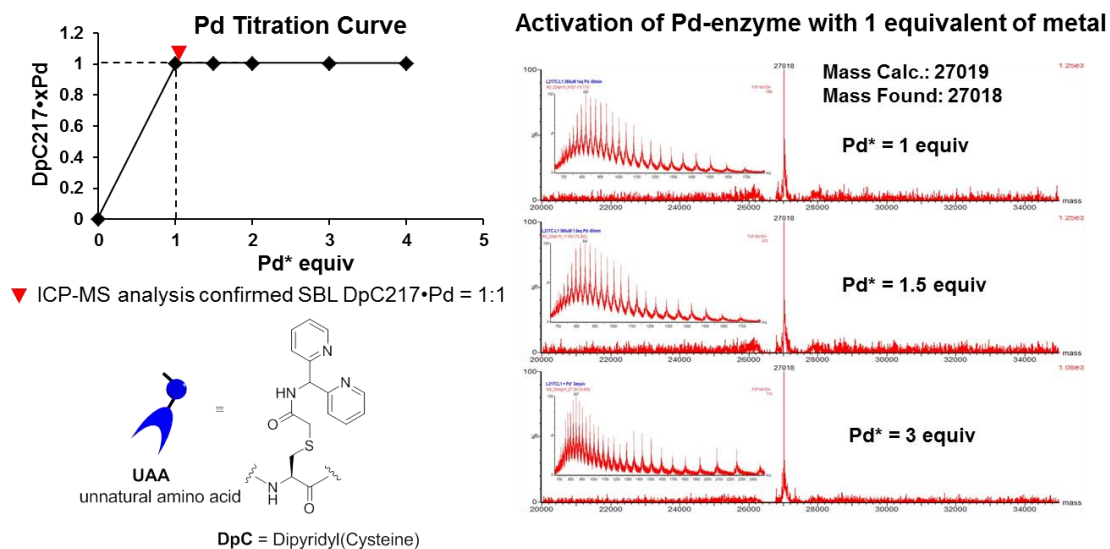
2.5.14 Palladium tolerance and titration studies

A 1 mL sample of SBL-Dpc217 (or SBL-Bpc217) was prepared at 0.8 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.01 M stock solution of Na_2PdCl_4 (Pd^{2+} source) was prepared in MilliQ water. To 100 μL aliquots of the above protein solution, was added 1-20 equiv. Pd^{2+} source according to the following table.

Table 2-2. BCA analysis to confirm loss in soluble protein content with SBL-Dpc217•Pd and SBL-Bpc217•Pd on incubation with increasing equiv. of Pd^{2+} precursor

Entry	Starting Concentration [Protein] μM	Starting Concentration [Protein] mg/mL	Pd^{2+} source equiv.	Protein concentration of SBL-Dpc217•Pd post incubation with Pd^{2+} source [Protein] mg/mL	Protein concentration of SBL-Bpc217•Pd post incubation with Pd^{2+} source [Protein] mg/mL
1	28	0.8	1	0.78	0.80
2	28	0.8	2	0.73	0.80
3	28	0.8	3	0.71	0.78
4	28	0.8	4	0.66	0.71
5	28	0.8	5	0.56	0.61
6	28	0.8	10	0.26	0.29
7	28	0.8	15	0.07	0.08
8	28	0.8	20	0.01	0

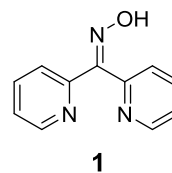
The reactions were incubated at 37 °C for 1 h, after which protein samples from entries 1-5 were analysed by LC-MS for titration studies. The protein samples were centrifuged at 13,000g for 2 min. After centrifugation, the protein content in each sample was analysed by BCA assay using BSA standards according to the standard procedure (See Section 2.4.5). 10 μL aliquot of each was analysed by non-reducing SDS-PAGE analysis after incubating samples with 1 μL 1 M HCl for 30 minutes at 4 °C.

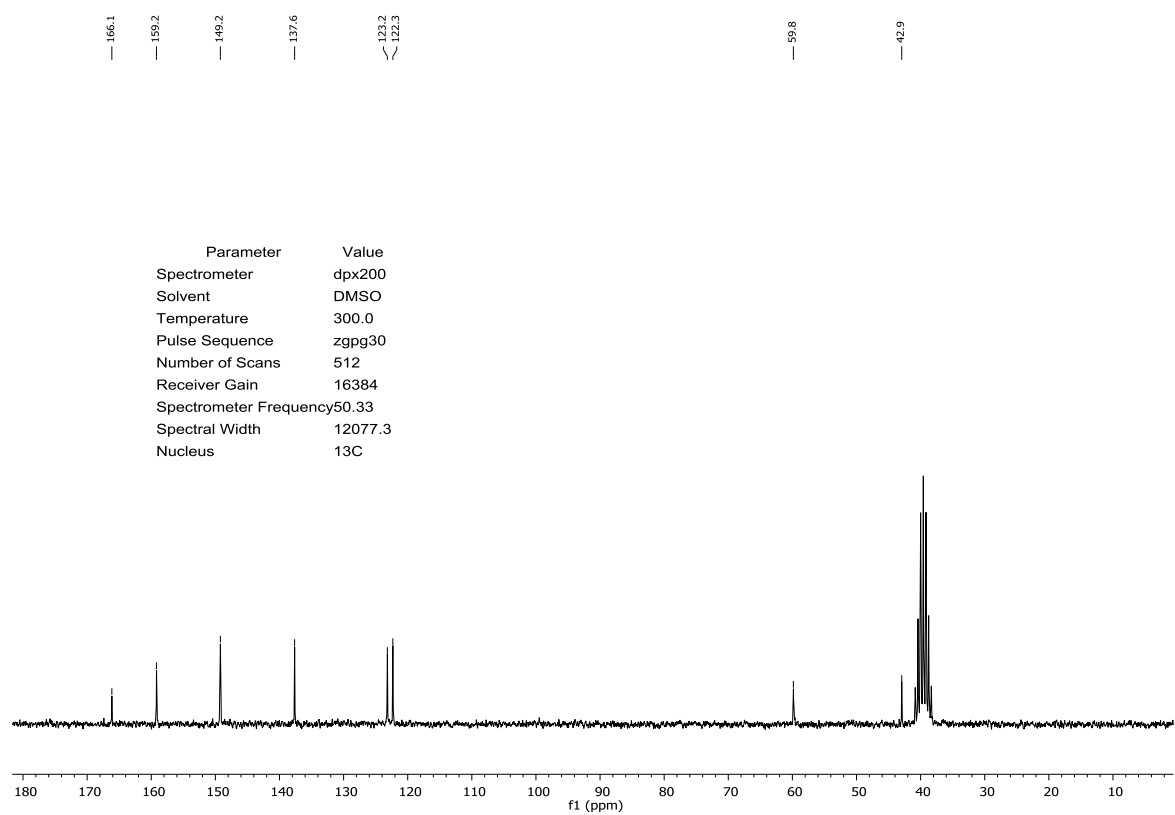
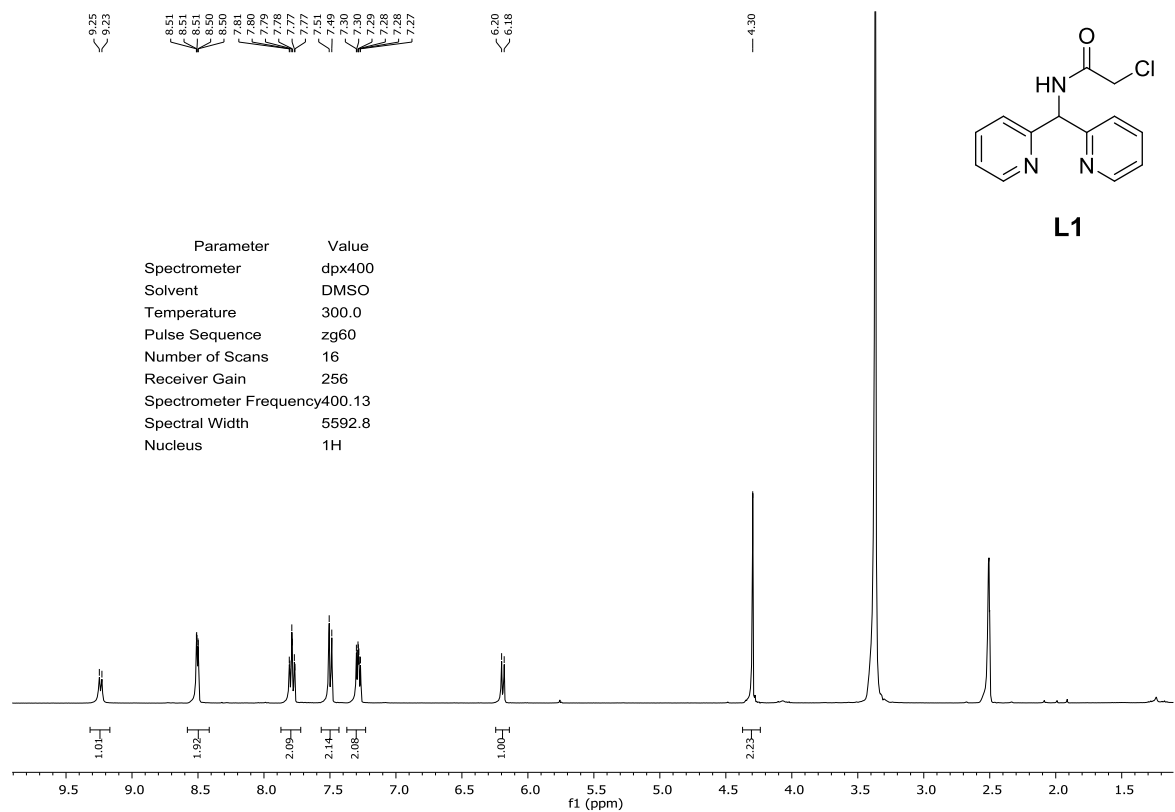


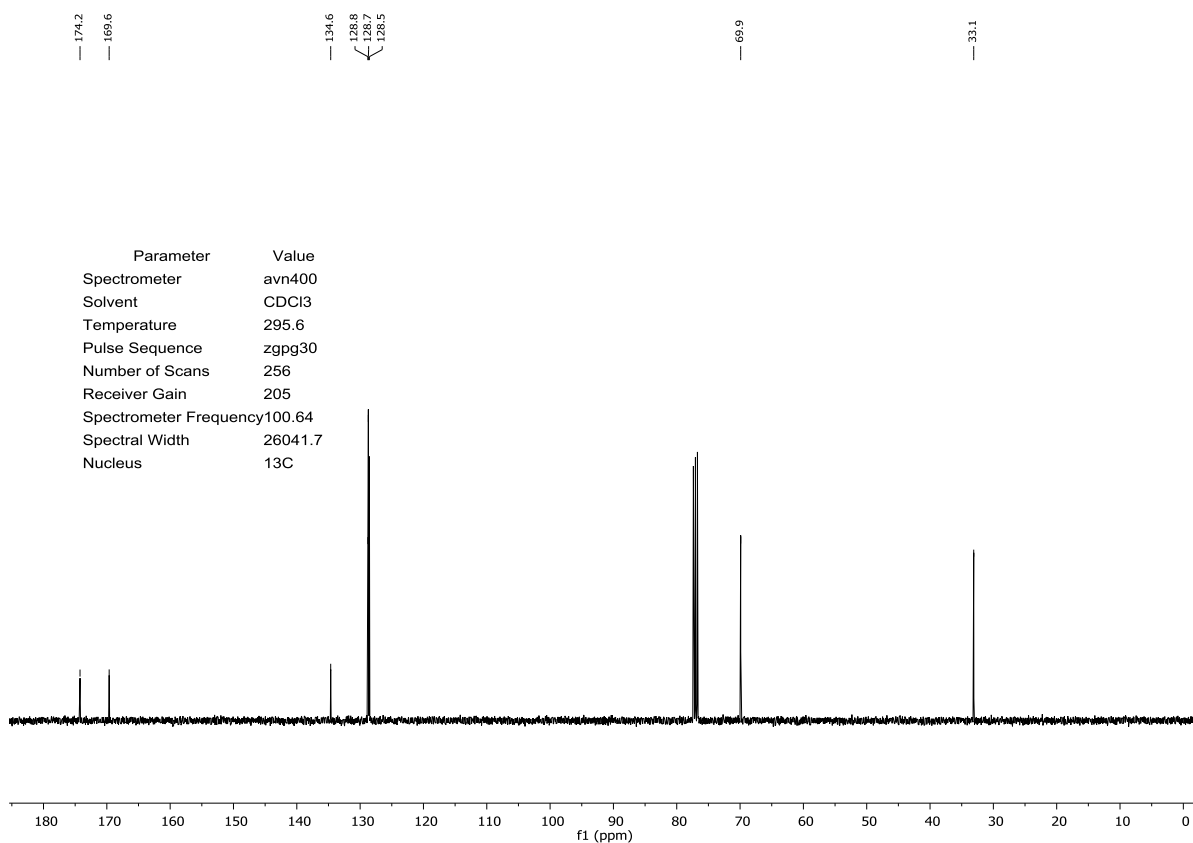
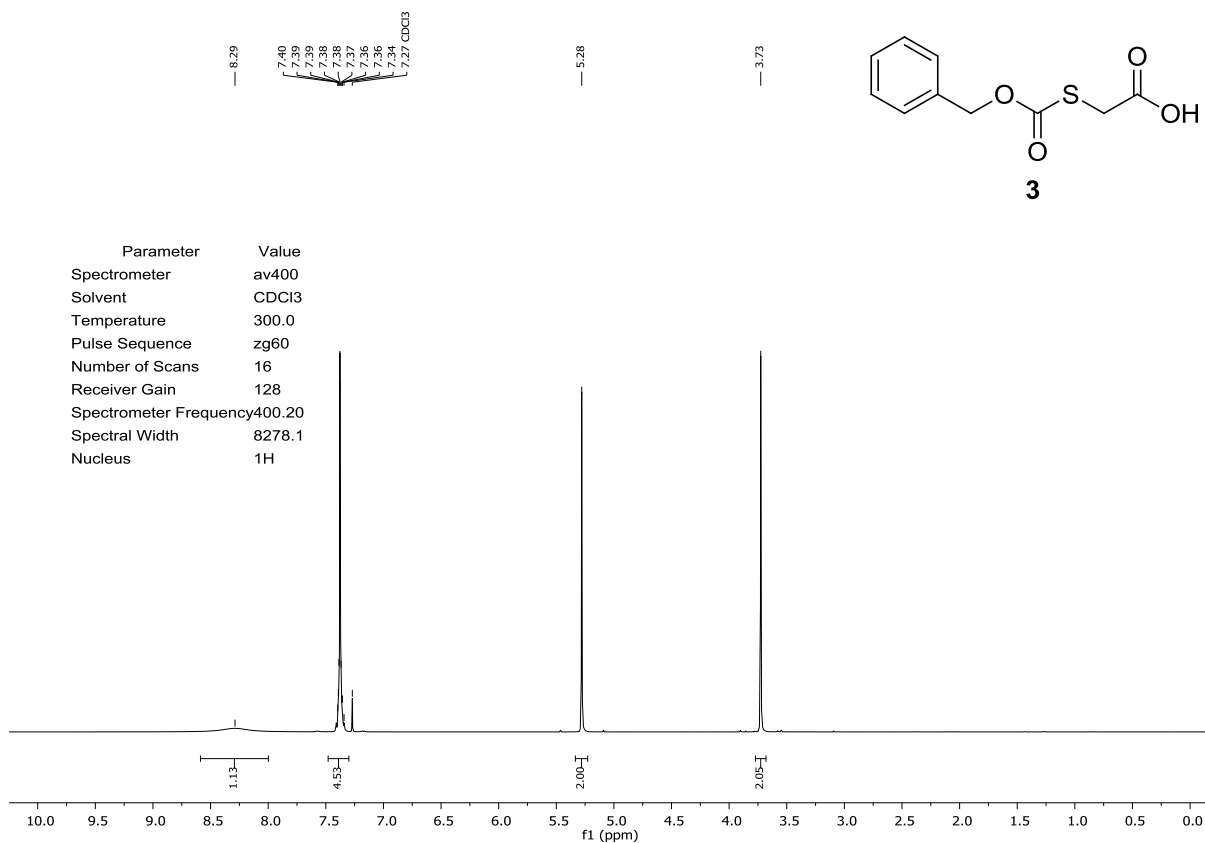
Scheme 2-29. Titration studies with SBL-Dpc217; clean Pd titration on modified SBL-Dpc217 with 1-5 equiv. of Pd precursor giving SBL-Dpc217•Pd = 1:1, confirmed by LC-MS and ICP-MS analyses

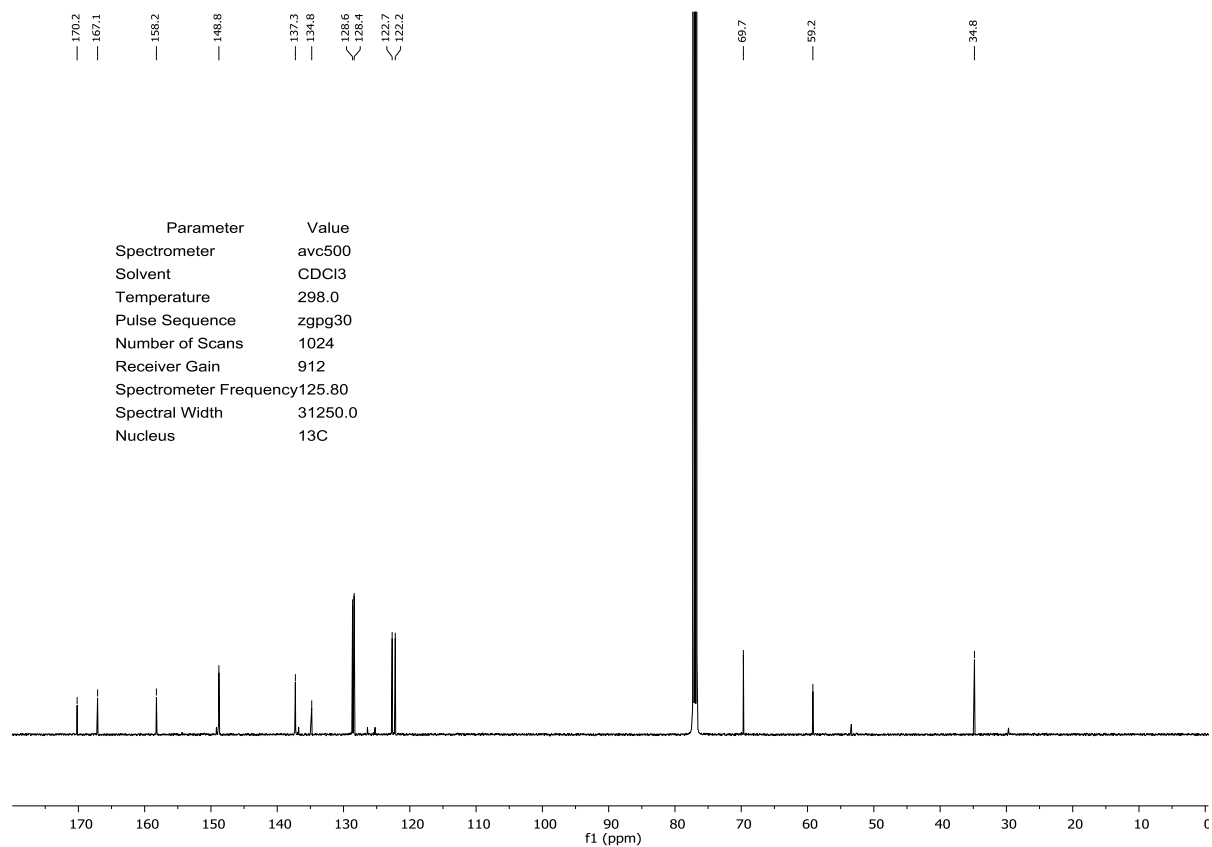
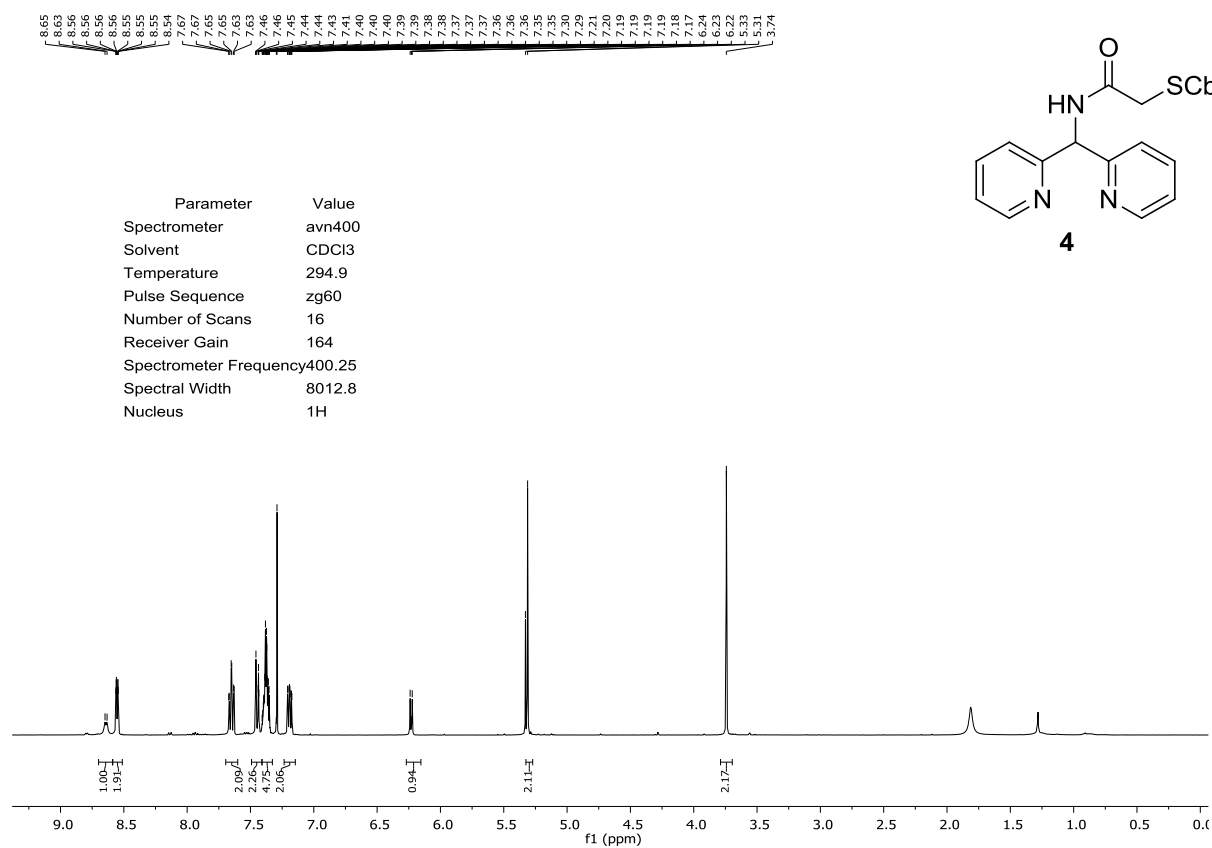
For titration experiments with SBL-Dpc156, the protein was borrowed from Dr. Alba Díaz-Rodríguez. The experiment was done with Dr. Maria Espling using the same protocol optimized previously by the author.

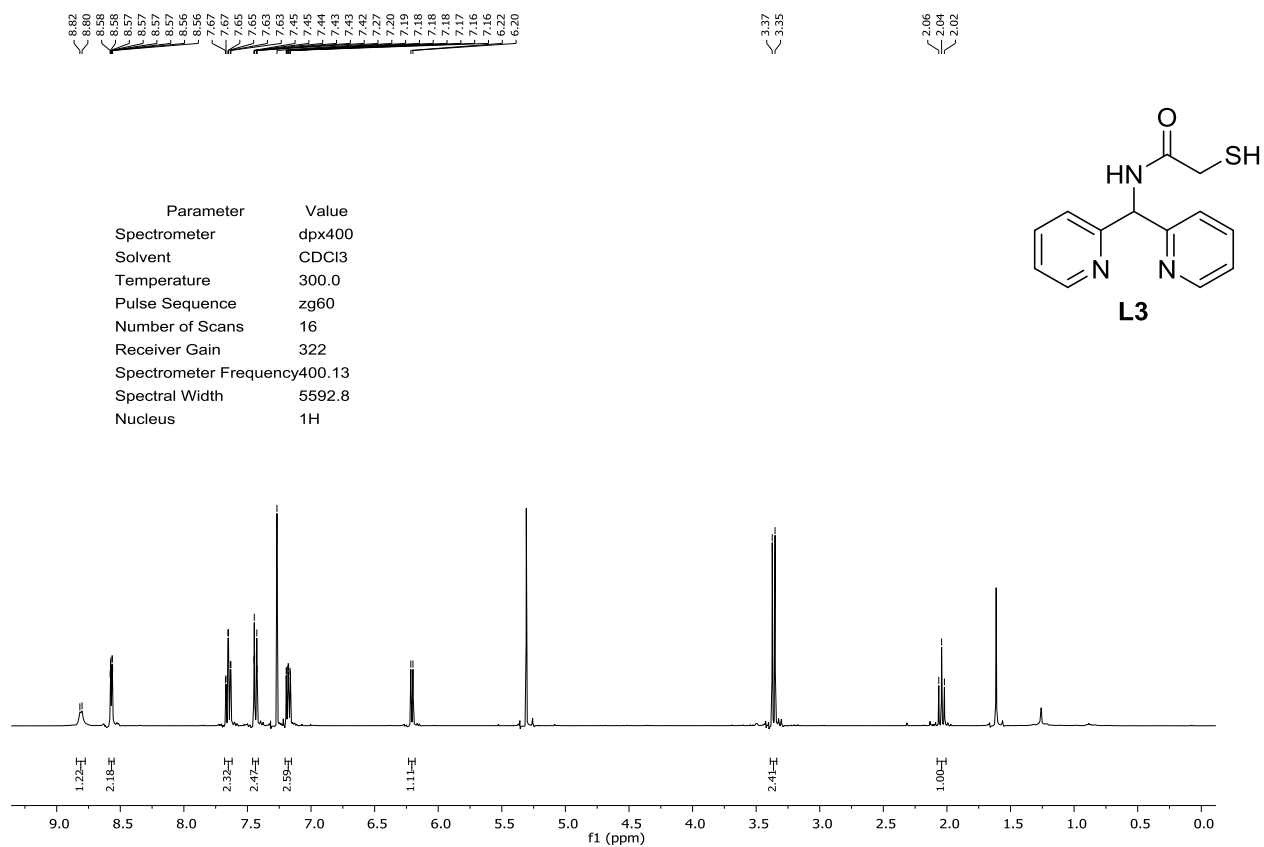
— 16.31

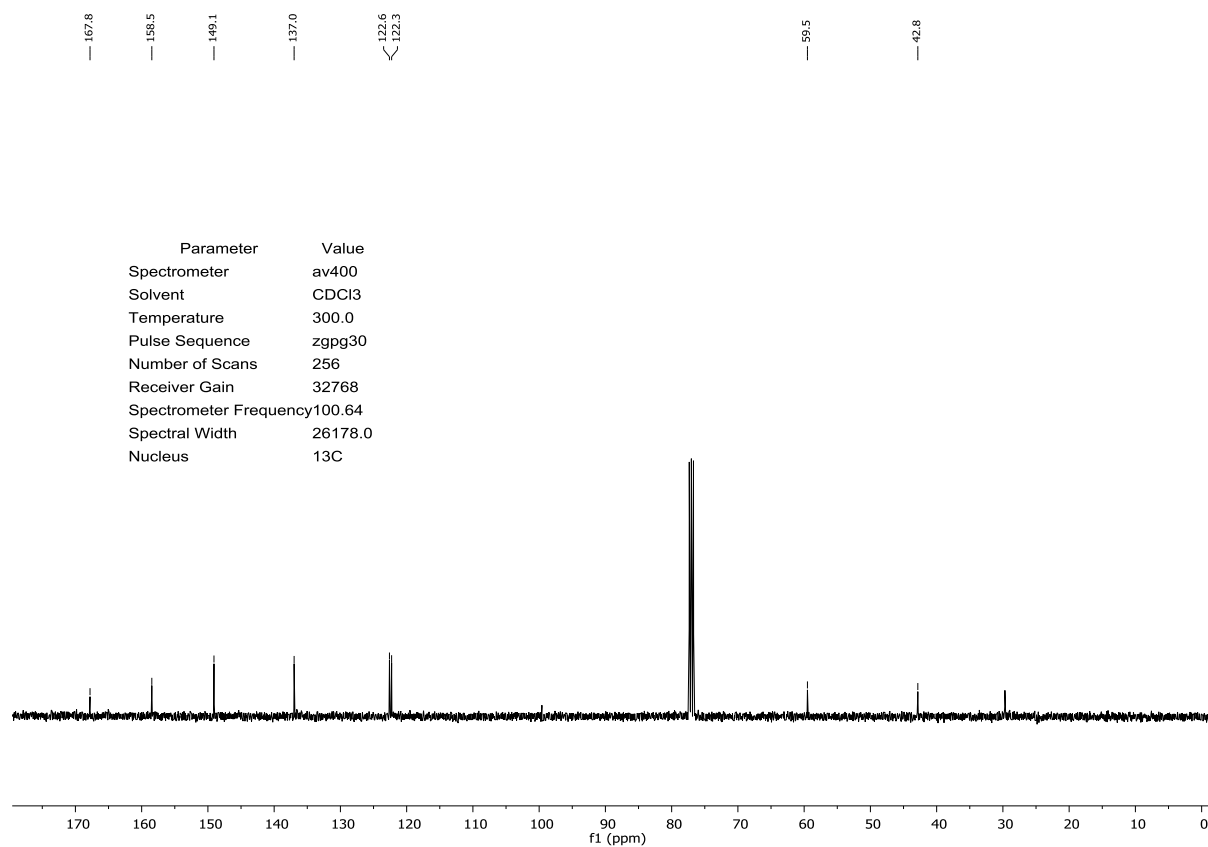
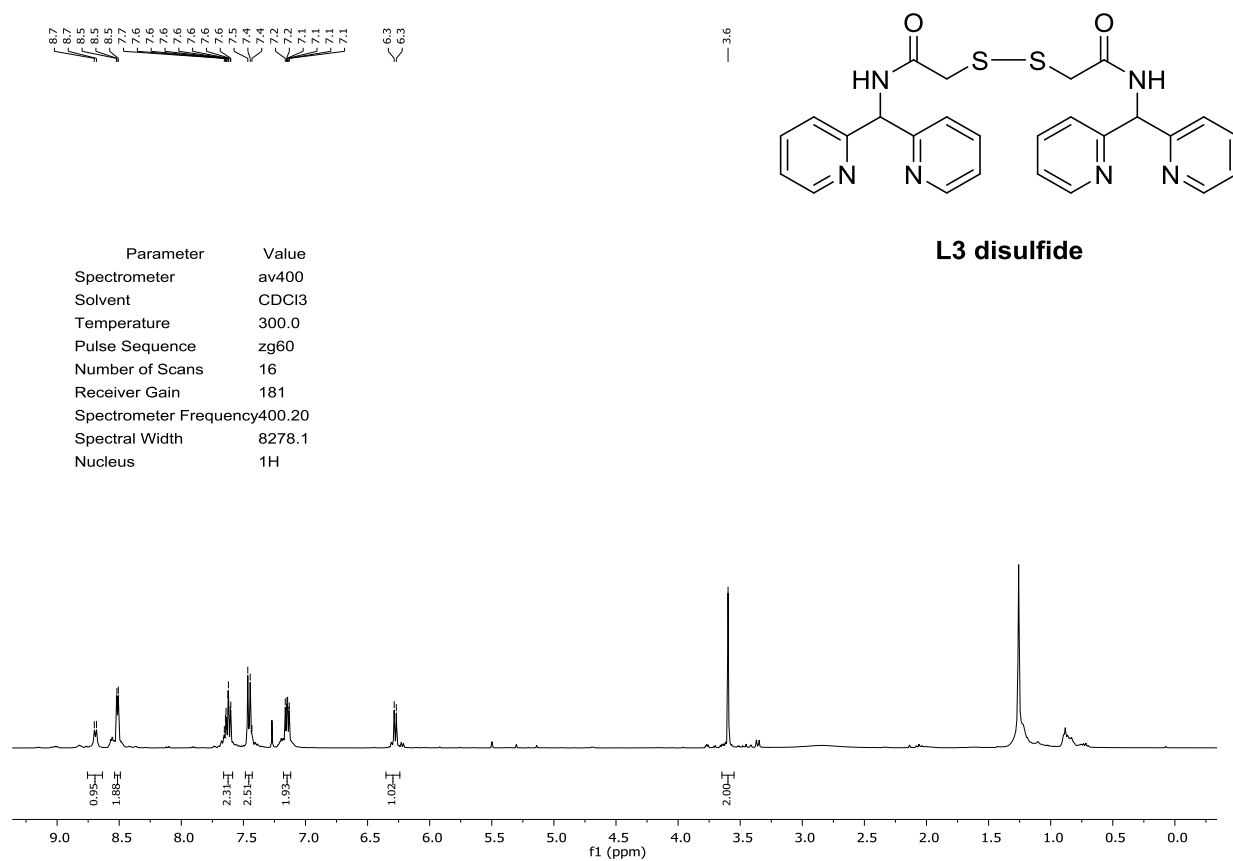


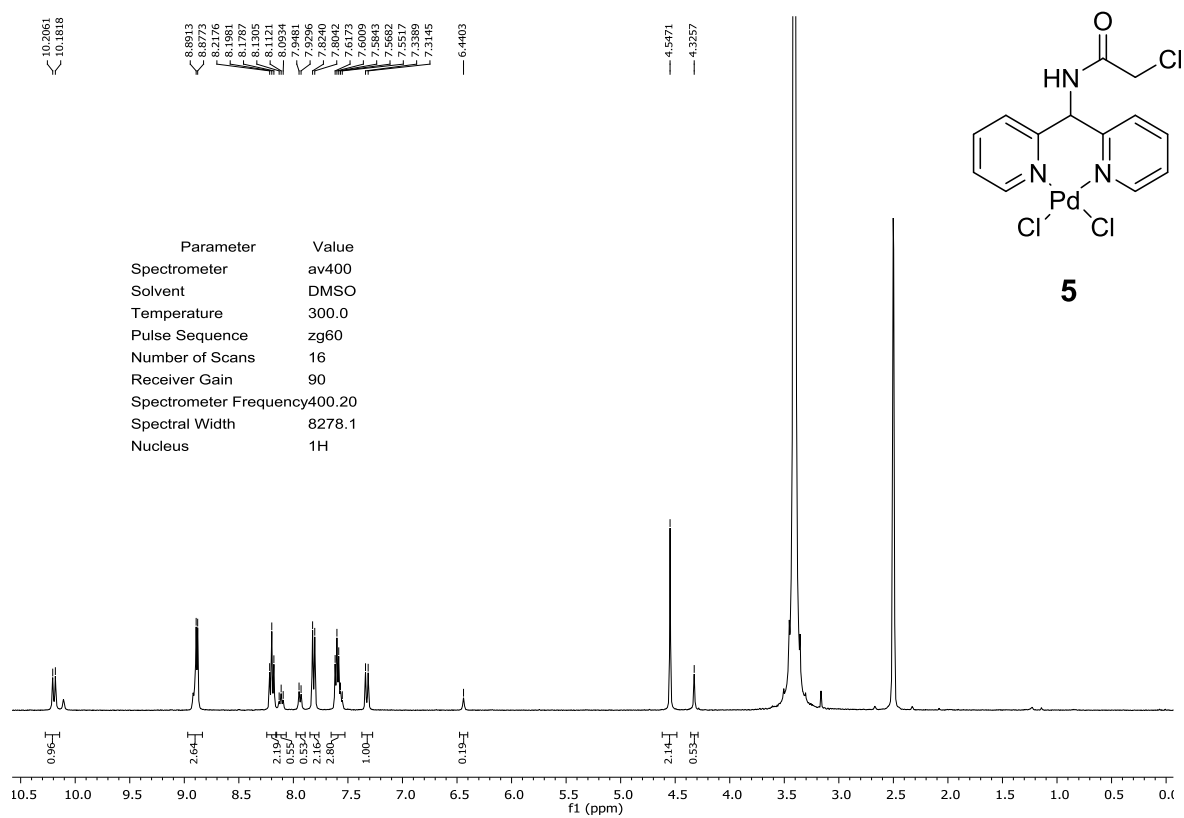


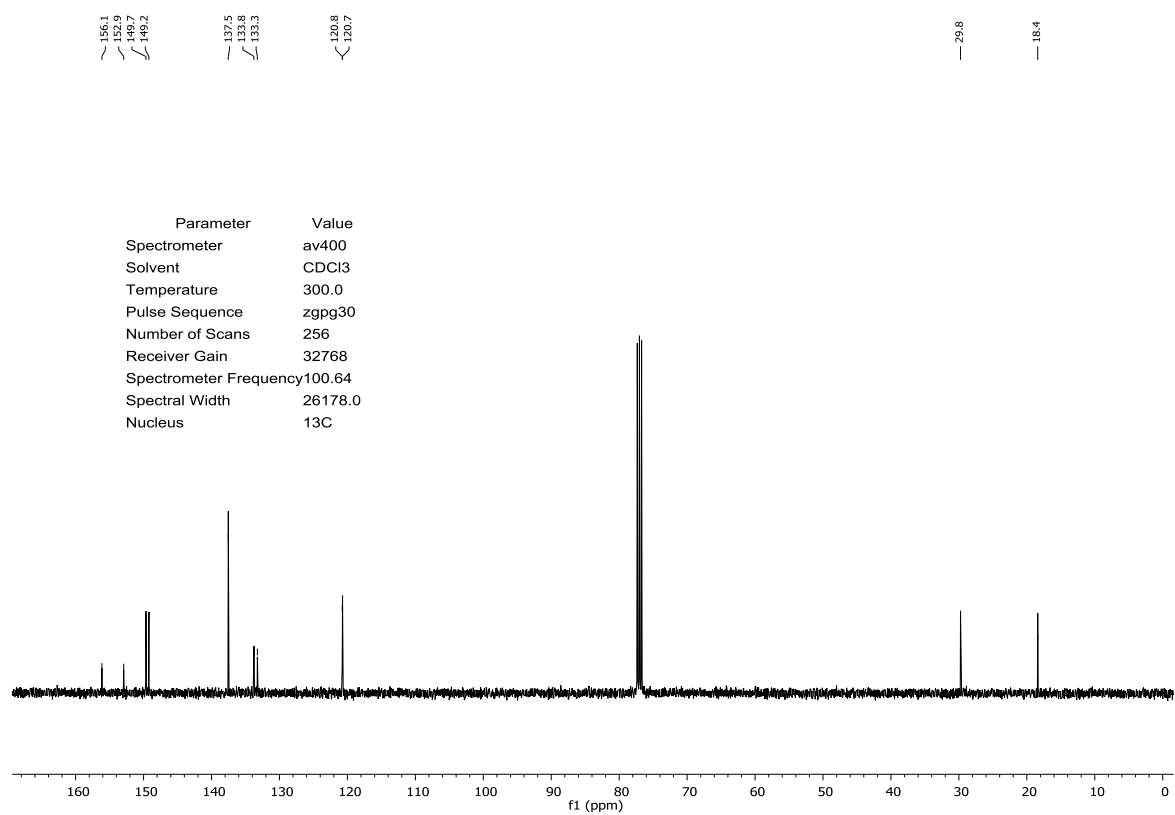
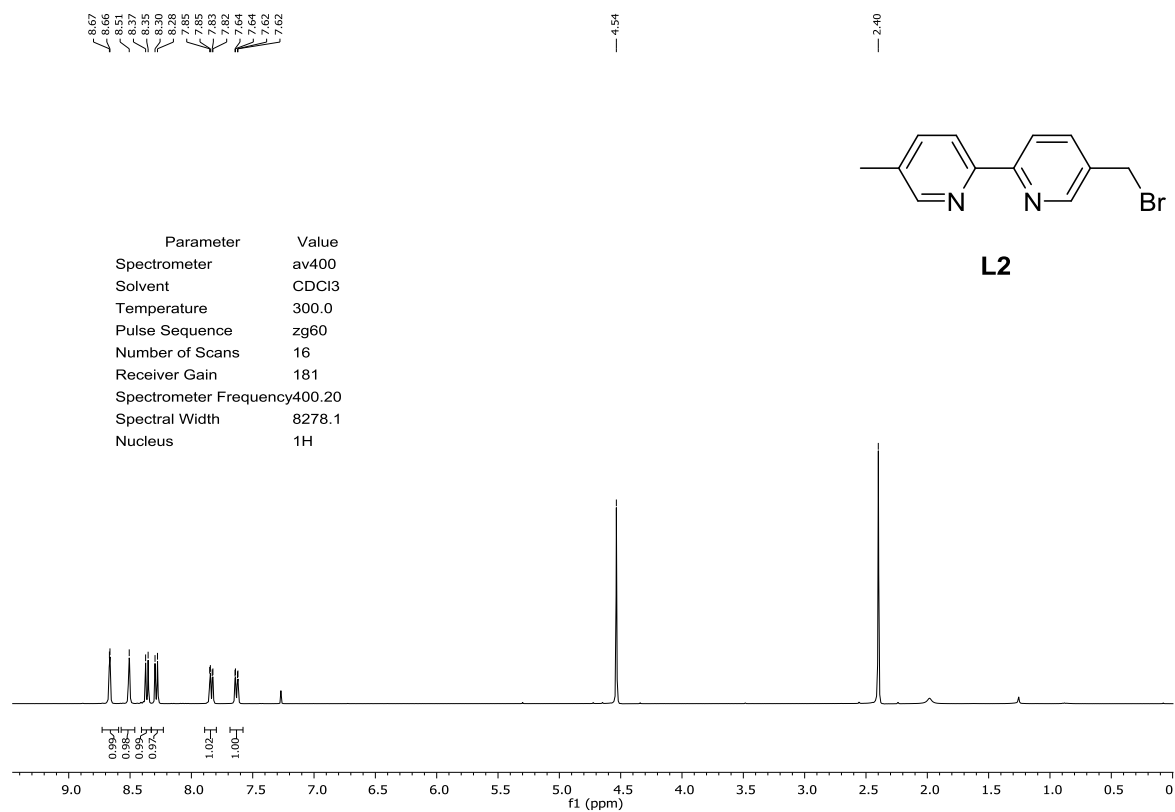


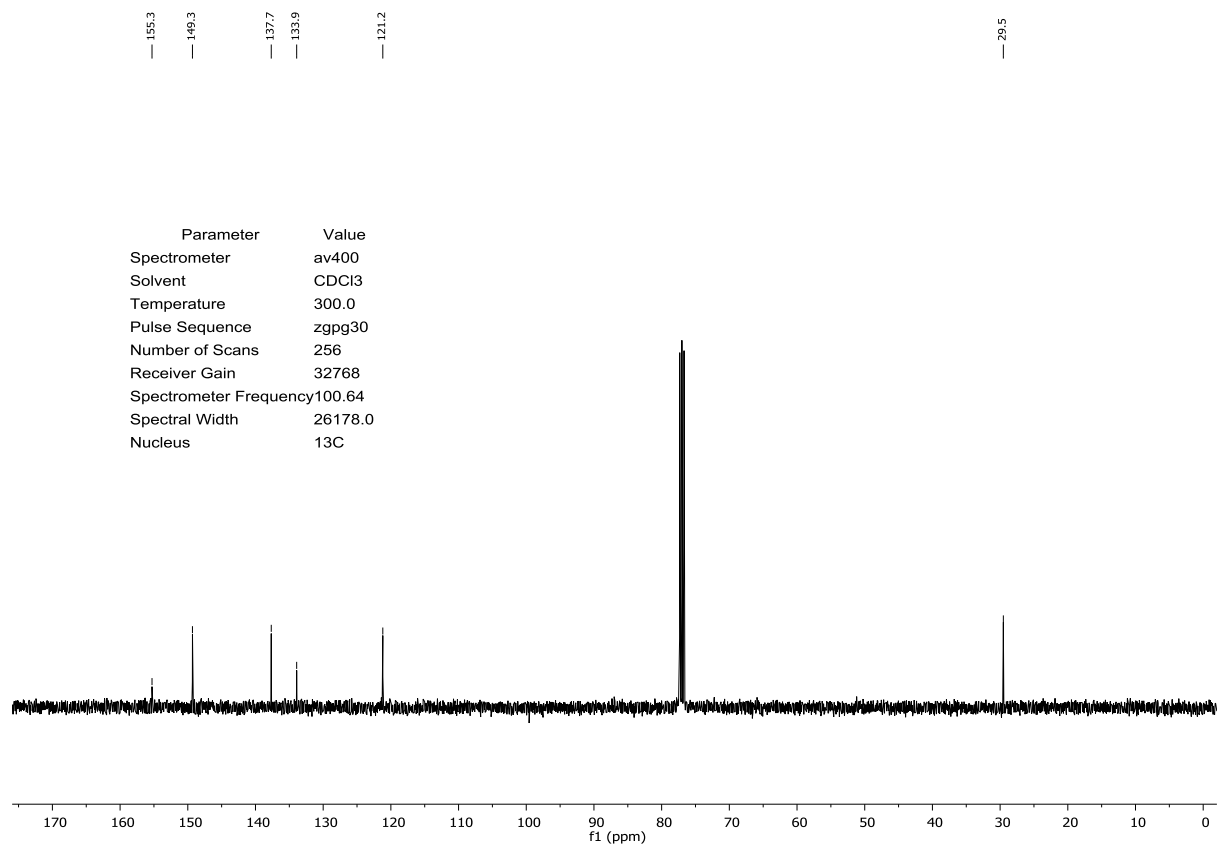
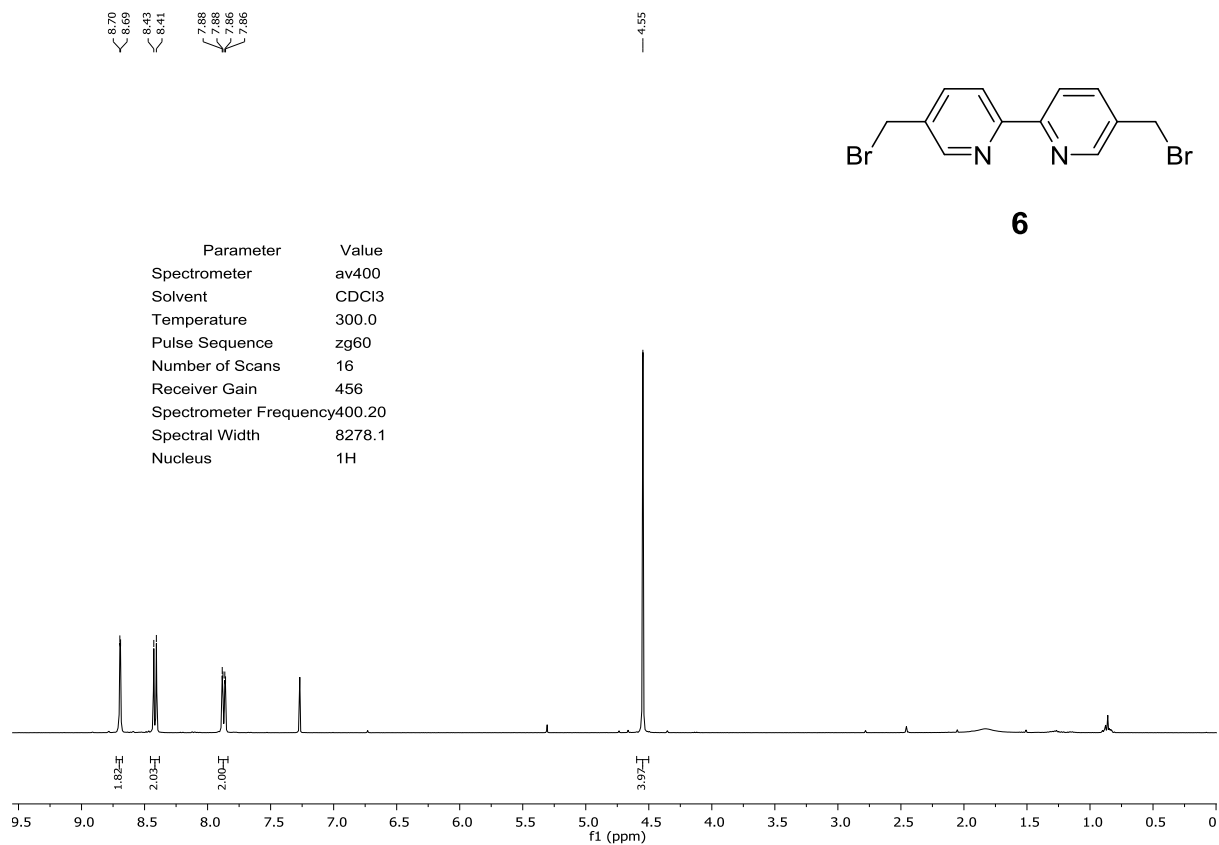


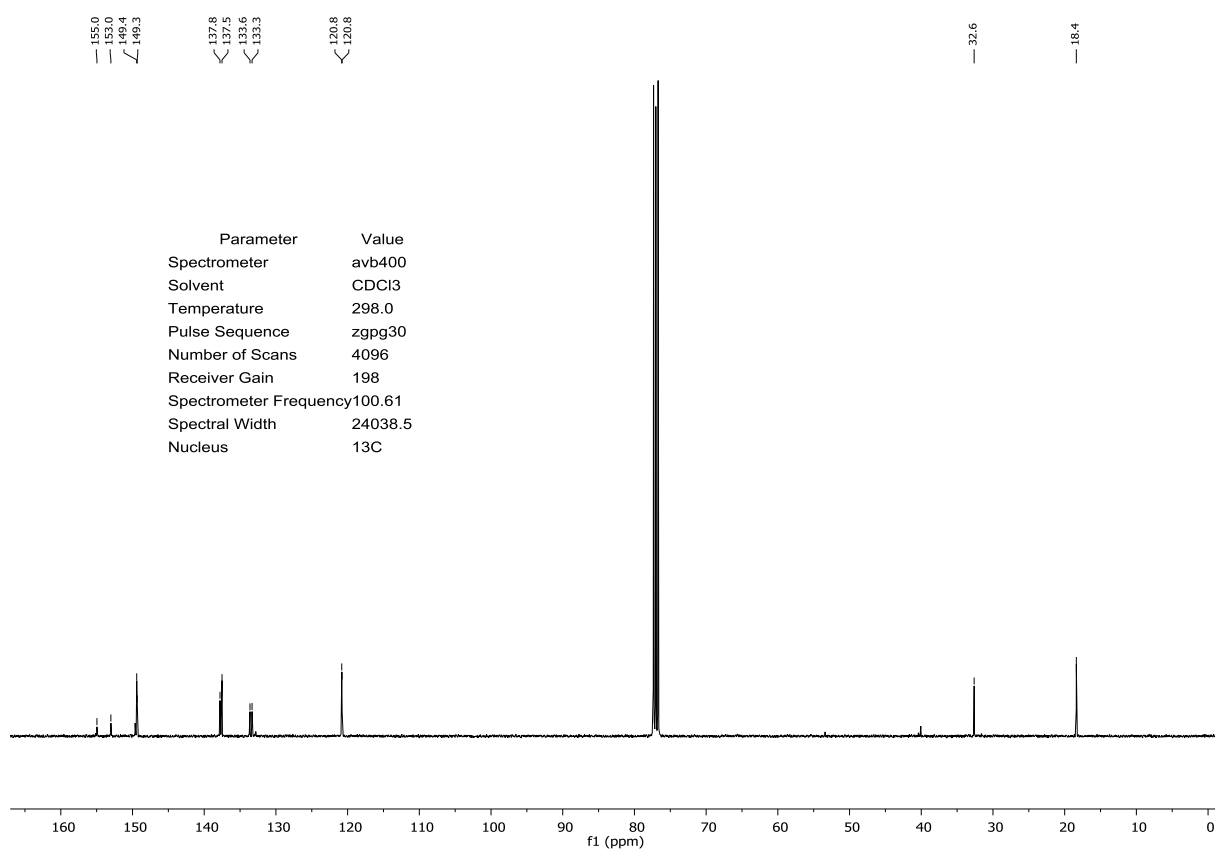
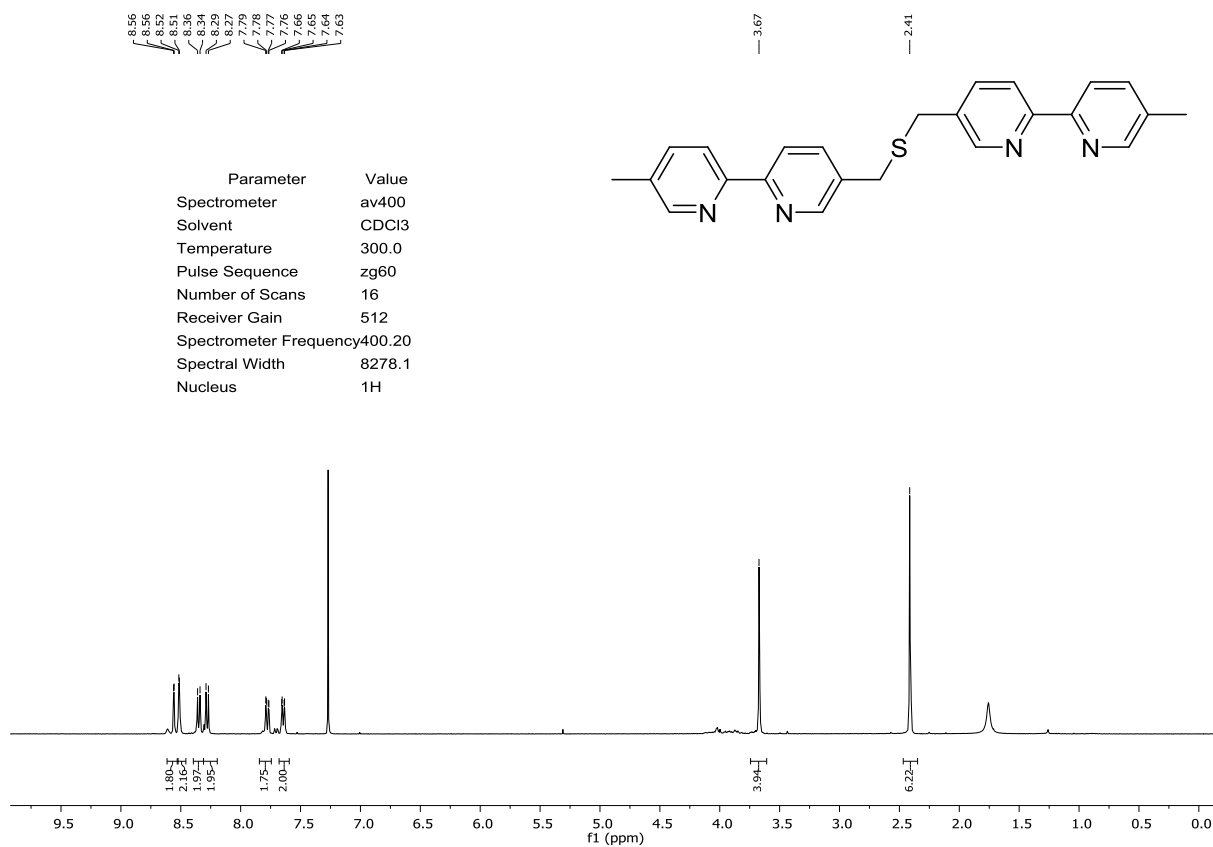


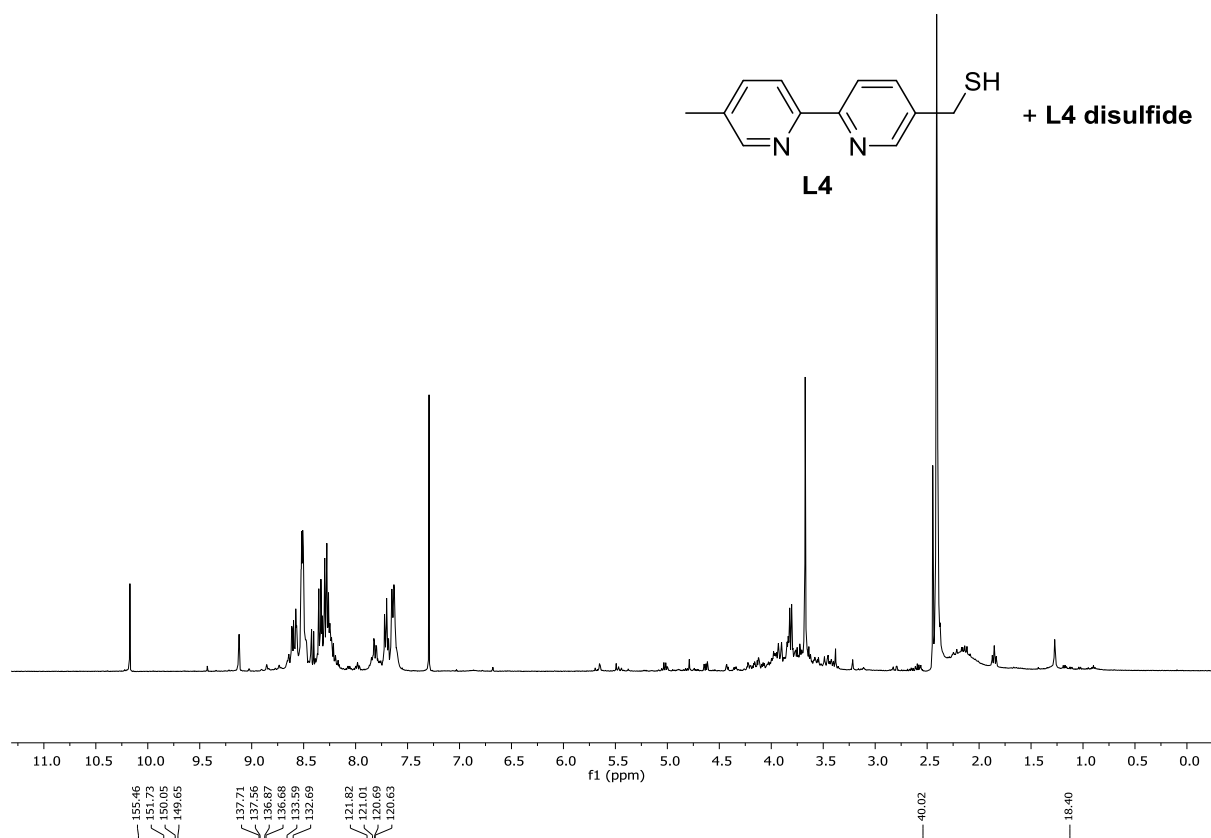




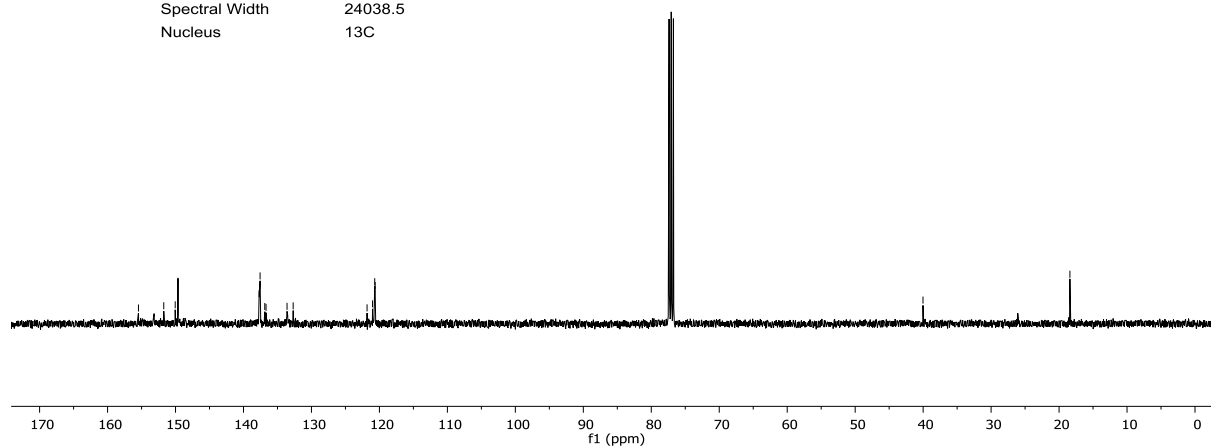


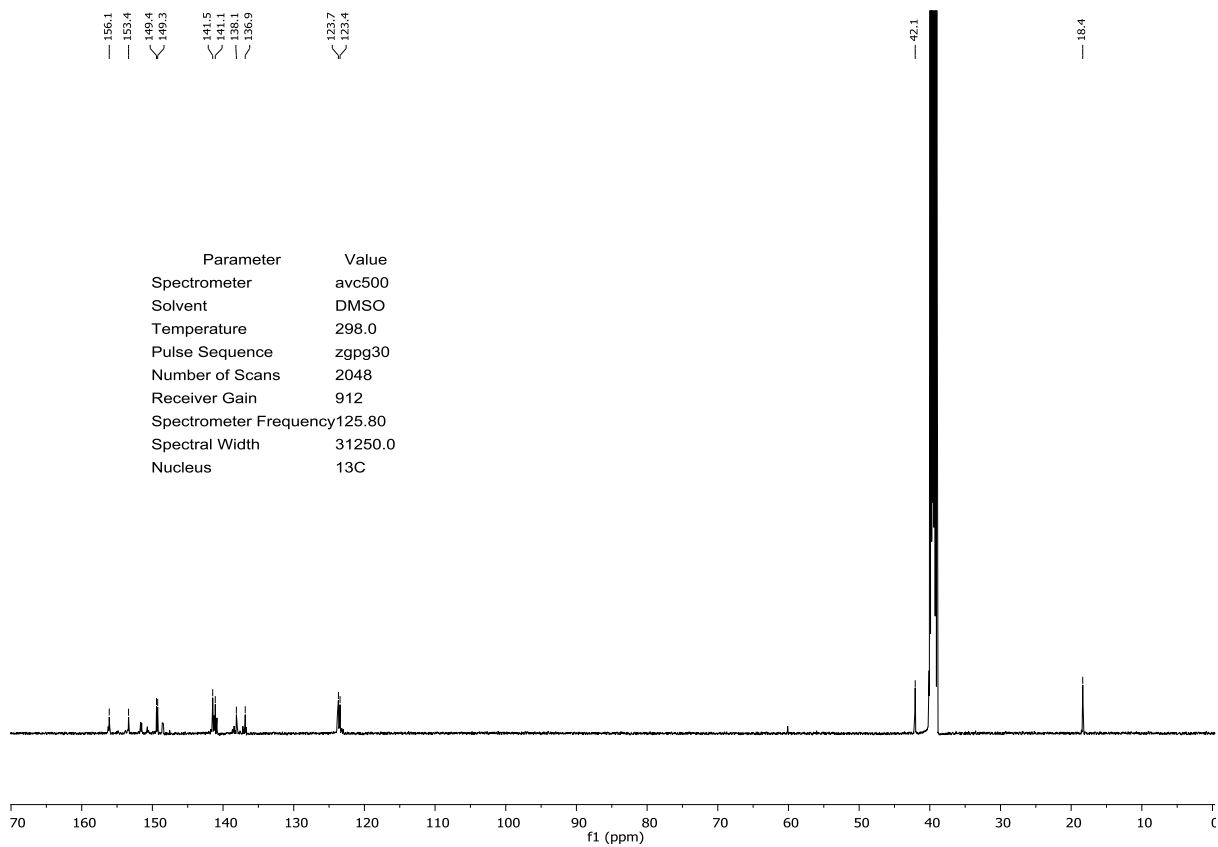
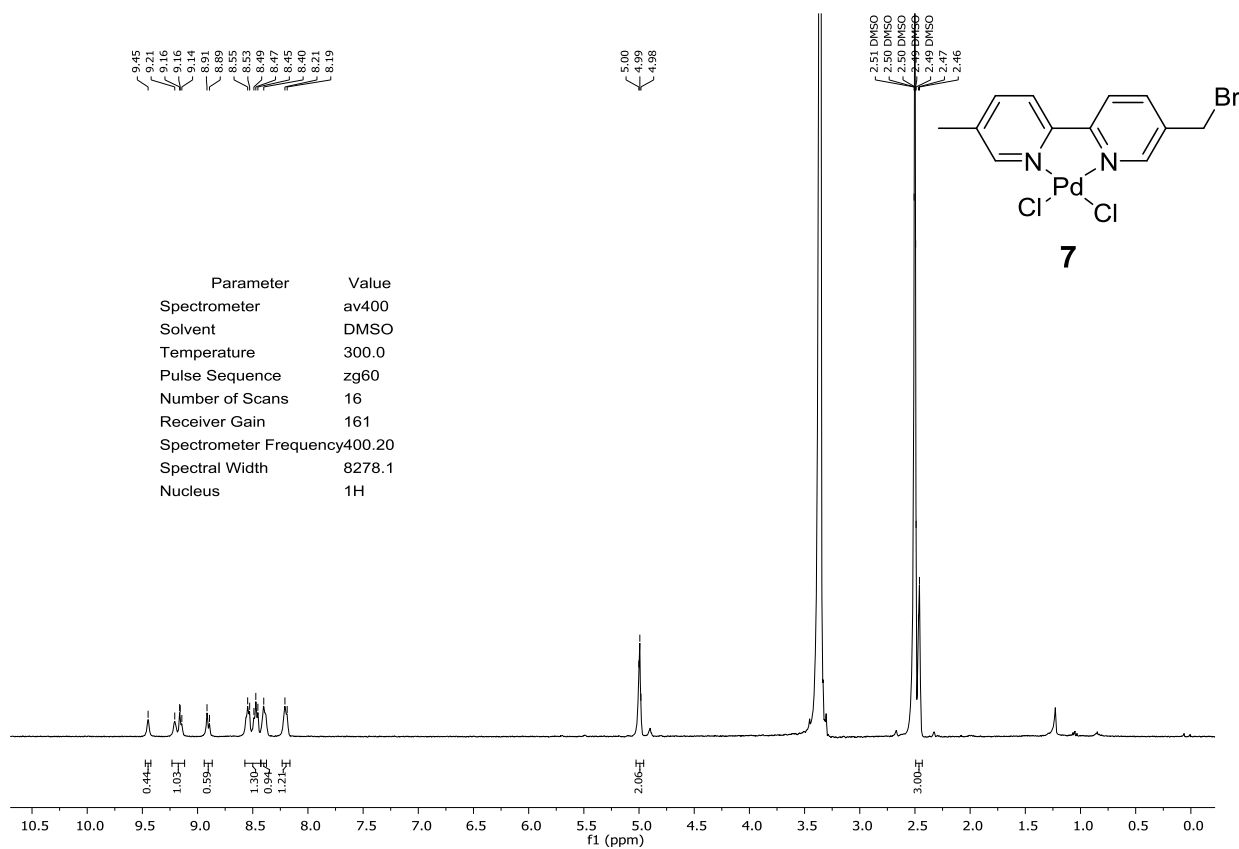


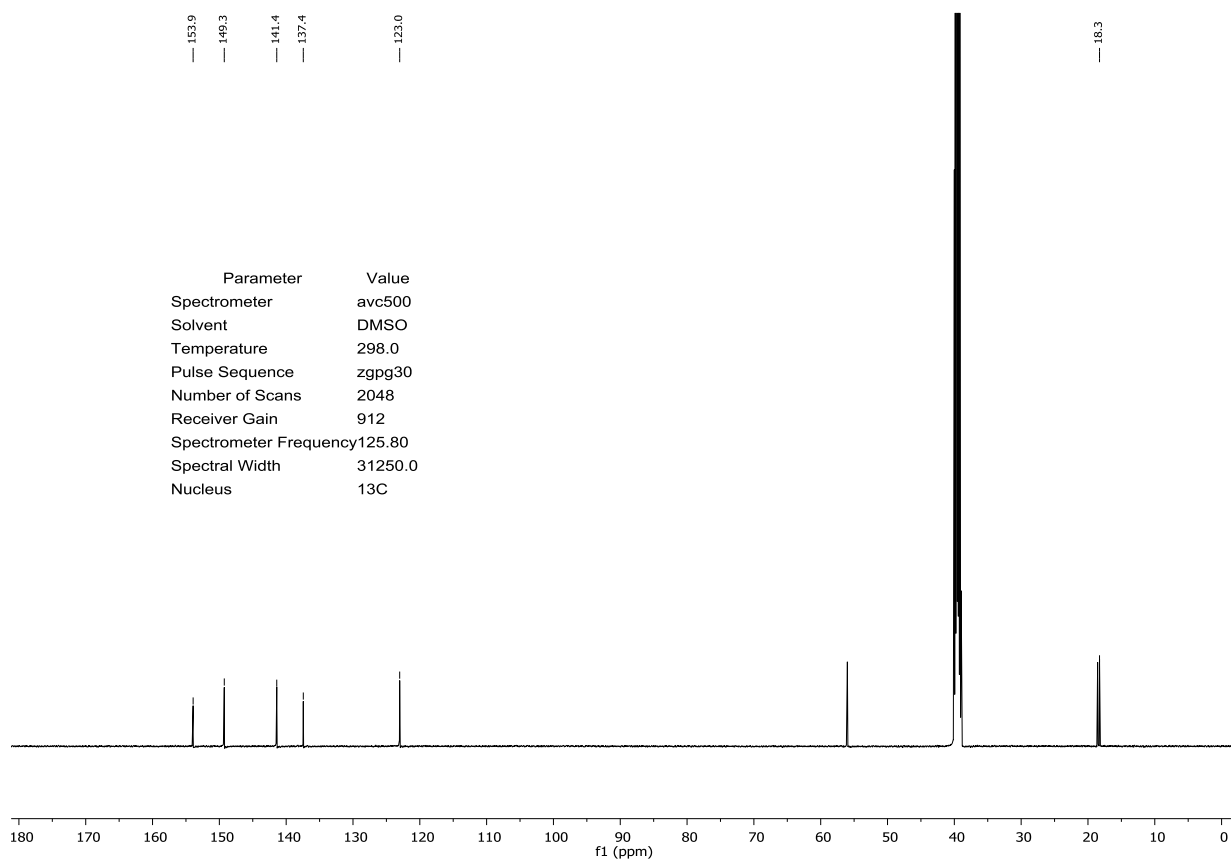
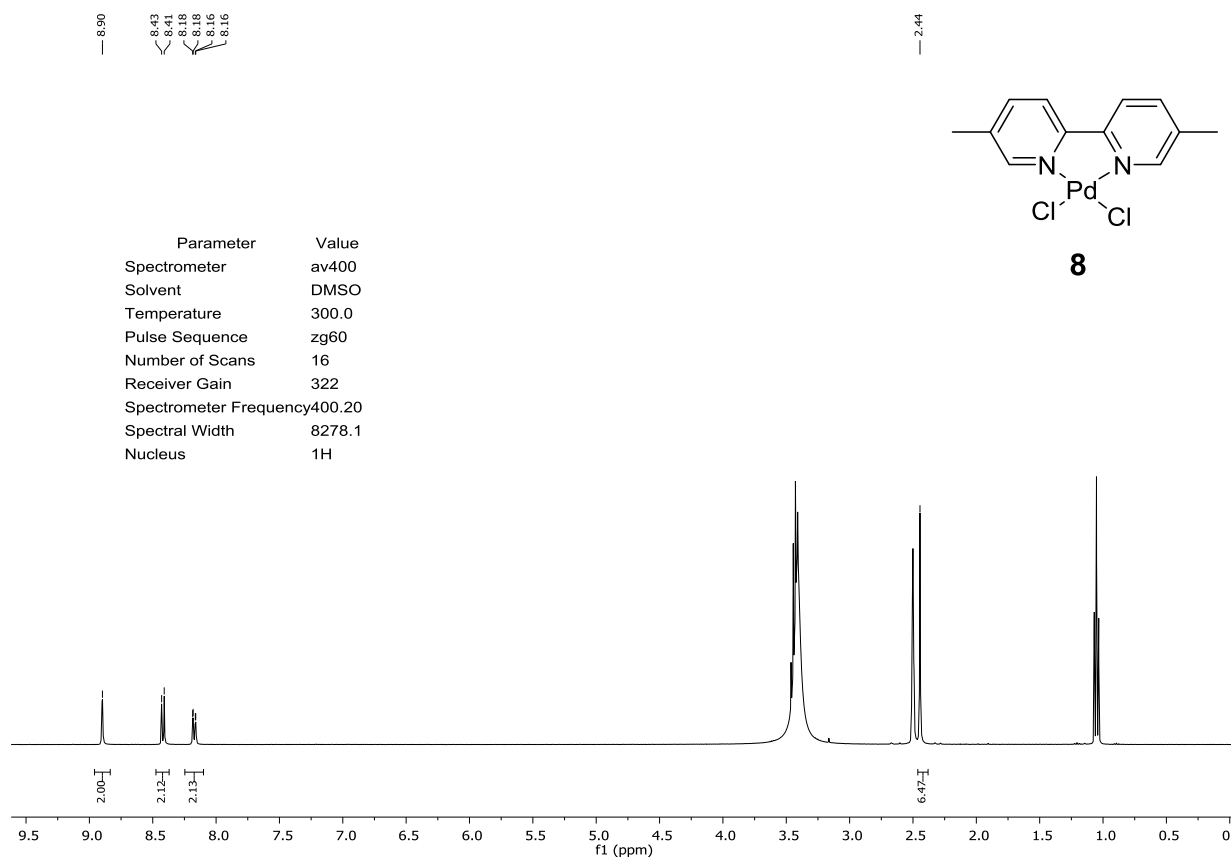




Parameter	Value
Spectrometer	avg400
Solvent	CDCl3
Temperature	294.4
Pulse Sequence	zgpg30
Number of Scans	256
Receiver Gain	2050
Spectrometer Frequency	100.64
Spectral Width	24038.5
Nucleus	¹³ C







Chapter 3

Oxidative Heck cross-coupling reaction using SBL-based artificial metalloenzymes

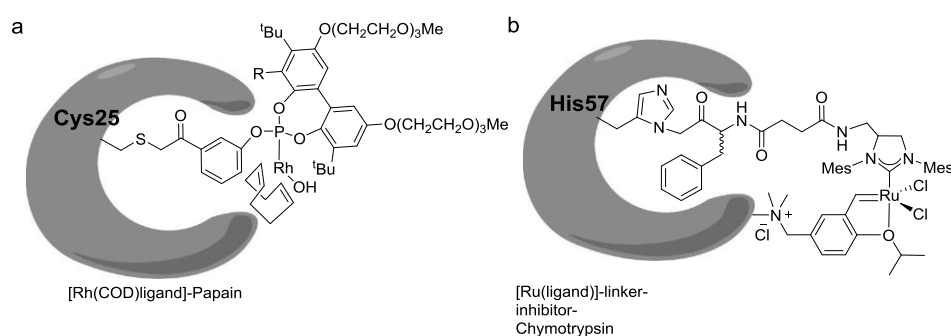
3.1 Introduction

3.1.1 C-C bond forming reactions by artificial metalloenzymes

Several artificial metallo-bioconjugates for important C-C bond forming reactions have been explored.¹⁻¹³ In general, non-covalent supramolecular approaches like the biotin-Strept(avidin) technology^{1-3,14-17} and DNA intercalation⁴⁻⁶ have been shown to catalyse a range of reactions with high conversions and moderate-to-excellent enantioselectivities.

These methodologies have been used to complement metal activity with secondary coordination sphere interactions such that the biomolecular framework provides for asymmetric induction. Often a spacer/ linker group is attached to the metal ligand moieties to distance the metal from the protein pocket, thereby increasing metal accessibility and activity.^{2-4,9} This results in the creation of a biomolecular ligand serving as an elegant chiral auxiliary for complex organic reactions. In such systems, the activity of the metal and the selectivity conferred by the biomolecule essentially remain deconvoluted.¹⁷ For instance, covalent attachment of organometallic metal complexes to catalytic residues of papain¹⁸⁻²¹ and lipase¹⁸ has been used to generate artificial metalloproteins (Scheme 3-1a). Although such an anchoring strategy leads to inactivation of the enzyme, the non-native metal installed serves as new reaction centre and enzymes can be evolved in a

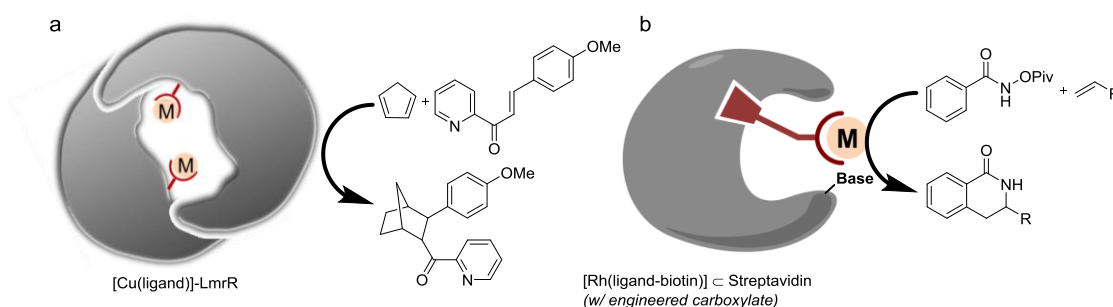
‘Darwinistic’ manner to yield asymmetric catalysts.²² Similar to this approach, the inhibition mechanism of a *N*-tosyl phenylalanyl chloromethylketone (TPCK)-like serine protease inhibitor was used to install Grubbs catalyst (containing a linker moiety) along the line of substrate binding pockets (S1-S4) in α -Chymotrypsin (Scheme 3-1b). This active site modification rendered the hydrolytic enzyme inactive yet the metal displayed ring closing metathesis (RCM) activity.⁹



Scheme 3-1. Artificial metalloenzymes based on proteases created using (a) Covalent attachment of a phosphite ligand to catalytic Cys25 of papain¹⁹ (b) Covalent attachment of Hoveyda-Grubbs catalyst using intrinsic inhibition mechanism of a TPCK-like inhibitor⁹

Recently, an interesting design was based on covalent anchoring of metal binding co-factors on LmrR (Lactococcal multidrug resistance Regulator) at the interface of this homodimeric protein (Scheme 3-2a).⁷ This new [Cu(II)-ligand]-LmR metalloenzyme catalysed Diels-Alder reactions with high enantioselectivity (up to 97% ee). It was shown that rate acceleration was possible, as evidenced by higher conversions to the desired product using this artificial metalloenzyme when compared to the free Cu-complex in aqueous solution.⁷ Excitingly in another design, the synergistic behaviour between the protein and the metal catalyst was highlighted where a carefully positioned basic residue (aspartate) acts as an activator for a Rh-catalysed C-H activation step and helps to tune the properties of the metal ion (Scheme 3-2b).¹ The enantioselective coupling of a range of

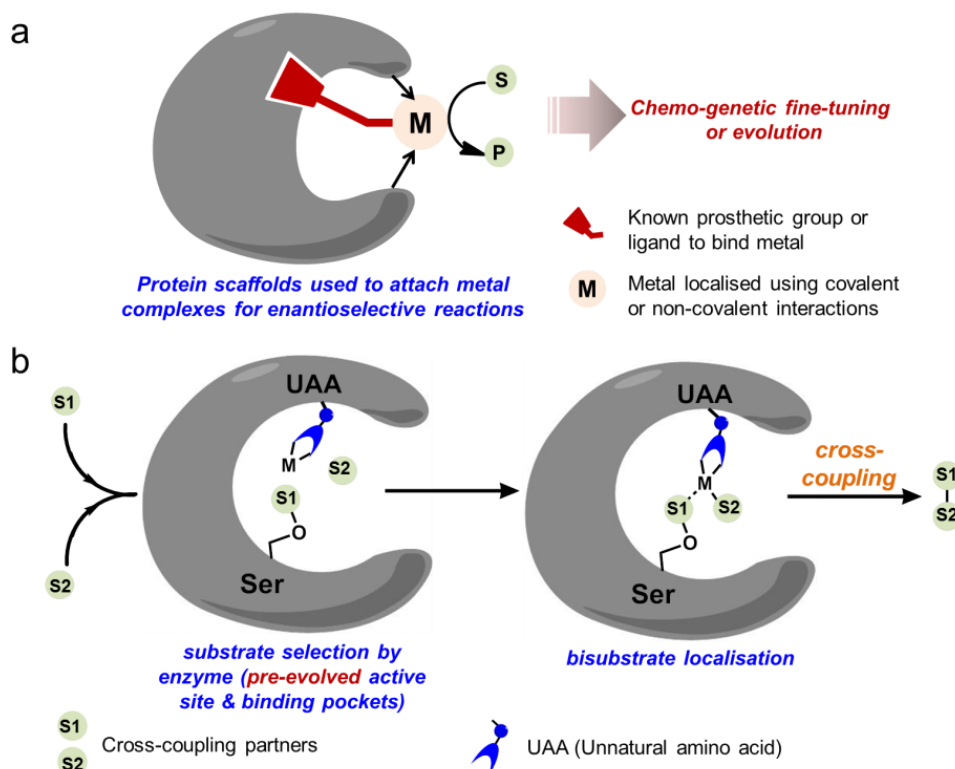
benzamides and alkenes was demonstrated with 92-fold rate acceleration and enantiomeric ratios up to 93:7.¹



Scheme 3-2. (a) Diels-Alder reaction at the dimer interface of artificial metalloenzyme created using a homodimeric protein scaffold⁷ (b) Benzannulation reaction proceeding *via* C-H activation using a Rh(III)-based artificial metalloenzyme¹

Overall, both chemogenetic fine-tuning and evolutionary approaches have been used to generate artificial metalloenzymes (Chapter 1 Section 1.3.2). In this way, the attractive and complementary features of both homogeneous and enzyme catalysis have been combined to introduce unnatural reactivities into biomolecules. However, an important natural feature of enzyme catalysis, namely naturally evolved substrate specificity, has been overlooked. Our aim was to employ the pre-evolved enzyme pockets to localise the cross-coupling partners and carry out non-native transformations (Scheme 3-3). We envisioned that a hydrophobic substrate binding pocket of a serine protease could help to recruit the aryl coupling partners (bisubstrate localisation), while the redox metal centre installed in the vicinity could enable the cross-coupling reaction. Previously, the chemical installation of *N,N*-chelating moieties such as found in the unnatural amino acids (UAAs) bipyridine-cysteine (Bpc) and dipyrityl-cysteine (Dpc) was successfully demonstrated at two sites on the model protein to yield tailored enzymes, namely SBL-Bpc217, SBL-Dpc217, SBL-Bpc156 and SBL-Dpc156 (Chapter 2 Section 2.2.2.4). Fundamentally, this approach excludes any linker moieties that could remove the metal catalyst away from the pocket. In

this chapter, we describe our efforts to study C-C cross-coupling reactions proceeding via elaborate sequence of steps in the enzyme active site.

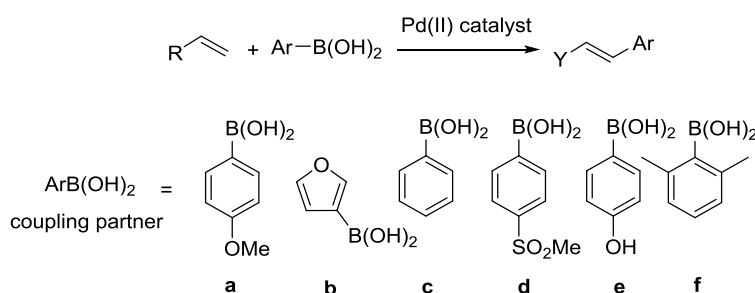


Scheme 3-3. (a) General representation of artificial metalloenzymes where protein scaffolds serve as ‘hosts’ for catalytic metal or tune the properties of the metal ion through second coordination sphere interactions (b) Design of artificial metalloenzymes based on bisubstrate localisation strategy using convoluted enzyme and metal manifolds

3.2 Results and discussion

3.2.1 Design of substrates and modelling studies

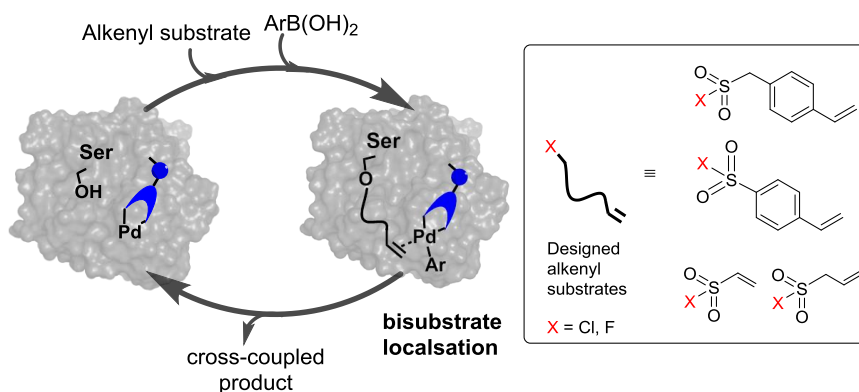
We envisioned that the enzyme selectivity could be combined with a catalytically potent metal to facilitate Pd-catalysed cross-coupling reactions (Chapter 1 Section 1.5). In order to demonstrate the synergy between two potent biological and chemical catalysts (SBL and palladium), oxidative Heck cross-coupling²³⁻²⁷ between an olefin and phenylboronic acids was chosen as the model reaction (Scheme 3-4).



Scheme 3-4. Oxidative Heck cross-coupling between olefin and arylboronic acid substrates

3.2.1.1 Design and synthesis of substrates

To begin with, the enzyme-substrate binding interactions of SBL were revisited (See Chapter 2 Section 2.2.1.2). The S1 pocket of SBL has an innate specificity for hydrophobic substrates. The total binding volume available for P1 side chains within this S1 pocket is $169 \pm 30 \text{ \AA}^3$, which fits benzyl and phenyl groups comfortably and selectively.²⁸ In general, the S1 pocket accommodates the side-chains of amino acids with a preference for Phe, Tyr > Leu, Met, Lys > His, Ala, Gln, Ser >> Glu, Gly.²⁸⁻³⁰ With an aim to exploit this S1 selectivity and natural chemo-selection facilitated by the catalytic machinery (Ser221-His64-Asp32 triad), nucleophilic localization of phenylmethanesulfonyl fluoride (PMSF) based alkenyl substrates was envisaged (Scheme 3-5).



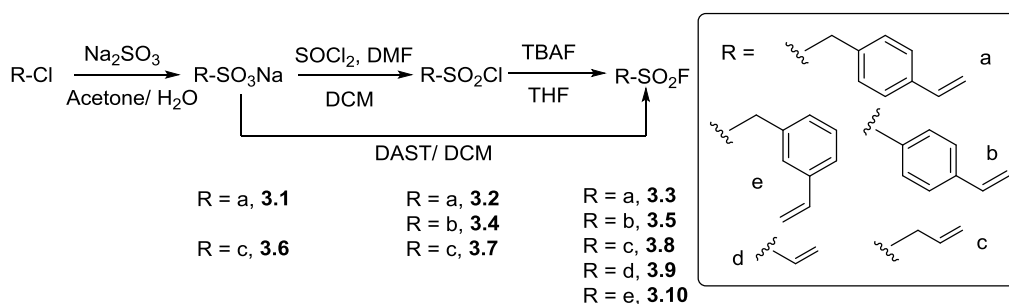
Scheme 3-5. Proposed oxidative Heck cross-coupling reaction in the enzyme active site, facilitated by bisubstrate localisation of the alkene and arylboronic acid substrates

PMSF-based derivatives were chosen as substrates to establish proof-of-concept for our bisubstrate localisation strategy, mainly due to the following reasons, (i) high selectivity of serine proteases like SBL towards hydrophobic substrates with phenyl rings could be exploited, (ii) strong covalent interaction of the PMSF with the active site serine residue would enable the detection of the enzyme intermediates of the cross-coupling reaction using LC-MS analysis, (iii) the alkane sulfonyl group is a good leaving group towards S_N2 displacement reactions, so that release of the final product by the enzyme could be promoted. Thus particularly for these substrates, it was speculated that the enzymic selection would exploit both complementary fit (S1 pocket) and chemoselectivity (specific reactivity of Ser221) of the active site.

The designed PMSF derivatives were synthesised by using known chemical routes for synthesizing sulfonyl fluorides[†] (Scheme 3-6).³¹⁻³³ In general, a non-oxidative synthesis of such PMSF derivatives was carried out which involved refluxing benzyl halides with sodium sulphite, followed by treatment with a chlorinating agent like thionyl chloride ($SOCl_2$). Subsequent fluoride exchange gave the desired product. Alternatively, sulfonyl fluorides were obtained in a single step from the corresponding sodium sulfonates using diethylaminosulfur trifluoride (DAST) as the fluorinating agent. Using this established procedure, sulfonyl halides **3.2**, **3.3** and **3.5**[†] were obtained in moderate yields (See Experimental Section 3.5.2). However, the synthesis of substrates **3.8**[†] and **3.10**[†] using this

[†]Dr. Alba Díaz-Rodríguez conducted exploratory work in our laboratory to synthesize PMSF-based substrates (including **3.3** and **3.10**) to study cross-coupling reactions by mass spectrometry. Synthesis of compound **3.3** and its interaction with SBL was confirmed by her. The compound **3.3** was re-synthesized for the work reported here using the same route to continue the investigation of cross-coupling reactions using substrate **3.3**. Syntheses of compounds **3.5** and **3.8** were done by Dr. Maria Espling. Experimental results that were obtained by them are indicated in the body of this chapter by reference to this footnote and their contributions are stated explicitly in the experimental sections of this chapter.

approach was not successful. Ethenesulfonyl fluoride **3.9** was obtained from commercial sources.

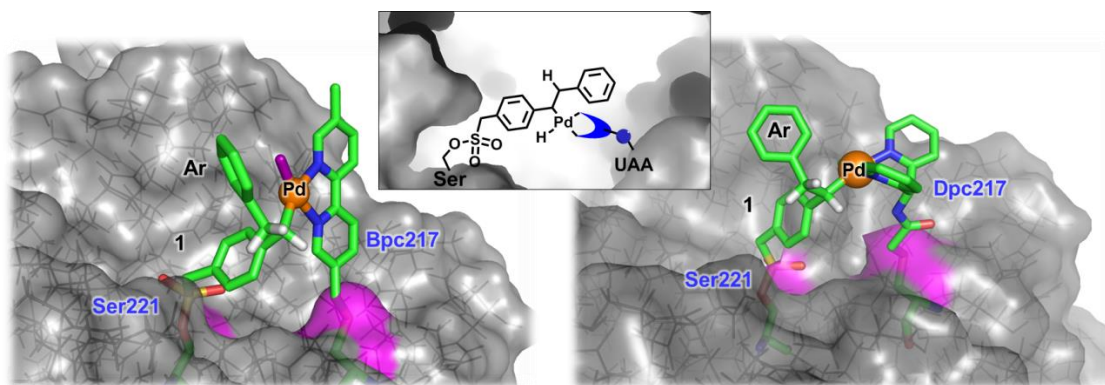


Scheme 3-6. Synthetic routes to obtain sulfonyl halide substrates

3.2.1.2 Modelling studies[†]

Preliminary modelling studies were performed to model the proposed ‘bisubstrate localisation’ intermediate in SBL-Bpc217 and SBL-Dpc217 that contain the metal binding amino acids installed in the S1' pocket. By using the SBL-WT crystal structure (PDB ID: 1GCI),³⁴ *p*-vinylPMSF **3.3** (PMSv) substrate was modelled at Ser221 and cysteine at position 217 was converted to Bpc or Dpc residue. Next, palladium was modelled on Bpc/Dpc in a fixed square planar coordination geometry. Molecular mechanics suggested that the key C-C bond forming step between the alkene moiety and an aryl boronic acid was plausible. This implies that the cross-coupling intermediate bridged between the S1 and S1' pockets of SBL could be accommodated in the active site as shown (Scheme 3-7). Hence, modelling indicated that this interplay of S1 and S1' pockets could be utilized to create an activating microenvironment for cross-coupling reactions.

[†] Modelling experiments were carried out by Dr. Mark Wormald



Scheme 3-7. Cross-coupling intermediate for oxidative Heck reaction between substrate **3.3** and phenylboronic acid, bridged between S1 and S1' pockets, in SBL-Bpc217 (left) and SBL-Dpc217 (right)

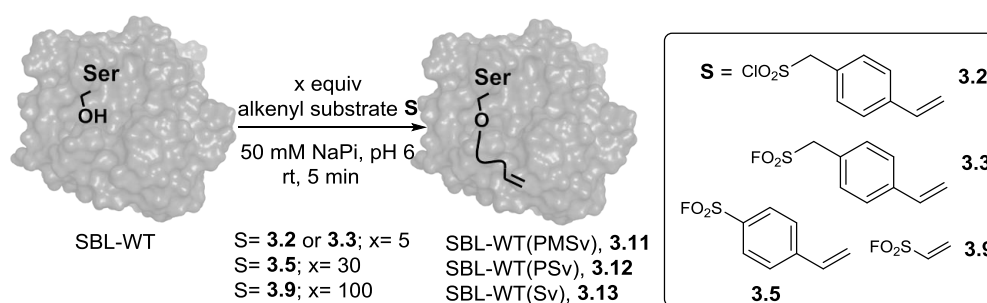
In addition to this, smaller substrates including allylSO₂F **3.8**, vinylSO₂F **3.9** and *m*-vinylPMSF **3.10**, were modelled onto Ser221 to check their suitability as coupling partners for the target reaction (See Experimental Section 3.5.3). Modelling studies pointed out that cross-coupling between aryl boronic acids and *m*-vinylPMSF **3.10** was plausible. However, based on the relative distances and orientations of the substrates and the metal catalyst in the active site, the lack of phenyl moiety in substrates **3.8** and **3.9** may pose a challenge in realising cross-coupling reaction using the proposed strategy. Given the difficult synthesis of compound **3.8** and **3.10**, sulfonyl halides **3.2**, **3.3**, **3.5** and **3.9** were used for studying the enzyme-substrate interactions described in section 3.2.2.

3.2.2 Enzyme-substrate interactions

3.2.2.1 Interaction of new PMSF analogues with SBL

The highly specific interaction of nucleophilic active site Ser221 towards sulfonyl halides, coupled with the selectivity of the S1 pocket for hydrophobic moieties, are both central to our approach. We first assessed the nucleophilic interaction of sulfonyl halides **3.2**, **3.3**, **3.5** and **3.9** with SBL-WT (Scheme 3-8).

The nucleophilic attack by Ser221 on designed *p*-vinylphenylmethanesulfonyl (PMSv) substrates **3.2** and **3.3** (analogous to phenylalanine side-chain of P1 residue in native substrates) was successfully confirmed by LC-MS analysis (See Experimental Section 3.5.4). The ESI-MS spectrum revealed complete uptake of substrate to give SBL-WT(PMSv) adduct **3.11** with as low as 5 equiv. of the sulfonyl chloride **3.2** or sulfonyl fluoride **3.3**. Given the high nucleophilicity of Ser221, structurally similar substrates, including *p*-vinylphenylsulfonyl (PSv) fluoride **3.5** and vinylsulfonyl (Sv) fluoride **3.9** were also successfully loaded on SBL-WT to give corresponding enzyme-substrate adducts SBL-WT(PSv) **3.12** and SBL-WT(Sv) **3.13**, albeit only when high molar excess of these substrates with respect to the enzyme was used (See Experimental Section 3.5.4).



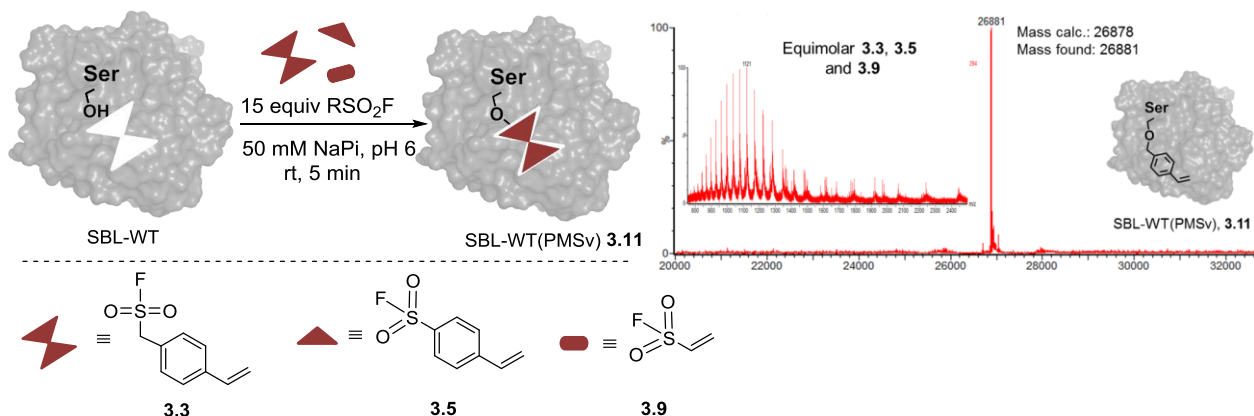
Scheme 3-8. Evaluation of enzyme-substrate interactions to assess the enhanced nucleophilicity of Ser221 in enzyme active site

3.2.2.2 Competition loading of PMSF analogues with SBL

To further investigate the enzymic selection for hydrophobic substrates, a panel of alkene substrates **3.3**, **3.5** and **3.9** was screened in competition loading tests using SBL-WT[†] (Scheme 3-9). When equimolar amounts of each substrate was present in the solution, only loading of PMSv substrate on SBL-WT was observed by LC-MS analysis. Remarkably, the competition loading assay confirmed the exquisite selectivity of SBL for PMSv-based

[†] This experiment was done with Dr. Maria Espling.

alkene substrate **3.3** over vinylphenylsulfonyl/ vinylsulfonyl fluorides **3.5** and **3.9** (See Experimental Section 3.5.5). This provides the supporting evidence that the enzymic selection is derived from the shape complementarity (of the S1 binding pocket and the substrates) as well as the enhanced nucleophilicity of the serine protease.



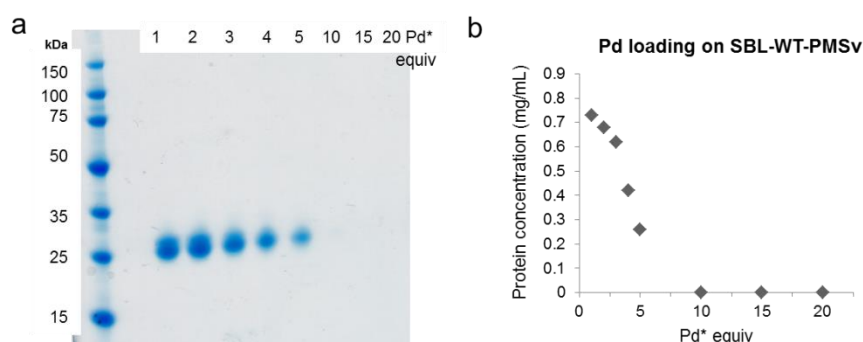
Scheme 3-9. Evaluation of enzyme-substrate interactions to assess the selectivity of S1 pocket of SBL for sulfonyl halide substrates **3.3**, **3.5** and **3.9**

Together taken with the results in section 3.2.2.1, it is expected that hydrophobic moieties of the substrate interact with the binding pocket, followed by nucleophilic attack by the activated Ser221 of the catalytic triad. This highly specific interaction could be used to precisely localise the substrates in the desired enzyme pocket, thereby setting the stage for the subsequent cross-coupling reaction. The selectivity of the mutant enzyme SBL-L217C was also confirmed and similar results were obtained when 10 equiv. of substrate **3.3** was used which led to the formation of enzyme-substrate adduct SBL-L217C(PMSv) **3.14** (See Experimental Section 3.5.6).

3.2.2.3 Palladium loading on SBL-WT(PMSv) species

In our design, installed chelating amino acids Dpc and Bpc play a crucial role in binding the metal in the tailored active site (See Chapter 2 Section 2.2.3.3). However, olefins are also known ligands in transition metal complexes.³⁵ To investigate whether the alkenyl

moiety in modified SBL-WT(PMSv) could load palladium, we incubated SBL-WT(PMSv) with 1-5 equiv. of Pd^{2+} precursor at 37 °C for 1 h. LC-MS analysis revealed that palladium loading on the vinyl moiety of our substrate was ineffective at such low equiv. of Na_2PdCl_4 used as Pd(II) source (See Experimental Section 3.5.7). In this context, we foresee Bpc and Dpc might play an indispensable role in localising the metal to bring it together with the tethered substrates in the active site. Additionally, we also observed significant protein precipitation on incubating SBL-WT with increasing equiv. of Pd(II) source. Using the established procedure (Section 2.2.3.2), the loss of protein could be monitored by non-reducing PAGE gel analysis and bicinchoninic acid (BCA) assay to measure the concentration of soluble protein (Scheme 3-10). This further confirmed the tolerance levels of the protein for a non-native metal (such as palladium) and was found in keeping with our previous results (Chapter 2 Section 2.2.3.2).



Scheme 3-10. (a) Non-reducing PAGE analysis of SBL-WT(PMSv) to confirm protein precipitation; (b) Loss of soluble protein content from the solution on increased loading of palladium measured by BCA assay

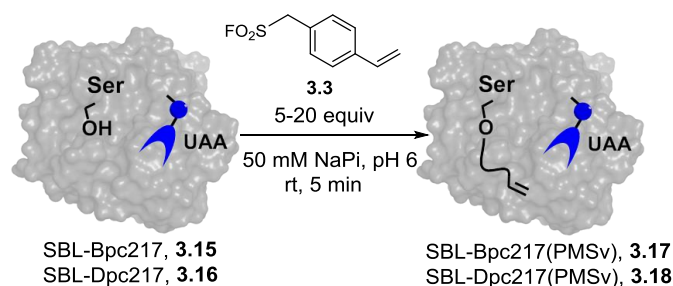
3.2.3 Engineered cycle

Following the construction of the metal-free proteins SBL-Bpc217 (**3.15**) and SBL-Dpc217 (**3.16**) as putative enzymes (See Section 2.2.2.4), we proceeded to test their activity towards the model oxidative Heck cross-coupling reaction between the alkene substrate **3.3** and representative boronic acids (Scheme 3-4). The proposed oxidative Heck

cross-coupling of olefins and aryl boronic acids involves transmetalation, alkene coordination, β -hydride elimination, reductive elimination and metal reoxidation sequences^{36,37} (also see Chapter 1 Section 1.5.1). In order to follow the oxidative Heck cross-coupling in the active site of SBL, the various steps were first optimized individually. Our efforts for the optimization of these steps are described in Section 3.2.3.1 – 3.2.3.4. This enabled us to follow the reaction in a systematic manner using SBL-based metalloenzymes as discussed in section 3.2.3.5.

3.2.3.1 Establishing an LC-MS based methodology to monitor intermediates

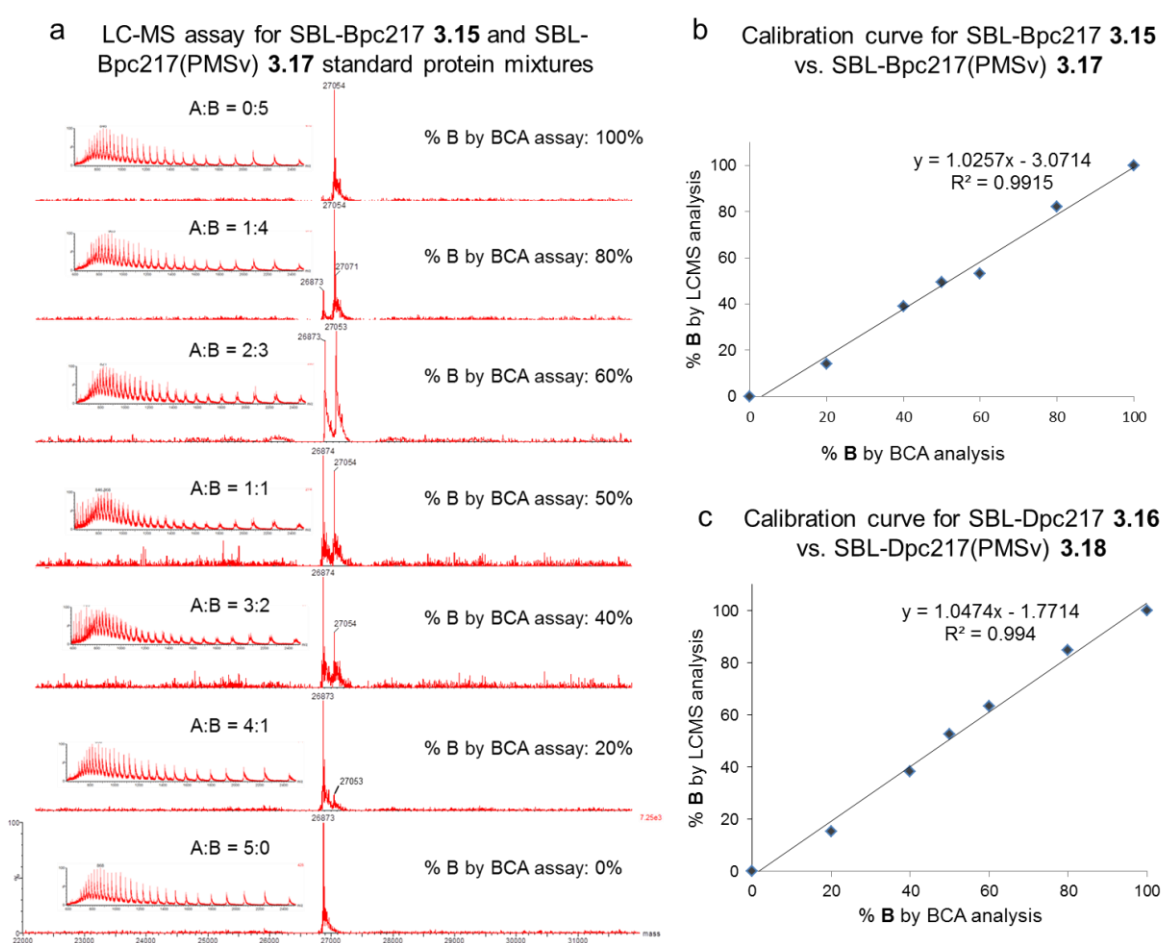
The substrate access and interaction of SBL-Bpc217 and SBL-Dpc217 was first assessed by reacting 20 equiv. of PMSv substrate **3.3** with these tailored enzymes (Scheme 3-11). Pleasingly, highly specific interaction of these substrates was confirmed with both enzymes resulting in the formation of enzyme-substrate adducts SBL-Bpc217(PMSv) **3.17** and SBL-Dpc217(PMSv) **3.18**. The loading of this substrate displayed similar selectivity as the SBL-WT, as described in Section 3.2.2.



Scheme 3-11. Enzyme-substrate interactions of SBL-Bpc217 and SBL-Dpc217 with alkene **3.3**

Having confirmed this highly selective interaction of the substrates with the tailored enzymes, a reliable methodology to monitor the target cross-coupling reaction was needed. For this, an LC-MS based assay was performed by comparing relative MS peak intensities arising from standard mixtures of SBL-Bpc217 **3.15** and substrate bound SBL-

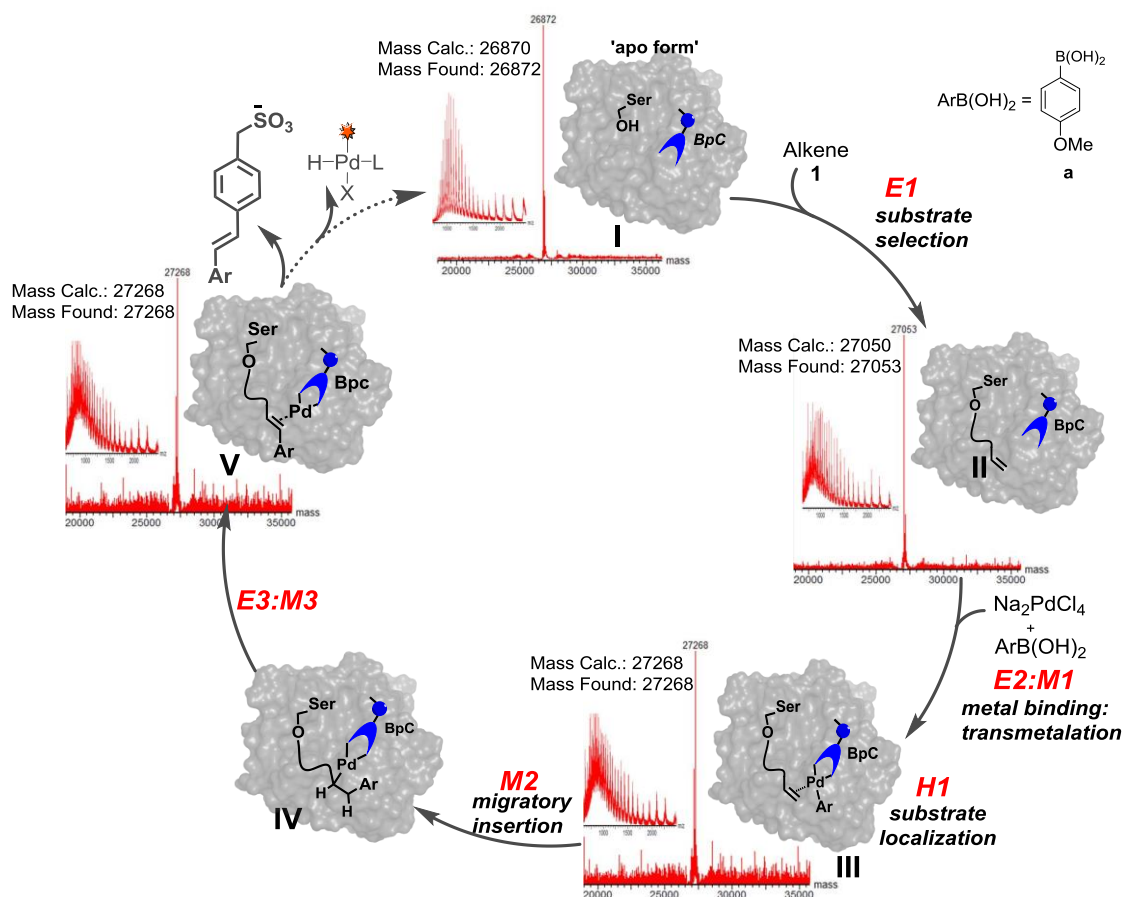
Bpc217(PMSv) **3.17** proteins (Scheme 3-12a). The resultant calibration curve indicated that both protein species ionised with similar efficiency (Scheme 3-12b). This confirmed that ESI-MS intensities provide a valid method to assess on-protein conversions for coupling reactions. Similar behaviour was observed for mixtures of proteins containing SBL-Dpc217 **3.16** and substrate bound SBL-Dpc217(PMSv) **3.18**, providing evidence in support of validity and applicability of this LC-MS assay for different protein species (Scheme 3-12c).



Scheme 3-12. (a) LC-MS assay to compare relative MS peaks of standard protein mixtures of SBL-Bpc217 and SBL-Bpc217(PMSv), (b) Calibration curve for SBL-Bpc217 and SBL-Bpc217(PMSv), (c) Calibration curve for SBL-Dpc217 and SBL-Dpc217(PMSv)

3.2.3.2 Oxidative Heck cross-coupling reaction using SBL-Bpc217•Pd

With these results in hand, cross-coupling reaction between alkene substrate **3.3** and arylboronic acid using SBL-Bpc217•Pd was explored. ESI-MS has previously been used to investigate the mechanism of oxidative Heck cross-coupling by detecting the different cationic-palladium complexes present in the solution for cross-coupling of small molecule substrates³⁶ (also see Chapter 1 Section 1.5.1). We envisaged that the enzyme- (E1, E2 etc.) and metal-mediated steps (M1, M2 etc.) can be monitored by LC-MS analysis for an oxidative Heck cross-coupling reaction between enzyme-bound alkene substrate **3.3** and *p*-methoxyphenylboronic acid substrate **a** (Scheme 3-13).



Scheme 3-13. Oxidative Heck cross-coupling reaction taking place in the active site of SBL-Bpc217•Pd

The cycle started with apoenzyme **I** (SBL-Bpc217) that served to both select and covalently bind the alkene substrate **3.3** giving enzyme-substrate SBL-Bpc217(PMSv) adduct **II** (E1 – substrate selection). This was followed by simultaneous addition of 1 equiv. of Pd²⁺ precursor and of the phenylboronic acid **a**. This second step involved specific metal binding (E2) by the Bpc/ Dpc moieties and the transmetalation of arylboronic acid by palladium (M1). A mass corresponding to the arylpalladium intermediate **III** was observed by the LC-MS analysis and hence the bisubstrate localization was facilitated in the enzyme active site. Once localised, the arylpalladium intermediate **III** would coordinate the alkene before the C-C bond formation step. We considered that this step corresponds to a hybrid step H1, where the alkene coordination to the localised palladium is ensured by the successful positioning of the alkene substrate in the binding pocket by the enzyme. The next step involved the key C-C bond formation (M2) to give species **IV** after the migratory insertion of the alkenyl substrate. This followed by β -hydride elimination (M3) to give enzyme-product intermediate **V** SBL-Bpc217(P)•Pd **3.19** (cross-coupled product referred to as 'P'). As a result of the dissociation of the product from palladium, a palladium hydride species is believed to be formed. The active Pd(II) catalyst can be regenerated by using the molecular oxygen as the reoxidant.^{36,38,39}

Excited with the trapping of the intermediate **V** that corresponds to the tethered cross-coupled product to the enzyme, it was realised further optimization to facilitate product release was required. At this stage, several methods to promote product release were considered, as discussed in the next section.

3.2.3.3 Hydrolysis of final cross-coupled product: optimization

In order to release the final cross-coupled product 'P', different methods to cleave off the alkane sulfonyl group from the enzyme were considered (Scheme 3-14). Initial optimizations of hydrolysis conditions were performed using SBL-WT(PMSv) protein.

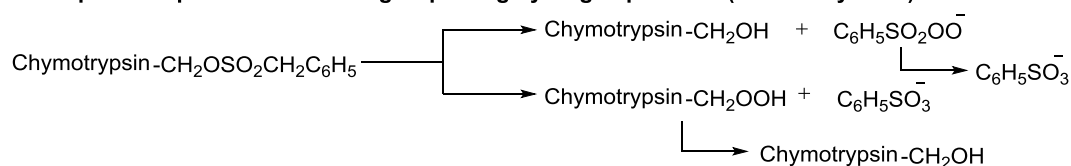
In 1973, Radomsky *et al.* reported nucleophilic displacement of the phenylmethanesulfonyl (PMS) group covalently bound to a serine protease (α -Chymotrypsin-PMS) using hydroperoxide/methylhydrogen peroxide (Scheme 3-14a).⁴⁰ The displacement of the PMSv group from SBL-WT(PMSv) was attempted with H₂O₂. While some hydrolysis of the PMSv group to give SBL-WT was detected, non-specific oxidation of other protein residues could not be discounted while using H₂O₂ as a nucleophile (See Experimental Section 3.5.8). Further, it has been previously shown that the oxidative inactivation of subtilisin can be caused by millimolar concentrations of H₂O₂ due to Met222 sulfoxidation in the active site (next to Ser221).⁴¹ Hence, this approach was not considered suitable for our purposes.

Polgar *et al.* have previously shown that an S_N2 displacement reaction at the active site serine residue using thiol acetate could convert subtilisin to its mercaptan form 'thio-subtilisin' (Scheme 3-14b).⁴² To test this route, SBL-WT(PMSv) **3.11** was incubated with 500 equiv. thiol acetate at 37 °C. We observed a mass peak corresponding to the serine-SBL-WT (26698 Da) by LC-MS analysis[‡] (See Experimental Section 3.5.8). This was perhaps due to hydrolysis of the sulfonyl group due to a pH change during the reaction (possibly due to contamination of reagent with acetic acid). Furthermore, this route was

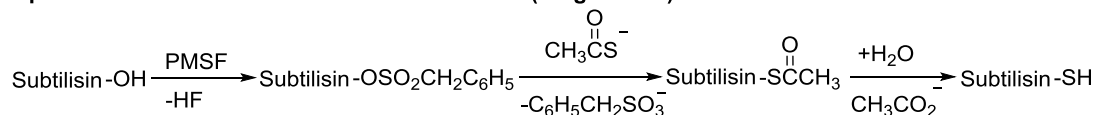
[‡] a mass corresponding to 26740 Da was also observed by LC-MS analysis

previously used to demonstrate chemical mutagenesis of the enzyme by turning the active site serine to a cysteine residue.⁴² Hence, this route was considered unsuitable.

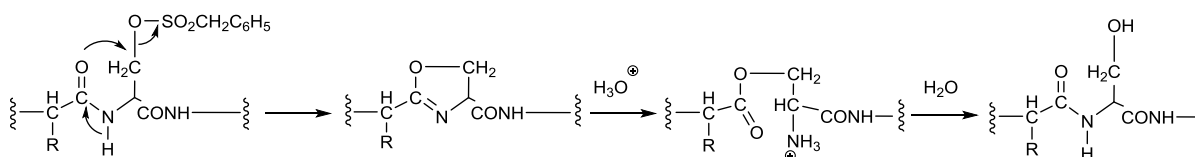
a Nucleophilic displacement of PMS group using hydrogen peroxide (Radomsky *et al.*)



b Reported chemical route to obtain thio-subtilisin (Polgar *et al.*)



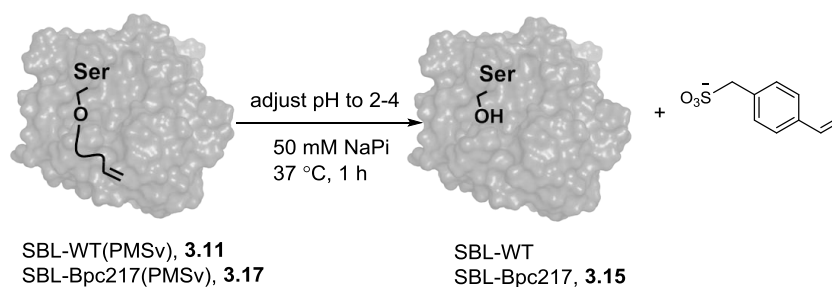
c Desulfonylation of PMS group in acidic conditions (Gold *et al.*)



Scheme 3-14. Optimization of conditions for chemical hydrolysis of the alkane sulfonyl group tethered to serine proteases

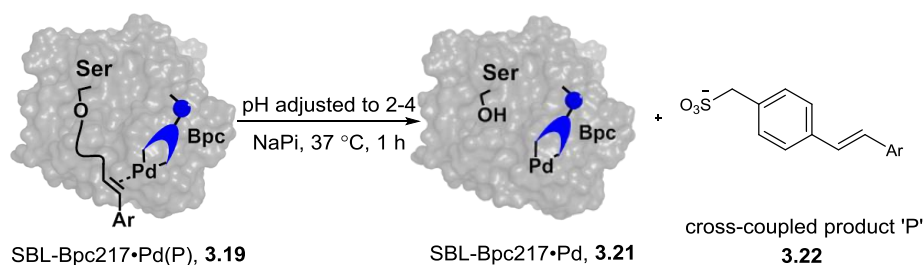
Encouraged by signs of release of PMSv group from SBL-WT(PMSv) under acidic conditions (as observed earlier while using thiol acetate as the nucleophile), we focussed our attention on facilitating desulfonylation in aqueous solutions by modulating pH conditions. Gold *et al.* reported the hydrolysis of alkane sulfonyl group on α -Chymotrypsin-PMS at pH above 8.5 and below 3 at 40 °C (Scheme 3-14c).⁴³ In this case, it was proposed that the desulfonylation proceeds by the formation of the oxazoline protein intermediates, followed by rapid hydrolysis to the *O*-acylserine derivative in acidic conditions.⁴³ At neutral/basic pH, an *O,N*-acyl shift would lead to the native peptide backbone structure (Scheme 3-14c). We screened different buffer systems over a range of pH (NaCl-HCl pH 2, Citrate buffer pH 3-5, Tris-HCl pH 8, Na₂CO₃ pH 10) using SBL-WT(PMSv) protein (Scheme 3-15). LC-MS analysis revealed rapid hydrolysis of the PMSv group from SBL-WT(PMSv) was possible under acidic conditions (See

Experimental Section 3.5.8 for details). When this acidic hydrolysis was attempted on SBL-Bpc217(PMSv) protein, in-solution hydrolysis could be successfully promoted by modulating pH to 2-4 by the addition of HCl (Scheme 3-15, see Experimental Section 3.5.9).



Scheme 3-15. Optimization of conditions for hydrolysis of tethered PMSv group by desulfonylation under acidic conditions

Going back to the SBL-Bpc217•Pd cycle (Scheme 3-13), chemical hydrolysis of cross-coupled product ‘P’ from intermediate **V** SBL-Bpc217(P)•Pd **3.19** was facilitated by carefully adjusting the pH to 2-4. On incubation of the reaction at 37 °C for 1 h, the holoenzyme SBL-Bpc217•Pd **3.21** could be detected successfully, after the release of cross-coupled product **3.22** (Scheme 3-16). The small molecule cross-coupled product **3.22** was subsequently isolated from the solution after the hydrolysis step (See Experimental Section 3.5.13). Characterization of **3.22** was accomplished by reverse phase HPLC-MS analysis and further confirmed by HRMS and MS/MS analyses.

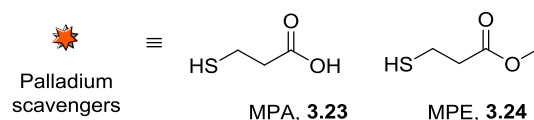


Scheme 3-16. Hydrolysis of tethered cross-coupled product ‘P’ by desulfonylation under acidic conditions from SBL-Bpc217(P)•Pd intermediate **3.19**

Further, to ensure the re-entry of the enzyme in the cycle *via* the apo form **I**, scavenging of the metal from species **VI** was also envisaged. More importantly, it was also noted that the mass difference between the cycle intermediates **III** or **V** is only 1 unit (after β -hydride elimination). Selective metal scavenging from these intermediates would allow accurate structural assignment of the mass observed by LC-MS analysis, as will be discussed in next section.

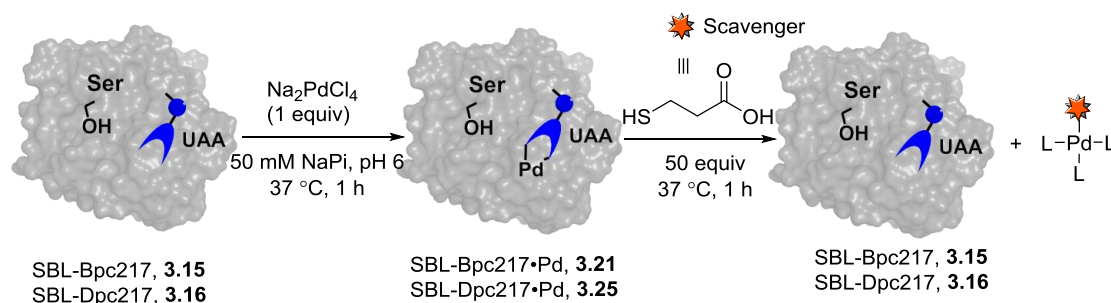
3.2.3.4 Palladium scavenging: optimization

In order to ascertain that the migratory insertion of the alkenyl substrate has taken place to form the C-C bond, we investigated methods to selectively scavenge palladium from the enzyme intermediate **V**. Previously, Pd^{2+} ions from aqueous and non-aqueous solutions were captured by using mercapto-propionic acid (MPA) **3.23** as scavenger, hinged onto magnetic beads for subsequent recovery.⁴⁴ By using this thiol-based scavenger MPA, non-specifically bound palladium to protein surface was successfully scavenged (presumably by forming six-membered chelates with MPA in its free form under the mildly basic conditions used).⁴⁵ In addition to this, we proposed that the corresponding ester form **3.24**, ie. 3-methyl, mercaptopropionate ester (MPE) would also be an effective palladium scavenger. More importantly, the pH of the reaction solution could be carefully controlled while using MPE **3.24** as scavenger. This would ensure that the tethered alkanesulfonyl group is not simultaneously hydrolysed under acidic conditions as seen in our earlier experiments. Hence both **3.23** and **3.24** were tested for scavenging palladium from different enzyme species (Scheme 3-17).



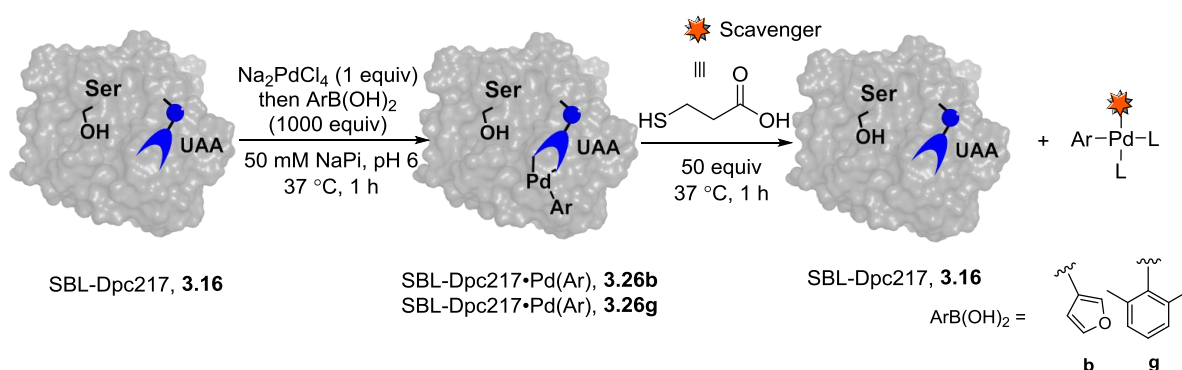
Scheme 3-17. Chemical structures of thiol-based palladium scavengers

To begin with, test reactions were carried out starting with SBL-Bpc217 which was first converted to the metal-containing holo-form SBL-Bpc217•Pd (**3.21**). Scavenging of palladium from metal-containing SBL-Bpc217•Pd was attempted with 50 equiv. of MPA **3.23** (Scheme 3-18). Upon incubating the protein solution for 1 h at 37 °C, the scavenging of palladium was complete as full conversion to SBL-Bpc217 **3.15** was observed. Similar results were obtained when metal removal from SBL-Dpc217•Pd **3.25** was carried out, thus showing that palladium scavenging was feasible and can be monitored by appropriate shift in mass by LC-MS analysis (See Experimental Section 3.5.10).



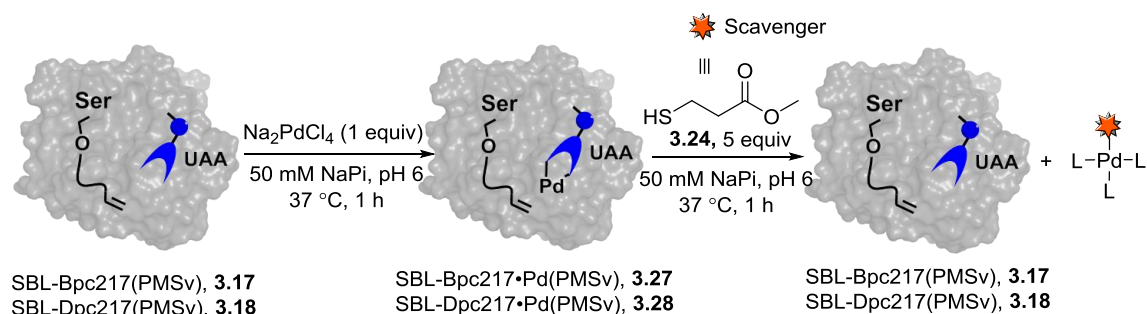
Scheme 3-18. Test reaction for palladium scavenging for SBL-Bpc217•Pd and SBL-Dpc217•Pd

To test whether distinction between **III** and **V** from the cycle was indeed possible by removing metal selectively from these enzyme species, initial tests were performed with SBL-Dpc217•Pd(Ar) adduct **3.26** (Scheme 3-18). For this, first the SBL-Dpc217 **3.16** was treated with one equiv. of Pd²⁺ source (Na₂PdCl₄), followed by subsequent addition of arylboronic acid substrate (**b** or **g**). Gratifyingly, transmetalation adduct **3.26** was observed upon incubation for 1 h at 37 °C. On addition of MPA scavenger **3.23**, the palladium along with the aryl group was removed and SBL-Dpc217 **3.16** was recovered (See Experimental Section 3.5.11).



Scheme 3-18. Test reaction for palladium scavenging for SBL-Dpc217•Pd(Ar) adduct using MPA

Finally, we focussed our efforts on optimizing conditions for selective scavenging from species **V** without any hydrolysis of alkane sulfonyl group as the side reaction. Because of acidic nature of MPA scavenger **3.23**, we decided to use the corresponding ester MPE (**3.24**). Test reaction with substrate bound SBL-Bpc217•Pd(PMSv) species **3.27** confirmed palladium scavenging was possible by using just 5 equiv. of MPE **3.24** as metal scavenger (Scheme 3-19). On incubating the reaction at 37 °C for 1 h, complete scavenging of palladium was observed from the protein species giving SBL-Bpc217(PMSv) **3.17**. As proposed, no hydrolysis of the bound PMSv substrate was observed from the above protein species when using MPE as scavenger, as judged by LC-MS analysis. Similar results were obtained when SBL-Dpc217•Pd(PMSv) **3.28** was used for metal scavenging test reaction (See Experimental Section 3.5.12).



Scheme 3-19. Test reaction for palladium scavenging for SBL-Bpc217•Pd(PMSv) and SBL-Dpc217•Pd(PMSv) adduct using MPE

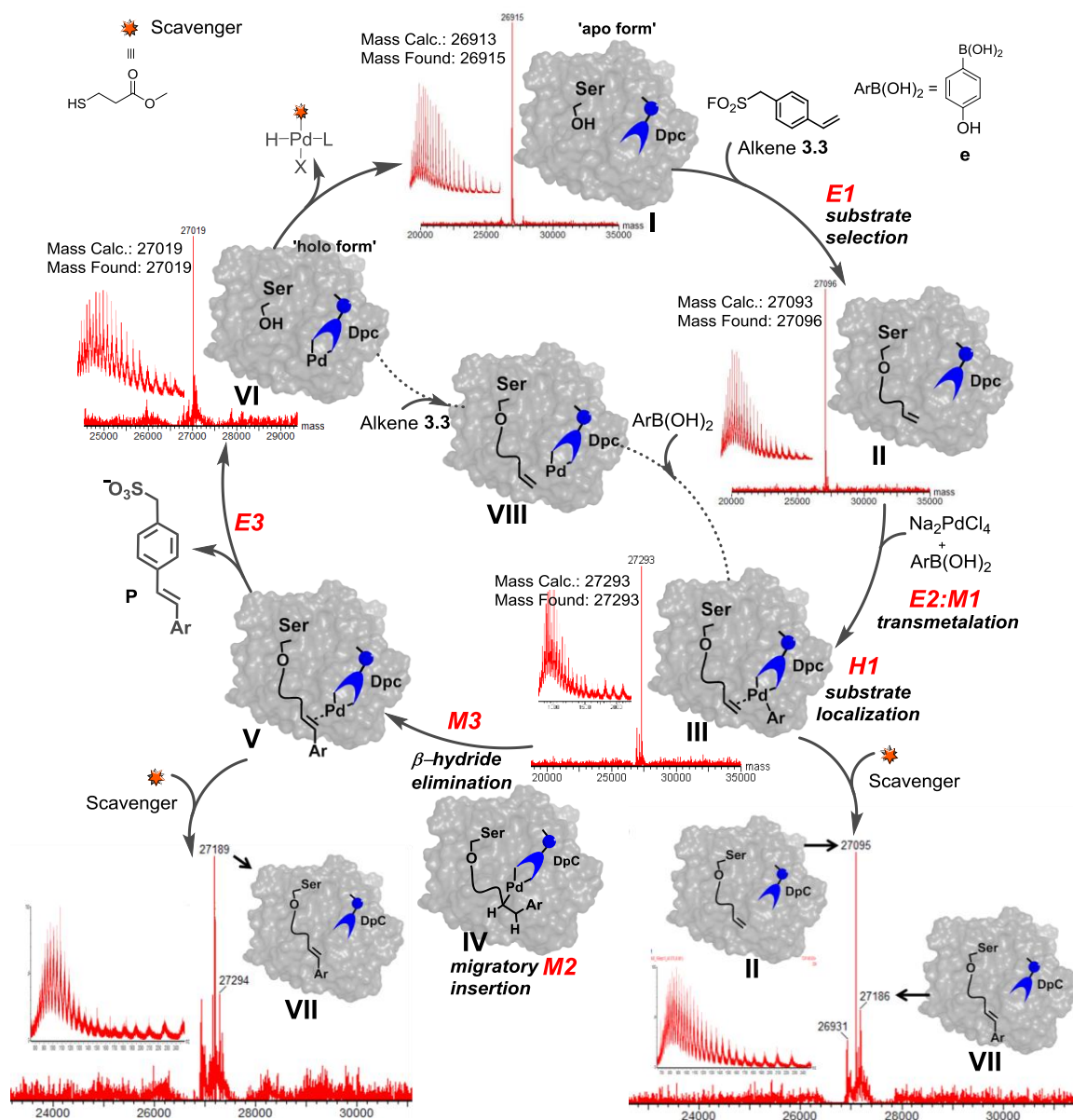
3.2.3.5 Oxidative Heck cross-coupling reaction using SBL-Dpc217•Pd

Having established product release and metal scavenging procedures, we employed the previously engineered SBL-Dpc217 enzyme in an oxidative Heck cross-coupling reaction between the tethered alkene **3.3** and *p*-hydroxyboronic acid substrate **e** (Scheme 3-20).

Substrate selection by the enzyme binding pocket (E1 step) resulted in the formation of enzyme-substrate SBL-Dpc217(PMSv) adduct **II**. Upon the addition of 1 equiv. of Pd²⁺ precursor and 1000 equiv. of the boronic acid substrate, transmetalation of arylboronic acid (M1) and palladium binding (E2) with the enzyme facilitated bisubstrate localization (hybrid step H1) as discussed previously. This hybrid step proceeds by productive binding of the substrate with the metal (alkene coordination) as can be seen by the trapping of intermediate **III**, as observed by the LC-MS analysis. Next, the C-C cross-coupling step proceeds via migratory insertion of aryl group (M2) leading to formation of species **IV**, followed by β -hydride elimination (M3) that leads to the formation of intermediate **V** (Scheme 3-20).

Importantly, the distinction between species **III** and **V** was accomplished by selective scavenging of palladium using 3-methyl, mercaptopropionate ester (MPE), which yielded the enzyme-alkene (species **II**) or the enzyme-product (species **VII**) respectively. The addition of 5 equiv. of MPE scavenger to the protein solution after 45 min resulted in quenching of the reaction by the removal of metal catalyst. At this stage, SBL-Dpc217(PMSv) adduct (enzyme species **II**) was obtained as the major enzyme species along with formation of 26% of the cross-coupling product. Crucially, intermediate **VII** SBL-Dpc217(P) (SBL-Dpc217 containing tethered product) was observed upon palladium scavenging on completion of the reaction after 2.5 h (See Scheme 3-20). 77% mass

corresponding to SBL-Dpc217(P) and 23% SBL-Dpc217 was observed (perhaps after hydrolysis of final product and oxidation), as confirmed by LC-MS analysis (Experimental Section 3.5.22 for details).



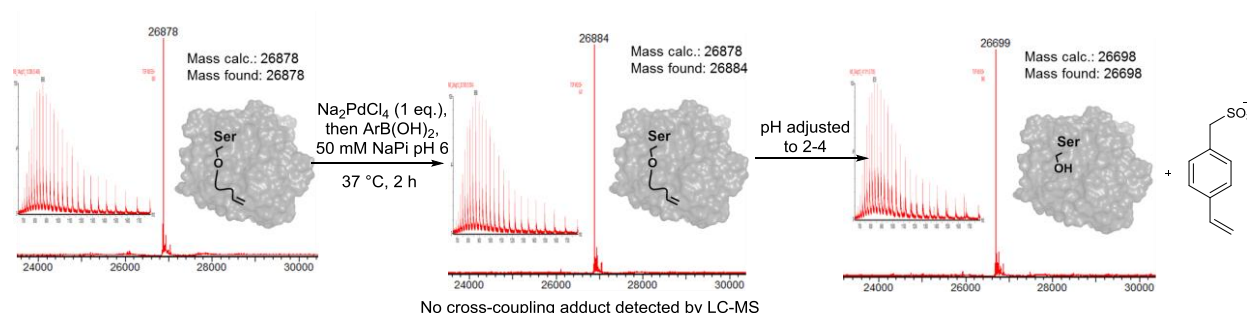
Scheme 3-20. Oxidative Heck cross-coupling reaction taking place in the active site of SBL-Dpc217•Pd; the various steps in the biotransformation pathway facilitated by the artificial metalloenzyme are depicted namely, E1 - substrate selection, E2 - specific metal binding, M1 - transmetalation, hybrid step H1 - bisubstrate localization, M2 - migratory insertion, M3 - β hydride elimination, E3 - product release

Having successfully convoluted the metal and enzyme steps that allowed for the direct observance of **V**, we next enabled the controlled release of product 'P'. Advantageously, the on-protein hydrolysis of sulfonyl esters can be precisely modulated by pH alteration based on our previous results. Thus, we devised this final hydrolysis step to release the coupled product **P** from species **V** by simply altering the buffer conditions, with concomitant formation of holo-enzyme **VI** (Scheme 3-20). By adjusting the pH to 2-4, the final coupled product **3.22e** was obtained and its identity was confirmed by reverse phase HPLC-MS analysis and further confirmed by HRMS and MS/MS analyses (See Experimental Section 3.5.15, 3.5.16). Further reloading of the enzyme using different strategies has been discussed in section 3.2.5.

3.2.4 Control reactions

Control reactions for the cross-coupling of alkene **3.3** and arylboronic acid using SBL-L217C or SBL-WT that lack metal binding amino acids Bpc or Dpc, gave negative results (Scheme 3-21). No formation of either intermediate **V** by LC-MS analysis after 2 h of the reaction at 37 °C or final coupled product after the hydrolysis step was observed under the same conditions (See Experimental Section 3.5.17). This confirmed that the presence of chelating UAAs (Dpc or Bpc) is critical for the localisation of the metal inside the enzyme. When the reaction was prolonged for 6 h, only 26% SBL-WT(P) **3.30** cross-coupled product tethered to the enzyme was observed after incubation at 37 °C. This confirmed that in presence of nucleophilic enzyme and metal at the active site, the cross-coupling reaction is faster than the background reaction due to free Pd²⁺ catalyst in the solution. The unnatural amino acids (Dpc or Bpc) along with other protein residues provide for specific chelation site for the metal that helps in facilitating the reaction. In this case, the negative

control reaction with SBL-L217C for the cross-coupling of alkene **3.3** and arylboronic acid shows that in our system Cys217 is also not a good ligand for Pd²⁺ species.



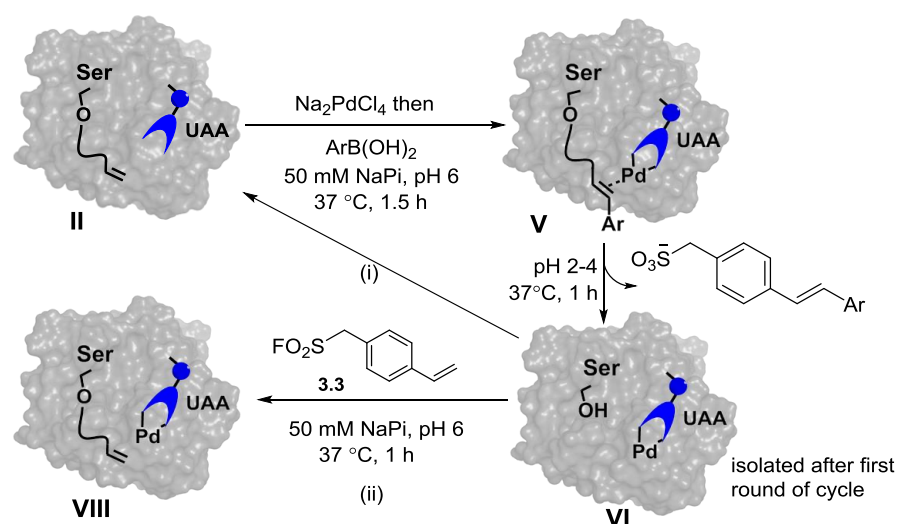
Scheme 3-21. Oxidative Heck cross-coupling reactions of SBL-WT(PMSv) with 4-hydroxyphenyl boronic acid substrate **d** followed by hydrolysis of the tethered alkanesulfonyl group (control reaction)

In another control reaction, no cross-coupling was seen in the absence of metal catalyst when both coupling partners were incubated with the enzyme. Further, the possibility of Pd-catalysed coupling of boronic acid to another non-specific protein site was also ruled out by a negative control reaction in the absence of any olefinic substrate, during which no aryl adduct was observed by LC-MS in the absence of alkene moiety (See Experimental Section 3.5.17).

3.2.5 Reloading

Having demonstrated the biotransformation pathway for oxidative Heck cross-coupling reaction using SBL-Bpc217•Pd and SBL-Dpc217•Pd, we next investigated the reloading strategies to access further rounds of cycle. First, the reloading of substrate **3.3** for a second round of the cycle was attempted *via* intermediate **VIII**, as depicted in the cycle (Scheme 3-22). The loading of the alkene substrate with 5-10 equiv. was found unsuccessful at first, as no enzyme-substrate adduct was detected by LC-MS analysis. It was speculated that the ‘pH toggle’ introduced in the first round of the cycle may be the cause for this reduced selectivity of the serine protease (most likely due to enzyme

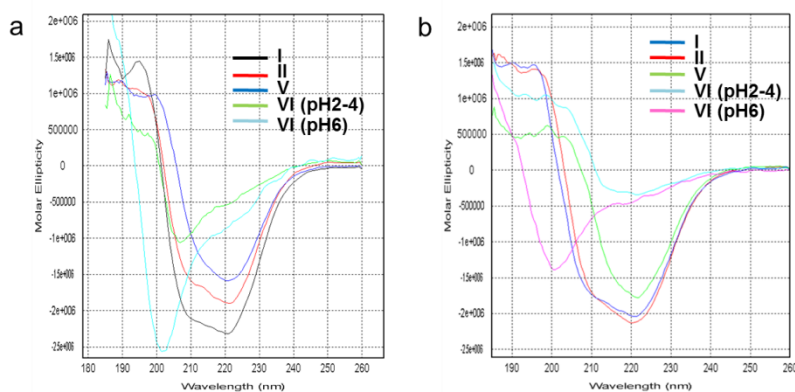
denaturation under acidic conditions). With the hope that the protein may still have retained some nucleophilicity due to robust nature of SBL, we increased the substrate concentration and were pleased to see some substrate loading in the SBL-Dpc217•Pd enzyme, to give intermediate **VIII** (SBL-Dpc217•Pd(PMSv)). However, loading of the second batch of substrate **3.3** proved most successful for SBL-Bpc217•Pd, albeit with concomitant scavenging of palladium by excess alkene in the solution (Scheme 3-22). Treatment of species **VI** (SBL-Bpc217-Pd) with 50 equiv. substrate **3.3** allowed the successful reloading of alkene to give SBL-Bpc217(PMSv) (**II**), where the palladium was removed from the protein due to excess alkene in the solution.



Scheme 3-22. Reloading of alkene substrate **3.3** on (i) SBL-Bpc217•Pd and (ii) SBL-Dpc217•Pd isolated after first round of cycle

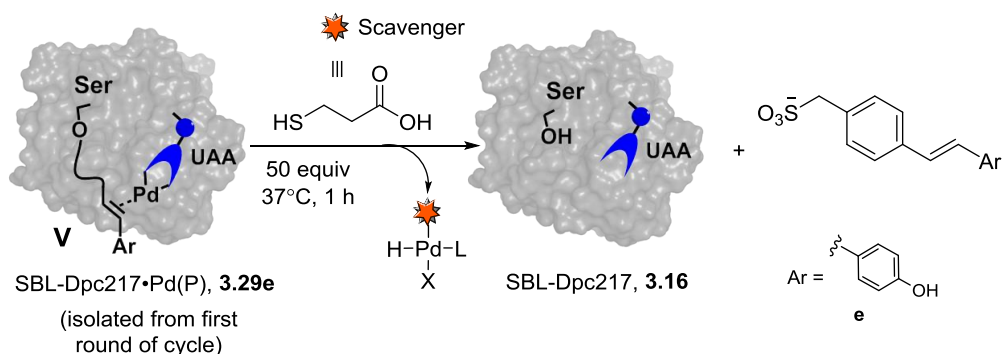
Intrigued by the findings from our reloading trials, we further investigated changes in the secondary structure of the subtilisin during and after the first round of the cycle. The ‘pH toggle’ used to facilitate product release (due to extremely slow release of alkane sulfonyl substrates from the enzyme) is an evident bottleneck in our approach. This was confirmed by far-UV circular dichroism (CD) spectra for the engineered cycle with SBL-Dpc217•Pd and SBL-Bpc217•Pd. Systematic monitoring of far-UV CD after each step from the cycle

revealed significant overall changes in the secondary structure after introducing the ‘pH toggle’ (Scheme 3-23).



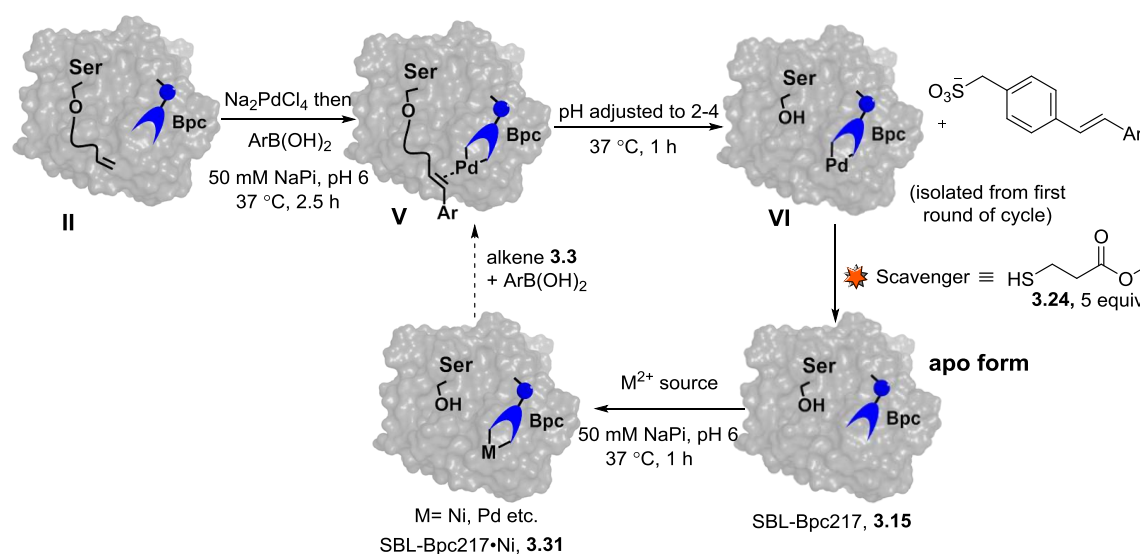
Scheme 3-23. Far-UV CD spectra of various intermediates from the cycle (a) for SBL-Dpc217•Pd (b) for SBL-Bpc217•Pd; intermediates **I–VI** from the engineered cycle (Scheme 3-20) were analysed

Interestingly, simultaneous hydrolysis of the final cross-coupled product ‘P’ and scavenging of palladium from the species **V** in one step also proved possible by using an acidic scavenger (Scheme 3-24). By directly introducing a pH change caused by the addition of 3-mercaptopropionic acid (MPA), apo-enzyme **I** (SBL-Dpc217) was isolated and the cross coupled product was cleaved off from the enzyme (See Experimental Section 3.5.20).



Scheme 3-24. Simultaneous hydrolysis and palladium scavenging from species **V** from the cycle using MPE as the scavenger

On the other hand, we investigated if selective palladium scavenging from enzyme species **VII** (SBL-Bpc217•Pd) using MPE that gave apo-enzyme **I**, could be reloaded with another potential metal catalyst (such as nickel). The apoenzyme species **I** obtained in this way was successfully reloaded with 1 equiv. Ni^{2+} precursor (Scheme 3-25). On incubating the apo-enzyme SBL-Bpc217 (isolated after the first round of the cycle) with 1 equiv. of $\text{Ni}(\text{OAc})_2$ for 1 h at 37 °C, the corresponding holo-form SBL-Bpc217•Ni **3.31** could be obtained as confirmed by LC-MS analysis (See Experimental Section 3.5.21). This suggests this enzymatic catalyst can potentially be re-used, thereby making a wide variety of Pd/Ni-catalysed reactions accessible. Unfortunately, further substrate reloading on this nickel-based metalloenzyme with 5 equiv. of alkene **3.3** was unsuccessful, perhaps due to denaturation of the enzyme after several chemical steps in the first cycle.

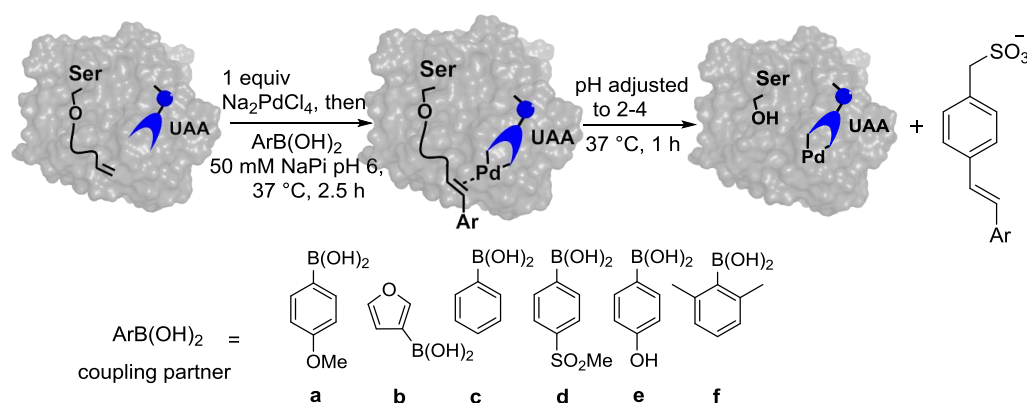


Scheme 3-25. Reloading of species **VI** isolated from the first cycle with a different metal via the apoenzyme

3.2.6 Substrate scope of oxidative Heck cross-coupling reaction

The selectivity of the S1 pocket to harbour hydrophobic substrates was successfully demonstrated in Section 3.2.2. On the other hand, localisation of a range of arylboronic acids was enabled due to broader specificity of the S1' pocket and subsequent palladation,

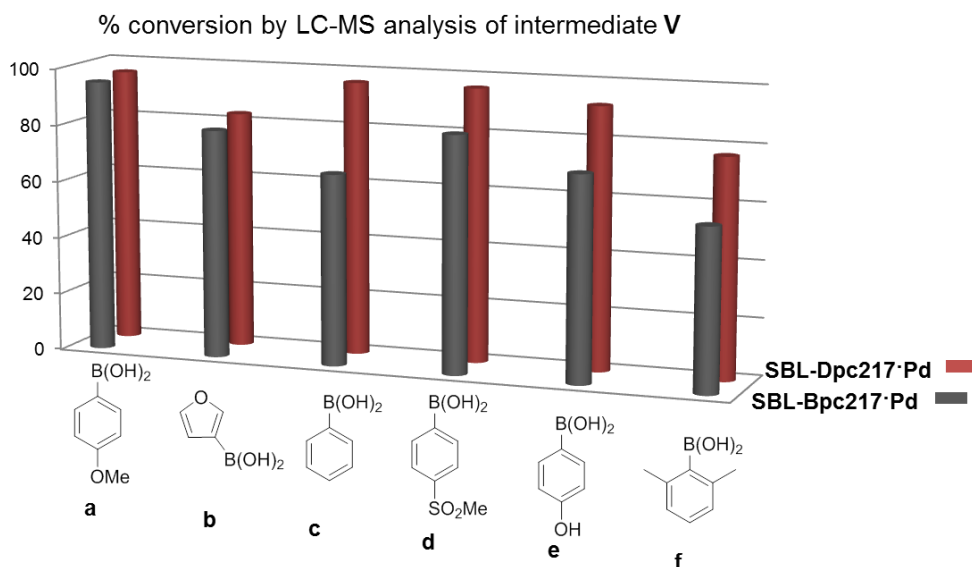
reportedly proceeding *via* transmetalation to oxopalladium species.²⁶ While exploring the substrate breadth, various (hetero)aromatic boronic acids with different substituents were used as coupling partners (Scheme 3-26).



Scheme 3-26. Oxidative Heck cross-coupling reaction for alkene **3.3** and different arylboronic acids using SBL-based metalloenzymes

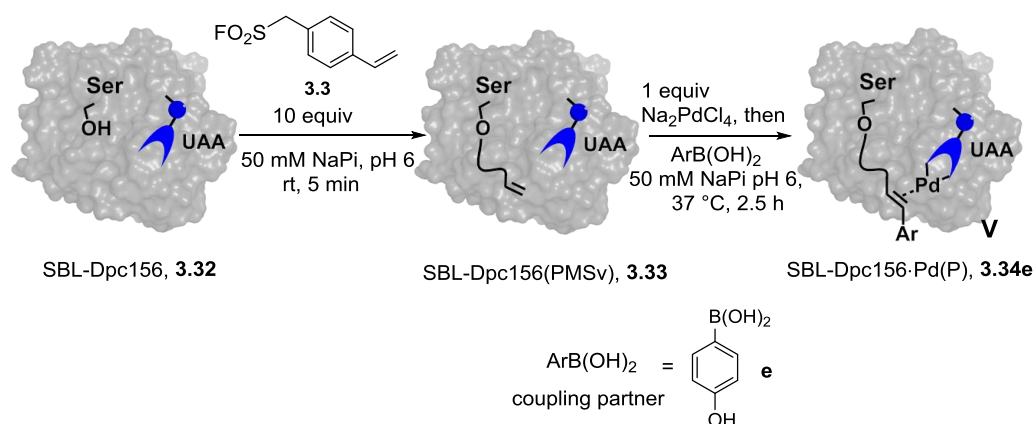
Transmetalation of several arylboronic acids and subsequent oxidative Heck cross-coupling product tethered to the enzyme was observed using engineered SBL enzymes. Notably, on-protein conversions could be directly determined from these observed species **V** by LC-MS using calibration curves previously derived by using synthesized standard protein samples (Section 3.2.3 and Experimental Section 3.5.13-14). The enzyme-product adduct **V** was detected by LC-MS for a series of boronic acid substrates **a-f** in good to excellent on-protein conversions for both SBL-Bpc217•Pd and SBL-Dpc217•Pd (Scheme 3-27). Relative conversions of SBL-Dpc217 vs. SBL-Bpc217 revealed different selectivities by the two UAAs perhaps due to different modes of metal binding. The reactivities of different aromatic boronic acids also differed perhaps owing to their electronic or steric properties. The two-dimensional activity plot constructed in this manner suggests scope for subsequent screening and optimization of these designed artificial metalloenzymes. The chemical optimization involves using different metal

binding moieties (such as Bpc and Dpc) that can be easily introduced using our tag-and-modify approach (Chapter 2 Section 2.2.1).



Scheme 3-27. Plot demonstrating substrate scope for oxidative Heck cross-coupling reaction of alkene **3.3** with various arylboronic acids using SBL-Bpc217•Pd and SBL-Dpc217•Pd

To analyse the effect of changing the site of installation of metal binding amino acid, we decided to use another SBL variant that contains the Dpc at position 156 (other than at position 217). To test this, previously synthesized SBL-Dpc156 was considered suitable as it contains an installed metal binding site on a flexible loop in the S1 pocket (Scheme 3-28). We were pleased to see that localisation of alkene substrate **3.3** was achieved by using 10 equiv. of the substrate (the presence of UAA on a flexible surface exposed loop in the same S1 pocket did not preclude the binding of the alkene substrate in the S1 pocket). Moreover, subsequent oxidative Heck cross-coupling was also accomplished when one equiv. of Pd²⁺ precursor was added along with model phenylboronic acid **e**, when the mass corresponding to SBL-Dpc156•Pd(P) (species **V** from the cycle) was observed by LC-MS analysis (See Experimental Section 3.5.22).



Scheme 3-28. Oxidative Heck cross-coupling reaction for alkene **3.3** and arylboronic acid using engineered SBL-Dpc156•Pd

3.2.7 Investigation of the pH dependence of cross-coupling reaction

Further, different buffer systems and pH conditions were screened for the enzyme-mediated oxidative Heck cross-coupling reaction between alkene **3.3** and arylboronic acid to validate the effect of different pH on the target cross-coupling reaction. The following buffer systems at different pH were investigated:

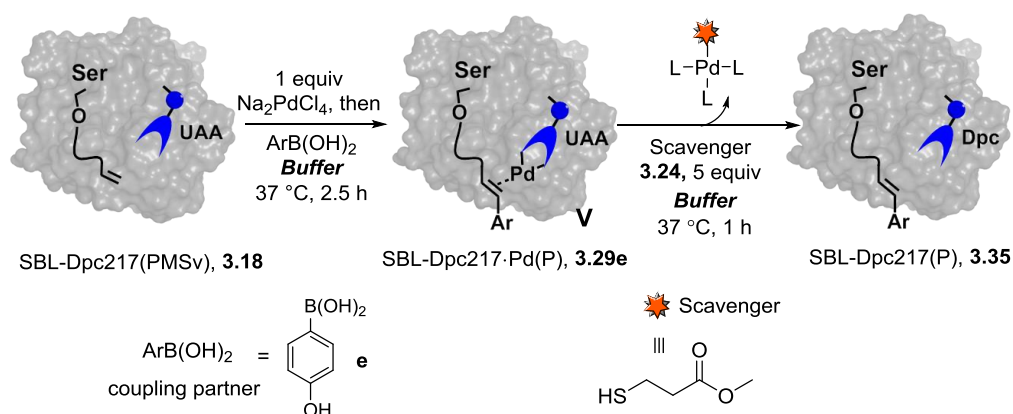
50 mM sodium acetate buffer (NaOAc), pH 5.5;

50 mM sodium phosphate buffer (NaPi), pH 6;

50 mM 3-(*N*-morpholino)propanesulfonic acid buffer (MOPS), pH 7;

50 mM sodium phosphate buffer (NaPi), pH 8.

The reaction started with SBL-Dpc217 enzyme that serves to uptake alkenyl substrate **3.3** (Scheme 3-29). On addition of 1 equiv. of palladium precursor and arylboronic acid, on-protein cross-coupling was detected as could be seen by the formation of species **V** containing the tethered product. It was found that the cross-coupling was independent of the nature of the buffer system or the pH based on the observance of the intermediate **V** from the engineered cycle.



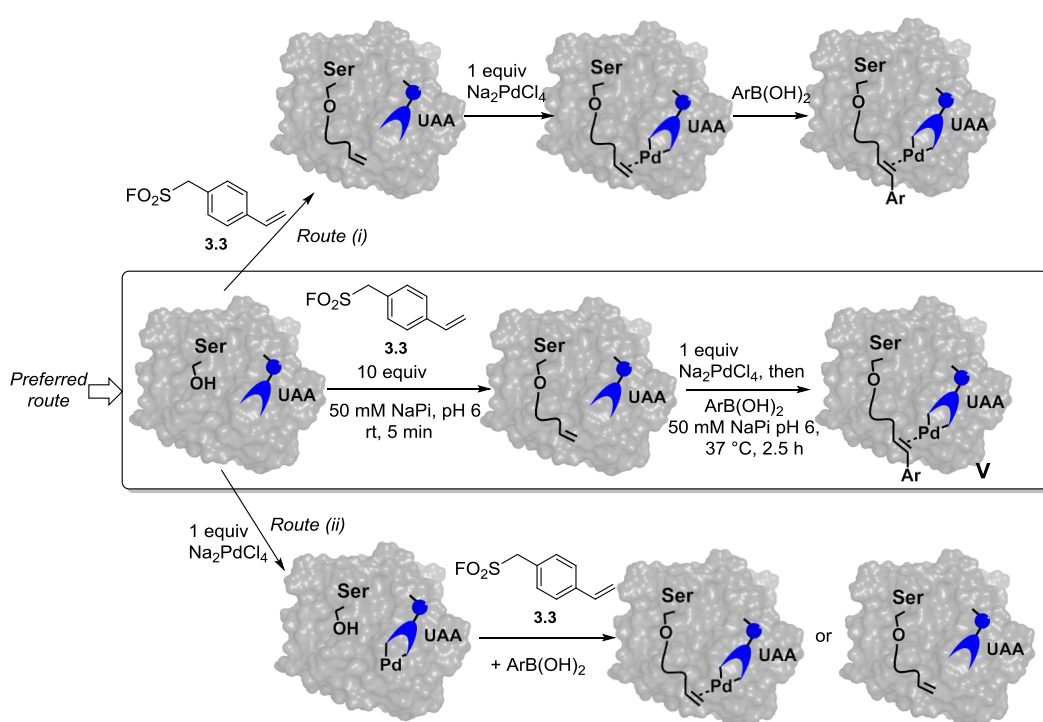
Scheme 3-29. Oxidative Heck cross-coupling reaction for alkene **3.3** and arylboronic acid using engineered SBL-Dpc217•Pd to investigate effect of different buffer systems

Further, full conversion to the tethered cross product adduct was accomplished by selective metal scavenging on incubating the reaction at 37 °C after 2.5 h; species **VI** was obtained from reactions in all four buffer systems. Some amount of species **I** was seen too, perhaps due to simultaneous hydrolysis and scavenging (See Experimental Section 3.5.23 for details).

3.2.8 Compulsory order of palladation and on-protein cross-coupling reaction

In order to demonstrate the on-protein cross-coupling and estimate conversions for the first round of cycle, we discovered that it was important to optimize stoichiometric reaction of the enzyme with the alkene substrate **3.3**. This was because the excess alkene in the solution could also coordinate metal and was found to interfere with the specific complexation of palladium with engineered enzymes SBL-Dpc217 and SBL-Bpc217 for the subsequent reaction (preferred route in Scheme 3-30). This was additionally ensured by removing any excess alkene substrate in the reaction by desalting or dialysis (only slight excess of alkene was added to ensure complete uptake of alkene substrate as monitored by LC-MS). Other alternative routes were also explored. For example, addition of alkene substrate to SBL-Dpc217(PMSv) followed by complexation of palladium (after removing

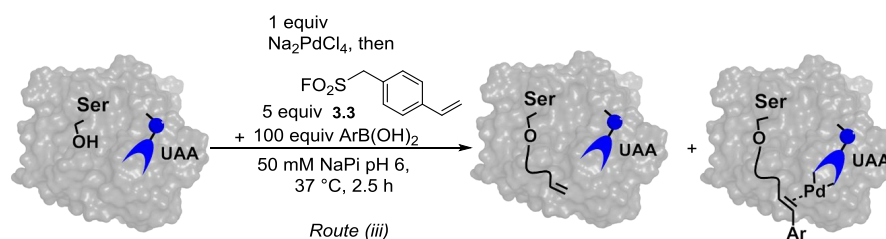
excess alkene from the solution) led to species SBL-Dpc217•Pd(PMSv), which was subsequently found to load the second arylboronic acid substrate (Route (i) in Scheme 3-30). However, the cross-coupling reaction was found to proceed much slower in this case (even after 4 h only <20% product formation was detected by LC-MS). Alternatively, simultaneous addition of both the coupling partners to SBL-Dpc217•Pd or SBL-Bpc217•Pd, only gave adducts SBL-Dpc217•Pd(PMSv) or SBL-Bpc217(PMSv) respectively with no formation of the cross-coupled product (Route (ii) in Scheme 3-30).



Scheme 3-30. Establishing compulsory order for oxidative Heck cross-coupling reaction for alkene **3.3** and arylboronic acid using engineered SBL-enzymes

As a third option, simultaneous addition of metal catalyst and the cross-coupling partners was also investigated with SBL-Dpc217. While the uptake of alkene substrate by the enzyme was found to be successful with SBL-Dpc217(PMSv) detected as the major adduct by LC-MS, at best only 20% of on-protein cross-coupled product formation was detected

using route (iii) in Scheme 3-31. In this case, it is likely that the scavenging of Pd^{2+} ions by excess alkene substrate may have precluded full conversion to the cross-coupled product.



Scheme 3-31. Establishing compulsory order for oxidative Heck cross-coupling reaction for alkene **3.3** and arylboronic acid using engineered SBL-enzymes

3.3 Conclusions

In this chapter, we used previously synthesized artificial metalloenzymes, SBL-Dpc217•Pd and SBL-Bpc217•Pd to demonstrate cross-coupling reaction in the active site of SBL. Serine protease such as SBL was used as a model system to gain leverage from the natural substrate selectivity of its binding sites and enhanced nucleophilicity of its catalytic machinery. In this proof-of-concept study, oxidative Heck cross-coupling reaction of an alkenyl substrate and phenylboronic acid was chosen as the model reaction.

First, enzyme selectivity was demonstrated by exploring a set of several substrates differing in their molecular shape and volume, which were screened using an ESI-MS based competition assay. A similar semi-quantitative ESI-MS assay has been previously used to analyse tightly bound non-covalent protein-DNA⁴⁶ and protein-ligand composites.⁴⁷ Using this, a general trend of substrate selectivity for benzyl>phenyl>alkyl substrates was established. Notably, the striking discrimination between substrates **3.3** and **3.8**, only differing in methylene linker, indicates that the enzyme not only serves to recognise the sulfonyl moiety but there is also significant effect of shape complementarity

in our system. We decided to exploit these characteristic structural features of a natural enzyme pocket to carry out a metal catalysed cross-coupling reaction.

Owing to the natural substrate specificity of the enzyme, we demonstrate oxidative Heck cross-coupling reaction^{27,48,49} by merging two distinct catalytic mechanisms, i.e. nucleophilic activation derived from the enzyme (SBL) and the migratory redox activity of a non-native metal (palladium). The various steps of the biotransformation pathway of this ‘Heck-ase’ have been studied closely by trapping different reactive enzyme intermediates. The first coordination sphere interactions of the metal localised inside the enzyme were studied in a systematic manner and the role of the ‘intrinsic’ metal in the cross-coupling reaction was highlighted by the observance of reaction intermediates using mass spectrometry. Interweaving of the various steps for nucleophilic enzyme catalysis (E1, E2, E3) with metal catalysis (M1, M2, M3) necessitated both convolution of these steps and the creation of a hybrid step (H1) leading to a synergistic bio-metallic mechanism (see cycle).

The presence of a chelating amino acid (Bpc or Dpc) was an important determinant for site-specific palladium complexation and metal activity. Through the rigorous reaction monitoring and step-wise study of the complex cycle, a guided approach can be taken for proper ligand design, location and accessibility of the metal and optimal reaction conditions. While the coupling reaction was carried out in 50 mM NaPi, pH 6 under mild conditions at 37 °C, the release of cross-coupled product and regeneration of the metalloenzyme **V** required a ‘pH toggle’. In this way, we were able to also demonstrate different methods for successful entry into a second round of the cycle. Moreover, a

selective metal scavenging technique was highlighted to access different reloading pathways.

The methodology described here converted a ‘washing-powder enzyme’ such as SBL into a cross-coupling biocatalyst and could serve as a prototypical design for future design of specific and functional metalloenzymes. The engineered reaction pathway shown in **Scheme 3-20** gives a snapshot of the catalytic cycle depicting interlaced nucleophilic catalysis (by activated serine hydrolase) and migratory redox catalysis (by Pd^{II}), serving as a roadmap for further screening of substrates. Next, it is tempting to characterize the local environment around the metal centre through X-ray crystallography and even crystallise the intermediates of the engineered cycle to gain mechanistic insights, as will be attempted in Chapter 4. Although the first signs of the desired catalytic turnover were observed, the choice of modified PMSF-based substrates used here has resulted in the covalent tethering of substrates and slow hydrolysis to release of final cross-coupled product. In this context, covalent yet reversible interactions of the ester/amide substrates with the serine proteases could be exploited to realise multi-turnover cross-coupling cycle. We believe such a substrate localisation approach based on high specificity of enzymes could be a powerful strategy for biocatalysis of important organic reactions, in particular bisubstrate cross-coupling reactions.

3.4 References

- (1) Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. *Science* **2012**, 338, 500.
- (2) Lo, C.; Ringenberg, M. R.; Gnanndt, D.; Wilson, Y.; Ward, T. R. *Chem. Commun.* **2011**, 47, 12065.
- (3) Pierron, J.; Malan, C.; Creus, M.; Gradinaru, J.; Hafner, I.; Ivanova, A.; Sardo, A.; Ward, T. R. *Angew. Chem. Int. Ed.* **2008**, 47, 701.
- (4) Roelfes, G.; Feringa, B. L. *Angew. Chem.* **2005**, 117, 3294.

- (5) Boersma, A. J.; Feringa, B. L.; Roelfes, G. *Angew. Chem.* **2009**, *121*, 3396.
- (6) Coquière, D.; Feringa, B. L.; Roelfes, G. *Angew. Chem.* **2007**, *119*, 9468.
- (7) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. *Angew. Chem.* **2012**, *124*, 7590.
- (8) Mayer, C.; Gillingham, D. G.; Ward, T. R.; Hilvert, D. *Chem. Commun.* **2011**, *47*, 12068.
- (9) Matsuo, T.; Imai, C.; Yoshida, T.; Saito, T.; Hayashi, T.; Hirota, S. *Chem. Commun.* **2012**, *48*, 1662.
- (10) Coquière, D.; Bos, J.; Beld, J.; Roelfes, G. *Angew. Chem. Int. Ed.* **2009**, *48*, 5159.
- (11) Reetz, M. T. *The Chemical Record* **2012**, *12*, 391.
- (12) Jing, Q.; Kazlauskas, R. J. *ChemCatChem* **2010**, *2*, 953.
- (13) Bertucci, C.; Botteghi, C.; Giunta, D.; Marchetti, M.; Paganelli, S. *Adv. Synth. Catal.* **2002**, *344*, 556.
- (14) Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**, *100*, 306.
- (15) Ward, T. R. *Acc. Chem. Res.* **2011**, *44*, 47.
- (16) Köhler, V.; Wilson, Y. M.; Dürrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knörr, L.; Häussinger, D.; Hollmann, F.; Turner, N. J.; Ward, T. R. *Nat Chem* **2013**, *5*, 93.
- (17) Skander, M.; Humbert, N.; Collot, J.; Gradinaru, J.; Klein, G.; Loosli, A.; Sauser, J.; Zocchi, A.; Gilardoni, F.; Ward, T. R. *J. Am. Chem. Soc.* **2004**, *126*, 14411.
- (18) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *CHIMIA International Journal for Chemistry* **2002**, *56*, 721.
- (19) Panella, L.; Broos, J.; Jin, J.; Fraaije, M. W.; Janssen, D. B.; Jeronimus-Stratingh, M.; Feringa, B. L.; Minnaard, A. J.; de Vries, J. G. *Chem. Commun.* **2005**, 5656.
- (20) Haquette, P.; Talbi, B.; Canaguier, S.; Dagorne, S.; Fosse, C.; Martel, A.; Jaouen, G.; Salmain, M. *Tetrahedron Lett.* **2008**, *49*, 4670.
- (21) Haquette, P.; Dumat, B.; Talbi, B.; Arbabi, S.; Renaud, J.-L.; Jaouen, G.; Salmain, M. *J. Organomet. Chem.* **2009**, *694*, 937.
- (22) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M.; Maiwald, P.; Peyralans, J. J. P.; Maichele, A.; Fu, Y.; Jiao, N.; Hollmann, F.; Mondière, R.; Taglieber, A. *Tetrahedron* **2007**, *63*, 6404.
- (23) Jung, Y. C.; Mishra, R. K.; Yoon, C. H.; Jung, K. W. *Org. Lett.* **2003**, *5*, 2231.
- (24) Andappan, M. M. S.; Nilsson, P.; Larhed, M. *Chem. Commun.* **2004**, 218.
- (25) Andappan, M. S.; Nilsson, P.; Larhed, M. *Mol Divers* **2003**, *7*, 97.
- (26) Yoo, K. S.; Yoon, C. H.; Jung, K. W. *J. Am. Chem. Soc.* **2006**, *128*, 16384.
- (27) Yoo, K. S.; Park, C. P.; Yoon, C. H.; Sakaguchi, S.; O'Neil, J.; Jung, K. W. *Org. Lett.* **2007**, *9*, 3933.
- (28) Estell, D. A.; Graycar, T. P.; Miller, J. V.; Powers, D. B.; Wells, J. A.; Burnier, J. P.; Ng, P. G. *Science* **1986**, *233*, 659.
- (29) Perona, J. J.; Craik, C. S. *Protein Sci.* **1995**, *4*, 337.

- (30) Wells, J. A.; Estell, D. A. *Trends Biochem. Sci* **1988**, *13*, 291.
- (31) Alapafuja, S. O.; Nikas, S. P.; Shukla, V. G.; Papanastasiou, I.; Makriyannis, A. *Tetrahedron Lett.* **2009**, *50*, 7028.
- (32) Brouwer, A. J.; Ceylan, T.; Jonker, A. M.; Linden, T. v. d.; Liskamp, R. M. J. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 2397.
- (33) Nishiguchi, A.; Maeda, K.; Miki, S. *ChemInform* **2007**, *38*, no.
- (34) Kuhn, P.; Knapp, M.; Soltis, S. M.; Ganshaw, G.; Thoene, M.; Bott, R. *Biochemistry* **1998**, *37*, 13446.
- (35) Albano, V. G.; Natile, G.; Panunzi, A. *Coord. Chem. Rev.* **1994**, *133*, 67.
- (36) Enquist, P.-A.; Nilsson, P.; Sjöberg, P.; Larhed, M. *The Journal of Organic Chemistry* **2006**, *71*, 8779.
- (37) Gligorich, K. M.; Sigman, M. S. *Chem. Commun.* **2009**, 3854.
- (38) Stahl, S. S.; Thorman, J. L.; Nelson, R. C.; Kozee, M. A. *J. Am. Chem. Soc.* **2001**, *123*, 7188.
- (39) Thiel, W. R. *Angew. Chem. Int. Ed.* **1999**, *38*, 3157.
- (40) Radomsky, N. A.; Gibian, M. U.; Elliott, D. L. *J. Am. Chem. Soc.* **1973**, *95*, 8717.
- (41) Stauffer, C. E.; Etson, D. J. *J. Biol. Chem.* **1969**, *244*, 5333.
- (42) Polgar, L.; Bender, M. L. *J. Am. Chem. Soc.* **1966**, *88*, 3153.
- (43) Gold, A. M.; Fahrney, D. *Biochemistry* **1964**, *3*, 783.
- (44) Rossi, L. M.; Vono, L. L. R.; Silva, F. P.; Kiyohara, P. K.; Duarte, E. L.; Matos, J. R. *Applied Catalysis A: General* **2007**, *330*, 139.
- (45) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, *47*, 1698.
- (46) Kapur, A.; Beck, J. L.; Brown, S. E.; Dixon, N. E.; Sheil, M. M. *Protein Sci.* **2002**, *11*, 147.
- (47) Satake, Y.; Abe, S.; Okazaki, S.; Ban, N.; Hikage, T.; Ueno, T.; Nakajima, H.; Suzuki, A.; Yamane, T.; Nishiyama, H.; Watanabe, Y. *Organometallics* **2007**, *26*, 4904.
- (48) Cho, C. S.; Uemura, S. *J. Organomet. Chem.* **1994**, *465*, 85.
- (49) Du, X.; Suguro, M.; Hirabayashi, K.; Mori, A.; Nishikata, T.; Hagiwara, N.; Kawata, K.; Okeda, T.; Wang, H. F.; Fugami, K.; Kosugi, M. *Org. Lett.* **2001**, *3*, 3313.

3.5 Experimental procedures

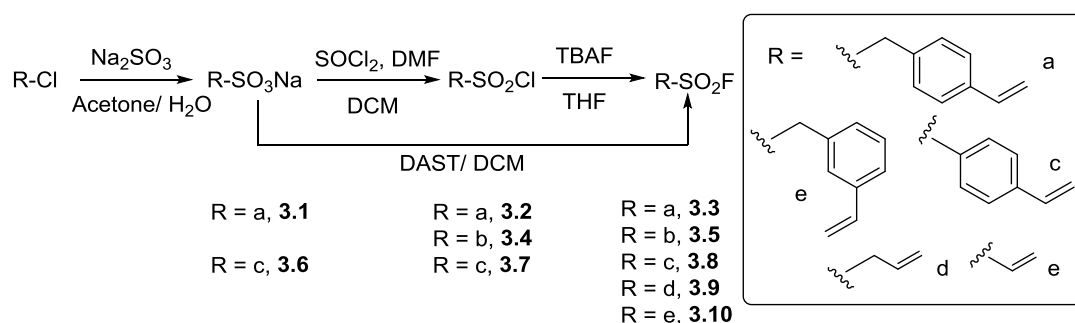
3.5.1 General considerations

See Chapter 2 Section 2.5.1 for general considerations on chemical synthesis. For protein analysis, liquid chromatography-mass spectrometry (LC-MS) was performed on a Micromass LCT (ESI-TOF-MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C4 column (250 x 4.6 mm x 5 μ m). Water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase at a flow rate of 1.0 mL min⁻¹. The gradient was programmed as follows: 95% A (5 min isocratic) to 95% B over 15 min then isocratic for 5 min. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V for Subtilisin *Bacillus lentus* (SBL). Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 l/hr. Spectra were calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25V. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.0 from Waters) according to manufacturer's instructions.

Modified and un-modified proteins were not resolved during chromatography and typically eluted between 13 and 16 minutes. MS data therefore reflects the total protein content of the reaction mixture. Conversions were calculated from the relative intensities of mass peaks after MaxEnt processing. Additionally, using known protein standards, we showed that both modified and unmodified proteins are ionized with similar efficiency (See Section 3.1.6).

Design of substrates and modelling studies

3.5.2 Chemical Synthesis



General procedure for synthesis of sodium sulfonates

Sodium sulfonate salts were obtained from corresponding benzyl bromides by following a published procedure.³¹ Typically, to a solution of benzyl halide (3 mmol) in acetone (7 mL) was added sodium sulfite (5.5 mmol) dissolved in water (7 mL). The reaction mixture was refluxed for 18-24 h. Alternatively, the reaction mixture was heated for 15 min at 160 °C under microwave irradiation. The reaction mixture was cooled to room temperature and solvent was evaporated *in vacuo*. The solid obtained was washed thoroughly with EtOAc to obtain the white salts used in the next step without further purification.

General procedure for synthesis of sulfonyl chlorides^{32,33}

To a suspension of the sodium sulfonate (2 mmol) in anhydrous toluene/ DMF (7:1 mL), was added thionyl chloride (3 mmol). The mixture was heated to 50 °C for 3 h. The reaction was quenched with water and the product was extracted with CH₂Cl₂, washed with brine and dried with MgSO₄. The solvent and volatiles were evaporated under reduced pressure and the crude product was used for the next step without any further purification.

General procedure for synthesis of sulfonyl fluorides^{32,33}

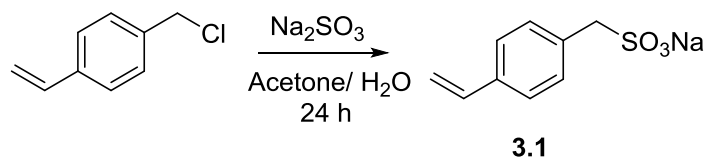
To a solution of the sulfonyl chloride (2 mmol) in THF (15 mL) was added tetrabutylammonium fluoride solution in THF. The reaction mixture was stirred at room

temperature for 2 h. Subsequently, 1 M HCl was added and THF was evaporated *in vacuo*. The product was extracted with CH₂Cl₂, washed with brine and dried with MgSO₄. Sulfonyl fluorides were purified by column chromatography.

General procedure for synthesis of sulfonyl fluorides (using DAST)

To a stirred suspension of sodium sulfonate salt (1.2 mmol) in DCM (7 mL), in an ice bath was added (diethylamino)sulfur trifluoride (2 mmol). The reaction mixture was stirred overnight at room temperature. The product was extracted with ethyl acetate, washed with brine and dried with MgSO₄. The solvent was evaporated and the product was purified by column chromatography.

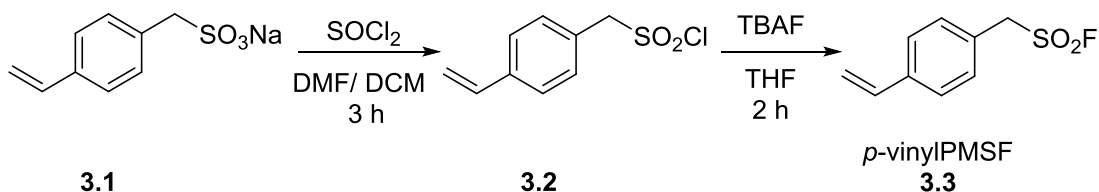
*Sodium (4-vinylphenyl)methanesulfonate **3.1***



To a solution of 4-vinylbenzyl chloride (912 mg, 6 mmol) in acetone (10 mL) was added sodium sulfite (1.2g, 9 mmol) dissolved in water (10 mL). The reaction mixture was refluxed for 24 h. The reaction mixture was cooled to room temperature and solvent was evaporated *in vacuo*. The solid obtained was washed thoroughly with EtOAc to obtain the white salt **3.1** (1.2 g, 91%) used in the next step without further purification. IR (neat): 3429.26, 1629.04, 1510.97, 1412.07, 1190.24, 1145.24, 1045.36, 963.85, 898.95, 835.65; ¹H NMR (400 MHz, D₂O): δ = 4.10 (s, 2H, CH₂S), 5.27 (d, *J*=11.0, 1H, CH=CH₂), 5.81 (d, *J*=17.7, 1H, CH=CH₂), 6.74 (dd, *J*=17.7, 11.0, 1H, CH=CH₂), 7.33 (d, *J*=8.2, 2H, SCH₂CCH_{ar}), 7.44 (d, *J*=8.2, 2H, CH_{ar}CCH=CH₂); ¹³C NMR (101 MHz, D₂O): δ = 57.0 (CH₂S), 115.0 (CH=CH₂), 126.7 (CH_{ar}CCH=CH₂ x 2), 131.1 (SCH₂CCH_{ar} x 2), 131.7

(SCH₂C_q), 136.7 (C_qCH=CH₂), 137.5 (CH=CH₂); HRMS m/z (ESI⁺): Found 243.0057 [M+Na]⁺; C₉H₉Na₂O₃S requires 243.0062.

(4-vinylphenyl)methanesulfonyl fluoride 3.3

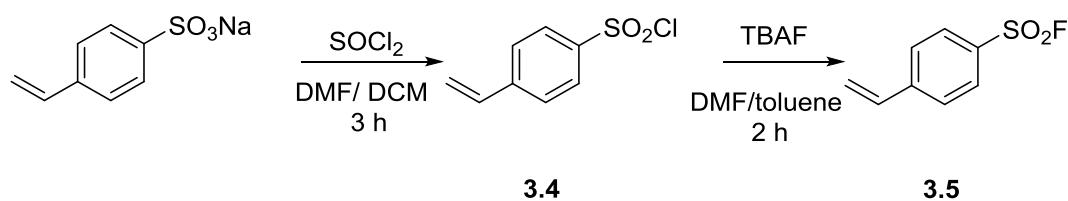


To a suspension of the sodium sulfonate **3.1** (500 mg, 2.27 mmol) in anhydrous DMF/DCM (300 μ L DMF in 10 mL DCM), was added thionyl chloride (574 mg, 3.6 mmol). The mixture was stirred at room temperature for 3 h, after which the solvent was evaporated and crude sulfonyl chloride **3.2** was used without purification in the subsequent halide exchange step. ¹H NMR (400 MHz, CDCl₃) δ = 4.89 (s, 2H, CH₂S), 5.39 (d, *J*=10.9, 1H, CH=CH₂), 5.86 (d, *J*=17.6, 1H, CH=CH₂), 6.77 (dd, *J*=17.6, 10.9, 1H, CH=CH₂), 7.47 (d, *J*=8.3, 2H, SCH₂CCH_{ar}), 7.52 (d, *J*=8.3, 2H, CH_{ar}CCH=CH₂). ¹³C NMR (176 MHz, CDCl₃) δ = 70.8 (CH₂S), 116.0 (CH=CH₂), 125.3 (SCH₂C_q), 127.0 (CH_{ar}CCH=CH₂ x 2), 131.6 (SCH₂CCH_{ar} x 2), 135.8 (CH=CH₂), 139.6 (C_qCH=CH₂).

To a solution of sulfonyl chloride **3.2** (200 mg, 0.93 mmol) in THF (10 mL) was added tetrabutylammonium fluoride solution in THF (3.6 mmol). The mixture was stirred at room temperature for 2 h, after which the reaction was quenched by addition of water and the product was extracted from by DCM (10 mL x 3). Solvent was evaporated and the crude product was purified by flash chromatography on silica gel using CH₂Cl₂, to yield 120 mg of a white solid (26% over two steps). *R*_f = 0.71 (SiO₂, 33% EtOAc/Petrol); IR (neat): 2992.27, 1510.54, 1391.06, 1364.75, 1259.76, 1209.54, 1187.73, 1161.87, 1138.76, 1019.65, 990.94, 924.96, 879.02, 852.13, 836.44, 808.19, 763.67, 738.93, 704.34, 650.07.

^1H NMR (400 MHz, CDCl_3) δ = 4.59 (s, 2H, CH_2S), 5.27 (d, J =10.9, 1H, $\text{CH}=\text{CH}_2$), 5.77 (d, J =17.6, 1H, $\text{CH}=\text{CH}_2$), 6.72 (dd, J =17.6, 10.9, 1H, $\text{CH}=\text{CH}_2$), 7.36 (d, J =8.3, 2H, SCH_2CCHar), 7.41 (d, J =8.3, 2H, $\text{CHarCCH}=\text{CH}_2$); ^{13}C NMR (101 MHz, CDCl_3) δ = 53.4 (CH_2S), 113.2 ($\text{CH}=\text{CH}_2$), 114.6 (SCH_2C_q), 126.5 ($\text{CHarCCH}=\text{CH}_2 \times 2$), 128.8 ($\text{SCH}_2\text{CCHar} \times 2$), 136.2 ($\text{CH}=\text{CH}_2$), 136.9 ($\text{C}_q\text{CH}=\text{CH}_2$). ^{19}F NMR (377 MHz, CDCl_3) δ = 51.3. LRMS m/z (Cl^+): 200.03 $[\text{M}]^+$, 218.06 $[\text{M}+\text{NH}_4]^+$.

4-vinylbenzenesulfonyl fluoride **3.5**



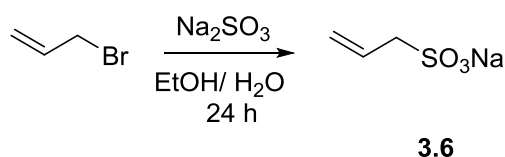
Synthesis of compound **3.5** was done by Dr. Maria Espling.

To a suspension of the vinylbenzene sulfonate (2 mmol) in anhydrous DMF/toluene (1 mL DMF in 7 mL toluene), was added SOCl_2 (3 mmol). The reaction was stirred at 50 °C for 3 h under nitrogen. The reaction mixture was quenched by addition of H_2O , and the sulfonyl chloride was extracted with DCM. The organic phase was collected and washed with brine, dried over MgSO_4 and solvent was evaporated to yield a pale yellow solid **3.4** that was used in the next step without purification.

To a solution of **3.4** in 15 mL THF, was added TBAF (4 mmol) dissolved in THF (1 M solution). The reaction was stirred in room temperature for 1 h, and then quenched with 1M HCl. THF was removed by evaporation and the remaining mixture was extracted by DCM. The organic phase was collected, washed with brine and dried over MgSO_4 . Solvent was evaporated and the crude product was purified by flash chromatography on silica gel to yield a white solid in 17 % yield over two steps. R_f = 0.7 (SiO_2 , 1:2 EtOAc/Petrol); ^1H

NMR (400 MHz, CDCl₃) δ = 5.53 (d, J =10.9, 1H, CH=CH₂), 5.96 (d, J =17.6, 1H, CH=CH₂), 6.77 (dd, J =17.6, 10.9, 1H, CH=CH₂), 7.61 (d, J =8.3, 2H, CH_{ar}CCH=CH₂), 7.95 (d, J =8.5, 2H, SCH₂CCH_{ar}). ¹³C NMR (101 MHz, CDCl₃) δ = 119.3, 127.0, 128.7, 131.2, 131.5, 134.7, 144.7. ¹⁹F NMR (376 MHz, CDCl₃) δ = 66.25. LRMS m/z (ESI⁺): 187.0 [M+H]⁺.

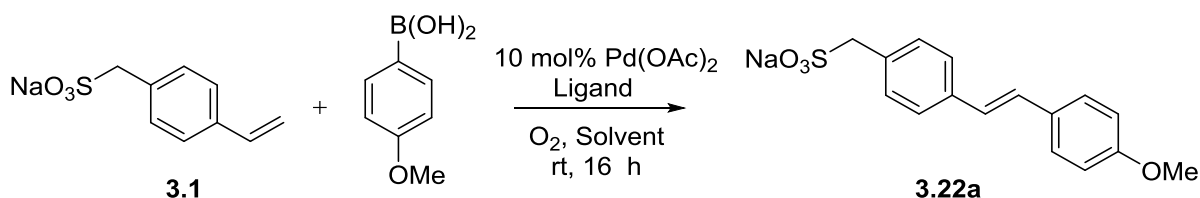
Sodium prop-2-ene-1-sulfonate **3.6**



To a solution of 3-bromoprop-1-ene (2.9 mL, 33 mmol) in 95% ethanol (20 mL) was added sodium sulfite (5.4g, 43 mmol) dissolved in water (10 mL). The reaction mixture was refluxed for 16 h. The reaction mixture was cooled to room temperature and solvent was evaporated *in vacuo*. The solid obtained was washed thoroughly with EtOAc and hexane to obtain the white solid **3.6** (2.8 g, 59%) used in the next step without further purification. IR (neat): 3465.93, 1222.92, 1180.00, 1050.92, 989.35, 920.63, 797.09, 654.40. ¹H NMR (400 MHz, D₂O) δ = 3.55 (dt, J =7.3, 1.1, 2H), 5.22 – 5.33 (m, 2H), 5.82 (ddt, J =17.4, 10.2, 7.4, 1H). ¹³C NMR (101 MHz, D₂O) δ = 55.5, 122.3, 127.5.

Preparation of final coupled product standards

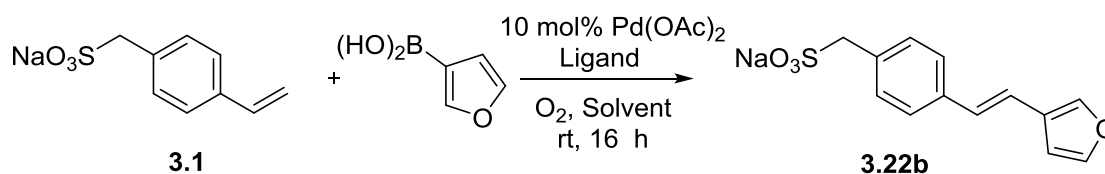
Sodium (*E*)-(4-(4-methoxystyryl)phenyl)methanesulfonate **3.22a**



To a pre-mixed solution of palladium acetate (10.2 mg, 0.045 mmol) and 5,5'-dimethyl-2,2'-bipyridine (10 mg, 0.054 mmol) in water (10 mL) for 30 min, were added olefin **3.1**

(170 mg, 0.7 mmol) and 4-methoxyphenylboronic acid (243 mg, 1.6 mmol) along with (180 mg, 2.2 mmol) NaOAc as solid. The reaction mixture was stirred under O₂ balloon at room temperature for 16 h. The solvent was evaporated and crude product was re-dissolved in water, filtered through 0.45µm filtered and purified by rpHPLC using a Phenomenex Synergi-Hydro-RP C18 column to obtain **3.22a** as white solid. Solvent System: 0.1% formic acid in H₂O (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow rate of 1.0 mL min⁻¹. The multi-step gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 20% B over 1 min followed by a gradual linear gradient to 38% B over next 18 minutes and then 100% B over 2 minutes. The column was flushed with 100% B for an additional 2 min, followed by a linear gradient over 2 minutes back to 100% A. The column was re-equilibrated with 100% A for additional 5 minutes. t_R = 14.0. Yield = 5.2%. ¹H NMR (500 MHz, D₂O) δ = 3.11 (s, 3H, CH₃), 4.01 (s, 2H, CH₂S), 6.86 (d, J =16.4, 1H, CH=CH), 6.98 (d, J =16.4, 1H, CH=CH), 7.15 (s, 4H, CHar), 7.35 (d, J =8.2, 2H, CHar), 7.54 (d, J =8.2, 2H, CHar). ¹³C NMR (126 MHz, D₂O) δ = 43.2 (CH₃), 56.5 (CH₂S), 126.6, 126.7, 126.8, 127.4, 130.8, 132.1, 132.3, 135.3, 136.8, 142.7. HRMS m/z (ESI): Found 303.0683 [M-Na]⁺; C₁₆H₁₅O₄S requires 303.0697.

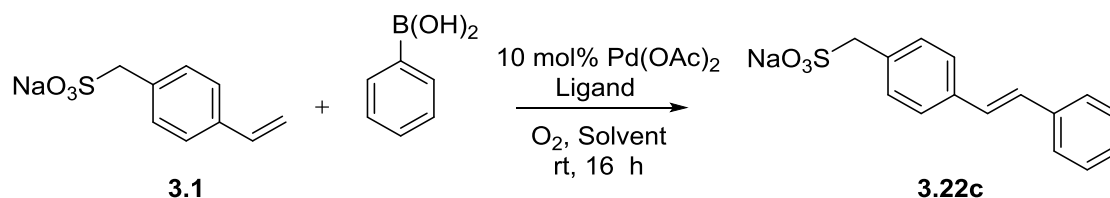
Sodium (E)-4-(2-(furan-3-yl)vinyl)phenyl)methanesulfonate 3.22b



To a pre-mixed solution of palladium acetate (10.2 mg, 0.045 mmol) and 5,5'-dimethyl-2,2'-bipyridine (10 mg, 0.054 mmol) in DMF (7 mL) for 30 min were added olefin **3.1** (100 mg, 0.455 mmol) and furan-3-ylboronic acid (101 mg, 0.9 mmol). The reaction mixture was stirred under O₂ balloon at room temperature for 16 h. The solvent was

evaporated and crude product was re-dissolved in water, filtered through 0.45µm filtered and purified by rpHPLC using a Phenomenex Synergi-Hydro-RP C18 column to obtain **3.22b** as the product (7 mg, 5.4%). Solvent System: 0.1% formic acid in H₂O (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow rate of 1.0 mL min⁻¹. The gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 100% B over 24 min followed by 100% B for an additional 2 min. A linear gradient over 2 minutes back to 100% A was used to bring the column back to initial conditions. The column was re-equilibrated with 100% A for additional 5 minutes. t_R = 11.95. ¹H NMR (400 MHz, D₂O) δ = 4.06 (s, 2H, CH₂SO₃Na), 6.72 (s, 1H), 6.87 (d, J =16.5, 1H, CH=CH), 7.07 (d, J =16.5, 1H, CH=CH), 7.30 (d, J =7.7, 2H), 7.44 (d, J =8.8, 3H), 7.58 (s, 1H); ¹³C NMR (126 MHz, D₂O) δ = 56.6, 107.1, 118.9, 124.2, 126.1, 127.6, 130.7, 136.8, 141.6, 144.3; HRMS m/z (ESI): Found 263.0374 [M-Na]⁻; C₁₃H₁₁O₄S requires 263.0384.

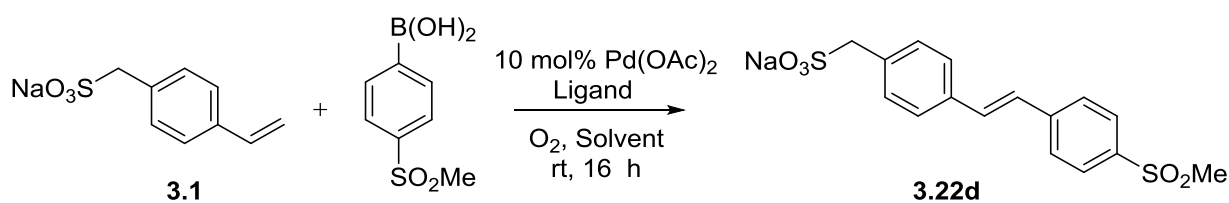
Sodium (4-styrylphenyl)methanesulfonate **3.22c**



To a mixture of palladium acetate (10.2 mg, 0.045 mmol) and 5,5'-dimethyl-2,2'-bipyridine as a ligand (10 mg, 0.054 mmol) at room temperature was successively added water (15 mL), (4-hydroxyphenyl)boronic acid (101 mg, 0.9 mmol), Na₂CO₃ (100 mg, 0.9 mmol) and olefin **3.1** (100 mg, 0.455 mmol). The reaction mixture was heated to 50 ° C for 3 h and then stirred under O₂ balloon at room temperature for 16 h. The crude product was filtered through 0.45 µm filtered, concentrated and purified by rpHPLC using a Phenomenex Synergi-Hydro-RP C18 column to give **3.22c** (20 mg, 15%). Solvent System: 0.1% formic acid in H₂O (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow

rate of 1.0 mL min⁻¹. The gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 100% B over 24 min followed by 100% B for an additional 2 min. A linear gradient over 2 minutes back to 100% A was used. The column was re-equilibrated with 100% A for additional 5 minutes. t_R = 13.15; ¹H NMR (500 MHz, D₂O) δ = 4.11 (s, 2H, CH₂S), 7.22 (s, 2H, CHar), 7.28 (t, J =7.4, 1H, CHar), 7.35-7.40 (m, 4H, CHar), 7.57 (t, J =7.4, 4H, CHar); ¹³C NMR (126 MHz, D₂O) δ = 56.6, 126.5, 126.63, 128.0, 128.2, 128.8, 129.0, 130.8, 131.1, 136.8, 137.1; HRMS m/z (ESI): Found 273.0594 [M-Na]⁻; C₁₅H₁₃O₃S requires 273.0591.

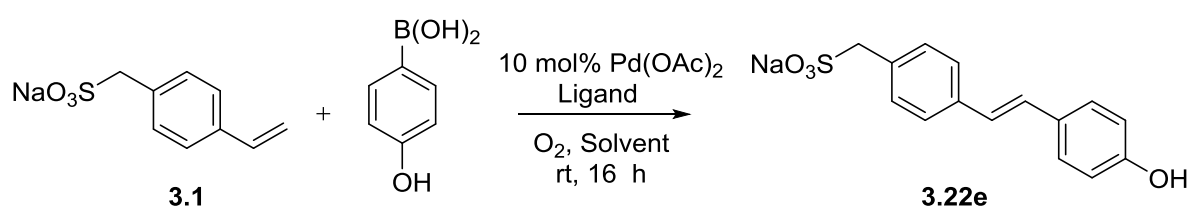
Sodium (E)-(4-(4-(methylsulfonyl)styryl)phenyl)methanesulfonate 3.22d



To a pre-mixed solution of palladium acetate (10.2 mg, 0.045 mmol) and 5,5'-dimethyl-2,2'-bipyridine (10 mg, 0.054 mmol) in water (10 mL) for 30 min, were added olefin **3.1** (170 mg, 0.7 mmol) and (4-(methylsulfonyl)phenyl)boronic acid (320 mg, 1.6 mmol) along with (180 mg, 2.2 mmol) NaOAc as solid. The reaction mixture was stirred under O₂ balloon at room temperature for 16 h. The solvent was evaporated and crude product was re-dissolved in water, filtered through 0.45 μ m filtered and purified by rpHPLC using a Phenomenex Synergi-Hydro-RP C18 column to obtain **3.22d** as white solid. Solvent System: 0.1% formic acid in H₂O (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow rate of 1.0 mL min⁻¹. The multi-step gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 20% B over 1 min followed by a gradual linear gradient to 38% B over next 18 minutes and then 100% B over 2 minutes. The column was flushed with 100% B for an additional 2 min, followed by a linear gradient over 2 minutes

back to 100% A. The column was re-equilibrated with 100% A for additional 5 minutes. $t_R = 8.5$. Yield = 2%. ^1H NMR (400 MHz, D_2O) $\delta = 3.13$ (s, 3H), 4.06 (s, 2H), 7.12 (d, $J=16.6$, 1H), 7.24 (d, $J=16.4$, 1H), 7.31 (d, $J=8.2$, 2H), 7.48 (d, $J=7.8$, 2H), 7.65 (d, $J=8.2$, 2H), 7.78 (d, $J=8.2$, 2H). HRMS m/z (ESI): Found 351.0369 $[\text{M}-\text{Na}]^-$; $\text{C}_{16}\text{H}_{15}\text{O}_5\text{S}_2$ requires 351.0366.

Sodium (E)-(4-(4-hydroxystyryl)phenyl)methanesulfonate 3.22e

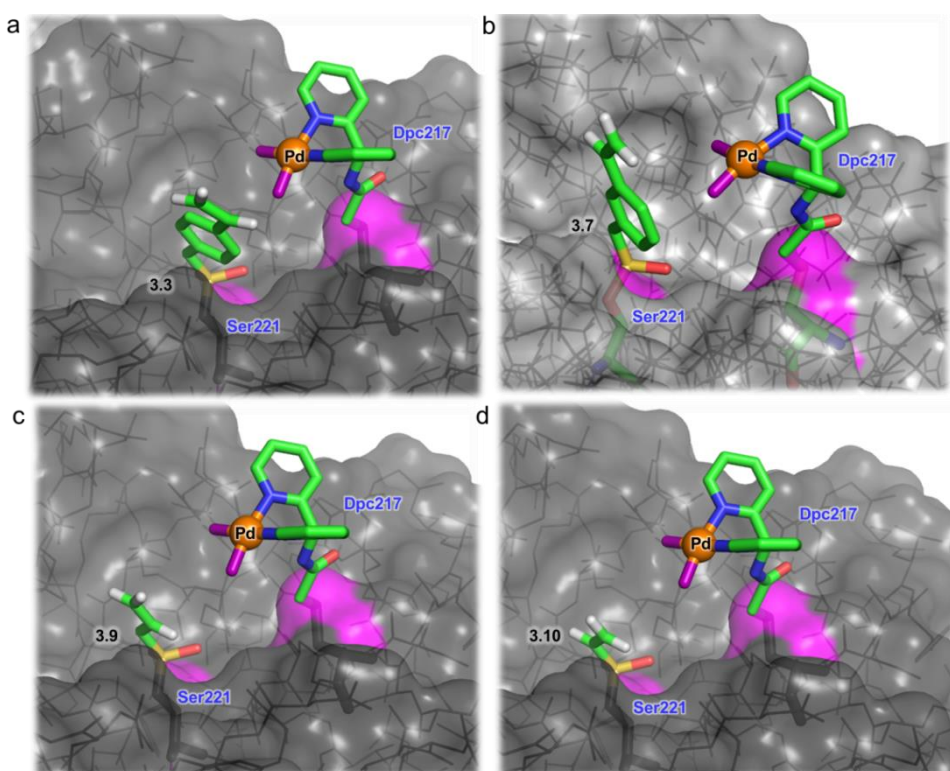


A mixture of phenylboronic acid (1 mmol), olefin **3.1** (0.5 mmol), palladium acetate (0.05 mmol), and sodium acetate (4 mmol) was stirred in acetic acid (15 mL) under O_2 balloon at room temperature for 16 h. The solvent was evaporated and crude product was re-dissolved in water, filtered through 0.45 μm filter and purified by rpHPLC using a Phenomenex Synergi-Hydro-RP C18 column to give trans isomer, **3.22e** (13 mg, 9.2%). Solvent System: 0.1% formic acid in H_2O (solvent A) and 0.1% formic acid in CH_3CN (solvent B) at a flow rate of 1.0 mL min^{-1} . The multi-step gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 20% B over 1 min followed by a gradual linear gradient to 38% B over next 18 minutes and then 100% B over 2 minutes. The column was flushed with 100% B for an additional 2 min, followed by a linear gradient over 2 minutes back to 100% A. The column was re-equilibrated with 100% A for additional 5 minutes. $t_R = 9.25$. ^1H NMR (500 MHz, D_2O) $\delta = 4.10$ (s, 2H, $\text{CH}_2\text{SO}_3\text{Na}$), 6.85 (d, 2H, $J = 8.6$ Hz, Ar), 7.05 (d, 1H, $J = 16.4$, $\text{CH}=\text{CH}$), 7.15 (d, 1H, $J = 16.4$, $\text{CH}=\text{CH}$), 7.34 (d, 2H, $J = 8.1$ Hz, Ar), 7.49 (d, 2H, $J = 8.7$ Hz, Ar), 7.51 (d, 2H, $J = 7.9$ Hz, Ar) ppm; ^{13}C NMR (126 MHz, D_2O) $\delta = 56.6$, 115.8, 125.9, 126.3, 128.1, 128.4, 129.7,

130.7, 130.8, 137.1, 155.4 ppm; HRMS m/z (ESI): Found 289.0536 $[M-Na]^+$; $C_{15}H_{13}O_4S$ requires 289.0540.

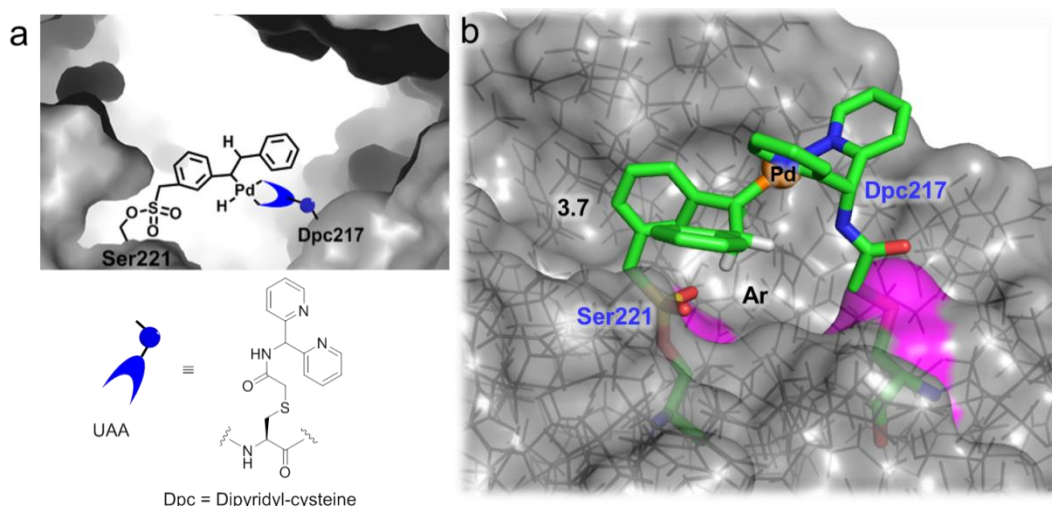
3.5.3 Modelling studies

Modelling experiments were done by Dr. Mark Wormald. A preliminary modelling study was performed using Insight II and Discover **software** (Accelrys) on a Silicon Graphics Indigo 2 workstation. The SBL-WT crystal structure (PDB ID: 1GCI) was obtained from the Protein Data Bank.³⁴ Residue 217 was converted to a cysteine. The Dpc•Pd or Bpc•Pd complexes was installed at position 217 by adding corresponding ligands **L1** or **L2** to Cys217, and *p*-vinylPMSF was added to Ser221. Models for *m*-vinylPMSF **3.10** and alkyl-chain substrates **3.8** and **3.9** installed at Ser221 were also built (Scheme 3-32a-d).



Scheme 3-32. Models for installation of unnatural amino acid Dpc at position 217 and alkene substrates (a) *p*-vinylPMSF **3.3** (b) *m*-vinylPMSF **3.10** (c) allylsulfonyl fluoride **3.8** (d) vinylsulfonyl fluoride **3.9** installed at Ser221 in active site of SBL

Models of the reaction intermediates were built for the transmetalation of phenylboronic acid on to the palladium. These were energy minimised using the CVFF force field in Discover, keeping the peptide backbone completely rigid. Only relaxation of the ligands and the amino acid side-chains in the immediate vicinity of the ligands was allowed, and the coordination geometry of the palladium was fixed as square planar. The intermediates were then inspected for any steric clashes or high energy interactions. Modelling study suggested that the migratory insertion step was feasible and the cross-coupling may be facilitated within the enzyme active site when *m*- or *p*-vinylPMSF was chosen as the coupling partner for oxidative Heck cross-coupling reaction.



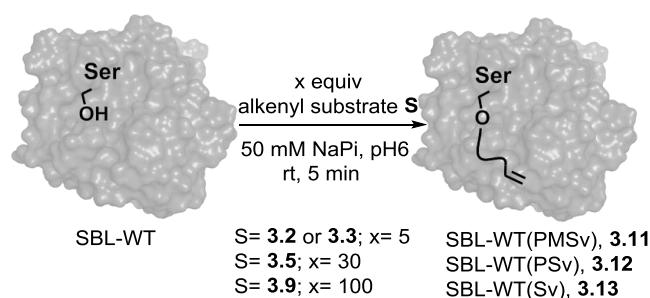
Scheme 3-33. (a) Schematic diagram for cross-coupling intermediate for oxidative Heck reaction between *m*-vinylPMSF **3.7** and phenylboronic acid in enzyme active site (b) Model for cross-coupling intermediate for oxidative Heck reaction between phenylboronic acid and substrate **3.10** installed in SBL-Dpc217 (See Scheme 3-7 for similar models for cross-coupling of *p*-vinylPMSF substrate **3.3** and phenylboronic acid in the active site of SBL-Bpc217 and SBL-Dpc217)

Enzyme-Substrate Interactions

3.5.4 Interaction of PMSF analogues with SBL-WT

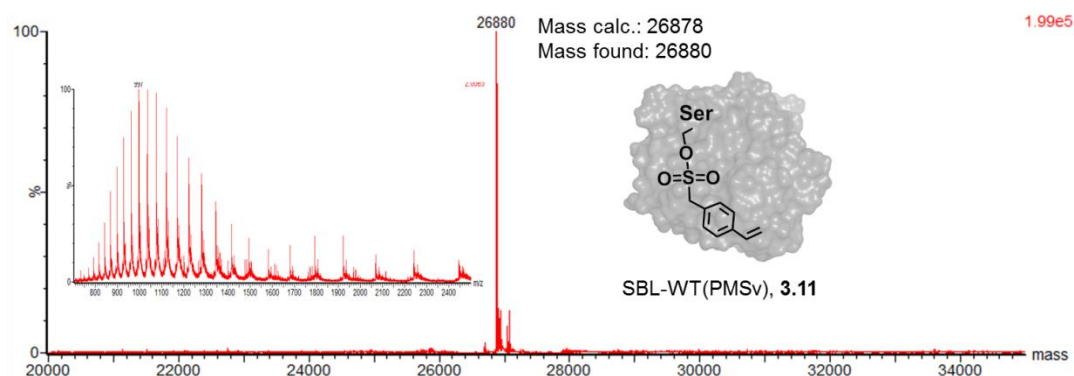
Interaction of novel PMSF-based substrates **3.3** or (**3.2**), **3.8** and **3.10** with SBL-WT was confirmed when complete conversion to the SBL-WT(PMSv), SBL-WT(PSv) and SBL-WT(Sv) species was observed. A 50 μ L sample of SBL-WT was prepared at 0.73 mg/mL in sodium

phosphate buffer (50 mM, pH 6). A 0.1 M stock solution of **3.3** (or **3.2** or **3.8** or **3.10**) was prepared in DMF.



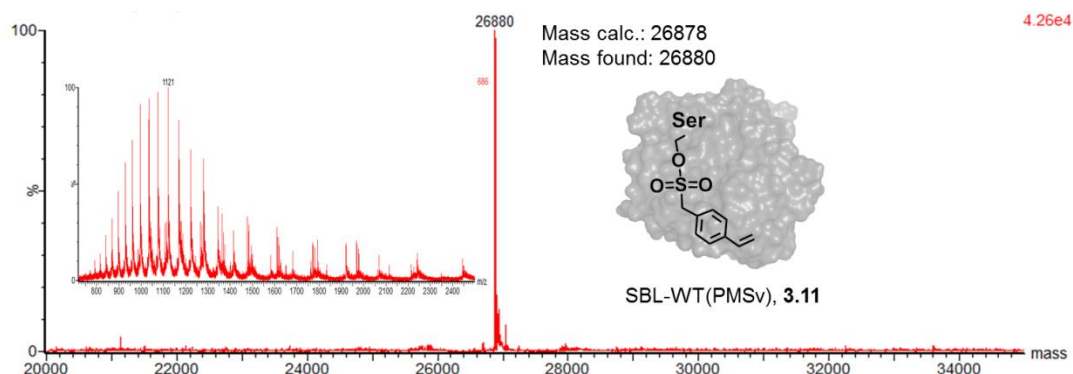
Substrate interaction with *p*-vinylphenylmethylsulfonyl chloride **3.2**:

To 50 μ L sample of SBL-WT (1.35 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.2 μ L of **3.2** in DMF (20.2 nmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) **3.11** (calculated mass 26878, mass found 26878). The ESI-MS for the modification is shown below.



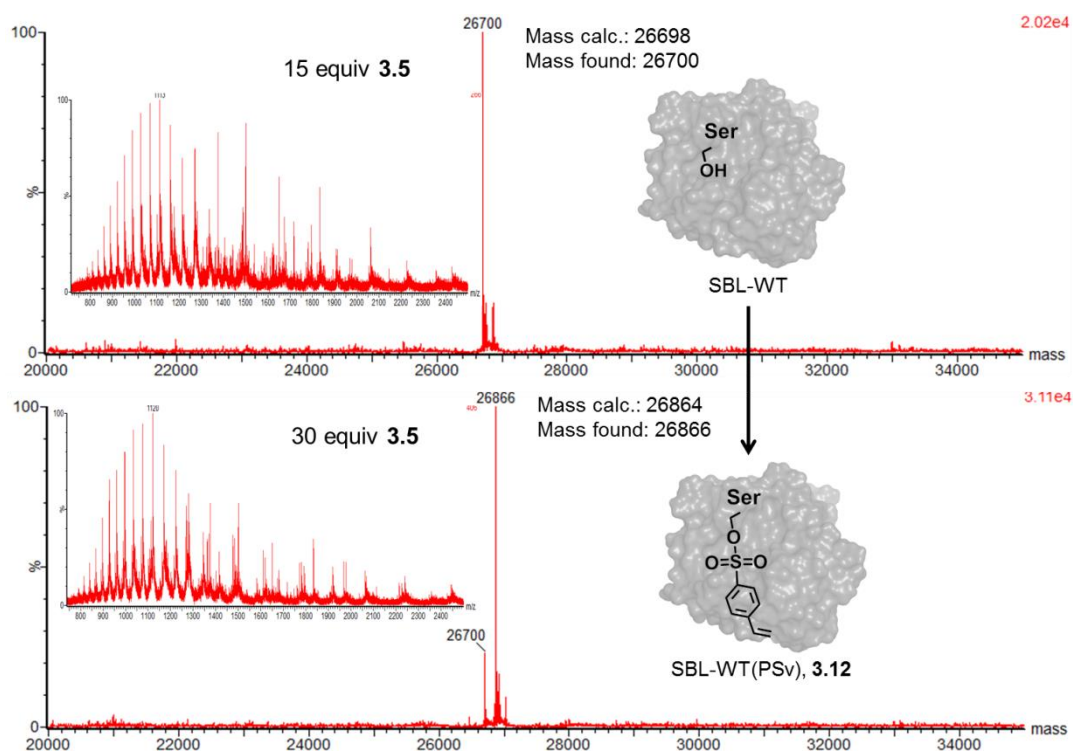
Substrate interaction with *p*-vinylphenylmethylsulfonyl fluoride **3.3**:

To 50 μ L sample of SBL-WT (1.35 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.2 μ L of **3.3** in DMF (20.2 nmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) **3.11** (calculated mass 26878, mass found 26880). The ESI-MS for the modification is shown below.



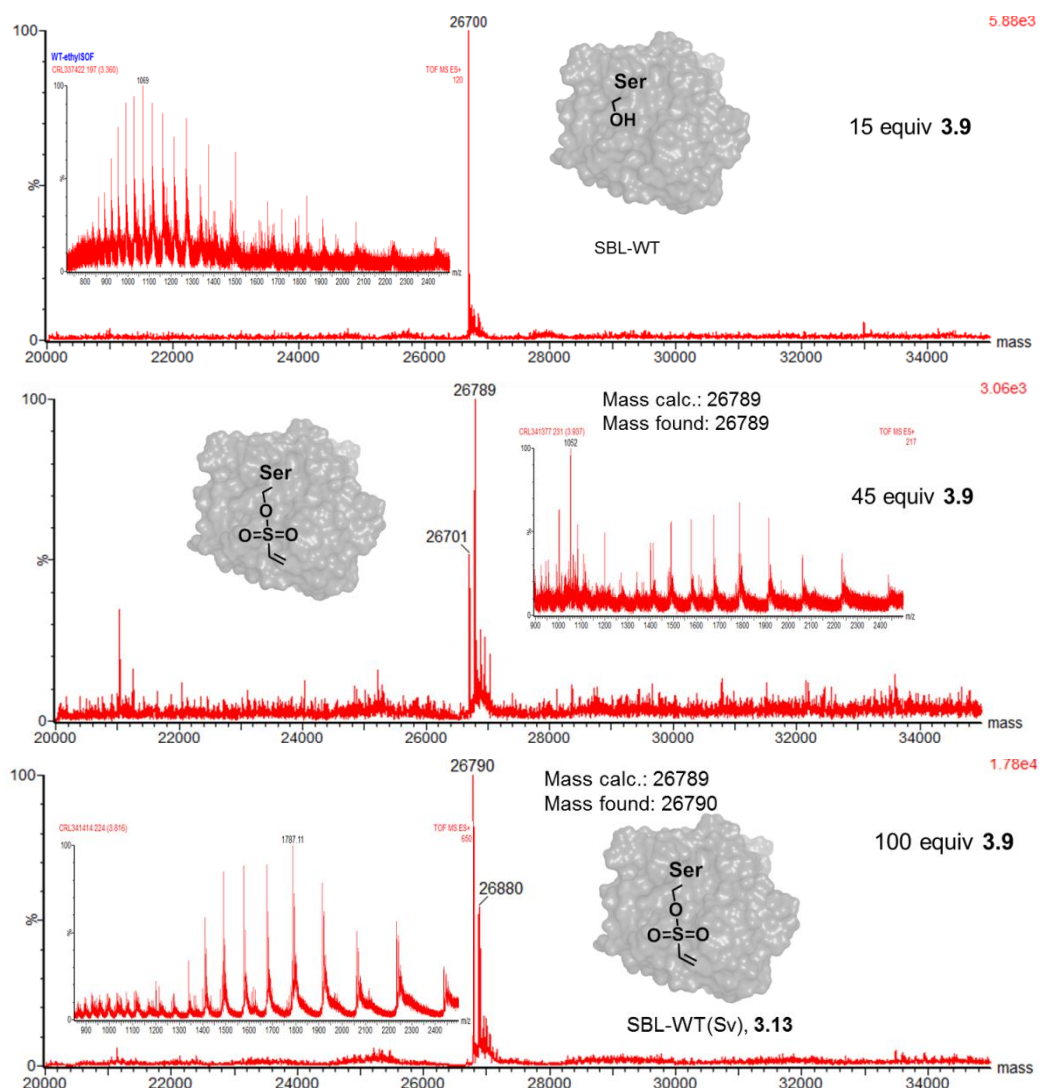
Substrate interaction with *p*-vinylphenylsulfonyl fluoride **3.5**:

To 50 μ L sample of SBL-WT (1.35 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.2 μ L of **3.5** in DMF (20.2 nmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. ESI-MS spectrum of the protein revealed no uptake of the substrate **3.5** by SBL-WT. another 0.2 μ L aliquot of **3.5** was added to the protein solution and the reaction was analysed by LC-MS after 5 min incubation at room temperature. A single protein species was detected with a mass corresponding to the SBL-WT(PSv) **3.12** (calculated mass 26864, mass found 26866). The ESI-MS for the modification is shown below.

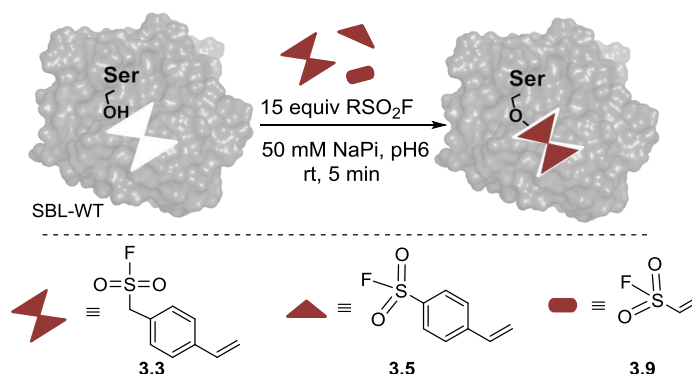


Substrate interaction with vinylsulfonyl fluoride **3.9**:

To 50 μL sample of SBL-WT (1.35 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 0.2 μL of **3.9** in DMF (20.2 nmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(Sv) **3.13** (calculated mass 26789, mass found 26790). The ESI-MS for the modification is shown below.



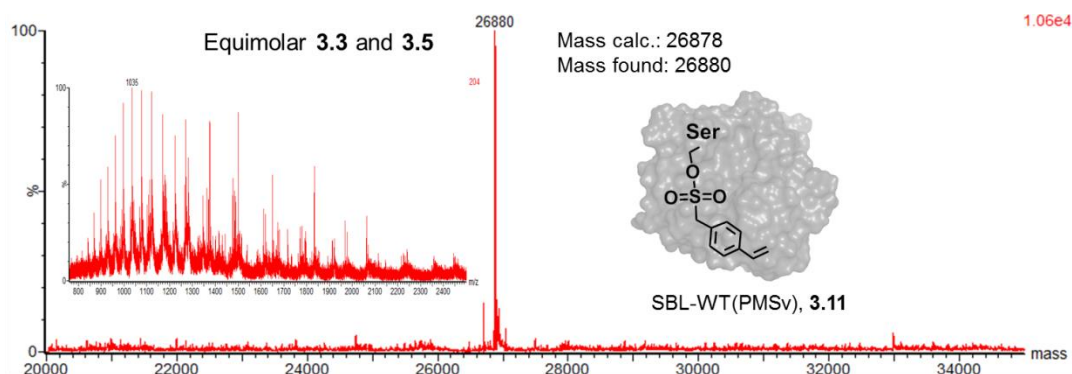
3.5.5 Competition loading assay of substrates 3.3, 3.5 and 3.9 using SBL-WT



SBL-WT was prepared at 0.73 mg/mL in sodium phosphate buffer (pH 6.0, 50 mM). 0.1 M stock solution of sulfonyl fluorides **3.3**, **3.5** and **3.9** in DMF were prepared.

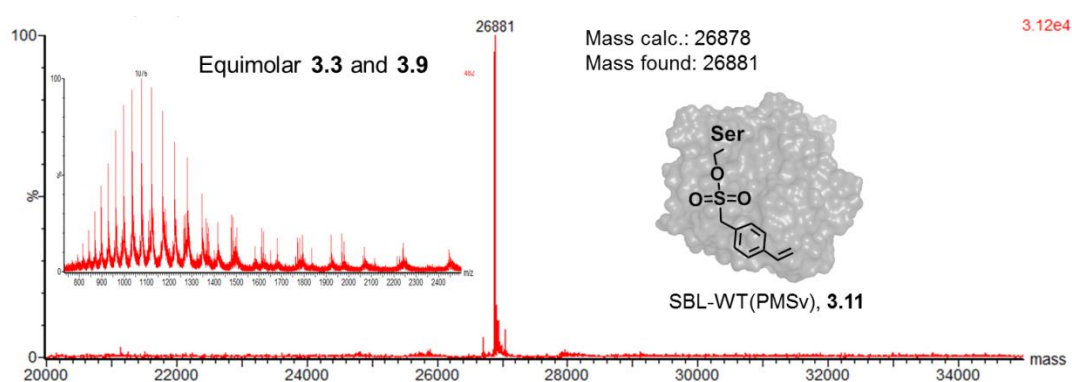
Competition loading test of substrate 3.3 and 3.5:

An equimolar mixture of substrates **3.3** and **3.5** (20.2 nmol each) in DMF was prepared for the competition assay. To a 50 μ L aliquot of the protein solution (1.35 nmol) in 50 mM NaPi buffer pH 6, the above mixture was added. The reaction solution was vortexed to homogenize and after 5 minutes of incubation at room temperature, an aliquot of the reaction mixture was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) **3.11** (calculated mass 26878, mass found 26880). The ESI-MS for the modification is shown below.



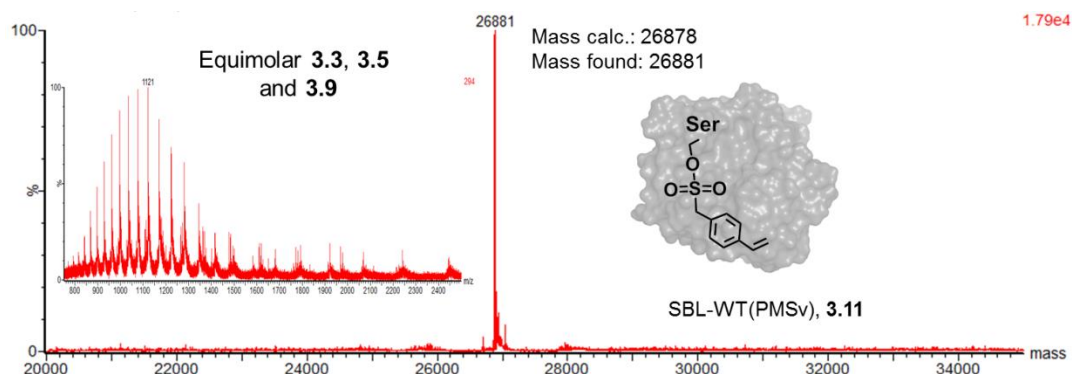
Competition loading test of substrate 3.3 and 3.9:

An equimolar mixture of substrates **3.3** and **3.9** (20.2 nmol each) in DMF was prepared for the competition assay. To a 50 μ L aliquot of the protein solution (1.35 nmol) in 50 mM NaPi buffer pH 6, the above mixture was added. The reaction solution was vortexed to homogenize and after 5 minutes of incubation at room temperature, an aliquot of the reaction mixture was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) **3.11** (calculated mass 26878, mass found 26881). The ESI-MS for the modification is shown below.



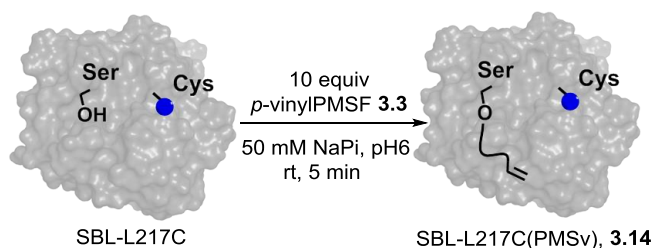
Competition loading test of substrate 3.3, 3.5 and 3.9:

An equimolar mixture of substrates **3.3**, **3.5**, **3.9** (20.2 nmol each) in DMF was prepared for the competition assay. To a 50 μ L aliquot of the protein solution (1.35 nmol) in 50 mM NaPi buffer pH 6, the above mixture was added. The reaction solution was vortexed to homogenize and after 5 minutes of incubation at room temperature, an aliquot of the reaction mixture was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) **3.11** (calculated mass 26878, mass found 26881). The ESI-MS for the modification is shown below.

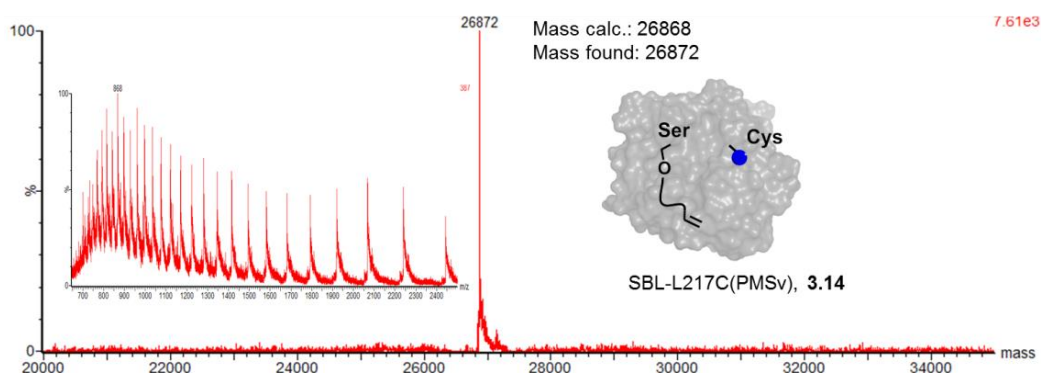


3.5.6 Interaction of substrate 3.3 with SBL-L217C

Interaction of *p*-vinylPMSF **3.3** with SBL-L217C was confirmed when complete conversion to the SBL-L217C(PMSv) species was observed.



A 100 μ L sample of SBL (mutant) L217C was prepared at 2 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.5 M stock solution of *p*-vinylPMSF **3.3** was prepared in DMF. To 100 μ L sample of SBL-L217C (7.4 nmol) in 50 mM sodium phosphate buffer, pH6 at 4°C, was added 0.15 μ L of **3.3** in DMF (74 nmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-L217C(PMSv) **3.14** anchoring *p*-vinylPMSF substrate (calculated mass 26868, mass found 26872). The ESI-MS for the modification is shown below.



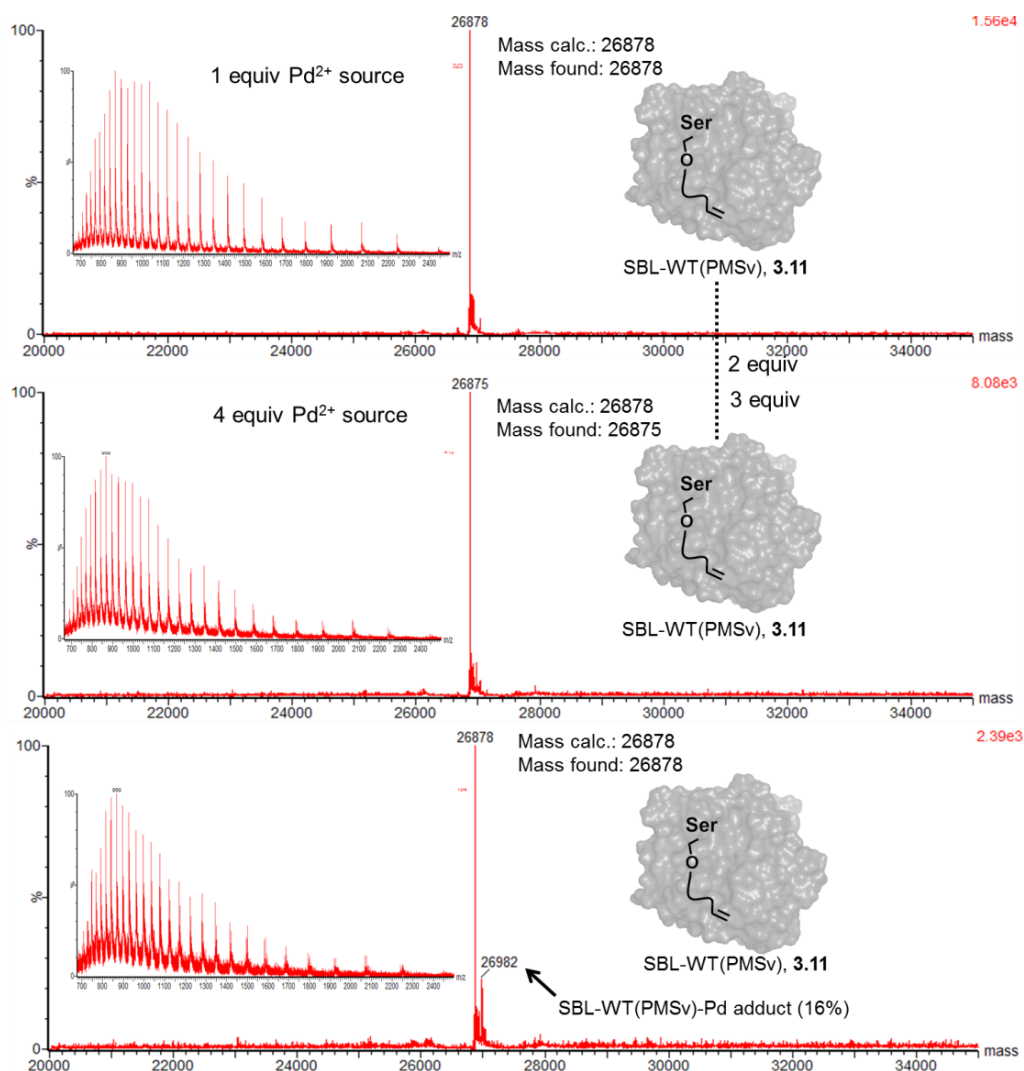
3.5.7 Palladium loading tests on SBL-WT(PMSv)

A 1 mL sample of SBL-WT(PMSv) **3.11** was prepared at 0.75 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.01 M stock solution of Na₂PdCl₄ (Pd²⁺ source) was prepared in Milli-Q® H₂O. To 100 µL aliquots of the above protein solution, was added 1-20 equiv. Pd²⁺ source according to the following table.

Table 3-1. BCA analysis to confirm loss in soluble protein content with SBL-WT(PMSvinyl) on incubation with increasing equiv. of Pd²⁺ precursor

Entry	Starting Concentration [Protein] µM	Starting Concentration [Protein] mg/mL	Pd ²⁺ source equiv.	Protein concentration of post incubation with Pd ²⁺ source [Protein] mg/mL
1	28	0.75	1	0.73
2	28	0.75	2	0.68
3	28	0.75	3	0.62
4	28	0.75	4	0.42
5	28	0.75	5	0.26
6	28	0.75	10	0
7	28	0.75	15	0
8	28	0.75	20	0

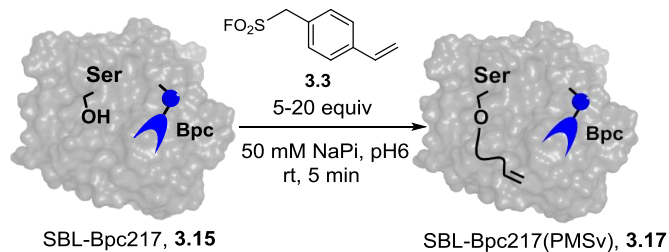
The reactions were incubated at 37 °C for 1 h, after which protein samples from entries 1-5 were analysed by LC-MS. The protein samples were centrifuged at 13,000g for 2 min. After centrifugation, the protein content in each sample was analysed by BCA assay using BSA standards according to the standard procedure (See Section 2.5.5). 10 µL aliquot of each was analysed by non-reducing SDS-PAGE analysis after incubating samples with 1 µL 1M HCl for 30 minutes at 4 °C. The ESI-MS for the modification are shown below.



Engineered cycle

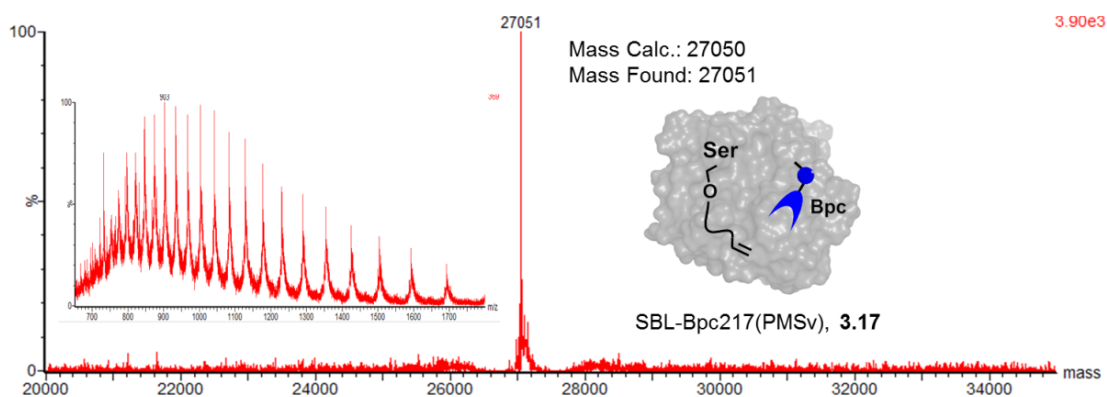
3.5.8 LC-MS assay

Enzyme-substrate interaction of **3.3** with SBL-Bpc217:

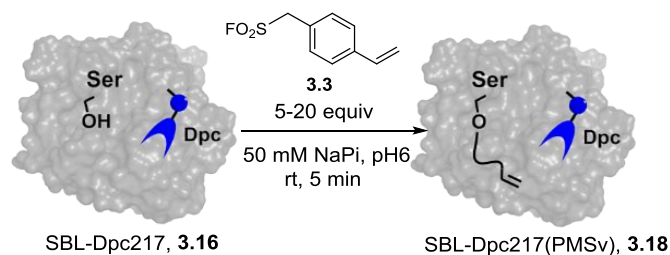


A 2 mL sample of SBL-Dpc217 **3.15** was prepared at 1.5 mg/mL in sodium phosphate buffer (50 mM, pH 6), as described before (See Chapter 2 Section 2.5.7). A 0.5 M stock solution of *p*-vinylPMSF **3.3** was prepared in DMF. To 2 mL sample of SBL-Dpc217 (111

nmol) in 50 mM sodium phosphate buffer, pH 6 at 4°C, was added 2.9 µl of **3.3** in DMF (1.45 µmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217(PMSv) **3.17** (calculated mass 27050, mass found 27051). The excess *p*-vinylPMSF was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was stored on ice until needed and concentration was measured using BCA assay.

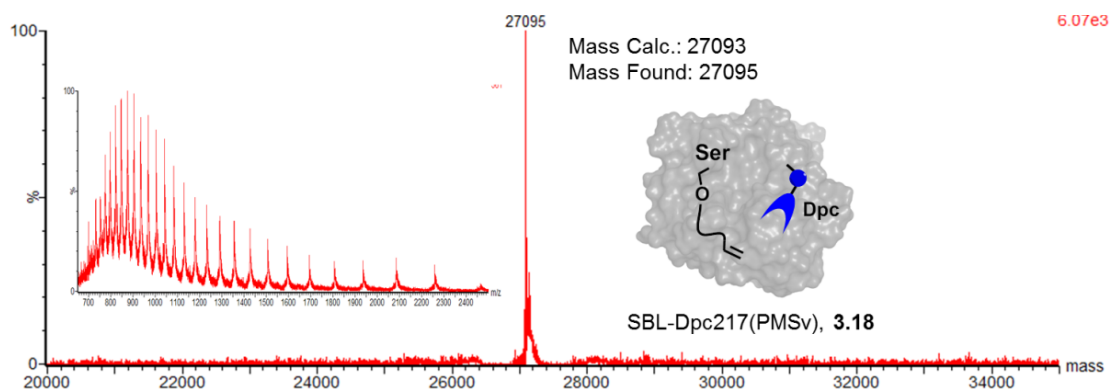


Enzyme-substrate interaction of **3.3** with SBL-Dpc217:



A 2 mL sample of SBL-Dpc217 **3.16** was prepared at 1.5 mg/mL in sodium phosphate buffer (50 mM, pH 6), as described before (See Chapter 2 Section 2.5.7). A 0.5 M stock solution of *p*-vinylPMSF **3.3** was prepared in DMF. To 2 mL sample of SBL-Dpc217 (111 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4°C, was added 2.9 µl of **3.3** in DMF (1.45 µmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-

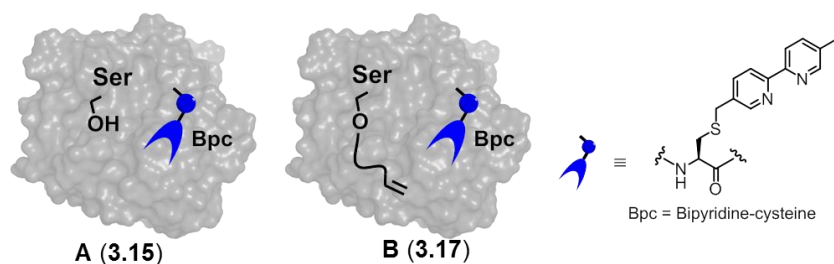
Dpc217(PMSv) **3.18** (calculated mass 27093, mass found 27096). The excess *p*-vinylPMSF was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was stored on ice until needed and concentration was measured using BCA assay.



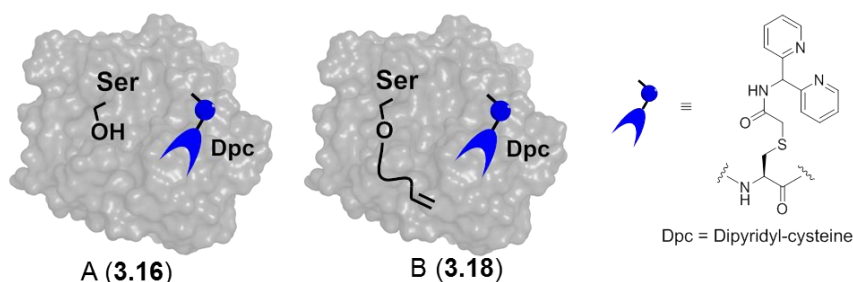
The method for calculating reaction conversions by LC-MS analysis was validated by correlating relative MS peak intensities to mixtures of proteins with known relative concentrations.

This assay is described below:

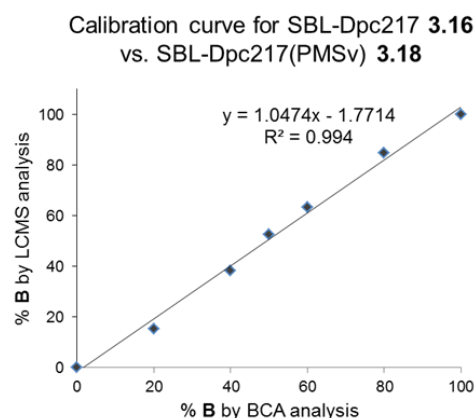
SBL-Bpc217 vs. SBL-Bpc217(PMSv): Seven samples containing a mixture of two different protein species **A** (**3.15**) and **B** (**3.17**) ($c = 0.19$ mg/ mL, measured by standard bicinchoninic acid (BCA) assay from Thermo Scientific) were prepared according to the following **A:B** ratios (5:0, 4:1, 3:2, 1:1, 2:3, 1:4 and 0:5). 20 μ L of each sample was analyzed by LC-MS. The LC-MS spectra of these mixtures are shown below. The calculated masses for **A** and **B** are 26913 and 27093, respectively. A graph of the percentage of **B** by MS against percentage of **B** measured was then plotted (See Scheme 3-12 for plot and ESI-MS spectra).

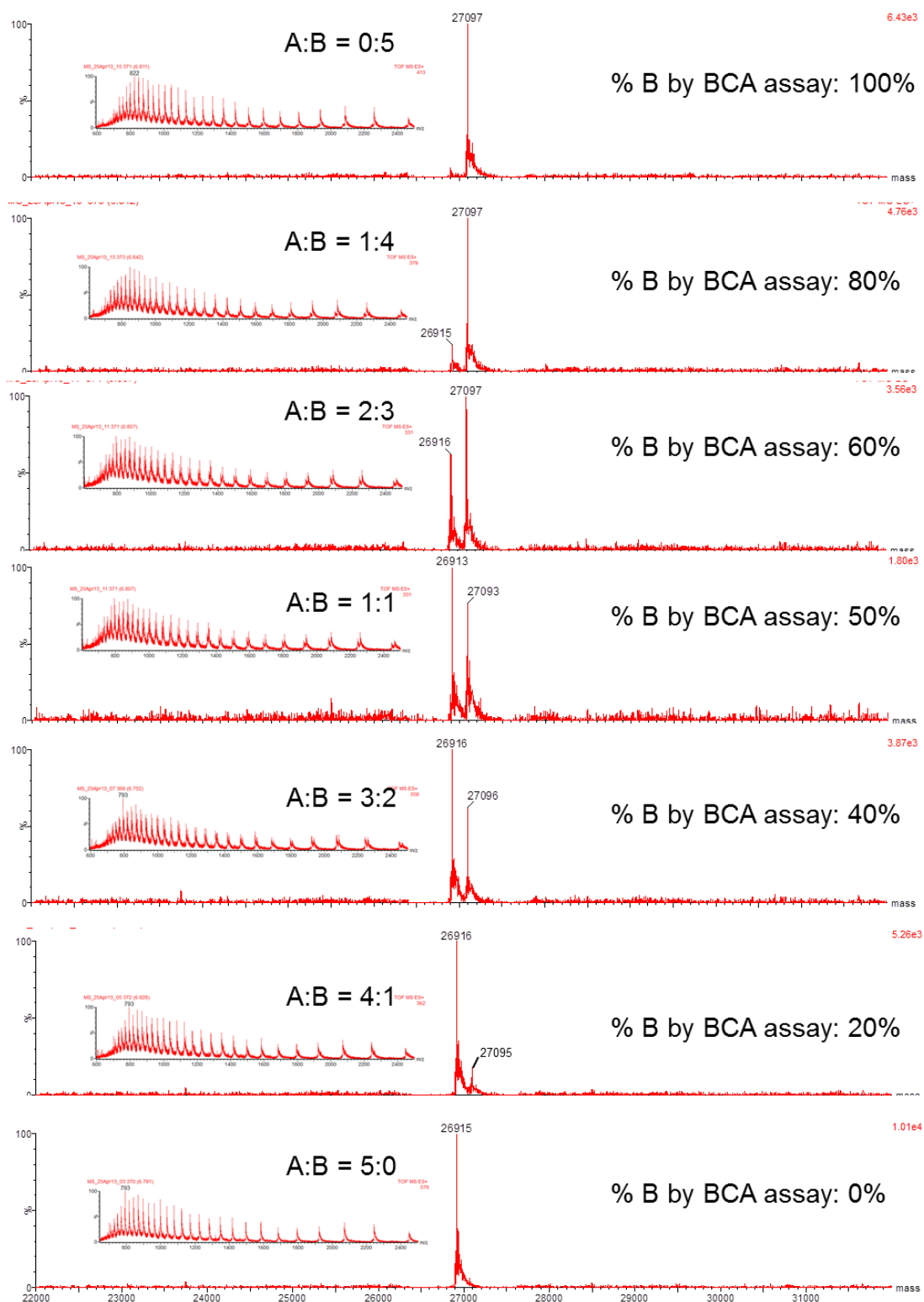


SBL-Dpc217 vs. SBL-Dpc217(PMSv): Seven samples containing a mixture of two different protein species **A (3.16)** and **B (3.18)** ($c = 0.19 \text{ mg/mL}$, measured by standard bicinchoninic acid (BCA) assay from Thermo Scientific) were prepared according to the following **A:B** ratios (5:0, 4:1, 3:2, 1:1, 2:3, 1:4 and 0:5). 20 μL of each sample was analyzed by LC-MS. The LC-MS spectra of these mixtures are shown below. The calculated masses for **A** and **B** are 26913 and 27093, respectively. A graph of the percentage of **B** by MS against percentage of **B** measured was then plotted.



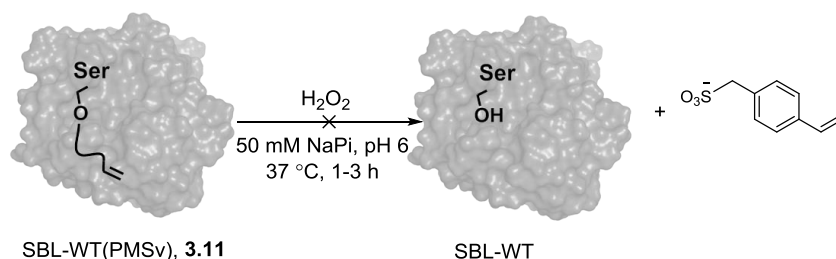
% B by LCMS analysis	% B by BCA analysis
0	0
20	15.3
40	38.3
50	52.6
60	63.3
80	84.7
100	100



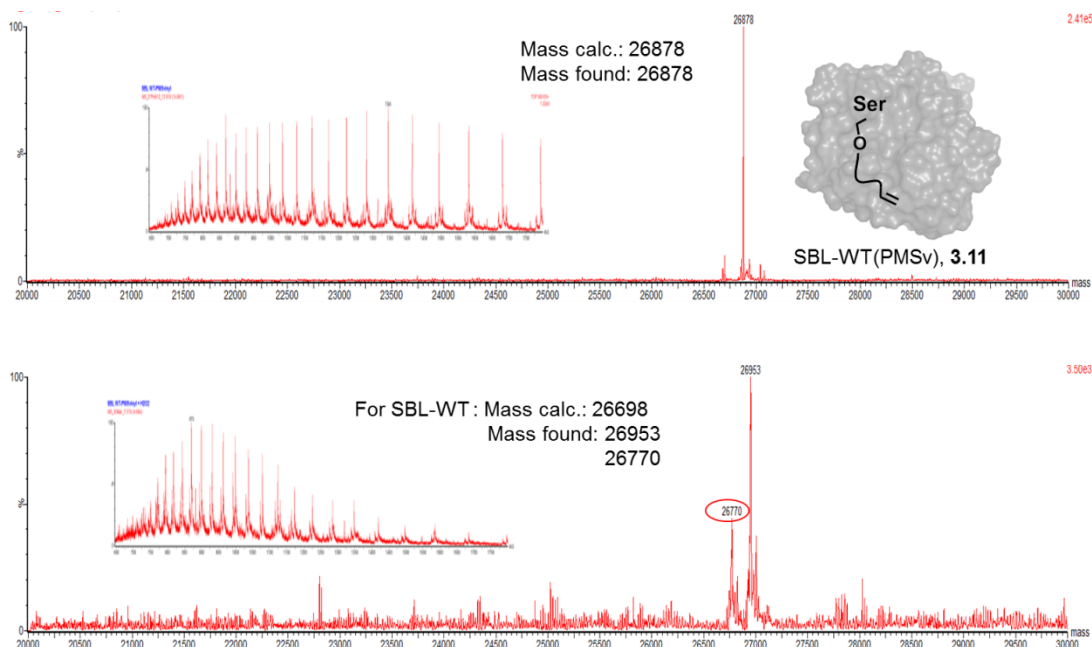


3.5.8 Hydrolysis of alkanesulfonyl group from SBL-WT(PMSv)

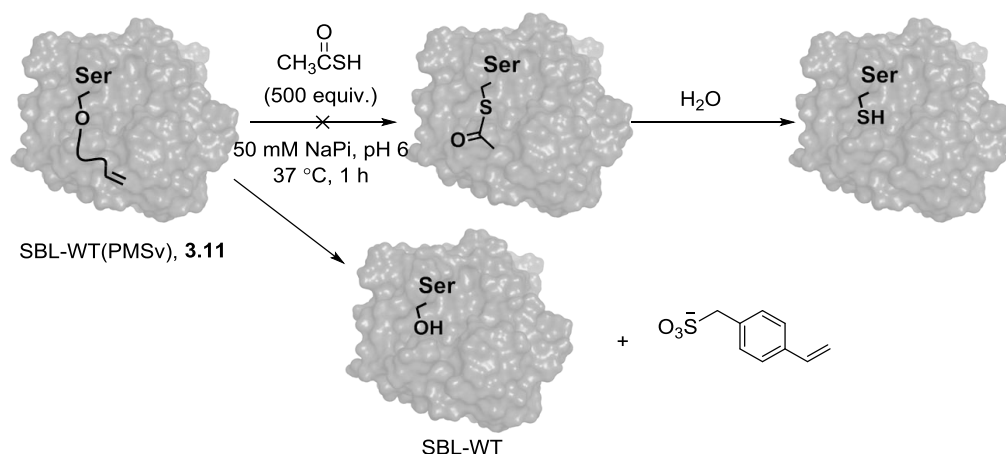
3.5.8.1 Hydrolysis test using Hydrogen peroxide⁴⁰



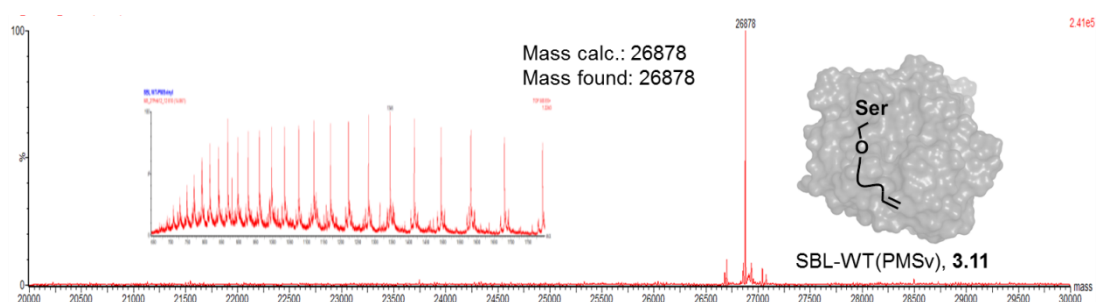
To 50 μL sample of SBL-WT(PMSv) (4.1 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.62 μL of H_2O_2 solution (0.31 μmol). The reaction solution was vortexed to homogenize and the reaction was incubated at 37 $^\circ\text{C}$ for 1 h. After reaction, 5 μL aliquot of the reaction was analysed by LC-MS. A mixture of protein species was detected by LC-MS analysis corresponding to the masses 26953 and 26770 Da (for SBL-WT mass calculated: 26698). Partial hydrolysis of the substrate seen with +72 adduct perhaps due to non-specific oxidation.

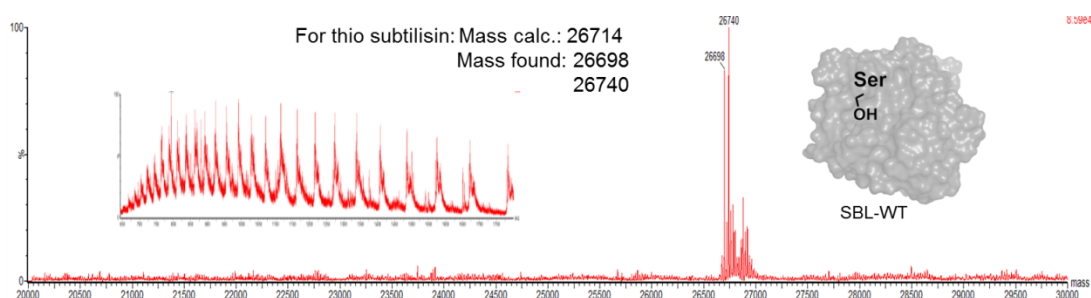


3.5.8.2 Hydrolysis test using thiol acetate⁴²



A 2 mL sample of SBL-WT(PMSv) **3.11** was prepared at 2.2 mg/mL in sodium phosphate buffer (50 mM, pH 6), as described before. A 0.5 M stock solution of thiolacetate was prepared in DMF. To 50 μL sample of SBL-WT(PMSv) (4.1 nmol) in 50 mM sodium phosphate buffer pH 6, was added 4.1 μL of thiol acetate in DMF (2.05 μmol). The reaction solution was vortexed to homogenize and the reaction was incubated at 37 °C for 1 h. After reaction, 5 μL aliquot of the reaction was analysed by LC-MS. A mixture of protein species was detected by LC-MS analysis corresponding to the mass of SBL-WT (mass calculated; 26698, mass found: 26698 Da) and +42 adduct that could not be assigned (mass calculated; 26698, mass found: 26740 Da).





3.5.8.3 Hydrolysis test in aqueous solutions⁴³

Optimisation of conditions for hydrolysis of the alkane sulfonyl group to release the active site serine protease was accomplished by performing test hydrolysis on SBL-WT(PMSv). A screening of different buffers over a range of pH was performed to hydrolyse the sulfonyl group from the enzyme. A complete set of conditions and results obtained by using this method has been summarised in Table below.

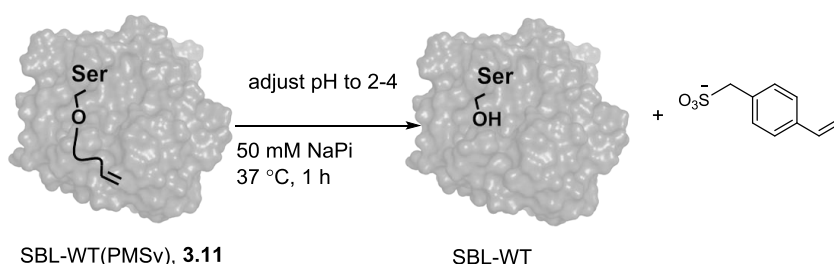


Table 3-2. Conditions for optimization of hydrolysis of alkane sulfonyl group from Ser221

Entry	Buffer Solution	pH	% Conversion		
			After buffer exchange column, 4°C	37 °C, 1.5 h	37 °C, 3 h
1	NaCl-HCl	2	77%	100%	-
2	Citrate (Na)	3	100%	-	-
3	Citrate (Na)	4	15%	-	-
4	Citrate (Na)	5	20%	72%	100%
5	TRIS-HCl	8	-	7%	10%
6	Carbonate (Na)	10	-	-	-

Typically, 100 μ L sample of SBL-WT(PMSv) was prepared in 50 mM sodium phosphate buffer, pH 6. A PD10 G-25 spin trap column was pre-equilibrated with the desired buffer solution. SBL-WT(PMSv) was passed through the pre-equilibrated PD10 G-25 spin trap

column to exchange the buffer. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT (calculated mass 26698, mass found 26699) showing release of the bound PMSv substrate. The ESI-MS for different buffer conditions are shown below.

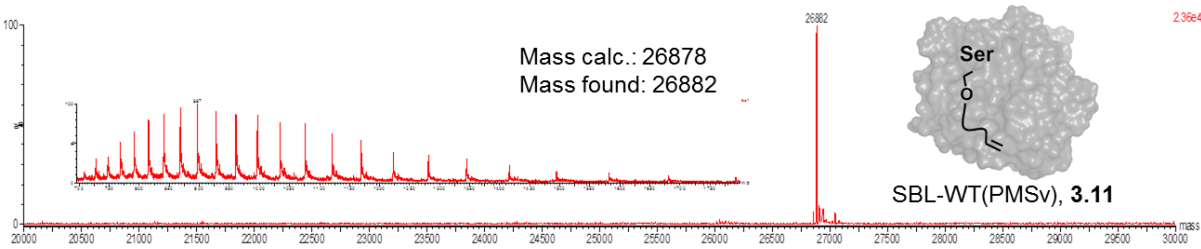


Table 3-2, Entry 1: NaCl-HCl solution pH 2

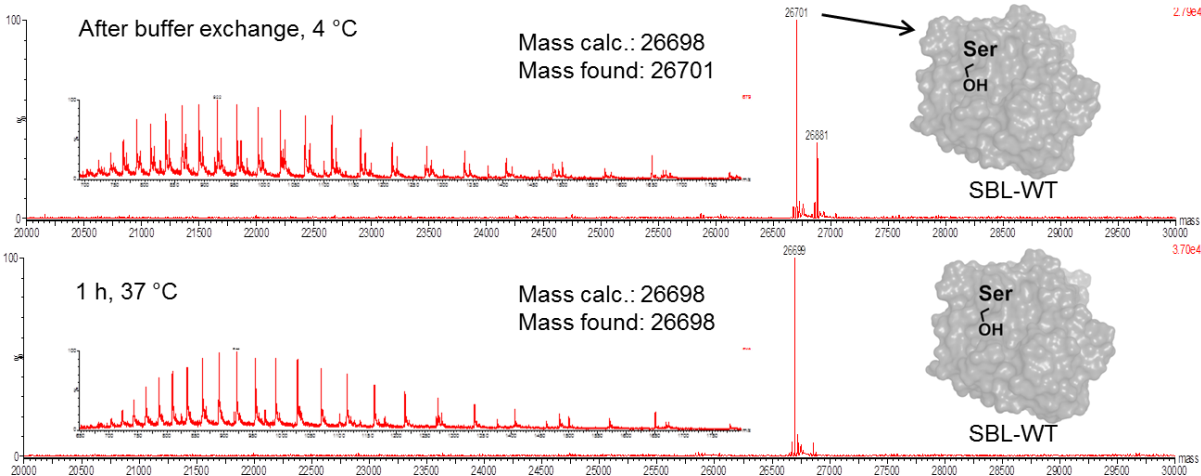


Table 3-2, Entry 2: Citrate (Na) solution pH 3

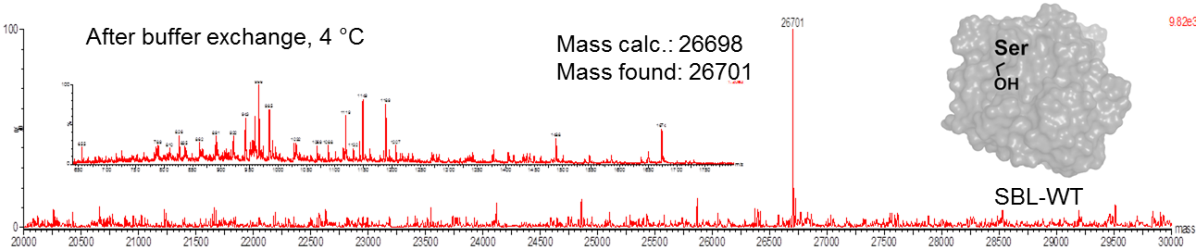


Table 3-2, Entry 4: Citrate (Na) solution pH 5

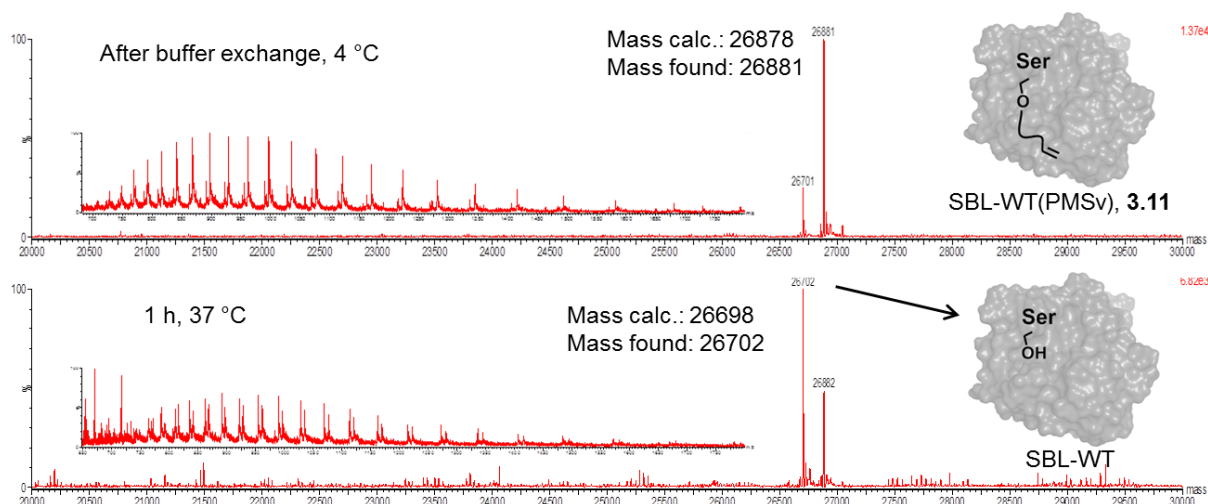


Table 3-2, Entry 5: TRIS-HCl solution pH 8

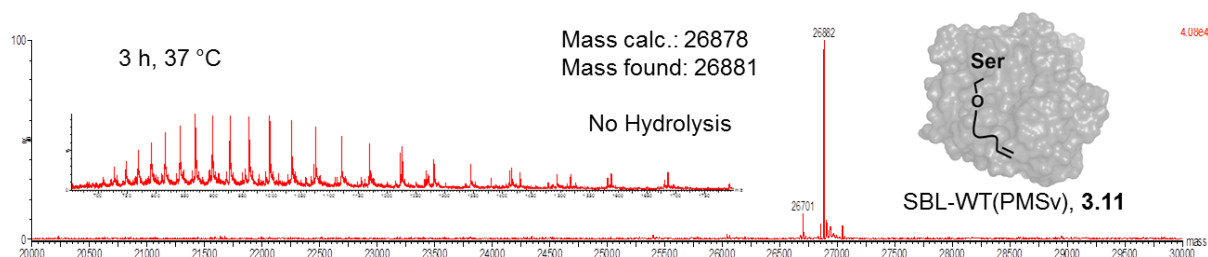
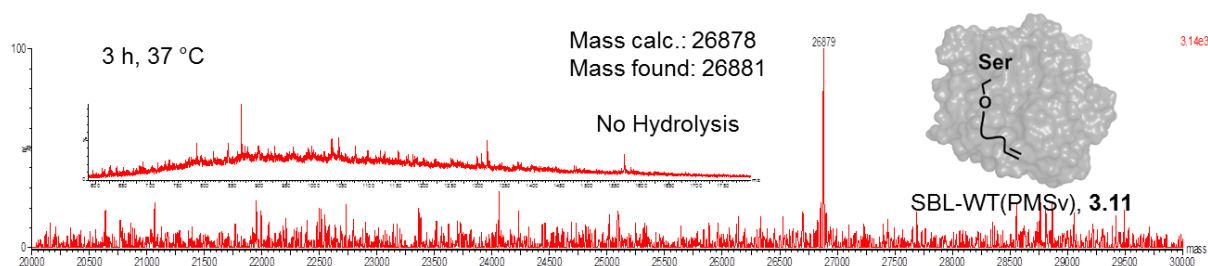
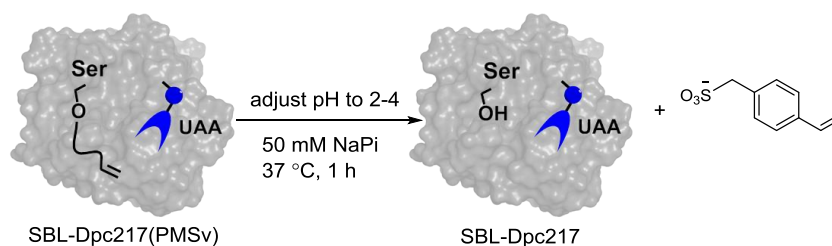


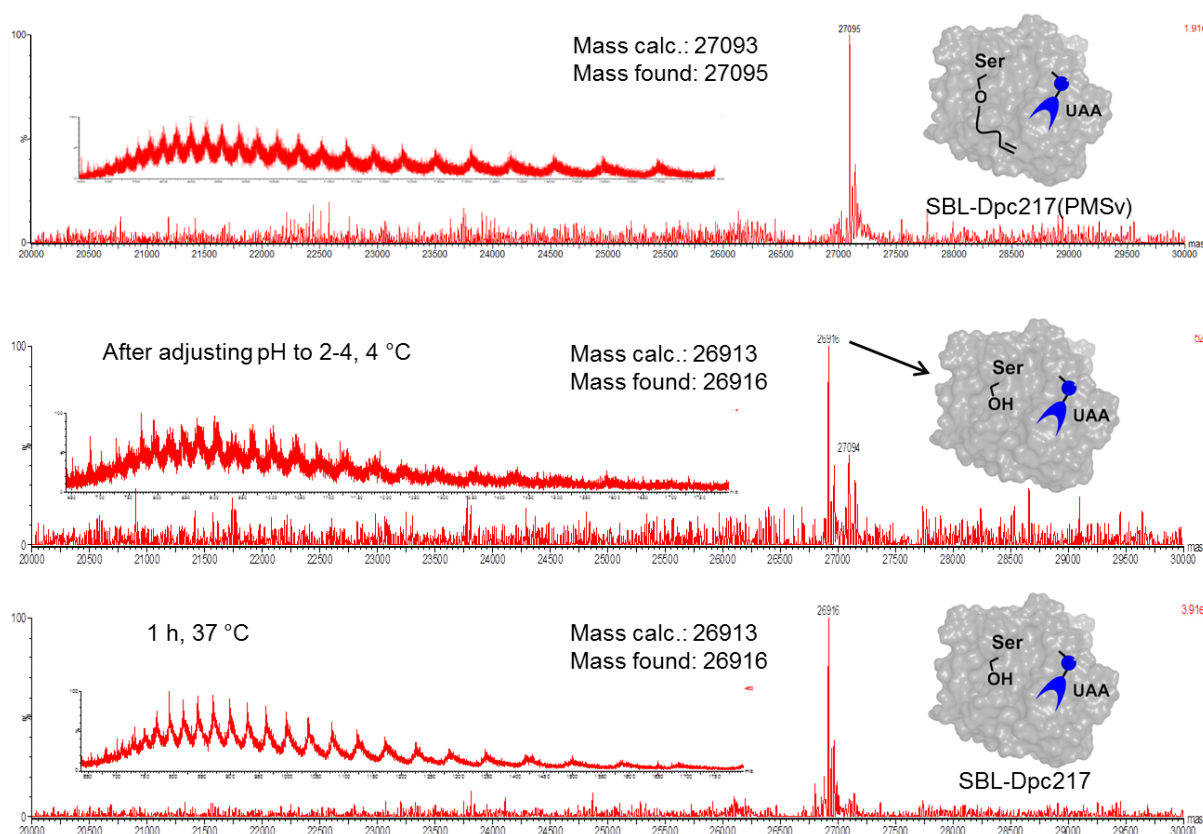
Table 3-2, Entry 6: Carbonate (Na) solution pH 10



3.5.9 Hydrolysis of SBL-Dpc217(PMSv) by adjusting pH to 2-4

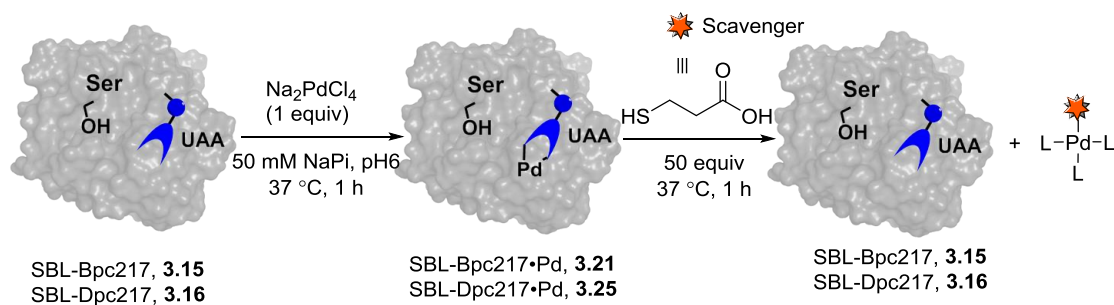


A 100 μL sample of SBL-Dpc217(PMSv) was prepared in 50 mM sodium phosphate buffer, pH 6. The pH of the solution was adjusted to 2-4 using 1 M HCl. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}\text{C}$ for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217 (calculated mass 26913, mass found 26916) showing release of the bound PMSv substrate. The ESI-MS for the modification are shown below.



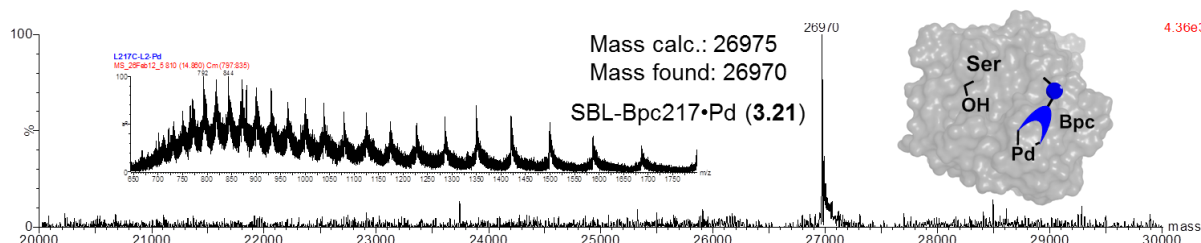
Metal Scavenging: optimization

3.5.10 Palladium Scavenging test using SBL-Bpc217•Pd and SBL-Dpc217•Pd



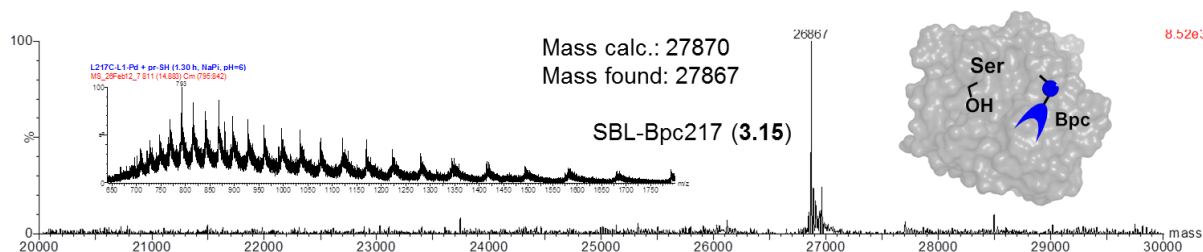
Palladium Scavenging test using SBL-Bpc217•Pd

A 100 μL sample of chemically modified SBL-Bpc217 **3.15** was prepared at 1.0 mg/mL in 50 mM sodium phosphate buffer pH 6, and the enzyme was stored on ice until needed. A 10 mM stock solution of Na_2PdCl_4 was prepared in Milli-Q® water and 0.5 M stock solution of 3-mercaptopropionic acid (MPA) was prepared in DMF. SBL-Bpc217•Pd **3.21** was prepared by following standard procedure (See Experimental Section 2.5.9). To 100 μL sample of SBL-Bpc217 (3.7 nmol) in 50 mM phosphate buffer, pH 6 at 4 °C, was added 0.37 μL of Na_2PdCl_4 (3.7 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. The reaction was monitored LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26976) showing uptake of palladium.



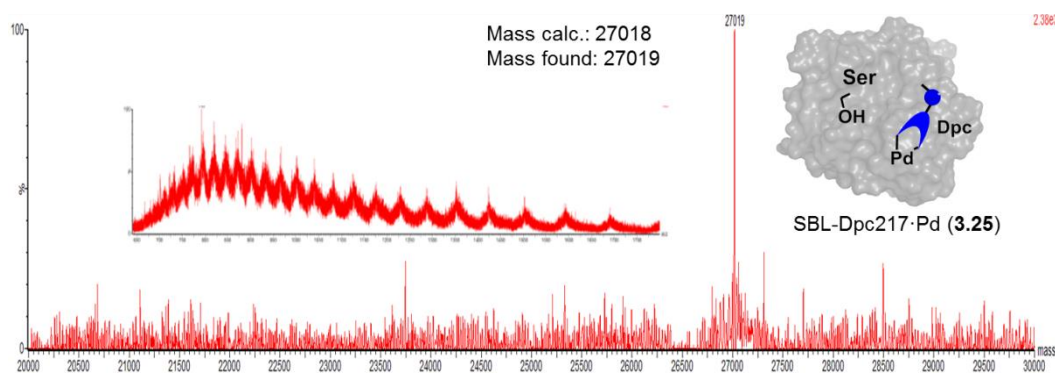
To 70 μL SBL-Bpc217•Pd (2.28 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.25 μL of MPA solution (114 nmol). The reaction was incubated at 37 °C for 1 h

and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Bpc217 **3.15** (calculated mass 26870, mass found 26871) showing scavenging of palladium from the active site.

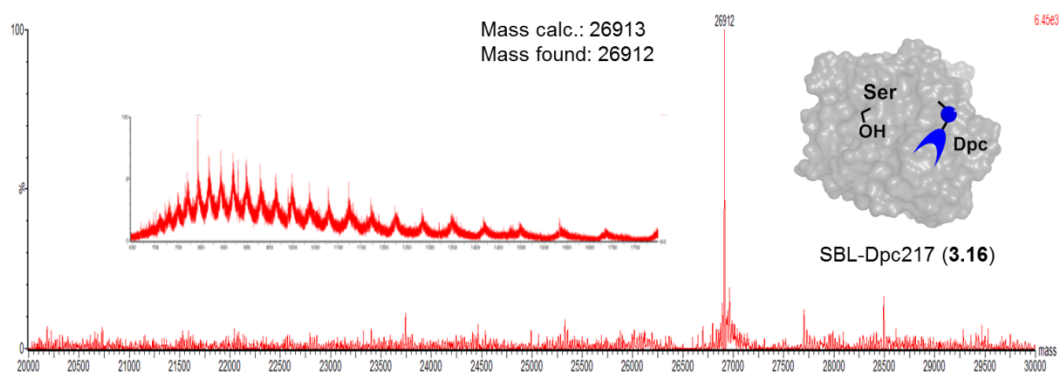


Palladium Scavenging test using SBL-Dpc217•Pd

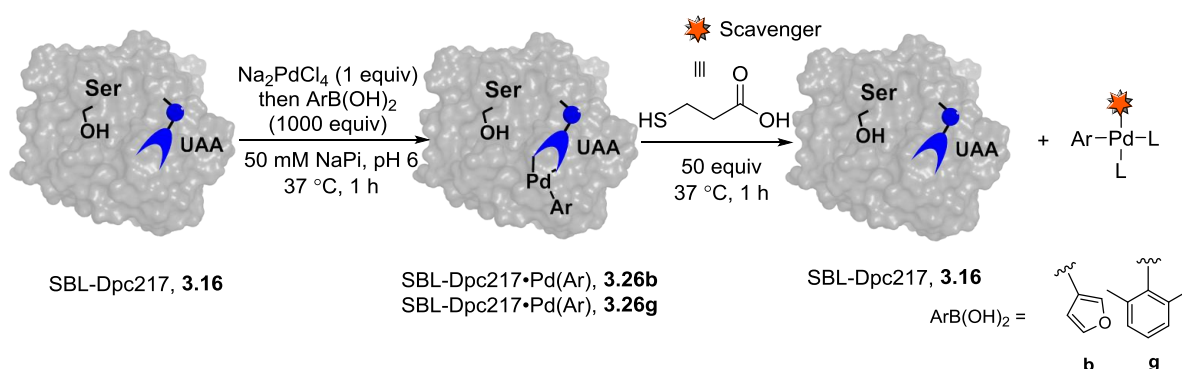
A 100 μ L sample of chemically modified SBL-Dpc217 **3.16** was prepared at 1.0 mg/mL in 50 mM sodium phosphate buffer pH 6, and the enzyme was stored on ice until needed. A 10 mM stock solution of Na_2PdCl_4 was prepared in Milli-Q[®] water and 0.5 M stock solution of 3-mercaptopropionic acid (MPA) was prepared in DMF. SBL-Dpc217•Pd **3.25** was prepared by following standard procedure (See Experimental Section 2.5.9). To 100 μ L sample of SBL-Dpc217 (3.7 nmol) in 50 mM phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 0.37 μ L of Na_2PdCl_4 (3.7 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 $^{\circ}\text{C}$ for 1 h. The reaction was monitored LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27019, mass found 27019) showing uptake of palladium.



To 70 μL SBL-Dpc217•Pd (2.28 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.25 μL of MPA solution (114 nmol). The reaction was incubated at 37 $^{\circ}\text{C}$ for 1 h and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Dpc217 **3.16** (calculated mass 26913, mass found 26912) showing scavenging of palladium from the active site.



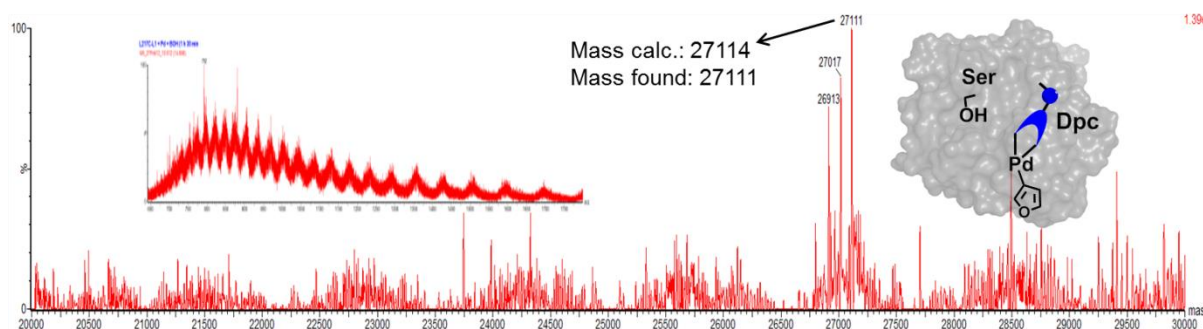
3.5.11 Palladium scavenging test using SBL-Dpc217•Pd(Ar)



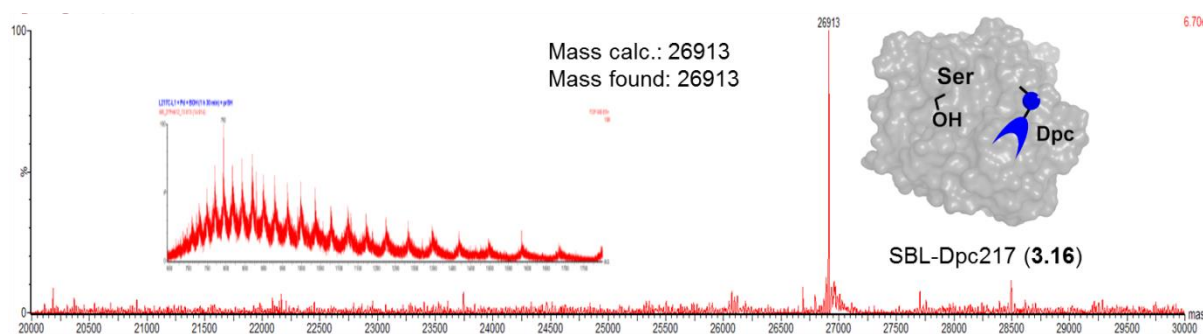
Palladium Scavenging test using SBL-Dpc217•Pd(Ar); [Ar = **b**]

A 100 μL sample of chemically modified SBL-Dpc217 **3.16** was prepared at 0.87 mg/mL in 50 mM sodium phosphate buffer pH 6, and the enzyme was stored on ice until needed. A 10 mM stock solution of Na_2PdCl_4 was prepared in Milli-Q® water and 0.5 M stock solution of arylboronic acid **a** and 3-mercaptopropionic acid (MPA) was each prepared in DMF. To 100 μL sample of SBL-Dpc217 (3.2 nmol) in 50 mM phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 0.32 μL of Na_2PdCl_4 (3.2 nmol), followed by addition of 6.4 μL of

arylboronic acid **b**. The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1.5 h. The reaction was monitored LC-MS. A mixture protein species was detected with a mass corresponding to the SBL-Dpc217•Pd(Ar) **3.26b** (calculated mass 27124, mass found 27122) and SBL-Dpc217•Pd **3.25** (calculated mass 26018, mass found 26018) along with unreacted starting material.



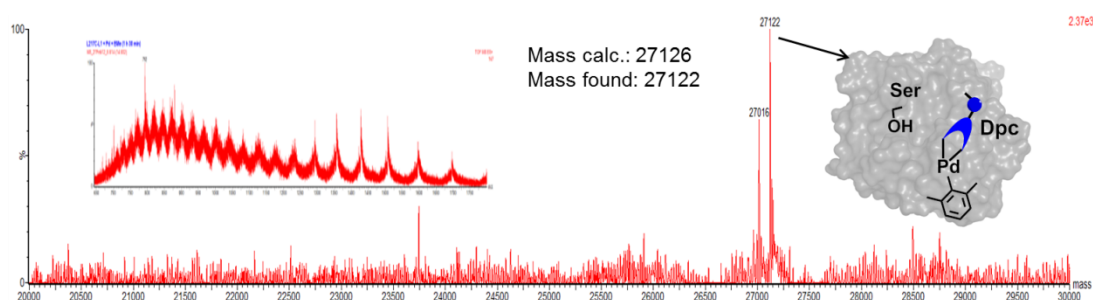
To 70 μ L SBL-Dpc217•Pd(Ar) (2.28 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.25 μ L of MPA solution (114 nmol). The reaction was incubated at 37 °C for 1 h and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Dpc217 **3.16** (calculated mass 26913, mass found 26913) showing scavenging of palladium from the active site.



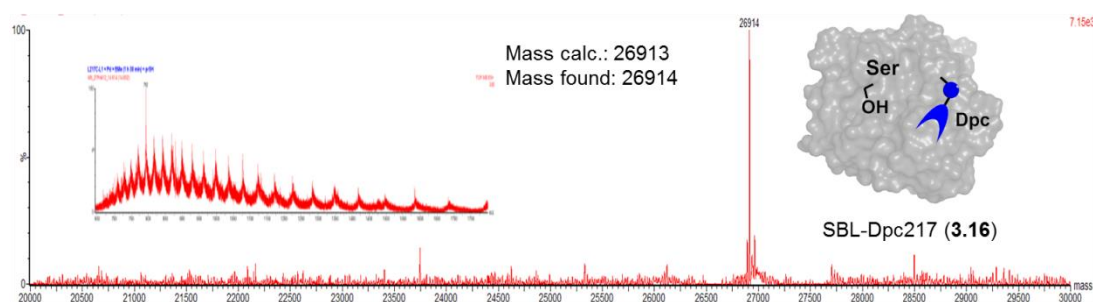
Palladium Scavenging test using SBL-Dpc217•Pd(Ar); [Ar = g]

SBL-Dpc217•Pd(Ar) **3.26g** was prepared by following the same procedure as above by using arylboronic acid **g**. LC-MS analysis revealed a mixture of SBL-Dpc217•Pd(Ar)

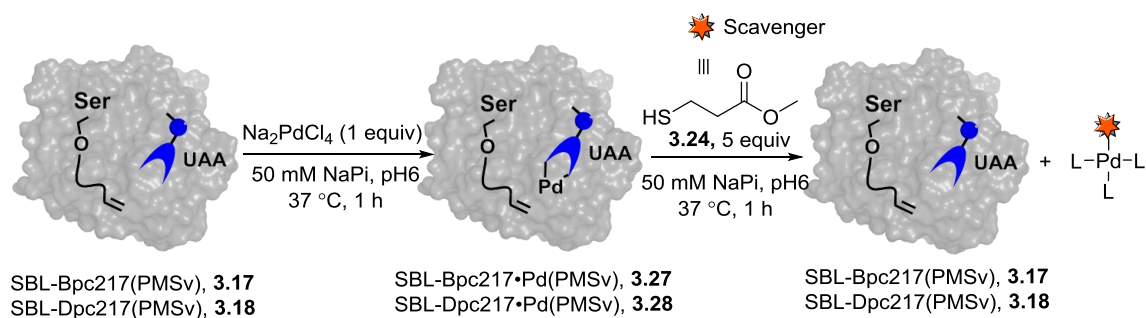
3.26g (calculated mass 26122, mass found 26126) and SBL-Dpc217•Pd **3.25** (calculated mass 26018, mass found 26018) along with unreacted starting material.



To 70 μ L SBL-Dpc217•Pd(Ar) mixture (2.28 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.25 μ L of MPA solution (114 nmol). The reaction was incubated at 37 $^{\circ}$ C for 1 h and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Dpc217 **3.16** (calculated mass 26913, mass found 26913) showing scavenging of palladium from the active site.

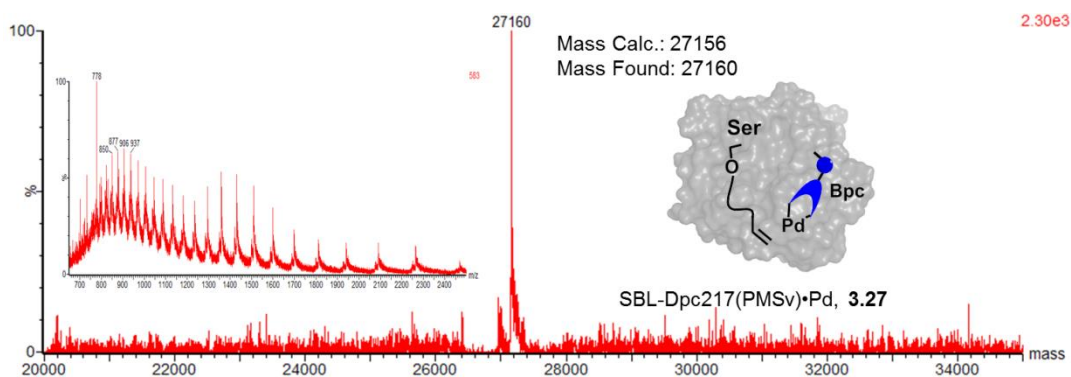


3.5.12 Selective metal scavenging test using SBL-Dpc217•Pd(PMSv), without hydrolysis of alkanesulfonyl group

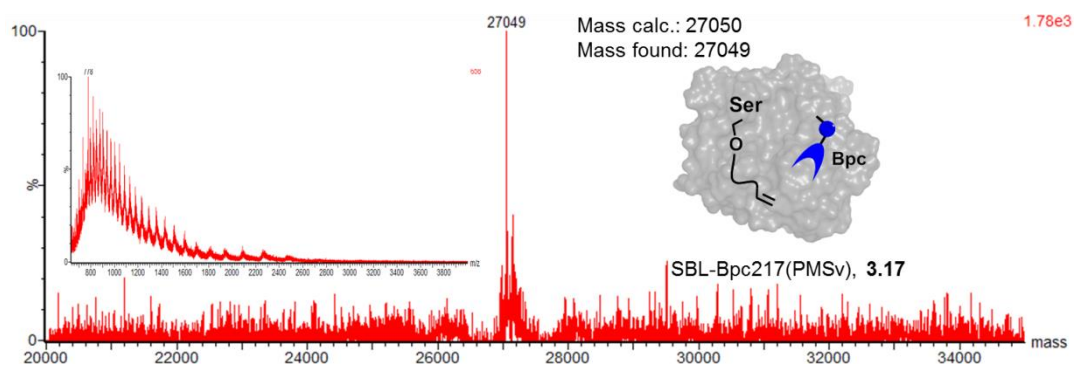


Palladium Scavenging test using SBL-Bpc217•Pd(PMSv)

A 200 μL sample of chemically modified SBL-Bpc217(PMSv) **3.17** was prepared at 0.8 mg/mL in 50 mM sodium phosphate buffer pH 6, and the enzyme was stored on ice until needed. A 10 mM stock solution of Na_2PdCl_4 was prepared in Milli-Q® water and 0.1 M stock solution of Methyl, 3-mercaptopropionate ester (MPE) was prepared in DMF. To 200 μL sample of SBL-Bpc217(PMSv) (6 nmol) in 50 mM phosphate buffer, pH 6 at 4 °C, was added 0.6 μL of Na_2PdCl_4 (6 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. The reaction was monitored LC-MS and a single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd(PMSv) **3.27** (calculated mass 27156, mass found 27160).

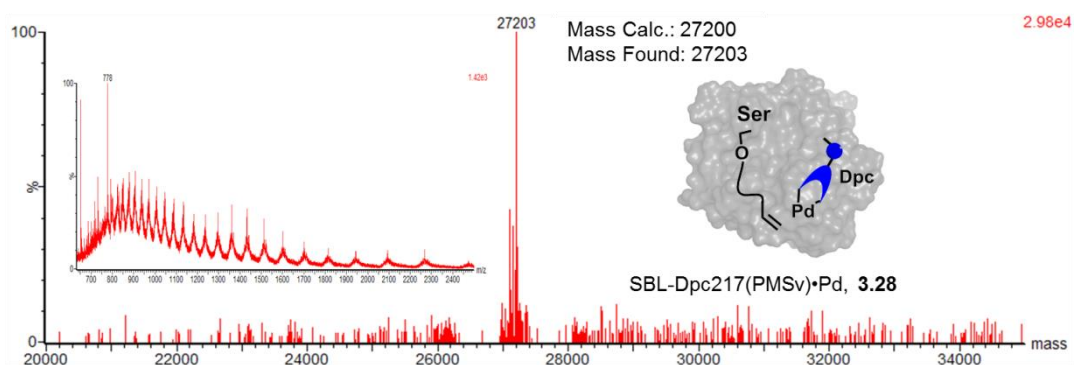


To 200 μL SBL-Bpc217•Pd(PMSv) (6 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.6 μL of MPE solution (60 nmol). The reaction was incubated at 37 °C for 1 h and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Bpc217(PMSv) **3.17** (calculated mass 27050, mass found 27049) showing selective scavenging of palladium from the active site, without hydrolysis of tethered PMSv substrate.



Palladium Scavenging test using SBL-Dpc217•Pd(PMSv)

SBL-Dpc217•Pd(PMSv) **3.28** was prepared by following the same procedure as above. Upon palladium complexation, LC-MS analysis revealed a single protein species with mass corresponding to SBL-Dpc217•Pd(PMSv) **3.28** (calculated mass 27200, mass found 27203). ESI-MS for the modification is shown below.



To 200 μ L SBL-Dpc217•Pd(PMSv) (6 nmol) in 50 mM sodium phosphate buffer pH 6, was added 0.6 μ L of MPE solution (60 nmol). The reaction was incubated at 37 $^{\circ}$ C for 1 h and was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Dpc217(PMSv) **3.18** (calculated mass 27093, mass found 27094) showing selective scavenging of palladium from the active site, without hydrolysis of tethered PMSv substrate.

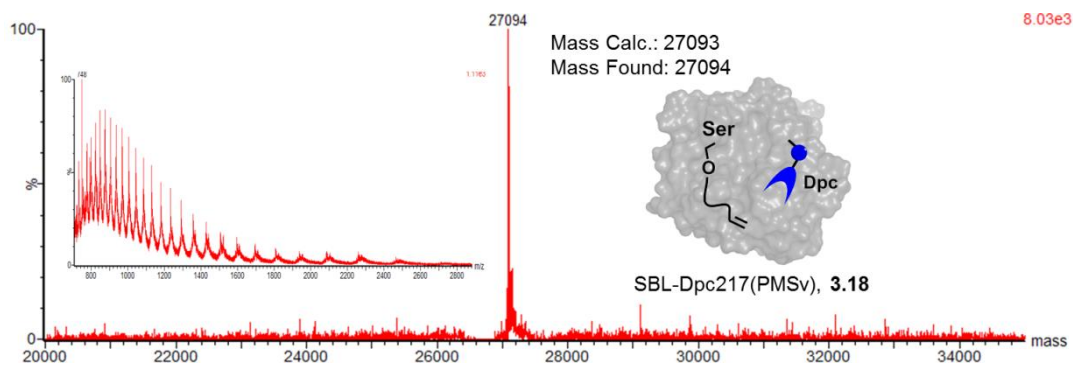
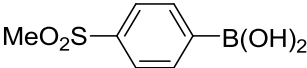
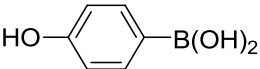


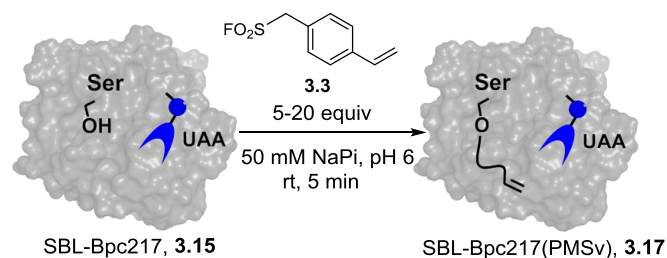
Table 3-3. Oxidative Heck-type cross-coupling reaction with different boronic acid substrates^[a]

Entry	Enzyme SBL-	Alkene substrate	ArB(OH) ₂	Enzyme-Product Intermediate [b]	% conversion [c]
1	Bpc217	3.3		3.19a	>95
2	Bpc217	3.3		3.19b	80
3	Bpc217	3.3		3.19c	67
4	Bpc217	3.3		3.19d	83
5	Bpc217	3.3		3.19e	72
6	Bpc217	3.3		3.19f	57
7	Dpc217	3.3		3.29a	>95
8	Dpc217	3.3		3.29b	83
9	Dpc217	3.3		3.29c	>95

10	Dpc217	3.3		3.29d	>95
11	Dpc217	3.3		3.29e	>90

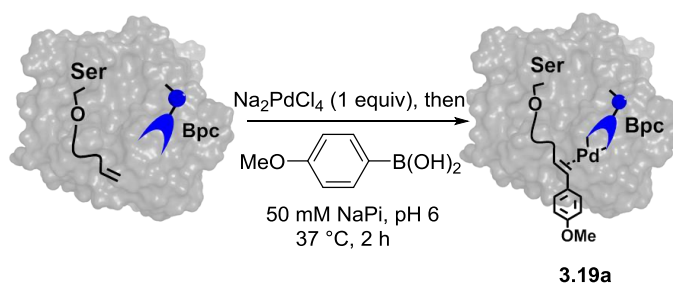
^[a]Cross coupling reactions were carried out using typical conditions: 30 μ M SBL-UAA217 tethering the alkenyl substrate **3.3**, 30 mM ArB(OH)_2 and 30 μ M Na_2PdCl_4 in 50 mM NaPi (pH 6) at 37 $^\circ\text{C}$ for 2 h. ^[b]Enzyme reactions were monitored by LC-MS using a Phenomenex Jupiter 5 μ m C4 (250 $\times$ 4.6 mm) column. ^[c]% conversions as observed by LC/MS analysis. ^[d]Final coupled product was isolated after a hydrolysis step. ^[e]Retention time of the final coupled product as analysed by rpHPLC using a Phenomenex Synergi 4 μ m Hydro-RP C18 (100 $\times$ 4.6 mm) column.

3.5.13 Oxidative Heck-type cross-coupling reaction using SBL-Bpc217

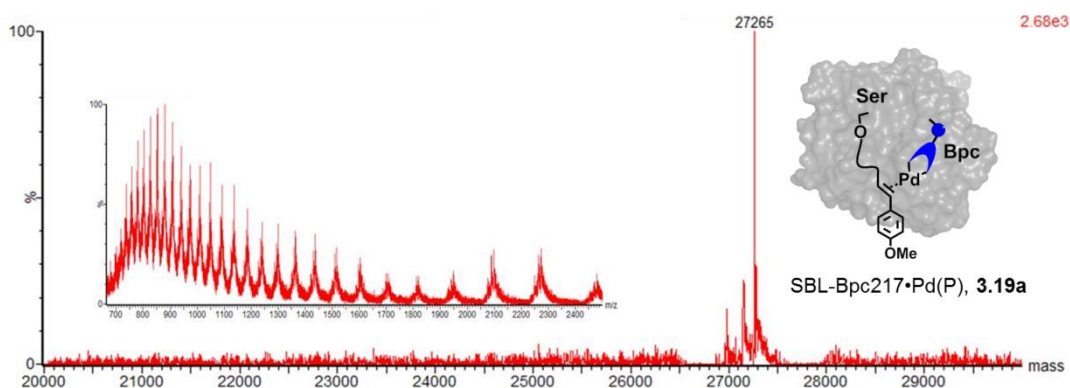


A 800 μ L sample of SBL-Bpc217 was prepared at 1.75 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.5 M stock solution of alkene **3.3** was prepared in DMF. To 800 μ L sample of SBL-Bpc217 (52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^\circ\text{C}$, was added 5.2 μ L of alkene **3.3** in DMF (2.6 μ mol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217(PMSv) **3.17** (calculated mass 27050, mass found 27051). The excess alkene was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was dialysed against 50 mM NaPi pH 6 for 2 h and subsequently stored on ice until needed and concentration was measured using BCA assay.

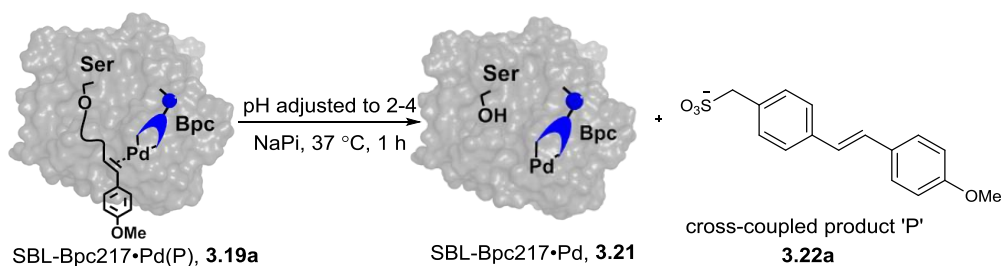
Table 3-3, Entry 1: using 4-methoxyphenylboronic acid



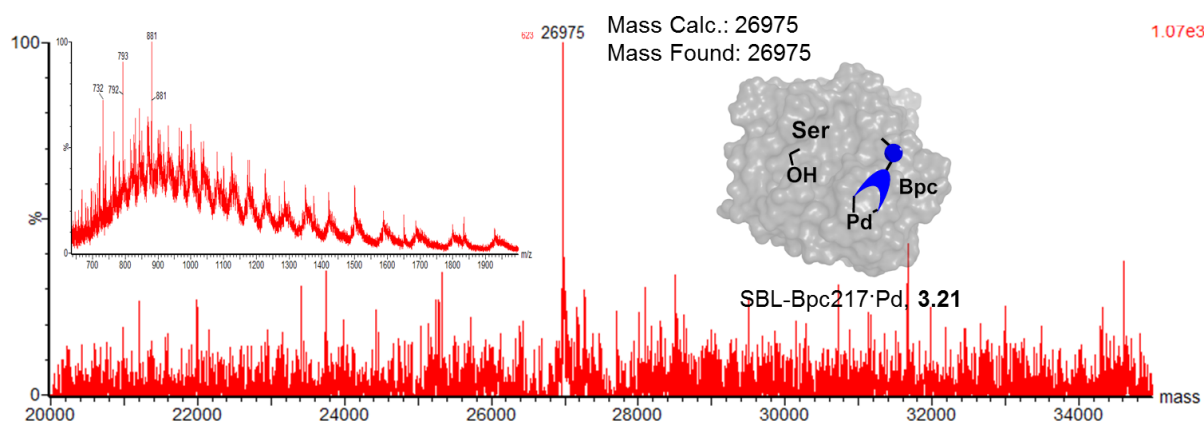
A 200 μL sample of SBL-Bpc217(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μL sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μL of 4-methoxyphenylboronic acid in DMF (5.5 μmol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19a** was detected with a mass corresponding to the SBL-Bpc217•Pd(P) anchoring coupled product (mass calculated: 27262, mass found: 27265). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



Hydrolysis of final product 3.22a



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19a** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26974) showing release of the final coupled product **3.22a**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22a** was isolated and analysed by rpHPLC from a combined volume of two reactions (total volume 300 μ L). The freeze dried sample was taken and re-dissolved in 30 μ L Milli-Q® water. The final product sample was vortexed, sonicated gently and centrifuged. The sample (20 μ L) was separated on a Phenomenex Synergi-Hydro-RP C18 column using HPLC method A. The final product was analysed by

measuring UV absorbance at $\lambda=319$ nm. The standard chromatogram for the separation is given below. $t_R=15.025$; HRMS m/z (ESI): Found 303.0702 $[M-Na]^+$; $C_{15}H_{13}O_4S$ requires 303.0697.

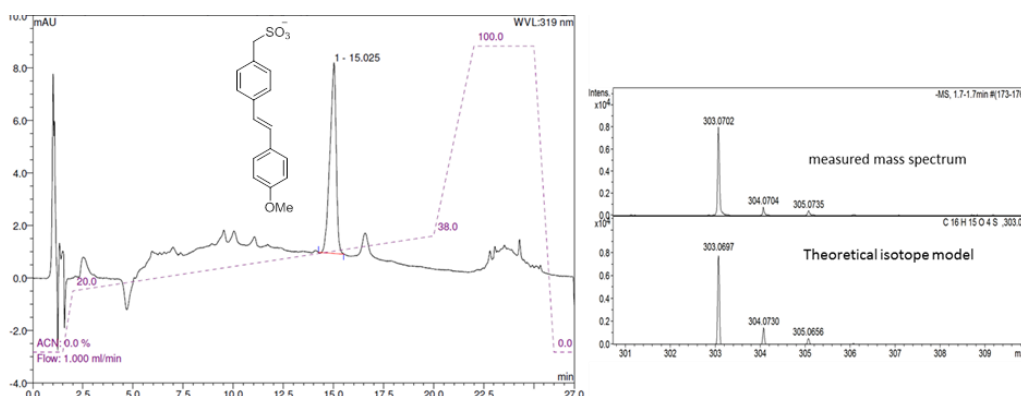
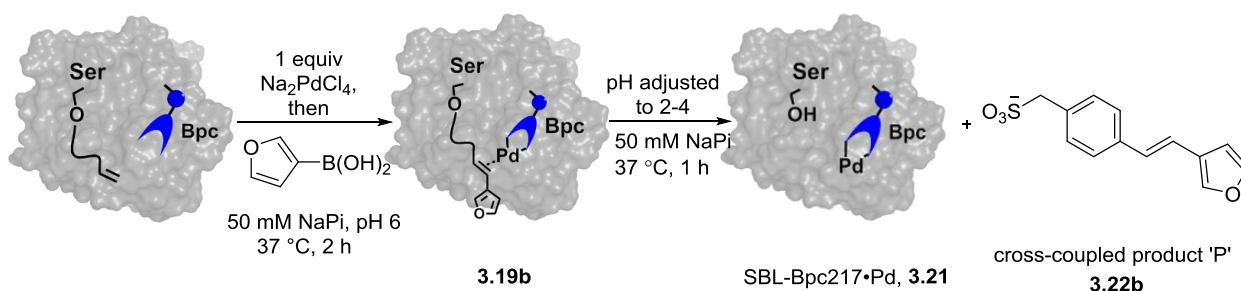
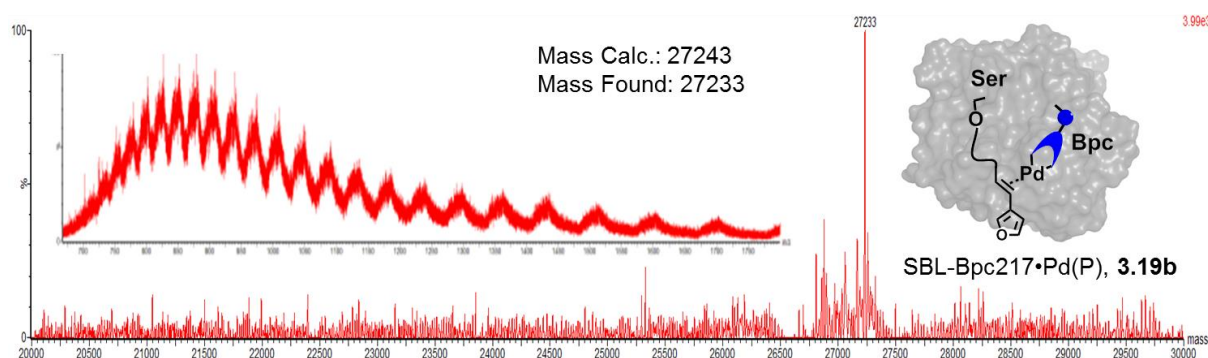


Table 3-3, Entry 2: using 3-furanboronic acid

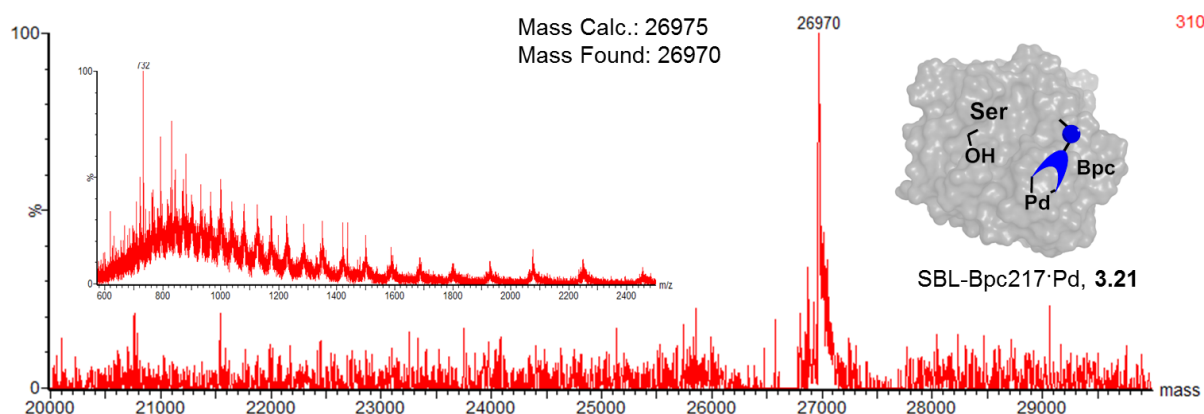


Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of 3-furanboronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19b** was detected (mass calculated: 27224, mass found: 27231). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The

protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19b** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26970) showing release of the final coupled product **3.22b**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 306 nm) required a scale-up of the protein reactions. Final coupled product **3.22b** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below. t_R = 13.37.

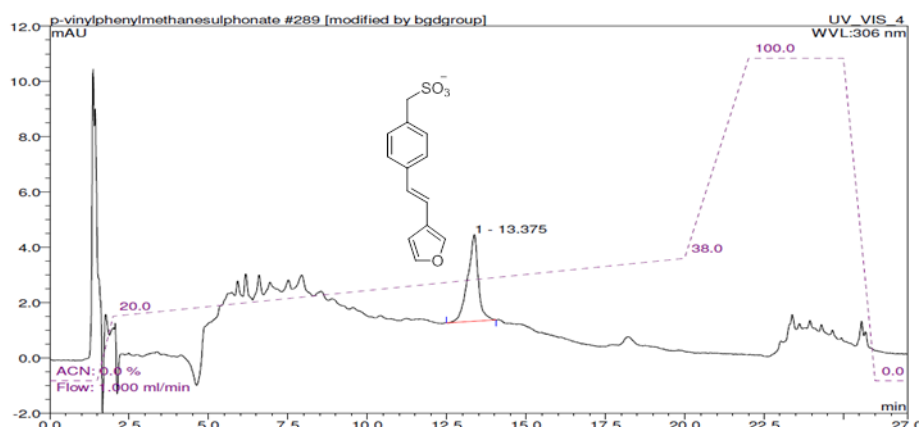
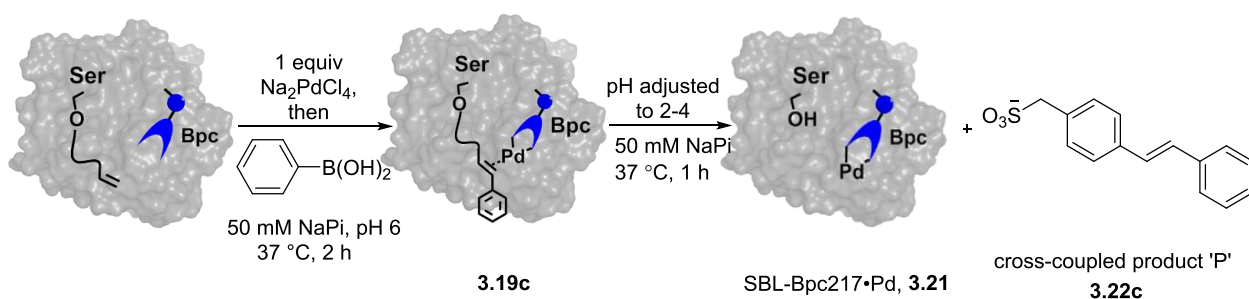
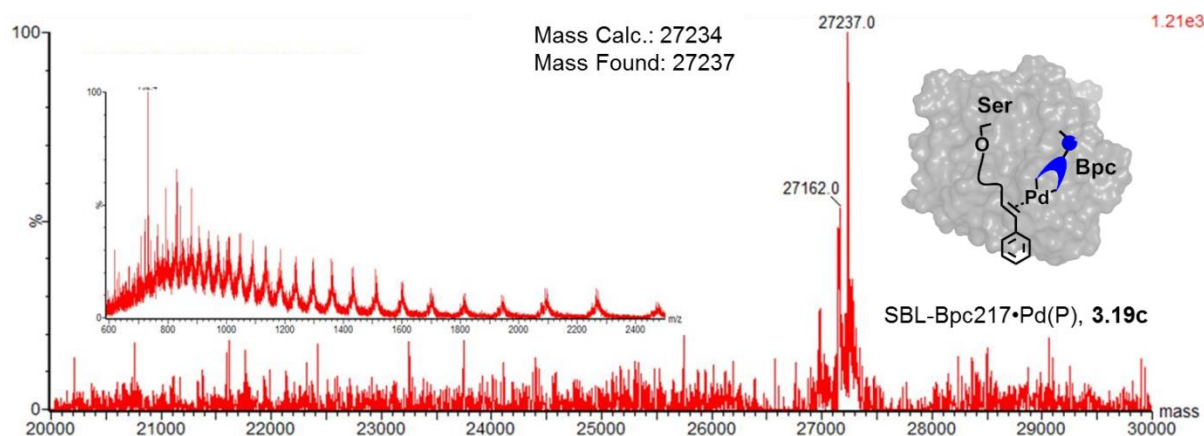


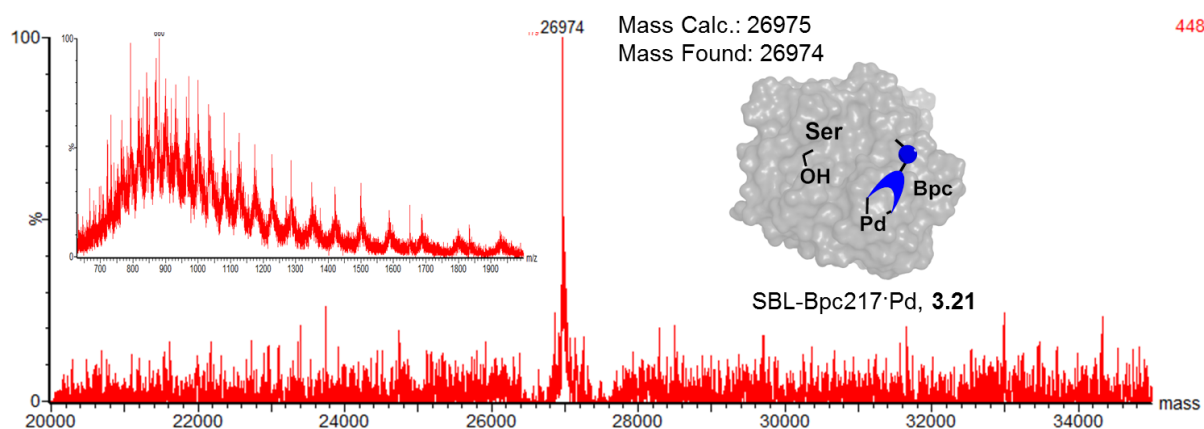
Table 3-3, Entry 3: using phenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.56 μ L of Na₂PdCl₄ (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of phenylboronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19c** was detected (mass calculated: 27234, mass found: 27237). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19c** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26974) showing release of the final coupled product **3.22c**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 311 nm) required a scale-up of the protein reactions. Final coupled product **3.22c** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below.
 t_R = 12.91.

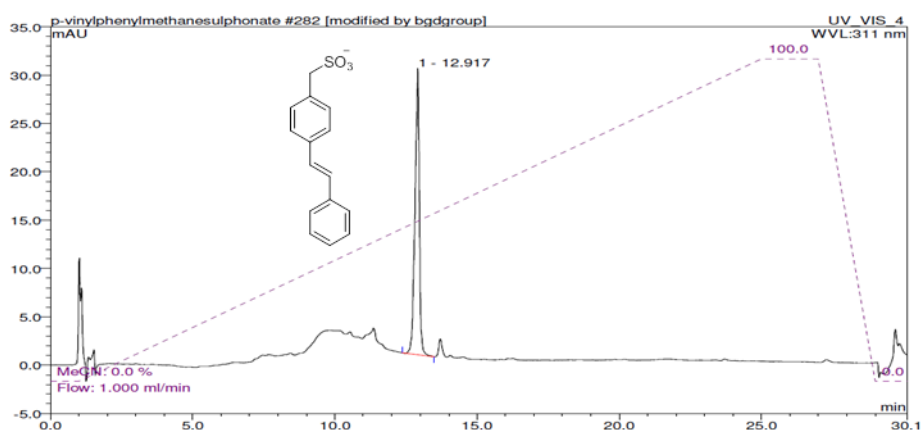
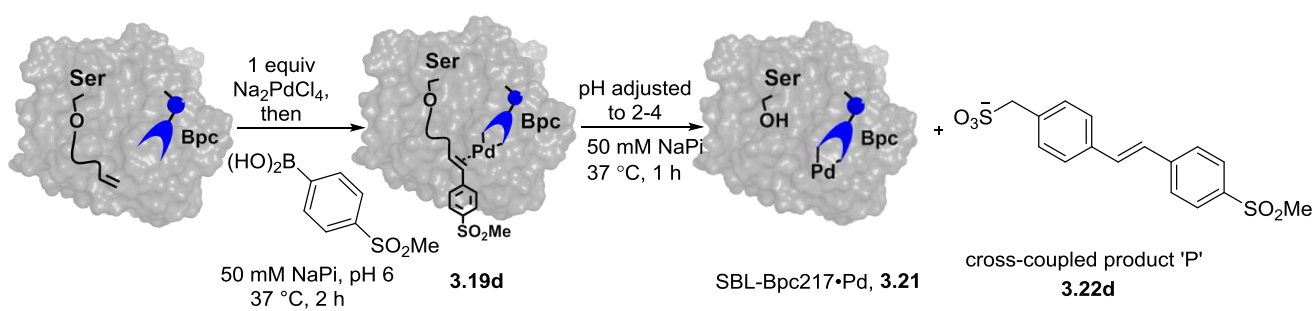
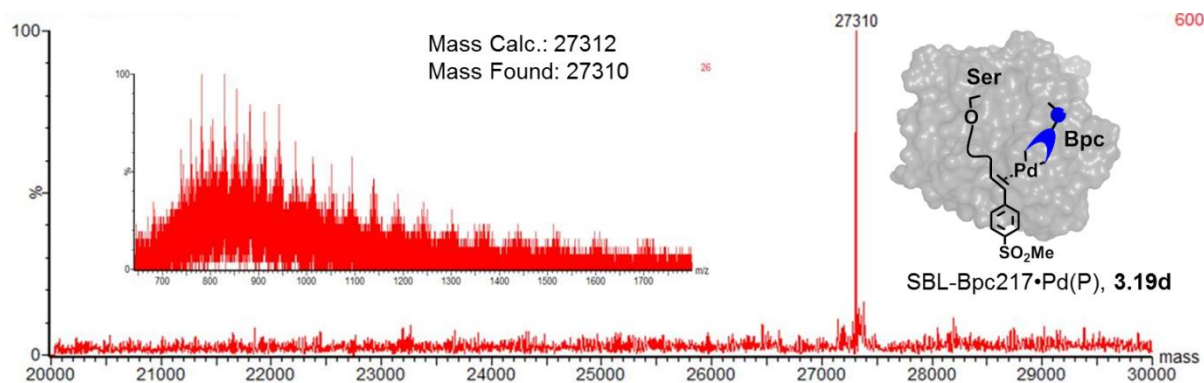


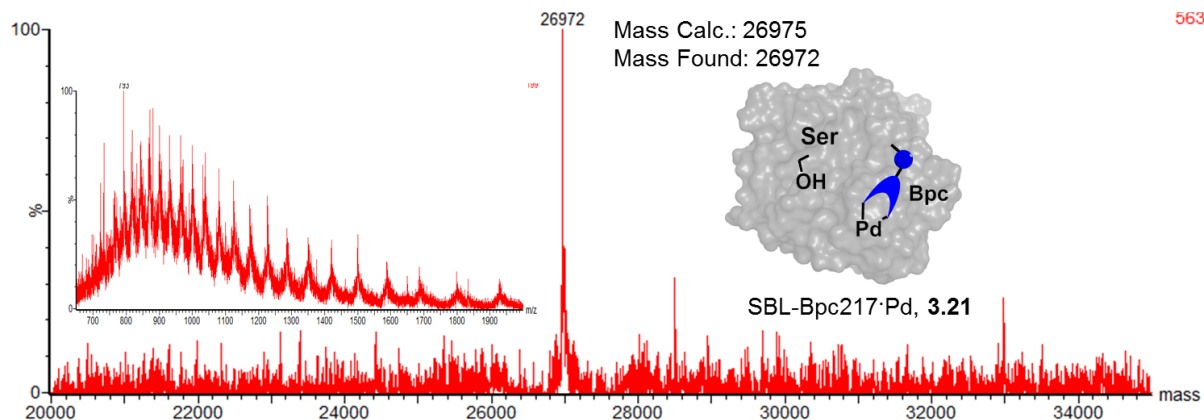
Table 3-3, Entry 4: using 4-tosylphenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of 4-tosylphenylboronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19d** was detected (mass calculated: 27312, mass found: 27310). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19d** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26972) showing release of the final coupled product **3.22d**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22d** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below. t_R = 8.87.

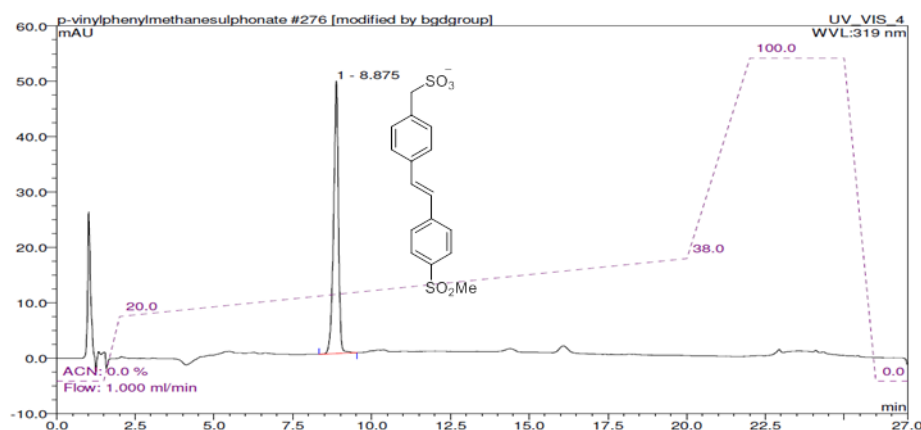
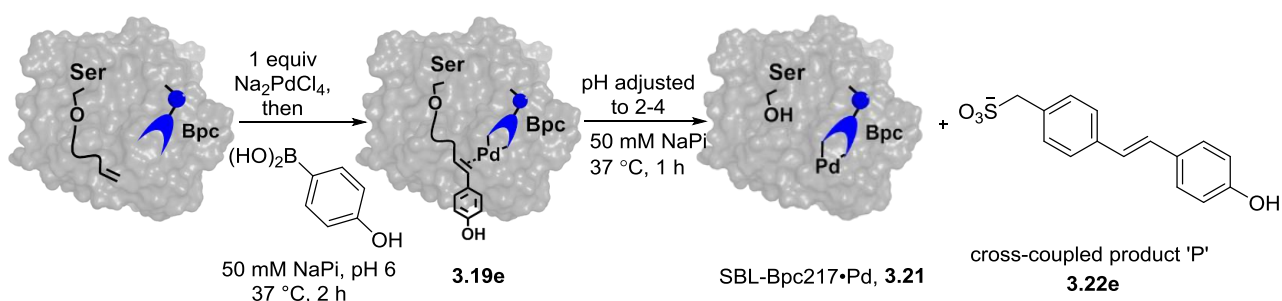
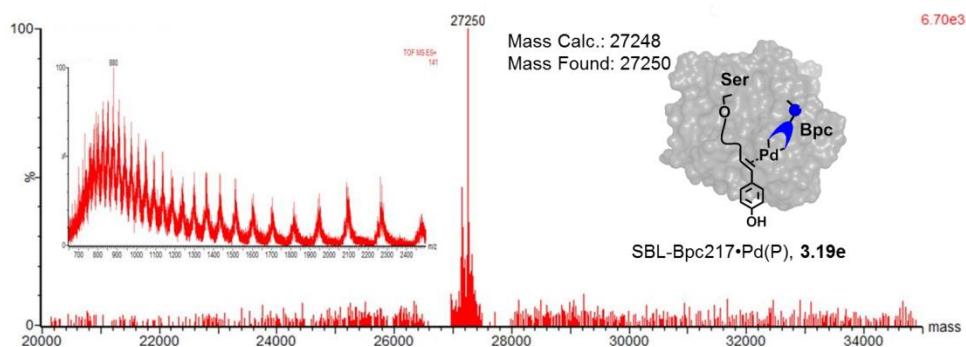


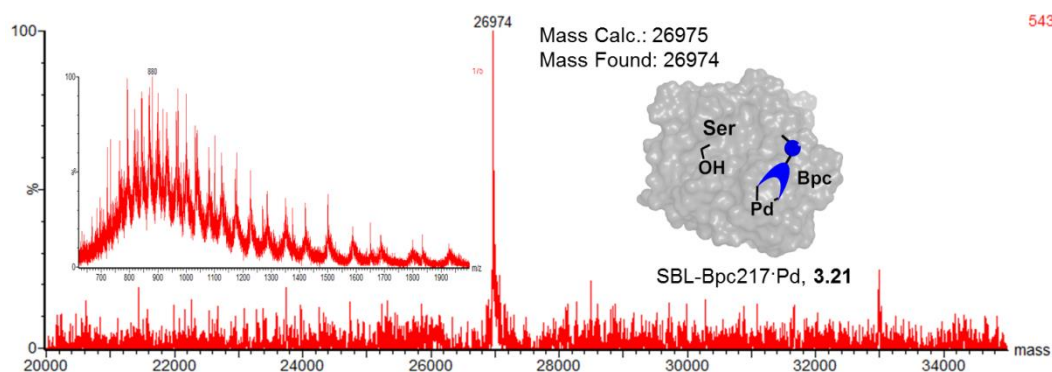
Table 3-3, Entry 5: using 4-hydroxyphenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of 4-hydroxyphenylboronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19e** was detected (mass calculated: 27250, mass found: 27250). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19e** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26974) showing release of the final coupled product **3.22e**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22e** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below.

t_R = 8.4.

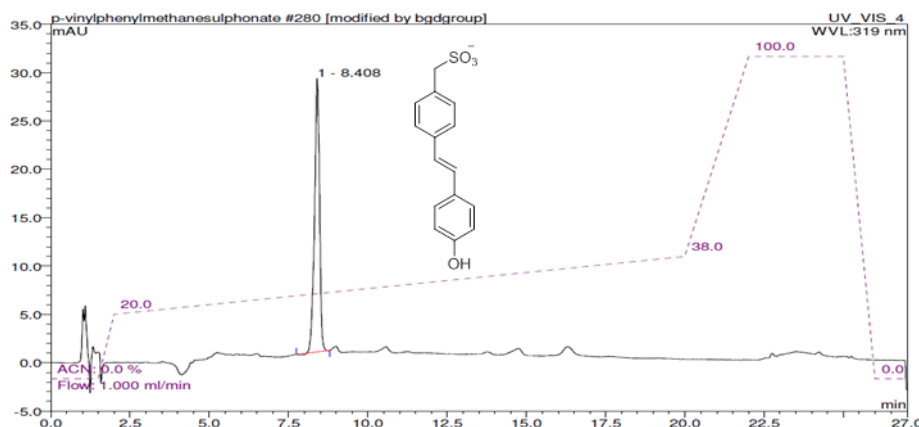
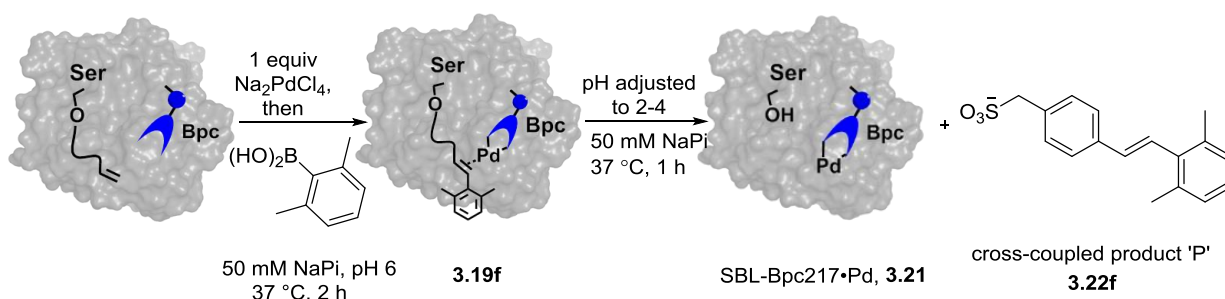
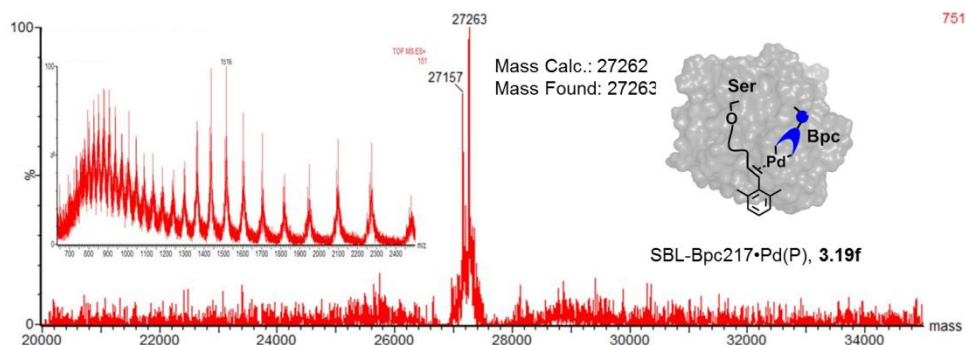


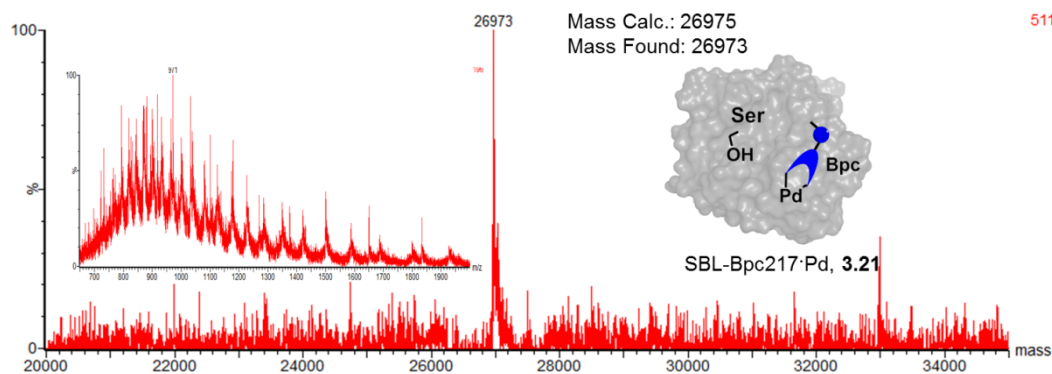
Table 3-3, Entry 6: using (2,6-dimethylphenyl)boronic acid



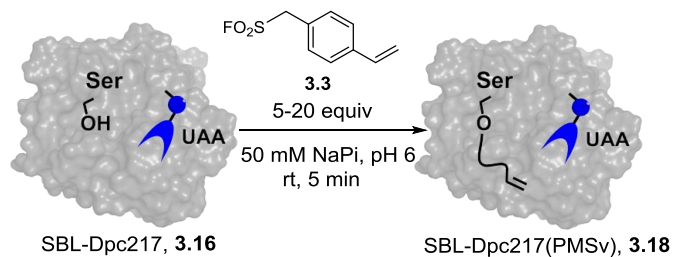
Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Bpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of (2,6-dimethylphenyl)boronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.19f** was detected (mass calculated: 27262, mass found: 27263). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Bpc217•Pd(P) **3.19f** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd **3.21** (calculated mass 26975, mass found 26970) showing release of the final coupled product **3.22f**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).

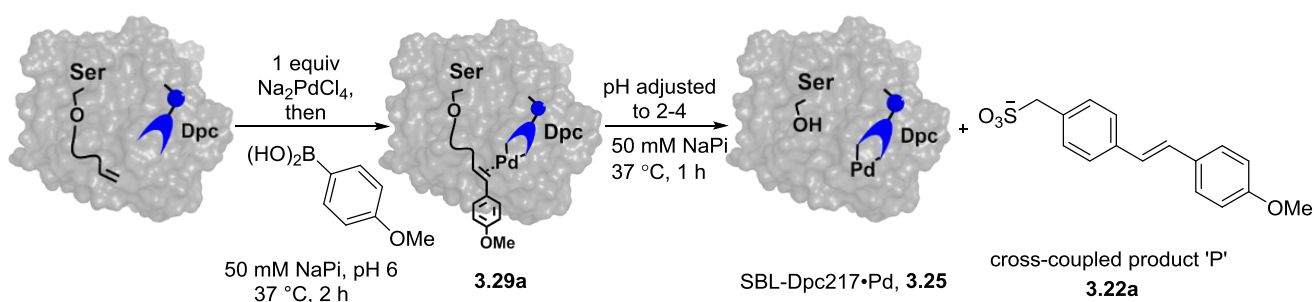


3.5.14 Oxidative Heck-type cross-coupling reaction using SBL-Dpc217



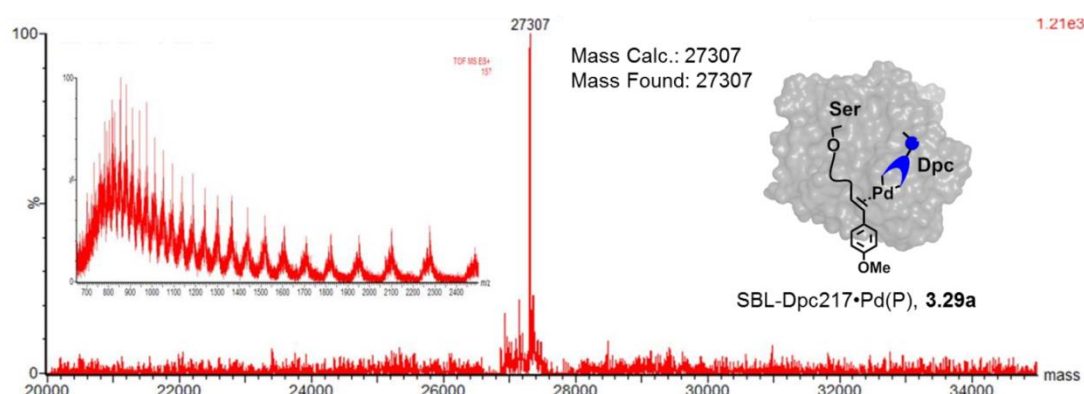
A 800 μL sample of SBL-Dpc217 **3.16** was prepared at 1.75 mg/mL in sodium phosphate buffer (50 mM, pH 6). A 0.5 M stock solution of alkene **3.3** was prepared in DMF. To 800 μL sample of SBL-Dpc217 (52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 5.2 μL of alkene **3.3** in DMF (2.6 μmol). The reaction solution was vortexed to homogenize and subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217(PMSv) **3.18** (calculated mass 27093, mass found 27096). The excess alkene was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). The enzyme was dialysed against 50 mM NaPi pH 6 for 2 h and subsequently stored on ice until needed and concentration was measured using BCA assay.

Table 3-3, Entry 7: using 4-methoxyphenylboronic acid

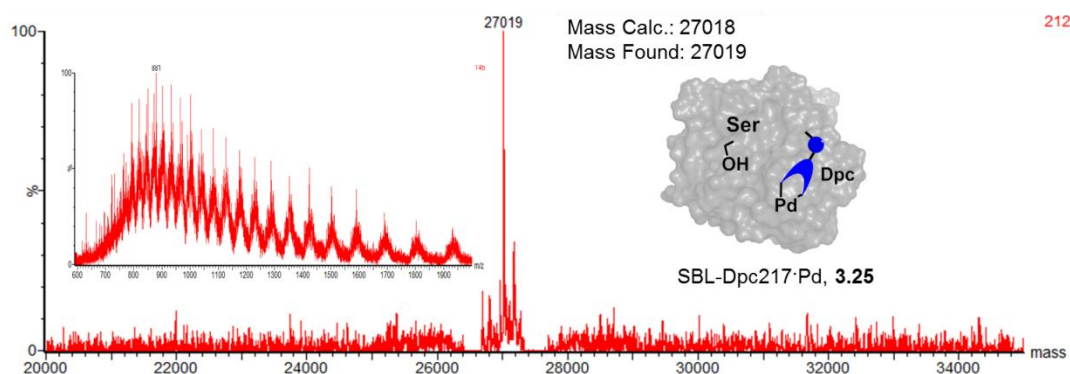


Stock solutions of reagents were prepared as before. To 200 μL sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 0.56 μL of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μL of 4-methoxyphenylboronic acid in DMF (5.5 μmol). The reaction mixture was incubated at 37 $^{\circ}\text{C}$ for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29a** was detected (mass calculated: 27307, mass found: 27312). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham

Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Dpc217•Pd(P) **3.29a** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27018, mass found 27019) showing release of the final coupled product **3.22a**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22a** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below.

$t_R = 19.4$. NB the HPLC solvent A was different from standard conditions (99.9% water + 0.1% formic acid was used).

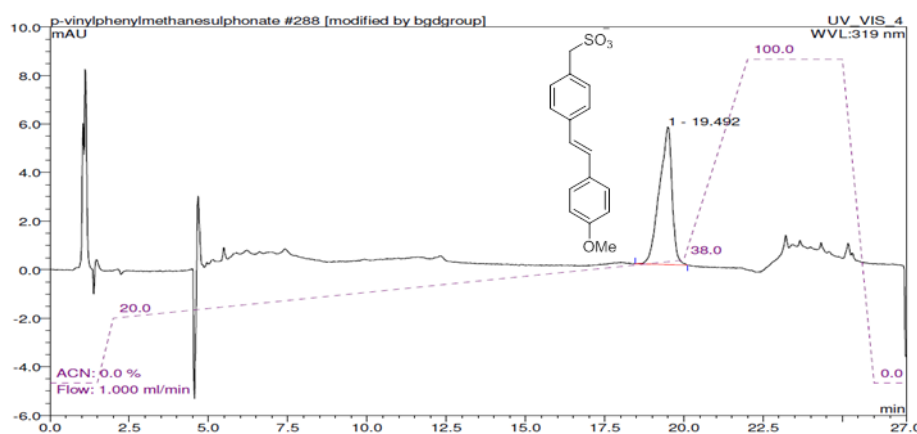
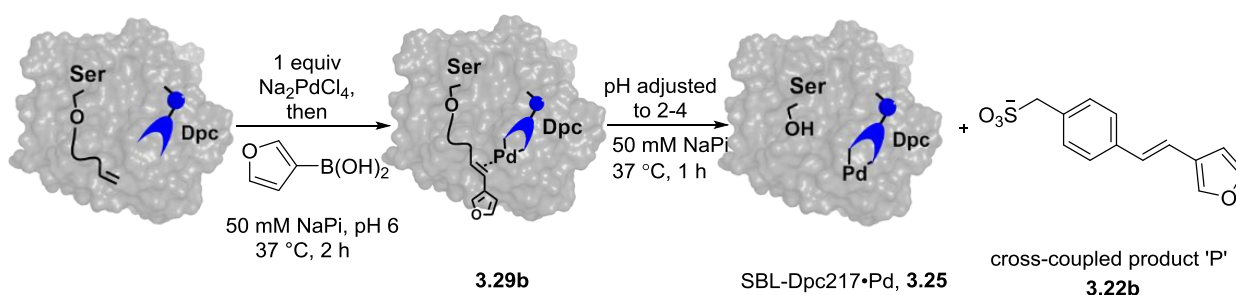
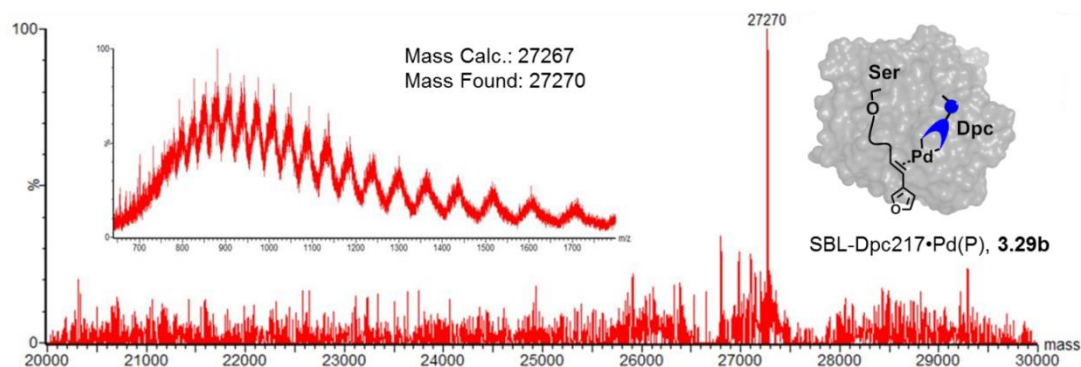


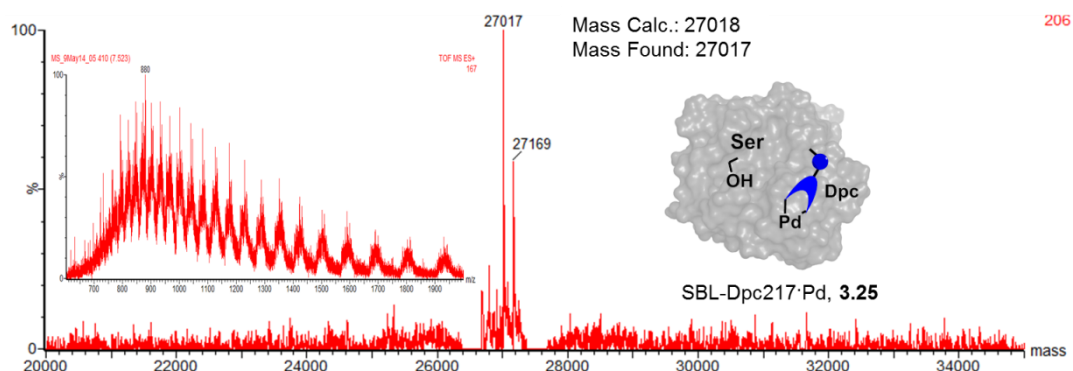
Table 3-3, Entry 8: using 3-furanboronic acid



Stock solutions of reagents were prepared as before. To 200 μL sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μL of 3-furanboronic acid in DMF (5.5 μmol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29b** was detected (mass calculated: 27268, mass found: 27270). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Dpc217•Pd(P) **3.29b** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27018, mass found 27017) showing release of the final coupled product **3.22b**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 306 nm) required a scale-up of the protein reactions. Final coupled product **3.22b** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below.
 t_R = 13.19.

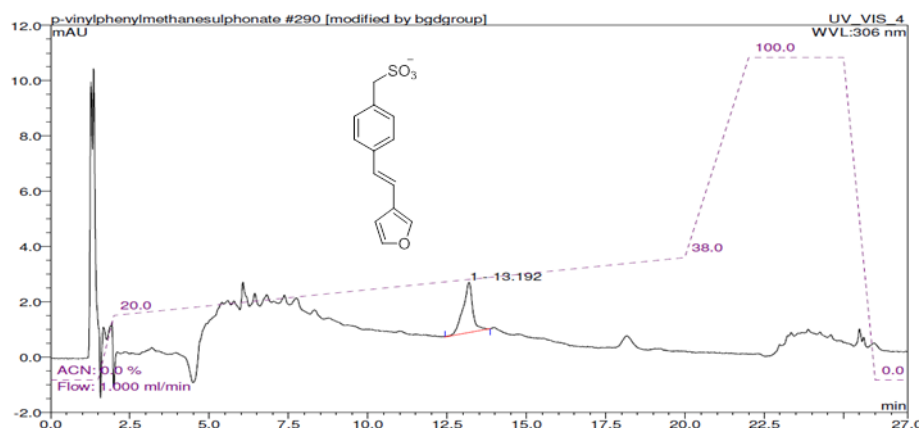
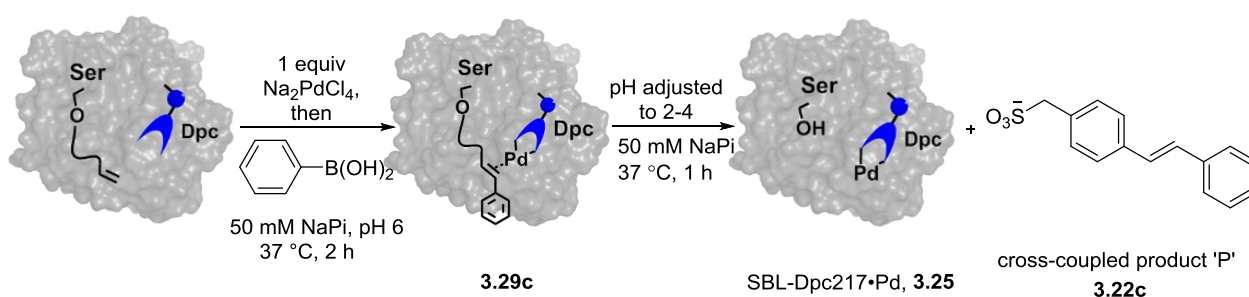
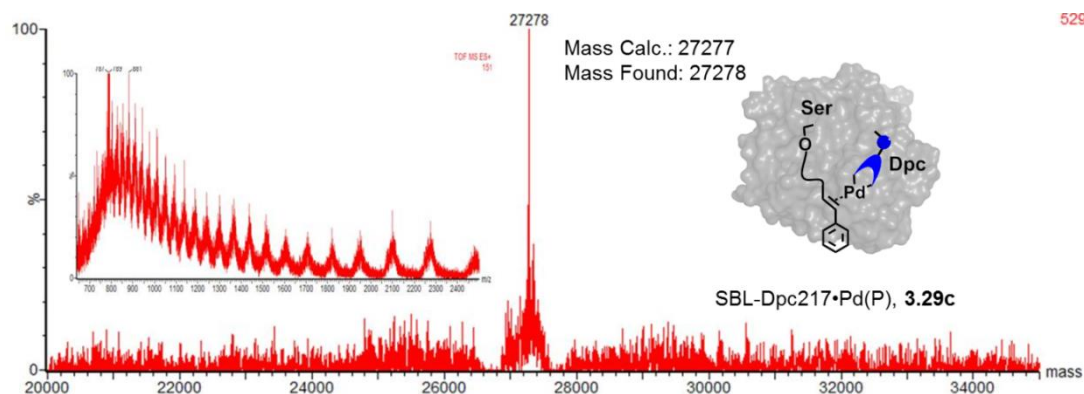


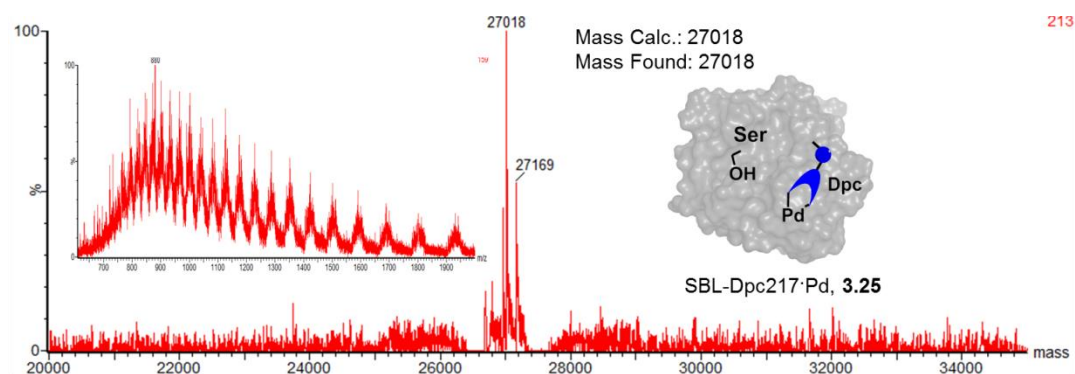
Table 3-3, Entry 9: using phenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of phenylboronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29c** was detected (mass calculated: 27277, mass found: 27278). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Dpc217•Pd(P) **3.29c** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27018, mass found 27017) showing release of the final coupled product **3.22c**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 311 nm) required a scale-up of the protein reactions. Final coupled product **3.22c** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below. t_R = 12.94.

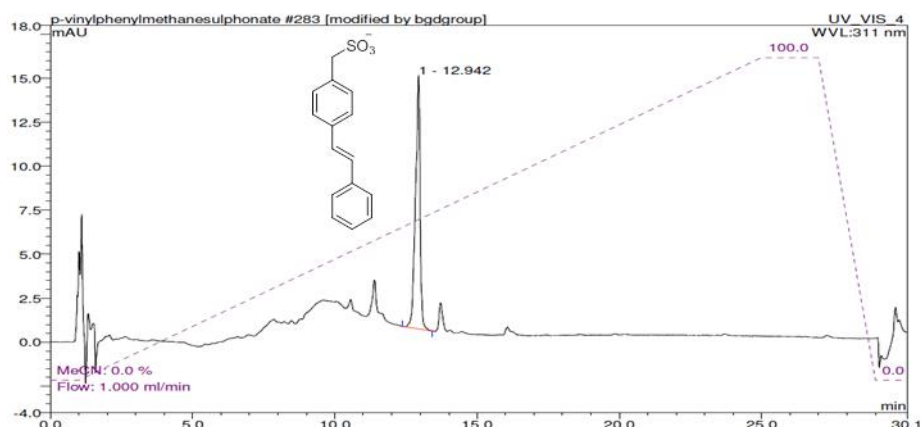
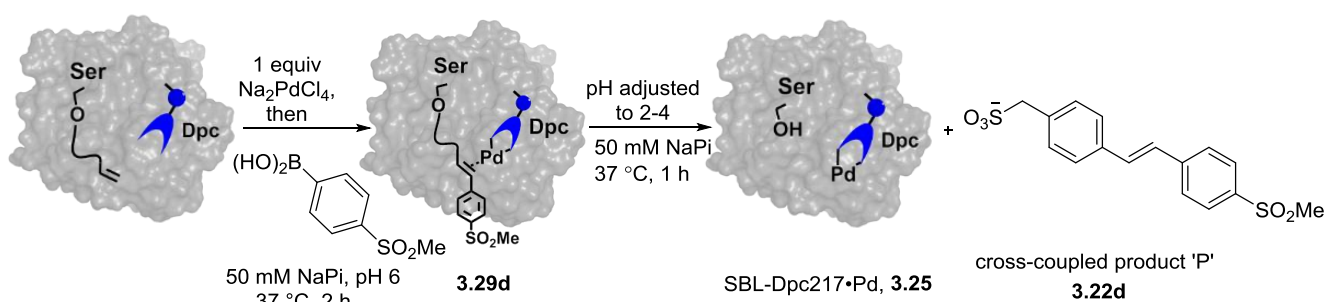
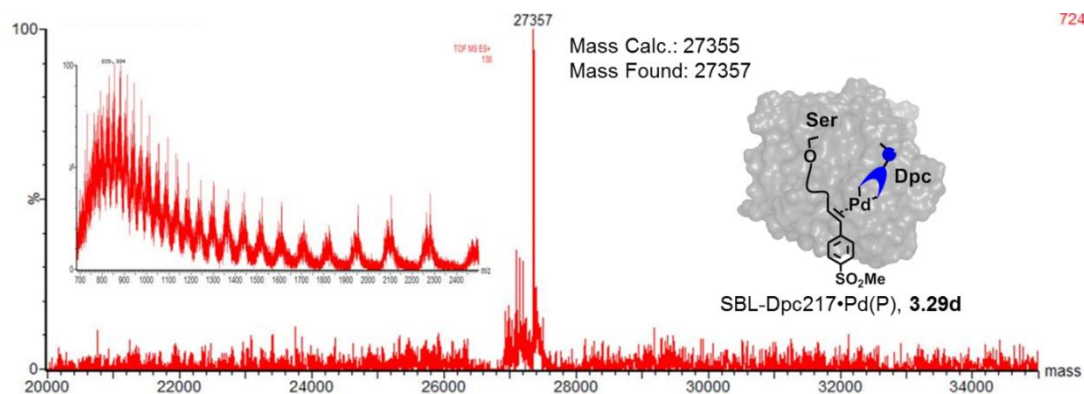


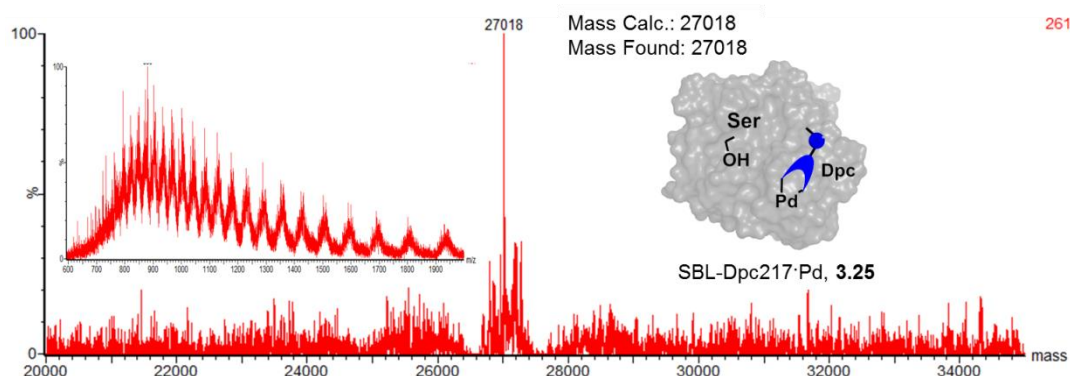
Table 3-3, Entry 10: using 4-tosylphenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μL sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μL of 4-tosylphenylboronic acid in DMF (5.5 μmol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29d** was detected (mass calculated: 27355, mass found: 27357). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Dpc217•Pd(P) **3.29d** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27018, mass found 27018) showing release of the final coupled product **3.22d**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22d** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below. t_R = 12.0. NB the HPLC solvent A was different from standard conditions (100% water + .1% FA was used).

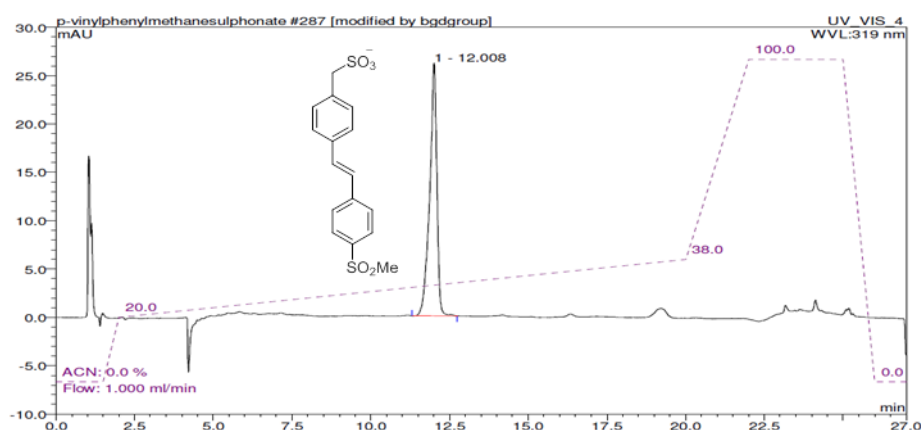
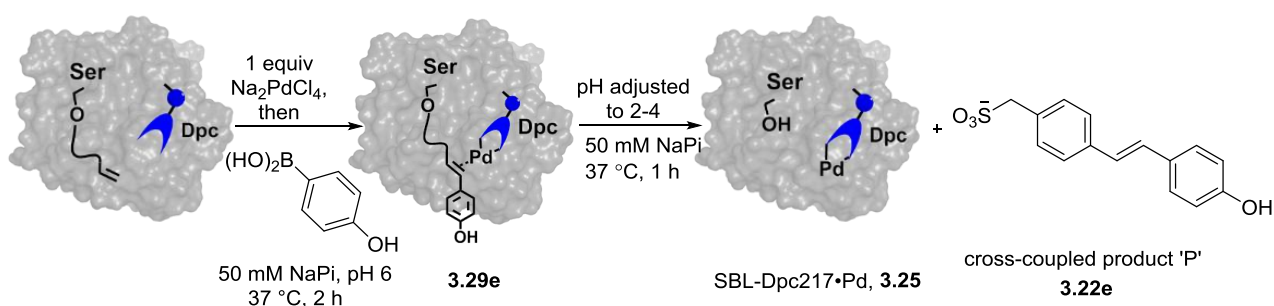
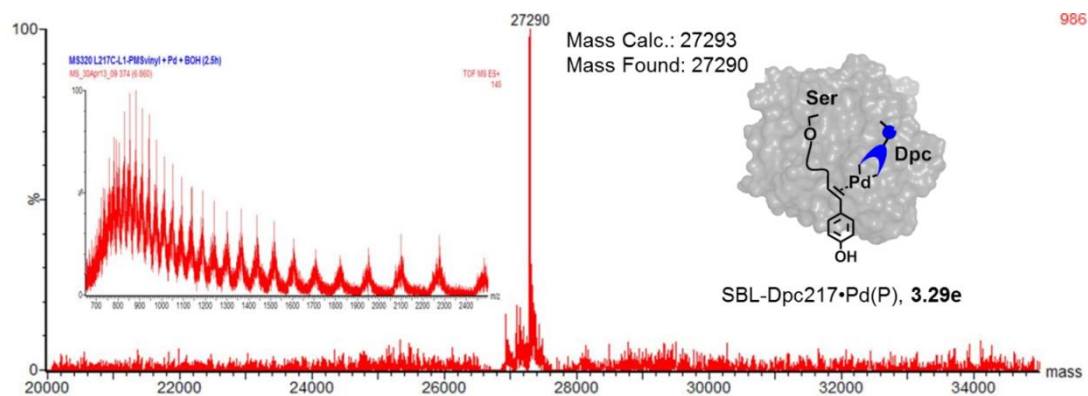


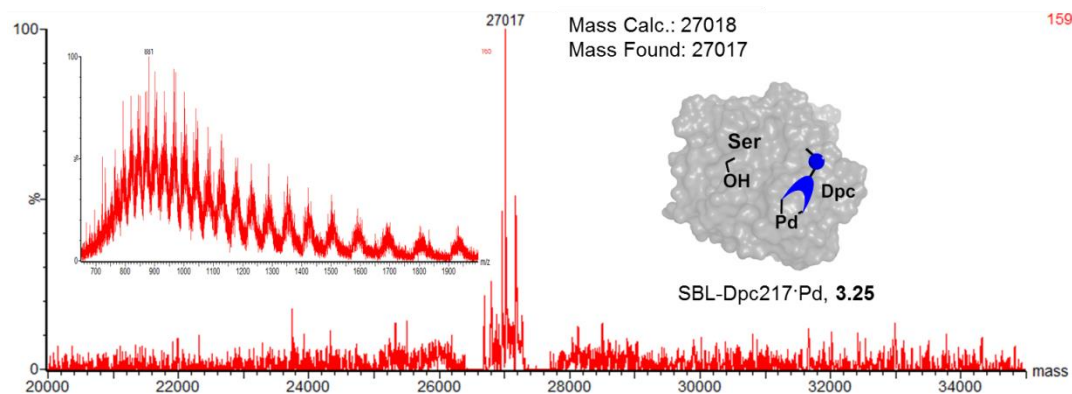
Table 3-3, Entry 11: using 4-hydroxyphenylboronic acid



Stock solutions of reagents were prepared as before. To 200 μL sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μL of 4-hydroxyphenylboronic acid in DMF (5.5 μmol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29e** was detected (mass calculated: 27293, mass found: 27290). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



To 200 μ L sample of SBL-Dpc217•Pd(P) **3.29e** (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd **3.25** (calculated mass 27018, mass found 27017) showing release of the final coupled product **3.22e**. The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for quantitative analysis of small molecule by reverse phase HPLC (rpHPLC).



Detection of the final product by UV-HPLC (UV= 319 nm) required a scale-up of the protein reactions. Final coupled product **3.22e** was isolated using the same procedure as described above (Entry 1). The standard chromatogram for the separation is given below. t_R = 8.95.

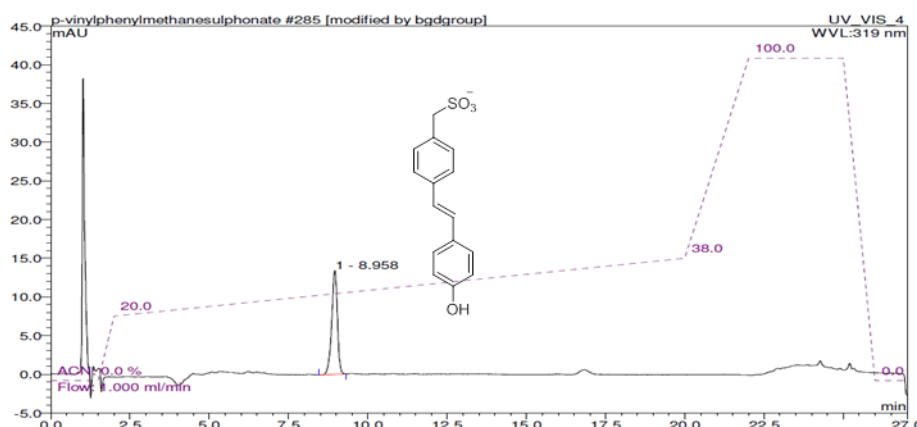
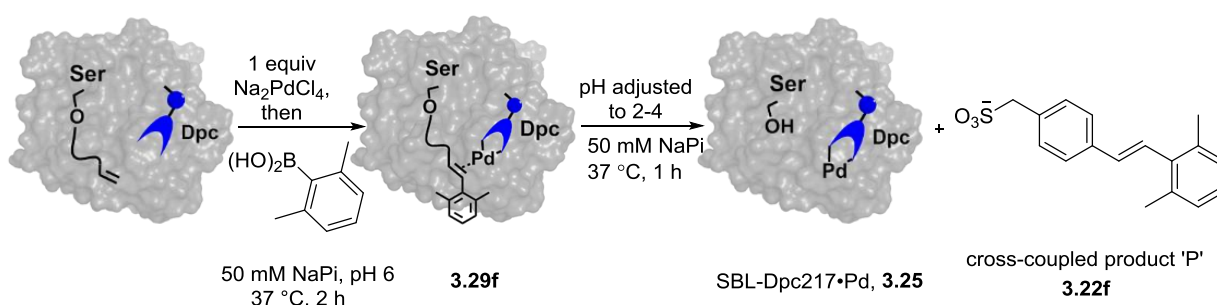
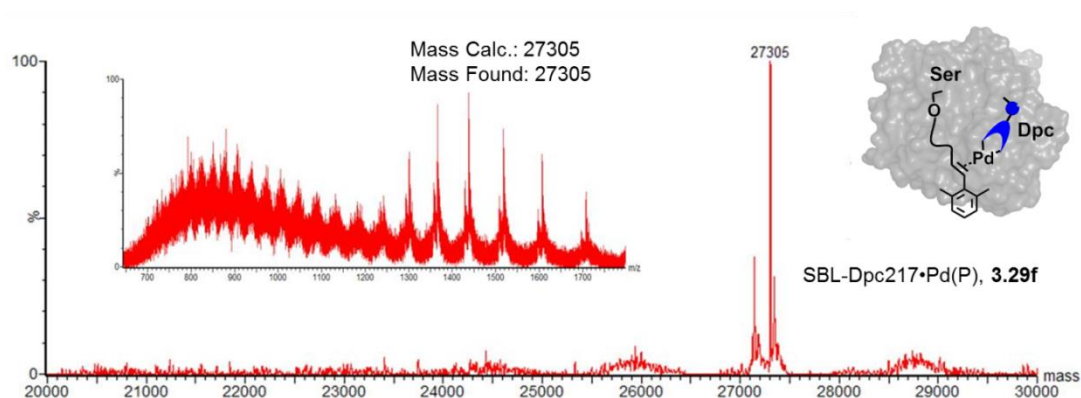


Table 3-3, Entry 12: using (2,6-dimethylphenyl)boronic acid



Stock solutions of reagents were prepared as before. To 200 μ L sample of SBL-Dpc217(PMSv) (5.52 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11.2 μ L of (2,6-dimethylphenyl)boronic acid in DMF (5.5 μ mol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was monitored by LC-MS. A single protein species **3.29f** was detected (mass calculated: 27305, mass found: 27305). The ESI-MS for the modification is shown below. The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed.



Isolation and characterization of final coupled product

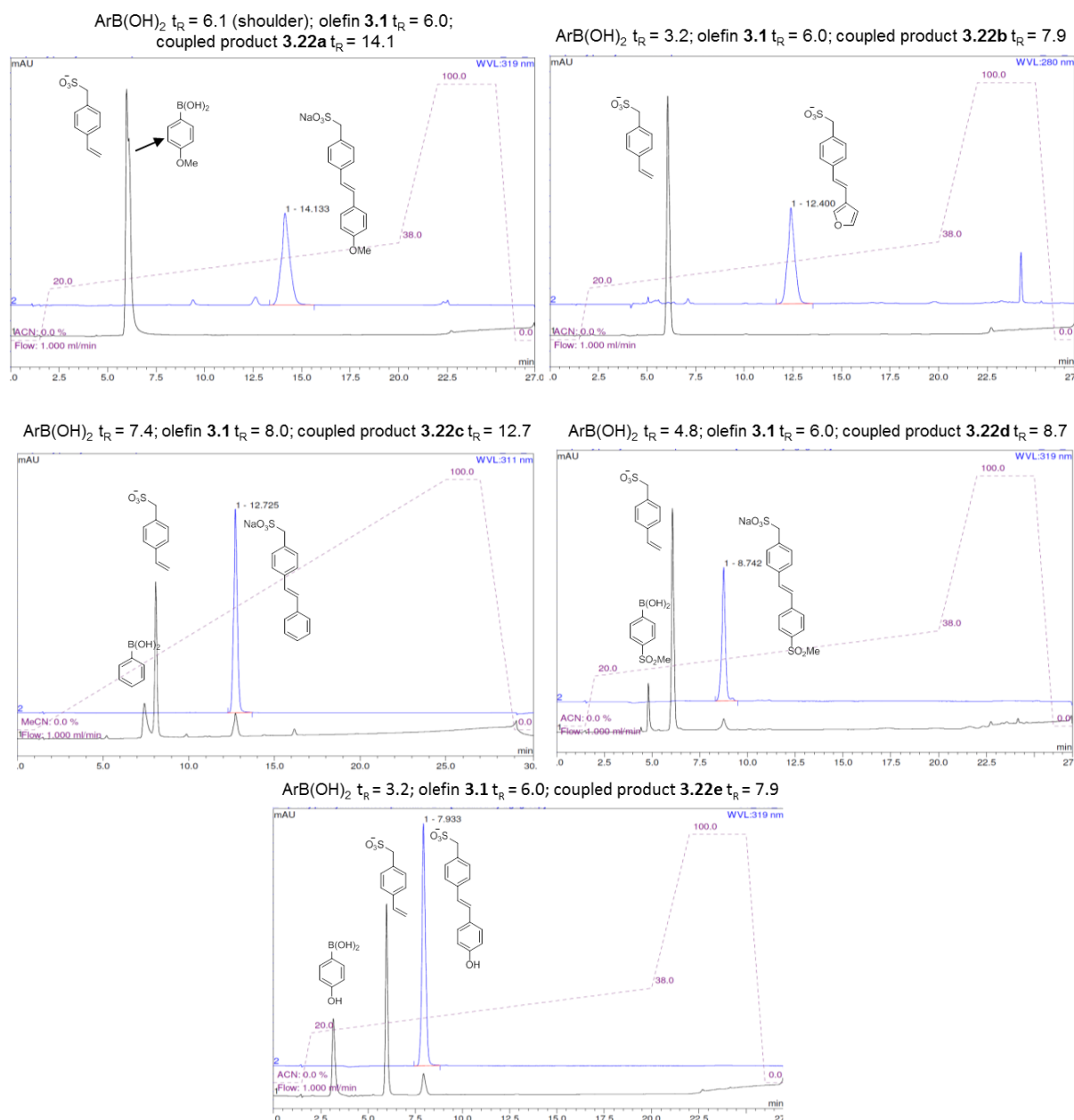
3.5.15 Standard chromatograms for rpHPLC separation

HPLC analysis was performed using a Dionex UltiMate3000 Pump and Variable Wavelength Detector, controlled by Chromeleon v6.8 software. The coupled product and starting materials were separated on a Phenomenex Synergi-Hydro-RP C18 column under reversed-phase chromatographic conditions. Solvent System: 0.1% formic acid in H₂O: CH₃CN 95:5 (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow rate of 1.0 mL min⁻¹.

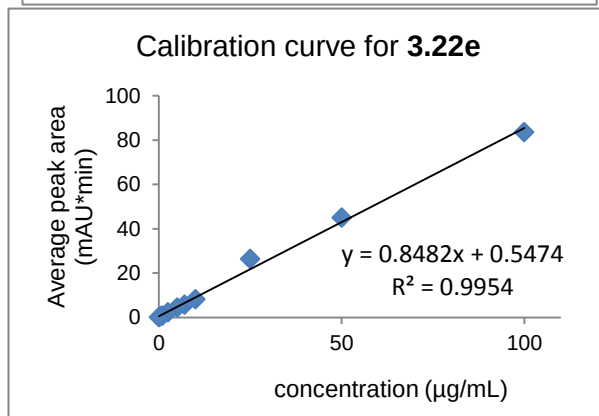
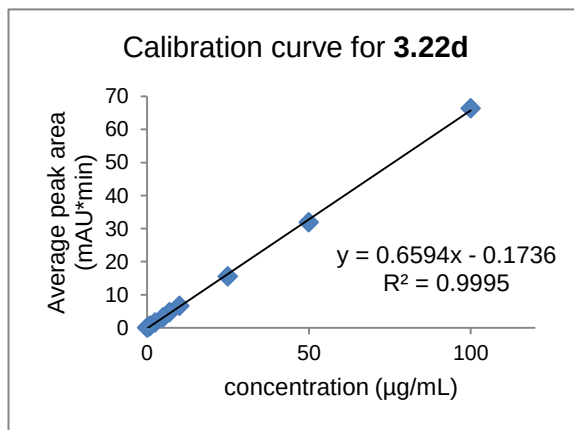
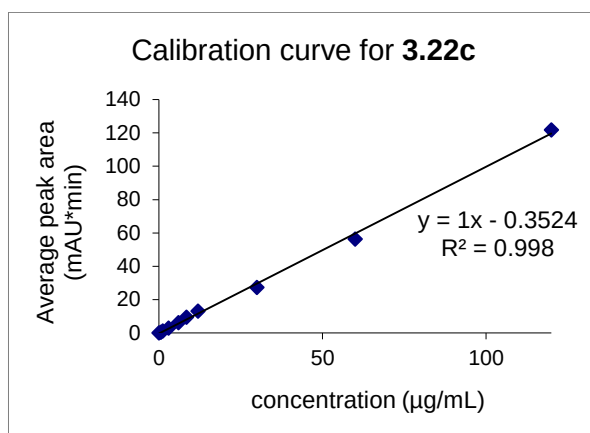
HPLC Method A. The multi-step gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 20% B over 1 min followed by a gradual linear gradient to 38% B over next 18 minutes and then 100% B over 2 minutes. The column was flushed with 100% B for an additional 2 min, followed by a linear gradient over 2 minutes back to 100% A. The column was re-equilibrated with 100% A for additional 5 minutes.

HPLC Method B. The gradient was programmed as follows: 100% A for 1 min and then a linear gradient to 100% B over 24 min followed by 100% B for an additional 2 min. A linear gradient over 2 minutes back to 100% A was used. The column was re-equilibrated with 100% A for additional 5 minutes.

A 100 $\mu\text{g/mL}$ solution of the arylboronic acid, compound **3.1** and corresponding final coupled product **3.22a-e** each was prepared in Milli-Q® water. The three components were mixed and diluted to a final volume of 50 μL . 20 μL of the mixture was loaded on a Phenomenex Synergi 4 μm Hydro-RP C18 (100 $\times$ 4.6 mm) column. The three components of the mixture were separated under reverse-phase chromatographic conditions using HPLC method A (except for **3.22c** for which HPLC method B was used) using the appropriate wavelength for detection ($\lambda = 319 \text{ nm}$ for detection of compound **3.22a-e** and $\lambda = 254 \text{ nm}$ for ArB(OH)_2 and olefin **3.1**) and retention time was recorded.

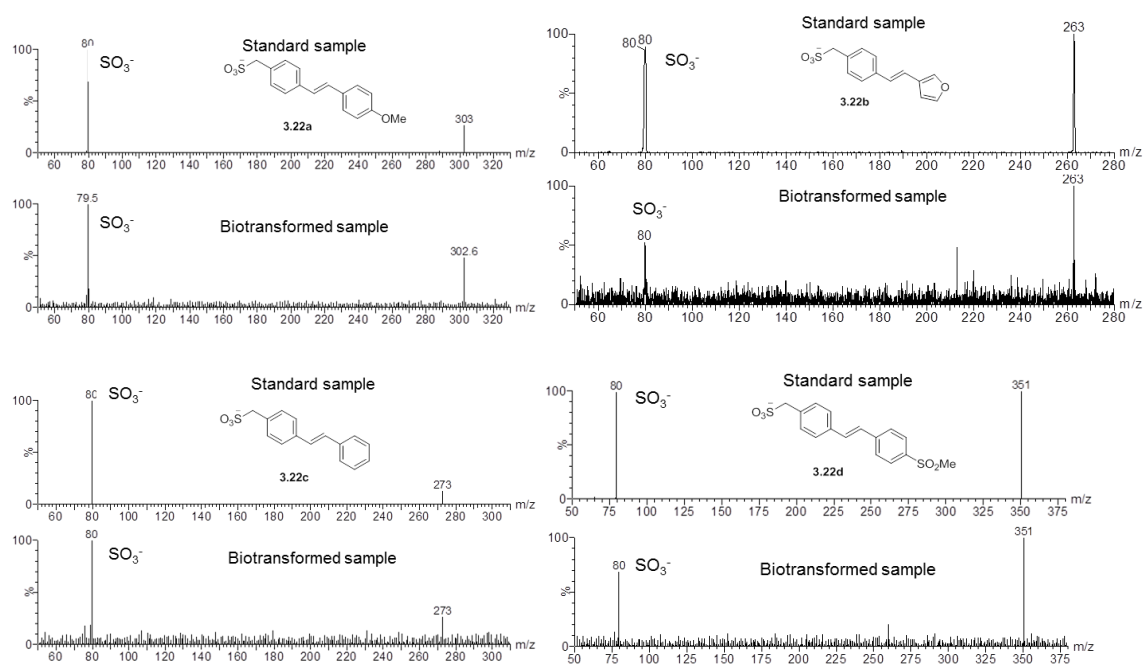


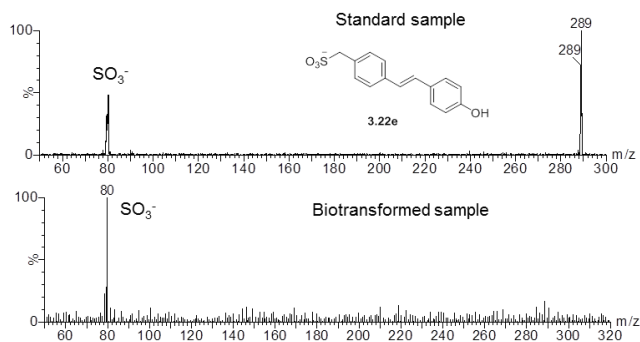
Calibration curves for HPLC standards. Typically, a 1 mg/mL stock solution of final product standard **3.22x** was prepared. Standard solutions for the calibration curves were prepared by serial dilutions to 120, 60, 30, 12, 8.4, 6, 3, 1.2, 0.6 µg/mL. 20 µL was loaded on a Phenomenex Synergi-Hydro-RP C18 column under the same rpHPLC conditions described above. Calibration curves were plotted using average area under the peak from data collected in two successive runs. Standard calibration curves for final products **3.22c**, **3.22d** and **3.22e** are shown below.



3.5.16 Tandem MS analysis of Final Coupled product 3.22a-e

HPLC-ESI-MS/MS analysis was carried out using a Waters Binary HPLC pump interfaced to a Micromass[®] Quattro[™] micro atmospheric pressure ionization (API) mass spectrometer. Chromatography was performed on a Phenomenex Synergi 4 μ m Hydro-RP C18 (100 $\times$ 4.6 mm) column under reversed-phase conditions using the multi-step gradient HPLC Method A, as described above. Solvent System: 0.1% formic acid in H₂O:CH₃CN 95:5 (solvent A) and 0.1% formic acid in CH₃CN (solvent B) at a flow rate of 1.0 mL min⁻¹. The mass spectrometer was set up in a negative-ion electrospray mode with the following parameters: capillary voltage 2.3 kV, source temperature 120°C, cone gas 40 L/hr, desolvation gas flow 600 L/hr, desolvation gas temperature 280 °C, cone voltage 35 V, collision energy 20 V and gas cell pressure at $\sim 4.5 \times 10^{-3}$ mbar. The synthesized final coupled products were used to obtain the standard fragmentation pattern. The biotransformed final product samples were subsequently analysed by rpHPLC-MS/MS to generate the fragmentation spectra for the coupled products obtained from the enzyme reactions.

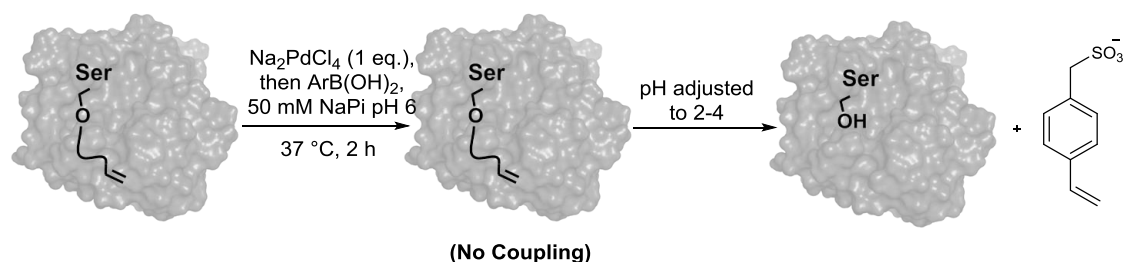




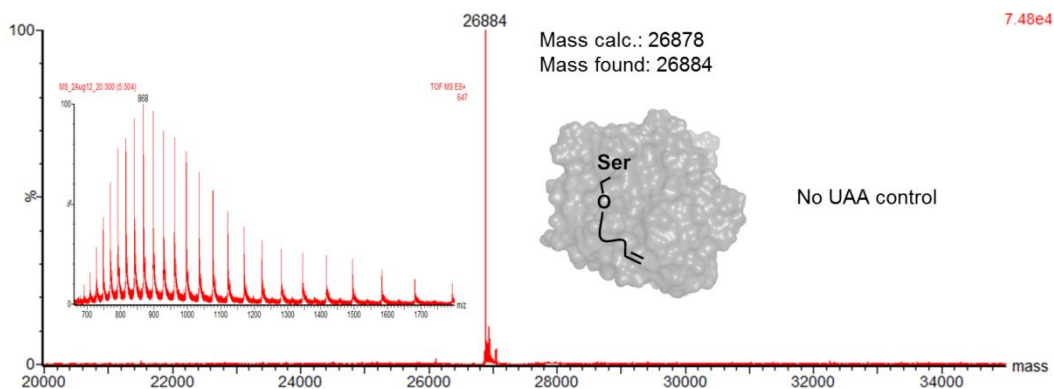
Control Reactions

3.5.17 Cross-coupling control reactions

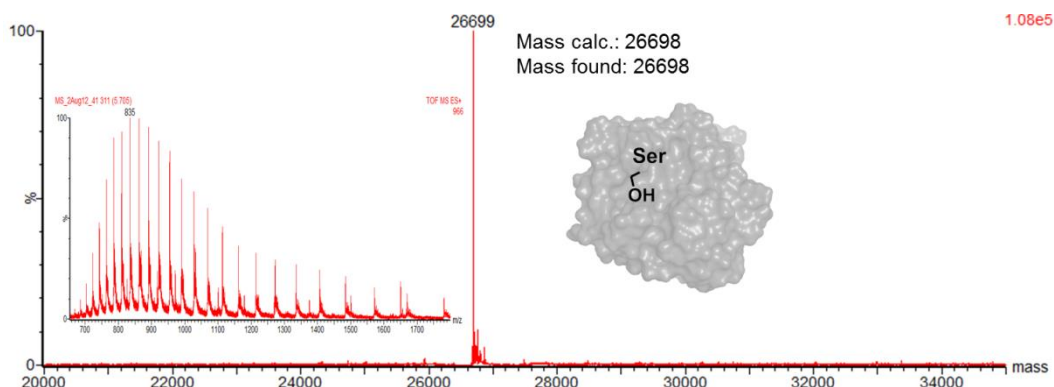
Controls using SBL-WT (No UAA control): Boronic acid and Palladium only



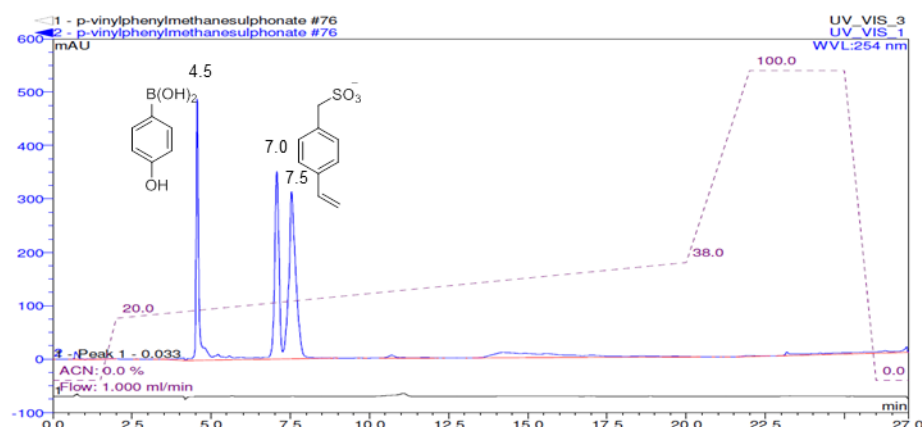
A 350 μ L sample of SBL-WT(PMSv) was prepared at 0.78 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 350 μ L sample of SBL-WT(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 1.1 μ L of Na₂PdCl₄ (11 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 20.2 μ L of 4-hydroxyphenylboronic acid in DMF (10.1 μ mol). The reaction mixture was incubated at 37 °C for 2 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) (calculated mass 26878, mass found 26884). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control is shown below.



To 340 μ L sample of SBL-WT(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL WT (calculated mass 26698, mass found 26699). The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for analysis of small molecule was by rpHPLC.

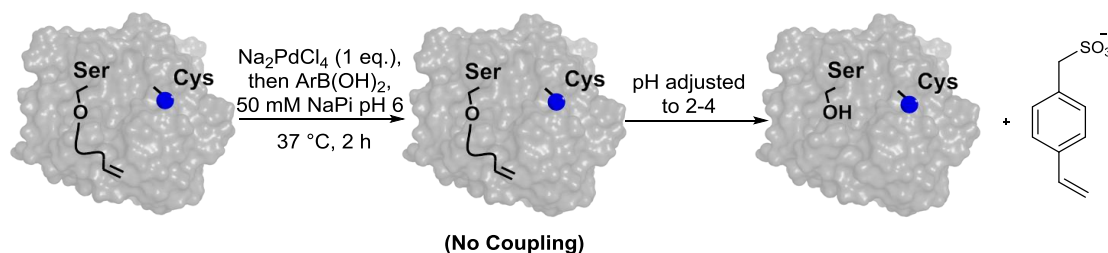


Hydrolysis product was isolated and analysed by rpHPLC. The freeze dried sample was taken and re-dissolved in 20 μ L MilliQ water. The final product sample was vortexed, sonicated gently and centrifuged. The sample was separated on a Phenomenex Synergi-Hydro-RP C18 column using HPLC method A. The final product was analysed by measuring UV absorbance at $\lambda=319$ nm. The standard chromatogram for the separation is given below. $t_R=7.5$; HRMS m/z (ESI): Found 197.03 $[M-H]^-$.



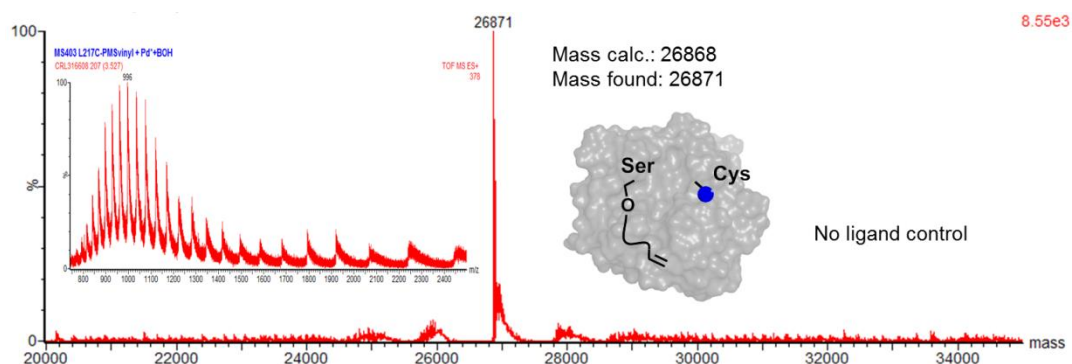
Controls using SBL-L217C, No ligand Lx

Boronic acid and Palladium only:

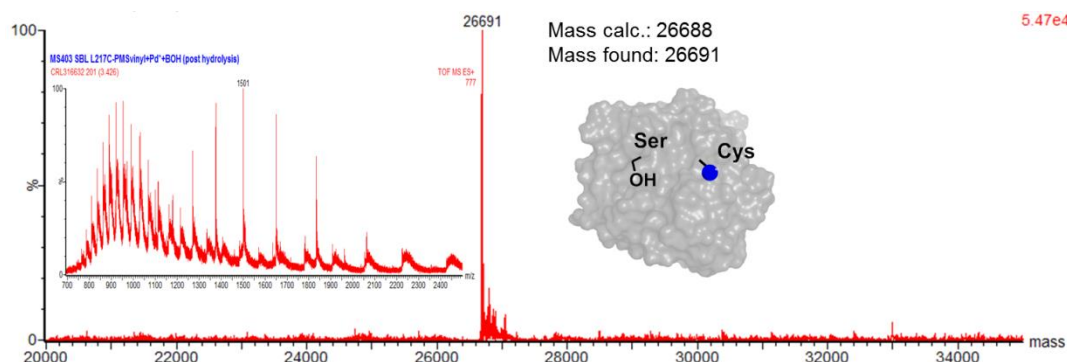


A 200 μ L sample of SBL-L217C(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 5 mM stock of Na_2PdCl_4 was made in MilliQ® and 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μ L sample of SBL-L217C(PMSv) (5.6 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^\circ\text{C}$, was added 1.1 μ L of Na_2PdCl_4 (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11 μ L of 4-hydroxyphenylboronic acid in DMF (5.6 μ mol). The reaction mixture was incubated at 37 $^\circ\text{C}$ for 2 h, after which 10 μ L aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-L217C(PMSv) (calculated mass 26868, mass found 26871). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate

buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control reaction is shown below.



To the remaining sample of SBL-L217C(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-L217C (calculated mass 26688, mass found 26690). The ESI-MS is shown below.

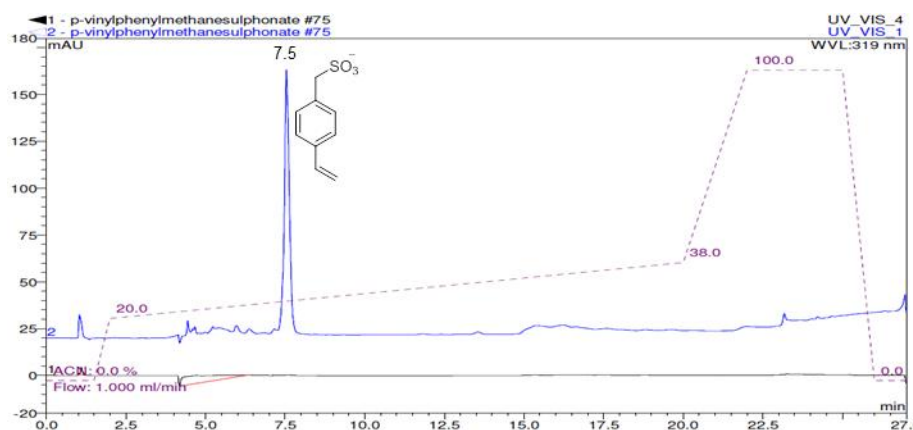


For rpHPLC analysis:

A 350 μ L sample of SBL-L217C(PMSv) was prepared in sodium phosphate buffer (pH 6, 50 mM). A 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 350 μ L sample of SBL-L217C(PMSv) (10.1 nmol) at 0.78 mg/mL in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 1.1 μ L of Na_2PdCl_4 (11.0 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 20.2 μ L of 4-hydroxyphenylboronic acid in DMF (10.1 μ mol). The reaction mixture was incubated at 37 °C for 2 h, after which 10 μ L aliquot of the reaction was analysed by LC-MS. A single

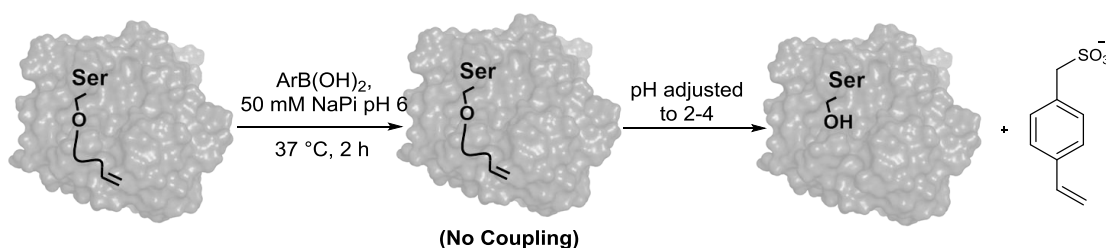
protein species was detected with a mass corresponding to the SBL-L217C(PMSv) (calculated mass 26868, mass found 26874). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6). To the remaining sample of SBL-L217C(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. Hydrolysis product was isolated and analysed by rpHPLC. The freeze dried sample was taken and re-dissolved in 20 μ L MilliQ water. The final product sample was vortexed, sonicated gently and centrifuged. The sample was separated on a Phenomenex Synergi-Hydro-RP C18 column using HPLC method A (Solvent A = 0.1% formic acid in 99.9% water). The final product was analysed by measuring UV absorbance at $\lambda=319$ nm. The standard chromatogram for the separation is given below.

$t_R=7.5$; LRMS m/z (ESI): Found 197.02 [M-H]⁻

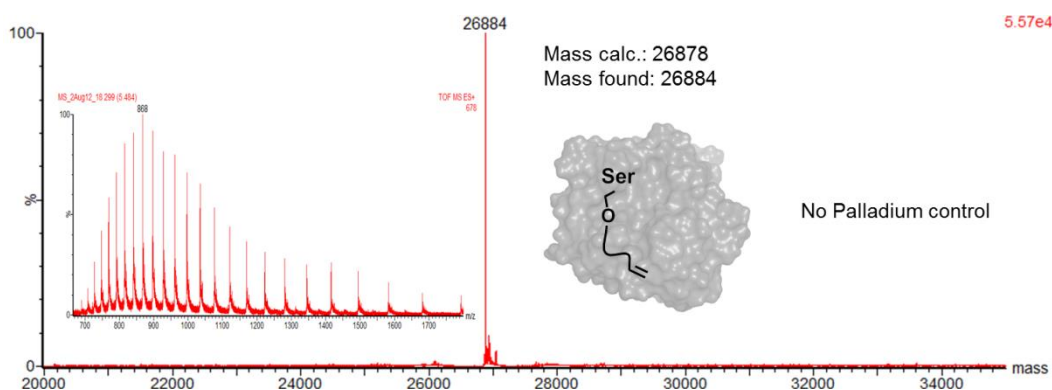


No Palladium controls:

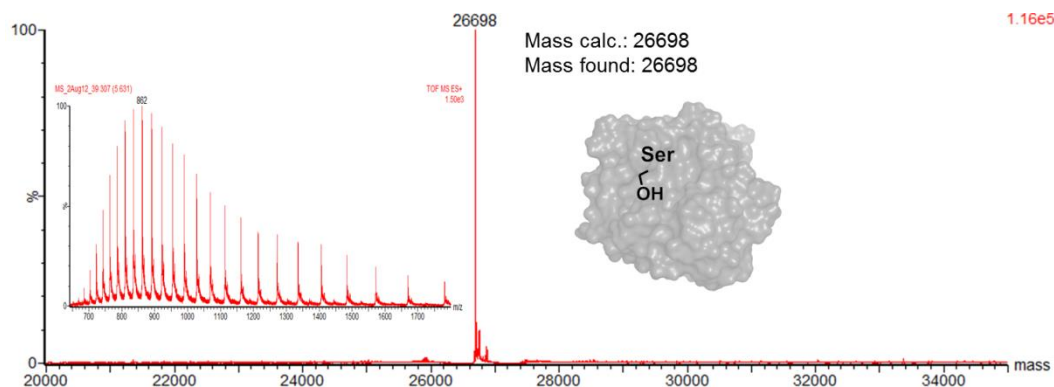
Using SBL-WT:



A 350 μ L sample of SBL-WT(PMSv) was prepared at 0.78 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 0.5 M stock solution of 3-furanboronic acid was prepared in DMF. To 350 μ L sample of SBL-WT(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH6 at 4 $^{\circ}$ C, 20.2 μ L of 3-furanboronic acid in DMF (10.1 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h, after which 10 μ L aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) (calculated mass 26878, mass found 27884). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control is shown below.

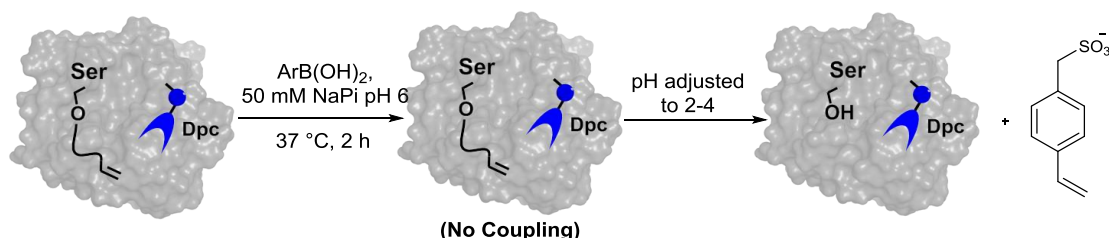


To the remaining sample of SBL-WT(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT (calculated mass 26698, mass found 26698). The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for analysis of small molecule was by rpHPLC.



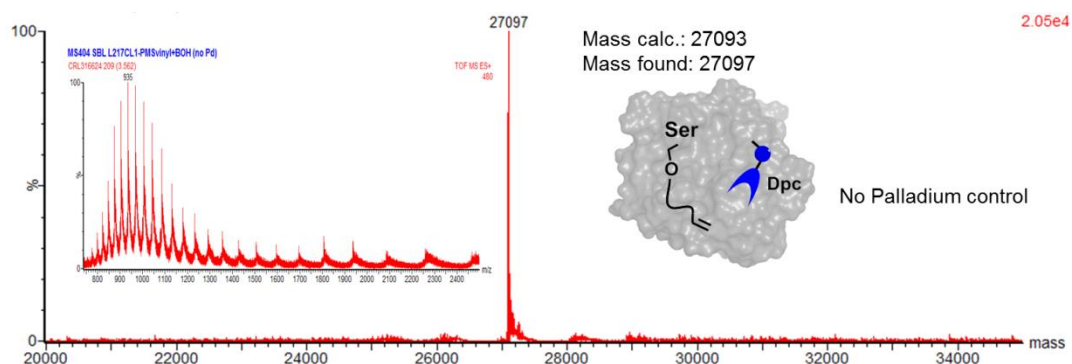
Hydrolysis product was isolated and analysed by rpHPLC. The freeze dried sample was taken and re-dissolved in 20 μ L MilliQ water. The final product sample was vortexed, sonicated gently and centrifuged. The sample was separated on a Phenomenex Synergi-Hydro-RP C18 column using HPLC method A (Solvent A = 0.1% formic acid in 99.9% water). The final product was analysed by measuring UV absorbance at $\lambda=319$ nm. The chromatogram for the separation is given below, $t_R=7.5$.

Using SBL-Dpc217:

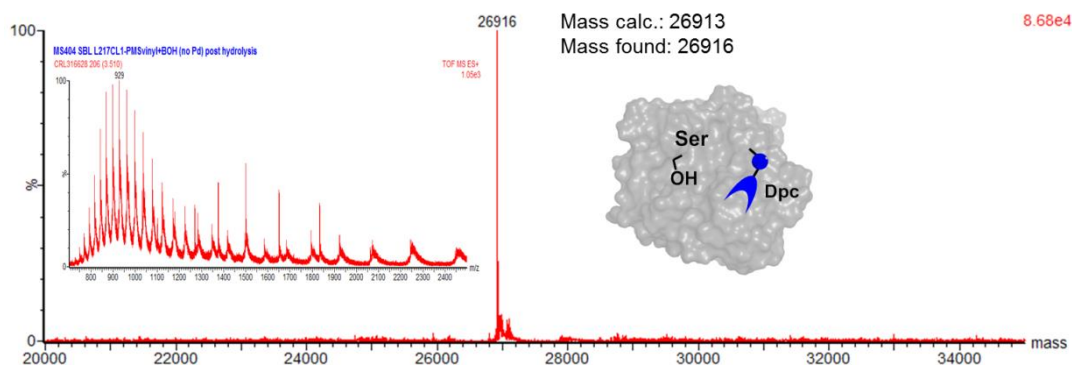


A 200 μ L sample of SBL-Dpc217(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μ L sample of SBL-Dpc217(PMSv) (5.6 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}$ C, was added 11 μ L of 4-hydroxyphenylboronic acid in DMF (5.6 μ mol). The reaction mixture was incubated at 37 $^{\circ}$ C for 2 h, after which 10 μ L aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217(PMSv) (calculated mass 27093, mass found 27097). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium

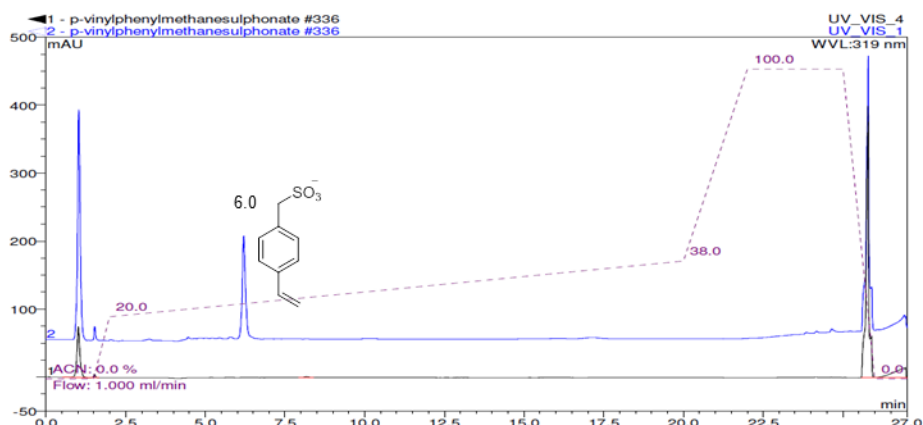
phosphate buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control reaction is shown below.



To the remaining sample of SBL-Dpc217(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217 (calculated mass 26913, mass found 26916). The ESI-MS is shown below. The remaining sample was freeze dried and stored until needed for analysis of small molecule was by rpHPLC.

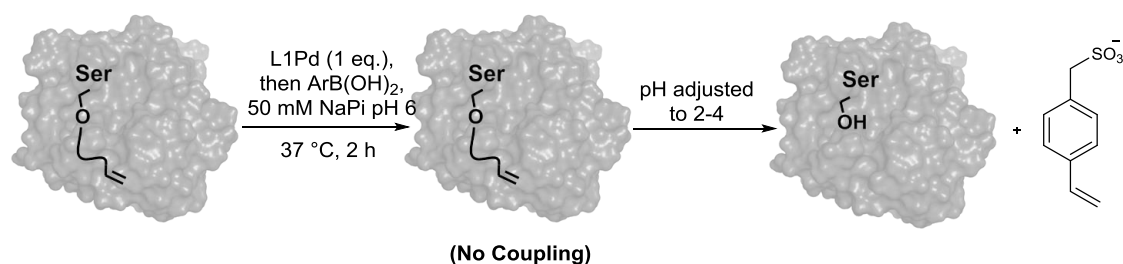


Hydrolysis product was isolated and analysed by rpHPLC. The freeze dried sample was taken and re-dissolved in 20 μ L MilliQ water. The final product sample was vortexed, sonicated gently and centrifuged. The sample was separated on a Phenomenex Synergi-Hydro-RP C18 column using HPLC method A. The final product was analysed by measuring UV absorbance at $\lambda=319$ nm. The chromatogram for the separation is given below.



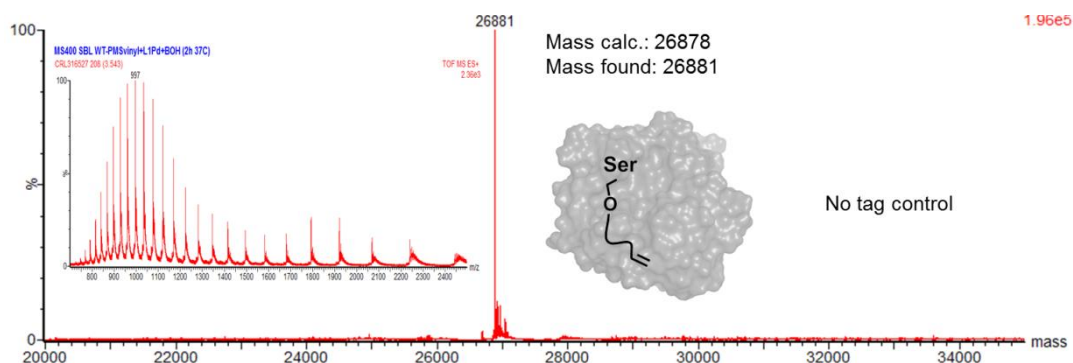
Controls using SBL-WT (No tag control)

Boronic acid and L1Pd complex only:

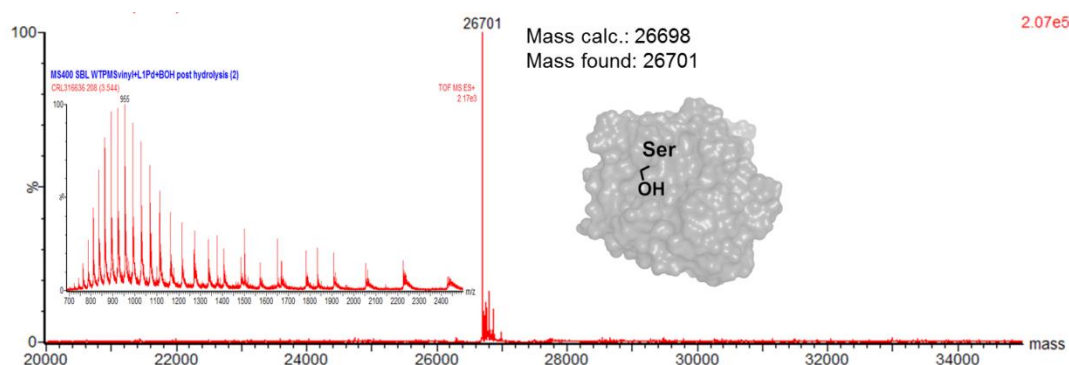


A 200 μL sample of SBL-WT(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 10 mM stock of L1Pd complex was made in MilliQ® and 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μL sample of SBL-L217C(PMSv) (5.6 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of L1-Pd solution (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11 μL of 4-hydroxyphenylboronic acid in DMF (5.6 μmol). The reaction mixture was incubated at 37 °C for 2 h, after which 10 μL aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT(PMSv) (calculated mass 26878, mass found 26881). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer

(50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control reaction is shown below.

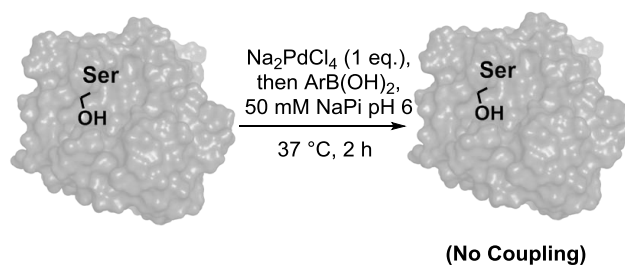


To the remaining sample of SBL-WT(PMSv) (10.1 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added HCl (1 M) to adjust pH to 2-4. The reaction solution was vortexed to homogenize and incubated at 37 °C for 1 h. The reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT (calculated mass 26698, mass found 26701). The ESI-MS is shown below.

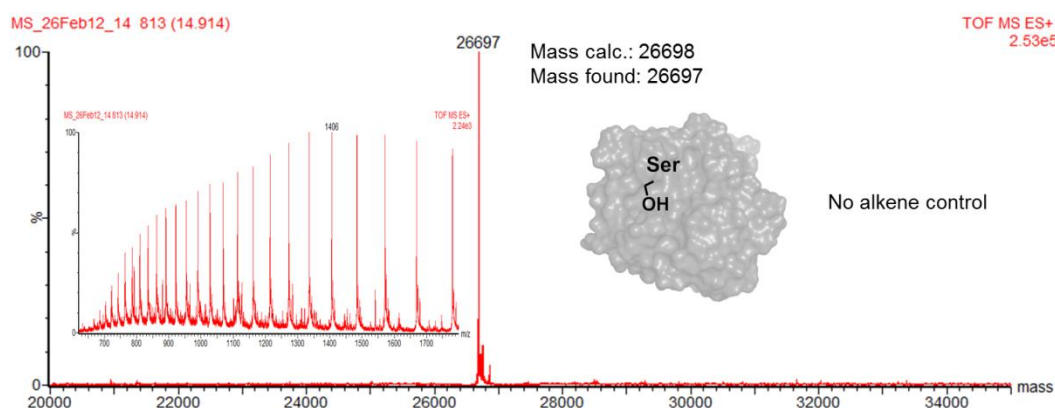


Controls using SBL-WT (No alkene control)

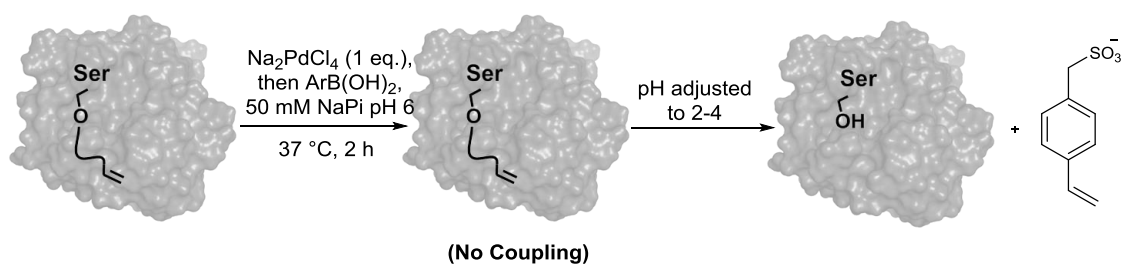
Boronic acid and palladium only:



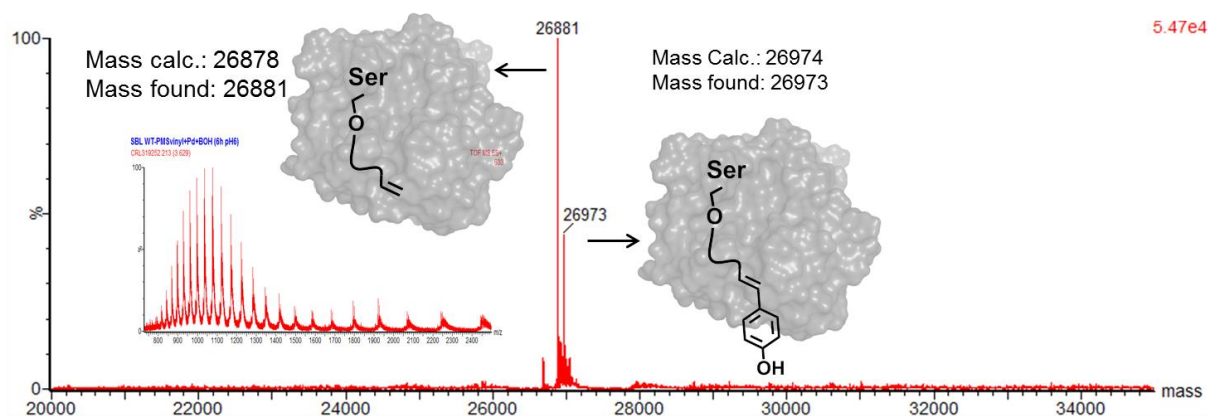
A 50 μL sample of SBL-WT was prepared at 1.0 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 10 mM stock of Na_2PdCl_4 was made in MilliQ® and 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 50 μL sample of SBL-WT (1.9 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^\circ\text{C}$, was added 0.2 μL of Na_2PdCl_4 solution (1.9 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 3.8 μL of 4-hydroxyphenylboronic acid in DMF (1.9 μmol). The reaction mixture was incubated at 37 $^\circ\text{C}$ for 2 h, after which 10 μL aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-WT (calculated mass 26698, mass found 26695). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control reaction is shown below.



Controls using SBL-WT (No UAA control): Boronic acid and Palladium only for prolonged time (6 h)



A 200 μL sample of SBL-WT(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 10 mM stock of Na_2PdCl_4 was made in MilliQ® and a 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μL sample of SBL-WT(PMSv) (5.6 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 °C, was added 0.56 μL of Na_2PdCl_4 solution (5.6 nmol). The reaction solution was vortexed to homogenize, followed by subsequent addition of 11 μL of 4-hydroxyphenylboronic acid in DMF (5.6 μmol). The reaction mixture was incubated at 37 °C for 6 h, after which 10 μL aliquot of the reaction was analysed by LC-MS. A mixture of protein species was detected with a mass corresponding to the SBL-WT(P) 26% (calculated mass 26974, mass found 26973) and rest starting material SBL-WT(PMSv) (calculated mass 26878, mass found 26881). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The ESI-MS for the control reaction is shown below.



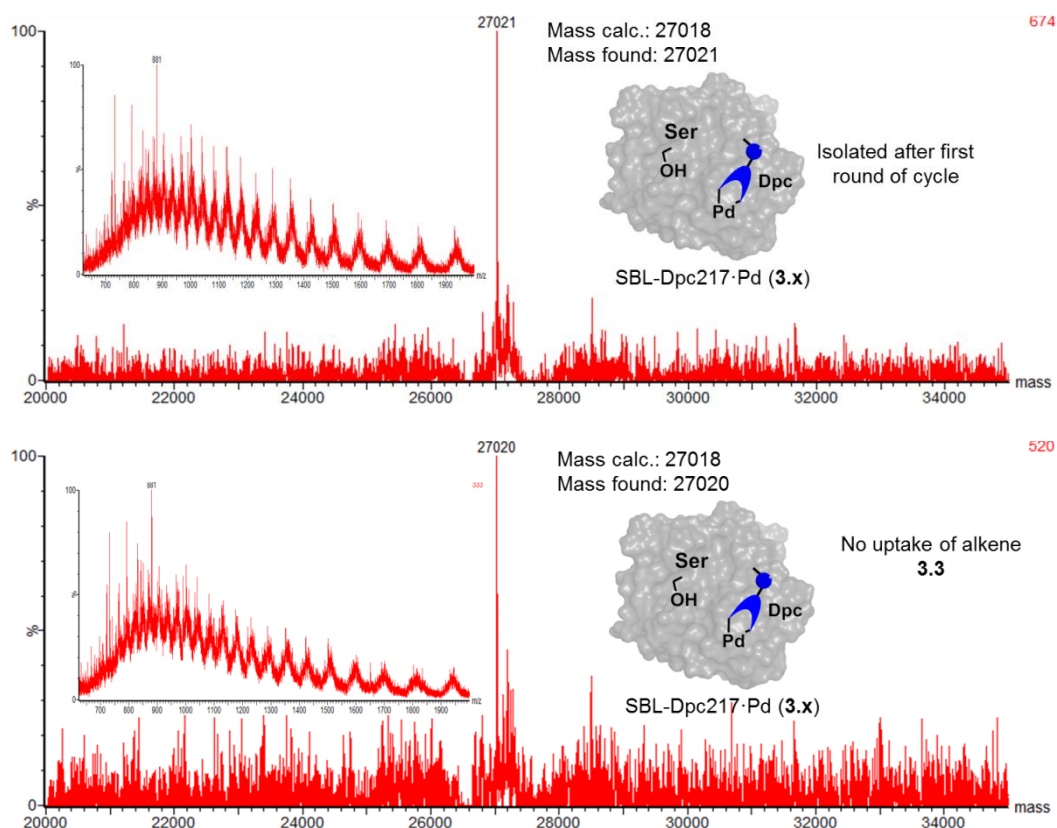
Reloading

3.5.18 Reloading of substrates

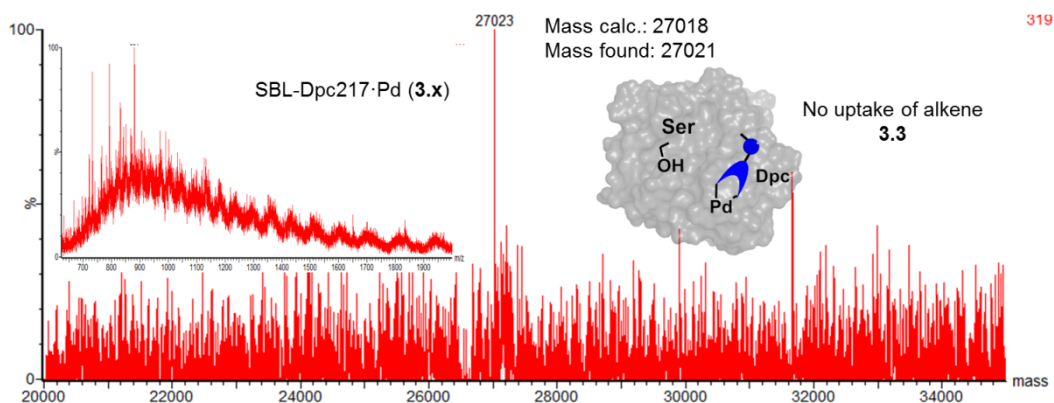
Loading of alkene substrate 3.3 on SBL-Dpc217•Pd:

A 180 μL sample of SBL-Dpc217•Pd **3.25** was prepared at 0.46 mg/mL in 50 mM sodium phosphate (NaPi) buffer, pH 6, after isolation from the first round of the cycle (Section

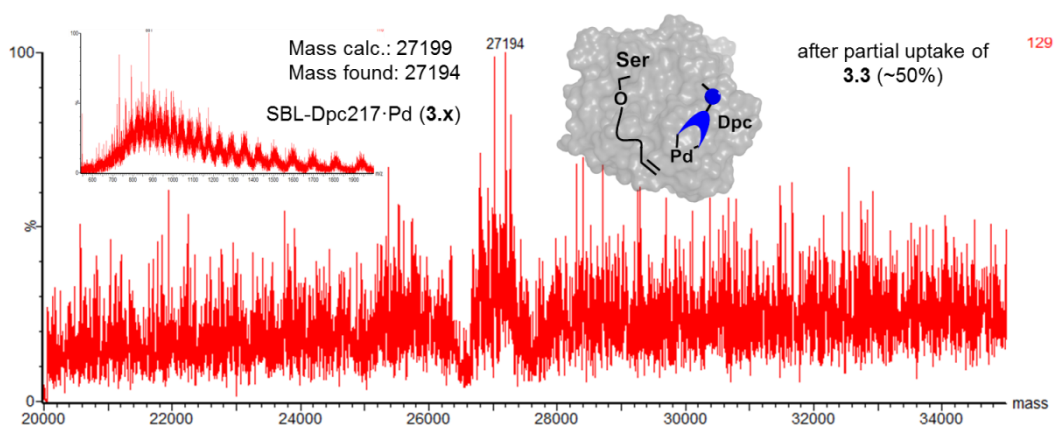
3.2.3.5). The excess reagents were desalted by passing the protein through Sephadex G-25 medium Spintrap column (GE Healthcare) and dialysis against 50 mM NaPi pH 6. A 0.1 M stock solution of *p*-vinylPMSF substrate (**3.3**) was prepared in DMF. To 180 μ L sample of SBL-Dpc217•Pd (3.2 nmol) in 50 mM NaPi pH 6, was added 0.32 μ L **3.3** (32 nmol). The reaction solution was vortexed to homogenize and incubated at room temperature for 10 minutes. LC-MS analysis revealed no uptake of substrate **3.3**, as mass corresponding to SBL-Dpc217•Pd was observed (mass calculated: 27018, mass found: 27020).



Another 0.32 μ L **3.3** (32 nmol) substrate was added and an aliquot was analysed by LC-MS. A single protein species was observed with mass corresponding to SBL-Dpc217•Pd (mass calculated: 27018, mass found: 27023), showing no uptake of substrate **3.3**. ESI-MS for the modification is shown below.



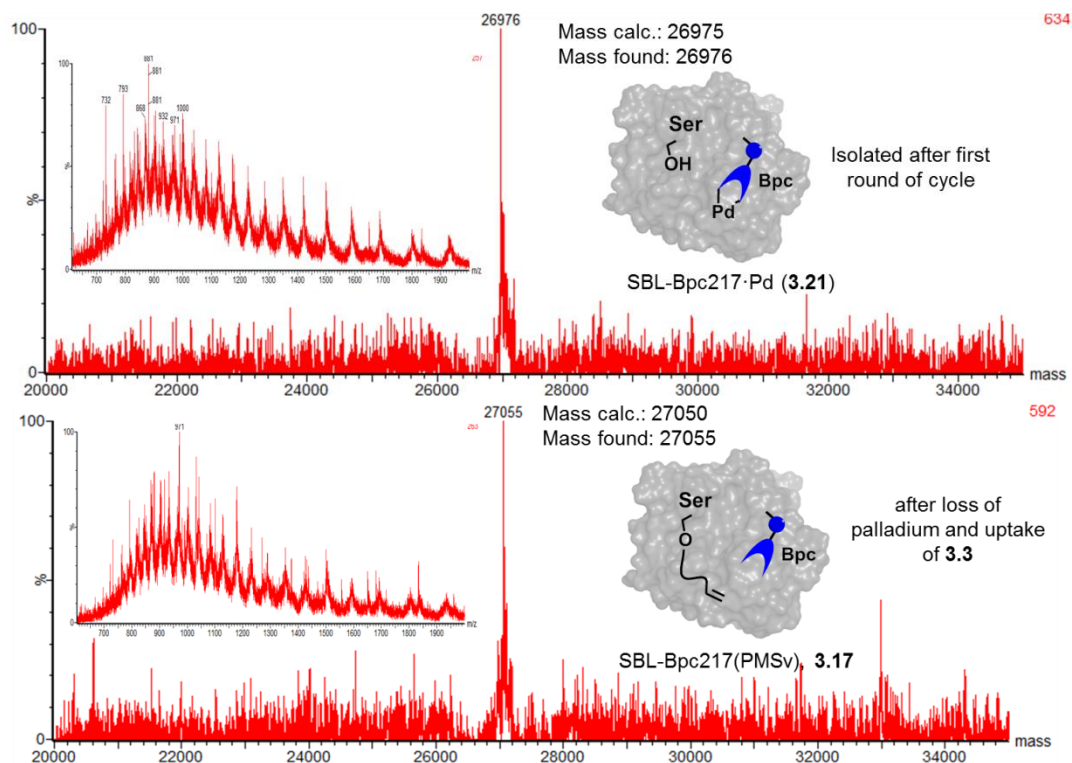
A further aliquot of **3.3** (0.5 μ L, total of 50 equiv.) was added and the reaction was subsequently analysed by LC-MS. Only partial uptake of alkene substrate was observed (50% by LC-MS) giving mass corresponding to SBL-Dpc217-Pd(PMSv) **3.28** (mass calculated: 27199, mass found: 27194).



*Loading of alkene substrate **3.3** on SBL-Bpc217-Pd:*

A 50 μ L sample of SBL-Bpc217-Pd **3.21** was prepared at 0.54 mg/mL in 50 mM sodium phosphate (NaPi) buffer, pH 6, after isolation from the first round of the cycle (section 3.2.3.3). The excess reagents were desalted by passing the protein through Sephadex G-25 medium Spintrap column (GE Healthcare) and dialysis against 50 mM NaPi pH 6. A 0.1 M stock solution of *p*-vinylPMSF substrate (**3.3**) was prepared in DMF. To 50 μ L sample of SBL-Bpc217-Pd (1 nmol) in 50 mM NaPi pH 6, was added 0.5 μ L **3.3** (50 nmol). The reaction solution was vortexed to homogenize and incubated at room temperature for 10

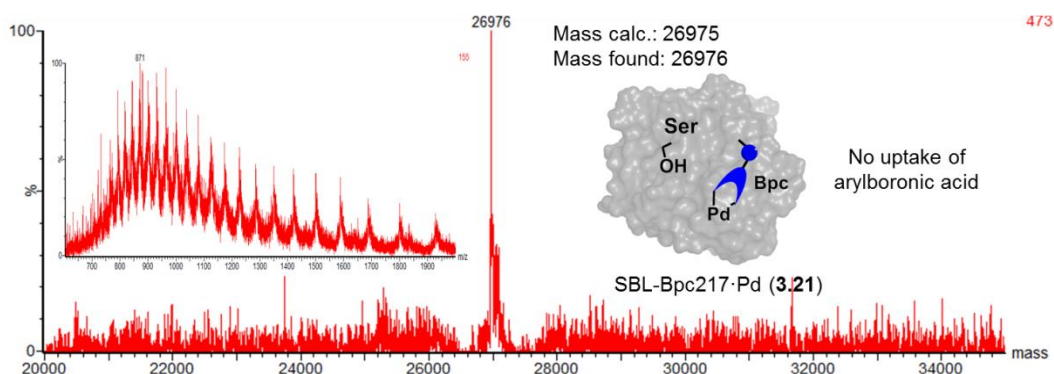
minutes. An aliquot of the reaction was subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217(PMSv) **3.17** after loss of palladium (mass calculated: 27050, mass found: 27055).



*Loading of arylboronic acid **b** on SBL-Bpc217•Pd:*

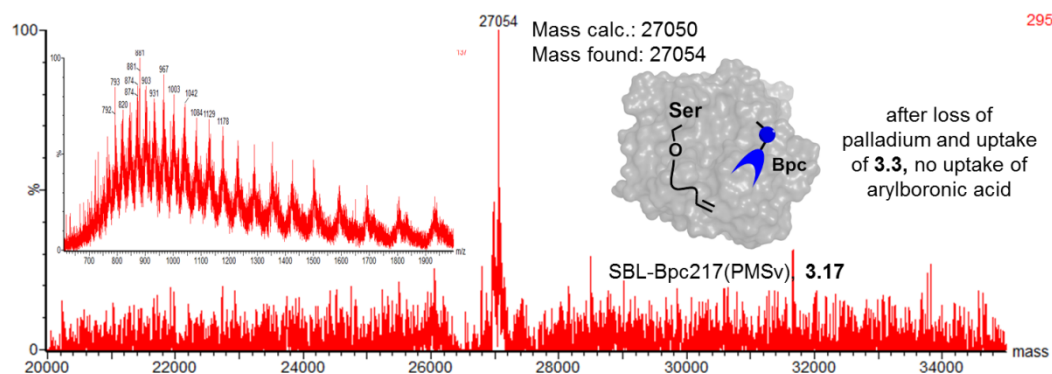
A 50 μ L sample of SBL-Bpc217•Pd **3.21** was prepared at 0.22 mg/mL in 50 mM sodium phosphate (NaPi) buffer, pH 6, after isolation from the first round of the cycle (Section 3.2.3.3). The excess reagents were desalted by passing the protein through Sephadex G-25 medium Spintrap column (GE Healthcare) and dialysis against 50 mM NaPi pH 6. A 0.5 M stock solution of 4-methoxyphenylboronic acid **b** was prepared in DMF. To 50 μ L sample of SBL-Bpc217•Pd (0.5 nmol) in 50 mM NaPi pH 6, was added 1 μ L of 4-methoxyphenylboronic acid **b**. The reaction solution was vortexed to homogenize and incubated at 37 $^{\circ}$ C for 1 h. An aliquot of the reaction was subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-

Bpc217•Pd **3.21** showing no uptake of boronic acid substrate **b** (mass calculated: 27975, mass found: 27976).



*Simultaneous loading of alkene substrate **3.3** and arylboronic acid on SBL-Bpc217•Pd:*

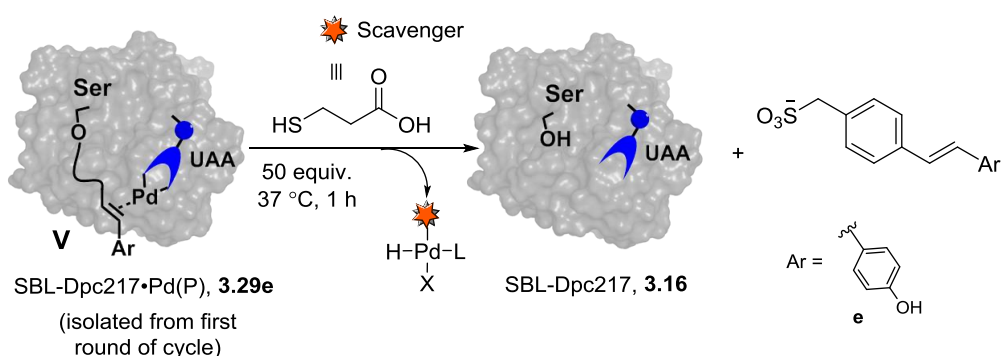
A 50 μ L sample of SBL-Bpc217•Pd **3.21** was prepared at 0.22 mg/mL in 50 mM sodium phosphate (NaPi) buffer, pH 6, after isolation from the first round of the cycle (section 3.2.3.3). The excess reagents were desalted by passing the protein through Sephadex G-25 medium Spintrap column (GE Healthcare) and dialysis against 50 mM NaPi pH 6. A 0.1 M stock solution of *p*-vinylPMSF substrate (**3.3**) and 0.5 M stock solution of 4-methoxyphenylboronic acid **b** was prepared in DMF. To 50 μ L sample of SBL-Bpc217•Pd (0.5 nmol) in 50 mM NaPi pH 6, was added 0.25 μ L **3.3** (25 nmol) and 1 μ L of 4-methoxyphenylboronic acid **b**. The reaction solution was vortexed to homogenize and incubated at 37 °C for 30 minutes. An aliquot of the reaction was subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217(PMSv) **3.17** after loss of palladium (mass calculated: 27050, mass found: 27054).



3.5.19 CD studies

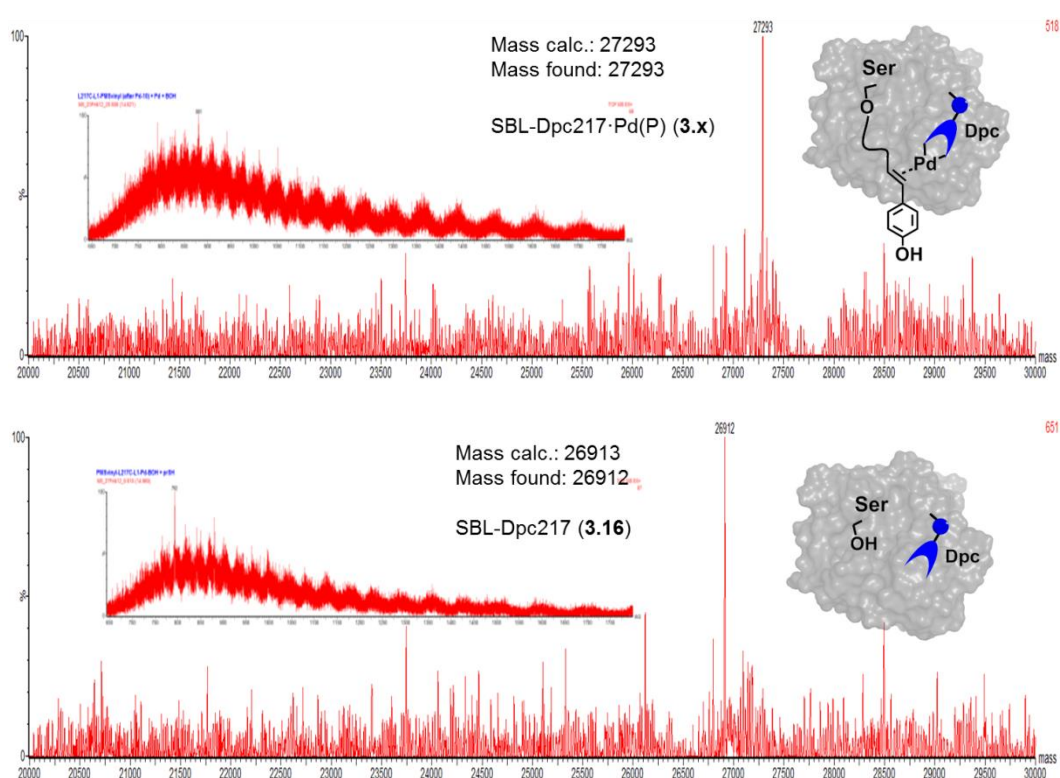
CD measurements were made using a Chirascan spectropolarimeter fitted with a Peltier temperature controller. SBL-WT and modified SBL-Bpc217 and SBL-Dpc217 were each made at 0.5 mg/mL solutions each in 50 mM sodium phosphate, pH 6. CD spectra were measured for each intermediate of the engineered cycle (Section 3.2.3.3), in a 1 mm quartz cuvette at room temperature using a scan rate of 50 nm/ min, 1 nm interval, 1 nm bandwidth and a response time of 0.5 s. After baseline correction, ellipticities in deg were converted to molar ellipticities ($\text{deg cm}^2 \text{ dmol}^{-1}$) by normalizing for the concentration of peptide bonds and path length.

3.5.20 Simultaneous final product hydrolysis and metal scavenging

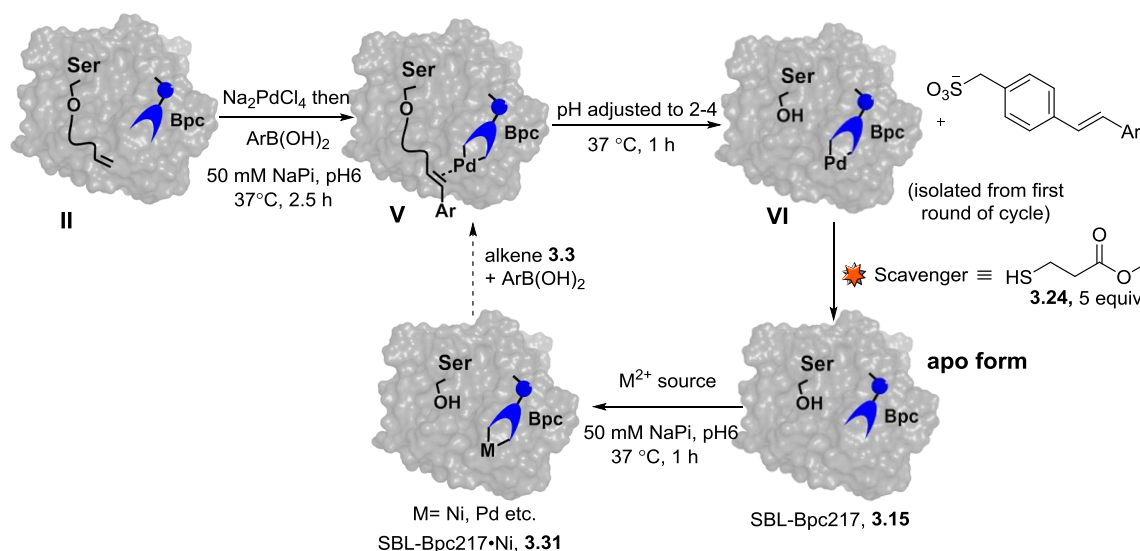


A 100 μL sample of SBL-Dpc217•Pd(P) **3.29e** was prepared at 0.73 mg/mL in 50 mM sodium phosphate (NaPi) buffer, pH 6, after isolation from the first round of the cycle (Section 3.2.3.3). To 70 μL sample of SBL-Dpc217•Pd(P) **3.29e** (1.9 nmol) in 50 mM

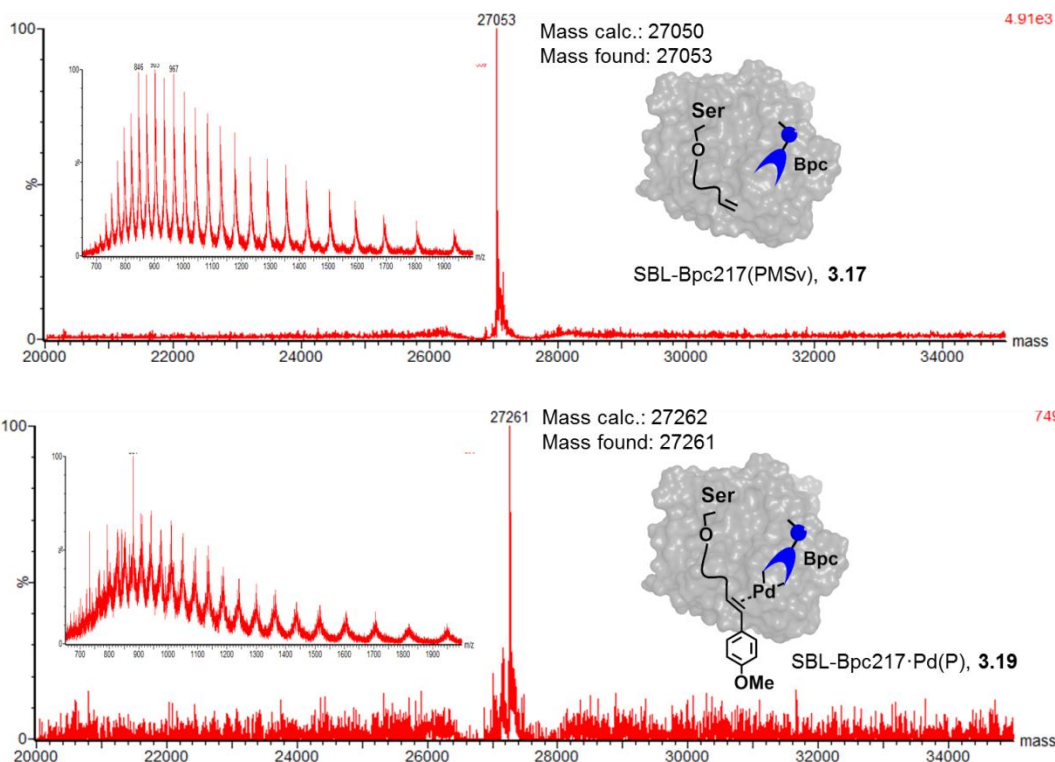
sodium phosphate buffer, pH 6 at 4°C, was added 0.29 μL (143 nmol) of mercaptopropionic acid. The reaction solution was vortexed to homogenize and was incubated at 37 °C for 1 h. The reaction was analysed by LC-MS on completion. A single protein species was detected with a mass corresponding to the SBL-Dpc217 **3.16** (calculated mass 26913, mass found 26912) showing simultaneous scavenging of palladium from the active site and hydrolysis of the final product due to pH change. ESI-MS for the modification is shown below.



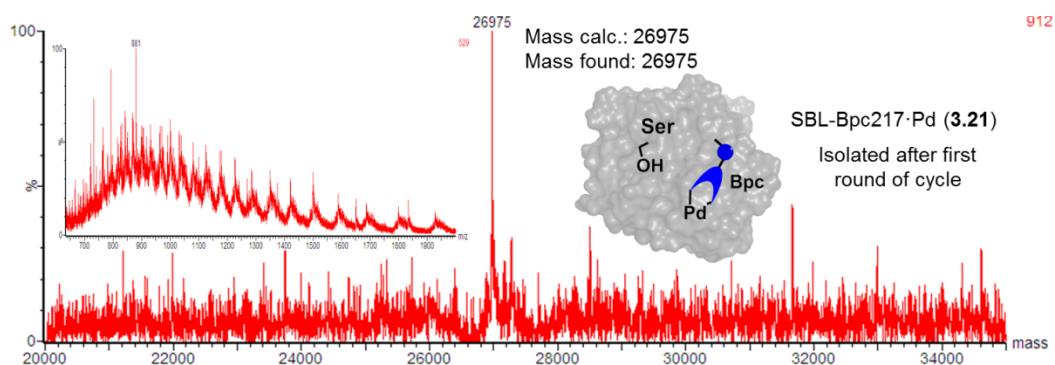
3.5.21 Reloading with a different metal



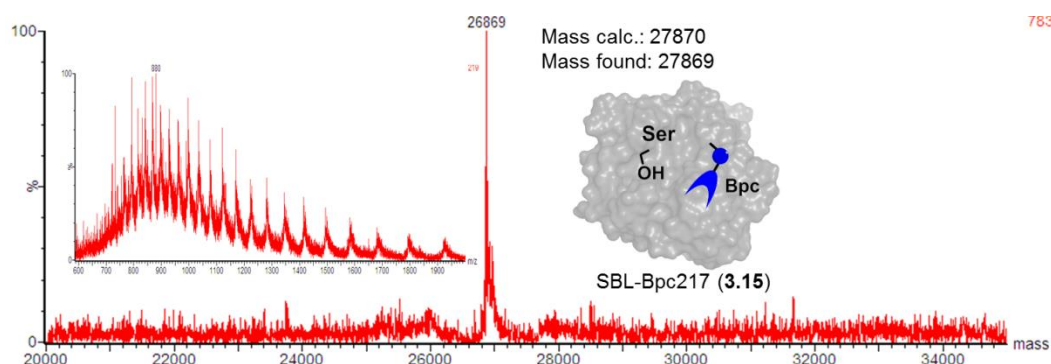
Cross-coupling reaction: A 360 μL sample of SBL-Bpc217(PMSv) **3.17** was prepared at 0.8 mg/mL in 50 mM sodium phosphate buffer, as described before (Section 3.2.3). A 0.5 M stock solution of 4-methoxyphenylboronic acid (**b**) was prepared in DMF and 0.01 M Na_2PdCl_4 was prepared in Milli-Q® water. To 360 μL sample of SBL-Dpc217(PMSv) (10.6 nmol) in 50 mM sodium phosphate buffer at 4 $^\circ\text{C}$, was added 1.0 μL of Na_2PdCl_4 , followed by 20 μL of boronic acid **b** in DMF (1.06 μmol). The reaction solution was vortexed to homogenize and incubated at 37 $^\circ\text{C}$ for 2 h. An aliquot of the reaction was subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Bpc217•Pd(P) **3.19** tethering the cross-coupled product (mass calculated: 27262, mass found: 27261). The excess reagents were desalted by passing the protein through Sephadex G-25 medium spintrap column (GE Healthcare). The protein solution was stored at 4 $^\circ\text{C}$ until further use.



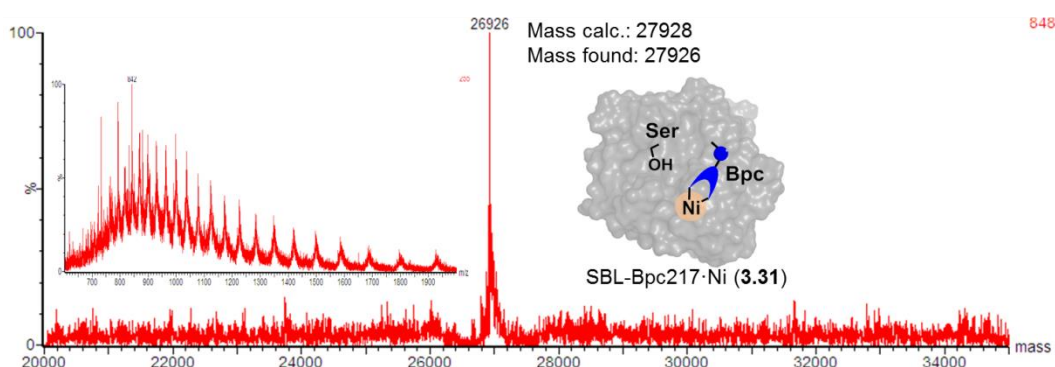
Hydrolysis of final product 'P': A 1 M stock of HCl was prepared. To a 180 μ L sample of SBL-Dpc217•Pd(P) (6 nmol), was added 12 μ L of HCl to adjust pH to 2-4. The reaction was incubated at 37 °C for 1 h, after which an aliquot was analysed by LC-MS. A single protein species was observed with mass corresponding to SBL-Bpc217•Pd (mass calculated: 26975, mass found: 26975) confirming hydrolysis of the cross-coupled product. ESI-MS of the modification is shown below. The protein was dialysed against 50 mM NaPi pH 6 for 2 h and stored at 4 °C until further use.



Metal scavenging: A 0.01 M stock solution methyl, 3-mercaptopropionate ester (MPE) was prepared in Milli-Q® water. To a 180 µL sample of SBL-Dpc217•Pd(P) (6 nmol), was added 1.2 µL of MPE (12 nmol). The reaction was incubated at 37 °C for 30 minutes, after which an aliquot was analysed by LC-MS. A single protein species was observed with mass corresponding to SBL-Bpc217•Pd (mass calculated: 26870, mass found: 26869), confirming scavenging of palladium. ESI-MS of the modification is shown below.

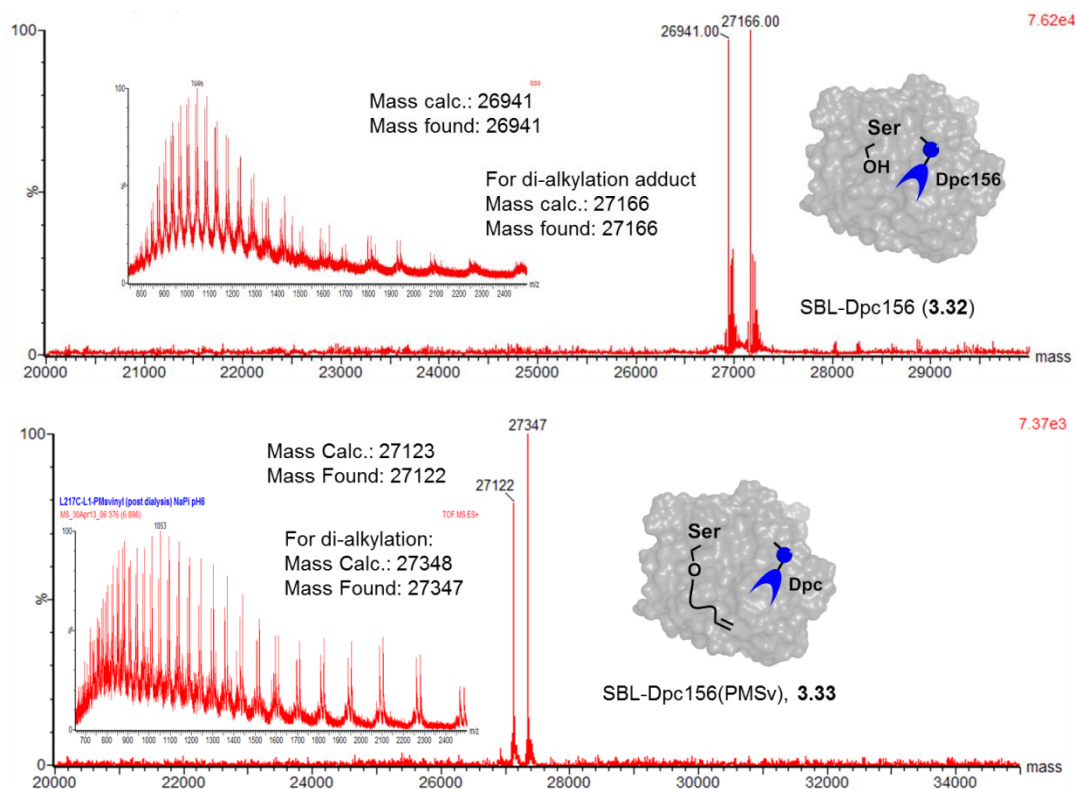


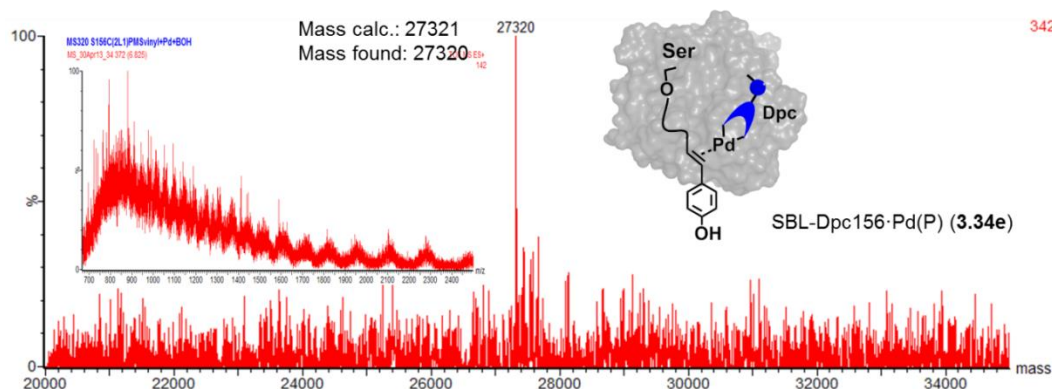
Metal reloading: A 0.01 M stock solution of Ni(OAc)₂ was prepared in Milli-Q® water. To a 90 µL sample of SBL-Dpc217•Pd(P) (3 nmol), was added 0.3 µL of Ni(OAc)₂ solution (3 nmol). The reaction was incubated at 37 °C for 30 minutes, after which an aliquot was analysed by LC-MS. A single protein species was observed with mass corresponding to SBL-Bpc217•Ni (mass calculated: 26928, mass found: 26926), confirming reloading of nickel. ESI-MS of the modification is shown below.



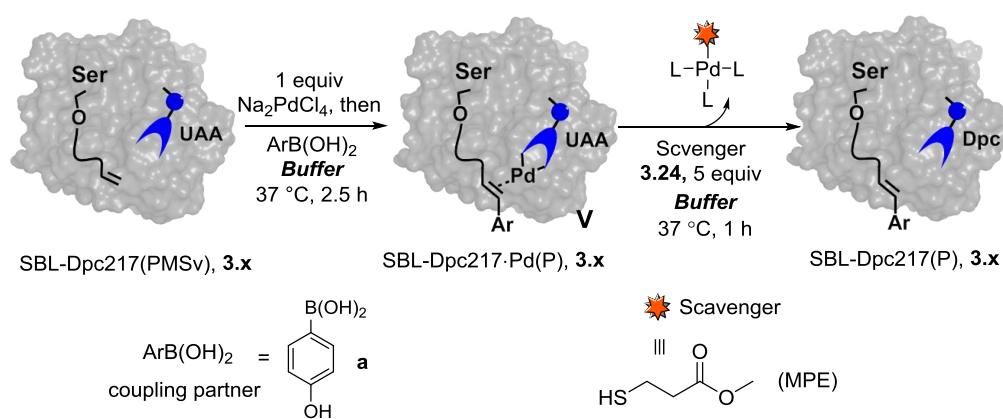
3.5.22 Oxidative Heck cross-coupling reaction using SBL-Dpc156

A 200 μL sample of SBL-Dpc156(PMSv) was prepared at 0.75 mg/mL in sodium phosphate buffer (pH 6, 50 mM). A 0.5 M stock solution of 4-hydroxyphenylboronic acid was prepared in DMF. To 200 μL sample of SBL-Dpc156(PMSv) (5.6 nmol) in 50 mM sodium phosphate buffer, pH 6 at 4 $^{\circ}\text{C}$, was added 11 μL of 4-hydroxyphenylboronic acid in DMF (5.6 μmol). The reaction mixture was incubated at 37 $^{\circ}\text{C}$ for 2 h, after which 10 μL aliquot of the reaction was analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217(PMSv) (calculated mass 27321, mass found 27320). The excess boronic acid was desalted by passing the protein through a Sephadex G-25 medium (Amersham Biosciences). The protein was eluted with sodium phosphate buffer (50 mM, pH 6) and stored on ice until further needed. The ESI-MS for the control reaction is shown below.





3.5.23 Investigation of pH dependence of cross-coupling reaction



General procedure:

Cross-coupling reaction (Step i): A 360 μL sample of SBL-Dpc217(PMSv) **3.18** was prepared at 0.8 mg/mL in 50 mM **buffer**, as described before (Section 3.5.8). A 0.5 M stock solution of 4-hydroxyphenylboronic acid (**a**) was prepared in DMF and 0.01 M Na_2PdCl_4 was prepared in Milli-Q® water. To 360 μL sample of SBL-Dpc217(PMSv) (10.6 nmol) in 50 mM **buffer** at 4 °C, was added 1.063 μL of Na_2PdCl_4 , followed by 21.2 μL of boronic acid **a** in DMF (1.45 μmol). The reaction solution was vortexed to homogenize and incubated at 37 °C for 2.5 h. an aliquot of the reaction was subsequently analysed by LC-MS. A single protein species was detected with a mass corresponding to the SBL-Dpc217•Pd(P) **3.29e** tethering the cross-coupled product (mass calculated: 27293, mass found: 27295). The excess reagents were desalted by passing the protein through spintrap Sephadex G-25 medium column (GE Healthcare). The sample was dialysed

against respective buffer for 2.5 h and protein solution was stored at 4 °C until further use.

Following four buffer systems were used:

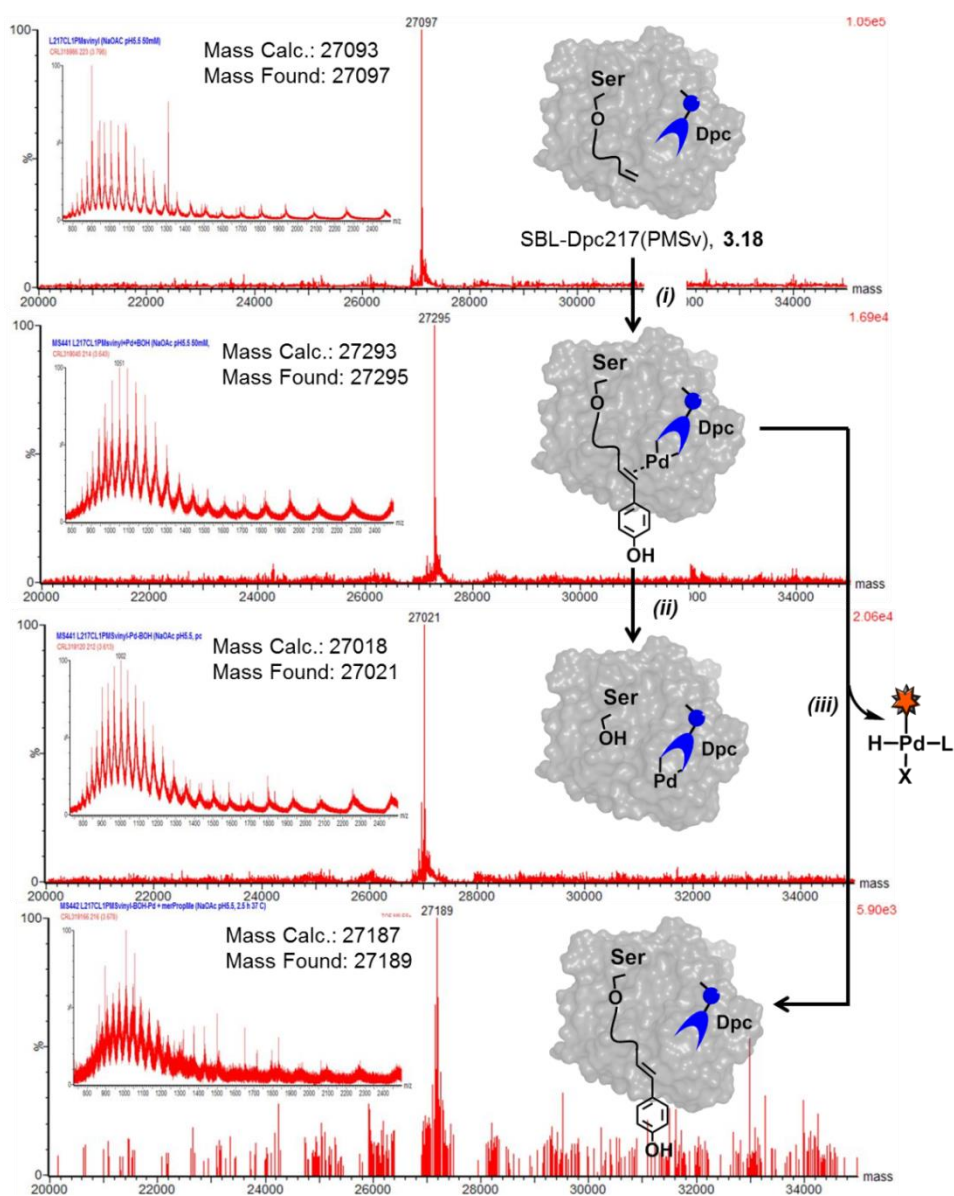
- (i) 50 mM sodium acetate buffer (NaOAc), pH 5.5
- (ii) 50 mM sodium phosphate buffer (NaPi), pH 6
- (iii) 50 mM 3-(N-morpholino)propanesulfonic acid buffer (MOPS), pH 7
- (iv) 50 mM sodium phosphate buffer (NaPi), pH 8

Hydrolysis of final product 'P' (Step ii): A 180 µL sample of SBL-Dpc217•Pd(P) **3.29e** was prepared at 0.25 mg/ml in 50 mM **buffer**. 1 M stock of HCl was prepared. To a 180 µL sample of SBL-Dpc217•Pd(P) (0.95 nmol), was added 5 µL of HCl to adjust pH to 2-4. The reaction was incubated at 37 °C for 1 h, after which an aliquot was analysed by LC-MS.

Palladium Scavenging (Step iii): A 100 µL sample of SBL-Dpc217•Pd(P) **3.29e** was prepared at 0.25 mg/ml in 50 mM **buffer** and 0.005 M stock solution methyl, 3-mercaptopropionate ester (MPE) was prepared in Milli-Q® water. To a 100 µL sample of SBL-Dpc217•Pd(P) (0.95 nmol), was added 0.76 µL of MPE (3.78 nmol). The reaction was incubated at 37 °C for 1 h, after which an aliquot was analysed by LC-MS.

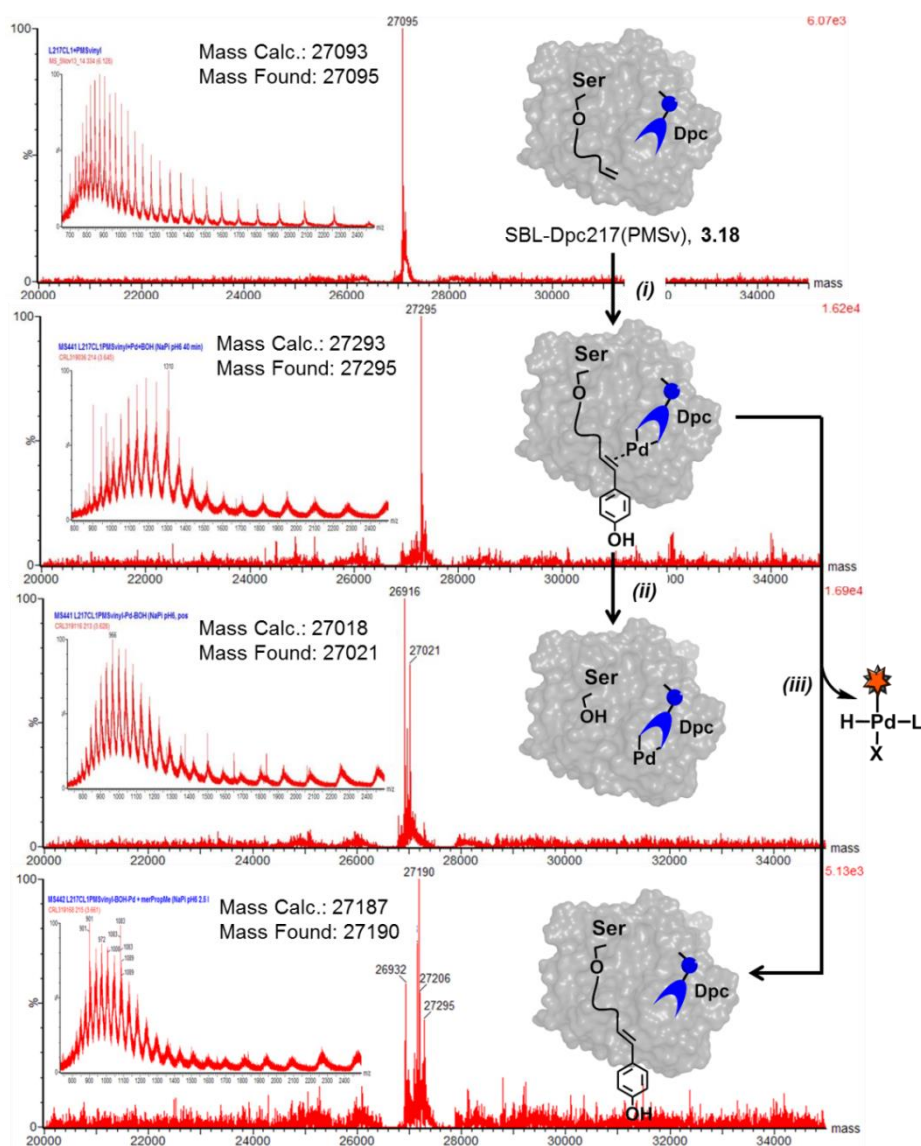
Oxidative Heck cross-coupling reaction for alkene **3.3** and 4-hydroxyphenylboronic acid using engineered SBL-Dpc217•Pd in different buffer systems was carried out using general procedure (*Step i*). After reaction the reaction was divided into two parts and discussed below.

50 mM NaOAc, pH 5.5: To a 180 μ L sample of SBL-Dpc217•Pd(P) (1.69 nmol), was added 5 μ L of HCl to adjust pH to 2-4 for hydrolysis of final product (*Step ii*). LC-MS analysis revealed conversion to SBL-Dpc217•Pd (mass calculated: 27018, mass found: 27021). To a 100 μ L sample of SBL-Dpc217•Pd(P) (0.95 nmol), was added 0.76 μ L of MPE (3.78 nmol) for palladium scavenging (*Step iii*). LC-MS analysis revealed conversion to SBL-Dpc217(P) (mass calculated: 27187, mass found: 27189).



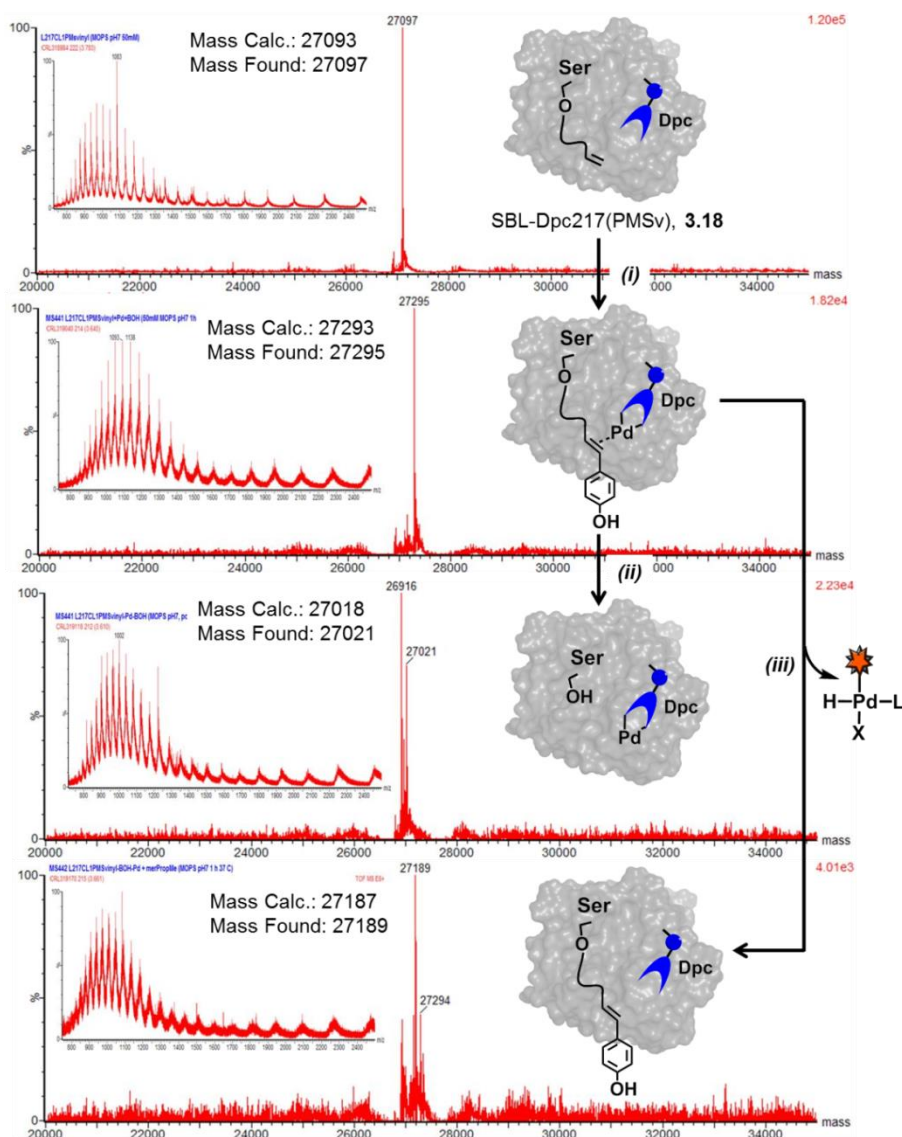
Scheme 3-34. Conditions: (i) 1 equiv. Na₂PdCl₄, 100 equiv. ArB(OH)₂ (ii) 4 equiv. MPE, 1 h 37 °C (iii) pH adjusted to 2-4, 1 h 37 °C

50 mM NaPi, pH 6: To a 180 μ L sample of SBL-Dpc217•Pd(P) (2.18 nmol), was added 5 μ L of HCl to adjust pH to 2-4 for hydrolysis of final product (*Step ii*). LC-MS analysis revealed conversion to SBL-Dpc217•Pd (mass calculated: 27018, mass found: 27021) and SBL-Dpc217 (mass calculated: 26913, mass found: 29916). To a 100 μ L sample of SBL-Dpc217•Pd(P) (1.21 nmol), was added 0.97 μ L of MPE (4.84 nmol) for palladium scavenging (*Step iii*). LC-MS analysis revealed conversion to SBL-Dpc217(P) (mass calc.: 27187, mass found: 27190) and some hydrolysis of final product was also observed (mass found: 26932).



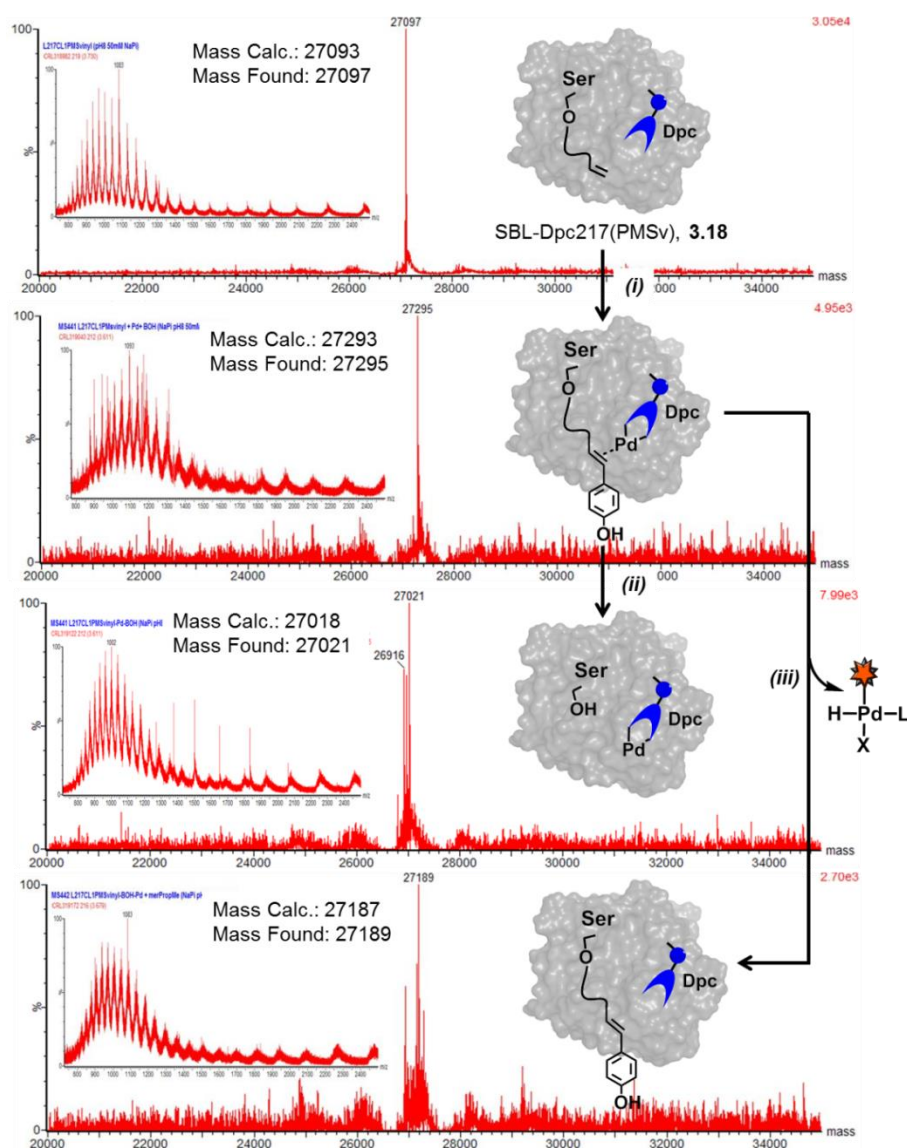
Scheme 3-35. Conditions: (i) 1 equiv. Na_2PdCl_4 , 100 equiv. ArB(OH)_2 (ii) 4 equiv. MPE, 1 h 37 $^\circ\text{C}$ (iii) pH adjusted to 2-4, 1 h 37 $^\circ\text{C}$

50 mM MOPS, pH 7: To a 180 μ L sample of SBL-Dpc217•Pd(P) (1.98 nmol), was added 5 μ L of HCl to adjust pH to 2-4 for hydrolysis of final product (*Step ii*). LC-MS analysis revealed conversion to SBL-Dpc217•Pd (mass calculated: 27018, mass found: 27021) and SBL-Dpc217 (mass calculated: 26913, mass found: 29916). To a 100 μ L sample of SBL-Dpc217•Pd(P) (1.08 nmol), was added 0.88 μ L of MPE (4.34 nmol) for palladium scavenging (*Step iii*). LC-MS analysis revealed conversion to SBL-Dpc217(P) (mass calc.: 27187, mass found: 27189) and some hydrolysis of final product was also observed (mass found: 26932).



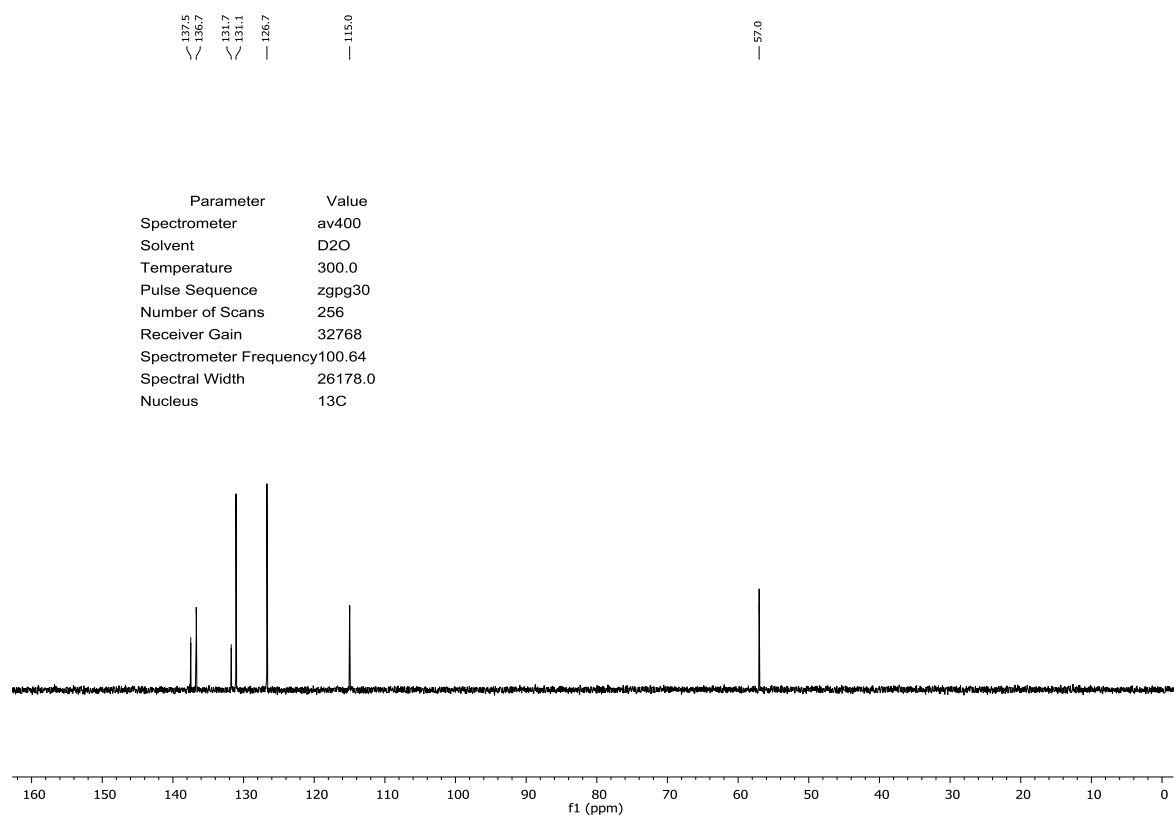
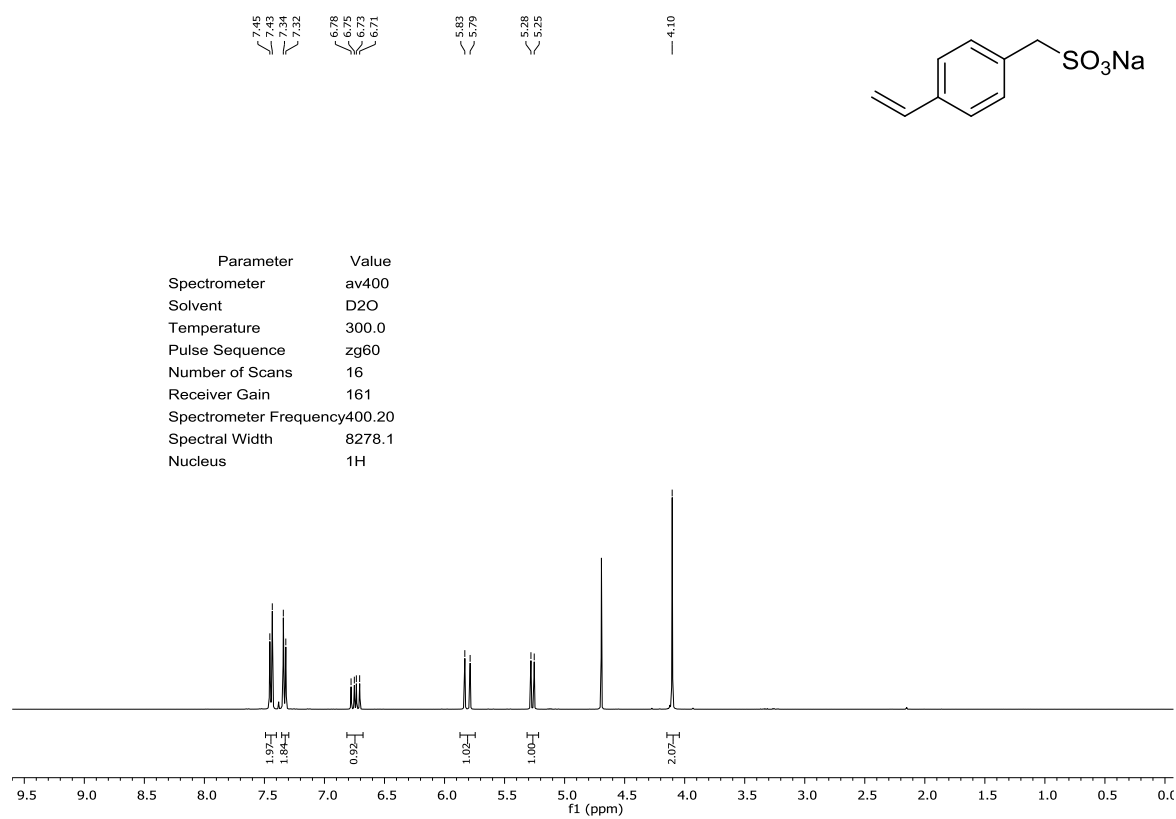
Scheme 3-36. Conditions: (i) 1 equiv. Na_2PdCl_4 , 100 equiv. $\text{ArB}(\text{OH})_2$ (ii) 4 equiv. MPE, 1 h 37 $^\circ\text{C}$ (iii) pH adjusted to 2-4, 1 h 37 $^\circ\text{C}$

50 mM NaPi, pH 6: To a 180 μ L sample of SBL-Dpc217•Pd(P) (2.42 nmol), was added 5 μ L of HCl to adjust pH to 2-4 for hydrolysis of final product (*Step ii*). LC-MS analysis revealed conversion to SBL-Dpc217•Pd (mass calculated: 27018, mass found: 27021) and SBL-Dpc217 (mass calculated: 26913, mass found: 29916). To a 100 μ L sample of SBL-Dpc217•Pd(P) (1.35 nmol), was added 1.08 μ L of MPE (5.4 nmol) for palladium scavenging (*Step iii*). LC-MS analysis revealed conversion to SBL-Dpc217(P) (mass calc.: 27187, mass found: 27189) and some hydrolysis of final product was also observed (mass found: 26932).



Scheme 3-37. Conditions: (i) 1 equiv. Na_2PdCl_4 , 100 equiv. $\text{ArB}(\text{OH})_2$ (ii) 4 equiv. MPE, 1 h 37 $^\circ\text{C}$ (iii) pH adjusted to 2-4, 1 h 37 $^\circ\text{C}$

NMR spectra



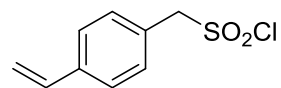
7.53
7.51
7.48
7.46

6.80
6.78
6.76
6.73

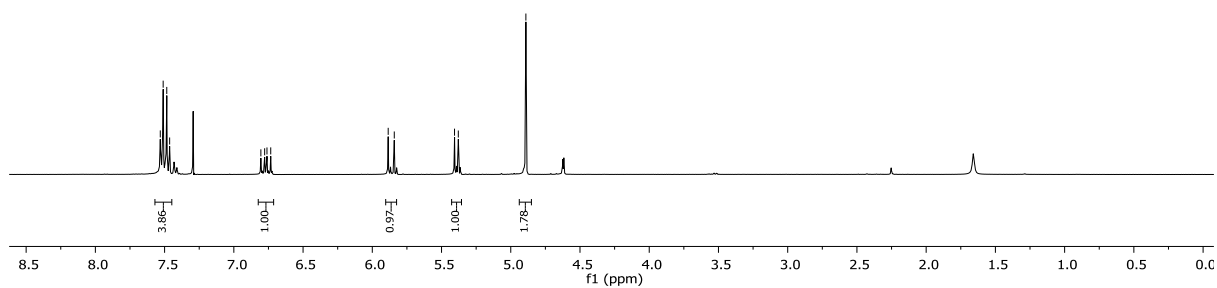
5.89
5.84

5.41
5.38

4.89



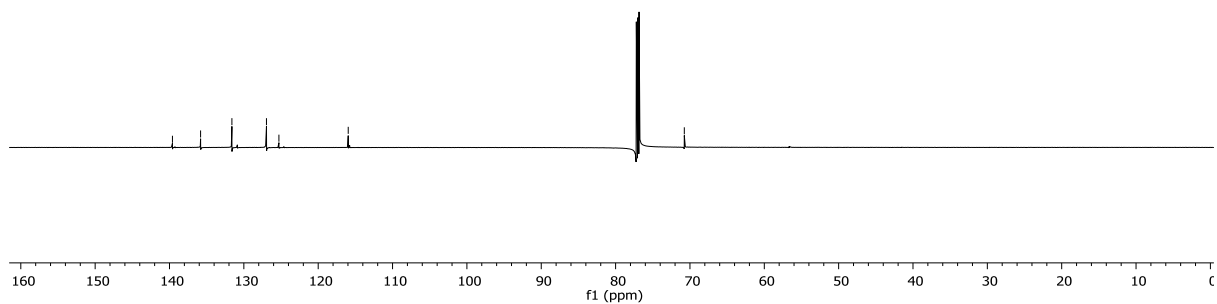
Parameter	Value
Spectrometer	avg400
Solvent	CDCl ₃
Temperature	294.1
Pulse Sequence	zg30
Number of Scans	16
Receiver Gain	412
Spectrometer Frequency	400.20
Spectral Width	10000.0
Nucleus	¹ H

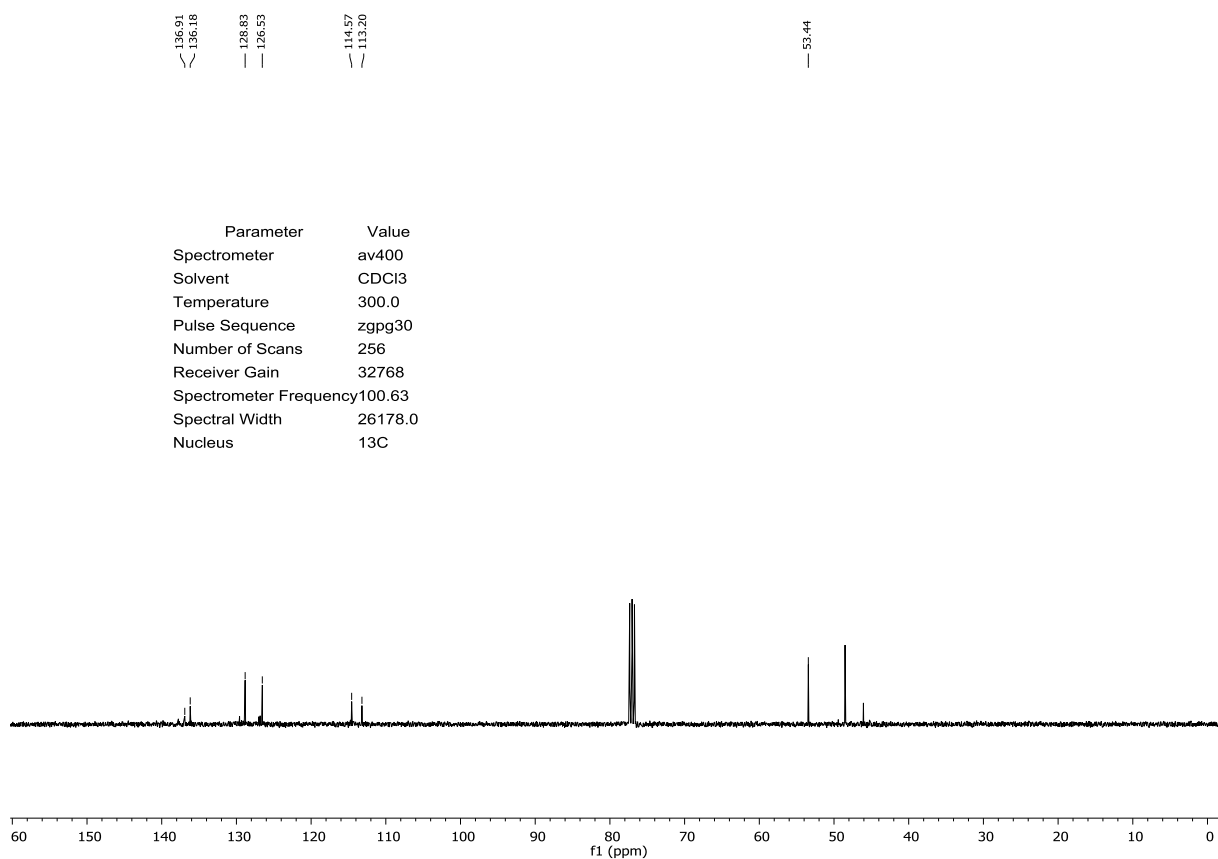
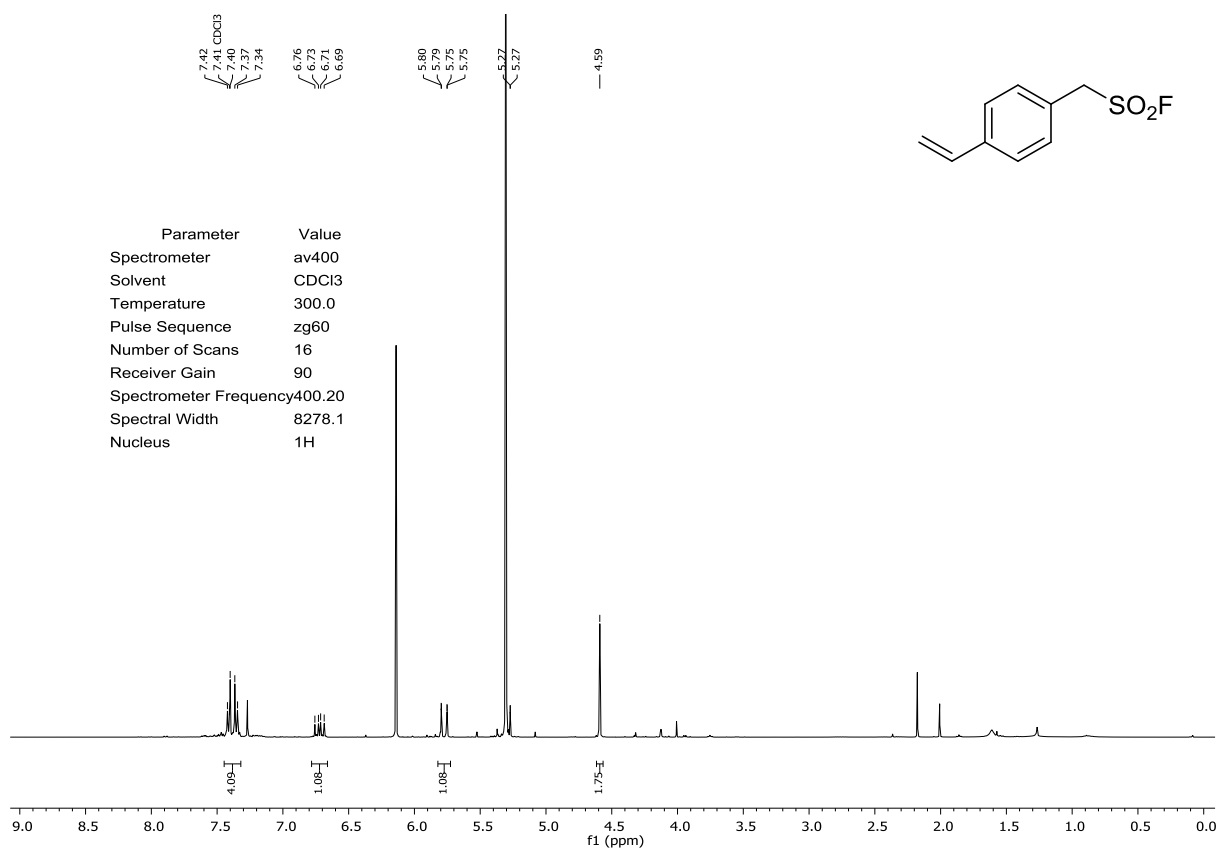


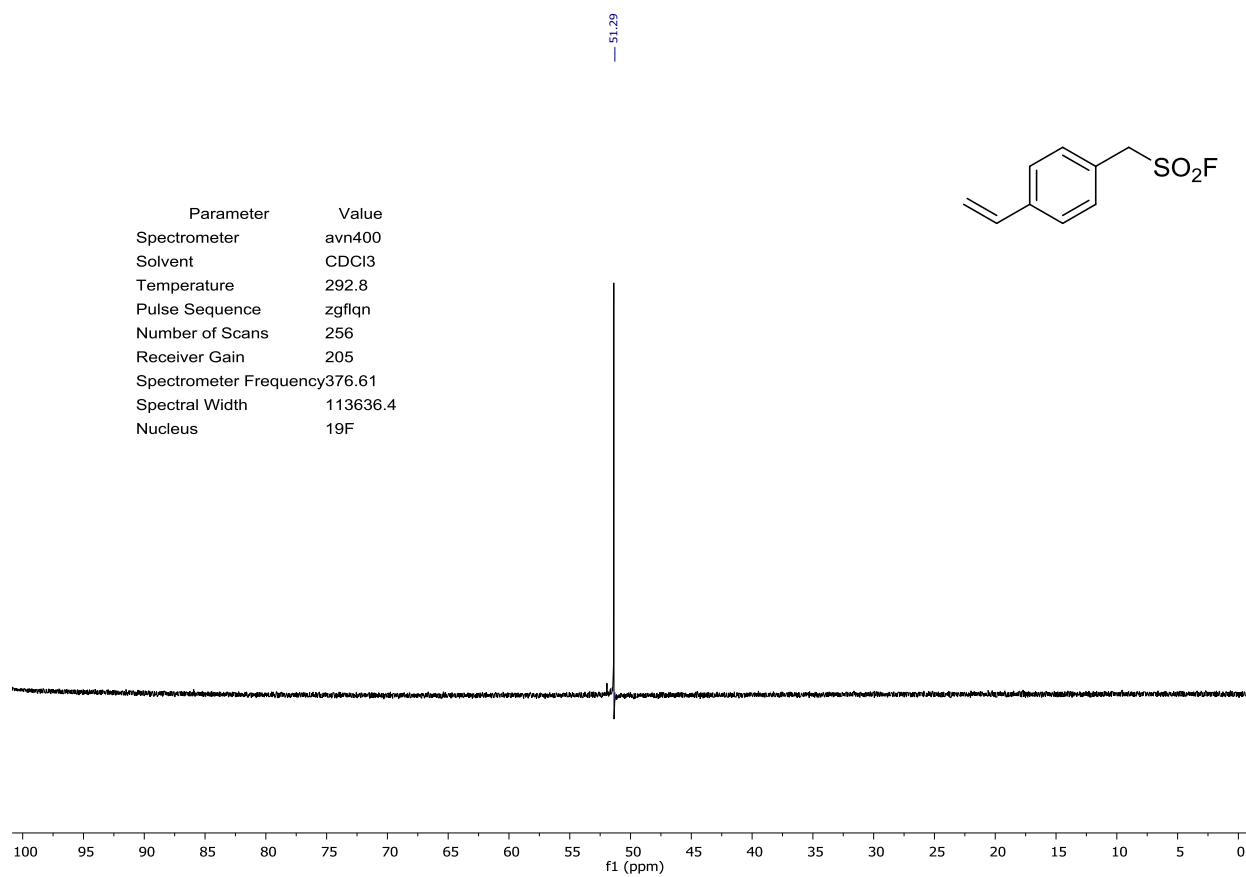
139.60
135.82
131.61
126.96
125.28
115.96

70.76

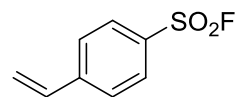
Parameter	Value
Spectrometer	av700
Solvent	CDCl ₃
Temperature	298.0
Pulse Sequence	zgpg30
Number of Scans	1486
Receiver Gain	2050
Spectrometer Frequency	176.01
Spectral Width	41666.7
Nucleus	¹³ C



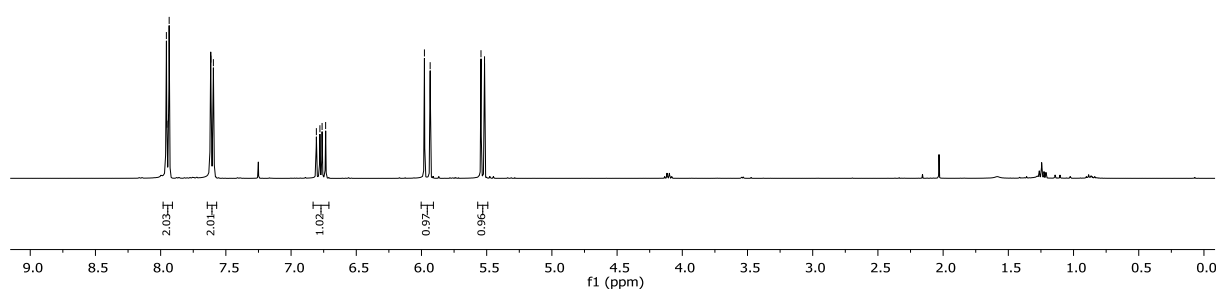




7.96
7.95
7.94
7.93
7.62
7.61
7.60
6.81
6.78
6.76
6.73
5.98
5.93
5.54
5.52

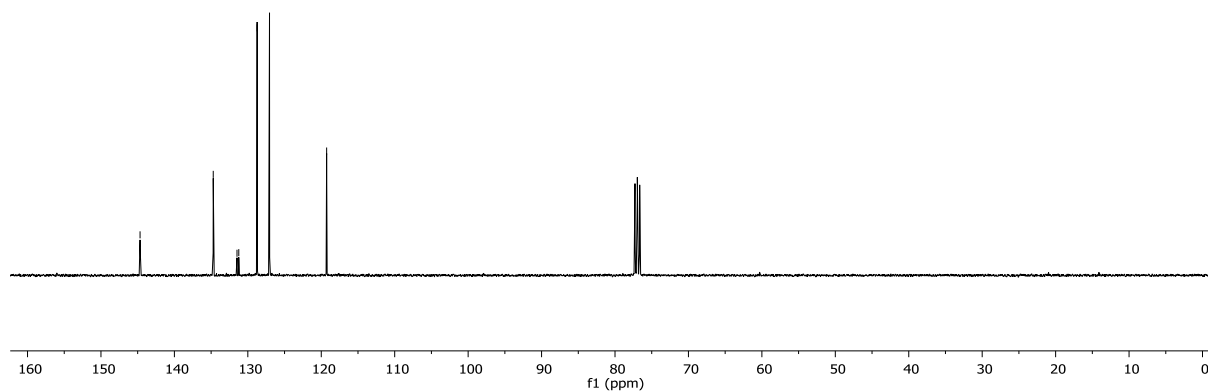


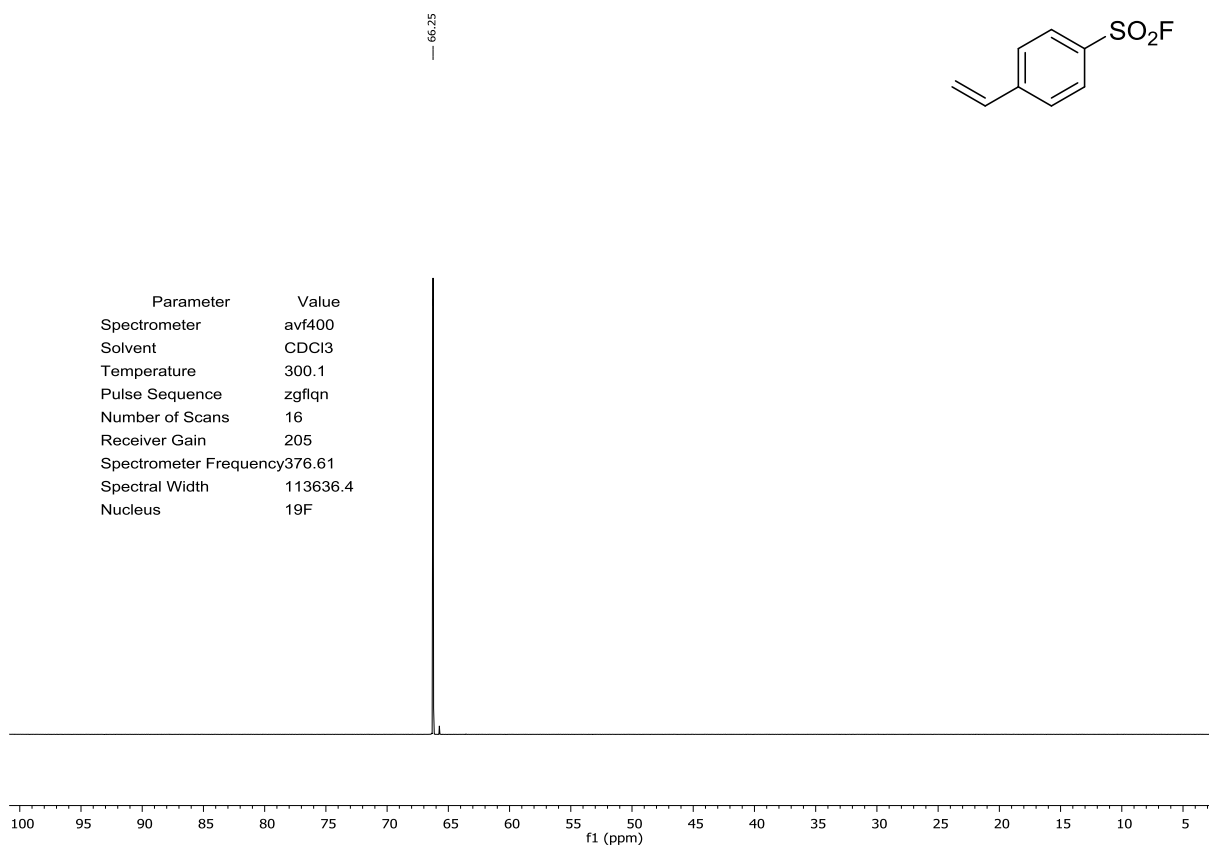
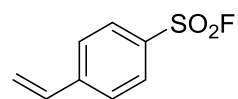
Parameter	Value
Spectrometer	avg400
Solvent	CDCl ₃
Temperature	294.6
Pulse Sequence	zg30
Number of Scans	16
Receiver Gain	142
Spectrometer Frequency	400.20
Spectral Width	10000.0
Nucleus	¹ H

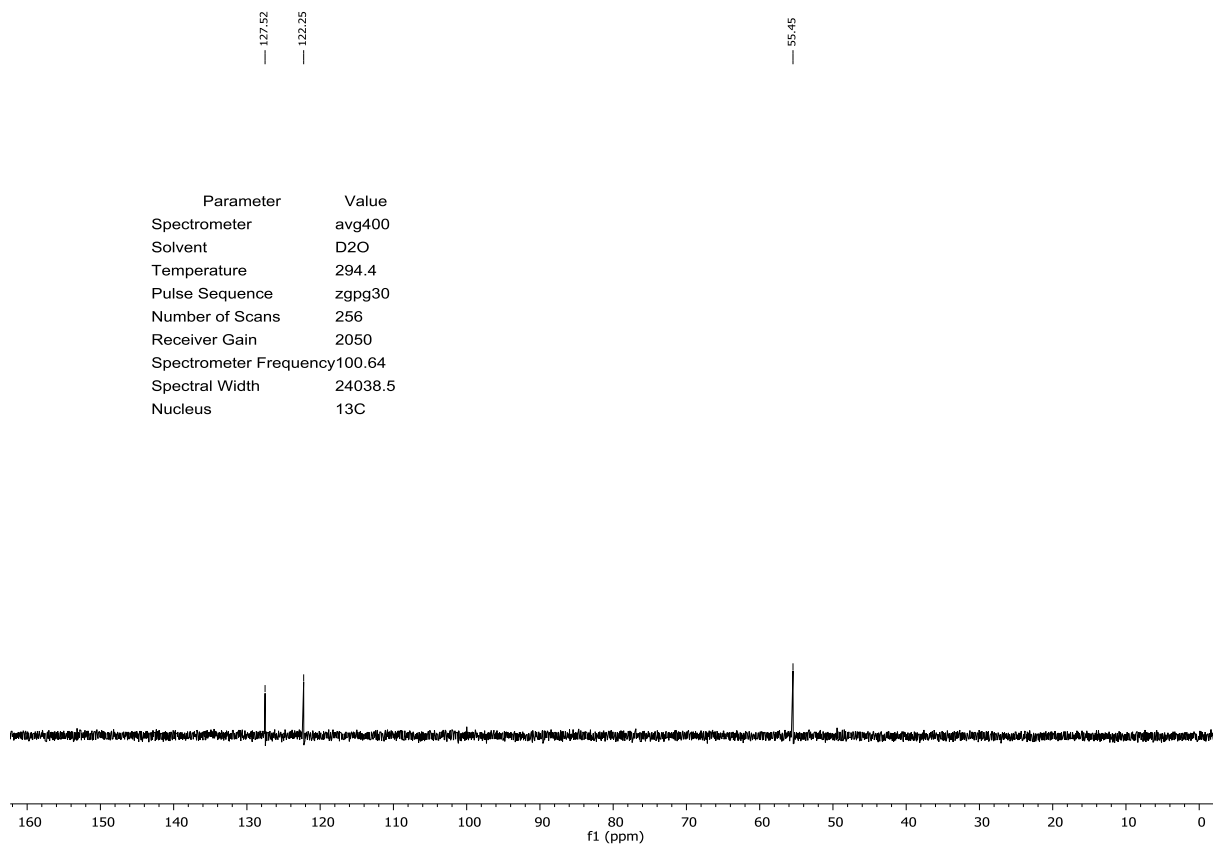
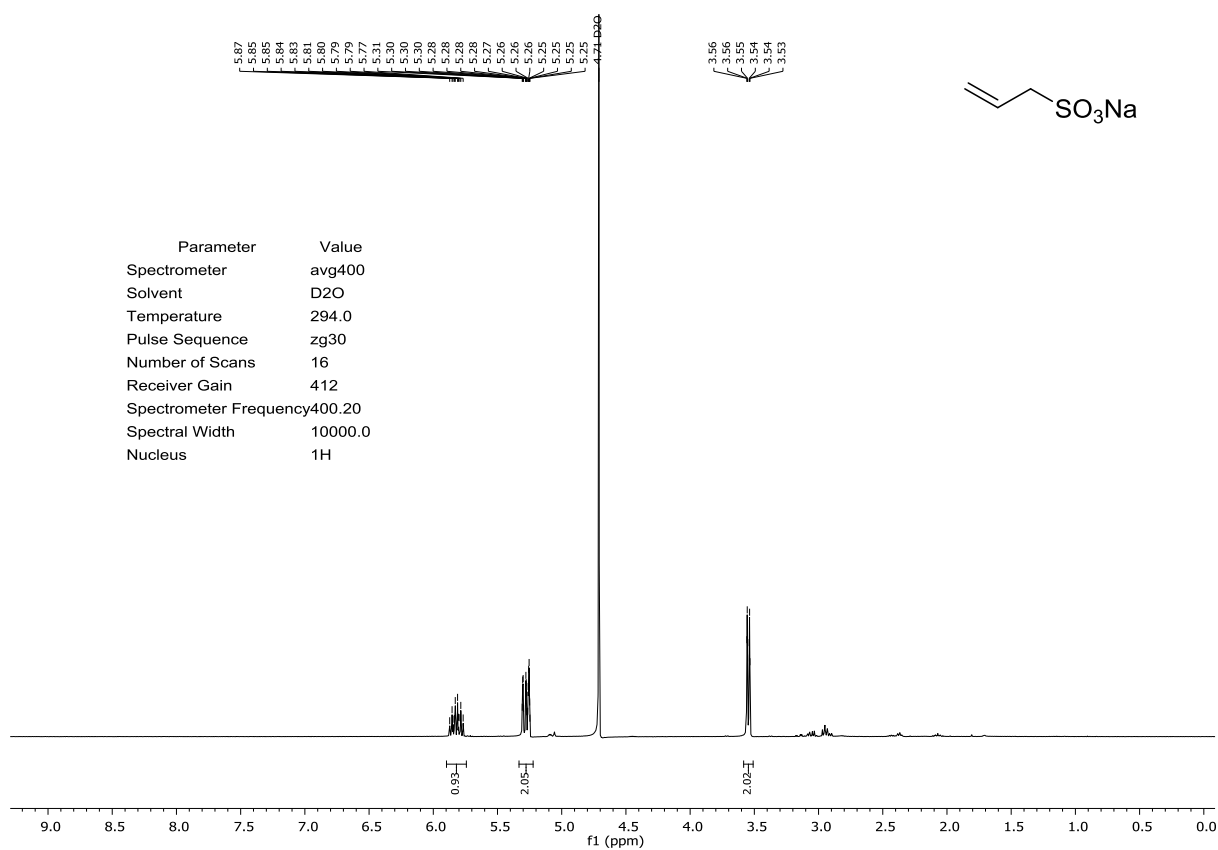


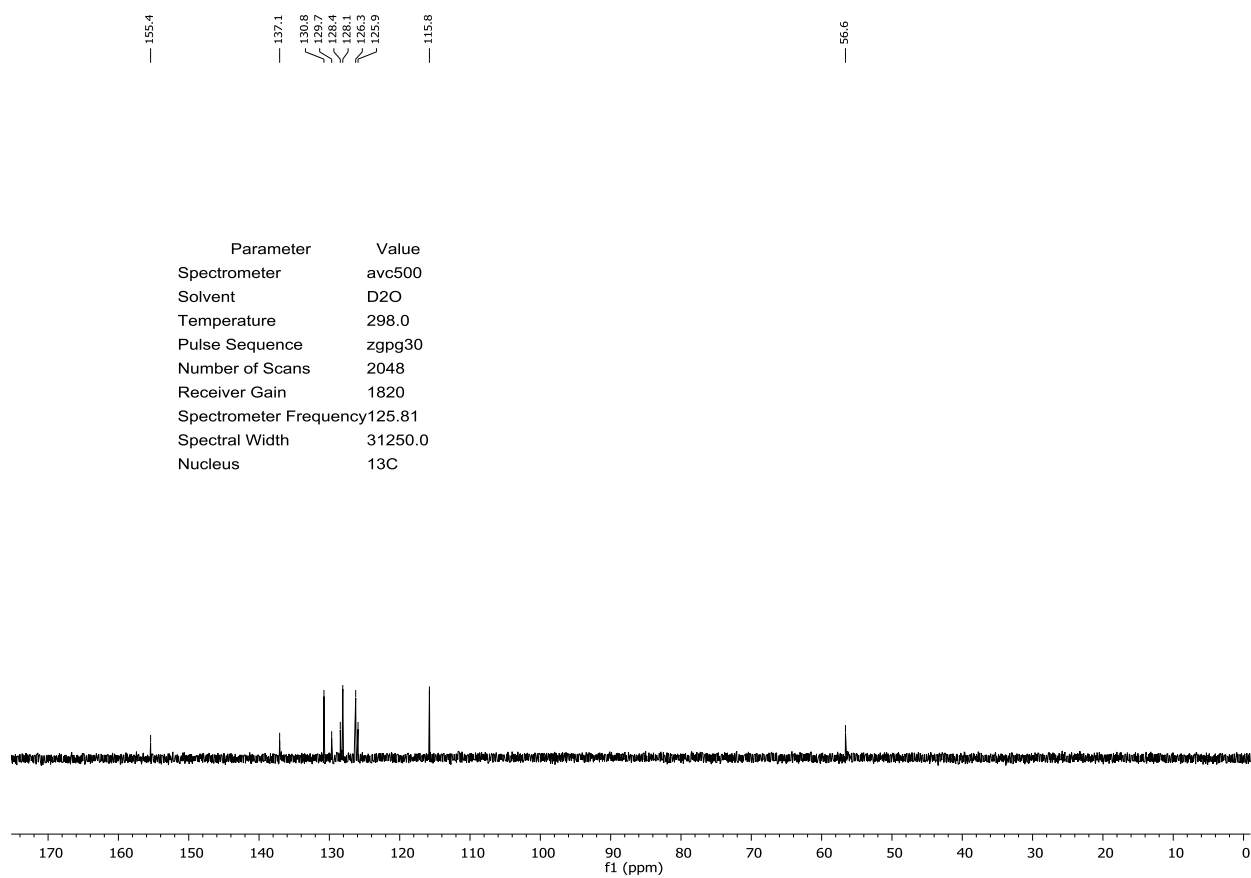
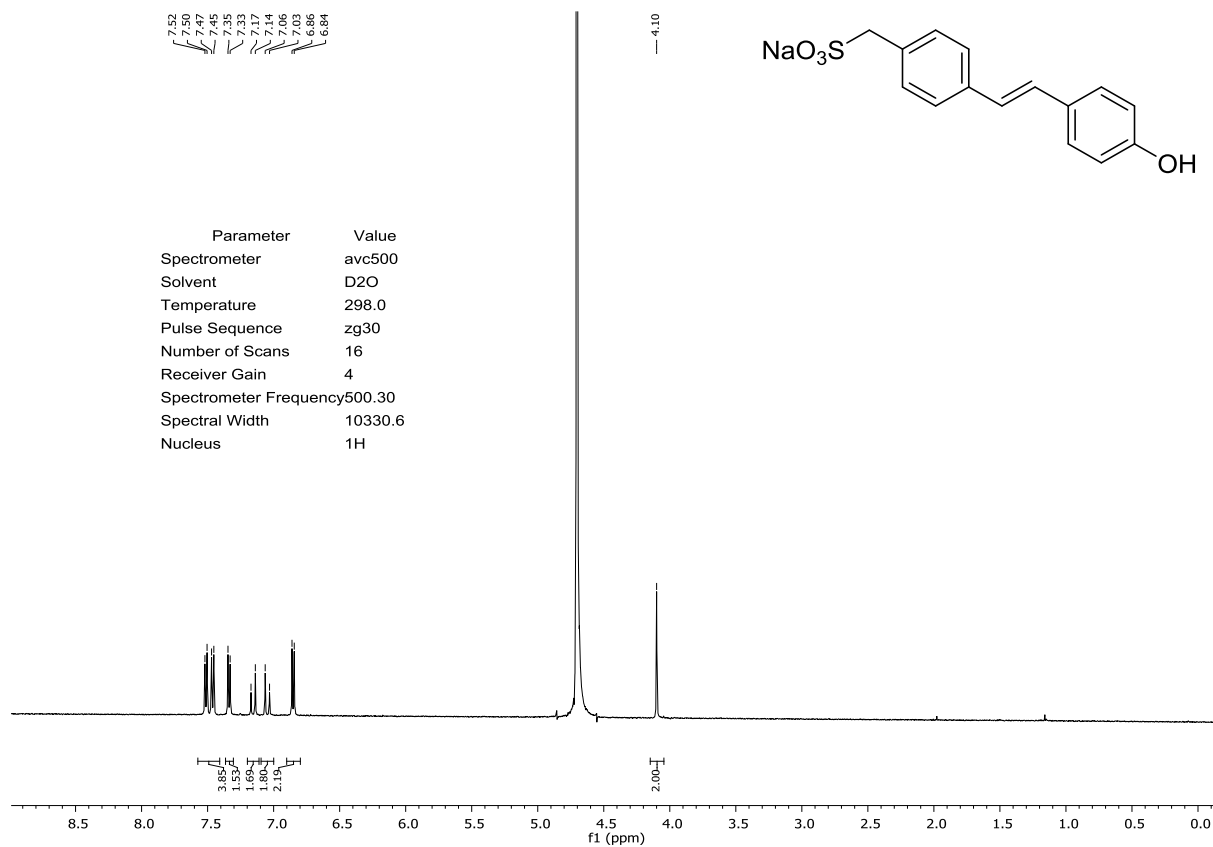
144.68
134.71
131.47
131.23
128.73
127.05
119.27

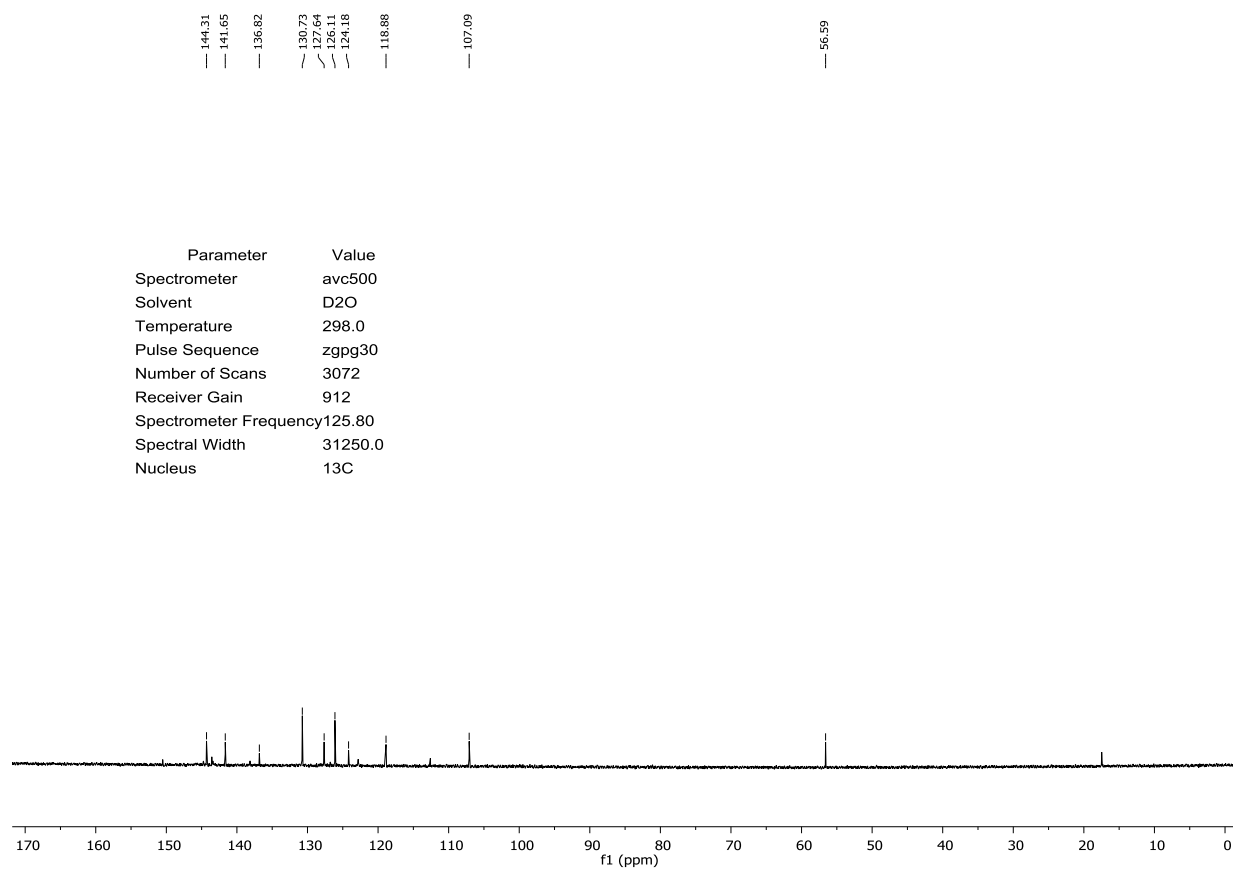
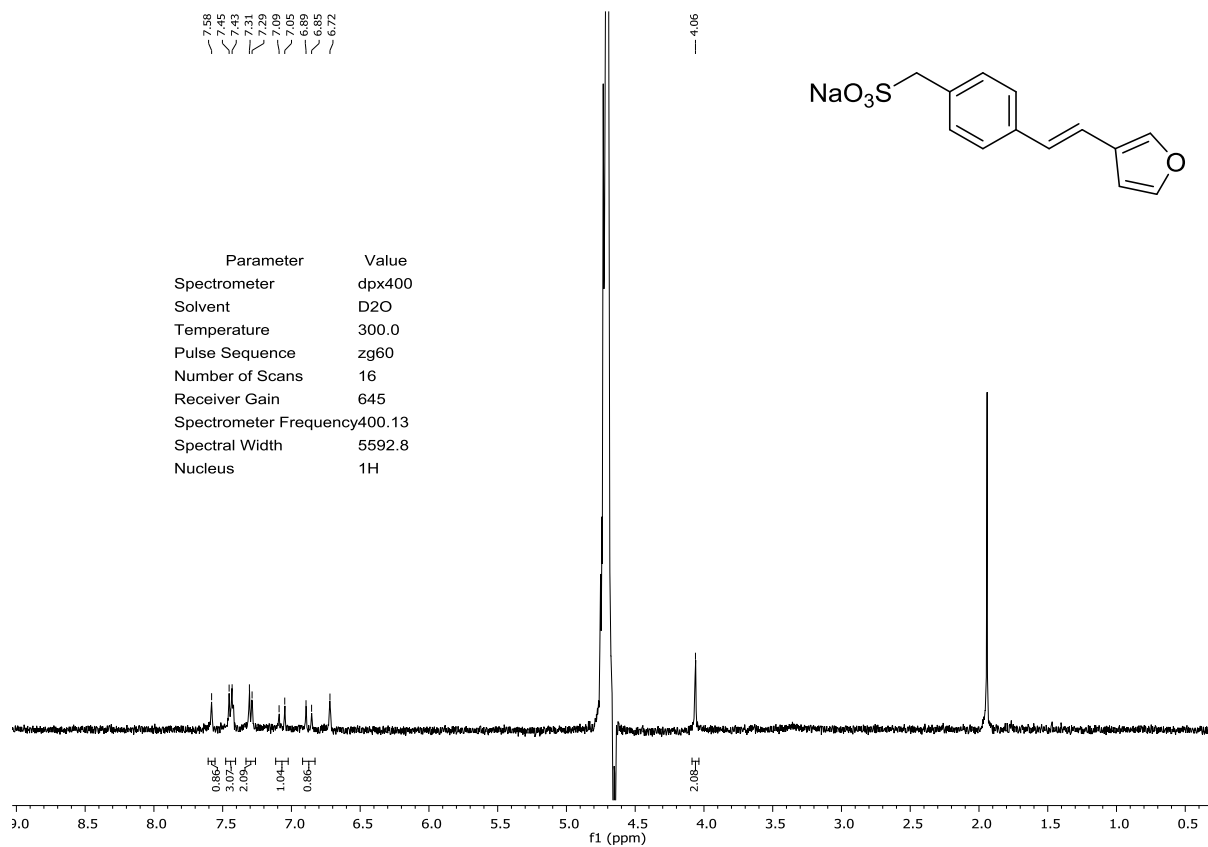
Parameter	Value
Spectrometer	avg400
Solvent	CDCl ₃
Temperature	294.8
Pulse Sequence	zgpg30
Number of Scans	256
Receiver Gain	2050
Spectrometer Frequency	100.63
Spectral Width	24038.5
Nucleus	¹³ C

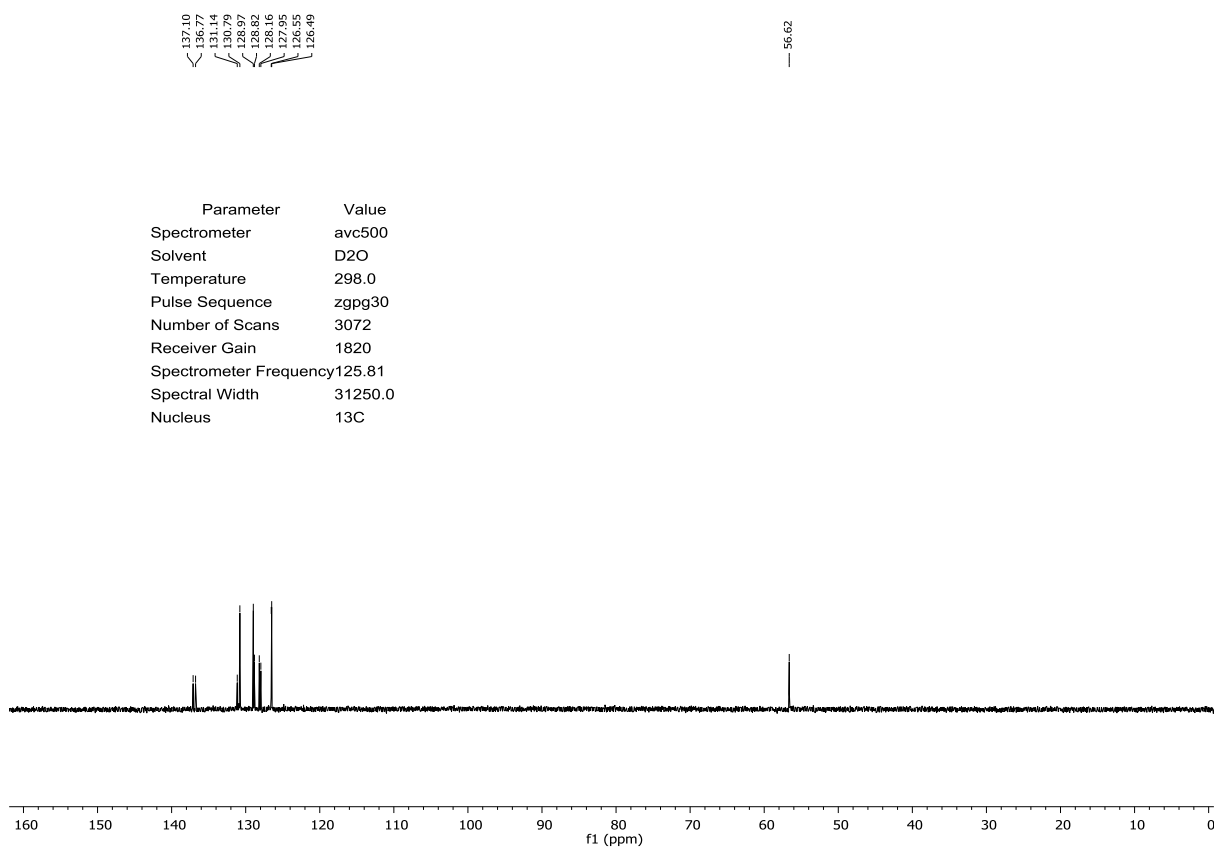
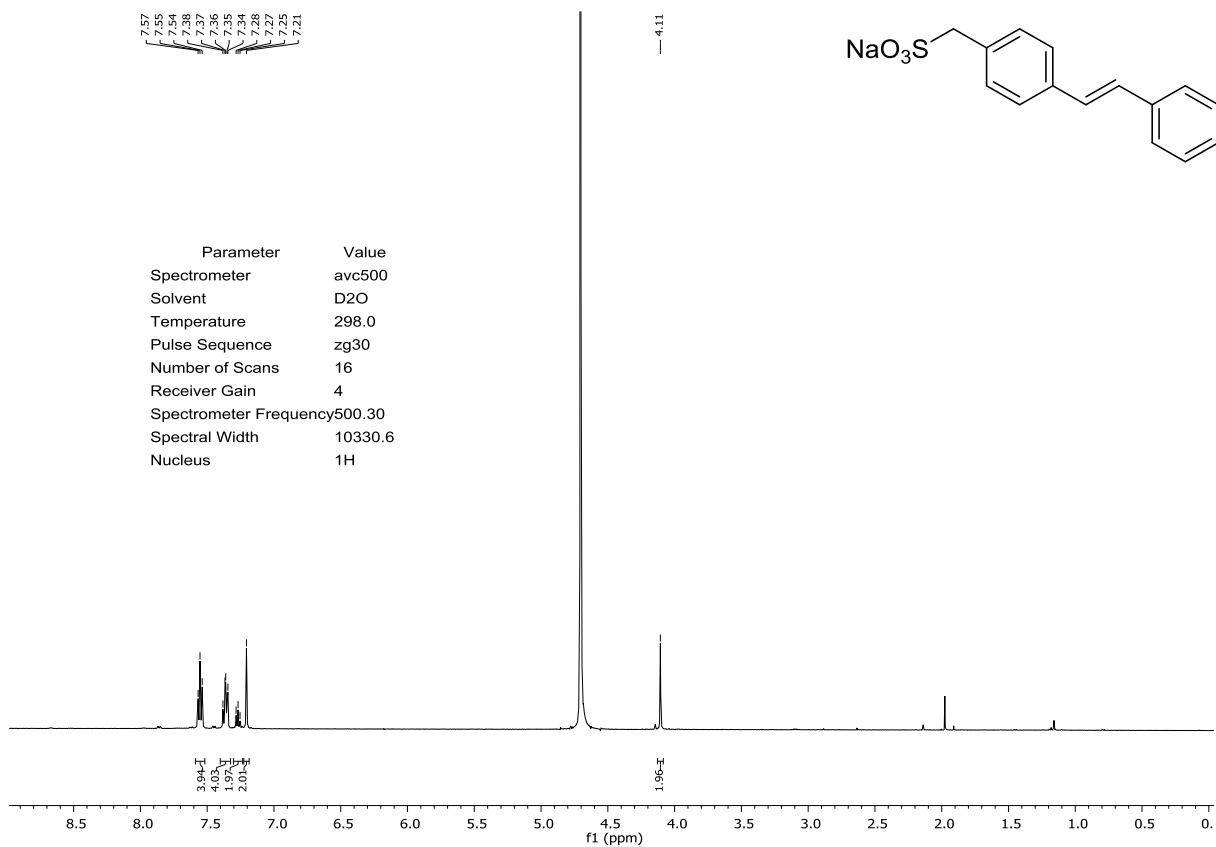


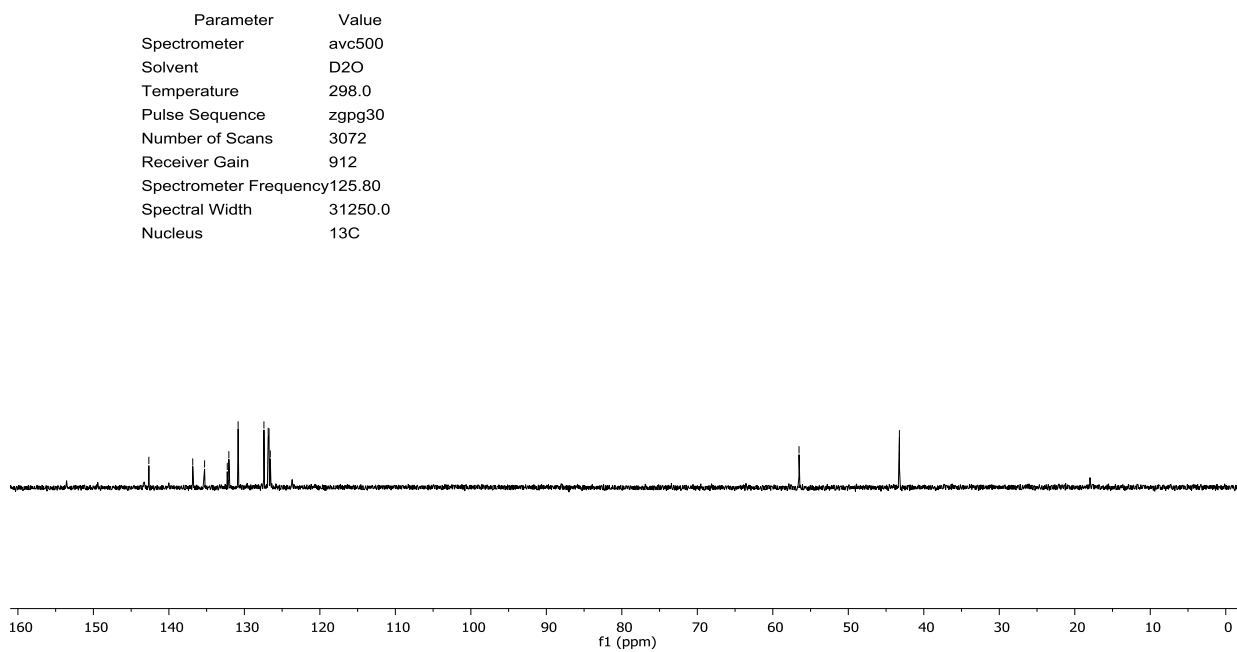
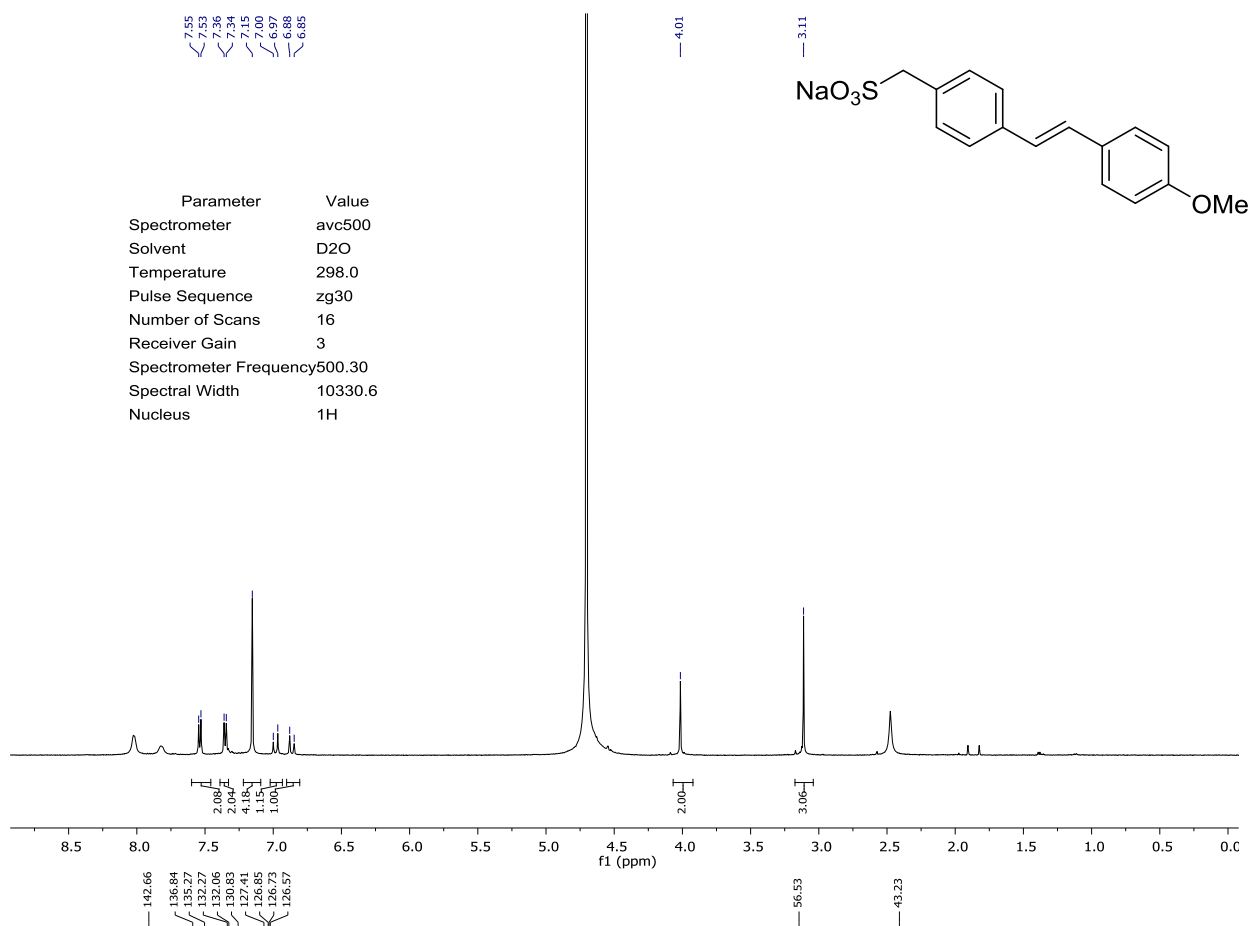


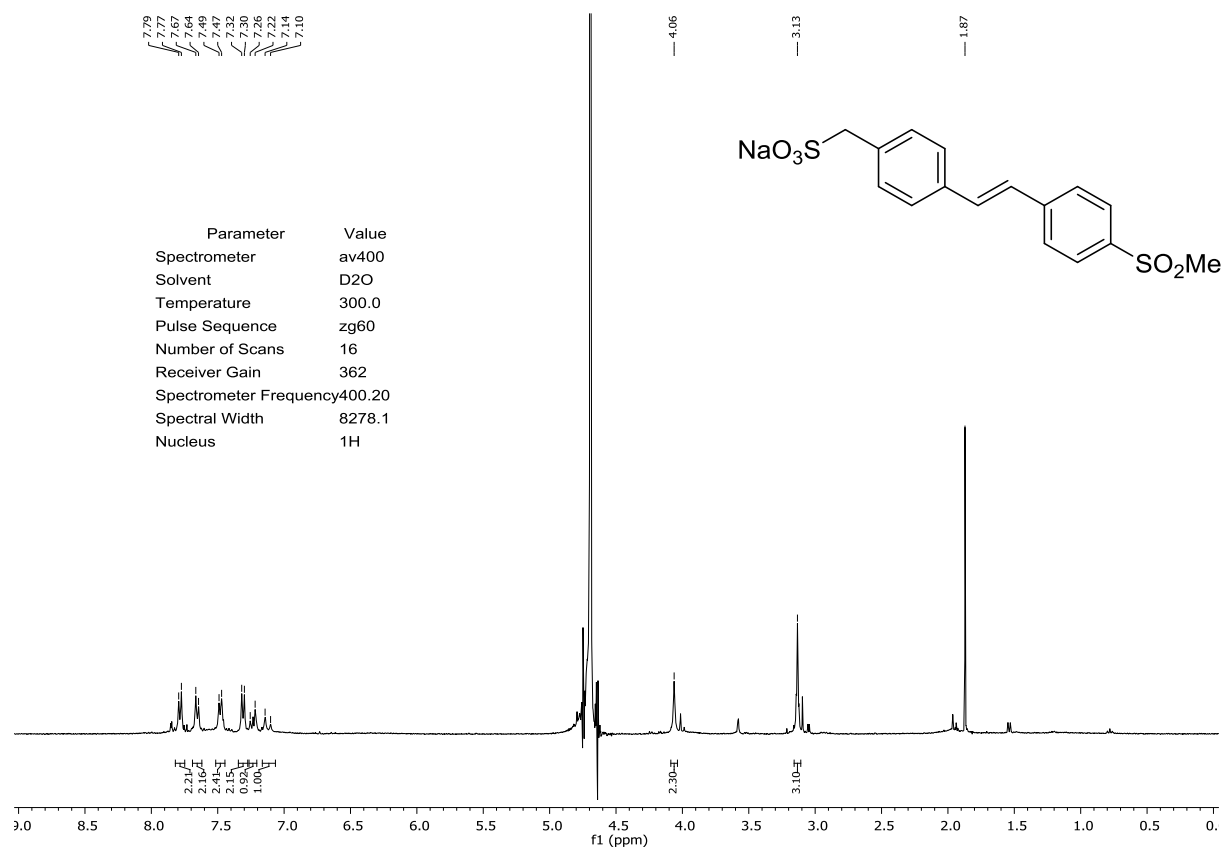












Chapter 4

Structural characterization of SBL-based artificial metalloenzymes

4.1 Introduction

4.1.1 Crystal structures of artificial metalloenzymes

Structural characterization of artificial metalloenzymes by exploiting naturally evolved metal binding sites (by replacing the native metal ions)^{1,2} or by accumulation of metal ions in protein scaffolds to form metal-nanoparticles^{3,4} has been most successful (see Chapter 1 Section 1.3.5). Precise, site-specific installation of metal centres by using unnatural cofactors is a powerful technique to construct artificial metalloenzymes. X-ray crystal structures of artificial metalloenzymes created by using natural prosthetic groups or their analogues known to bind tightly to the hydrophobic enzyme pockets, such as biotin-binding pocket or heme cavity, have also been reported.⁵⁻⁷ Crystal structures of non-covalent insertion of unnatural cofactors has indeed provided invaluable insights into the precise positioning of the metal and substrates in the protein cavity, with implications for the observed activity and selectivity.⁶

However, in some of these structures the partial electron density of the cofactors was an issue either due to disorder⁷ or low occupancy.^{8,9} For example, structural characterization of an Iridium-based artificial transfer hydrogenase was also achieved, however occupancy of the iridium atom was set to 50% perhaps due to partial dissociation of metal complex

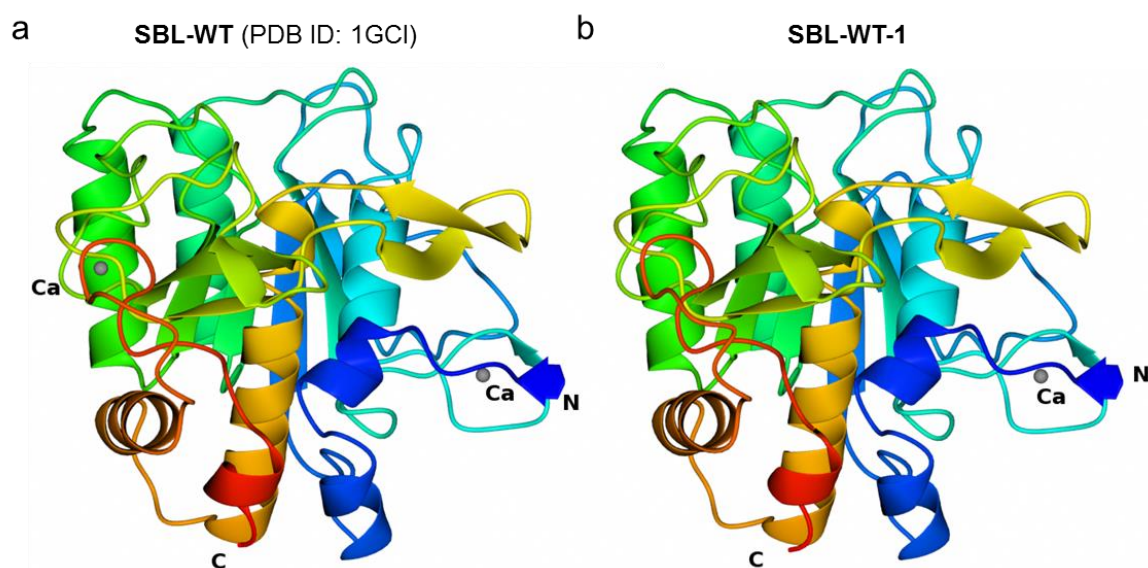
upon soaking the crystals with the metal precursor.⁹ In this regard, dual anchoring of the organometallic ligand based on supramolecular and dative binding, has proved especially successful for structural analysis of artificial metalloenzymes.¹⁰ Notably, there are limited examples of crystal structures of artificial metalloenzymes with covalently bound, unnatural metal cofactors introduced into proteins.^{11,12}

4.1.2 Model protein scaffold: Subtilisin

Subtilisins are globular proteins with highly conserved three dimensional folds comprising seven core β -strands flanked by α -helices on both sides ($\alpha/\beta/\alpha$ tertiary structure). The three representative subtilisins from *Bacillus* species, i.e. BPN' from *Bacillus amylofaciens*, Carlsberg from *Bacillus subtilis* and Savinase from *Bacillus lentus*, have been extensively studied and three-dimensional structures have been determined.¹³⁻²⁰ Despite the huge variations in their respective amino acid sequences, subtilisins are commonly numbered according to BPN' sequence because of this characteristic conserved tertiary structure.²¹ Another conserved structural feature of these enzymes is the Ca^{2+} binding sites.^{22,23}

In this chapter we report the X-ray crystal structures of SBL-based artificial metalloenzymes. In order to structurally characterize the different modifications that were previously installed on SBL (Chapter 2 and Chapter 3), we first crystallised wild-type SBL structure (**SBL-WT-1**). This **SBL-WT-1** and a published **SBL-WT** (PDB ID: 1GCI)¹⁷ were used as models for solving structures (Scheme 4-1). In contrast to the two conserved calcium binding sites reported in literature,^{17,24} only one Ca^{2+} binding site was observed in **SBL-WT-1** structure obtained here (and other SBL structures described in section 4.2.5). It should be noted that in the structural characterization of other reported subtilisin

variants, the protein was co-crystallized with CaCl_2 (10 mM). However, we obtained **SBL-WT-1** crystal without the addition of any extra calcium ions in the crystallization buffer.



Scheme 4-1. Ribbon diagrams of structure of (a) **SBL-WT** (PDB ID: 1GCI)¹⁷ (b) **SBL-WT-1**, the Ca atom(s) and the N and C termini are labelled; SBL-WT (PDB ID:1GCI) has two conserved Ca^{2+} binding sites, whereas we obtained SBL structures with only one Ca^{2+} binding site as shown; the figures were made with ccp4mg²⁵ by Wendy Offen

The active site of subtilisin comprises an Asp32-His64-Ser221 catalytic triad and an oxyanion hole formed by the backbone NH of Ser221 and Asn155 (BPN' numbering).^{18,19} The catalytic residues are flanked by binding pockets on both sides Sn, S2, S1 and S1', Sn' that determine the substrate specificity of these enzymes (See Chapter 2 Section 2.2.1.2). The specificities of these pockets combined with the catalytic machinery are responsible for the remarkable hydrolytic activity of subtilisins (Chapter 1 Section 1.6).²⁶ Mechanistically, Ser221 acts as the nucleophile and His64 is poised to play a dual role in hydrolysis, functioning as a general base catalyst to enhance the nucleophilicity of Ser221, while also acting as a general acid catalyst to expel the amine of the hydrolysed substrate.²⁶⁻²⁸ Asp32 helps to ensure the positioning of the correct tautomeric form of His64 during catalysis and provides electrostatic stabilization.²⁶

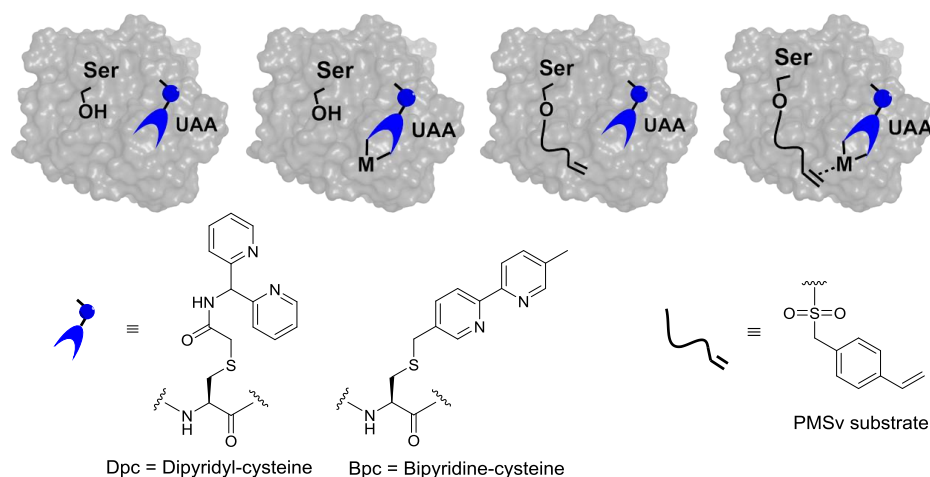
4.2 Results and discussion[†]

4.2.1 Target structures

Subtilisin from *Bacillus lentus* (SBL) was previously employed as a model protein for chemical installation of the metal binding amino acids bipyridine-cysteine (Bpc) and dipyriddy-cysteine (Dpc), that were used for chelating different metals to generate artificial metalloenzymes (as discussed in Chapter 2 Section 2.2.2.4). These tailored enzymes were subsequently used for oxidative Heck cross-coupling reactions between a tethered alkene substrate and various arylboronic acids (See Chapter 3 Section 3.2.6). Having demonstrated that a successful reaction cycle took place in the enzyme active site, we embarked on structural characterization of these metalloenzymes in order to gain insights about the local environment around the installed metal centres.

The key intermediates from the engineered cycle of SBL-based metalloenzymes discussed in Chapter 3 Section 3.2.3.3 were chosen as target structures for crystallization (Scheme 4-2). These intermediates contained different modifications incorporated in the enzyme active site, including Bpc and Dpc that were installed at position 217 in mutant SBL-L217C, turning it into the metal-free apo-enzymes SBL-Bpc217 and SBL-Dpc217. In this way, metals such as palladium (Pd) and nickel (Ni) were coordinated to SBL enzymes leading to the corresponding metalloenzymes. Further, an alkene substrate *p*-vinylphenylmethanesulfonyl fluoride (PMSv) was covalently attached to active site Ser221 to give an enzyme-substrate complex (Scheme 4-2).

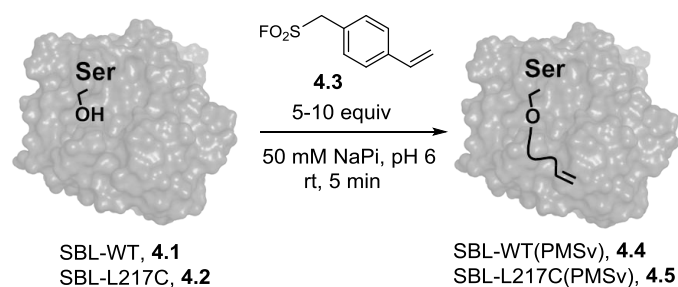
[†] This chapter was carried out in collaboration with Wendy Offen and Prof. Gideon J Davies, York Structural Biology Laboratory, University of York. Screening of crystallization conditions was done with Wendy Offen. Further optimization of crystallization conditions, data processing and refinement was done by Wendy Offen as indicated in the respective section. Experimental results that were obtained by Wendy Offen are indicated in the body of this chapter by reference to this footnote and her contributions are stated explicitly in the experimental sections of this chapter.



Scheme 4-2. Key intermediates from the engineered cycle of the SBL-based artificial metalloenzymes, resulting from different covalent modifications installed in the enzyme active site as depicted; metal binding unnatural amino acids Dpc and Bpc are installed at position 217 and a PMSF-based alkene substrate forms a covalent adduct with the active site residue Ser221

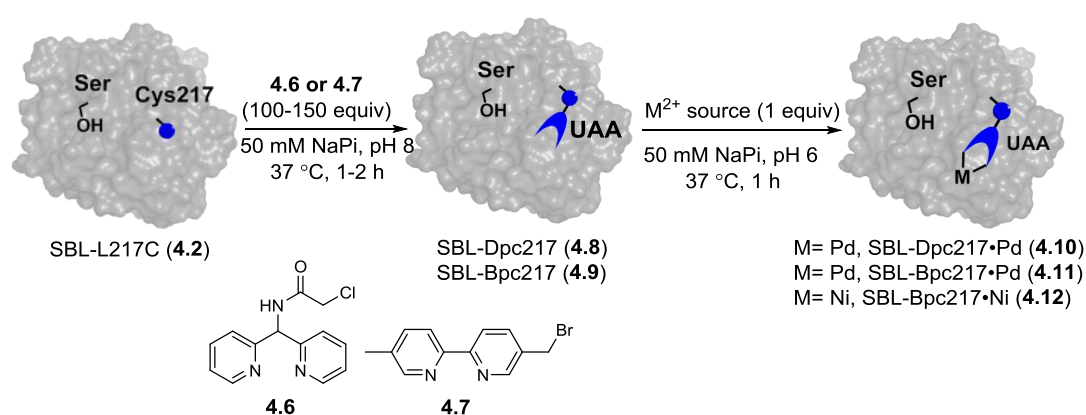
4.2.2 Protein modifications and characterization

SBL-WT (**4.1**) and SBL-L217C (**4.2**) were used as precursor proteins to prepare the chemically modified protein targets for crystallization studies. Initially, SBL-WT and SBL-L217C were used for screening and optimization of the crystallization conditions, which were subsequently applied to protein samples containing the various modifications of interest. To determine the binding interactions of the bound substrate with the protein, SBL-WT and SBL-L217C were treated with 5-10 equiv. of PMSv substrate **4.3** to form modified proteins SBL-WT(PMSv) **4.4** and SBL-L217C(PMSv) **4.5** as shown in Scheme 4-3.



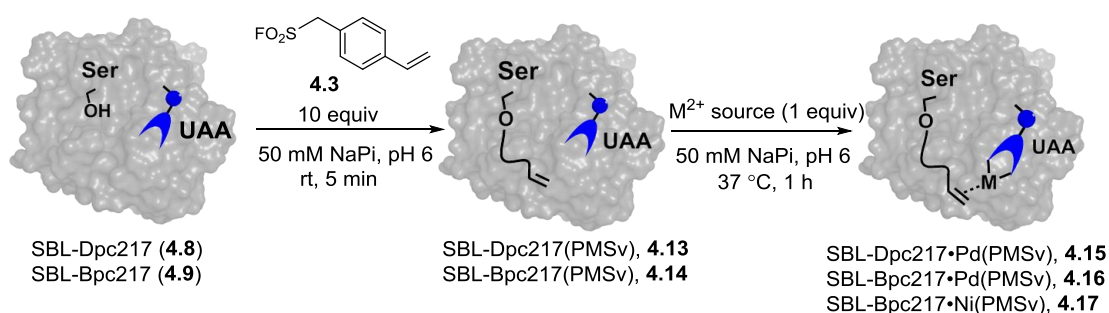
Scheme 4-3. Preparation of substrate bound samples by incubating SBL-WT and SBL-L217C with PMSv substrate **4.3**

Further, to obtain structural evidence for the successful incorporation of the unnatural metal binding amino acids and the metal centres, samples corresponding to the apo-proteins and metal-containing proteins were prepared. In order to undertake the screening of crystallization conditions, a scale-up of purification and protein modification reactions was required (See Experimental Section 4.5.2-4.2.5). Direct alkylation of SBL-L217C (**4.2**) with the dipyridyl and bipyridine ligands (**4.6** and **4.7**) successfully yielded the metal-free apo-enzymes SBL-Dpc217 (**4.8**) and SBL-Bpc217 (**4.9**) as shown in Scheme 4-4. We have previously shown that metal centres (such as palladium and nickel) can be successfully installed using these chelating amino acids, Dpc and Bpc (See Section 2.2.2.4). Previously, it was shown that SBL-Bpc217 could coordinate nickel, while complexation of nickel with SBL-Dpc217 did not show the desired metal adduct by LC-MS analysis (See Chapter 2 Section 2.2.2.4). In particular, selective and stoichiometric interaction of the metal ions with Bpc confirmed its high affinity for divalent transition metals. Thus in a similar way, we achieved metal complexation with M^{2+} precursors to form holo-enzymes SBL-Dpc217•Pd (**4.10**), SBL-Bpc217•Pd (**4.11**) and SBL-Bpc217•Ni (**4.12**), which were used in crystallization experiments.



Scheme 4-4. Preparation of apo- and holo-forms of SBL-based metalloenzymes

Additionally, crystallization samples corresponding to the intermediates from the engineered cycle were also synthesized. To characterize the substrate binding interactions, apo-enzymes **4.8** and **4.9** were loaded with alkene substrate **4.3** to form species SBL-Dpc217(PMSv) **4.13** and SBL-Bpc217(PMSv) **4.14**, which subsequently were converted into the metal bound species SBL-Dpc217•Pd(PMSv) **4.15** and SBL-Bpc217•Pd(PMSv) **4.16** upon palladium complexation. The synthetic route followed for the preparation of these enzymes is shown in Scheme 4-5. Because of its versatility in catalysing a range of reactions including C-C cross-coupling reactions, we reasoned nickel could be an alternative to palladium. Hence, nickel was used as a surrogate metal for palladium and synthetic enzymes SBL-Bpc217•Ni **4.12** and SBL-Bpc217•Ni(PMSv) **4.17** were also screened in crystallization trials.



Scheme 4-5. Preparation of crystallization samples corresponding to cycle intermediates; substrate **4.3** forms a covalent adduct with Ser221, followed by metal complexation by installed UAAs. All of the above synthesized proteins were used for screening of crystallization conditions after adequate protein purification and sample conditioning. In general, chemical modification of precursor proteins (SBL-WT and mutant SBL-L217C) was carried out as described above, following which all protein samples were concentrated (typically to >12 mg/mL). Crystallization samples were purified twice by size exclusion chromatography prior to screening. All the modifications and the identity of the protein in each sample were successfully confirmed by LC-MS analysis (See Experimental Section 4.5.2-4.5.4 for details). No presence of any other contaminant was detected in the same molecular mass

range. The purity and homogeneity was confirmed by non-reducing PAGE gel and the protein content of each sample was determined using the bicinchoninic acid (BCA) assay, according to a standard procedure (See Experimental Section 4.5.5).

4.2.3 Initial Screening

Initial crystallization screening[†] was carried out using synthetic proteins SBL-Dpc217 and SBL-Bpc217. First, these samples were incubated with a 3 molar excess of Pd²⁺ precursor (Na₂PdCl₄) for 1 h at 37 °C prior to screening, and alkene **4.3** or PMSF were added as substrates upon plate set-up. The first round of screening led us to some promising conditions and a few crystals were obtained. Data were collected from a single crystal of **SBL-Dpc217(Pd) with PMSF** (added on set-up) grown using a commercial screen for crystallization trials based on polyethylene glycol (PEG)/ion conditions (PACT *premier*TM screen, Molecular Dimensions). The crystal grew from a solution containing 0.2 M potassium thiocyanate (KSCN), 0.1 M Bis-Tris propane pH 8.5, 20% PEG 3350, and was in space group C2, with unit cell dimensions 40.1, 74.2, 73.9 Å and angles 90, 95.1, 90 degrees. Data were collected to 1.35 Å, at Diamond beamline io2. The structure was solved by molecular replacement using Phaser, with SBL-WT PDB ID: 1GCI as the model. There was 1 molecule in the asymmetric unit, and the Matthew's coefficient was 2.1. There was difference density for 2 or 3 atoms adjacent to the Cys217 sulphur atom. An anomalous map showed peaks for 2 calcium atoms and the Met S atoms (but the Cys217 sulfur peak was not observable). Surprisingly, an anomalous signal was also seen for an atom occupying density 2.20 Å from Cys217 sulfur in the difference map. Unfortunately, the Dpc modification was not visible in the Fo-Fc map, and therefore the structure was not refined further.

[†] Initial screening experiments were done with Wendy Offen

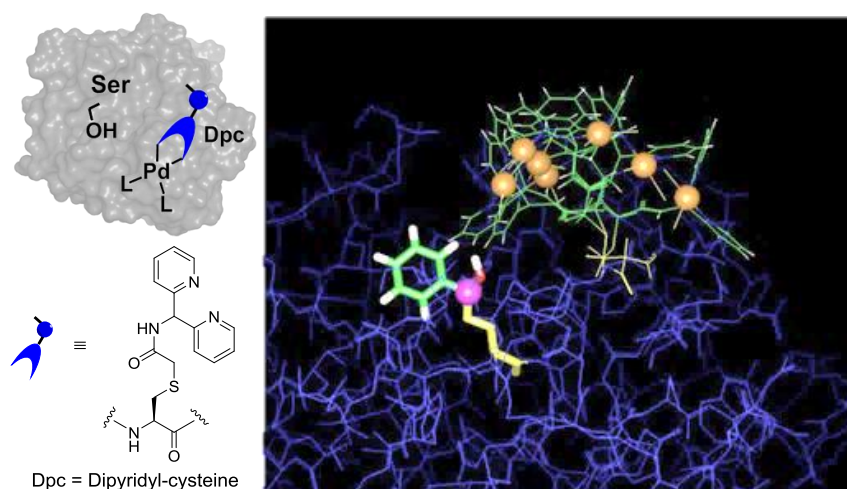
These observations were unexpected as LC-MS analysis of the protein solution had confirmed the identity of the protein when a mass corresponding to the **SBL-Dpc217•Pd(PMS)** species was observed. Moreover, an X-ray fluorescence (XRF) scanning of the crystal could not detect the presence of palladium in the crystal since the experimental range used at the synchrotron could not cover the wavelengths required for the detection of palladium. While the first round of screening did not give us the desired results, two key issues were identified which were borne in mind in future crystallization trials.

(a) Poor ligand density due to inherent flexibility of the modifications

In the protein crystal containing Cys217 modified with dipyriddy ligand **4.6** to give SBL-Dpc217, only poor ligand density was detected around cysteine (crystals with ligand Dpc were obtained with a resolution of 1.35 Å, but there was no complete or even partial ligand difference density visible in Fo-Fc maps contoured at 3.0 rmsd, or when the contour level was decreased to 2.0 rmsd). The b-factors of the probable neighbouring residues Asn62, His64, Tyr209, and Met222 were all less than 10 Å², which indicated that the pocket which the modification would be expected to occupy is relatively ordered in this structure. It must be noted that this unnatural amino acid Dpc was installed at the base of the S1' subsite of SBL in our design (Section 2.2.1.2). We speculated that incomplete electron density of this modification may be due to its inherent flexibility such that several conformers of this ligand may be possible. In this context, preliminary modelling studies[†] suggested that multiple sterically allowed orientations of the Dpc•Pd complex are possible in the absence

[†] Modelling studies were done by Dr. Mark Wormald

of any other additional interactions of the 3rd or 4th coordination sites of the metal with other neighbouring protein residues (Scheme 4-6).



Scheme 4-6. Model[†] showing several possible orientations of Dpc installed in S1' site of SBL-Dpc217•Pd; the protein (in blue) contains a molecule of phenylboronic acid attached to Ser221 (in green with boron atom in pink) and Dpc•Pd complex at position 217 (green with metal depicted as orange sphere); modified side chains of Ser221 (in S1 pocket) and Cys217 (in S1' pocket) are shown in yellow; model was built using PDB ID: 1GCI¹⁷

(b) Palladium scavenging

Previously, LC-MS analysis confirmed that thio reagents, such as 3-mercaptopropionic acid (MPA) and thiol acetate, are potent metal scavengers (See Section 3.2.3.2). With as little as the 3 equiv of palladium source used during crystallization, it is highly likely that the metal may be removed from the protein by such compounds present in the crystallization solution. It was thought in retrospect that potassium thiocyanate (KSCN) or PEG used in our crystallization conditions could scavenge palladium from the protein, and therefore this crystallization condition was unlikely to yield the desired metal-containing protein complex. In this regard, we have previously found that an excess of any good metal coordinating reagents (even vinyl groups) could remove palladium from the modified protein (See Section 3.2.5).

4.2.4 Optimization of crystallization conditions

Three promising conditions from the first round of screening were followed up in subsequent rounds:

- 1.8-3.2 M NaCl, 0.1 M NaOAc buffer, pH 4.6
- 13% PEG 3350, 0.1 M ammonium sulfate (AS), 0.1 M Bis-Tris buffer, pH 5.5
- 0.2 M lithium chloride (LiCl), 20% PEG 6K, 0.1 M NaOAc buffer, pH 5

Different strategies adopted for optimization of crystallization conditions are discussed below.

(a) Co-crystallization of different modifications

In initial optimization experiments, only SBL-Dpc217 and SBL-Bpc217 were used and were incubated with metal just prior to plate set-up for crystallization screening. However, considerable protein precipitation was seen with samples that were pre-incubated in this manner. Hence, in future trials the ligands, metals (nickel or palladium) and substrate **4.3** were co-crystallized (where the substrate binding and metal attachment was confirmed prior to crystallization trials). This was found to be beneficial not only for the stability of the protein but also to ensure solubility of the ligands and the substrates. Advantageously, the stoichiometry of the metal:protein composites could also be precisely verified by LC-MS analysis. Further rounds of crystallization were performed with the aim of characterizing the key intermediates of the engineered cycle. These intermediates contain combinations of three possible modifications on the protein (the chelating amino acids Dpc or Bpc, alkene substrate PMSv and metal). The preparation of 17 such variants containing the target modification has been described in Section 4.2.2.

(b) Screening different Pd(II) sources and pH conditions

In the first round of screening, Na_2PdCl_4 was used as the Pd^{2+} source. Considering NO_3^- is a weaker coordinating ion than Cl^- ,^{29,30} we reasoned that the coordination of metal due to faster release of soluble palladium species under crystallization conditions may be possible. Thus, the use of a different palladium salt ($\text{Pd}(\text{NO}_3)_2 \cdot x\text{H}_2\text{O}$) was also explored. The crucial aspect of the stability of different modifications under crystallization conditions was also taken into account. We have shown that PMSv substrate can be hydrolysed from the protein under acidic conditions (See Chapter 3 Section 3.2.3.1). Thus in addition to crystallization screening using ammonium sulphate/PEG condition in Bis-Tris at pH 5.5 and the NaCl conditions in NaOAc at pH 4.6, protein crystallization was also attempted at higher pH (eg. at pH 7.5-8.5) to avoid any possible hydrolysis of the bound PMSv substrate.

(c) X-Ray damage

The possible instability of the palladium metal centre on exposure to X-ray during data collection was also speculated. X-ray induced radiation damage to macromolecular structures can be mediated by hydroxyl radicals that are formed as a result of irradiation of the crystals.^{31,32} Proteins containing redox-active metal species are susceptible to radiation damage causing reduction of metal centres.³³ Yano et al. reported similar reduction of high-valent Mn_4Ca cluster to $\text{Mn}(\text{II})$ similar to that found in aqueous solution.³⁴ In the case of X-ray damage, a strategy to collect a minimal amount of data or using short wavelengths could be envisaged. Moreover, from our initial round of screening it was realized that the detection of palladium in protein crystals grown from the corresponding samples by X-ray fluorescence (XRF) was not possible. This is because the experimental range for XRF scans does not cover the wavelengths necessary for detecting palladium.

Hence, crystallization of both palladium- and nickel-containing protein samples was attempted in further rounds of screening.

(d) Soaking experiments

Soaking small molecules into protein crystals provides a simple and straightforward method to visualize modifications as an alternative to co-crystallization techniques, provided that the ligands have high affinity for the protein binding sites. Therefore, taking into account the relatively shallow active site of SBL, soaking of the crystals with metal, alkene and arylboronic acid substrates was also performed with different soak times.

Overall, the crystallization of different SBL-variants was accomplished in six rounds of screening as summarized below.

I round: Broad screening of crystallization conditions using different commercially available and optimized screens (See Appendix 4-1).

II round: SBL-WT(PMSv) and SBL-L217C(PMSv) samples were crystallized using two optimized screens based on successful conditions from the first round: 1.8-3.2 M NaCl, 0.1 M NaOAc buffer pH 4.6; 0.2 M LiCl, 20% PEG 6K, 0.1 M NaOAc buffer pH 5 (See Appendix 4-2).

III round: Scavenger free conditions were optimized (using SBL-WT sample as positive control). Other variables, such as different additives and pH conditions were also investigated. One promising and reproducible crystallization condition was identified using seeding to induce crystal growth: 13% PEG 3350, 0.1 M AS, 0.1 M Bis-Tris buffer, pH 5.5 (See Appendix 4-2). Crystals from this round were used to obtain a seed stock for

further crystallization trials.

IV round: Crystals obtained for different modifications optimizing around two conditions (13% PEG 3350, 0.1 M ammonium sulfate (AS), 0.1 M Bis-Tris buffer pH 5.5; 0.2 M LiCl, 20% PEG 6K, 0.1 M NaOAc buffer pH 5). A range of pH conditions was also screened. Flexibility of Dpc modification and possible palladium scavenging by PEG were interfering factors (See Appendix 4-3).

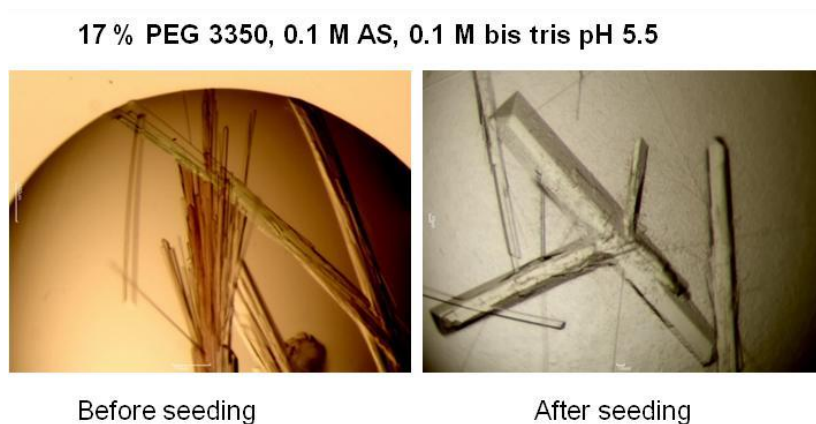
V round[†]: Crystals grown using 1.8-3.2 M NaCl, 0.1 M NaOAc buffer pH 4.6 with SBL-Bpc217 samples with different modifications (palladium and PMSv substrate). Soaking experiments were performed using alkene and ArB(OH)₂ substrates (See Appendix 4-4).

VI round[†]: Crystals grown using 0.1 M AS, 0.1 M Bis-Tris buffer pH 7.5-8.5 with SBL-Bpc217 samples and nickel as a surrogate for palladium. Crystal growth was slower at higher pHs. A seeding stock from thin long crystals (grown at pH 8.5) was used to facilitate crystal growth. A range of pHs was also screened in order to optimize crystallization conditions at higher pH (See Appendix 4-5).

4.2.5 Structure descriptions

The structural determination of relative orientations of the installed UAAs (Bpc or Dpc) and the PMSv substrate **4.3**, and the precise location of the metal in the enzyme active site would provide insights into the specific binding interactions and inform future catalyst design. To this end, the three major modifications (ligand, alkenyl substrate and metal) engineered into SBL as part of our design were structurally characterized using crystals grown in the different rounds of screening described above (See Section 4.2.4). A seeding

stock was used to facilitate crystal growth and SBL crystals were grown from two conditions in different space groups. SBL crystals typically diffracted with a resolution of 1.0-1.6 Å. The space group of PEG-grown SBL crystals was $P2_12_12_1$, with cell dimensions of $a = 53$ Å, $b = 61$ Å, $c = 75$ Å, $\alpha = \beta = \gamma = 90^\circ$ while the space group of SBL crystals grown in NaCl conditions were $P2_1$ with cell dimensions of $a = 40$ Å, $b = 63$ Å, $c = 42$ Å; $\alpha = 90^\circ$, $\beta = 118^\circ$, $\gamma = 90^\circ$.



Scheme 4-7. Images of SBL crystals grown in PEG conditions before seeding (left) and after seeding (right)

Four crystal structures were refined and key findings from these crystal structures have been described in this section. Data collection and refinement statistics for these SBL structures are summarized in Table 4-1. Some common features were observed in several of our SBL structures. Only one Ca^{2+} binding site was observed in our SBL structures (as discussed in section 4.1.2). There was unmodelled density observed between the side-chains of Asn155, Asn218, and Phe189. There was a sulphate ion in all four structures, modelled at an occupancy of 0.75 to 0.85, hydrogen bonded to Thr134. There are 1-3 molecules glycerol (all modelled at 0.5, or 0.4). In addition, in the NaCl-grown structures (although unrefined) density for typically 1 Ca^{2+} , 2 Cl^- and several Na^+ ions were observed in the maps.

Table 4-1. Data collection and refinement statistics for SBL structures[†]

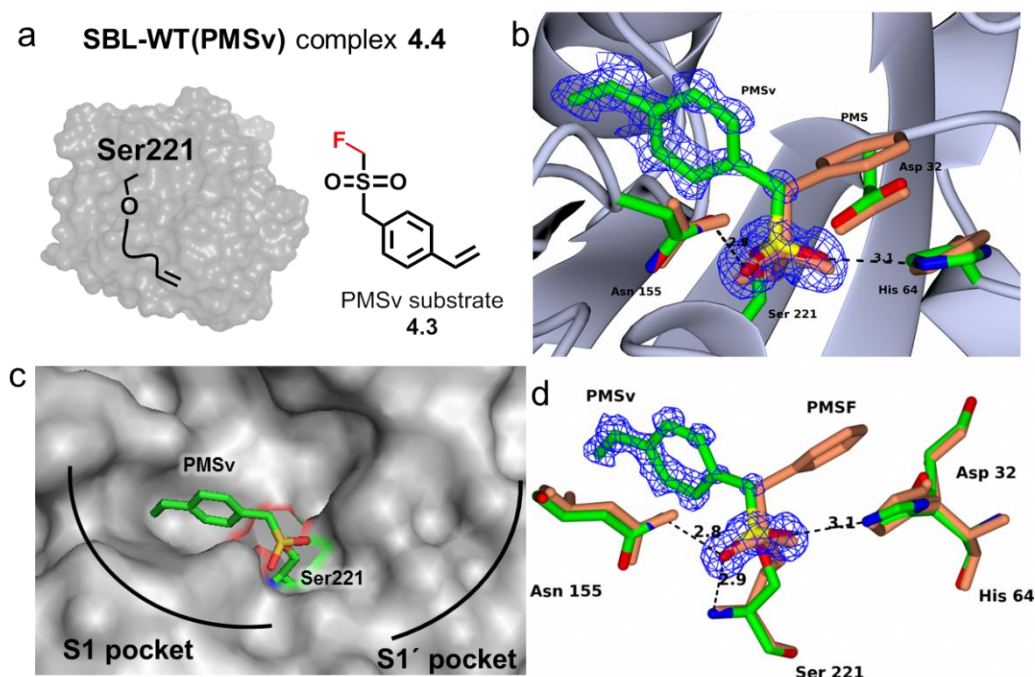
	SBL- WT(PMSv)	SBL-Bpc217 (pH 5.5)	SBL-Bpc217 (pH 8.5)	SBL-Bpc217•Ni
Data collection				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	52.6,61.2,74.6	52.6,61.1,74.8	52.4,61.0,74.7	52.0,61.3,74.7
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	43.02-1.10	47.33-1.15	47.26-1.10	47.37-1.55
<i>R</i> _{sym} or <i>R</i> _{merge}	0.036 (0.059)	0.084 (0.356)	0.061 (0.543)	0.081 (0.498)
<i>I</i> / σ <i>I</i>	33.3 (17.9)	11.6 (2.4)	16.2 (2.4)	13.4 (3.3)
Completeness (%)	98.0 (91.7)	97.0 (72.8)	99.9 (98.7)	99.9 (100.0)
Redundancy	6.3 (4.5)	5.8 (2.7)	6.1 (4.1)	5.8 (5.5)
Refinement				
Resolution (Å)	1.10	1.15	1.10	1.55
No. reflections	96307	83629	97753	35362
<i>R</i> _{work} / <i>R</i> _{free}	0.080/0.093	0.104/0.121	0.109/0.123	0.134/0.154
No. atoms				
Protein	2057	1996	2064	1950
Ligand/ion	37	39	26	33
Water	475	360	319	230
<i>B</i> -factors (Å ²)				
Protein	5	8	9	13
Ligand/ion	10	19	18	20
Water	15	21	22	26
R.m.s deviations				
Bond lengths (Å)	0.013	0.011	0.016	0.019
Bond angles (°)	1.7	1.6	1.7	1.9
Ramachandran plot				
Favoured	98 %	98 %	98 %	97 %
Allowed	2 %	2 %	2 %	3 %
Outliers	0 %	0 %	0 %	0 %
PDB code	5AQE	5ARB	5ARC	5ARD

4.2.5.1 SBL-WT(PMSv) crystal structure

To elucidate the contributing enzyme-substrate interactions forming the key features of our design, **SBL-WT(PMSv)** complex (**4.4**) was crystallized by the sitting drop method using ammonium sulphate as precipitant. The high-resolution crystal structure of **SBL-WT(PMSv)** was refined to 1.1 Å resolution (Scheme 4-8; PDB ID: 5AQE). A molecule of alkene **4.3** was modelled into density in Fo-Fc and 2Fo-Fc maps calculated using Refmac, with phases from a model consisting of chain A of 1GCI.pdb. Alkene **4.3** was found

covalently bound to O γ Ser221 (in the active site of SBL) and is modelled with an occupancy of 0.5, while the sulfonyl group (-SO₂) is modelled at a higher occupancy of 0.75 (Scheme 4-8a and b). As the model comprised only the protein chain, map bias was avoided, allowing unprejudiced judgement of the ligand position. The X-ray data and refinement statistics are summarized in Table 4-1.

In the **SBL-WT(PMSv)** structure reported here, the substrate PMSv occupies a similar position when compared with the partially bound PMS group in another reported structure of a mutant SBL-S87N (PDB ID: 1IAV).²¹ N δ 1 His64 in the latter is hydrogen bonded to the carboxyl group of catalytic triad residue Asp32, and N ϵ 2 is 3.3 Å from the side chain O of Ser221 and 2.8 Å from a sulfonyl O atom. However, His64 in **SBL-WT(PMSv)** mainly exhibits a different conformation and has been modelled in two different conformations of occupancies of 0.8 and 0.2 (Scheme 4-8b). For the predominant conformation, N ϵ 2 His64 is within hydrogen bonding distance of one of the sulfonyl O atoms (3.1 Å) of the PMSv substrate, and N δ 2 is hydrogen bonded to O δ 1 Asn62 (3.1 Å). Importantly, for the substrate bound **SBL-WT(PMSv)** structure, the other sulfonyl O atom is hydrogen bonded to NH Ser221 (2.9 Å) and N δ 2 Asn155 (2.9 Å), in the ‘oxyanion hole’ site of SBL.



Scheme 4-8. (a) Graphical representation of the enzyme-substrate complex **SBL-WT(PMSv)** showing the structure of alkene **4.3** attached to Ser221, (b) Close view of the active site structure of substrate bound **SBL-WT(PMSv)**; PMSv and the residues in the catalytic triad (Ser221, His 64 and Asp32) and oxyanion site (Asn155) are shown as cylinders, with C atoms in green, N in blue, O in red and S in yellow, (c) Surface view of **SBL-WT(PMSv)** structure depicting PMSv substrate localised in the hydrophobic S1 pocket of SBL, (d) Overlay of **SBL-WT(PMSv)** structure (with residues coloured as in (b)) on subtilisin *Bacillus lentus* variant complexed with PMS (PDB ID: 1C9N; with atoms in coral) showing omit map density for PMSv substrate (modelled at occupancy 0.5) contoured at 0.8 rmsd; hydrogen bonds are represented as dashed lines with distances given in Å; (b,d) the figures were made with ccp4mg²⁵ by Wendy Offen

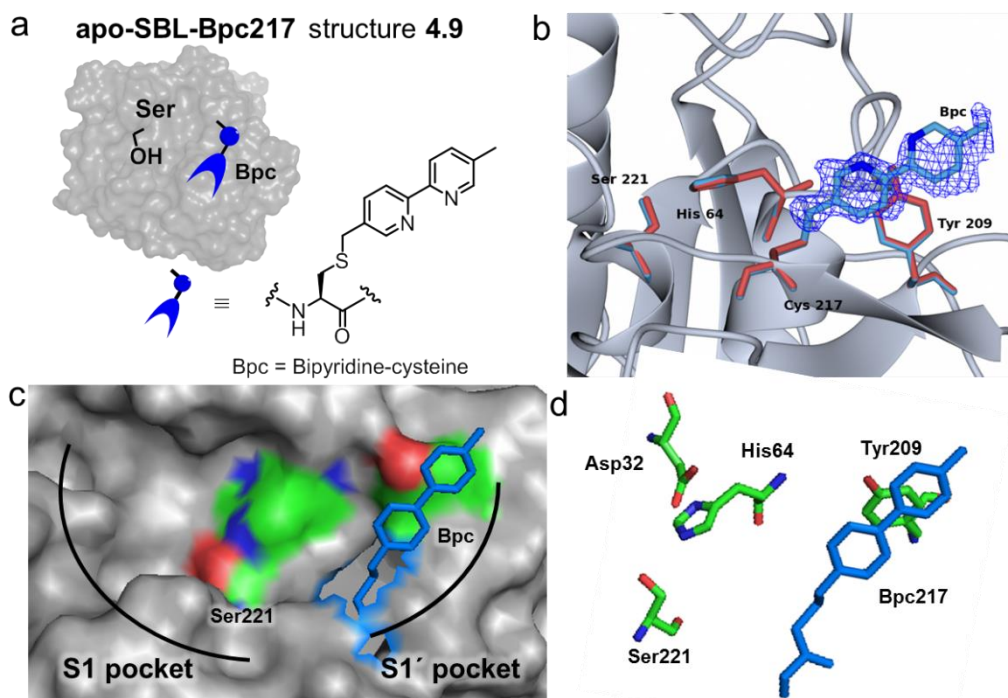
Interestingly, the vinylphenylmethyl moiety of the bound PMSv substrate points towards Ser166 and sits in the hydrophobic S1 pocket lined by residues 125-127 (Scheme 4-8c). Similar to the structure reported here, there is previous evidence of the phenylmethyl moiety of PMSF or side-chain phenylalanine of a chloromethylketone inhibitor fitting in this hydrophobic cleft in other SBL crystal structures.^{21,35} The crystal structure here further confirms our view of binding determinants such as this characteristic hydrophobicity of S1 pocket that gives rise to the substrate selectivity for the engineered SBL enzymes (as discussed in Chapter 3 Section 3.2.2). Notably, the orientation of the vinylphenylmethyl

moiety in the **SBL-WT(PMSv)** structure differs from that of the PMS group in two previously reported structures of a mutant SBL co-crystallised with PMSF.²¹ The aryl groups in these reported structures (1C9M and 1C9N.pdb) point towards and stack over His64 of the catalytic triad. An overlay of **SBL-WT(PMSv)** with 1C9N.pdb illustrates at least two possible binding conformations for the bound substrate (Scheme 4-8d).

4.2.5.2 SBL-Bpc217 structure [orientation 1]

Successful positioning of the unnatural amino acid Bpc in the S1' subsite was demonstrated in the structure obtained from a crystal of **apo-SBL-Bpc217** protein (**4.9**), using the sitting drop method with PEG 3350 and ammonium sulphate as precipitants in Bis-Tris buffer pH 5.5. The crystal structure was determined to 1.15 Å resolution (PDB ID: 5ARB). The structure showed the bipyridine moiety and its covalent attachment to the cysteine at position 217, modelled at an occupancy of 0.7, with b-factors varying from 13 to 30 Å² (Scheme 4-9). The X-ray data and refinement statistics are summarized in Table 4-1.

Additional π - π stacking interactions of Bpc were evident with residue Tyr209. Tyr209 is aligned in a similar plane to the pyridine rings, with the hydroxyl O atom 3.3 Å away from both of the N atoms of Bpc as shown (Scheme 4-9b and d). In this structure the metal binding amino acid Bpc is orientated away from the active site residue Ser221 in the S1 pocket. However, there is a region of unmodelled density between the modelled Bpc and His64, located close enough to Cys217 to represent a covalently attached moiety, but not sufficiently complete to enable the fitting of a second Bpc conformation at a lower occupancy. There is also an extra region of unmodelled density close to the side chain of Asn155, which we were unable to assign.

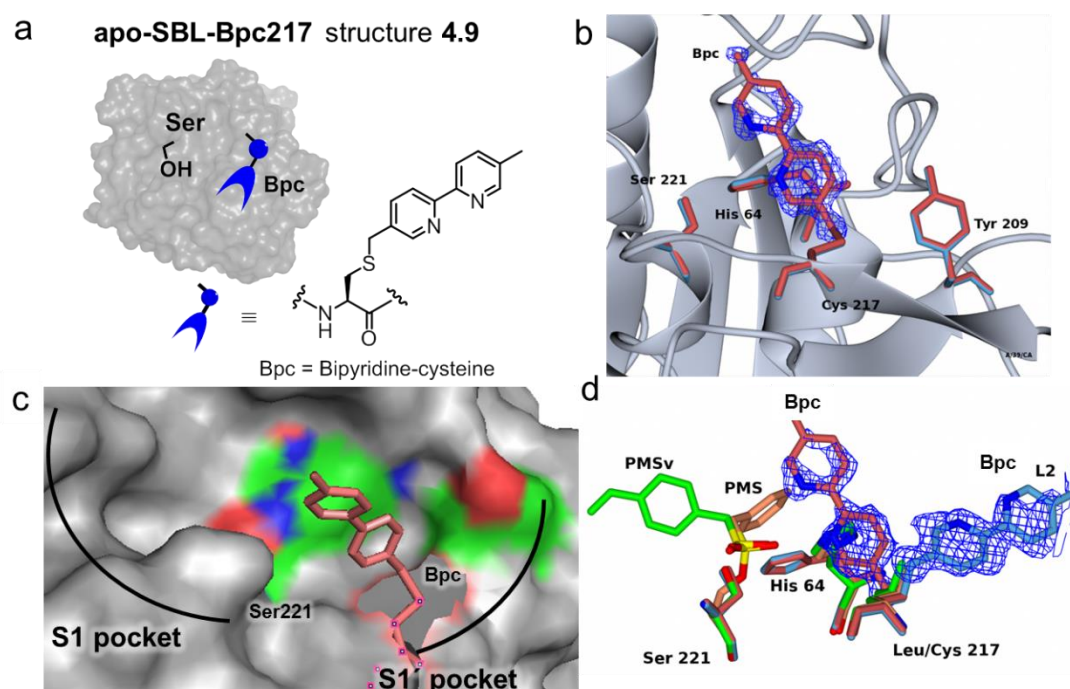


Scheme 4-9. (a) Graphical representation of the **apo-SBL-Bpc217** showing chemical structure of Bpc, (b) Close view of the active site structure of **apo-SBL-Bpc217** showing Bpc in a structure of a crystal grow at pH 5.5 (in pale blue, occ 0.7) contoured at 2.4 rmsd, (c) Surface view of **apo-SBL-Bpc217** structure depicting Bpc installed in S1' pocket pointing away from active site Ser221, (d) Relative orientations of active site residues and Tyr209 with respect to Bpc; residues in the catalytic triad (Ser221, His 64 and Asp32) and other pertinent residues are labelled (with residues coloured as in (Scheme 4-8b)); (b,d) the figures were made with ccp4mg²⁵ by Wendy Offen

4.2.5.3 SBL-Bpc217 structure [orientation 2]

Interestingly, we find an alternative orientation of Bpc in another structure obtained from an **apo-SBL-Bpc217** crystal soaked with PMSv substrate. A single crystal of **apo-SBL-Bpc217** protein (4.9) was obtained by the sitting drop method using PEG and ammonium sulphate as precipitant in Bis-Tris buffer pH 8.5, and was soaked with 3 mM PMSv substrate (4.3) for 10 min at room temperature. The crystal structure of **apo-SBL-Bpc217** (with PMSv soak) was refined to 1.1 Å resolution (PDB ID: 5ARC). The X-ray data and refinement statistics are summarized in Table 4-1.

The structure showed the bipyridine moiety pointing towards active site Ser221 and S1 pocket (Scheme 4-10), modelled at an occupancy of 0.5 (b-factors vary from 14 to 22 Å²). However, there is no density for the PMSv substrate visible in the maps calculated for this structure. There are additional density differences which could not be modelled, but which lie over the equivalent location of the first Bpc ring in the pH 5.5 structure described earlier in Section 4.2.5.2.



Scheme 4-10. (a) Graphical representation of the **apo-SBL-Bpc217** showing structure of Bpc, (b) Close view of the active site of **apo-SBL-Bpc217** showing Bpc in a structure grown at pH 8.5, soaked with PMSv for 10 min (PMSv not visible) (occ 0.5 in pale crimson); residues in the catalytic triad (Ser221, His 64 and Asp32) and other pertinent residues are labelled (coloured as in Scheme 4-8b above), (c) Surface view of **apo-SBL-Bpc217** depicting Bpc installed in S1' pocket pointing towards the active site Ser221, (d) Bpc in a structure of a crystal grown at pH 5.5 (in pale blue, occ 0.7) and at pH 8.5, soaked with PMSv for 10 min (PMSv not visible) (occ 0.5 in pale crimson), WT-PMSv (occ 0.5, green) and 1C9N.pdb (occ 1, coral); omit electron density Fo-Fc maps for Bpc contoured at 2.4 rmsd; (b,d) the figures were made with ccp4mg²⁵ by Wendy Offen

In the two Bpc structures reported here (PDB ID: 5ARB and 5ARC, obtained from crystals grown at pHs 5.5 and 8.5 respectively), His64 occupies a similar position to the less

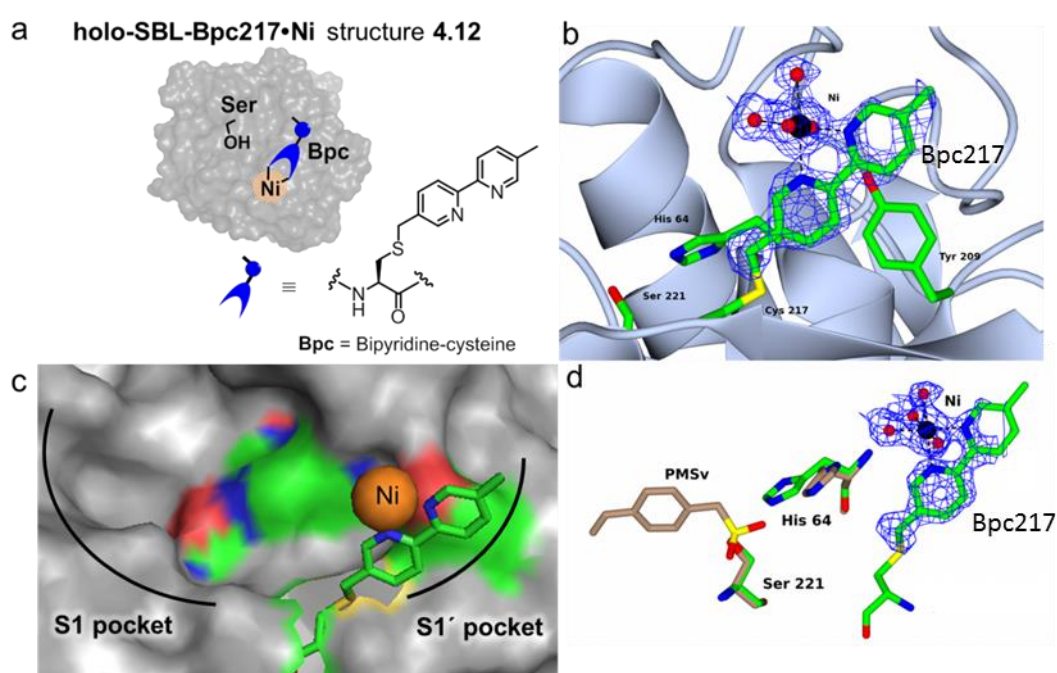
occupied conformation seen in the SBL-WT(PMSv) structure. In both structures, Nδ1 His64 forms hydrogen bonds with Oδ1 and Oδ2 Asp32 and Nε2 with OH Ser221 and a sulfonyl O atom. This orientation is similar to that in 1IAV.pdb, but different to the conformation observed in 1C9M and 1C9N.pdb, two structures in which PMS is oriented towards Leu217. This alternative histidine side-chain position circumvents likely steric clashes with the active site oriented bipyridine moiety in **apo-SBL-Bpc217** structure grown at pH 8.5, as demonstrated in an overlay of 1C9N.pdb and 5ARC.pdb (scheme 4-10d).²¹ In contrast, the Bpc conformation observed at pH 5.5 (5ARB.pdb; Section 4.2.5.2) directed away from the active site could coexist with either histidine orientation.

Graycar *et al.* observed that PMSF is visible at higher occupancies in structures for which the phenyl methyl group lies stacked over His64,²¹ i.e. for a histidine orientation the same as that in 1C9N.pdb (hydrogen bonded to Asp32) but dissimilar to that in the Bpc structures. This would suggest that it is unlikely that Bpc217 and His64 could occur in a suitable arrangement which would allow the concurrent observation of PMS bound at high occupancy in a crystal structure. It is interesting to note that the PMS moiety in 1IAV.pdb has atoms modelled at varying occupancies, with only the sulfonyl atoms at full occupancy. His64 was oriented towards Asp32 in all of the Bpc structures studied here, which might provide a reason why PMSv was not observable in any of the Bpc structures.

4.2.5.4 SBL-Bpc217•Ni structure

The location of the metal was confirmed using nickel as a metal surrogate. A crystal of **holo-SBL-Bpc217•Ni** protein (4.12) was obtained using PEG and ammonium sulphate as precipitant as above. The crystal structure of **holo-SBL-Bpc217•Ni** was determined to 1.55 Å resolution (PDB ID: 5ARD). The structure showed one atom of nickel was

coordinated to Bpc and four associated waters in a distorted octahedral coordination geometry (all at an occupancy of 0.5), installed at position 217 in the S1' subsite as originally designed (Scheme 4-11). This Bpc217•Ni complex is in a similar orientation to that in the structure of **apo-SBL-Bpc217** grown at pH 5.5 (Section 4.2.5.2). The X-ray data and refinement statistics are summarized in Table 4-1. Further, an X-ray fluorescence (XRF) scan confirmed the presence of nickel in the crystal (See Experimental Section 4.5.8).



Scheme 4-11. (a) Graphical representation of the **holo-SBL-Bpc217•Ni** showing structure of Bpc•Ni complex at position 217, (b) Close view of the active site structure of **holo-SBL-Bpc217•Ni**; showing the residues of the catalytic triad (Ser221, His 64 and Asp32) and Tyr209 and Bpc217 (coloured as in Scheme 4-8b above), (c) Surface view of **holo-SBL-Bpc217•Ni** structure depicting Bpc•Ni complex installed in S1' pocket pointing away from the active site Ser221, (d) Difference density omit map for Bpc and its associated Ni ion and waters contoured at 2.3 rmsd. The SBL-Bpc217•Ni structure is shown with C atoms in green, SBL-WT(PMSv) structure in brown. Bpc217, the Ni ion (black) and waters (red) are modelled at occupancy of 0.5. Ni coordination is shown as dashed lines of 2.0 – 2.2 Å; (b,d) the figures were made with ccp4mg²⁵ by Wendy Offen

The distances between the metal ion and its O (from water) and N ligands (from Bpc) is in 2.0-2.2 Å range. The two bipyridine nitrogens form a square equatorial plane with Ni-N distances of 2.1 Å, the other two O atoms in this plane are at 2.0 and 2.2 Å. The other two O atoms in the axial plane have bond distances of 2.0 and 2.2 Å. Notably, the side-chain of Tyr209 is aligned in a similar plane to the bipyridine rings (forming π - π stacking interactions as shown earlier in **apo-SBL-Bpc217** structure in Section 4.2.5.2), and its OH group is 3.5 Å from both of the N atoms of Bpc. There was additional unmodelled planar ring-shaped density close to Cys217 which may be occupied by Bpc oriented towards Ser221, but which was not clear enough to allow fitting of a whole alternative Bpc moiety.

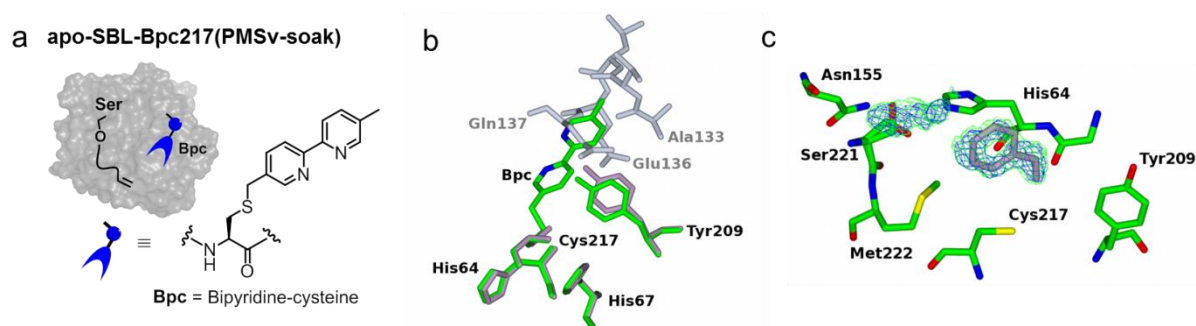
The crystal structure of **SBL-Bpc217•Ni** contains four associated water molecules, even though nickel acetate was used as the Ni²⁺ source during the metal coordination reaction. Additionally, the LC-MS analysis confirmed the mass of different metal coordinated enzyme species (SBL-Bpc217•Ni, SBL-Dpc217•Pd and SBL-Bpc217•Pd), after the loss of associated chloride/acetate ligands (See Chapter 2 Section 2.2.2). This confirmed our speculation of relatively fast exchange of acetate or chloride ligands to form corresponding aqua complexes which are perhaps dehydrated under LC-MS conditions. Datasets collected from other crystals grown from SBL-Bpc217•Ni samples showed clear electron density for nickel[†] (See Appendix 4-5). Overall, the orientations of Bpc217•Ni complex in these crystals were same as reported for the **SBL-Bpc217•Ni** structure and they exhibited similar interactions with neighbouring residues to the ones described above.

4.2.5.5 SBL-Bpc217 structure (with PMSv soak)

Additional crystallization conditions were also optimized to work under scavenger free

[†] Personal observation Wendy Offen

conditions based around NaCl/NaOAc conditions (1.8-3.2 M NaCl, 0.1 M sodium acetate pH 4.6). This was used to produce a stock for seeding, which was used to grow better quality crystals of SBL-WT(PMSv) **4.4** and SBL-L217C(PMSv) **4.5**. Using these promising scavenger free conditions (that did not contain PEG or KSCN), data were collected from **apo-SBL-Bpc217** crystals grown in NaCl (1.4 M NaCl, 0.1 M sodium acetate pH 4.6) conditions. SBL crystals in NaCl conditions grew in a different space group as described in Section 4.2.5. Unfortunately, clear density for Bpc was not seen in the electron density map. Overlaying the structure with that of a PEG grown crystal showed that the position of Gln137 from a symmetry-related molecule in the NaCl condition (in grey) would interfere with Bpc in the position in the **apo-SBL-Bpc217** [orientation 1] structure (Scheme 4-12b).



Scheme 4-12. (a) Graphical representation of the **apo-SBL-Bpc217(PMSv-soak)** showing structure of Bpc at position 217 and covalent attachment of PMSv substrate with Ser221, (b) Diagram to show position of residues around and including Cys217 in SBL-Bpc (pH5.5) overlaid on the equivalent residues in a crystal structure grown with NaCl as precipitant (in lilac). Residues in a symmetry related molecule of the NaCl-grown structure, which would clash with the Bpc modification, are shown in grey, (c) diagram to show extra unmodelled Fo-Fc and 2Fo-Fc density (omit maps contoured at 3 and 1 sigma respectively) in the active site of an apo-SBL-Bpc217 crystal grown in NaCl condition and soaked overnight with PMSv. Possible position of vinylphenyl group is shown in lilac; a region of electron density close to Ser221 is not modelled. Ser221 occupies two side-chain conformations in this structure; (b-c) the figures were made with ccp4mg²⁵ by Wendy Offen

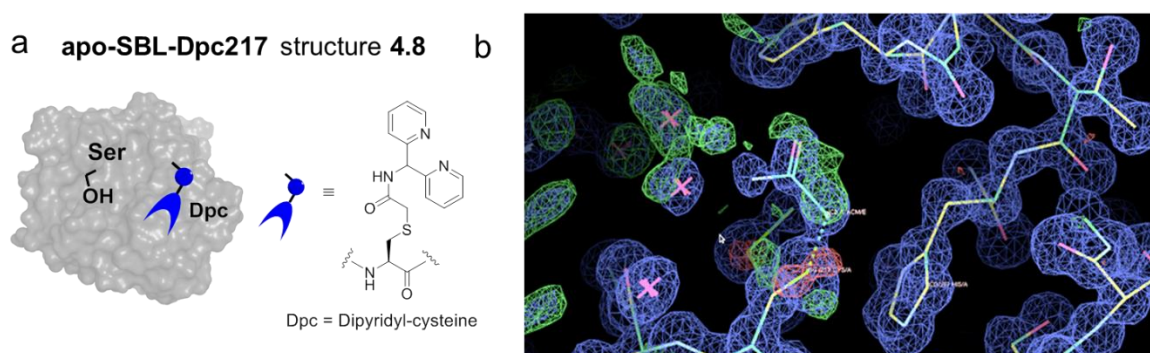
For soaking experiments, a single **apo-SBL-Bpc217** crystal grown by the sitting drop

method with NaCl as precipitant was soaked overnight with PMSv substrate **4.3**. Data were collected to 1.2 Å on beamline io4 at Diamond. The protein chain of 1SVN.pdb was used as the model for molecular replacement with Phaser, and an electron density map was calculated using the protein molecule with one calcium ion, 2 sodium ions, 2 chloride ions and water molecules added. The map of this **apo-SBL-Bpc217(PMSv-soak)** showed density which could be occupied by an aryl ring with two C atoms attached (like the vinyl-aryl ring of PMSv) at an occupancy of 1.0 (Scheme 4-12c). However, the density was not contiguous with an adjacent region of unmodelled density, where the methylene carbon atom and sulfonyl group were expected. The latter unmodelled density was also close to, but was not contiguous with density for the Ser221 side-chain (seen in two side-chain conformations), so covalent linkage with Ser221 was not observed. Considering these issues, it was not possible to model in a molecule of PMSv, and the structure was not refined.

If there was a molecule of partially occupied PMSv in this **apo-SBL-Bpc217(PMSv-soak)** structure, it would be interesting to note that the aryl group would be oriented towards Cys217 (NB. The density for Bpc was not visible in NaCl-grown crystals for the same reason as discussed above). Interestingly, orientation of the PMSv group observed in this structure was found to be opposite to that in the SBL-WT(PMSv) structure (See Section 4.2.5.1) as demonstrated in the overlay (Scheme 4-8b). Importantly, the side-chain of Asn155 (in the oxyanion site) was also found oriented towards the PMSv group (Scheme 4-12c). Although this crystal structure suggests flexibility of PMSv modification in the SBL active site, our further attempts to validate this observation by co-crystallization experiments were unsuccessful (See Appendix 4-4).

4.2.5.6 SBL-Dpc217 structure

A crystal of **SBL-Dpc217** protein (**4.8**) was obtained by the sitting drop method using ammonium sulphate and PEG as a precipitant. At best, a density for 2 or 3 atoms bound to the sulfur of Cys217 were visible in the 2Fo-Fc map. The possible position of the amide moiety of Dpc is modelled at occupancy of 0.5 and it was not possible to model the complete Dpc structure (Scheme 4-13). This was a consistent trend in all the crystals grown with Dpc modification perhaps due to the inherent flexibility of this modification containing two pyridyl rings attached with a methylene bridge (discussed in Section 4.2.3).



Scheme 4-13. (a) Graphical representation of the **apo-SBL-Dpc217** showing structure of unnatural amino acid Dpc at position 217, (b) 2Fo-Fc map (rmsd 1.0) for Dpc217 modification with the -C(=O)NH- of Dpc modelled at an occupancy of 0.5; the screenshot of **apo-SBL-Dpc217** as represented in Coot is shown

4.2.6 Microbeam Proton Induced X-ray Emission[†]

Microbeam Proton Induced X-ray Emission (MicroPIXE) is a useful technique for determining the elemental composition of liquid and crystalline protein samples, with relative accuracy of $\pm 10\text{-}20\%$.^{36,37} Systemic errors are minimized using this technique and intrinsic sulfur atoms provide for internal calibration making quantitative analysis possible. The detection limits have been defined as $1\text{-}10\ \mu\text{g g}^{-1}$. In our samples, the stoichiometry of

[†] MicroPIXE analysis on protein crystals grown with palladium-containing samples was carried out in collaboration with Professor Elspeth Garman.

the metal:protein composite was confirmed by LC-MS analysis as 1:1 and this confirmed that the sample contained palladium prior to crystallization screening. Furthermore, ICP-MS analysis had shown the protein:Pd stoichiometry to be 1:1 (Chapter 2 Section 2.2.3). However, none of the crystals grown with SBL-Bpc217•Pd or SBL-Dpc217•Pd showed any electron density for palladium. Moreover, an XRF scan could not be used to confirm the presence of palladium as described above (See Section 4.2.4). In this context, we used microPIXE to analyse these crystals to compare metal stoichiometry in the crystals grown with palladium with the results previously obtained by titration studies, LC-MS and ICP-MS analyses of these solutions. For our purposes, this technique was considered useful as it can also help to determine if there is any radiation damage in the SBL-Bpc217•Pd or SBL-Dpc217•Pd crystals that have been used to collect data previously.

SBL-Bpc217•Pd crystal was washed prior to analysis by microPIXE (see Experimental Section 4.5.12). In addition, the protein solution (from which the crystals were grown) was also analyzed. The SBL-Bpc217•Pd protein solution showed a strong signal for palladium. Importantly, the ratio of protein:Pd was found to be 1:0.9 for the protein solution. Considering the limits of experimental error, these results further support our results from ICP-MS analysis that showed the giving values corresponding to the presence of just one palladium atom per molecule of enzyme (Chapter 2 section 2.2.3). Unfortunately, the washed SBL-Bpc217•Pd protein crystal that was tested did not show a signal for palladium.[†]

[†] We were unable to obtain quantification data for this crystal due to technical difficulty; however no signal for palladium was detected in the crystal. Additionally, data could not be fully collected on an unwashed SBL-Bpc217•Pd crystal, but there appeared to be no palladium (Personal communication, Glyn Hemsworth)

4.3 Conclusions

Crystallization of the apo-SBL and holo-SBL enzymes was undertaken to elucidate the local environment around the installed metal centre in the active site. Even though the identity and homogeneity of these modified SBL-enzymes has been successfully confirmed by LC-MS analysis, X-ray crystallographic studies were performed to unambiguously confirm the site of installation of these modifications. Furthermore, we were also interested in the structural characterization of the different intermediates from the engineered cycle (Chapter 3) to study the various interactions of the SBL constructs with the bound substrate and metal. These interactions could potentially influence the activity, and the crystal structures of the intermediates would thus inform the future design and optimization of such biocatalysts. With this aim, the structural analysis of a substrate bound crystal **SBL-WT(PMSv)** was obtained and binding determinants that govern the specificity of the S1 pocket for such substrates were illustrated. The PMSv substrate was localised in the hydrophobic S1 pocket and various enzyme-substrate interactions were confirmed by X-ray crystallography, which were found to be in agreement with our previous results based on LC-MS (See Chapter 3 Section 3.2.2).

The X-ray crystal structures of **apo-SBL-Bpc217** were obtained in two different orientations confirming the successful installation of the unnatural amino acid residue in the S1' pocket of SBL. The root-mean-square deviation (rmsd) values of the C α atoms of **apo-SBL-Bpc217** grown at pH 5.5 and pH 8.5 from those of **SBL-WT** (PDB ID:1GCI)¹⁷ are 0.07 and 0.23 Å respectively. The overall structure of **SBL-Bpc217** was identical to that of **SBL-WT**, showing that the overall fold of SBL is preserved despite the incorporation of the unnatural amino acid Bpc in the active site of the serine protease. This tailored SBL active site was shown to harbour a non-native metal by structural

characterization of **SBL-Bpc217•Ni**, confirming the creation of a specific metal binding site. The rmsd value of the Ca's when the structure of **SBL-Bpc217•Ni** is superposed on SBL-WT (PDB ID:1GCI)¹⁷ structure is 0.521 Å , showing that the incorporation of metal in the active site did not alter the structure of SBL. The metal moiety was not coordinated to any other close lying amino acids and the occupancy of the Bpc217•Ni complex in this structure was found to be 50%, which is indicative of the conformational flexibility of the installed unnatural amino acid in the active site. We predict that this flexibility may be favourable for entry and accommodation of cross-coupling partners. The overlays of SBL structures obtained here clearly show the locations and possible relative orientations of the Bpc-metal complex and the substrate (Scheme 4-11d).

The lower occupancy of alkene substrate seen in **SBL-WT(PMSv)** and in other SBL structures (where electron density for only the covalently attached -SO₂ group to Ser221 was observed) indicates considerable conformational flexibility of the bound substrate at 100 K (Appendix 4-2 and 4-3). In addition, there is also evidence in the literature for the disorder or low occupancy of the phenyl ring of PMSF.²¹ This flexibility of the bound PMS groups shows that the enzyme not only serves to select, but may also influence the productive binding of alkene substrates to orient it towards the S1' subsite where the metal is installed. Furthermore, it seems unlikely that Bpc could exist oriented towards the active site (as in the pH 8.5 **apo-SBL-Bpc217 structure**) concurrently with the PMSv substrate being bound, with the vinyl group pointing towards the Bpc group (in a position similar to that of the PMS group in, for example, 1C9N.pdb), as would be expected for intermediate **III** in the engineered cycle (Chapter 3). Nonetheless, the caveat that these protein crystals structures could be affected by intermolecular contacts (which are not a limiting factor in solution under the reaction conditions) holds true. Having confirmed the location and close

proximity of the metal and bound substrate in the SBL active site through these structures, we conclude that the oxidative Heck cross-coupling reaction is plausible in the SBL active site.

It has been previously shown that different ligand-protein combinations position the metal catalyst in a unique microenvironment within the protein pocket, which influences the activity of artificial metalloenzymes.^{8,38} In this regard, the crystal structure of **SBL-Bpc217•Ni** provides information about the chiral environment around the metal which may be further used for ‘designed evolution’ of artificial metalloenzyme activity in the future to explore different metal-catalysed reactions.

4.4 References

- (1) Hakansson, K.; Wehnert, A.; Liljas, A. *Acta Crystallographica Section D* **1994**, *50*, 93.
- (2) Abe, S.; Hirata, K.; Ueno, T.; Morino, K.; Shimizu, N.; Yamamoto, M.; Takata, M.; Yashima, E.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, *131*, 6958.
- (3) Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, *130*, 10512.
- (4) Ueno, T.; Abe, M.; Hirata, K.; Abe, S.; Suzuki, M.; Shimizu, N.; Yamamoto, M.; Takata, M.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, *131*, 5094.
- (5) Satake, Y.; Abe, S.; Okazaki, S.; Ban, N.; Hikage, T.; Ueno, T.; Nakajima, H.; Suzuki, A.; Yamane, T.; Nishiyama, H.; Watanabe, Y. *Organometallics* **2007**, *26*, 4904.
- (6) Ueno, T.; Koshiyama, T.; Ohashi, M.; Kondo, K.; Kono, M.; Suzuki, A.; Yamane, T.; Watanabe, Y. *J. Am. Chem. Soc.* **2005**, *127*, 6556.
- (7) Ueno, T.; Ohashi, M.; Kono, M.; Kondo, K.; Suzuki, A.; Yamane, T.; Watanabe, Y. *Inorg. Chem.* **2004**, *43*, 2852.
- (8) Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; LeTrong, I.; Stenkamp, R. E.; Ward, T. R. *Angew. Chem.* **2008**, *120*, 1422.
- (9) Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angew. Chem. Int. Ed.* **2011**, *50*, 3026.

- (10) Zimbron, J. M.; Heinisch, T.; Schmid, M.; Hamels, D.; Nogueira, E. S.; Schirmer, T.; Ward, T. R. *J. Am. Chem. Soc.* **2013**.
- (11) Ory, J. J.; Mazhary, A.; Kuang, H.; Davies, R. R.; Distefano, M. D.; Banaszak, L. J. *Protein Eng.* **1998**, *11*, 253.
- (12) Fukumoto, K.; Onoda, A.; Mizohata, E.; Bocola, M.; Inoue, T.; Schwaneberg, U.; Hayashi, T. *ChemCatChem* **2014**, *6*, 1229.
- (13) Drenth, J.; Hol, W. G. J.; Jansonius, J. N.; Koekoek, R. *Eur. J. Biochem.* **1972**, *26*, 177.
- (14) Wright, C. S.; Alden, R. A.; Kraut, J. *Nature* **1969**, *221*, 235.
- (15) Bott, R.; Ultsch, M.; Kossiakoff, A.; Graycar, T.; Katz, B.; Power, S. *J. Biol. Chem.* **1988**, *263*, 7895.
- (16) Gallagher, T.; Oliver, J.; Bott, R.; Betzel, C.; Gilliland, G. L. *Acta Crystallographica Section D* **1996**, *52*, 1125.
- (17) Kuhn, P.; Knapp, M.; Soltis, S. M.; Ganshaw, G.; Thoene, M.; Bott, R. *Biochemistry* **1998**, *37*, 13446.
- (18) Bode, W.; Papamokos, E.; Musil, D. *Eur. J. Biochem.* **1987**, *166*, 673.
- (19) McPhalen, C. A.; James, M. N. G. *Biochemistry* **1988**, *27*, 6582.
- (20) Goddette, D. W.; Paech, C.; Yang, S. S.; Mielenz, J. R.; Bystroff, C.; Wilke, M. E.; Fletterick, R. J. *J. Mol. Biol.* **1992**, *228*, 580.
- (21) Graycar, T.; Knapp, M.; Ganshaw, G.; Dauberman, J.; Bott, R. *J. Mol. Biol.* **1999**, *292*, 97.
- (22) Micheelsen, P. O.; Vévodová, J.; De Maria, L.; Østergaard, P. R.; Friis, E. P.; Wilson, K.; Skjøt, M. *J. Mol. Biol.* **2008**, *380*, 681.
- (23) Bryan, P. N. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **2000**, *1543*, 203.
- (24) Betzel, C.; Klupsch, S.; Papendorf, G.; Hastrup, S.; Branner, S.; Wilson, K. S. *J. Mol. Biol.* **1992**, *223*, 427.
- (25) McNicholas, S.; Potterton, E.; Wilson, K. S.; Noble, M. E. M. *Acta Crystallographica Section D* **2011**, *67*, 386.
- (26) Wells, J. A.; Estell, D. A. *Trends Biochem. Sci* **1988**, *13*, 291.
- (27) Bott, R. R.; Chan, G.; Domingo, B.; Ganshaw, G.; Hsia, C. Y.; Knapp, M.; Murray, C. J. *Biochemistry* **2003**, *42*, 10545.
- (28) Robertus, J. D.; Kraut, J.; Alden, R. A.; Birktoft, J. J. *Biochemistry* **1972**, *11*, 4293.
- (29) Thathagar, M. B.; Kooyman, P. J.; Boerleider, R.; Jansen, E.; Elsevier, C. J.; Rothenberg, G. *Adv. Synth. Catal.* **2005**, *347*, 1965.
- (30) Li, J.; Lin, S.; Wang, J.; Jia, S.; Yang, M.; Hao, Z.; Zhang, X.; Chen, P. R. *J. Am. Chem. Soc.* **2013**, *135*, 7330.
- (31) Helliwell, J. R. *J. Cryst. Growth* **1988**, *90*, 259.
- (32) Garman, E. *Acta Crystallographica Section D* **2010**, *66*, 339.

- (33) Hough, M. A.; Antonyuk, S. V.; Strange, R. W.; Eady, R. R.; Hasnain, S. S. *J. Mol. Biol.* **2008**, *378*, 353.
- (34) Yano, J.; Kern, J.; Irrgang, K.-D.; Latimer, M. J.; Bergmann, U.; Glatzel, P.; Pushkar, Y.; Biesiadka, J.; Loll, B.; Sauer, K.; Messinger, J.; Zouni, A.; Yachandra, V. K. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 12047.
- (35) Robertus, J. D.; Alden, R. A.; Birktoft, J. J.; Kraut, J.; Powers, J. C.; Wilcox, P. E. *Biochemistry* **1972**, *11*, 2439.
- (36) Garman, E. *Structure* **1999**, *7*, R291.
- (37) Garman, E. F.; Grime, G. W. *Prog. Biophys. Mol. Biol.* **2005**, *89*, 173.
- (38) Pordea, A.; Ward, T. R. *Chem. Commun.* **2008**, *0*, 4239.

4.5 Experimental procedures

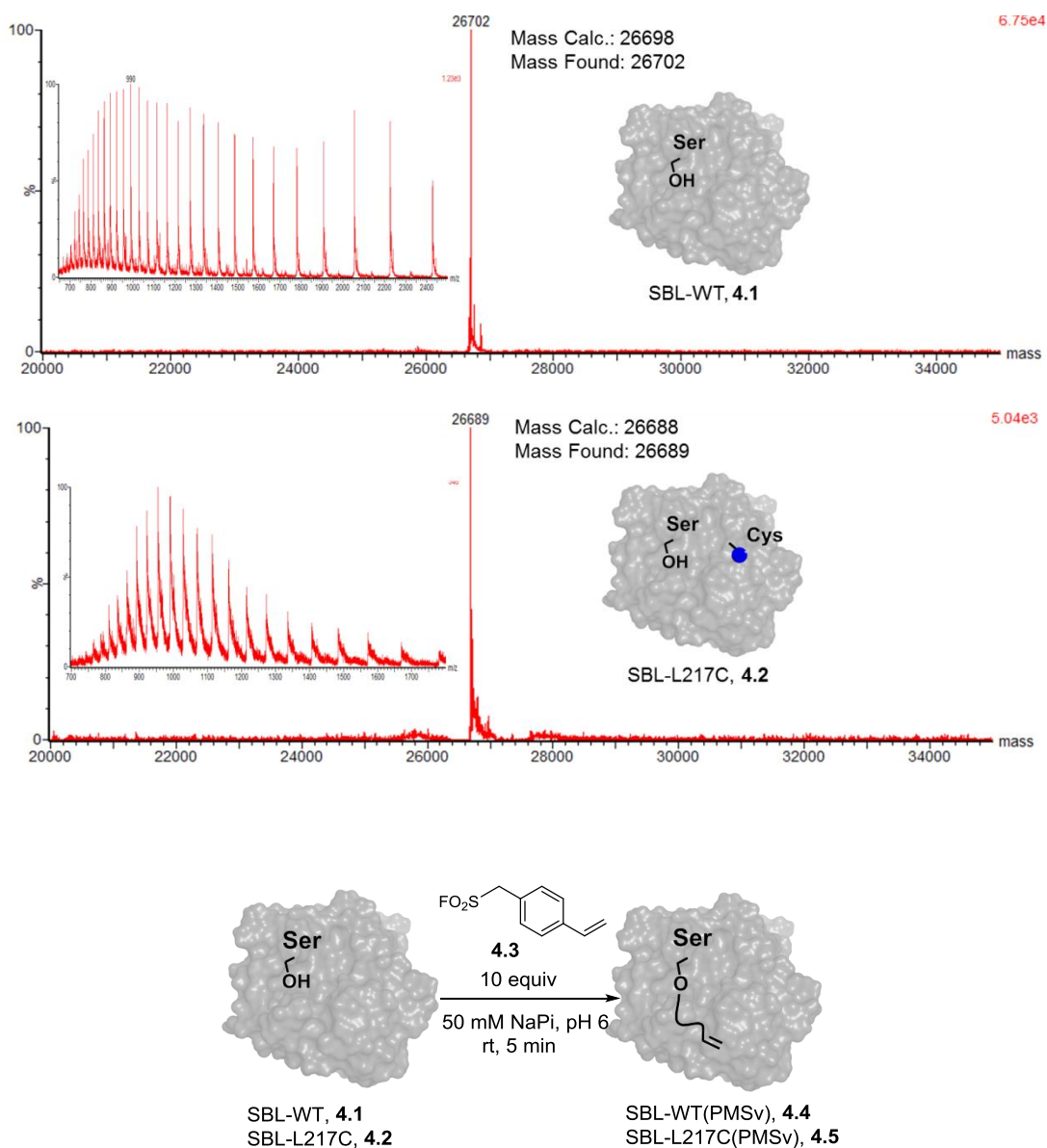
4.5.1 General considerations

Syntheses of the *p*-vinylPMSF substrate **4.3** (PMSv) and ligands **4.6** and **4.7**, used to install metal binding amino acids Bpc and Dpc, was carried out as described previously (Chapter 2 Section 2.5.3 and Chapter 3 Section 3.5.2). Protein samples corresponding to the target structures for crystallization were prepared using optimized chemical modification procedures described previously (for details See Chapter 2 Section 2.5.7 and Chapter 3 Section 3.5.8). Sample preparation for crystallization was carried out by scaling up of the optimized procedures by 10-12 times. Prior to setting up of the plates for screening, the samples were conditioned according to an established procedure described in Section 4.5.5.

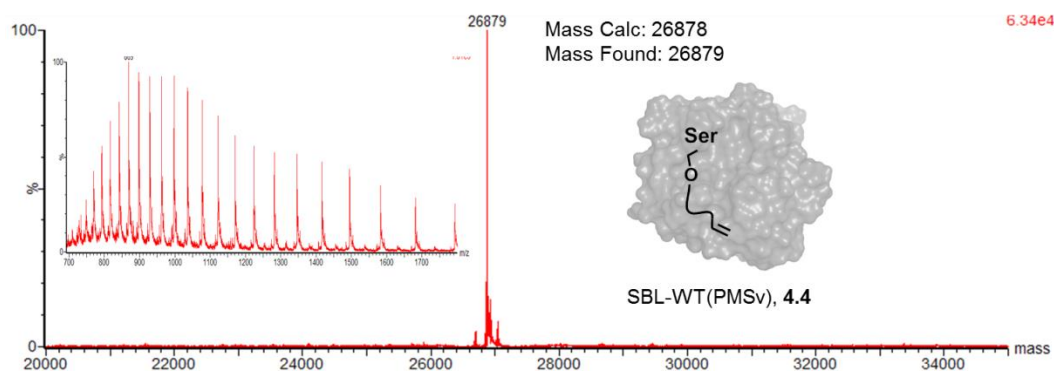
Protein modifications and characterization

4.5.2 Preparation of substrate bound SBL-WT(PMSv) and SBL-L217C(PMSv)

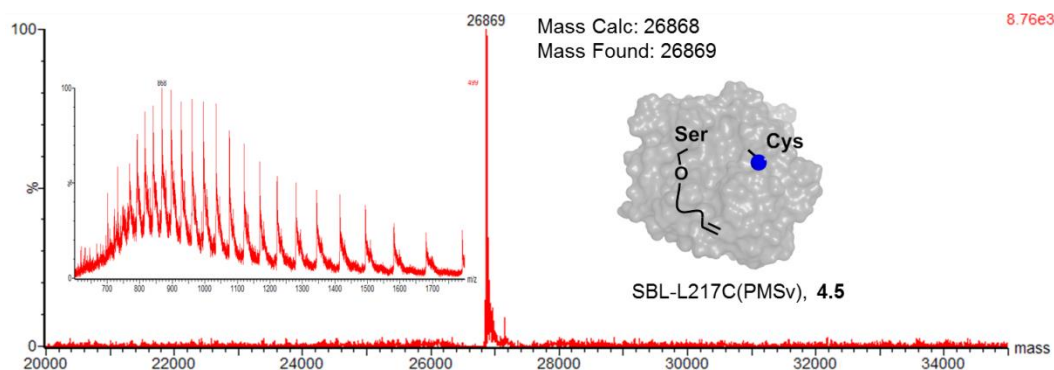
The purification of SBL-WT **4.1** and SBL-L217C **4.2** was accomplished by fast protein liquid chromatography using HiTrap SP sepharose HP columns (GE Healthcare) as described in Section 2.5.5.



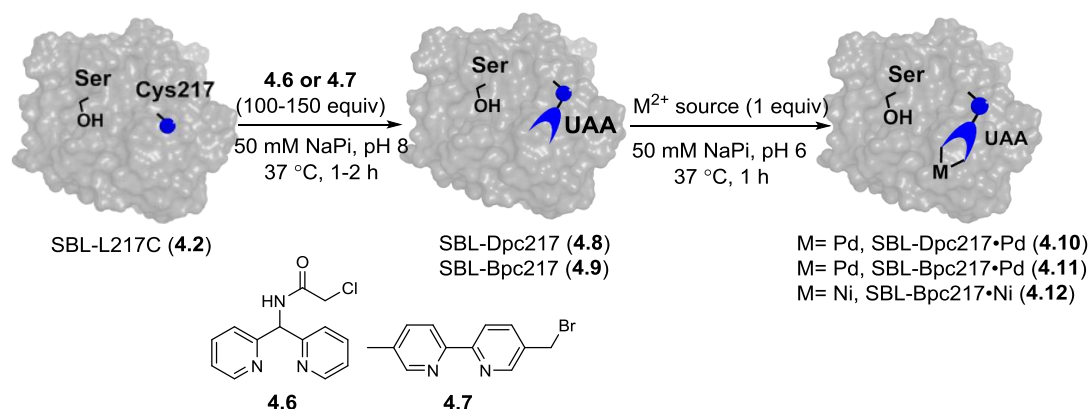
A 3 mL sample of SBL-WT **4.1** (90 μ M) was prepared at 2.4 mg/mL in 50 mM sodium phosphate buffer, pH 6. A 0.5 M stock solution of **4.3** was prepared in DMF. To a solution of SBL-WT (270 nmol) in 50 mM NaPi, pH 6, was added 9.4 μ L PMSv substrate **4.3** and the reaction was shaken at room temperature for 10 minutes. LC-MS analysis of an aliquot of the sample revealed >95% conversion to the substrate bound form, SBL-WT(PMSv) **4.4** (mass calculated: 26878, mass found: 26881). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.



A 4 mL sample of SBL-L217C **4.2** (90 μ M) was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 8 and was pre-treated with DTT (2.1 μ mol) for 10 minutes at room temperature. A 0.5 M stock solution of **4.3** was prepared in DMF. To a solution of SBL-L217C (245 nmol) in 50 mM NaPi, pH 6, was added 5 μ L PMSv substrate **4.3** and the reaction was shaken at room temperature for 10 minutes. LC-MS analysis of an aliquot of the sample revealed full conversion to the substrate bound form, SBL-WT(PMSv) **4.5** (mass calculated: 26868, mass found: 26871). The sample was conditioned according to standard procedure outline in Section 4.5.5. ESI-MS of the modification is shown below.

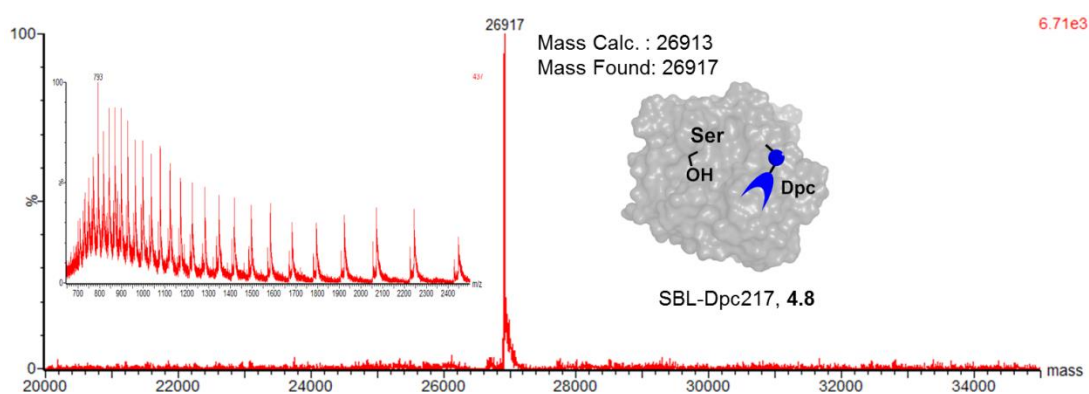


4.5.3 Preparation of apo- and holo- forms of SBL-based metalloenzymes



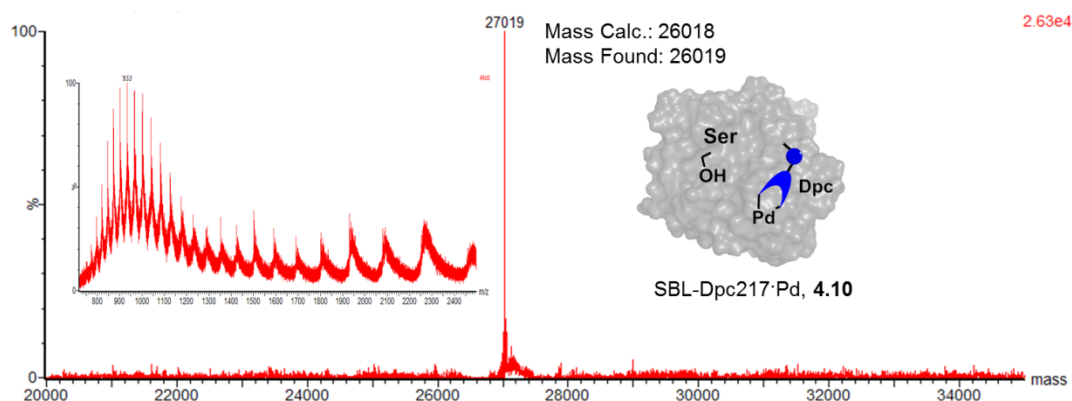
SBL-Dpc217 and palladium complexation:

A 3.5 mL sample of SBL-L217C **4.2** was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 8 and was pre-treated with DTT (2.1 μmol) for 10 minutes at room temperature. A 1 M stock solution of **4.6** was prepared in DMF. To a solution of SBL-L217C (215 nmol) in 50 mM NaPi, pH 8, was added 32 μL of **4.6** (32.3 μmol) and the reaction was shaken at 37 $^\circ\text{C}$ for 2 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Dpc217 **4.8** (mass calculated: 26914, mass found: 26917). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.



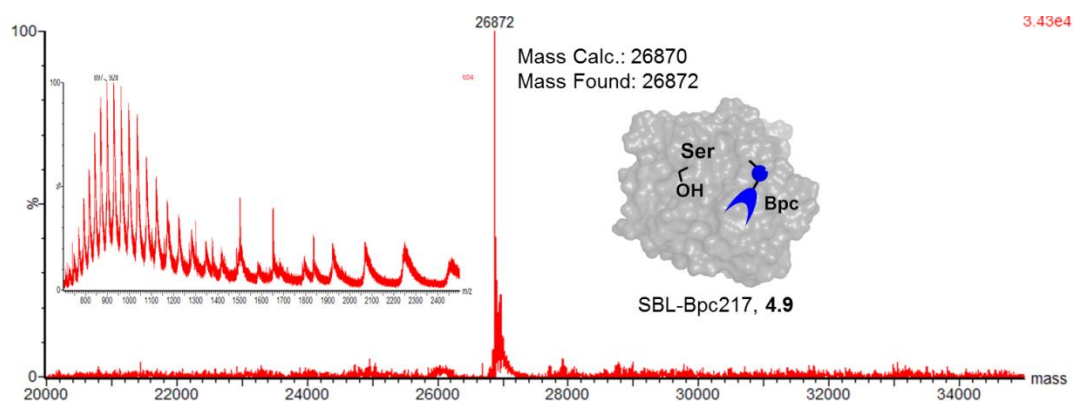
A 3.5 mL sample of SBL-Dpc217 **4.8** was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 6. The protein solution was buffer exchanged into 50 mM NaOAc,

pH 5.9 using PD-10 G-25 column (GE Healthcare). A 0.01 M stock solution of Na₂PdCl₄ was prepared in Milli-Q® water. To a 5 mL solution of SBL-Dpc217 (215 nmol) in 50 mM NaPi, pH 8, was added 22 µL of Na₂PdCl₄ (220 nmol) and the reaction was shaken at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Dpc217•Pd **4.10** (mass calculated: 27018, mass found: 27019). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

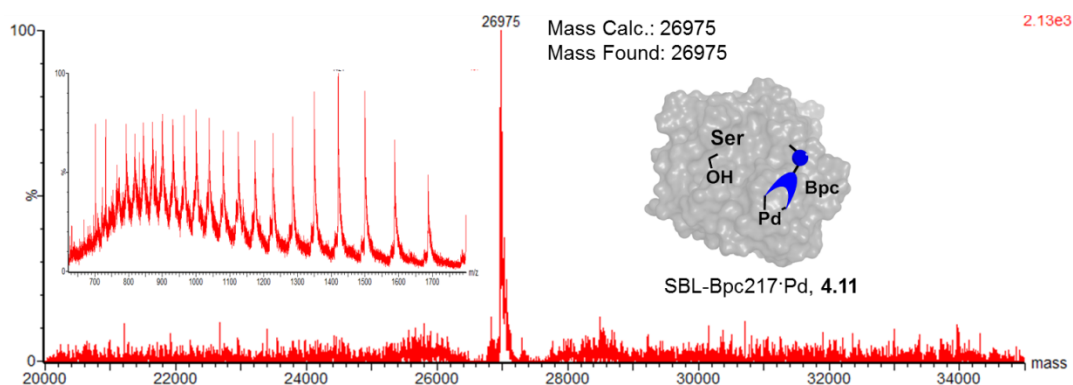


SBL-Bpc217 and metal complexation:

A 3 mL sample of SBL-L217C **4.2** was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 8 and was pre-treated with DTT (2.1 µmol) for 10 minutes at room temperature. A 0.5 M stock solution of **4.7** was prepared in DMF. To a solution of SBL-L217C (184 nmol) in 50 mM NaPi, pH 8, was added 80 µL of **4.7** (36.8 µmol) and the reaction was shaken at 37 °C for 2 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217 **4.9** (mass calculated: 26870, mass found: 26917). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

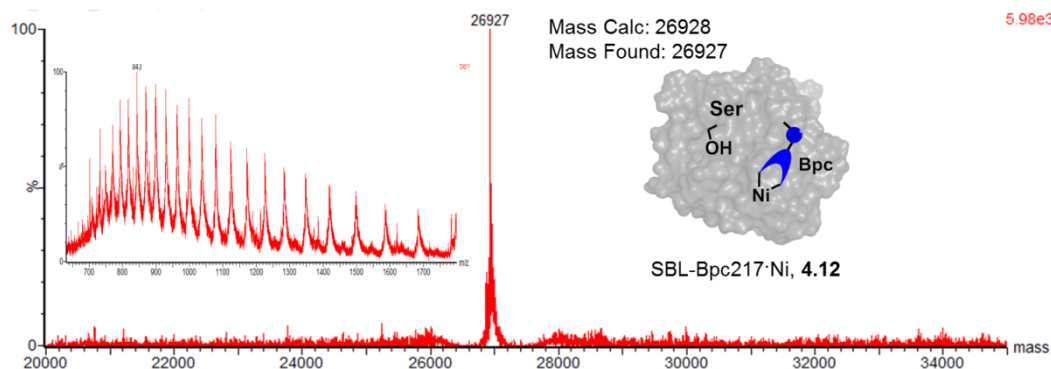


A 7 mL sample of SBL-Bpc217 **4.9** was prepared at 1.45 mg/mL in 50 mM sodium phosphate buffer, pH 6. The protein solution was buffer exchanged into 50 mM NaOAc, pH 5.9 using PD-10 G-25 column (GE Healthcare). A 0.01 M stock solution of Na_2PdCl_4 was prepared in Milli-Q® water. To a 10.5 mL solution of SBL-Bpc217 (380 nmol) in 50 mM NaPi, pH 8, was added 40 μL of Na_2PdCl_4 (399 nmol) and the reaction was shaken at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217•Pd **4.11** (mass calculated: 26975, mass found: 26976). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

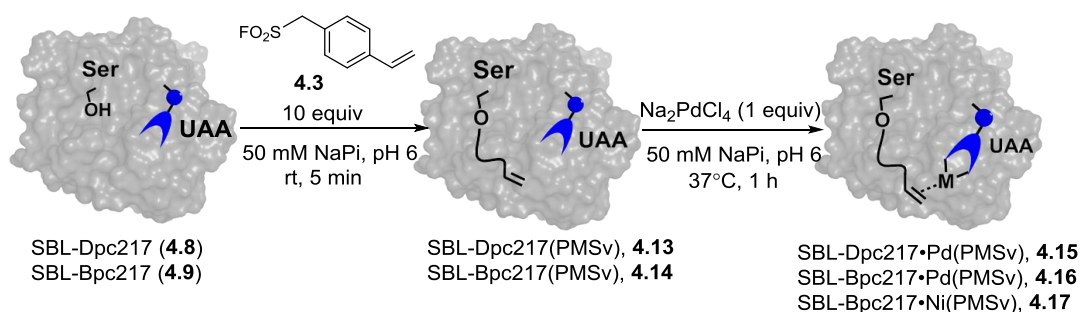


A 10.5 mL sample of SBL-Bpc217 **4.9** was prepared at 0.95 mg/mL in 50 mM sodium acetate buffer, pH 5.9. A 0.01 M stock solution of $\text{Ni}(\text{OAc})_2$ was prepared in Milli-Q® water. To a 10.5 mL solution of SBL-Bpc217 (380 nmol), was added 38 μL of $\text{Ni}(\text{OAc})_2$

(380 nmol) and the reaction was shaken at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217•Ni **4.12** (mass calculated: 26928, mass found: 26927). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

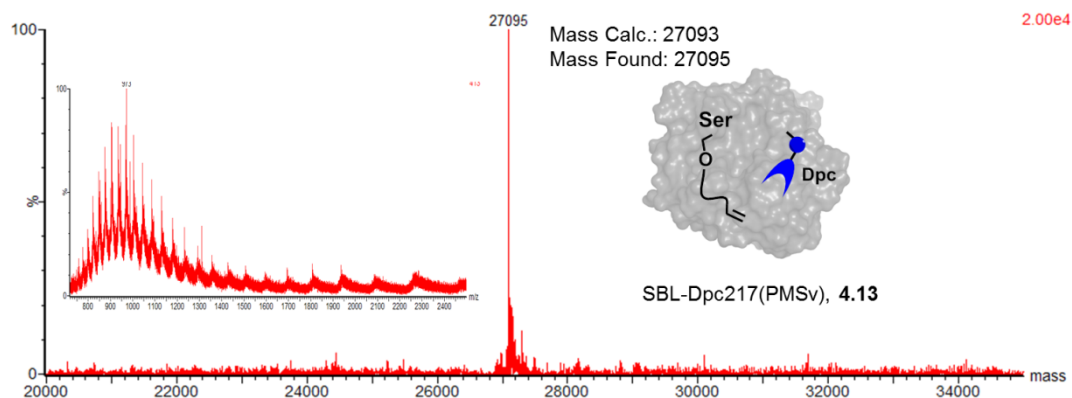


4.5.4 Preparation of crystallization samples for cycle intermediates

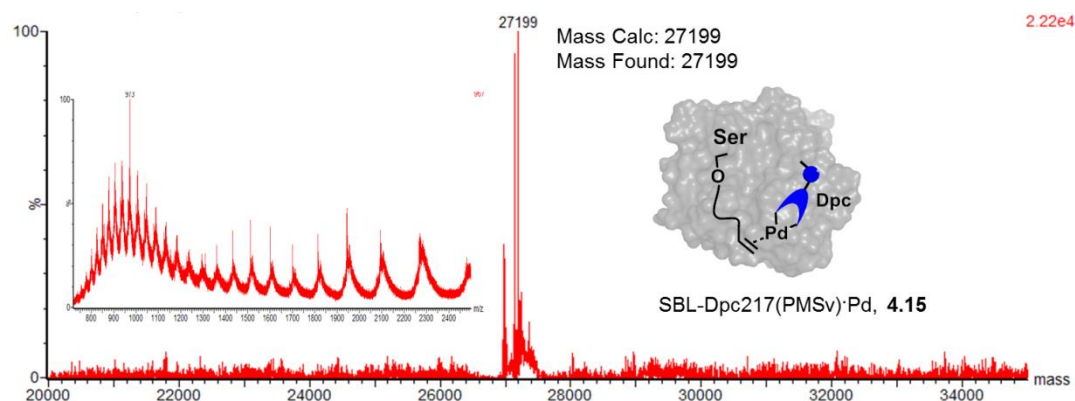


*SBL-Dpc217 with *p*-vinylPMSF **4.3** and palladium complexation:*

A 3.5 mL sample of SBL-Dpc217 **4.8** was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 6. A 0.5 M stock solution of *p*-vinylPMSF **4.3** was prepared in DMF. To a 3.5 mL solution of SBL-Dpc217 (215 nmol) in 50 mM NaPi, pH 6, was added 5 μ L of **4.3** (430 nmol) and the reaction was shaken at room temperature for 10 minutes. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Dpc217(PMSv) **4.13** (mass calculated: 27093, mass found: 27095). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

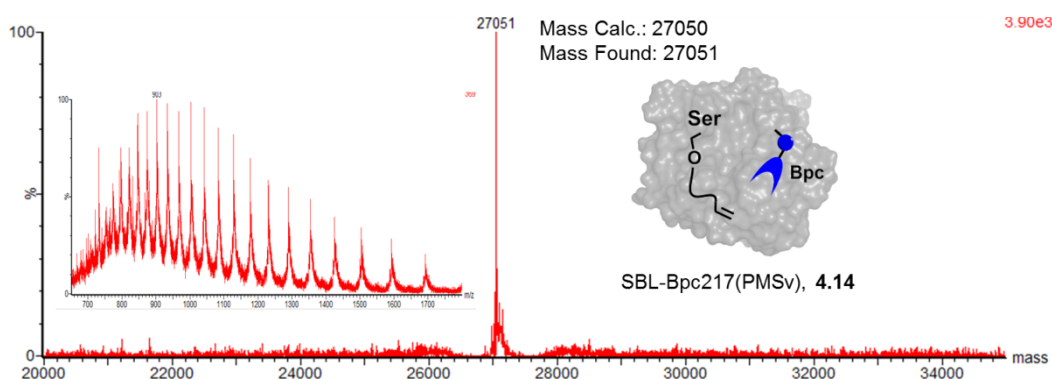


A 5 mL sample of SBL-Dpc217(PMSv) **4.13** was prepared at 0.78 mg/mL in 50 mM sodium phosphate buffer, pH 6. The excess *p*-vinylPMSF was removed by dialysis against 50 mM sodium phosphate buffer, pH 6. A 0.01 M stock solution of Na₂PdCl₄ was prepared in Milli-Q® water. To 5 mL sample of SBL-Dpc217(PMSv) (144 nmol) in 50 mM phosphate buffer, pH 6 at 4°C, was added 18 µl of Na₂PdCl₄ (172 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed a mixture of masses corresponding to SBL-Dpc217•Pd(PMSv) **4.15** (mass calculated: 27199, mass found 27199) and unreacted SBL-Dpc217(PMSv) **4.13** (mass calculated: 27093, mass found: 27095). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.

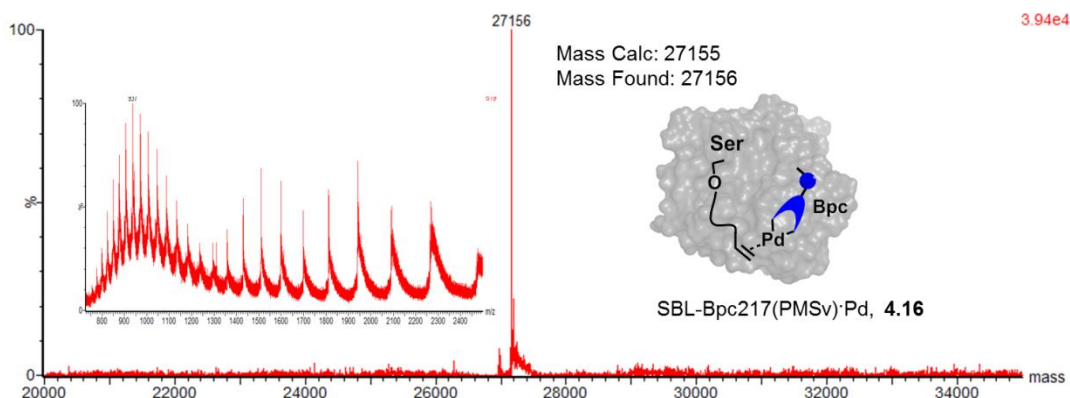


*SBL-Bpc217 with *p*-vinylPMSF **4.3** and metal complexation:*

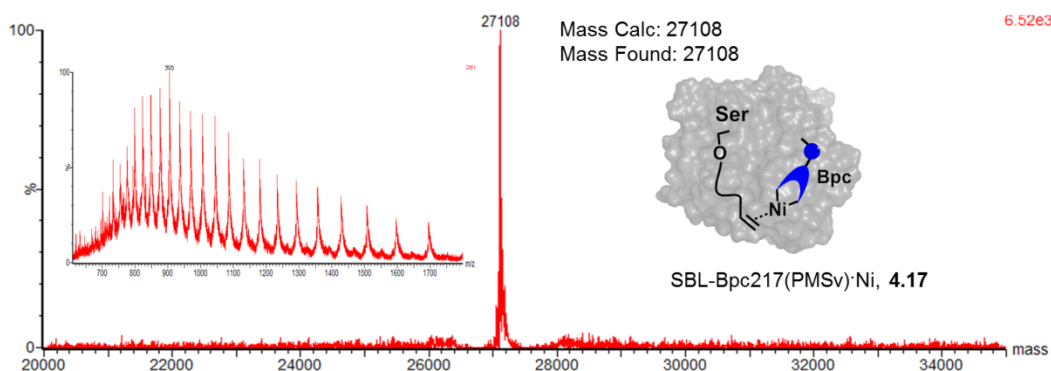
A 3 mL sample of SBL-Bpc217 **4.9** was prepared at 1.64 mg/mL in 50 mM sodium phosphate buffer, pH 6. A 0.5 M stock solution of *p*-vinylPMSF **4.3** was prepared in DMF. To a 3 mL solution of SBL-Bpc217 (185 nmol) in 50 mM NaPi, pH 6, was added 4 μ L of **4.3** (370 nmol) and the reaction was shaken at room temperature for 10 minutes. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217(PMSv) **4.14** (mass calculated: 27050, mass found: 27051). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.



A 6.7 mL sample of SBL-Bpc217(PMSv) **4.13** was prepared at 0.75 mg/mL in 50 mM sodium phosphate buffer, pH 6. The excess *p*-vinylPMSF was removed by dialysis against 50 mM sodium phosphate buffer, pH 6. A 0.01 M stock solution of Na₂PdCl₄ was prepared in Milli-Q® water. To 5 mL sample of SBL-Bpc217(PMSv) (184 nmol) in 50 mM phosphate buffer, pH 6 at 4°C, was added 18 μ L of Na₂PdCl₄ (184 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217•Pd(PMSv) **4.16** (mass calculated: 27155, mass found 27156). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.



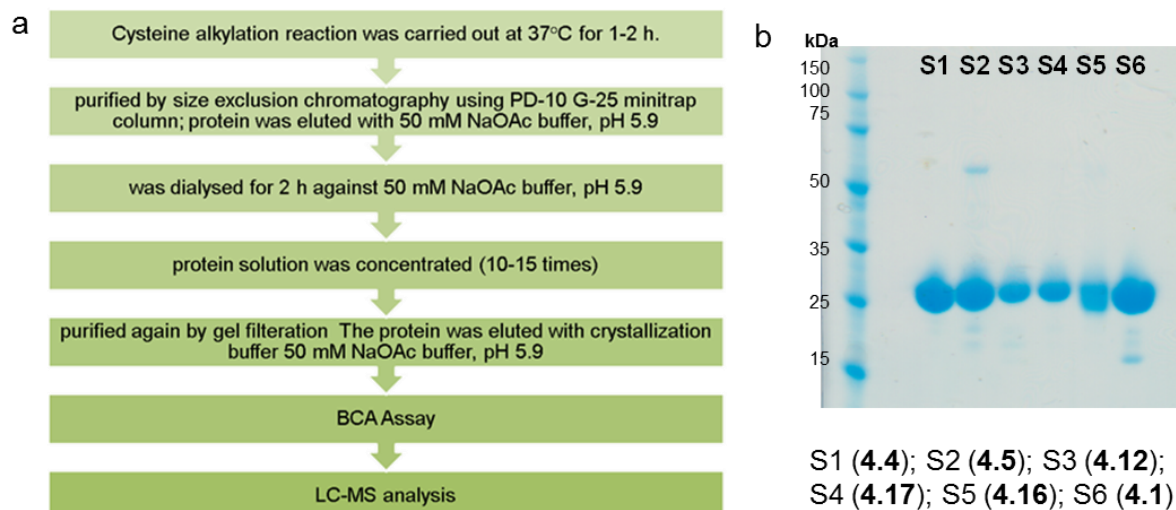
A 6.7 mL sample of SBL-Bpc217(PMSv) **4.13** was prepared at 0.75 mg/mL in 50 mM sodium phosphate buffer, pH 6. The excess *p*-vinylPMSF was removed by dialysis against 50 mM sodium phosphate buffer, pH 6. A 0.01 M stock solution of Ni(OAc)₂ was prepared in Milli-Q® water. To 5 mL sample of SBL-Bpc217(PMSv) (184 nmol) in 50 mM phosphate buffer, pH 6 at 4°C, was added 18 µL of Ni(OAc)₂ (184 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1 h. LC-MS analysis of an aliquot of the sample revealed full conversion to SBL-Bpc217•Ni(PMSv) **4.17** (mass calculated: 27108, mass found 27108). The sample was conditioned according to the standard procedure outlined in Section 4.5.5. ESI-MS of the modification is shown below.



4.5.5 Standard procedure for sample conditioning

The crystallization samples were conditioned using the following standard procedure:

- (i) The purification of SBL-WT and SBL-L217C was accomplished by using HiTrap SP HP sepharose column chromatography as described in Chapter 2 Section 2.5.5.
- (ii) Cysteine alkylation reaction was carried out at 37 °C for 1-2 h using 150-200 equiv ligands **4.6** and **4.7**.
- (iii) Where applicable, the protein was incubated with 1 equiv metal precursor for 1 h at 37 °C. Alternatively the protein was treated with 20 equiv *p*-vinylPMSF **4.3** for 10 minutes at room temperature.
- (iv) The reaction was purified by size exclusion chromatography using PD-10 G-25 minitrap column (GE Healthcare). The protein was eluted with 50 mM sodium acetate (crystallization buffer), pH 5.9 at 4 °C.
- (v) The chemically modified protein was dialysed for 2 h against 50 mM NaOAc buffer, pH 5.9 (buffer was exchanged twice).
- (vi) The protein solution was concentrated (10-12 times) using Vivaspin® (Sartorius Stedim, MW 5000 cut-off).
- (vii) The protein was purified again by gel filtration using PD-10 G-25 minitrap column (GE Healthcare). The protein was eluted with 50 mM NaOAc buffer, pH 5.9.
- (viii) A 1 µL aliquot of the crystallization sample was analysed by BCA assay to determine the final concentration of the protein sample prior to setting up the crystallization plates.
- (ix) 1 µL of the crystallization sample was used to prepare an LC-MS sample to confirm the identity of the protein and for non-reducing PAGE gel analysis.
- (x) Typically, the final concentrations of the protein samples used for crystallizations was 12-30 mg/ mL.



Scheme 4-14. (a) Standard procedure for sample conditioning prior to screening of crystallization conditions; (b) Representative non-reducing PAGE gel analysis of crystallization samples

4.5.6 Modelling studies

Modelling experiments were done by Dr. Mark Wormald, as described in Chapter 2 Section 2.5.9.

4.5.7 Crystallisation

These experiments were done with Wendy Offen. In general, crystals were grown using commercial or optimized screens (See Appendices 4-1 to 4-5) in 96- or 48-well sitting drop trays (MRC 96- and MRC Maxi 48-well crystallization plates (Swissci) Molecular Dimensions). After optimization, all crystallizations were seeded using a stock made with a SeedBead (Hampton Research Ltd) and SBL crystals grew within 2-5 days at 18 °C.

SBL-WT(PMSv) crystal: Substrate bound SBL-WT(PMSv) (**4.4**) at a concentration of 19 mg/ mL in 50 mM sodium acetate, pH 5.9, was crystallised using a drop containing 0.5 µl protein: 0.1 µl seed solution: 0.4 µl mother liquor (the latter comprising 9% PEG

(polyethylene glycol) 3350, 50 mM ammonium sulphate, 0.1 M Bis-Tris, pH 5.5). The seed stock (I) was made from crystals grown over 19% PEG 3350, 0.2 M ammonium sulphate, 0.1 M Bis-Tris pH 5.5. The crystal was harvested via mother liquor supplemented with 20% (v/v) glycerol as a cryoprotectant, using a nylon CryoLoop (Hampton Research), into liquid nitrogen.

SBL-Bpc217 [orientation 1] crystal: SBL-Bpc217 (**4.9**) was grown from 12 mg/ mL enzyme mixed 1:0.1:0.4 with the same seed stock (I) and 15% PEG 3350, 75 mM ammonium sulphate, 0.1 M Bis-Tris, pH 5.5 mother liquor. The crystal was harvested via mother liquor supplemented with 20% (v/v) glycerol as a cryoprotectant.

In addition, further crystals were grown over 27% PEG 3350, 0.1 M ammonium sulphate, 0.1 M Bis-Tris-propane pH 8.5 were used to make a seed stock (II).

SBL-Bpc217 [orientation 2] crystal: SBL-Bpc217 (**4.9**) grew from 12 mg/ mL enzyme with seed stock (II) using 25% PEG 3350, 0.1 M ammonium sulphate, 0.1 M Bis-Tris-propane, pH 8.5 as the mother liquor. The crystals were soaked with 3 mM PMSv substrate **4.3** in mother liquor for 10 min prior to fishing via cryoprotectant, but PMSv was not observed in the structure obtained.

SBL-Bpc217•Ni crystal: A crystal of SBL-Bpc217•Ni (**4.12**) was grown over 33% PEG 3350, 0.1 M ammonium sulphate, 0.1 M Tris pH 7.5.

4.5.8 Soaking Experiments

These experiments were done by Wendy Offen on the crystals grown previously. Crystals

of **SBL-Bpc217** had grown easily with NaCl as precipitant (1.4 M NaCl, 0.1 M sodium acetate pH 4.6), soaking experiments with *p*-vinylPMSF **4.3** (PMSv) and Na₂PdCl₄ were attempted with these NaCl-grown crystals.

Soak 1: A 0.5 M solution of *p*-vinylPMSF **4.3** was made in DMF, which was further diluted 100x with mother liquor solution. The solution looked cloudy due to precipitation of PMSv in aqueous solutions (the cloudiness had disappeared within 1 h). 1 µL was added to a 1.5 µl crystallization drop, and the drop was left for 20 h, after which the crystal was fished via a cryoprotectant comprising 4.5 M NaCl, 0.1 M Na acetate pH 4.6.

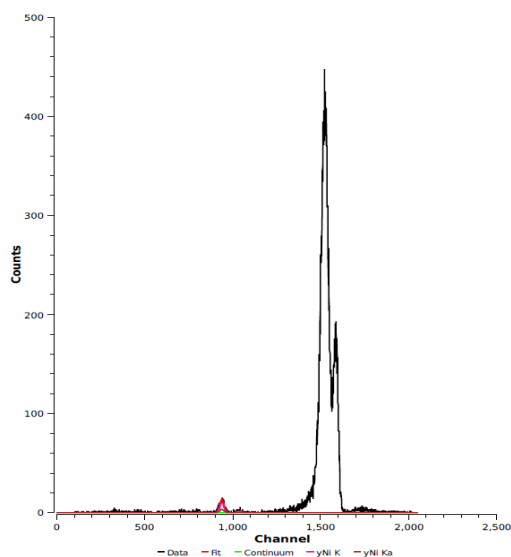
Soak 2: Another crystal from the same drop was rinsed through 6 x 20 µl well solution to wash off unbound PMSv, and then incubated in 20 µl well solution with 1 µl 10 mM Na₂PdCl₄ added. The drop was incubated at room temperature for 12 min, before fishing via the cryoprotectant.

Soak 3: A further soak experiment with PMSv substrate with a shorter incubation time, and an overnight Na₂PdCl₄ soak, was performed as follows. A crystal of SBL-Bpc217 grown in NaCl was left for 30 min after adding 2 µl 5 mM PMSv, prepared as above, to the 1 µl drop. The crystal was rinsed by fishing through 6 drops of 20 µl well solution, and incubated overnight in a 1 µl drop in the original crystallisation tray, to which 2 µl 1mM Na₂PdCl₄ (made by diluting a 10 mM stock 10 fold in well solution) was added. The crystal was fished via cryoprotectant as above.

4.5.9 Data collection

Data were collected at Diamond Light Source Ltd., on beamlines io3 (**SBL-WT(PMSv)**)

and **SBL-Bpc217•Ni**, to 1.0 Å and 1.55 Å respectively) and io4 (**SBL-Bpc217 [orientation 1]** grown at pH 5.5 and **SBL-Bpc217 [orientation 2]** grown at 8.5, to 1.2 Å and 1.1 Å respectively). An X-ray fluorescence scan was performed on a further crystal grown in the same drop as SBL-Bpc217•Ni crystal described above.



Scheme 4-15. X-Ray fluorescence scan for SBL-Bpc217•Ni crystal

Additional datasets were collected for each crystal from soaking experiments (sub-projects **soak 1** and **soak 2** collected on beamlines io4 and io2 respectively), which had space group $P2_1$ and cell 40, 63, 42 Å; 90 118 90°. Data was collected from **soak 3** crystal on beamline io2.

4.5.10 Data Processing and Refinement

Data processing and refinement of crystal structures was done by Wendy Offen. Data processing was carried out using XDS. The crystals grown in PEG/ammonium sulphate conditions were of space group $P2_12_12_1$, with unit cell dimensions of 52.65 Å, 61.25 Å, 74.61 Å (See Table 4-1). For the SBL-Bpc217•Ni crystal, fewer batches of data were used as the Rmerge started to increase towards the end of data collection. The structures were

solved using the isomorphous structure PDB entry 1GCI (without waters or ions) as the model. Initial cycles of Refmac gave a reasonable R-factor, and water molecules and calcium and sulfate ions were added manually to the structure in Coot, followed by further cycles of refmac before fitting in a molecule of PMSv (**4.3**) into partial density next to Ser221 in the WT structure at an occupancy of 0.5, or ligand **4.7** into partial density at Cys217 for the modified protein structures (at occupancies of 0.7 for SBL-Bpc217 [orientation 1] grown at pH 5.5, 0.5 for SBL-Bpc217 [orientation 2] grown at pH 8.5 and 0.5 for L2-Ni). In addition, Ni was modelled at occupancy of 0.5 in the SBL-Bpc217•Ni structure.

4.5.11 Amino acid sequence for PDB deposition

The sequence information for the structures is same as Uniprot P29600, however the numbering for the latter is non-contiguous as it is based on alignment to Uniprot sequence P00782 for subtilisin BPN' from *Bacillus amyloliquefaciens*. Thus there are no residues 37, 58 or 159-162. The protein sequence is deposited as:

```
AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPTQDGN
GHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNGMHVA
NLSLGSPSPSATLEQAVNSATSRGVLVVAASGNSGAGSISYPARYANAMAVGATDQNNNR
ASFSQYGAGLDIVAPGVNVQSTYPGSTYASCNGTSMATPHVAGAAALVKQKNPSWSNVQI
RNHLKNTATSLGSTNLYGSGLVNAEAATR
```

4.5.12 MicroPIXE analysis

MicroPIXE analysis was done in collaboration with Prof. Elspeth Garman. The crystal sample was washed by Wendy Offen prior to analysis. Two crystals grown in crystallisation condition is 20% PEG 3350, 0.1 M Tris pH 7.5 were washed prior to analysis with a solution that contained 21% PEG 3350, 0.1 M ammonium sulfate, 0.1 M tris pH 7.5 (which was the mother liquor) and one with 21% PEG 3350, 0.1 M tris pH 7.5

(i.e. without ammonium sulfate to avoid any possible interference due to NH_4^+ and SO_4^- ions). Protein solutions of SBL-Bpc217•Pd in 50 mM NaOAc buffer, pH 5.9 at 23 mg/mL were also analysed. The sulphur content in protein is 4 as given in the protein sequence above (3 methionines, 1 cysteine), that provides for internal calibration.

For measurements, a 2.5 MeV energy proton beam with a current between 100–200 pA was used. The distance of the sample to Si(Li) detector was 60 mm. A beam of approximate size of 2.5 μm and 130 μm Be filter was used. A SILI10-41 detector was used (with same characteristics as the database detector SILI80_3). Measurement from glass standard was run and validated prior to recording data from protein solution and crystal.

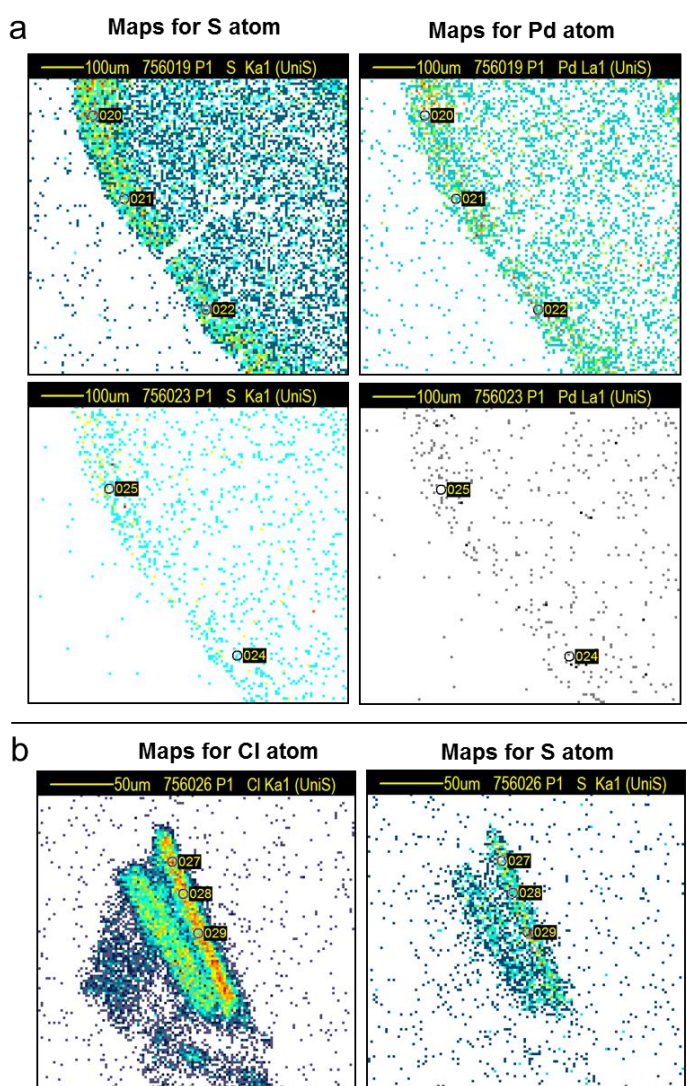
Table 4-2. Elemental analysis from SBL-Bpc217•Pd protein solutions

	Run	Element		
		16 S K 1	46 Pd K 1	46 Pd LA 1
Concentration	756020	30913	20466.4	18616.3
	756024	22472.1	20170.9	17277.3
	756025	29518.7	18988.4	17771.8
Atoms per molecule	756020	4	0.80	0.73
	756024	4	1.08	0.93
	756025	4	0.78	0.73
	Average	4	0.89	0.79
LOD (atoms per molecule)	756020	0.0017	0.035	0.0013
	756024	0.0029	0.050	0.0018
	756025	0.0019	0.029	0.0014
	Average	0.0022	0.038	0.0015
Standard deviation		0	0.14	0.09
Average fit error		0.00020	0.00074	0.000091
Quadratic deviation		0.00020	0.14	0.095

Table 4-3. Elemental analysis from SBL-Bpc217•Pd protein crystal

	Run	Element		
		16 S K 1	46 Pd K 1	46 Pd LA 1
Concentration	756027	23818.4	0	1691.1
	756028	23389.9	0	2140.7
	756029	28565.5	1105.3	1953.6

Atoms per molecule	756027	4	0	0.086
	756028	4	0	0.11
	756029	4	0.047	0.083
	Average	4	0.016	0.093
LOD (atoms per molecule)	756027	0.0065	0.030	0.017
	756028	0.0062	0.028	0.018
	756029	0.0065	0.023	0.019
	Average	0.0064	0.027	0.018
Standard deviation		0	0.022	0.012
Average fit error		0.026	0.11	0.072
Quadratic deviation		0.026	0.11	0.073



Scheme 4-16. Elemental maps collected from (a) SBL-Bpc217•Pd protein solutions showing location of sampling points on the protein drop; maps from run 756019 (left), maps from run 756023 (right), (b) SBL-Bpc217•Pd protein crystal showing location of sampling points

Appendix 4-1

Optimization of crystallization conditions: Round I (sitting drop plates)

PLATE	SBL PROTEIN	SUBSTRATE	SCREEN	METHOD	REMARKS
<i>P#1 drop 1</i>	Dpc217(Pd)	PMSF	Ammonium sulphate (Qiagen)	3 equiv Na ₂ PdCl ₄ added on plate set-up	
<i>P#1 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#2 drop 1</i>	Dpc217(Pd)	PMSF	SaltRx	3 equiv Na ₂ PdCl ₄ added on plate set-up	Needle cluster in position H12
<i>P#2 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#3 drop 1</i>	Bpc217(Pd)	PMSF	Ammonium sulphate (Qiagen)	3 equiv Na ₂ PdCl ₄ added on plate set-up	
<i>P#3 drop 2</i>	Bpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#4 drop 1</i>	Bpc217(Pd)	PMSF	SaltRx	3 equiv Na ₂ PdCl ₄ added on plate set-up	Needle cluster in C7 (drop 1) and H12 (drop 1)
<i>P#4 drop 2</i>	Bpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#5 drop 1</i>	Dpc217(Pd)	PMSF	Ammonium sulphate (Qiagen)	No palladium (control)	Needle cluster in G8 (drop 2) and G10 (drop 2)
<i>P#5 drop 2</i>	Bpc217(Pd)	PMSF			
<i>P#6 drop 1</i>	Dpc217(Pd)	PMSF	JCSG+	3 equiv Na ₂ PdCl ₄ added on plate set-up	Crystal cluster in B2 (drop 2)
<i>P#6 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#7 drop 1</i>	Dpc217(Pd)	PMSF	PACT	3 equiv Na ₂ PdCl ₄ added on plate set-up	Single crystal in H4 (drop 1) 2D crystal in F4 (drop 1)
<i>P#7 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#8 drop 1</i>	Dpc217(Pd)	PMSF	Ammonium sulphate (Qiagen)	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
<i>P#8 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#9 drop 1</i>	Dpc217(Pd)	PMSF	SaltRx	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
<i>P#9 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#10 drop 1</i>	Dpc217(Pd)	PMSF	JCSG+	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
<i>P#10 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#11 drop 1</i>	Dpc217(Pd)	PMSF	PACT	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
<i>P#11 drop 2</i>	Dpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#12 drop 1</i>	Bpc217(Pd)	PMSF	Ammonium sulphate (Qiagen)	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
<i>P#12 drop 2</i>	Bpc217(Pd)	<i>p</i> -vinylPMSF			
<i>P#13 drop 1</i>	Bpc217(Pd)	PMSF	SaltRx	Prior incubation of protein with 3 equiv	

			Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	
P#13 drop 2	Bpc217(Pd)	<i>p</i> -vinylPMSF		

Appendix 4-1 contd.

Optimization of crystallization conditions: Round I (24-well hanging drop plates)

PLATE	SBL PROTEIN	SUBSTRATE	SCREEN	METHOD	REMARKS
T#1 drop 1	Dpc217	PMSF	Ammonium sulphate (20-35%)	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	Varying ratio of protein to precipitant in the drop (2.5:1, 2:1, 1:1); pH maintained constant at 4.9
T#1 drop 2	Dpc217(Pd)	PMSF			
T#1 drop 3	Bpc217(Pd)	PMSF			
T#2 drop 1	Dpc217	PMSF	Ammonium sulphate (20-35%)	Prior incubation of protein with 3 equiv Na ₂ PdCl ₄ for 1 h at 37°C, 50 rpm	Varying pH 4-8; constant protein to precipitant ratio in the drop (2:1)
T#2 drop 2	Dpc217(Pd)	PMSF			
T#2 drop 3	Bpc217(Pd)	PMSF			

Appendix 4-2

Optimization of crystallization conditions: Round II and III

PLATE	SBL PROTEIN	SAMPLE CONC(mg/mL)	SCREEN	METHOD	REMARKS
T#1	WT	20	Index Screen	SBL-WT dialysed into 50 mM NaOAc buffer pH 5.9	Protein to precipitant ratio 1:1; crystals grew in F6, F7 and F9) G2, G3, G4, G5
T#2 T#3	WT	20	48-well format around F9 Index Screen	SBL-WT dialysed into 50 mM NaOAc buffer pH 5.9	Two duplicate trays set up; one seeded and other with additives
T#4 T#5	WT	20	48-well format around G4 Index Screen	SBL WT dialysed into 50 mM NaOAc buffer pH 5.9	Two duplicate trays set up; one seeded and other with additives
T#6	WT	20	PACT Premier Screen	SBL WT dialysed into 50 mM NaOAc buffer pH 5.9	Single crystals in A8, A9, A10
T#7	WT(PMSv)	19	NaOAc/ NaCl Opt. Screen	SBL WT dialysed into 50 mM NaOAc buffer pH 5.9; 7.3 mM <i>p</i> -vinylPMSF added on set up	Varying protein to precipitant ratio in the drop (2:1, 1:1, 1:2)
T#8	WT(PMSv)	19	PACT Premier Screen	SBL WT dialysed into 50 mM NaOAc buffer pH 5.9; 7.3 mM <i>p</i> -vinylPMSF added on set up	Varying protein to precipitant ratio in the drop (2:1, 1:1, 1:2); partial density of PMSv seen (only -SO ₂ visible)
T#9	L217C(PMSv)	14.7	NaOAc/ NaCl Opt. Screen	SBL L217C modified with <i>p</i> -vinylPMSF; dialysed into 50 mM NaOAc buffer pH 5.9	Varying protein to precipitant ratio in the drop (2:1, 1:1, 1:2); partial density of PMSv seen (only -SO ₂ visible)
T#10	L217C(PMSv)	14.7	PACT Premier Screen	SBL L217C modified with <i>p</i> -vinylPMSF; dialysed into 50 mM NaOAc buffer pH 5.9	Varying protein to precipitant ratio in the drop (2:1, 1:1, 1:2); partial density of PMSv seen (only -SO ₂ visible)

Appendix 4-3

Optimization of crystallization conditions: Round IV

PLATE	SBL PROTEIN	SAMPLE CONC. (mg/mL)	SCREEN	METHOD	REMARKS
<i>T#1</i> 15-05-2013	WT(PMSv)	12	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Single crystal harvested; data collected 28-05-13, density of PMSv seen
<i>T#2</i> 15-05-2013	L217C(PMSv)	12	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Single crystal harvested; data collected 28-05-13, partial density of PMSv seen (only -SO ₂ visible)
<i>T#3</i> 15-05-2013	Dpc217	12	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Single crystal harvested; data collected 28-05-13
<i>T#4</i> 15-05-2013	Bpc217	9.5	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9;	Single micro-crystal grown
<i>T#5</i> 15-05-2013	Dpc217(PMSv)	15	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9, 3 equiv Pd(NO ₃) ₂ added on set up	Single crystal harvested; data collected 28-05-13
<i>T#6</i> 15-05-2013	Bpc217(PMSv)	13	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9, 3 equiv Pd(NO ₃) ₂ added on set up	Crystal Cluster seen
<i>T#7</i> 15-05-2013	Dpc217•Pd	12	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits
<i>T#8</i> 15-05-2013	Bpc217•Pd	10.3	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits
<i>T#9</i> 15-05-2013	Dpc217(PMSv) (Pd)	15	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9; incubated with 3 equiv Pd(NO ₃) ₂ at 37 C for 1 h prior to set up	One single crystal
<i>T#10</i> 15-05-2013	Bpc217(PMSv) (Pd)	13	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9; incubated with 3 equiv Pd(NO ₃) ₂ at 37 C for 1 h prior to set up	

Appendix 4-3 contd.

Optimization of crystallization conditions: Round IV

PLATE	SBL PROTEIN	SAMPLE CONC. (mg/mL)	SCREEN	METHOD	REMARKS
<i>T#1</i> 16-05-2013	Dpc217	12	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Mainly precipitated
<i>T#2</i> 16-05-2013	Bpc217	9.5	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9;	Globular matter seen- precursor to crystals?
<i>T#3</i> 16-05-2013	Dpc217(PMSv)	15	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9, 3 equiv Pd(NO ₃) ₂ added on set up	Needle cluster
<i>T#4</i> 16-05-2013	Bpc217(PMSv)	13	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9, 3 equiv Pd(NO ₃) ₂ added on set up	Needle cluster
<i>T#5</i> 16-05-2013	Dpc217•Pd	12	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Globular matter seen- precursor to crystals?
<i>T#6</i> 16-05-2013	Bpc217•Pd	10.3	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Crystalline matter seen
<i>T#7</i> 16-05-2013	Dpc217(PMSv) (Pd)	15	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9; incubated with 3 equiv Pd(NO ₃) ₂ at 37 C for 1 h prior to set up	
<i>T#8</i> 16-05-2013	Bpc217	13	Optimized PACT Screen	dialysed into 50 mM NaOAc buffer pH 5.9; incubated with 3 equiv Pd(NO ₃) ₂ at 37 C for 1 h prior to set up	Several hits
<i>T#9</i> 07-06-2013	Bpc217	14	Optimized PACT Screen (around E3 E4)	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits, density for Bpc seen at occupancy 0.7
<i>T#10</i> 07-06-2013	Bpc217 (PMSv)•Pd	12.4	Optimized Pact Screen (around E3 E4)	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits

Appendix 4-4

Optimization of crystallization conditions: Round V

PLATE	SBL PROTEIN	SAMPLE CONC. (mg/mL)	SCREEN	METHOD	REMARKS
<i>T#1</i> 07-06-2013	Bpc217	12	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	
<i>T#2</i> 07-06-2013	Bpc217•Pd	14	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	
<i>T#3</i> 07-06-2013	Bpc217 (PMSv)•Pd	12.4	Optimized Index Screen	dialysed into 50 mM NaOAc buffer pH 5.9	
<i>T#4</i> 10-06-2013	Bpc217	12	Optimized NaCl/NaOAc Screen	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits, no density for Bpc seen
<i>T#5</i> 10-06-2013	Bpc217	12	Optimized NaCl/NaOAc Screen	dialysed into 50 mM NaOAc buffer pH 5.9, crystal soaked with PMSv substrate	No density for Bpc, density for phenyl+2C atoms visible
<i>T#6</i> 10-06-2013	Bpc217•Pd	14	Optimized NaCl/NaOAc Screen	dialysed into 50 mM NaOAc buffer pH 5.9	
<i>T#7</i> 10-06-2013	Bpc217 (PMSv)•Pd	12.4	Optimized NaCl/NaOAc Screen	dialysed into 50 mM NaOAc buffer pH 5.9	

Appendix 4-5

Optimization of crystallization conditions: Round VI

PLATE	SBL PROTEIN	SAMPLE CONC. (mg/mL)	SCREEN	METHOD	REMARKS
<i>T#1</i> 07-06-2013	SBL- WT(PMSv)	35	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9	Single crystal, density for a sulfonyl group on Ser221 (attachment visible at a contour level of 0.8)
<i>T#2</i> 07-06-2013	SBL-Bpc217	30	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9, soaked with PMSv substrate for 10 min	Single crystal analysed, density for Bpc visible pointing towards Ser221, but no density for PMSv seen
<i>T#3</i> 07-06-2013	SBL-Bpc217	30	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9, soaked with PMSv substrate overnight	Partial density for maybe 1 atom near Ser221, no clear density for Bpc seen
<i>T#4</i> 07-06-2013	SBL- Bpc217•Ni	12.4	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9	Several hits, clear density for Ni in several crystals seen
<i>T#5</i> 10-06-2013	SBL-Bpc217 (PMSv)•Ni	12	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9	very slow crystal growth at pH 8.5
<i>T#6</i> 10-06-2013	SBL- Bpc217(PM Sv)•Pd	14	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9	very slow crystal growth at pH 8.5
<i>T#7</i> 10-06-2013	SBL-Bpc217 (PMSv)•Pd	12.4	0.1 M AS, 0.1 M bis-tris buffer pH 8.5	dialysed into 50 mM NaOAc buffer pH 5.9	

Chapter 5

Design of β -lactamase-based artificial metalloenzymes

5.1 Introduction

Previously, metal binding sites were introduced in the active site of a model enzyme, subtilisin from *Bacillus lentus* (SBL), to generate artificial metalloenzymes SBL-Dpc217•Pd, SBL-Bpc217•Pd and SBL-Bpc217•Ni (Chapter 2 Section 2.2.2.4). It was also found useful to vary the location of the metal binding amino acid to a different site on the same protein, resulting in the creation of holo-enzyme SBL-Dpc156•Pd (Chapter 3 Section 3.2.6). These enzymes were then employed in oxidative Heck cross-coupling reactions. Furthermore, prior work in our group on catalysis of Suzuki-Miyaura cross-coupling reaction using SBL-based artificial metalloenzymes did not give the desired results (See Chapter 1 Section 1.6). Given the limited success in the catalysis trials using SBL as model protein which has a relatively shallow active site, we decided to utilize a different protein scaffold (β -lactamase) that could provide a pre-evolved enzyme pocket for substrate localisation and metal binding. Furthermore, the rational design of artificial metalloenzymes for chemical cross-coupling reactions using our previously established ‘tag-and-modify’ methodology will be described in this chapter.

We aim at creating an artificial metalloenzyme for the Suzuki-Miyaura cross-coupling reaction between organoboronic acids and aryl halides (or triflates), involving the oxidative addition, transmetalation and reductive elimination sequences (Chapter 1 Section 1.5.2). The utility and popularity of the Suzuki-Miyaura cross-coupling reaction for

efficient synthesis of biaryl compounds is undebatable.^{1,2} Previously, Watanabe and co-workers demonstrated the Suzuki-Miyaura cross-coupling of 4-iodoaniline and phenylboronic acid by employing ferritin-based metal nanoparticles (Pd(allyl)•apo-Frs).³ The turnover frequencies (TOF) were determined to be 3500 ± 400 [TOF defined as product (mol) per Pd(allyl)•apo-Fr per hour], which were about half that of the free catalyst in the solution, indicating that the protein environment influences the activity of the metal catalyst. A properly designed artificial metalloenzyme for the Suzuki-Miyaura cross-coupling would add this important cross-coupling reaction to the repertoire of artificial metalloenzymes.

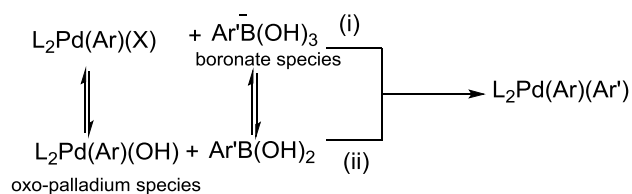
5.2 Results and discussion[†]

5.2.1 Design of system

Arylboronic acids are widely used substrates in synthetic organic chemistry, mainly due to their oxidative and water stability, and generally non-toxic nature. Moreover, these serve as coupling partners in several C-C cross-coupling reactions, including the Suzuki-Miyaura cross-coupling reaction. The transmetalation step that involves transfer of an organic group from organoboron compound to palladium complex, is often proposed to be the rate limiting step along the catalytic pathway for this reaction (Scheme 5-1).⁴ This step is facilitated by the presence of a base that promotes the formation of boronate species. These boronates are proposed to be the reactive species in the transmetalation step, that can enable the transfer of an aryl group on to the palladium halide species *via* the boronate pathway (Scheme 5-1 (i)).^{5,6} However, it has been shown that the reaction of the oxo-

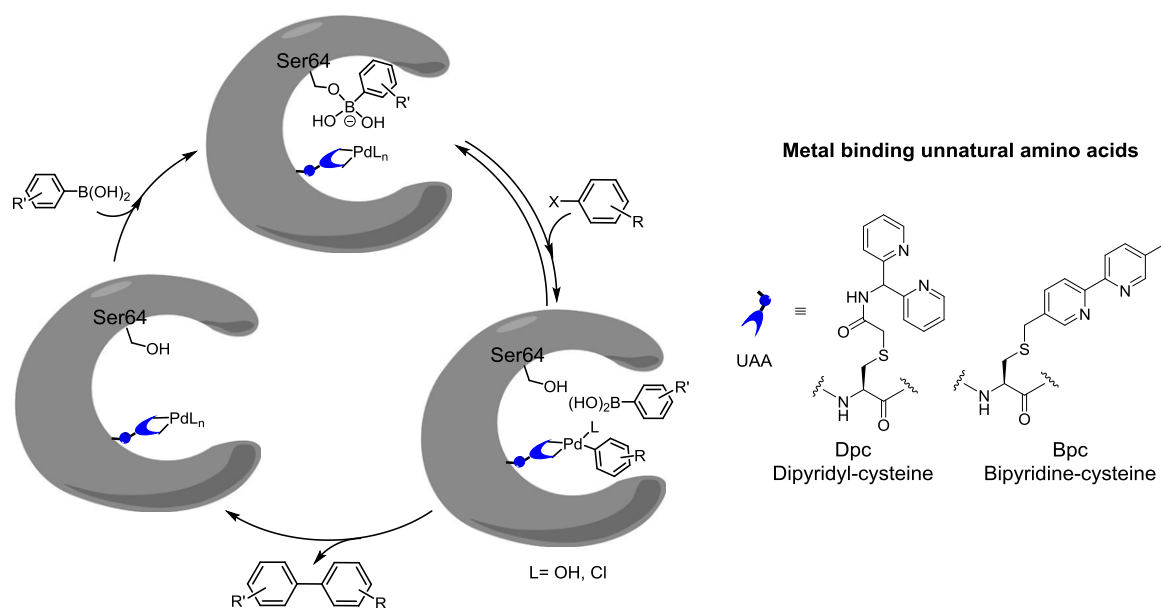
[†]Future work using the protein constructs reported in this chapter is being carried on by Dr. Maria Espling in our group. Experimental results that were obtained by Maria Espling are indicated in the body of this chapter with reference to this footnote and her contributions are stated explicitly in the experimental sections of this chapter.

palladium species with the aryl boronic acid is kinetically favoured (Scheme 5-1 (ii)).^{4,7,8} It was shown that the oxo-palladium pathway is in fact faster than the boronate pathway by four orders of magnitude.⁴



Scheme 5-1. Two proposed transmetalation pathways (i) boronate pathway (ii) oxo-palladium pathway

Phenylboronic acids are competitive and reversible inhibitors of serine proteases⁹ and serine β -lactamases.¹⁰⁻¹² For our purposes, a serine hydrolase capable of localising arylboronic acids was considered as a suitable candidate (Scheme 5-2). We reasoned that the enzyme active site would provide an activating micro-environment by hosting a metal catalyst. Moreover, the selective and reversible binding of the arylboronic acid and aryl halide substrates with the enzyme would serve to localise the substrates. Thus, we envisioned that the well-established ‘oxyanion hole’ of a serine hydrolase could help in driving the transmetalation step along the catalytic pathway. The design of the proposed metalloenzyme was based on the combination of rational design and chemical modification strategies to ensure the site-selective installation of metal binding amino acids.



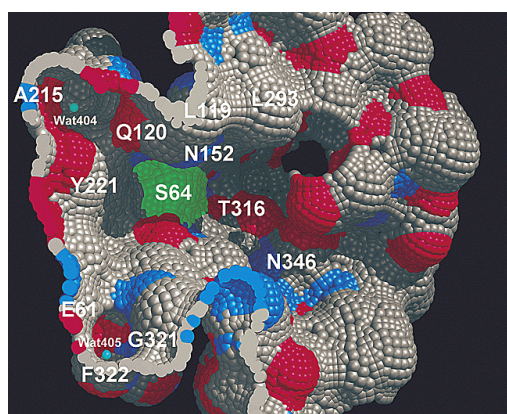
Scheme 5-2. Representative model of the proposed metalloenzyme depicting installation of an unnatural amino acid (UAA) and localization of the cross-coupling substrates within the active site; chemical structures of the designed metal binding UAAs are shown

5.2.1.1 Rational design of AmpC-based artificial metalloenzymes

β -lactamases are a class of enzymes which cause the hydrolysis of the lactam bond in β -lactam antibiotics, rendering them inactive. In the first step, the catalytic Ser64 attacks the lactam carbonyl carbon, opening the lactam ring to give an acylated intermediate. In the second step, this acyl intermediate is attacked by an activated water molecule, releasing the hydrolysed β -lactam. Catalytic roles have been confirmed for Ser64, Lys67, Tyr150 and Lys315.¹³ Additionally, participation of other residues such as Asn152, Thr316 and Ser/Ala318 have also been suggested.¹³ AmpC, a class C β -lactamase, was chosen as the model enzyme mainly because of the enhanced nucleophilicity of the active site serine residue (Ser64) and its high affinity for phenylboronic acids. Moreover, availability of extensive structural and functional information on these enzymes makes them a particularly attractive scaffold for employing rational design methodologies. The cysteine-free wild-type AmpC protein allows for the chemogenetic tailoring of an existing protein

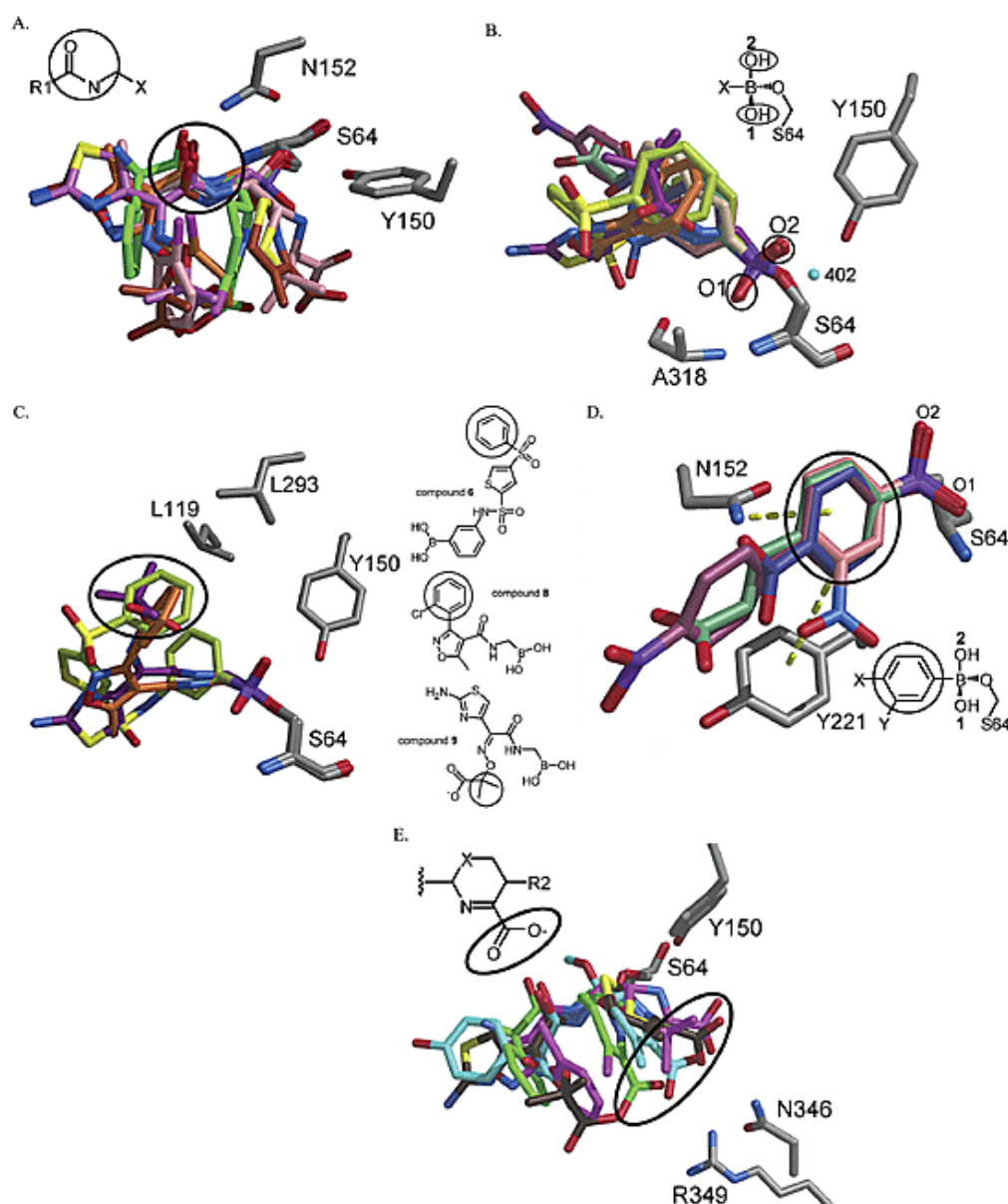
scaffold by the installation of a metal binding centre at different locations within its active site. Bipyridine- and dipyridyl-based compounds have been used as ligands to mediate efficient aqueous Suzuki-Miyaura cross-coupling reactions.¹⁴⁻¹⁷ *N,N*-chelating amino acids, Dpc and Bpc, previously installed in the SBL scaffold (Chapter 2) were considered ideal to bind the metal catalyst in the active site.

AmpC β -lactamase from *E. coli*,¹⁰ has a mixed α/β secondary structure, with nine anti-parallel stranded β sheets in the middle flanked by a row of three α - helices on one side and helical domain with eleven helices on the other side. Several phenylboronic acids have been found to be effective transition state inhibitors of β -lactamases¹⁰ and the structural data on the tetrahedral intermediates formed with the active site serine is well documented.^{11,12} Binding site ‘hotspots’ on the AmpC were identified using experimental and computational approaches (Scheme 5-3).^{10-12,18} This study involved structure-based probing of the active site to assess the ability of a particular enzyme site to bind different functionalities (such as amide or aryl groups).



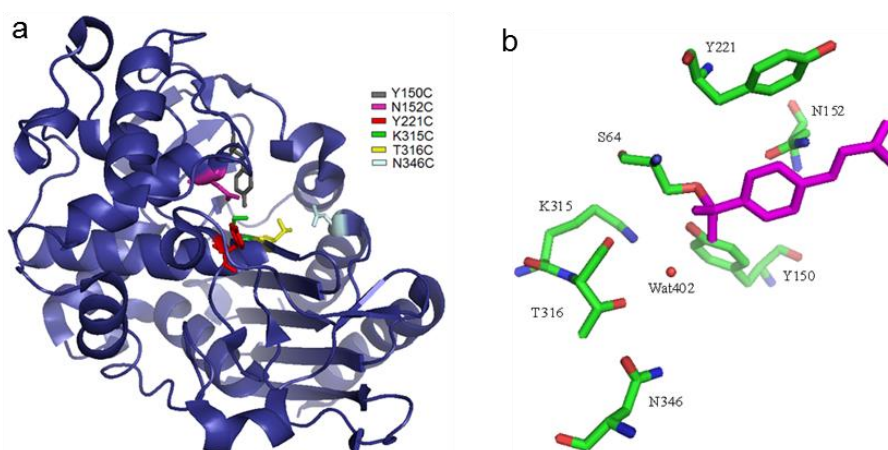
Scheme 5-3. (a) Molecular surface of the active site region of AmpC. The catalytic residue Ser64 is colored green. Several conserved residues are labeled. [Reprinted (adapted) with permission from Powers, R. A.; Shoichet, B. K. J. Med. Chem. 2002, 45, 3222. Copyright © 2002 American Chemical Society]

Among the experimentally determined binding sites in AmpC¹² was the amide-binding site identified by superimposing all amide containing substrates for this enzyme, recognised closely by N152 and Y150 residues (Scheme 5-4a). Additionally, Y150 and A318 were found to recognise the hydroxyl oxygens on the boron atom, one of which also showed interactions with the conserved water molecule Wat402 (Scheme 5-4b). This is the presumed deacylating water that gives the direction of the hydrolytic attack for the acyl adduct in the native hydrolysis reaction. This site also corresponds to the ‘oxyanion hole’ or ‘electrophilic hole’ of AmpC. We reasoned that these residues may play an important role in the activation of the boronic acid and contribute towards the activity of the variant enzyme towards the cross-coupling reaction. Furthermore, N152 and Y221 were found to make quadrupole-dipole and quadrupole-quadrupole interactions with the aryl ring present off the boronic acid group, as judged from the geometry of approach in the crystal structures (Scheme 5-4d).¹² We speculated that mutations at these residues may also prove to be useful for the transmetalation step in the catalytic cycle of this coupling reaction. K315, T316 and N346 were chosen because of their suggested role in the catalysis of the native substrates.¹³



Scheme 5-4. Binding sites identified experimentally on AmpC. (a) An amide binding site identified by superimposing all amide-containing compounds; the amide group is circled in black. (b) Two hydroxyl binding sites identified by superimposing all boronic acid inhibitors. The hydroxyl sites are circled and labeled O1 and O2. (c) A hydrophobic binding site identified by superimposing the relevant structures; the hydrophobic binding site is composed of residues Leu119 and Leu293, and the hydrophobic portions of the compounds interacting with these residues are circled. (d) An aryl binding site identified. Dashed yellow lines indicate the quadrupole interactions with Asn152 and Tyr221. (e) The carboxylate binding site identified. The C3(4) carboxylate group is circled. Nitrogen atoms are colored blue, oxygens are colored red, sulfurs are colored yellow, chlorines are colored magenta, borons are colored purple, and carbons are colored gray, except for carbon atoms of the different ligands. [Reprinted (adapted) with permission from Powers, R. A.; Shoichet, B. K. *J. Med. Chem.* 2002, 45, 3222. Copyright © 2002 American Chemical Society]

We reasoned that the installation of the metal centre at these highlighted sites could provide close proximity and favourable orientations of the cross-coupling substrates with respect to the metal. Based on the locations and distances of these residues with respect to active site Ser64 bonded with a phenylboronic acid substrate,¹² Y150, N152, Y221, K315, T316 and N346 were selected as target residues for site directed mutagenesis (Scheme 5-5).



Scheme 5-5. (a) Whole structure of AmpC-WT (in blue) showing the relative locations of the target residues for mutagenesis; side chains of the residues are shown in different colours as indicated, (b) close view of the relative orientations of the target residues for site-directed mutagenesis with respect to active site Ser64 bound to a phenylboronic acid (shown in pink); PDB ID: 1KE0¹²

5.2.2 Subcloning and site-directed mutagenesis

Efforts to tailor the enzyme active site were undertaken by using two different mutagenesis approaches to install the desired point mutations on AmpC, namely site-directed mutagenesis and Overlap Extension PCR. In essence, both these techniques rely on synthetic oligonucleotides which contain the desired mutation and are used to extend along the length of the parent base sequence to yield a mutagenized gene sequence. Plasmid vector p0G0295 bearing *ampC* gene for AmpC β -lactamase was borrowed from the Shoichet group, University of California, San Francisco (UCSF) (See Experimental Section 5.5.1).

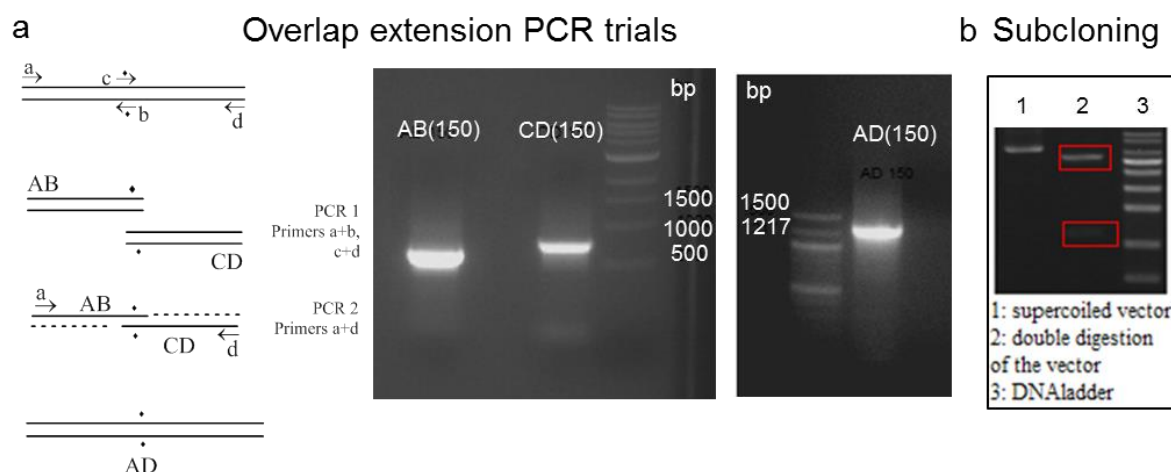
By employing the QuikChange® site-directed mutagenesis protocol using a commercial kit, the desired mutation can be introduced in a single PCR run starting with a dsDNA plasmid containing target gene isolated from *dam*⁺ *E. coli* strain. Two complementary oligonucleotides with desired base substitution are used for PCR with the DNA template to obtain a mutated plasmid with staggered nicks. This is followed by DpnI digestion (an endonuclease specific for methylated and hemi-methylated DNA) to enrich mutated plasmid as against the parent DNA plasmid. Several attempts to generate genes encoding for single cysteine variants were made using this approach. While the PCR reaction proceeded smoothly, subsequent transformation in competent cells was unsuccessful as no bacterial colonies could be obtained.

A range of conditions were tested in order to obtain the desired results, like employing different annealing temperatures, using additives like DMSO or different *E. coli* strains, DNA polymerases etc. (See Experimental Section 5.5.2). In general, our mutagenesis trials were found to be problematic because the vector p0G0295 (that carries tetracycline resistance) was found to be incompatible with the competent cells (XL1 or XL-10 supercompetent cells). Screening of different cell strains (JM109, TOP10) did not lead to the desired results. Later, it was realised that the commercial QuikChange® protocol is only tested and optimized for its use with certain cell strains (XL1 or XL-10 supercompetent cells) and this procedure may not be generally suitable for use with alternative competent cells.

Thus, we decided to use the conventional Overlap Extension PCR technique¹⁹ (OE-PCR) for site-directed mutagenesis (Scheme 5-6a). In the Overlap Extension PCR reaction, the mutant gene AD can be synthesized starting with four suitably designed primer sets

(primers a-d) in two rounds of PCR cycles. The ‘flanking primers a and d’ bind at the beginning and the end of the target gene, while the ‘internal primers b and c’ containing the modified base information, bind to the site where mutation is desired. The first PCR amplifies the DNA fragment from the mutation site to one end of the target gene, yielding PCR products with overlapping regions. After purification of both the PCR products, the two amplicates are mixed and the second PCR is started. The expected end product is the mutagenized target gene incorporating desired mutation and two appropriate restriction enzyme sites for further engineering purposes. Thus, by employing the standard molecular cloning procedures for OE-PCR,^{20,21} mutant genes encoding for AmpC variants Y150C, N152C, Y221C, K315C and T316C were synthesized as confirmed by visualization of the genes by DNA-Agarose gel electrophoresis (See Experimental Section 5.5.2). The prepared genes were digested with NdeI and BamHI-HF enzymes to introduce restriction sites and these double digestion reactions were found successful by DNA gel analysis.

Unfortunately, subsequent ligation reactions to insert the gene in the p0G0295 vector did not lead to the desired results. Moreover, the N346C mutant gene could not be synthesized by this method with the primer pair used for this reaction as no amplification could be seen by the DNA gel. In our hands, the vector p0G0295 was found to be quite difficult to manipulate for mutagenesis reactions. Thus, the need for subcloning the gene into a commonly used pET vector was realised.



Scheme 5-6. (a) Overlap extension PCR trial to construct Y150C mutant gene; general scheme (left) DNA gel visualization of PCR products ‘AB’ and ‘CD’ (middle) and PCR product ‘AD’ (right); (b) DNA gel of pET-28a(+)-ampC construct[†]

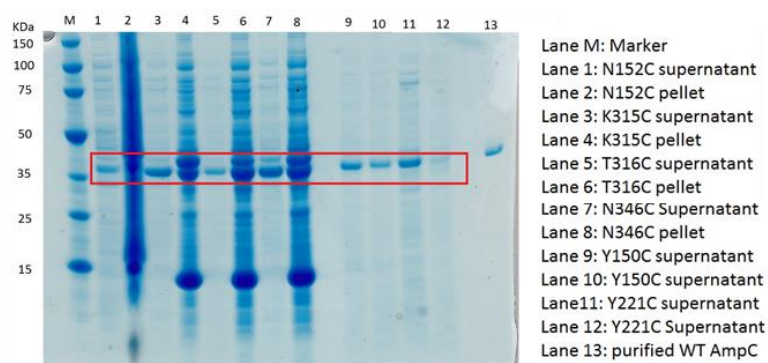
For ease of mutagenesis experiments, *ampC* gene sub-cloned into pET-28a(+) between NdeI and BamHI restriction sites was ordered through GenScript USA Inc. Double digestion of the vector bearing *ampC* gene confirmed the success of the subcloning experiment (Scheme 5-6b). The plasmid pET-28a(+)-ampC was subjected to two rounds of site-directed mutagenesis to substitute two cysteine residues occurring in the N-terminal of the protein sequence in order to obtain a cysteine-free wild-type (See Experimental Section 5.5.3). With *ampC* gene inserted in a convenient pET vector in hand, mutagenesis experiments led to the generation of the desired single cysteine variants, Y150C, N152C, Y221C, K315C, T316C and N346C (See Experimental Section 5.5.4).

5.2.3 Protein expression and purification: optimization

AmpC-WT could be expressed from a p0G0295-ampC plasmid construct bearing a *ampC* gene. The *ampC* gene in p0G0295 vector is under the control of a temperature-sensitive

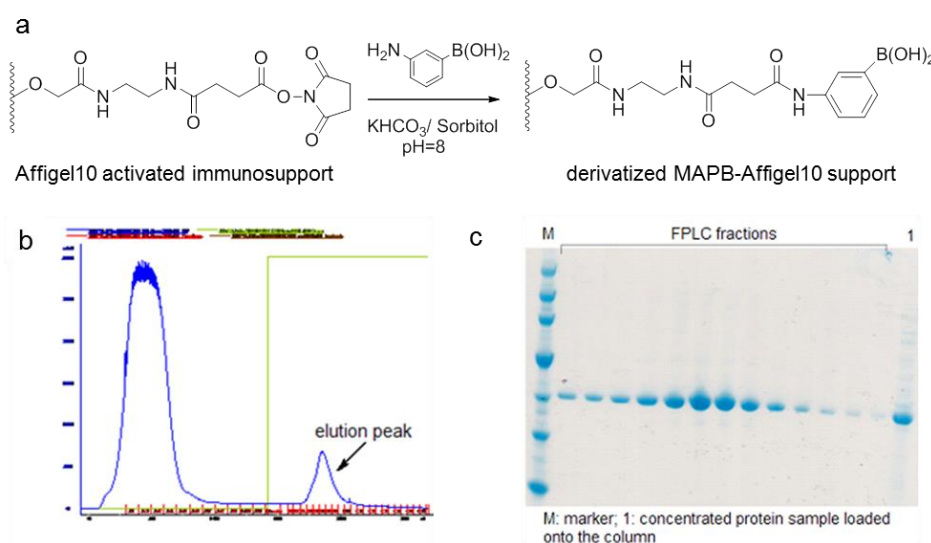
[†]DNA gel of sub-cloning of *ampC* gene into pET-28a(+)-ampC construct was ordered through GenScript Inc. USA

repressor (c1857) which is rendered inactive on raising the temperature to 42 °C, thereby inducing expression of the protein (See Experimental Section 5.5.1). The enzyme is released out into the cytosol on stressing the cells at higher temperatures, thereby enabling isolation of soluble protein from the supernatant (culture medium). However after subcloning, the expression of the AmpC-variants in BL21(DE3) cells required optimization. The expression of AmpC variant proteins from pET-28a(+)-ampC in *E. coli* BL21 (DE3) strain was induced by Isopropyl β -D-1-thiogalactopyranoside (IPTG). Although preliminary expression trials showed that the pET-28a(+) system provided a suitable vector enabling convenient over-expression of AmpC β -lactamase in *E. coli*, it was important to ensure that good amounts of soluble protein could be obtained for our purposes. Pleasingly, significant protein levels were observed in both supernatant (culture medium ~800 mL) and cell pellet for each variant (Scheme 5-7). The secretion of the desired protein to the culture medium helped to get rid of the unwanted proteins and other cytoplasmic contaminants. We reasoned that the isolation of the protein from the supernatant would minimize purification efforts as compared to that from cell lysate, so the supernatant was concentrated to a reduced volume and used further for purification trials.



Scheme 5-7. SDS PAGE gel for expression of AmpC variant enzymes using pET-28a(+)-ampc plasmid construct

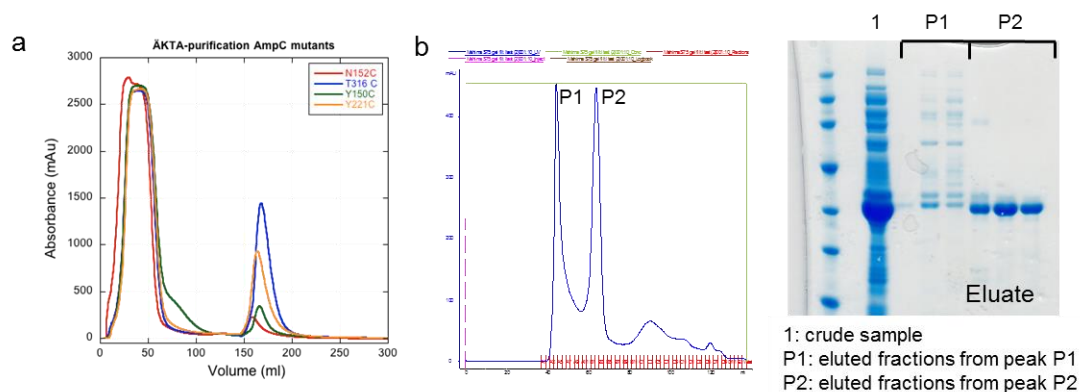
An efficient protein purification method is important for ensuring access to good amounts of proteins for subsequent chemical modification and catalysis experiments. Previously, it has been shown that phenylboronic acid can serve as a reversible and inert active site affinity tag for purification of AmpC β -lactamase.^{10,22} Considering this, activated immunoaffinity support Affi-Gel10 (Bio-Rad) was coupled to *m*-aminophenyl boronic acid (MAPB), thereby providing a high-affinity chromatography method for these proteins (Scheme 5-8a). The purification of AmpC-WT was accomplished by employing an MAPB-Affi-Gel10 based affinity column, where the protein was selectively bound to the phenylboronic acid tag and subsequently eluted by passing 0.5 M boric acid buffer pH 7, over the column. Thus, an effective one-step purification method was optimized for the isolation and purification of the AmpC proteins from the supernatant concentrate (Scheme 5-8b and c). Typically, 15-20 mg protein could be purified from 800 mL culture medium.



Scheme 5-8. (a) Coupling reaction of Affi-Gel10 immunosupport derivatised with MAPB ligand, (b) FPLC trace for purification of AmpC-K315C using MAPB-Affi-Gel10 column (c) SDS-PAGE gel for FPLC fractions collected after one-step purification of AmpC-K315C

The purification of AmpC-WT, AmpC-K315C and AmpC-N346C using an MAPB-Affi-Gel10 column was quite effective. However, other variant enzymes did not bind very well

with the column under the same conditions. Limited binding of AmpC-Y221C and AmpC-N152C was evident from the unbound proteins observed in the flow-through fractions during purification (Scheme 5-9a). We speculated that certain active site mutations could disrupt the interactions of MAPB with the enzyme pocket leading to loss in binding efficiency. In this context, pre-incubation of the protein sample with the solid support or purification at a reduced flow-rate would allow for maximum contact time for MAPB binding, which proceeds *via* the formation of a covalent intermediate with the active site serine residue. To avoid purification problems due to different binding affinities of AmpC variants, we employed size-exclusion chromatography as an alternative purification strategy. This would also circumvent any possible contamination by MAPB in future catalysis experiments caused due to possible column degradation. We were pleased to find that the purification of AmpC-N152C by gel filtration proved to be an effective method, and purity of the protein was confirmed by SDS-PAGE analysis (Scheme 5-9b).



Scheme 5-9. (a) Overlay of FPLC traces for purification of AmpC variants using MAPB-affigel column showing different binding affinity of different active site AmpC mutant proteins (b) FPLC trace (left) and SDS PAGE gel (right) for purification of AmpC-N152C; Figure (a) overlay traces were borrowed from by Dr. Maria Espling

Thus purified, wild-type AmpC (AmpC-WT) and AmpC variants (N152C, Y221C, K315C, T316C and N346C)[†] were analysed by SDS-PAGE and bicinchoninic acid (BCA) assay to confirm their purity, homogeneity and concentration (See Experimental Section 5.5.7). The identity of AmpC-WT and active site variants was confirmed by LC-MS analysis (Table 5-1). Further, the presence of a single cysteine in AmpC variants N152C, Y221C, K315C, T316C and N346C was confirmed by Ellman's assay (Table 5-1). Notably, LC-MS analysis of purified AmpC-Y150C showed a single peak corresponding to a +45 adduct and Ellman's test showed no corresponding increase in mass. This reveals that either there is a post translational modification on cysteine, or the cysteine in this variant occupies a hindered position preventing its accessibility by Ellman's reagent.

Table 5-1. Characterization of AmpC-WT and AmpC variants by LC-MS analysis and Ellman's test

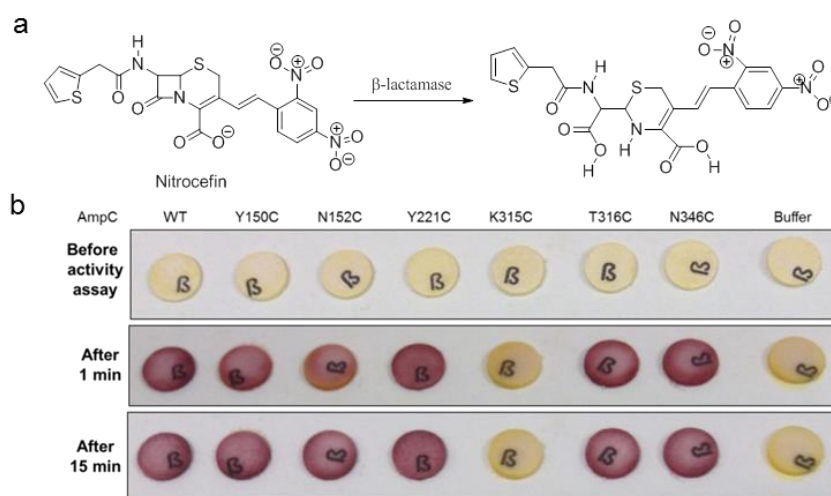
Entry	AmpC protein	Expression levels	Mass calculated (Da)	Mass found (Da)	Ellman's adduct (Da)
1	WT	good	39551	39551	No reaction
2	Y150C	low	39491	39536 ^a	No reaction
3	N152C	low	39540	39542	39740
4	Y221C	good	39491	39494	39687
5	K315C	good	39526	39527	39726
6	T316C	good	39553	39554	39753
7	N346C	good	39547	39555	39754

^a+45 mass adduct was observed by LC-MS analysis

Nitrocefin is a chromogenic cephalosporin substrate routinely used for rapid detection of β -lactamase activity. Hydrolysis of this native substrate (yellow) can be detected by an immediate colour change to red (Scheme 5-10a). β -lactamase activity of AmpC-WT and its variants (except AmpC-K315C) was confirmed using this qualitative activity test

[†] Purified proteins Y150C, N152C, Y221C and T316C were borrowed from Dr. Maria Espling

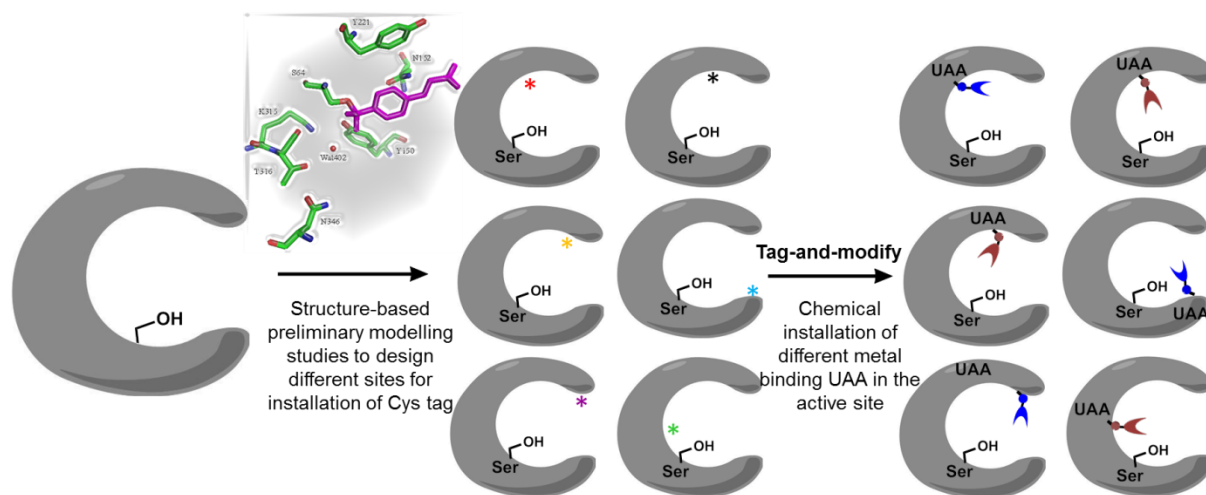
(Scheme 5-10b). Active site residue K315 plays an important role in the hydrolytic mechanism. Hence, point mutation at this site may have resulted in this loss of activity.



Scheme 5-10. (a) Hydrolysis of cephalosporin substrate, nitrocefina, by AmpC β -lactamase (b) Nitrocefina test to detect β -Lactamase activity in AmpC variants

5.2.4 Chemogenetic design: ligand incorporation

Six active-site AmpC variants and two dipyrrolyl- or bipyridine-based ligands **Lx** (to install metal-binding amino acids, Bpc and Dpc) were prepared using genetic and chemical methods. Having successfully confirmed the identity and purity of AmpC variants, we proceeded with the installation of the metal-binding amino acids to obtain the various possible metal-free apo-enzymes (Scheme 5-11). This would ensure the positioning of metal catalyst at different locations in the active site relative to the nucleophilic serine residue.

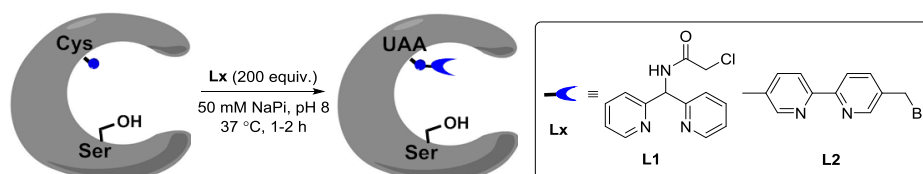


Scheme 5-11. General strategy for rational design and creation of AmpC-based artificial metalloenzymes

The direct alkylation of these single cysteine variants was carried out with an excess of dipyrrolyl- or bipyridine-based ligands **Lx** to ensure a rapid and complete reaction (Table 5-2). Consistent with our previous observations, AmpC-Y150C (with mass corresponding to +45 adduct) showed no reaction with either of the halogenated ligands (Table 5-2, entry 1-2). Pleasingly, direct alkylation of AmpC-N152C, AmpC-Y221C, AmpC-T316C and AmpC-N346C proved most successful and corresponding apo-enzymes were obtained in >95% conversions by LC-MS analysis (Table 5-2, entry 3-10). Alkylation with **Lx** at position 315 proceeded quite slowly, perhaps due to relatively hindered location of this residue inside the enzyme active site (Table 5-2, entry 11). On the other hand, the reaction of a large excess of **L2** with AmpC-Y221C and AmpC-K315C showed multiple adducts corresponding to double and triple modification perhaps due to reaction with the surface residues (Table 5-2, entry 5, 11). In this regard, although reaction at the nucleophilic sulfhydryl group of cysteine residues is the most common method to modify proteins, the competing alkylation at surface exposed nucleophilic lysine and histidine residues cannot be disregarded for macromolecular systems such as proteins. The reaction with **L2** for variant AmpC-Y221C could be further optimized by reducing the number of equivalents

of the ligand and shorter reaction times. In addition to this, AmpC-T316C and AmpC-N346C showed a DTT adduct (+154) with **L1** (Table 5-2, entry 6). This can be optimized by desalting excess DTT after reduction of the disulfides.

Table 5-2. Chemical installation of UAAs, Bpc and Dpc, in AmpC variants^a



Entry	AmpC protein	Lx	Mass calculated	Mass found	% Conversion ^b
1	Y150C-Bpc	L2	39672	-	-
2	Y150C-Dpc	L1	39716	-	-
3	N152C-Bpc	L2	39723	39726	>95%
4	N152C-Dpc	L1	39766	39767	>95%
5	Y221C-Bpc	L2	39736	39736 ^c	>95%
6	Y221C-Dpc	L1	39674	39676	>90%
7	T316C-Bpc	L2	39736	39736	>95%
8	T316C-Dpc	L1	39779	39933 ^c	35%
9	N346C-Bpc	L2	39737	39738	>95%
10	N346C-Dpc	L1	39782	39933 ^d	>95%
11	K315C-Bpc ^e	L2	39708	39711 ^c	>95%
12	WT	L1	-	No reaction	-

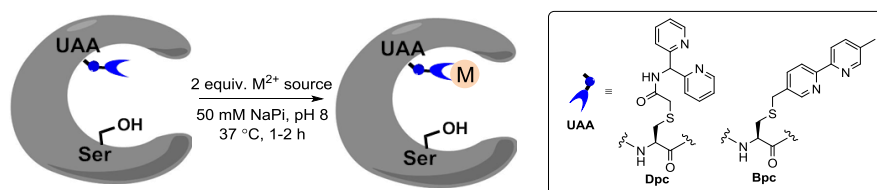
^aAmpC variants were pre-treated with DTT (10 equiv.) and reacted with 200 equiv. of **Lx** at 37 °C for 1-2 h, ^b % conversions were calculated based on LC-MS analysis, ^cmultiple peaks were observed, ^dmass corresponding to DTT adduct were observed by LC-MS analysis, ^eonly 35% conversion after 3.5 h; complete conversion after 16 h

In general, the AmpC variants N346C, N152C, T316C and Y221C were successfully used for incorporation of unnatural metal binding amino acids, which could be used for site selective metal complexation. The ESI-MS confirming the identities of these protein conjugates would provide reference for future work using these metalloenzymes (See Experimental Section 5.5.8).

5.2.5 Metal complexation

Metal complexation was attempted with AmpC-Dpc221, AmpC-Bpc316, AmpC-Dpc346 and AmpC-Bpc346 protein conjugates. Both nickel and palladium complexation was carried out using Na_2PdCl_4 and $\text{Ni}(\text{OAc})_4$ as Pd(II) and Ni(II) sources, respectively. While AmpC-Dpc221 showed about 20% palladium loading, nickel loading on this protein could not be observed by LC-MS analysis (Table 5-3, entry 1-2). AmpC-Bpc316 metal complexation with 2 equiv. of the respective metal sources showed only 29% palladium loading and 44% nickel loading by LC-MS analysis (Table 5-3, entry 3-4). Incubation of AmpC-Bpc346 with Na_2PdCl_4 resulted in only partial uptake of palladium (Table 5-3, entry 6). Promising results were obtained by employing AmpC-Dpc346 protein, shown to have a mass corresponding to a DTT adduct, which gave >95% conversion to the corresponding metalloenzyme species (Table 5-3, entry 5).

In these metal loading attempts, the protein solutions were dialysed extensively to get rid of any excess ligand that may be remaining in the solution. Size exclusion chromatography could be performed prior to metal complexation to ensure absence of any exogenous ligand. Nonetheless from these trials, the palladium loading on AmpC-Dpc221, AmpC-Dpc346 and AmpC-Bpc346 and nickel loading on T316C seems promising. Non-specific metal binding to surface exposed nucleophilic amino acids has also been previously acknowledged.^{23,24} Such non-specific coordination of the palladium to nucleophilic residues on the protein surface could generate a large number of isoforms. Thus, the control reactions with 2-3 equiv. of Pd(II) precursor with the AmpC-WT were carried out. Non-specific palladium binding was not observed by LC-MS analysis even after 3 h incubation at 37 °C (Table 5-3, entry 7).

Table 5-3. Metal complexation reaction with AmpC proteins^a

Entry	AmpC protein	M	Mass calculated	Mass found	% Conversion ^b
1	AmpC-Dpc221	Pd	39822	39821 ^c	57%
2	AmpC-Dpc221	Ni	39775	-	-
3	AmpC-Bpc316	Pd	39841	39841	29%
4	AmpC-Bpc316	Ni	39794	39793	44%
5	AmpC-Dpc346	Pd	39887	40037 ^d	>95%
6	AmpC-Bpc346	Pd	39843	39842	57%
7	AmpC-WT	Pd	39551	-	-

^aAmpC proteins were incubated with 2 equiv. of M²⁺ precursor at 37 °C for 1-2 h, ^b%

conversions were calculated based on LC-MS analysis, ^cmultiple peaks were observed,

^dmass corresponding to DTT adduct observed by LC-MS analysis

5.3 Conclusions

Structural information can rationally guide the design of new metalloenzymes. In this preliminary work, a structure guided approach was taken for chemogenetic design of AmpC-based artificial metalloenzymes. Mutagenesis experiments led to a library of active site variants and convenient expression and protein purification procedures were established. Of the six distinct active site variants screened, K315C and Y150C gave unsatisfactory results in chemical modification reactions. While the identity of the Y150C variant expressed seems ambiguous because of the +45 adduct detected consistently by LC-MS and its unreactivity towards Ellman's and other alkylating reagents. K315C mutation showed that some key active site interactions may be disrupted, leading to loss of hydrolytic activity of this variant. Alkylation at position 315 proceeded quite slowly, although complete alkylation could be seen after an overnight reaction. The location of this residue is deeper inside the active site and the hindered location of this residue could explain the slow reaction rates. In this context, the Dha strategy could be employed for

obtaining chemically modified K315C variants (Chapter 2 Section 2.2.2.3). Nevertheless, variant proteins N152C, Y221C, T316C and N346C could be expressed and purified in good levels and reaction conditions for chemical incorporation of unnatural amino acids for these proteins have been optimized.

We propose that the AmpC enzyme active site would be sufficiently large to accommodate the transition metal complex and the two coupling partners, thereby providing an activating environment for cross-coupling reactions. Initial results have been obtained for metal complexation by these apo-enzymes. The local environment of the AmpC active site may also be responsible for different metal affinities as compared to SBL protein, as the steric and electronic properties of the two proteins are different. This reveals that further optimization for metal loading experiments, and perhaps longer incubation times with the desired metal would ensure synthesis of homogeneous AmpC-metalloenzymes. Further, metal loading for AmpC-Bpc346 (with Pd) and AmpC-Bpc316 (with Ni) seems most promising and should be further explored. These metalloenzymes can be further screened for their activity towards Pd- and Ni-catalysed bisubstrate cross-coupling reactions.

5.4 References

- (1) Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, 95, 2457.
- (2) Kotha, S.; Lahiri, K.; Kashinath, D. *Tetrahedron* **2002**, 58, 9633.
- (3) Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, 130, 10512.
- (4) Carrow, B. P.; Hartwig, J. F. *J. Am. Chem. Soc.* **2011**, 133, 2116.
- (5) Braga, A. A. C.; Morgon, N. H.; Ujaque, G.; Maseras, F. *J. Am. Chem. Soc.* **2005**, 127, 9298.
- (6) Braga, A. A. C.; Ujaque, G.; Maseras, F. *Organometallics* **2006**, 25, 3647.
- (7) Amatore, C.; Jutand, A.; Le Duc, G. *Chem. Eur. J.* **2011**, 17, 2492.
- (8) Schmidt, A. F.; Kurokhina, A. A.; Larina, E. V. *Russ. J. Gen. Chem.* **2011**, 81, 1573.

- (9) Stoll, V. S.; Eger, B. T.; Hynes, R. C.; Martichonok, V.; Jones, J. B.; Pai, E. F. *Biochemistry* **1998**, 37, 451.
- (10) Usher, K. C.; Blaszcak, L. C.; Weston, G. S.; Shoichet, B. K.; Remington, S. J. *Biochemistry* **1998**, 37, 16082.
- (11) Weston, G. S.; Blázquez, J.; Baquero, F.; Shoichet, B. K. *J. Med. Chem.* **1998**, 41, 4577.
- (12) Powers, R. A.; Shoichet, B. K. *J. Med. Chem.* **2002**, 45, 3222.
- (13) Oefner, C.; D'Arcy, A.; Daly, J. J.; Gubernator, K.; Charnas, R. L.; Heinze, I.; Hubschwerlen, C.; Winkler, F. K. *Nature* **1990**, 343, 284.
- (14) Wu, W.-Y.; Chen, S.-N.; Tsai, F.-Y. *Tetrahedron Lett.* **2006**, 47, 9267.
- (15) Yang, J.; Li, P.; Wang, L. *Synthesis* **2011**, 2011, 1295.
- (16) Lv, G.; Mai, W.; Jin, R.; Gao, L. *Synlett* **2008**, 2008, 1418.
- (17) Nájera, C.; Gil-Moltó, J.; Karlström, S. *Adv. Synth. Catal.* **2004**, 346, 1798.
- (18) Tondi, D.; Powers, R. A.; Caselli, E.; Negri, M. a.-C.; Blázquez, J.; Costi, M. P.; Shoichet, B. K. *Chemistry & Biology* **2001**, 8, 593.
- (19) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, 77, 51.
- (20) Heckman, K. L.; Pease, L. R. *Nat. Protocols* **2007**, 2, 924.
- (21) Sambrook, J. *Molecular cloning : a laboratory manual / Joseph Sambrook, David W. Russell*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N.Y, 2001.
- (22) Cartwright, S. J.; Waley, S. G. *Biochem. J* **1984**, 221, 505.
- (23) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.* **2006**, 10, 253.
- (24) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, 47, 1698.

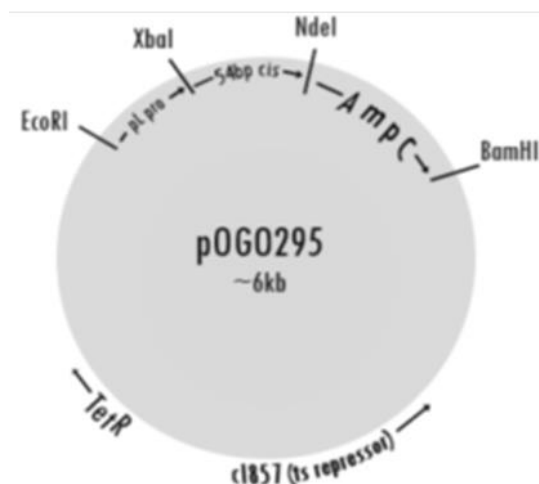
5.5 Experimental procedures

5.5.1 General considerations

See Chapter 2 experimental procedures for general considerations in chemical synthesis and protein analysis. In addition to these, general procedures used for molecular cloning are described in Section 5.5.11. Plasmid vector p0G0295 bearing *ampC* gene for protein synthesis of AmpC β -lactamase was borrowed from the Shoichet group, University of California, San Francisco (UCSF).

Introductory note about the plasmid vector p0G0295 bearing *ampC* gene

The vector used here is called p0G0295, which is a λ expression vector with promoter pL of coliphage λ , which is commonly used as the promoter system for expression vectors because of its high efficiency and ease of regulation. The *ampC* gene has been inserted in the multiple cloning site (MCS) region of the vector between the commonly used restriction sites NdeI and BamHI. The plasmid carries tetracycline resistance due to the presence of transcriptional regulator Tet^R which regulates the expression of tetracycline resistance genes. The *ampC* gene is under the control of a temperature sensitive repressor (c1857) which is active at 28 °C and represses transcription. The cells can be grown at 37 °C and the expression of the protein can be induced by increasing the temperature to 42 °C, which renders the repressor inactive and de-represses transcription. The cells are induced to release the enzyme into the supernatant by heat stress (42 °C) and the protein is extracted from the supernatant.



Scheme 5-12. Vector map for pOG0295-ampC construct (as borrowed from the Shoichet lab, UCSF)

DNA Base sequence of WT *ampC* gene

GAATTC TGCATTAGAANCANCCNNGTAGTAGGTTNNGNCCGGTGAGCACANCNCCGCAAGGAANGTGCA
TGCAAGGAGATGGCNNCCAACAGTCCCCNGGNCACGGNNCCTGCCACCATACCCACGCCGAAACAAGC
GNTTCATGAGCCNGAANGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGC
AACCGCACCTGTGGCGCCGGTGATGCCGNCCACGATGCGTCCGGCGTAGAGGATCTCTCACCTACCAA
ACAATGCCCCCCTGCAAAAAATAAATTCATATAAAAAACATACAGATAACCATCTGCGGTGATAAATTATC
TCTGGCGGTGTTGACATAAATACCACTGGCGGTGATACTGAGCACATCAGATCTATTAACCTCAA TCTAGA
GGGTATTAATAATGTATCGCGATTTAAATAAGGAGGAATAA CATATG TTCAAAACGACGCTCTGCGCCTTA
TTAATTACCGCCTCTTGCTCCACATTTGCTGCCCCCTCAACAAATCAACGATATTGTGCATCGCACAATTAC
CCCGCTTATAGAGCAACAAAAGATCCCGGGTATGGCGGTGGCGGTAATTTATCAGGGTAAACCTTATTAC
TTTACCTGGGGCTATGCGGACATCGCCAAAAGCAGCCCGTCACACAGCAAACGTTGTTTGAGTTAGGTT
CGGTCAGCAAAACATTTACTGGCGTGCTTGTTGGCGACGCTATTGCTCGAGGGGAAATCAAGTTAAGCG
ATCCCAACAACAAAATACTGGCCTGAACCTTACCGCTAAACAGTGGAATGGGATCACACTATTACATCTCGC
AACCTACACTGCTGGCGGCCTGCCATTGCAGGTGCCGGATGAGGTGAAATCCTCAAGCGACTTGCTGCG
CTTCTATCAAACTGGCAGCCTGCATGGGCTCCAGGAACACAACGTCTGTATGCCAACTCCAGTATCGGT
TTGTTGCGGCGCACTGGCTGTGAAGCCGTCTGGTTTGAGTTTTGAGCAGGCGATGCAAACTCGTGTCTTC
CAGCCACTCAAACTCAACCATACGTGGATTAATGTACCGCCCGCAGAAGAAAAGAATTACGCCTGGGGA
TATCGCGAAGGTAAGGCAGTGCATGTTTCGCCTGGGGCGTTAGATGCTGAAGCTTATGGTGTGAAGTCG
ACCATTAAGATATGGCCCGCTGGGTGCAAAGCAATTTAAACCCCTTGATATCAATGAGAAAACGCTTC
AACAAGGGATACAACCTGGCACAATCTCGCTACTGGCAAACCGGCGATATGTATCAGGGCCTGGGCTGGG

AAATGCTGGACTGGCCGGTAAATCCTGACAGCATCATTAAACGGCAGTGACAATAAAATTGCACTGGCAGC
 ACGCCCCGTAAAAGCGATTACGCCCCCAACTCCTGCAGTACGCGCATCATGGGTACATAAAACAGGGGC
 GACCGGCGGATTTGGTAGCTATGTCGCGTTTATTCCAGAAAAAGAGCTGGGTATCGTGATGCTGGCAAA
 CAAAAACTATCCCAATCCAGCGAGAGTCGACGCCGCCTGGCAGATTCTTAACGCTCTACAGTAAATTCC
 ATCGGGTCCGAATTTTCGGACCTTTTCTCCGCTTTTCTTGCTGTCATCTACACTTAGGATCCCGCGAGG
 CTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGGTGTCAAAAA

- start/ stop codon for the *ampC* gene have been highlighted in red
- restriction sites have been highlighted in green (used for future molecular cloning purposes)

5.5.2 Site-directed mutagenesis

Standard procedures were followed for all molecular biology manipulations (as described in Experimental Section 5.5.11).²¹ Some of the selected mutagenesis attempts have been described in the following section, while all the attempts have been listed in a Table 5-4.

Primer designing. Primers were designed by using a web-based program, PrimerX. Primer pairs used for the molecular cloning experiments have been listed in Table 5-4.

Table 5-4. List of primers used for molecular cloning experiments

Target gene or purpose	Primer ID	Base Sequence 5' to 3'
Primers for Site Directed Mutagenesis		
Y150C	Y150C-Fwd	acaacgtctgtgtccaactccagtatcg
	Y150C-Rev	gatactggagttggcacacagacgttgt
N152C	N152C-Fwd	acgtctgtatgcctgtccagtatcg
	N152C-Rev	caaaccgatactggagcaggcatacagac
Y221C	Y221C-Fwd	gcgttagatgctgaagcttgggtgtgaagtc
	Y221C-Rev	tcgacttcacaccacaagcttcagcatc
K315C	K315C-Fwd	tcatgggtacattgcacagggggcgac
	K315C-Rev	ggtcgcccctgtgcaatgtacctatgatg
T316C	T316C-Fwd	tcatgggtacataaatgtggggcgaccg
	T316C-Rev	gccggtcggccacatttatgtacca
N346C	N346C-Fwd	aaaaactatccctgtccagcgagagtcg
	N346C-Rev	gactctcgctggacagggatagttttg
Sequencing primers		

Begin Forward	ttcaaaacgacgctctgc
End Reverse	actctcgctggattggg
Mid Forward	agtatcggttgttcggc
Mid Reverse	catcgctgctcaaaactca
Primers for overlap extension PCR (containing restriction sites on 5')	
Flanking primer-a	ggaacacaacgtctgtgtgccaactccagtatcg
Flanking primer-d	accgatactggagttggcacacagacgttgtgttc

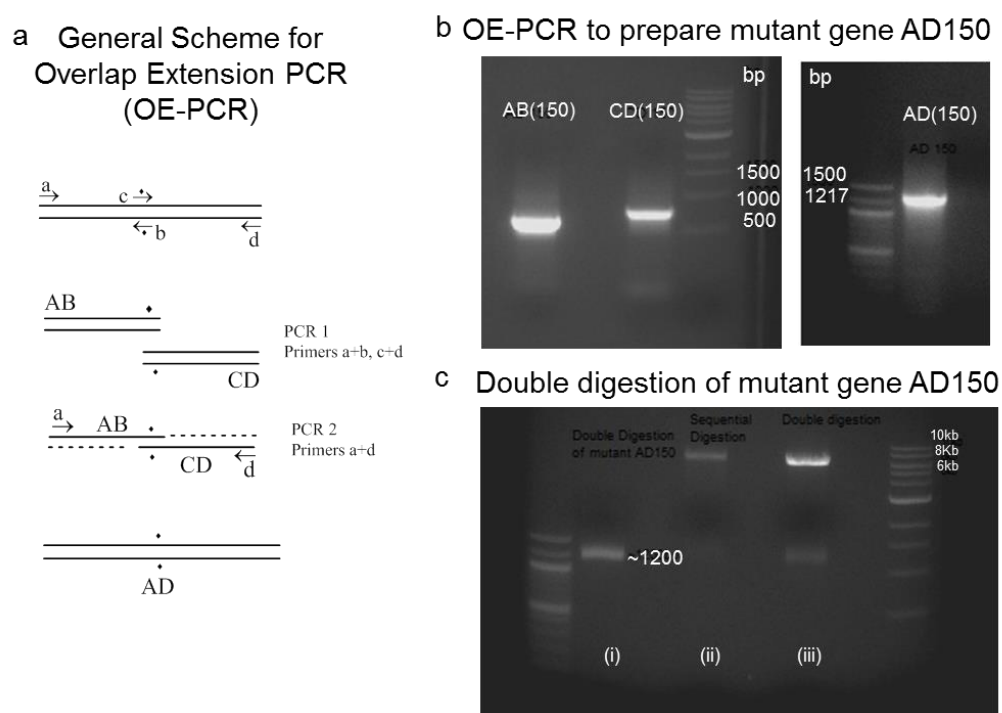
Mutagenesis. Mutagenesis experiments were carried out by using the QuikChange[®] Site-Directed Mutagenesis Kit (Agilent). Several attempts were made to obtain the mutated DNA for Y150C, N152C, Y221C, K315C, T316C and N346C. The PCR reactions proceeded smoothly as visualised by DNA Agarose gel electrophoresis. Table 5-5 gives the list of conditions attempted to carry out site-directed mutagenesis.

Table 5-5. List of conditions employed to optimize mutagenesis experiments using QuikChange[™] Site-Directed Mutagenesis Kit (Agilent)

Entry	Conditions used				Remarks
	PCR Cycle (Annealing Temperature)	Transformation Procedure	<i>E. coli</i> Strain used	DNA Polymerase used	
1	55 °C	Heat shock	Oneshot Top10	Pfu Ultra	No amplification seen
2	60 °C	Heat shock	Oneshot Top10	Pfu Ultra	Amplification seen, but no colonies were obtained
3	60 °C	Heat shock	JM109	Pfu Ultra	Amplification seen, but no colonies were obtained
4	60 °C	Heat shock	Oneshot Top10	Pfu Ultra	Amplification seen, but no colonies were obtained
5	60 °C	Electroporation	Oneshot Top10	Pfu Ultra	Amplification seen, but no colonies were obtained
6	60 °C	Electroporation	Oneshot Top10	KOD	Amplification seen, but no colonies were obtained

Overlap Extension PCR. Scheme 5-13 was followed to make the mutant genes using Overlap Extension PCR. In this technique, internal primers ‘b’ and ‘c’ used in two separate reactions give hybrid products AB and CD. During the second PCR, strands with 3'

complementary ends created by mutagenic primers ‘b’ and ‘c’ will anneal. Primers ‘a’ and ‘d’ will amplify one of these hybridization products (sense strand from AB and antisense strand from CD) to generate the predominant product AD, whereas the other hybridization product (sense strand CD and antisense strand from AB) will not be amplified in the absence of any driving primers. Similarly, although the sense and antisense strands of each product AB and CD will also re-hybridize, without primers band c, they are not major products of this PCR. Thus, primers ‘a’ and ‘d’ will preferentially amplify mutant gene AD.



Scheme 5-13. (a) General scheme followed for site-directed mutagenesis by Overlap Extension PCR (OE-PCR), (b) DNA gel visualization of PCR products ‘AB150’ and ‘CD150’ (left) and PCR product ‘AD150’ (right) after OE-PCR reaction, (c) DNA gel visualization of (i) double digestion (ii) sequential double digestion reactions of mutant gene Y150C and (iii) double digestion of vector p0G0295 containing AmpC WT gene used as positive control

Synthesis of mutant gene AD150 and introduction of restriction sites by double digestion. Mutant gene corresponding to Y150C mutation was synthesized according to Scheme 5-13 and subsequently digested with NdeI and BamHI-HF to introduce restriction sites on the gene. 5 µg WT p0G0295 vector containing *ampC* gene was also digested with these enzymes as a control and to obtain the cut vector p0G0295. The DNA gels depict successful synthesis of AD150 by PCR and double digestions. Mutant gene Y150C and the digested vector were gel-purified for their use in subsequent ligation reaction.

Ligation reaction. Ligation of the double-digested product AD of the mutant genes corresponding to Y150C, N152C, Y221C, K315C and Y316C and the cut vector DNA was attempted using the Quick LigationTM Kit (NEB) according to the manufacturer's protocol. The amount of insert for the reaction with 50 ng vector was calculated using the following equation.

$$\text{ng insert} = (\text{insert size (kb)} / \text{vector size (kb)}) * \text{ng vector}$$

Optimization of Ligation conditions.

By varying insert to vector ratio. Several ligation reactions with varying amount of insert to vector (molar ratios 1:1, 3:1, 6:1 and 10:1) were tried out. This was done to saturate the vector with excessive amount of insert to increase the probability of successful ligation.

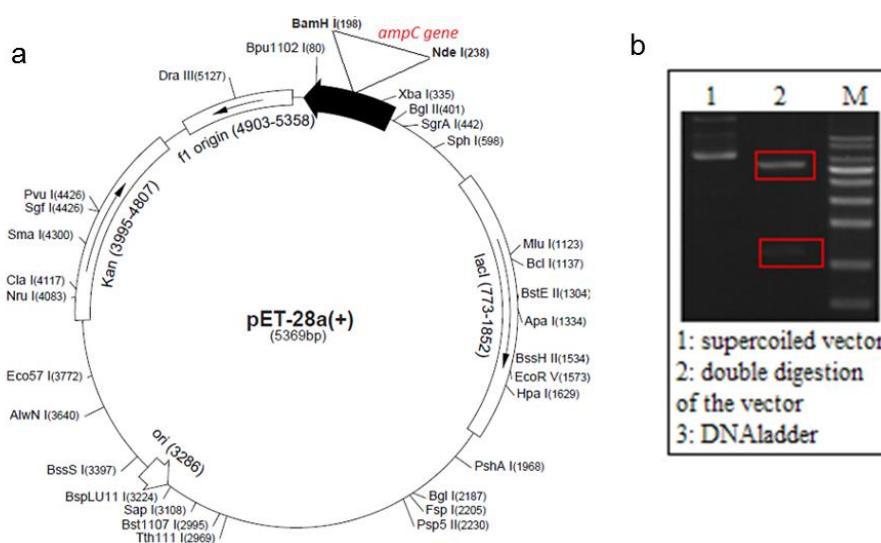
By varying incubation time. Ligation reactions were also tried with increased incubation time (up to 16 h), followed by transformation into DH5α cells along with necessary controls. No colonies could be obtained.

Transformation into subcloning efficiency cells. DH5α bacterial cells were transformed with ligation reactions (~4 µL), as per the manufacturer's protocol. Plasmid DNA included

with DH5 α cells was used as positive control; double-digested vector only positive control was also used. No colonies were obtained in either case, except for the positive control (showing that the competent cells were viable).

5.5.3 Subcloning of *ampC* gene into pET28a(+)

The *ampC* gene introduced in pET-28a(+) vector between NdeI and BamHI restriction sites was ordered through GenScript Inc. USA. Double digestion of the vector bearing *ampC* gene confirmed the success of the subcloning experiment. The plasmid thus obtained was transformed into XL10-Gold[®] Ultracompetent cells for mutagenesis purposes and correct gene sequence was confirmed by sequencing experiments ordered through Geneservice, University of Oxford.



Scheme 5-14. Vector map for pET-28a(+)-*ampC* construct; *ampC* gene has been inserted between BamH I and Nde I sites in the commercial pET-28a(+) vector (b) DNA gel of pET28a(+)-*ampC* construct showing successful insertion of *ampC* gene (as received from GenScript Inc. USA)

5.5.4 Site-directed mutagenesis using pET-28a(+)-*ampC* construct

Two rounds of mutagenesis were carried out to substitute the cysteine residues occurring at the 5' end of the gene of interest in this construct. Successful mutagenesis reactions gave the following wild-type *ampC* gene (WT-*ampC* gene).

WT-*ampC* gene

MFKTTLSALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFETWG
YADIAKKQPV TQQTLEFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
TLLHLATYTA GGLPLQVPDE VKSSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
STIEDMARWV QSNLKPLDIN EKTLQQGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
NGSDNKIALA ARPVKAITPP TPAVRASWVH KTGATGGFGS YVAFIPEKEL GIVMLANKNY
PNPARVDAAW QILNALQ

Six successful site-directed mutagenesis reactions using this WT-*ampC* gene sequence led to generation of the following active site mutants of *ampC* gene in pET28a(+). The sites of mutation have been highlighted in red. The correct sequence was confirmed by sequencing experiments ordered through Geneservice, University of Oxford.

Y150C

MFKTTLSALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFETWG
YADIAKKQPV TQQTLEFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
TLLHLATYTA GGLPLQVPDE VKSSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
STIEDMARWV QSNLKPLDIN EKTLQQGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
NGSDNKIALA ARPVKAITPP TPAVRASWVH KTGATGGFGS YVAFIPEKEL GIVMLANKNY
PNPARVDAAW QILNALQ

N152C

MFKTTLSALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFETWG
YADIAKKQPV TQQTLEFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
TLLHLATYTA GGLPLQVPDE VKSSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
STIEDMARWV QSNLKPLDIN EKTLQQGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
NGSDNKIALA ARPVKAITPP TPAVRASWVH KTGATGGFGS YVAFIPEKEL GIVMLANKNY
PNPARVDAAW QILNALQ

Y221C

MFKTTLSALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFETWG
YADIAKKQPV TQQTLEFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
TLLHLATYTA GGLPLQVPDE VKSSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK

STIEDMARWV QSNLKPLDIN EKTLLQGGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
 NGSDNKIALA ARPVKAITPP TPAVRASWVH KTGATGGFGS YVAFIPEKEL GIVMLANKNY
 PNPARDAAW QILNALQ

K315C

MFKTTLALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFTWG
 YADIAKKQPV TQQTLLFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
 TLLHLATYTA GGLPLQVPDE VKSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
 PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
 STIEDMARWV QSNLKPLDIN EKTLLQGGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
 NGSDNKIALA ARPVKAITPP TPAVRASWVH KGTGATGGFGS YVAFIPEKEL GIVMLANKNY
 PNPARDAAW QILNALQ

T316C

MFKTTLALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFTWG
 YADIAKKQPV TQQTLLFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
 TLLHLATYTA GGLPLQVPDE VKSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
 PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
 STIEDMARWV QSNLKPLDIN EKTLLQGGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
 NGSDNKIALA ARPVKAITPP TPAVRASWVH KGTGATGGFGS YVAFIPEKEL GIVMLANKNY
 PNPARDAAW QILNALQ

N346C

MFKTTLALL ITASSSTFAA PQQINDIVHR TITPLIEQQK IPGMAVAVIY QGKPYFTWG
 YADIAKKQPV TQQTLLFELGS VSKTFTGVLG GDAIARGEIK LSDPTTKYWP ELTAKQWNGI
 TLLHLATYTA GGLPLQVPDE VKSSDLLRF YQNWQPAWAP GTQRLYANSS IGLFGALAVK
 PSGLSFEQAM QTRVFQPLKL NHTWINVPPA EEKNYAWGYR EGKAVHVSPG ALDAEAYGVK
 STIEDMARWV QSNLKPLDIN EKTLLQGGIQL AQSRYWQTGD MYQGLGWEML DWPVNPDSII
 NGSDNKIALA ARPVKAITPP TPAVRASWVH KTGATGGFGS YVAFIPEKEL GIVMLANKNY
PNPARDAAW QILNALQ

5.5.5 Protein Expression

Expression of AmpC β -lactamase wild-type protein (AmpC-WT) from p0G0295-*ampC* plasmid DNA in JM109 *E. coli* cells was accomplished according to a published procedure.¹⁹ Two overnight cultures were set up in 5 mL 2xYT media each containing the appropriate antibiotic (5 μ l of 10 mg/mL tetracycline stock in EtOH was added) using a

single colony of wild-type *ampC*-p0G0295 in JM109 cells. The cultures were incubated at 37 °C, 200 rpm for 16 h. The overnight cultures were used to inoculate 1 L of 2xYT media containing tetracycline (1 mL of a 10 mg/mL stock solution in EtOH; final concentration 10 µg/ mL) and placed in an incubator shaker at 37 °C, 200 rpm. The OD₆₀₀ was recorded every 45 min until it reached to 0.3-0.5. Cells were then induced by heat shock by shaking the flask at 40.5 °C (up to 41 °C) at 210 rpm. The cell growth was monitored every 4 h by measuring OD₆₀₀ throughout the day. Heat shock was continued for another 24 h.

Harvesting. At the end of heat shock, the β-lactamase activity was detected through Nitrocefin activity assay (described below) with a 10 µL aliquot of the supernatant. β-lactamase was released out into the cytosol so it ended up in the supernatant and not the pellet. The cells were separated from the supernatant by two centrifugation steps at 5,000 g in a JLA 9.1 rotor (20 min at 4 °C). The centrifugation step was repeated twice for ease of the subsequent filtration step. The supernatant was separated and was filtered through a 0.45 µm filter, followed by filtration of flow-through using a 0.2 µm filter. The activity of the filtered supernatant was tested by BCA assay to assess the amount that could be loaded over the column. 0.01% (g/mL) sodium azide was added to supernatant as a preservative. The filtered supernatant was concentrated to reduce the volume to load over the column, using an ultra-filtration set-up at 4°C. pH of supernatant was adjusted to 7 with solid Tris-HCl. The supernatant was stored at 4 °C until further purification.

Preparation of AmpC β-lactamase wild-type (AmpC-WT) and AmpC variant proteins. Expression of these variant enzymes from pET-28a(+)-ampC in BL21(DE3) *E. coli* cells was accomplished by using the standard expression conditions for pET vectors in this strain of cells (according to the manufacturer's instructions).

Pre-culture and Innoculation. In a typical experiment, *E. coli* BL21 (DE3) cells harbouring the desired plasmid were grown in 10 mL of freshly prepared and autoclaved LB medium (BD 244620, Difco) containing a final concentration of 50 µg/mL of kanamycin. The cells were incubated overnight at 37 °C. The cells were next transferred into 800 mL of fresh LB broth (containing same antibiotic concentration described above) and were incubated at 37 °C with vigorous shaking until reaching an OD₆₀₀ 0.6 - 0.8 (typically 3-4 h).

Induction and harvesting. To the cells was added 800 µL of a previously prepared 1M IPTG (isopropyl-β-Dithiogalactoside) stock solution. The cells were stirred at 16 °C for 12 h after which time they were harvested via centrifugation (8,000 rpm, 10 min, 4 °C).

Cell Lysis. The *E. coli* cells containing overexpressed 6X His tagged protein were disrupted upon addition of DNase (0.1 mg/mL in binding buffer), egg white lysozyme (1 mg/mL in binding buffer) and protease inhibitor cocktail tablet (1 tablet/ 8 g cell pellet, Complete EDTA free, Roche) with vigorous stirring overnight at 4 °C, followed by pulse sonication (30 s on, 60 s off, 0.5 s pulses). Sonication was carried out at 0 °C for a total of 6 cycles. The lysate was transferred to a cooled centrifuge tube and the cell debris was removed via centrifugation (20,000 rpm, 40 min, 4 °C).

5.5.6 Protein Purification

Preparation of MAPB.¹⁹ 62.5 mL of 1 M KHCO₃/ 0.5 M sorbitol was stirred in a 500 mL flask, with slow addition of 6.25 g *m*-aminophenylboronic acid (MAPB). The final solution was heated at 37 °C for 2 h. MAPB was found to be quite insoluble so about 10 mL more of buffer solution was added and heating was continued at 42 °C for another 2 h.

MAPB was dissolved in the buffer overnight at room temperature, and then stored at 4 °C. The pH of the resulting solution was found to be ~8.

Derivatization of Affi-Gel 10 with MAPB. Affi-Gel10 contains a derivatized, cross-linked agarose gel bead support with a neutral 10-atom spacer arm and a terminal *N*-hydroxysuccinimide group which can be displaced by ligands containing primary amine groups (such as MAPB). The MAPB solution was allowed to get to room temperature. 25 mL Affi-Gel 10 was washed with 5 volumes of chilled isopropanol (125 mL) and then with 5 volumes of cold milliQ water (125 mL) under vacuum (maintained at low suction in order to avoid the gel from getting dry). The last water wash was drained by gravity. When grainy-wet, Affi-Gel 10 was scooped into the flask containing MAPB solution. The mixture was left on a rocking table overnight at room temperature until it turned translucent (pH ~8). Affi-Gel 10-MAPB mixture was packed into an Econo-column (Bio-Rad). The column was washed with 10 volumes (250 mL) of wash buffer A (1 M NaCl/ 0.5 M Sorbitol, pH 7) and then with 250 mL wash buffer B (0.5 M Boric acid, pH 7). The column was equilibrated with 10 volumes of load buffer (50 mM NaPi/ 0.5 M NaCl, pH 7-7.5) at 4 °C.

Protein purification. Purification of AmpC-WT and AmpC variants (Y150C, N152C, Y221C, K315C, T316C and N346C) was carried out using Affigel-MAPB column. Expression and purification of AmpC variants Y150C, N152C, Y221C and T316C was carried out by Dr. Maria Espling.

The concentrated supernatant was loaded onto the column (pre-equilibrated with 2-3 column volumes of load buffer) at a rate of 1-2 mL/ min. The flow-through was collected

in a labeled vessel. Ideally, the column should bind 1-2 mg protein per mL of Affi-Gel 10. The column was loaded with slightly more than the maximum binding capacity. The activity of the flow-through was tested by Nitrocefin spot test to assess protein binding by the column. The column was washed with ~4 column volumes of load buffer and the flow through was collected. Subsequently, the protein was eluted by washing the column with 50-100 mL of the elution buffer (0.5 M Boric acid/ 0.5 M NaCl, pH 7) and fractions were collected. The enzyme was dialyzed into 50 mM KPi phosphate buffer, pH 7 for storage at 4 °C. The dialysis buffer was switched twice to make elution buffer reach picomolar concentration. Typically, 15-20 mg protein was obtained from 1 L culture.

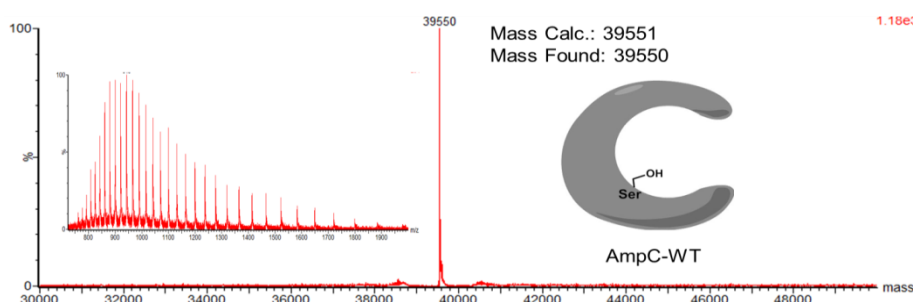
Alternatively, gel filtration of AmpC152 was carried out using the Superdex16/60 gel filtration column. The column was equilibrated with 4 volumes of load buffer (50 mM NaPi/ 0.5 M NaCl, pH 7-7.5) at 4 °C. The concentrated supernatant (~2 mL) was loaded onto the column at a rate of 1-2 mL/ min. The column was washed with ~4 column volumes of load buffer and fractions were collected and analysed by SDS-PAGE analysis.

5.5.7 Characterization of AmpC-WT and AmpC-variants

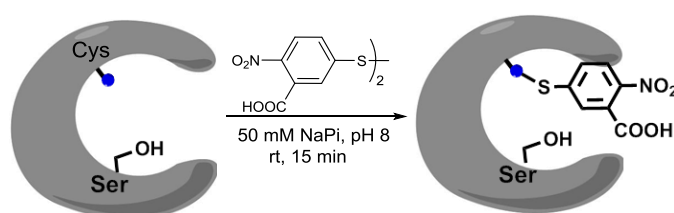
Activity. Nitrocefin disks (Sigma) were employed for this qualitative enzyme assay. A nitrocefin disk (6 mm diameter filter paper disks impregnated with a chromogenic cephalosporin substrate, nitrocefin) was placed in a petri dish. The disk was moistened with one drop of de-ionized water, followed by application of a 10 µL aliquot of the supernatant. β -lactamase activity was indicated by an immediate colour change to a red coloured hydrolyzed compound.

Concentration. Enzyme concentration was measured by BCA assay. BCA assay was carried out in a 96-well format employing BSA standards and assay reagents A and B (mixed in a ratio 50:1). 200 μ L of the reagent solution was taken in each well and 10 μ L of protein sample/ BSA standards were added successively. The plate was incubated at 37 °C for 30 min. Absorbance was measured at 562 nm using a UV/ Visible spectrophotometer. Linearity was observed for the BSA standards from which unknown concentration was calculated.

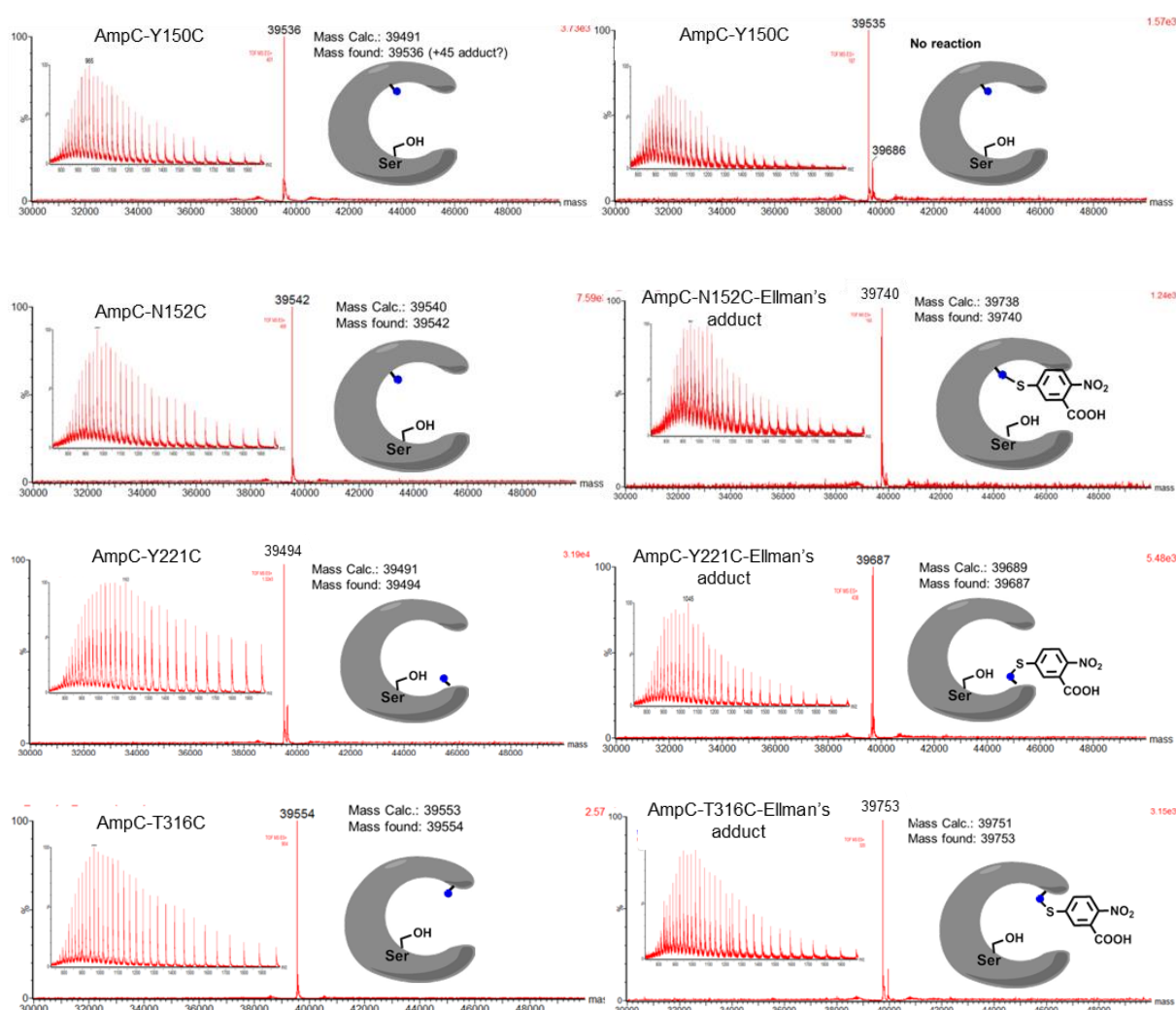
Purity and homogeneity. AmpC was found to be homogeneous by SDS PAGE. Novex Bis-Tris gels (4-12%, Invitrogen) were used. Gels were placed in the gel tank, and the gel tank was filled with MES running buffer (800 mL, prepared from 20x Invitrogen concentrate). Samples for the gel were prepared by adding sample (15 μ L) to 6x dye (3 μ L) and denaturing at 90 °C for 5-10 min. 10 μ L of this denatured sample was loaded onto the gel. The gel was run at 200 V for 50 min or the time taken for the dye to reach the bottom of the well. Coomassie Brilliant Blue Stain (20 mL) was added to the gel. It was allowed to develop overnight on an orbital rocker, after which it was washed with water twice. LC-MS analysis confirmed the correct mass of the enzyme.

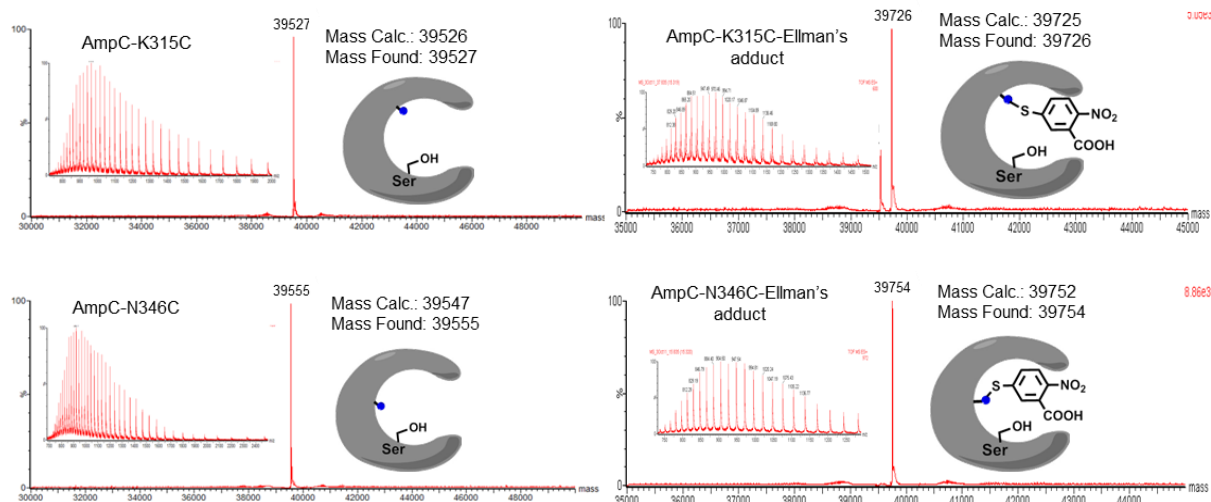


Ellman's test for single cysteine mutants



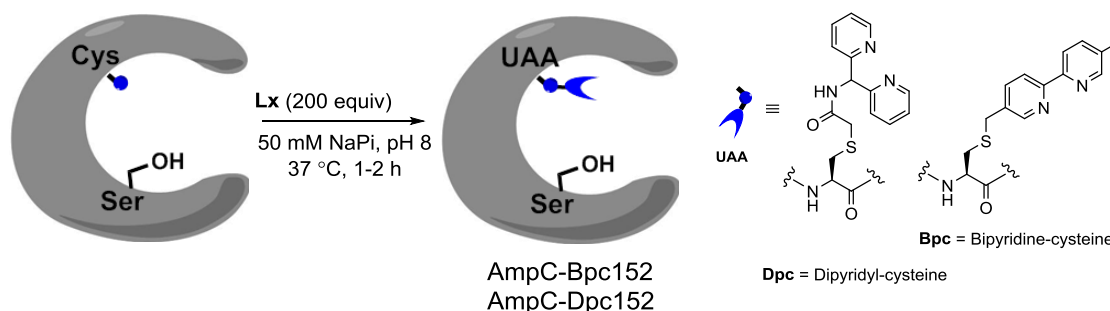
Ellman's test. Presence of a single cysteine was confirmed by treatment of the AmpC-variants with 50 equiv. of Ellman's reagent. Typically, a 10 μ L aliquot of AmpC-variant was prepared at 0.4 mg/mL in sodium phosphate buffer (50 mM, pH 8). A 25 mM stock solution of Ellman's reagent was prepared in 50 mM sodium phosphate buffer, pH 8. To the 10 μ L sample of protein (0.1 nmol) was added 0.2 μ L of the Ellman's solution (5 nmol). The reaction was shaken for 1 h at room temperature and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the Ellman's disulfides. ESI-MS for modifications are shown below.





5.5.8 Direct alkylation of AmpC-variants with ligands Lx

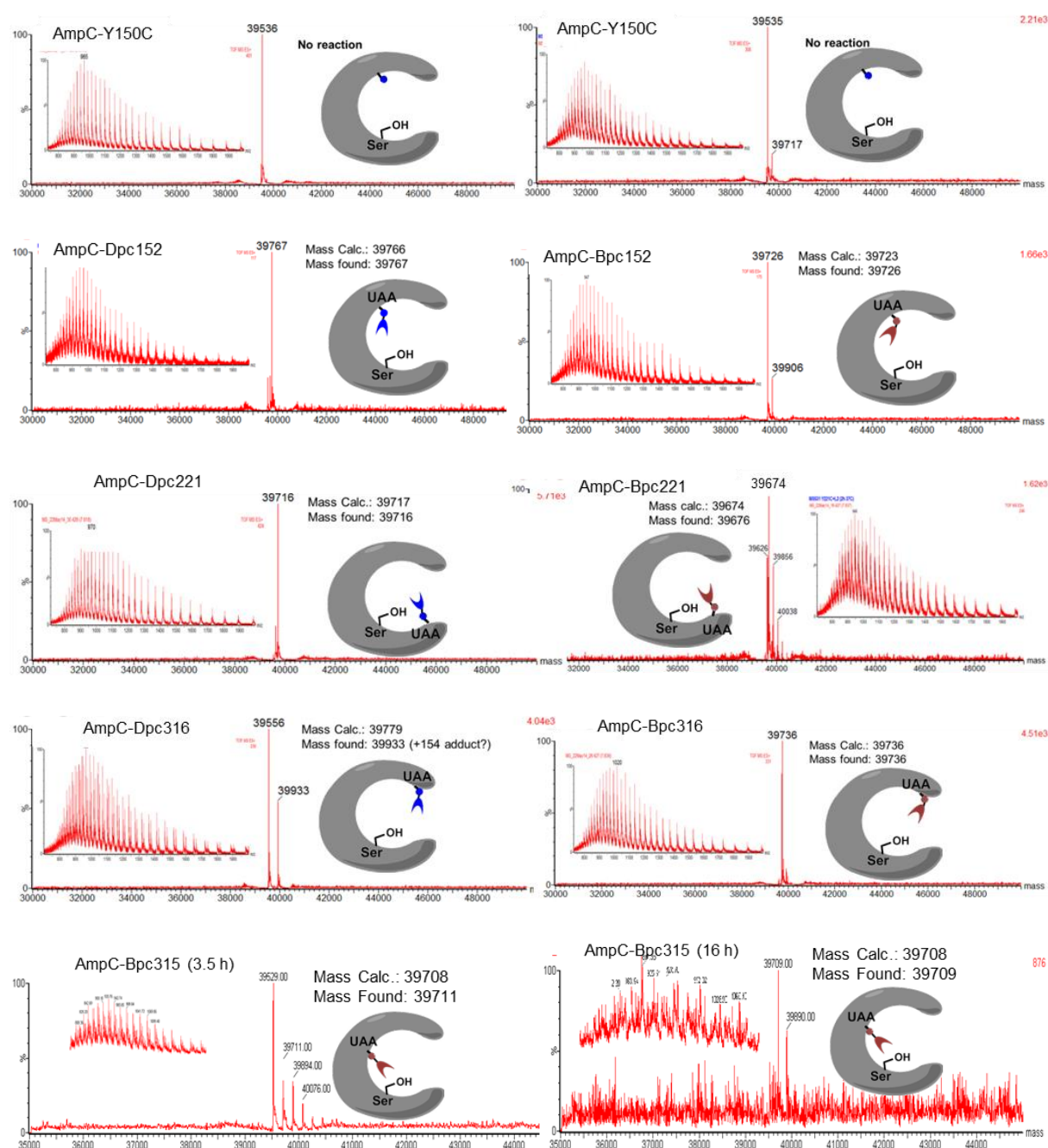
Chemical modification of AmpC-variants was accomplished by direct alkylation using halogenated ligands **L1** and **L2** using the following standard procedure.

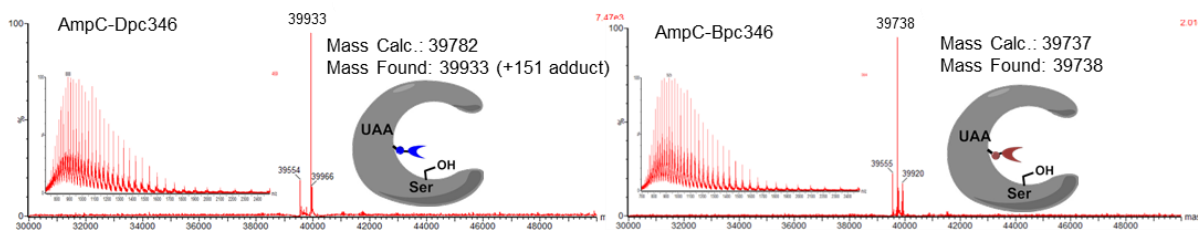


A 100 μ L sample of AmpC-N152C was prepared at 0.6 mg/mL in sodium phosphate buffer (50 mM, pH 8) and the enzyme was stored on ice until needed. The concentration was determined by BCA assay as described (Section 5.5.7). A 1 M stock solution of DTT was prepared in MilliQ water and a 1 M solution of **L1** in DMF (0.5 M stock solution of **L2** was used because of its limited solubility in water).

To 50 μ L of AmpC-N152C (0.75 nmol) in 50 mM phosphate buffer, pH 8 at 4 $^{\circ}$ C, was added 10 equiv. of DTT to reduce the possible disulfide bonds. The reaction was shaken for 15 min at room temperature. The protein solution was reacted with **Lx** (200 equiv.).

The reaction mixture was incubated at 37 °C for up to 2 h. The reaction was monitored at intervals of 1 h by LC-MS. A single protein species was detected showing a mass corresponding to the apo-enzyme species (For AmpC-Dpc152: mass calculated 39766, mass found 39767; For AmpC-Bpc152: mass calculated 39723, mass found 39726). PD-10 spin trap G-25 column were pre-equilibrated with 50 mM sodium phosphate buffer, pH 6 at 4 °C. The reaction mixture was dialysed against sodium phosphate buffer (50 mM, pH 6) for 2 h. ESI-MS for the modifications is shown below.

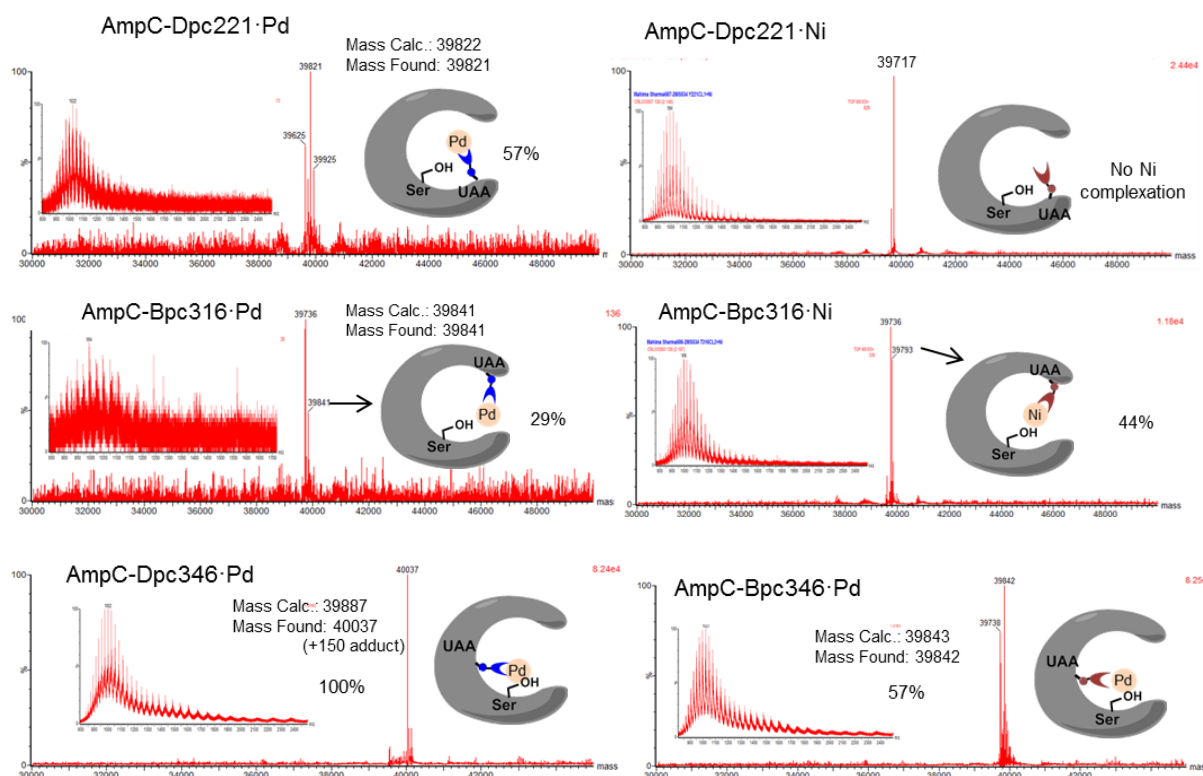




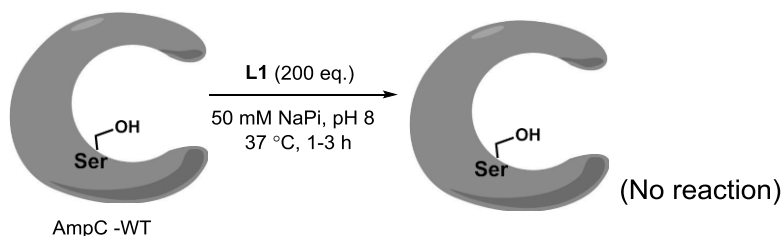
5.5.9 Metal incorporation

Stock solution of Na_2PdCl_4 and $\text{Ni}(\text{OAc})_2$. 0.5 M final stock solution was made by suspending NaCl (58.5 mg, 1 mmol) and PdCl_2 (89 mg, 0.5 mmol) in 1 mL distilled H_2O . The mixture was stirred in a sealed flask at 45 °C until PdCl_2 dissolved completely. This stock solution was diluted to 5 mM or 10 mM stock solution using MilliQ water. A 10 mM stock solution of $\text{Ni}(\text{OAc})_2$ in MilliQ water was used for complexation reactions.

Typically, a 100 μL sample of AmpC-Dpc221 was prepared at 0.38 mg/mL in sodium phosphate buffer (50 mM, pH 6) and the enzyme was stored on ice until needed. A 10 mM stock solution of metal precursor (Na_2PdCl_4 or $\text{Ni}(\text{OAc})_2$) was prepared in MilliQ water as described above. To 100 μL sample of AmpC-Dpc221 (0.9 nmol) in 50 mM phosphate buffer, pH 6 at 4 °C, was added 0.2 μL of metal precursor (2 nmol). The reaction solution was vortexed to homogenize and subsequently incubated at 37 °C for 1-3 h. The reaction was monitored LC-MS. The ESI-MS on completion of reaction after 3 h are shown below.

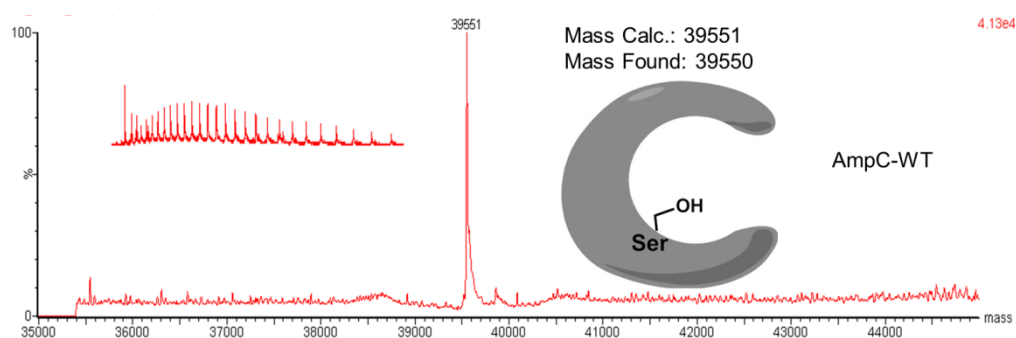


5.5.10 Control reactions using AmpC-WT

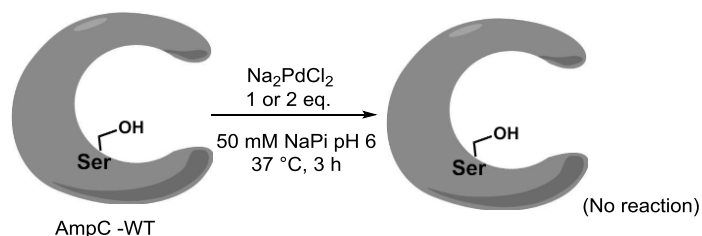


Treatment of AmpC-WT with ligand L1. A 1 ml sample of AmpC-WT was prepared at 1 mg/mL in sodium phosphate buffer (50 mM, pH 8) and the enzyme was stored on ice until needed. To 200 μ l (5 nmol) of AmpC-WT in 50 mM phosphate buffer, pH 8 at 4 $^{\circ}$ C, 0.2 μ l of DTT (1 M, 25 equiv.) was added and the reaction was incubated at room temperature for 15 min. 1.01 μ l of **L1** (1.01 μ mol, 1 M solution in DMF) was added to the protein solution. The reaction mixture was incubated at 37 $^{\circ}$ C for 3 h. The reaction was monitored at an interval of 1 h by LC-MS. PD-10 spin trap G-25 column was pre-equilibrated with 50 mM acetate buffer, pH 5 at 4 $^{\circ}$ C. The reaction mixture was de-salted using pre-equilibrated PD-10 spin trap G-25 column after 3 h, followed by dialysis in water for 2 h. No reaction

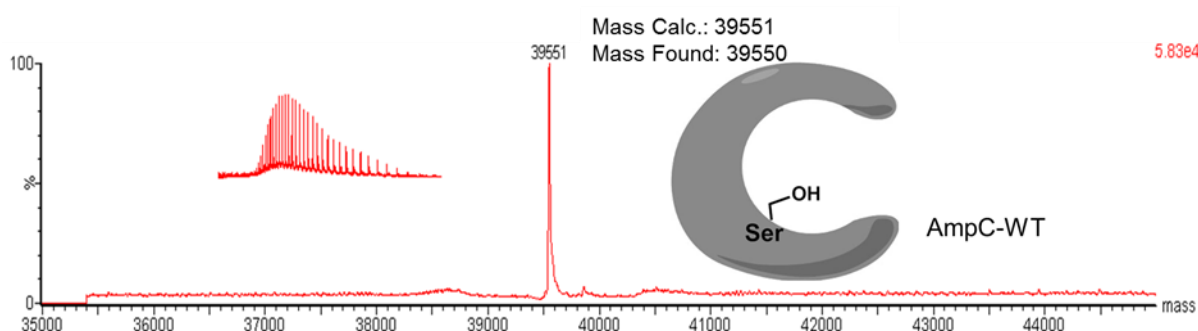
was observed even after incubating the reaction mixture for 3 h. The ESI-MS of the reaction is shown below.



Pd Complexation Reaction.



To 100 μ L (0.002 mmol) of AmpC-WT in 50 mM phosphate buffer, pH 6 at 4 $^{\circ}$ C (0.8 mg/mL; concentration determined by BCA assay), 0.34 μ L or 0.64 μ L of Na_2PdCl_4 (5 mM, 1 or 2 eq.) were added to AmpC-WT in two separate tubes. The reaction mixture was incubated at 37 $^{\circ}$ C for 3 h, and aliquots were analysed by LC-MS after an interval of 1 h. PD-10 spin trap G-25 column was pre-equilibrated with 50 mM acetate buffer, pH 5. The reaction mixture was de-salted using pre-equilibrated PD-10 spin trap G-25 column, followed by dialysis in water for 2 h. No reaction was observed even after incubating the reaction mixture for 3 h. The ESI-MS of the reaction is shown below.



5.5.11 Standard Operating Procedures (for Molecular Cloning)

DNA Isolation using QIAprep[®] Spin Miniprep Kit. Miniprep Plasmid DNA isolation and purification was carried out using the QIAprep[®] Spin Miniprep Kit according to the manufacturer's protocol. The overnight culture was centrifuged and the supernatant was poured off. Pelleted bacterial cells were resuspended in buffer P1 (250 µL) and transferred to a microcentrifuge tube. Buffer P2 (250 µL) was added, and the sample was mixed thoroughly by inverting the tube 6 times. Buffer N3 (350 µL) was added and mixed immediately and thoroughly by inverting the tube 6 times. The mixture was centrifuged at 13,000 rpm (17,900 x g) for 10 min in a table-top microcentrifuge. The supernatant was pipetted into the QIAprep spin column, and centrifuged for 1 min. The flow-through was discarded. Buffer PE (750 µL) was added to the spin column, and centrifuged for 1 min. This wash step was repeated twice. The flow-through was discarded, and the spin column was centrifuged for 1 min to remove residual wash buffer. To elute DNA, the spin column was placed in a clean 1.5 mL microcentrifuge tube. Water (50 µL) was added to the centre of the QIAprep spin column, and it was let to stand for 1 min, and then centrifuged for 1 min (13,000 rpm). Pure DNA was stored at -20 °C.

Transformation. Transformation of One Shot[®] Top10, DH5α, JM109, BL21(DE3) and XL10-GOLD Ultracompetent chemically competent cells was carried out by following manufacturer's instructions. Typically, the competent cells were gently thawed on ice, and then 50 µL aliquoted into a pre-chilled 14 mL BD falcon polypropylene round-bottomed tube. 1 µl DNA was transferred to the competent cells. The reaction was gently swirled to mix, and then incubated on ice for 30 min. The reaction was heat pulsed for 30 or 40 seconds (depending on the strain of cells) at 42 °C and then placed on ice for 2 min. Sterile SOC media (250 µL) preheated to 42 °C was added, and then the reaction was incubated at

37 °C for 1 h (225 rpm). The transformation reaction was plated onto agar plates containing appropriate resistance. The transformation plates were incubated at 37 °C for a minimum of 16 h. The plates containing colonies were stored at 4 °C.

Preparation of 2xYT-Agar Plates. 1.6 g tryptone, 1.0 g yeast extract, 0.5 g NaCl were added to 100 mL milliQ water and the pH was adjusted to 7. Next, 1.5 g agar was added to the solution. The solution was autoclaved. Once it had cooled down slightly, 100 µL tetracycline (10 mg/ mL) was added. Approximately 20 mL of the solution was poured into petri dishes. They were allowed to set, and then stored at 4 °C until needed.

Making Glycerol Stock. An overnight culture was set up in a 5 mL falcon tube with 5 µL tetracycline (10 mg/ mL) with a single colony. Since tetracycline is toxic for the cells during long term storage, the cells from the overnight culture were spun down and resuspended in a tetracycline-free sterile 2xYT media. 500 µL of this resuspended culture was added to 500 µL 50% glycerol solution under sterile conditions. The resulting glycerol stock was stored at -80 °C. A portion from the top of the frozen glycerol stock was streaked out on a 2xYT-agar plate containing tetracycline for selection. The plate was incubated overnight and a mini culture (10 mL) was done to ascertain that the glycerol stock had the desired cells growing.

Primer designing. Primer designing was done by using a web-based program, PrimerX, based on the DNA base sequence confirmed through the Geneservice, Department of Biochemistry, Oxford and subsequently checked using the Primer Premier software & design tool (Sigma) for potential secondary structure or primer dimer. Since the design of

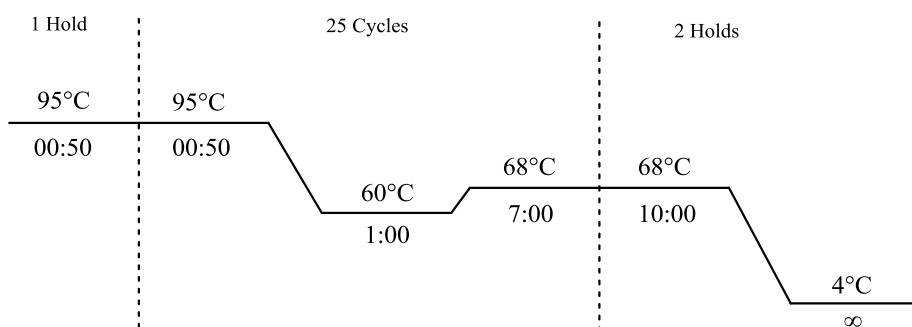
primers is crucial to subsequent cloning steps, following points were borne in mind while designing the oligos.

- GC content maintained between 40-60%
- T_m was calculated using the design tool. For sequencing purposes, T_m was maintained between 60-70 and length of oligos being 18-22 bp, while oligos for PCR mediated reactions had $T_m > 75$ and length of primers between 25-45 bp
- Primers were required to have weak/very weak secondary structure and no primer dimer

PCR-mediated Mutagenesis using QuikChange[®] Site-Directed Mutagenesis Kit (Agilent). Each reaction was prepared as per the manufacturer's instructions.

5 μ L 10x reaction buffer
1 μ L (50 ng) ds DNA template
125 ng forward oligonucleotide primer
125 ng reverse oligonucleotide primer
1 μ L dNTP mix
ddH₂O to 50 μ L

The reaction mixture was vortexed and centrifuged. PfuUltra DNA polymerase (1 μ L, 2.5 U/ μ L) was added to the final volume of 50 μ L and the mixture was pipetted to ensure homogeneity. The polymerase chain reaction (PCR) was carried out using the following cycling parameters.



Scheme 5-15. *PCR cycle used for mutagenesis reaction*

Success of the PCR was confirmed by DNA agarose gel electrophoresis. Dpn I restriction enzyme (1 μ L, 10 U/ μ L) was added to the amplification reaction. The reaction mixture was mixed gently and thoroughly by pipetting the solution up and down several times. The reaction was incubated at 37°C for 1 h to digest the parental (the non-mutated) supercoiled dsDNA.

PCR-mediated mutagenesis using Overlap Extension PCR. Overlapping primers that introduce the base pair substitution were designed. These primers needed to be longer than usual (55-65 base pairs). Appropriate restriction enzyme sites were introduced on the 5' ends of the primers (making sure the directionality was correct). Mutagenesis PCR reaction was set up (2 subsequent cycles). PCR products from the first cycle were used as templates for second PCR. Typically, following composition of reaction mixture and PCR cycle was used, keeping in mind the size of the final product.

Step 1. PCR set up for AB and CD fragments

Primers a, b, c, d were diluted to 20 pmol/ μ L. Reagents were combined in two PCR tubes as follows.

Table 5-6. List of components in the reaction mixture for PCR reaction to generate fragments AB and CD

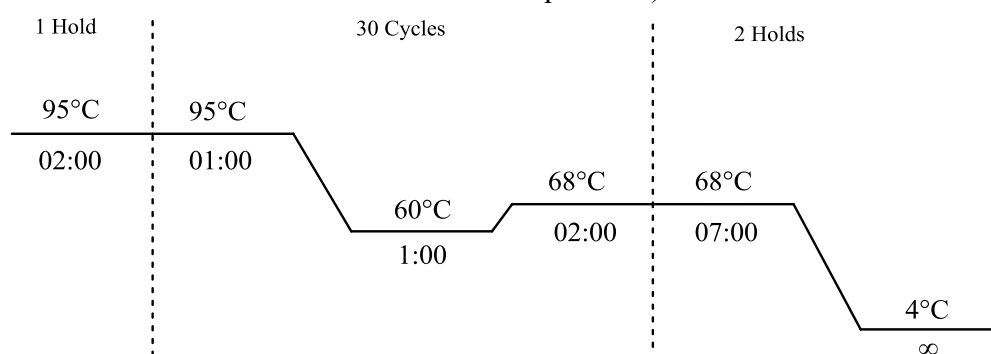
Component	Amount per reaction (μL)		
	AB	CD	Final
Template DNA	1	1	50-125 ng
10x Buffer (pH=7.5)	5	5	1x
MgCl ₂ (25 mM)	3	3	2-3 mM
DMSO	1.5	1.5	3%
dNTP (2 mM)	5	5	20 mM
Primer a	2.5	-	50 pmol
Primer b	2.5	-	50 pmol
Primer c	-	2.5	50 pmol
Primer d	-	2.5	50 pmol
Water	29.5	29.5	To 50 μl
KOD or PfuUltra DNA polymerase (2.5 U/μl)	1	1	2.5 U

Step 2. PCR run for AD

Table 5-7. List of components in the reaction mixture for PCR reaction to prepare fragments AB and CD

Component	Amount per reaction (μL)	
	AD	Final
AB	variable	100 ng
CD	variable	100 ng
10x Buffer (pH=7.5)	5	1x
MgCl ₂ (25 mM)	3	2-3 mM
DMSO	1.5	3%
dNTP (2 mM)	5	20 mM
Primer a	2.5	50 pmol
Primer d	2.5	50 pmol
Water	29.5	To 50 μl
KOD or PfuUltra DNA polymerase (2.5 U/μl)	1	2.5 U

(For KOD polymerase annealing and extension time was set according to the manufacturer's protocol).



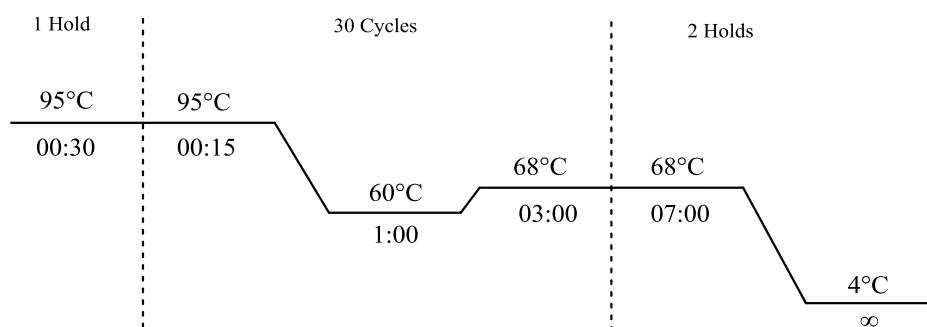
Scheme 5-16. Typical PCR Cycle used for overlap extension PCR

The vector and the PCR product were double digested with the correct restriction enzymes. The digested PCR product and the digested vector were gel purified. Mutated gene was ligated in the vector. This was followed by transformation in the appropriate strain of *E. coli* competent cells.

PCR Amplification. PCR amplification was done by using One Taq® Hot Start DNA polymerase (NEB). Following reaction conditions and amplification cycle were used according to the manufacturer's protocol.

Table 5-8. List of components in the reaction mixture for PCR amplification

Reaction Component	Amount (μL)
5x Reaction Buffer	10
10mM dNTP	1
10 μM Forward primer	1
10 μM Reverse primer	1
DNA polymerase (10U)	1
MilliQ	to 50



Scheme 5-17. PCR cycle used for amplification of PCR products

PCR Clean-up. PCR clean-up was accomplished by using QIAquick® Gel Extraction Kit. Five volumes of buffer PB was added to 1 volume of the PCR sample and was mixed. A QIAquick® spin column was placed in the provided 2 ml collection tube. To bind the DNA, the sample was applied to the column and was centrifuged for 30-60 s. The flow-through was discarded. The column was placed back into the same tube. 0.75 ml Buffer PE was applied to the column and was centrifuged for 30-60 s. The flow-through was discarded and the column was placed back in the same tube. The column was centrifuged for an additional 1 min to get rid of the residual ethanol from PE buffer. QIAquick® column was placed in a clean 1.5 mL microcentrifuge tube. To elute DNA, 50 µL of de-ionised water (pH 7.0-8.5) was applied to the center of the QIAquick® membrane and was centrifuged for 1 min.

DNA agarose Gel Electrophoresis. Agarose (0.5 g) was added to TAE buffer (50 mL) in a conical flask. The solution was heated in the microwave until the agarose was dissolved. Ethidium bromide (6 µL per litre of gel) was added once the mixture had cooled to 55 °C. The solution was poured into the gel holder, allowed to set and then placed into the well. Once set, TAE buffer was poured into the well, and then the samples were loaded. The gel

was run at 80 V (~90 mA), and then viewed under UV transmitter. The gel fragments were excised using a blade.

Gel Extraction of PCR product. The DNA fragment was excised from the agarose gel with a clean, sharp scalpel. The gel slice was weighed in a colourless tube. 3 volumes of buffer QG were added to 1 volume of gel (100 mg ~ 100 μ L). The gel slice was dissolved by incubating the tube at 50°C for 10 min. To help dissolve the gel, the contents were mixed by vortexing the tube every 2-3 min during the incubation. After the gel slice had dissolved completely, the colour of the mixture was found to be yellow (similar to buffer QG without dissolved agarose). If the colour of the mixture was orange or violet, 10 μ L of 3 M sodium acetate was added, pH 5.0, and mixed until the colour of the mixture became yellow. 1 gel volume of isopropanol was then added to the sample and mixed. A QIAquick spin column was placed in a provided 2 mL collection tube. To bind DNA, the sample was applied to the QIAquick column, and centrifuged for 1 min. The flow-through was discarded and QIAquick column was placed back in the same collection tube. 0.5 mL of buffer QG was applied to QIAquick column and centrifuged for 1 min for additional wash step to remove all traces of agarose. To wash, 0.75 mL of Buffer PE was added to QIAquick column and was centrifuged for 1 min. The flow-through was discarded and the tube was centrifuged for an additional 1 min at $\geq 10,000 \times g$ (~13,000 rpm). The flow-through was discarded before this additional centrifugation step. The QIAquick column was placed into a clean 1.5 mL microcentrifuge tube. To elute the DNA, 50 μ L of ddH₂O was applied to the center of the QIAquick membrane and the column was centrifuged for 1 min at maximum speed.

Restriction site Double Digestion. Typically, the reaction was set up as follows and incubated for 3-4 h in a 37 °C water bath. The reaction conditions were checked by using the Double Digest Finder Tool provided by NEB.

Table 5-9. List of components in the reaction mixture for double digestion reactions

Component	Recommended conditions
DNA template	1 µg
10x NEB Buffer (3 or 4)	1x final conc.
BSA	100 µg/ mL
NdeI	10 U
BamHI	10 U
MilliQ	To 50 µL

Ligation using Quick Ligation™ Kit. Ligation reactions were set up according to the manufacturer's instructions. Typically, 50 ng of vector was used for ligation with the desired gene. The amount of insert for the reaction with 50 ng vector was calculated using the following equation: [ng insert = (insert size (kb)/ vector size (kb))* ng vector]

Table 5-10. List of components in a typical ligation reaction

Component	Conditions used
Vector	50 ng
Gene of interest	3 molar excess w.r.t. vector
Quick Ligase buffer	1x
Quick Ligase	1 µL
MilliQ	To 20 µL

Chapter 6 Summary

Incorporation of metal catalysts into biomolecules, such as proteins, DNA etc., to perform non-native organic transformations underpins the concept of artificial metalloenzymes. The merger of complementary properties of enzyme and organometallic catalysts can lead to new strategies in synthetic chemistry. Our goal was to create artificial metalloenzymes for chemical cross-coupling reactions. Serine protease subtilisin from *Bacillus lentus* (SBL) was chosen as the model enzyme. The key enzyme-substrate (ES) intermediate is often regarded as an important mechanistic element in enzyme catalysis. Inspired by natural enzyme catalysis, a bisubstrate localisation strategy was used to demonstrate non-natural cross-coupling activity of SBL-based artificial metalloenzymes. We envisaged that the active serine could act as a nucleophile and help bring together the substrates with the metal centre facilitating the cross-coupling reaction. More importantly, our aim was to utilise the enzyme-substrate interactions of the hydrophobic S1 pocket of SBL to localise the coupling partners. This necessitated interlacing of the two catalytically active manifolds – one derived from the enzyme scaffold (SBL) and the other from a metal catalyst (palladium) incorporated into the enzyme.

In Chapter 2, we describe the design and construction of SBL-based artificial metalloenzymes. We explored two different approaches for chemical installation of unnatural amino acids in the enzyme, namely, direct alkylation of a cysteine (Cys) residue and conjugate addition at a dehydroalanine (Dha) residue. The Dha route, i.e, chemical conversion of Cys to Dha followed by installation of *N,N*-chelating moieties by subsequent addition of thiol ligands, was used for the first time to create artificial metalloenzymes.

Furthermore, direct alkylation of single cysteine residue was also successfully employed to install different metal binding amino acids into SBL to synthesize two apoenzymes, SBL-Bpc217 and SBL-Dpc217. As part of our interest to explore Pd(II)-catalysed cross-coupling reactions, palladium complexation with these apoenzymes was undertaken. This led to the creation of two model SBL-based metalloenzymes, SBL-Bpc217•Pd and SBL-Dpc217•Pd. Crucially, LC-MS, ICP-MS and microPIXE analyses along with the negative control reactions confirmed the metal binding selectivity and stoichiometry. Further, the properties of these metal binding sites were studied by using different metal precursors and titration studies.

Based on preliminary modelling studies, a novel *p*-vinylphenylmethanesulfonyl analogue (PMSv) was synthesized to act as the coupling partner in an oxidative Heck cross-coupling reaction. Enzyme selectivity for this PMSv substrate due to the strong nucleophilic interactions as well as substrate specificity of the hydrophobic S1 pocket was demonstrated by competition loading tests in the presence of similar sulfonyl substrates (Chapter 3). Given these results, we began our investigation into monitoring cross-coupling reactions in the enzyme active site by mass spectrometry. The mechanism of action of the SBL-based metalloenzymes was built up step-by-step by detecting various intermediates for nucleophilic enzyme catalysis and metal catalysis. The final coupled product was isolated after a hydrolysis step and was further characterized by HPLC, MS and MS/MS analyses of the small molecule. We have engineered a single turnover cross-coupling reaction using SBL-based metalloenzymes, and on-protein conversions for oxidative Heck cross-coupling reaction were ascertained based on a LC-MS assay. Several reloading strategies for re-entry into a second round of cycle were also explored.

Furthermore, circular dichroism studies were carried out to analyse the secondary structures of the different intermediates along the engineered cycle.

Crystallization experiments for several intermediates of oxidative Heck cross-coupling cycle of the two SBL-metalloenzymes were undertaken, along with soaking experiments with the metal and both the coupling partners (alkenyl substrate and arylboronic acids). Extensive screening led to optimized conditions for protein crystallization. The crystal structures of **SBL-WT(PMSv)** and **SBL-Bpc217·Ni** confirmed the successful positioning of the metal and the substrates as designed. In the crystal structure of **SBL-WT(PMSv)**, the vinyl moiety of the bound PMSv substrate was observed pointing at the back of the S1 pocket in the opposite direction from the S1' pocket (Chapter 4). Moreover, X-ray crystal structure analysis also revealed some flexibility of the bound alkene substrate in the enzyme active site. This indicated that the enzyme and the metal catalyst might together play a role in orienting the substrates in a productive conformation for the subsequent M-catalysed reaction. The alkyl chain moiety (that tethers these amino acids to the protein) in Dpc differs from Bpc by two atoms. Our X-ray crystallographic studies indicate that this longer tether group may be responsible for increased flexibility of Dpc. This could account for the incomplete or lack of Dpc density in our crystal structures.

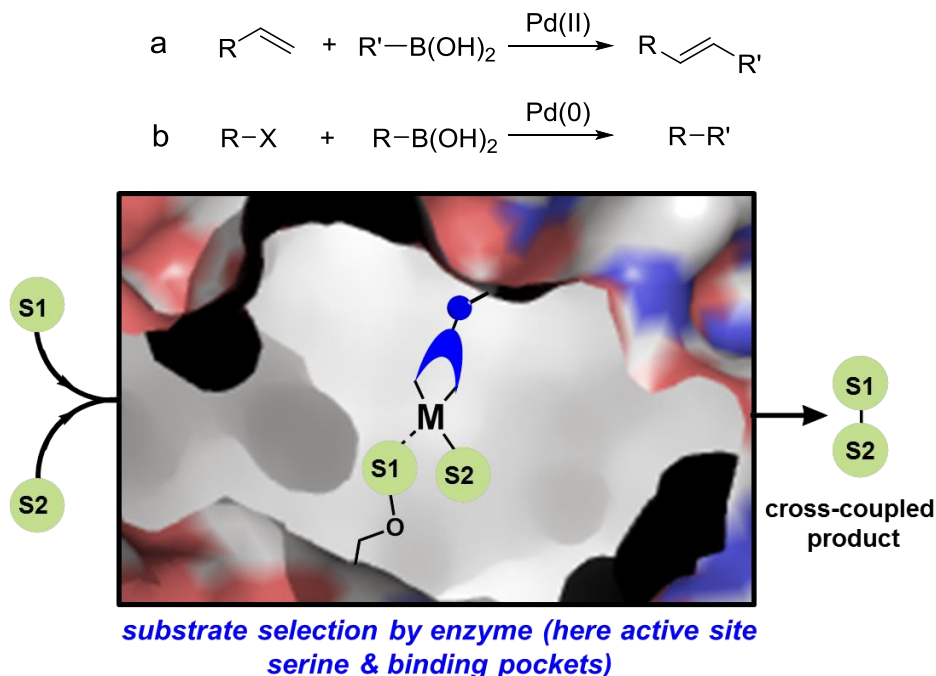
Having delineated a proposed cycle along with a design/construction methodology, we embarked on designing artificial metalloenzymes using another well-characterized serine hydrolase, AmpC β -lactamase (Chapter 5). Phenylboronic acids (that act as coupling partners in several metal-catalysed coupling reactions) have been found to be competitive and reversible inhibitors of serine proteases and serine β -lactamases. We reasoned that the artificial metalloenzyme active site could provide for an activating microenvironment by

reversibly binding with the boronic acid substrates for the chemical cross-coupling reactions. Six active site residues were selected as targets for site-directed mutagenesis for our studies. The *ampC* gene was inserted in a convenient pET vector and mutagenesis experiments were successfully carried out that led to the generation of the desired single cysteine AmpC-variants, Y150C, N152C, Y221C, K315C, T316C and N346C. Protein expression and purification of these active-site variants and the wild-type enzyme was optimized. A ‘tag-and-modify’ methodology was employed for stepwise assimilation of various components (enzyme, metal binding amino acid, metal catalyst) to ensure variability at each level for future fine-tuning and activity optimization. The chemical modification of these mutants to install a metal coordinating ligand was successful in yielding a library of metal-free apoenzymes.

Future directions

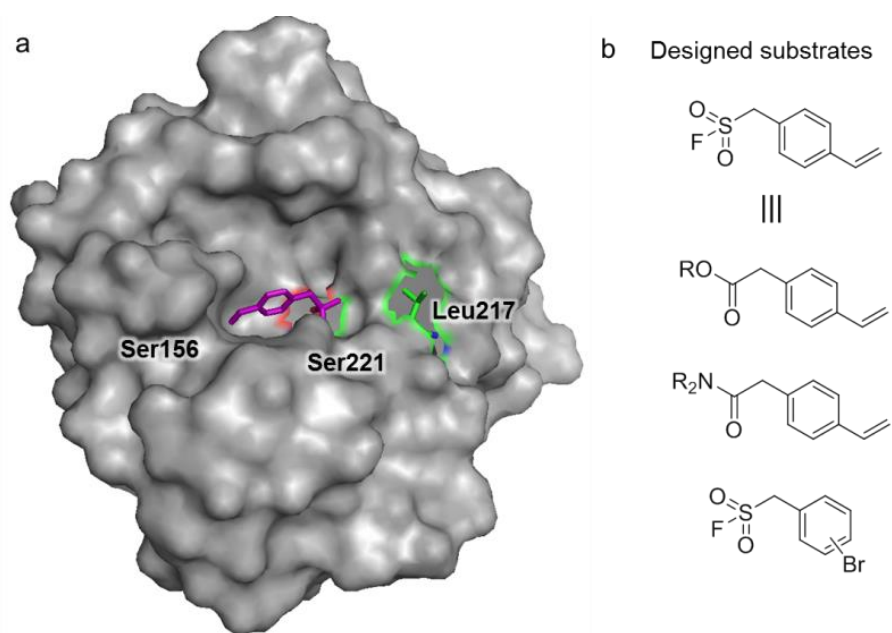
Overall, active-site engineering allowed for the installation of chelating amino acids for site-selective metal coordination leading to creation of artificial metalloenzymes. In order to exploit the enzyme’s affinity for its substrates and carry out metal-catalysed cross-coupling reactions, bisubstrate localisation has been introduced for the design of functional metalloenzymes. Such bio-cooperative metal catalysis presents an attractive route for performing important organic reactions enzymatically. The inherent substrate selectivity of the enzyme forms a constitutive element in our design of SBL or AmpC-based metalloenzymes. Taking advantage of the natural serine hydrolase activity, multi-turnover cross-coupling reactions using the SBL and AmpC-based artificial metalloenzymes can be explored. The ‘oxyanion hole stabilisation’ along with the innate hydrolytic activity of these enzymes could help achieve a full catalytic turnover. In particular, Pd(II) catalysed

oxidative Heck cross-coupling, and Pd(0) catalysed Mizoroki-Heck and Suzuki-Miyaura cross-coupling reactions can be explored.



Scheme 6-1. Palladium catalysed cross-coupling reactions (a) Oxidative Heck (b) Suzuki-Miyaura

Specifically, the design of SBL-metalloenzymes is based on tailoring the S1 and S1' enzyme pockets for substrate localisation and metal installation, respectively. The catalytic machinery of SBL (Asp32-His64-Ser221) is flanked by these substrate binding pockets on both sides. The S1 pocket contains the hydrophobic crevice giving rise to selectivity for PMS like substrates, whereas the S1' pocket contains the metal catalyst. Based on the crystal structure of **SBL-WT(PMSv)** reported in this thesis, we propose covalent yet reversible interactions of designed ester/amide substrates can be exploited to carry out multi-turnover oxidative Heck cross-coupling reactions. Another interesting direction would be to study the different reactive intermediates in a Suzuki-Miyaura cross-coupling reaction using LC-MS using covalent interactions of a PMSF-based aryl bromide substrate (Scheme 6-2).



Scheme 6-2. (a) Crystal structure of **SBL-WT(PMSv)**; SBL surface in grey depicting position 217 and 156 and the bound substrate (pink) are shown (b) Substrate analogues for oxidative Heck and Suzuki-Miyaura cross-coupling reactions



Title: Control of the Coordination
Structure of Organometallic
Palladium Complexes in an apo-
Ferritin Cage

Author: Satoshi Abe, Jochen Niemeyer,
Mizue Abe, et al

Publication: Journal of the American
Chemical Society

Publisher: American Chemical Society

Date: Aug 1, 2008

Copyright © 2008, American Chemical Society

Logged in as:
Mahima Sharma
Account #:
3000833816

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW



Title: Coordinated Design of Cofactor and Active Site Structures in Development of New Protein Catalysts

Author: Takafumi Ueno, Tomomi Koshiyama, Masataka Ohashi, et al

Publication: Journal of the American Chemical Society

Publisher: American Chemical Society

Date: May 1, 2005

Copyright © 2005, American Chemical Society

Logged in as:
Mahima Sharma
Account #:
3000833816

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW



Title: Incorporation of a Phebox Rhodium Complex into apo-Myoglobin Affords a Stable Organometallic Protein Showing Unprecedented Arrangement of the Complex in the Cavity

Author: Yuh Satake, Satoshi Abe, Seiji Okazaki, et al

Publication: Organometallics

Publisher: American Chemical Society

Date: Sep 1, 2007

Copyright © 2007, American Chemical Society

Logged in as:
Mahima Sharma
Account #:
3000833816

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW



Title: A Dual Anchoring Strategy for the Localization and Activation of Artificial Metalloenzymes Based on the Biotin–Streptavidin Technology

Author: Jeremy M. Zimbron, Tillmann Heinisch, Maurus Schmid, et al

Publication: Journal of the American Chemical Society

Publisher: American Chemical Society

Date: Apr 1, 2013

Copyright © 2013, American Chemical Society

Logged in as:
Mahima Sharma
Account #:
3000833816

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW

**Title:** Structure-Based Approach for
Binding Site Identification on
AmpC β -Lactamase**Author:** Rachel A. Powers, Brian K.
Shoichet**Publication:** Journal of Medicinal Chemistry**Publisher:** American Chemical Society**Date:** Jul 1, 2002

Copyright © 2002, American Chemical Society

Logged in as:
Mahima Sharma

LOGOUT

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

BACK

CLOSE WINDOW