

Structure and Activity of Amide Bond Synthetases

Benjamin Rowlinson

M.Sc. by Research

University of York

Chemistry

October 2018

Abstract

The synthesis of amide bonds is one of the most common processes in pharmaceutical chemistry. Despite this, the production of amide bonds remains a very wasteful process, relying on coupling reagents that lead to poor atom economies. An alternative method for the production of amide bonds with high atom economies may be found in enzymatic biocatalysis. Exploration of natural paths to amide synthesis is key to the development of this. McbA is an ATP-dependent amide bond synthetase from *Marinactinospora thermotolerans*. McbA has been found to catalyse the production of amide bonds in β -carboline amides. The ability of McbA to accept a wide range of substrates has previously been demonstrated making it a promising candidate for the development of a broad-range amide bond forming catalyst. In this work is described the crystallisation of McbA and the determination of an McbA structure, as well as the identification and characterisation of a new homologue of McbA from *Streptoalloteichus hindustanus*.

Crystals of both a K483A mutant of McbA and the wild type McbA were produced and structures were determined. From the structure it was seen that McbA is remarkably similar to the ANL superfamily of adenyating enzymes with two distinct domains; a larger N-terminal domain and a smaller C-terminal domain. Two conformations of McbA were seen in the asymmetric unit, which have been termed the adenylation conformation and amidation conformation based on their similarity to the adenylation and thiolation conformations seen in the ANL enzymes. The structures offer potential insights for future engineering work.

ShABS was identified as a 51% homologue of McbA. The substrate specificity of this protein was screened and revealed a broad specificity similar to McbA but with some complementary activity.

The discovery and characterisation of these enzymes may represent the beginnings of an enzymatic toolbox for amide bond synthesis.

Contents

Abstract	2
List of Tables	5
Table of Figures	6
List of Ligands and Their Numbering in This Work	11
Acknowledgements	12
Declaration.....	13
1. Introduction	14
1.1. The Problem with Current Methods of Amide Bond Formation	14
1.2. How Does Nature Make Amide Bonds?.....	17
1.2.1. Industrial Amide Synthesis Using Biocatalysis	17
1.2.2. Non-Ribosomal Peptide Synthetases (NRPSs).....	19
1.2.3. The ANL Superfamily of Adenylating Enzymes.....	20
1.2.4. Amide Bond Formation Catalysed by Carboxylic Acid Reductases (CARs)	24
1.2.5. The Amide Bond Synthetases	25
1.3. The Amide Bond Synthetase McbA Could Offer a Potential Solution to the Problems of Amide Bond Synthesis	27
1.4. AIMS.....	30
2. Methods	31
2.1. Expression and Purification of McbA and Homologues	31
2.1.1. Transformation into BL21 (DE3) Cells.....	31
2.1.2. Gene Expression.....	31
2.1.3. Immobilised Metal Affinity Chromatography (IMAC) Purification of McbA and Homologues.....	32
2.1.4. Size Exclusion Chromatography (SEC) Purification of McbA and Homologues .	32
2.1.5. Cleavage of McbA His-tag	33
2.2. Characterisation McbA and homologues.....	33

2.2.1. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)	33
2.2.2. Analytical Size Exclusion Chromatography Analysis of McbA	34
2.2.3. Polyacrylamide Gel Electrophoresis (PAGE) Analysis of McbA with Different Ligands	35
2.3 Crystallisation of McbA	36
2.3.1. 96-well 2-Drop MRC Crystallisation Plates	36
2.3.2. Crystal optimisation	36
2.3.3. Seeding	37
2.3.4. Crystal Testing, Data Collection and Structure Building	37
2.4. Producing a D201A McbA mutant	38
2.5 Screening <i>ShABS</i> Substrates	41
2.5.1. Small scale screens	41
2.5.2. 10 mg scale up reaction	41
3. Results and Discussion	43
3.1. McbA Structure Determination	43
3.1.1. Expression and Purification of McbA and Homologues	43
3.1.2. Crystallisation of Mutant and Wild-Type McbA	47
3.1.3. Structure Building and Refinement	55
3.1.4. Investigation of the structure of McbA	59
3.1.5. Producing a D201A McbA Mutant	68
3.2. Characterisation of a New Amide Bond Synthetase, <i>ShABS</i>	71
3.2.1. Expression and Purification of <i>ShABS</i>	73
3.2.2. Probing the Substrate Scope of <i>ShABS</i>	74
3.2.3. Investigating the Enantioselectivity of <i>ShABS</i>	76
4. Conclusion	79
References	81

List of Tables

Table 1. Protocol for resolving and stacking layers of a 12% SDS-PAGE.....	33
Table 2. Recipe for resolving and stacking layers of an 8.75% Native-PAGE.....	35
Table 3. Contents of PCR mix used for QuickChange mutagenesis to generate D201A McbA mutant.....	39
Table 4. PCR conditions for the first phase of QuickChange PCR for generation of the D201A McbA mutant. Annealing temperature starts at 10 °C above T_m and is decreased by 1 °C each cycle for 10 cycles.	40
Table 5. PCR conditions for the second phase of QuickChange PCR for generation of the D201A McbA mutant. The annealing temperature is equal to the calculated T_m for the primers.....	40
Table 6. Refinement statistics for the K483A McbA structure in complex with AMP and 3 . Values in brackets refer to data for the highest resolution shells.	56
Table 7. Refinement statistics for the wt-McbA structure in complex with 3 and AMP. Values in brackets refer to data for the highest resolution shells.	58
Table 8. Conversions after 24 h with 1: 1.2 molar ratio of acid to amine. Conversions are shown for McbA, <i>AcABS</i> and <i>ShABS</i> . McbA and <i>AcABS</i> conversions are shown as a comparison and had been previously determined.....	75

Table of Figures

Figure 1. Examples of major pharmaceuticals containing amide bonds with the amide bonds highlighted in red. Atorvastatin, Lidocain and Imatinib are shown which are used to lower cholesterol, treat cancer and as an anaesthetic respectively.	14
Figure 2. Amide couplings using common coupling reagents. (A) scheme for amide coupling using EDC and HOBT (B) scheme for amide coupling using Propylphosphonic anhydride (T3P).....	16
Figure 3. Hydrolytic amide formation catalysed by NHase.	18
Figure 4. Synthesis of Bz-Ala-Tyr-NH ₂ from Bz-Ala-OMe and Tyr-NH ₂ using a proteinase from the <i>Thermus</i> Rt41A.	18
Figure 5. Synthesis of ampicillin catalysed by a penicillin acylase.....	18
Figure 6. Amide bond synthesis catalysed by the lipase CAL-B aided by the removal of water using low pressure.....	19
Figure 7. Schematic representation of a generic three module NRPS and their actions. M1- initiation module M2 and M3- Elongation modules. The first amino acid is coupled to the PCP domain via an adenylate intermediate generated by the adenylation domain using ATP. The condensation domain then forms the amide bond between the next amino acid and the activated thioester. Amino acid couplings are terminated by the thioesterase domain that releases the peptide from the NRPS.....	20
Figure 8. Reactions catalysed by different members of the ANL superfamily. Each reaction begins with the formation of an adenylate with ATP. The activated adenylate then proceeds to the next step in the reaction. NRPSs couple the adenylate to a PCP domain in a thiolation reaction. Acyl-CoA synthetases couple the adenylate to CoA in a thiolation reaction. Firefly luciferase catalyses a decarboxylation of the adenylate to form oxyluciferin.....	21
Figure 9. Adenylation domains of ANL enzymes showing the different conformational states. N-terminal domain is shown with a green surface and C-terminal domain is shown with a coral coloured surface. A. The crystal structure of firefly luciferase in the adenylation conformation (PDB:1LCI) ^[27] . B. The crystal structure of gramicidin synthetase S in the thiolation conformation with AMP and Phenylalanine bound in the active site (PDB code:1AMU) ^[33]	22
Figure 10. The active site of gramicidin synthetase S with AMP and phenylalanine bound. The interaction of Lys517 with both AMP and phenylalanine is highlighted (PDB code: 1AMU) ^[33] , pH 6.5.....	23

Figure 11. Alternate reactions catalysed by the carboxylic acid reductases. The adenylation ‘A’ domain is shown in red with the large N-terminal domain and smaller C-terminal domain. The peptidyl carrier protein ‘PCP’ domain is shown in blue and the reductase ‘R’ domain is shown in pink.....	25
Figure 12. The native reactions carried out by the amide synthetases NovL, CouL and CloL in the synthesis of novobiocin, coumermycin A and clorobiocin. Only the first coupling reaction carried out by CouL is depicted. Amide bonds are highlighted in red.	26
Figure 13. Pictet-Spengler cyclisation catalysed by McbB.....	27
Figure 14. The native reaction carried out by McbA. The acid is first activated by adenylation and the adenylate is then able to undergo amidation.	28
Figure 15. Some of the carboxylic acid substrates accepted by McbA.	29
Figure 16. Primers used for QuickChange mutagenesis to produce the D201A McbA mutant.	39
Figure 17. A comparison of McbA purity using different expression and purification methods analysed by SDS-PAGE. A. SDS-PAGE analysis of McbA after SEC expressed in LB media and with a protease inhibitor used for lysis. 1: molecular weight marker. 2-6: fractions from SEC purification. B. SDS-PAGE analysis of McbA expressed in TB media. 1: molecular weight marker. 2-12: fractions from SEC purification. Calculated MW for McbA is 53 kDa	44
Figure 18. Chromatogram from SEC McbA purification with aggregate and monomer peaks labelled as determined by analytical SEC	45
Figure 19. Analytical size exclusion chromatogram of McbA A : Directly after SEC B : 3 d after SEC. The first peak to elute from SEC was run and is shown in orange and the second peak is shown in blue. Standards are shown in grey. Calculated MW for McbA is 53 kDa. The second peak (blue) corresponds to the mass for the monomeric form of McbA. The first peak (orange) has a peak for the monomeric form of McbA as well as a peak at >700 kDa. B shows very little aggregation occurs at 4°C. Standards: 1 . Thyroglobulin bovine MW: 670 kDa 2 . γ -globulins from bovine blood MW: 150 kDa 3 . Albumin chicken egg grade VI MW: 44.3 kDa 4 . Ribonuclease A type I-A from bovine pancreas MW: 13.7 kDa 5 . P-aminobenzoic acid	46
Figure 20. Summary of co-crystallisation attempts with each construct and each ligand used. Attempted conditions are shown with a green tick and conditions not attempted are shown with a red cross. The Carboxylic acid substrate used was 3 . For the product both 3a and 9a were attempted. The ATP carba-analogue was α,β -Methyleneadenosine 5'-triphosphate 18 .	48
Figure 21. Structure of α,β -Methyleneadenosine 5'-triphosphate used as an ATP mimic.	49

Figure 22. Native-PAGE analysis of McbA in complex with the carboxylic acid substrate 3 , the naphthoic acid product 10a , ATP, AMP, ATP with the carboxylic acid substrate 3 , β -mercaptoethanol (β -Me) and dithiothreitol (DTT).	49
Figure 23. Examples of salt crystals obtained in initial screens. A: CSS F8 condition (1.8 M Li ₂ SO ₄ , MES pH 5.5) with wt-McbA. B: Morpheus C4 condition with K483A McbA. C: AMP aluminium fluoride contamination. D: Index D9 condition with his cleaved McbA, AMP and 3	50
Figure 24. Diffraction pattern from a salt crystal grown in Salt RX screen.....	51
Figure 25. K483A McbA crystals in complex with the carboxylic acid substrate 3 and AMP. Grown in PDB minimal H12 condition (1.6 M Na citrate pH 7.5)	52
Figure 26. Diffraction patterns from K483A McbA crystals in complex with the carboxylic acid substrate 3 and AMP. A : In house diffraction showing only a few peaks at low resolution indication the presence of protein. B : Diffraction pattern from The Diamond Light Source synchrotron showing diffraction to ~ 2.8 Å.	52
Figure 27. A . small needle crystals of wt-McbA with 3 and AMP in 1.6 M Na citrate pH 7.5 B . wt-McbA crystals with 3 and AMP in 6 % Ethylene Glycol v/v, 1.6 M Na citrate pH 7.5	53
Figure 28. Diffraction patterns of wt-McbA crystals A :in house diffraction B : Diffraction pattern from The Diamond Light Source synchrotron showing diffraction to > 3 Å.....	54
Figure 29. The asymmetric unit from the wt-McbA crystal containing five chains with a P2 ₁ 2 ₁ 2 ₁ space group. Chains 1-4 are in the adenylation conformation and chain 5 is in the amidation conformation. Each chain has both AMP and 3 bound in the active site. The structure was refined to a resolution of 2.7 Å.....	59
Figure 30. Crystal structure of wt-McbA at 2.7 Å in A: The adenylation conformation, and B: The amidation conformation. The C-terminal domain is shown in coral and rotates 142° with respect to the N-terminal domain, shown in green, between the adenylation and amidation conformations.	60
Figure 31. A : The structure of wt-McbA in complex with AMP and 3 in the adenylation conformation shown in blue superimposed (using secondary structure match) with the structure of DhbE shown in green (2.28 Å over 450 C α atoms. PDB code: 1MDF). ^[36] B : The structure of the amidation conformation of McbA show in blue with 3 and AMP bound superimposed with the thiolation conformation of 4CBL (2.03 Å over 415 C α atoms PDB code: 3CW9). ^[37] C : The structure of the amidation conformation of McbA show in blue with 3 and AMP bound superimposed with a CAR adenylation domain shown in red (2.55 Å over 411 C α atoms. PDB code: 5MSS). ^[41]	62

Figure 32. **A:** The structure of 4CBL (4CW9) in the thiolation conformation with 4-chlorophenacyl-CoA bound (shown in green).^[37] **B:** The structure of McbA in the amidation conformation. **3** shown in green can be seen through the putative amine tunnel. Comparing the structures shows a similar tunnel in McbA to the 4-chlorophenacyl-CoA binding site in 4CBL but 4CBL contains positive residues (shown in blue) around this site whereas McbA contains negative residues (shown in red)..... 63

Figure 33. The structure of wt-McbA with AMP and **3** shown in green and D201 shown at the end of the putative amine tunnel..... 64

Figure 34. **A:** The interactions between K483 and the carboxylic acid substrate and AMP in the adenylation conformation. **3** is shown in blue and AMP is shown in yellow. The electron density for the 2mF_o-DF_c map of **3** is shown in blue at a level of 1 σ . The N-terminal domain is shown in green and the C-terminal domain in coral. **B:** The structure of wt-McbA showing the movement of K483 with the conformational change. The two conformations are overlaid with the adenylation conformation shown in green and the amidation conformation shown in coral..... 66

Figure 35. Active site of wt-McbA in the adenylation conformation with N-terminal domain residues shown in green and C-terminal domain residues shown in coral. The carboxylic acid substrate **3** is shown in blue and AMP is shown in yellow. The interactions between the carboxylic acid substrate **3** with G295 and F301 are shown. 67

Figure 36. Agarose gel analysis of QuickChange mutagenesis showing the plasmid between 6 and 8 kb, plasmid with insert is ~ 7 kb. 1: Quick-Load purple 1 kb DNA Ladder 2: empty lane 3-6: Quick change PCR products. 69

Figure 37. SDS-PAGE analysis of the purification of D201A McbA after SEC. 1: Broad range molecular weight marker 2-6: Fractions from SEC containing D201A McbA. Expected mass = 53 kDa 70

Figure 38. Sequence alignment of ShABS and McbA. Sequence identity 51%. Matched residues are shown between the sequences with similar residues shown as a +. 72

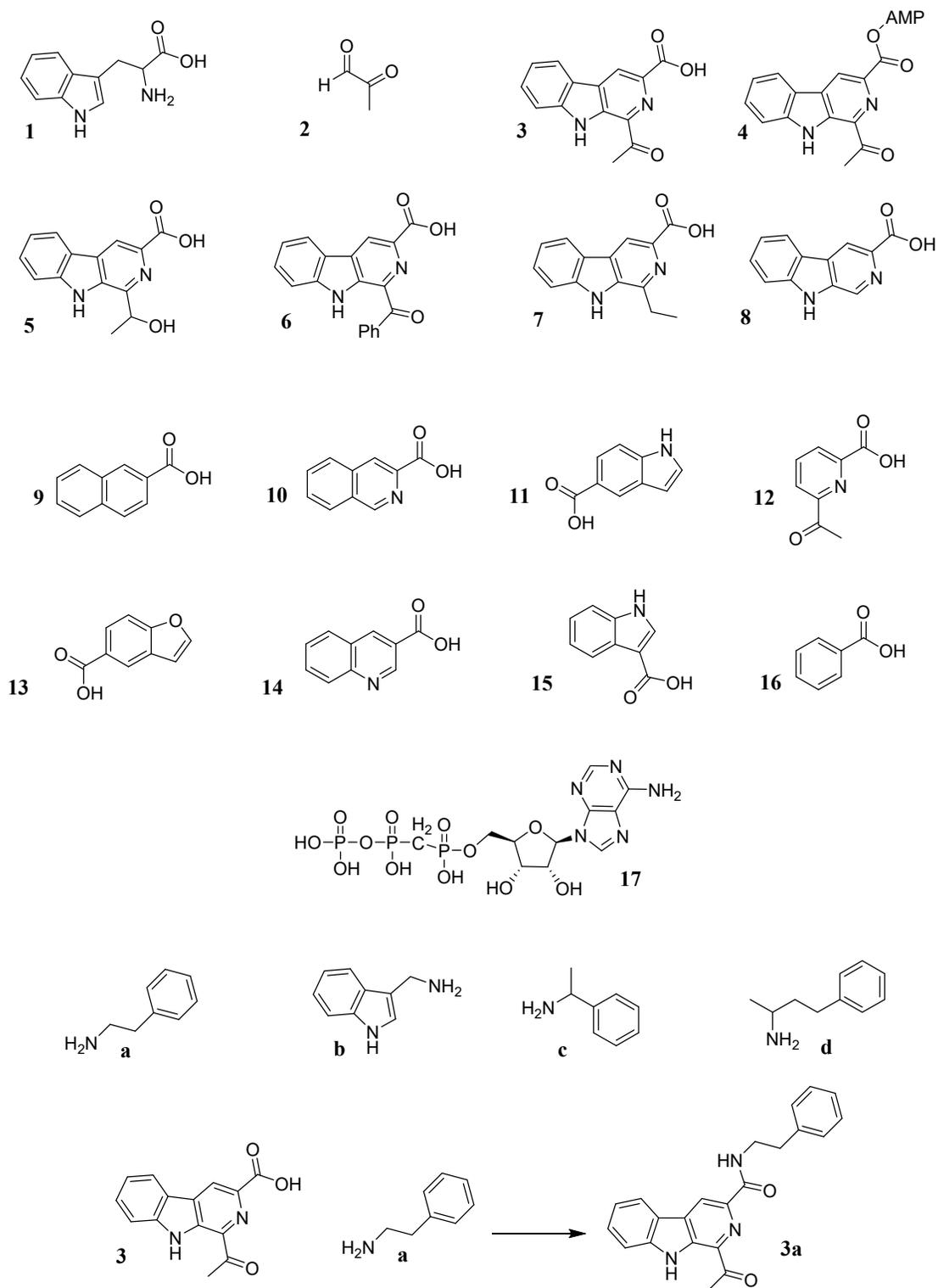
Figure 39. SDS-PAGE analysis of ShABS purifications. **A:** ShABS purification by IMAC 1: broad range molecular weight ladder. 2-9: Fractions from IMAC containing protein **B:** ShABS purification by SEC. 1: broad range molecular weight ladder 2: Pure ShABS protein from SEC. Calculated MW for ShABS= 54 kDa 73

Figure 40. Chromatogram from SEC purification of ShABS two peaks, the second of which has a prominent shoulder..... 74

Figure 41. HPLC analysis of the reaction of **3** and **d** with ShABS to form **3d** showing 100% conversion. Retention time for **3d**= 9.2 min..... 76

Figure 42. **A**: Chiral HPLC of the 10 mg scale up reaction of **3** and **d** with ShABS. **B**: Chiral HPLC of a racemic mixture of **3d**. Comparing **A** and **B** shows that ShABS catalyses the reaction with a 93% ee. 77

List of Ligands and Their Numbering in This Work



Amide products are given a number and letter based on their corresponding acids and amines

Acknowledgements

Thank you to Gideon Grogan for his excellent support and supervision during the project and to Mark Petchey for both production of McbA constructs, substrates and product standards as well as his help and encouragement during the project.

I would also like to thank Stephanie Dannevald and Amina Frese for the cloning of McbA. As well as Anibal Cuetos for the production of many of the constructs used and for initial screening of McbA.

Thanks to the YSBL support staff for keeping the labs running and for making this possible and to all of YSBL and the Grogan group for being so friendly and welcoming and making this project such an enjoyable experience.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

The initial cloning of McbA was carried out by Stephanie Dannevald and Amina Frese. Cloning of *AcABS* and generation of the K483A mutant was achieved by Mark Petchey. All other constructs were produced by Anibal Cuetos. Previous crystallisation attempts had been made by Anibal Cuetos and Mark Petchey. Homologues of McbA were identified by Professor Gideon Grogan. Screening of McbA with β -carboline acids was done by Anibal Cuetos. Further McbA substrate screening and screening of *AcABS* was done by Mark Petchey.

1. Introduction

1.1. The Problem with Current Methods of Amide Bond Formation

Amide bonds are highly prevalent amongst both pharmaceuticals and agrochemicals. They are found in a number of major pharmaceutical chemicals, such as atorvastatin, lidocaine and imatinib (Figure 1). The importance of bioactive amides means that amide bond formations account for 16% of all reactions carried out in medicinal chemistry laboratories.^[1]

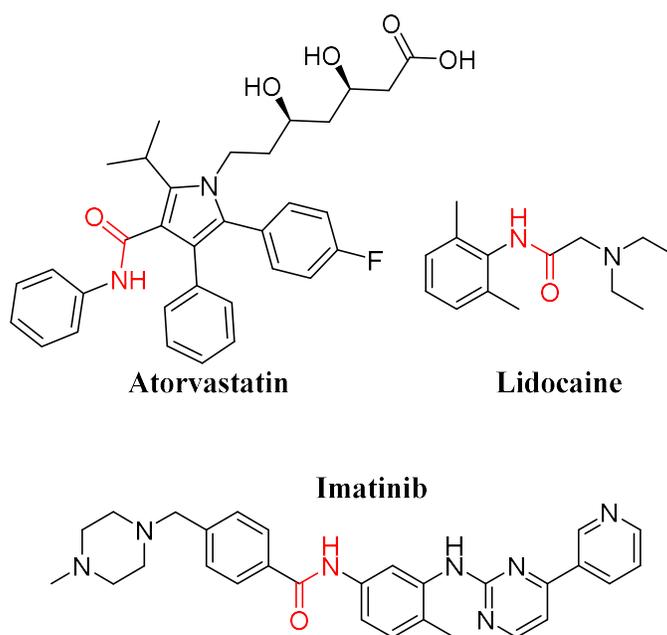


Figure 1. Examples of major pharmaceuticals containing amide bonds with the amide bonds highlighted in red. Atorvastatin, Lidocaine and Imatinib are shown which are used to lower cholesterol, treat cancer and as an anaesthetic respectively.

The synthesis of amide bonds is a relatively simple yet costly process and is usually achieved through the coupling of a carboxylic acid and an amine. This reaction is not spontaneous and requires very high heat (>200 °C) or, as is usually the case, activation of the carboxylic acid using coupling reagents.^[2-4] Coupling reagents tend to produce an acid chloride, anhydride or active ester intermediate. Acid chloride coupling has been used in the production of amide

bonds for many years and was the method adopted in the production of the first synthetic dipeptide.^[5] The acid chloride is traditionally generated using a highly toxic chlorinating agent such as thionyl chloride. This method is not compatible with many protecting groups and uses organic solvents such as dimethylformamide (DMF). Other coupling methods are now generally favoured over acid chloride coupling for most purposes. 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and hydroxybenzotriazole (HOBt) are commonly employed for the production of amide bonds.^[6] This reaction proceeds *via* the production of an EDC ester intermediate that goes on to form the HOBt activated ester with the carboxylic acid (Figure 2A). The activated ester is then free to undergo nucleophilic attack by the amine. The process also frequently involves the use of DMF, which the European chemicals agency have recently classified as a substance of very high concern.^[7] In addition to this, the explosion risk associated with HOBt, as well as the poor atom economy for the reaction as a result of the urea side product, means that this process is less than ideal.

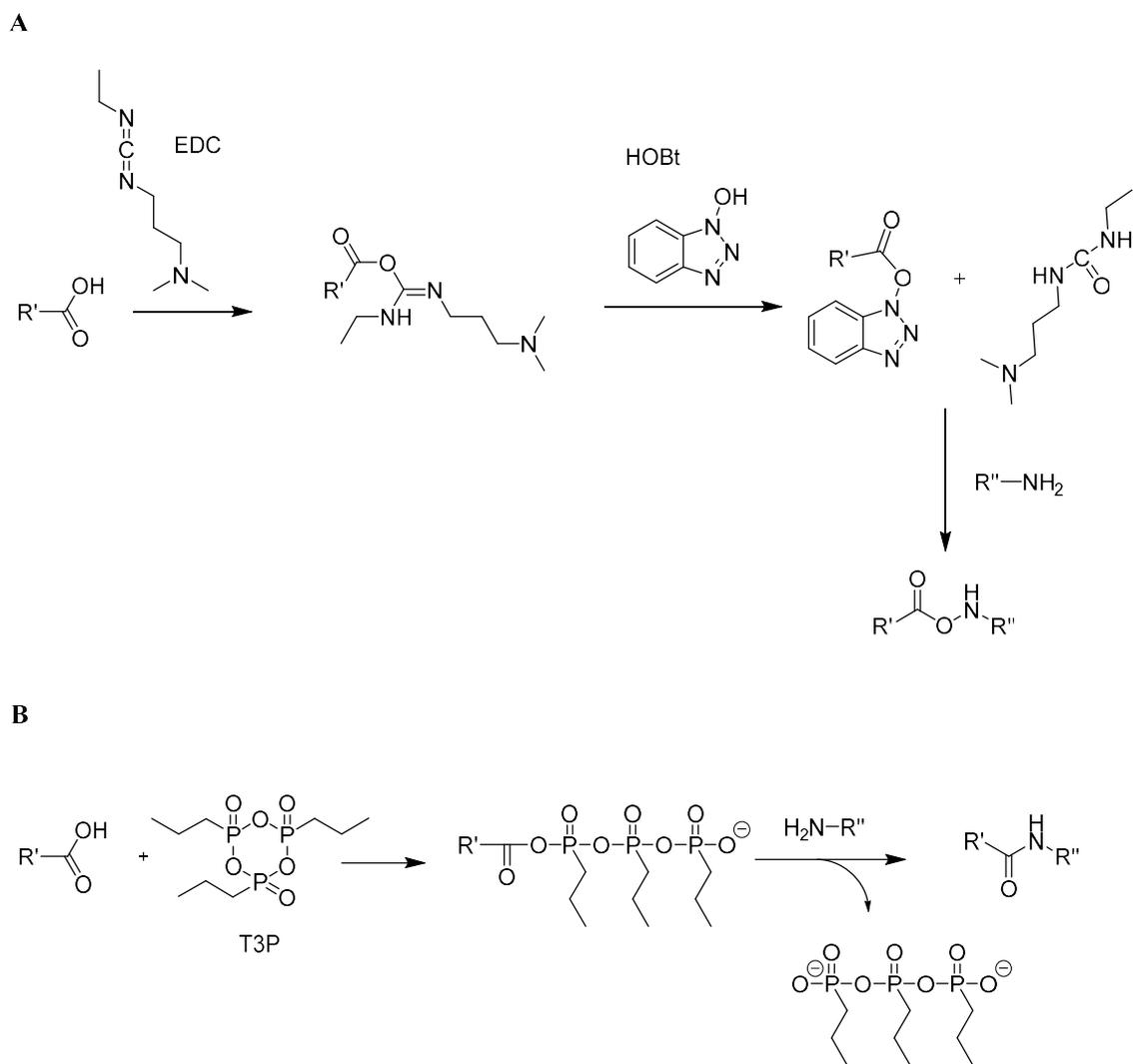


Figure 2. Amide couplings using common coupling reagents. **(A)** scheme for amide coupling using EDC and HOBT **(B)** scheme for amide coupling using Propylphosphonic anhydride (T3P).

An alternative to this is Propylphosphonic anhydride (T3P) coupling (Figure 2B), which has recently seen widespread industrial use due to its ability to suppress epimerization, its low toxicity and easy handling.^[4, 8] T3P reacts with the acid substrate to form an activated carboxylic-phosphoric-anhydride that can then undergo nucleophilic attack by the amine substrate. This process leads to the release of environmentally damaging phosphate waste which also means the reaction has a very poor atom economy.

In 2007 the issues associated with these coupling reagents led the American Chemical Society Green Institute to vote for amide bond formation using poor atom economy reagents as the number one issue facing organic chemistry.^[9] With this in mind potential solutions to this problem are now being explored. One potential solution to this issue lies in biocatalysis. An enzymatic catalyst for the synthesis of amides would be highly desirable and would solve many of the problems associated with traditional chemical methods. A biocatalytic approach would reduce the need for damaging solvents and dangerous coupling reagents. It would also be able to reduce the problem of epimerisation.

1.2. How Does Nature Make Amide Bonds?

Nature is able to make amide bonds for the production of proteins, peptides and small molecules. A range of enzymes are employed for this purpose. This provides a good starting point in the search for a broadly specific amide bond forming catalyst. One of the most obvious examples of amide bond formation in nature is found in ribosomal protein and peptide production. The ribosome catalyses the formation of the majority of amide bonds found in nature. This is a highly specific process involving the coupling of amino acids to tRNA by aminoacyl-tRNA synthetases. The ribosome which is made up of both protein and RNA then carries out the amide bond formation between the amino acids. The very high specificity and complexity of this system severely limits the use of the ribosome for biocatalytic purposes, however it may find some use in the production of ribosomal peptides with antibiotic activity.^[10, 11]

1.2.1. Industrial Amide Synthesis Using Biocatalysis

Some amide synthesis is already being carried out using enzymatic catalysis. This includes the synthesis of acrylamide from acrylonitrile that utilises a nitrile hydratase (NHase) (Figure 3). The use of NHase negates the need for high temperatures which are commonly employed in other methods of acrylonitrile synthesis.^[12] However, as a hydrolytic reaction this method is only able to produce primary amides and so is not applicable for amide bond formation.

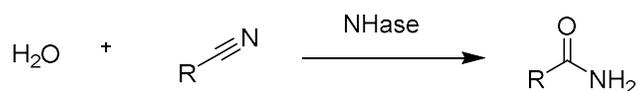


Figure 3. Hydrolytic amide formation catalysed by NHase.

Amide couplings have however been achieved through the use of proteases.^[13-16] This usually involves the removal of water from the system, which can be achieved through immobilisation of the enzymes such as in the synthesis of Bz-Ala-Tyr-NH₂ by a proteinase from a thermophilic organism *Thermus* strain Rt41A (Figure 4).^[14] This method for amide bond formation is limited by the selectivity of the proteases for amino acid substrates. Although some variability exists as shown by the ability to synthesise isopeptide bonds.^[17]

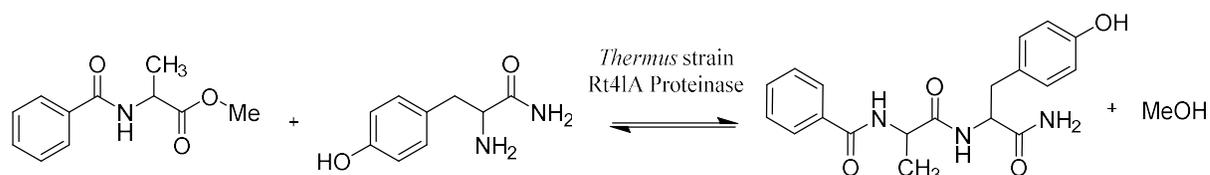


Figure 4. Synthesis of Bz-Ala-Tyr-NH₂ from Bz-Ala-OMe and Tyr-NH₂ using a proteinase from the *Thermus* Rt41A.

Another industrial process that has adopted enzymatic amide bond synthesis is in the production of penicillin and penicillin derivatives, such as ampicillin (Figure 5).^[18] This uses penicillin acylase to catalyse the amide coupling resulting in the loss of water.

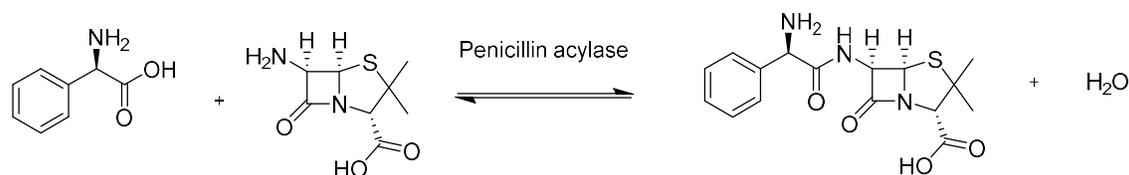


Figure 5. Synthesis of ampicillin catalysed by a penicillin acylase.

Amide bond formation has also been achieved through the use of lipases which have been shown to accept a wider range of substrates and to have a high enantioselectivity.^[10] Lipase B from *Candida antarctica* (CAL-B) was shown to catalyse amide bond formation without the use of activating agents or harsh solvents. This was achieved by pushing the equilibrium towards amidation through the removal of water with reduced pressure (Figure 6).^[19]

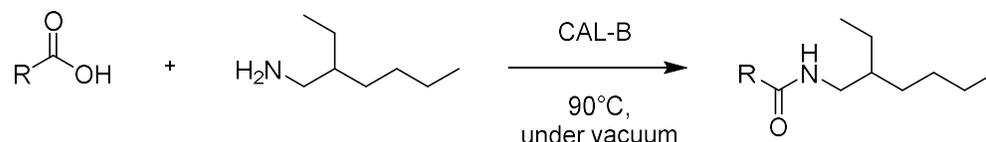


Figure 6. Amide bond synthesis catalysed by the lipase CAL-B aided by the removal of water using low pressure.

1.2.2. Non-Ribosomal Peptide Synthetases (NRPSs)

Non-ribosomal peptide synthetases (NRPSs) are protein complexes that feature enzymes able to form amide bonds for the production of peptides. These non-ribosomal peptides have a broad range of biological roles. Of particular note are their antibiotic roles, which means that these peptides may soon see widespread pharmaceutical use.^[20] Examples include vancomycin, which is currently employed as a drug of last resort against the growing threat of drug-resistant bacteria.^[21-23] NRPSs catalyse the formation of peptides but do not include many of the proofreading steps and specificity regulators associated with ribosomal peptide synthesis, such as the coupling by aa-tRNA synthetases. NRPSs contain three domains for amide bond formation, an adenylation domain, a peptidyl carrier protein domain and a condensation domain.^[21] The adenylation domain activates the acid by forming an acyl adenylate using ATP. The adenylate then forms a thioester with the pantetheine cofactor of the peptidyl carrier protein which transports the acid to the condensation domain where amide bond formation takes place (Figure 7).

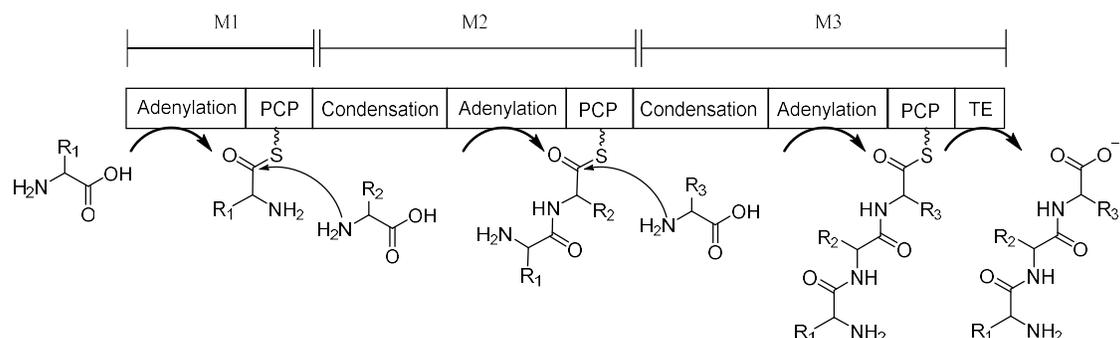


Figure 7. Schematic representation of a generic three module NRPS and their actions. M1- initiation module M2 and M3- Elongation modules. The first amino acid is coupled to the PCP domain *via* an adenylate intermediate generated by the adenylation domain using ATP. The condensation domain then forms the amide bond between the next amino acid and the activated thioester. Amino acid couplings are terminated by the thioesterase domain that releases the peptide from the NRPS.

The non-ribosomal peptides generated are often subjected to a range of post translational modifications such as crosslinking and glycosylation. The complexity and specificity of these systems mean NRPSs are poor candidates for broad spectrum *in vitro* biocatalysts. However, some promising engineering work summed up in Winn *et al.* may mean these enzymes can be used for industrial synthesis of peptide antibiotics.^[24]

1.2.3. The ANL Superfamily of Adenylating Enzymes

NRPSs can feature domains that are part of a larger family of enzymes known as the ANL superfamily or class 1 adenylate-forming enzymes. The ANL superfamily is named after the enzyme families it encompasses: acyl-CoA synthetases,^[25] NRPSs and luciferases.^[26, 27] The superfamily is made of these enzymes as they all share a degree of sequence similarity (~20%) and structural homology, as well as all catalysing an adenylation step as part of their mechanism.^[28-30] This does not include unrelated adenylating enzymes such as the NRPS-independent siderophore (NIS) adenylating enzymes or amino-acyl-tRNA synthetases.^[31, 32] All of the enzymes in this superfamily act *via* a two-step mechanism in which an adenylate is formed from ATP which activates the acid for coupling (Figure 8).^[28] In NRPSs and acyl-CoA

synthetases the adenylate is activated for the formation of a thioester with a pantetheine cofactor. In the luciferases, the adenylate is able to undergo an oxidative decarboxylation. The intermediate formed can then decompose to release a photon.

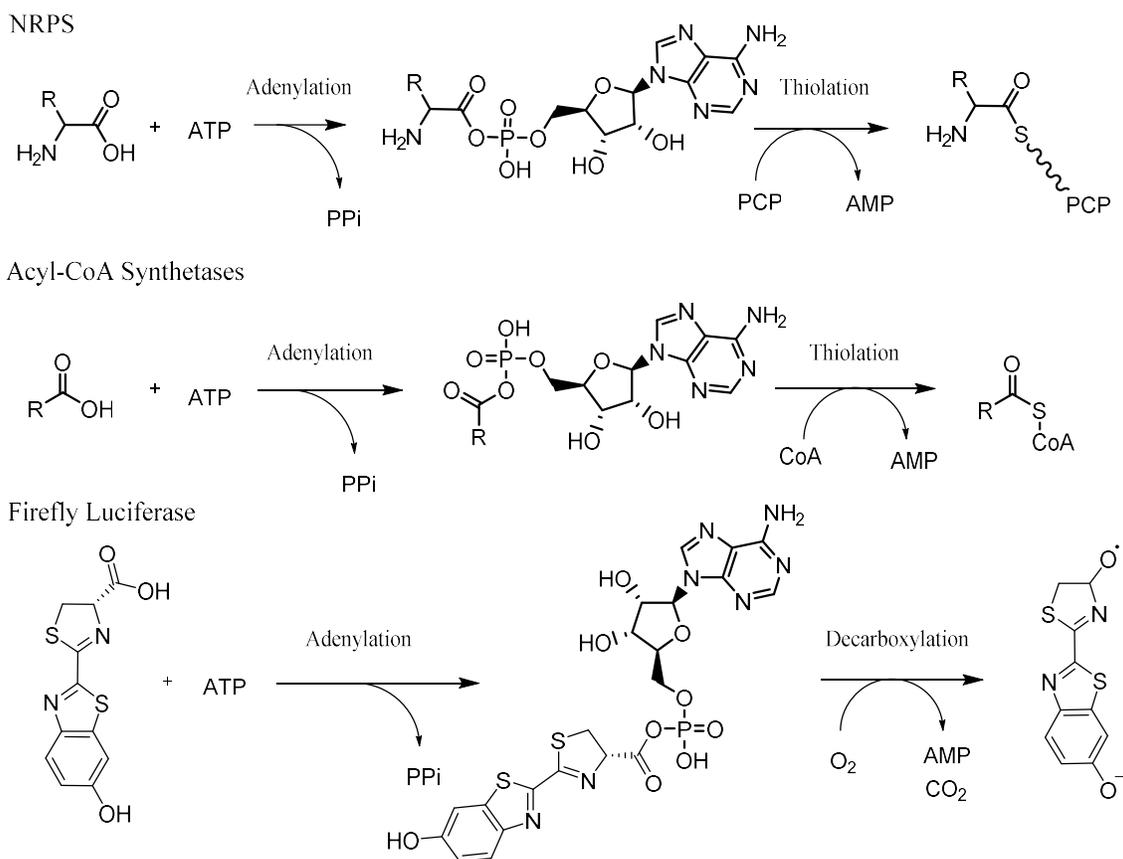


Figure 8. Reactions catalysed by different members of the ANL superfamily. Each reaction begins with the formation of an adenylate with ATP. The activated adenylate then proceeds to the next step in the reaction. NRPSs couple the adenylate to a PCP domain in a thiolation reaction. Acyl-CoA synthetases couple the adenylate to CoA in a thiolation reaction. Firefly luciferase catalyses a decarboxylation of the adenylate to form oxyluciferin.

Structural studies of these proteins have provided a good understanding of how they function. The first ANL enzyme to be crystallised was firefly luciferase which was detailed by Conti *et al.* in 1996 (Figure 9A).^[27] This allowed identification of two domains, a large N-terminal domain (4-436) and a smaller C-terminal domain (440-544). The unstructured region

connecting the two domains is referred to as the A8 loop.

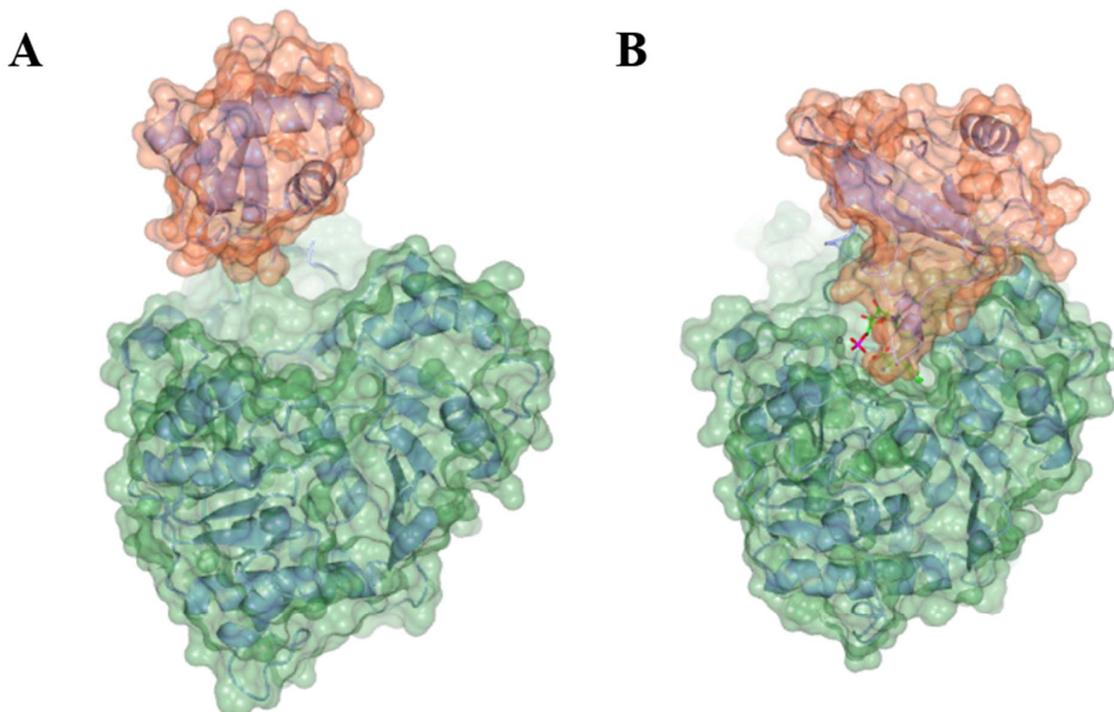


Figure 9. Adenylation domains of ANL enzymes showing the different conformational states. N-terminal domain is shown with a green surface and C-terminal domain is shown with a coral coloured surface. **A.** The crystal structure of firefly luciferase in the adenylation conformation (PDB:1LCI)^[27]. **B.** The crystal structure of gramicidin synthetase S in the thiolation conformation with AMP and Phenylalanine bound in the active site (PDB code:1AMU)^[33].

The structure did not contain a ligand and it was thought that the binding site, predicted by conserved sequence motifs,^[34] was too large to accommodate the substrate. This led to the idea that a large domain movement must occur upon substrate binding. This supported earlier research on adenyating enzymes that also predicted domain movement.^[28] Evidence for this domain movement came from the crystal structure of the NRPS enzyme gramicidin synthetase S which had the L-phenylalanine substrate and AMP bound in the active site (Figure 9B).^[33] The structure also showed a 94° rotation of the smaller C-terminal domain which demonstrates the domain movement that occurs on substrate binding (Figure 9). The domain movement

brings C-terminal domain residues into the active site. This allows interaction of Lys517 with both ligands, indicating a potential catalytic role (Figure 10).

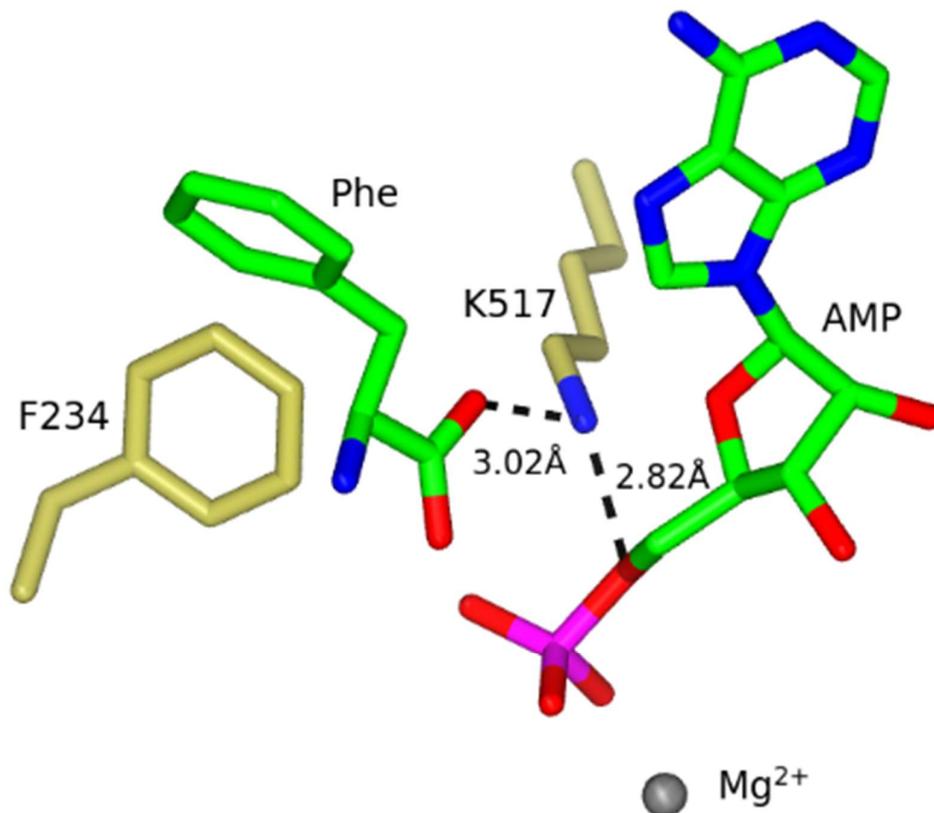


Figure 10. The active site of gramicidin synthetase S with AMP and phenylalanine bound. The interaction of Lys517 with both AMP and phenylalanine is highlighted (PDB code: 1AMU)^[33], pH 6.5

This finding was reinforced by the work of Horswill *et al.* who demonstrated that a Lys592 (corresponding to Lys517 in gramicidin synthetase S) knock out in propionyl CoA-synthetase resulted in a loss of activity.^[35] Additional structural studies have now confirmed these characteristics as common among all the ANL superfamily.^[36, 37] Along with these structural studies, kinetic studies have also helped to build the idea of an adenylation and thiolation conformation. The mutagenesis studies conducted by Wu *et al.* led to the identification of His207 as a key residue of interest.^[38] It was seen that in the adenylation conformation, His207

is able to interact with the A8 loop. It was also demonstrated that a mutation at this point causes a large decrease in activity (100-fold decrease in both adenylation and thiolation rates). It is believed that His207 blocks the CoA binding site in the adenylation conformation and is also important in stabilisation of the thiolation conformation.

1.2.4. Amide Bond Formation Catalysed by Carboxylic Acid Reductases (CARs)

Carboxylic acid reductases catalyse the reduction of a carboxylic acid to its corresponding aldehyde.^[39] Sequence analysis has shown CARs to be closely related to the ANL superfamily of enzymes, sharing ~20% sequence homology with some members of the ANL superfamily.^[40] This is not too surprising considering the similarity in the mechanisms seen between CARs and ANL enzymes. The CARs are multidomain enzymes comprising a reduction domain, a peptidyl carrier domain, and crucially an adenylation domain.^[41] In much the same way as ANL enzymes, the acid is first activated by reaction with ATP to form an adenylate. The adenylate can then proceed *via* the normal route, in which the adenylate forms a thioester with the phosphopantetheine (PPant) group, which transfers the acid to the reduction domain where it is reduced using NADPH. Alternatively, Flitsch *et al.* demonstrated that providing CARs with amine rather than NADPH allowed for the synthesis of amide bonds (Figure 11).^[42] It was then shown that this ability could be exploited to synthesise a range of amide products including ilepcimide, an anticonvulsant drug. A truncated form of CAR was then produced lacking the reduction domain.^[42] This new protein retained amidation activity but lost the reductase activity resulting in a complete change in function. This demonstrated that the reductase domain is not required for activity. A mutant CAR was generated without the PPant attachment site.^[42] This enzyme also retained amidation activity whilst losing reductase activity. These enzymes also avoid the issue of amidation activity being suppressed by NADPH as would be the case with other CARs.^[42] The potential of these enzymes as biocatalysts is currently restricted by the use of large amounts of the amine substrate, up to 100-fold excess, in an effort to drive the reaction towards the amide product.^[42] This may however be overcome by enzyme engineering.

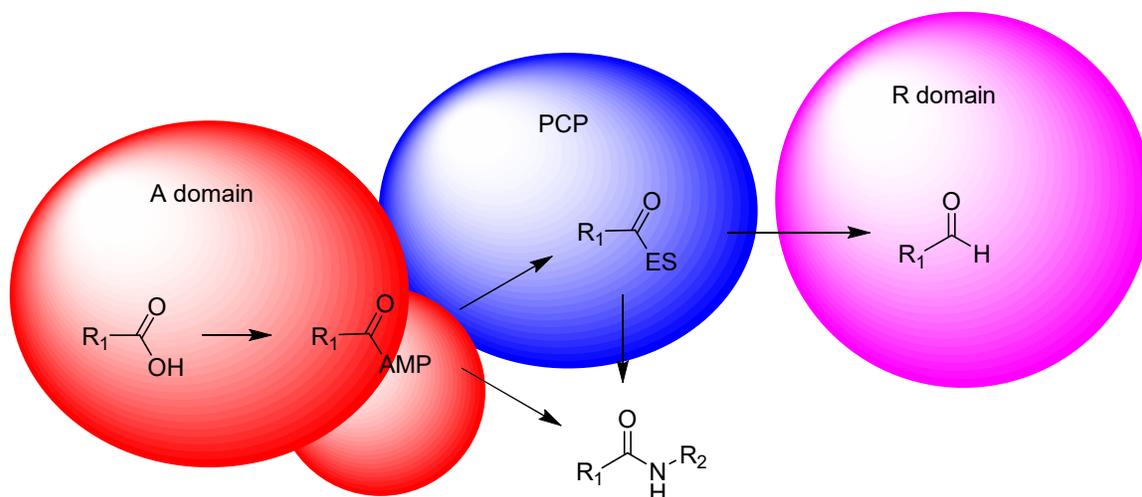


Figure 11. Alternate reactions catalysed by the carboxylic acid reductases. The adenylation ‘A’ domain is shown in red with the large N-terminal domain and smaller C-terminal domain. The peptidyl carrier protein ‘PCP’ domain is shown in blue and the reductase ‘R’ domain is shown in pink.

1.2.5. The Amide Bond Synthetases

A number of enzymes related to the ANL superfamily have been reported to directly catalyse amide bond formation *via* an adenylate intermediate. These enzymes have been labelled amide bond synthetases (ABSs). ABSs have been identified in the biosynthesis of the aminocoumarin antibiotics novobiocin,^[43, 44] clorobiocin,^[45, 46] and coumermycin A.^[47] These are synthesised by NovL, CloL and CouL respectively (Figure 12).

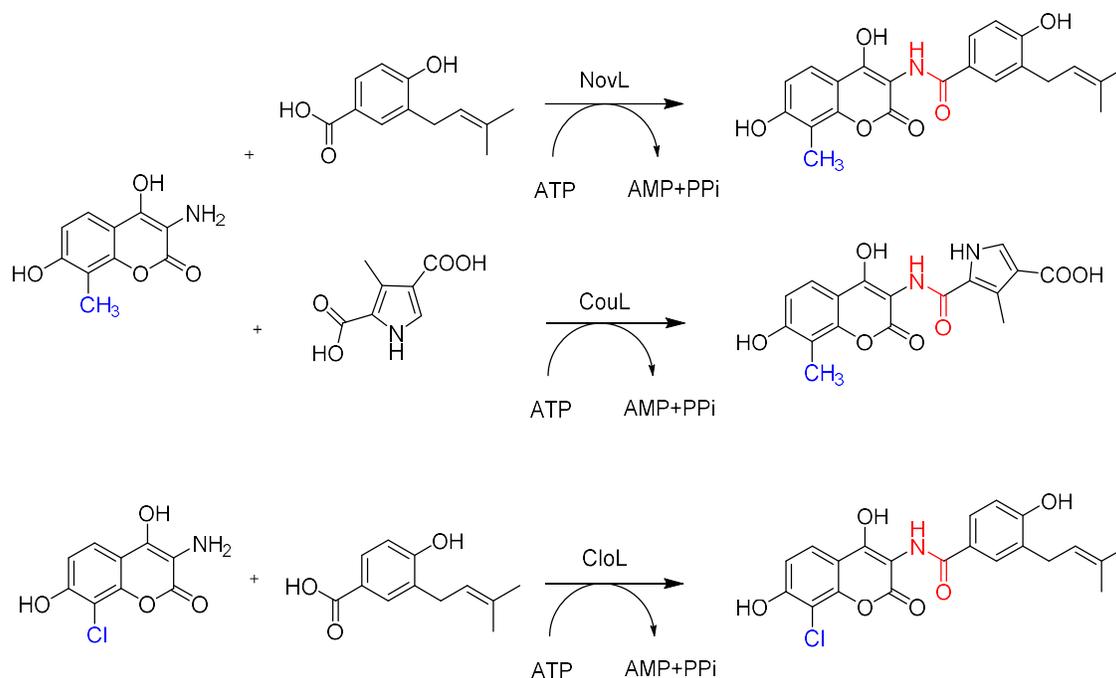


Figure 12. The native reactions carried out by the amide synthetases NovL, CouL and CloL in the synthesis of novobiocin, coumermycin A and clorobiocin. Only the first coupling reaction carried out by CouL is depicted. Amide bonds are highlighted in red.

Sequence analysis of these proteins reveals a striking similarity to the adenylation domain of the ANL superfamily. A number of conserved sequences indicate that these proteins consist of a larger N-terminal domain connected to a smaller C-terminal domain by an A8 loop, although no structural data are available to confirm or deny this hypothesis. Compared to other adenylation enzymes the ABSs are relatively simple and are able to catalyse both the adenylation and amidation steps in the same active site using only ATP and without coupling to a pantetheine cofactor or CoA. This simplicity is very promising for the development of a new biocatalyst for amide bond formation. In addition to this CloL has been shown to accept a wide range of substrates making these very promising candidates for engineering of a new biocatalyst.^[46] The use of these enzymes is however hampered by the lack of crystal structures available, making engineering more challenging, and by the relatively high specificity for the amine substrate.

1.3. The Amide Bond Synthetase McbA Could Offer a Potential Solution to the Problems of Amide Bond Synthesis

Recently a new ABS enzyme has been reported for the synthesis of β -carboline derivatives produced by the actinomycete *Marinactinospora thermotolerans* isolated from the South China Sea.^[48-50] Studies of actinomycetes have led to the discovery of many bioactive compounds and the expanded search into marine actinomycetes offers even more opportunities for new drug discoveries.^[51] Three genes were identified in the gene cluster responsible for β -carboline synthesis, *mcbA*, B and C. McbB was found to catalyse a Pictet-Spengler cyclisation (Figure 13).

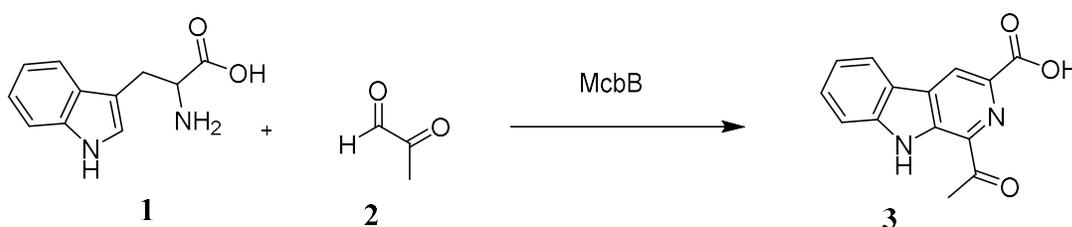


Figure 13. Pictet-Spengler cyclisation catalysed by McbB.

McbC and McbA were initially identified using sequence homology as a decarboxylase and CoA ligase respectively.^[49] However, it was demonstrated that a gene knock-out of McbA led to a build-up of 1-acetyl-3-carboxy- β -carboline **3** and failure to produce the amide product.^[49] This indicated that McbA played a role in amide bond synthesis. Building upon this finding Ji *et al.* were able to identify McbA as an ATP-dependent amide bond synthetase and demonstrated its ability to synthesise a diverse range of β -carboline derivatives.^[50] It is believed that McbA acts *via* an adenylate intermediate and both adenylation and amidation steps occur in the same active site (Figure 14), much like the aminocoumarin amide bond synthetases.

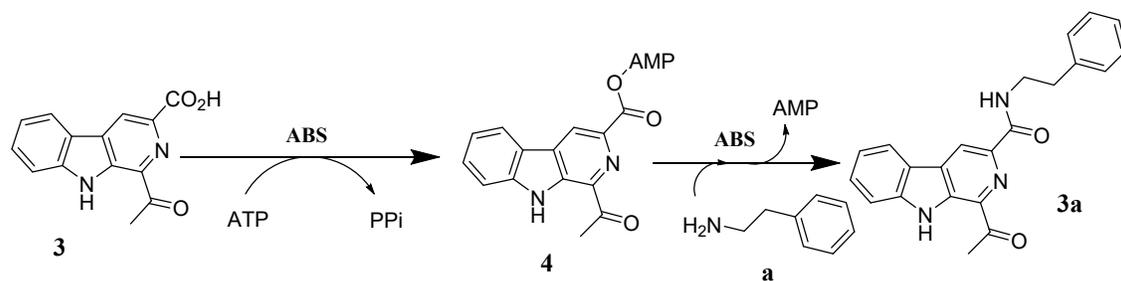
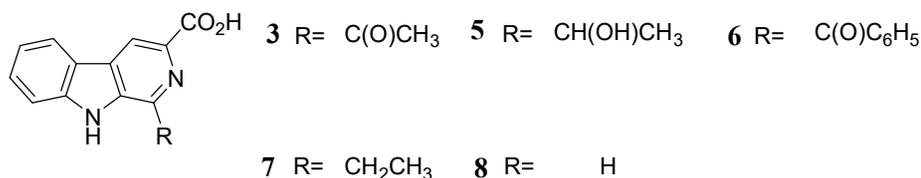


Figure 14. The native reaction carried out by McbA. The acid is first activated by adenylation and the adenylate is then able to undergo amidation.

The ability of McbA to produce amide bonds from a broad range of starting materials using only ATP, and without forming a coenzyme A thioester intermediate, makes it a very promising platform for the development of a new catalyst for amide bond formation. Our group purchased the gene for McbA, which was then successfully cloned and expressed. Further screening of substrates (carried out by Anibal Cuetos and Mark Petchey) has only shown greater promise for McbA as a general biocatalyst for amide bond synthesis, with an even broader range of substrates accepted (Figure 15). The acceptance of benzoic acid **16** is of particular interest and highlights the potential of McbA as a broad range catalyst.

β -carboline derivatives



Other carboxylic acid substrates

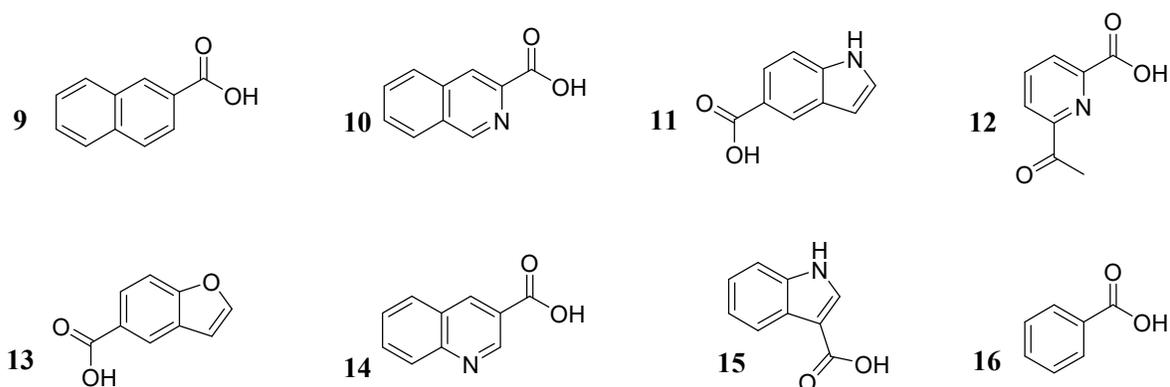


Figure 15. Some of the carboxylic acid substrates accepted by McbA.

Four homologues of McbA have also been identified by BLAST search. A 55% homologue, *AcABS*, from *Actinoalloteichus cyanogriseus* was also cloned and expressed by our group and was also shown to have amide bond synthetase activity. Another two homologues, *ShABS* and *NfABS* from *Streptoalloteichus hindustanus* and *Nocardioopsis flavescens* respectively, were also acquired and attempts to express and characterise these are described within. For McbA and homologues to be applied industrially, engineering of the proteins will need to be carried out to improve their activity towards a desired substrate. To this end a crystal structure of the enzyme would be extremely valuable and would allow for a better understanding of the mechanism as well as identification of key residues of interest for mutation.

Previous attempts by our group to crystallise McbA have proved to be unsuccessful. These have included extensive screening of crystallisation conditions, surface entropy reduction,^[52]

and separate domain crystallisation. A reduced activity mutant was also produced by mutating K483 to an alanine. K483 was believed to stabilise the phosphates of ATP, enabling attack by the carboxylic acid. These constructs were designed based on a model generated by the Robetta server.^[53] In this thesis, is described the crystallisation of McbA and the K483A mutant as well as the characterisation of the McbA homologue *ShABS*

1.4. AIMS

In order to inform the engineering of McbA and homologues toward the industrial production of amide bonds these aims were set:

1. Determine the crystal structure of McbA.
2. Use the structure of McbA to design mutants to assay to aid in the elucidation of the enzymatic mechanism.
3. Express the homologues of McbA, *NfABS* and *ShABS*, and assay these for amide Bond synthetase activity.

If these aims are met it would give a greater understanding of the McbA family of adenylating enzymes. This would be of great value for potential engineering work to enable these enzymes to be used industrially.

2. Methods

2.1. Expression and Purification of McbA and Homologues

The genes for McbA, *AcABS* and *NsABS* had previously been cloned into the pET-YSBLIC3C(+) vector. The genes encoding for *NfABS* and *ShABS* were purchased already ligated into the pET28a(+) vector.

2.1.1. Transformation into BL21 (DE3) Cells

BL21 (DE3) cells were used for the production of all McbA constructs and homologues. The plasmids were used to transform BL21 (DE3) *E. coli* cells for gene expression. 1 µl of the plasmid was added to 50 µl of cells under aseptic conditions. The cells were then left to incubate on ice for 30 min before undergoing a heat shock at 42 °C for 10 s. This induces uptake of the DNA by the bacteria. After another 5 min incubation on ice, 150 µl of super optimal broth with catabolite repression (SOC) media were added to the cells. The culture was then incubated at 37 °C, with shaking at 220 rpm, for 1 h before being spread onto an agar plate containing 30 mg ml⁻¹ kanamycin. This allows for selection of colonies containing the plasmid.

2.1.2. Gene Expression

Pre-cultures were made by adding one colony of transformed BL21 (DE3) *E. coli* to 10 ml of LB media containing 30 mg ml⁻¹ kanamycin. This was then grown overnight (16-18 h) at 37 °C with shaking at 220 rpm. 5 ml of preculture was used to inoculate 1 l of LB media containing 30 mg ml⁻¹ kanamycin. LB media was chosen as it produced very pure protein despite TB media giving a higher yield. This was because protein quality was deemed to be more important than yield for crystallisation. The large-scale cultures were grown at 37 °C, with shaking at 180 rpm, until an OD₆₀₀ of 0.7 was reached. Expression was then induced by addition of 0.1 mM IPTG and incubation at 16 °C with shaking at 180 rpm overnight (16-18 h). The cells were

pelleted by centrifugation at 5422 g for 20 min. The supernatant was discarded and the pellet was recovered and either frozen in liquid nitrogen or used for purification without freezing.

2.1.3. Immobilised Metal Affinity Chromatography (IMAC) Purification of McbA and Homologues

The pellet was resuspended in resuspension buffer [50 mM HEPES, 300 mM NaCl, 10 mM MgCl₂, 10% glycerol (v/v), 20 mM imidazole]. Previously purifications had been carried out with either no imidazole in the resuspension buffer, or low imidazole running buffer. It was then found that adding 20 mM imidazole improved both the purity and yield of the protein after IMAC. In a further effort to further improve upon the purity of the enzyme, a protease inhibitor tablet was added to the resuspension. The cells were lysed by sonication for 10 min with 30 s on and 30 s off. The lysate was then pelleted by centrifugation at 41657 g for 20 min. The supernatant was collected and loaded at a rate of 5 ml min⁻¹ onto a GE HisTrap Ni NTA column pre-equilibrated with binding buffer (50 mM HEPES, 300 mM NaCl, 10 mM MgCl₂, 10% glycerol (v/v), 20 mM imidazole). Either an AKTA Pure or Start FPLC system was then used to elute the protein using an imidazole gradient of 20 mM – 300 mM over 30 column volumes.

2.1.4. Size Exclusion Chromatography (SEC) Purification of McbA and Homologues

The protein was further purified by size exclusion chromatography which uses porous beads to separate proteins based on size. The protein was first concentrated to below 2 ml using a 30 kDa cut off centricon. The protein was then manually loaded onto a HiLoad 16/600 Superdex 200 pg column pre-equilibrated with SEC buffer [50 mM HEPES, 300 mM NaCl, 10 mM MgCl₂, 10% glycerol (v/v), pH 8]. The protein was eluted over 1 column volume. using an AKTA Pure or Start FPLC. Fractions under the peak on the chromatogram were analysed by SDS-PAGE.

2.1.5. Cleavage of McbA His-tag

The plasmid containing the gene for wild-type McbA encoded for a human rhinovirus 3C protease (HRV-3C) cleavage site before the histidine tag, allowing for His-tag cleavage using HRV-3C protease. The reaction was carried out after IMAC purification and was dialysed overnight into low imidazole buffer. A 1:100 ratio of HRV-3C:McbA was used. The reaction was removed from the dialysis bag and was passed through a Ni-NTA column. The cleaved protein passed through the column and the un-cleaved protein, and the HRV-3C protease, bound to the column. The cleaved protein was then further purified by SEC.

2.2. Characterisation McbA and homologues

2.2.1. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

In order to analyse the purity of the protein a 12% SDS-PAGE was made up using the recipe in Table 1.

Resolving Gel	Stacking Gel
<ul style="list-style-type: none">• 3.2 ml deionised water• 2.5 ml resolving gel buffer (1.5 M Tris, 0.4% SDS, pH 8.8)• 4.2 ml acrylamide stock• 50 μl 10% ammonium persulfate (APS)• 8 μl Tetramethylethylenediamine (TEMED)	<ul style="list-style-type: none">• 3.2 ml deionised water• 1.3 ml stacking gel buffer (0.4 M Tris, 0.4% SDS, pH 6.8)• 0.5 ml acrylamide stock• 10 μl bromophenol blue solution 0.1%• 25 μl 10% ammonium persulfate (APS)• 8 μl Tetramethylethylenediamine (TEMED)

Table 1. Protocol for resolving and stacking layers of a 12% SDS-PAGE.

A 12% gel was chosen to give good separation for the size of the proteins. The SDS denatures proteins and provides a negative charge relative to their molecular weight allowing them to be separated based on their mass to charge ratio. This separation allows for the purity of the protein to be determined.

The resolving gel layer was first poured between two plates and levelled out with butan-1-ol. This was left to set for 20 min. The butan-1-ol was then poured off and the stacking layer was added. A comb inserted to introduce wells for the samples. The stacking layer was then left to set for 20 min.

Samples after IMAC and SEC purification were made up by diluting the sample 3 μ l in 10 μ l and then adding 10 μ l of 2x SDS sample buffer. Sample buffer was made up with 4.8 ml of deionised water, 1.2 ml of 0.5 M Tris pH 6.8, 1 ml of glycerol, 2 ml of 10 % (w/v) SDS-solution, 0.5 ml of 1 % (w/v) bromophenol blue, 0.5 ml of β -mercaptoethanol. To run the gel the samples were further denatured by heated at 95 °C for 5 min. The samples were then mixed and loaded onto the gel. The gel was run at 200 mV for 45 min.

The gel was stained by microwaving with Coomassie stain and left to stain for at least 2 h. The stain was then changed for water in order to remove excess staining solution from the gel.

2.2.2. Analytical Size Exclusion Chromatography Analysis of McbA

Analytical size exclusion was performed on an AKTA Pure FPLC system using a Superdex™ 200 Increase 10/300 GL column. 200 μ l of sample at a concentration of 1 mg ml⁻¹ was injected and the column was run with 0.001 M phosphate buffer, 0.14 M NaCl, pH 7.4 at 0.75 ml min⁻¹. A standard curve was produced by running a premix of proteins with known molecular weights.

2.2.3. Polyacrylamide Gel Electrophoresis (PAGE) Analysis of McbA with Different Ligands

A native PAGE was used to analyse the different oligomeric states and conformations of the protein. This technique is similar to SDS-PAGE but the gel and buffers do not contain SDS meaning the protein is not denatured. The gel was made up using the recipe in table 2.

Resolving Gel	Stacking Gel
<ul style="list-style-type: none">• 4.4 ml deionised water• 2.5 ml resolving gel buffer (1.5 M Tris, pH 8.8)• 2.9 ml acrylamide stock• 50 μl 10% ammonium persulfate (APS)• 8 μl Tetramethylethylenediamine (TEMED)	<ul style="list-style-type: none">• 3.2 ml deionised water• 1.3 ml stacking gel buffer (0.4 M Tris, pH 6.8)• 0.5 ml acrylamide stock• 25 μl 10% ammonium persulfate (APS)• 8 μl Tetramethylethylenediamine (TEMED)

Table 2. Recipe for resolving and stacking layers of an 8.75% Native-PAGE

The resolving layer was poured and levelled out using butan-1-ol. This was left to set for 20 min before the butan-1-ol was poured off and the stacking layer was added. This was left to set for 20 min. Samples were made up by adding 10 μ l sample buffer to 10 μ l of sample. The sample buffer was made up with 4.8 ml of deionised water, 1.2 ml of 0.5 M Tris pH 6.8, 1 ml of glycerol, 0.5 ml of 1 % (w/v) bromophenol blue.

Reducing agents used were at a 5 mM concentration and ligands were at a 10 mM concentration. The gel was run at 200 mV for 40 min with running buffer not containing any SDS. The gel was stained using the same procedure as with SDS-PAGE.

2.3 Crystallisation of McbA

Crystallisation conditions for all proteins were initially screened in 96 well 2-drop plates using commercially available screens. Hits were then optimised using either 96-well 2-drop plates or 48-well MRC MAXI plates. The concentration of each protein was varied, between 2 – 40 mg/ml, until a good amount of precipitation was seen after 24 hours (about 50% of drops containing precipitate).

2.3.1. 96-well 2-Drop MRC Crystallisation Plates

Initial screens were set up in 96-well -2 drop crystallisation plates using a Mosquito robot. 56 µl of the screen condition was added to the reservoir using either a multi-channel pipette or a hydra robot. Wells were set up by dispensing 150 nl or 100 nl of protein and adding 150 nl or 200 nl of the reservoir solution to produce a 1:1 or 1:2 screen: protein ratio. Commercial screens used included JCSG+, PACT, Index, Clear strategy screen I+II, PDB minimal, Crystal screen HT, MPD Screen, Salt RX, PGA and Morpheus.

2.3.2. Crystal optimisation

Optimisation of crystal hits was carried out using 48-well MRC MAXI plates. These allow for a larger drop volume than the 96-well plates which can often lead to higher quality and larger crystals. Optimisations involved varying precipitant, salt and protein concentrations as well as the pH. Various additives were also added to the reservoir, such as poly-alcohols and potential ligands (e.g. AMP and **17**), in an effort to produce higher quality crystals by slowing down crystal growth. To set up the plates 100 µl of buffer solution, containing any additives, was added to the wells. A Mosquito robot was then used to dispense 250 nl of protein and 250 nl of buffer into the wells for a 1:1 protein to buffer ratio. Alternatively, different concentrations were sometimes screened by using 0.8:1 and 1.2:1 protein to buffer ratio in alternating wells.

For ligand co-crystallisation, the protein was incubated with 4 mM of ligand (10 mM for ATP, AMP or ADP) for 30 min on ice before crystal screens were set up. All β -carbolines were in 10 mM DMSO stock solutions. All other ligand stocks were made up with ultra-pure water. For co-crystallisation with the ADP-aluminium fluoride complex, the complex was produced by incubating 0.3 M ADP with 0.2 M AlCl_3 , 0.4 M NaF and 0.3 M MgCl_2 .^[54] This was then left to incubate for 1 h before being used for co-crystallisation with McbA.

2.3.3. Seeding

Seeding was attempted in order to improve crystal quality. Seeding involves adding some protein crystal to a crystallisation drop in order to provide a site of nucleation. Skipping the nucleation step allows for screening of conditions that favour crystal growth rather than having to find conditions that favour both nucleation and growth. Seed stock was produced from poor quality crystals by adding both reservoir buffer and seeds to a Hampton Research seed bead and vortexing for 30 s. An Oryx 8 liquid handling robot was used to set up seeding experiments. These were set up in 96-well plates with a two drop format. The top drop contained 0.3 μl of protein solution, 0.05 μl of seed stock and 0.25 μl of the screen buffer. The bottom drops contained 0.3 μl of protein solution and 0.3 μl of screen buffer. Because the Oryx uses only one needle to dispense the screen, seed stock left in the needle after dispensing the top drop is dispensed in the bottom drop.

2.3.4. Crystal Testing, Data Collection and Structure Building

Crystals were washed in 15% ethylene glycol cryoprotectant before being flash frozen in liquid nitrogen. Initial data were collected in house using a Rigaku MicroMax-007 HF X-ray Generator. Images were taken at 0° & 90° with a 0.5° oscillation, 5 min exposure at 2.5\AA resolution. For the K483A crystals a longer exposure time (10 min) was used.

Data for the K483A crystal were collected on beamline i03 at the Diamond Light Source and data for wt-McbA were collected on beamline i04. The data for both structures were processed and integrated using XDS and scaled using SCALA within the Xia2 processing system.^[55-57] Molecular replacement using MOLREP with RpPAT from *Rhodopseudomonas palustris* (28% sequence identity, PDB code 4GXQ) as the molecular replacement model was successful in solving the structure of K483A McbA.^[58, 59] The K483A structures were refined by using repeated cycles of the COOT and REFMAC5 programmes.^[60, 61] Once the protein backbone, side chains and water molecules were built, residual density was present in the omit map in all five active sites.

In Chains A-D found in the adenylation conformation this could be modelled as AMP and **3**. In the E subunit found in the amidation conformation **3** could be modelled but the density for the AMP phosphate was missing. ACEDRG was used to prepare coordinate and library files for **3**.^[62] The K483A structure was then used to solve the wt-McbA structure. The K483A structure was refined against the density for wt-McbA by using repeated cycles of COOT and REFMAC5.^[60, 61] The same ligands were found in both structure and so were modelled into the wt-MCBA structure in the same way as for the K483A mutant structure.

2.4. Producing a D201A McbA mutant

The aspartate in position 201 in McbA was seen to be conserved across all ABS enzymes. This aspartate corresponds to a conserved Histidine in ANL enzymes (H207 in 4CBL).^[37] The structure of McbA revealed that this residue sits at the end of the amine binding tunnel, leading us to consider the possibility that D201 is reasonable for deprotonating the amine substrate. In order to investigate the role of D201 a D201A mutant was produced. The McbA D201A mutant was generated using QuickChange site directed mutagenesis. Primers were designed to change the aspartate in position 201 to an alanine (Figure 16).

Fwd: 5'- ATTCCGCTGAGCGCACTGGGTGGTGAA-3'

Rev: 3'- AAGGCGACTCGCGTGACCCACCACTTG-5'

Figure 16. Primers used for QuickChange mutagenesis to produce the D201A McbA mutant.

A 50 μ l reaction was set up with these primers containing the components in Table 3.

Component	Volume	Final concentration
10x Buffer	5 μ l	1x
25 mM MgSO ₄	3 μ l	1.5 mM
dNTPs 2 mM each	5 μ l	0.2 mM each
Forward primer 10 mM	0.75 μ l	0.3 μ M
Reverse primer 10 mM	0.75 μ l	0.3 μ M
Template DNA 100 ng/ μ l	1 μ l	2 ng/ μ l
KOD polymerase	1 μ l	0.02 μ l
H ₂ O	33.5 μ l	N/A

Table 3. Contents of PCR mix used for QuickChange mutagenesis to generate D201A McbA mutant

Touchdown PCR was then used in order to improve the chance of success. For this the PCR reaction was set up with an annealing step that was 10 °C above the calculated T_m which decreased by 1 °C every cycle until the T_m was reached (Table 4). Amplification was then continued for 25 more cycles using the T_m as the annealing temperature (Table 5).

Phase 1

Phase	Temp	Time
Denature	95 °C	3 min
Denature	95 °C	30 s
Annealing	$T_m + 10$ °C	45 s
Extension	72 °C	6 min

Table 4. PCR conditions for the first phase of QuickChange PCR for generation of the D201A McbA mutant. Annealing temperature starts at 10 °C above T_m and is decreased by 1 °C each cycle for 10 cycles.

This was repeated for 10 cycles decreasing by 1 °C annealing temperature each cycle.

Phase 2

Phase	Temp	Time
Denature	95 °C	30 s
Annealing	T_m	45 s
Extension	72 °C	6 min

Table 5. PCR conditions for the second phase of QuickChange PCR for generation of the D201A McbA mutant. The annealing temperature is equal to the calculated T_m for the primers.

This was then repeated for 25 cycles before the reaction was terminated by incubation at 72 °C for 5 min.

The reaction was analysed by running 10 µl on a 1% agarose gel. If amplification had been successful, DpnI digest was then performed to remove the template DNA. 1 µl of DpnI was added to the PCR product. This was centrifuged for 1 min before incubation at 37 °C for 1 h. The resulting product was then transformed into stellar competent *E. coli* cells which repaired nicks in the plasmid. A 10 ml pre-culture was then grown of these cells at 37 °C with shaking at 220 rpm overnight. A monarch plasmid miniprep kit was then used to extract the mutant

plasmid. The plasmid was then sent for sequencing by GATC to confirm the mutation was successful.

2.5 Screening *ShABS* Substrates

2.5.1. Small scale screens

Small scale (100 μ l) reactions were carried out for substrate screens in 1.5 ml Eppendorf tubes. The reactions were carried out in potassium phosphate buffer (50 mM, pH 7.5), with 0.4 mM acid, 0.6 mM amine (1.2 equivalents), 2 mM ATP (5 equivalents) and 0.02 mM of the enzyme. The acid was added from a 10 mM DMSO stock leading to a final DMSO concentration of 4% (v/v). The reactions were carried out at 37 °C with shaking at 180 rpm for 24 h before being stopped by extraction with ethyl acetate (4 x 250 μ l). The ethyl acetate was left to evaporate and the extracted compounds were dissolved in 100 μ l of methanol for HPLC.

The HPLC was run using a CNW Athena C18-WP, 5 μ l, 100Å, column with eluents; A) H₂O, 0.1% formic acid, B) MeCN 0.1% formic acid. The method used was 40% of B to 100% B in 5 min followed by 10 min at 100% B with a flowrate of 1 ml min⁻¹. The column was kept at a constant 35 °C.

Peaks were assigned based on acid and amide standards and conversions were calculated based on the difference between acid and amide peak areas.

2.5.2. 10 mg scale up reaction

In a 250 ml conical flask 10 mg of **3** were dissolved in 400 μ l of DMSO. To this was added 0.06 mmol of **d** and 0.2 mmol ATP. This was made up to 9 ml with 50 mM KP_i pH 7.5 before 1 ml of 0.2 mM *ShABS* was added. This reaction was left at 37 °C with shaking at 180 rpm for 24 h. The reaction was stopped by extraction with 100 ml ethyl acetate. This was washed with HCl 0.1 N (50 ml), NaOH 0.1 N (50 ml), and brine (50 ml), before being dried over Na₂SO₄,

filtered and the solvent evaporated under reduced pressure. This resulted in the production of the product as a yellow powder. The product was then characterised by NMR and mass spectrometry.

^1H NMR (400 MHz, Chloroform- d) δ 10.41 (s, 1H), 9.13 (s, 1H), 8.23 (d, J = 7.9 Hz, 1H), 7.87 (d, J = 8.9 Hz, 1H), 7.69 – 7.56 (m, 2H), 7.42 – 7.34 (m, 1H), 7.30 – 7.24 (m, 3H), 7.19 – 7.04 (m, 1H), 4.35 (m, 1H), 2.92 (s, 3H), 2.78 (t, J = 8.4 Hz, 2H), 1.99 (m, 2H), 1.39 (d, J = 6.5 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform- D) δ 202.4, 164.2, 142.0, 141.7, 139.6, 136.5, 133.6, 132.8, 129.9, 128.6, 128.5, 126.0, 122.58, 121.7, 121.2, 118.6, 112.3, 45.4, 39.2, 32.8, 25.8, 21.4. HRMS (ESI+, m/z): calculated for $(\text{C}_{24}\text{H}_{24}\text{N}_3\text{O}_2)^+$ [($M+H$) $^+$] 386.1863; found: 386.1851.

The chiral HPLC was run using a Chiralpak ID4 HPLC column with 20% isopropanol: 80% hexane as eluent with a flow rate of 1 ml min^{-1} .

3. Results and Discussion

3.1. McbA Structure Determination

Previous attempts had been made by our group to crystallise McbA, but no crystals had been obtained. Prior to this project a number of constructs of McbA were produced. These included both His-tagged and non-His-tagged variants, a surface entropy reduction mutant (McbA-SER) and a K483A mutant. In order to improve the chances of crystallising McbA attempts were made to improve the purity of the enzymes. This purer enzyme was then used to set up crystallisation screens for each construct and homologue, with a range of different ligand complexes. Crystallisation was eventually achieved with K483A McbA in complex with **3** and AMP allowing for a structure to be produced. The same conditions that produced crystals of K483A McbA were then used to produce a crystal of the wild-type McbA. The structure of the wild-type McbA was successfully solved using the K483A McbA structure.

3.1.1. Expression and Purification of McbA and Homologues

Before attempting crystallisation of McbA the purity of then enzyme first had to be improved in order to improve the chances of successful crystallisation. In an effort to improve the quality of the protein a protease inhibitor was introduced to the re-suspended bacteria pellet before sonication. This should prevent the production of fragments of the protein that can be difficult to remove. Based on previous expression tests it was also noted that using LB media for bacterial growth and gene expression produced purer protein than TB media which had been used previously.

The plasmid for N-terminally His-tagged McbA was first used to transform BL21 (DE3) chemical competent *E. coli* cells. A pre-culture was set up using a colony of these cells and was used to inoculate 2 l of LB media. The cultures were grown using previously generated protocols. After the cultures were grown the protein was purified by two step purification using

IMAC and SEC. The protein was analysed by SDS-PAGE and it was seen that the protein was significantly purer when using the modified protocol (Figure 17).

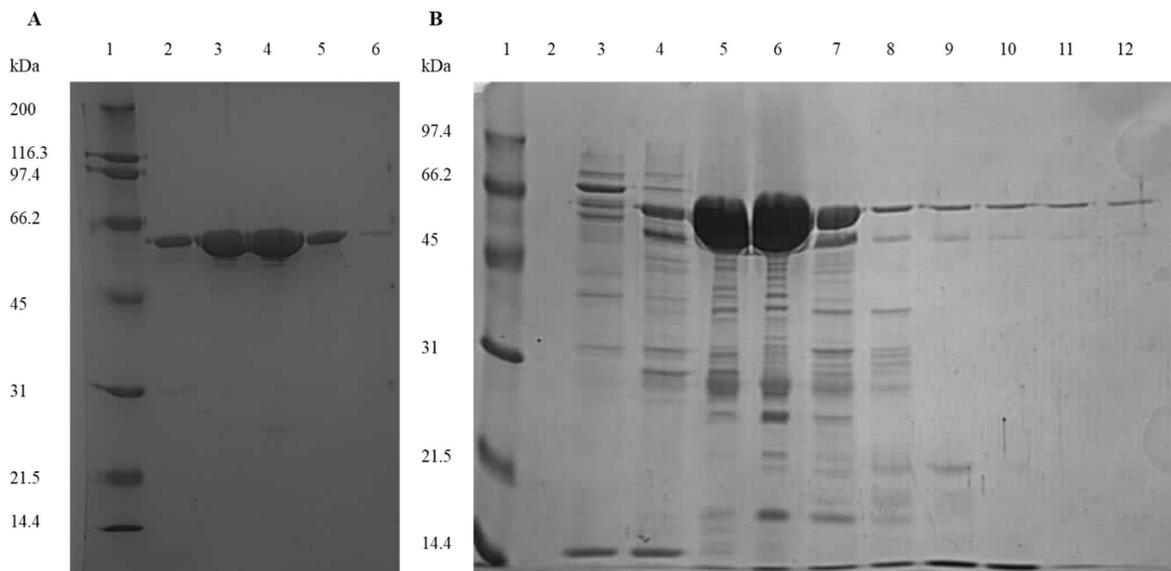


Figure 17. A comparison of McbA purity using different expression and purification methods analysed by SDS-PAGE. A. SDS-PAGE analysis of McbA after SEC expressed in LB media and with a protease inhibitor used for lysis. 1: molecular weight marker. 2-6: fractions from SEC purification. B. SDS-PAGE analysis of McbA expressed in TB media. 1: molecular weight marker. 2-12: fractions from SEC purification. Calculated MW for McbA is 53 kDa

After size exclusion two peaks were seen on the chromatogram (Figure 18). This was true for all of the constructs and homologues of McbA.

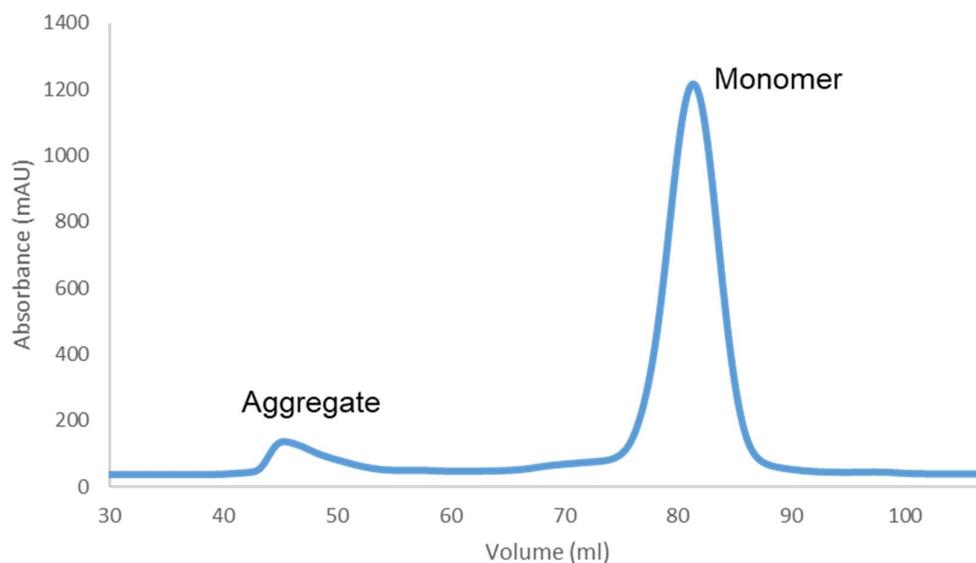


Figure 18. Chromatogram from SEC McbA purification with aggregate and monomer peaks labelled as determined by analytical SEC

Previous work had shown the second peak to elute, and so the smaller protein form, was active and the first peak to elute to be inactive. When analysed by SDS-PAGE these peaks both have a band at about 53 kDa leading to the belief that the larger molecule must be an alternative oligomeric state of McbA. To investigate this further analytical size exclusion chromatography was used (Figure 19A).

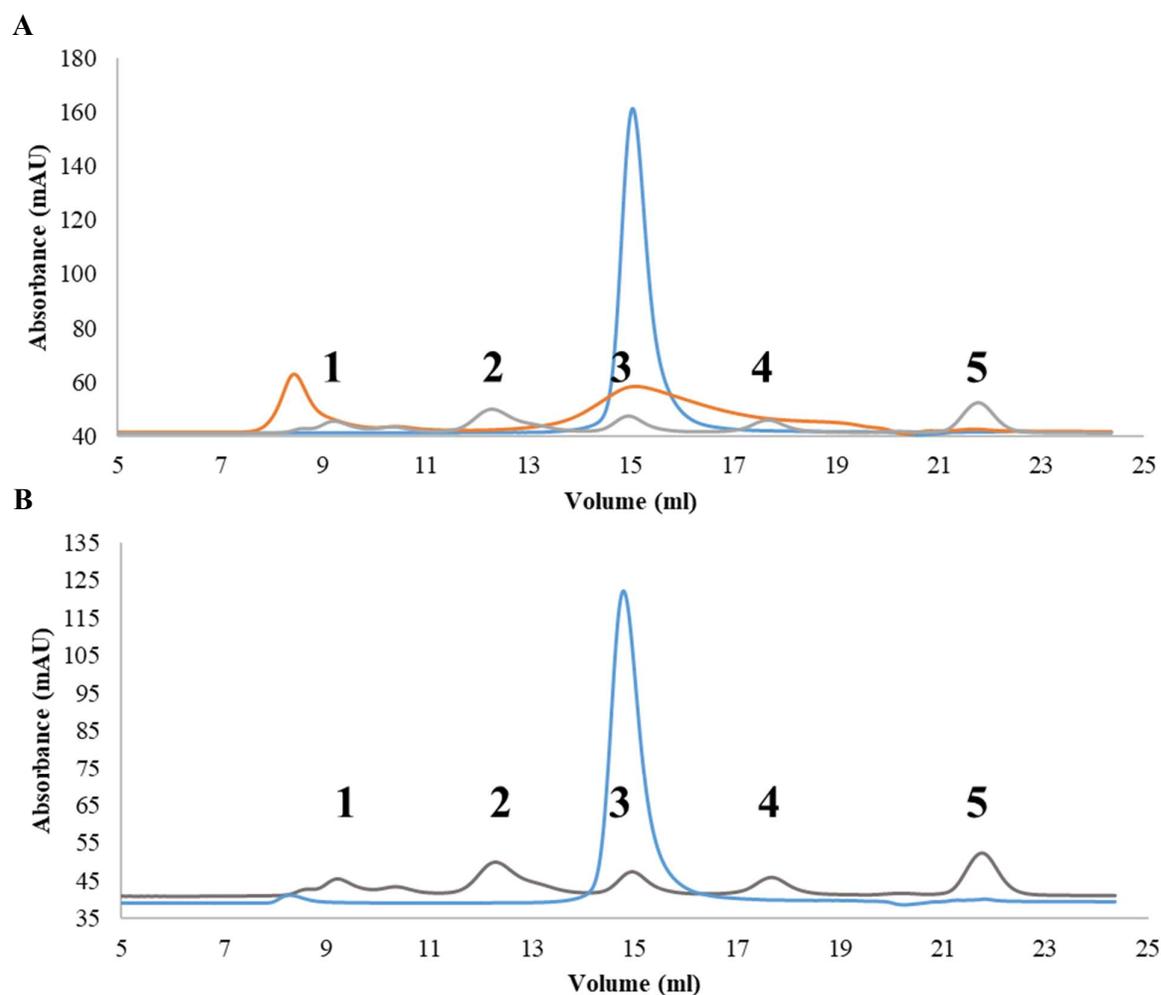


Figure 19. Analytical size exclusion chromatogram of McbA **A**: Directly after SEC **B**: 3 d after SEC. The first peak to elute from SEC was run and is shown in orange and the second peak is shown in blue. Standards are shown in grey. Calculated MW for McbA is 53 kDa. The second peak (blue) corresponds to the mass for the monomeric form of McbA. The first peak (orange) has a peak for the monomeric form of McbA as well as a peak at >700 kDa. **B** shows very little aggregation occurs at 4°C. Standards: **1**. Thyroglobulin bovine MW: 670 kDa **2**. γ -globulins from bovine blood MW: 150 kDa **3**. Albumin chicken egg grade VI MW: 44.3 kDa **4**. Ribonuclease A type I-A from bovine pancreas MW: 13.7 kDa **5**. P-aminobenzoic acid

The smaller molecule eluted at a molecular weight that corresponded to the monomeric form of McbA and the larger species eluted at a very high molecular weight (>700 kDa). Due to the very high mass of the larger species it was believed that this was a product of aggregation. It was thought that aggregation of the monomer may be the reason for the difficulty in crystallisation of McbA. In order to investigate the propensity of McbA to aggregate another analytical size exclusion was run with the monomeric protein sample after 3 d at 4°C (Figure

19B). After this time a very small amount of aggregation was seen but this was not deemed to be large enough to be of concern. The aggregation previously seen may be due to denaturation of the protein during expression or lysis.

The quality of the protein obtained was sufficient to use for screening of crystallisation conditions. The same purification protocol was able to produce similar quality protein for all constructs of McbA and *AcABS*.

In order to expand the number of proteins that could be screened His tag cleavage was performed. The pETYSBLIC3C(+) vector allows for highly specific his tag cleavage using HRV-3C protease.

3.1.2. Crystallisation of Mutant and Wild-Type McbA

In order to produce crystals of McbA a systematic approach was used in which it was attempted to co-crystallise all available constructs of McbA with all available ligands. These attempts are summed up in Figure 20.

Enzyme (construct)	Ligand						
	No ligand	Substrate Carboxylic acid (CA)	CA plus ATP	CA plus AMP	Product	ATP carba- analogue	AMP plus metal fluoride
McbA (LIC- 3C – N-term his-tag)	✓	✓	✓	✓	✓	✓	✓
McbA (LIC- 3C – Cleaved his- tag)	✓	✓	✓	✓	✗	✓	✗
McbA K438A	✓	✓	✓	✓	✗	✓	✗
McbA-SER	✓	✗	✗	✗	✗	✗	✗
<i>AcABS</i>	✓	✓	✓		✗	✓	✗
<i>ShABS</i>	✓	✓	✓	✓	✗	✗	✗

Figure 20. Summary of co-crystallisation attempts with each construct and each ligand used. Attempted conditions are shown with a green tick and conditions not attempted are shown with a red cross. The Carboxylic acid substrate used was **3**. For the product both **3a** and **9a** were attempted. The ATP carba-analogue was α,β -Methyleneadenosine 5'-triphosphate **17**.

The constructs used were His-tagged wt-McbA, His-tag cleaved wt-McbA, K483A McbA, SER-McbA and later D201A McbA. In addition to this, crystallisation of the homologues *AcABS* and *ShABS* was also attempted. The ligands used for co-crystallisation were **3**, **3a**, **10a**, an uncleavable ATP analogue **17** (Figure 21) and an ADP-aluminium fluoride complex.

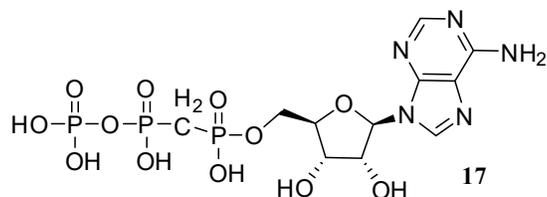


Figure 21. Structure of α,β -Methyleneadenosine 5'-triphosphate used as an ATP mimic.

The effect of ligand binding and reducing agents on conformation and oligomeric state was investigated by Native-PAGE (Figure 22). It was hoped that one of these ligands might stabilise the conformation and thus aid crystallisation, however no effect was seen on the native-PAGE.

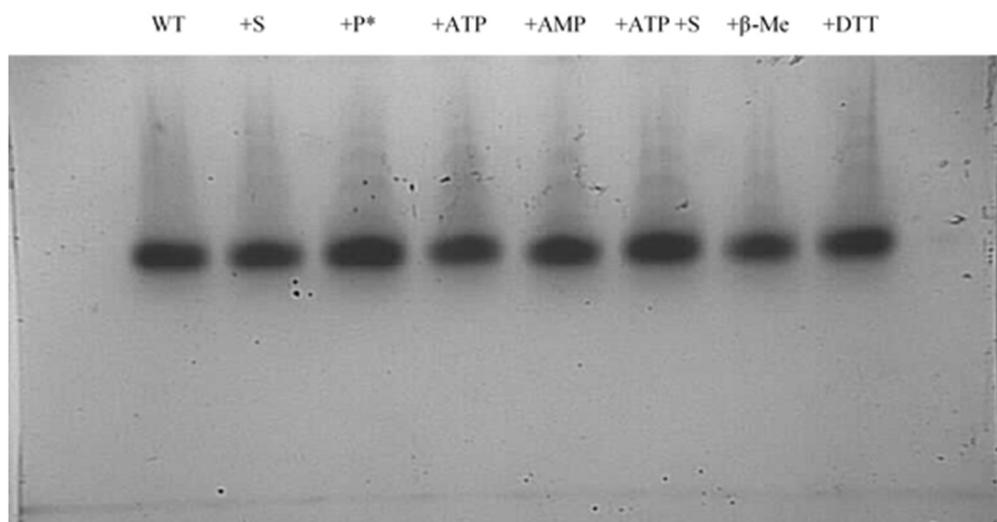


Figure 22. Native-PAGE analysis of McbA in complex with the carboxylic acid substrate **3**, the naphthoic acid product **10a**, ATP, AMP, ATP with the carboxylic acid substrate **3**, β -mercaptoethanol (β -Me) and dithiothreitol (DTT).

The ADP-aluminium fluoride complex had been shown to mimic transition states in F1-ATPases and so it was hoped this would act as a transition state mimic for McbA.^[63] However crystallisation screens set up with the complex produced solid non-protein crystal material in the drops (Figure 23C). McbA catalyses the cleavage of ATP to AMP and PPi, meaning the complex may not have been mimicking the correct transition state and may still have been cleavable by McbA. Because of this crystallisation attempts with the ADP-aluminium fluoride

complex were abandoned. Crystallisation with the native product **3a** was attempted but the compound was found to be insoluble at any reasonable DMSO concentration. Crystallisation was also attempted with **10a** however this did not yield any crystals.

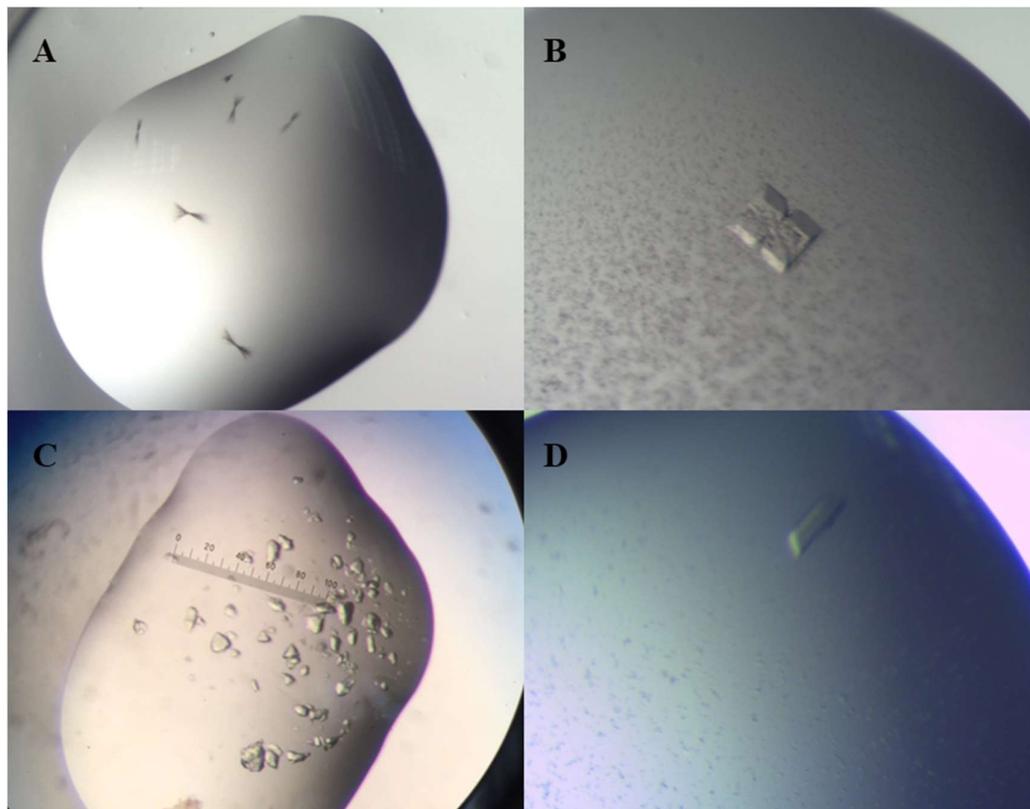


Figure 23. Examples of salt crystals obtained in initial screens. A: CSS F8 condition (1.8 M Li_2SO_4 , MES pH 5.5) with wt-McbA. B: Morpheus C4 condition with K483A McbA. C: AMP aluminium fluoride contamination. D: Index D9 condition with his cleaved McbA, AMP and **3**.

Initial crystallisation screening was carried out using sitting drop vapour diffusion in 96-well plates. At least three commercial screens were used for each of the protein-ligand combinations including Index CSS I+II and one other one. Most of the protein-ligand combinations were screened using more than three screens. During the initial screening a number of non-protein crystals were found (Figure 23). The absence of protein was confirmed by X-ray diffraction patterns. Salt crystals gave a distinctive diffraction pattern with intense reflections at high resolution (Figure 24).

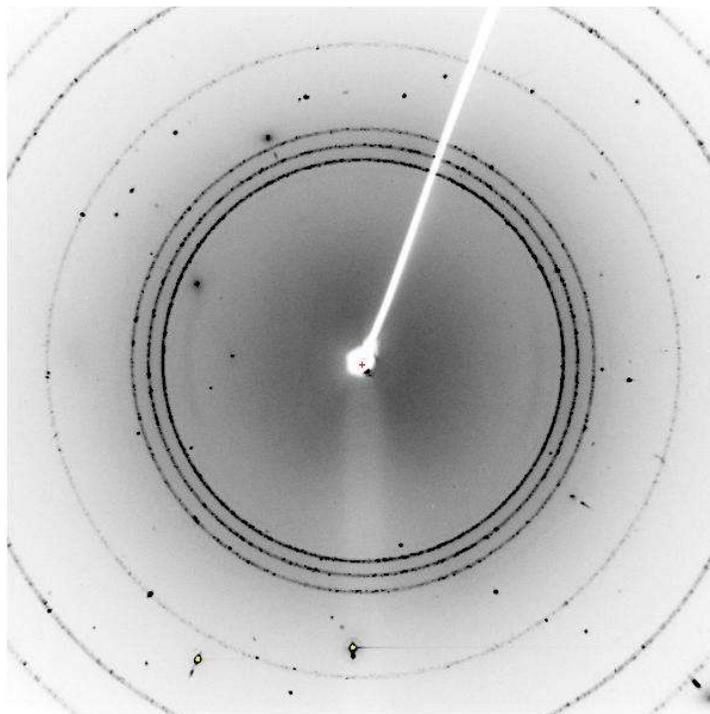


Figure 24. Diffraction pattern from a salt crystal grown in Salt RX screen

Eventually crystals were obtained in the PDB minimal screen for the K483A mutant with the carboxylic acid substrate **3** and AMP (Figure 25).

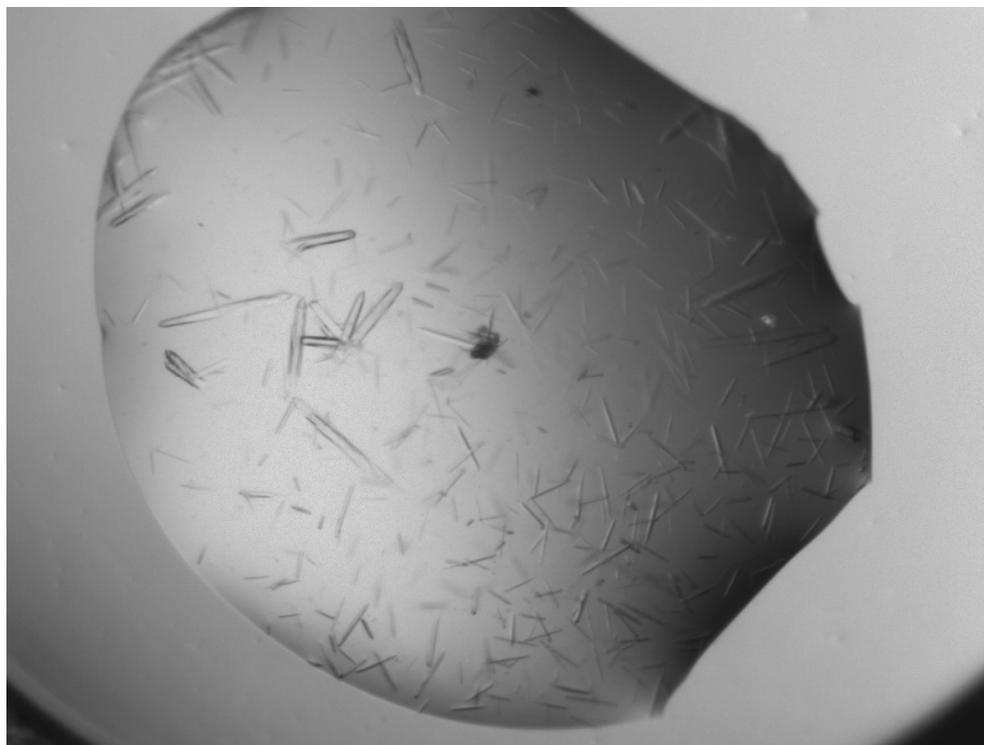


Figure 25. K483A McbA crystals in complex with the carboxylic acid substrate **3** and AMP. Grown in PDB minimal H12 condition (1.6 M Na citrate pH 7.5)

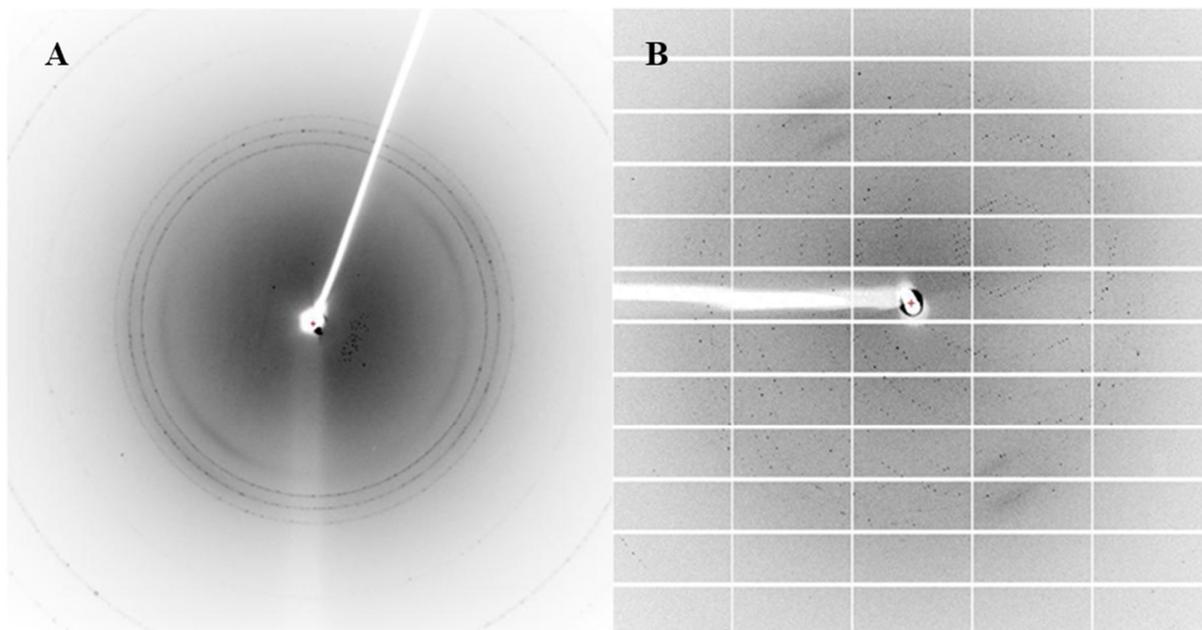


Figure 26. Diffraction patterns from K483A McbA crystals in complex with the carboxylic acid substrate **3** and AMP. **A**: In house diffraction showing only a few peaks at low resolution indication the presence of protein. **B**: Diffraction pattern from The Diamond Light Source synchrotron showing diffraction to $\sim 2.8 \text{ \AA}$.

This crystal was fished and tested in house and showed some diffraction at low resolution (~ 10 Å) that indicated the presence of protein (Figure 26A). In order to confirm the presence of protein the crystal was sent to the Diamond Light Source synchrotron. It was also hoped that the stronger and more focused beam would give a higher resolution than the in house diffraction, especially considering the small size of the crystal. The results from Diamond showed diffraction to 2.8 Å (Figure 26B) which confirmed the presence of protein. The K483A mutant has reduced activity compared to wt-McbA. This may have an effect on substrate binding affinity and thus may have allowed stabilisation of the structure which improved the mutant's ability to crystallise.

After the successful crystallisation of the K483A McbA mutant, crystallisation of wt-McbA was attempted using the same conditions. Using these conditions resulted in showers of very small needle crystals to form (Figure 27A). These were then optimised in 48-well MAXI plates. Different pH, protein concentration and Na citrate concentrations were all tested as well as poly-alcohol additives which may slow down the crystal growth. This resulted in the production of larger wt-McbA crystals in 6% ethylene glycol (v/v), 1.6 M Na citrate pH 7.5 (Figure 27B)

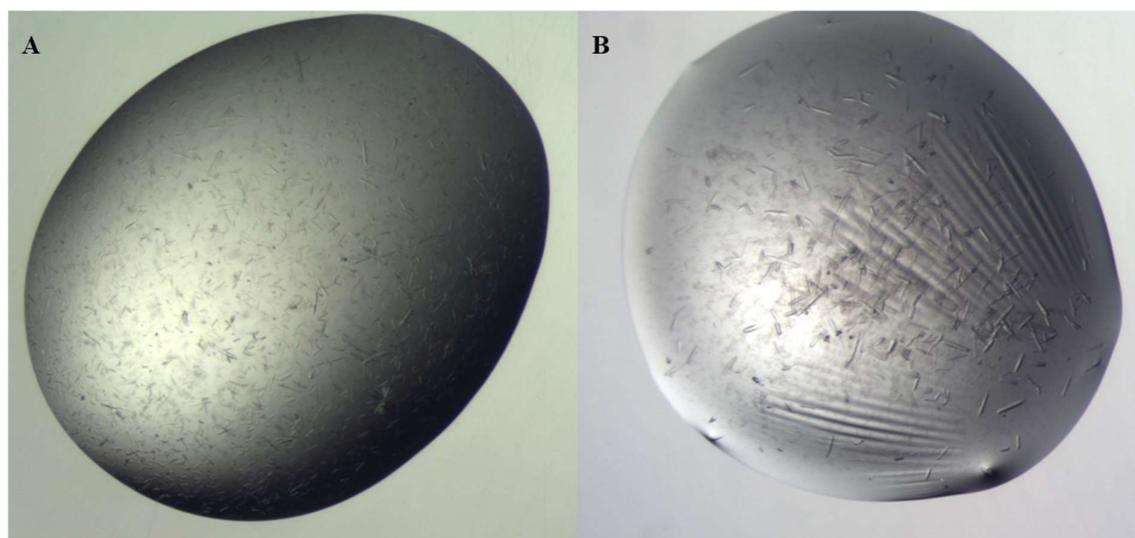


Figure 27. **A.** small needle crystals of wt-McbA with **3** and AMP in 1.6 M Na citrate pH 7.5 **B.** wt-McbA crystals with **3** and AMP in 6 % Ethylene Glycol v/v, 1.6 M Na citrate pH 7.5

These crystals were fished and tested in house (Figure 28A) this confirmed the presence of protein in the crystals and indicated a potentially higher resolution than was achieved for K483A McbA.

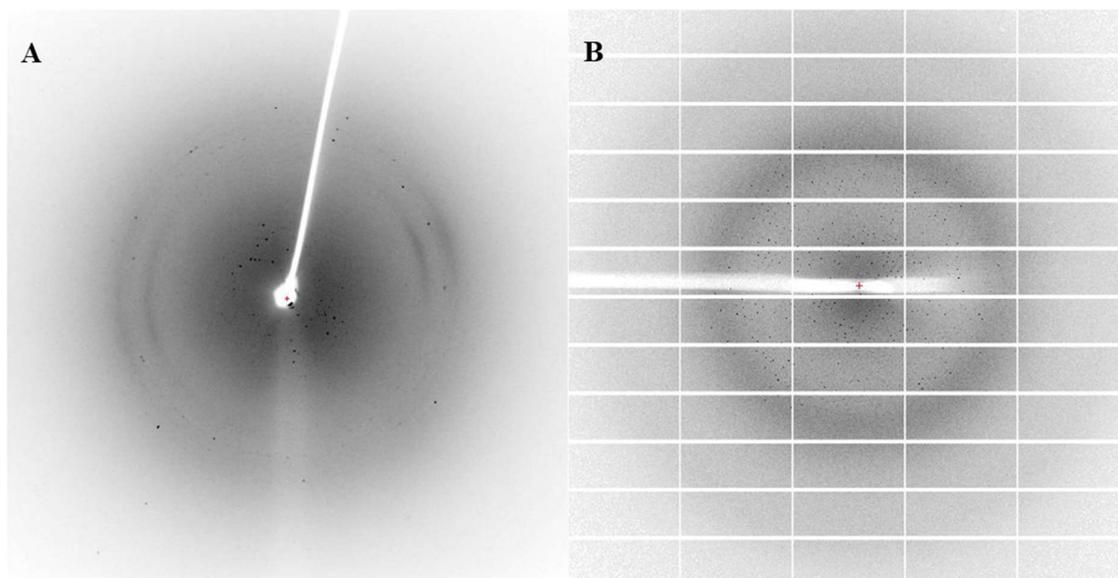


Figure 28. Diffraction patterns of wt-McbA crystals **A**:in house diffraction **B**: Diffraction pattern from The Diamond Light Source synchrotron showing diffraction to $> 3 \text{ \AA}$

The crystals were then sent to The Diamond Light Source synchrotron for testing which produced diffraction data of $\sim 2.7 \text{ \AA}$ (Figure 28B).

Attempts to improve upon these crystals to produce a higher resolution structure were made. Optimisation of these crystals with various additives only succeeded in producing similar quality crystals. These crystals were then used as seed stock for later experiments.

3.1.3. Structure Building and Refinement

The K483A McbA structure was solved using molecular replacement with RpPAT from *Rhodopseudomonas palustris* (28% sequence identity, PDB code 4GXQ) as the molecular replacement model.^[59] RpPAT is an acetyl transferase that is able to bind ATP, explaining its similarity to McbA. The structure was built and refined using repeated cycles of COOT and REFMAC5.^[60, 61] Five protein chains were found in the asymmetric unit with one of these in a different conformation to the other four. AMP and **3** were successfully modelled into the active sites of all five chains. Data collection and refinement statistics for the K483A structure can be seen in Table 6.

Table 6. Refinement statistics for the K483A McbA structure in complex with AMP and **3**. Values in brackets refer to data for the highest resolution shells.

K483A McbA in complex with 3 and AMP	
Beamline	Diamond I03
Wavelength (Å)	0.97624
Resolution (Å)	108.83-2.80 (2.86-2.80))
Space Group	$P2_12_12_1$
Unit cell (Å)	a = 118.23; b = 130.74; c = 196.38 $\alpha = \beta = \gamma = 90^\circ$
No. of molecules in the asymmetric unit	5
Unique reflections	75674 (4421)
Completeness (%)	100.0 (100.0)
R_{merge} (%)	0.26 (1.13)
$R_{\text{p.i.m.}}$	0.11 (0.46)
Multiplicity	13.2 (13.4)
$\langle I/\sigma(I) \rangle$	8.3 (2.6)
Overall B factor from Wilson plot (Å ²)	41
$CC_{1/2}$	0.99 (0.67)
$R_{\text{cryst}}/R_{\text{free}}$ (%)	20.5/24.5
r.m.s.d 1-2 bonds (Å)	0.010
r.m.s.d 1-3 angles (°)	1.41
Avge main chain B (Å ²)	52
Avge side chain B (Å ²)	54
Avge water B (Å ²)	29
Avge ligand B for AMP (Å ²)	54
Avge ligand B for 3 (Å ²)	45

The wt-McbA crystal was found to have the same the same space group as for the K483A structure. This allowed the structure to be solved by refining the K483A structure against the wt-McbA data. AMP and **3** were also successfully modelled into this structure. A higher resolution was achieved for the wt-McbA structure than the K483A structure. Data collection and refinement statistics can be found in Table 7.

Table 7. Refinement statistics for the wt-McbA structure in complex with **3** and AMP. Values in brackets refer to data for the highest resolution shells.

Wt-McbA in complex with 3 and AMP	
Beamline	Diamond I04
Wavelength (Å)	0.97950
Resolution (Å)	65.54-2.59 (2.66-2.59)
Space Group	$P2_12_12_1$
Unit cell (Å)	a = 119.05; b = 131.07; c = 196.04 $\alpha = \beta = \gamma = 90^\circ$
No. of molecules in the asymmetric unit	5
Unique reflections	95977 (6982)
Completeness (%)	100.0 (99.9)
R_{merge} (%)	0.26 (0.86)
$R_{\text{p.i.m.}}$	0.13 (0.47)
Multiplicity	8.1 (8.2)
$\langle I/\sigma(I) \rangle$	6.2 (2.2)
Overall B factor from Wilson plot (Å ²)	33
$CC_{1/2}$	0.99 (0.49)
$R_{\text{cryst}}/R_{\text{free}}$ (%)	21.3/24.9
r.m.s.d 1-2 bonds (Å)	0.010
r.m.s.d 1-3 angles (°)	1.33
Average main chain B (Å ²)	40
Average side chain B (Å ²)	42
Average water B (Å ²)	28
Average ligand B for AMP (Å ²)	75
Average ligand B for 3 (Å ²)	42

3.1.4. Investigation of the structure of McbA

The structure discussion will focus on the wt-McbA structure due to its improved resolution and the presence of the K483 residue. Five McbA molecules were found in the asymmetric unit (Figure 29), one of these was in a different conformation to the other four. Comparison of these structures to that of the ANL enzymes allowed identification of the two conformations as matching the adenylation and thiolation conformations of NRPSs and acyl-CoA synthetases. The thiolation conformation has been termed the amidation conformation for McbA.

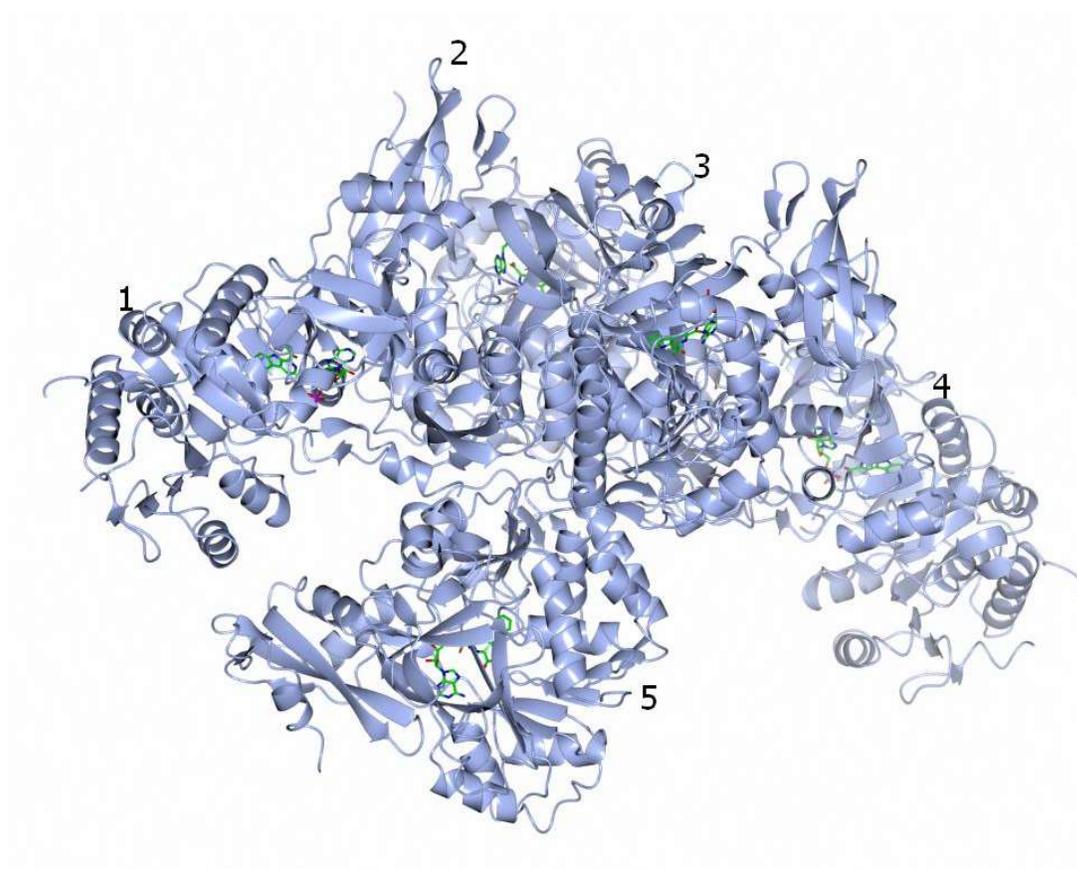


Figure 29. The asymmetric unit from the wt-McbA crystal containing five chains with a $P2_12_12_1$ space group. Chains 1-4 are in the adenylation conformation and chain 5 is in the amidation conformation. Each chain has both AMP and **3** bound in the active site. The structure was refined to a resolution of 2.7 Å.

The structure of McbA is consistent with the two-domain structure of the ANL class of adenylating enzymes. It shows a large N-terminal domain (1-394) and a smaller C-terminal domain (395-494) with the hinge residue at Q394. Comparison of the two conformations shows a 142° rotation of the C-domain in relation to the N-domain between the two conformations (Figure 30). This flexibility was likely the cause in the difficulty in crystallising McbA.

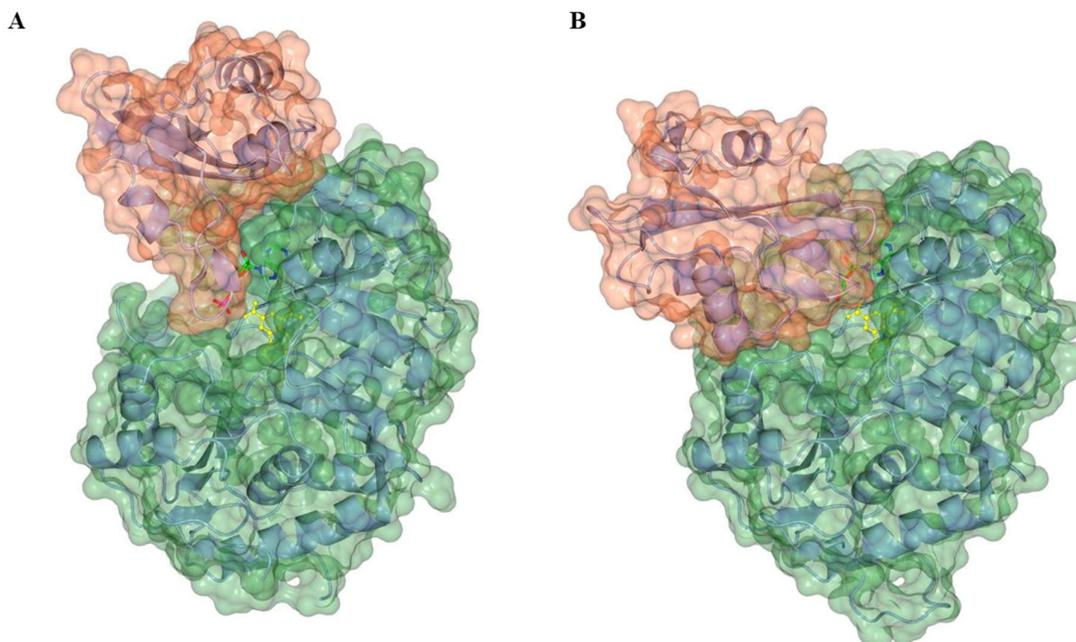


Figure 30. Crystal structure of wt-McbA at 2.7 Å in A: The adenylation conformation, and B: The amidation conformation. The C-terminal domain is shown in coral and rotates 142° with respect to the N-terminal domain, shown in green, between the adenylation and amidation conformations.

Comparison of the different conformations with other ANL enzymes allowed the different conformations to be assigned as the adenylation conformation (Figure 30A) and the amidation conformation (Figure 30B). The structures superimpose (using secondary structure match) very well with both ANL and CAR adenylating domains (Figure 31). The adenylation conformation can be seen to superimpose well with other ANL adenylation conformations such as the NRPS DhbE standalone adenylation domain (2.28 Å over 450 C α atoms. PDB code: 1MDF) (Figure 31A).^[36] The amidation conformation can be seen superimposed well with the thiolation conformations of 4-chlorobenzoyl-CoA ligase (4CBL) (2.03 Å over 415 C α atoms PDB code:

3CW9) (Figure 31B) and a CAR adenylation domain (2.55 Å over 411 C α atoms. PDB code: 5MSS) (Figure 31C).^[37, 41]

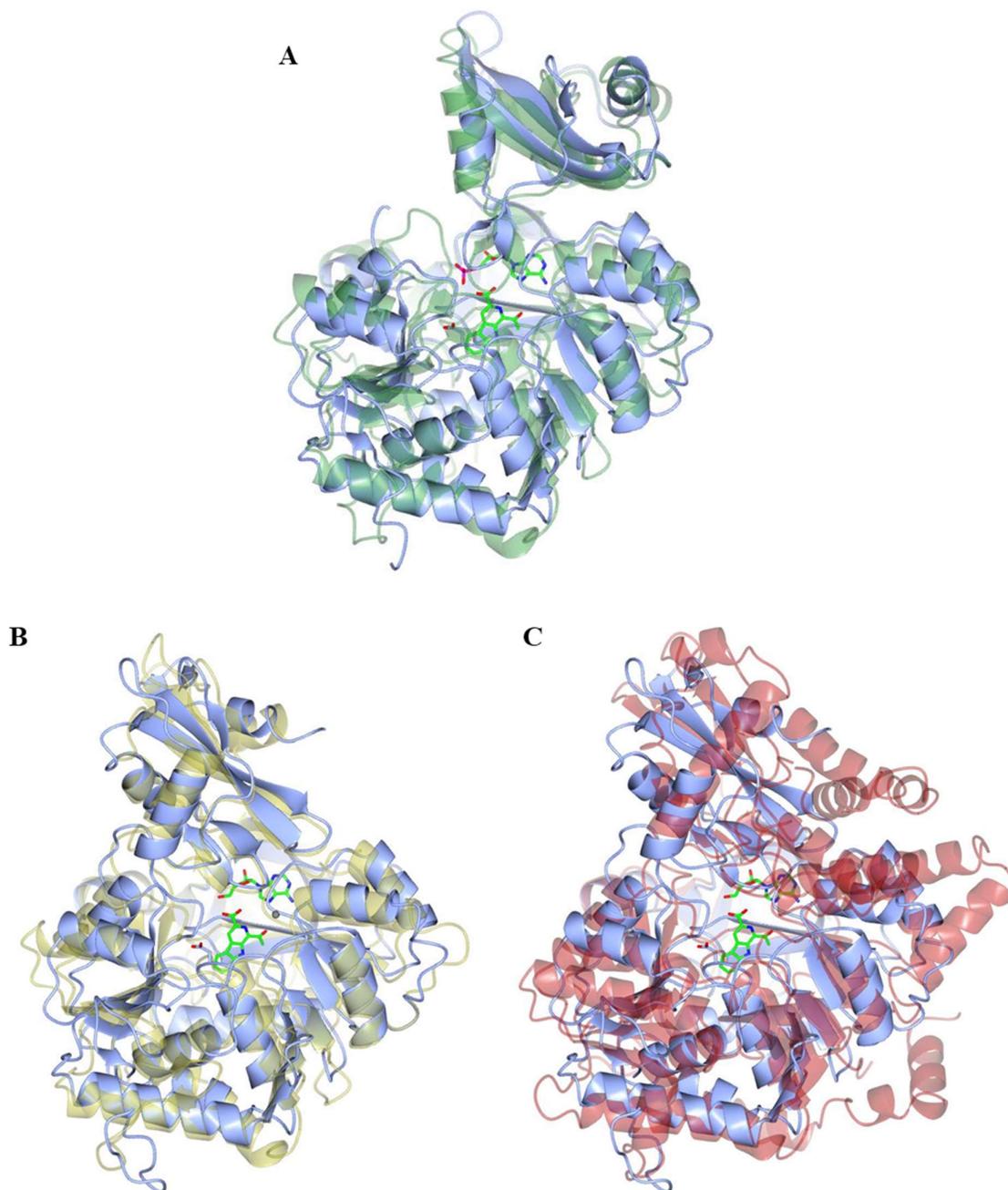


Figure 31. **A:** The structure of wt-McbA in complex with AMP and **3** in the adenylation conformation shown in blue superimposed (using secondary structure match) with the structure of DhbE shown in green (2.28 Å over 450 C α atoms. PDB code: 1MDF).^[36] **B:** The structure of the amidation conformation of McbA show in blue with **3** and AMP bound superimposed with the thiolation conformation of 4CBL (2.03 Å over 415 C α atoms PDB code: 3CW9).^[37] **C:** The structure of the amidation conformation of McbA show in blue with **3** and AMP bound superimposed with a CAR adenylation domain shown in red (2.55 Å over 411 C α atoms. PDB code: 5MSS).^[41]

The similarity of the two conformations to those seen in ANL and CAR enzymes as well as a relatively high sequence similarity between McbA and some ANL enzymes (>20%) indicate a shared evolutionary path. This is probably the result of a divergent evolution that led to the development of the CARs, NRPSs, Acyl-CoA synthetases and ABSs. The comparison with ANL enzymes also allowed identification of a potential amine tunnel. This can be seen in the comparison with 4CBL which had 4-chlorophenacyl-CoA bound (Figure 32).^[37]

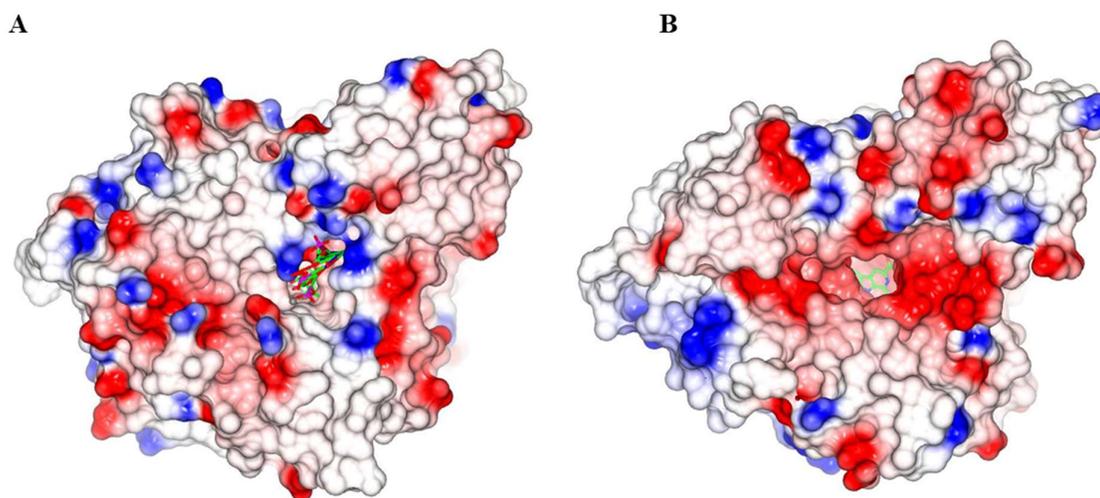


Figure 32. **A:** The structure of 4CBL (4CW9) in the thiolation conformation with 4-chlorophenacyl-CoA bound (shown in green).^[37] **B:** The structure of McbA in the amidation conformation. **3** shown in green can be seen through the putative amine tunnel. Comparing the structures shows a similar tunnel in McbA to the 4-chlorophenacyl-CoA binding site in 4CBL but 4CBL contains positive residues (shown in blue) around this site whereas McbA contains negative residues (shown in red).

The 4CBL structure shows a number of positive residues around the binding tunnel which are absent in McbA such as R87, R475 and K477. McbA even contains some negative residues around the putative amine binding tunnel such as D201, D463, E221 and E400 which may aid amine binding. One of these residues D201 is conserved across all of the McbA homologues and corresponds to a conserved histidine in ANL enzymes (H207 in 4CBL).^[37] This residue

sits at the interface between the putative amine tunnel and the active site (Figure 33). It is believed that D201 may be responsible for the deprotonation of the amine substrate for attack on the adenylate. To test this hypothesis a D201A mutant was generated (see section 6.1.5).

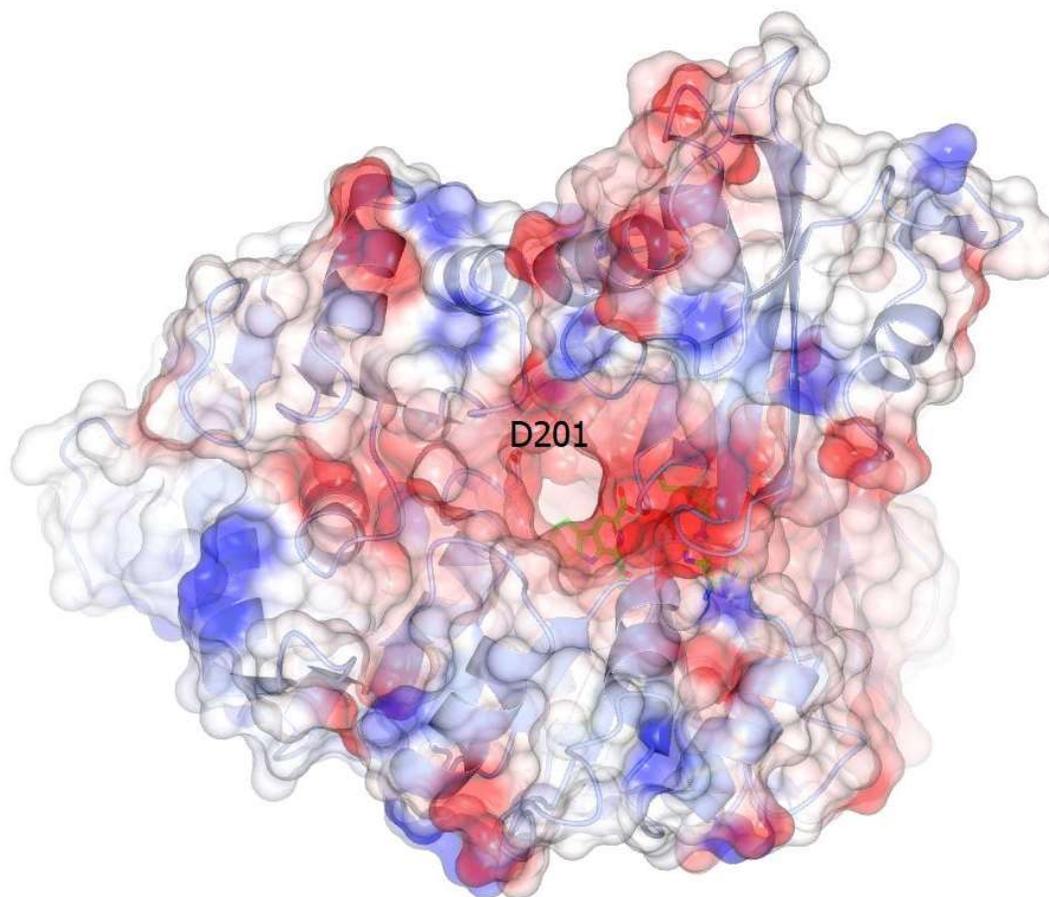


Figure 33. The structure of wt-McbA with AMP and **3** shown in green and D201 shown at the end of the putative amine tunnel.

The wt-McbA structure contained density for K483 that was missing in the K483A structure. This lysine is conserved among the ANL enzymes and is believed to be involved in the adenylation step.^[28] It can be seen that K483 is able to interact with both AMP and **3**, and may

act to stabilise the adenylate intermediate (Figure 34A). Comparing the location of K483 in both conformations (Figure 34B) shows that in the adenylation conformation K483 is in the active site and it is able to interact with both AMP and **3** but in the amidation conformation K483 is moved away from the active site which may promote attack of the adenylate and release of the amide product. This may explain the reduced activity of the K483A mutant which may not bind both substrates as effectively and they therefore may not be positioned optimally in the active site.

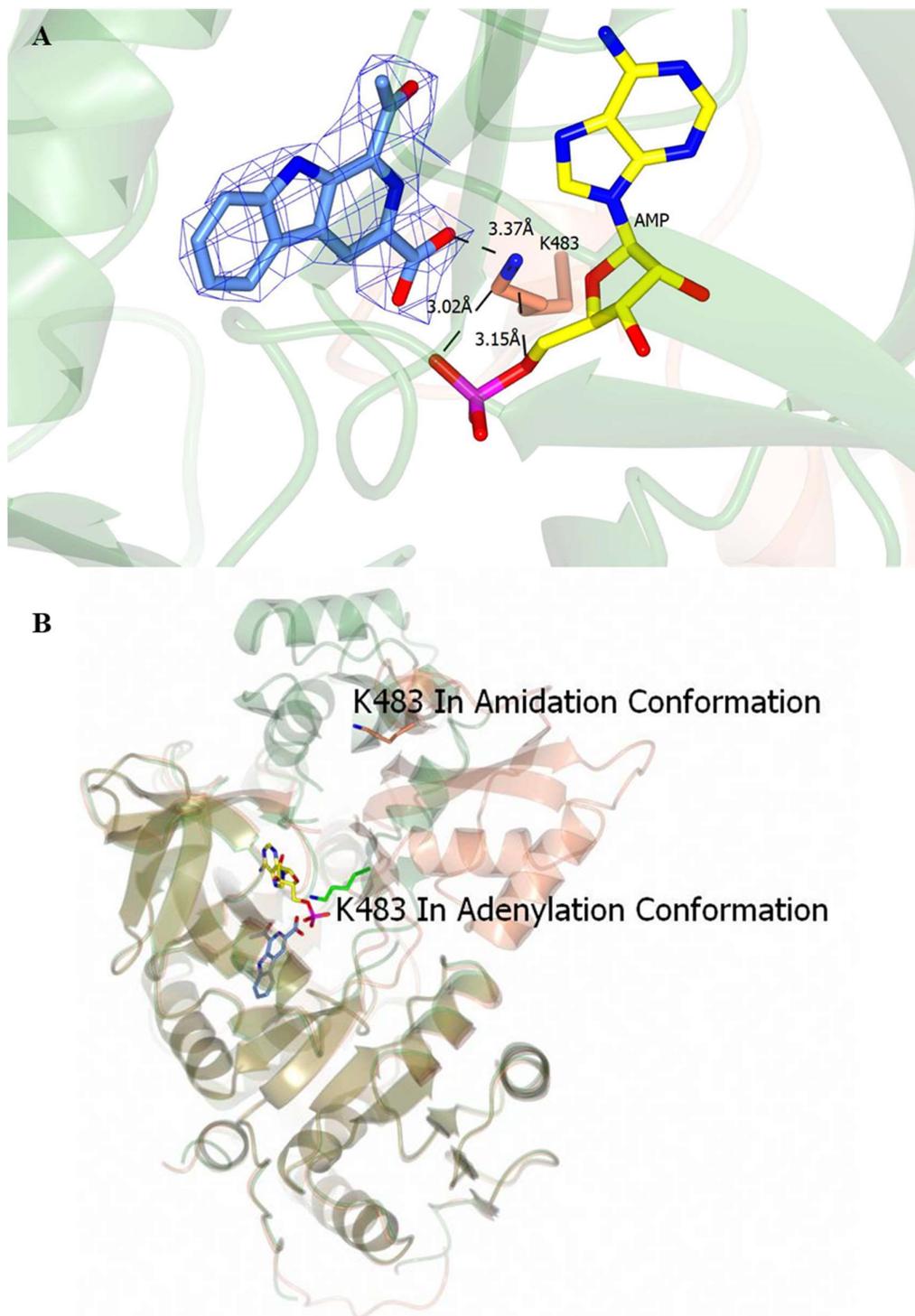


Figure 34. **A:** The interactions between K483 and the carboxylic acid substrate and AMP in the adenylation conformation. **3** is shown in blue and AMP is shown in yellow. The electron density for the $2mF_o-DF_c$ map of **3** is shown in blue at a level of 1σ . The N-terminal domain is shown in green and the C-terminal domain in coral. **B:** The structure of wt-McbA showing the movement of K483 with the conformational change. The two conformations are overlaid

with the adenylation conformation shown in green and the amidation conformation shown in coral.

The structure of the active site reveals some reasoning for the broad substrate specificity of McbA (Figure 35). It can be seen that only a few residues are able to interact with the carboxylic acid substrate. The backbone of G295 is able to interact with the ketone oxygen and the pyrrole nitrogen at 3.9 Å and 3.5 Å respectively. There is also potential for π -H interaction with F301 which is 3.81 Å away. The hydrophobic pocket produced by F301 and L202 also contributes to substrate binding. K483 is only able to interact with the carboxylic acid of **3** and thus would not discriminate between different carboxylic acids. The relatively few interactions between McbA and **3** is likely the cause of the broad substrate specificity.

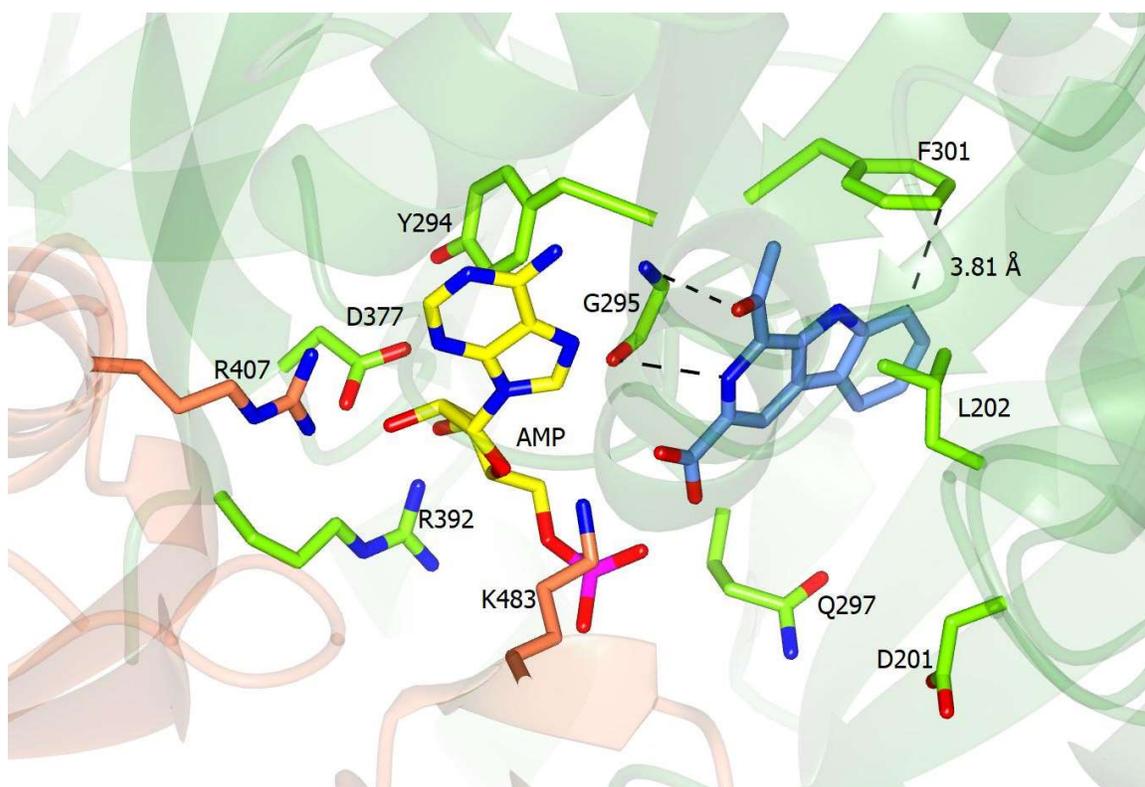


Figure 35. Active site of wt-McbA in the adenylation conformation with N-terminal domain residues shown in green and C-terminal domain residues shown in coral. The carboxylic acid substrate **3** is shown in blue and AMP is shown in yellow. The interactions between the carboxylic acid substrate **3** with G295 and F301 are shown.

The AMP can be seen to have potential interactions with K483, R392, R407 and D377. The domain movement from adenylation to amidation conformation moves both K483 and R407 out of the active site and so allows for dissociation of the AMP. D201A can also be seen pointed towards the putative amine tunnel.

3.1.5. Producing a D201A McbA Mutant

In order to investigate the mechanism of McbA further a D201A point mutation was generated. D201 sits at the interface of the putative amine tunnel and the active site (Figure 33) and it was thought that D201 may be responsible for deprotonation of the amine substrate. D201 corresponds to a conserved histidine in ANL enzymes and CARs. It is important for gating the thiolation tunnel and for stabilising the thiolation conformation. The D201A mutant was generated using QuickChange mutagenesis. The mutagenesis PCR product was analysed by agarose gel (Figure 36) which showed a band at the expected weight for the intact plasmid with the gene insert. This demonstrated that the primers were successful in amplifying the plasmid. The mutation was then confirmed by DNA sequencing.

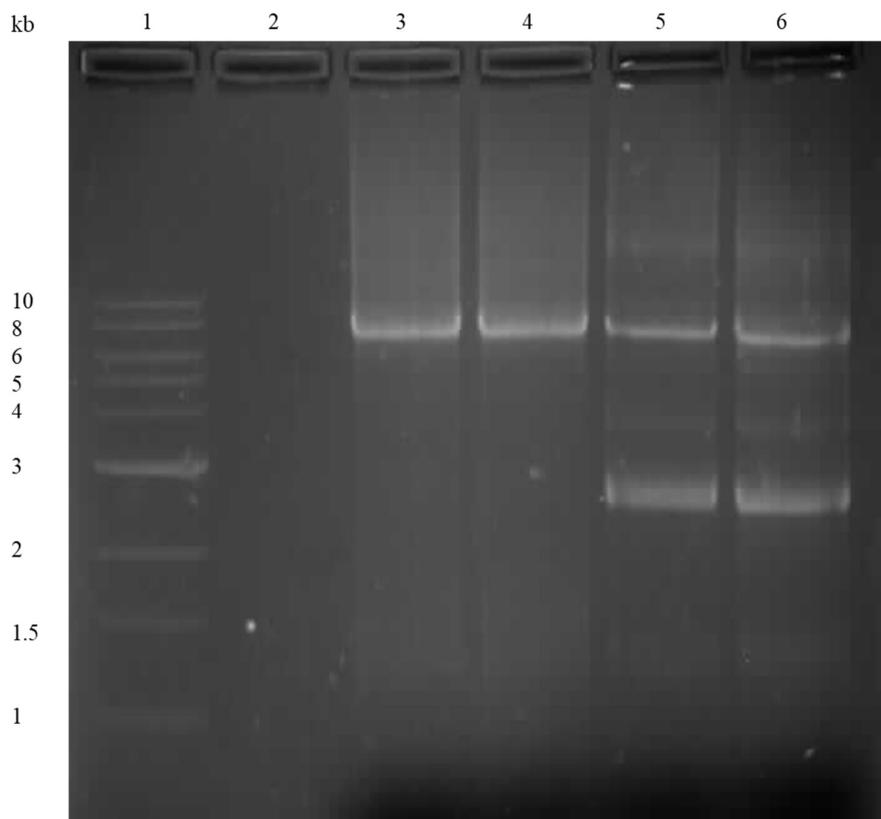


Figure 36. Agarose gel analysis of QuickChange mutagenesis showing the plasmid between 6 and 8 kb, plasmid with insert is ~ 7 kb. 1: Quick-Load purple 1 kb DNA Ladder 2: empty lane 3-6: Quick change PCR products.

The plasmid was transformed into BL21 (DE3) *E. coli* and the gene was expressed using the same protocol as for the other McbA constructs. The protein was then purified by both IMAC and SEC using the standard procedure which resulted in highly pure protein (Figure 37).

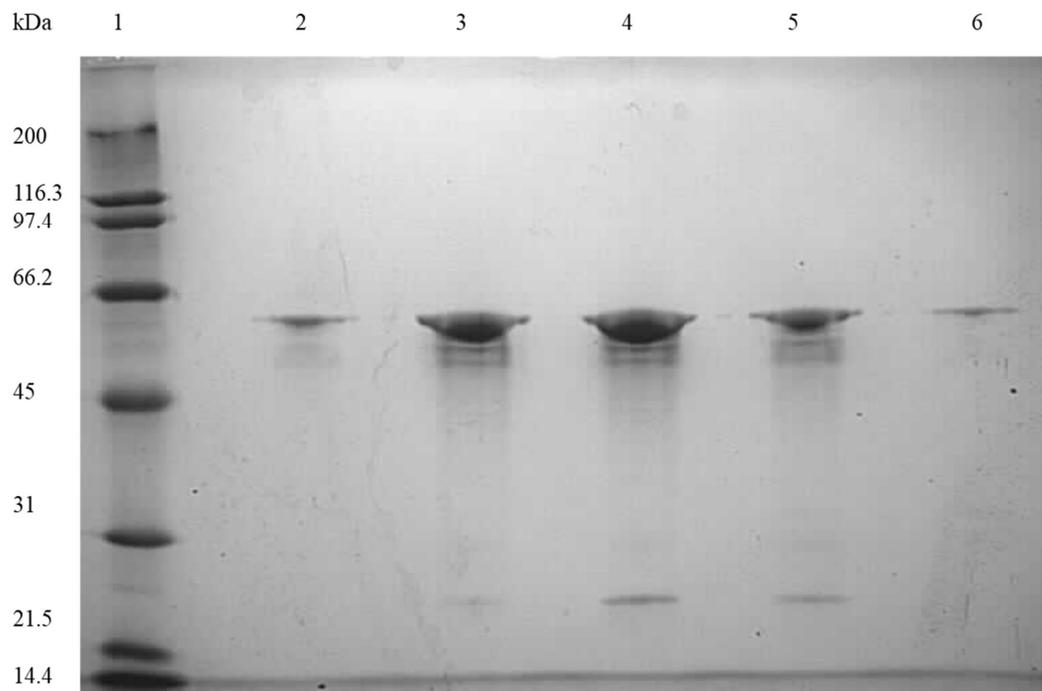


Figure 37. SDS-PAGE analysis of the purification of D201A McbA after SEC. 1: Broad range molecular weight marker 2-6: Fractions from SEC containing D201A McbA. Expected mass = 53 kDa

3.2. Characterisation of a New Amide Bond Synthetase, *ShABS*

Identification of new amide bond synthetases allows for a greater coverage of substrates for enzymatic amide couplings. This is important in order to allow for production of the highly diverse range of amides used in the pharmaceutical industry. In order to identify additional ABSs, homologues of McbA were investigated. It is hoped that comparison of these homologues may also provide additional mechanistic information. So far of the four homologues of McbA identified two have been successfully expressed and purified by our group. These homologues are *AcABS* and *ShABS*. *AcABS* has been previously characterised by our group as an amide bond synthetase with similar activity to McbA. *ShABS* was identified by BLAST search as a 51% homologue of McbA from *Streptoalloteichus hindustanus* (Figure 38).

McbA	MGYARRVMDGI----GEVAVTGAGGSVTGARLRHQVRLLAHALTEAGIPPGRGVACLHAN	56
	MGY RV++G+ G A+ A +TGA ++ A AL G+ PG GV LH N	
ShABS	MGYLHRVVEGLKANAGGEALVSADRRLTGAETLEEIHRTARALAAQGLRPGDGVVTLHGN	60
McbA	TWRAIALRLAVQAIGCHYVGLRPTAAVTEQARAIAAADSAALVFEPSVEARAADLLERVS	116
	A+ LR+AVQ +GC Y GLRP A E+A +A A++AA V++P + AA+LL V	
ShABS	GVEAVVLRIVQQLGCRYAGLRPVFATREKANFLAEAEAAAFVYQPDMADEAAELLREVP	120
McbA	VPVVLSLGPTSRGRDILA-ASVPEGTPLRYREHPEGIADVAVFTSGTTGTPKGVVHSSSTAM	175
	P VLSLGP G D++A A P+ + V FT GTTG KGV + +	
ShABS	TPRVLSLGPAPLGEDLVALAGAQSAEPVEFTADERAATAVGFTGTTGTRAKGVCRAPFDL	180
McbA	SACVDAAVSMYGRGPWRFLIPIPLSDLGGELAQCTLATGGTVVLLLEEFQPDVLEAIERE	235
	AC+DA++++G GPWRFL+ IP++DLGGE+A+ TLA GGTVV L E+F+P +L I E	
ShABS	EACLDASLTIFGEGPWRFLVCIPIADLGGEMAEWTLAAGGTVVLRDFEPADILATIGAE	240
McbA	RATHVFLAPNWLYQLAEHPALPRSDLSSLRRVVYGGAPAVPSRVAAARERMG-AVLMQNY	294
	R THVF AP W+YQLAEHPAL +DLSSL ++ YGGAP+ P+R+A A E++G +L+ Y	
ShABS	RTHVFCAPGWVYQLAEHPALADADLSSLTQIPYGGAPSTPARIADALEKLRPLLHVCY	300
McbA	GTQEAAAFIAALTPDDHAR--RELLTAVGRPLPHVEVEIRDDSGGTLPRGAVGEVWVRSPM	352
	G+QE ++ L+ +DH R R LL +VG+ LP E+ IRD G LP G VGEV VRS M	
ShABS	GSQEGGWMTWLSAEDHVRADRYLLNSVGKALPGTEIAIRDQDGADLPVGTVGEVCVRSTM	360
McbA	TMSGYWRDPERTAQVLSGGWLRGTGDVGTDFEDGHLHLTDRLQDIIIVEAYNVYSRRVEHV	412
	M GYWR PE TA+ + GWL TGD+G D +G+L+L DR +D+IIVEAYNVYS+ VEHV	
ShABS	LMRGYWRLPELTAKTVRDGWLHTGDLGRLDTEGYLYLVDRAKDVIIVEAYNVYSQEVEHV	420
McbA	LTEHPDVRAAAVVGVPDPDSGEAVCAAVVVADG-ADPDPEHLRALVRDHLGDLHVPRRVE	471
	LT HPDVR AAVVGVPD D+ EAV AAVV A+G + D + LRALVR LG +H P+ ++	
ShABS	LTGHPDVRYAAVVGVPDHDTEAVYAAVVAEGVGEIDVDELRALVVRTTLGPVHEPKHLD	480
McbA	FVRSIPVTPAGKPKVKVRTWFTD	495
	V +IP TP GKPKD +RT	
ShABS	VVDTIPTTPRGKPKDSALRTRWRAAQRA	508

Figure 38. Sequence alignment of *ShABS* and *McbA*. Sequence identity 51%. Matched residues are shown between the sequences with similar residues shown as a +.

3.2.1. Expression and Purification of *ShABS*

The gene for *ShABS* was purchased already cloned into the pET-28a vector. This construct was used to transform BL21 (DE3) *E. coli*. *ShABS* was then expressed and purified by IMAC and SEC using the standard methods. This resulted in very pure protein (Figure 39).

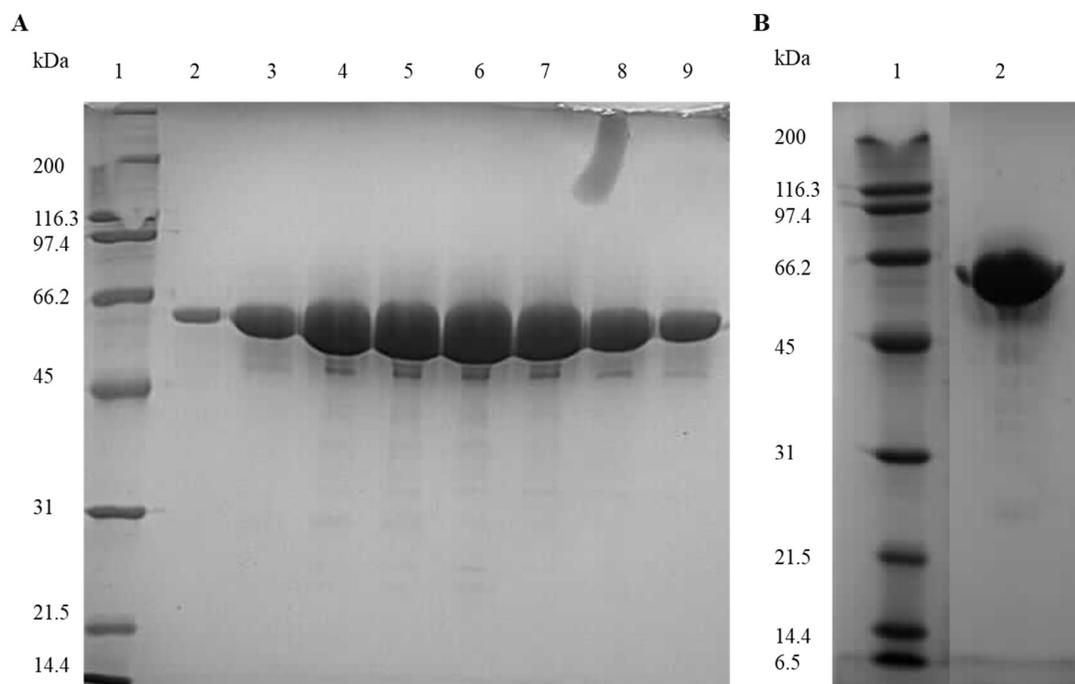


Figure 39. SDS-PAGE analysis of *ShABS* purifications. **A:** *ShABS* purification by IMAC 1: broad range molecular weight ladder. 2-9: Fractions from IMAC containing protein **B:** *ShABS* purification by SEC. 1: broad range molecular weight ladder 2: Pure *ShABS* protein from SEC. Calculated MW for *ShABS*= 54 kDa

As with *McbA* and *AcABS*, two peaks were seen on the chromatogram after SEC (Figure 40). The smaller mass peak was assumed to be the monomer and so was used for assays.

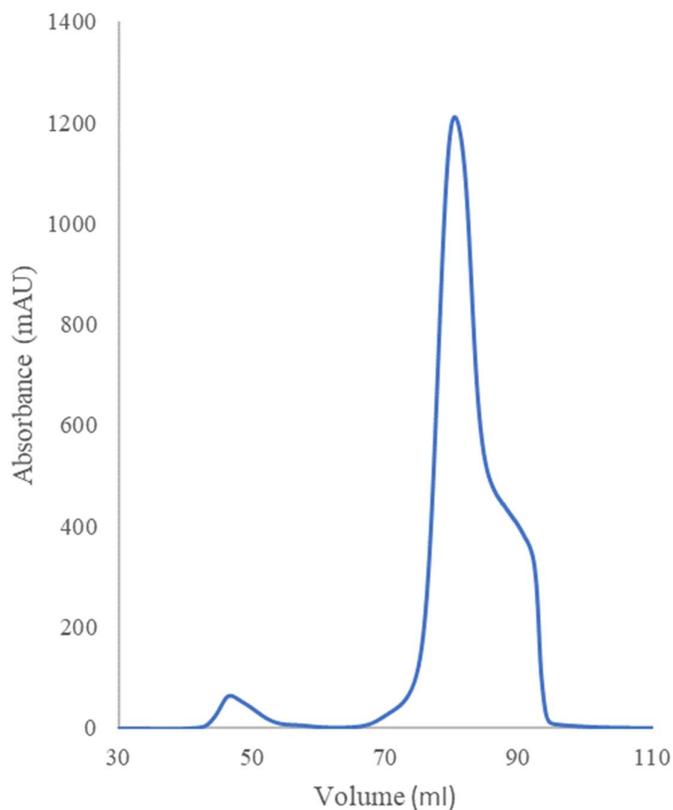


Figure 40. Chromatogram from SEC purification of *ShABS* two peaks, the second of which has a prominent shoulder.

3.2.2. Probing the Substrate Scope of *ShABS*

The activity of *ShABS* was then investigated. First *ShABS* was tested with the native reaction of McbA with **3** and **a**. The reaction was analysed by HPLC and showed 100% conversion after 24 h. Encouraged by this result the ability of *ShABS* to act on various β -carboline substrates was investigated. The results of this can be seen in Table 8.

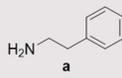
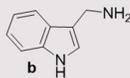
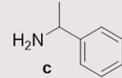
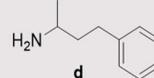
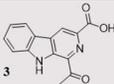
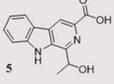
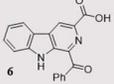
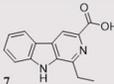
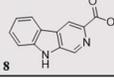
	 a			 b			 c			 d		
	McbA	AcABS	ShABS	McbA	AcABS	ShABS	McbA	AcABS	ShABS	McbA	AcABS	ShABS
 3	3a			3b			3c			3d		
	100%	100%	100%	81%	57%	100%	6%	6%	95%	26%	n.d	100%
 5	5a			5b			5c			5d		
	96%	100%	100%	100%	100%	92%	0%	0%	47%	17%	0%	99%
 6	6a			6b			6c			6d		
	82%	28%	97%	0%	0%	64%	0%	0%	2%	2%	0%	20%
 7	7a			7b			7c			7d		
	100%	100%	100%	91%	92%	54%	0%	0%	85%	7%	0%	100%
 8	8a			8b			8c			8d		
	6%	46%	87%	0%	14%	88%	0%	0%	49%	0%	0%	79%

Table 8. Conversions after 24 h with 1: 1.2 molar ratio of acid to amine. Conversions are shown for McbA, *AcABS* and *ShABS*. McbA and *AcABS* conversions are shown as a comparison and had been previously determined.

The results show that *ShABS* has a similar substrate preference to McbA and *AcABS*, showing an apparent 100 % conversion of **3a** and **7a**, the same as McbA and *AcABS*. It is however notable that in many cases, such as **5c**, **6b**, **7c**, **8c** and **8d**, *ShABS* is able to carry out conversions not possible with McbA or *AcABS*. This activity complements the activities of McbA and *AcABS* and allows for the conversion of a greater range of substrates. *ShABS* appears to accept **8** and **c** much more readily than the other homologues. It is also notable that to some extent *ShABS* is able to carry out all of the amide bond formations in this screen. This indicates a broader substrate scope of *ShABS* compared to McbA. The phenylalanine in position 301 in McbA is conserved in *AcABS* but replaced by a tryptophan in *ShABS*. The bulkier tryptophan may reduce the size of the substrate binding site and thus improve the affinity of *ShABS* for some of the smaller substrates such as **8** and **c**. Other residues that interact with the acid substrate such as L202 and G295 are conserved. The diversity in selectivity between the homologues allows a greater coverage of substrates and forms the beginnings of an enzymatic toolbox for amide bond synthesis.

An interesting result is the apparent 100 % conversion of *Sh*ABS with **3** and **d** (Figure 41). This indicates a lack of enantioselectivity which contrasts the 96% ee of McbA for the *S* enantiomer.

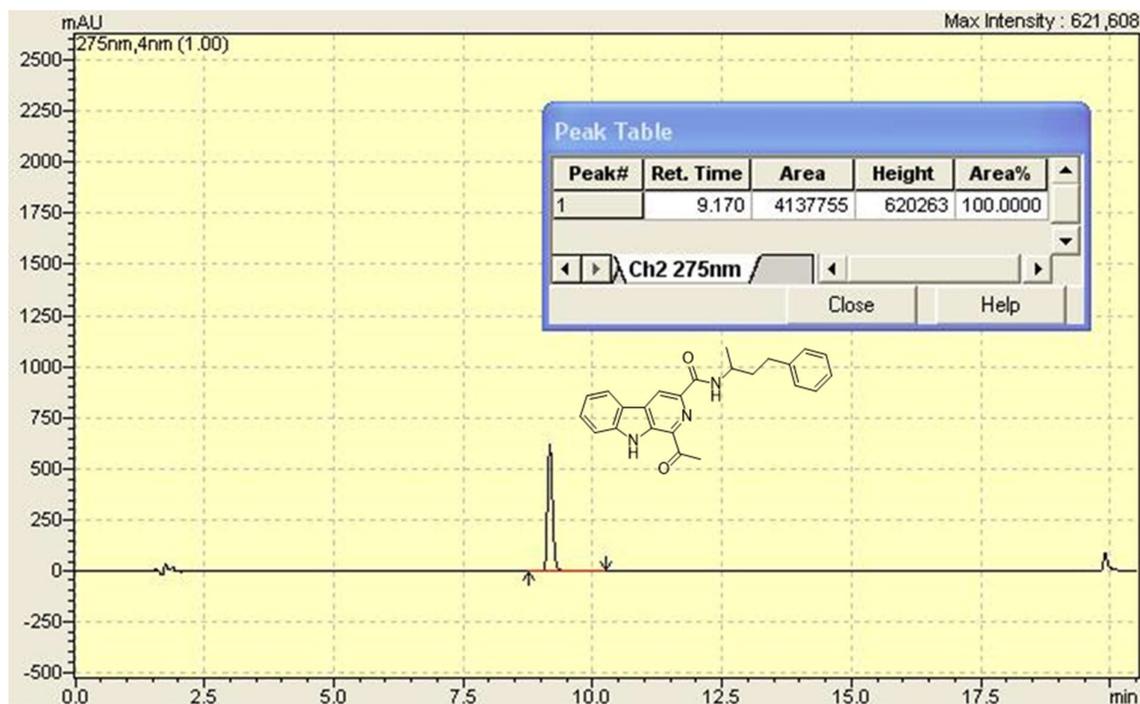


Figure 41. HPLC analysis of the reaction of **3** and **d** with *Sh*ABS to form **3d** showing 100% conversion. Retention time for **3d**= 9.2 min

3.2.3. Investigating the Enantioselectivity of *Sh*ABS

In order to investigate the enantioselectivity of *Sh*ABS a 10 mg scale up reaction was carried out. The product of the reaction was characterised by NMR spectrometry and mass spectroscopy. The enantiomers were successfully separated by chiral HPLC and the reaction was compared to a racemic mixture (Figure 42).

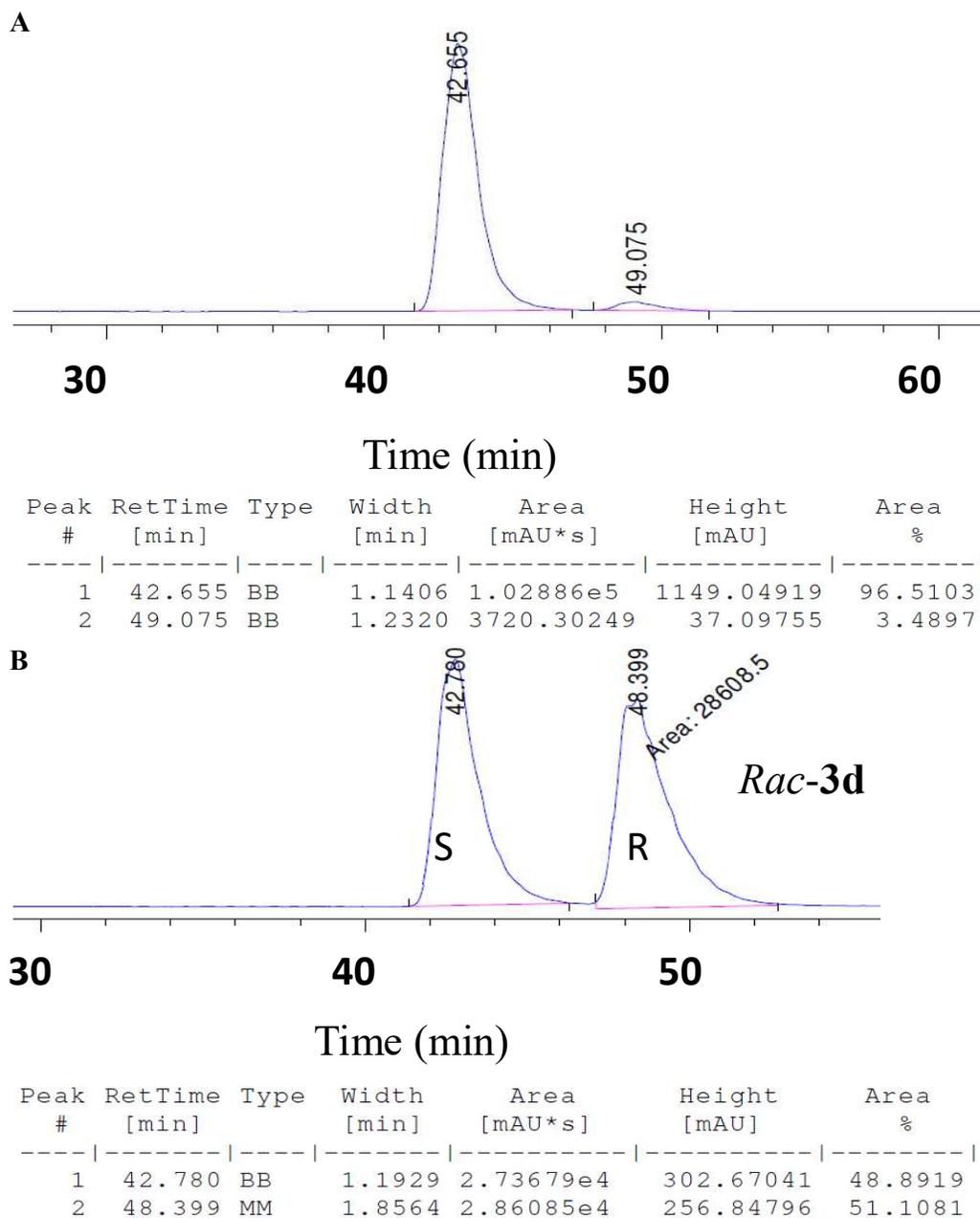


Figure 42. **A**: Chiral HPLC of the 10 mg scale up reaction of **3** and **d** with *ShABS*. **B**: Chiral HPLC of a racemic mixture of **3d**. Comparing **A** and **B** shows that *ShABS* catalyses the reaction with a 93% ee.

This showed that *ShABS* catalysed the reaction with 97% ee which was highly unexpected given the 100% conversion seen on a small scale. The scale up reaction gave an isolated yield of 79%. This result indicated that the amount of amine present in the small scale reactions must

be greater than thought as all of the acid was converted to amide. The good yield from the scale up reaction does however indicate a promising scalability for *Sh*ABS catalysed reactions.

4. Conclusion

The difficulty in crystallising McbA was overcome by utilising a systematic approach which resulted in the production of two crystal structures of McbA with AMP and **3** bound in the active site. These structures have proved to be very informative and have allowed the identification of McbA as an ANL-like adenylating enzyme. They have also shed light on the substrate specificity of McbA showing the large active site with few specific interactions to the carboxylic acid substrate. This may explain the reason for the broad selectivity of McbA. The structures also contained two conformations of McbA, that has provided increased mechanistic understanding of ABSs by allowing us to see changes to the active site with the conformational change. A putative amine binding channel was also identified by comparison to 4CBL (Figure 32). The structure was then used to design a mutant for activity assays that will add additional insight into the mechanism of these enzymes. In addition to this a new homologue of McbA, *ShABS*, was identified. *ShABS* was successfully expressed and purified and substrate screening was carried out. Initial screening has been very positive for *ShABS* which is able to fill in gaps in specificity not covered by McbA or *AcABS*.

Attempts have been made to co-crystallise McbA with the amine substrate **a**. Poor quality crystals were produced in the same conditions that generated the wt-McbA crystals but these were not fishable. Attempts to optimise these with seeding of K483A crystals was unsuccessful. Further attempts to optimise these crystals with different additives and amine substrates may be able to produce high quality crystals of the amine complex. This would be extremely valuable for our understanding of the enzymatic mechanism. Crystallisation of the homologues of McbA would also further our understanding of these enzymes, allowing identification of conserved structural motifs and reasoning for differences in activity. This knowledge could be used to inform future engineering work towards the production of an effective broad range amide bond forming catalyst. The current structures could be used to design more rigid variants of these proteins by truncating the C-terminal domain or by mutating the hinge residue. This would however make these structures less valuable as it appears that the flexibility is key for activity. Assay of the D201A mutant's activity will be undertaken which will hopefully further our understanding of the mechanism of these enzymes with respect to the amidation step. This

would be useful as the current structures have not provided as much insight into the amidation step as the adenylation step due to the lack of bound amine substrate.

Due to the promising outlook of *ShABS*, screening of substrates should be extended beyond β -carboline acids and with a wider range of amines. Eventually it is hoped that these enzymes can be engineered towards the production of industrially relevant compounds. The complementary substrate scope of *ShABS* to *AcABS* and *McbA* may make it a more suitable candidate for engineering towards the production of some substrates than *McbA* or *AcABS*.

McbA and homologues offer exciting new potential for biocatalytic amide bond formation. Their broad substrate scope, activity on bulky aromatic compounds and high enantioselectivity adds up to a very promising starting point for the development of a new enzymatic toolbox for amide bond formation. The structures of *McbA* provide a wealth of information on the mode of action of these proteins which will be vitally important for the engineering of these ABSs.

References

- [1] S. D. Roughley, A. M. Jordan, *Journal of Medicinal Chemistry* **2011**, *54*, 3451-3479.
- [2] J. S. Carey, D. Laffan, C. Thomson, M. T. Williams, *Organic & Biomolecular Chemistry* **2006**, *4*, 2337-2347.
- [3] E. Valeur, M. Bradley, *Chemical Society Reviews* **2009**, *38*, 606-631.
- [4] J. R. Dunetz, J. Magano, G. A. Weisenburger, *Organic Process Research & Development* **2016**, *20*, 140-177.
- [5] E. Fischer, E. Fourneau, *Berichte Der Deutschen Chemischen Gesellschaft* **1901**, *34*, 2868-2877.
- [6] L. A. Carpino, *Journal of the American Chemical Society* **1993**, *115*, 4397-4398.
- [7] L. Bergkamp, N. Herbatschek, *Review of European, Comparative & International Environmental Law* **2014**, *23*, 221-245.
- [8] S. D. Boggs, J. D. Cobb, K. S. Gudmundsson, L. A. Jones, R. T. Matsuoka, A. Millar, D. E. Patterson, V. Samano, M. D. Trone, S. P. Xie, X. M. Zhou, *Organic Process Research & Development* **2007**, *11*, 539-545.
- [9] D. J. C. Constable, P. J. Dunn, J. D. Hayler, G. R. Humphrey, J. L. Leazer, R. J. Linderman, K. Lorenz, J. Manley, B. A. Pearlman, A. Wells, A. Zaks, T. Y. Zhang, *Green Chemistry* **2007**, *9*, 411-420.
- [10] J. Pitzer, K. Steiner, *Journal of Biotechnology* **2016**, *235*, 32-46.
- [11] G. Bierbaum, H. G. Sahl, *Current Pharmaceutical Biotechnology* **2009**, *10*, 2-18.
- [12] H. Yamada, M. Kobayashi, *Bioscience Biotechnology and Biochemistry* **1996**, *60*, 1391-1400.
- [13] A. Goswami, S. G. Van Lanen, *Molecular Biosystems* **2015**, *11*, 338-353.
- [14] S. A. Wilson, R. M. Daniel, K. Peek, *Biotechnology and Bioengineering* **1994**, *44*, 337-346.
- [15] J. U. Klein, V. Cerovsky, *International Journal of Peptide and Protein Research* **1996**, *47*, 348-352.
- [16] I. Gill, R. Lopezfandino, E. Vulfson, *Journal of the American Chemical Society* **1995**, *117*, 6175-6181.
- [17] N. Wehofskey, M. Alisch, F. Bordusa, *Chemical Communications* **2001**, 1602-1603.
- [18] A. Pessina, P. Luthi, P. L. Luisi, J. Prenosil, Y. S. Zhang, *Helvetica Chimica Acta* **1988**, *71*, 631-641.

- [19] A. K. Prasad, M. Husain, B. K. Singh, R. K. Gupta, V. K. Manchanda, C. E. Olsen, V. S. Parmar, *Tetrahedron Letters* **2005**, *46*, 4511-4514.
- [20] P. K. Linden, A. W. Pasculle, D. McDevitt, D. J. Kramer, *Journal of Antimicrobial Chemotherapy* **1997**, *39*, 145-151.
- [21] R. Finking, M. A. Marahiel, *Annual Review of Microbiology* **2004**, *58*, 453-488.
- [22] B. K. Hubbard, C. T. Walsh, *Angewandte Chemie-International Edition* **2003**, *42*, 730-765.
- [23] H. C. Neu, *Science* **1992**, *257*, 1064-1073.
- [24] M. Winn, J. K. Fyans, Y. Zhuo, J. Micklefield, *Natural Product Reports* **2016**, *33*, 317-347.
- [25] P. Berg, *Journal of Biological Chemistry* **1956**, *222*, 991-1013.
- [26] B. Diez, S. Gutierrez, J. L. Barredo, P. Vansolingen, L. H. M. Vandervoort, J. F. Martin, *Journal of Biological Chemistry* **1990**, *265*, 16358-16365.
- [27] E. Conti, N. P. Franks, P. Brick, *Structure* **1996**, *4*, 287-298.
- [28] A. M. Gulick, *ACS Chemical Biology* **2009**, *4*, 811-827.
- [29] J. A. Gerlt, P. C. Babbitt, *Annual Review of Biochemistry* **2001**, *70*, 209-246.
- [30] S. Schmelz, J. H. Naismith, *Current Opinion in Structural Biology* **2009**, *19*, 666-671.
- [31] G. L. Challis, *ChemBiochem* **2005**, *6*, 601-611.
- [32] C. S. Francklyn, *Biochemistry* **2008**, *47*, 11695-11703.
- [33] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO Journal* **1997**, *16*, 4174-4183.
- [34] M. Gocht, M. A. Marahiel, *Journal of Bacteriology* **1994**, *176*, 2654-2662.
- [35] A. R. Horswill, J. C. Escalante-Semerena, *Biochemistry* **2002**, *41*, 2379-2387.
- [36] J. J. May, N. Kessler, M. A. Marahiel, M. T. Stubbs, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12120-12125.
- [37] A. M. Gulick, X. F. Lu, D. Dunaway-Mariano, *Biochemistry* **2004**, *43*, 8670-8679.
- [38] R. Wu, J. Cao, X. F. Lu, A. S. Regeer, A. M. Gulick, D. Dunaway-Mariano, *Biochemistry* **2008**, *47*, 8026-8039.
- [39] P. Venkitasubramanian, L. Daniels, J. P. N. Rosazza, *Journal of Biological Chemistry* **2007**, *282*, 478-485.
- [40] W. Finnigan, A. Thomas, H. Cromar, B. Gough, R. Snajdrova, J. P. Adams, J. A. Littlechild, N. J. Harmer, *Chemcatchem* **2017**, *9*, 1005-1017.

- [41] D. Gahloth, M. S. Dunstan, D. Quaglia, E. Klumbys, M. P. Lockhart-Cairns, A. M. Hill, S. R. Derrington, N. S. Scrutton, N. J. Turner, D. Leys, *Nature Chemical Biology* **2017**, *13*, 975-+.
- [42] A. J. L. Wood, N. J. Weise, J. D. Frampton, M. S. Dunstan, M. A. Hollas, S. R. Derrington, R. C. Lloyd, D. Quaglia, F. Parmeggiani, D. Leys, N. J. Turner, S. L. Flitsch, *Angewandte Chemie-International Edition* **2017**, *56*, 14498-14501.
- [43] M. Steffensky, A. Muhlenweg, Z. X. Wang, S. M. Li, L. Heide, *Antimicrobial Agents and Chemotherapy* **2000**, *44*, 1214-1222.
- [44] M. Steffensky, S. M. Li, L. Heide, *Journal of Biological Chemistry* **2000**, *275*, 21754-21760.
- [45] F. Pojer, S. M. Li, L. Heide, *Microbiology-SGM* **2002**, *148*, 3901-3911.
- [46] U. Galm, M. A. Dessoy, J. Schmidt, L. A. Wessjohann, L. Heide, *Chemistry & Biology* **2004**, *11*, 173-183.
- [47] E. Schmutz, M. Steffensky, J. Schmidt, A. Porzel, S. M. Li, L. Heide, *European Journal of Biochemistry* **2003**, *270*, 4413-4419.
- [48] H. B. Huang, Y. L. Yao, Z. X. He, T. T. Yang, J. Y. Ma, X. P. Tian, Y. Y. Li, C. G. Huang, X. P. Chen, W. J. Li, S. Zhang, C. S. Zhang, J. H. Ju, *Journal of Natural Products* **2011**, *74*, 2122-2127.
- [49] Q. Chen, C. T. Ji, Y. X. Song, H. B. Huang, J. Y. Ma, X. P. Tian, J. H. Ju, *Angewandte Chemie-International Edition* **2013**, *52*, 9980-9984.
- [50] C. T. Ji, Q. Chen, Q. L. Li, H. B. Huang, Y. X. Song, J. Y. Ma, J. H. Ju, *Tetrahedron Letters* **2014**, *55*, 4901-4904.
- [51] W. Fenical, P. R. Jensen, *Nature Chemical Biology* **2006**, *2*, 666-673.
- [52] D. R. Cooper, T. Boczek, K. Grelewska, M. Pinkowska, M. Sikorska, M. Zawadzki, Z. Derewenda, *Acta Crystallographica Section D-Biological Crystallography* **2007**, *63*, 636-645.
- [53] D. E. Kim, D. Chivian, D. Baker, *Nucleic Acids Research* **2004**, *32*, W526-W531.
- [54] M. A. Ponomarev, V. P. Timofeev, D. I. Levitsky, *FEBS Letters* **1995**, *371*, 261-263.
- [55] W. Kabsch, *Acta Crystallographica Section D-Biological Crystallography* **2010**, *66*, 125-132.
- [56] P. Evans, *Acta Crystallographica Section D-Biological Crystallography* **2006**, *62*, 72-82.
- [57] G. Winter, *Journal of Applied Crystallography* **2010**, *43*, 186-190.
- [58] A. Vagin, A. Teplyakov, *Journal of Applied Crystallography* **1997**, *30*, 1022-1025.

- [59] H. A. Crosby, K. C. Rank, I. Rayment, J. C. Escalante-Semerena, *Journal of Biological Chemistry* **2012**, 287, 41392-41404.
- [60] P. Emsley, K. Cowtan, *Acta Crystallographica Section D-Biological Crystallography* **2004**, 60, 2126-2132.
- [61] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallographica Section D-Biological Crystallography* **1997**, 53, 240-255.
- [62] F. Long, R. A. Nicholls, P. Emsley, S. Grazulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallographica Section D-Structural Biology* **2017**, 73, 112-122.
- [63] J. P. Issartel, A. Dupuis, J. Lunardi, P. V. Vignais, *Biochemistry* **1991**, 30, 4726-4733.