

**Studies on the Biosynthetic Pathways
of Clavulanic Acid and Cephamicin C
in *Streptomyces clavuligerus***

A.K. Mackenzie

*Faculty of Natural Resources and Agricultural Sciences
Department of Molecular Biology
Uppsala*

**Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2007**

Acta Universitatis Agriculturae Sueciae

2007: 19

ISSN 1652-6880
ISBN 978-91-576-7318-3
© 2007 Alasdair Mackenzie, Uppsala
Tryck: SLU Service/Repro, Uppsala 2007

Abstract

Mackenzie, A., Studies on the Biosynthetic Pathways of Clavulanic Acid and Cephameycin C in *Streptomyces clavuligerus*. Doctoral dissertation.
ISSN 1652-6880 ISBN 978-91-576-7318-3

The discovery of penicillin, a β -lactam antibiotic, changed the way humans thought about infectious disease. Unfortunately, the widespread use of antibiotics has led to a concomitant increase in bacterial resistance to these drugs. One of the most common mechanisms of bacterial resistance to β -lactam antibiotics is mediated through the hydrolysis of the β -lactam ring by β -lactamases. In order to combat resistance two strategies have been employed: identification of antibiotics resistant to hydrolysis by β -lactamases, and the development of β -lactamase inhibitors. This thesis describes studies on enzymes/proteins in the biosynthetic pathways of β -lactam antibiotics and β -lactamase inhibitors.

DAOCS is a non-heme Fe(II) dioxygenase that catalyses the oxidative ring expansion of the penicillins to cephalosporins. The expansion of the five-membered penicillin ring to a six-membered cephem ring provides increased resistance to β -lactamases. The work described here led to the production of crystals with an alternate packing of molecules (belonging to a new space group), which did not show twinning, an anomaly hampering previous structural work.

Clavulanic acid is a potent inhibitor of class A bacterial β -lactamases. CAD is a short chain reductase responsible for the catalysis of the penultimate step in clavulanic acid biosynthesis: the NADPH-dependent reduction of the unstable intermediate clavulanate-9-aldehyde to clavulanic acid. Structures of CAD in complex with the NADPH co-factor, and clavulanic acid, are described here, leading to a proposed reaction mechanism, and an increased understanding of how the enzyme is able to catalyse a reaction involving such a labile intermediate.

Approximately half of the genes in the clavulanic acid biosynthesis gene cluster are open reading frames, without a known function. The gene product of *Orf15* has been shown to be essential for clavulanic acid production. The structure reported here reveals that it shares similarity with substrate-binding proteins. The complex of ORF15 with L-arginine, a precursor of clavulanic acid, suggests multiple roles for the protein.

Keywords. β -lactam antibiotics, *Streptomyces clavuligerus*, 2-oxoglutarate dependent oxygenase, short chain dehydrogenase, oligopeptide-binding protein, clavulanic acid biosynthesis, protein crystallography, deacetoxycephalosporin C synthase, clavulanic acid dehydrogenase, arginine,

Authors address: Alasdair Mackenzie, Department of Molecular Biology, SLU, Box 590, S-751 24, Uppsala, Sweden. alasdair@xray.bmc.uu.se

Dedicated to:

Mr R. Morris who planted the seed.

Dr J. Ross, who added the fertilizer and said “grow”.

Contents

1 INTRODUCTION	11
1.1 A BRIEF HISTORY OF ANTIBIOTICS	12
1.2 ANTIBIOTIC CLASSES	13
1.2.1 β -Lactam Antibiotics	14
1.2.2 Penicillin; the First β -Lactam Antibiotic	15
1.2.3 Cephalosporins - 1 st to 4 th Generation	15
1.2.4 Other β -Lactams	17
1.3 MECHANISM AND ACTION OF β -LACTAM ANTIBIOTICS	18
1.4 MECHANISMS OF BACTERIAL RESISTANCE TO β -LACTAM ANTIBIOTICS	19
1.4.1 Resistance by Modifying Existing Systems	20
1.4.2 Protein Mediated Destruction of the β -lactam ring	20
1.5 HOW TO COMBAT ANTIBIOTIC RESISTANCE?	21
1.6 THE β -LACTAMASE INHIBITOR CLAVULANIC ACID	22
1.6.1 Mechanism of β -Lactamase Inhibition by Clavulanic Acid.	23
2 β-LACTAM BIOSYNTHESIS IN <i>STREPTOMYCES CLAVULIGERUS</i>	25
2.1 ARRANGEMENT AND REGULATION OF THE CEPHAMYCIN C/CLAVULANIC ACID SUPER CLUSTER	25
2.2 BIOSYNTHESIS OF PENICILLIN N, CEPHALOSPORIN C AND CEPHAMYCIN C	27
2.3 THE BIOSYNTHESIS OF CLAVULANIC ACID	29
2.4 AIMS OF THE THESIS	33
3 DEACETOXYCEPHALOSPORIN C SYNTHASE, DAOCS (PAPER I)	35
3.1 BACKGROUND	35
3.2 AN ALTERNATE CRYSTAL PACKING OF DAOCS IS REQUIRED	36
3.3 SUB-CLONING, AND INITIAL EXPRESSION AND PURIFICATION TRIALS	37
3.4 A SUMMARY OF SUBSEQUENT RESULTS	37
3.4.1 Effect of His ₍₆₎ -Tag on Crystallization	38
3.4.2 Effect of His ₍₆₎ -Tag on Crystal Packing	38
4 CLAVULANIC ACID DEHYDROGENASE, CAD (PAPER II)	40
4.1 BACKGROUND	40
4.2 EXPRESSION, PURIFICATION, AND CRYSTALLIZATION OF CAD	41
4.2.1 Initial Trials Using Recombinant Fusion Protein	41
4.2.2 Initial Trials Using Native Protein	41
4.2.3 Expression, Purification, and Crystallization of SeMet Protein	42
4.3 STRUCTURE DETERMINATION AND MODEL BUILDING	43
4.3.1 Location of the Selenium Atom Sites	43
4.3.2 Model Building	43
4.3.3 Model Refinement	44
4.4 STRUCTURE OF CAD	44
4.4.1 Secondary and Tertiary Structure	44
4.4.2 Quaternary Structure	45
4.5 SIMILARITY WITH OTHER SDRS	47
4.5.1 Sequence Similarity	47
4.5.2 Structural Similarity	48

4.6 THE COFACTOR AND SUBSTRATE BINDING SITES	50
4.6.1 <i>NADPH Cofactor Binding in the Binary Complex</i>	50
4.6.2 <i>Substrate Binding Site</i>	51
4.6.3 <i>Reaction Stereochemistry and Mechanism</i>	53
4.7.1 <i>Stabilizing the Unstable?</i>	55
5 ORF15 (PAPER III)	57
5.1 BACKGROUND	57
5.2 EXPRESSION, PURIFICATION AND CRYSTALLIZATION OF ORF15	58
5.3 STRUCTURE DETERMINATION	59
5.4 THE STRUCTURE OF ORF15	60
5.4.1 <i>Secondary and Tertiary Structure</i>	60
5.4.2 <i>Structural Similarity to Other SBPs</i>	62
5.4.3 <i>Domain Movements in ORF and Other Related SBPs</i>	63
5.4.4 <i>Crystal Soaks with Potential Substrates</i>	66
5.4.5 <i>Arginine Binding and the Binding Cleft</i>	66
5.5 THE FUNCTION OF ORF15	68
5.5.1 <i>Is ORF15 a Potential Cytoplasmic “Intermediate Chaperone”?</i>	69
5.5.2 <i>Is ORF15 a Potential Periplasmic Oligopeptide Binding Protein?</i>	70
5.5.3 <i>Structural Similarity Between ORF15 and ORF7?</i>	71
6. FUTURE PROSPECTIVES	73
6.1 DAOCS	73
6.2 CAD	73
6.3 ORF15	74
7 REFERENCES	75
8 ACKNOWLEDGMENTS	87

Appendix

Papers I-III

This thesis is based on the papers I-III, which will be referred to by their Roman numerals:

- I. Öster, L. M., Terwisscha van Scheltinga, A. C., Valegård, K., **MacKenzie-Hose, A.**, Dubus, A., Hajdu, J. & Andersson, I. (2004) Conformational Flexibility of the C terminus with Implications for Substrate Binding and Catalysis Revealed in a New Crystal Form of Deacetoxycephalosporin C synthase. *J. Mol. Biol.* 343, 157–171
- II. **MacKenzie, A. K.**, Kershaw, N. J., Hernandez, H., Robinson, C.V., Schofield, C.J. & Andersson, I. (2007) Clavulanic Acid Dehydrogenase: Structural and Biochemical Analysis of the Final Steps in the Biosynthesis of the β -Lactamase Inhibitor Clavulanic Acid. *Biochemistry* 46, 1523–1533
- III. **Mackenzie, A.K.**, Caines, M. E. C., Kershaw, N. J., Schofield, C. J., Andersson, I. & Valegård, K. The Crystal Structure of ORF15 from the Clavulanic Acid Biosynthesis Gene Cluster Suggests a Potential Role as a Carrier Protein for Biosynthetic Intermediates. *In Manuscript*.

Additional Paper

Johansson M, **Mackenzie-Hose A**, Andersson I, Knorpp C. (2004) Structure and Mutational Analysis of a Plant Mitochondrial Nucleoside Diphosphate Kinase. Identification of Residues Involved in Serine Phosphorylation and Oligomerization. *Plant Physiol.* 136, 3034–3042.

Papers I, II, are reproduced by permission of the journal.

Abbreviations

2ODD	2-oxoglutarate dependent dioxygenase
AAA	L-a-aminoadipic acid
ACV	δ (L-a-aminoadipyl)-L-cysteinyl-D-valine
BLS	β -lactam synthase
CAD	clavulanic acid dehydrogenase
CAS	clavaminic acid synthase
CEAS	N ² -(2-carboxyethyl)arginine synthase
DAC-AT	deacetylcephalosporin acetyl transferase
DAOC	deacetoxycephalosporin C
DAOCS	deacetoxycephalosporin C synthase
DAOCS/DACS	deacetoxy/deacetylcephalosporin synthase
IAT	acyl coenzyme A:isopenicillin N acyltransferase
IPNS	isopenicillin N synthase
OAT2	ornithine acetyl transferase
MIR	multiple isomorphous replacement
MAD	multiple-wavelength anomalous diffraction
NADPH	nicotinamide adenine dinucleotide (reduced form)
NAM	N-acetylmuramic acid
NAG	N-acetyl-glucosamine
NCS	non-crystallographic symmetry
ORF	open reading frame
MIR	multiple isomorphous replacement
PAH	proclavamate amidinohydrolase
PBP	penicillin binding protein
PEG	polyethylene glycol
r.m.s	root mean square
SAD	single-wavelength anomalous diffraction
SBP	substrate-binding protein
SDR	short chain dehydrogenase/reductase
Å	ångström

Foreword. Poor Old Noah

Noah is one of the few distinguished people acknowledged as a prophet in three of the world's religions. With such high accolades we can assume that Noah was a wise man. Noah is accredited with the rescue of all animal and bird species, after a cataclysmic deluge of rain lasting 40 days and 40 nights (Gen. 7:4 New International Version) purportedly flooded the earth for one hundred and fifty days (Gen. 7:24). While, there is no specific reference to plant species, one can assume that a diligent Noah was also responsible for plant conservation, as he is reported to have grown the first vineyard upon returning to *terra firma* (Gen. 9:20). If this is true for plants, then one wonders if Noah was also responsible for the preservation of microorganisms, although anaerobic bacteria, and viral species may have struggled to justify their apparent journey on the Ark (Gen. 7:15).

Judging from the number of species present on earth today it is easy to see that Noah took his job seriously, obviously undergoing great hardship in order to collect some of these species. The preservation of bacterial species such as *Neisseria gonorrhoeae* may have been a relatively easy, *albeit* a rather uncomfortable discharge of ones duties. The preservation of other diseases such as Kuru required brains, a strong stomach, and questionable ethics. One could argue of course, that not all the species Noah preserved were intended to survive the great flood. There may have been one or two hitchhikers who slipped onboard unnoticed, right under Noah's nose.

The bacterium *Staphylococcus aureus* can be found living happily, amongst other places, in the nasal passages of humans. *S. aureus* is responsible for a range of medical conditions ranging from minor skin infections, such as pimples, to more serious conditions including pneumonia, and septicemia. If the story of Noah is true, then the strain of *S. aureus* that slipped onboard the Ark was no ordinary bacterium. This strain was a methicillin-resistant *Staphylococcus aureus* (MRSA) which, somehow, had achieved resistance to all know penicillins of the 20th century. Some of these penicillins are completely unnatural, produced in a laboratory by semi-synthesis. This begs the question: how could a bacterium residing in Noah's nose achieve resistance to the semi-synthetic penicillin's of the 20th century, some 4000 years prior to their creation?

Indeed, it might be difficult to believe that all the MRSA alive today are direct descendants from the bacteria in Noah's nose after a lengthy excursion at sea. Perhaps we should turn to the wisdom of another seafaring man, who undertook an even longer expedition on HMS Beagle. Charles Darwin, like Noah, also had an interest in animals and plant species. If we take on board Mr. Darwin's theory of evolution through natural selection, then perhaps we might be forced to conclude that the phenomenon of bacterial resistance to antibiotics actually represents living proof of his theory; that antibacterial resistance to antibiotics is still, unfortunately, a case of evolution in action.

1 Introduction

Imagine how your life would have been without being able to turn to the aid of antibiotics; how a case of pneumonia could be fatal, or how a simple cut while shaving could turn into lethal septicemia. For my generation it is difficult to imagine, almost unfathomable. Even for my parents' generation, who were born at a time when the commercial production of penicillin was being developed, antibiotics have always been a magic bullet, killing the microbes without poisoning the infected individual. Yet if we go back to my grandfather's generation there were no antibiotics, no miracle cures. Treatments involved the use of poisons like strychnine, or arsenic, and the use of antiseptics. Unfortunately for my grandfather, who succumbed to tuberculosis, streptomycin had not yet become available.

The word antibiotic is derived from the Greek *anti* (against) and *bios* (life). Paul Vuillemin, a pupil of Louis Pasteur, coined the term "antibiosis" in 1889 to describe the process by which life could be used to destroy life. However, it was Selman Waksman who coined the term "antibiotic" in 1945. Originally the term only applied to compounds derived from living organisms capable of killing or inhibiting bacterial growth. The term has now been expanded to include synthetic antimicrobial drugs as well. Natural antibiotics belong to the group of compounds known as secondary metabolites. They are generally produced at low specific rates, and are not considered essential to the organism producing them, at least in pure culture. In their natural environment these antibiotics are essential to the organisms that produce them. Antibiotics can be considered the chemical weapons of the microbial world, were they are used for the survival and competitive advantage within the microenvironment (Demain & Fang, 2000).

When antibiotics, such as penicillin, were first introduced in the 1940's a victory was declared against infectious disease. A panacea had been found, which would change the way humans thought about infectious disease. The free love movement of the "swinging sixties" might not have been so free if it weren't for the relatively simple treatment of many sexually transmitted diseases with antibiotics. Unfortunately the euphoria of the antibiotic "magic bullet" has dissipated. The huge consumption of antibiotics, and their widespread use, and abuse, has led to the concomitant emergence of bacterial resistance to antibiotics. Over-optimism and complacency in the pharmaceutical industry, even as recently as the 1980's, saw little development towards new antibiotic drugs. At the same time the spread of antibiotic resistance has accelerated. This has led to the re-emergence of diseases once controlled, such that bacterial infections represent, once again, one of the greatest threats to human health (Cohen, 2000).

In order to address the problem of antibiotic resistance in bacteria we must first understand how antibiotics work, and the mechanisms bacteria employ in order to avoid their effects.

1.1 A Brief History of Antibiotics

Although the name “Alexander Fleming” is synonymous with penicillin and antibiotics, he was certainly not the first person to observe the inhibition of bacterial growth by fungi. The antibacterial properties of various moulds have been used since the times of ancient Greece and China, and Russian peasants were known to use warm soil as a treatment for wounds. However, primitive technology did not permit the isolation of the active compounds to improve their efficiency.

The use of antibiotics in modern medicine perhaps began with Joseph Lister, who in 1871 used mould contaminated urine to treat infected wounds. His inspiration was Sir John Scott Burdon-Sandersson, who observed that bacteria were unable to grow in culture fluid covered with mould. Lister also went on to describe the antibacterial action of *Penicillium glaucum*, a fungus used in French blue cheese production. Six years later Louis Pasteur and Jules Joubert observed the inhibition of growth of anthrax bacteria using soil microorganisms (Pasteur & Joubert, 1877), while E. von Freudenreich observed the antibacterial properties of culture filtrates from *Pseudomonas aeruginosa* (Freudenreich, 1888). In 1897 the thesis of medical student Ernest Duchesne described the partial purification of a compound from a *Penicillium* mould which he used to cure a guinea pigs of typhoid. The Pasteur Institute ignored his results, and Duchesne’s achievements passed into obscurity after his early death. It is interesting to note that penicillin is ineffective against typhoid. The mould used by Duchesne was not preserved, and thus we will never know which species, or antibiotic, he had discovered.

In the early 1920’s Andre Gratia and Sara Dath observed that their *Staphylococcus aureus* cultures were inhibited by fungal contamination. They successfully identified the species as a *Penicillium*, but they did not pursue a further investigation of the moulds antibacterial properties. Around the same time Alexander Fleming observed the weak antibacterial effect of lysozyme. It was not until 1928 when Fleming made his serendipitous observation while looking through old culture plates of *S. aureus*. Fleming observed that bacterial growth was inhibited around a contaminating mould, and the subsequent isolation of the mould demonstrated that it was *Penicillium notatum*. The antibacterial agent was named penicillin (Fleming, 1929), although Fleming’s attempts to isolate the active compound proved unsuccessful, with his work remaining largely unnoticed for another nine years.

It was not long after Fleming’s discovery the first successful oral antibiotics were introduced. Gerhard Domagk had discovered the antibacterial properties among certain dyes, one of which was a sulfonamide with the trade name prontosil. The synthetic sulfonamides played a central role in preventing infection during World War II. It was the German discovered sulfonamides, and not English penicillin, which cured an ill Winston Churchill in Carthage, Tunisia, in 1943, despite reports to the contrary by the British press. Incidentally, at the same time of Domagk’s discovery, a chemist by the name of Ernst Chain left Germany to work in Oxford. Chain was soon to become intertwined in the penicillin success story.

Throughout the 1930's Fleming continued to work on penicillin. Initial tests were inconclusive, but this was due to its use as an antiseptic. In 1933 disaster struck, when fellow rifle club member, and team captain, Keith Rogers contracted conjunctivitis during a competition. Fleming made a penicillin ointment, cured his captain, and the match was won. More importantly Fleming had won over any doubts that penicillin would be unsuitable for clinical use.

In 1938 in Oxford, Howard Florey, who had previously worked with lysozyme, read Fleming's paper, and tried to purify three promising substances. One of these substances was penicillin. Ernst Chain and Norman Heathley determined how penicillin could be isolated and concentrated, as well as correctly predicting its structure, and by 1940 the Oxford team was able to produce enough penicillin to begin testing. With World War II raging the development and production of penicillin continued in the USA. Fleming's original strain only produced trace amounts of penicillin, insufficient for large-scale production, however a mould isolated from a rotting cantaloupe in a Peoria market (Raper, 1946) proved a serendipitous ally, capable of producing up to $60 \mu\text{g ml}^{-1}$ penicillin. This began a history of 'strain improvement' through brute force genetic manipulation, eventually leading to the 'Wisconsin family' of superior strains, some of which produced over $1800 \mu\text{g ml}^{-1}$ (Backus & Stauffer, 1955). By 1945 the problems with rapid secretion in the urine had been overcome, and with advances in production and semi-synthesis, penicillin was successfully mass-produced.

Around the same time Selman Waksman discovered streptomycin, which represented a new class of antibiotics. The mechanism by which streptomycin inhibits bacteria growth is quite different to that of penicillin, and it later proved successful in the treatment of tuberculosis. Since then a plethora of antibacterial agents belonging to different antibiotic classes have been identified.

1.2 Antibiotic Classes

Antibiotics can be organized into different classes according to their effect and mode of action. Their effect is either bactericidal or bacteriostatic, where the former kills the bacteria, and the latter inhibits the growth of the bacteria, allowing the immune system to deal with the infection. Antibiotics can be described as possessing different "spectra of activity", where "broad spectrum" antibiotics are effective against a large number of infectious bacterial species. The opposite is true for narrow-spectrum antibiotics, which are only active against a specific family, or genus of bacteria. The different classes of antibiotics, their effect on bacteria and principal mechanisms have been summarized in Table 1.

Table 1 Representative classes of antibiotics, their effect, and mode of action.

Class	Effect ^a	Action
Aminocyclitols	b.s	Protein synthesis
Aminoglycosides	b.c	Protein synthesis
β -lactams	b.c	Cell wall
Fluoroquinolones	b.c	Genetic replication
Lincosamides	b.s	Protein synthesis
Macrolides	b.s	Protein synthesis
Sulfonamides	b.s	Metabolic processes
Tetracyclines	b.s	Protein synthesis
Phenicol	b.c	Protein synthesis

^a. The effect of antibiotics defined as either bacteriostatic (b.s) or bacteriocidal (b.c).

1.2.1 β -Lactam Antibiotics

Penicillin belongs to the class of β -lactam antibiotics. The β -lactams possess a bacteriocidal effect, except against *Enterococcus sp.*, where they act as a time dependent killer. They are among the most successful antibiotics, and account for 60-65% of the total world market, generating revenues in the order of \$US 15 - 23 billion, per annum (Demain & Elander, 1999; Elander, 2003).

The β -lactam antibiotics share a common chemical motif, which consists of a strained, four-membered, heterocyclic ring, known as the β -lactam ring. The β -lactam ring is typically fused to a second ring forming a bicyclic ring structure. The second ring may be five-membered (penam/penem) or six-membered (cepham/cephem), where the nomenclature used reflects the ring saturation (Figure 1.1).

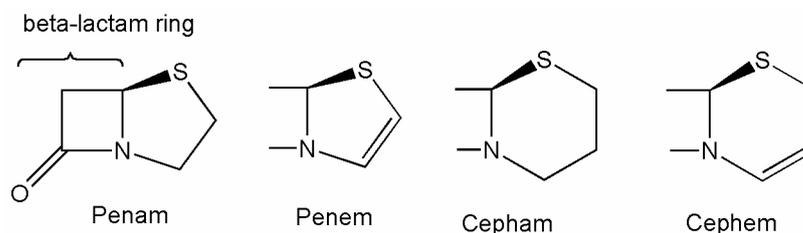


Figure 1.1 Structure of the β -lactam ring, the common motif of the β -lactam antibiotics, and the nomenclature of the second ring fused to the β -lactam ring

The β -lactams can be sub-divided into a number of different groups, in part defined by the chemistry of the second ring, and by group substitution on the heterocyclic ring. Significant variation can be found in the chemical structures of the different β -lactams. Such variation is important as it creates a range of characteristics defining the physical properties of the antibiotic, e.g. solubility, stability, as well as the spectrum of activity. While some of these variations are synthesized naturally, many are introduced through the methods of semi-synthesis of fermented compounds.

1.2.2 Penicillin; the First β -Lactam Antibiotic

Although Chain and Abraham had predicted the structure of penicillin it was the pioneering work of Dorothy Crowfoot Hodgkins who determined the three dimensional structure of penicillin using X-ray crystallography (Crowfoot *et al.*, 1949). The penicillins consist of a bicyclic ring, where the β -lactam ring is fused to a five-membered thiazolidine (penam) ring containing a sulphur atom (Figure 1.2). The bicyclic rings are arranged such that they possess $3R,5R$ stereochemistry. The penicillin nucleus consists of 6-aminopenicillanic acid (6-APA), which can be modified at the R_1 position to introduce a range of different side chains of different chemical properties. More recently advances in the chemical synthesis now permit the synthesis of a number of other bicyclic β -lactam penams/penems, (see section 1.2.4).

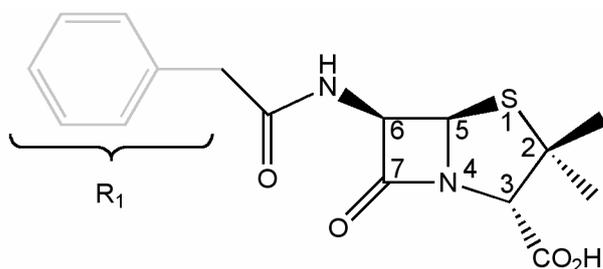


Figure 1.2 The structure of the benzylpenicillin (PenG). The penicillin 6-APA core is shown in black, while the R_1 group is shown in light grey. The numbering of the carbons across the bicyclic ring is shown.

1.2.3 Cephalosporins - 1st to 4th Generation

Not long after the successful mass production of penicillin, a new β -lactam antibiotic was discovered. Cephalosporins were produced by a fungus originally identified at in a Sardinian sewer (Brotzu, 1948), and showed increased activity towards Gram-negative bacteria. The cephalosporins differ from the penicillins in that the β -lactam ring is fused to a six-membered (cephem) dihydrothiazine ring (Figure 1.3). The expansion of the five-membered ring to a six-membered ring also affords an increased resistance against bacterial β -lactamases (see section 1.4). The cephalosporin nucleus consists of 7-aminocephalosporanic acid (7-ACA), which can be modified at the R_1 and R_2 positions to produce a variety of different cephalosporins.

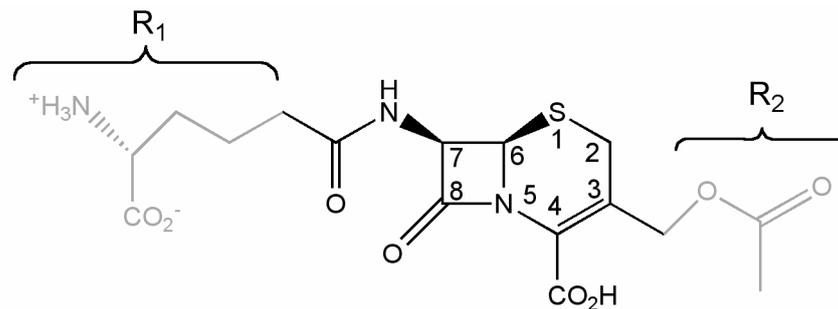


Figure 1.3 Structure of cephalosporin C. The 7-ACA nucleus is shown in black, while the R1 and R2 groups are shown in light grey. The numbering of carbons across the bicyclic ring is shown.

Both penicillin and cephalosporin C were isolated from filamentous fungi, *e.g.* *Penicillium chrysogenum* or *Cephalosporium acremonium*. However, the synthesis of β -lactam antibiotics is not exclusively limited to fungi. Bacteria, such as *Streptomyces clavuligerus* and *Amycolatopsis lactamdurans*, are capable of producing an even broader range of β -lactam antibiotics. It is these “bacterial antibiotics” which helped to give rise to new generations of cephalosporins.

The cephalosporins can be subjectively divided into “generations” based on their antimicrobial activity and resistance to β -lactamases. In general each successive cephalosporin generation is less active against Gram-positive bacteria, however the activity against Gram-negative bacteria has increased, particularly against aerobes. Their discovery was an important solution to the emerging problem of bacterial resistance to penicillins.

For historical reasons the cephameycins and carbacephems have been classified with the second generation cephalosporins, although technically speaking they are not cephalosporins, since they lack the 7ACA core. However, like the cephalosporins they consist of a β -lactam ring fused to a cephem ring. In the carbacephems a carbon is substituted for the sulphur at position 1, while the β -lactam ring of the cephameycins has a methoxyl group at the C7-position (see Figure 1.4). The third generation cephalosporins include the oxacephems, although they are sometimes regarded as fourth generation antibiotics. In the oxacephems oxygen is substituted at position 1 of the cephem ring instead of carbon or sulphur. The fourth generation cephalosporins show an extended spectrum of activity, with increased activity towards bacterial β -lactamases.

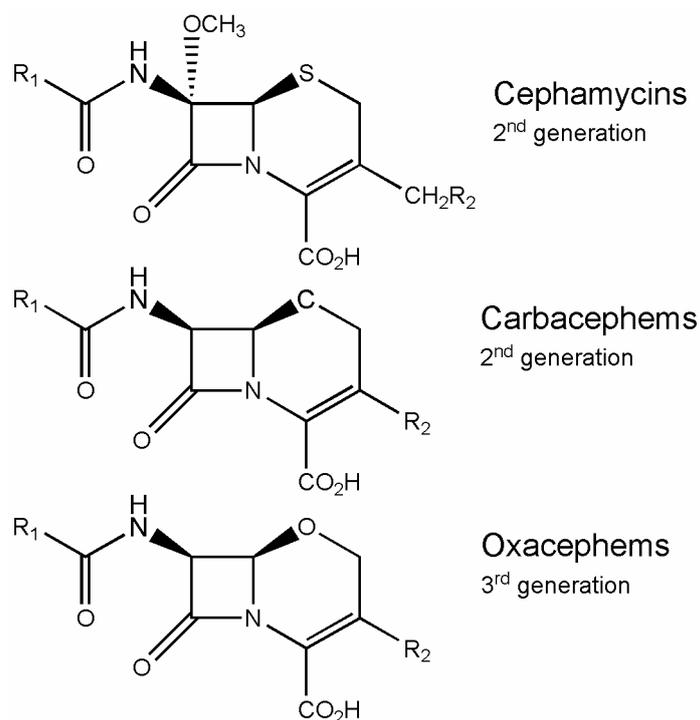


Figure 1.4 Core structures of cephamecins, carbacephems, and oxacephems representing the second and third generation cephalosporins. Note that the cephem ring of carbacephems and oxacephems may also be saturated resulting in the carbacephams or oxacephams, respectively.

1.2.4 Other β -Lactams

Other β -lactam structures have been isolated and characterized from bacterial sources. These include the monolactams, penems and carbapenems (Figure 1.5), as well as the clavams (Figure 1.6). The monolactams represent perhaps the most literal definition of a β -lactam antibiotic, consisting of a naked β -lactam ring. A second ring may be present as the R₁ or R₂ side chains, *e.g.* the side-chain of the monobactam Aztreonam (see Figure 1.5).

The core penem structure resembles penicillin, except the second ring possesses an unsaturated penem ring, and the R₁ side chain has moved to the C2 position of the penem ring. In the carbapenems, the sulphur at position one of the penem ring is substituted for a carbon atom. Using chemical synthesis it is possible to replace the sulphur for a selenium, nitrogen, or oxygen giving rise to selenapenam, azapenam or oxapenam, respectively. Of all the β -lactam antibiotics, it is the carbapenems that possess the broadest spectrum activity (Kropp *et al.*, 1985).

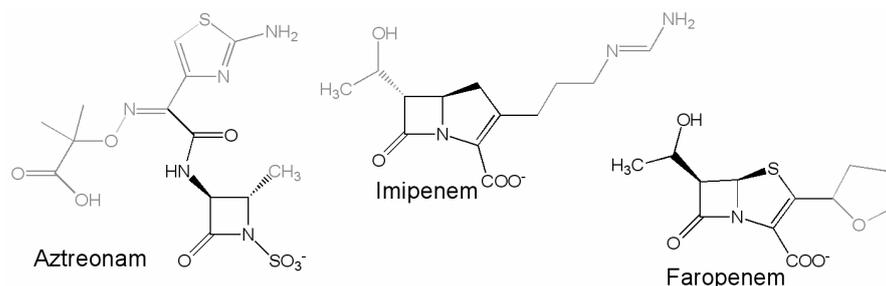


Figure 1.5 Examples of the monolactams (Aztreonam), penems (Faropenem) and carbapenems (Imipenem). The core structure is shown in black, and the R_1 and/or R_2 side chain in grey.

The clavam nucleus is also similar to that of penicillin, except that the penam ring contains oxygen instead of sulphur at position one. The R_1 side chain is located off the C2 position on the penam ring, with an additional R_2 side chain at the C3 position. Unlike the other β -lactams, the clavams show variation in the stereochemical arrangement of the penam ring with respect to the β -lactam ring (Figure 1.6). The significance of this stereochemical difference is discussed later in section 1.6).

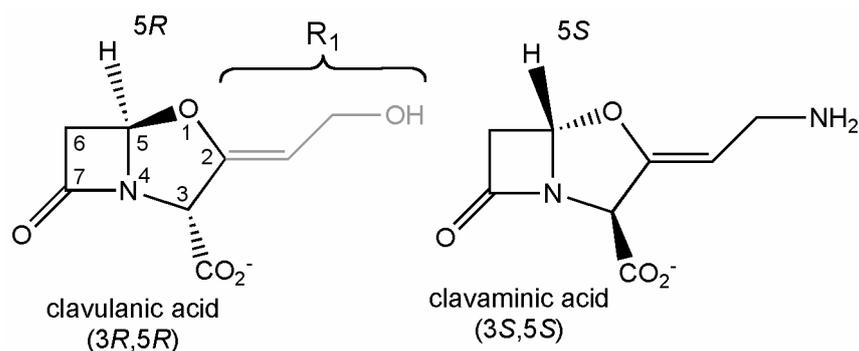


Figure 1.6. The stereochemical arrangement of the bicyclic rings in $3R,5R$ -clavulanic acid and $3S,5S$ -clavaminic acid. The hydrogen of the C5 carbon can be either pro- R or pro- S . Likewise the stereochemistry of the C3 carboxyl group can be pro- R or pro- S . The clavam R_1 side-chain is shown in grey.

1.3 Mechanism and Action of β -Lactam Antibiotics

The β -lactam antibiotics act by interfering with the enzymes involved in the growth of the bacterial cell wall, and only exert a bactericidal effect on growing cells. The cell wall not only determines the shape of the organism, but also protects it as well. The cell wall is composed of, amongst other things, a peptidoglycan layer which forms a crystalline lattice. The peptidoglycan layer is formed from linear chains of two alternating amino sugars, N-acetyl-glucosamine (NAG) and N-acetyl-muramic acid (NAM). The NAM/NAG-peptide subunits are then cross-linked together by a short amino acid chain containing D-alanine, D-glutamine and mesodiaminopimelic acid (or lipid II), although the composition varies between

species. The final transpeptidation step is facilitated by membrane bound DD-transpeptidases, also known as penicillin binding proteins (PBPs) (Fig 1.7a). The β -lactam antibiotics are analogues of D-alanyl-D-alanine, and thus act as pseudo-substrates of the PBPs. Once bound in the active site the β -lactam ring is hydrolysed, acylating the active site serine to form a stable covalent acyl-enzyme complex (Fig 1.7b).

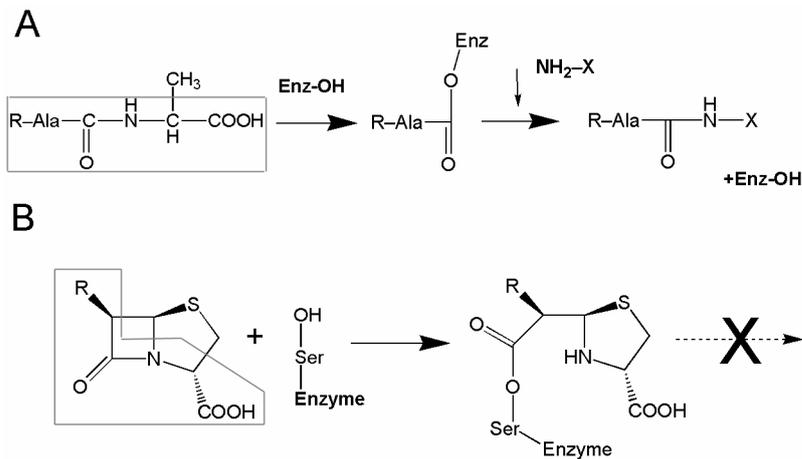


Figure 1.7 The reaction catalysed by PBPs, and their inhibition by β -lactam antibiotics. (A) The transpeptidase reaction catalysed by PBPs. (B) The inhibition of PBPs by β -lactam antibiotics. The structural analogy between D-Ala-D-Ala and the β -lactam antibiotics are shown with a grey box.

For a long time it was thought that the inactivation of the PBP prevented the cross-linking of the NAG/NAM subunits in the peptidoglycan layer, in turn weakening the cell wall, eventually leading to bacteriolysis (Tipper & Strominger, 1965). However the lysis of the bacterial cell wall has been shown to be enzymatically mediated (Rogers & Forsberg, 1971; Tomasz *et al.*, 1970). Thus, while the inhibition of the cell wall construction results in the lysis of the cell, the mechanism by which the cell dies is not fully understood (Bayles, 2000).

1.4 Mechanisms of Bacterial Resistance to β -Lactam Antibiotics

Almroth Wright, whom Fleming worked for at St Marys, had predicted bacterial resistance to penicillin, even before it was observed in some strains of *E. coli* (Abraham & Chain, 1940). By the 1950's resistance to penicillin in *Staphylococcus* had become a major problem, and today once a drug enters widespread therapeutic use, the clock is ticking.

Bacterial resistance to antibiotics illustrates a wonderful, if not irksome example of evolution through mutation and (un)natural selection. Given the large number of bacteria in an infection cycle, and an intrinsic mutation rate of *ca.* 10^7 it is only a matter of time before a mutation inexorably arises conferring some form of resistance. The resistant mutant soon becomes the dominant variant in the bacterial

population, especially if the antibiotic is given at sub-therapeutic levels, since the survival of the resistant bacteria is almost guaranteed. Many of the genes conferring antibiotic resistance are borne on plasmids, which can be rapidly spread between bacterial cells and species. The activity of transposons' can later lead to the incorporation of the resistance genes into the bacterial genome.

1.4.1 Resistance by Modifying Existing Systems

There are several different mechanisms by which bacteria can protect themselves against antibiotics (Poole, 2004; Walsh, 2000), where more than one may be simultaneously employed. General mechanisms, such as improved efflux pumps, prevent the accumulation of therapeutic drug levels within the cell. Similarly, the channels in the outer-membrane formed by porins can be altered, changing the permeability of the membrane, thus preventing the antibiotic from accumulating in the cell. Such mechanisms can be effective against antibiotics from different classes, targeting different processes, *e.g.* tetracyclines (protein synthesis) and β -lactams (cell wall synthesis). Alternatively, the target of the antibiotic can also be modified, or "reprogrammed", in order to reduce its susceptibility to the drug. For example, mutations in the active site of a PBP has resulted in a β -lactam antibiotic-insensitive transpeptidase (Mainardi *et al.*, 2002).

1.4.2 Protein Mediated Destruction of the β -lactam ring

The most prevalent mechanism of antibiotic resistance employed by bacteria involves the enzymatic destruction of the β -lactam ring by a group of enzymes known as β -lactamases (penicillinases). Thus, while the β -lactam ring acts as a champion inhibiting PBPs, it is also the Achilles heel when challenged by β -lactamases. The β -lactamases are structurally related to PBPs (Massova & Mobashery, 1998), and are capable of hydrolysing the β -lactam ring (Figure 1.8); thus they are potentially active upon a number of different β -lactam antibiotics. Secreted into the periplasm these enzymes can destroy the β -lactam antibiotics before they can even reach their target.

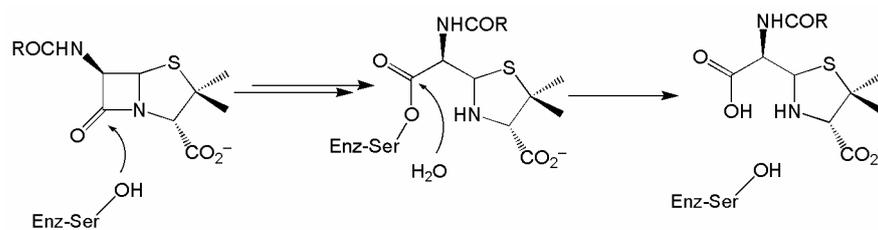


Figure 1.8 The hydrolysis of the labile acyl-enzyme complex catalysed by a serine β -lactamase.

The β -lactamases have been classified according to their sequence similarities, substrate preference, and whether they are encoded for by genes on a plasmid or chromosome. They can also be classified by their mechanism of action, such that class A (penicillinases), class C (cephalosporinases) and class D (oxacillinases) β -lactamases are metal-independent enzymes, utilizing a catalytic serine. The class B (metallo-enzymes) enzymes employ a metal ion, *e.g.* Zn^{2+} , at their active site, and

can be inhibited by chelating agents. Despite the growing importance of metallo- β -lactamases, it is the serine β -lactamases that represent the greatest clinical challenge. The classical extended-spectrum- β -lactamases, which have evolved from class A TEM and SHV enzymes, represent the most prevalent types, with over 150 TEM/SHV-derived enzymes reported.

1.5 How to Combat Antibiotic Resistance?

Answering the above question with a pithy “Stop using antibiotics!” might be considered a stupid reply by some people. Then again, if you consider that the problem of bacterial antibiotic resistance arose due to our wide spread use/abuse of antibiotics, then removing the selection pressure would relieve the need for bacteria to evolve mechanisms to circumvent the effect of antibiotics. Of course, such a solution would be rather impractical. If we elect to stop using antibiotics our fate would be the same as if we continued to over-use antibiotics, *i.e.* we are left with no effective treatment against infectious bacteria, a situation reminiscent of the pre-1930’s. So how can we rationally approach this problem of antibiotic resistance?

A less extreme strategy is to extend the lifespan of current antibiotics by restricting their use. In the developed world it has been estimated that 50% of prescribed antibiotics are for viral infections, which do not respond to antibiotics. In some developing countries antibiotics can be obtained without prescription at the pharmacy, in others, drug availability is limited and of dubious quality or potency. Either way this tends to create a strong tendency towards self-medication, resulting in sub-therapeutic doses of antibiotics; an outstanding method for selecting drug resistant bacteria. Restricting the use of antibiotic not only applies to humans. Take for example the consumption of the antibiotic vancomycin in Denmark in 1994; the human consumption (24 kg) was dwarfed by the amounts used for animal health (24,000 kg). Thankfully, as of the first of January, 2006, the EU has banned the use of antibiotics for growth promotion in animal feed. Unfortunately, the situation appears unlikely to be mirrored in the U.S.A, where such practice is widespread and encouraged.

Simultaneously, the development of new antibiotic classes will provide new treatment against the current multi-resistant bacteria. The use of genomics may help in the identification of essential bacterial genes, which are appropriate targets for inhibition by new drugs. For example, the enzyme peptide deformylase is an essential gene in many bacteria, to which new potent inhibitors have been developed (Chen *et al.*, 2000).

Another strategy involves the development of new antibiotics, which circumvent resistance by targeting the mechanisms of resistance, *e.g.* β -lactamases. In order to combat β -lactamase mediated resistance two strategies have been used. The first involves the modification to the β -lactam core, or variation in the R₁/R₂ side-chains. In this way the antibiotic is no longer recognized by the β -lactamase. These modifications can be achieved using chemical modifications *via* semi-synthesis, or by the identification and development of new antibiotic compounds amongst the

secondary metabolites of other bacteria. Unfortunately, only a few amino acid substitutions, at a limited number of positions, are required before resistance is once again acquired. Like the PBPs, the β -lactamases are capable of altering/expanding their binding sites, to allow the binding of the bulky oxyimino side-chains of broad spectrum β -lactams, resulting once again in the hydrolysis of the β -lactam antibiotic.

An alternate strategy involves the development of β -lactamase inhibitors. Some compounds possess both antibiotic and β -lactamase inhibitory properties, *e.g.* broad-spectrum carbapenems. Others are potent β -lactamase inhibitors, but lack sufficient antibacterial activity *per se*, and are therefore administered in combination with more potent antibiotics. The use of such inhibitors potentially allows the use of β -lactam antibiotics to which bacteria have previously developed resistance, since the inhibitor effectively “knocks out” the bacterial resistance mechanism. Several serine β -lactamase inhibitors are in clinical use, *e.g.* see Figure 1.13. Of these two are unnatural compounds (sulbactam and tazobactam) produced by semi-synthesis. A third inhibitor, and by far the most common, is clavulanic acid, which is produced by anaerobic fermentation of the actinomycete, *Streptomyces clavuligerus*.

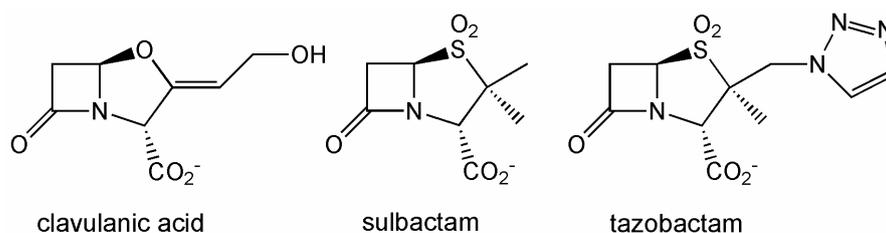
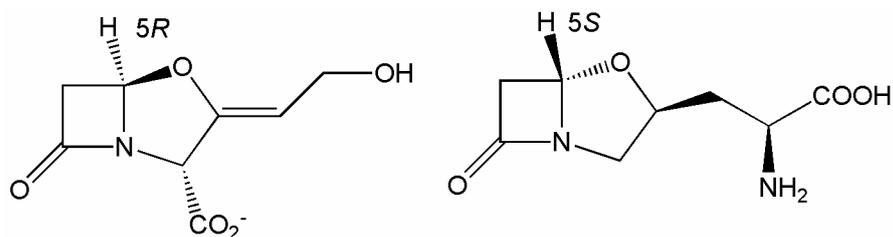


Figure 1.9 The structure of some of the serine β -lactamase inhibitors in clinical use.

1.6 The β -Lactamase Inhibitor Clavulanic Acid

As mentioned previously, the clavams can be classified according to the stereochemical arrangement of their bicyclic ring, *i.e.* the position of the C5 hydrogen can be either pro-*R* or pro-*S* (see Figure 1.10). The X-ray structure of clavulanic acid (Brown *et al.*, 1984; Howarth *et al.*, 1976) reveals that the β -lactamase inhibitor possesses 3*R*,5*R* stereochemistry across the strained bicyclic ring, which it shares with penicillins. In this respect clavulanic acid acts as a very slow substrate for β -lactamases. On the other hand, the 5*S*-clavams do not show β -lactamase inhibitory activity, although several possess antibacterial and antifungal activity (Baggaley *et al.*, 1997). Because the stereochemistry of the 5*S*-clavams is different to that of clavulanic acid and the other β -lactam antibiotics, the mechanism by which the 5*S*-clavams inhibit bacterial growth is also different. For example, valclavam and hydroxyethylclavam act by the non-competitive inhibition of bacterial homoserine-O-succinyl transferase (Röhl *et al.*, 1987).



clavulanic acid

alanylclavam

Figure 1.10 The structure of (3*R*,5*R*)-clavulanic acid, and the antibacterial (5*S*)-alanylclavam, which demonstrate the two different stereochemical arrangements of the oxazolidinone ring in clavams. Note that the C3 side chain is absent in the 5*S*-clavams, which possess antibacterial activity.

1.6.1 Mechanism of β -Lactamase Inhibition by Clavulanic Acid.

Clavulanic acid, or clavulanate, (Figure 1.10) is a weak antibiotic, but a potent inhibitor of class A, and some class D, serine β -lactamases. It acts as a pseudo-substrate, occupying the active site of the β -lactamase for sufficiently long time to prevent the degradation of co-administrated β -lactam antibiotics. At higher concentrations, *e.g.* at 115 fold excess, clavulanic acid binds irreversibly (Fisher *et al.*, 1978) to the β -lactamase through a complex mechanism which results in the formation of a stable acyl-enzyme complex (Brown *et al.*, 1996) shown schematically in figure 1.11.

The binding of clavulanic acid to SHV-1, a typical class A serine β -lactamase, is coordinated in part by the hydrogen bonding of the C7 carboxyl group of clavulanic acid to the catalytic serine (Ser70) and a main-chain nitrogen, *e.g.* Ala²³⁷, in an "oxyanion hole". This subsequently leaves the carbonyl carbon (C7) vulnerable to nucleophilic attack by Ser⁷⁰, resulting in the formation of a covalent acyl-intermediate, and the opening of the β -lactam ring. This is followed by the opening of the five-membered oxazolidinone ring, resulting in the linearization of the inhibitor, as an imine intermediate. The imine intermediate subsequently undergoes isomerization, resulting in the formation of a *cis*-enamine, which can isomerise to the more stable *trans*-enamine (Figure 1.11). After several hours the enzyme is irreversibly inhibited via the covalent acylation at Ser¹³⁰ (Chen & Herzberg, 1992; Padayatti *et al.*, 2005).

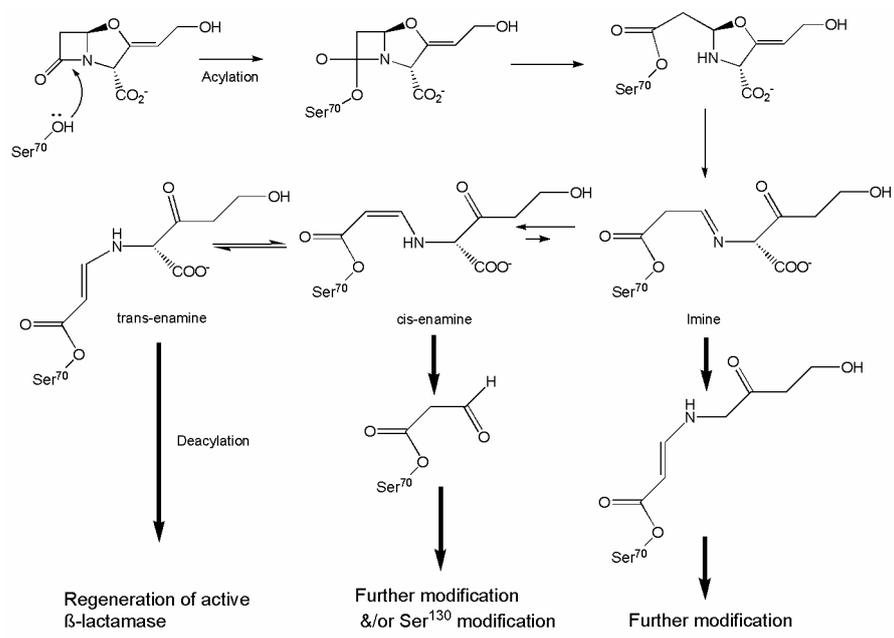


Figure 1.11 A general scheme demonstrating the complexity of the inhibition of β -lactamases by clavulanic acid, adapted from Padayatti *et al.* (Padayatti *et al.*, 2005).

2 β -Lactam Biosynthesis in *Streptomyces clavuligerus*

Take a deep breath next time you are outdoors after it has been raining. Can you “smell” the rain? That smell is created by the spores of the Actinomycetes. The Actinomycetes are a group of Gram-positive bacteria found inhabiting the soil, and are renowned for their ability to produce secondary metabolites of pharmaceutical interest. Members of the genus *Streptomyces* account for approximately 70–80% of the production of potential pharmaceutical compounds. The species *Streptomyces clavuligerus* is of particular interest, as it produces several β -lactam compounds. Initially described as a cephamycin C producing species (Higgins & Kastber, 1971; Nagarajan *et al.*, 1971), it also produces deacetylcephalosporin C, as well as different clavam compounds, including the medicinally important β -lactamase inhibitor clavulanic acid (Brown *et al.*, 1976).

The ability to synthesize clavams is confined to the genus *Streptomyces*, with the production of clavulanic acid restricted to only four species, *S. clavuligerus*, *S. jumonjinensis*, *S. katsurahamanus* and one unidentified species. *S. lipmanii* produces many of the intermediates involved in clavulanic acid biosynthesis, but does not produce the inhibitor itself. It is interesting to note that the species producing clavulanic acid also produce cephamycin C, and that none of the actinomycetes produce clavulanic acid alone (Challis & Hopwood, 2003). This situation is mirrored in the clinical application of clavulanic acid, which is co-administered with more potent antibiotics, *e.g.* Augmentin TM(amoxicillin and clavulanic acid).

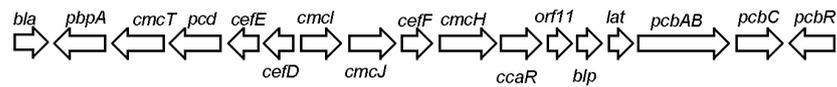
Only two *Streptomyces* genomes have been sequenced, *S. coelicolor* (A3) (Bentley *et al.*, 2002) and *S. avermitilis* (Ikeda *et al.*, 2003), although neither of these species produce clavulanic acid. The sequencing reveals a large genome, by bacterial standards (8.8 Mbp and 9.0 Mbp, respectively). The protein coding sequences (ORFs) are typified by possessing a high G + C content (>70%). Approximately 4.5% and 6% of the genes are predicted to involved in the production of secondary metabolite, where these genes are often arranged in clusters.

2.1 Arrangement and Regulation of the Cephamycin C/Clavulanic Acid Super Cluster

In *S. clavuligerus* the genes for the biosynthesis of cephamycin C and clavulanic acid are arranged in a super cluster (Figure 2.1). The cluster includes genes for the biosynthesis of the β -lactam compounds, as well as accessory genes. The latter are responsible for gene regulation within the super cluster, the export of the β -lactam compounds from the cell, as well as conferring self-resistance. Furthermore, several genes involved in the early biosynthetic steps of clavulanic acid are duplicated. A gene cluster containing a minimum of four genes homologous to *ceaS*, *bls*, *pah* and *oat* (Jensen *et al.*, 2004b; Tahlan *et al.*, 2004b) is located

elsewhere in the genome. This paralogous gene cluster is differentially regulated, and only expressed in soy growth medium (Tahlan *et al.*, 2004a), as opposed to starch asparagine medium. In addition, a homologue to *cas2* is located elsewhere in the genome (Marsh *et al.*, 1992), and is flanked by genes associated with the production of other 5S-clavams metabolites (Mosher *et al.*, 1999).

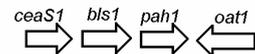
cephamycin C gene cluster



clavulanic acid gene cluster



paralogue gene cluster



clavam gene cluster



Figure 2.1 The arrangement of genes within the cephamycin C/ clavulanic acid super cluster, where *pcbR* is shown in common to both the cephamycin C and clavulanic acid gene clusters. The paralogous gene cluster of the early clavulanic acid genes is shown, as well as the *cas2* homologue flanked by the *cvm* genes important to 5S-clavam biosynthesis. Note that the cluster is not drawn to scale.

The life cycle of *Streptomyces sp.* involves a number of morphological changes, and the production of secondary metabolites, including antibiotics, are associated with these changes. The life cycle starts with the germination of a spore to produce hyphal filaments, which form an elongated and repeatedly branched network. This creates a flat colony of multinucleoid hyphae, known as the substrate mycelium. After several days, the substrate mycelium produces aerial hyphae, which stand up in the air, forming a white fuzzy layer on the mycelium. The aerial hyphae undergo extensive cellular division, eventually dividing into uni-nucleoidal compartments, which then mature into a spore.

The morphological processes of aerial hyphae production appear to be initially regulated by the *Bld* genes. Mutants of these genes are unable to produce aerial hyphae, nor spores, and are thus described as possessing a "bald" (*bld*) phenotype. Initially identified in *S. coelicolor*, several *Bld* genes have been identified in *S. clavuligerus*, including *BldA* and *BldG*. Both *bldA* and *bldG* mutants possess a bald phenotype (Bignell *et al.*, 2005; Trepanier *et al.*, 2002), however the *bldG* mutant is completely blocked in the synthesis of clavulanic acid, 5S-clavams (*e.g.*

clavam-2-carboxylate and alanylclavam) as well as cephamycin C. The *BldG* gene is thought to encode an anti-anti-sigma-factor (Bignell *et al.*, 2000), regulating the activity of a target sigma-factor. *BldG* appears to act upstream of the *CcaR* transcriptional activator of cephamycin C and clavulanic acid biosynthesis (Perez-Llarena *et al.*, 1997). *CcaR* is located in the cephamycin C gene cluster and encodes a protein with similarity to *ompR* DNA regulatory proteins, binding to the bidirectional promoter of *cefD* and *cmcI* in the cephamycin C gene cluster (Santamarta *et al.*, 2002). Furthermore, *CcaR* binds to the promoter of *Clar* in the clavulanic acid gene cluster, which in turn encodes a protein with homology to the *LysR*-type DNA regulatory proteins. *Clar* is a transcriptional activator of the genes within the clavulanic acid gene cluster including itself. Mutants of *Clar* produce penicillin/cephamycin, but do not produce clavulanic acid. The genes involved in the biosynthesis of early intermediate are expressed normally in the *clar* mutant (*ceaS*, *bls*, *pah*, *cas*), while genes upstream in the cluster are not expressed *e.g.* *orf7*, *Clar*, *cad*, *orf10* (Jensen & Paradkar, 1999). A more detailed understanding of the processes involved in the genetic regulation the β -lactam antibiotic genes in *S. clavuligerus* is still required.

2.2 Biosynthesis of Penicillin N, Cephalosporin C and Cephamycin C

Penicillin N, cephalosporin C and cephamycin C share a common biosynthetic origin. The pathway is summarized in Figure 2.2. L- α -amino adipic acid (L- α -AAA), L-cysteine, and L-valine are condensed to form δ (L- α -aminoadipyl)-L-cysteinyl-D-valine (ACV), where L-valine undergoes a configurational inversion as part of the reaction. The condensation of these amino acids is catalysed by a single multifunctional enzyme, ACV synthase, in an Mg^{2+} /ATP-dependent reaction (Byford *et al.*, 1997). ACV is then cyclised by isopenicillin N synthase (IPNS), which catalyses the formation of the β -lactam ring. IPNS is an iron (II) oxygen-dependent enzyme, catalysing the removal of four electrons and four hydrogens from ACV, to form water and isopenicillin N. Isopenicillin N serves mainly as a precursor for penicillins as well as cephalosporins and cephamycins, although it does possess weak antibacterial activity itself.

In the fungi producing only penicillin (*e.g.* *Penicillium chrysogenum*) the final reaction is catalysed by acyl coenzyme A:isopenicillin N acyltransferase (IAT) (Tobin *et al.*, 1990). An equivalent enzyme has not been identified in bacteria. The enzyme IAT catalyses the removal of the hydrophilic L- α -AAA side chain, in exchange for a hydrophobic acyl group. Since the substrate specificity of IAT is rather unspecific (Luengo, 1995) the synthesis of particular penicillins can be directed by the addition of appropriate precursor molecules during fermentation, *e.g.* phenoxyacetic or phenylacetic acid to produce penicillin V or penicillin G.

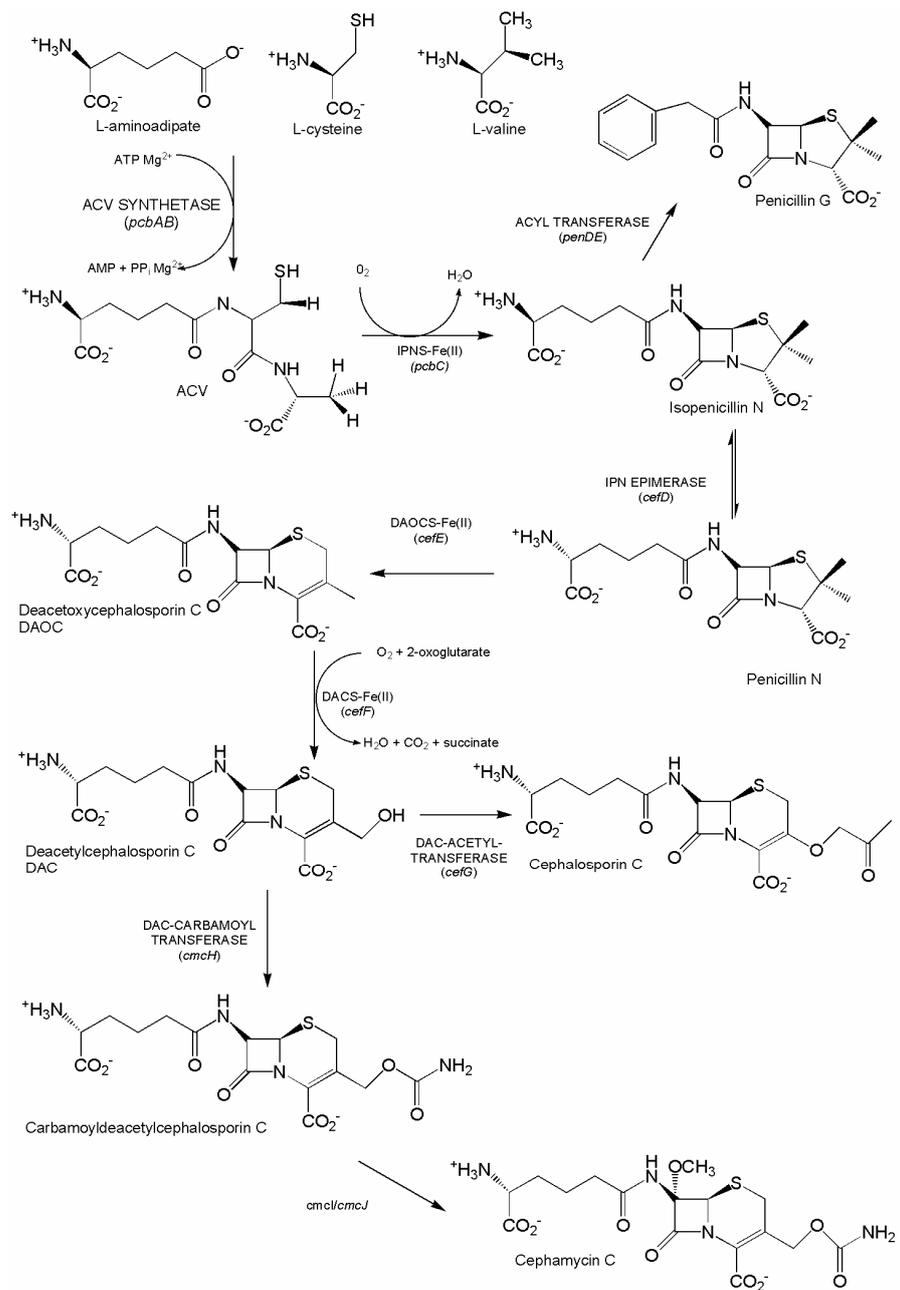


Figure 2.2 The biosynthetic pathway of penicillin N, cephalosporin C and cephamycin C in *Streptomyces clavuligerus*. The enzymes catalysing the reactions (and the genes encoding them) are shown.

Isopenicillin N serves as the branch point between penicillins and cephalosporins, where the enzyme isopenicillin epimerase is responsible for the conversion of isopenicillin N to penicillin N (Jensen *et al.*, 1983). The five membered

thiazolidine ring of penicillin N undergoes expansion to create the six-membered dihydrothiazine ring of deacetoxycephalosporin C (DAOC). DAOC is subsequently hydroxylated at the C-3' position to form deacetylcephalosporin C (DAC). In prokaryote cephalosporin producers, *e.g.* *S. clavuligerus*, these reactions are performed by two separate enzymes (Jensen *et al.*, 1985), such that deacetoxycephalosporin C synthase (DAOCS) expands the thiazolidine ring and deacetylcephalosporin C synthase (DACS) hydroxylates the C-3' position. In eukaryotic cephalosporin producers, *e.g.* *Cephalosporium acremonium*, both steps of the reaction are catalysed by the bi-functional enzyme deacetoxy/deacetylcephalosporin synthase (DAOCS/DACS). The enzymes catalysing the conversion of penicillin N to DAC, *i.e.* DAOCS/DACS, DAOCS and DACS, are all Fe(II), 2-oxoglutarate-dependent dioxygenases (2ODDs).

The final step of cephalosporin C synthesis is assisted by acetyl-CoA:DAC acetyltransferase (Brakhage, 1998), which catalyses the transfer of an acetyl moiety from acetyl-CoA to the hydroxyl group of DAC. DAC represents the second branch point in the synthesis of β -lactam antibiotics in *S. clavuligerus*, dividing the cephalosporin C and cephamycin C pathways. To date cephamycin C has only been observed in bacteria.

The C-3 acetoxy group of DAC is exchanged for a carbamoyl group, to form carbamoyldeacetylcephalosporin C. The reaction is performed by DAC-carbamoyl transferase, and requires carbomoyl phosphate, ATP and Mn^{2+} and Mg^{2+} ions (Brewer *et al.*, 1980; Coque *et al.*, 1995). The final step of cephamycin C biosynthesis requires the methylation of the C7 carbon of the β -lactam ring. The gene products of *CmcI* and *CmcJ* are responsible for these final steps, but the mechanism by which these enzymes methylate their substrate is yet to be fully elucidated.

A number of the enzymes involved in penicillin, cephalosporin C and cephamycin C biosynthesis have been structurally determined. The structure of IPNS from *Aspergillus nidulans* was the first to be solved (Roach *et al.*, 1995), resulting in a detailed knowledge of the reaction mechanism (Burzlaff *et al.*, 1999; Roach *et al.*, 1997). The structure of DAOCS from *S. clavuligerus* represents the first 2ODD to be solved (Valegård *et al.*, 1998), and is discussed further in Chapter 3. More recently the structure of *CmcI* (Öster *et al.*, 2006) and DAC-AT (S. Lejon, *Per. Comm.*) have been solved in the β -lactam antibiotic group in Uppsala.

2.3 The Biosynthesis of Clavulanic Acid

The total synthesis of clavulanic acid is hampered by its dense functionality and lability (Baggaley *et al.*, 1997; Townsend, 2002), and therefore commercial production is by fermentation in submerged cultures of *S. clavuligerus*. As a result, the regulation of the genes involved in its synthesis, and characterization of the proteins they encode, has been studied extensively for the past 30 years. Approximately eighteen genes (*orf2-19*) appear to be responsible for the biosynthesis, transport, resistance to, and regulation of, clavulanic acid (Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Li *et al.*, 2000; Mellado *et al.*, 2002), although a

defined function has been demonstrated for only half the genes in the cluster (see table 2.1). The remaining genes have a function tentatively assigned on the basis of sequence similarity with proteins of known function.

Table 2.1 A Summary of genes and their function in the biosynthesis of clavulanic acid. Proteins of known function are in normal text, and in bold text if structurally determined. Proteins for which no specific function in clavulanic acid biosynthesis has been demonstrated are in italics.

Gene	Encodes/putative function
<i>Orf2</i>	N²-(2-carboxyethyl)arginine synthase (CEAS2)
<i>Orf3</i>	β-lactam synthase (BLS2)
<i>Orf4</i>	proclavamate amidinohydrolase (PAH2)
<i>Orf5</i>	clavaminic acid synthetase (CAS2)
<i>Orf6</i>	ornithine acetyltransferase (OAT2)
<i>Orf7</i>	<i>peptide binding protein</i>
<i>Orf8</i>	ClaR regulatory protein (DNA binding protein)
<i>Orf9</i>	clavulanic acid dehydrogenase (CAD)
<i>Orf10</i>	<i>P450 mono-oxygenase</i>
<i>Orf11</i>	<i>ferredoxin</i>
<i>Orf12</i>	acetyl transferase
<i>Orf13</i>	<i>efflux pump</i>
<i>Orf14</i>	<i>acetyltransferase</i>
<i>Orf15</i>	peptide binding protein
<i>Orf16</i>	<i>hypothetical protein</i>
<i>Orf17</i>	N-glycyl-clavaminic acid synthetase (GCAS)
<i>Orf18</i>	PBP
<i>Orf19</i>	PBP

Clavulanic acid and the 5*S*-clavams share a common biosynthetic precursor, clavaminic acid (Egan *et al.*, 1997). The steps leading to the biosynthesis of clavaminic acid have been characterized (Figure 2.3) and the enzymes catalyzing them identified. In the first step L-arginine and D-glyceraldehyde-3-phosphate are condensed to form N²-(2-carboxyethyl)arginine in a thiamine diphosphate-dependent reaction, which is catalysed by the enzyme N²-(2-carboxyethyl)arginine synthase (CEAS). This is followed by the formation of the β-lactam ring, by β-lactam synthase (BLS), to yield deoxyguanidinoproclavaminic acid. The mechanism by which BLS forms the β-lactam ring is quite different to that in penicillin/cephalosporin biosynthesis. Infact BLS shows more sequence and structural homology to class B asparagine synthetases (Miller *et al.*, 2001) than it does to IPNS. Deoxyguanidinoproclavaminic acid is subsequently converted in four steps to 3*S*,5*S*-clavaminic acid. Three of these steps are catalysed by the 2ODD clavaminic acid synthase (CAS2). The first step catalysed by CAS2 involves the hydroxylation to give a secondary alcohol (Baldwin *et al.*, 1993), which is neither an inhibitor, nor a substrate of CAS2. Proclavamate amidinohydrolase (PAH) hydrolyses the guanidino group to yield proclavaminic acid, the second substrate of CAS. In a two-step reaction CAS catalyses the formation of the bicyclic clavam ring, resulting in the formation of clavaminic acid.

The crystal structures of the first four enzymes of the pathway have been reported, *i.e.* CEAS (Caines *et al.*, 2004), BLS (Miller *et al.*, 2001), PAH (Elkins *et al.*, 2002), CAS2 (Zhang *et al.*, 2000). An additional structure, ornithine acetyl transferase (OAT2) has also been determined (Elkins *et al.*, 2005). While not directly involved in the synthesis of clavulanic acid, this enzyme is involved in the biosynthesis of L-arginine from ornithine, which is required for clavulanic acid synthesis. It has been demonstrated that tritium-labelled ornithine is incorporated into clavulanate (Townsend *et al.*, 1986).

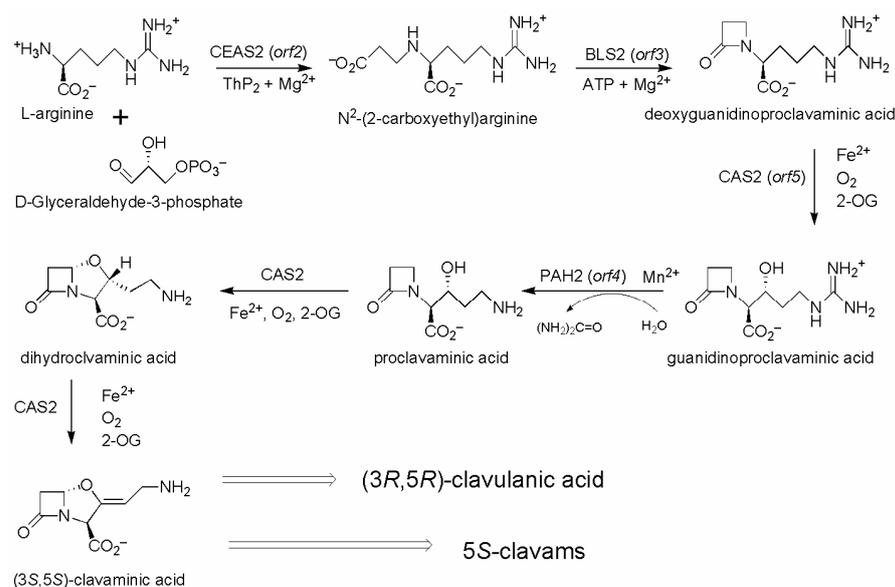


Figure 2.3 The biosynthesis of (3S,5S)-clavaminic acid from L-arginine and D-glyceraldehyde-3-phosphate.

Clavaminic acid serves as the branch point between the synthesis of the 5S-clavams and clavulanate (Figure 2.4). The biosynthesis of clavulanate from clavaminic acid requires a double epimerisation and oxidative deamination to yield (3R,5R)-clavulanate-9-aldehyde (clavaldehyde). While this is thought to be performed by the gene products of *orf10* and *orf11*, there is little mechanistic information on how this is achieved at present. Recently it was shown that *orf17* encodes a N-glycyl-clavaminic acid synthetase (Arulanantham *et al.*, 2006), capable of converting clavaminic acid to N-glycyl-clavaminic acid in an ATP dependent reaction. Whether N-glycyl-clavaminic acid represents an additional step in the synthesis of clavulanic acid remains unclear at present. The final step in clavulanic acid biosynthesis results in the reduction of clavulanate-9-aldehyde (clavaldehyde) catalysed by clavulanic acid dehydrogenase (CAD) (Fulston *et al.*, 2001; Nicholson *et al.*, 1994)). The structure of CAD, and the reaction catalysed by it are discussed in further detail in Chapter 4.

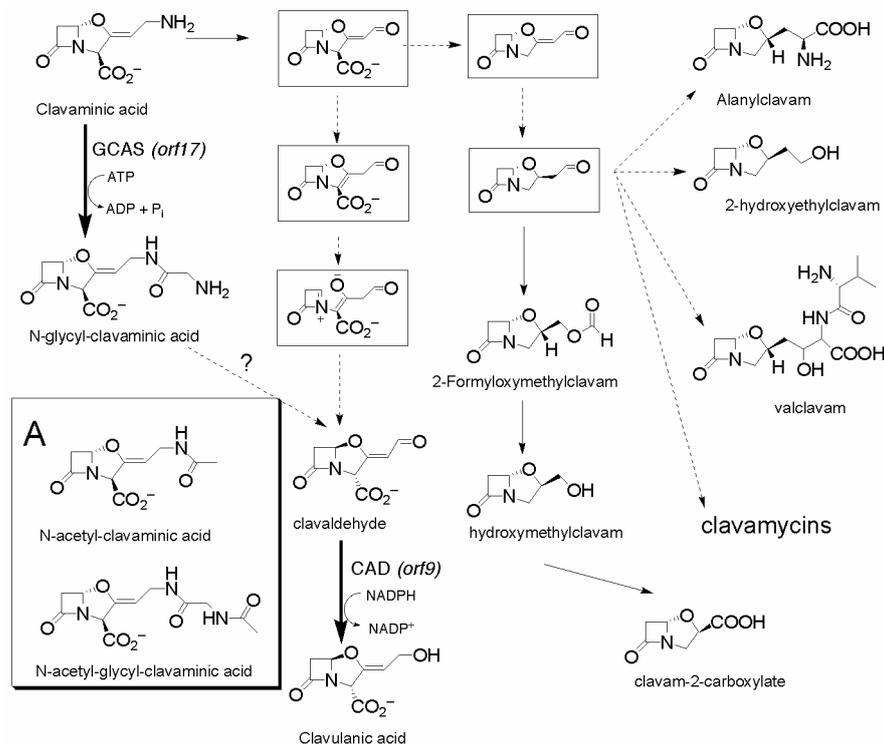


Figure 2.4 Potential biosynthetic pathway from (3*S*,5*S*)-clavaminic acid to both the 5*S*-clavams and (3*R*,5*R*)-clavulanic acid (Baldwin *et al.*, 1994; Liras & Rodriguez-Garcia, 2000). Proposed intermediates are in boxes. The N-acetyl derivatives of clavaminic acid accumulating in the *del8* mutant of *S. clavuligerus* (Elson *et al.*, 1998) are shown in Box A.

Recently, the gene products of *orf18* and *orf19* were shown to encode the high molecular weight membrane bound PBPs PbpA and Pbp2, respectively (Ishida *et al.*, 2006). Both PBPs showed a reduced affinity for β -lactam antibiotics and clavulanic acid, where Pbp2 possessed the lowest affinity for clavulanic acid between the two PBPs. The synthesis of PBPs resistant to β -lactam antibiotics and clavulanic acid affords a means by which *S. clavuligerus* is able to achieve self-resistance.

The exact role of the remaining nine ORFs (*orf12–16*, *orf7*) in clavulanic acid biosynthesis remains unclear at present. Knockout disruption-mutant studies suggested that *orf7*, *orf10*, *orf12*, *orf15*, *orf16* and *orf17* were essential for the synthesis of clavulanic acid. Mutants of *orf11*, and *orf13* were severely compromised, but did produce detectable levels of products (Jensen *et al.*, 2004a). The reported effects of disrupting *orf14* vary from partial loss (Mellado *et al.*, 2002) to almost complete loss of clavulanic acid production (Jensen *et al.*, 2004a).

The disruption mutations also had an effect on the production of 5*S*-clavams, among these 2-hydroxymethylclavam, clavam-2-carboxylate and alanylclavam. The production of 5*S*-clavams are typically more variable, even in wild-type strains

(Thai *et al.*, 2001), however the disruption of *orf11* and *orf13* resulted in a 50-90% and 90-100% reduction of C-2-C and 2HMC, respectively. The levels of 5S-clavams in the other mutants varied, although detectable levels were produced (Jensen *et al.*, 2004a). In addition the knockout-disruption of *orf15* or *orf16* resulted in the production of N-acetyl-glycyl-clavaminic acid (NAG-clavam) and N-glycyl-clavaminic acid, two 5S-clavam previously detected in the *dcl8* mutant (Elsom *et al.*, 1998).

The role of the remaining ORFs is further complicated by the observation that the gene disruption of several ORFs (*orf10*, *orf12*, and *orf15*) resulted in the synthesis of the non- β -lactam antibiotic holomycin (de la Fuente *et al.*, 2002; Lorenzana *et al.*, 2004).

2.4 Aims of the Thesis

In order to further study the enzymes involved in β -lactam antibiotic production we have used a structural approach, through the technique of X-ray crystallography. The shape of a protein determines, in part, its function. Thus an understanding of the three-dimensional structure of a protein can lead to detailed knowledge at the molecular level of how a particular protein performs its role.

The genes involved in the biosynthesis of penicillin N, cephalosporin C, and cephamycin C have been identified and characterised. However, the mechanisms by which many of the enzymes catalyse their reactions remains to be determined. For example, several of the genes involved encode non-heme iron oxygenases *e.g.* IPNS; DAOCS, DACS, and DAOC/DAC, most of which are 2ODDs. These enzymes are thought to perform catalysis using a highly oxidising iron(IV) species, however this ferryl intermediate has only been observed *via* spectroscopic methods (Proshlyakov *et al.*, 2004).

The number of genes known to be involved in the biosynthesis of clavulanic acid has increased dramatically in the last few years, initially six ORFs, then nine ORFs, and now eighteen ORFs. Approximately half of the functional assignments are based on sequence similarity to other proteins rather than a known function. The aim of this thesis is to further examine the mechanism of action of enzymes and proteins involved in the synthesis of β -lactam antibiotics in *S. clavuligerus*.

The story of penicillin provides a wonderful example of how basic research lead to the serendipitous discovery of a compound that ultimately changed human history. However, the subsequent research and development of penicillin represented a concerted scientific effort to understand not only how penicillin works, but how it could be improved. The research contained within this thesis does not contain serendipitous discoveries that will be as herald as penicillin. However, this thesis contributes to the research and knowledge required to understand how β -lactam compounds are synthesised in nature's "chemical kitchen". Ultimately this boils down to a simple desideratum; new β -lactam antibiotics with which to fight our battle against pathogenic bacteria. Through a greater understanding of the underlying processes we can improve the yields using the existing techniques, and

hopefully in the future identify more environmentally friendly and inexpensive synthetic routes to new more powerful antibiotics.

3 Deacetoxycephalosporin C Synthase, DAOCS (Paper I)

3.1 Background

The gene *CefE* in the cephamycin gene cluster of *S. clavuligerus* encodes for the enzyme deacetoxycephalosporin C synthase (DAOCS). DAOCS is responsible for the expansion of the five-membered thiazolidine ring of penicillin N to the six-membered dihydrothiazine ring of deacetoxycephalosporin (DAOC), illustrated in figure 3.1. DAOCS is also capable of expanding the thiazolidine ring of various other penicillins, albeit with a reduced efficiency (Lloyd *et al.*, 1999).

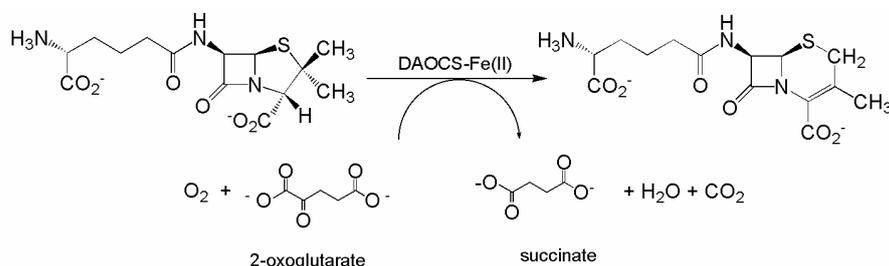


Figure 3.1 The ring expansion reaction catalysed by DAOCS.

The enzyme belongs to a class of enzymes known as 2-oxoglutarate dependent dioxygenases (2ODDs), which require a non-heme iron, 2-oxoglutarate and dioxygen (Prescott, 1993). DAOCS was the first 2ODD structure to be determined (Valegård *et al.*, 1998), and was found to possess a fold similar to that of IPNS (Roach *et al.*, 1995). The structure of DAOCS consisted of a distorted jelly roll fold, flanked by α -helices, where Fe(II) is coordinated by two histidines and an aspartic acid. Since then several other 2ODD structures have been determined, *e.g.* CAS (Zhang *et al.*, 2000), CAR-C (Clifton *et al.*, 2003), ACCO (Zhang *et al.*, 2004), proline-3-hydroxylase (Clifton *et al.*, 2001). Despite low sequence similarity within the 2ODD family, their structures reveal a common fold, *i.e.* the distorted jelly roll. The 2ODD structures also reveal the similar nature of Fe(II) coordination, where two histidines and one aspartic acid or glutamic acid create the 2-His-carboxylate facial triad motif (Hegg & Que, 1997).

In order to bind/react with oxygen, Fe(II) must first be activated. 2-Oxoglutarate binds to the iron and activates it for the binding of oxygen. In the process of this activation 2-oxoglutarate is decarboxylated to succinate, generating CO₂ and an oxidising intermediate, *i.e.* a ferryl-species Fe(IV). Succinate is ligated to the iron, and most probably serves to stabilise the reactive intermediate prior to the binding of penicillin. In order for penicillin to bind, succinate must first be expelled from the active site. This triggers oxidative attack on penicillin, ultimately leading to the expansion of the penam ring, resulting in the formation of cephalosporin and water in the active site. The high resolution structure of DAOCS led to the

proposal of a mechanism for the formation of the ferryl intermediate (Valegård *et al.*, 1998), however the later catalytic steps, resulting in the formation of the cephalosporin, are yet to be fully elucidated. These latter steps are of particular interest since the identification of the residues involved in binding and catalysis would be invaluable for future genetic redesigning of DAOCS.

Further insight into the catalytic reaction performed by DAOCS could be gleaned by the capture of the ferryl-intermediate using protein crystallography. The ferryl-intermediate is most likely reduced by the X-ray beam during data collection, preventing it from being seen in the crystal structure. By using the techniques of time resolved crystallography, in a manner similar to that used by Berglund *et al.*, (2002), it could be possible to capture the ferryl intermediate in short time slices, before the high oxidation state is reduced. This would hopefully lead to an enhanced understanding of the reaction mechanism.

3.2 An Alternate Crystal Packing of DAOCS is Required

The crystallization of native DAOCS protein resulted in the formation of crystals belonging to the space group *R3* (Valegård *et al.*, 1998). Structure determination using the *R3* crystals was, however, hampered by merohedral twinning (Terwisscha van Scheltinga *et al.*, 2001). Merohedral twinning is a packing anomaly, where distinct domains of the protein can pack in alternate orientations within the three dimensions of the crystal. The degree of twinning can range from no twinning (twin fraction 0) to almost perfect twinning (twin fraction 0.5). The twin fraction of the data can be estimated, allowing the data to be “detwinned”. However, such methods amplify experimental errors, where the experimental error introduced by detwining is proportional to the twin fraction. To capture the DAOCS ferryl-intermediate using protein crystallography would require the collection of data on crystals with a twin fraction of less than 0.3.

The structure of the apo enzyme revealed that the monomers are arranged in a trimer (Figure 3.2), where the carboxyl-terminal arm (residues 308-311) binds in the active site of a neighbouring molecule. If the packing of DAOCS within the crystal could be significantly altered, it may be possible to overcome the technical difficulties associated with merohedral twinning. Crystallization conditions had been extensively screened for conditions conducive to an alternate crystal form, however none were identified. Thus, modifications to the enzyme preventing the formation of the *R3* trimer were required. This could be achieved by the deletion of the C-terminal arm, since it is involved in trimer formation. However the C-terminal arm is important for enzymatic activity, as demonstrated by mutation and truncation studies (Chin & Sim, 2002; Lee *et al.*, 2001; Lee *et al.*, 2002). Instead, DAOCS was expressed with an N-terminal His₍₆₎-tag, where the resulting fusion protein produced crystals belonging to an alternate space group.

3.3 Sub-cloning, and Initial Expression and Purification Trials

The *daocs/cefE* gene from *S. clavuligerus* was amplified by PCR from the pML1 plasmid (Lloyd *et al.*, 1999) using the PCR primers: forward primer 5' – CAT GGC AGC CAT ATG GAC ACG ACG GTG CCC ACC TTC AG – 3' ; reverse primer 5'–AGC AGC CGG ATC CTC CTA TGC CTT GGA TGT GCG GCG GAT G – 3'. The PCR product was cloned into the pET15b expression vector using the *Nde* I and *Bam*H I restriction sites. Positive clones of the *daocs/pET15b* construct were selected on LB-agar plates using ampicillin (100 µg ml⁻¹) as a selection marker, and confirmed with restriction enzyme digests and directional PCR. The *daocs/pET15b* construct was then transformed into *E. coli* BL21 (DE3) cells. Initial expression trials were conducted at 30 °C and 37 °C using 100 ml *E. coli* cultures grown in LB medium containing ampicillin. Recombinant protein expression was induced with 0.5 mM IPTG at an OD_{600nm} = 1.00. The reduction of growth temperature to 30 °C resulted in increased expression of soluble recombinant protein (data not shown).

The pET15 expression vector contained an N-terminal poly His₍₆₎-tag and a thrombin protease site, thus the protein could be separated from the lysed bacterial culture using immobilised-metal-affinity-chromatography (IMAC). The initial purification trial used a 1 L culture of *E. coli* grown at 30° C. Protein expression was induced as described above, and cells harvested 2 hours after induction by centrifugation at 5,000 g for 15 minutes at 4 °C. The bacterial pellet was resuspended in lysis buffer consisting of 50 mM NaH₂PO₄ (pH 8.0), 500 mM NaCl, and 10 mM imidazole. Total cells were disrupted by sonication and the lysate cleared by centrifugation at 12,000 g for 30 minutes at 4 °C. The cleared lysate was loaded onto a Ni-NTA column and subsequently eluted using a imadazole gradient of 10–300 mM imidazole. Fractions containing protein were analysed by SDS-PAGE, with appropriate fractions pooled and concentrated. The concentrated protein was subjected to size exclusion chromatography using a Superdex75 gel filtration column, before being concentrated once again. The resulting protein was deemed very pure (>95%) as judged by SDS.

The purification of the His₍₆₎-DAOCS protein was more rapid in comparison to the procedures for the native protein (Lloyd *et al.*, 1999). This is important when considering that mononuclear ferrous enzymes are sensitive to oxygen, degrading under aerobic conditions (Barlow *et al.*, 1997).

3.4 A Summary of Subsequent Results

(The subsequent purification, crystallization, and structural determination of the His₍₆₎-DAOCS protein was performed by L. Öster (Paper I). The effect of the His₍₆₎-tag on crystallization, and crystal packing will be discussed here. The results from the complexes of *DAOCS-Fe(II)-2-oxoglutarate*, *DAOCS-Fe(II)-succinate*, *DAOCS-Fe(II)-ampicillin*, and *DAOCS-Fe(II)-2-cephalosporin C* will not.)

3.4.1 Effect of His₆-Tag on Crystallization

The His-tagged DAOCS crystals were obtained in conditions similar to the native protein, except that the concentration of ammonium sulphate was reduced from 1.75 M to 0.9–1.1 M and the pH was reduced from 7.0–7.5 to 8.0. The optimum temperature for crystallization was dramatically altered from 20 °C for the native protein, to 4 °C for the His-tagged protein. Furthermore, the addition of 2-oxoglutarate was no longer required to obtain crystals of the new His-tagged protein. Previously the technique of seeding was essential to obtaining native crystals, however crystals of the His-tagged protein grew spontaneously. In fact, it was necessary to slow the rate of vapour diffusion using Al's oils (Hampton), reducing the formation of microcrystals.

3.4.2 Effect of His₆-Tag on Crystal Packing

The resulting crystals of His-tagged protein were hexagonal rods, belonging to the space group $P3_121$, which differs to the previous rhombohedrally shaped crystals belonging to space group $R3$. Most importantly the new (His-tagged) crystals did not show merohedral twinning. This is due to a considerable difference in the packing of molecules within the crystal (see Figure 3.2). Although the molecules of DAOCS are still arranged as trimers in the new crystal form, the symmetry of the $P3_121$ trimer is quite different to that of the $R3$ trimer. This difference is most likely due to disruption of the N-terminal interaction between adjacent molecules seen in the $R3$ trimer. As a result of the new trimer symmetry the C-terminal arm no longer binds within the active site. Thus the N- and C-terminal residues are no longer involved in any intermolecular contacts in the new crystal form.

Unfortunately the packing of molecules within the new crystal form created large solvent channels, increasing the solvent content of the crystals to *ca.* 70%, much higher than the previous *ca.* 40% ($R3$ crystals). This resulted in fragile, and difficult to handle, crystals. Such a high solvent content would also allow for increased flexibility in the crystal lattice, where this may partially account for the poorer diffraction quality of the new crystals (2.3 Å compared to 1.3 Å).

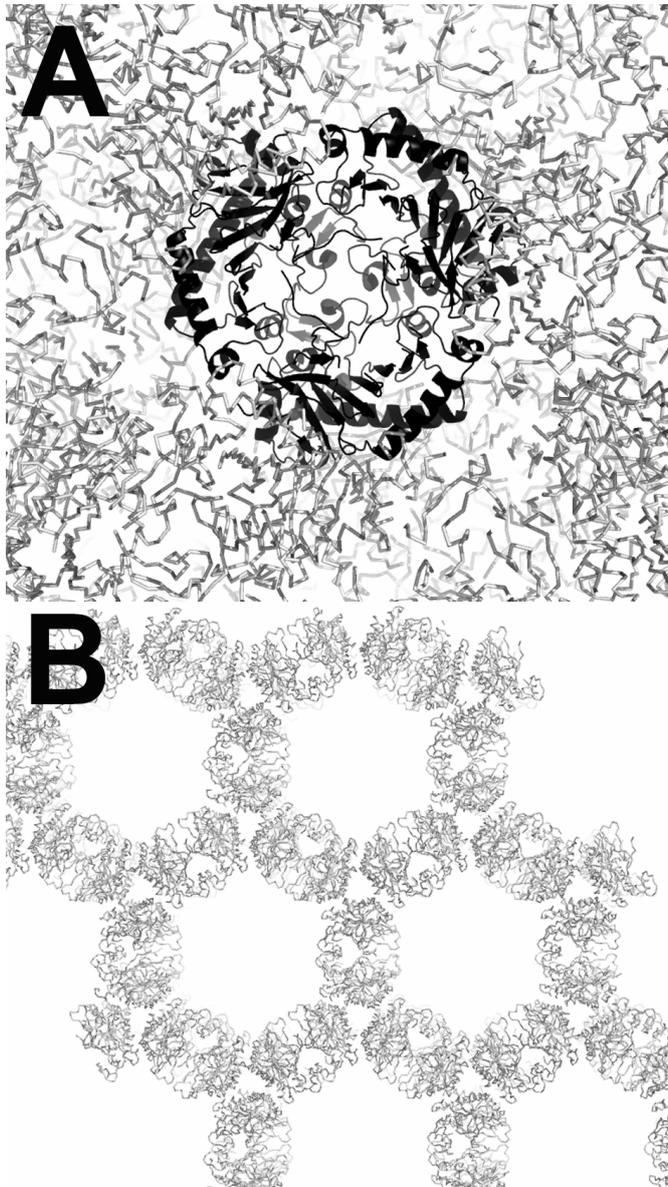


Figure 3.2 Packing of molecules within the two different crystal forms. In both cases DAOCS forms trimers, although the interactions forming these trimers are mediated by different regions of the protein. (A) The $R3$ crystal form as viewed down the three-fold symmetry axis, note the dense packing of molecules contributing to a low solvent content. (B) The $P3_121$ crystal form, viewed down the three-fold axis, note the large solvent channels formed between the trimers.

4 Clavulanic Acid Dehydrogenase, CAD (Paper II)

4.1 Background

The clavam 3*R*,5*R*-clavulanate-9-aldehyde (clavaldehyde) is a potent inhibitor of β -lactamases, unfortunately the electron drawing properties of the α/β -unsaturated allylic aldehyde make it an unstable molecule. It is susceptible to hydrolysis in solution, with a half-life of approximately one hour (Fulston *et al.*, 2001; Nicholson *et al.*, 1994), resulting in fragmentation of the bicyclic ring (Baggaley *et al.*, 1997). The reduction of the C9-aldehyde to an alcohol increases the stability of the β -lactamase inhibitor in solution, and thus clavaldehyde represents the final intermediate in the biosynthesis of clavulanic acid.

The penultimate step (Figure 4.1) is catalyzed by clavulanic acid dehydrogenase (CAD), also known as clavulanate-9-aldehyde reductase (CAR) (Perez-Redondo *et al.*, 1998). The reaction catalyzed by CAD is dependent upon the nucleotide co-factor NADPH (Nicholson *et al.*, 1994, Fulston *et al.*, 2001).

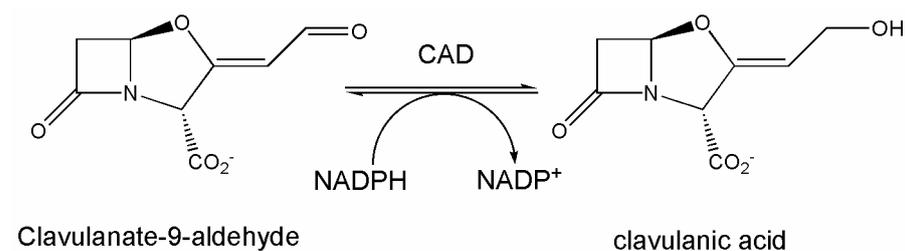


Figure 4.1 The CAD catalysed reduction of clavulanate-9-aldehyde to clavulanic acid using the cofactor NADPH.

The enzyme CAD is encoded by *orf9* in the clavam gene cluster, and is expressed as a monostrophic transcript (Perez-Redondo *et al.*, 1998), under the regulation of the ClaR promoter. To date no homolog has been observed elsewhere in the genome. Sequence analysis of CAD reveals similarity to the Short Chain Dehydrogenase/Reductase (SDR) family (Nicholson *et al.*, 1994). Over 3000 SDR members have been identified, with diverse substrates, ranging from alcohols and sugars to steroids and aromatic compounds. SDRs are formed by subunits of approximately 250 amino acids. The sequence identity within the SDR family is typically low, in the order of 15-30%. There are very few highly conserved residues characterizing the family, although a number of important motifs can be identified. Despite the low identity within the family, the SDRs show a conserved structural motif ($\beta/\alpha/\beta$), *i.e.* the Rossmann fold, which is common for enzymes binding nucleotide coenzymes (Rossmann *et al.*, 1975).

An important question in clavulanic acid biosynthesis is how the highly labile intermediates are efficiently processed through the pathway. Of particular interest is how CAD can catalyse the reduction of the unstable clavuldehyde without fragmentation of the bicyclic ring. In order to further investigate the nature of the CAD catalysed reaction, the structure of CAD was determined, along with CAD in complex with clavulanic acid.

4.2 Expression, Purification, and Crystallization of CAD

4.2.1 Initial Trials Using Recombinant Fusion Protein

Cad/orf9 was amplified by PCR from a genomic DNA preparation from *S. clavuligerus* (courtesy of M. Svenda) and cloned into a pET30 (N-terminal His₆-Tag) and pET32 (Trx-His₆-Tag) vector, using ligation independent cloning (Novagen). The *cad/pET* constructs were transformed into *E. coli* pLysS (DE3), and grown in LB medium at 18 °C using appropriate selection markers. Recombinant protein expression was induced at a cell density of OD₆₀₀ = 0.5, by the addition of 0.5 mM isopropyl IPTG, and the temperature raised to 27 °C. Cells were harvested by centrifugation five to six hours after induction, and stored at -80 °C.

The recombinant protein was purified using a combination of affinity chromatography (Ni-NTA), anion exchange chromatography (mono-Q) and size exclusion chromatography (Sephadex 200) all at pH 8.0. Fractions were checked for purity using SDS-PAGE. Purified protein from the *cad/His₆* construct was light brown in colour, precipitated readily upon concentration, and was partially truncated when visualized on SDS-PAGE (*ca.* 26 kDa.). No further work continued using this construct. The *cad/Trx-His₆* construct (*ca.* 54 kDa.) resulted in the production of considerable quantities of super soluble protein, *ca.* 80 – 100 mg ml⁻¹. Despite numerous crystallization trials, conditions suitable to crystallization could not be identified. Removal of the thioredoxin/His₆ tag using the protease Factor Xa resulted in the production of native CAD, which when concentrated (*ca.* 10 mg ml⁻¹), resulted in the formation of crystals overnight (Crystal screen I, Hampton Ltd). The crystals appeared as small overlapping needles, fused together, and could not be successfully separated. The resulting diffraction pattern could not be processed, due to very high mosaicity and split reflections.

4.2.2 Initial Trials Using Native Protein

In order to avoid cleavage of the Trx/His₆ fusion protein further work continued using the native *cad/pET24(+)* (*cad/Native*) construct obtained from Oxford (courtesy of N. Kershaw, Oxford). The construct was transformed into *E. coli* BL21 (DE3) and the cells grown in LB media as previously described. The initial purification protocol consisted of anion and cation exchange chromatography, using mono-Q (Tris-HCl pH 8.0) and mono-S (MES pH 5.5). The changing of pH by dialysis resulted in partial precipitation of CAD (pI 6.5), although significant quantities of protein remained soluble, allowing continued purification using gel filtration (Tris-HCl pH 8.0). The purified protein was over 95% pure, as judged by SDS-PAGE and was concentrated to *ca.* 10 mg ml⁻¹.

Small crystals were obtained using Clear Screen I (Molecular Dimensions) with cacodylate (pH 6.5) as the buffering agent. The inclusion of 2 mM NADPH improved the visual appearance of the crystals, and seeding improved crystal size. A needle-shaped crystal was prized apart, and frozen in liquid nitrogen using glycerol as a cryo-protectant. A complete data set to 1.8 Å was collected on a crystal at the ESRF (Grenoble, France). The crystals belonged to space group $P2_1$, with unit cell dimensions of $a = 91.8 \text{ \AA}$, $b = 87.5 \text{ \AA}$, $c = 91.9 \text{ \AA}$, $\beta = 119^\circ$. The Matthews coefficient indicated an estimated four, or six, molecules in the asymmetric unit. Despite extensive attempts, the structure could not be solved using molecular replacement. An alternate protocol devised by colleagues in Oxford resulted in the formation of crystals belonging to an orthorhombic space group. Crystals obtained from either of the two purification protocols were unsuitable for phase determination by the method of Multiple Isomorphous Replacement (MIR) due to a lack of isomorphism between the crystals.

4.2.3 Expression, Purification, and Crystallization of SeMet Protein

The *cad*/Native construct was transformed into *E. coli* 834 (DE3). Cells were grown in SeMet medium, using L-seleno-methionine (Molecular Dimensions), and the same temperature regime and procedures as previously described. The recombinant protein was purified as described above, except that 5 mM β -mercaptoethanol was added to all the solutions. It is interesting to note that the SeMet-labelled CAD did not bind to the Mono-Q column, possibly as a result of an altered isoelectric point due to SeMet incorporation. Instead, an equal volume of 2M $(\text{NH}_4)_2\text{SO}_4$ was added to the flow-through, and purification continued as usual. The purified protein was concentrated (*ca.* 8 mg ml⁻¹) and stored in 5 mM HEPES (pH 7.5), 1 mM EDTA and 5 mM β -mercaptoethanol.

Previous CAD crystals had been obtained using cacodylate (pH 6.5) as a buffering agent. However, cacodylate contains arsenic, which gives an anomalous signal close to the selenium absorption edge, thus to reduce potential problems with data collection the crystallization buffer was changed to Bis-Tris (pH 6.5). Small crystals were obtained using equal volumes of a reservoir solution consisting of 10 % (w/v) PEG 1000, 10 % (w/v) PEG 8000, and 0.3 M sodium acetate, and protein (*ca.* 6 mg ml⁻¹). The small crystals were crushed, and used for streak seeding into a drop consisting of 2 μ l of protein (*ca.* 3 mg ml⁻¹ + 2 mM NADPH) and 2 μ l of the same reservoir solution. This resulted in the growth of crystals within 1 to 3 days at 22 °C. The crystals were prepared for data collection at cryogenic temperatures by transfer to a cryo-solution consisting of the reservoir solution containing 25% PEG 400, followed by submersion in liquid nitrogen.

4.3 Structure Determination and Model Building

4.3.1 Location of the Selenium Atom Sites

A single-wavelength anomalous diffraction (SAD) data set was collected on a SeMet crystal at a fixed wavelength beam line ID14-3 (ESRF, Grenoble, France). The crystal diffracted to 1.8 Å, and belonged to an orthorhombic space group ($P222$), with cell dimensions $a = 58$ Å, $b = 122$ Å, $c = 126$ Å. The dimensions of the a -axis were initially misinterpreted as being 116 Å, due to the presence of a 'ghost' diffraction pattern. The space group could not be confidently assigned, as the intensities along the h axis were incomplete, *i.e.* the $k + l = 0$ systematic absences had not been measured. The data were initially processed to 1.9 Å (360°) using DENZO and SCALEPACK (Otwinowski & Minor, 1997) in the space groups $P2_12_12$ due to isomorphism with a previous Pt cyanide MAD data set. Assuming the presence of four molecules in the asymmetric unit gave a V_M of 2.2 Å³ Da⁻¹ (Matthews, 1968) with 42.7% solvent. The protein sequence for CAD predicted seven methionines, although ESI mass spectrometry and Edman sequencing confirmed the loss of the N-terminal methionine (Paper II). The SHELX suit (Schneider & Sheldrick, 2002) was used to locate the 24 selenium sites. Initial SHELX trials indicated sufficient anomalous signal, but the 20 heavy atom sites located were of dubious quality, due to low occupancy. The reprocessing of the data in the space group $P2_12_12_1$, and the re-running of SHELX using the options "NO PATS" and "WEED" resulted in improved heavy atom site occupancies, with a clear difference in hand after solvent flattening. However, a model could not be built into the electron density. Attempts to identify the non-crystallographic symmetry (NCS) arrangement of the selenium sites in O (Jones *et al.*, 1991), revealed the presence of an almost pure translation vector, confirmed by FFT (CCP4, 1994). This accounted for the improvement of heavy atom sites using the "NO PATS" option in SHELX. The refinement of the heavy atom sites using Auto-Sharp (La Fortelle & Bricogne, 1997) resulted in the identification of three more selenium sites, and solvent flattening with SOLOMON (Abrahams & Leslie, 1996) showed a clear distinction between the solvent and the four molecules of the asymmetric unit. The running of Auto-Sharp using "auto-build" (ARP-wARP (Perrakis *et al.*, 1999)) resulted in the building of several β -strands, arranged in manner similar to a related SDR, Rv2002 from *Mycobacterium tuberculosis* (pdb code 1NFF) (Yang *et al.*, 2003).

4.3.2 Model Building

The monomer of the related SDR from *Mycobacterium* was used as a search model to locate the four molecules in the electron density maps from Auto-Sharp using MolRep (Vagin & Teplyakov, 1997). Rigid body refinement of the phased molecular replacement solution showed the clear presence of the co-substrate, NADPH, in the electron density. A poly-alanine model was prepared, and further building continued using ARP-wARP. A mask was prepared using MAMA and the NCS symmetry operators determined in O (Jones *et al.*, 1991), allowing electron density averaging using the RAVE suite of programs (Kleywegt & Jones, 1999; Kleywegt & Read, 1997). The averaged maps, in combination with the heavy atom sites, allowed manual docking of the sequence to the partial model. Further

building cycled between ARP-wARP and RAVE, with manual rebuilding of the model using O. This resulted in a complete model from Ser³ to Val²⁴⁷, where additional electron density for Pro² could be seen in one of the molecules. No electron density was observed for the Met¹, consistent with the EPI-MS and Edman sequencing data.

4.3.3 Model Refinement

The data were re-processed to 1.8 Å using MOSFLM (CCP4, 1994; Leslie, 1992) and SCALA (CCP4, 1994), using the first 137 ° of rotation, thus omitting data most severely affected by radiation damage during collection. The model was initially refined in CNS (Brunger *et al.*, 1998) using simulated annealing with 4-fold NCS constraints, and unbiased HL coefficients from auto-SHARP. The co-factor could be clearly seen in both the $2F_o$ and $F_o - F_c$ maps (contoured at 1.0 δ , and 4.0 δ , respectively), and was subsequently modelled as NADPH, even though the puckering of the reduced nicotinamide ring could not be seen at 1.8 Å resolution. Refinement continued using CNS with gradual relaxation of the NCS parameters, from four-fold constraints, to two-fold constraints, and finally to four-fold restraints. The final stages of refinement were performed using REFMAC5 (Murshudov *et al.*, 1997) where TLS refinement (Winn *et al.*, 2001) with each monomer as a TLS group resulted in a significant drop in both R_{crys} and R_{free} . Solvent molecules were added using ARP-wARP (Lamzin *et al.*, 2001) and were manually inspected in O. The final model has an R_{crys} of 18.2 % and an R_{free} of 20.9%.

4.4 Structure of CAD

4.4.1 Secondary and Tertiary Structure

Previous sequence analysis of CAD revealed homology to enzymes belonging to the SDR family (Nicholson *et al.*, 1994). Members of the SDR family possess very few highly conserved residues, and generally show low sequence homology, in the order of 15-30%. The three-dimensional structures of SDRs show a conserved Rossmann fold, common to nucleotide-binding enzymes (Rossmann *et al.*, 1974). The structure of CAD reveals that the enzyme consists of a single α/β domain, composed of seven β -strands (β A, β B, β C, β D, β E, β F, β G) and eight α -helices (α B, α C, α D, α E, α F, α FG1, α FG2, α G), consistent with the Rossmann fold. Seven β -strands form a parallel β -sheet, which is flanked by helices α C, α B, and α G on one side, and helices α D, α E, and α F on the other. Helices α FG1 and α FG2 are positioned to one side over the top edge of the C-terminal end of the β -sheet, the co-factor binding site. Helix α FG1 lies in a similar plane to helix α B, except that the helical axis is pointing towards the pyrophosphate moiety of the nucleotide cofactor, in a manner similar to helix α B. Helix α FG2 lies on the top edge of the β -sheet, and assists in the creation of the substrate-binding site.

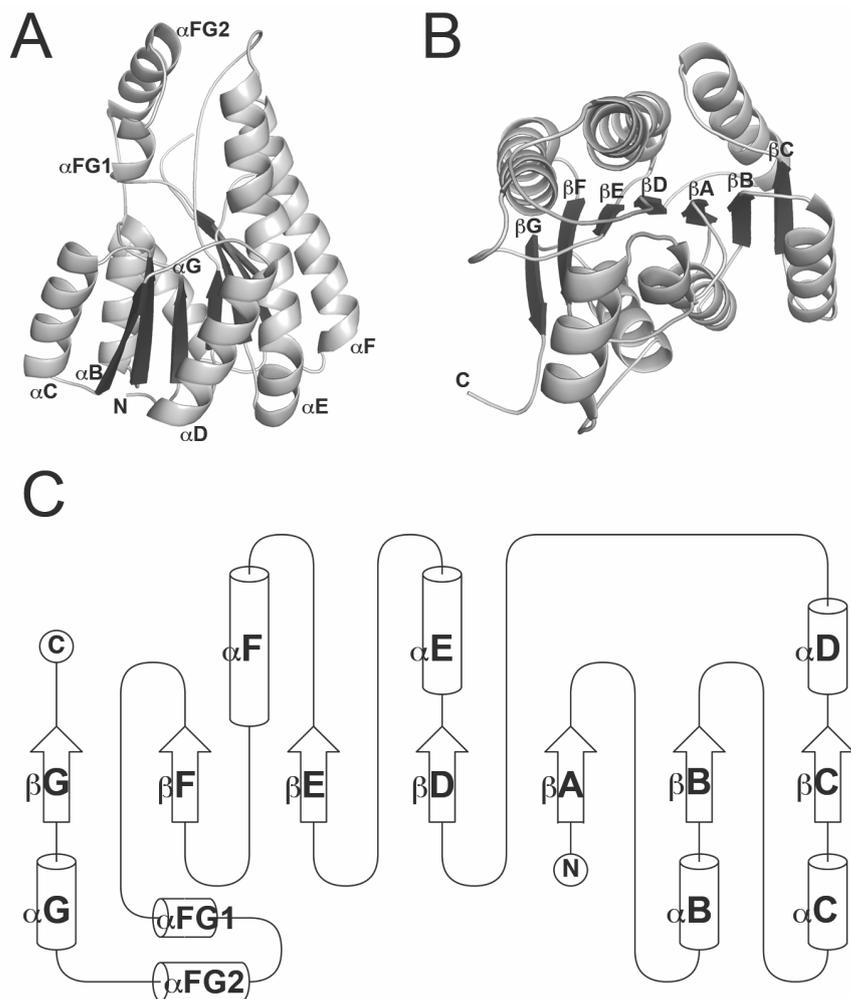


Figure 4.2 The Rossmann fold of CAD. (A) The tertiary structure of CAD, where the N-terminus and α -helices are labelled. (B) Tertiary structure of CAD as viewed perpendicular to the parallel β -sheet, where the C-terminus and β -strands are labelled. (C) The secondary structure elements and their topological relationship to each other.

4.4.2 Quaternary Structure

The four molecules (A–D) of the asymmetric unit are arranged to form a tetramer, possessing 222 point symmetry (Figure 4.3). The arrangement is similar to that observed in other tetrameric SDRs. This results in the formation of three two-fold axes, which have been labelled P, Q, and R (after Rossmann *et al.*, 1973). These axes represent the interfaces of the tetramer interactions. The interactions along the P and Q axes comprise the greatest interaction surface area, with 1390 \AA^2 and 1631 \AA^2 per monomer, respectively. The R axis contributes little to tetramer formation, with 80 \AA^2 per monomer. The interactions of the P axis consist of the hydrophobic interactions of helix αG between monomers A and B, and by the hydrogen bonding

of β -sheet extension (β G). The Q interface is characterized by the hydrophobic interaction of two helices, α E and α F, from both the A and C monomers, which results in the formation of a four-helix bundle.

The interactions along the P and Q axes are important for defining the oligomeric state of SDRs. For example the rat liver dihydropteridine reductase contains an eighth parallel β -strand, preventing tetramer interactions along the P axis (Varughese *et al.*, 1992). Alternatively, the 3α -hydroxysteroid dehydrogenase/carbonyl reductase possesses two additional α -helices prior to α F, which prevent dimer interactions along the Q axis, preventing the formation of the four-helix bundle (Grimm *et al.*, 2000). Dimer interactions along the Q axis are favoured, due to a reduction in the hydrophobic surface exposed to the solvent (Sawicki *et al.*, 1999), thus the dimeric form of CAD observed in solution most likely consists of this Q-axis dimer (molecules A and C).

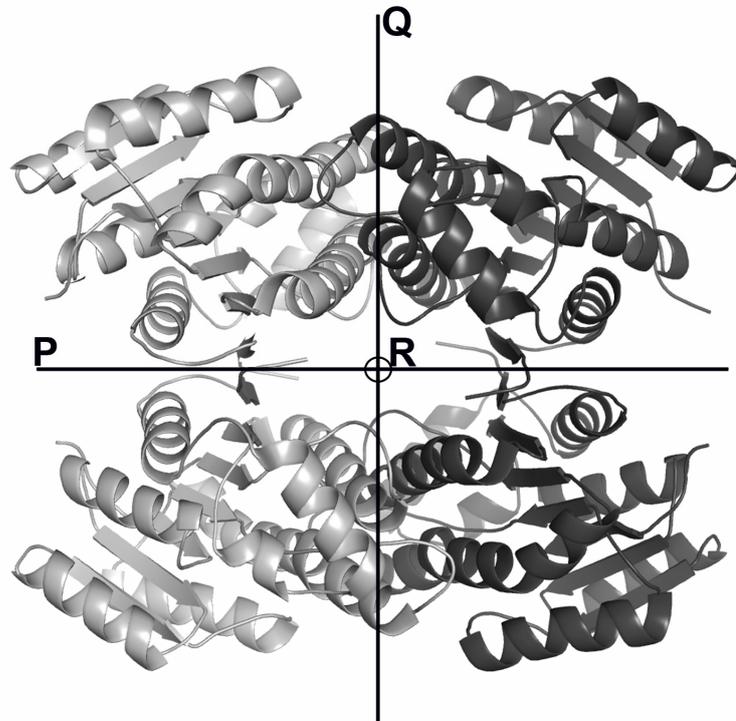


Figure 4.3 Quaternary structure of CAD. The P, Q, and R axes are indicated, where the R axis is parallel to the point of view. The A/B dimer (light grey) forms the P axis, and the A/C dimer (light grey| dark grey) forms the Q axis.

In comparison to many other tetrameric SDRs CAD shows reduced interactions along the R axis, due to a shorter C-terminal tail, which is held in a more compact position. The C-terminal tail is tucked in close to α FG2, where this interaction is aided by a salt bridge between Arg²¹³ and Asp²⁴⁵. The hydrogen bonding

interactions between Gln²⁴⁶ and Thr¹⁸⁷ may also contribute, although these residues are also involved in the binding of the substrate.

In between the four CAD monomers a tunnel is formed, running along the R-axis, with approximate dimensions of 8 Å x 13 Å x 34 Å (Figure 4.4). The tunnel is filled with solvent, allowing for hydrogen bonding interactions between all four monomers *via* extended solvent bridges. Similar tunnels can be found in other tetrameric SDRs, although the C-terminal tail frequently forms a cap over the tunnel, closing it to the external solvent. In CAD, the substrate binding sites from two monomers related by the R-axis (*e.g.* monomers B and C) are located in close proximity to the tunnel entrance. The tunnel is of sufficient size to permit movement of the substrate through the tunnel, although this is unlikely to happen in a solvent filled tunnel.

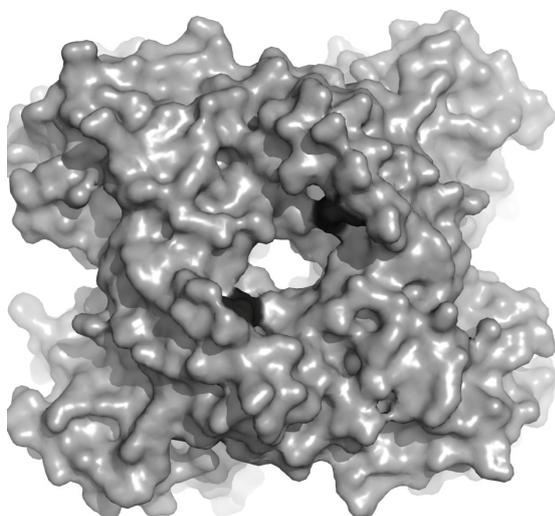


Figure 4.4 A surface projection showing the tunnel formed between the four monomers of the tetramer, as viewed along the R axis. The position of the substrate binding sites are shown in black.

4.5 Similarity with Other SDRs

4.5.1 Sequence Similarity

The SDR superfamily can be classified into five different families based on conserved sequence motifs (Kallberg *et al.*, 2002). Few residues are strictly conserved throughout the superfamily, not even the residues of the catalytic triad (S, Y, K), which define reductase activity (Jornvall *et al.*, 1995). For example, in human mitochondrial 2,4-dienoyl-coA reductase the typical YxxxK motif is replaced by the unique YxxxN motif (Alphey *et al.*, 2005), while in the enol-ACP reductase from *Mycobacterium tuberculosis* the conserved tyrosine is substituted for methionine, YxxMxxxK (Rozwarski *et al.*, 1999).

Several of the motifs defining classical SDRs are present in CAD. The Asn⁹⁰-Asn⁹¹-Ala⁹²-Gly⁹³ motif is important in stabilizing the central β -sheet (Filling *et al.*, 2002), while the glycine rich motif, Thr¹³-Gly¹⁴-X₃-Gly¹⁸-X-Gly²⁰, assists the binding of the dinucleotide cofactor (Jornvall *et al.*, 1995). Asp⁶⁴ has been proposed to assist stabilization of a pocket binding the cofactor adenine ring (Filling *et al.*, 2002), however in CAD it appears to also interact directly with the adenine ring. The catalytic triad has been extended to include a conserved asparagine and water molecule, giving rise to the catalytic tetrad (Filling *et al.*, 2002; Oppermann *et al.*, 2003). Both the catalytic triad (Ser¹⁴², Tyr¹⁵⁵, and Lys¹⁵⁹) and Asn¹¹⁵ are conserved in CAD, as is the water molecule associated with Asn¹¹⁵ (see Section 4.7.) Another conserved water molecule found linking the glycine rich nucleotide binding motif, the cofactor pyrophosphate, and the C-terminal residue of β D (*i.e.* Asn⁹¹) in other SDRs (Bottoms *et al.*, 2002) is also conserved in CAD.

Members of the classical SDR family can be classified further based on their cofactor dependence (Kallberg *et al.*, 2002). The fact that CAD is dependent on NADPH as a cofactor, rather than NADH (Fulston *et al.*, 2001) Paper II) is consistent with its structure. Classical SDRs binding NAD(H) possess an acidic residue at the end of strand β B, which interacts with the adenosine-ribose hydroxyl groups, *e.g.* Asp³⁸ in the 3 α ,20 β -hydroxysteroid dehydrogenase (Ghosh *et al.*, 1994). In CAD this position is occupied by the non-polar residue, Ala³⁸, preventing any electrostatic repulsion of the negatively charged 2'-phosphate group of the cofactor. Classical SDR enzymes binding NADP(H) generally possess two characteristic basic residues: the first immediately preceding the second glycine of the glycine-rich cofactor-binding motif, and the second following directly after the crucial acidic residue of NAD(H)-preferring enzymes (Kallberg *et al.*, 2002; Tanaka *et al.*, 2001). In CAD these positions correspond to residues Ser¹⁷ and Arg³⁹, both of which coordinate phosphate oxygen's from the pyrophosphate, and adenine-phosphate respectively. The substitution of Ser¹⁷ for a basic residue demonstrates that the requirement for basic residues is not absolute, and thus CAD can be classified into the cP2 subfamily (Kallberg *et al.*, 2002).

4.5.2 Structural Similarity

A C α alignment of SDRs reveals the conserved nature of the α/β fold (Figure 4.5), where significant variation can be found between the structures around the C-terminus, and around helices α FG1 and α FG2. The variation in the C-terminal region is important for interactions along the P and R interfaces (Figure 4.5f). The variation found in the α FG1- α FG2 region is due to its proximity to the binding site, where these helices form a lid over the substrate-binding site (Duax *et al.*, 1996; Tanaka *et al.*, 2001). Given the diverse range of substrate specificities of SDR enzymes, the variation in this region is hardly surprising. Since the substrate of CAD is a relatively small molecule, the α FG1- α FG2 lid adopt a rather closed conformation in comparison to other SDRs. The closed conformation of the lid region is stabilised by the interactions between the α FG1- α FG2 loop and the β E- α E loop, where hydrogen bonds between Thr¹⁹⁷ and Asp¹⁰⁷ and the carbonyl group

of Leu⁹⁷ link the two loops *via* a solvent bridge (e.g. water Z⁶¹ in the binary complex).

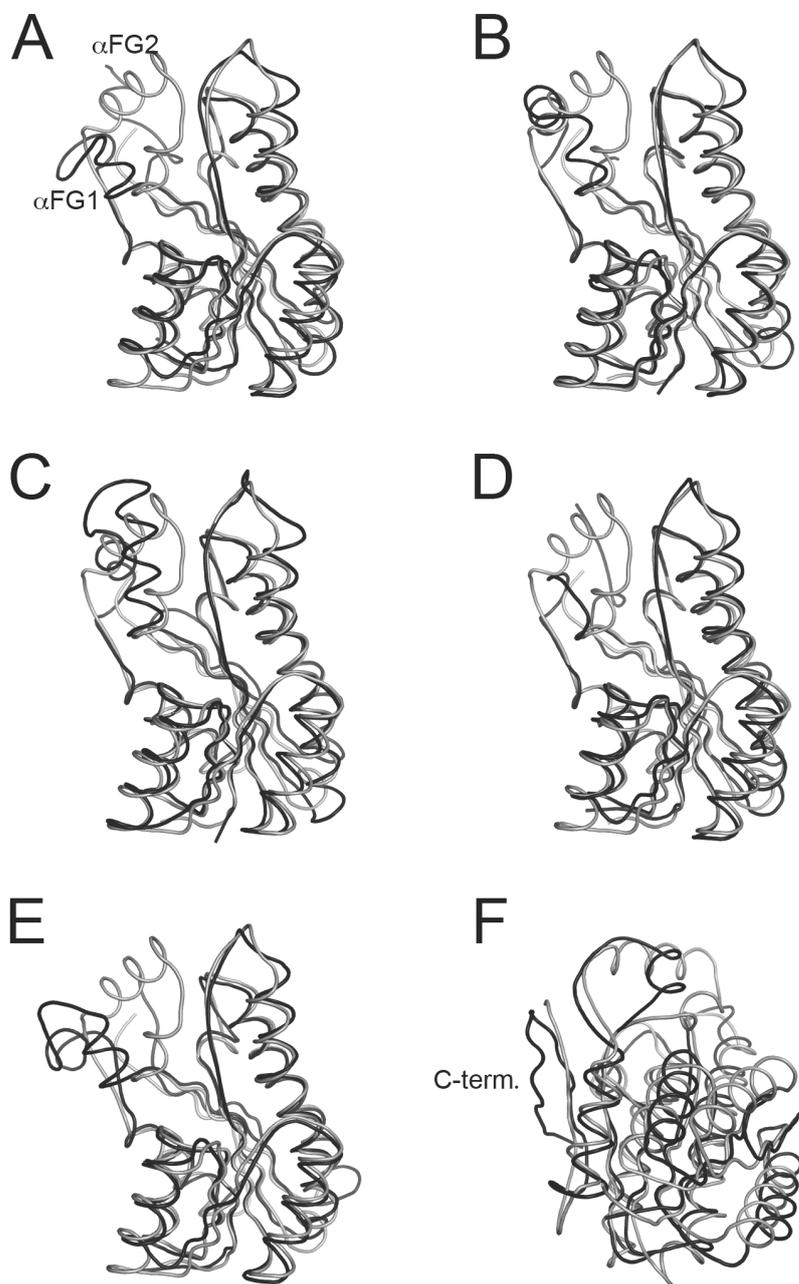


Figure 4.5 (A) – (F) Legend on next page.

Figure 4.5 Conservation of the Rossmann-fold in CAD (grey) and other SDR enzymes (black). The C α alignment of CAD (2JAH) with (A) 3 α ,20 β -hydroxysteroid dehydrogenase (2HSD, r.m.s.d. = 1.21 Å for 207 atoms). (B) β -keto acyl carrier protein reductase (1EDO, r.m.s.d. = 1.26 Å for 217 atoms). (C) meso-2,3-butanediol dehydrogenase (1GEG, r.m.s.d. = 1.24 Å for 215 atoms). (D) Rv2002 (1NFF, r.m.s.d. = 1.31 Å for 212 atoms). (E) 3 β /17 β hydroxysteroid dehydrogenase (1HXH, r.m.s.d. 1.23 Å for 206 atoms. (F) Rat liver dihydropteridine reductase (1DHR, r.m.s.d. = 1.88 Å for 166 atoms), demonstrating the elongated C-terminal region which prevents the formation of a tetramer.

4.6 The Cofactor and Substrate Binding Sites

4.6.1 NADPH Cofactor Binding in the Binary Complex

The cofactor is bound at the C-terminal end of the central β -sheet (strands β B, β A, β D and β E) in a binding pocket with a solvent-accessible buried surface of *ca.* 1325 Å² (Brunger *et al.*, 1998). Strands β C and β F form the two sides of the pocket, binding the adenine and nicotinamide moieties, respectively. The cofactor is bound in an extended conformation. The degree of extension has been measured in other SDRs using the distance between the C6 carbon of the adenine ring, and the C2 carbon of the nicotinamide group in a *syn* conformation (Rossmann *et al.*, 1975). In CAD this distance is 13.8 Å, a value similar to other SDRs.

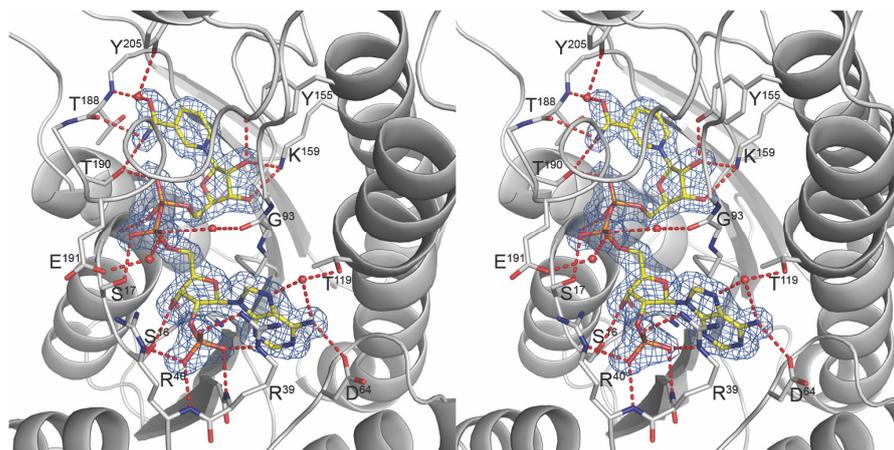


Figure 4.6 Stereo-figure depicting the binding of NADPH (yellow) in CAD (grey). Water molecules are depicted as red spheres, and putative hydrogen bonding interactions as red dashes. The $2mF_{obs} - DF_{calc}$ electron density map (blue) at NADPH is contoured at 1δ , where δ is the root-mean square electron density for the unit cell. The Helix α FG1 has been represented as a loop for clarity.

The NADPH cofactor is composed of three different parts, the adenine moiety, the pyrophosphate group and the nicotinamide moiety. The adenine moiety consist of an adenine nucleotide connected to a ribose, where the 2'-hydroxyl group is substituted for a phosphate group. The adenine nucleotide is bound in a hydrophobic pocket (Gly¹⁴, Ala³⁸, Val⁶⁵, Ala⁹², Ile⁹⁴), as well as forming a potential hydrogen bond to the conserved Asp⁶⁴. Arg³⁹ assists in the closure of cofactor binding site, and interacts with the 2'-phosphate group, along with Arg⁴⁰. The conserved T¹³-G¹⁴-X₃-G¹⁸-X-G²⁰ motif is located in the loop between β A and α B.

Ser¹⁶ interacts with the ribose hydroxyl group, while Ser¹⁷ interacts with the pyrophosphate moiety of NADPH, hydrogen bonding to one of the phosphate oxygens. The other oxygen is coordinated by the carbonyl of Gly⁹³ *via* a solvent bridge (*e.g.* water W⁷³ in the A molecule of the binary complex). The dipole charge of helix α B is likely to assist in the stabilization of the pyrophosphate group. Helix α FG1 may also stabilize the pyrophosphate group in a similar way, albeit to a lesser degree due to its decreased length. The nicotinamide moiety consists of a ribose, with a nicotinamide ring bound to it. The nicotinamide ring is bound in a pocket consisting of a hydrophobic “floor” (residues Ile¹⁹, Met¹⁴⁰, Pro¹⁸⁵, Thr¹⁸⁸), and a polar “ceiling”. The polar residues assist in coordinating substrate binding, as well as catalyzing the reduction of the aldehyde. Lys¹⁵⁹ forms bifurcated hydrogen bonds to both the 2' and 3' hydroxyl groups of the nicotinamide ribose, assisting in coordination of the cofactor, as well as lowering the pKa of the catalytic base Tyr¹⁵⁵ (Chen *et al.*, 1993) allowing proton transfer as part of the catalytic mechanism.

4.6.2 Substrate Binding Site

Clavuldehyde is an unstable molecule in solution, and thus to circumvent problems associated with the preparation and storage of the aldehyde, clavulanic acid was used for crystallographic analyses. SDRs often catalyse reversible reactions, and this is true of CAD. In this respect clavulanic acid can be also considered a substrate of CAD, although the equilibrium of the reaction appears to lie firmly on the side of clavulanic acid and NADP⁺ (Fulston *et al.*, 2001); Paper II). The “soaking in” of clavulanic acid into a CAD crystal co-crystallized with NADPH ensured that a non-productive complex was formed, since clavulanic acid cannot be oxidized to clavuldehyde in the presence of NADPH.

Like other SDR enzymes the substrate binding pocket is located in the region around helices α FG1 and α FG2. The binding pocket is formed by a small invagination of *ca.* 10 Å, with a surface interaction area of *ca.* 480 Å² (Brunger *et al.*, 1998). The binding site is created by the residues in the loops β D– α E (Met⁹⁵, Leu⁹⁷), β E– α F (Ser¹⁴²–Ala¹⁵²), and β F– α FG1 (Gly¹⁸⁶, Thr¹⁸⁷). Residues from helices α F (Asn¹⁵¹, Ala¹⁵², Tyr¹⁵⁵, Gln¹⁵⁶) and α FG2 (Thr²⁰¹, Tyr²⁰⁵, Arg²⁰⁸, Ile²⁰⁹) also contribute to the formation of the binding site, as does the C-terminal Gln²⁴⁶. Although the binding site is located at the periphery of the molecule, a small tunnel extends towards the hydrophobic core of the protein, where the catalytic residues and cofactor are located.

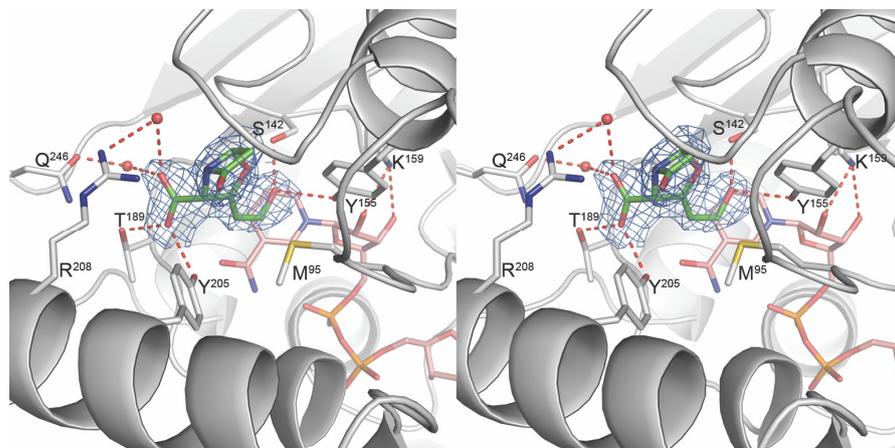


Figure 4.7 Stereo-figure depicting the binding of clavulanic acid (green) and NADPH (salmon) in the CAD ternary complex (grey). Water molecules are depicted as red spheres, and putative hydrogen bonding interactions as red dashes. The $2mF_{obs} - DF_{calc}$ electron density map (blue) at clavulanic acid is contoured at 1σ , where σ is the root-mean square electron density for the unit cell.

The binding of clavulanic acid results in the displacement of four water molecules. For example, the waters Z^{83} , Z^{110} , Z^{95} , and Z^{23} in the binary complex are replaced by the C9 alcohol moiety, the O5 oxygen of the C3 carboxyl group, the β -lactam ring C4 carbon, and the C7 carbonyl group, of clavulanic acid respectively. Several hydrogen bonds secure clavulanic acid within the binding site, such that the C9 alcohol forms hydrogen bonds to both Ser¹⁴² and the catalytic base, Tyr¹⁵⁵, while the C3 carboxyl moiety interacts with Thr¹⁸⁷ and Tyr²⁰⁵ (Figure 4.7). Additional coordination of clavulanic acid is achieved via solvent interactions with Gln²⁴⁶ and Arg²⁰⁸ (e.g. waters Z^{58} and Z^{68} in the ternary complex).

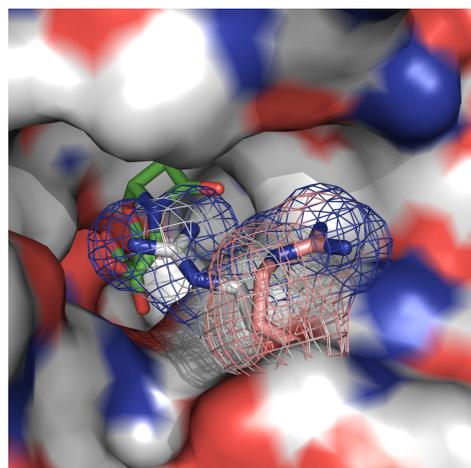


Figure 4.8 The movement of Arg²⁰⁸ associated with the binding of clavulanic acid. Arg208 moves from a side pocket in the binary complex (salmon) to a position directly above clavulanic acid (green) in the ternary complex (grey).

Conformational changes associated with substrate binding have been observed in other SDRs, *e.g.* 7α -hydroxysteroid dehydrogenase reductase (Tanaka *et al.*, 1996). However, in CAD no significant conformational changes were observed between the binary and ternary complexes (R.M.S.D = 0.189 Å for the 981 C α atoms of the tetramer). Within the substrate binding site only Arg²⁰⁸ appears to undergo any conformational change upon the binding of clavulanic acid (Figure 4.7). In the binary complex the Arg²⁰⁸ side chain lies predominantly to one side, leaving a tunnel leading to the substrate binding pocket. In the ternary complex the Arg²⁰⁸ side chain lies directly over the substrate-binding pocket, essentially closing the binding site to the external solvent (Figure 4.8). In both the binary and ternary complexes the side chain of Arg²⁰⁸ appears to be quite motile, indicated by the higher *B*-factor (average *B*-factor = 40.6 Å² and 45.4 Å², respectively) and poorly defined electron density (RSCC = 0.823 and 0.877 respectively).

4.6.3 Reaction Stereochemistry and Mechanism

The binding of clavulanic acid in the active site positions the C9 carbon is *ca.* 2.5 Å above the pro-*S* hydrogen of the nicotinamide ring. An angle of 123 ° formed between the C4 hydrogen of NADPH and carbon C9 of clavulanic acid, consistent with the proposed common mechanism for SDR oxidoreductases (Jornvall *et al.*, 1995; Tanaka *et al.*, 2001) where the 4-pro-*S* hydrogen of NADPH is transferred to the substrate's carbonyl group. Clavulanic acid adopts an eclipsed conformation within the active site, such that the C9-OH bond projects towards the clavam oxygen, being almost coplanar with the O1-C2 and exocyclic alkene bonds (Figure 4.9a). This is in contrast to the conformation observed in small molecule crystal structures of clavulanic acid, where the C9-OH bond points away from the clavam ring nucleus (Figure 4.9b)

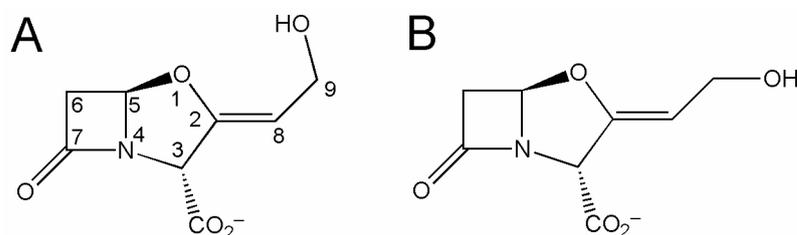


Figure 4.9 Clavulanic acid in an eclipsed conformation (A) seen in the binding site of CAD, and in an "open" conformation (B) observed in the small molecule structure (Brown *et al.*, 1984).

Ornithine is a direct precursor of arginine, such that radiolabelled ornithine is incorporated into clavulanic acid. Previous studies using tritium labelled ornithine demonstrated that the pro-*R*, but not the pro-*S* hydrogen, was incorporated at the C9 position of clavulanic acid (Townsend *et al.*, 1986). Furthermore, the stereochemistry of the ornithine C5 pro-*R* hydrogen was inverted, such that it occupied the pro-*S* position (Figure 4.10). The adoption of an eclipsed conformation within the active site is necessary for the hydride transfer to the re-

face of the aldehyde. In this way the hydrogen transferred from NADPH occupies the pro-*R* methylene position of clavulanic acid.

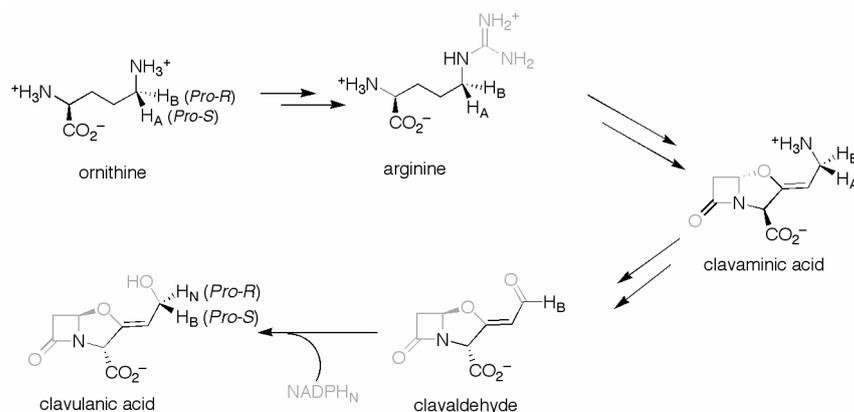


Figure 4.10 A summary of previous labelling studies showing the incorporation of tritiated ornithine clavulanic acid via arginine (Townsend *et al.*, 1986).

The hydrogen bonding of the C9-aldehyde to Ser¹⁴² and Tyr¹⁵⁵ probably promotes nucleophilic attack by the hydride. The transfer of the hydride ion results in a positive charge on the newly oxidised nicotinamide ring. Such a charge would assist the lowering of the pK_a of the Tyr¹⁵⁵ hydroxyl group to around pH 6.5–7.0, allowing the transfer of a proton from the tyrosine hydroxyl group to the alkoxide forming during the reduction step. A mechanism leading to the reprotonation of the catalytic Tyr¹⁵⁵ has been proposed for SDRs, where protons are extracted and transferred from surrounding water molecules in a small hydrophilic pocket (Filling *et al.*, 2002; Oppermann *et al.*, 2003). The proposed proton relay system in CAD is shown in figure 4.11b, where the residues Thr¹³, Asn⁹⁰, Ala⁹², Thr¹¹⁴, and Asn¹¹⁵ may play a role in stabilising the water molecules, facilitating proton transfer in an otherwise hydrophobic environment. This is similar to that proposed for the 3β/17β-hydroxysteroid dehydrogenase (3β/17β-HSD) (Filling *et al.*, 2002), except that the residue corresponding to Ser¹¹⁴ in 3β/17β-HSD is a glycine residue (Gly¹¹⁸) in CAD. Such a mechanism requires the displacement of the NADP⁺ from the co-substrate binding pocket, raising the pK_a value of the Tyr¹⁵⁵ to allow proton extraction. Since large conformational changes are required for the binding of NADPH it is also possible that Tyr¹⁵⁵ is reprotonated directly from the solvent.

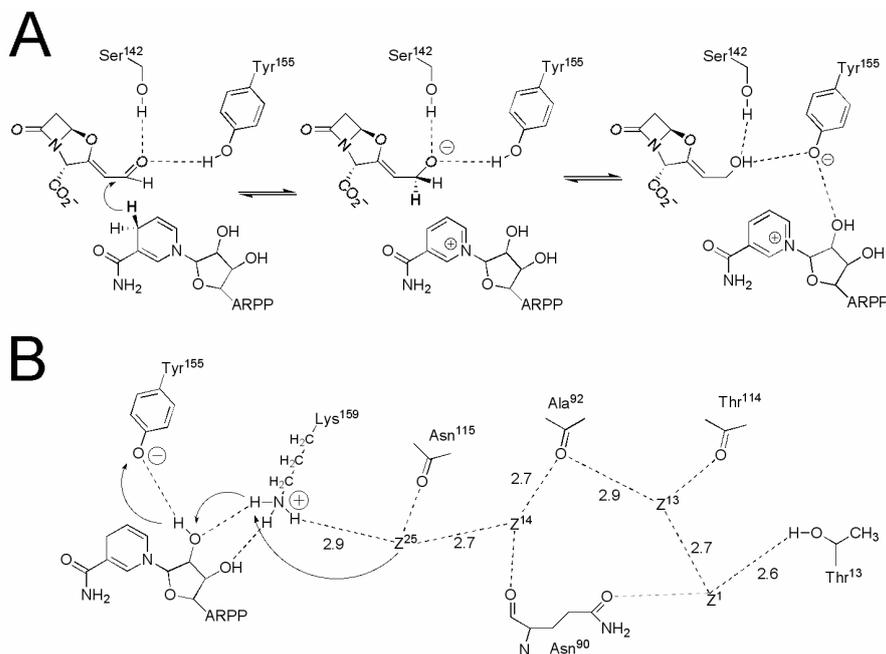


Figure 4.11 CAD catalysed reduction of clavialdehyde. (A) Proposed reaction mechanism. (B) Potential proton relay based on the mechanism proposed for the 3 β /17 β -HSD (Filling *et al.*, 2002), where the arrows indicate the path of protons, rather than electrons. Residues coordinating waters are shown, as well as distance between water molecules (in Å). ARPP is the adenosine ribose pyrophosphate moiety of NADP(H).

4.7.1 Stabilizing the Unstable?

The clavam bicyclic ring is more strained than the bicyclic ring of the penicillins. This is due to the substitution of an oxygen atom for the sulphur atom, the absence of a C6 side-chain, and the presence of the exo- β -hydroxyethylidene function at C2 (Baggaley *et al.*, 1997). The dipole moment of the C9 aldehyde leaves clavialdehyde prone to nucleophilic attack, making it an unstable molecule in solution; to paraphrase J. Hajdu "the alkene side chain is like a straw, if you attack the C9-aldehyde you suck electrons out of the bicyclic ring, and break it open". So how does CAD catalyse the reduction of such an unstable molecule?

One potential mechanism by which CAD protects the reactive intermediate is in the procurement of clavialdehyde away from the aqueous environment of the surface. The substrate binding site is located close the surface of the molecule, however the reactive C9-aldehyde is buried in the hydrophobic core, thus preventing adventitious hydrolysis by water molecules. Another potential mechanism may be the shielding of the β -lactam ring. In the binding pocket the β -lactam C7 carbonyl is pointing towards a hydrophobic pocket (residues Leu⁹⁷, Val¹⁴⁹, Ala¹⁵², Tyr²⁰⁵), and is thus unable to participate in hydrogen bonding. The opening of the β -lactam ring by β -lactamases involves the nucleophilic attack of the C7 carbonyl by a serine residue. By shielding the C7 carbonyl in the binding site CAD may prevent the nucleophilic attack by opportunistic water molecules. Interestingly a similar

situation is found in CAS (Zhang *et al.*, 2000), where the monocyclic intermediate, proclavaminic acid, is also shielded in a small hydrophobic pocket.

Additional stability during the reduction of clavaldehyde could be achieved by the interaction of additional residues during catalysis. Methionines possess a lone electron pair, which are capable of hydrogen bonding as well as assisting chemical catalysis. Met⁹⁵ lies in relatively close proximity to the C9 carbon of clavulanic acid, *ca.* 3.6 Å away, and could easily adopt a favourable conformation allowing the lone pair to move closer, *ca.* 2.7 Å. Such a mechanism may stabilise clavaldehyde during hydride transfer, as well as the alkoxide reaction intermediate.

Another potential mechanism of preventing the degradation of clavaldehyde in solution is by limiting the time the aldehyde is exposed to the general solvent of the cell. It is possible that a CAD tetramer interacts directly with the enzyme(s) responsible for the synthesis of clavaldehyde to facilitate a rapid transfer of the unstable intermediate. At present the gene products of *orf10* and *orf11* are thought to be responsible for catalysing the formation of clavaldehyde, however little is known about how this is achieved, nor whether they interact with other enzymes (or proteins) from the biosynthesis pathway.

5 ORF15 (Paper III)

5.1 Background

Approximately half of the genes in the clavulanic acid gene cluster are open reading frames (ORFs), with putative functions assigned based on their sequence similarity to proteins of known function (Section 2.3). The role of these ORFs in the biosynthesis of clavulanic acid biosynthesis remains unclear (section 2.3.1). *Orf15* is one of these genes.

Orf15 is thought to be expressed as a polycistronic transcript along with *orf16*, *orf17* and *orf18* (Mellado *et al.*, 2002). It encodes a 562 amino acid protein, with an estimated molecular weight of 61.9 kDa. Sequence comparisons using BLAST revealed that ORF15 shares similarity with members of the bacterial periplasmic protein-like II superfamily. These bacterial substrate-binding proteins (SBPs) are associated with the binding of substrates, which are imported into the bacterial cell via the membrane bound permeases of the ATP-binding cassette (ABC) transport pathway (Ames, 1986; Higgins, 1992). ORF15 shows the greatest similarity to members of the family five SBPs (Tam & Saier, 1993), possessing the characteristic sequence motif (PROSITE PS01040), which is represented by residues Gly⁸⁸ to Tyr¹⁰⁹ (Mellado *et al.*, 2002). ORF15 shows the highest sequence similarity with oligopeptide-binding proteins, with *ca.* 20% sequence identity and *ca.* 40% sequence similarity, including DppA from *E. coli* and OppA from *Salmonella typhimurium* (Abouhamad *et al.*, 1991; Olson *et al.*, 1991).

ORF15 also shows a 48% amino acid sequence identity with ORF7, which is located upstream of *Orf15* in the clavulanic acid gene cluster (see Figure 2.1). Oligopeptide transporters have not been reported in other antibiotic gene clusters, and the apparent requirement for two putative oligopeptide-binding proteins in clavulanic acid biosynthesis is rather intriguing. Both *orf7* and *orf15* appear to play an essential role in the production of clavulanic acid (de la Fuente *et al.*, 2002; Jensen *et al.*, 2000; Jensen *et al.*, 2004a; Lorenzana *et al.*, 2004) with a complete loss of clavulanic acid production associated with the gene-disruption mutants. Despite their similarity the two genes show no cross complementation (Jensen *et al.*, 2004a; Lorenzana *et al.*, 2004). The gene disruption of *orf15* also affected the production of 5S-clavams. Jensen *et al.* (2004a) observed detectable levels of alanylclavam in *orf15* mutants, while Lorenzana *et al.* (2004) did not. N-acetyl-glycyl-clavaminic acid (NAG-clavam) and N-glycyl-clavaminic acid, which have only been detected previously in the *dcl8* mutant (Elson *et al.*, 1998), were also produced by the *orf15* mutant, *albeit* at a much lower levels (Jensen *et al.*, 2004a). The non β -lactam antibiotic holomycin was also produced by the *orf15* mutants (de la Fuente *et al.*, 2002; Lorenzana *et al.*, 2004).

The results of Lorenzana *et al.* (2004) are consistent with the assigned function for both ORF7 and ORF15 as oligopeptide-binding proteins, based on sequence similarity. Gene disruption mutants of *orf15* mutants were completely inhibited in their ability to utilise a nonapeptide (bradykinin) present in the growth medium.

This suggests that ORF15 is involved in the uptake of oligopeptide nutrient sources, and that both ORF7 and ORF15 possess different substrate preferences. Furthermore, mutants of *orf7* and *orf15* were partially inhibited in their ability to transport of the phosphinothricyl-alanyl-alanine (Pt-AA) into the cell. Pt-AA is a toxic tripeptide known to enter the cell *via* the oligopeptide transport system (Diddens *et al.*, 1976), and is again consistent with the proposed role of both ORF7 and ORF15 as oligopeptide-binding proteins.

ORF15 also appears to be involved in additional cellular processes, as mutants of *orf15* were unable to form aerial mycelium, and thus could not produce spores (Lorenzana *et al.*, 2004). This so-called "bald phenotype" has been observed in the *Bld* gene mutants (section 2.1) which have been well characterised in *S. coelicolor* A3. It is interesting to note that a bald phenotype was not reported for *orf15* disruption mutants in other studies (de la Fuente *et al.*, 2002; Jensen *et al.*, 2004a).

Without a doubt the big question remains: What is the function of ORF15 in clavulanic acid biosynthesis?

5.2 Expression, Purification and Crystallization of ORF15

Orf15 from *S. clavuligerus* was initially cloned into a pET28a vector, containing an N-terminal His₆-tag with the sequence MGSSHHHHHHSSGLVPRGS. A purification protocol for the his-tagged recombinant protein was developed by collaborators in Oxford. It soon became apparent that the purified protein did not tolerate freezing at -80 °C, as its ability to crystallize diminished rapidly with each successive freeze/thaw cycle. The freezing of the bacterial pellet prior to purification also resulted in reduced yields of ORF15, thus the expression and purification of the recombinant protein was performed consecutively using a (more or less) non-stop protocol described below.

The *orf15*/His₍₆₎ construct was transformed into *E. coli* BL21 (DE3) cells. Cells were grown at 18 °C in 2YT media containing kanamycin (30 µg ml⁻¹). Expression was induced by the addition of 1 mM IPTG when the culture reached a cell density of OD_{600nm} = 0.6–0.8. Initially the temperature was raised to 30 °C, and the cells grown for 6 hours. However, for the expression of the Se-Met protein, and subsequent native protein, the cells were grown at a constant 18 °C for 18 hours after induction. Cells were resuspended in 50 mM phosphate buffer (pH 7.5) and lysed by the addition of lysozyme and sonication. The lysate was cleared by centrifugation, and the supernatant was loaded onto a Ni-NTA column (GE Healthcare). The column was washed with a step-wise gradient of 5% and 10% imidazole, before elution of the recombinant protein using a 10%-100% imidazole gradient. Samples eluted from the column were frequently collected in small vials containing stabilisation buffer, which consisted of 50 mM Tris-HCl (pH 7.5), 5 mM EDTA and 100 mM NaCl. This was required to prevent partial precipitation of the purified protein after elution. The peak fractions were analysed by SDS-PAGE and pooled before being dialysed overnight in 50 mM Tris-HCl (pH 7.5), 2 mM EDTA and 100 mM NaCl. Protein was concentrated and subjected to size exclusion chromatography (Superdex S200) to yield ORF15 of >90% purity by

SDS-PAGE analysis. The purified ORF15 protein was stored in 5 mM HEPES (pH 8.0), 1 mM EDTA and 5 mM β -mercaptoethanol. The purified recombinant protein corresponded to a band of 64 kDa on SDS-PAGE, consistent with the predicted size of the his-tagged protein. Gel filtration suggested that ORF15 was present predominantly as a monomer in solution, although aggregates of higher molecular weight could be observed eluting in the void volume.

The expression of seleno-methionine protein was performed as above, with the following exceptions: the *orf15/His₍₆₎* construct was transformed into *E. coli* 834 (DE3) cells, and grown in SeMet medium™ (Molecular Dimensions Ltd, U.K) using L-seleno-methionine, with 5 mM β -mercaptoethanol was included in all the purification buffers.

Crystallisation experiments were performed at 18 °C using the sitting-drop vapour diffusion method. Crystals were obtained using 20-23% (w/v) PEG6000 in 0.1M NaCl and 0.1M Tris/HCl. The crystallization conditions frequently produced a mass of small crystals with a morphology resembling a "sea-urchin". A seed stock was prepared by crushing the sea-urchin crystals, and three-dimensional crystals were obtained by seeding of the crystallization drop of the same conditions with the seed-stock using a whisker obtained from *Felis silvestris sp. Catus*. Diffracting crystals of Se-Met protein were obtained by seeding into a drop consisting of 1.9 μ L of protein (*ca.* 9 mg ml⁻¹) and 1.9 μ L of reservoir solution consisting of 14 % (w/v) PEG 6000, 0.05 M Tris/HCl (pH 8.0), 0.32 M sodium acetate and 2.5 mM MgCl₂.

5.3 Structure Determination

The native crystals varied in diffraction quality, the majority of crystals were highly mosaic, with split reflections, and diffracting to only low resolution. However, screening of the crystals for improved diffraction identified a crystal diffracting to 1.45 Å. A data set was collected at ID14-1 (ESRF, Grenoble, France) at cryogenic temperatures. Data were initially processed using DENZO/SCALEPACK (Otwinowski & Minor, 1997), where the crystal belonged to space group *P*2₁, with unit cell dimensions *a*= 53.6 Å, *b*= 84.4 Å, *c*= 62.4 Å, β = 91.7 °. Attempts to solve the structure using molecular replacement (MR, AMORE, PHASER) were unsuccessful.

A single-wavelength anomalous diffraction (SAD) data set was collected on a Se-Met crystal diffracting to 1.7 Å, at a remote wavelength for selenium (ID14-3, ESRF, Grenoble, France). The data were collected using an oscillation angle of 0.5 ° with 720 images collected. The data were initially processed to 1.8Å using DENZO/SCALEPACK(Otwinowski & Minor, 1997). The crystal belonged to space group *P*2₁, with unit cell dimensions *a* = 53.3 *b* = 120.8 *c* = 80.0 β = 94.1 °. The inversion of the *b* and *c* axes gave an asymmetric unit twice the size of the native data. Assuming the presence of two molecules in the asymmetric unit gave a *V_M* of 2.1 Å³ Da⁻¹ (Matthews, 1968) with 40.1 % solvent. Initially confusion about the amount of data collected resulted in the processing of only the first 200 images (100 ° rotation). The selenium sites could be located using the SHELX suite

(Schneider & Sheldrick, 2002), and the sites refined using AUTO-SHARP (La Fortelle & Bricogne, 1997). The NCS between the two molecules could be identified in the heavy atom sites using O (Jones *et al.*, 1991), and was apparently imperfect preventing electron density averaging. The electron density maps produced after solvent flattening in AUTO-SHARP certainly looked like protein, however the building a sensible main-chain trace seemed impossible. It appeared as if the Se-Met derivative would not deliver the required phasing power. Such derivatives are infrequently referred to as "Kleijwegt" derivatives.

The missing diffraction images were located and all of the data re-processed. Solvent flattening and density modification in SOLOMON (Abrahams & Leslie, 1996), and the use of ARP-wARP (Perrakis *et al.*, 1999) both implemented in AUTO-SHARP, resulted in the tracing of an almost complete main-chain trace from Arg⁷ to Ala⁵⁵⁹ in both molecules of the asymmetric unit. No electron density could be seen for a small loop (Gly³⁴⁰-Gly³⁴⁵) in either of the two molecules. The protein sequence could be easily docked by the manual insertion of residues into the electron density maps from AUTO-SHARP using O. Attempts to use electron density averaging were hampered by an imperfect two-fold non-crystallographic symmetry between the two molecules, as a result domain movements.

The data were reprocessed with MOSFLM (CCP4, 1994; Leslie, 1992) and scaled with SCALA (CCP4, 1994) to 1.7 Å prior to refinement. The model was initially refined in CNS (Brunger *et al.*, 1998) using simulated annealing. Difference maps clearly showed the presence of additional density in the binding site of both molecules, which was later attributed to the cryo-protectant PEG 400. Refinement continued using REFMAC5 (Murshudov *et al.*, 1997), with TLS refinement (Winn *et al.*, 2001) used in the final stages with each domain as a TLS group. Solvent molecules were added using ARP-wARP (Lamzin *et al.*, 2001) and were manually inspected in O. The Se-Met structure was used as a template for molecular replacement, however initial attempts were unsuccessful. The re-processing of the native data using MOSFLM, and scaling in SCALA, resulted in a structural solution. In the native (and substrate complex) electron density for the small loop Gly³⁴⁰-Gly³⁴⁵ was clearly visible, as was an additional C-terminal residue, Glu⁵⁶⁰.

5.4 The Structure of ORF15

5.4.1 Secondary and Tertiary Structure

The structure of ORF15 reveals a bi-lobed, pear shaped molecule with approximate dimensions of 72 Å x 60 Å x 40 Å (Figure 5.1a). A voluminous cavity is formed in between the two lobes (Figure 5.1b), and creates the substrate binding cleft (described in section 5.4.2). The protein is composed of 16 α -helices (α A- α P) and 13 β -strands (β 1- β 13), which form three β -sheets, arranged in three separate domains. The three domains are all members of the α/β -class. Domains I and II form the larger lobe at the "bottom" of the pear shaped molecule. Domain III forms the second lobe, on the "top" of the molecule. The two lobes are linked by an interdomain linker, which permits movements between the two lobes. The structure is characterised by a series of meandering loops link the secondary structure elements, and account for *ca.* 60% of the structure.

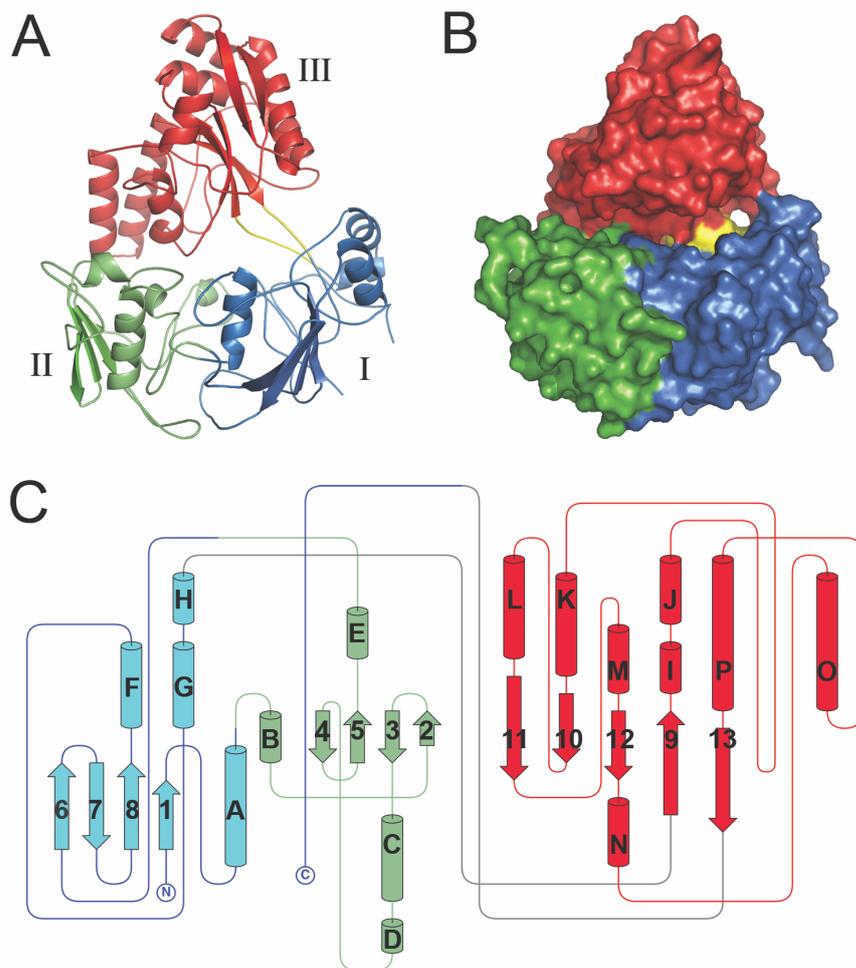


Figure 5.1 Overall structure of ORF15. (A) The tertiary structure of ORF15 with domains indicated with roman numerals and coloured accordingly: domain I (blue), domain II (green) and domain III (red), the interdomain linker region is shown in yellow. (B) Surface projection of ORF15 showing the large cleft formed between the two lobes. (C) Topology diagram showing the arrangement of the secondary structure elements in ORF15, with the same domain colouring scheme as in 5.1a.

A topology diagram indicating the arrangement of the secondary structure elements is shown in Figure 5.1c Domain I consists of a four stranded mixed β -sheet, and four helices, although only helix αA flanks the β -sheet. The domain is formed by three polypeptide segments from separate regions of the sequence (residues 7-64, 214-297, 533-561), and begins with $\beta 1$ and αA . After helix αA the chain trace departs from domain I to form domain II (residues 65-213), which is composed of two β -hairpins (βB , βC , βD), and four α -helices (αB – αE). The β -strands form an antiparallel β -sheet, which is only flanked by one helix, αC . From domain II, the chain returns to domain I to form the three additional antiparallel β -strands,

($\beta 6$ – $\beta 8$), completing the mixed β -sheet. The three additional helices (αF – αH) of domain I are also formed, before the chain trace departs domain I to form the first of two inter-domain linkers (residues 298–302) linking domains I and III. Domain III (residues 303–529) is composed of five β -strands ($\beta 9$ – $\beta 13$) creating a mixed β -sheet, which is surrounded by seven α -helices (αI – αP). The β -sheet is flanked by helices αJ and αK on one side, and αM , αN and αP on the other, while αK and αL lie over the top of the β -sheet. The chain trace departs domain III to form the second inter-domain linker region (residues 530–532) before returning to complete domain I.

5.4.2 Structural Similarity to Other SBPs

Previous sequence analysis (Mellado *et al.*, 2002) revealed similarity to bacterial SBPs, in particular family five SBPs (Tam & Saier, 1993) which include oligopeptide-binding proteins. A search of the structural database using DALI (Holm & Sander, 1995) and DEJAVU (Kleywegt & Jones, 1997) confirmed that ORF15 is structurally related to the family five SBPs (Table 5.1). These include: the dipeptide-binding protein DppA from *E. coli* (1dpe), and the oligopeptide-binding proteins, OppA from *Salmonella typhimurium* (1rkm) and AppA from *Bacillus subtilis* (1xoc), and the nickel binding protein, NikA from *E. coli* (1zlq). ORF15 also showed similarity to several structures obtained in structural genomics projects, including a chitin-binding protein from *Vibrio cholerae* (1zty), the lipoprotein LpqW from *Mycobacterium* (2grv), Ylib from *E. coli* (1uqw), and the periplasmic oligopeptide-binding protein (TM1223) from *T. Maritima* (1vr5). While there is structural information for these proteins, a defined function may not yet have been demonstrated.

Table 5.1 Sequence and structural similarity between ORF15 and some representative SBPs of known structure. A dash line delineates the SBPs possessing three domains, and those possessing two domains.

Pdb	S.I. ^a	S.S. ^b	R.M.S.D. ^c	Lali ^d	Lseq2 ^e
1dpe	17.6%	38%	3.9 Å	474	507
1rkm	18.9%	41.5%	3.9 Å	483	517
1xoc	17.3%	37.7%	3.9 Å	468	504
1zlq	17.5%	36.6%	3.2 Å	465	496
1zty	16.3%	35.2%	4.9 Å	466	528
2grv	18.9%	34.0%	5.2 Å	452	516
1lst	10.3%	18.5%	4.5 Å	83	239
1r9l	12.5%	24.2%	3.5 Å	92	309

^a percentage sequence identity.

^b percentage sequence similarity.

^c positional r.m.s.d of superimposed C α atoms from the Dali search.

^d total number of equivalence residues.

^e length of the entire chain of the equivalent structure.

ORF15 shares very little structural similarity with the smaller periplasmic SBPs, *e.g.* the lysine/arginine/ornithine-binding protein (LAO) from *Salmonella typhimurium* (1lst), and is due to the absence of domain II in these proteins. The

sequence similarity between ORF15 and these proteins is quite low. Some similarity can be found between domain III of ORF15 and domain II of the betaine binding protein, ProX from *E. coli* (1r9l). The global comparative statistics can be improved by the alignment of individual domains. The arrangement of the three domains in ORF15 and the other family five SBPs is essentially the same, although the architecture and folding of the main-chain trace comprising the respective domains does differ.

5.4.3 Domain Movements in ORF and Other Related SBPs

Although many SBPs are co-purified with a ligand this was not the case for the recombinant ORF15. However, additional electron density could be seen in the cleft for both molecules in the Se-Met structure. This extra density has been attributed to the cryo-protectant PEG 400, although the possibility that a PEG molecule from the precipitant solution has co-crystallized with ORF15 cannot be excluded. The binding of PEG molecules in the cleft gave an initial indication as to the location of the substrate-binding site. Furthermore, the native structure revealed several glycerol molecules also bound within the binding cleft, in close proximity to the arginine binding site (see section 5.4.5).

Solute-binding proteins capture their respective ligands in the cleft formed by the two lobes. The binding of ligands by SBPs is associated with domain movements. Lobe 1 closes over lobe 2 trapping the ligand in the binding cleft, and thus SBPs can be in either an "open" (unligated) conformation, or a "closed" (ligated) conformation. For example, the dipeptide binding protein (DppA), undergoes a 55 ° rigid body rotation when moving between the open and closed conformations (Dunten & Mowbray, 1995; Nickitenko *et al.*, 1995), whereas the oligopeptide binding protein (OppA) undergoes a 26 ° rigid body rotation (Sleigh *et al.*, 1997; Tame *et al.*, 1994). In their closed conformations both DppA and OppA are essentially equally "closed", and therefore the difference in rotation, *i.e.* 55° vs 26 ° reflects the relative degree of "openness" in the crystal structures. Such conformational changes have been likened to a Venus fly trap (Quijcho & Ledvina, 1996), and firmly ensconce the ligand within the SBP preventing its escape. The domain movements associated with substrate binding also offer a means by which the ABC transport machinery can recognize the "loaded" protein. The rigid-body rotations between the lobes are mediated by changes in the ϕ and ψ angles of residues in the two inter-domain regions linking the two lobes.

A C α alignment of the ORF15 native (ORF15_{NAT}) structure with the ORF15 Se-Met "A molecule" (ORF15_{MSEA}) and ORF15 Se-Met "B-molecule" (ORF15_{MSEB}) revealed a small domain movements up to 14.3° between the structures (Figure 5.2). ORF15_{MSEB} is most likely to represent the most closed conformation of ORF15. While the electron density for several of the side-chains lining the binding cleft is marginal, it does appear that further closure of the binding cleft is unlikely. Several close van der Waals contacts are formed between the two lobes for example Tyr⁵³ and Ala⁴²⁸, Tyr⁵² and Trp⁴⁸⁰, Gln²⁷⁹ and Lys⁴⁰¹, His²⁸⁰ and Gly³³⁹, Gln²⁸³ and Gln³⁴³. The disorder observed in the small loop of domain III (Gly³⁴⁰-Gly³⁴⁵) is also probably as a result of the loops close proximity to domain I.

Apart from the relatively small rigid-body domain movements, the side-chain of Trp⁵⁷ is displaced upon binding of PEG, or arginine (section 5.4.5). As a result of this "Trp flip" the side chain of Gln¹⁸⁹ is also displaced. These two residues remain undisturbed in the native structure, despite the presence of glycerol in the binding cleft. The binding of PEG also results in a planar rotation of the Tyr⁵² side-chain, compared to ORF15_{NAT} (and the arginine complex), although this is most likely the result of domain movements, rather than directly related to the binding of PEG.

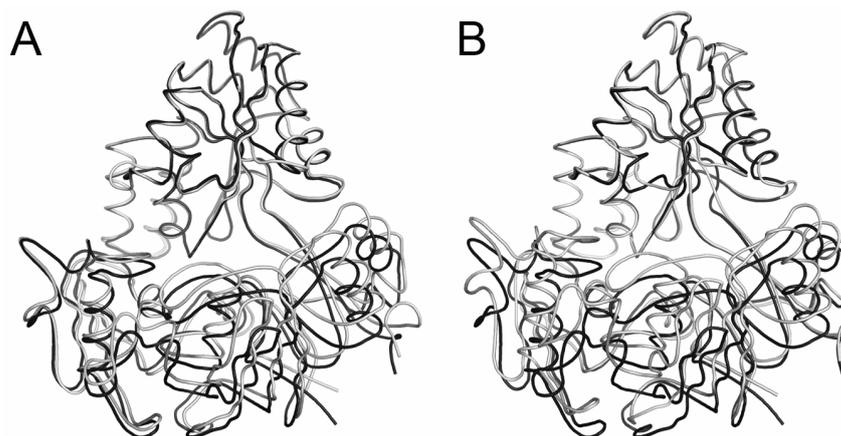


Figure 5.2 Domain movements between the various molecules of ORF15. (A) C α alignment between the ORF15_{NAT} (black) and ORF15_{MSEA} (grey) structures, where lobe 1 rotates 8.9° towards lobe 2. (B) C α alignment of ORF15_{NAT} (black) and ORF15_{MSEB} (grey) where lobe 1 moves 14.3° towards lobe 2.

Given that there is only a *ca.* 14° rotation of lobe 1 onto lobe 2 this would suggest that the "open" conformation of ORF15_{NAT} is already in a predominantly closed state. A comparison of ORF15 with other SBPs in their open and closed conformations (see Figure 5.3) revealed that ORF15_{NAT} is more closed than DppA, and more similar to OppA. On the other hand, the closed conformation of ORF15_{MSEB} is more open than the closed conformations of DppA and OppA. Interestingly the lipoprotein LpqW from *M. tuberculosis* appears to have crystallized in an already "closed" conformation. A ₃₁₀ helix in the binding groove appears to prevent further closure of the binding cleft, and an inflexible inter-domain linker region is likely to impair a more open conformation (Marland *et al.*, 2006). In ORF15 there do not appear to be any factors that would hinder the flexibility of a inter-domain region, and it is likely that the relatively closed conformation of ORF15_{NAT} is the result of the packing of the molecules in the crystal.

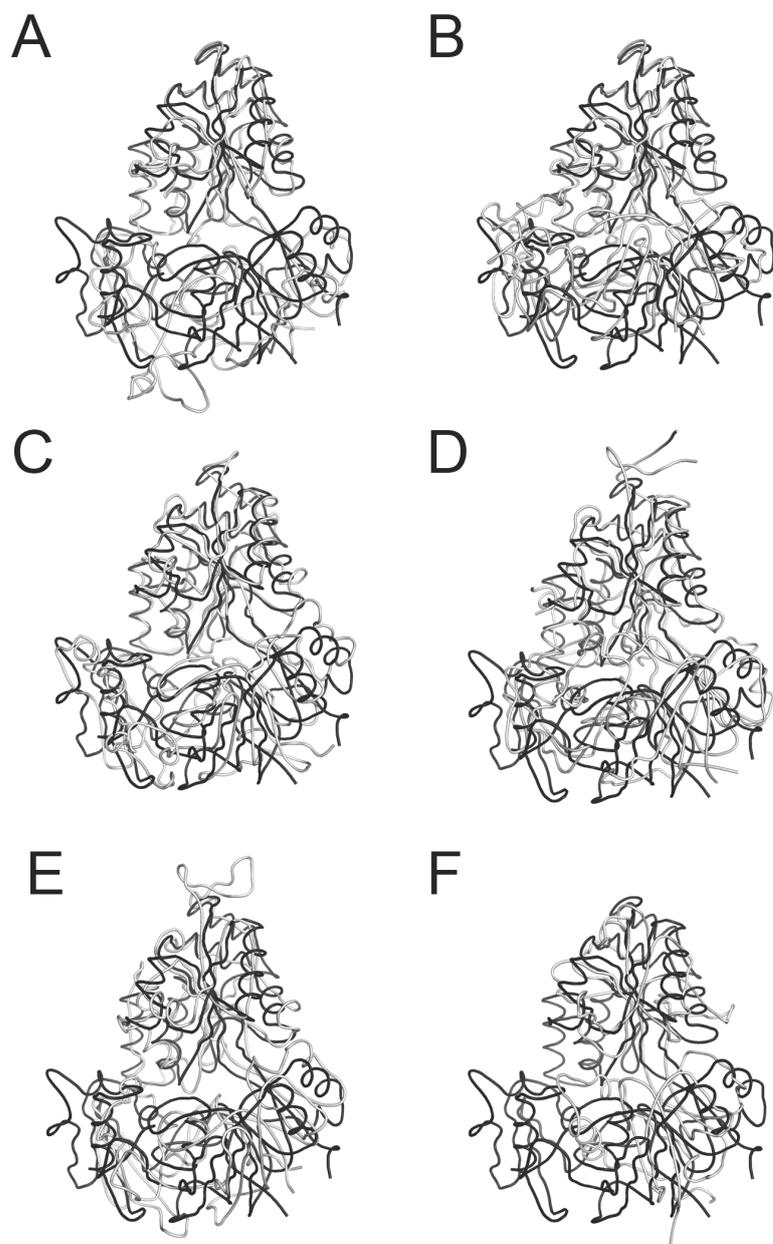


Figure 5.3 Structural similarities between ORF15 (black) and other SBPs (grey) in open and closed conformations (OC and CC respectively). Statistics refer to the alignment of C α atom for domain III. (A) ORF15 and DppA OC (1DPE, r.m.s.d. = 1.62 Å for 200 atoms). (B) ORF15 and DppA CC (1DPP). (C) ORF15 and OppA OC (1RKM, r.m.s.d. = 1.68 Å for 187 atoms). (D) ORF15 and LpqW CC (2GRV, r.m.s.d. = 1.83 Å for 178 atoms). (E) ORF15 and Nika OC (1UIU, r.m.s.d. = 1.78 Å over 181 atoms). (F) ORF15 and ProX CC (1R9L).

5.4.4 Crystal Soaks with Potential Substrates

Orf15 is likely to be expressed in a polycistronic transcript along with *orf16–18* (Mellado *et al.*, 2002). *Orf18* encodes a PBP (PbpA), which possesses a reduced binding capacity for β -lactam antibiotics and clavulanic acid (Ishida *et al.*, 2006), forming an important part of the self-resistance mechanism. Furthermore, *Orf14* is located directly downstream of *Orf15*. *Orf14* encodes a putative acetyltransferase, sharing sequence similarity with an acetyltransferase from *Pseudomonas syringae*, which is involved in tabtoxin resistance, and may be an important factor in conferring antibiotic resistance in *S. clavuligerus*. In order to see if ORF15 was potentially involved in the binding/transport of, or resistance to, clavulanic acid, the β -lactamase inhibitor was soaked into crystals at 50 mM for *ca.* 1 minute.

The production of NAG-clavam, in the gene disruption mutants of *orf15* suggests that ORF15 is somehow involved in the production of this 5*S*-clavam. A potential substrate analogue of NAG-clavam, N-acetyl-ornithine-glycine, was soaked into crystals at 25 mM for *ca.* 1 minute.

Previous sequence analysis together with the observations of Lorenzana *et al.*, (2004) suggests that ORF15 is a potential oligopeptide-binding protein. To test the proposed role of ORF15 as an unspecific oligopeptide-binding protein the oligopeptides Gly₄ and Gly₆ were soaked into crystals (20 mM for 5 minutes), as were the amino acids phenylalanine, and proline.

The structures of the complexes with potential substrates described above revealed no electron density for their respective ligands. Furthermore, no conformational changes were observed in the side-chains of Trp⁵⁷ or Gln¹⁸⁹, unlike that seen in the Se-Met structure. The cryo-protectant glycerol could clearly be seen in the binding cleft, occupying positions similar to those observed in the native structure.

5.4.5 Arginine Binding and the Binding Cleft

The amino acid arginine, a precursor of clavulanic acid, was soaked into crystals, and a data set collected on a crystal diffracting to *ca.* 1.6 Å. The structure was solved by molecular replacement and revealed arginine bound in a negatively charged pocket in the cleft between the two lobes. The arginine was bound in close proximity to the intersection of the three domains, and did not result in any significant domain movements, with an r.m.s.d. of 0.13 Å between the ORF15_{ARG} and ORF15_{NAT} structures, consistent with the proposed "closed" conformation in the crystal. This is in contrast to the binding of arginine by LAO (Oh *et al.*, 1994), where arginine is firmly encapsulated within the binding cleft by a 52 ° rotation between the two domains. Furthermore, LAO forms numerous hydrogen bonds with arginine, compared to ORF15 where there are relatively few direct hydrogen bonding interactions between arginine and the protein (Figure 5.4a).

The carboxyl oxygen of arginine forms a hydrogen bond to the main-chain nitrogen of Gly⁴⁵⁴, and a nearby solvent acetate ion, which also forms interacts with the hydroxyl oxygen. The positively charged amide nitrogen of arginine interacts with

Asp⁴⁵⁶, an interaction frequently seen in oligopeptide binding proteins in order to stabilise the amide charge. The charge is further stabilised by the interaction of the amide with the main-chain carbonyl of Gly⁴⁵⁴ and through the interaction of the π -electrons from Trp⁴⁵³. These interactions are mediated through the residues in domain III, while the positively charged guanidinium moiety is co-ordinated through the residues on lobe 1. The binding of the arginine guanidinium group requires the displacement of the Trp⁵⁷ side-chain, which in turn displaces the Gln¹⁸⁹ side-chain. The charge on the NH2 nitrogen is stabilized by the interaction with the π -electrons of Trp⁵⁷. The NH1 nitrogen is hydrogen bonded to the carbonyl oxygen of Ala⁴⁵⁵ and Ala⁴⁵⁷ *via* a solvent bridge (W¹ and W² respectively). The NH1 nitrogen forms additional interactions with nearby glycerol (shown in Figure 5.4b), which occupies an invagination extending from the binding cleft, towards the core of lobe 1, nestled between domains I and II. A second invagination can be found in the opposite direction, nestled between domains II and III. The NE nitrogen interacts with the main-chain carbonyl of Tyr⁵².

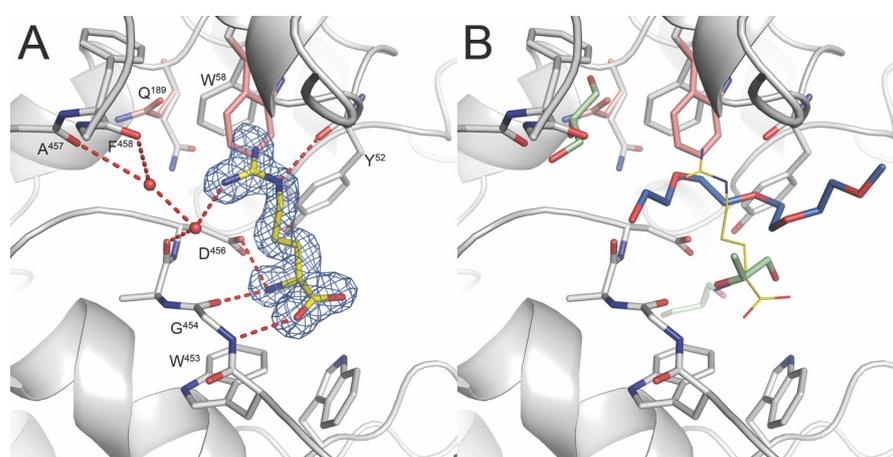


Figure 5.4 The binding of various ligands in cleft of ORF15. (A) The binding of L-arginine. The $2mF_{obs}-DF_{calc}$ electron density is shown in blue, and is contoured at 1σ (where σ is the root mean square deviation for the electron density for the unit cell). Hydrogen bonding interactions are shown as red dashes, and waters as red spheres. The interactions with a nearby glycerol and acetate molecule are not shown. (B) A superposition of additional ligands bound in the ORF15 binding cleft. Part of a PEG 400 molecule (blue) bound in ORF15_{MSEA} and three glycerol molecules (green) bound in ORF15_{NAT}. The arginine is shown for reference.

The interaction of arginine with tryptophan side-chains is reminiscent of the "tryptophan box" found in the glycine betaine-binding protein, ProX from *E. coli* (Schiefner *et al.*, 2004). In ProX, three tryptophan residues are arranged in close proximity creating a "Trp box", such that the π -electrons interact, and stabilize, the charge of the ligand. In ORF15 three tryptophan residues (Trp⁵⁷, Trp⁴⁵³, and Trp⁴⁸⁰) can be found in the binding cleft, although only two appear to participate in the binding of arginine. In ORF15 the tryptophan residues are more spatially separated, forming less of a "Trp box", but contributing to a "Trp trap" with which to bind arginine.

In the ORF15_{NAT} structure a glycerol could be identified in two conformations bound in close proximity to the position occupied by arginine (Figure 5.4b). A glycerol molecule can also be found at an additional position in the binding cleft, away from the arginine binding site, in a positively charged area of the binding cleft.

5.5 The Function of ORF15

At the beginning of this chapter I posed a question. Now that we know the structure of ORF15, do we know its function? Unfortunately the answer is no, but the structure does give some important information, and clues as to a potential function.

In trying to understand the function of ORF15 we first need to consider its location in the cell. In Gram-negative bacteria SBPs are secreted into the periplasm space, which is located in between the inner and outer membranes. Potential substrates, such as oligopeptides, enter the periplasm through porin proteins in the outer membrane, and are subsequently bound by the SBPs. The binding of the substrate by the SBP results in a conformational change between the two lobes, which assists the ABC transporters (permeases) recognise the loaded SBPs. The permease then actively imports the substrate into the cell, where ATP is hydrolysed in the process. In Gram-positive bacteria, SBPs must first cross the cell wall, before being tethered to the cell wall in close proximity to the integral components of the transport system. In both cases the protein requires an N-terminal signal peptide indicating that it should be secreted outside the cell. In the case of Gram-positive bacteria the SBP also requires a means by which it can be attached to the outer membrane, either by a trans-membrane N-terminal helix, or by a lipid attachment site (PROSITE PDOC00013). In the latter the secretion signal peptide is cleaved, and the protein is covalently attached to a lipid molecule via a cysteine residue.

Sequence analysis reveals that ORF15 lacks an N-terminal signal peptide for secretion *via* the Sec pathway (Bendtsen *et al.*, 2004). Furthermore, ORF15 lacks an N-terminal trans-membrane helix, lipid attachment motif, or sufficiently large hydrophobic patch with which the protein could interact with the lipids of the outer cell wall. Folded proteins can be secreted from the cell *via* the twin arginine translocase (Tat) pathway (Berks *et al.*, 2003), which possesses an N-terminal signal peptide with a RRxΦΦ motif, where Φ denotes a hydrophobic residue. While ORF15 does possess several twin arginines, in particular R⁶R⁷ and R²⁹R³⁰R³¹ in the N-terminal, it does not appear to possess the entire signal peptide (Bendtsen *et al.*, 2005). Taken together the sequence data suggests that ORF15 is localized in the cytoplasm. However SBPs are not localized within the cytoplasm, they are found either attached to the cell wall or in the periplasmic space. Recent observations using cryo-electron microscopy (Matias & Beveridge, 2005; Matias & Beveridge, 2006) suggest that, at least some, Gram-positive bacteria possess a small “periplasmic space” associated with the cell wall. Although this remains to be determined for *S. clavuligerus*, should such a periplasm exist, then it may be possible that ORF15 is located in this space. This would negate the need for a

means by which it attaches to the intracellular cell wall, although how ORF15 would be transported across the cytoplasmic membrane remains unclear. Since the cellular location of ORF15 requires further investigation, the potential role of ORF15 in either the cytoplasm or potential periplasm/ or cell wall will be discussed.

5.5.1 Is ORF15 a Potential Cytoplasmic "Intermediate Chaperone"?

The binding of arginine by ORF15 is of great interest since arginine is a precursor of clavulanic acid. The presence of three ornithine acetyl transferases, *Oat2* in the clavulanic acid gene cluster (Kershaw *et al.*, 2002), *Oat1* in a paralogous gene cluster (Tahlan *et al.*, 2004b) and *ArgJ* elsewhere in the genome (de la Fuente *et al.*, 2004), suggests that the accumulation of arginine is important in *S. clavuligerus*. It is therefore possible that ORF15 is involved in the procurement, and possible transport, of arginine within the cytoplasm for its subsequent conversion into clavulanic acid and 5S-clavams. Glycerol, which was found in several locations in the binding cleft, bears some resemblance to the other clavulanic acid precursor, D-glyceraldehyde-3-phosphate (G3P). There does not appear to be any volume constraints preventing the binding of G3P, however residues capable of coordinating the negatively charged phosphate group appear to be lacking, at least in the arginine binding site. It is interesting to note that on either side of the negatively charged arginine binding site, there are positively charged surfaces, where a cryo-protectant molecule can be found bound, although this glycerol appears to only be partially bound, as demonstrated by elevated *B*-factors.

Extending this potential role, ORF15 may act as an "intermediate chaperone" assisting the production of clavulanic acid, by preserving the pool of intermediates. Many of the intermediates in clavulanic acid are unstable molecules in solution, in part due to the β -lactam ring. Since arginine does not occupy the entire binding cleft, there may be additional space for the "side-chains" of the first intermediate of the biosynthesis pathway, *e.g.* N^2 -(2-carboxyethyl)-arginine, and deoxyguanidinoproclavaminc acid. In the process of binding of arginine (or PEG) the side chain of Trp⁵⁷ is flipped to an alternate conformation to make sufficient space for the guanidinium group, as well as stabilise one of the positively charged nitrogens. It is therefore possible that other side-chains within the binding cleft may permit the binding of other clavam intermediates, for example NAG-clavam. The potential binding of NAG-clavam requires further investigation, as N-acetyl-ornithine-glycine may be an unsuitable substrate analogue.

It is difficult to reconcile the differences between a complete loss of clavulanic acid production observed in *orf15* mutants, and the proposed role of ORF15 as an intermediate chaperone. Indeed, if ORF15 were to act as a chaperone, then a reduction in clavulanic acid might be expected, due to a reduction in the pool of clavam intermediates. Detectable levels of 5S-clavams have been observed in mutants of *orf15* mutant, *albeit* at reduced levels (Jensen *et al.*, 2004a), although Lorenzana *et al.*, (2004) were unable to detect 5S-alanylclavam. While the production of 5S-clavams is often more varied, if ORF15 acted as an intermediate chaperone a reduction in 5S-clavams would be expected, since both 5S-clavams

and clavulanic acid share a common biosynthesis pathway, at least until clavaminic acid (Egan *et al.*, 1997).

5.5.2 Is ORF15 a Potential Periplasmic Oligopeptide Binding Protein?

The increased resistance to the toxic tripeptide Pt-AA and the inability to utilize a nonapeptide (bradykinin) in the growth medium in mutants of *orf15* (Lorenzana *et al.*, 2004) appears consistent with the proposed role for ORF15 as an oligopeptide-binding protein. The structure of ORF15 is also consistent with this proposed role, sharing structural similarity with other oligopeptide binding proteins.

The size of the peptide substrate bound by oligopeptide-binding proteins, *e.g.* dipeptide *versus* pentapeptide, is dependent on the size of the binding pocket, rather than the sequence of the substrate peptide. Polar side-chains lining the substrate-binding pocket coordinate the hydrogen bonding of the peptide's main-chain atoms. The positive charge of the peptide's N-terminal amide is frequently stabilized by the interaction with an aspartic acid residue, while the C-terminal carboxyl residue is often found interacting with a charged glutamic acid side-chain. Water molecules move in and out of the binding pocket depending on the size, and hydrophobicity, of the peptide's side-chains. The peptides polar side-chains are coordinated *via* solvent bridges to polar residues lining the binding pocket which occupy the substrate-binding cleft.

ORF15 possesses a large binding cleft, potentially capable of binding oligopeptides of several residues. Several aspartic acid residues, can be found within the binding cleft, and other polar residues capable of coordinating peptide main-chain residues are present. The observed binding of arginine within the cleft may explain the apparent import of bradykinin, since the nonapeptide possesses both an N- and C-terminal arginine in its sequence (RPPGFSPFR). The binding of bradykinin is most likely non-specific, since the oligopeptide-binding protein AppA has also been observed to bind longer peptides, including bradykinin (Picon & van Wely, 2001). The attempts to bind the oligopeptides Gly₄ and Gly₆ in ORF15 crystals were unsuccessful, but this does not necessarily exclude a role for ORF15 as an oligopeptide-binding protein. Indeed, the relatively closed conformation of ORF15 in the crystals may restrict the binding of small oligopeptides, but favour the binding of amino acids, although ORF15 does not appear to be a non-specific amino acid binding protein. Furthermore, no ligands were co-purified with ORF15, suggesting that there is some specificity in substrate preference.

Periplasmic and extracellular SBPs associated with the ABC transport system are responsible for the import of nutrients into the bacterial cell; they also participate in additional cellular processes, including the initiation of chemotaxis, and the binding of signalling molecules stimulating morphological development (Koide & Hoch, 1994; Lazazzera & Grossman, 1998; Solomon *et al.*, 1996). For example, a number of genes have been identified in a signalling cascade ultimately leading to the formation of aerial mycelia in *S. coelicolor* A3 (Nodwell *et al.*, 1999). Mutants of the *Bld* genes are unable to develop aerial mycelia, and thus possess a bald

phenotype. The *BldK* operon in *S. coelicolor* A3 encodes five genes (*A–E*) with homology to the ABC transporters (Nodwell *et al.*, 1996). One of these genes, *BldKB* encodes a putative oligopeptide-binding protein involved in signal reception and/or transport. BldKB apparently binds a 665 oligopeptide, containing serine and glycine residues (Nodwell & Losick, 1998), which in turn stimulates a signalling cascade, ultimately leading to the formation of aerial mycelia.

The reported “bald” phenotype associated with the *orf15* mutant (Lorenzana *et al.*, 2004) implies that ORF15 may also be involved in similar processes. Both ORF15 and BldKB share *ca.* 44% sequence similarity, although *BldKB* possesses both a N-terminal signal peptide and lipid attachment motif (PROSITE PDOC00013) associated with extracellular oligopeptide-binding proteins. Furthermore, BldKB possesses a potential ATP/GTP “P-loop” (PROSITE PS00017), which provides a potential means for propagating a signal upon binding of the signalling peptide. While ORF15 lacks such a nucleotide-binding site, it may still be involved in potential signalling pathways. For example, the small changes observed between the open and closed conformations in ORF15 might be significantly larger in solution, and thus permit interaction in signalling pathways *via* Venus fly trap movements, like other SBPs. Peptides accumulating in culture broths have been shown to stimulate antibiotic production in fresh medium cultures (Sanchez & Brana, 1996). Thus it may well be that morphological development and β -lactam antibiotic/ clavulanic acid production is stimulated by oligopeptide signalling molecules. This then raises further question: why are there two apparent oligopeptide-binding proteins in the clavulanic acid gene cluster?

5.5.3 Structural Similarity Between ORF15 and ORF7?

Proteins with a similar sequence often possess a similar structure, where higher sequence similarity/identity implies higher the structural similarity. Such rationale allows for the predicted three-dimensional modelling of a protein structure based on its sequence similarity to known structures. The reliability of such methods are of course dependent on the amount of information available and the degree of sequence similarity identity, with the axiom being “the higher the better”. ORF7 and ORF15 share 48% sequence identity, and 63% sequence similarity. Does the structure of ORF15 tell us anything about the structure of ORF7? Perhaps this is a rhetorical question, since we have not yet predicted a model for ORF7 based on the structure of ORF15.

ORF7 is most likely a bi-lobed, pear shaped molecule, composed of three domains. Two of the tryptophan residues located in the ORF15 binding cleft are conserved in ORF7 (Trp⁵⁷, Trp⁴⁵³), as is the aspartic acid binds the arginine amide nitrogen (Asp⁴⁵⁶). The third tryptophan is replaced by a glutamine, providing a residue with hydrogen bonding potential. A data set has been collected on an ORF7 crystal, and attempts to solve the structure using ORF15 as a molecular replacement template have so far been unsuccessful (K. Valegård, S. Lejon, Per. Comm.). This is most likely due to ORF7 being in an alternate conformation with respect to the two lobes, although other factors may complicate the molecular replacement procedure.

```

1 MTTAARRPAPT TAGAGWDAGV GALVNP SRRRGGTLRLVSSADVDSLDPARTYYVWVLLQ 60
2 ----METTRSTTADEGFDAGVRGVVAF TDAPGGTLRLVRTDDFDSLDPGNTYYAYTWNFL
                                     *
1 RLLNRTLMA YPTDPGPAGLVPAPDLAEGPGEVSDGGGTWTYRLRRGLRYDDGTPITSDDV 120
2 RLIGRTLVTFD TAPGKAGQRLV PDLAESLGE SSEDEGRVWVTYRLREGLRYEDGTPVVSADI

1 RHAVQR -VFAQDVLPGGPTYLIPLLDDPERPYPGPYRTDEPLRSVLTPEHTIVFRLTRP 180
2 KHAIARSNYGTDV L GAGPTYFRHLLG -TEYGGPWREFDADGPVTLETPEDETLVFRLLREP

1 FSDFDHLM AQC CAAPVPRRS DTGADYGRDPRSSGPYRVARHEPDTLLHLERNPHWDRATD 240
2 FAGMDLLATMPST TPVPRDRDTGA EYRLRPVATGPYRIVSYTRGELAVLEPNPHWDPETD

1 EIRPALPDRVELTIGLDVDVLDARLIAGEFDINLEGRGLQHAAQRRADEVLRSHTDNP 300
2 EVRVQRASRIE VHLGKDPHEVDRMLLAGEAHVDLAGFGVQPA AQERILAEPELRAHADNP

1 RTSFLHFVAMOPHI PPFDNVHVRRAVQYAADKILLQDARGGFPVNGGDLTTALFPPTLPAH 360
2 LTGFTWIYCLSSRIAPFDNVHCRRAVQFATDKAAMQEAYGG -AVGGDIATLLPPTLDGY

1 QDDL LYPTGPD LRGDLDAARAELAAGL PDGFRAVIGTQ--RGKFRLVADAVVESLARVG 418
2 KHFD RYPVGP EGTGDL EAARAELKLAGMPDGFRTRIAARKDRLKEYRAAEAL AAGLARVG

1 IELTVKELDVATYFSLGAGHPETVREHGLGLLVTDWGADFPTEYGF LAPLV DGRQIKRNG 478
2 IEAEVLD FPGSDYFD RYGGCPEYLREHGIGIIMFGWGADFPDGYGFLQQITDGRAIKERG
                                     *
1 GNWNLP ELDDPEVNALIDETLHTTDP AARAELWRAVERRVMEHAVLLEPLVHDKTLHFRNP 538
2 -NQNMGE LDDPEINALLDEGAQCADPARRAEIWHRIDQLTMDHAVIVPYLYPRSLLYRHP
                                     *
1 WVTNVYVHPAFGLYDIQAMGLAEED
2 DTRNAFVTGSFGMYDYVALGAK---

```

Figure 5.5 The sequence similarity between ORF15 (1) and ORF7 (2), where identical residues are highlighted in grey. A (*) indicates the position of the three tryptophan residues in the ORF15 binding cleft.

No additional ligands were co-purified with ORF15, which suggests that there is some relative specificity in substrate preference. The opposite appears to be true for ORF7, which is purified with additional, as yet unidentified ligands (S. Lejon, Per. Comm.). This suggests that ORF15 is a substrate specific oligopeptide-binding protein, while ORF7 is more unspecific. Furthermore, while ORF15 contains twin arginines in its N-terminal sequence, ORF7 does not. Is it possibly that ORF7 is localised in the cytoplasm, and that ORF15 is transported across the cell *via* the Tat pathway?

6. Future Prospectives

6.1 DAOCS

The initial task of altering the packing of protein molecules in the crystal using a His₍₆₎ tag was successful. The crystals of the His-tagged protein showed a different crystal packing, resulting in a non-twinned crystal. Unfortunately, the increased solvent content of the crystals probably contributed to a significant reduction in the diffraction of the crystals. The new crystals did not diffract at sufficient resolution to detect the small differences caused during ferryl formation. The addition of different purification tags in order to disrupt the packing of the protein molecules in the crystal is still an interesting option. Using the techniques developed in high through-put crystallography it would be possible to screen a large number of different constructs simultaneously. Perhaps the hardest decision boils down to the choice of which purification tag to add. Both the His-tagged DAOCS, and Trx-His-tagged CAD provide a nice example of how purification tags can affect the crystallization (or lack there of) of a protein.

The challenge of trying to capture the ferryl intermediate still remains for any soul brave enough to accept. It may still be possible to use the twinned *R3* crystals using the philosophy that sheer force by numbers will eventually identify crystals with a sufficiently low twin fraction. Given the recent advances in automation for both crystallization experimental set up and screening of crystals at synchrotrons, it should not be too long before an armchair crystallographer can sit at home, and click their way through a thousand crystals grown essentially a robot. Of course this is little conciliation to the biochemist who still has to purify DAOCS using the extended protocol, or the joy of preparing crystals under anaerobic conditions.

6.2 CAD

Now that we have the structure of CAD can we expect an intensive project of protein engineering to improve the enzymes performance? From an industrial point of view this is unlikely to occur in the near future. However, in the long term it is possible to envisage engineering of the CAD binding site. If the mechanism by which the 3S,5S-clavams undergo enantiomerisation can be identified, and the process industrialised, then a plethora of new potential β -lactamase inhibitors could be made. By the modification of the CAD binding site, new potential substrates could be accommodated and reduced to more stable alcohols.

From a purely academic point of view the structure of CAD is interesting. The observed binding of clavulanic acid in an eclipsed conformation within the binding site was unexpected, and rationalises previous labelling experiments. Furthermore, the potential mechanisms by which CAD protects the labile substrate are of great interest, given the instability of many of the clavam intermediates.

6.3 ORF15

Where to even begin? The structure has revealed some tantalizing hints as to the function of ORF15, but no clear role in clavulanic acid biosynthesis has yet emerged. The two proposed roles require further investigation using a combination of different techniques. The location of both ORF7 and ORF15 in the cell will be important in determining the function of the proteins. The observed "Trp flip" seen in the binding of arginine or PEG can be used to screen the binding of potential substrates using tryptophan fluorescence. Although the signal may be diluted by the other eight tryptophans, potential intermediates could be added to the protein in solution, and their binding monitored *via* the changes in fluorescence. Furthermore, our Oxford collaborators have expressed and purified CEAS, BLS, PAH and CAS and mixed them *in vitro* with glyceraldehyde-3-phosphate and arginine in an attempt to produce clavaminic acid, with reasonable success. It would be interesting to add ORF15 to the enzyme cocktail and see if the clavaminic acid production was improved. Alternatively, ORF15 could be immobilized on an Ni-NAT column and fractions isolated from lysed *S. clavuligerus* cells wash through the column. The protein could be eluted and ligand binding monitored using Mass-spectrometry.

Comparing the structure of ORF15 with ORF7 is an exciting prospect. How similar are the structures? From outward appearances ORF7 looks to be an unspecific peptide-binding protein. How do the binding sites differ? Why is ORF7 less selective than ORF15? Where are the ORFs located in the cell? Why are there no apparent ABC permease genes in the clavulanic acid gene cluster? So many questions remain.

7 References

- Abouhamad, W.N., Manson, M., Gibson, M.M. & Higgins, C.F. 1991. Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella typhimurium*: characterization of the dipeptide permease (*Dpp*) and the dipeptide-binding protein. *Mol. Microbiol.* **5**, 1035-1047.
- Abraham, E.P. & Chain, E. 1940. An enzyme from bacteria able to destroy penicillin. *Nature* **146**, 837.
- Abrahams, J.P. & Leslie, A.G. 1996. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D Biol. Crystallogr.* **52**, 30-42.
- Alphey, M.S., Yu, W., Byres, E., Li, D. & Hunter, W.N. 2005. Structure and reactivity of human mitochondrial 2,4-dienoyl-CoA reductase: enzyme-ligand interactions in a distinctive short-chain reductase active site. *J. Biol. Chem.* **280**, 3068-3077.
- Ames, G.F. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. *Annu. Rev. Biochem.* **55**, 397-425.
- Arulanantham, H., Kershaw, N.J., Hewitson, K.S., Hughes, C.E., Thirkettle, J.E. & Schofield, C.J. 2006. ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid. *J. Biol. Chem.* **281**, 279-287.
- Backus, M.P. & Stauffer, J.F. 1955. The production and selection of a family of strains in *Penicillium chrysogenum*. *Mycologia* **47**, 429-463.
- Baggaley, K.H., Brown, A.G. & Schofield, C.J. 1997. Chemistry and biosynthesis of clavulanic acid and other clavams. *Nat. Prod. Rep.* **14**, 309-333.
- Baldwin, J.E., Goh, K.-C. & Schofield, C.J. 1994. Biosynthetic Precursors of Valclavam. *Tetrahedron Lett.* **35**, 2779-2782.
- Baldwin, J.E., Lloyd, M.D., Wha-Son, B., Schofield, C.J., Elson, S.W., Baggaley, C.H. & Nicholson, N.H. 1993. A substrate analogue study on clavaminic acid synthase: possible clues to the biosynthetic origin of proclavaminic acid. *J. Chem. Soc., Chem. Comm.*, 500-502.
- Bayles, K.W. 2000. The bactericidal action of penicillin: new clues to an unsolved mystery. *Trends Microbiol.* **8**, 274-278.
- Bendtsen, J.D., Nielsen, H., von Heijne, G. & Brunak, S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783-795.
- Bendtsen, J.D., Nielsen, H., Widdick, D., Palmer, T. & Brunak, S. 2005. Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**, 167.
- Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M.A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. & Hopwood,

- D.A. 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.
- Berglund, G.I., Carlsson, G.H., Smith, A.T., Szoke, H., Henriksen, A. & Hajdu, J. 2002. The catalytic pathway of horseradish peroxidase at high resolution. *Nature* **417**, 463-468.
- Berks, B.C., Palmer, T. & Sargent, F. 2003. The Tat protein translocation pathway and its role in microbial physiology. *Adv. Microb. Physiol.* **47**, 187-254.
- Bignell, D.R., Tahlan, K., Colvin, K.R., Jensen, S.E. & Leskiw, B.K. 2005. Expression of *ccaR*, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus*, is dependent on *bldG*. *Antimicrob. Agents. Chemother.* **49**, 1529-1541.
- Bignell, D.R., Warawa, J.L., Strap, J.L., Chater, K.F. & Leskiw, B.K. 2000. Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. *Microbiology* **146** (Pt 9), 2161-2173.
- Bottoms, C.A., Smith, P.E. & Tanner, J.J. 2002. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.* **11**, 2125-2137.
- Brakhage, A.A. 1998. Molecular regulation of beta-lactam biosynthesis in filamentous fungi. *Microbiol. Mol. Biol. Rev.* **62**, 547-585.
- Brewer, S.J., Taylor, P.M. & Turner, M.K. 1980. An adenosine triphosphate-dependent carbamoylphosphate-3-hydroxymethylcephem O-carbamoyltransferase from *Streptomyces clavuligerus*. *Biochem J* **185**, 555-564.
- Brotzu, G. 1948. Recherche su di un nuovo antibiotico. *Lavori dell'Istituto d'Igiene di Cagliari*, 1-11.
- Brown, A.G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J.D., Reading, C. & Rolinson, G.N. 1976. Naturally-occurring beta-lactamase inhibitors with antibacterial activity. *J. Antibiot. (Tokyo)* **29**, 668-669.
- Brown, A.G., Corbett, D.F., Goodacre, J., Harbridge, J.B., Howarth, T.T., Ponsford, R.J., Stirling, I. & King, T.J. 1984. Clavulanic acid and its derivatives. Structure elucidation of clavulanic acid and the preparation of dihydroclavulanic acid, isoclavulanic acid, esters and related oxidation products. *J. Chem. Soc., Perkin Trans. I*, 635-650.
- Brown, R.P., Aplin, R.T. & Schofield, C.J. 1996. Inhibition of TEM-2 beta-lactamase from *Escherichia coli* by clavulanic acid: observation of intermediates by electrospray ionization mass spectrometry. *Biochemistry* **35**, 12421-12432.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. & Warren, G.L. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921.
- Burzlaff, N.I., Rutledge, P.J., Clifton, I.J., Hensgens, C.M., Pickford, M., Adlington, R.M., Roach, P.L. & Baldwin, J.E. 1999. The reaction cycle of isopenicillin N synthase observed by X-ray diffraction. *Nature* **401**, 721-724.

- Byford, M.F., Baldwin, J.E., Shiau, C.Y. & Schofield, C.J. 1997. The Mechanism of ACV Synthetase. *Chem Rev* **97**, 2631-2650.
- Caines, M.E., Elkins, J.M., Hewitson, K.S. & Schofield, C.J. 2004. Crystal structure and mechanistic implications of N²-(2-carboxyethyl)arginine synthase, the first enzyme in the clavulanic acid biosynthesis pathway. *J. Biol. Chem.* **279**, 5685-5692.
- CCP4 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763.
- Challis, G.L. & Hopwood, D.A. 2003. Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc. Natl. Acad. Sci. U.S.A* **100**, 14555-14561.
- Chen, C.C. & Herzberg, O. 1992. Inhibition of beta-lactamase by clavulanate. Trapped intermediates in cryocrystallographic studies. *J. Mol. Biol.* **224**, 1103-1113.
- Chen, D.Z., Patel, D.V., Hackbarth, C.J., Wang, W., Dreyer, G., Young, D.C., Margolis, P.S., Wu, C., Ni, Z.J., Trias, J., White, R.J. & Yuan, Z. 2000. Actinonin, a naturally occurring antibacterial agent, is a potent deformylase inhibitor. *Biochemistry* **39**, 1256-1262.
- Chen, Z., Jiang, J.C., Lin, Z.G., Lee, W.R., Baker, M.E. & Chang, S.H. 1993. Site-specific mutagenesis of *Drosophila* alcohol dehydrogenase: evidence for involvement of tyrosine-152 and lysine-156 in catalysis. *Biochemistry* **32**, 3342-3346.
- Chin, H.S. & Sim, T.S. 2002. C-terminus modification of *Streptomyces clavuligerus* deacetoxycephalosporin C synthase improves catalysis with an expanded substrate specificity. *Biochem. Biophys. Res. Commun.* **295**, 55-61.
- Clifton, I.J., Doan, L.X., Sleeman, M.C., Topf, M., Suzuki, H., Wilmouth, R.C. & Schofield, C.J. 2003. Crystal structure of carbapenem synthase (CarC). *J. Biol. Chem.* **278**, 20843-20850.
- Clifton, I.J., Hsueh, L.C., Baldwin, J.E., Harlos, K. & Schofield, C.J. 2001. Structure of proline 3-hydroxylase. Evolution of the family of 2-oxoglutarate dependent oxygenases. *Eur. J. Biochem.* **268**, 6625-6636.
- Cohen, M.L. 2000. Changing patterns of infectious disease. *Nature* **406**, 762-767.
- Coque, J.J., Perez-Llarena, F.J., Enguita, F.J., Fuente, J.L., Martin, J.F. & Liras, P. 1995. Characterization of the cmcH genes of *Nocardia lactamdurans* and *Streptomyces clavuligerus* encoding a functional 3'-hydroxymethylcephem O-carbamoyltransferase for cephamycin biosynthesis. *Gene* **162**, 21-27.
- Crowfoot, D., Bunn, C.W., Rogers-Low, B. & Turner-Jones, A. 1949. *X-ray Crystallographic Investigation of the Structure of Penicillin* Princeton University press.
- de la Fuente, A., Lorenzana, L.M., Martin, J.F. & Liras, P. 2002. Mutants of *Streptomyces clavuligerus* with disruptions in different genes for clavulanic acid biosynthesis produce large amounts of holomycin: possible cross-regulation of two unrelated secondary metabolic pathways. *J. Bacteriol.* **184**, 6559-6565.
- de la Fuente, A., Martin, J.F., Rodriguez-Garcia, A. & Liras, P. 2004. Two proteins with ornithine acetyltransferase activity show different functions in

- Streptomyces clavuligerus*: Oat2 modulates clavulanic acid biosynthesis in response to arginine. *J. Bacteriol.* **186**, 6501-6507.
- Demain, A.L. & Elander, R.P. 1999. The beta-lactam antibiotics: past, present, and future. *Antonie Van Leeuwenhoek* **75**, 5-19.
- Demain, A.L. & Fang, A. 2000. The natural functions of secondary metabolites. *Adv. Biochem. Eng. Biotechnol.* **69**, 1-39.
- Diddens, H., Zahner, H., Kraas, E., Gohring, W. & Jung, G. 1976. On the transport of tripeptide antibiotics in bacteria. *Eur. J. Biochem.* **66**, 11-23.
- Duax, W.L., Griffin, J.F. & Ghosh, D. 1996. The fascinating complexities of steroid-binding enzymes. *Curr. Opin. Struct. Biol.* **6**, 813-823.
- Dunten, P. & Mowbray, S.L. 1995. Crystal structure of the dipeptide binding protein from *Escherichia coli* involved in active transport and chemotaxis. *Protein Sci.* **4**, 2327-2334.
- Egan, L.A., W., B.R., Iwata-Reuyl, D. & Townsend, C.A. 1997. Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. *J. Am. Chem. Soc.*, **119**, 2348-2355.
- Elander, R.P. 2003. Industrial production of beta-lactam antibiotics. *Appl. Microbiol. Biotechnol.* **61**, 385-392.
- Elkins, J.M., Clifton, I.J., Hernandez, H., Doan, L.X., Robinson, C.V., Schofield, C.J. & Hewitson, K.S. 2002. Oligomeric structure of proclavaminic acid amidino hydrolase: evolution of a hydrolytic enzyme in clavulanic acid biosynthesis. *Biochem. J.* **366**, 423-434.
- Elkins, J.M., Kershaw, N.J. & Schofield, C.J. 2005. X-ray crystal structure of ornithine acetyltransferase from the clavulanic acid biosynthesis gene cluster. *Biochem. J.* **385**, 565-573.
- Elson, S.W., Gillett, J., Nicholson, N.H. & Tyler, J.W. 1998. N-acyl derivatives of clavaminic acid produced by a mutant of *Streptomyces clavuligerus*. *J. Chem. Soc., Chem. Comm.*, 979-980.
- Filling, C., Berndt, K.D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jornvall, H. & Oppermann, U. 2002. Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.* **277**, 25677-25684.
- Fisher, J., Charnas, R.L. & Knowles, J.R. 1978. Kinetic studies on the inactivation of *Escherichia coli* RTEM beta-lactamase by clavulanic acid. *Biochemistry* **17**, 2180-2184.
- Fleming, A. 1929. On the antibacterial action of *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *Brit. J. Exp. Pathol.* **10**, 226-236.
- Freduenreich 1888. De l'antagonisme et de l'immunité, qu'il confère au milieu de culture. *Annales de l'Institut Pasteur* **2**, 200-206.
- Fulston, M., Davison, M., Elson, S.W., Nicholson, N.H., Tyler, J.W. & Woroniecki, S.R. 2001. Clavulanic acid biosynthesis; the final steps. *J. Chem. Soc., Perkin Trans. I*, 1122-1130.
- Ghosh, D., Wawrzak, Z., Weeks, C.M., Duax, W.L. & Erman, M. 1994. The refined three-dimensional structure of 3 α , 20 β -hydroxysteroid dehydrogenase and possible roles of the residues conserved in short-chain dehydrogenases. *Structure* **2**, 629-640.

- Grimm, C., Maser, E., Mobus, E., Klebe, G., Reuter, K. & Ficner, R. 2000. The crystal structure of 3 α -hydroxysteroid dehydrogenase/carbonyl reductase from *Comamonas testosteroni* shows a novel oligomerization pattern within the short chain dehydrogenase/reductase family. *J. Biol. Chem.* **275**, 41333-41339.
- Hegg, E.L. & Que, L., Jr. 1997. The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **250**, 625-629.
- Higgins, C.E. & Kastber, R.E. 1971. *Streptomyces clavuligerus* sp. nov., a \square beta-lactam antibiotic producer. *Int. J. Syst. Evol. Microbiol.*, 326-331.
- Higgins, C.F. 1992. ABC transporters: from microorganisms to man. *Ann. Rev. Cell. Biol.* **8**, 67-113.
- Holm, L. & Sander, C. 1995. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478-480.
- Howarth, T.T., Brown, A.G. & King., T.J. 1976. Clavulanic acid, a novel β -lactam isolated from *Streptomyces clavuligerus*; X-ray crystal structure analysis. *J. Chem. Soc., Chem. Comm.*, 266-267.
- Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. 2003. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* **21**, 526-531.
- Ishida, K., Hung, T.V., Liou, K., Lee, H.C., Shin, C.H. & Sohng, J.K. 2006. Characterization of *pbpA* and *pbp2* encoding penicillin-binding proteins located on the downstream of clavulanic acid gene cluster in *Streptomyces clavuligerus*. *Biotechnol. Lett.* **28**, 409-417.
- Jensen, S.E., Elder, K.J., Aidoo, K.A. & Paradkar, A.S. 2000. Enzymes catalyzing the early steps of clavulanic acid biosynthesis are encoded by two sets of paralogous genes in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* **44**, 720-726.
- Jensen, S.E. & Paradkar, A.S. 1999. Biosynthesis and molecular genetics of clavulanic acid. *Antonie Van Leeuwenhoek* **75**, 125-133.
- Jensen, S.E., Paradkar, A.S., Mosher, R.H., Anders, C., Beatty, P.H., Brumlik, M.J., Griffin, A. & Barton, B. 2004a. Five additional genes are involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* **48**, 192-202.
- Jensen, S.E., Westlake, D.W. & Wolfe, S. 1983. Partial purification and characterization of isopenicillin N epimerase activity from *Streptomyces clavuligerus*. *Can. J. Microbiol.* **29**, 1526-1531.
- Jensen, S.E., Westlake, D.W. & Wolfe, S. 1985. Deacetoxycephalosporin C synthetase and deacetoxycephalosporin C hydroxylase are two separate enzymes in *Streptomyces clavuligerus*. *J. Antibiot. (Tokyo)* **38**, 263-265.
- Jensen, S.E., Wong, A., Griffin, A. & Barton, B. 2004b. *Streptomyces clavuligerus* has a second copy of the proclavaminic amidinohydrolyase gene. *Antimicrob. Agents Chemother.* **48**, 514-520.
- Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47** (Pt 2), 110-119.

- Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. & Ghosh, D. 1995. Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34**, 6003-6013.
- Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. 2002. Short-chain dehydrogenases/reductases (SDRs). *Eur. J. Biochem.* **269**, 4409-4417.
- Kershaw, N.J., McNaughton, H.J., Hewitson, K.S., Hernandez, H., Griffin, J., Hughes, C., Greaves, P., Barton, B., Robinson, C.V. & Schofield, C.J. 2002. ORF6 from the clavulanic acid gene cluster of *Streptomyces clavuligerus* has ornithine acetyltransferase activity. *Eur. J. Biochem.* **269**, 2052-2059.
- Kleywegt, G.J. & Jones, T.A. 1997. Detecting folding motifs and similarities in protein structures. *Methods in Enzymology* **277**, 525-545.
- Kleywegt, G.J. & Jones, T.A. 1999. Software for handling macromolecular envelopes. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 941-944.
- Kleywegt, G.J. & Read, R.J. 1997. Not your average density. *Structure* **5**, 1557-1569.
- Koide, A. & Hoch, J.A. 1994. Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol. Microbiol.* **13**, 417-426.
- Kropp, H., Gerckens, L., Sundelof, J.G. & Kahan, F.M. 1985. Antibacterial activity of imipenem: the first thienamycin antibiotic. *Rev. Infect. Dis.* **7 Suppl 3**, S389-410.
- La Fortelle, E. & Bricogne, G. 1997. Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology*, pp472-492.
- Lamzin, V.S., Perrakis, A. & Wilson, K.S. 2001. The ARP/WARP suite for automated construction and refinement of protein models. In *International Tables for Crystallography*. Edited by M.G. Rossmann & E. Arnold. Kluwer Academic Publishers, Dordrecht. pp 720-722.
- Lazazzera, B.A. & Grossman, A.D. 1998. The ins and outs of peptide signaling. *Trends Microbiol.* **6**, 288-294.
- Lee, H.J., Lloyd, M.D., Harlos, K., Clifton, I.J., Baldwin, J.E. & Schofield, C.J. 2001. Kinetic and crystallographic studies on deacetoxycephalosporin C synthase (DAOCS). *J Mol Biol* **308**, 937-948.
- Lee, H.J., Schofield, C.J. & Lloyd, M.D. 2002. Active site mutations of recombinant deacetoxycephalosporin C synthase. *Biochem. Biophys. Res. Commun.* **292**, 66-70.
- Leslie, A.G.W. 1992. Recent changes to the MOSFLM package for processing film and image plate data. In *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, No. 26
- Li, R., Khaleeli, N. & Townsend, C.A. 2000. Expansion of the clavulanic acid gene cluster: identification and in vivo functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. *J. Bacteriol.* **182**, 4087-4095.
- Liras, P. & Rodriguez-Garcia, A. 2000. Clavulanic acid, a beta-lactamase inhibitor: biosynthesis and molecular genetics. *Appl. Microbiol. Biotechnol.* **b**, 467-475.

- Lloyd, M.D., Lee, H.J., Harlos, K., Zhang, Z.H., Baldwin, J.E., Schofield, C.J., Charnock, J.M., Garner, C.D., Hara, T., Terwisscha van Scheltinga, A.C., Valegard, K., Viklund, J.A., Hajdu, J., Andersson, I., Danielsson, A. & Bhikhabhai, R. 1999. Studies on the active site of deacetoxycephalosporin C synthase. *J Mol Biol* 287, 943-960.
- Lorenzana, L.M., Perez-Redondo, R., Santamarta, I., Martin, J.F. & Liras, P. 2004. Two oligopeptide-permease-encoding genes in the clavulanic acid cluster of *Streptomyces clavuligerus* are essential for production of the β -lactamase inhibitor. *J Bacteriol* 186, 3431-3438.
- Luengo, J.M. 1995. Enzymatic synthesis of hydrophobic penicillins. *J Antibiot (Tokyo)* 48, 1195-1212.
- Mainardi, J.L., Morel, V., Fourgeaud, M., Cremniter, J., Blanot, D., Legrand, R., Frehel, C., Arthur, M., Van Heijenoort, J. & Gutmann, L. 2002. Balance between two transpeptidation mechanisms determines the expression of beta-lactam resistance in *Enterococcus faecium*. *J Biol Chem* 277, 35801-35807.
- Marland, Z., Beddoe, T., Zaker-Tabrizi, L., Lucet, I.S., Brammananth, R., Whisstock, J.C., Wilce, M.C., Coppel, R.L., Crellin, P.K. & Rossjohn, J. 2006. Hijacking of a substrate-binding protein scaffold for use in mycobacterial cell wall biosynthesis. *J. Mol. Biol.* 359, 983-997.
- Marsh, E.N., Chang, M.D. & Townsend, C.A. 1992. Two isozymes of clavamate synthase central to clavulanic acid formation: cloning and sequencing of both genes from *Streptomyces clavuligerus*. *Biochemistry* 31, 12648-12657.
- Massova, I. & Mobashery, S. 1998. Kinship and diversification of bacterial penicillin-binding proteins and beta-lactamases. *Antimicrob Agents Chemother* 42, 1-17.
- Matias, V.R. & Beveridge, T.J. 2005. Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. *Mol. Microbiol.* 56, 240-251.
- Matias, V.R. & Beveridge, T.J. 2006. Native cell wall organization shown by cryo-electron microscopy confirms the existence of a periplasmic space in *Staphylococcus aureus*. *J. Bacteriol.* 188, 1011-1021.
- Matthews, B.W. 1968. Solvent content of protein crystals. *J. Mol. Biol.* 33, 491-497.
- Mellado, E., Lorenzana, L.M., Rodriguez-Saiz, M., Diez, B., Liras, P. & Barredo, J.L. 2002. The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: genetic organization of the region upstream of the *car* gene. *Microbiology* 148, 1427-1438.
- Miller, M.T., Bachmann, B.O., Townsend, C.A. & Rosenzweig, A.C. 2001. Structure of beta-lactam synthetase reveals how to synthesize antibiotics instead of asparagine. *Nat Struct Biol* 8, 684-689.
- Mosher, R.H., Paradkar, A.S., Anders, C., Barton, B. & Jensen, S.E. 1999. Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavamate synthase 1 in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 43, 1215-1224.

- Murshudov, G.N., Vagin, A.A. & Dodson, E.J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240-255.
- Nagarajan, R., Boeck, L.D., Gorman, M., Hamill, R.L., Higgins, C.E., Hoehn, M.M., Stark, W.M. & Whitney, J.G. 1971. Beta-lactam antibiotics from *Streptomyces*. *J. Am. Chem. Soc.* **93**, 2308-2310.
- Nicholson, N.H., Baggaley, K.H., Cassells, R., Elson, S.W., Fulston, M., Tyler, J.W. & Woroniecki, S.R. 1994. Evidence that the immediate biosynthetic precursor of clavulanic acid is its N-aldehyde analogue. *J. Chem. Soc., Chem. Comm.*, 1281-1282.
- Nickitenko, A.V., Trakhanov, S. & Quioco, F.A. 1995. 2 A resolution structure of DppA, a periplasmic dipeptide transport/chemosensory receptor. *Biochemistry* **34**, 16585-16595.
- Nodwell, J.R. & Losick, R. 1998. Purification of an extracellular signaling molecule involved in production of aerial mycelium by *Streptomyces coelicolor*. *J. Bacteriol.* **180**, 1334-1337.
- Nodwell, J.R., McGovern, K. & Losick, R. 1996. An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. *Mol. Microbiol.* **22**, 881-893.
- Nodwell, J.R., Yang, M., Kuo, D. & Losick, R. 1999. Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genetics* **151**, 569-584.
- Oh, B.H., Ames, G.F. & Kim, S.H. 1994. Structural basis for multiple ligand specificity of the periplasmic lysine-, arginine-, ornithine-binding protein. *J. Biol. Chem.* **269**, 26323-26330.
- Olson, E.R., Dunyak, D.S., Jurss, L.M. & Poorman, R.A. 1991. Identification and characterization of *dppA*, an *Escherichia coli* gene encoding a periplasmic dipeptide transport protein. *J. Bacteriol.* **173**, 234-244.
- Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jornvall, H. 2003. Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem. Biol. Interact.* **143-144**, 247-253.
- Oster, L.M., Lester, D.R., Terwisscha van Scheltinga, A., Svenda, M., van Lun, M., Genereux, C. & Andersson, I. 2006. Insights into cephamycin biosynthesis: the crystal structure of Cmcl from *Streptomyces clavuligerus*. *J. Mol. Biol.* **358**, 546-558.
- Otwinowski, Z. & Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326.
- Padayatti, P.S., Helfand, M.S., Totir, M.A., Carey, M.P., Carey, P.R., Bonomo, R.A. & van den Akker, F. 2005. High resolution crystal structures of the trans-enamine intermediates formed by sulbactam and clavulanic acid and E166A SHV-1 β -lactamase. *J. Biol. Chem.* **280**, 34900-34907.
- Pasteur, L. & Joubert, J. 1877. Charbon et septicémie. *Compt. Rend.* **85**, 101-105.
- Perez-Llarena, F.J., Liras, P., Rodriguez-Garcia, A. & Martin, J.F. 1997. A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in

- overproduction of both beta-lactam compounds. *J. Bacteriol.* **179**, 2053-2059.
- Perez-Redondo, R., Rodriguez-Garcia, A., Martin, J.F. & Liras, P. 1998. The *claR* gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. *Gene* **211**, 311-321.
- Perrakis, A., Morris, R. & Lamzin, V.S. 1999. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458-463.
- Picon, A. & van Wely, K.H.M. 2001. Peptide binding to the *Bacillus subtilis* oligopeptide-binding proteins OppA and AppA. *Mol. Biol. Today* **2**, 21-25.
- Poole, K. 2004. Resistance to β -lactam antibiotics. *Cell Mol. Life Sci.* **61**, 2200-2223.
- Prescott, A.G. 1993. A dilemma of dioxygenases (or where biochemistry and molecular biology fail to meet). *J. Exp. Bot.* **44**, 849-861.
- Proshlyakov, D.A., Henshaw, T.F., Monterosso, G.R., Ryle, M.J. & Hausinger, R.P. 2004. Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/ α -ketoglutarate dioxygenase. *J. Am. Chem. Soc.* **126**, 1022-1023.
- Quioco, F.A. & Ledvina, P.S. 1996. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20**, 17-25.
- Raper, K.B. 1946. The development of improved penicillin-producing moulds. *Annals of the New York Academy of Sciences* **48**, 41-56.
- Roach, P.L., Clifton, I.J., Fulop, V., Harlos, K., Barton, G.J., Hajdu, J., Andersson, I., Schofield, C.J. & Baldwin, J.E. 1995. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**, 700-704.
- Roach, P.L., Clifton, I.J., Hensgens, C.M., Shibata, N., Schofield, C.J., Hajdu, J. & Baldwin, J.E. 1997. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830.
- Rogers, H.J. & Forsberg, C.W. 1971. Role of autolysins in the killing of bacteria by some bactericidal antibiotics. *J. Bacteriol.* **108**, 1235-1243.
- Röhl, F., Rabenhorst, J. & Zahner, H. 1987. Biological properties and mode of action of clavams. *Arch. Microbiol.* **147**, 315-320.
- Rossmann, M.G., Adams, M.J., Buehner, M., Ford, G.C., Hackert, M.L., Liljas, A., Rao, S.T., Banaszak, L.J., Hill, E., Tsernoglou, D. & Webb, L. 1973. Letter: Molecular symmetry axes and subunit interfaces in certain dehydrogenases. *J. Mol. Biol.* **76**, 533-537.
- Rossmann, M.G., Liljas, A., Brändén, C. & Banaszak, L.J. 1975. Evolutionary and structural relationships among dehydrogenases. In *The Enzymes*. Academic Press. New York. pp 61-101.
- Rossmann, M.G., Moras, D. & Olsen, K.W. 1974. Chemical and biological evolution of nucleotide-binding protein. *Nature* **250**, 194-199.
- Rozwarski, D.A., Vilcheze, C., Sugantino, M., Bittman, R. & Sacchettini, J.C. 1999. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP

- reductase, InhA, in complex with NAD⁺ and a C16 fatty acyl substrate. *J. Biol. Chem.* **274**, 15582-15589.
- Sanchez, L. & Brana, A.F. 1996. Cell density influences antibiotic biosynthesis in *Streptomyces clavuligerus*. *Microbiology* **142** (Pt 5), 1209-1220.
- Santamarta, I., Rodriguez-Garcia, A., Perez-Redondo, R., Martin, J.F. & Liras, P. 2002. *CcaR* is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcl* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. *J. Bacteriol.* **184**, 3106-3113.
- Sawicki, M.W., Erman, M., Puranen, T., Vihko, P. & Ghosh, D. 1999. Structure of the ternary complex of human 17beta-hydroxysteroid dehydrogenase type 1 with 3-hydroxyestra-1,3,5,7-tetraen-17-one (equilin) and NADP⁺. *Proc. Natl. Acad. Sci. U.S.A* **96**, 840-845.
- Schiefner, A., Breed, J., Bosser, L., Kneip, S., Gade, J., Holtmann, G., Diederichs, K., Welte, W. & Bremer, E. 2004. Cation-pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. *J. Biol. Chem.* **279**, 5588-5596.
- Schneider, T.R. & Sheldrick, G.M. 2002. Substructure solution with SHELXD. *Acta Cryst. D Biol. Crystallogr.* **58**, 1772-1779.
- Sleigh, S.H., Tame, J.R., Dodson, E.J. & Wilkinson, A.J. 1997. Peptide binding in OppA, the crystal structures of the periplasmic oligopeptide binding protein in the unliganded form and in complex with lysyllsine. *Biochemistry* **36**, 9747-9758.
- Solomon, J.M., Lazazzera, B.A. & Grossman, A.D. 1996. Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in *Bacillus subtilis*. *Genes Dev.* **10**, 2014-2024.
- Tahlan, K., Anders, C. & Jensen, S.E. 2004a. The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. *J. Bacteriol.* **186**, 6286-6297.
- Tahlan, K., Park, H.U., Wong, A., Beatty, P.H. & Jensen, S.E. 2004b. Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* **48**, 930-939.
- Tam, R. & Saier, M.H., Jr. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* **57**, 320-346.
- Tame, J.R., Murshudov, G.N., Dodson, E.J., Neil, T.K., Dodson, G.G., Higgins, C.F. & Wilkinson, A.J. 1994. The structural basis of sequence-independent peptide binding by OppA protein. *Science* **264**, 1578-1581.
- Tanaka, N., Nonaka, T., Nakamura, K. & Hara, A. 2001. SDR: structure, mechanism of action, and substrate recognition. *Curr. Org. Chem.* **5**, 89-111.
- Tanaka, N., Nonaka, T., Tanabe, T., Yoshimoto, T., Tsuru, D. & Mitsui, Y. 1996. Crystal structures of the binary and ternary complexes of 7 α -hydroxysteroid dehydrogenase from *Escherichia coli*. *Biochemistry* **35**, 7715-7730.

- Terwisscha van Scheltinga, A.C., Valegard, K., Ramaswamy, S., Hajdu, J. & Andersson, I. 2001. Multiple isomorphous replacement on merohedral twins: structure determination of deacetoxycephalosporin C synthase. *Acta. Cryst. D Biol. Crystallogr.* **57**, 1776-1785.
- Thai, W., Paradkar, A.S. & Jensen, S.E. 2001. Construction and analysis of β -lactamase-inhibitory protein (BLIP) non-producer mutants of *Streptomyces clavuligerus*. *Microbiology* **147**, 325-335.
- Tipper, D.J. & Strominger, J.L. 1965. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc. Natl. Acad. Sci. U.S.A* **54**, 1133-1141.
- Tobin, M.B., Fleming, M.D., Skatrud, P.L. & Miller, J.R. 1990. Molecular characterization of the acyl-coenzyme A:isopenicillin N acyltransferase gene (*penDE*) from *Penicillium chrysogenum* and *Aspergillus nidulans* and activity of recombinant enzyme in *Escherichia coli*. *J. Bacteriol.* **172**, 5908-5914.
- Tomasz, A., Albino, A. & Zanati, E. 1970. Multiple antibiotic resistance in a bacterium with suppressed autolytic system. *Nature* **227**, 138-140.
- Townsend, C.A. 2002. New reactions in clavulanic acid biosynthesis. *Curr. Opin. Chem. Biol.* **6**, 583-589.
- Townsend, C.A., Ho, M.-F. & Mao, S.-S. 1986. The stereochemical fate of (2*RS*,5*R*) and (2*RS*,5*S*)-[5-³H]ornithine in clavulanic acid biosynthesis. *J. Chem. Soc. Chem. Comm.* **8**, 638-639.
- Trepanier, N.K., Jensen, S.E., Alexander, D.C. & Leskiw, B.K. 2002. The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. *Microbiology* **148**, 643-656.
- Vagin, A. & Teplyakov, A. 1997. MOLREP: an Automated Program for Molecular Replacement. *J. App. Crystallogr.* **30**, 1022-1025.
- Valegard, K., van Scheltinga, A.C., Lloyd, M.D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H.J., Baldwin, J.E., Schofield, C.J., Hajdu, J. & Andersson, I. 1998. Structure of a cephalosporin synthase. *Nature* **394**, 805-809.
- Varughese, K.I., Skinner, M.M., Whiteley, J.M., Matthews, D.A. & Xuong, N.H. 1992. Crystal structure of rat liver dihydropteridine reductase. *Proc Natl. Acad. Sci. U.S.A* **89**, 6080-6084.
- Walsh, C. 2000. Molecular mechanisms that confer antibacterial drug resistance. *Nature* **406**, 775-781.
- Winn, M.D., Isupov, M.N. & Murshudov, G.N. 2001. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 122-133.
- Yang, J.K., Park, M.S., Waldo, G.S. & Suh, S.W. 2003. Directed evolution approach to a structural genomics project: Rv2002 from Mycobacterium tuberculosis. *Proc. Natl. Acad. Sci. U.S.A* **100**, 455-460.
- Zhang, Z., Ren, J., Stammers, D.K., Baldwin, J.E., Harlos, K. & Schofield, C.J. 2000. Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat. Struct. Biol.* **7**, 127-133.

Zhang, Z., Ren, J.S., Clifton, I.J. & Schofield, C.J. 2004. Crystal structure and mechanistic implications of 1-aminocyclopropane-1-carboxylic acid oxidase--the ethylene-forming enzyme. *Chem. Biol.* **11**, 1383-1394.

8 Acknowledgments

An almost chronological list of thanks goes as follows:

The greatest of thanks go to my supervisor, Inger, the master of juggling- how you keep all those balls in the air I'll never know. Your generosity and enthusiasm know no bounds. Some how you found the time to take on a plant physiologist and turn him into a protein crystallographer. Thank you for passing on your scientific knowledge, for pointing me in the right direction, and for keeping me on track. Also big thank you for not just focusing on work! Thanks for all the fun things in Oxford, the parties and important discussions on ice-skating, skiing and sailing.

Thanks to Janos, my co-supervisor, for being one of the few remaining mad scientists. Your enthusiasm is contagious, and if we could harness your energy we could power the entire nation for years. Thanks for illuminating discussions on the boat over G&T, pointing the way to the Oxford pubs, an elegant bathing style at Tällberg, and being a patron of all things fun.

A big thank you to Diane, who brought me to Sweden in the first place. Thank you trying to teach me molecular biology, and for all those euro-land adventures.

Thanks to Karl, who brought me orange juice after trying to poison me with snaps. For the all those late nights at beer club, classic quotes and that characteristic laugh. For all those crazy mid-summer parties and for that rather enjoyable day trip through the Stockholm archipelago by kayak, foot and bike.

Thanks to Mark for turning me into an obsessive climber. Thanks for all those adventures on ice (horizontal & vertical), snow (flat & steep), water (sail & paddle) and for trying to kill me by sending me up apple cake arete (5+) in the Lofoten Islands. Thanks for the dinners, discussions, being computer nerd and for photography tips.

To all the old members of the Inger/Janos corridor of B10:1:
Richard for a classic laugh and jovial sense of humour, and for skiing in Åre. Tove, for computer handling technique and people handling skills. Gunilla, for the fine art of knitting, and driving your own race. Gunnar, for being such a character. Adam, for that strange party in Oxford, and even stranger pub afterwards. S. Iwata, for great office organisation skills, and for being so focused at the beamlines. Hanna and Abe for awful jokes, entertaining puns and for your amazing energy and lust for life- you guys are an inspiration

Thanks to Gisela, for all the dinners/parties and for general energy. For a fun time in Tassie land, and for being so positive and entertaining.

Thanks to Gösta, for being efficient, organised and punctual. For being such an enthusiast- always tagging along, but complaining how you would like to be

somewhere else. Thanks for all the iceskating, climbing (plastic & rock) expeditions, and for headless chicken complaints at boulder open. For all those parties and late night snusing of Cuban cigars.

Thanks to Talal for being the social butterfly. For your passion for life, all those dinners, parties and discussions on anything and everything. For your cheeky antics and sense of humour. For slaying me in squash, for climbing and being a fashion critic (at the same time).

A big thanks to my old office mate Anke, for your fantastic laugh, and for interesting dutch expressions, for teaching me crystallography, explaining how things work, and being amazingly patient. For skating trips, and late night talisman sessions, for being so enthusiastic, and a wonderful person.

A big thank you to the SLU crowd:

Kenth for a wicked session of late night quake and beer, and general fun and entertainment. Deva and Rams for giving me so much grief about cricket, and for being so damn entertaining. Andreas K (and Miss G) for the days of liver damage, music, no sleep and hamburgers, music, music, and watching crazy Finns. Thanks for the Britney fan mail too. M. Forstner for interesting climbing attire, and for “umpa umpa” beer music to a techno beat in an audi at 120 kmph on E4. Fredrik, for climbing adventures (plastic, rock, ice), dinners/parties and the “cheese and scotch” incident. Thanks for being so passionate about science. Andrea for being such a cool Grrrl, for letting me try the cello, the parties, music, ski trips to Romme, and for humouring Fredrik by pretending to be interested in Dicty. Lotta for always being up for something fun, for coming to the synchrotron and for looking after beer club. Professor Urszula for the parties, for falling asleep on the couch at BM14, for all the fun on the ESRF trips, and for introducing me to the obsession with small pickles. Rosie and Malin, for karaoke at Tällberg, for outstanding entertainment late at night at the ESRF, for the ski trip, and some inspirational rally driving in the mountains of France. How many guys on a rugby team? Magareta for advice on seeding and crystallization. Nisse for entertainment on synchrotron trips. Christer for fixing the TV, for mean go-kart tactics, and for wearing a tie. Martin for getting us organised with group meetings, for previous president duties, and for entertainment late at night on the beamlines. Stefan for being such a cool professor, for late night discussions on science and world travel, and for keeping the department running. Wimal for being such a nice guy, for making sure everyone is OK, for late night chats at Tällberg. Sanjeewani for keeping me company on Sherry’s course. Saeid for entertainment at the ESRF, and for always saying “I’m too old for this” and then staying up later than the youngsters anyway. Ellenor for keeping my paper work under control and things running smoothly around the department. Ulla for entertainment at the ESRF, for departmental organisation, and making sure everyone is happy. Jonas for the way in which your scary minds works, for being obsessed with bizarre internet things, and for all those crazy stories. Agata for being a beer queen, and such a social person, for keeping me company in Sherry’s course. Anton for being working interesting hours, and for being such a smart guy, for interesting stories, and for always listening to any protein problems. Sherry for running a great protein course.

Anatoly for being friendly and positive, for getting the discussions going at group meetings, and for asking questions. Hasse for being so relaxed, for your advice and comments. Welcome to the group Rie. Jerry for music, dancing, and being a party animal. Mats for coffee room entertainment.

Thanks to my first collaborators out at SLU, Monika and Carina. Thanks for all the help with NDPK, sorry we didn't get that S69 mutant to work.

Thanks to my second collaborators in Oxford. Chris for all those helpful comments, and for looking after us all in Oxford. Nadia for help with CAD, for taking really good care of me in Oxford, and for all your useful comments and advice. Matts and Mark for looking after us.

To the wonderful computer nerds that make the 0's and 1's go around. Erling, David, and Remco - you are the heads of department, it is chaos without you.

Thanks to the upstairs people:

To the Åqvist lads (Johan, FredrikViktor, Jens, Martin, Martin and Martin, Lars and Stefan) for ultimate frisbee and general entertainment in the coffee room. Sinisa for help with enantiomer docking, and for finding outrageous bouldering problems. Jimmy for being probably the best beer club co-president in the world, for your driving skills in a go-kart or out on the open road at 1 am with no lights and determination. For saving me from the water at Karls mid-summer party. Seved for Tippet, Daniel for being just a little bit crazy- in a good way, and for being such a gadget nerd. Patrik for never saying no to any form of adventure. For all the skating, skiing, climbing trips, added value on synchrotron trips, discussion on photography, and for trying to teach me obscure norrland phrases. Torsten for all the singing, chuckles and generally festive atmosphere. Alina for Grrrr! power, energy, and for being crazy in a good way. Lena for lunchroom entertainment, and for getting people organised when it comes to singing, or party like activities. Annette and Nisse- for all the parties/BBQs, late night beer clubs, climbing trips (plastic and rock). For always being up for an adventure in any form, and for being so damn entertaining. Tex for being so damned organised, and for teaching me a trick or two about crystals. Christofer for the obsession with food, and for interesting theories on what should be done with GM virus'. Eva Lena for black, and popping around late at night to entertain at synchrotrons. Wojciech, Pavel, Magnus for interesting lunch time conversations. Nina, for trying to teach me Swedish all those years ago, and for great cloning advice. Henke for being a walking wikipedia, for all the stories and ice tips, and for driving me off the track at the Go-kart race. Smiley Anna for being such so much fun. Cuba Anna for climbing (rock and ice) oh so long ago, for parties and slam dancing- where's me jumper, and late night discussions about Cuba. Adrian for being such a friendly guy, and wicked dancer. Gerard (aka DVD) for all the red dwarf, and for helping me rediscover my inner punk rocker, for the interesting music, for answering about 10,000 stupid crystallography questions, and for the nerdy stuff like USF and EDS. Lars for trying to teach me crystallography, and for lunch time discussions. Alwyn for singing in the corridor, amusing stories, terrorising people and for trying to teach me O. Marian for skating adventures, and lunch room chats. Kaspars for

lunch room chats. Smart Martin 1 and smart Martin 2 for being so smart, and general entertainment at coffee breaks. Fariborz for being such excellent value, for the belays at the climbing wall, and for all those discussions. Emma, for trying to teach me Swedish so long ago, for all those dinners and parties, for nude drawing, discussions about music, telemark adventures, and iceskating, for always looking on the positive side, and for your energy and determination. Bror for being a legend.

To the clav acid mini-structural genomics team:

Martin, for always being so willing to help, for Wednesday nights at Fellini's and METAL! For being the most evil quake player ever, and for help with the molecular biology. Karin, huge thanks for ORF15 and all the support that has come with it. For being such a data collection wizard, and for keeping us under control while in France. Sara, for your passion for life, for all those late night laughs on the beamlines, and for a fantastic sense of humour. Linda, for being so damn organised, and excellent entertainment on synchrotron trips,, for being such a nice person. Tom for being such a braniac when it comes to CNS, averaging, and well..... just about all things crystallographic.....and everything else for that matter.

Thanks to the poeple in Janos corridor:

Calle for slam dancing, frisbee and general amusement. Magnus for discussions on guitar hero's, music and always joining the party. Michiel for being such a photography nerd, for lunch time chats, and assistance with docking. Alexandra for being such an enthusiast, for all your energy, and chats on skype. Filipe for ultimate frisbee, and always up for fun. Nicusor for vampires, lard, politics, quantum physics, and all the other nonsense we talk about, for your insightful views, and for making me almost understand physics.

A big hello to all the new members at the Structural Biology groups at the BMC- Enjoy your stay at one of the most fun places to work.

To all the people outside of work who have made the years in Uppsala very entertaining. All the climbers at the Stallet climbing wall, a big thank you for always being so friendly and sharing the obsession. Shawn, Kristin and Damian for all your energy, adventures and for being truely wonderful people, and for discussing socialism for 3 hours in a sauna (Shawn). Tina, Emil and Linus for not letting the little one stop your adventures. Susan for looking after us in København, for being a fashion queen, party animal, and always being willing to try something else. Mathias (and family) for looking after us in Grenoble, and pointing us to the direction of rock. Keep up the adventures. Dave for all those NZ adventures, and for the Oxford pints. Pia and Per, and miss V for looking after us in Göteborg.

To the Tassie crew. Rob, Hugh and Evan- miss your crazy antics. The surfing lads- miss those waves and brews. Alex for looking after us in London, and for your passion for life. The 346 sandy bay crew: Kate, Jenny, Chris, Jimmy and Campbell- thank for the e-mails and keeping in touch.

To my parents Ken and Alison, for their love and support over the years- miss you guys. To my brother Bogle, mad man and maniac for all sorts of mischief, miss the adventures. To Kajsa and Bullen, my other parents, thanks for taking such good care of me. Oskar, Kajsa and little Saga, Kalle and Poie and Pucko-Nico- thanks for all the fun and entertainment you guys bring with you.

The greatest of all thanks goes to Jenny. Through out the highs and lows of my PhD you have provided such love and support. You are my absolute hero. Without you this book would not have been possible - and finally it is done! I am all yours baby!

Finally, to dear little Rasmus - sorry that you have such nerdy parents. You are too young to understand how difficult the last few months have been, but I hope that when you grow up you do. If you should do a PhD- don't make it as difficult as I did. I am looking forward to having some time to really PLAY.

