# The Biosynthesis of β-Lactams

*A thesis submitted to*

*the Board of the Faculty of Physical Sciences*

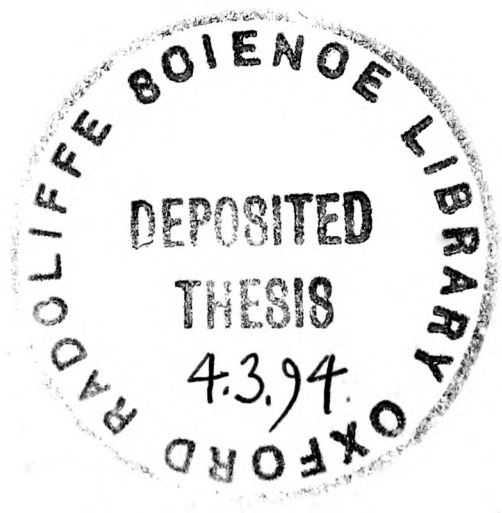
*University of Oxford*

*in partial fulfilment of the requirements for*

*the Degree of Doctor of Philosophy*


***Kee Chuan Goh***

***The Dyson Perrins Laboratory***


*Michaelmas Term*
*1993*


*Linacre College*
*Oxford*

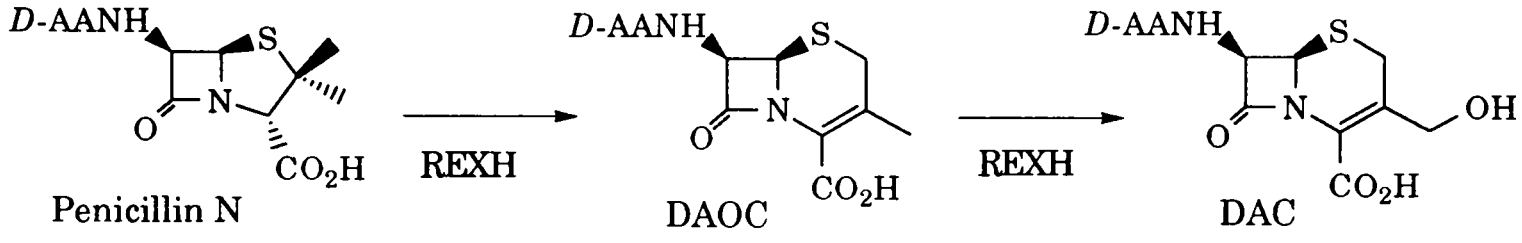
# ABSTRACT
# The Biosynthesis of β-Lactams

***Kee Chuan Goh***
***Linacre College***

***Michaelmas 1993***
***D. Phil.***

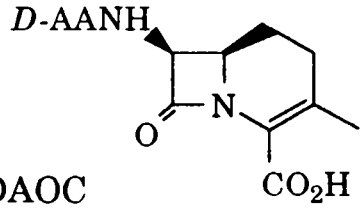
This thesis reports the work done on two research projects which were carried out independently of each other but converge on the central theme of β-lactam biosynthesis. Chapter 1 provides an overview of biosynthesis in secondary metabolism, with special emphasis on current knowledge about the β-lactams.

The first project, covered from Chapters 2 to 5, was part of our group's continuing effort to understand the structure and mechanism of Ring Expandase-Hydroxylase (REXH), an enzyme involved in the biosynthesis of cephalosporin C in *Cephalosporium acremonium*. REXH is a bifunctional enzyme, converting penicillin N to DAOC and thence to DAC.



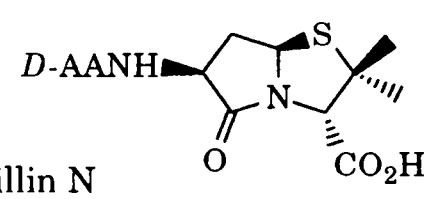

Penicillin N DAOC DAC

Chapter 3 discusses the investigation of purification protocols for native REXH and soluble recombinant REXH, as well as an improved refolding method for recombinant REXH expressed as inclusion bodies. Chapter 4 describes two new alternative substrates for REXH, *viz.* carba-DAOC and DAC, whilst the γ-lactam analogue of penicillin N was not found to be a substrate for REXH. Chapter 5 summarises some structural investigations of REXH employing methods such as electrospray mass spectrometry, selective proteolysis and inhibition kinetics.

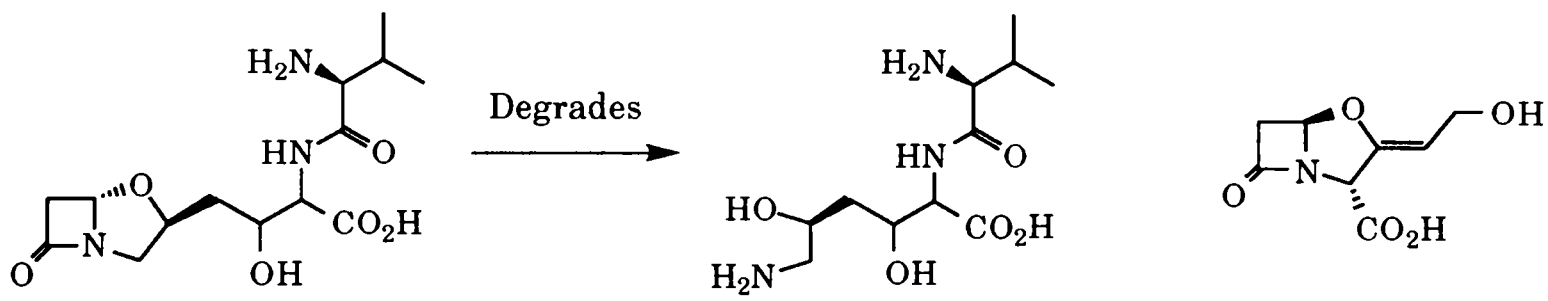



Carba-DAOC

γ-lactam analogue of penicillin N




The second project, covered from Chapters 6 to 9, represents the first biosynthetic studies on valclavam, an antifungal produced by *Streptomyces antibioticus*. Valclavam belongs to the family of clavams which includes clavulanic acid as its most well studied member.




Valclavam A fragment of valclavam Clavulanic acid

Chapter 7 details the development of methods for the bioassay, fermentation and isolation of valclavam. It also describes the isolation of a stable degradation fragment of valclavam which led to the revision of the structures of valclavam and Tü 1718B (another metabolite from the same organism). Chapter 8 gives an account of the whole-cell feeding experiments which strongly suggest that the primary metabolic precursors for valclavam are *L*-valine, *L*-arginine, *L*-methionine and glycerol. Chapter 9 reports the discovery of two enzymic activities, belonging to those of clavaminic acid synthase and proclavaminic acid amidino hydrolase, which are likely to be involved in the biosynthesis of valclavam. Together, the results of Chapters 8 and 9 point to an extensive overlap between the clavulanic acid pathway in *Streptomyces clavuligerus* and the valclavam pathway in *Streptomyces antibioticus*.

# ACKNOWLEDGEMENT

This thesis is made possible with the generous help of the following individuals and organisations, to whom I owe immense gratitude.

Professor **Jack Baldwin** is thanked for kindly accepting me into an excellent laboratory where the expertise and facilities available offer tremendous scope for scientific training. His inspirational leadership and unflagging interest in my work were largely instrumental in sustaining my morale throughout the entire project.

Dr **Christopher Schofield** conducted his day-to-day supervision with contagious enthusiasm and energy, providing plenty of encouragement along the way. His personal interest in my general well-being and things Singapore cemented the foundation for a promising life-long friendship. In short, a fine mentor and a good friend.

Dr **Robert Field** was a charismatic lab manager during my early days in Oxford, and was responsible for imparting a lot of new techniques and concepts. Dr **Yoshiyuki Fujishima**, Dr **Matthew Lloyd**, Mr **Christopher Lawrence**, Mr **Shiau Chia-Yang** and Mr **Roland Brown** were excellent company throughout my time in Enzymology. Dr **Michael Byford**, Dr **Stephen Cole**, Dr **Peter Roach** and Dr **Mark Wood** are gratefully acknowledged for their assiduous proof-reading, but all mistakes remaining were wholly mine. The presence of Mrs **Wendy Sobey**, Dr **Zhang Zhi-Hong**, Mr **John Barlow** , Mr **Timothy Hitchman** and Mr **Justin Withers** helped contribute to a vibrant lab atmosphere.

Technical support for my work was expertly rendered by the following people: Mr **Jonathan Pitt** (for store service and odd jobs), Mr **Anthony Willis** (for amino acid analysis), Dr **Timothy Claridge** (for NMR), Dr **Robin Aplin** (for mass spectrometry) and Dr **John Sutherland** (for recombinant REXH). A special note of thanks is reserved for Mr **John Keeping** for his genial disposition demonstrated throughout my request for fermentations.

Professor Dr **Hans Zähner** (University of Tübingen, Germany) gave us invaluable help which enable us to initiate the valclavam project. Dr **Yeh Wu-Kuang** (Eli **Lilly** & Co, USA) facilitated some of our work on REXH and my first visit to USA. SmithKline **Beecham** (Surrey, UK), Hoffmann-La **Roche** (Nutley, USA) and **Ciba-Geigy** (Basel, Switzerland) were helpful in no small way, ever so responsive with our requests for chemicals and data.

My time in Oxford has been made much more enjoyable by solid friendship with some people. They include **Eric Yap**, **Wong Heng-Vee**, **Wong Heng-Yu**, **Thomas Leung**, **Tan Juan-Boon**, **Simon Griffin**, **Keiko Murata**, **Yuri Nakagawa**, **Richard Werner** and **Tseng Wang-Ruu**. **Peter** and **Marjorie Curtis** are a superb couple to live with, especially during the stressful period of thesis-writing.

My study in Oxford was made possible by a handsome scholarship from the Singapore **Economic Development Board** and **Glaxo** Holdings PLC. I am especially indebted to Ms **Lam Yeen-Lan** for administering the scholarship in a most professional and compassionate manner. The **Committee of Vice-Chancellors and Principals**, United Kingdom, is also thanked for the award of the Overseas Research Student scholarship.

My **family**, especially **Ma** and **Pa**, has given me strength in everything I do, especially over the last three years of my intensive study in a foreign land.

# ABBREVIATIONS

| | |
|---|---|
| ACV | δ-(*L*-α-**A**minoadipoyl)-*L*-**C**ysteinyl-*D*-**V**aline |
| ACVS | δ-(*L*-α-**A**minoadipoyl)-*L*-**C**ysteinyl-*D*-**V**aline **S**ynthetase |
| aka | **a**lso **k**nown **a**s |
| α-KG | α-**K**eto**G**lutarate |
| ATCC | **A**merican **T**ype **C**ulture **C**ollection |
| CAS | **C**lavaminic **A**cid **S**ynthase |
| CD | **C**ircular **D**ichroism |
| CFE | **C**ell-**F**ree **E**xtract |
| COSY | **CO**rrelated Spectroscop**Y** |
| cpm | **c**ounts **p**er **m**inute |
| Da | **Da**lton (unit of atomic mass) |
| DAC | **DeA**cetylcephalosporin **C** |
| DACS | **DeA**cetylcephalosporin **C** **S**ynthase |
| DAOC | **DeA**cet**O**xycephalosporin **C** |
| DAOCS | **DeA**cet**O**xycephalosporin **C** **S**ynthase |
| DNA | **D**eoxyribo**N**ucleic **A**cid |
| DTT | **DiT**hio**T**hreitol |
| EPR | **E**lectron **P**aramagnetic **R**esonance |
| ESMS | **E**lectro**S**pray **M**ass **S**pectrometry |
| EXAFS | **E**xtended **X**-ray **A**bsorption **F**ine **S**tructure |
| FD-MS | **F**ield **D**esorption - **M**ass **S**pectrometry |
| FPLC | **F**ast **P**rotein **L**iquid **C**hromatography |
| HMBC | **H**eteronuclear **M**ultiple-**B**ond **C**onnectivity |
| HMQC | **H**eteronuclear **M**ultiple **Q**uantum **C**oherence |
| HPLC | **H**igh **P**erformance **L**iquid **C**hromatography |
| $IC_x$ | **I**nhibitory **C**oncentration (for **x**% inhibition of the enzyme) |
| IPNS | **I**so**P**enicillin **N** **S**ynthase |
| IPTG | **I**so**P**ropyl**T**hio**G**alactoside |
| IR | **I**nfra-**R**ed |
| kDa | **k**ilo**Da**lton |
| $K_m$ | **M**ichaelis Constant |
| LSC | **L**iquid **S**cintillation **C**ounting |
| MW | **M**olecular **W**eight |
| NMR | **N**uclear **M**agnetic **R**esonance |
| NRRL | **N**orthern **R**egional **R**esearch **L**aboratories |
| PAH | **P**roclavaminic acid **A**midino **H**ydrolase |
| PCA | **Pro**C**l**avaminic **A**cid |

| | |
|---|---|
| PMSF | **P**henyl**M**ethane**S**ulfonyl **F**luoride |
| PVDF | **P**oly **V**inylidene **D**i**F**luoride |
| REXH | **R**ing **EX**pandase-**H**ydroxylase |
| $R_f$ | **R**atio of **f**rontiers |
| rpm | **r**evolutions **p**er **m**inute |
| $R_t$ | **R**etention **t**ime |
| RT | **R**oom **T**emperature |
| SDS-PAGE | **S**odium **D**odecyl **S**ulfate-**P**oly**A**crylamide **G**el **E**lectrophoresis |
| ssp | **ssp**ecies |
| TSP | 3-**T**rimethyl**S**ilyl-[2,2,3,3-$d_4$]-**P**ropionic acid |
| UV | **U**ltra-**V**iolet |
| $V_{max}$ | **Max**imal **V**elocity |
| v/v | **v**olume by **v**olume |
| w/v | **w**eight by **v**olume |
| w/w | **w**eight by **w**eight |

## Single-letter code for common amino acids

| | |
|---|---|
| A | Alanine |
| R | Arginine |
| N | Asparagine |
| D | Aspartic acid |
| C | Cysteine |
| Q | Glutamine |
| E | Glutamic acid |
| G | Glycine |
| H | Histidine |
| I | Isoleucine |
| L | Leucine |
| K | Lysine |
| M | Methionine |
| F | Phenylalanine |
| P | Proline |
| S | Serine |
| T | Threonine |
| W | Tryptophan |
| Y | Tyrosine |
| V | Valine |

# TABLE OF CONTENTS

# 9 Biosynthetic Enzymes for Valclavam

# CHAPTER 1

# GENERAL INTRODUCTION : AN OVERVIEW OF SECONDARY METABOLITE BIOSYNTHESIS

## 1.1 Primary versus secondary metabolism

The term metabolism denotes the totality of all chemical transformations occurring within an organism. Primary metabolism consists of those processes, shared largely by all living organisms, which are fundamental to their energetics and growth.[1] Examples include the glycolytic pathway, citric acid cycle, photosynthesis and DNA replication. The sugars, amino acids, fatty acids, organic bases (purines and pyrimidines) and macromolecules participating therein are known as primary metabolites. Secondary metabolism, most prevalent among plants and microorganisms, is often very specific to a small subset of the biological world (usually at the taxonomic level of a genus, species or strain). Secondary metabolites have no apparent role in the internal economy of the producer[2]. A universally acceptable definition of secondary metabolites is really quite elusive. One can at best attempt a description of some common observations about these interesting molecules[3]: 'Secondary metabolites are those metabolites which are often produced in a phase subsequent to growth, have no function in growth (although they may have a survival function), are produced by certain taxonomic groups of microorganisms, have unusual chemical structure, and are often formed as mixtures of closely related members of a chemical family. Whereas primary metabolism is basically the same for all living systems, secondary metabolism is restricted to plants and microorganisms and is often species or strain specific.'

Due partly to the fact that secondary metabolic pathways must have their starting points in the primary metabolic network and partly to our superficial understanding of the natural functions of secondary metabolites, there is no sharp demarcation between the two classes of metabolism. Like many biological classification schemes, this division has its own grey area. Despite their dispensability in the organic growth of the producer, secondary metabolites are generally acknowledged to have the general effect of enhancing the evolutionary fitness of the producer by eliciting specific physiological and behavioural responses from another organism[4]. Many suggestions have been made as to how selective advantages can be conferred on the producers, some of which are outlined below[5]:

1. Chemical warfare against competing organisms
2. Chemical deterrence against potential predators
3. Maintenance of interspecific symbiotic relationships
4. Intraspecific signalling functions
5. Nutrient-transporting agents
6. Reserve pool of new pathways - raw materials for inventive evolution

## 1.2 A new frontier for chemists

For many years, primary metabolism had been in the mainstream of intensive basic research, while secondary metabolism was relatively neglected[2] (except possibly by organic chemists who were intrigued by the synthetic talent of Nature, and industrial scientists hunting for more 'wonder' drugs like penicillin). The situation today has changed. Virtually all of the primary metabolic pathways have been elucidated and their regulatory mechanisms understood to a large extent. Many of the enzymes and genes involved in the processes have been manipulated directly and studied in precise molecular terms.

The frontier now belongs to secondary metabolism, where the shroud of mystery beckons to the curious scientists. What are the natural functions of secondary metabolites? How are these molecules made? What is the rationale for such an enormous structural diversity to cope with a much narrower range of perceived functions? Do secondary metabolic enzymes differ from their primary counterparts in any fundamental ways? Are there any significant differences in the genetic regulation of the two classes of metabolism? How did secondary metabolites evolve? Are they products of 'inventive' evolution, whereby the haphazard diversion from a primary pathway yields a molecule that improves the evolutionary fitness of its producer by serendipity? If so, does inventive evolution occur by the duplication and divergence of genes originally coding for primary metabolism? Does horizontal gene transfer also play a role in the acquisition of secondary pathways? What are the origins of the resistance genes which producers use to prevent suicide? Are they related phylogenetically to the resistance genes found in clinically resistant microorganisms?

The multitude of interesting questions above have been gradually addressed as the most powerful methods in chemistry, microbiology, enzymology and molecular biology are being brought to bear on the vast domain of secondary metabolism. The challenges posed by the field make it certain to be a rich provider of scientific rewards in the future.

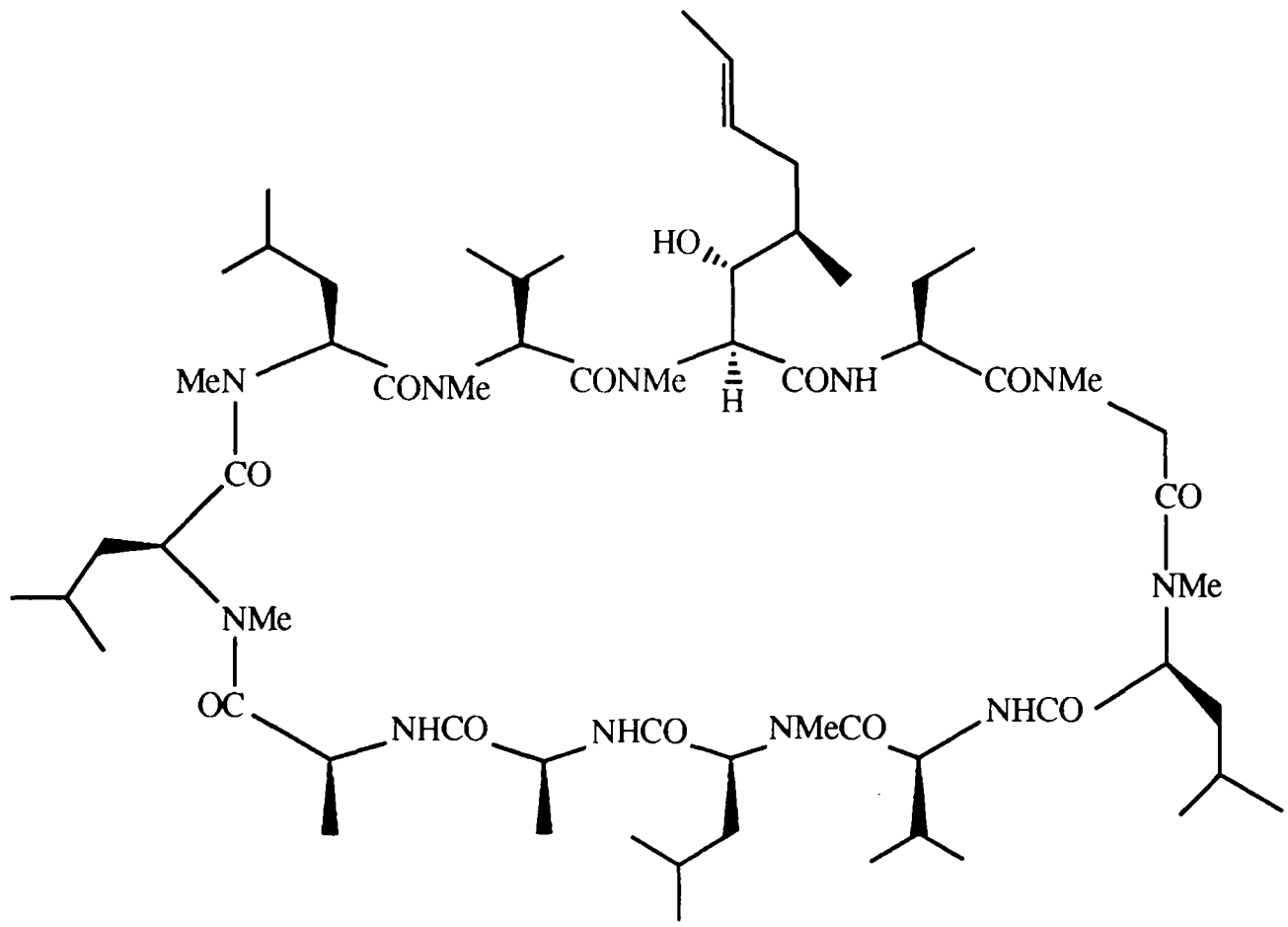
## 1.3 A new opportunity for industry

In this age of exacerbated competition for resources, economics and social conscience dictate that a scientific investigation should be justified not by its intrinsic intellectual challenge alone, but also by its relevance to current problems and its potential dividends to mankind. So, how can a better understanding of secondary metabolism contribute to the global welfare at large, now and in many years to come?

The story of penicillin has been told and retold countless times[6]. It represents one of the first commercial exploitation of a microbial secondary metabolite for chemotherapeutic purposes. Since then, the pharmaceutical industry has been the home of commercial secondary metabolism, conducting large-scale programs in natural product screening to discover novel therapeutic substances. Lately, however, a number of companies became discouraged with this

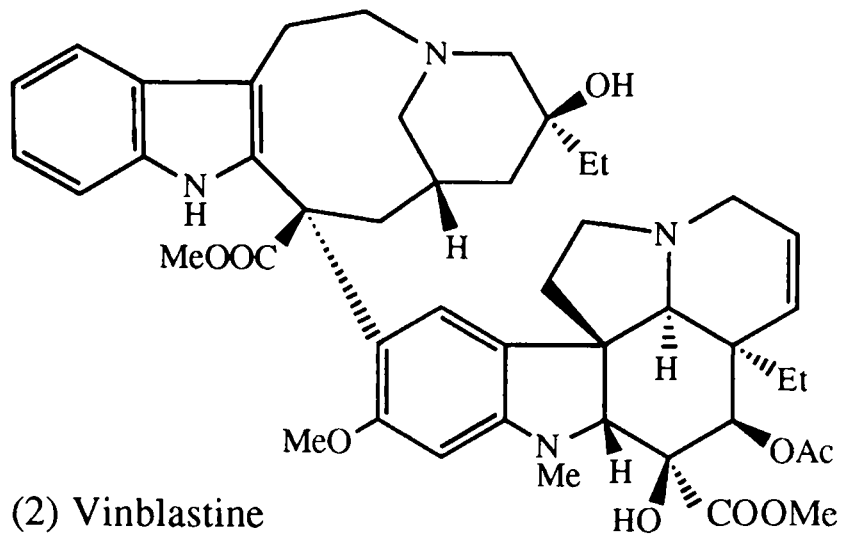
random drug discovery process and the proliferation of antibiotics in the market, and began to de-emphasize it in favour of *de novo* rational drug design and biotechnology products[2]. This is possibly a mistake as explained below.

Screening programs in the past were designed exclusively for the detection of antimicrobial and antitumour activities.[7] Such a biased approach naturally restricted the scope of clinical applications of secondary metabolites. Today, it is realised that the diverse range of structural variations found in these molecules is matched by a comparable diversity in their biological activities, implying a wider application not just in the clinical areas but also in other fields such as agrochemicals.[8] Examples[2] include the use of cyclosporin (1) as an immunosuppressant for transplant patients, vinblastine (2) as an antineoplastic agent to treat Hodgkin's disease and other lymphomas, and monensin (3) as an anti-coccidial agent.

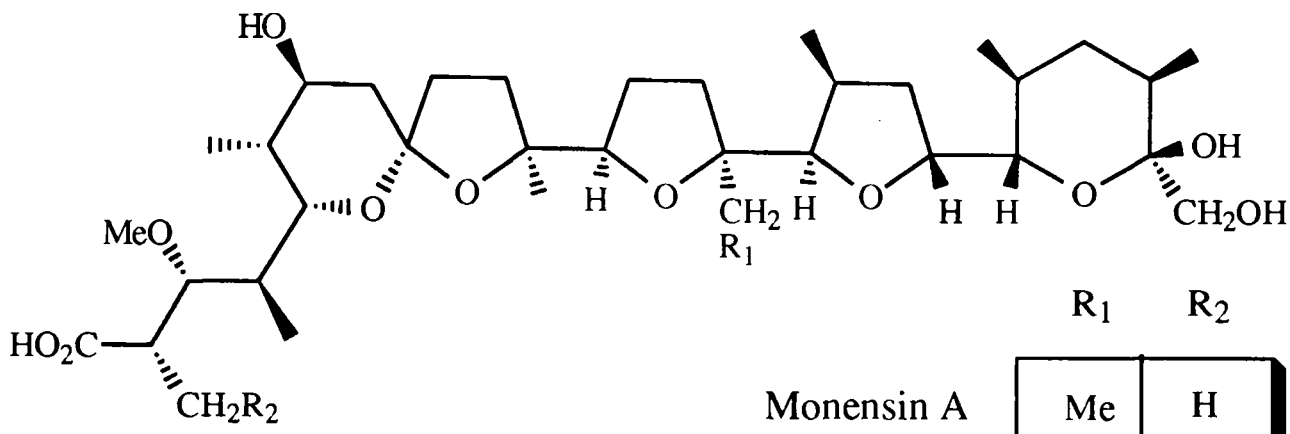



(1) Cyclosporin

Many secondary metabolites, both of plant and microbial origin, are in various stages of clinical trials and look set to play a big role in their own therapeutic areas. They hold great promise as antitumour, antiviral, antiparasitic, anti-inflammatory and hypocholesterolaemic agents. A remarkable example[9] lies in the potential use of taxol (4) to treat certain types of cancer. In the agrochemical business, a number of microbial products are being evaluated for their potent insecticidal and herbicidal properties.

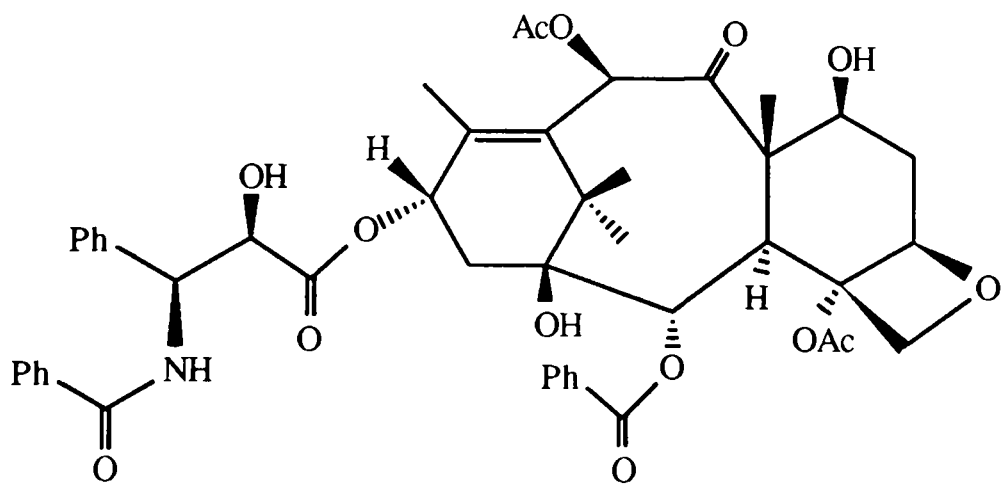



(2) Vinblastine




| | $R_1$ | $R_2$ |
|---|---|---|
| Monensin A | Me | H |
| Monensin B | H | H |
| Monensin C | Me | Me |

(3) The monensins




(4) Taxol

The era of antibiotic discovery based on the random screening of natural products from microbes is being replaced by sophisticated, target-directed, high throughput, mode-of-action screens[7]. This new approach is made possible by applying the latest techniques of molecular biology and biotechnology. In addition, a knowledge of ecology and an appreciation of biodiversity has been employed to source for suitable screen samples from both plants and microorganisms. Ultimately, it is believed that the amazing plethora of molecular architecture elaborated by Nature has a higher chance of enabling recognition processes involving disease-

implicated receptors, compared with the chemical repositories of pharmaceutical companies built by the template-driven imagination of scientists[10]. Natural product screening has become less of a hit-and-miss exercise, and more of a rational, coherent scientific endeavour.

Another great asset to industry arising from the applications of molecular biology is the ease of development of an improved producing strain[2]. This has been greatly facilitated by the clustering of biosynthetic, resistance and regulatory genes of secondary metabolism, such as in the biosynthesis of penicillins, cephalosporins, streptomycin, erythromycin and tetracycline. It is abundantly clear that pharmaceutical and agrochemical companies would stand to lose if they fail to continue tapping on the synthetic prowess of Nature.

## 1.4 Methodology in biosynthetic studies

In Section 1.2, the questions about various aspects of secondary metabolism were posed in what seems to be a logical progressive sequence to make a foray into this exciting interdisciplinary branch of science. The order of activities does not necessarily follow in a rigid manner and there is inherent opportunity for one fruitful piece of work to corroborate earlier results, instigate new lines of investigation in previously explored area, or even provide definitive answers to previous questions.

Among the barrage of questions presented, the one which probably belongs to the forte of chemists (in the broadest sense) concerns the specific enzymatic steps by which precursors from the primary metabolic pool are assembled into the final secondary metabolic product. Interest in the biosynthesis of natural products is nearly as old as the study of natural products themselves. However, it is the early Fifties that heralded the classical era of experimental biosynthetic investigations with the ready availability of radioisotopes of carbon and hydrogen. The results of numerous *in vivo* incorporation experiments with intact plants and microorganisms have largely borne out the early biogenetic hypotheses of Collie, Robinson, Ruzicka and other pioneers[1]. The development of high-field NMR, with its increasing sophistication and ever more powerful correlation methods, has come to dominate the structural field[11]. Its combination with stable isotopes like $^{13}C$, $^{2}H$, $^{18}O$ and $^{15}N$ spurred biosynthetic studies to greater progress by circumventing tedious degradative procedures formerly used to locate exact points of radioisotope incorporation. Mass spectrometry has also been influential in natural product studies, being increasingly exploited in biosynthetic work. The relatively new mode of electrospray MS, in particular, has proven its tremendous value in both small and large-molecule (*e.g.* enzyme) characterization by biosynthetic investigators[12, 13]. Finally, X-ray crystallography must deserve special mention in the structural field as the method provides absolute answers to structure and stereochemistry when data interpretation from other methods can often go awry.

Over the last two decades, incorporation experiments with intact organisms have been gradually superseded by the use of isolated enzyme systems to study individual steps in a

pathway[14]. This has enabled a detailed insight into the molecular events that occur during enzymatic catalysis, involving interaction of substrates, cofactors and reactive functionalities in the protein architecture. Also, the study of enzymes in secondary metabolism has presented new challenges and brought about new perspectives in protein science.

To obtain enough protein for more extensive structural and mechanistic studies, chemists have been turning to recombinant DNA techniques for the isolation, cloning and over-expression of the structural genes coding for the enzymes under investigation. Such endeavours in molecular biology can lead to immensely rich spin-offs, providing novel information about genetic regulation of the secondary metabolic pathways and the phylogenetic relationship among the biosynthetic enzymes.

All in all, a conventional methodology for a biosynthetic project can be outlined as follows:

1. Selecting an interesting secondary metabolite and a suitable producer for study
2. Devise a biosynthetic hypothesis based on compound structure, organism, habitat, *etc.*
3. Optimising production of the metabolite from the organism
4. Identify precursors from the primary metabolic pool by incorporation experiments with intact organisms, either by radioisotope tracing or stable isotope-NMR tracing
5. Seek out *in vivo* intermediates and establish the order of biosynthetic steps
6. Isolate and characterize enzyme systems mediating each discrete biosynthetic step
7. Cloning the structural genes for biosynthetic enzymes, and over-expression
8. Further structural and mechanistic studies on the biosynthetic enzymes, native as well as the recombinant species

## 1.5 Major secondary metabolites and their precursors

The presently known systematic distribution of the major classes of secondary metabolites in the entire biosphere is summarised in Table 1.1, following the classification of Cavalier-Smith[15]. Three groups of compound from the primary metabolic pool are responsible for the biosynthesis of an overwhelming majority of secondary metabolites, namely acetate, shikimate and amino acids[1]. Acetate, in its activated form as acetyl-CoA, is incorporated through two main pathways. The polyketide pathway leads to polyacetylenes, prostaglandins, macrocyclic polyketides and polyphenols, while the mevalonate pathway gives rise to isoprenoids (terpenes, steroids and carotenoids). Metabolism *via* the shikimate pathway affords a large number of aromatic compounds related to the aromatic amino acids, phenylalanine and tyrosine. Many of these compounds are polyphenolic, containing the *o*-dihydroxy- or 1,2,3-trihydroxy- substitution pattern. The phenolic groups are often methylated, or conjugated with non-shikimate derived molecules *via* an ether or ester linkage. Shikimate metabolites can be structurally classified into members of the $ArC_3$ (phenylpropanoids), $ArC_2$, $ArC_1$ and $ArC_0$ families, based on the number of carbons remaining in the substituent.

| | Derived from acetyl-CoA | | | | Derived from amino acid pathways | | | | | | Mixed pathways | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Poly-acetylenes | Prosta-glandins | Cyclic polyketides/ polyphenols | Isoprenoids | Derived from shikimate | β-lactams | Peptides | Alkaloids | Others | Derived from purines | Flavonoids | Other heterosides | Others |
| **EUKARYOTES** | | | | | | | | | | | | | |
| Animalia | + | + | + | + | | | | | + | + | | + | |
| Fungi | + | | + | + | + | + | + | + | + | + | | + | + |
| Plantae | + | | + | + | + | | + | + | + | + | + | | + |
| Chromista | | | + | + | | | | | + | | | | |
| Protozoa | | | | + | | | | | + | + | | | |
| Archezoa | | | | | | | | | | | | | |
| **BACTERIA** | | | | | | | | | | | | | |
| Archae-bacteria | | | | + | | | | | | | | | |
| Eubacteria: | | | | | | | | | | | | | |
| Gram-positive | | | + | | + | + | + | + | + | | | + | |
| Gram-negative | | | | + | | + | | | | | | | |

**Table 1.1** The systematic distribution of major classes of secondary metabolites[15]

It is probably fair to say that the secondary metabolic pathways for amino acids are more diverse than those for the preceding precursors, due in part to the variety of amino acids available as building blocks[1]. The alkaloids, produced principally in higher plants, constitute the largest and most heterogeneous class of amino acid-derived metabolites, if only because of the broad, encompassing nature of its definition. Alkaloids can be grouped into two classes, one group being derived from aliphatic amino acids (*e.g.* ornithine and lysine) and the other originating from aromatic amino acids (phenylalanine, tyrosine and tryptophan). The first group comprises structural types belonging to the families of pyrrolidine, piperidine, pyrrolizidine, quinolizidine and pyridine alkaloids. The second group includes the families of isoquinoline, benzylisoquinoline, aporphine, opium and amaryllidaceae alkaloids. The interesting pharmacological properties of alkaloids accounted for their prominent role in the early history of natural product chemistry. Morphine, strychnine, quinine, nicotine and codeine were isolated in a substantially pure form by the mid-nineteenth century[1].

In the lower organisms where alkaloid pathways are generally absent, secondary metabolism of amino acids is actively engaged through the production of peptide derivatives, both cyclic and linear, and the β-lactams. The peptide derivatives as a group are interesting partly for the following reasons:

1. The majority of these peptides are synthesized in a non-ribosomal manner[16], in contrast to normal protein biosynthesis (translation from mRNA to polypeptide on ribosomes, with decoding by tRNA). These peptides are synthesized on multi-enzyme complexes, assembling themselves from amino acids activated by adenylation. Such a synthetic machinery confers a lower fidelity than one involving decoding by tRNA and provides a rationale for the wide occurrence of homologous peptide structures.

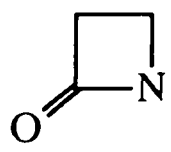
2. A high proportion of the constituent amino acids in peptide derivatives are of an unusual, non-proteinogenic nature and are often themselves derived from the twenty common proteinogenic amino acids. The lack of a tRNA to recognise these unusual amino acids explains why such peptides are not made by the normal translation process.
3. With few exceptions, these peptides have antibiotic properties, acting through various modes of action such as the inhibition of cell wall synthesis or ribosomal protein synthesis.
4. A large number of peptide derivatives are of mixed precursor origin, *i.e.* they contain not only amino acid-derived moieties in their structures. For cyclic peptides, these moieties may reside in the ring or an appendage to the ring. The common non-amino acyl units associated with peptide derivatives are sugars and lipid molecules.

Apart from peptide antibiotics incorporating sugars and lipid molecules, there are many other mixed-precursor secondary metabolites biosynthesized from certain combinations of acetate, mevalonate, shikimate and a few amino acids. The β-lactams are a major class of secondary metabolites derived from amino acids. As they constitute the central theme of this thesis, they will be described separately in the following sections.

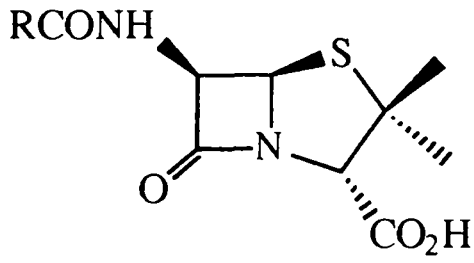
## 1.6 β-Lactams: Structural types and occurrence

The four-membered 2-oxoazetidine ring (5) is the structural motif that constitutes the basis of β-lactams as a family of compounds. The considerable strain in the β-lactam ring underlies the unique chemical reactivity and biological action of these compounds.[17] Naturally occurring β-lactams can be classified into the following groups according to the manner of elaboration of this basic unit:

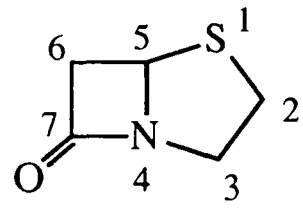
1. The penicillins (6) contain a 5-membered thiazolidine ring fused to the β-lactam ring. The bicyclic skeleton (7) is known as the penam nucleus.




(5)

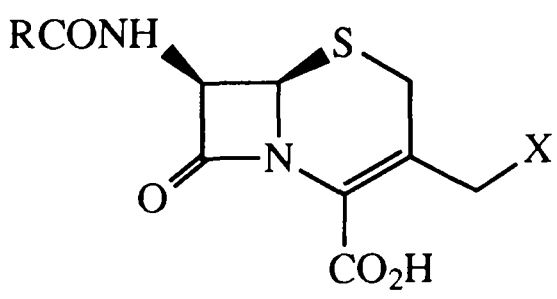



(6)

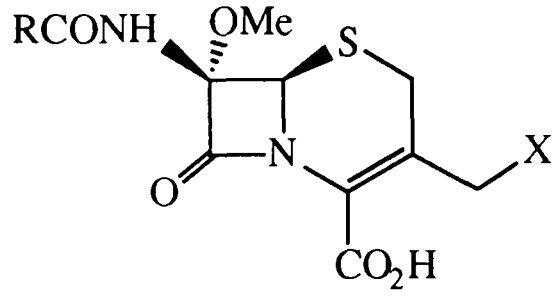



(7)

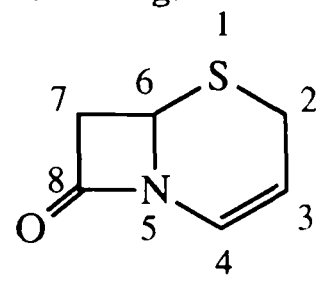
2. The cephalosporins (8) contain a dihydrothiazine ring fused to the β-lactam ring.




(8)



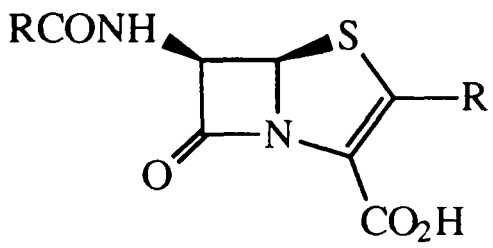

(9)



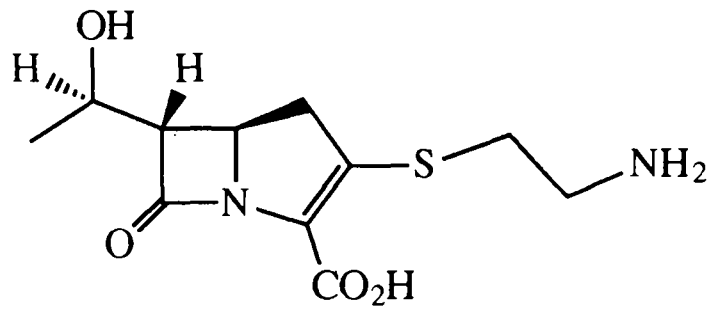

(10)

3. The cephamycins (9) are derivatives of cephalosporins containing a methoxy group at the 7α-position. The bicyclic skeleton (10) found in both cephalosporins and cephamycins is termed the cephem nucleus.

4. The penems (11) possess an unsaturation at the C-2/C-3 bond of the penam nucleus.



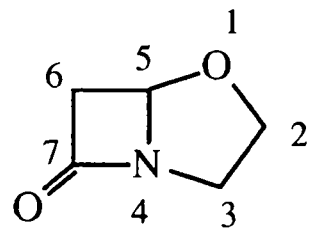

(11)



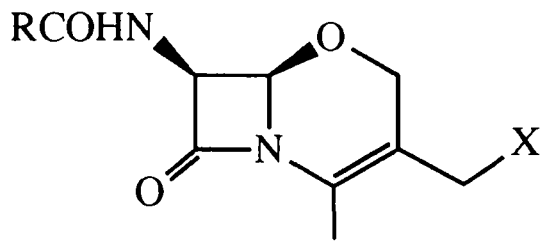

(12)

5. The carbapenems, as its name suggests, possess a nucleus that is the carbocyclic analogue of penems. An example is thienamycin (12).

6. The oxapenams possess the clavam nucleus (13) where there is a replacement of sulphur in penams with oxygen.




(13)

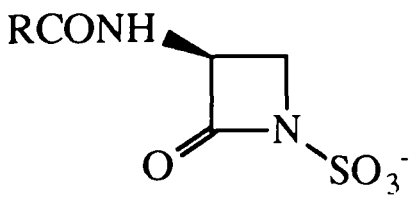



(14)

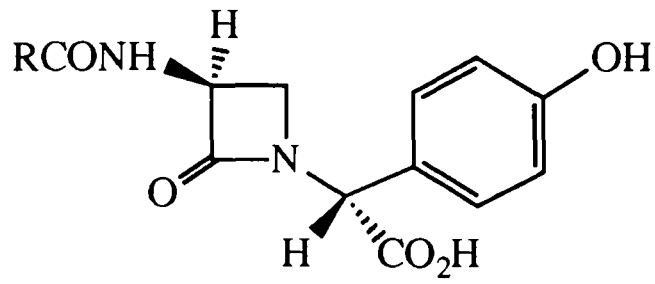
7. The oxacephems (14) are related to the cephems in the same manner as that between clavams and penams.

8. Monobactams (15) are monocyclic β-lactams with a sulfonate group attached to the nitrogen.

9. Nocardicins (16), the other class of monocyclic β-lactams, are substituted differently at the ring nitrogen.




(15)




(16)

As shown in Table 1.1, β-lactams have so far been isolated exclusively from the kingdoms of fungi and bacteria. The particular genera known to produce the various classes of β-lactams are given in Table 1.2.

| β-Lactam Nucleus | Fungi | Bacteria | |
|---|---|---|---|
| | | Actinomycetes | Eubacteria |
| Penam | *Aspergillus*<br>*Penicillium*<br>*Epidermophyton*<br>*Trichophyton*<br>*Polypaecilum*<br>*Malbranchea*<br>*Microsporum* | | |
| Cephem | *Cephalosporium*<br>*Paecilomyces*<br>*Scopulariopsis*<br>*Diheterospora*<br>*Spiroidium* | *Streptomyces*<br>*Nocardia* | *Flavobacterium*<br>*Lysobacter*<br>*Xanthomonas* |
| Clavam | | *Streptomyces* | |
| Carbapenem | | *Streptomyces* | *Erwinia*<br>*Serratia* |
| Monocyclic β-lactams | | *Nocardia* | *Pseudomonas*<br>*Chromobacterium*<br>*Acetobacter*<br>*Gluconobacter*<br>*Agrobacterium* |

**Table 1.2** Microorganisms producing β-lactams

Since the discovery of the therapeutic value of penicillin in 1941, the β-lactams as a group have emerged as the single most important class of agents in the clinical management of bacterial infections. Today, they command the largest market share for antibacterials, among other well known classes such as aminoglycosides, macrolides, tetracyclines and glycopeptides. β-Lactams which have been employed in clinical applications include some fifty penicillins, seventy cephalosporins, one monobactam and three carbapenems[18] and all exert bactericidal action through the inhibition of cell wall biosynthesis[19]. The majority of these clinically useful β-lactams are semi-synthetic in nature, being made commercially by the chemical modifications of fermentable products. Such modifications represent efforts to improve the efficacy of these antibiotics, through increasing their spectra of biological activity and stability to β-lactamases[20] (bacterial enzymes that inactivate β-lactams). With the bulk of the β-lactams being derived ultimately from fermentations, an understanding of the biosynthesis of these secondary metabolites is therefore crucial to all industrial scientists concerned with optimising their production. Interest among the academic scientists is fuelled by the novel and exciting chemistry employed by Nature in creating the strained β-lactam ring.

## 1.7 Biosynthetic routes to β-lactams

The biosynthesis of β-lactams, especially the penicillins and cephalosporins, has been reviewed on a frequent basis, the latest review appearing in 1993.[21] The extent of knowledge about the biosynthesis of each class of β-lactams is, not surprisingly, roughly proportional to their chronological age since discovery as well as their commercial value. Consequently, the penicillin-cephalosporin pathway is the most extensively studied, down to the detailed features of its enzymology[21] and molecular genetics,[22] while the monobactam pathway is only beginning to be understood. All β-lactams are derived from amino acids, either wholly or in part. A brief summary of the known biosynthetic routes to various classes of β-lactams is described below.

Biosynthetic pathways to the penicillins, cephalosporins and cephamycins are depicted in Figure 1.1. These pathways share in common the first two steps, starting from the condensation of three amino acids, *viz.* *L*-α-aminoadipic acid (17), *L*-cysteine (18), and *L*-valine (19), to give the tripeptide δ-(*L*-α-aminoadipoyl)-*L*-cysteinyl-*D*-valine (*LLD*-ACV, 20). The latter undergoes an oxidative bicyclisation to isopenicillin N (IPN, 21), mediated by the enzyme isopenicillin N synthase (IPNS).

The pathways for the penams and cephems diverge after IPN (21). In certain fungi including the *Penicillium* and *Aspergillus* spp., an acyltransferase enzyme exchanges the aminoadipoyl sidechain for a hydrophobic sidechain to give the hydrophobic penicillins, such as penicillin G (22), through the intermediacy of 6-aminopenicillanic acid (6-APA, 23). In other fungi such as *Cephalosporium* spp., as well as bacteria such as *Streptomyces* and *Flavobacterium* spp., an epimerase inverts the stereochemical configuration at the aminoadipoyl sidechain to produce penicillin N (24).[23, 24] The 5-membered thiazolidine ring in (24) is then transformed into the 6-membered dihydrothiazine ring in deacetoxycephalosporin C (DAOC, 25), catalysed by DAOC synthase, aka Ring Expandase. DAOC is next hydroxylated at the 3-methyl position to deacetylcephalosporin C (DAC, 26) by DAC synthase, aka DAOC hydroxylase.

DAOC synthase and hydroxylase are two discrete proteins in the prokaryotes,[25] but in *Cephalosporium* mould, the two sequential steps are mediated by a single bifunctional enzyme commonly known as Ring Expandase-Hydroxylase (REXH).[26-28] The evolutionary implication of this difference will be discussed in Chapter 2.

The second point of pathway divergence occurs at DAC (26). In *Cephalosporium acremonium*, an *O*-acetyltransferase acetylates the 3-hydroxymethyl group of DAC to yield the final biosynthetic product cephalosporin C (27).[29-31] In *Streptomyces clavuligerus*, DAC undergoes further elaboration to yield cephamycin C (28) as the final biosynthetic product. Current data support the conjecture that the latter transformation occurs through *O*-carbamoyl DAC (29) and 7α-hydroxy-*O*-carbamoyl DAC (30) successively.

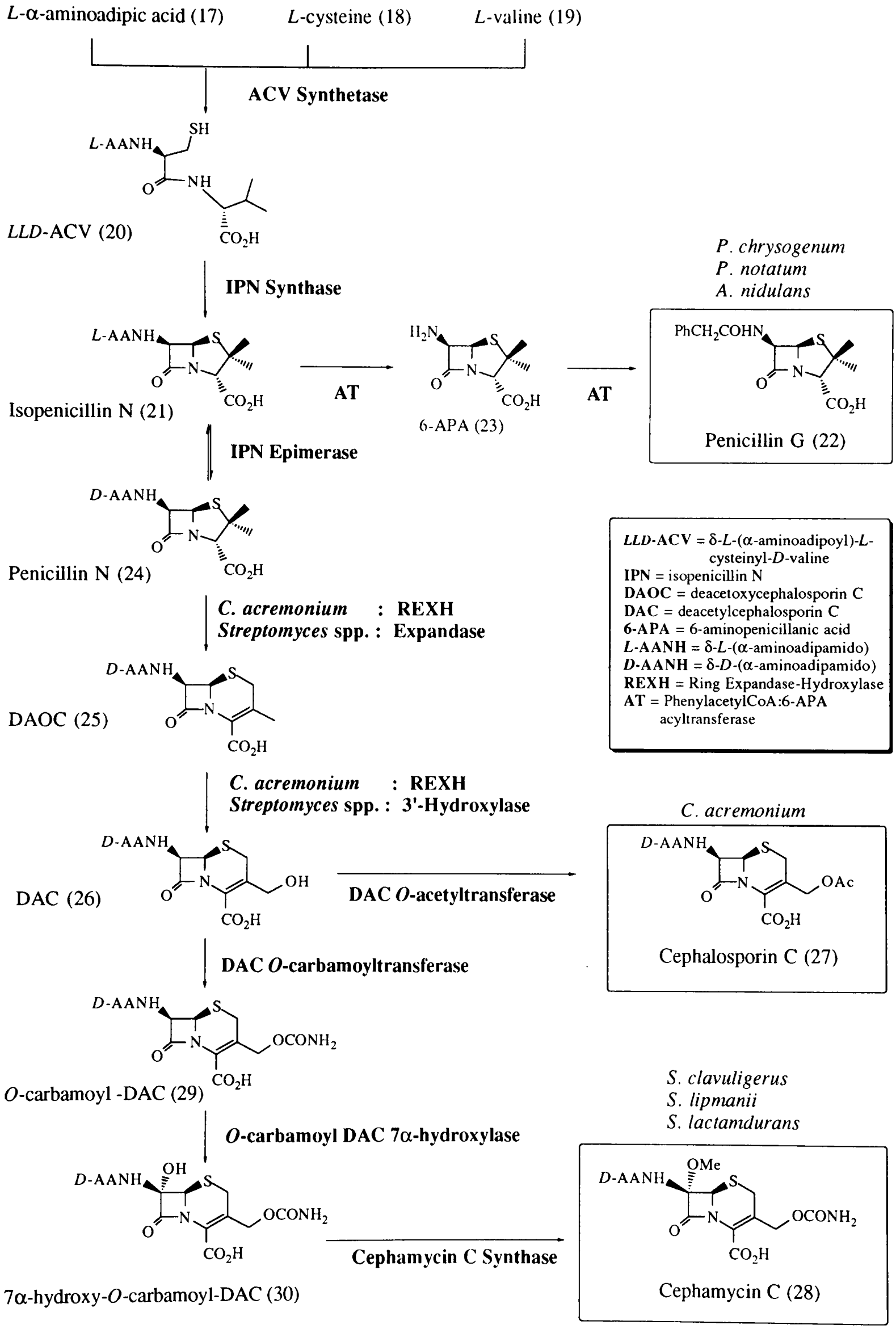



**Figure 1.1** The Biosynthesis of Penicillins, Cephalosporins and Cephamycins

Until now, the only clavam for which biosynthetic studies have been undertaken is clavulanic acid (31)[32], the biosynthesis of which will be treated in greater detail in Chapter 6. This substance possesses only weak antibacterial activity, but is a potent inhibitor of β-lactamases[33] and has been used in conjunction with some penicillins against penicillinase-producing bacteria. The biosynthetic scheme to clavulanic acid, which has not been completely elucidated, is depicted in Figure 6.3. Clavulanic acid (31) is a mixed-precursor metabolite, being derived from *D*-lactate (32), a $C_3$ intermediate from carbohydrate metabolism, and *L*-arginine (33).[34] The first biosynthetic intermediate to be discovered is proclavaminic acid (34), a monocyclic β-lactam.[21, 35, 36] This undergoes an oxidative cyclisation reaction to dihydroclavaminic acid (35) which is subsequently desaturated across the C-2/C-8 bond to generate clavaminic acid (36).[37, 38] Both steps are catalysed by an oxygenase called clavaminic acid synthase (CAS).[39]

The carbapenem nucleus is assembled from *L*-glutamate and acetate.[40] There is considerable structural diversity among the naturally occurring carbapenems, resulting in various other precursors being needed for the different substituents on the nucleus. These other precursors include *L*-cystine, β-alanine and *L*-methionine. It is remarkable that a carbapenem antibiotic, imipenem, is produced commercially by a > 20-stage total synthesis and sold at an exorbitant price due to its low titre in fermentation broths and high efficacy. This has been providing strong impetus for intensive research into carbapenem biosynthesis.

The two classes of monocyclic β-lactams, monobactams and nocardicins, derive their β-lactam carbons from *L*-serine.[41] The *N*-sulfonate group in monobactams originates from inorganic sulfur,[42] while the side-chains are elaborated from amino acids such as *D*-glutamate, *L*-alanine and *L*-methionine. The nocardicins contain side-chains which incorporated an unusual amino acid *L*-*p*-hydroxyphenylglycine. The detailed events in monobactam and nocardicin biosynthesis remain uncharacterised.

## 1.8 Oxygenases in β-lactam biosynthesis

Oxygenases are enzymes that utilise molecular dioxygen in their catalytic conversion of substrates.[43] They are ubiquitous in Nature, being extensively implicated in various aspects of primary and secondary metabolism. It is indeed surprising that the last comprehensive review on this immense and important group of enzymes dates back to about twenty years ago[43]. Since then, review articles tend to concentrate on specific sub-class of oxygenase enzymes. The oxygenases have been classified in several different ways depending on the emphasis of the investigators. Generally, they are divided into mono- and dioxygenases according to whether one or both atoms of the dioxygen molecule are incorporated into the substrates. This definition is somewhat arbitrary as there are dioxygenases with an oxygen atom ending up in water. Another classification scheme is based on the coenzyme/cofactor/cosubstrate requirements of the enzyme. Thus, oxygenases can be divided into the following categories[43]:

1. Haem-containing oxygenases
2. Non-haem iron oxygenases
3. α-Ketoglutarate-coupled dioxygenases
4. Flavoprotein oxygenases
5. Pterin-requiring aromatic amino acid hydroxylases
6. Copper-containing oxygenases, *etc.*

These categories are not mutually exclusive, nor are they representative of all oxygenases. Yet another empirical classification scheme is based on the type of reaction catalysed by the enzymes. Hence, oxygenases include among themselves hydroxylases, desaturases, cyclases, epoxidases and those that add oxygen atom to heteroatoms to give N-oxides, S-oxides and P-oxides. There has been no mechanistic classification scheme due to the generally poor understanding of the catalytic mechanisms of oxygenases.

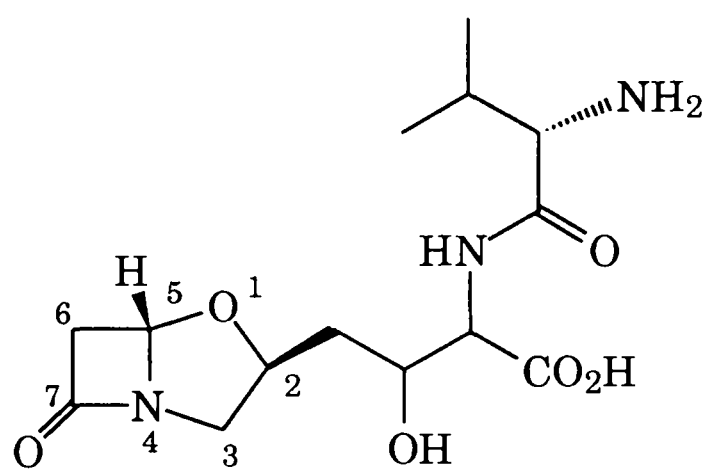
It is interesting to note the preponderance of oxygenase-mediated steps in the biosynthetic routes to various classes of β-lactams. Even more remarkably, the oxygenases known or speculated to be involved belong solely to the class of non-haem iron oxygenases. In the penicillin-cephalosporin-cephamycin pathway, the operative non-haem iron oxygenases are IPNS, DAOCS, DACS, DAOC/DACS (REXH), and cephalosporin 7α-hydroxylase. Except for IPNS, all of the above are dependent on α-KG as cosubstrate. In the pathway to clavulanic acid, the pivotal biosynthetic enzyme, CAS, is also an α-KG dependent oxygenase. In monobactam biosynthesis, it is believed that the methoxylation at C-3 is achieved by an initial hydroxylation mediated by a non-haem iron, α-KG dependent oxygenase[32]. Recently, gene sequences for IPNS, DAOCS, DACS and REXH have been reported.[21] Rigorous comparison showed a high degree of homology in some parts of the sequences, indicating a phylogenetic relationship among these enzymes. The significance of these observations will be elaborated in Chapter 2.

## 1.9 Aims of this project

In its broadest context, this project revolves around secondary metabolism in micro-organisms. It was intended to provide a wide exposure to the latest methodology in biosynthetic studies (as discussed in Section 1.4) that can be reasonably achieved within the time-scale of a D.Phil. candidature. In due course, two independent but related lines of research were adopted, which together provided useful training in various aspects of biosynthetic investigation.

The first piece of work concerns the biosynthesis of cephalosporin C (27) in *Cephalosporium acremonium* CO 728. It was intended to probe the structural and mechanistic features of Ring Expandase-Hydroxylase (REXH), the bifunctional non-haem iron, α-KG dependent oxygenase involved in the conversion of penicillin N (24) to DAC (26). It was envisaged that the studies would employ a range of techniques including protein separation technologies (FPLC and HPLC), selective cleavage, amino acid analysis, metal ion assay, mass spectrometry, inhibition kinetics, substrate analogue evaluation, *etc.*

The second line of research focused on the biosynthesis of valclavam (37), a clavam antibiotic produced by *Streptomyces antibioticus ssp. antibioticus* Tü 1718. There has not been any conclusive biosynthetic studies carried out on this compound since its isolation in 1985.[44] It was thus intended to 'start from scratch' to look for primary metabolites that serve as the basic building blocks in the assembly of this molecule. The work was envisaged to involve optimisation of fermentation conditions and isolation protocol, development of assay method for antibiotics, as well as whole-cell feeding of radiolabelled and other isotopically labelled compounds. It is obvious that the two independent lines of research converge on the central theme of β-lactam biosynthesis.




Valclavam (37)

# CHAPTER 2

# INTRODUCTION TO RING EXPANDASE-HYDROXYLASE

## 2.1 A brief history of REXH

The REXH-mediated stages in the biosynthetic route to cephalosporin C were established during the 1970's but homogeneous enzyme preparations were available only during the 80's. Thus, Kohsaka and Demain[45] were the first to describe a cell-free system from *C. acremonium* CW-19 that converted penicillin N (24) into a penicillinase-resistant, cephalosporinase-sensitive material which was subsequently identified by Yoshida *et al.* as DAOC (25).[46] That DAOC (25) is the precursor of DAC (26) was unequivocally demonstrated by the conversion of [3-*methyl*-$^3$H]-DAOC to [3-*hydroxymethyl*-$^3$H]-DAC with cell-free extracts of *C. acremonium*.[47] Further studies with crude preparations showed the stimulation of ring expansion by ferrous ions, α-KG and ascorbate.[48, 49] Similar cofactor/cosubstrate requirements were identified for 3'-hydroxylation activity in *C. acremonium*.[47] The direct involvement of molecular oxygen was first demonstrated by Stevens *et al.* who found partial incorporation of $^{18}O$ into cephalosporin C when *C. acremonium* was grown in an atmosphere of $^{18}O_2$.[50] Before its recognition as a bifunctional enzyme, the ring expansion and 3'-hydroxylation activities were purified independently in 1978[47] and 1983[48] respectively. In 1984, Scheidegger *et al.* speculated the involvement of a single bifunctional enzyme after failing to separate the two activities by a combination of chromatographic and isoelectric focusing techniques[28]. Purification of this enzyme to electrophoretic homogeneity was reported independently by two groups in 1987[26, 27], wherein the ring expandase activity was found to be associated with the 3'-hydroxylase activity in a constant ratio.[27] The bifunctionality of the enzyme was eventually established beyond doubt by the cloning and expression of its gene in an *E. coli* expression system[51]. The overexpressed protein was produced as inclusion bodies, from which active enzyme could be recovered by solubilisation with urea and subsequent refolding. Recently[52], a high-level soluble expression system has been developed which should pave the way for extensive biophysical studies into this enzyme.

## 2.2 A comparative survey of enzymes in the penicillin-cephalosporin-cephamycin biosynthetic pathway

The biosynthetic pathways to the classical β-lactams, *viz.* penicillins, cephalosporins and cephamycins, have been described briefly in Section 1.7. These pathways share a common starting point in three *L*-α-amino acids and diverge at two points, namely isopenicillin N (21) and DAC (26) (Fig. 1.1). At least eleven enzymes are known to be directly involved in the sum total of the pathways, and their salient features are summarised in Table 2.1.

| Enzyme | Organisms (where enzyme isolated) | MW range & structure | Gene cloned ? | Class of Enzyme |
|---|---|---|---|---|
| ACV Synthetase | *Penicillium chrysogenum*<br>*Cephalosporium acremonium*<br>*Streptomyces clavuligerus*<br>*Aspergillus nidulans* | 220 - 420 kDa<br>Possibly dimeric | (*pcb* AB)<br>Yes | Non-ribosomal peptide synthetase multi-enzyme complex, operating by the thiol-template mechanism |
| IPN Synthase | *Penicillium chrysogenum*<br>*Cephalosporium acremonium*<br>*Streptomyces clavuligerus*<br>*Streptomyces lactamdurans*<br>*Streptomyces lipmanii*<br>*Aspergillus nidulans*<br>*Lysobacter lactamgenus*<br>*Nocardia lactamdurans*<br>*Flavobacterium* | 36 to 39 kDa<br>Monomeric | (*pcb* C)<br>Yes | A unique oxygenase with unprecedented chemistry. Requires iron (II) for activity |
| Acyl-CoA : 6-APA acyltransferase | *Penicillium chrysogenum*<br>*Aspergillus nidulans* | 40 kDa<br>Heterodimeric | (*pen* DE)<br>Yes | ATP-dependent group transfer enzyme |
| IPN Epimerase | *Streptomyces clavuligerus*<br>*Streptomyces lipmanii*<br>*Nocardia lactamdurans* | 44 to 60 kDa<br>Monomeric | (*cef* D)<br>Yes | Pyridoxal-5'-phosphate dependent amino acid racemase |
| Ring Expandase (DAOCS) | *Streptomyces clavuligerus*<br>*Streptomyces lactamdurans* | 27 to 35 kDa<br>Monomeric | (*cef* E)<br>Yes | Non-haem iron, α-KG dependent dioxygenase |
| 3'-Hydroxylase (DACS) | *Streptomyces clavuligerus* | 34.6 kDa<br>Monomeric | (*cef* F)<br>Yes | Non-haem iron, α-KG dependent dioxygenase |
| Ring Expandase-Hydroxylase REXH (DAOC/DACS) | *Cephalosporium acremonium* | 36.5 kDa<br>Monomeric | (*cef* EF)<br>Yes | Non-haem iron, α-KG dependent dioxygenase |
| DAC *O*-acetyltransferase | *Cephalosporium acremonium* | 41 kDa<br>Monomeric | (*cef* G)<br>Yes | ATP-dependent group transfer enzyme |
| DAC *O*-carbamoyl transferase | *Streptomyces clavuligerus* | Not determined | (*cmc* H)<br>No | ATP-dependent group transfer enzyme |
| *O*-carbamoyl DAC 7α-hydroxylase | *Streptomyces clavuligerus* | 32 kDa<br>Monomeric | (*cmc* I)<br>No | Non-haem iron, α-KG dependent dioxygenase |
| Cephamycin C Synthase | *Streptomyces clavuligerus* | No MW reported<br>Monomeric | (*cmc* J)<br>No | Methyl transferase, utilising as donor S-adenosylmethionine |

**Table 2.1** Biosynthetic enzymes for penicillins, cephalosporins and cephamycins

From the table, it is evident that the biosynthesis of the classical β-lactam structures entails a rich variety of chemistry, ranging from more familiar types like peptide bond formation to the unique oxidative bicyclisation reaction of IPNS. Consequently, the biosynthetic enzymes cover a broad range of enzyme groups, including lyase, ligase, transferase, hydroxylase, cyclase and epimerase. It would be pertinent to make some comparisons between REXH, the subject of this thesis, and the other enzymes in the combined pathway (as listed in Table 2.1).

Whereas REXH is a bifunctional enzyme in *C. acremonium*, the same two steps are largely catalysed by two distinct and separable enzymes in *Streptomyces clavuligerus*[25]. The two enzymes, ring expandase (aka DAOC synthase) and 3'-hydroxylase (aka DAC synthase), belong to the same family of non-haem iron, α-KG dependent dioxygenases as REXH. The expandase exhibits hydroxylase activity at *ca.* 0.8% of its major activity, whereas the hydroxylase contains expandase activity at *ca.* 4% of its major activity. The degree of amino acid sequence identity among the three enzymes is very significant, being 57% between REXH and bacterial ring expandase, 54% between REXH and bacterial 3'-hydroxylase, and 59% between the two bacterial enzymes.[53] It has been postulated that their structural genes are products of divergent evolution from a common ancestral gene in the bacterium which coded for an ancient expandase with a minor hydroxylase activity[53] (Figure 2.1). A horizontal transfer to the fungus and subsequent modification resulted in a more balanced expandase and hydroxylase activities, leading to the modern bifunctional REXH. Meanwhile, duplication of the expandase gene in the bacterium enabled one copy of the gene to evolve and specialise in hydroxylase activity.

Another interesting relationship is found between REXH and IPNS. Whilst both are non-haem iron oxygenases of similar MW, requiring ascorbate and DTT for maximal activity, IPNS does not require α-KG as a cosubstrate. Instead, IPNS combines one dioxygen molecule with four hydrogen atoms to give two molecules of water, fully utilising the oxidative potential of dioxygen. Though REXH and IPNS share only a modest 21% amino acid sequence homology in *C. acremonium*, a 10-residue region containing Cys-100 of REXH is 50% homologous to the corresponding region containing Cys-106 of IPNS[51]. This region is of particular interest because Cys-106 in IPNS is important, though not absolutely essential, for substrate binding to IPNS[54]. It is proposed that the bacterial expandase gene was evolved from an ancient IPNS gene after the bacterium acquired an epimerase gene from a primary metabolic pathway[53]. The horizontal gene transfer from bacterium to fungus then involved a cluster containing genes for ACVS, IPNS, bacterial expandase and IPN epimerase. The epimerase bears no sequence homology to either REXH or IPNS, hence the speculation that it is 'borrowed' from a primary pathway. Figure 2.1 summarises the proposed evolutionary relationships discussed above.

Finally, it is worth mentioning that another non-haem iron, α-KG dependent dioxygenase, *viz. O*-carbamoyl DAC 7α-hydroxylase, appears in the pathway to cephamycin C. It would be of interest to obtain the sequence data for this enzyme and compare it with those of the oxygenases mentioned above. All in all, oxygenase-mediated transformations account for exactly half the number of biosynthetic steps to cephalosporin C (three out of six) as well as

those steps leading to cephamycin C (four out of eight), as depicted in Figure 1.1.

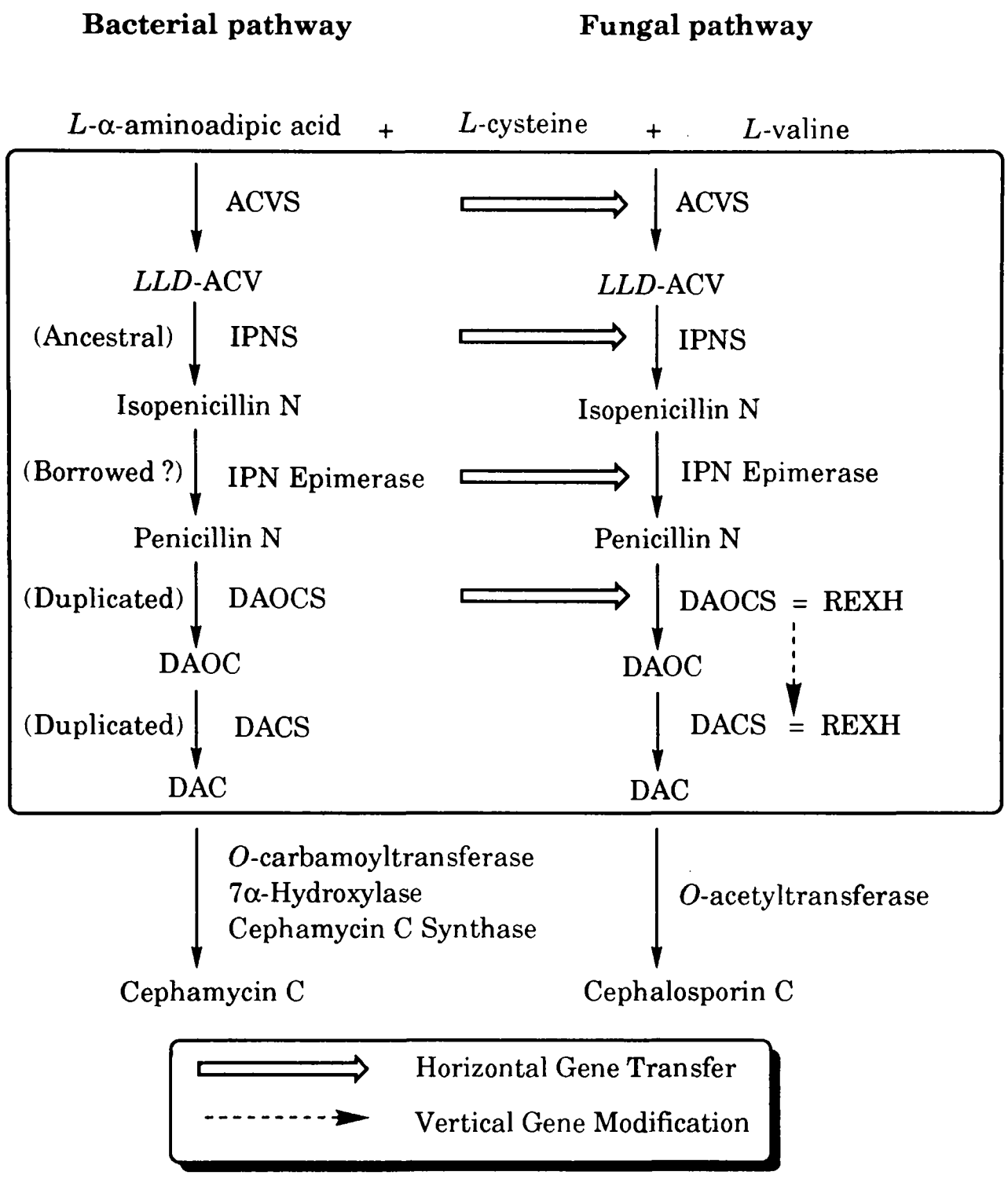



**Figure 2.1** Proposed metabolic evolution of genes coding for IPNS, IPN Epimerase, REXH, Bacterial Ring Expandase and Bacterial 3'-Hydroxylase

## 2.3 The non-haem iron, α-KG dependent dioxygenases

The occurrence of non-haem iron oxygenases among the known biosynthetic enzymes for various classes of β-lactams has been highlighted in Section 1.8. The probable evolutionary relationships among some of these enzymes were explored in the last section. This section is an attempt to examine the salient features that define this growing family of non-haem iron, α-KG dependent dioxygenases, of which REXH is a member. This family can be considered as the biggest subset of non-haem iron oxygenases, as shown in Figure 2.2.

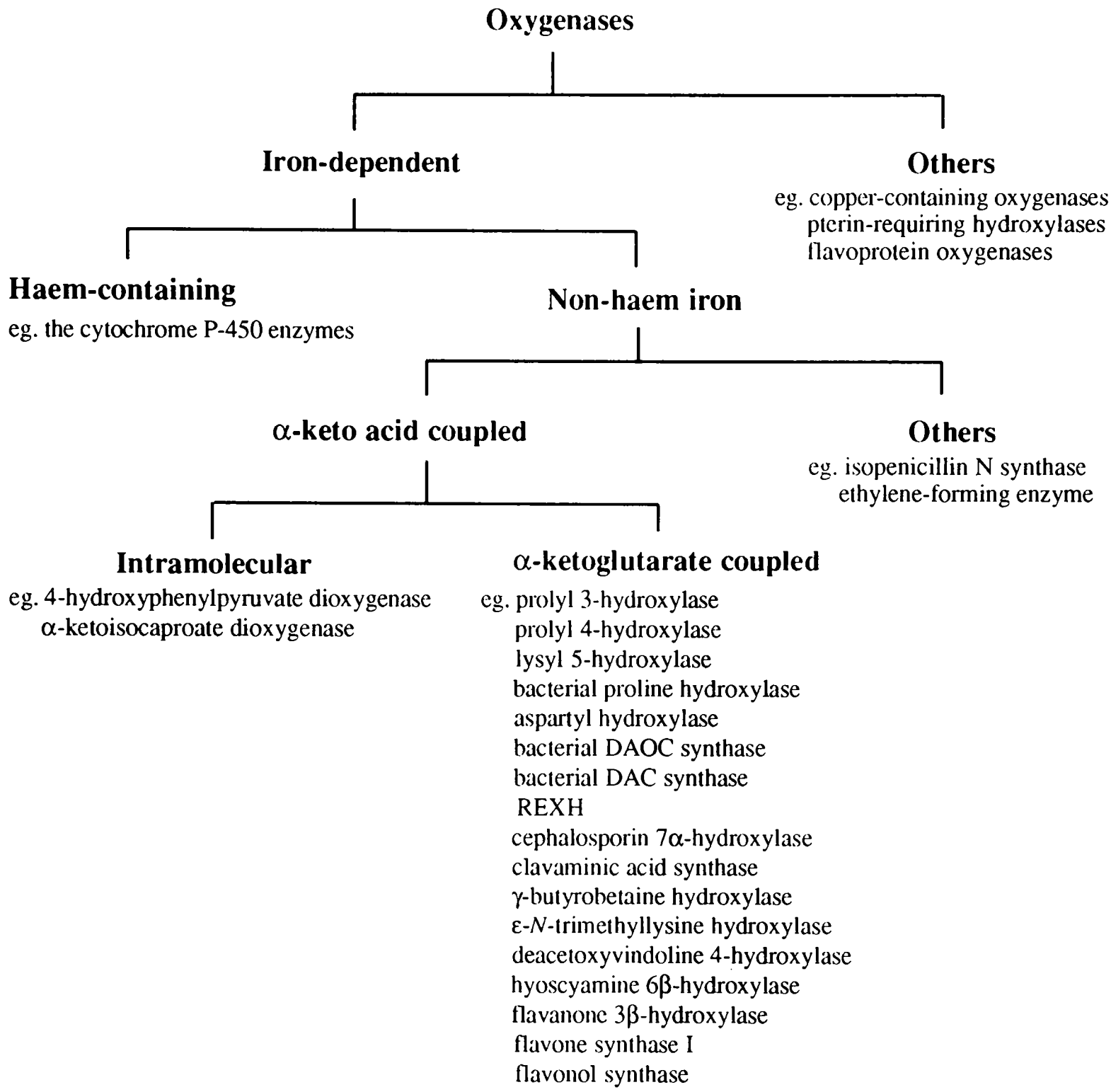



**Figure 2.2** Classification of the oxygenases

The last comprehensive review of α-KG dependent dioxygenases appeared about twenty years ago[55]. Since then, a lot of new knowledge about this group of enzymes has been obtained and is summarised here. Some characteristics of these enzymes are:

(1) The iron atom in the resting enzyme exists in the ferrous state, its coordination sphere including two or three histidyl nitrogen atoms[56, 57].

(2) Substrate turnover is accompanied by a stoichiometric consumption of α-KG and dioxygen in the ratio 1:1:1, with the release of product, succinate and carbon dioxide in the same stoichiometric amounts.[58]

(3) The kinetic scheme for these enzymes belong to an Ordered Ter Ter mechanism where the substrates bind in the order: α-KG, dioxygen, main substrate, and products are released in the order: oxidised product, carbon dioxide, succinate.[59, 60]

(4) A number of these enzymes have been able to demonstrate uncoupled decarboxylation of α-KG both in the absence and presence of the principal substrate.[58, 61]

(5) Ascorbate is usually required for optimal turnover of substrate *in vitro*, probably to keep the iron in the ferrous state.[26, 60, 61]

(6) An additional reducing agent, such as dithiothreitol (DTT), dithioerythritol (DTE) or glutathione (GSH) can stimulate these enzymes, probably by maintaining the integrity of some catalytically essential thiol groups. The enzymes are inactivated by thiol-quenching agents such as *N*-ethylmaleimide.[26, 62]

(7) Catalase enhances the activity of some of these enzymes and protects against inactivation.[63, 64]

(8) Divalent metal ions, such as $Cu^{2+}$, $Zn^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Ni^{2+}$, $Cd^{2+}$ and $Hg^{2+}$ can inhibit these enzymes.[27, 61, 62, 65] These divalent ions presumably compete for the ferrous binding site but are not capable of activating molecular oxygen.[58]

Since the first recognition of an α-KG dependent enzyme in 1967,[55] the list of identified members of this family of enzymes has grown continuously, occurring in animals, plants and microorganisms. They have been found to perform biologically important roles in the metabolism of amino acids, peptides, β-lactams, alkaloids, flavonoids, isoprenoids, macrolides, pyrimidines and nucleosides.

Recently, two papers[56, 57] reported some very insightful sequence comparisons among the α-KG dependent enzymes and other non-haem iron oxygenases such as IPNS and ethylene-forming enzyme. The most important observation is that all these enzymes possess two highly conserved regions, each containing a histidyl residue in the corresponding position. The authors rationalise that the ferrous ion is coordinated to at least two histidyl nitrogen atoms in the active site. They also suggest that the high sequence homology provides strong evidence for a common ferryl-oxo intermediate as the reactive iron species among these enzymes, a notion first proposed by Siegel[66]. At this stage, it is open to speculation whether the non-haem iron oxygenases as a whole, or even the smaller subset of α-KG dependent enzymes, arise through function-directed convergent evolution or divergent evolution from an ancestral oxygenase. This question stands apart from the rational proposal for metabolic evolution of the β-lactam biosynthetic genes described in the last section.

## 2.4 The place of REXH in industry

Granted that most β-lactam antibiotics in the market are manufactured by semi-synthetic processes (one exception being the 22-step total synthesis of imipenem), REXH as a biosynthetic enzyme commands notable industrial attention for the following reasons.

Firstly, it provides the biosynthetic entry point to the cephem antibiotics, which are clinically superior to the penicillins by virtue of their (a) greater stability to β-lactamases, (b) wider antibacterial spectra and (c) lower tendency to evoke allergic reactions. One has only to

refer to past issues of the British National Formulary to realise that the ratio of approved cephalosporins to penicillins has been on the rise in Britain. Another factor responsible for the increased clinical interest in the cephem ring system is the recent emergence of the so-called dual-action cephalosporins[67]. When cephalosporins exert their biological activity by reacting with bacterial enzymes, opening of the β-lactam ring can lead to expulsion of the 3'-substituent. If the released entity has its own antibacterial activity, the spectrum of the cephalosporin is effectively broadened. Studies have proven such a dual mode of action for these compounds[67].

Given the superior utility of the cephem ring system, it is an unacceptable fact that the most popular industrial production strain of *C. acremonium* excreted penicillins into the fermentation broth, thereby reducing the biosynthetic throughput to cephalosporin C[68]. This suggests that REXH is the major rate-limiting enzyme in cephalosporin C biosynthesis and hence a bottleneck to industrial production of antibiotics derived from cephalosporin C. With such a pivotal role played by REXH in the industrial production of β-lactams, it is not difficult to imagine the immense benefits that can be accrued from an elaborate genetic and protein engineering programme on the enzyme. Various possibilities are described below, with some already being pursued to quite an advanced stage:

(1) An obvious way to circumvent the problem of rate-limiting activity of REXH is to induce a higher rate of transcription in the fungus. This has been achieved by cloning back an extra copy of the REXH gene into a production strain of *C. acremonium*[69]. The doubling in gene dosage led to enhanced REXH activity, decreased penicillin excretion and up to 40% improvement in yield of cephalosporin C. Regulatory approval for using the engineered strain have been granted, and tests at industrial scale have begun[70].

(2) The compound 7-aminodeacetoxycephalosporanic acid (7-ADCA, 38) is an important intermediate useful in the manufacture of cephalexin and other oral cephalosporins. It is currently produced by a process involving four chemical and one enzymatic steps, as depicted in Figure 2.3. The chemical steps entail extensive solvent handling and usage of raw materials. However, production of 7-ADCA from a fermentation-derived deacetoxycephalosporin V (DAOV, 39) would be simple and efficient because (a) DAOV can be extracted readily from broth using protocols similar to those established for the hydrophobic penicillins and (b) enzymes exist that efficiently cleave the acyl group in DAOV to yield 7-ADCA. If the gene for REXH, *cef*EF, can be altered by site-directed mutagenesis to encode a form of REXH that efficiently converts penicillin V (40) to DAOV and not further, then expression of that modified gene in *P. chrysogenum* would be of interest. Efficient expression might allow direct production of DAOV by a *P. chrysogenum* transformant bearing, for example, a hybrid gene consisting of the promoter from the *P. chrysogenum pcb*C gene (*pcb*C encodes IPNS) and the open reading frame from the modified *cef*EF gene.

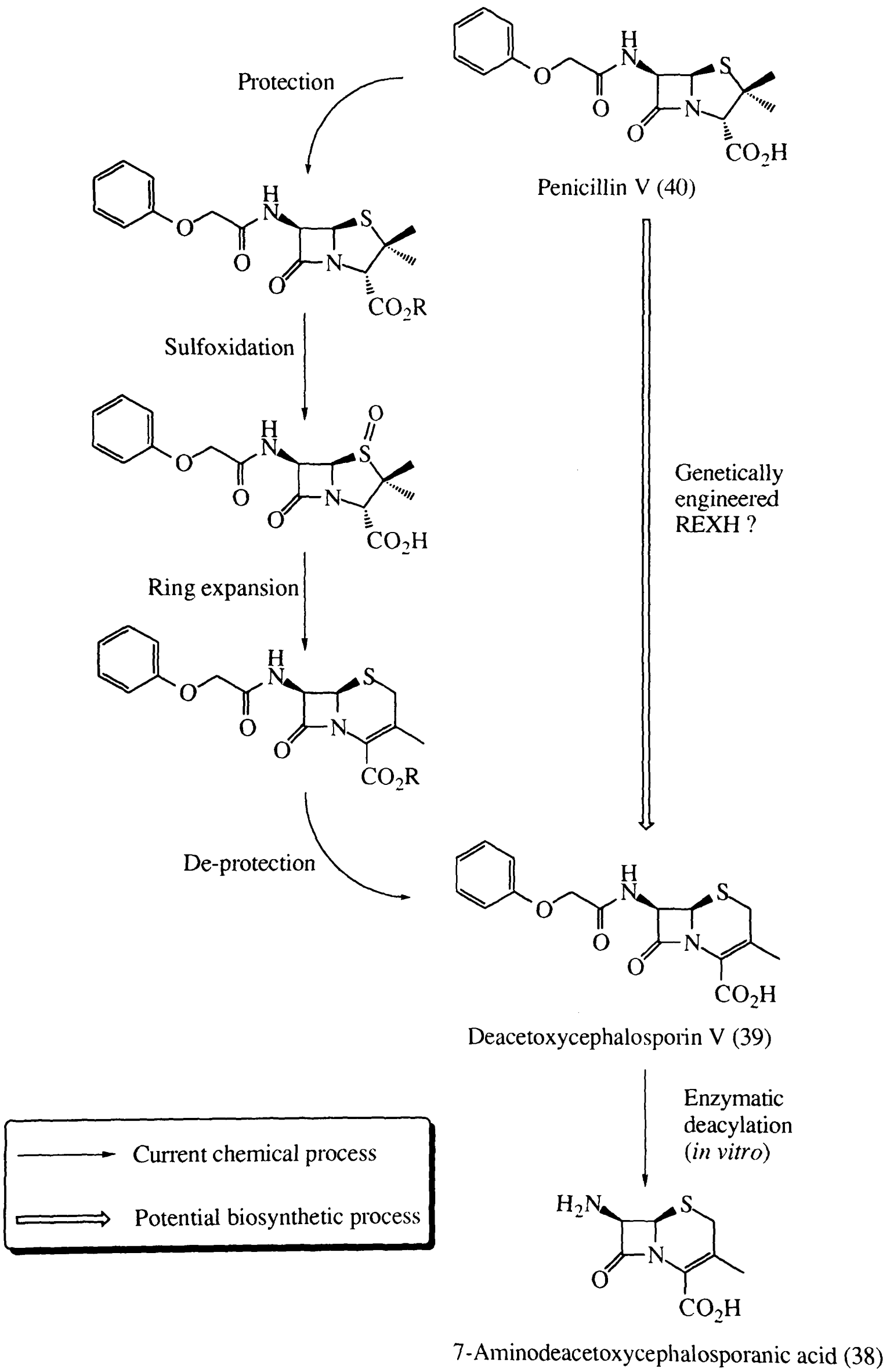
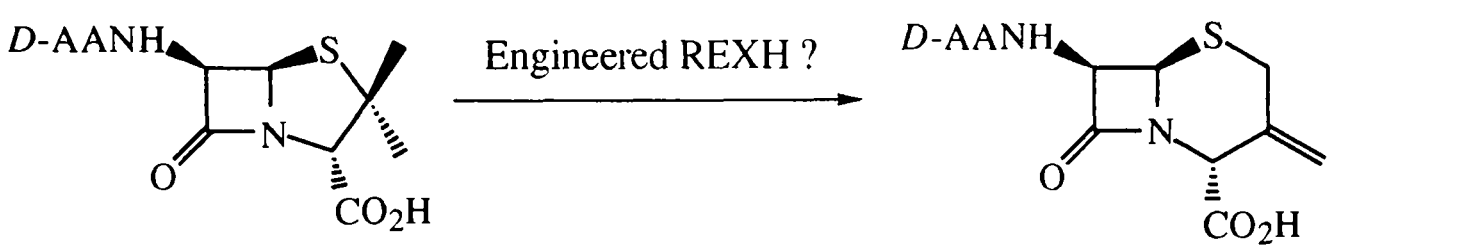



**Figure 2.3** Potential engineering of REXH to modify substrate specificity

(3) Another strategy for obtaining 7-ADCA would involve selective inactivation of the hydroxylase function of REXH to facilitate the accumulation of DAOC which can be enzymatically deacylated *in vitro* to 7-ADCA. Preliminary attempts by scientists at Eli Lilly utilising classical mutation-screening and site-directed mutagenesis have resulted in simultaneous loss of both functions[68].

(4) Another potential for industrial application comes from alteration of REXH for the conversion of penicillin N to exomethylene cephalosporin C (EMCC, 41), a key intermediate in the manufacture of cefaclor, an important oral cephalosporin.




Penicillin N (24) Exomethylene cephalosporin C (41)

Obviously, the sort of enzyme engineering discussed above is possible only if the active-site map and the molecular events in catalysis by REXH are understood in unambiguous terms. Insufficient knowledge in these areas provides the *raison d'être* for continued efforts in probing the structural and mechanistic aspects of this interesting enzyme, such as the work carried out in this project.

# CHAPTER 3

# PURIFICATION AND REFOLDING OF REXH

## 3.1 Introduction

"Don't waste clean thinking on dirty enzymes" - so goes a well-known dogma that lies at the heart of good enzymological practice[71]. In essence, it conveys the simple fact that meaningful conclusions about the individual properties of an enzyme can only be drawn from studies where the active enzyme has been well separated from other proteins and molecules in the crude cell extract. In all likelihood, the mixture of thousands of different enzymes released from a disrupted mammalian, yeast or bacterial cell will contain several that can act on the starting material or product of a particular enzyme's action. Only after the properties of the pure enzyme are known is it profitable to examine its behaviour in the crude state. In practice, no enzyme is purified to the point of absolute homogeneity. It is, however, possible to reach a state where the contaminants (other proteins and molecules) have only negligible effect on the enzyme activity.

Prior to the commencement of this project, native REXH from *Cephalosporium acremonium* was first purified partially by two groups[28, 72] in 1982/83, and subsequently to near homogeneity by another two groups.[26, 27] Protein separation technology is not a static science, and new chromatographic and electrophoretic methods are being continually developed to meet new challenges posed by advances in protein research. It is with a view to capitalising on the state-of-the-art technology that efforts were expended in this project to devise better means of purifying native REXH. Purification and refolding of recombinant REXH, overexpressed in *E. coli* as inclusion bodies,[51] was also subjected to optimisation studies for similar reasons. During the lifetime of this project, a high-level soluble expression system was developed for recombinant REXH by a collaborating group[52] and the soluble enzyme was subjected to its first purification studies. Purification studies in this project was aimed at furnishing enzyme of >95% electrophoretic purity by densitometric scanning, which would suffice for the investigations described in later chapters.

## 3.2 Experimental

### 3.2.1 Fermentation of *C. acremonium* CO 728 for the production of native REXH

Fermentation runs were routinely carried out by Mr John Keeping of the Microbiology Section, Dyson Perrins Laboratory. Glycerol freezes were kept at -80°C. The organism was grown at 27°C, for 10 days, on slopes of Complete Agar [Composition (g/l): sucrose, 20.0; $CaCO_3$, 5.0; $CaSO_4$, 12.5; ammonium acetate, 8.0; Oxoid L44 soya-peptone, 60.0; agar, 15.0;

pH 6.4]. Flasks of Seed Medium [$CaCO_3$, 3.0; ammonium sulfate, 1.0; corn steep liquor (Sigma 4648), 30.0; Oxoid L44 soya-peptone, 10.0; pH 7.0] were inoculated with mycelial suspension in sterile deionised water prepared from surface growth of well grown slopes. Flasks were incubated at 27°C and 200 rpm on a rotary shaker (2-inch throw) for 90 hr. This inoculum was then used to inoculate either further shake flasks or a fermenter.

(a) Shake flasks

Seed culture (5 ml) was used to inoculate each 500-ml Erlenmeyer flask containing 100 ml of Fermentation Medium [Composition (g/l): sucrose, 20.0; $CaCO_3$, 5.0; $CaSO_4$, 12.5; ammonium acetate, 8.0; Oxoid L44 soya-peptone, 30.0; corn steep liquor (Sigma 4648), 20.0; pH 7.0]. The flasks were then incubated at 27°C and 200 rpm on a rotary shaker (2-inch throw) for 54 hr.

(b) Fermenter

Ingredients for 25 litres of Fermentation Medium (composition as for shake flask growth) were dissolved in 23 litres of deionised water and sterilized *in situ* in a fermenter vessel. The contents of seed flasks were pooled to give 2 litres of inoculum. Temperature was controlled at 27°C, with an agitation speed of 600 rpm, and aeration rate of 20 litres of bacteriologically filtered air per min. Polypropylene glycol 2025 (BDH) was used as an anti-foam agent. Fermentations were run for 60 hr.

### 3.2.2 Purification of native REXH

All enzyme preparations were carried out between 0°C and 4°C throughout. Mycelia from the fermentations of *C. acremonium* CO 728 were harvested by centrifugation at 28 000 *g* for 1 hr. The compacted mycelia was weighed, packed into polythene bags and kept frozen at -80°C until further use. In a typical enzyme preparation, 60 g wet weight of frozen mycelia was thawed in 150 ml of the extraction buffer in a glass beaker. The extraction buffer contained 2 mM DTT, 0.015% (w/v) sodium azide and 1 mM PMSF in 50 mM Tris-HCl, pH 7.5. Once completely thawed, the mycelia was lysed by sonication using a 380 W sonicator operating with output control at level 5. The suspension was given 6 bursts of 30 sec each, with a 1-min cooling interval between consecutive bursts. After sonication, the cell debris was removed by centrifugation at 28000 *g* for 30 min. Aqueous protamine sulfate (5% w/v) was added to the supernatant (200 ml) over 10 min with stirring until the final concentration was 0.5%. The suspension was equilibrated for a further 10 min. The precipitated nucleic acids were removed by centrifugation at 28000 *g* for 20 min. About 220 ml of supernatant was obtained, to which 71.7 g solid ammonium sulfate was added over 10 min, with stirring, to give a 55% saturated solution. The suspension was equilibrated for another 20 min before centrifugation at 28000 *g* for 30 min. The precipitated protein contained only a negligible amount of REXH and was discarded. The 270 ml of supernatant was then raised to 75% saturation in ammonium sulfate by adding 34.3 g of the salt over 10 min with stirring, followed by 20 min of equilibration. The

precipitated protein was collected by centrifugation at 28000 *g* for 30 min. The pellet, containing REXH, was resuspended in buffer A [2 mM DTT, 0.015% (w/v) sodium azide in 50 mM Tris-HCl, pH 7.5] ready for chromatographic purification.

The pellet resuspension was filtered through a 0.22 μ filter and loaded directly onto a Superdex 200 prep grade 35/600 column, previously equilibrated with buffer A. Fractions of 3 ml were collected between elution volumes of 310 ml and 460 ml. Protein concentration was determined by the method of Bradford[73], while enzyme activity was measured by bioassay with *E. coli* X580. The fractions were also analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using the method of Laemmli[74]. Active fractions (as judged by bioassay) which were reasonably clean (as judged by SDS-PAGE) were pooled together and loaded onto a Mono Q HR 10/10 column, previously equilibrated using buffer A and buffer B (which is 1 M NaCl in addition to buffer A). The column was washed with 80 ml of buffer A, after which a linear gradient of 0% buffer B to 20% buffer B was applied over 100 ml. Fractions of 5 ml were collected, assayed for enzyme activity and protein content and analysed by SDS-PAGE. The active and satisfactorily purified fractions were pooled and concentrated using an Amicon stirred ultrafiltration cell followed by a Centricon microconcentrator. The protein concentrate was then loaded onto a Superdex 75 10/30 column, previously equilibrated with buffer A. Fractions of 0.5 ml were collected, assayed for protein content and enzyme activity and analysed by SDS-PAGE. The purest fraction at this point was > 95% pure by densitometric scanning.

### 3.2.3 Fermentation of *E. coli* NM 554 / pIT 511 for the production of insoluble recombinant REXH

Fermentation runs were routinely carried out by Mr John Keeping. Glycerol freezes were kept at -80°C. The seed inoculum was grown on 100 ml of LB Medium [Composition: Oxoid L42 Tryptone, 10.0; Oxoid L21 yeast extract, 5.0; NaCl, 10.0; pH 7.5] in 500-ml Erlenmeyer flasks. Each flask was supplemented with 0.1 ml of tetracycline (Sigma 10 mg/ml stock solution, in 50:50 EtOH/$H_2O$) to give a final concentration of 10 μg/ml. Flasks were incubated overnight at 27°C and 250 rpm on a rotary shaker (1-inch throw). Ingredients for 25 litres of LB Medium (composition as above) were dissolved in 22.5 litres of deionised water and sterilized *in situ* in fermenter vessel. After cooling, it was supplemented with 25 ml of tetracycline (composition as above) to give a final concentration of 10 μg/ml, and also with *D*-glucose (50 g in 250 ml, sterile-filtered) to give a final concentration of 0.2%.

Seed flasks were pooled to give 2 litres of inoculum. Fermentation was run at 30°C. bacteriologically filtered air was sparged at a rate of 10 litres/min with an agitation of 600 rpm. Polypropylene glycol 2025 (BDH) was used as an anti-foam agent. After 340 min of fermentation time, the temperature was switched to 42°C to trigger expression of the recombinant REXH gene.

The fermentation was run for a further 80 min and then harvested by means of a Sartocon II tangential flow concentrator, and then centrifuged at 9000 rpm for 40 min using a Beckman J2-21 centrifuge with a JA-10 rotor.

### 3.2.4 Purification and refolding of insoluble recombinant REXH

In a typical enzyme preparation, 5 g wet weight of frozen cells of *E. coli* NM 554 / pIT 511 were thawed in 10 ml of GEDA buffer (10% glycerol, 10% ethanol, 10 mM DTT and 10 mM ascorbate in 15 mM Tris-HCl, pH 7.5). The cells were sonicated four times for 20 sec each using a Soniprep 150 sonicator at maximum power, interspersed by 1 min of cooling between consecutive bursts. The cell debris was removed by centrifugation at 500 *g* for 30 min. The supernatant was subsequently centrifuged at 12000 *g* for 20 min to collect the REXH granules. The granular pellet (~ 2.4 g) was washed with 2 M urea (40 ml per g of granules) and centrifuged at 28000 *g* for 30 min. The supernatant, containing solubilised proteins but not REXH, was discarded. The pellet was resuspended in 5 M urea to solubilise the REXH present. The pH of the resuspension was adjusted to 11.5 with 1 M NaOH. The pH jump from *ca.* 8.0 - 8.5 enhanced solubilisation and unfolding of the REXH molecules. Any insoluble granular components were removed by centrifugation at 28000 *g* for 40 min. The resulting solution was diluted into a preliminary 'folding' buffer, consisting of 5 M urea and 5 mM DTT in 5 mM Tris-HCl, pH 9.5, to give a final protein concentration of 0.2 mg/ml. The solution was stirred at 4°C for 2 hr to equilibrate, after which further clarification was achieved by passing through a 0.22 µm filter. This solution was then loaded onto a Mono-Q HR 10/10 column, previously equilibrated with buffers C and D. Buffer C contained 5 M urea and 0.5 mM DTT in 5 mM Tris-HCl, pH 9.5, while buffer D contained 1 M KCl in addition to buffer C. The column was washed with 20 ml of buffer C, followed by a linear gradient of 0% to 25% buffer D over 200 ml. Fractions of 3 ml were collected, assayed for protein concentration and enzyme activity, and analysed by SDS-PAGE. The most active fractions, having greater than 90% purity, were pooled together for refolding. The refolding was achieved by stepwise dialysis through the following urea concentrations in 5 mM Tris-HCl, pH 9.5 with 5 mM DTT : 5 M to 3 M to 1 M to 0.5 M. Each dialysis step was carried out for 2 hr at 4°C. The activity was determined at each stage of the refolding to assess the efficiency of folding.

### 3.2.5 Fermentation of *E. coli* JM 109 / pRH 1091 for the production of soluble recombinant REXH

Shake-flask fermentations of *E. coli* JM 109 / pRH 1091 were devised and carried out by Mr Richard Heath (Dyson Perrins Laboratory) as part of his doctoral research[75]. Glycerol freezes of the bacterium were stored at -80°C. Starter cultures were prepared by using *ca.* 10 µl of these stocks to inoculate 5 ml of sterile LB medium [Composition (g/l): Oxoid L42 Tryptone, 10.0; Oxoid L21 yeast extract, 5.0; NaCl, 10.0] or 2x TY medium [Composition (g/l): Oxoid L42

Tryptone, 16.0; Oxoid L21 yeast extract, 10.0; NaCl, 5.0] containing chloramphenicol at 20 µg/ml. These cultures were grown overnight at 27°C, 250 rpm and used as 1% inocula for sterile 100 ml 2x $TY_{cam}$ medium, contained in 500 ml Erlenmeyer flasks. IPTG was added to a final concentration of 1 mM. Flasks were incubated at 30°C, 250 rpm. Growth was monitored by change in absorbance at 550 nm. Cells were harvested at the desired stage of growth ($OD_{550} \approx$ 6) by centrifugation (13 500 rpm, 15 min). Cell pellets were weighed, packed into polythene bags and stored at -80°C until further use.

### 3.2.6 Purification of soluble recombinant REXH

In a typical enzyme preparation using cells of *E. coli* JM 109 / pRH 1091, 6 g wet weight of frozen cells was thawed in 24 ml of buffer A [50 mM Tris-HCl, pH 7.5; 2 mM DTT]. The suspension was sonicated 5 times for 30 sec each, with a 1-min cooling interval in between. 8.7 mg of PMSF (in 1 ml ethanol) was added after the initial burst, giving a final concentration of 1 mM. After the third burst, 360 µg of DNase I (in 1 ml of buffer A) was added to reduce the viscosity caused by DNA release. The cell debris was removed by centrifugation at 46000 *g* for 30 min. The 25 ml of supernatant was filtered through a 0.22 µm filter and loaded onto a Q-Sepharose Fast Flow 16/10 column previously equilibrated with buffer A and buffer B. The column was washed with 80 ml of buffer A, after which a linear gradient of 0% to 25% buffer B was applied over 290 ml. Fractions of 10 ml were collected, assayed for protein content and enzyme activity, and analysed by SDS-PAGE. The active and reasonably pure fractions were pooled together and concentrated by ultrafiltration to about 10 ml. The concentrate was directly loaded onto a Superdex 200 prep grade 35/600 column, previously equilibrated with buffer A. Fractions of 15 ml were collected between elution volumes of 250 and 450 ml. The fractions were assayed for protein content and enzyme activity and analysed by SDS-PAGE. The purest fraction at this point was >95% pure by densitometric scanning.

### 3.2.7 General techniques

#### 3.2.7.1 Fast Protein Liquid Chromatography (FPLC)

All protein chromatographic procedures were performed with two FPLC systems from Pharmacia, a standard system and a Bio-Pilot system.

The standard system was configured with the following components: an LCC-500 (or LCC-500 Plus) controller, two P-500 pumps, a solenoid valve, a 0.6 ml fluoroplastic mixer, an MV-7 and two MV-8 control valves, a UV-1 (or UV-M) monitor and a Frac-100 (or Frac-200) fraction collector.

The Bio-Pilot system was configured with the following components: an LCC-500 Plus controller, two P-6000 pumps, a solenoid valve, a 6 ml mixer, an IMV-7 and four IMV-8 control valves, a UV-1 monitor and a Frac-100 (or Frac-200, or Frac-300) fraction collector.

Unless otherwise specified, all columns and resins were purchased from Pharmacia. Conditions for operating the systems and using the columns were made as close as possible to the recommendations in the commercial literature from Pharmacia.

### 3.2.7.2 Determination of protein concentration

Protein concentration was routinely determined by the method of Bradford[73], using a modified reagent made by mixing 100 mg of Coomassie Brilliant Blue G-250, 50 ml of 95% ethanol and 100 ml of 88% phosphoric acid, and made up to 1 litre with Milli-Q grade water. Protein solution (40 µl) was thoroughly mixed with 2 ml of the modified Bradford reagent in a 3-ml cuvette. After standing at room temperature for 10 min, the absorbance at 595 nm was measured. A blank reading was obtained by substituting buffer solution for the protein sample. Protein content was estimated by reference to a calibration curve obtained with standards of bovine serum albumin in the same buffer. Excessively concentrated protein samples were diluted to the calibration range to ensure higher accuracy.

### 3.2.7.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was carried out using a Bio-Rad Mini-Protean II system essentially according to the method of Laemmli[74]. REXH preparations were routinely analysed using a 3% stacking gel, a 12.5% (or 15%) resolving gel and a running buffer of 0.025 M Tris, 0.2 M glycine and 0.5% (w/v) SDS. The following table gives the composition for the requisite gels :

| Component | Final acrylamide concentration | | |
|---|---|---|---|
| | 3% (stacking) | 12.5% (resolving) | 15% (resolving) |
| 30 % (w/v) acrylamide<br>0.8% (w/v) bis-acrylamide | 0.32 ml | 2.5 ml | 3.0 ml |
| 0.75 M Tris-HCl, pH 8.8 | -- | 3.0 ml | 2.7 ml |
| 1M Tris-HCl, pH 6.8 | 0.40 ml | -- | -- |
| Milli-Q water | 2.32 ml | 0.2 ml | -- |
| TEMED | 5 µl | 10 µl | 10 µl |
| 10% SDS | 32 µl | 60 µl | 60 µl |
| 1% ammonium persulfate | 0.16 ml | 0.3 ml | 0.3 ml |

Resolving gels were overlaid with water to create a level gel slab, while stacking gels were fitted with either a 10- or 15-tooth comb for making sample wells. Protein samples were routinely prepared for gel electrophoresis by heating a mixture of protein and sample application

buffer [Composition: 0.625 M Tris-HCl, pH 6.8, 50% (v/v) glycerol, 10% (w/v) SDS, 0.01% (w/v) bromophenol blue and 1% (v/v) 2-mercaptoethanol] in a 4:1 ratio for 5 min at 100°C. To concentrate protein samples, an aqueous solution 100% (w/v) trichloroacetic acid was added to a final concentration of 10% (w/v). After centrifugation for 5 min at 13 000 rpm in a microfuge, the sample was dissolved in a 3:1 mixture of sample application buffer and saturated Tris base, and boiled for 10 min with periodic vortex mixing. Samples were electrophoresed at a constant current of 30 mA per gel.

Gels were routinely stained and fixed using 0.25% (w/v) Coomassie Brilliant Blue R-250 in aqueous 5% (v/v) acetic acid, 10% (v/v) ethanol at 50 to 70°C for 1 to 2 hr, and destained in the same solution lacking the dye. The following molecular weight standards were used as markers: hen egg ovotransferrin (77.0 kDa), bovine serum albumin (66.3 kDa), hen egg ovalbumin (45.0 kDa), bovine erythrocyte carbonic anhydrase (30.0 kDa), equine myoglobin (17.2 kDa) and equine cytochrome C (12.3 kDa).

### 3.2.7.4 Non-denaturing polyacrylamide gel electrophoresis (ND-PAGE)

To investigate the mass (and hence the aggregation state) of REXH in the native state, the technique of non-denaturing polyacrylamide gel electrophoresis (ND-PAGE) was used. The protocol employed was adapted from the commercial literature accompanying the 'Non-denaturing Protein Molecular Weight Markers Kit' supplied by Sigma Chemical Company. Gels were cast in the same manner as for SDS-PAGE, but according to the composition given in the table below.

| Component | Final acrylamide concentration | | | | |
|---|---|---|---|---|---|
| | Stacking | 7% | 8% | 9% | 10% |
| 3.0 M Tris-HCl, pH 8.9<br>0.23% (v/v) TEMED | -- | 6.0 ml | 6.0 ml | 6.0 ml | 6.0 ml |
| 0.5 M Tris-HCl, pH 6.7<br>0.46% TEMED | 8.0 ml | -- | -- | -- | -- |
| 28% (w/v) acrylamide<br>0.74% bis-acrylamide | -- | 12.0 ml | 13.7 ml | 15.4 ml | 17.1 ml |
| 10% (w/v) acrylamide<br>2.5% bis-acrylamide | 16.0 ml | -- | -- | -- | -- |
| 0.04 mg/ml riboflavin | 8.0 ml | -- | -- | -- | -- |
| 5% (w/v) sucrose | 32.0 ml | 27.0 ml | 25.3 ml | 23.6 ml | 21.9 ml |
| 8 mg/ml ammonium persulfate | -- | 3.0 ml | 3.0 ml | 3.0 ml | 3.0 ml |

The electrode buffer was identical to that used for SDS-PAGE except for the exclusion of SDS. Protein samples were mixed thoroughly with twice their volume of sample application

buffer [Composition in 3.25 ml: 0.5 M Tris-HCl (pH 6.7), 1 ml; glycerol, 1 ml; Bromophenol Blue 0.25 ml; Milli-Q water, 1 ml] before loading. The molecular weight markers used included lactate dehydrogenase (145 kDa) and β-amylase (200 kDa) in addition to those used for SDS-PAGE. Staining and destaining procedures were identical to SDS-PAGE.

Relative mobilities ($R_f$) for each of the standards and REXH were measured. A graph of [100 log (100$R_f$)] versus acrylamide concentration. A calibration curve was then obtained by plotting the slope of each standard against its molecular weight. The native MW of REXH was then estimated using this calibration curve.

### 3.2.7.5 Assay of enzyme activity

Unless otherwise specified, the assay for REXH expandase activity (by measuring the sum total of DAOC and DAC produced) was carried out using a modification of the published protocol[26]:

Reaction mixtures contained DTT (1 mM), ascorbate (1 mM), α-KG (1 mM), ferrous sulfate (50 μM), ammonium sulfate (50 mM), penicillin N (0.28 mM) and enzyme (*ca.* 0.03 to 0.3 mU) in 50 mM Tris-HCl, pH 7.5 at RT, in a total volume of 100 μl. Mixtures were incubated at 28°C in an orbital shaker operating at 200 rpm. After 20 min, the reaction was quenched with 100 μl of cold methanol and the precipitated protein removed by microfuge centrifugation. 100 μl of the supernatant was introduced into a 8-mm well in an agar plate seeded with *E. coli* X580 and penicillinase (DIFCO Laboratories, Bacto Penase Concentrate, 10 000 000 IU/ml). The plates were incubated at 37°C for 12 hr, after which the diameter of the inhibition zone was measured. The amount of cephalosporins (DAOC and DAC) was estimated by reference to a calibration graph obtained with standards of cephalosporin C.

The bioassay plates were routinely prepared by Mr John Keeping according to the following protocol :

Slopes of *E. coli* X580 were maintained at 4°C on 1x brain heart infusion agar [Composition of medium (g/l): Oxoid CM225 brain heart infusion, 18.5; Oxoid L11 bacteriological agar, 10.0; pH 7.4]. These slopes are viable for several months. A single loopful of this culture was transferred to 5 ml of sterile nutrient broth (Oxoid CM1 nutrient broth, 13.0 g/l) and incubated overnight at 37°C. 2 ml of this bacterial suspension was added to 200 ml of sterile half-strength brain heart infusion agar maintained in a liquid state at *ca.* 45°C in a water bath and the bacteria evenly distributed by swirling. For (+) plates, 0.2 ml of penicillinase was also added per 200 ml agar. 18 ml of molten agar was pipetted into a standard 90-mm triple vent Petri dish and allowed to harden. Once solidified, the plates were stored at 4°C and used within 2 weeks. 8-mm wells were made in the agar, using a cork borer, immediately before use. The test solution was then added, and the plates incubated at 37°C for *ca.* 12 hr. The (-) plates, without penicillinase, was prepared similarly for the purpose of quality control of the penicillin N used in the assay.

### 3.2.7.6 Equipment routinely used

*Spectrophotometry*
(1) Philips PU 8625 UV-Visible Spectrophotometer
(2) Philips Pye Unicam PU 8800 UV-Visible Spectrophotometer

*Gel scanning*
2202 Ultroscan Laser Densitometer (LKB Bromma, Sweden)

*Centrifugation*
(1) MSE Micro-Centaur Centrifuge
(2) Beckman Model J2-21 Centrifuge
(3) Beckman L8-70M Ultracentrifuge

*Ultrafiltration*
Stirred Ultrafiltration Cells, Amicon Models 8010, 8050, 8200 & 8400

*Sonication*
(1) MSE Soniprep 150
(2) Ultrasonic Processor W-380 (Heat Systems-Ultrasonics Inc., USA)

*Microconcentration*
(1) Centricon-100 microconcentrator (Amicon, USA)
(2) Centricon-300 microconcentrator (ditto)

*Plate Incubation*
Gallenkamp Duostat Incubator, Size 2

*Enzyme incubation*
G24 Environmental Incubator Shaker (New Brunswick , USA)

*Fermentation*
(1) G25 Controlled Environment Incubator Shaker (New Brunswick)
(2) LH-Fermentation Series 2000, 25-litre fermenter

*Weighing*
(1) Sartorius analytic model A 200 S
(2) Sartorius research model R 200 D
(3) Sartorius basic model S 3100 V

## 3.3 Results and Discussion

Routine fermentation runs, carried out by Messrs John Keeping and Richard Heath, presented no major problem in consistent production of the enzyme under study. This section deals only with purification studies performed in this project.

### 3.3.1 Purification of native REXH

Native REXH was consistently produced by *C. acremonium* CO 728 at a level of < 0.1% of total cell protein. The data for a typical purification (using 60 g wet weight of fungal mycelia) is summarised in Table 3.1, which shall be compared with the two earlier protocols reported by Yeh[27] and Baldwin[26].

In agreement with the Yeh protocol, sonication was found to be an efficient means of cell lysis to release intracellular proteins, and less cumbersome than the Dyno-Mill grinding method used in the Baldwin protocol. Working on the scale of 60-g mycelial extraction, the profile of protein release versus the extent of sonication is shown in Figure 3.1. Consequently, sonication was not necessary beyond the 6th 30-sec burst, since protein release was not significantly increased beyond that point and enzymes are well known to be susceptible to deactivation caused by excessive ultrasonic treatment[76]. It is important to note that an appropriate sonication procedure depends on the scale of extraction and hence the empirical determination of a suitable method is always necessary for every new situation.

| Step(a) | Total protein (mg) | Total activity(b) | | Specific activity (mU/mg) | Purification (fold) |
|---|---|---|---|---|---|
| | | mU | Recovery | | |
| Ammonium sulfate fractionate | 165 | 465 | - | 2.8 | - |
| Superdex 200 pool | 50.7 | 545 | 100 | 10.7 | 3.8 |
| Mono Q pool | 5.1 | 84.7 | 15.5 | 16.5 | 5.9 |
| Superdex 75 pool | 0.77 | 24.1 | 4.4 | 31.3 | 11.2 |

(a) Not all the enzyme eluted from a step is pooled for the next step.
(b) One milliunit (mU) of activity is defined as the amount of enzyme that produces one nanomole of product per minute under the assay conditions.

**Table 3.1** Purification of native REXH

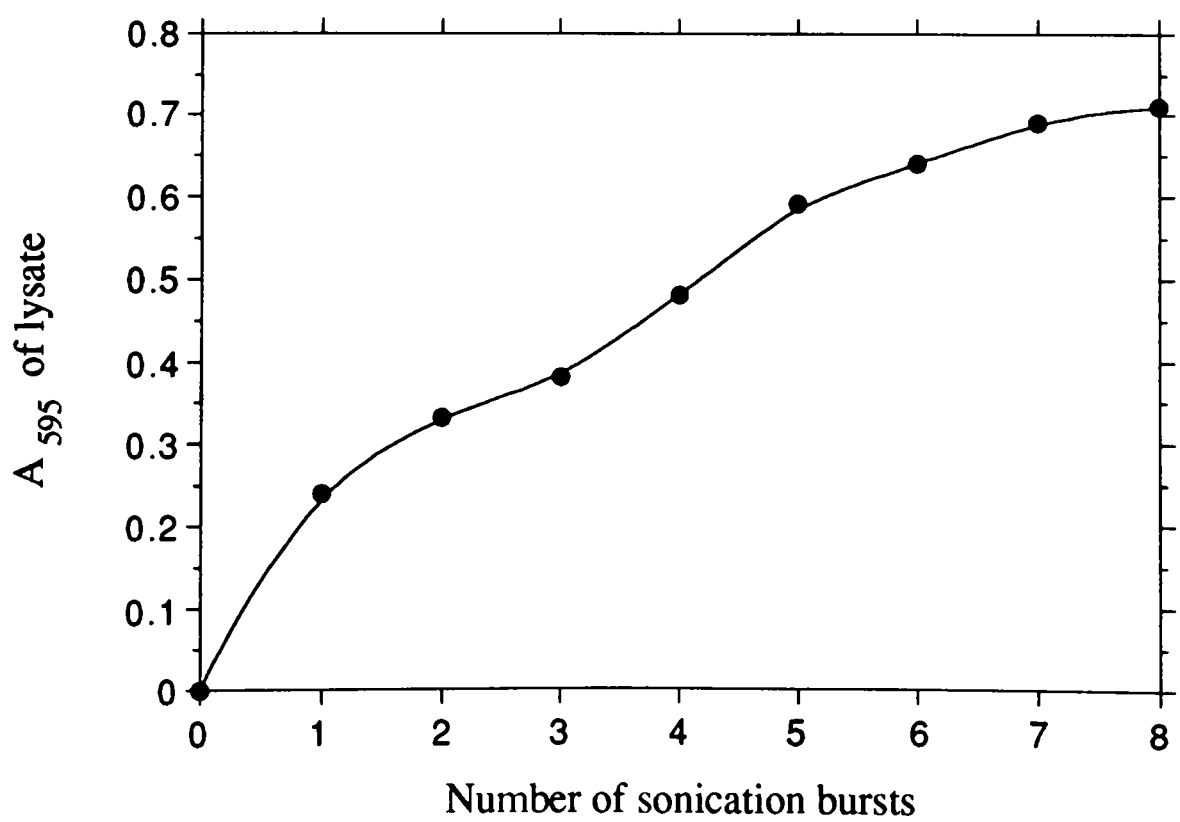



**Figure 3.1** Protein release profile upon sonication of *C. acremonium* mycelia

Both the protamine sulfate precipitation and ammonium sulfate fractionation were part of the original Baldwin protocol[26] and had been shown to be useful in respectively removing nucleic acids and some contaminating proteins. The first FPLC chromatographic step in the current work, utilising the new Superdex 200 gel-filtration resin, proved to be an improvement over the first step in the Yeh protocol, an ion-exchange process on DEAE Trisacryl that gave 1.6-fold purification.[27] Its advantages over the Sephadex G-75 gel filtration step in the Baldwin protocol (which gave similar purification yield) lies in (a) the use of a Bio-Pilot column (3.5 x 60 cm) that provided increased capacity and (b) the new medium which allowed increased flow rate without any compromise in resolution. The protein elution profile from the column is shown in Figure 3.2:

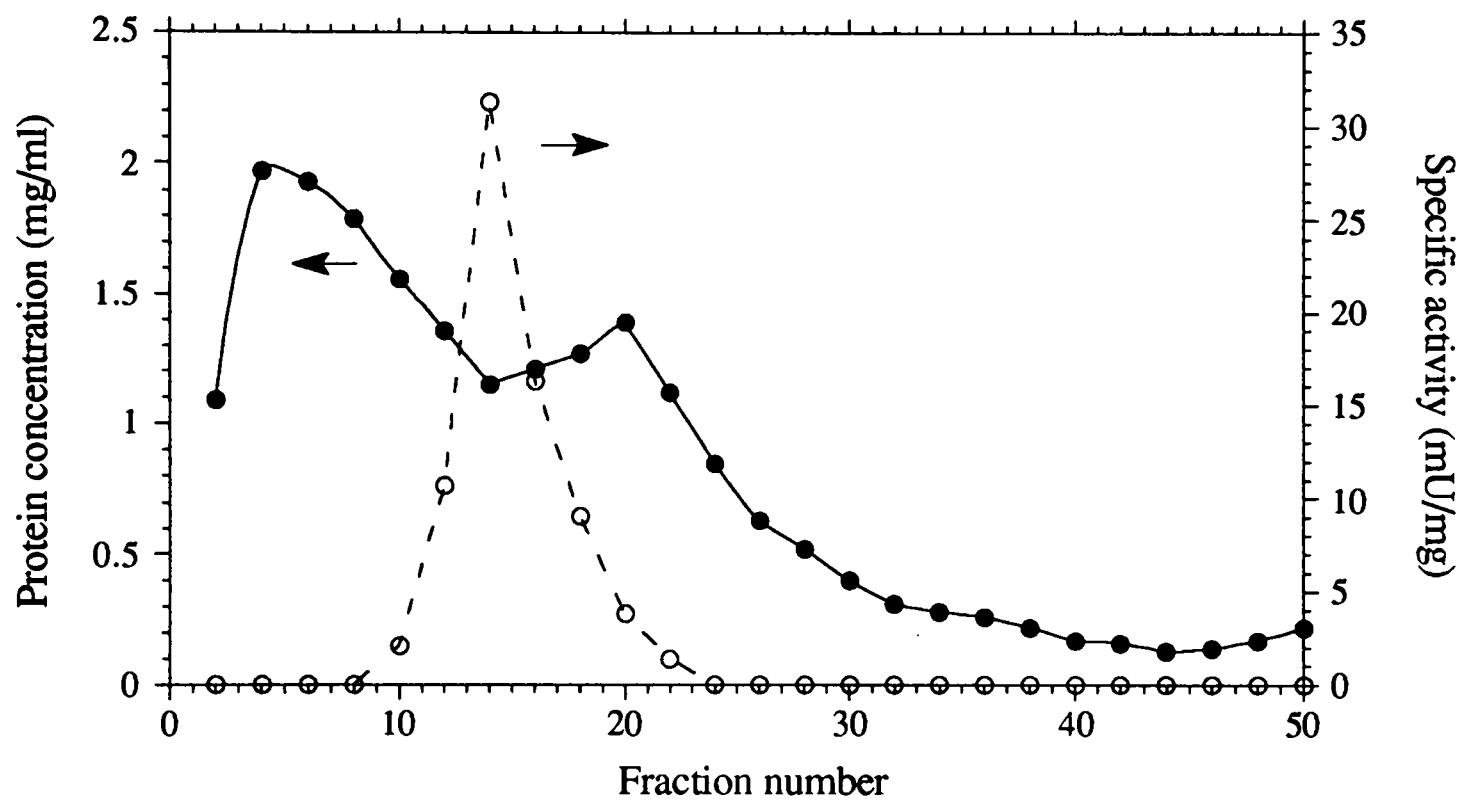



**Figure 3.2** Protein elution profile from a Superdex 200 run on native REXH

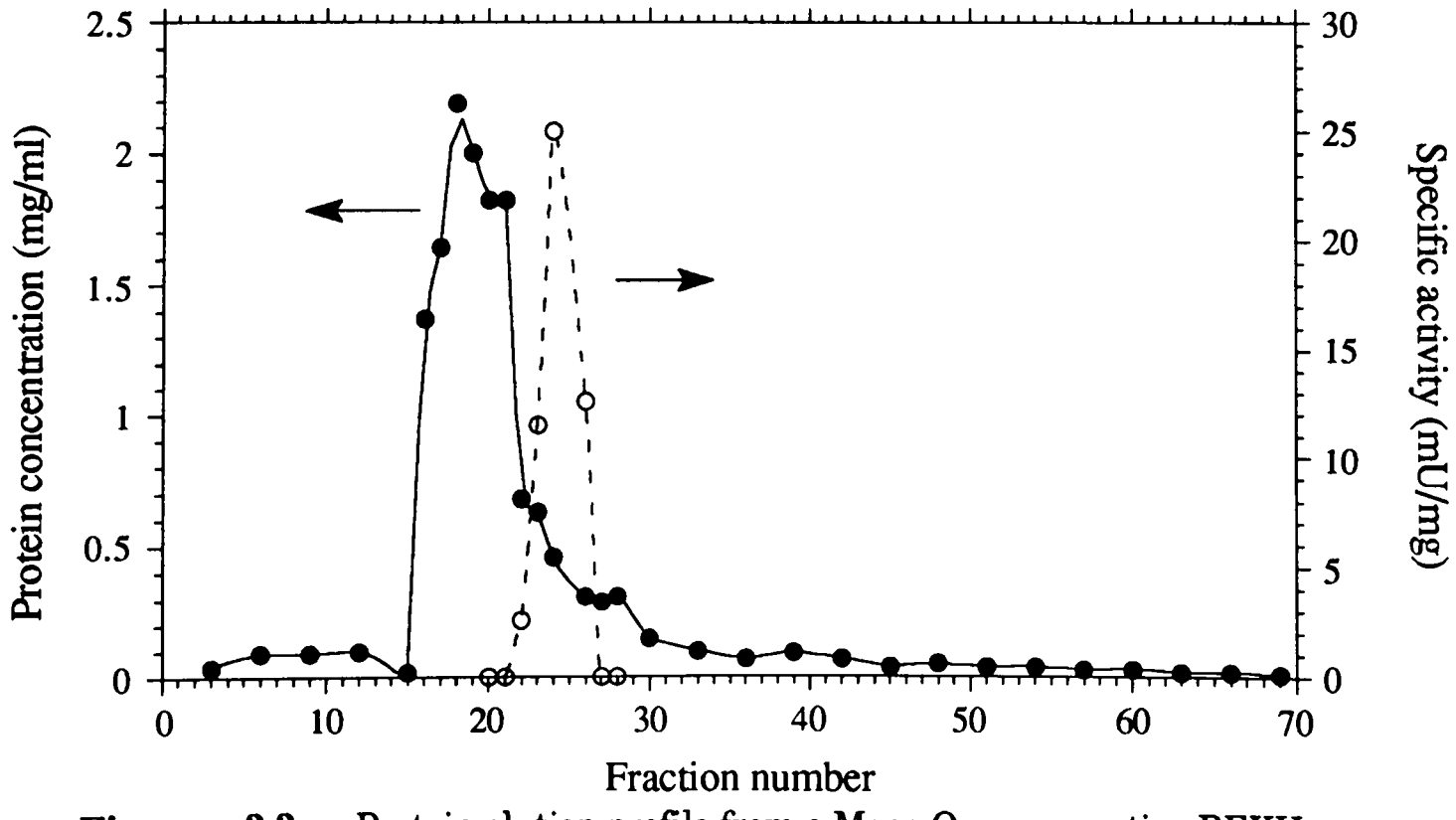



**Figure 3.3** Protein elution profile from a Mono Q run on native REXH

The second FPLC step, using the strong anion-exchange resin Mono Q, brought the purification level to 6-fold over the ammonium sulfate cut. The protein elution profile is shown in Figure 3.3. The final gel filtration step, employing the new Superdex 75 resin, removed major contaminating proteins present in the Mono Q eluate, as shown by the SDS-PAGE gel (Fig 3.4):

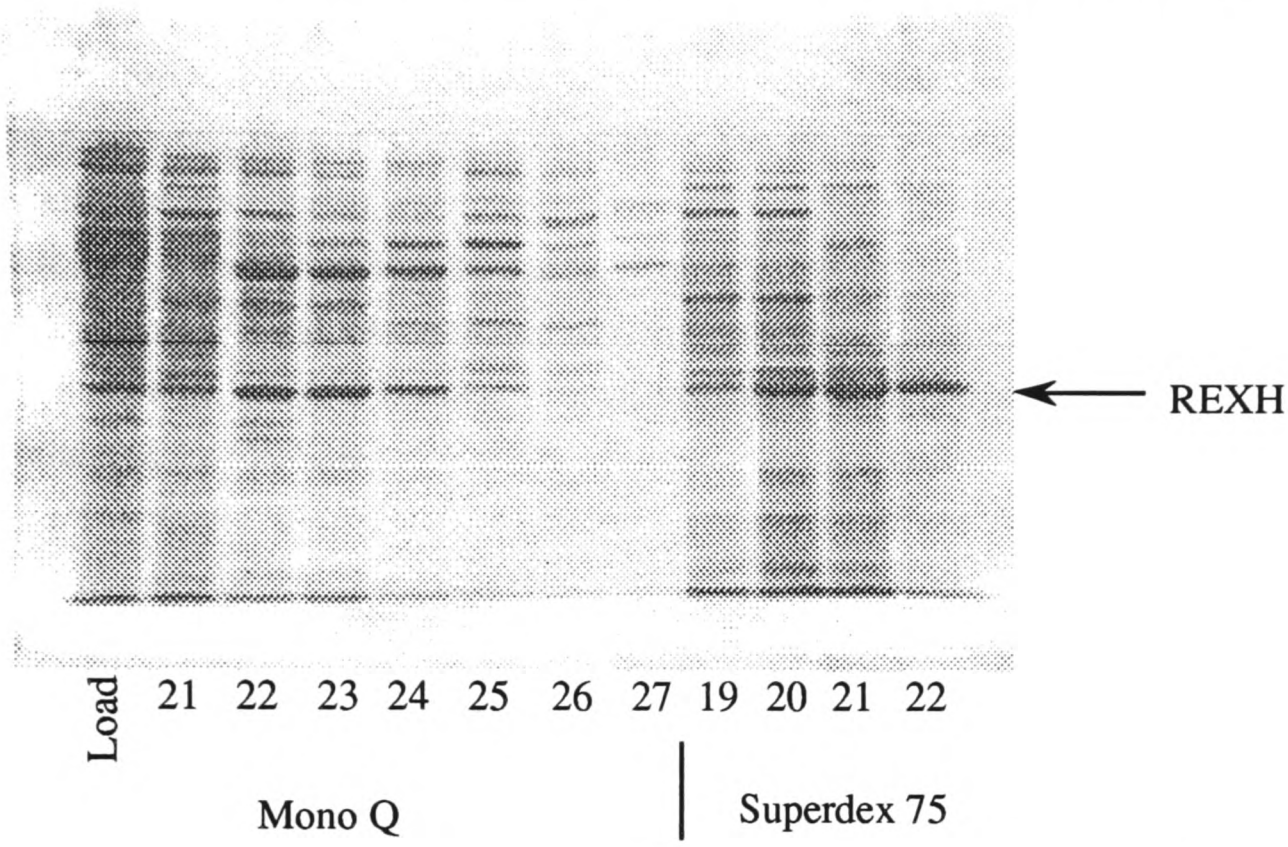



**Figure 3.4** SDS-PAGE gel of some fractions from the Mono Q and Superdex 75 eluates

The main advantages of the above protocol over the two published ones can be summarised as follows. Firstly, enzyme of comparable purity can be prepared in a shorter time, due to the use of high-capacity Superdex gel-filtration resins that allow faster flow rates. The final specific activities reported in the two previous protocols (*ca.* 35 - 40 mU/mg), obtained through four chromatographic steps were comparable to what was obtained here in three steps. The second advantage of this new protocol lies in the minimal treatment in between successive chromatographic steps, which is beneficial for maintaining the activity of the enzyme. Except for the concentration step required between the Mono Q and Superdex 75 runs, the pool from a preceding step is directly usable for the next step. For example, the high salt content in the ammonium sulfate resuspension is conveniently removed by Superdex 200 gel filtration, while the large volume of pooled eluate from Superdex 200 column is not a problem for loading onto the Mono Q column.

### 3.3.2 Purification and refolding of insoluble recombinant REXH

The work described in this Section was carried out in Lilly Research Laboratories (a Division of Eli Lilly and Company, Indianapolis, USA) in collaboration with Dr Wu-Kuang Yeh and Ms Sandhya Ghag of the Biochemical Development Department.

Since the knowledge of *E. coli* genetics was well advanced, it has been the focus for the

development of recombinant DNA techniques and historically, for the same reasons, *E. coli* was the most widely used host organism in which to express eukaryotic proteins.[77] While this imparts a clear advantage in that proteins can be produced in much higher levels compared to natural sources, it suffers from the drawback of often presenting the proteins in an inactive form due to its failure to adopt the native conformation. The inactive proteins often exist in the cytoplasm of *E. coli* as optically refractile aggregates called inclusion bodies (IB).[78] Indeed, the recovery of active enzymes from inclusion bodies has remained largely an empirical science among protein researchers ever since recombinant DNA technology delivered the first eukaryotic protein in *E. coli*. There is still no sure way to predict the solubility of a recombinant protein expressed in *E. coli*,[78] and no standard rules for recovering active enzymes from IBs.[79] The conventional strategy today still follows the pioneering work on *in vitro* refolding of denatured proteins initiated by Anfinsen,[80] whose seminal work led to the concept of the amino acid sequence as the single determinant of tertiary protein structure. Three essential steps are normally employed. In the first step, the aggregates are solubilised by a choice of several denaturing agents, such as pH changes, detergents and chaotropic agents, and partially purified. In the second step, the solubilised protein is brought into an environment which favours the native active conformation. Finally, since it is unrealistic to expect the folding to proceed in 100% yield, it is necessary to devise a way of removing the wrongly folded protein. Thus, the three steps consist of unfolding, refolding and isolation.

REXH, like many other eukaryotic proteins synthesized in *E. coli*, was initially produced in the form of inactive enzyme aggregated in cytoplasmic granules, or inclusion bodies (IB).[51] In our Oxford laboratory, refolding was carried out using Tris buffer of identical concentration and pH as that used for native REXH preparation, *i.e.* 50 mM and pH of 7.5. This was found to cause significant aggregation of REXH molecules, as determined by gel filtration on Superdex 200 resin and subsequent SDS-PAGE. As shown in the FPLC profile (Figure 3.5) and SDS gels (Figure 3.6), about 80% of REXH eluted in the void volume, indicating that the aggregated (but still soluble) protein molecules are of a MW higher than 600 kDa (the MW cut-off for Superdex 200 resin). Only about 20% eluted in the expected MW region and demonstrated catalytic activity. This result was interpreted as an inefficient refolding of REXH from the inclusion bodies. All attempts to salvage the activity of the wrongly folded protein molecules were unsuccessful. Further experiments showed that the enzyme molecules in the aggregates are held together by physical interactions only, which are dissociable by SDS. This is probably due to hydrophobic regions of partially folded molecules collapsing together before full refolding is achieved. Indeed, the concept of hydrophobic collapse is a notion used frequently in the study of 'the protein folding problem', *i.e.* in studying how the one-dimensional information in amino acid sequence codes for the dynamic three-dimensional behaviour of active proteins.[81] This assumption was used to guide further optimisation of refolding efficiency, which was carried out both at Indianapolis and Oxford.

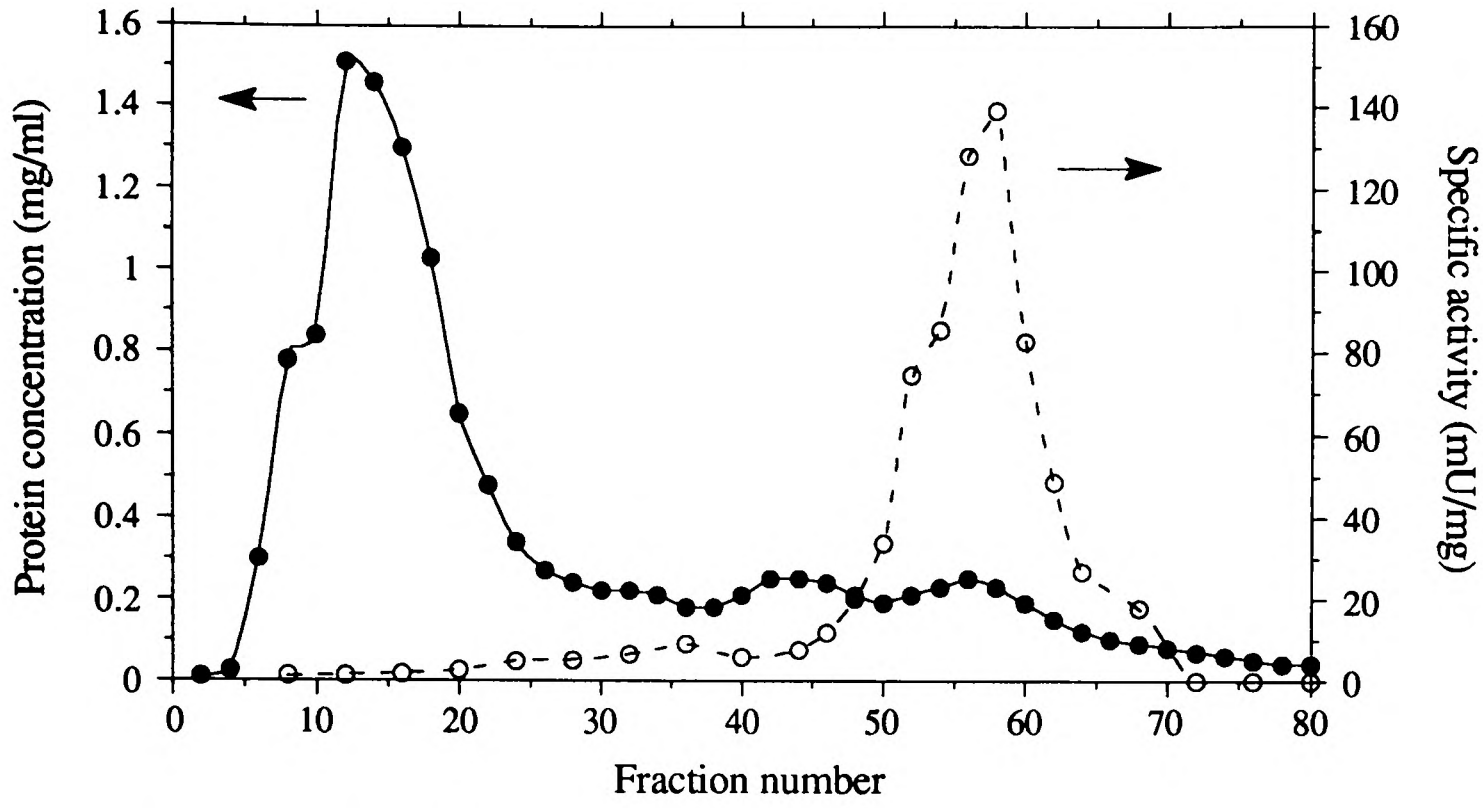



**Figure 3.5** Protein elution profile from a Superdex 200 run on insoluble recombinant REXH

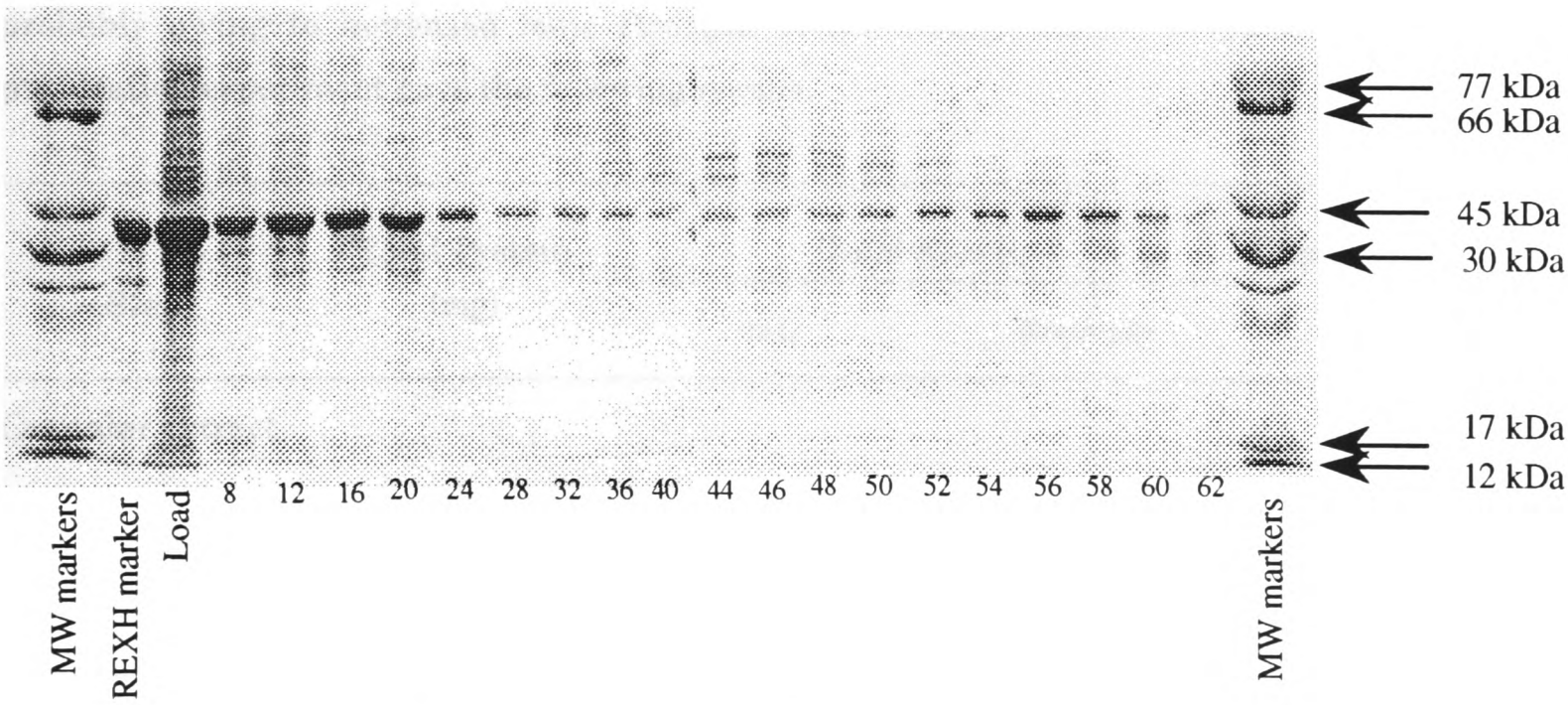



**Figure 3.6** SDS-PAGE gel for Superdex 200 fractions on insoluble recombinant REXH

For a start, the concentration of Tris in the preliminary 'folding' buffer and all subsequent chromatographic buffers was decreased 10-fold to 5 mM to reduce the ionic strength of the buffer. The concept of hydrophobic collapse is based on the idea that when an unfolded polypeptide with exposed hydrophobic side chains is placed in an aqueous solution, it will collapse to a state that shields hydrophobic groups from solvent. It is thought that lowering the ionic strength, and hence the polarity, of the medium will help to disfavour 'hydrophobic collapse' of the unfolded protein. Secondly, the buffer pH was increased from 7.5 to 9.5, designed to enhance the surface negative charges on REXH molecules. This should increase electrostatic

repulsions between REXH molecules and lead to greater stabilisation. Thirdly, the refolding was done in much lower protein concentration (< 0.2 mg/ml) so that intramolecular interactions occur in preference to intermolecular interactions. Previous studies on other systems have shown that protein concentration is an important parameter in optimising the yield of correctly folded protein.[82] Removal of urea by dialysis was done in three steps instead of one to provide a gradual change in environment for the refolding protein molecules. It was found that complete removal of urea led to severe activity loss, and 0.5 M was about the optimal for enzymic activity. Such a dependence on a low urea concentration for activity could be due to some kind of structural stabilisation provided by the molecule, and has also been observed for other recombinant proteins.[79]

Utilising the alterations prescribed above, the aggregation observed in earlier experiments were virtually eliminated in the pH-jumped and diluted solution, as verified by gel filtration on Superose 12 resin. However, it was still not possible to tell if all the monomeric enzyme obtained was correctly folded. The separation of REXH molecules with native conformation from those with inactive conformation can probably be achieved by an affinity chromatography[79] (*e.g.* using an α-KG-based resin), but this was not attempted. Subsequent FPLC on the Mono-Q column led to 40% aggregation, as judged by Superose 12 gel filtration, probably owing to increased ionic strength in the eluting buffer. A typical refolding and purification experiment gave the following results:

| Step | Total Protein (mg) | Total Activity | | Specific activity (mU/mg) |
|---|---|---|---|---|
| | | mU | Recovery | |
| Granular extract | | | | |
| pH-jumped | 100 | 8925 | – | 89.5 |
| Diluted | 100 | 18500 | – | 185 |
| Filtered | 100 | 16350 | 100 | 163.5 |
| Mono Q eluate | 47 | 8824 | 54 | 187 |

**Table 3.2** Refolding and purification of insoluble recombinant REXH

### 3.3.3 Purification of soluble recombinant REXH

The urea-solubilised recombinant REXH, prepared as in the last section, provided the main source of enzyme for structural and mechanistic studies at the beginning of this project. Hence, effort was initially targeted at optimising the recovery of enzyme from inclusion bodies.

This direction of work was abandoned immediately following the successful design of a soluble expression system for the recombinant enzyme by a collaborating group[52]. Although the refolding of REXH could have remained an interesting academic exercise, the pressing need then was to obtain large quantities of REXH in a short time.

A typical purification of soluble recombinant REXH, overexpressed in *E. coli* JM 109 / pRH 1091, on a 6-g cell pellet scale is shown in Table 3.3:

| Step(a) | Total protein (mg) | Total activity(b) | | Specific activity (mU/mg) | Purification (fold) |
|---|---|---|---|---|---|
| | | mU | Recovery | | |
| Crude sonicate | 383 | 8754 | 100 | 22.9 | - |
| Q-Sepharose Fast Flow pool | 33.3 | 6400 | 73 | 192 | 8.4 |
| Superdex 200 pool | 17.5 | 5230 | 60 | 299 | 13.1 |

(a) Not all the enzyme eluted from a step is pooled for the next step.
(b) One milliunit (mU) of activity is defined as the amount of enzyme that produces one nanomole of product per minute under the assay conditions.

**Table 3.3** Purification of soluble recombinant REXH

Thus, 5 sonication bursts of 30-sec duration were found to be sufficient for almost complete protein release at this scale, as shown in Figure 3.7.

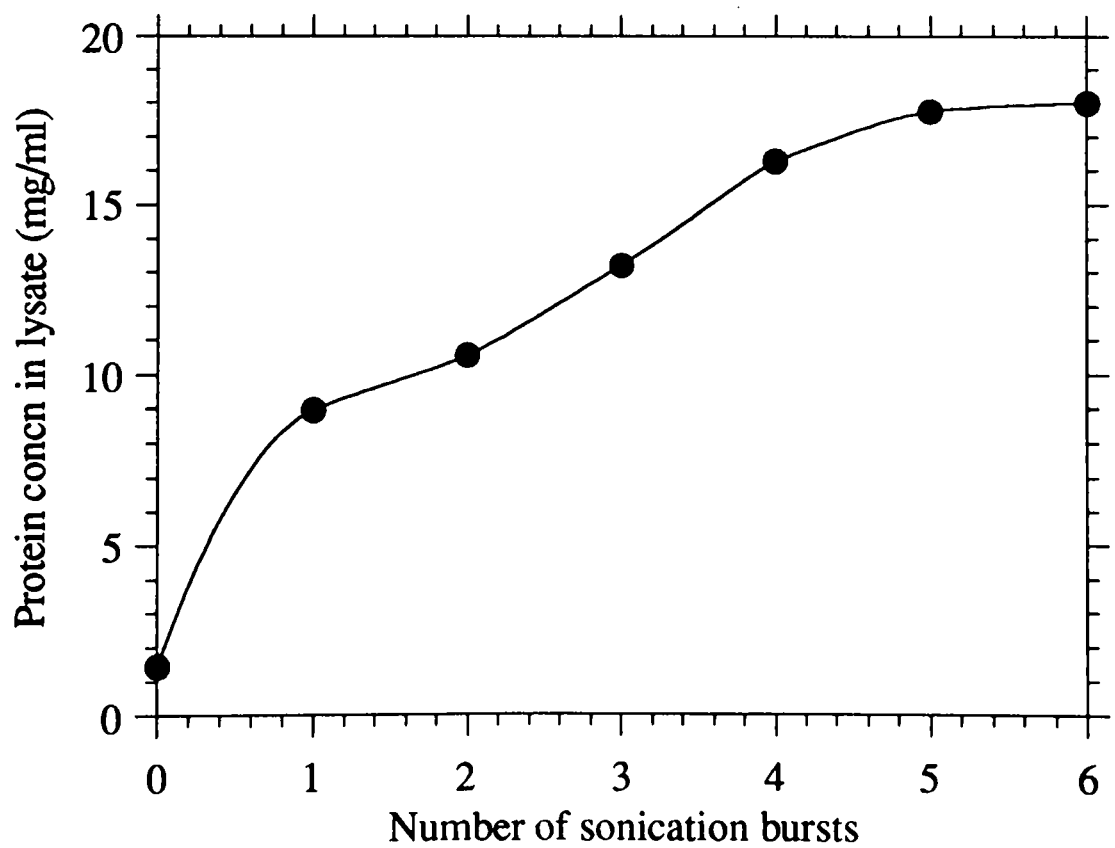



**Figure 3.7** Protein release profile upon sonication of *E. coli* JM 109 / pRH 1091

The proportion of REXH routinely obtained from the crude lysate was between 20 to 30 % of total cell protein[75]. Due to this high percentage, only two chromatographic steps are required to obtain the level of purity achieved in a 3-step protocol for native REXH. The first step, a strong anion exchange process on the Q-Sepharose Fast Flow resin, purified the enzyme 8-fold over the crude lysate. The protein elution profile is shown in Figure 3.8, and the SDS-PAGE results in Figure 3.9.

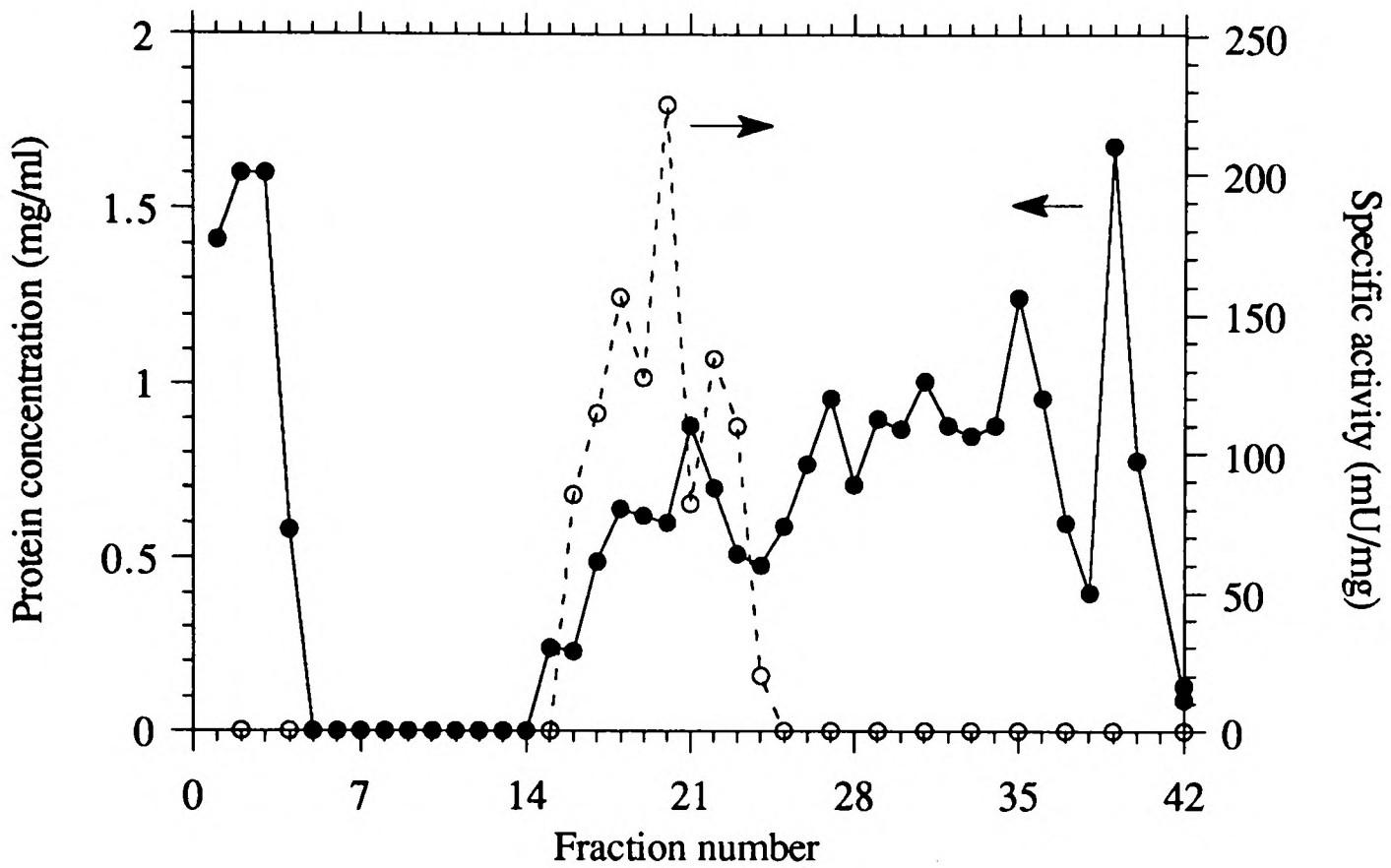



**Figure 3.8** Protein elution profile from a Q-Sepharose run on soluble recombinant REXH

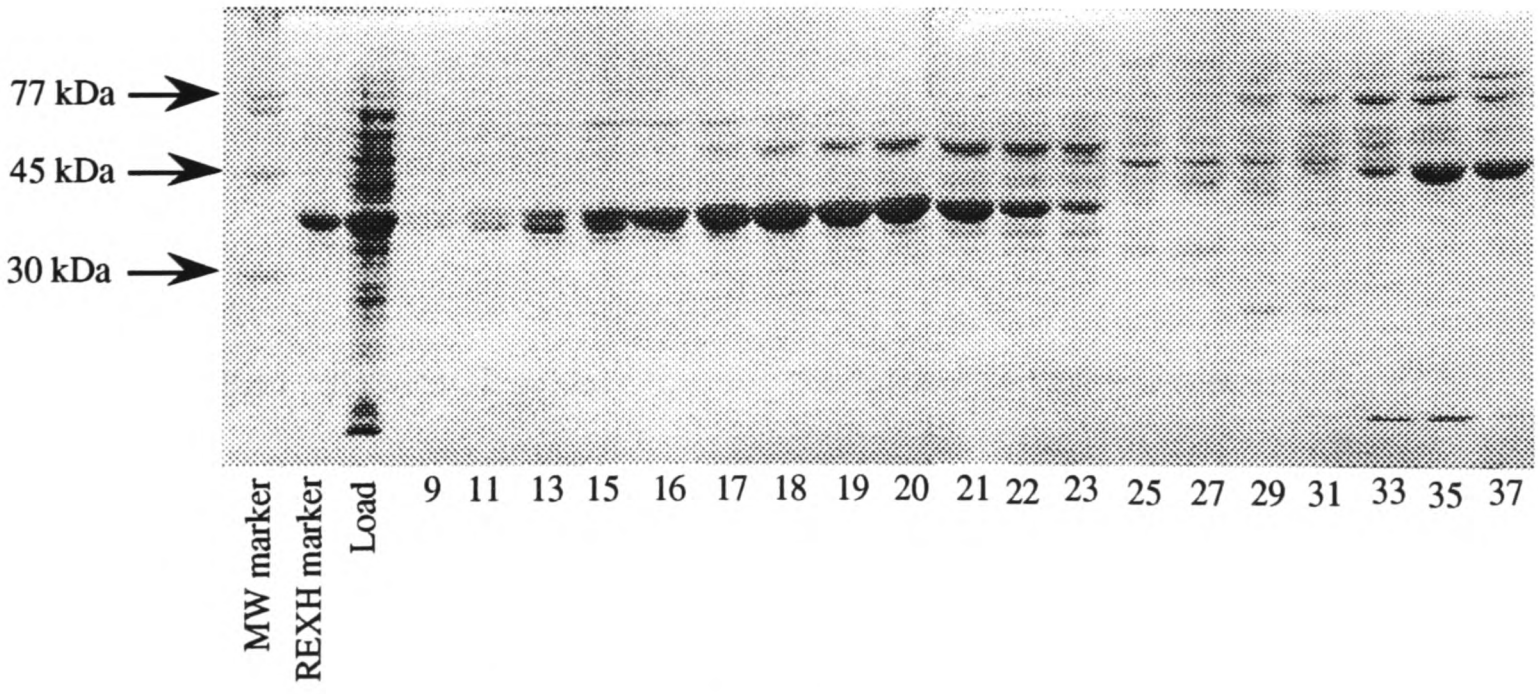



**Figure 3.9** SDS-PAGE gel for Q-Sepharose FF fractions on soluble recombinant REXH

The next step, gel filtration on the Superdex 200 column, removed a major contaminating protein with apparent MW between 45 and 66 kDa. The protein elution profile and the SDS-PAGE data are presented in Figures 3.10 and 3.11 respectively.

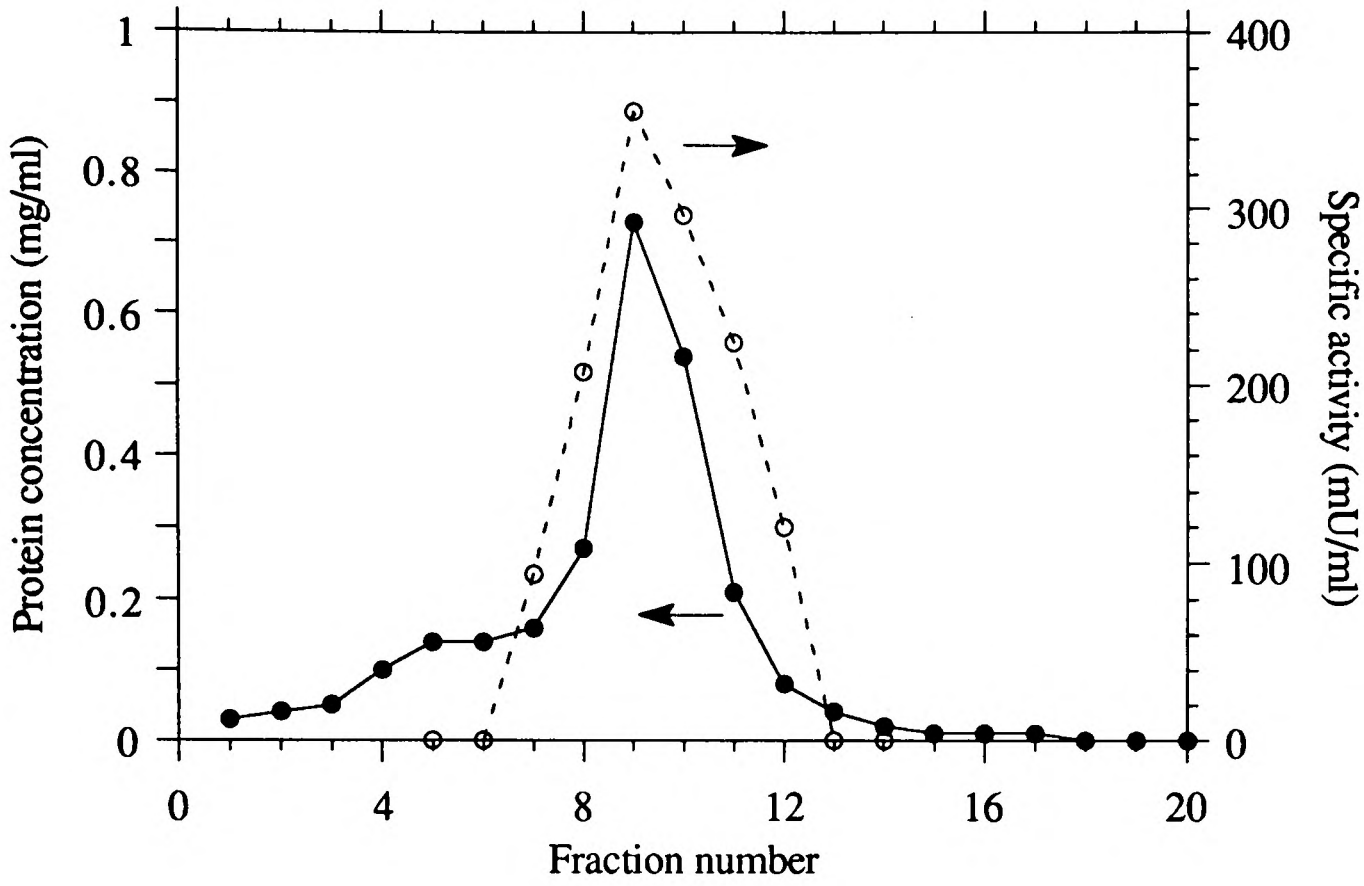



**Figure 3.10** Protein elution profile from a Superdex 200 run on soluble recombinant REXH

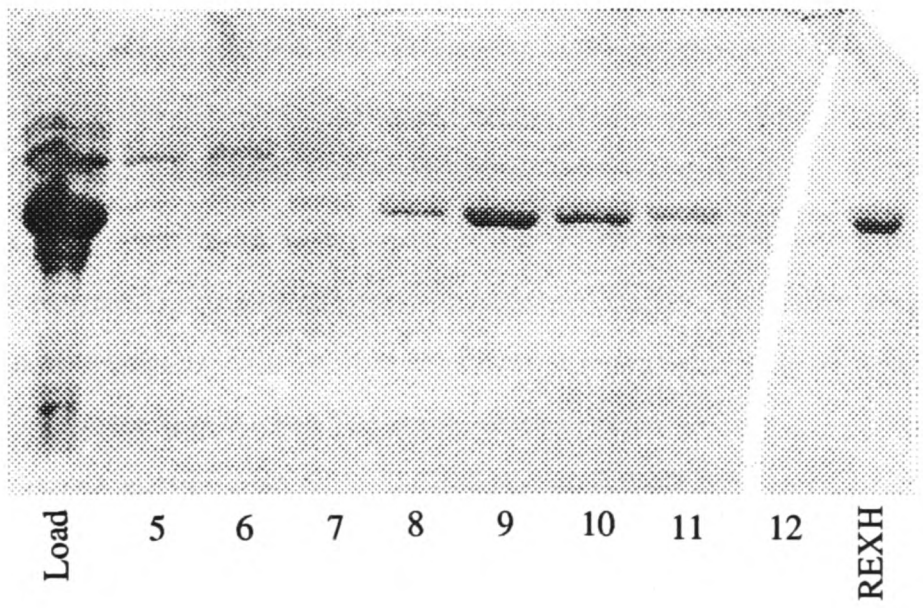



**Figure 3.11** SDS-PAGE gel for Superdex 200 fractions on soluble recombinant REXH

| Enzyme | Specific activity of fractions of comparable purity (mU/mg) |
|---|---|
| Native | 31.3 |
| Soluble recombinant | 356 |
| Insoluble recombinant | 139 |

**Table 3.4** Comparison of specific activities obtained with different forms of REXH

At this point, it is worth noting that the highest specific activities obtained with each form of REXH are significantly different, as shown in Table 3.4. The difference between the two forms of recombinant REXH is probably attributable to a discrepancy in the 'correctness' of the *in vivo* protein folding of the soluble form versus the *in vitro* folding of the insoluble form. It is possible that the environmental differences in the two folding conditions resulted in different folding intermediates, such that the active conformation is favoured as the final product in the *in vivo* process but is competing less effectively with inactive (incorrect) conformations in the *in vitro* protocol described under Section 3.2.4. If soluble recombinant REXH is assumed to be 100% correctly folded, then the efficiency of refolding for the insoluble recombinant REXH would be less than 40%.

The 11-fold difference in specific activity between native and soluble recombinant REXH is even more pronounced. In this case, one could perhaps postulate that, in *C. acremonium*, each enzyme molecule has a limited life-span, so that after a statistically determined number of catalytic turnovers, it becomes 'damaged' and must await degradation. Presumably, such dysfunctional molecules are not distinguishable from normal active ones by chromatographic techniques, resulting in their co-purification. Since native REXH has been performing its useful biological duty in the fungus, while recombinant REXH is derived in a virgin state, one would expect the latter to exhibit higher specific activity. If this conjecture is right, and assuming that soluble recombinant REXH contains 100% functional molecules, then 10 out of 11 REXH molecules extracted from the fungus have already served their due term and are 'tagged' for destruction. Indeed, since secondary metabolites are produced in a specific phase of the organism's life cycle, the requisite biosynthetic enzymes are expected to follow a pattern of expression and subsequent degradation.[83] Intracellular protein degradation has been recognised to play an important role in the modulation of the levels of specific proteins. One of the most studied eukaryotic pathways for such modulation is the ubiquitin pathway, where a 76-amino acid polypeptide called ubiquitin is ligated to a protein earmarked for 'retirement'.[84] This represents just one of the possible methods of biological tagging.

In summary, the purification studies carried out in this project have achieved the following goals:

1. A new protocol, employing new chromatographic media, has been developed for the purification of native REXH. The quality of enzyme obtained is comparable with published methods, while the time required is shorter.
2. The soluble recombinant REXH, expressed in *E. coli* JM 109 / pRH 1091, has been purified for the first time using a combination of strong anion exchange followed by gel filtration.
3. A refolding protocol was developed for solubilising recombinant REXH from inclusion bodies produced by *E. coli* NM 554 / pIT 511. The protocol removed the problem of molecular aggregation and enzyme wastage.

# CHAPTER 4

# SUBSTRATE ANALOGUE STUDIES WITH REXH

## 4.1 Methods in structural and mechanistic enzymology

The elucidation of structure-function relationship has always occupied a central position in protein research. In the case of enzymes, one is interested in defining the molecular events underlying catalysis - *i.e.* the mechanism - and how this is facilitated by the unique chemical environment - the structure - of the enzyme, especially within the active site. An increasing range of techniques has been brought to bear on these investigations. Structural questions have been addressed with physical methods such as X-ray crystallography, EXAFS spectroscopy, high-field NMR spectroscopy, mass spectrometry, CD spectroscopy and fluorescence spectroscopy. Chemical methods employed include chemical modification[85] and affinity labelling[86] of active site residues. The structural information thus obtained, together with kinetic experiments using natural substrate(s), substrate analogues and isotopically labelled substrates, provides the basis for unravelling the mechanism of the enzymic reaction. Recombinant DNA technology, especially through the use of site-directed mutagenesis, has already established itself as a powerful tool in mechanistic enzymology.[87] Increasingly, more and more enzymic processes have become amenable to direct probing by physical techniques, as exemplified by the recent observation of an acyl-enzyme intermediate in Class A β-lactamase catalysis using the technique of electrospray mass spectrometry.[88]

Kinetic evaluation with substrate analogues is a common method to map the active site topology of an enzyme and to elucidate its mode of catalysis. It is important to distinguish clearly between substrate analogues and isotopically labelled substrates. Whilst there is a tendency for some workers to treat the latter as a subset of the former, it is the intention of this thesis to use the following definition. Isotopically labelled substrates carry isotopic enrichment (*e.g.* $^{13}C$, $^{2}H$) at specific positions without any change in bond connectivity and stereochemistry. They can be very useful in mechanistic studies but rarely yield any direct structural information on the enzyme. On the other hand, substrate analogues incorporate modifications to the natural substrate including specific changes in bond connectivity and stereochemistry. They can contribute to the determination of both mechanism and active site topology. Despite the above distinction, it is possible for chemists to synthesize substrate analogues that contain specific isotopic enrichments, thus making compounds that come under both categories.

## 4.2 Mechanistic model for REXH

As mentioned in Section 2.3, the α-KG dependent dioxygenases are characterized by common features which include (1) reaction stoichiometry, (2) cofactor requirements, (3) susceptibility to certain inhibitors and (4) sequence homologies in two histidine-containing

regions. These important similarities point to a closely related mechanism for this family of enzymes. The most extensively studied member among them is the chick embryo prolyl 4-hydroxylase,[58, 66] which has become a prototype for mechanistic studies into this class of oxygenase. Over the last 30 years, several attempts have been made to provide a mechanistic model for the hydroxylation of peptidyl proline by this enzyme. The earlier models, which have failed to withstand the test of time and shall not be treated here, include those proposed by Fujita,[89] Lindstedt,[90] Hobza[91] and Hamilton.[92] Siegel[66] was the first to propose an iron-oxo intermediate as the active oxygen species, an idea based on an analogy with the proposed mechanism for haem-containing iron-dependent P-450 cytochromes.[93] Günzler[58] then provided a detailed stereochemical description for the generation of this reactive species in prolyl 4-hydroxylase action. Together, their combined strengths shall be used as the model for a proposed mechanism for REXH.

## 4.2.1 Activation of molecular oxygen

The Günzler mechanism[58] for prolyl 4-hydroxylase involves two steps for the overall catalytic process: the first step producing the active oxygen species and the second step utilizes this entity to hydroxylate the peptidyl proline. While most α-KG dependent dioxygenases mediate hydroxylation at unactivated carbon positions, notable exceptions exist where the enzymes carry out ring cyclisation, desaturation, epoxidation, alcohol oxidation and aldehyde oxidation.[94] It is envisaged that in all these different chemical outcomes, the strategy for oxygen activation is identical though the exact usage of the active oxygen species is somewhat different.

The working hypothesis for the activation of molecular oxygen in REXH is depicted in Figure 4.1. The catalytic ferrous ion, in an octahedral coordination geometry, is anchored within the active site by binding facially to three ligands from the protein architecture. The ligands are most likely to be two histidyl nitrogens and one tyrosyl oxygen, according to preliminary EXAFS experiments.[95] The remaining three coordination sites are occupied by oxygen atoms from C-1 and C-2 of α-KG and the dioxygen molecule in a bent conformation.

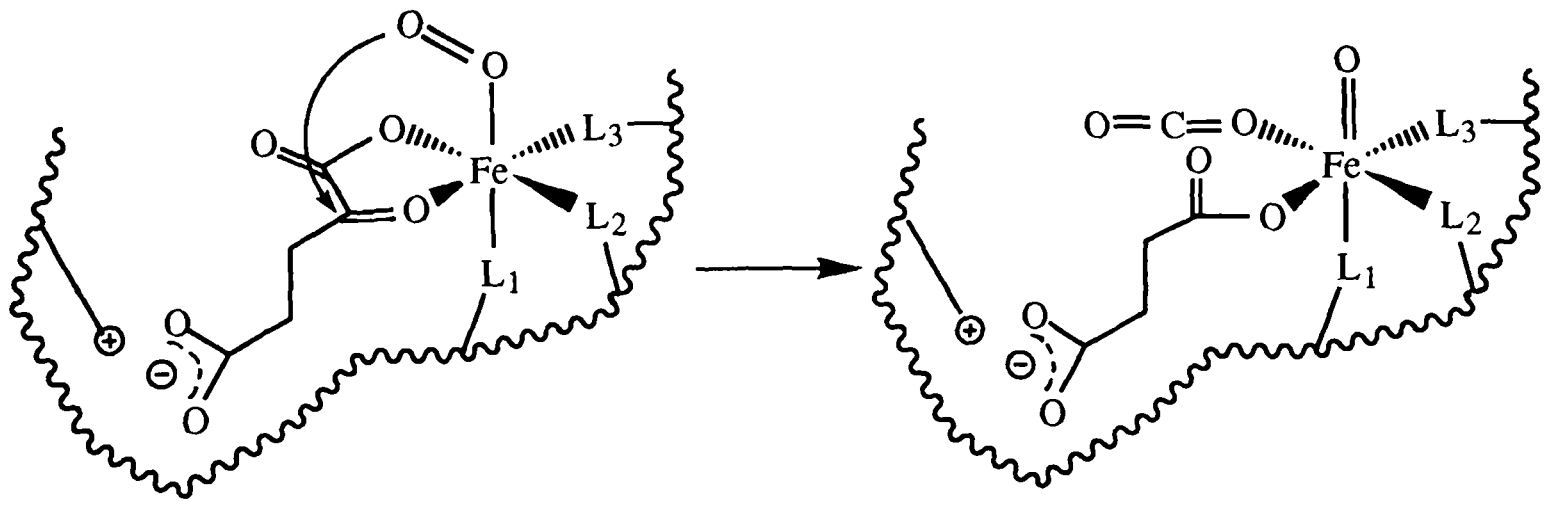



**Figure 4.1** Generation of the iron-oxo species

The substrates have been shown to bind in the following order: α-KG, molecular oxygen and then the main substrate. The uncoordinated atom of dioxygen then attacks the carbonyl carbon of α-KG, leading to a new C-O bond and the cleavages of the C1-C2 bond in α-KG and the O=O bond in dioxygen. The electronic rearrangement gives rise to carbon dioxide, succinate and a highly reactive iron-oxo moiety. The latter is the oxidising species which acts on the main substrate in the next stage of the enzymatic catalysis.

### 4.2.2 Ring expansion

In the ring expansion reaction (Fig. 4.2), it is proposed[97] that the iron-oxo species abstracts a hydrogen atom from the 3β-methyl group of penicillin N (24), giving a radical (42) which equilibrates freely with (43) and the tertiary radical (44), the latter being the most thermodynamically stable structure. Further abstraction of the hydrogen atom at C-4 of (44) results in the formation of DAOC (25) and water. A competing reaction of (44), whereby it undergoes recombination with the iron-hydroxo radical, gives rise to the 3β-hydroxy cepham (45) as a minor product. This shunt metabolite was first isolated from the culture broths of *C. acremonium* by Miller *et al.*[98] who found it not to be an intermediate in the ring expansion of penicillin N (24) to DAOC (25). This branched pathway is sensitive to a kinetic isotope effect, so that the ratio of radical recombination relative to hydrogen abstraction at C-4 can be increased by using 3-deuteriopenicillin N (46) as a substrate.[99, 100]

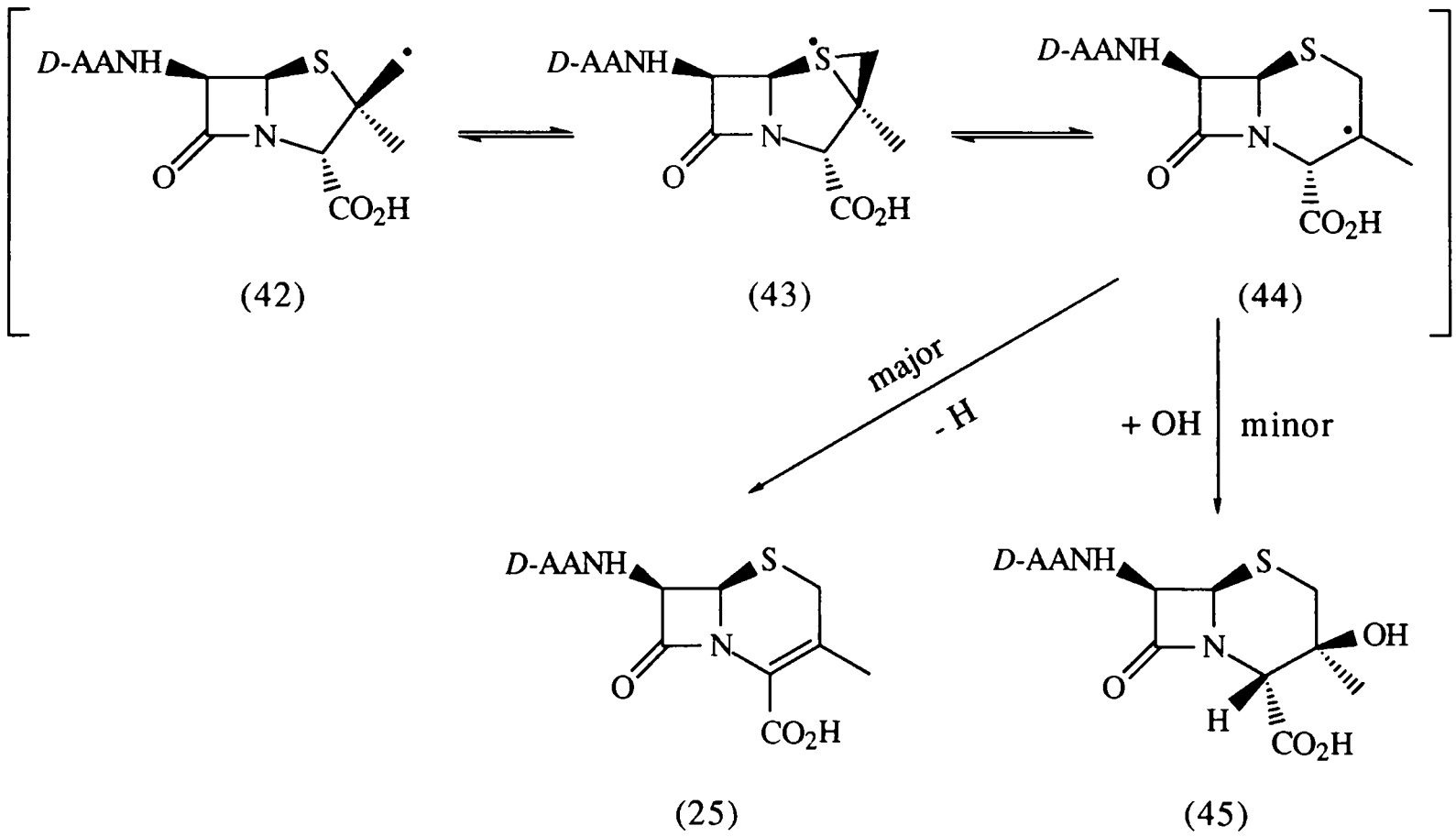



**Figure 4.2** Mechanism of ring expansion of penicillin N

The order of hydrogen abstraction outlined above was established by experiments using labelled substrates.[101] Incubation of a 1:1 mixture of penicillin N (24) and 2,2-

bis(trideuteriomethyl)penicillin N (47) with REXH resulted in preferential conversion of the protiated substrate, whereas in similar experiments with a 1:1 mixture of penicillin N (24) and 3-deuteriopenicillin N (46) both substrates were converted at the same rate (Fig. 4.3). These results accord with a stepwise process for hydrogen removal in which the methyl hydrogen is removed first. *In vitro* studies[102] using 2α-$^{13}CH_3$-penicillin N (48) have shown that the 2α-methyl group is incorporated into the methyl group of DAOC (25) (Fig. 4.4), confirming that the hydrogen abstraction occurs at the 2β-position, leaving the 2α-methyl group intact.

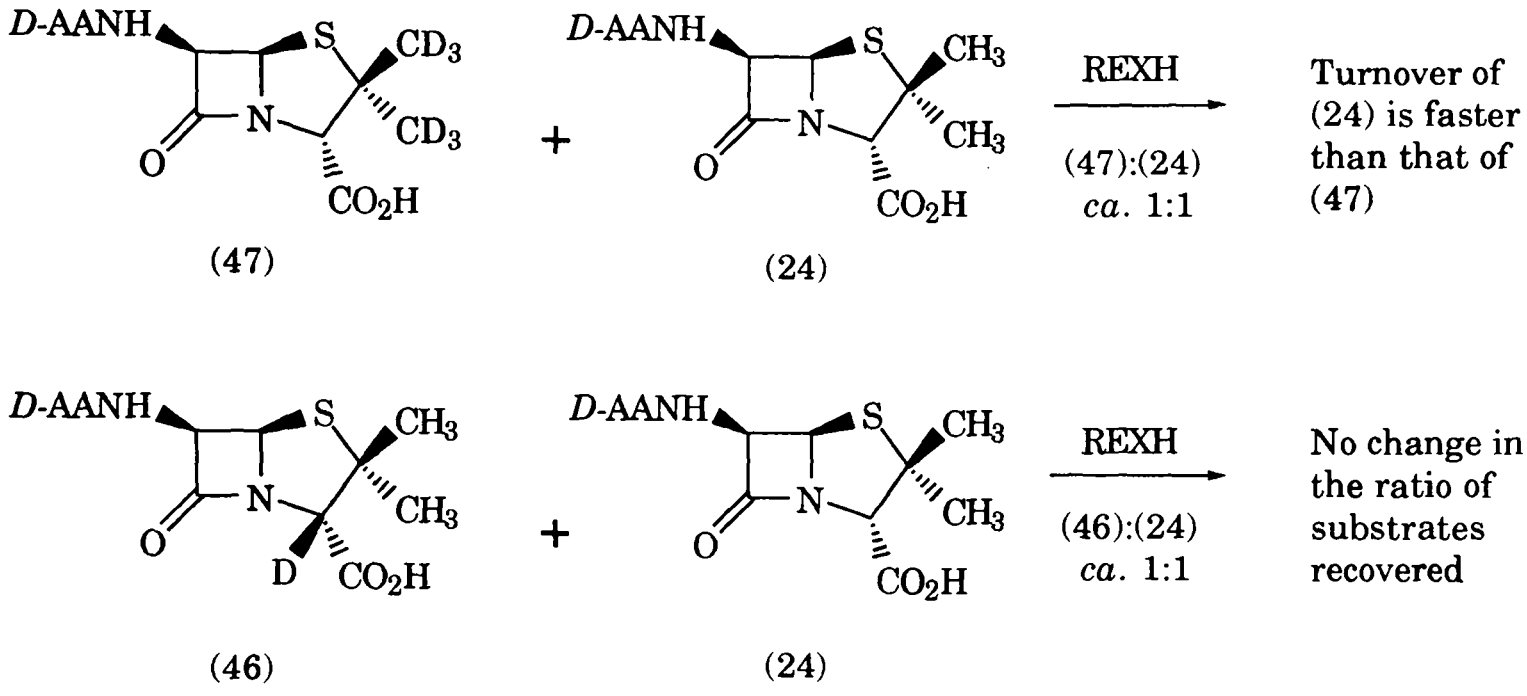



**Figure 4.3** Experiments to determine the order of hydrogen removal in ring expansion

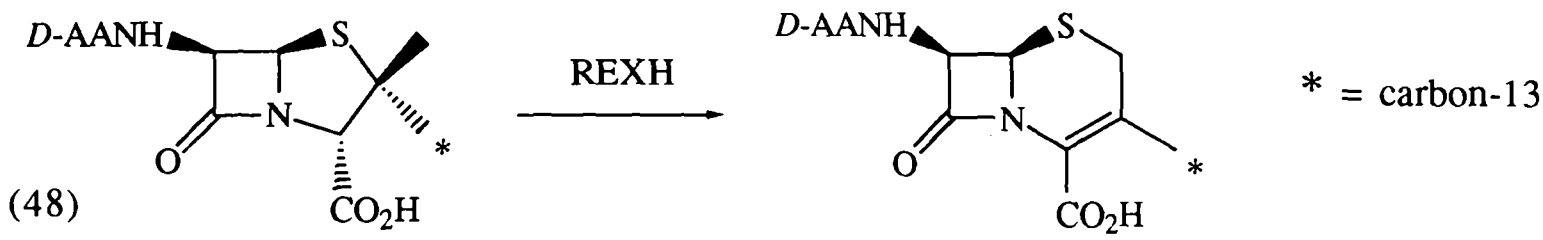



**Figure 4.4** Regiochemical fate of methyl groups in penicillin N

Townsend *et al.*[103] and Abraham *et al.*[104] have investigated the stereochemical course of the ring expansion process by feeding 'chiral-methyl' valines to *C. acremonium* broths. The independent conclusion of both groups was that the incorporation of the (3-*pro*-R) methyl group of valine into the endocyclic methylene group of cephalosporin C (27) occurs with complete loss of stereochemistry (Fig. 4.5). This is consistent with the presence of the radicaloid intermediate (42) during the ring expansion process, which can undergo rotation at the 2β-carbon faster than carbon-sulfur bond formation. The chemical feasibility of a ring expansion mechanism *via* a free radical process has been verified by a biomimetic study.[105]

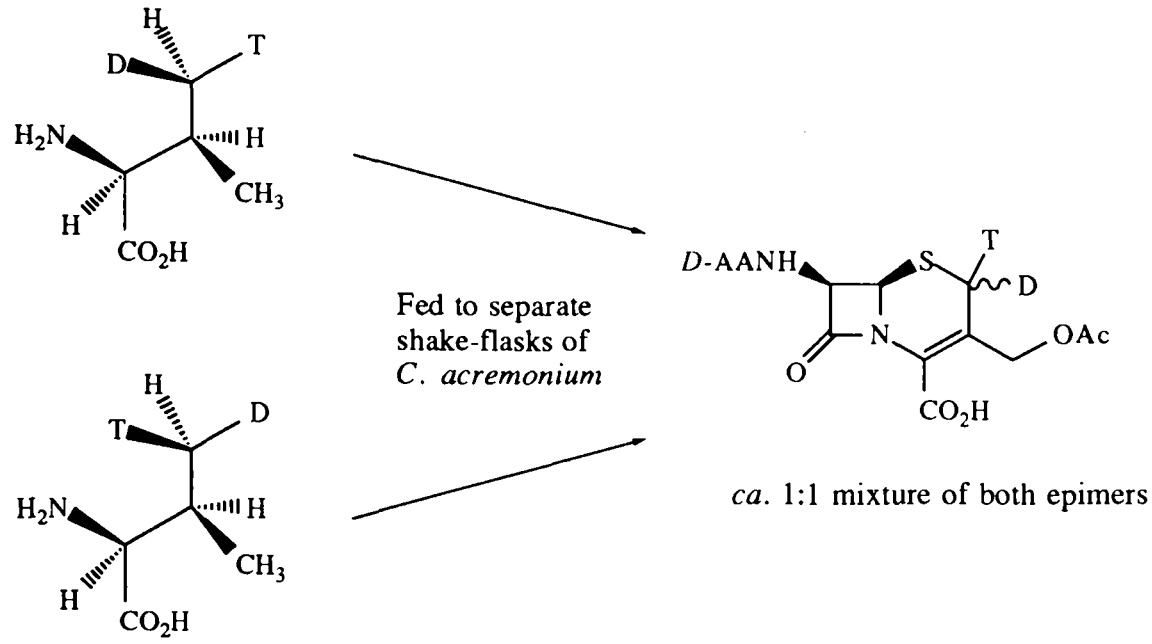



**Figure 4.5** Incorporation of 'chiral-methyl' valines into cephalosporin C

## 4.2.3 3'-Hydroxylation

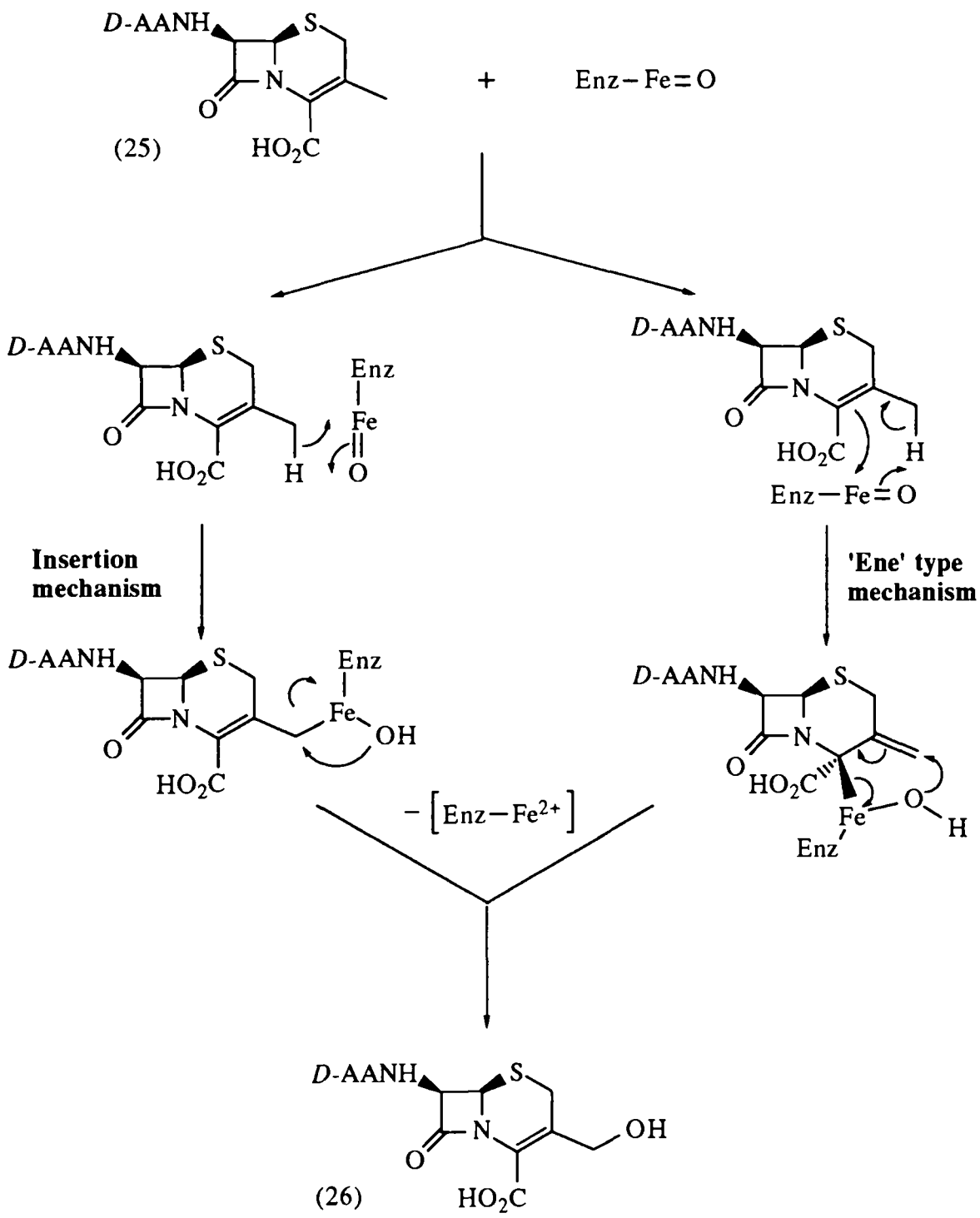



**Figure 4.6** Possible mechanisms for hydroxylation by REXH

A single active site is believed to be involved in both the ring expansion and 3'-hydroxylation processes, due mainly to the observation that substrates for one reaction exhibit competitive inhibition for the other[53]. Therefore, DAOC (25) produced by the ring expansion process has to be released from the enzyme before it can undergo hydroxylation. This will enable the regeneration of a new iron-oxo complex by exactly the same mechanism as described in Section 4.2.1. The mechanism of hydroxylation is less well studied than the ring expansion step. Two mechanisms are conceivable: (1) the direct insertion of an iron-oxo species into a C-H bond and (2) an 'ene' type mechanism (Figure 4.6).

That the first irreversible step in hydroxylation (with respect to DAOC) occurs at C-3' methyl group is confirmed by observation of a kinetic isotope effect on incubating a 1:1 mixture of DAOC (25) and the 3'-trideuteriated-DAOC (49),[106] as depicted in Fig. 4.7.

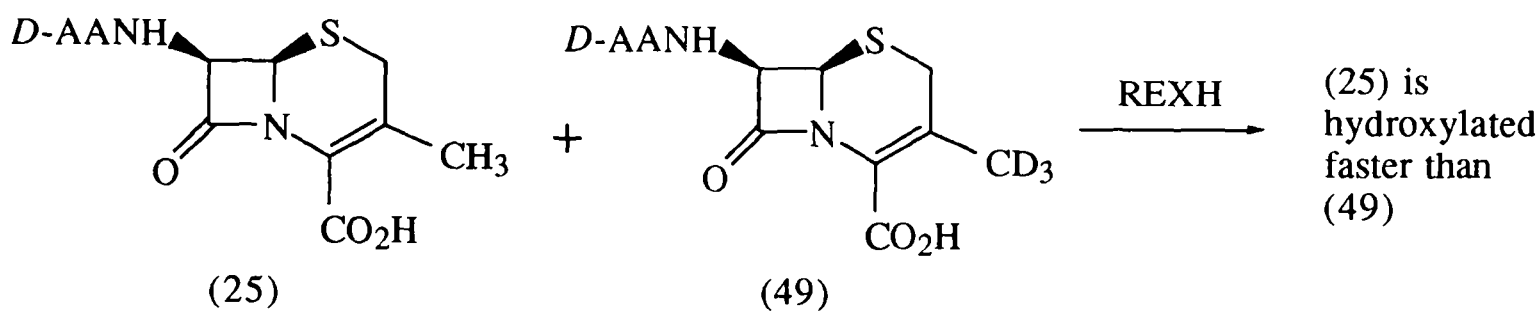



**Figure 4.7**

Townsend *et al.*[107] demonstrated that hydroxylation at the 3'-methyl group proceeds with overall retention of configuration (Fig. 4.8). This is rationalised by postulating that the iron-carbon species in hydroxylation (Fig. 4.6) cannot obtain radical stabilization from the ring sulfur, as in ring expansion step, and it rapidly collapses with retention of stereochemistry to the hydroxylation product. The observation of a net retention of stereochemistry is consistent with other α-KG dependent dioxygenases which carry out hydroxylation reactions.[94]

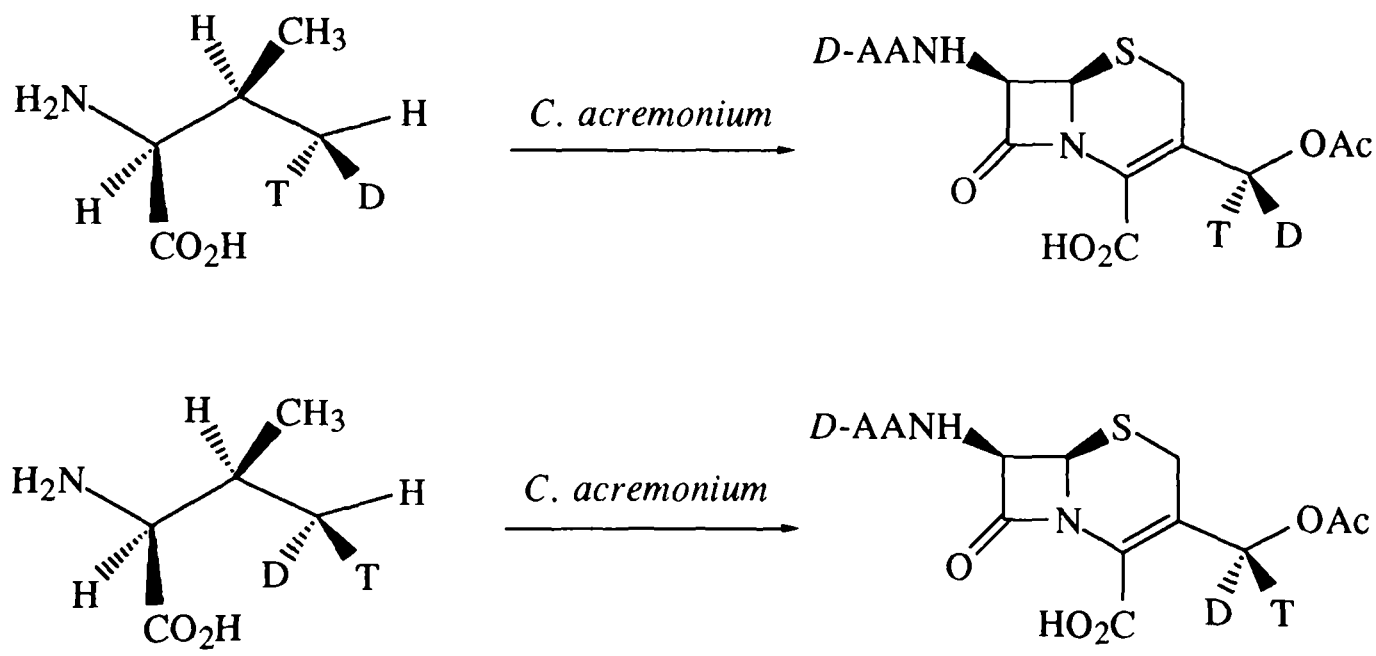



**Figure 4.8**

## 4.3 Previous substrate analogue studies with REXH

The success with which substrate analogue studies had been employed in the mechanistic studies of IPNS[108] prompted a similar line of investigation for REXH. A range of modified penicillins and cephalosporins have been synthesized and tested as substrates for ring expansion and 3'-hydroxylation.

### 4.3.1 Modification at the *N*-acyl side chain

Studies by Baldwin[109] and Yeh[53] indicated that a six-carbon *N*-acyl side chain, terminating in a carboxyl group, permits reasonable penam to cephem conversion by REXH (Table 4.1). This requirement is largely similar to that of IPNS,[108] with an important difference being the inability of REXH to process isopenicillin N (23) which bears the *L*-configured α-aminoadipoyl group.

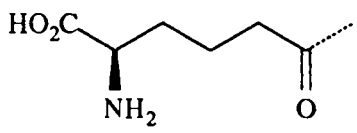
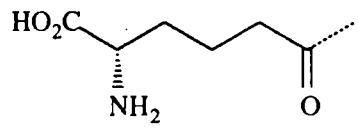
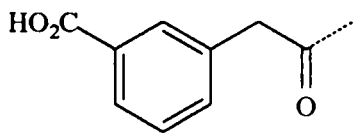
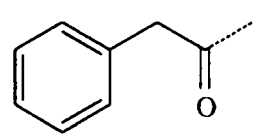
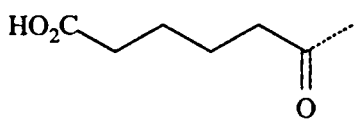
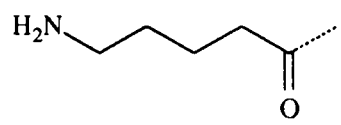
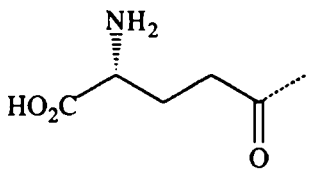
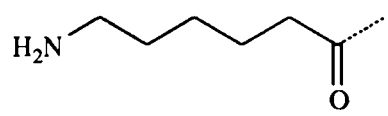
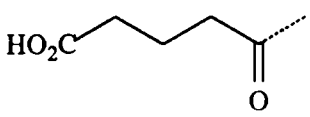
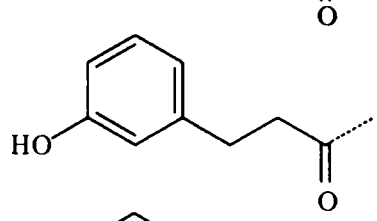
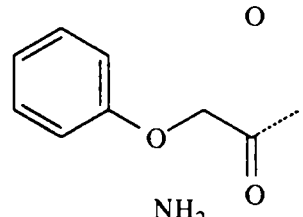
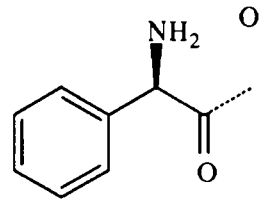
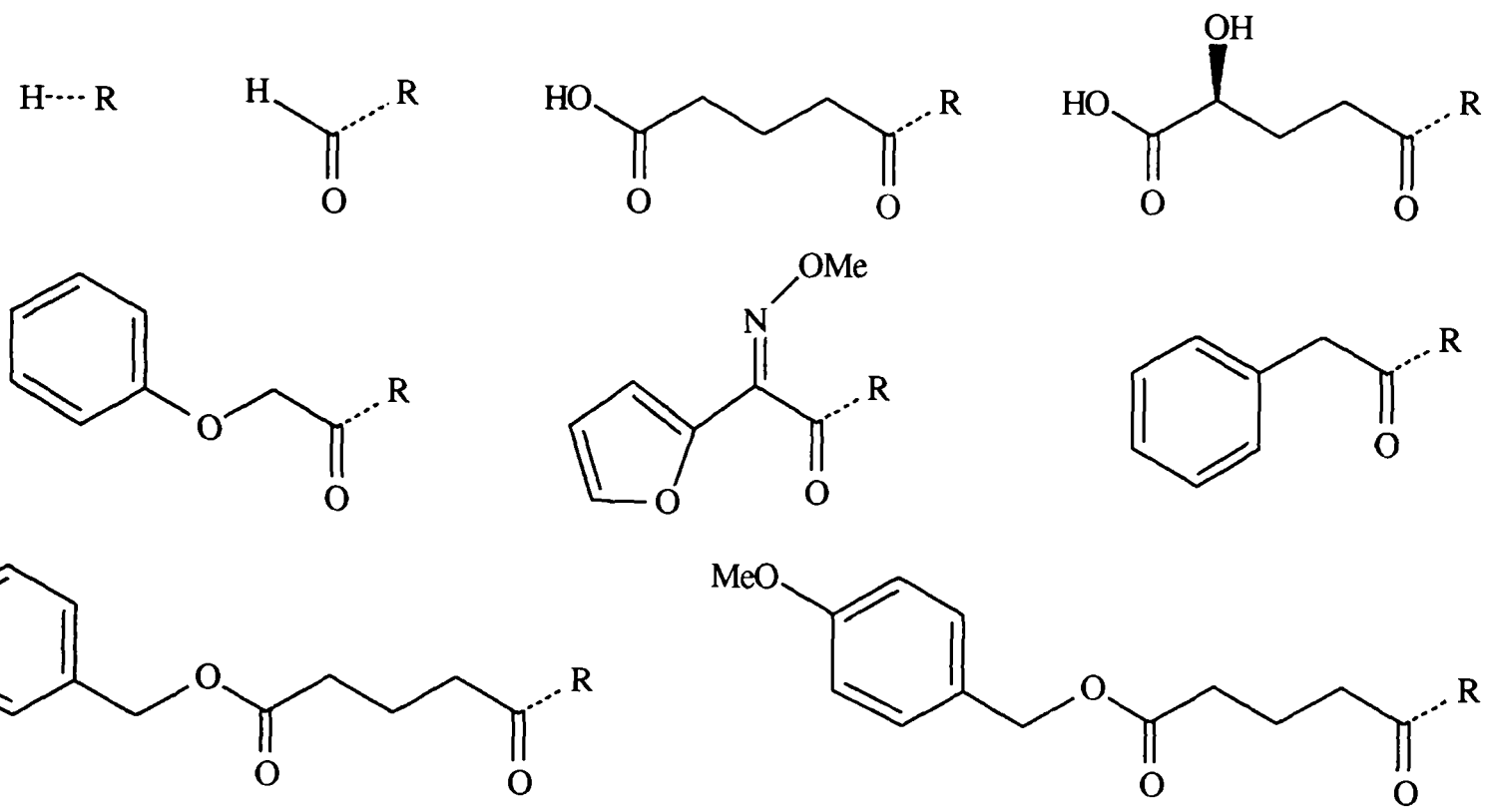
| Side-chains giving conversion | % velocity relative to (24) | Side-chains not giving conversion |
|---|---|---|
|  | 100 |  |
|  | 50 |  |
|  | 40 |  |
|  | 14 |  |
|  | 3 |  |
| | |  |
| | |  |
| | | H |

**Table 4.1** Side-chain variants of penicillin N (24) tested for ring expansion

Studies by Turner[47] and Yeh[53] revealed a high side-chain specificity for the hydroxylase activity of REXH. All the side chain variants evaluated (Figure 4.9) were found to be unacceptable as hydroxylation substrates. However, based on the assumption of a common active site for both the ring expansion and 3'-hydroxylation activities of REXH, the ring expanded products from substrates in Table 4.1 would have served as hydroxylation substrates too. Unfortunately, the possibility of such a further action by REXH was not examined by the workers.[109]



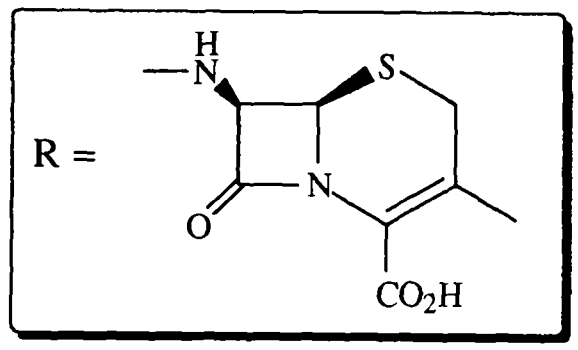
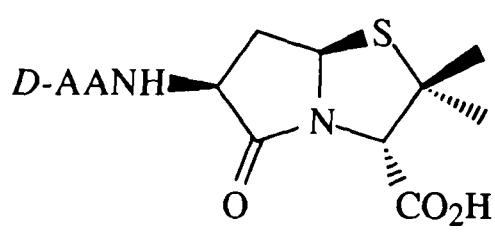




**Figure 4.9** Side-chain variants of DAOC (25) not hydroxylated by REXH

## 4.3.2 Modification at the β-lactam ring

The γ-lactam analogue (85) of penicillin N (24) was evaluated as a potential substrate for REXH in this project. The results are discussed in Section 4.4.

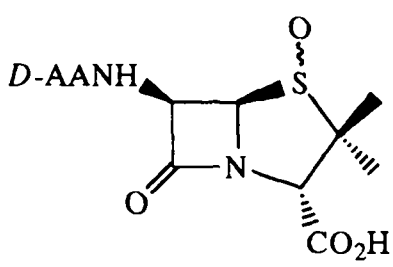



(85)

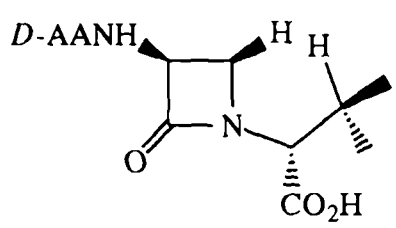
**γ-lactam analogue of penicillin N**

### 4.3.3 Modification at the sulfur atom

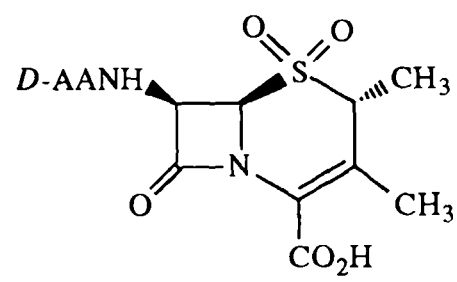
An earlier proposal[110] for the mechanism of ring expansion involved the intermediate formation of penicillin N sulfoxide (50) in analogy to the chemistry employed in the industrial manufacture of cephalosporins (Figure 2.3). Thus, the failure of REXH to act on both the α- and β-epimers of (50) provides evidence against their formation as free intermediates in the reaction, though the possibility of an enzyme-bound intermediate cannot be ruled out.[26]




(50)




(51)

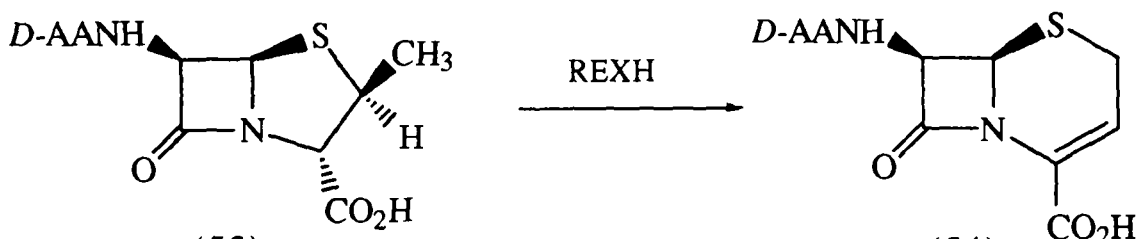



(52)

Studies by Yeh[53] found that dethia-penicillin N (51) and 2α-methyl-DAOC sulfone (52) are not substrates for REXH. The inability of (51) to undergo turnover could reflect either a crucial role for the sulfur atom during the enzymic events or a demanding geometrical requirements for ring expansion. The *S*-carbocyclic analogue of DAOC (25) was tested as a substrate in this project, with the results being discussed in Section 4.4.

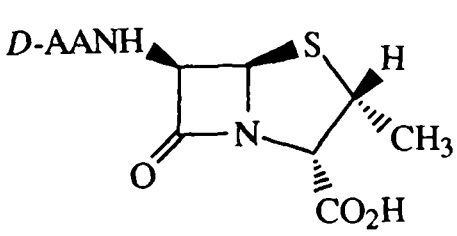
### 4.3.4 Modification at the C-2 methyl groups of penicillin N (24)

2α-Desmethyl penicillin N (53) was converted by REXH into 3-desmethyl-DAOC (54),[111] a result that can be predicted from previous knowledge that the 2α-methyl group of penicillin N (24) becomes the methyl group of DAOC (25) during ring expansion.[102] In the same vein, the 2β-desmethyl analogue (55) is not expected to be a substrate for REXH, though this is yet to be verified experimentally.

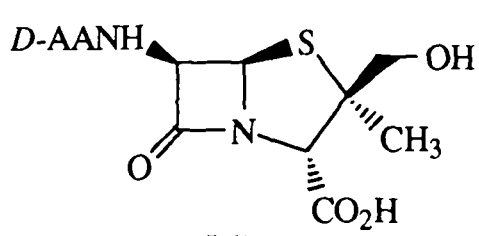



(53) (54)

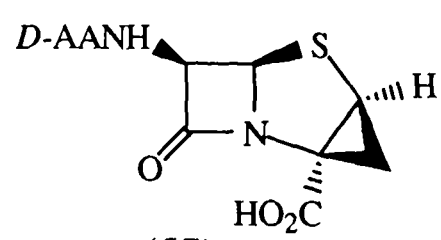
2β-hydroxymethyl penicillin N (56), previously proposed to be an intermediate in ring expansion,[110] was not found to be a substrate for REXH, thus dismissing its possibility as a free intermediate. The tricyclic penam (57) was found to be a potent reversible inhibitor of the ring expansion reaction.[26]




(55)



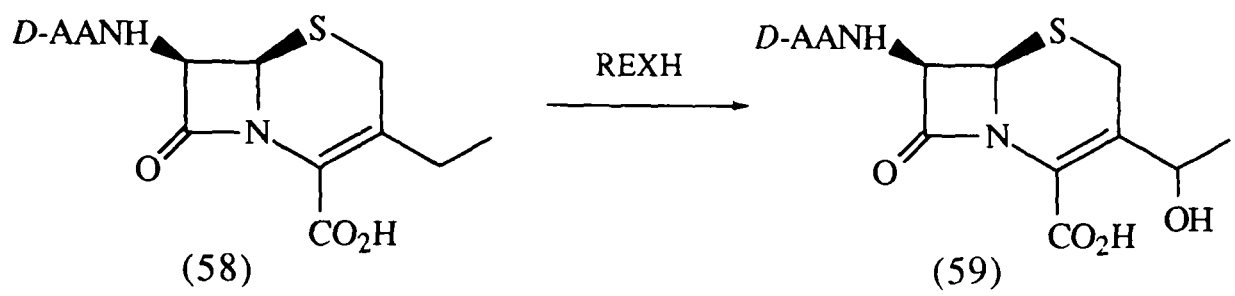

(56)




(57)

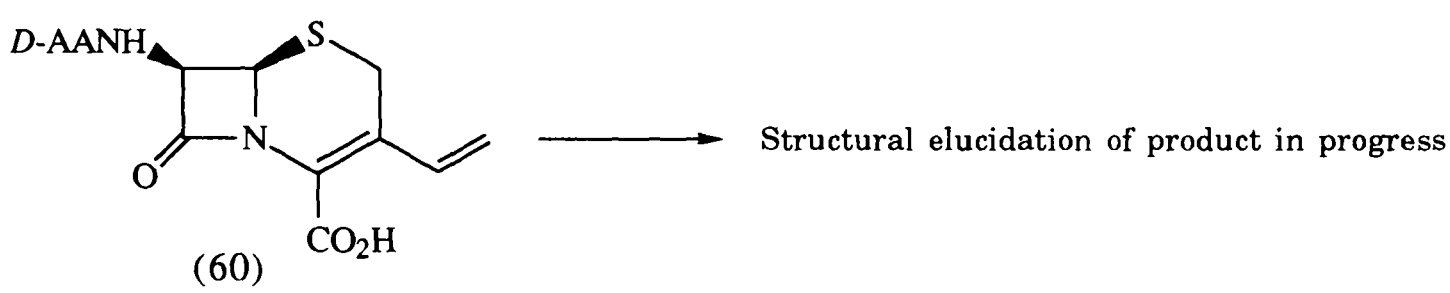
## 4.3.5 Modification at the methyl group of DAOC (25)

The 3-ethyl cephem (58) has been shown to be hydroxylated at the 3'-allylic position, giving (59) with yet undetermined stereochemistry at the new chiral centre.[112]



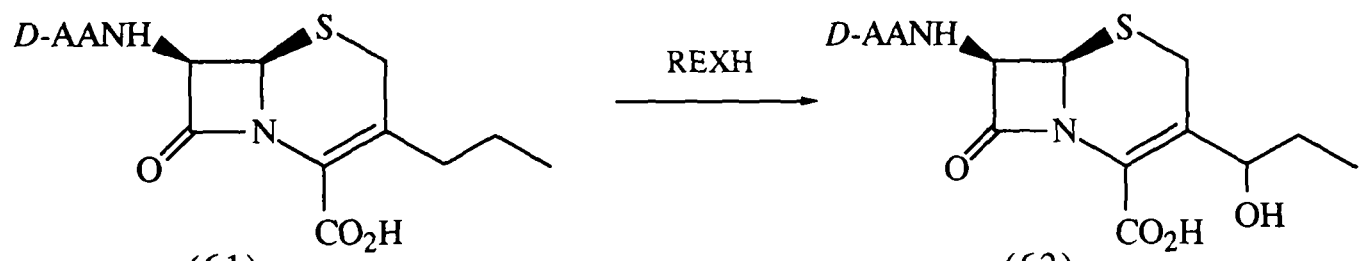

(58) (59)

The NMR spectra of the crude mixtures from incubation of 3-vinyl cephem (60) and 3-(*n*-propyl) cephem (61) indicated the formation of new products.[113] The incubation product from (60) has been isolated by HPLC and structural elucidation is under way.[114]



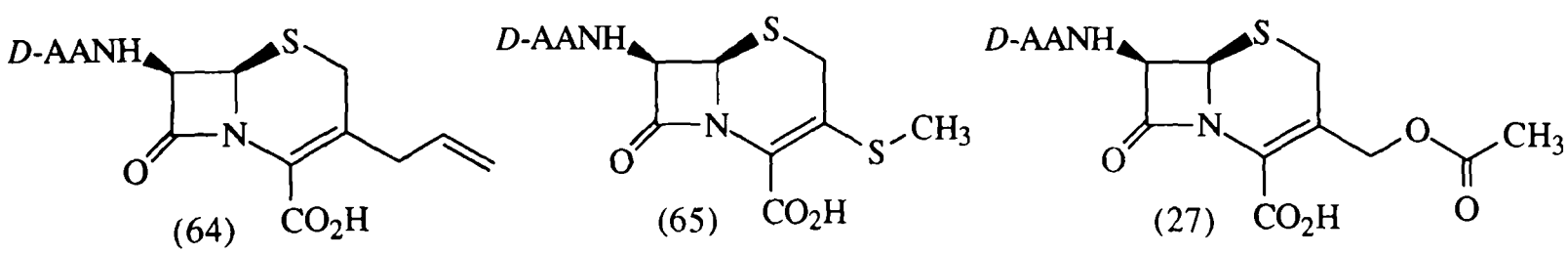

(60)

The product from incubation of (61), apart from its very low level present, has proven to be recalcitrant to isolation. The $^1H$ NMR spectrum of the crude product would suggest a hydroxylation at the 3'-allylic position to give (63).[114]




(61) (63)

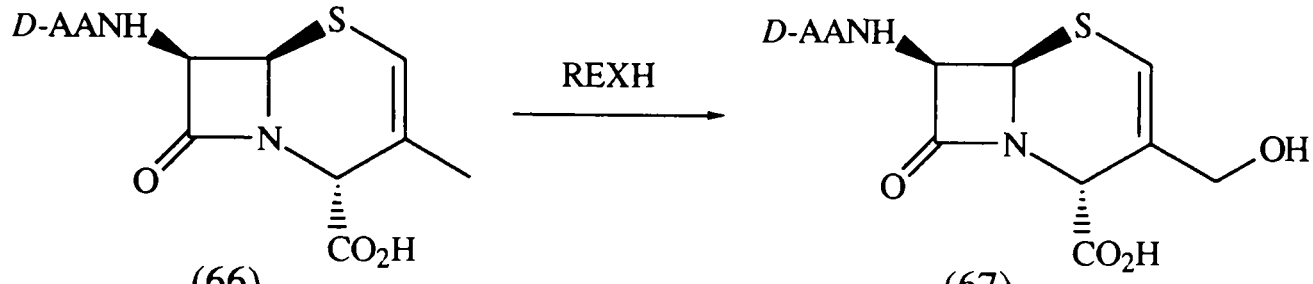
No turnover was detectable when the 3-allyl cephem (64)[115] and 3-thiomethyl cephem (65)[116] were incubated with REXH. The combined results of incubation of (61) and (64) suggested that a flexible 3-carbon unit has perhaps reached the spatial limit imposed by the corresponding domain of the active site. The failure of REXH to convert (65) demonstrated its lack of sulfoxidase activity, a mode of reactivity yet to be discovered among the α-KG dependent dioxygenases. Cephalosporin C (27), the biosynthetic end-product in *C. acremonium* , was not found to interact with REXH in any manner.[53]




(64) (65) (27)

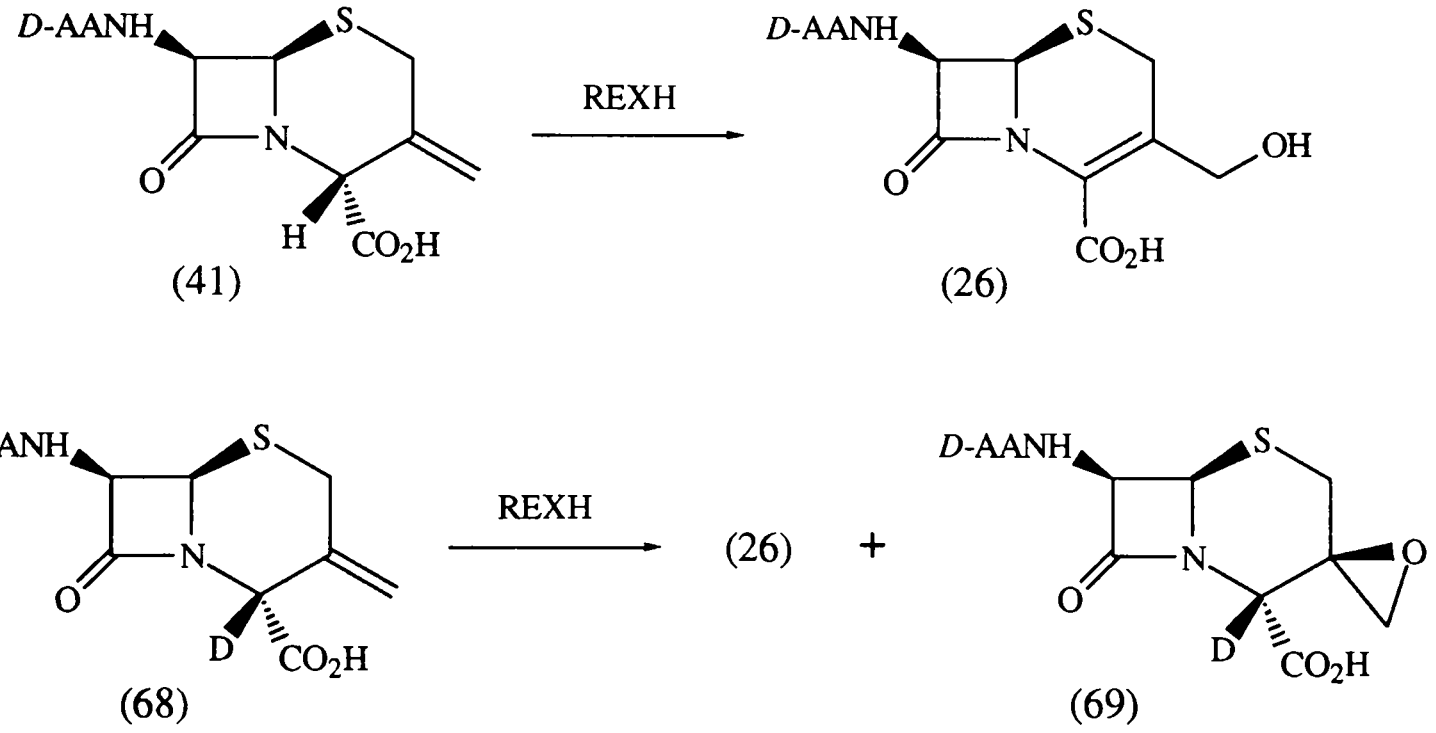
## 4.3.6 Double-bond isomers of DAOC (25)

The C3-C4 double bond in DAOC (25) can be isomerised across C2-C3 to give the Δ-2 cephem (66) or across C3-C3' to give exomethylene cephalosporin C (EMCC, 41). Both compounds have been synthesized and tested as substrates for REXH. Thus, incubation of (66) gave the hydroxylated Δ-2 cephem (67) as the single observable product (by 500 MHz $^1$H NMR).[117]




(66) (67)

EMCC (41) was first shown to be converted into DAC (26) by a cell-free extract of *C. acremonium*.[118] This conversion was later confirmed to be mediated by REXH, with *ca.* 40% incorporation of $^{18}O_2$ into the allylic alcohol of DAC (26).[119] Subsequently, the incubation of 4-deuterio-EMCC (68) was found to give, in addition to DAC (26), the spiro-epoxide cepham (69).[120-122]




(41) (26) (68) (69)

The epoxide formation is rationalised as the biasing of a shunt pathway through the operation of a deuterium kinetic isotope effect,[122] as depicted in Figure 4.10. This first demonstration of epoxidase activity in REXH is the third example of epoxidation carried out by an α-KG dependent dioxygenase, being precedented by thymine 7-hydroxylase[123] and hyoscyamine 6β-hydroxylase.[62] The epoxidation by REXH requires the *R*-configuration at C-4 since the C-4 epimer of EMCC (41) is not a substrate for REXH.[122]

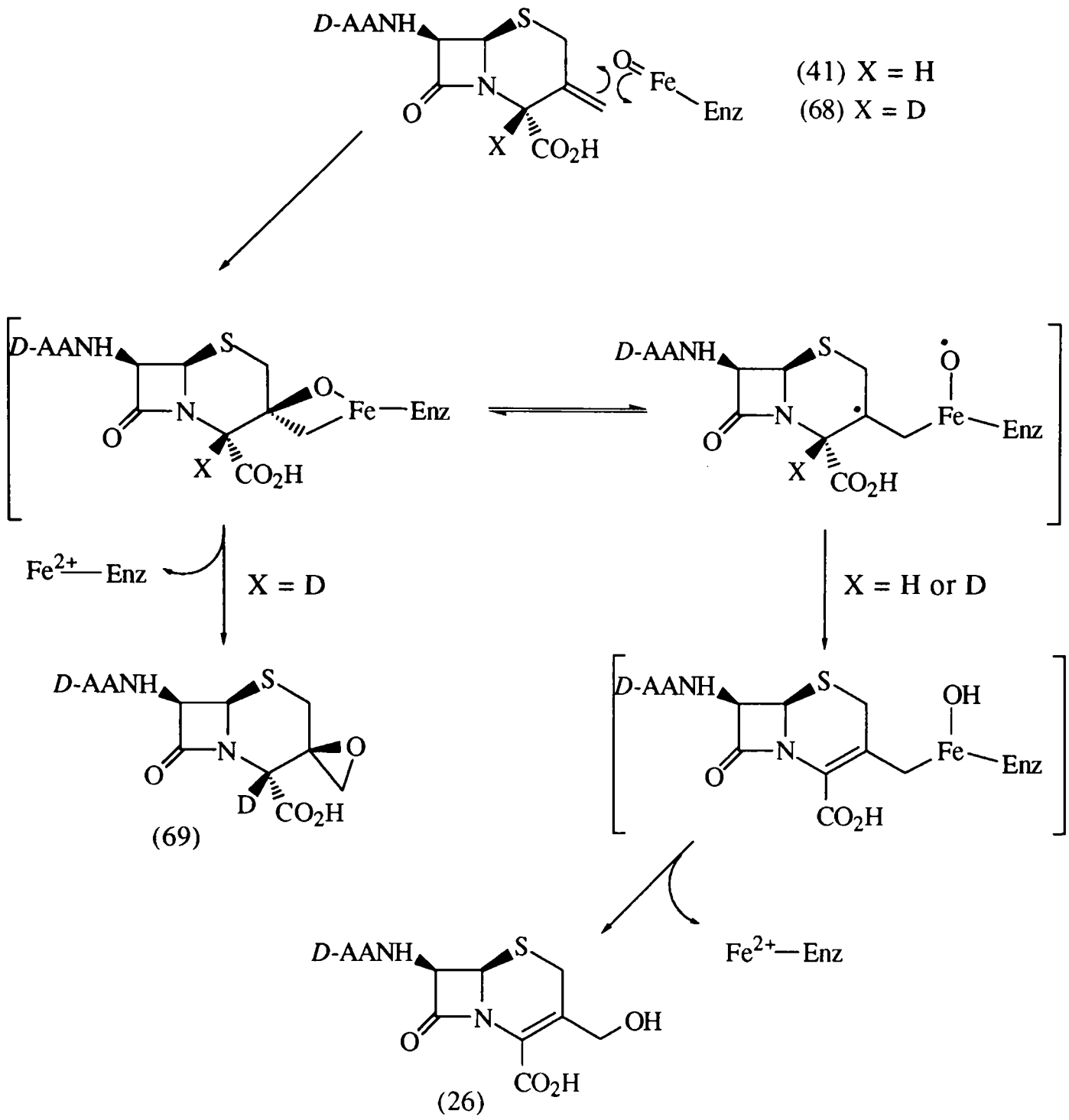
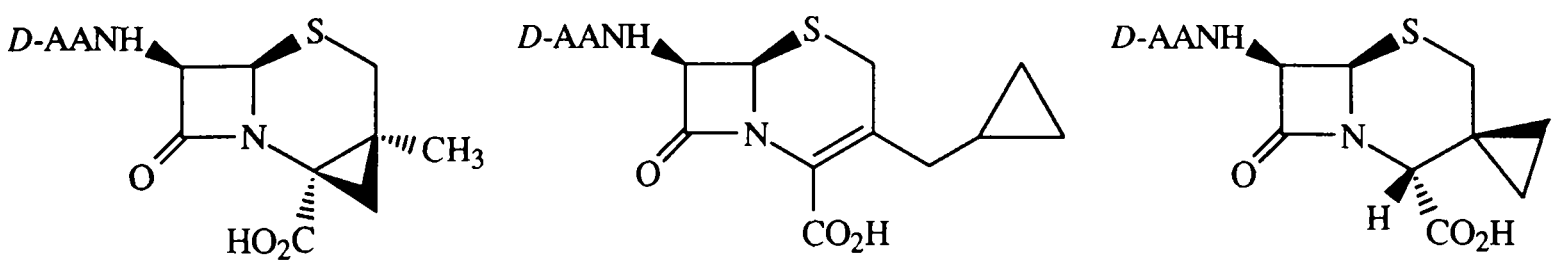



**Figure 4.10** Mechanism for the epoxidase activity of REXH

## 4.3.7 Cyclopropyl analogues of DAOC (25)




(70) (71) (72)

Substrate analogues containing a cyclopropyl ring have been used fruitfully to probe the lifetime and mechanism of decomposition of transient radicals produced in reactions catalysed by the P-450 cytochromes[124] and IPNS.[125] To investigate the hypothetical involvement of radicaloid intermediates in REXH catalysis, three cyclopropyl-containing analogues (70), (71) and (72) were synthesized and incubated with REXH.[126]

Analogues (70) and (71) were found to be neither substrates nor time-dependent inactivators of REXH, indicating their lack of interaction with the active site. The spirocyclic analogue (72) was, however, converted to the 3-hydroxyethyl cephem (73) which could be formed by either the ring opening of a radicaloid species or an 'ene' type mechanism (Figure 4.11).

The turnover of (72) represents the first example of catalytic product formation from a cyclopropyl ring cleavage pathway by an α-KG dependent dioxygenase, further demonstrating the versatility of REXH in its mode of reactivity. The multiple reactivity of REXH (ring expansion - *via* rearrangement and desaturation, hydroxylation, epoxidation, cyclopropyl ring opening) is apparently shared with other α-KG dependent dioxygenases (*e.g.* thymine 7-hydroxylase[123, 127] and hyoscyamine 6β-hydroxylase[62]), and is presumably a reflection of the high reactivity and concomitant low selectivity of the putative iron-oxo intermediate.

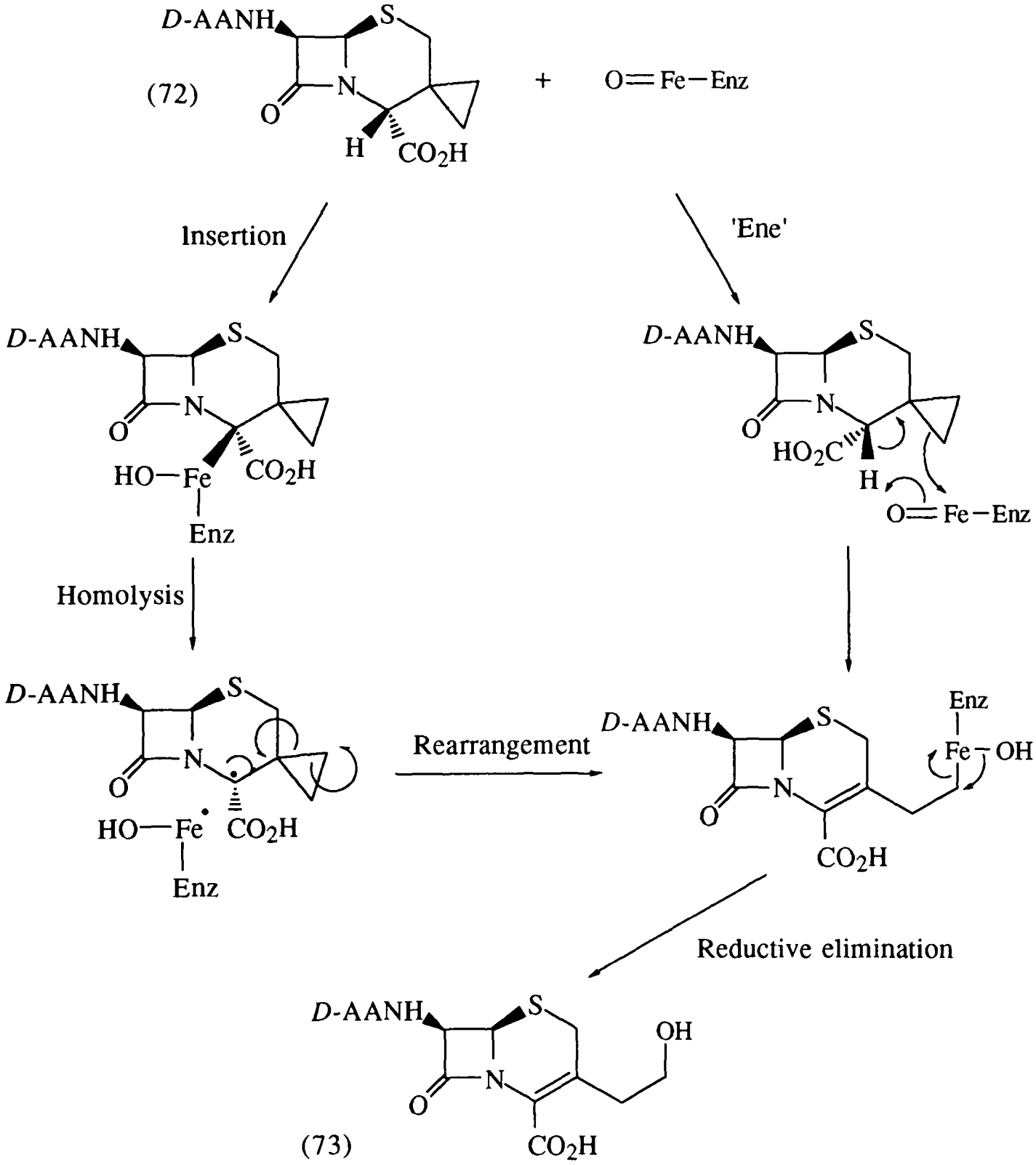



**Figure 4.11** Cyclopropyl ring opening catalysed by REXH

## 4.4 Results and Discussion

In the course of this project, there was opportunity to test three substrate analogues as potential substrates for REXH.

### 4.4.1 Incubation with carba-DAOC (75)

The 50% sequence homology in an important 10-residue region of REXH and IPNS and the possibility of a common evolutionary origin were discussed in Section 2.2. Mechanistic investigations outlined in the last two sections have indicated that the chemistry of REXH bears a strong similarity to that of IPNS,[108] and the substrate analogue and kinetic studies in both cases have been rationalised by invoking the intermediacy of a reactive iron-oxo species.[125] A major difference lies in the stoichiometry of dioxygen utilisation by the two enzymes. For IPNS, an overall four-electron oxidation of ACV(20) is achieved, concomitant with the conversion of one molecule of dioxygen to two molecules of water (Figure 4.12). In the case of REXH, only a two-electron oxidation of the β-lactam substrate is achieved for each molecule of dioxygen consumed during each reaction, since the cosubstrate α-KG undergoes simultaneous oxidative decarboxylation to succinate.[128]

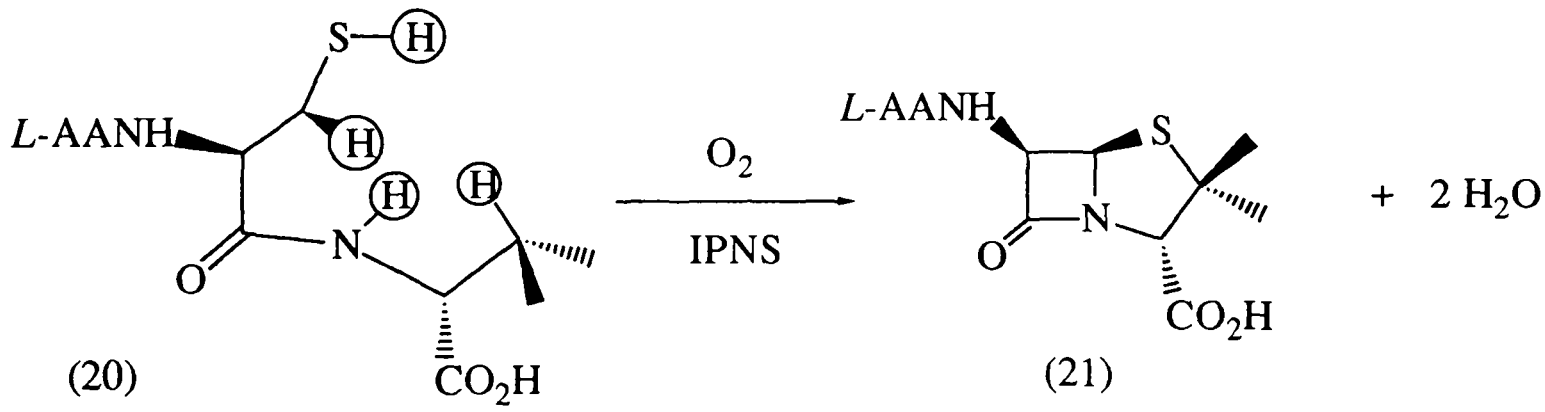
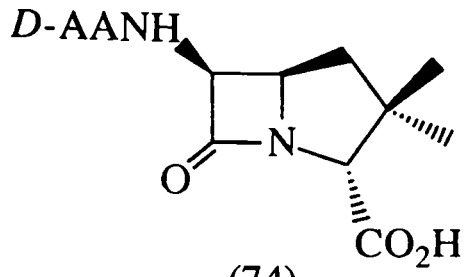


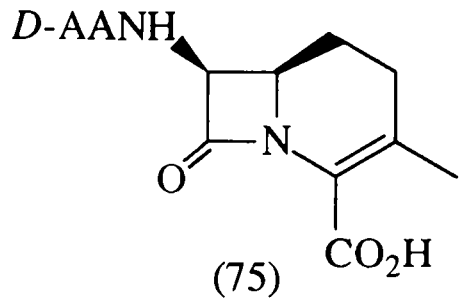

**Figure 4.12** 4-electron oxidation of ACV (20) by IPNS

Spectroscopic studies with IPNS have indicated that one equivalent of iron is bound per mole of enzyme and that in the resting enzyme the metal is in the high-spin ferrous state.[129] Further studies are consistent with the proposal[125] that binding of ACV (20) to IPNS results in the formation of a metal-sulfur bond.[130, 131] Three site-directed mutants of *C. acremonium* IPNS [Cys-106 → Ser-106, Cys-255 → Ser-255, (Cys-106, Cys-255) → (Ser-106, Ser-255)] were shown to retain some activity,[54] indicating that the formation of the metal-sulfur bond probably results from interaction of the thiol group of ACV (20) directly with metal. The formation of such a bond is an integral part of the proposed mechanism for IPNS.[125]

Given the pivotal role assigned to the thiol of ACV (20) in the action of IPNS, it would be interesting to explore the role of sulfur in the ring expansion and 3'-hydroxylation steps, using the carbocyclic analogues carba-penicillin N (74) and carba-DAOC (75) respectively.

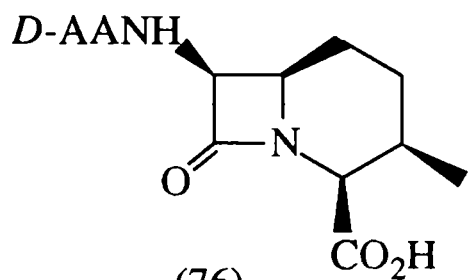



(74)

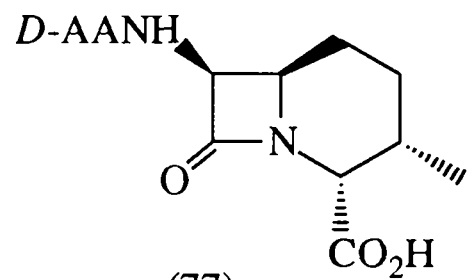



(75)

As yet, (74) has proved to be an unrealised synthetic target, possibly due to the intrinsic instability of the compound. Carba-cephem (75) was synthesized and found not to be a substrate of REXH by scientists at Eli Lilly.[53] Nevertheless, a gift sample of (75) was obtained for re-evaluation in this project. The sample was impure based on its HPLC trace [$C_{18}$ 250 x 10 mm, 25 mM ammonium bicarbonate, 4 ml/min, Detection at 220 & 260 nm] and was therefore purified in our laboratory. Carba-DAOC (75), constituting *ca.* 45% of the gift sample by weight, eluted at 13.0 min. The other two major 220 nm-absorbing peaks eluted at 10.7 and 17.8 min. Both contained 2 additional mass units (by electrospray MS) and were assigned as the isomeric cepham products (76) and (77) resulting from over-hydrogenation during the deprotection step.

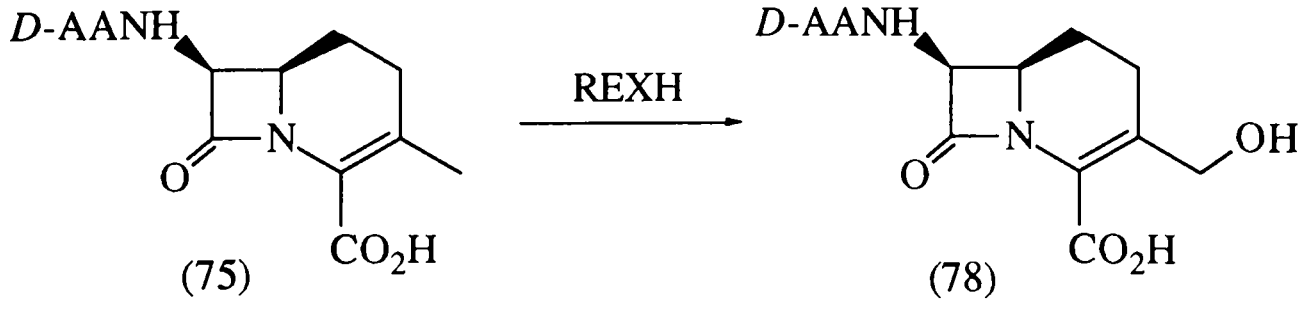



(76)




(77)

Purified (75) (2 mg) was incubated with refolded recombinant REXH under the conditions described in Section 4.5. The 500 MHz $^1$H NMR spectrum for the crude deproteinated mixture revealed the presence of an AB quartet, at *ca.* 4.1 ppm, which is reminiscent of the 3'-methylene protons in DAC (26). Upon HPLC [$C_{18}$ 250 x 4.6 mm, 25 mM ammonium bicarbonate, 1 ml/min, Detection at 220 & 260 nm], unreacted substrate eluted at 11.2 min while a new cephem product ($R_t$ = 4.9 min) was identified as carba-DAC (78) by mass spectrometry [m/z ($MH^+$) = 356] and NMR (Figure 4.13) [$\delta_H$ ($D_2O$, 500 MHz) ; 1.47-1.51 ( 1 H, m, 5-H ), 1.57-1.62 ( 2 H, m, C*H*$_2$CH$_2$CO ), 1.72-1.87 ( 2 H, m, C*H*$_2$CH$_2$CH$_2$CO ), 1.88-1.96 ( 1 H, m, 5-H ), 2.28-2.37 ( 4 H, C*H*$_2$CO, 2 x 4-H ), 3.62-3.67 ( 1 H, *m,* NC*H*CH$_2$CH$_2$ ), 3.81-3.86 ( 1 H, m, 6-H ), 4.13-4.18 ( 2 H, ABq, *J* = 13 Hz, C*H*$_2$OH ), 5.22 ( 1 H, d, *J* = 4.5 Hz, 7-H )]. A 2D-COSY experiment was consistent with the connectivity found in structure (78).




(75) (78)

Kinetic analysis of carba-DAOC as a substrate for refolded recombinant REXH gave $K_m$ = 710 ± 50 μM and $V_{max}$ = 129 ± 7 mU/mg-protein, compared with reported values of 37 ± 3 μM and 400 ± 20 mU/mg-protein respectively for the conversion of DAOC (25).[53] The significant increase in $K_m$, reflecting a lower affinity of the enzyme for the unnatural substrate, could probably be attributed to conformational changes caused by the substitution of the sulfur atom with a methylene group. The difference in $V_{max}$ was much less marked, indicating that the sulfur atom in DAOC (25) is not directly involved in any bonding changes during hydroxylation and hence exerts a smaller effect on the reaction kinetics. Compound (75) exhibited ~ 40% of the antibacterial activity of DAOC (25) against *E. coli* X580, while compound (78) exhibited ~ 75% of the activity of DAC (26). More experiments are required to find out if the inferior activity of the carba-cephems is caused by less efficient binding to the target bacterial enzymes or reduced permeability into the cells.

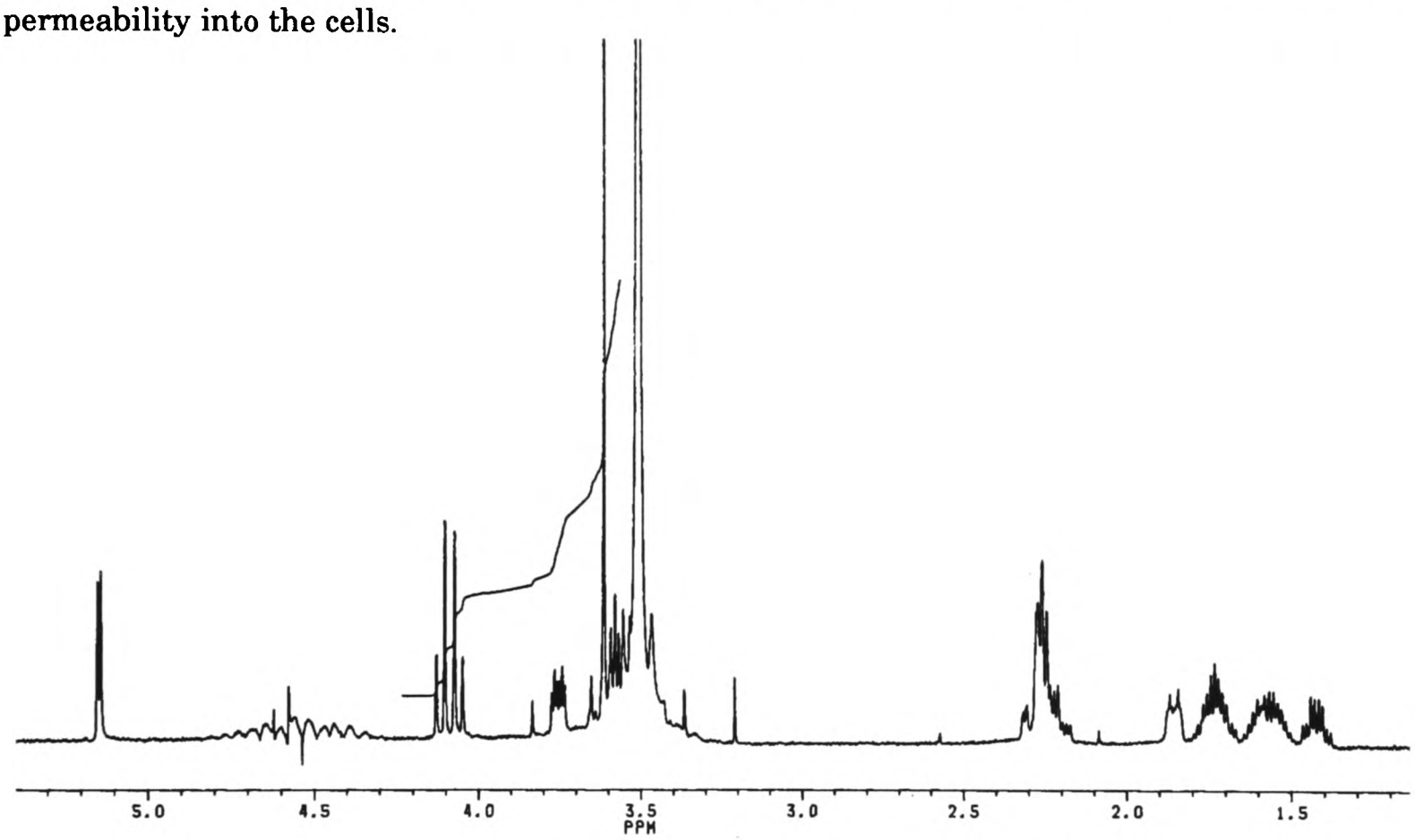



**Figure 4.13** $^1H$ NMR spectrum of carba-DAC (78)

Incubation of (75) under an atmosphere of $^{18}O_2$ produced (78) that contains about 70% of the heavier isotope at its allylic oxygen (Table 4.2).

| | | MH⁺ for (78) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Obtained in $^{16}O_2$ | **m/z** | 355 | 356 | 357 | 358 | 359 | 360 | 361 |
| | **%** | 0 | 100 | 19 | 6 | 2 | 2 | 3 |
| Obtained in $^{18}O_2$ | **m/z** | 355 | 356 | 357 | 358 | 359 | 360 | 361 |
| | **%** | 7 | 42 | 12 | 100 | 20 | 5 | 2 |

**Table 4.2** ESMS data for carba-DAC (78) from incubation of carba-DAOC (75) with REXH

This less-than-stoichiometric incorporation has several precedents, not only in other conversions carried out by REXH (Table 4.3), but also in the action of other α-KG dependent dioxygenases such as clavaminic acid synthase.[132]

While the above results have been attributed partly to the presence of residual $^{16}O_2$ in the enzyme solution, it is now unequivocal that the putative iron-oxo intermediate can undergo oxygen exchange with water, as demonstrated by the incorporation of $^{18}O$-label from $H_2{}^{18}O$ into the products[132, 133](Table 4.4). Therefore, it is speculated here that an incubation of (75) in $^{18}O$-labelled water would give results similar to those in Table 4.4.

| **Conversion** | **Level of $^{18}O$ incorporation from $^{18}O_2$** | **References** |
|---|---|---|
| DAOC (25) → DAC(26) | 30 - 40 % | [50, 99, 100] |
| EMCC (41) → DAC(26) | 30 - 40 % | [119, 134] |
| Penicillin N (24) → 3β-hydroxycepham | 50 - 60% | [99, 100] |

**Table 4.3** $^{18}O_2$ incorporation for REXH-catalysed reactions

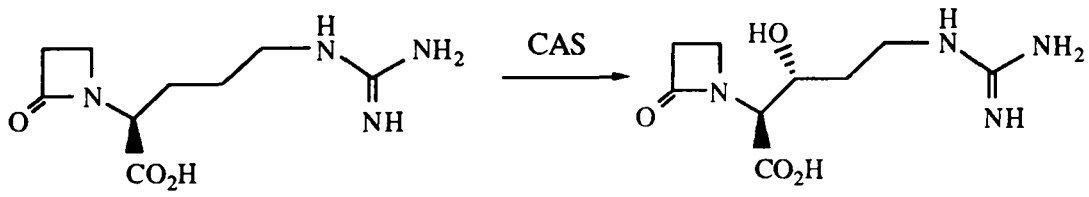
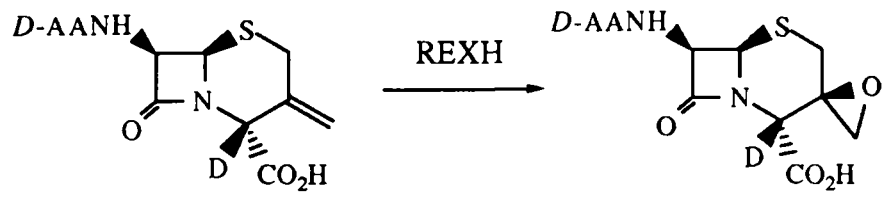
| **Enzymic conversion in $H_2{}^{18}O$** | **Level of incorporation** | **Ref.** |
|---|---|---|
|   | 9 % | [132] |
|   | 16 % | [133] |

**Table 4.4** $^{18}O$ incorporation from enzyme incubation in $H_2{}^{18}O$

The failure of Eli Lilly scientists to detect turnover of carba-DAOC (75) is puzzling. One possible explanation is that the 260 nm-transparent dihydro isomers, (76) and (77), in the impure sample could have been inhibitors of hydroxylase activity of REXH, and that their presence (~ 55%) prevented the hydroxylation of carba-DAOC (75) into carba-DAC (78). There has been no attempt to verify this conjecture. In any case, the evidence presented above showed unambiguously that unlike the case of IPNS, the sulfur moiety in DAOC (25) is not necessary for the hydroxylation mechanism of REXH.

## 4.4.2 Incubation with DAC (26)

The final step in the biosynthesis of cephalosporin C (27) in *C. acremonium* consists of the acetylation of DAC (26) by the enzyme DAC *O*-acetyltransferase (Figure 1.1). The enzyme activity in cell-free extracts was first reported independently by Fujisawa[30] and Liersch,[135] then purified to near homogeneity by Scheidegger[31] and the gene was ultimately cloned by a Japanese group.[29]

Fujisawa have also reported the generation of *C. acremonium* mutants capable of accumulating excess DAC (26) in the culture broth,[136] and subsequently showed that these mutants are deficient in the enzyme DAC *O*-acetyltransferase.[30] In the event, a new metabolite, 3-formyl cephalosporoate (79), was isolated from one of these mutants blocked in the terminal acetylation step.[137] It was proposed that (79) arises from an intracellular oxidation of DAC (26) to the bicyclic aldehyde (80), and then undergoes facile β-lactam cleavage to give the cephalosporoate[47, 137, 138] (Figure 4.14).

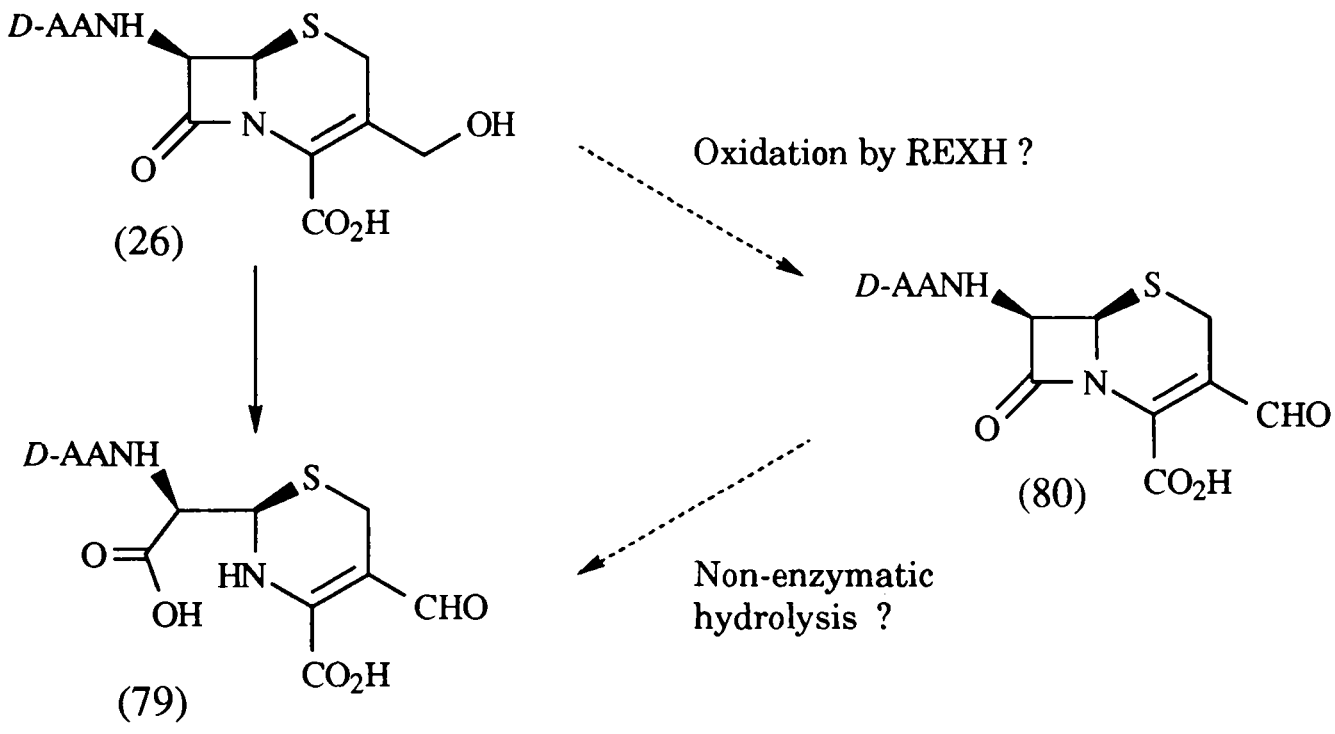



**Figure 4.14** Fate of DAC (26) in cephalosporin C-negative mutants of *C. acremonium*

There was, however, no previous speculation of REXH as the enzyme involved, though REXH is already responsible for the two successive oxidative steps that lead to DAC (26). As a result of this project being part of our larger research program on α-KG dependent dioxygenases, we were especially aware of the oxidative versatility of this class of enzymes and of the existence of a particular member capable of three consecutive oxidations (thymine 7-hydroxylase).[127] In this regard, we thought it worthwhile to evaluate DAC (26), an established product of REXH action, for its susceptibility to further oxidation by the enzyme.

Initially, we investigated the efficiency of DAC (26) as a substrate for soluble recombinant REXH, using standard incubation conditions described in Section 4.6. Thus, a typical reaction mixture (5 ml) contained the following components, in 50 mM Tris-HCl buffer (pH 7.5): 1 mg/ml enzyme, 1 mM DAC, 1 mM ascorbate, 2 mM DTT, 2 mM α-KG, 50 μM ferrous

sulfate and 50 mM ammonium sulfate. Examination of the $^1$H NMR (500 MHz) of the crude deproteinated mixture indicated the appearance of new resonances at *ca.* 9.05 ppm, pointing to the possible production of aldehydic products (Figure 4.15).

Initial purification of the crude incubation mixture by HPLC (Hypersil ODS column, 250 x 7 mm) was achieved using 0.1 % formic acid as eluant. The eluate was monitored at 260 nm [for unreacted DAC (26), retention volume *ca.* 39 ml] and at 300 nm [for the aldehyde product (79), retention volume *ca.* 21 ml].

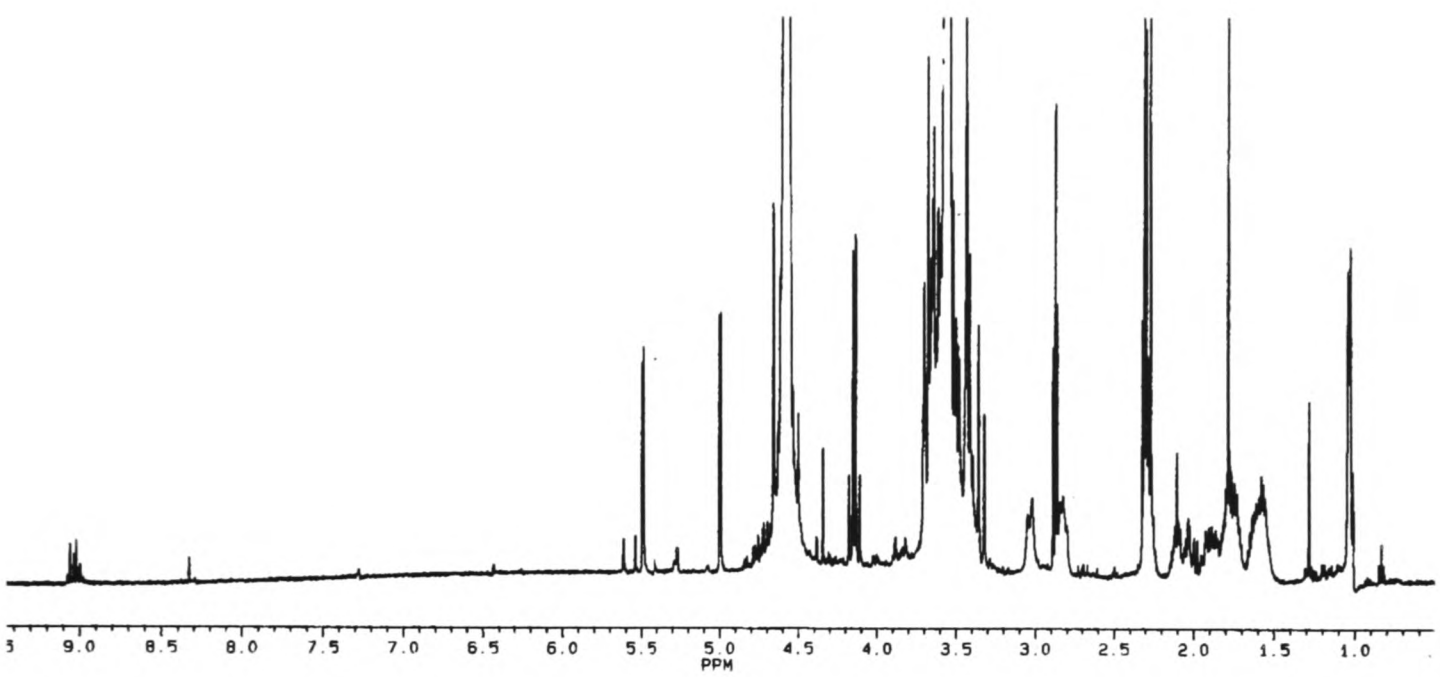


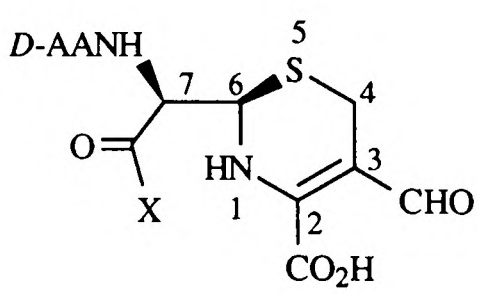

**Figure 4.15** $^1$H NMR spectrum for crude incubation mixture of DAC with REXH

Further HPLC purification on the same column, using 25 mM ammonium bicarbonate as eluant, gave an apparently pure [by $^1$H NMR (500 MHz)] product (retention volume, *ca.* 8.1 ml). The product gave UV and NMR spectra similar to those reported for the 3-formyl cephalosporoate (79).[137] However, electrospray mass spectrometry gave a value of m/z = 389 for the protonated molecular ion $MH^+$, as opposed to the anticipated value of 390 [for (79)], consistent with ring opening of the β-lactam by ammonia to give the amide (81) (Figure 4.16) ; $^1$H NMR (500 MHz, $D_2O$) $\delta_H$ 1.51 ~ 1.64 and 1.71 ~ 1.81 (4H, 2 x m, $CH_2CH_2CH_2CO$), 2.28 (2H, t, $J$ = 7.0 Hz, $CH_2CO$), 3.33 and 3.50 (2H, ABq, $J$ = 15.5 Hz, 2 x 4-H), 3.62 (1H, t, $J$ = 6.0 Hz, $HOOCCH(NH_2)CH_2$), 4.67 (1H, d, $J$ = 7.5 Hz, 6-H), 4.80 (1H, d, $J$ = 7.5 Hz, 7-H), 9.05 (1H, s, CHO); UV absorption spectrum gave a single maximum at 300 nm.

**Figure 4.16** Proposed structures of aldehyde products arising from action of REXH on DAC.

(79) X = OH (82) X = NH–$C(CH_2OH)_3$

(81) X = $NH_2$




Further incubations were then carried out in the absence of ammonium sulfate to avoid amidation. The product, isolated as before, was found to have a larger retention volume (*ca.* 38 ml using 0.1 % formic acid as eluant and *ca.* 17 ml using 25 mM ammonium bicarbonate as

eluant). ESMS and NMR data, however, indicated that opening of the β-lactam ring by Tris [$(HOCH_2)_3CNH_2$] had occurred, giving rise to (82) (Figure 4.16) ; $^1$H NMR (500 MHz, $D_2O$) $\delta_H$ 1.50 ~ 1.66 and 1.71 ~ 1.81 (4H, 2 x m, $CH_2CH_2CH_2CO$), 2.28 (2H, t, $J$ = 7.0 Hz, $CH_2CO$), 3.36 and 3.50 (2H, ABq, $J$ = 15.0 Hz, 2 x 4-H), 3.62 (6H, s, $NHC(CH_2OH)_3$) 3.66 (1H, t, $J$ = 6.0 Hz, $HOOCCH(NH_2)CH_2$), 4.69 (1H, d, $J$ = 8.0 Hz, 6-H), 4.77 (1H, d, $J$ = 8.0 Hz, 7-H), 9.08 (1H, s, CHO); electrospray MS positive mode m/z 493 $MH^+$.

Finally, incubations were carried out in 50 mM sodium phosphate buffer (pH 7.5), in the absence of ammonium sulfate, other conditions being identical to the previous incubations. Reverse phase HPLC (Hypersil ODS, 250 x 10 mm) gave a different aldehyde product [retention volume *ca.* 37 ml (0.1 % formic acid) and *ca.* 12 ml (25 mM ammonium bicarbonate]. $^1$H NMR and mass spectrometric data were consistent with the production of the previously[137] observed aldehyde (79) as the product : $^1$H NMR (500 MHz, $D_2O$) $\delta_H$ 1.54 ~ 1.64 and 1.75 ~ 1.82 (4H, 2 x m, $CH_2CH_2CH_2CO$), 2.28 (2H, t, $J$ = 7.0 Hz, $CH_2CO$), 3.39 and 3.46 (2H, ABq, $J$ = 16.0 Hz, 2 x 4-H), 3.66 (1H, t, $J$ = 6.0 Hz, $HOOCCH(NH_2)CH_2$), 4.59 (1H, d, $J$ = 6.0 Hz, 6-H), 4.84 (1H, d, $J$ = 6.0 Hz, 7-H), 9.04 (1H, s, CHO); electrospray MS positive mode m/z 390 $MH^+$.

Control experiments either using boiled enzyme or in the absence of α-KG did not result in the turnover of DAC (26) to any aldehydic products. Incubations under an $^{18}O_2$ atmosphere were conducted in order to determine the possible level of dioxygen incorporation into the aldehydic product (79). The mass spectrometric data revealed little (< 5 %) or no incorporation of the dioxygen label, in contrast to similar experiments for the conversion of DAOC (25) to DAC (26),[119] when significant incorporation of labelled oxygen from dioxygen into the product was observed. This apparent difference may be due either to the aldehyde undergoing exchange with $H_2{}^{16}O$, such that all initial label, if present, is eventually washed out, or may result from the mechanistic constraints of the reaction. All three aldehyde products, (79), (81) and (82), did not possess any biological activity against *E. coli* X580 (the routine test organism for cephalosporins in this project) which was not unexpected since the β-lactam ring is the basis for such activity. Preliminary kinetic analysis of DAC (26) as a substrate for the soluble recombinant enzyme in Tris buffer gave values of $K_m$ = 1.33 mM and $V_{max}$ = 180 mU/mg-protein. Incubations carried out with the native fungal enzyme also demonstrated turnover, further extending our previous observation of a similar substrate specificity for native and recombinant proteins.[117]

REXH is thus capable of further oxidising DAC (26) to cephem-3-carbaldehyde (80), which is probably hydrolysed non-enzymatically to give 3-formyl cephalosporoate (79). Again, this result stands in contrast to an earlier report by Lilly scientists[53] that DAC (26) is an inhibitor of expandase activity, but not a substrate, for REXH. Alcohol oxidation can now be added to the list of oxidative reactions carried out by REXH (desaturation, hydroxylation and epoxidation). The finding is also interesting in the following respects. Firstly, to our knowledge, thymine 7-hydroxylase[53] is the single precedent in the family of α-KG dependent dioxygenases to exhibit an ability to process a substrate through three successive oxidative stages. This enzyme converts the methyl group of thymine all the way through to a carboxylic acid, the

highest possible oxidation state (Figure 4.17). By analogy, REXH might be capable of further oxidising the bicyclic aldehyde (80) if the latter compound is not rapidly degraded by β-lactam ring opening. Compared with the above two enzymes, clavaminic acid synthase (CAS) is unusual in that it is capable of catalysing, *in vitro*, three oxidative steps in the clavulanic acid pathway with only two of them being consecutive in the biosynthetic sequence (Figure 6.3). Whether such an ability to execute consecutive oxidations is more widespread in the family of α-KG dioxygenases or restricted to these three particular members remains an open question. It is speculated here that a product can be oxidised further by the same α-KG coupled enzyme as long as it can fit back into its active site after the regeneration of the active iron-oxo species.

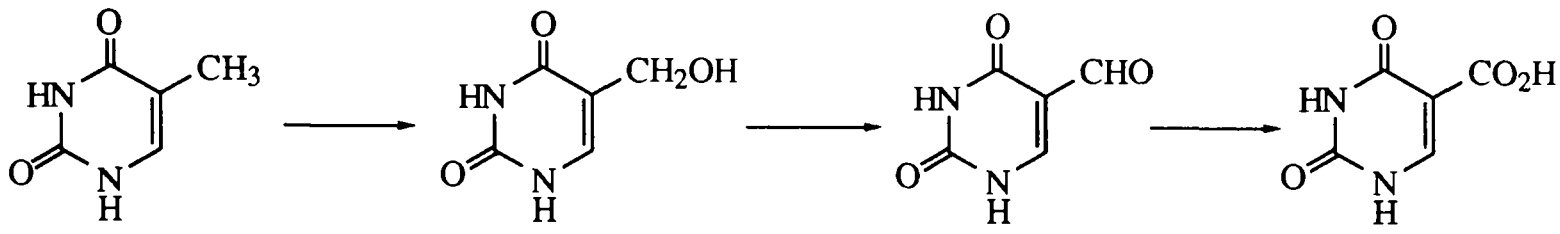



**Figure 4.17** Action of thymine 7-hydroxylase

In addition, it is of interest to compare the $K_m$ value obtained for the oxidation of DAC (26) with that reported for the acetylation of DAC (26) to cephalosporin C (27).[31] For the acetylation of DAC (26) as catalysed by DAC *O*-acetyltransferase, the reported $K_m$ value of 40 μM indicates that the acetyltransferase has a higher substrate affinity for DAC (26) than has REXH. Thus, it would seem reasonable to suppose that, *in vivo*, DAC (26) is largely processed through the acetylation pathway to cephalosporin C (26) before further oxidation can occur. Since the oxidation of DAC (26) to (79) results in rapid destruction of the β-lactam ring, it is possible that the acetylation step evolved to maintain the integrity of the cephem ring system.

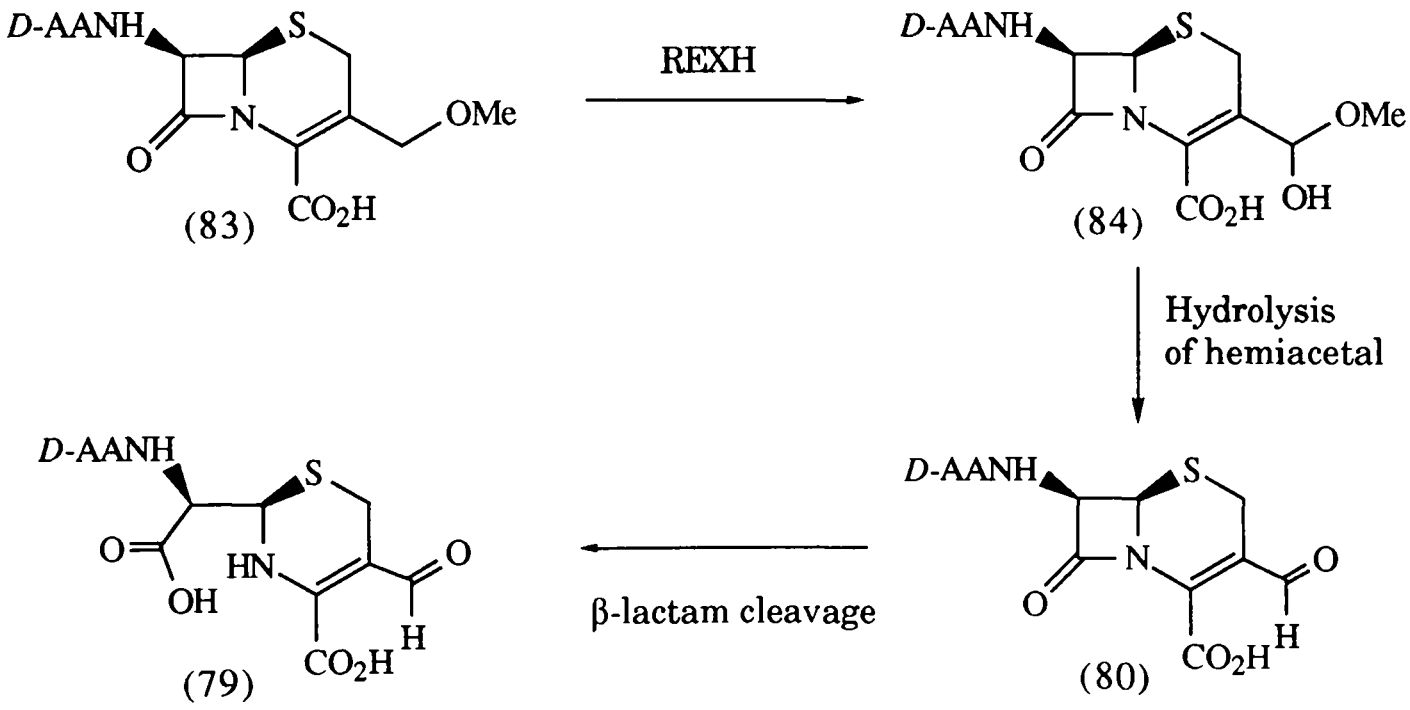


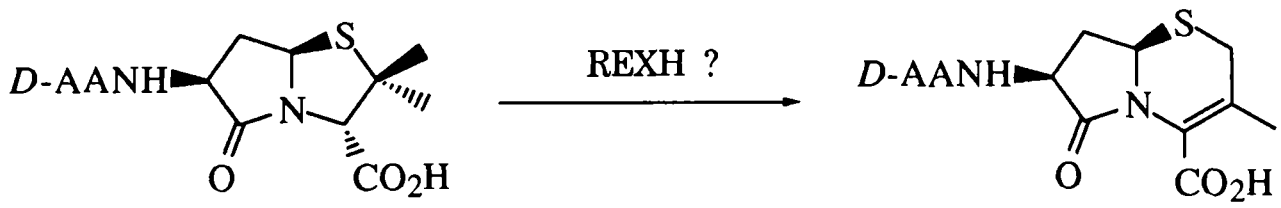

**Figure 4.18** Possible action of REXH on methyl-DAC (83)

Though biological methylation is a very common means of derivatising and deactivating a hydroxyl group, acetylation has been evolutionarily chosen in the case of cephalosporins probably owing to the fact that methyl-DAC (83) would still be susceptible to further oxidation by REXH (Figure 4.18). The resulting hemiacetal (84) would readily yield the aldehyde (80) and lose its β-lactam ring. Indeed, methyl-DAC bears a 3-substituent that is sterically comparable to the *n*-propyl group in (61), a proven unnatural substrate for REXH.

Lastly, the destruction of the β-lactam ring bears important considerations for the industrial production of cephalosporins. It is evident that DAC (26) levels should not be allowed to accumulate during fermentations in order to maximise biosynthetic throughput.

### 4.4.3 Incubation with γ-lactam analogue (85) of penicillin N (24)

The single most important threat to the continued use of β-lactams in anti-infective therapy lies in the endless emergence of resistant pathogens,[18] and one of the most common mechanisms for resistance manifests itself in the production of bacterial enzymes called β-lactamases which inactivate the antibiotics.[20] Medicinal chemists, therefore, have the task of finding new β-lactams that can evade the β-lactamases but exhibit similar antibiotic properties. The synthesis of biologically active non-β-lactam mimics of β-lactams, and the discovery of naturally occurring members, have since challenged the conventional wisdom that the β-lactam ring is a prerequisite for antibacterial action.[139] In the global hunt for biological surrogates of the β-lactam ring, efforts have been devoted to the design of γ-lactam analogues which could serve this role.[140] As such, it would be interesting to explore the possibility of exploiting the versatile oxidative capacity of REXH in γ-lactam synthesis *via* bio-transformation (Figure 4.19). An obvious choice, the γ-lactam analogue (85) of penicillin N was made by a member of our group and found to be inert towards REXH in preliminary experiments.[111] A fresh attempt of this incubation was made in the course of this project.




(85) (86)

**Figure 4.19**

Thus, 3.5 mg of the γ-lactam (85) was incubated with REXH under the conditions described in Section 4.5. The 500 MHz $^1$H NMR spectrum of the crude deproteinated mixture revealed no sign of any product formation, giving only the expected signals for intact (85), α-KG , DTT, ascorbate and Tris-buffer. Upon HPLC ($C_{18}$ 250 x 10 mm, 25 mM ammonium bicarbonate, 4 ml/min, Detection at 220 & 260 nm), the γ-lactam eluted at 43 min, while two minor peaks (220 nm-absorbing only) were collected at $R_t$ = 7.4 and 10.5 min after the solvent front. The absence of 260 nm-absorbing peaks, together with the NMR data, indicated that no products

resembling structure (86) were formed. Hence, the reaction envisaged in Figure 4.19 has not taken place. Surprisingly, the 220 nm-absorbing peak at $R_t$ = 10.5 min gave an NMR spectrum containing all the signals present in the γ-lactam (85), but with shifted $\delta_H$ values (Fig. 4.20).

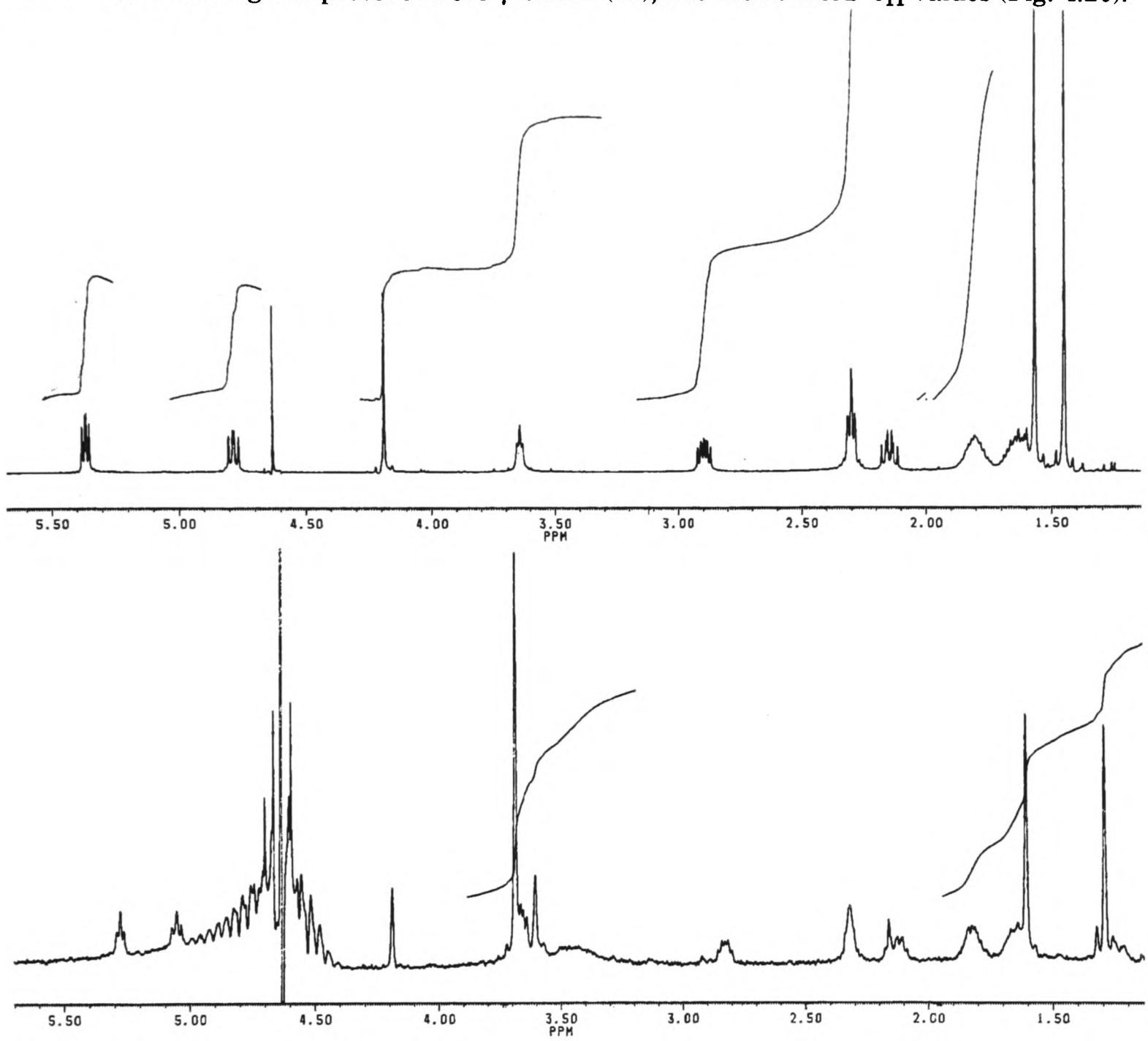
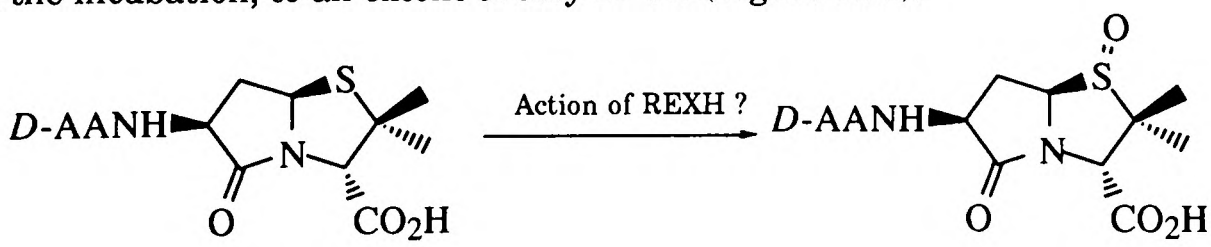



**Figure 4.20** NMR spectra for [top] the γ-lactam (85) and [bottom] a new incubation product

The greatest displacement of $\delta_H$ values occurred at one of the methyl groups (from 1.45 to 1.30 ppm) and the bridgehead proton (from 4.79 to 5.06 ppm), both of which are very near to the sulfur atom in terms of molecular geometry. The electrospray mass spectrum for this peak gave a molecular ion with 16 mass units heavier than the γ-lactam (85). These physical data led us to the tentative conclusion that (85) has been converted to one of the possible sulfoxides (87) during the incubation, to an extent of only 2 - 3% (Figure 4.21).




(85) (87)

**Figure 4.21** Production of a sulfoxide of γ-lactam (85) in incubation

This finding was interesting, as it would mean that REXH has further expanded its scope of oxidative power to include sulfoxidation, a realm that is yet to be encountered among the α-KG dependent dioxygenases. Within the larger family of α-keto acid coupled dioxygenases (Figure 2.2), however, there is an isolated precedent of an enzymic sulfoxidation catalysed by 4-hydroxyphenylpyruvate dioxygenase[141] (Figure 4.22), an intramolecular dioxygenase.

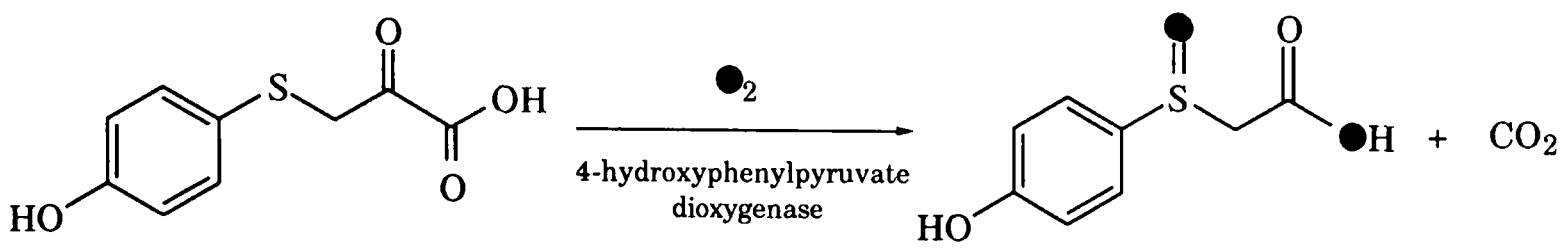
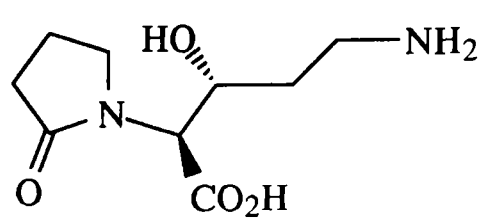


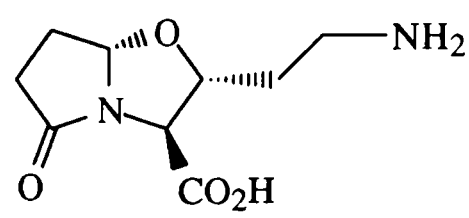

**Figure 4.22** The sole reported case of sulfoxidation by an α-keto acid coupled dioxygenase

The initial excitement was followed by an anti-climax when a similar level (2 - 3 %) of sulfoxidation of the γ-lactam (85) was observed in the control incubations, using either boiled or no enzyme. The sulfoxidation was thus a non-enzymatic process. Indeed, a literature search revealed that the system comprising ferrous ion and ascorbate (cofactors of REXH), in the presence of molecular oxygen, can hydroxylate aromatic rings at electronegative sites.[142, 143] The mechanism of such oxidations has not been studied. It is thus conceivable that the same mixture, with or without enzyme, can oxidise any electronegative centre such as sulfur in the γ-lactam (85). The stereochemical course of this sulfoxidation is a noteworthy point. The single set of signals in the NMR spectrum of the sulfoxide (87) indicated a very high degree of stereospecificity in this non-enzymatic reaction, though the configuration is not yet clear.

The possibility of (85) as an inhibitor of REXH was not investigated, so it was not clear whether or not the molecule is recognised by the enzyme active site. Considering the gross spatial changes caused by the γ-lactam ring, it is thought unlikely that any molecular recognition would take place. In this regard, CAS provides an interesting comparison, since it was found to accept (88) and (89), the γ-lactam analogues of proclavaminic acid (34) and dihydroclavaminic acid (35) respectively, as alternative substrates.[144] In this case, molecular recognition is still achieved, possibly owing to the much less severe spatial distortion arising from the lack of an *N*-acyl side chain on the γ-lactam ring.




(88)




(89)

Despite the above results, it is speculated here that REXH is potentially capable of sulfoxidation if the right molecular attributes are endowed on a penicillin N analogue to retain accessibility to the active site, yet preventing the radical abstraction step at the 2β-methyl group. The untested analogue, 2β-desmethylpenicillin N (55) could be just such a candidate.

## 4.5 Experimental

### 4.5.1 General techniques

(a) NMR

Proton NMR spectra were recorded on a Bruker AM 500 spectrometer at 500.13 MHz in $D_2O$ in a 5 mm NMR tube. Chemical shifts ($\delta_H$) are quoted in parts per million (ppm) and referenced to the sodium salt of 2,2,3,3-$d_4$-(trimethylsilyl)propionic acid (TSP), which was defined as having a chemical shift of 0 ppm exactly. Coupling constants ($J$) are quoted to the nearest 0.5 Hz. Spectra were obtained with suppression of the residual HOD signal using 22 to 27 L decoupling power. Samples containing less than 0.5 μmole of the desired component were routinely accumulated with 0.5 Hz line broadening to enhance sensitivity. $^1H$-$^1H$ COSY spectra were performed on the same machine by Mrs Elizabeth McGuinness of the Dyson Perrins Laboratory.

(b) Mass Spectrometry

The mass spectrometry service was provided by Dr Robin T. Aplin of the Dyson Perrins Laboratory. The analyses employed a Bio-Q electrospray mass spectrometer (VG Instruments) fitted with a quadrupole mass analyser with typical potential differences of 4 kV for the capillary, 1 kV for the electrode, 50 V for the skimmer and 250 V for the focusing lens. The freeze-dried samples were dissolved to *ca.* 20 μM in 1:1 methanol:water containing 1% acetic acid. Routinely, 50 μl of the sample solution was analysed at flow rate of 5 μl/min.

(c) HPLC

The HPLC system utilised was configured with the following components: a Waters 600E system controller, a Waters 600 multisolvent delivery system, a Waters 490E programmable multiwavelength detector, a BBC SE 120 chart recorder and a Rheodyne injector fitted with a 200 μl injection loop. Solvents were thoroughly degassed before use and sparged with helium at 15 ml/min during HPLC runs. Samples were clarified through a 0.22 μ membrane filter prior to injection with a Hamilton HPLC syringe. The HPLC columns were packed with Hypersil ODS resin, and fitted with guard columns containing Corasil packing. The column dimensions were: 250 x 4.6 mm (analytical), 250 x 7 mm (semi-preparative) and 250 x 10 mm (preparative).

(d) Others

Measurements of pH were made at ambient temperature using a KCl electrode and a digital pH-mV-temperature meter from Kent Electronics Instruments, calibrated with standards from BDH Laboratory Supplies. Acidification was achieved with 1 or 5 M HCl, while basification was done with 1 or 10 M NaOH.

UV-visible spectra were obtained with a Philips Pye Unicam PU 8800 UV-Visible Spectrophotometer. Enzyme incubations were conducted in a G24 Environmental Incubator Shaker (New Brunswick , USA). Samples were freeze-dried on an Edwards Micro Modulyo bench-top model. Procedures for the quantitative bioassay of carba-DAOC and carba-DAC were similar to those described in Section 3.2.7.5.

## 4.5.2 Incubation procedures

Enzymes used in incubations belonged to the purest fractions obtained according to purification procedures described in Chapter 3. Where ion exchange chromatography was the last purification step, desalting was carried out using PD-10 columns from Pharmacia. The various incubation protocols employed in the project are listed below:

(a) Standard protocol

Reaction mixtures contained α-KG (5 mM), ascorbate (1 mM), DTT (1 mM), ferrous sulfate (50 μM), ammonium sulfate (50 mM), substrate analogue (*ca.* 5 mM) and enzyme (*ca.* 1 to 30 mU) in 50 mM Tris-HCl, pH 7.5, in a total volume of 1 ml. A 10x-strength cofactor/cosubstrate solution was first prepared in 50 mM Tris and its pH readjusted to 7.5 with 10 M NaOH. 0.1 ml of this was then added to 0.8 ml enzyme fraction. The reaction was initiated by the addition of 0.1 ml of substrate analogue. Mixtures were incubated at 28°C in an orbital shaker operating at 200 rpm. After 2 hr, the reaction was quenched with 2 ml of acetone and centrifuged at 25 000 x *g* for 30 min to remove precipitated protein. Acetone was then removed by rotary evaporation at RT. The remaining aqueous portion was freeze-dried for NMR examination. The extent of turnover of the substrate analogue was determined by integration of the β-lactam ring proton signals in the NMR spectrum, recorded in $D_2O$.

(b) Without ammonium sulfate

Incubations in the absence of ammonium sulfate had been performed on DAC (26) as the potential substrate. In these experiments, ammonium sulfate was simply omitted from the cofactor/cosubstrate solution, with all other conditions identical to (a) above.

(c) Without ammonium sulfate, in phosphate buffer

Incubations in the absence of ammonium sulfate, and with Tris buffer replaced by phosphate buffer, had been performed on DAC (26) as the potential substrate. In these cases, the enzyme fractions were buffer-exchanged into 50 mM sodium phosphate, pH 7.5, using PD-10 columns from Pharmacia. The cofactor/cosubstrate solution, excluding ammonium sulfate, was prepared in the same phosphate buffer, and its pH readjusted to 7.5 with 10 mM NaOH. All other incubation parameters were identical to (a) above.

(d) Under $^{18}O_2$ atmosphere

Buffers used for the exchange on PD-10 columns and for the preparation of

cofactor/cosubstrate solution were thoroughly degassed *in vacuo* before use. Incubations were carried out in a sealed 100-ml bulb-shaped container of 97.7 % labelled oxygen, purchased from MSD Isotopes, Canada. The outer chamber was repeatedly filled with argon and evacuated several times to remove $^{16}O_2$. After rupture of the breakseal using a metal cylinder, the components of the incubation mixture were introduced into the $^{18}O_2$ atmosphere using an air-tight glass syringe and long needle (in the order cofactors, enzyme and substrate). Incubations were carried out at 28°C and 100 rpm. After 2 hr, the reaction was quenched by injecting 2 ml of acetone and shaked for 2 min. Centrifugation for protein sedimentation, followed by rotary evaporation to remove acetone were performed as described above.

### 4.5.3 Determination of Michaelis-Menten parameters

(a) For hydroxylation of carba-DAOC (75)

The refolded recombinant enzyme used belonged to the purest Mono Q fractions as obtained in Chapter 2. The range of concentrations used for carba-DAOC (75) was: 1.0 mM, 0.50 mM, 0.33 mM, 0.25 mM and 0.20 mM. All concentration points were done in triplicate.

Components of the reaction were added in the following sequence: (i) 120 µl of enzyme solution, (ii) 680 µl of Tris buffer, (iii) 100 µl of the 10x-strength cofactor/cosubstrate solution and (iv) 100 µl of a 10x-solution of the substrate. The reaction was incubated at 28°C and 200 rpm for exactly 10 min, after which it was quenched with 1 ml of cold methanol. The precipitated protein was removed by centrifugation and the supernatant used for HPLC analysis.

HPLC was performed on a Partisil 5 PAC column, with the mobile phase (2% acetic acid, 5 % acetonitrile, pH adjusted to 4.0 with $NH_4OH$) flowing at 1.5 ml/min. The eluate was monitored at 260 nm. Under these conditions, carba-DAOC and carba-DAC eluted at 8.0 min and 5.2 min respectively. Product formation was quantified by peak integration referenced to a calibration curve made for carba-DAOC. A direct linear plot[145, 146] was first obtained for all the data points to identify the bad points. The double reciprocal plot was then constructed with the acceptable points to obtain values for $K_m$ and $V_{max}$.

(b) For oxidation of DAC (26)

The soluble recombinant enzyme used belonged to the purest Superdex 200 fractions as obtained in Chapter 2. The range of concentrations used for DAC (26) was: 0.2 mM, 0.4 mM, 0.8 mM, 1.6 mM and 3.2 mM. All concentration points were done in triplicate.

Components of the reaction were added in the following sequence: (i) 100 µl of enzyme solution, (ii) 300 µl of Tris buffer, (iii) 50 µl of the 10x-strength cofactor/cosubstrate solution and (iv) 50 µl of a 10x-solution of the substrate. The reaction was incubated at 28°C and 200 rpm for exactly 10 min, after which it was quenched with 0.5 ml of cold methanol. The precipitated protein was removed by centrifugation and the supernatant used for HPLC analysis. HPLC analysis was identical to (a) above, but velocity of reaction was monitored by DAC depletion.

# CHAPTER 5

# STRUCTURAL STUDIES ON REXH

## 5.1 State of structural understanding on REXH

As a group, the α-KG dependent dioxygenases still possesses much to be explored in terms of their detailed structural features. For instance, there is as yet no report of an X-ray crystal structure for any of its members. Due to the lack of sequence homology between the primary structures of the α-KG enzymes and those of the 3-dimensionally defined families of proteins,[56, 57] it is difficult to make predictions about the structure of REXH at both the secondary and tertiary levels. Thus, even prevailing concepts about the ferrous active site of REXH (see Chapters 2 and 4) are not based on hard experimental data, but are inferred from two indirect means. The first relied on the analogy between the α-KG dependent dioxygenases and the more extensively studied class of cytochrome P-450 enzymes.[93, 96] The second relates to the primary sequence comparison among several α-KG dependent and related dioxygenases,[56, 57] which revealed two conserved regions each containing a histidyl residue that might be involved in ferrous ion coordination. In particular, the active site of IPNS (an enzyme which is structurally, mechanistically, functionally and probably evolutionarily related to REXH) has been probed by chemical techniques (photoaffinity labelling[147]) and physical methods like Mössbauer,[129, 130] EPR,[129, 130] electronic absorption,[129] NMR[131, 148] and EXAFS[149] spectroscopies. None of the above spectroscopic studies have been performed on REXH, nor any other α-KG enzyme. Indeed, structural probing of this class of enzymes have so far been limited to chemical approaches including the use of cosubstrate analogues[150] and affinity labelling.[57] It is therefore clear that a lot of scope exists for structural investigations into REXH, employing both chemical and physical methods. In this project, experiments were undertaken to lay the groundwork for more extensive work in this direction.

## 5.2 Results and Discussion

### 5.2.1 Electrospray mass spectrometry (ESMS)

Gel electrophoresis and size-exclusion chromatography have been the most extensively employed techniques in the MW determination of proteins.[151, 152] However, the margins of error for these two methods are large (~ 5 to 10%) due to substantial practical deviations from the assumptions made in their theoretical analyses. Since its inception in 1988,[153] the technique of electrospray mass spectrometry (ESMS) has been gaining popularity as the most accurate means of measuring the molecular weight of proteins, currently up to a limit of 130 kDa.[151] It has proven its value in various protein investigations including enzyme mechanistic studies,[13, 88] terminal processing analysis[154], chemical modification of proteins[155] and the study of

noncovalent protein-protein and protein-ligand interactions.[156, 157]

The principle underlying the application of ESMS for MW analysis of proteins is straightforward. The protein molecules are usually protonated by acidification with a weak organic acid, say 1% formic acid, in 50% aqueous methanol. The solution is sprayed into a dry atmosphere, whereby the solvent evaporates and the positively charged protein molecules are forced apart by electrostatic repulsion. The nebula of the protonated protein molecules so formed is accelerated towards the quadrupole mass analyser, where separation is achieved according to mass over charge (m/z) ratio. The 'softness' of the ionisation process results in little or no fragmentation of the intact protein. The extent of protonation of the basic residues on a single protein follows a bell-shaped distribution, such that several m/z species are present for a pure protein. The resulting mass spectrum consists of a set of peaks for the multiple charged ions, centred on the most abundant charged state, as shown in Figures 5.1 and 5.3 for the soluble recombinant and native forms of REXH respectively. The MW of the protein is determined by a computer algorithm. Figures 5.2 and 5.4 depict the transformed spectra after such an iteration, giving a more conventional looking spectrum. Samples containing two or more proteins will give spectra that exhibit the corresponding number of series of peaks. Table 5.1 summarises the ESMS results for REXH.

Enzymes used for ESMS experiments were purified as described in Chapter 3. During our work, it was found necessary to reduce the concentration of Tris buffer in the protein solution (from 50 mM to 5 mM) to prevent gradual 'clogging' of the spraying chamber. This was achieved by buffer-exchange using the PD-10 columns from Pharmacia.

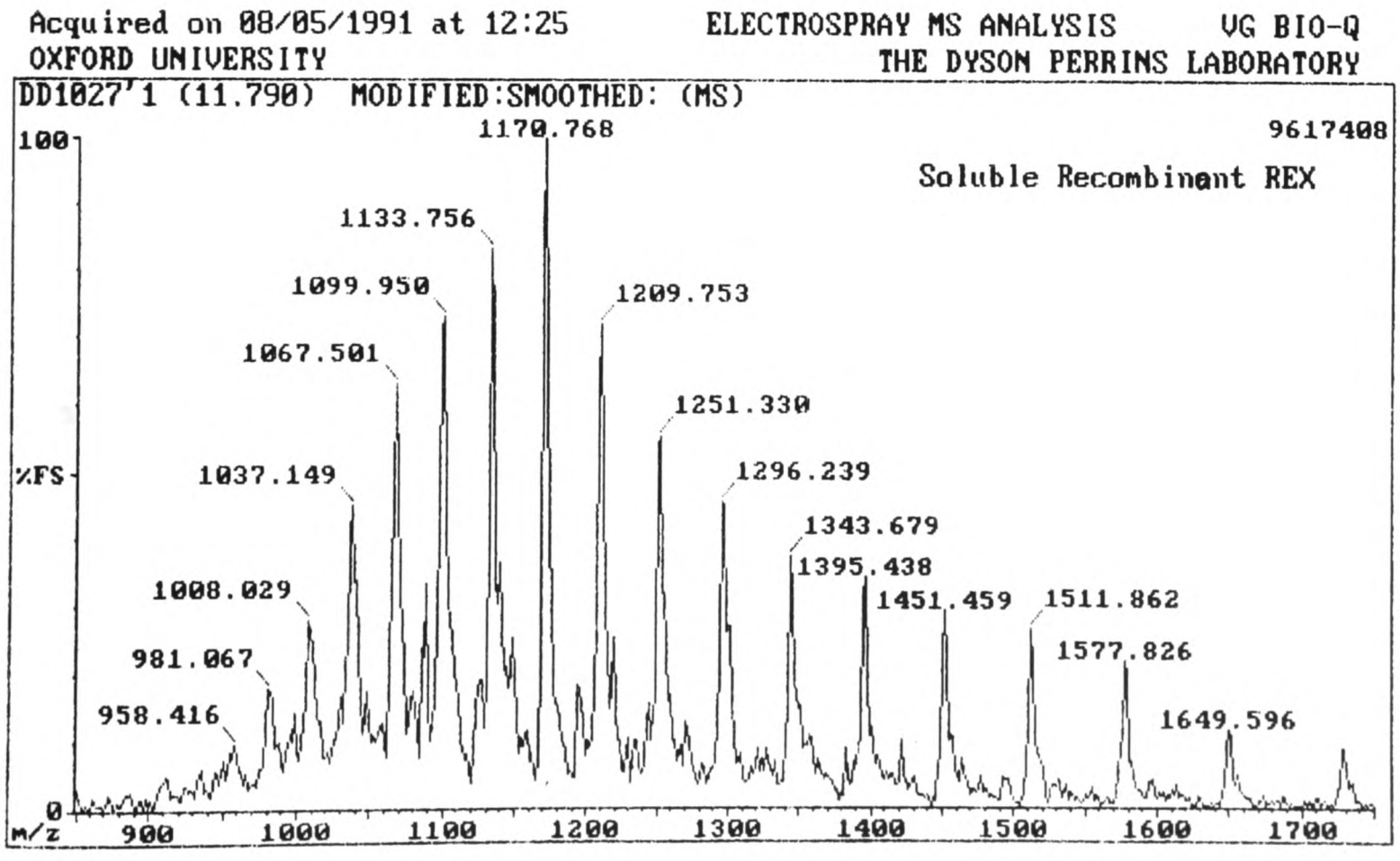



**Figure 5.1** Raw ESMS spectrum for soluble recombinant REXH

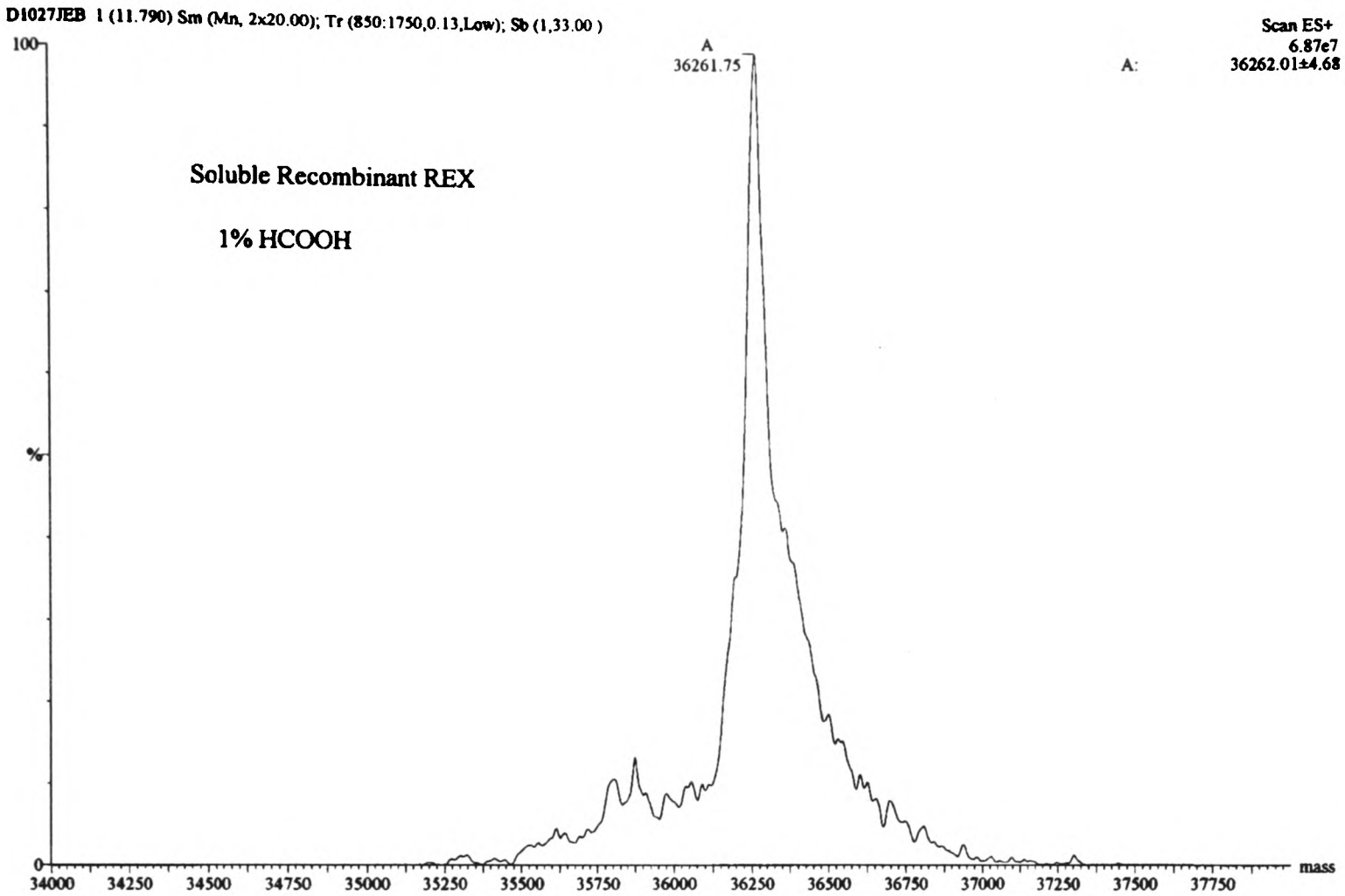



**Figure 5.2** Transformed ESMS spectrum for soluble recombinant REXH

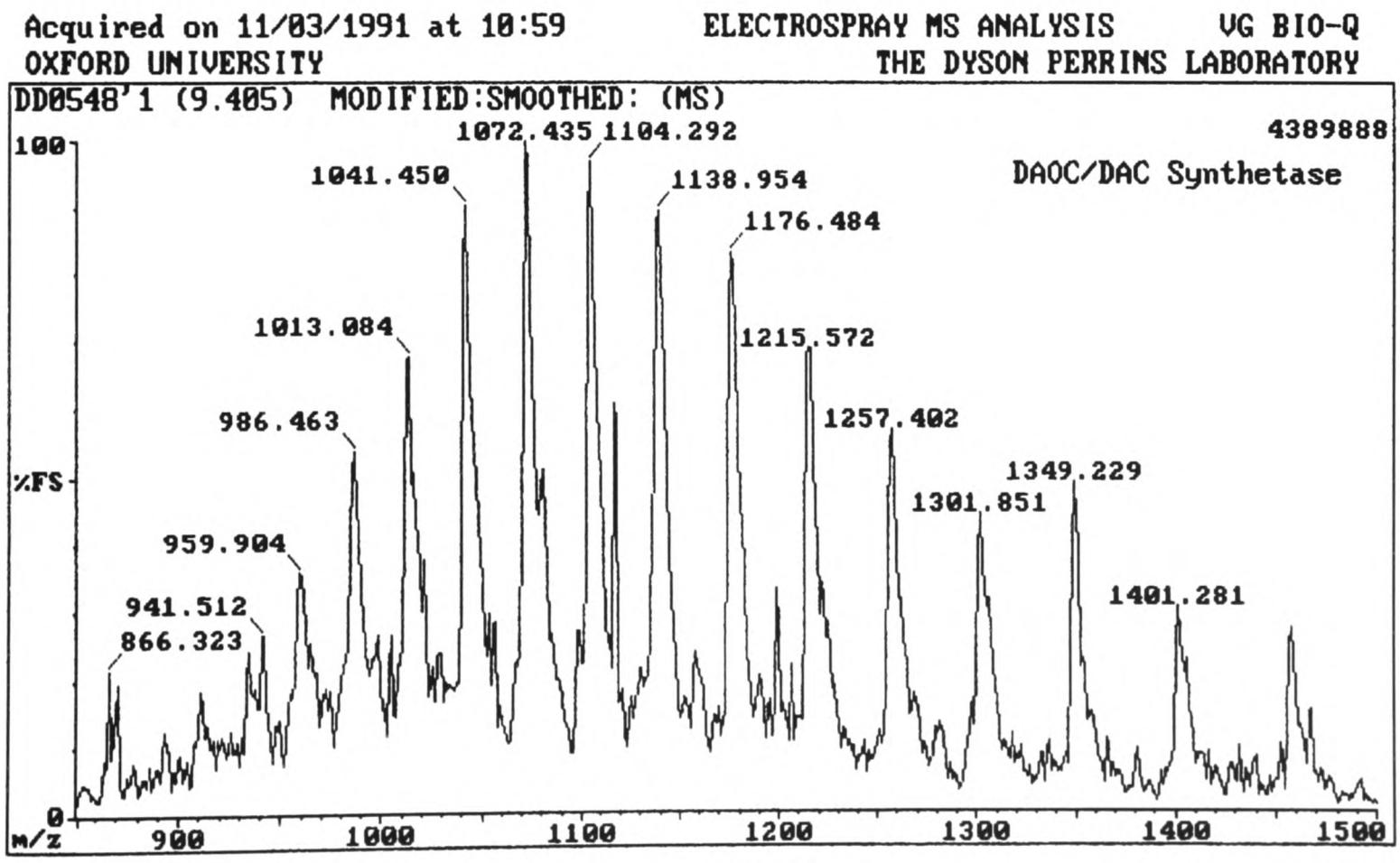



**Figure 5.3** Raw ESMS spectrum for native REXH

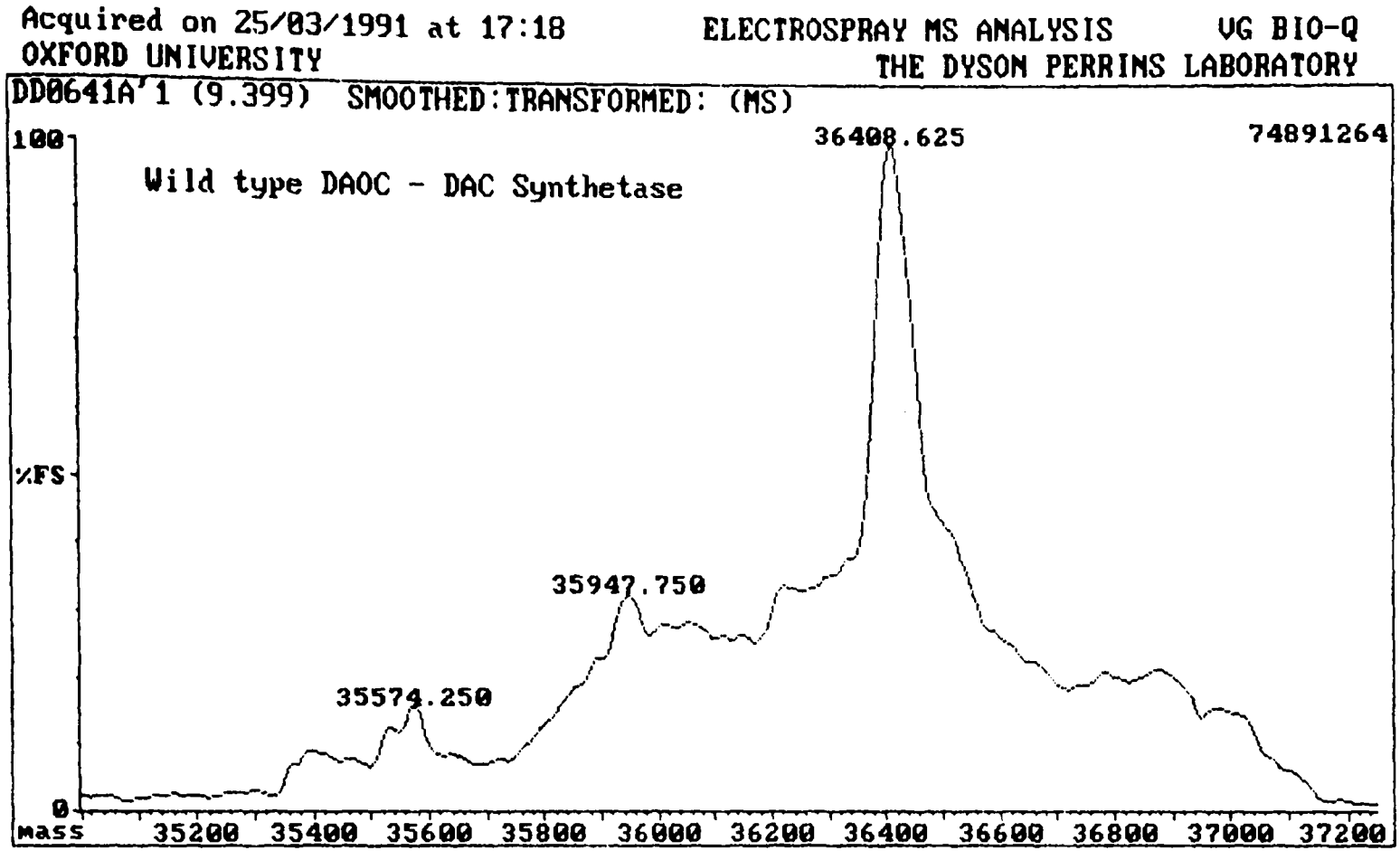



**Figure 5.4** Transformed ESMS spectrum for native REXH

| Enzyme | n | Observed MW | Sequence MW[a] | Difference[b] |
|---|---|---|---|---|
| Native | 3 | 36407.0 ± 5.1 | 36348.0 | 59.0 |
| Soluble recombinant | 3 | 36261.6 ± 9.9 | 36348.0 | 86.4 |
| Insoluble recombinant | 2 | 36357.0 ± 11.0 | 36348.0 | 9.0 |

**Legend**: n = number of determination, a = predicted MW based on ORF-derived amino acid sequence minus N-terminal methionine, b = difference between observed and sequence MW

**Table 5.1** Summary of ESMS data on the three different forms of REXH

The predicted MW of REXH based on the complete amino acid sequence deduced from its open reading frame[51] is 36479.0 Da. If the N-terminal methionine is excluded, the predicted MW would become 131 Da less, *i.e.* 36348.0 Da. Eukaryotic proteins usually lack the initiator methionine, which is removed by a specific methionine amino-terminal peptidase (MAP).[158-161] The observed value still leaves an excess mass of ~ 59 Da unexplained. The native enzyme has not been amenable to direct N-terminal sequencing by Edman methods, suggesting that the protein is blocked at the amino terminus.[26] The N-acetyl group usually observed when an initiator methionine is removed[162, 163] increases the MW by 42 Da. A methionine-truncated and N-terminally acetylated native REXH would still be 17 Da lighter than the observed value. Thus, the results obtained, while consistent with the presence of an N-terminal blockage, do not provide any definitive answer to the identity of the blocking group.

The insoluble recombinant REXH gave an observed MW which agrees with the predicted value within experimental error. It is therefore reasonable to suggest that this enzyme has no co- or post-translational modification apart from truncation at the N-terminal methionine. It is not surprising that *E. coli* does not possess the fungal cellular machinery used in co- or post-translational processing.

The soluble recombinant REXH gave results which are also not readily accounted for. N-terminal sequencing indicated that the methionine is processed off, but the observed MW is still 86.4 Da short of the sequence MW The loss of threonine from either the N- or C-terminus would result in a MW reduction of 101 Da, slightly more than the difference observed. It is difficult to infer anything with the available data.

If the speculation about insoluble recombinant REXH is correct, a logical corollary would be that REXH molecules do not bind any ferrous ion (a cofactor for catalysis, MW = 56) under the conditions of electrospray mass spectrometric determinations. This is not unexpected since the lower polarity (due to 50% methanol) and high acidity (due to 1% formic acid) of the analytical sample would have caused denaturation and consequent loss of iron-binding capability. To summarise, the results above demonstrate the potential utility of ESMS in analysing covalent modifications of proteins.

### 5.2.2 Peptide cleavage at Asp-Pro bond

As mentioned above, the N-terminus of native REXH has been consistently found to be refractory to Edman degradation by different groups,[26, 51] indicating that the blockage of the amino group may be due to intracellular processing rather than some artefacts arising from enzyme preparation. The exact nature of the N-terminal blockage was not determined, though the ESMS data above indicated that N-acetylation is not likely. As such, it would be useful to characterise more accurately and precisely the MW of the blocking group by mass spectrometry of a suitably-sized N-terminal fragment.

Thus, it is envisaged that a fruitful strategy might involve the application of ESMS in conjunction with traditional protein chemistry like selective proteolysis and peptide separation. The approach would then be to carry out repeated site-specific cleavages on successive N-terminal fragment, finally yielding an N-terminal peptide small enough to be accurately characterised by ESMS. Apparently, a compromise has to be made between the number of cleavage sites and ease of separation. Extensive cleavage at many sites produces smaller peptides, with the N-terminal fragment probably closer to the desired size for ESMS, but the subsequent isolation of a small N-terminal fragment could be difficult. Cleavage at a few sites simplifies purification because a smaller number of fragments is generated, though the larger N-terminal peptide formed probably requires further cleavage to achieve the desired size. The latter option seemed to be more attractive for REXH, after a detailed examination of the amino acid sequence of REXH (Figure 5.5) and the methods available for site-specific cleavage.[164]

```
MTSKVPVFRL  DDLKSGKVLT  ELAEAVTTKG  IFYLTESGLV
1           11          21          31
DDDHTSARET  CVDFFKNGSE  EEKRAVTLAD  RNARRGFSAL
41          51          61          71
EWESTAVVTE  TGKYSDYSTC  YSMGIGGNLF  PNRGFEDVWQ
81          91          101         111
DYFDRMYGAA  KDVARAVLNS  VGAPLAGEDI  DDFVEC***DP***LL
121         131         141         151
RLRYFPEVPE  DRVAEEEPLR  MGPHYDLSTI  TLVHQTACAN
161         171         181         191
GFVSLQCEVD  GEFVDLPTLP  GAMVVFCGAV  GTLATGGKVK
201         211         221         231
APKHRVKSPG  RDQRVGSSRT  SSVFFLRPKP  DFSFNVQQSR
241         251         261         271
EWGFNVRIPS  ERTTFREWLG  GNYVNMRRDK  PAAAEAAVPA
281         291         301         311
AAPVSTAAPI  AT
321         331
```

**Figure 5.5** Amino acid sequence of REXH deduced from cDNA[51]

Thus, the amino acid sequence of REXH (Figure 5.5) revealed the presence of a single Asp-Pro bond at position 157-158, which can be selectively cleaved by an acidic, non-enzymatic procedure in a chaotropic medium.[165, 166] The cleavage scheme was firstly tested on soluble recombinant REXH which can be purified in much larger amounts than native REXH. Thus, treatment of the 36.3 kDa soluble recombinant REXH with 7 M guanidinium hydrochloride in 30% (v/v) formic acid at 37°C for 5 days clearly produced two fragments, one band being very close to the 17 kDa MW marker and the other slightly larger (Figure 5.7). Indeed, a cleavage at the Asp-Pro bond of the sequence in Figure 5.5 would produce fragments of 17 and 19 kDa each. N-terminal sequencing of the two bands (Figure 5.6) confirmed the cleavage site to be at the lone Asp-Pro linkage. The progress of proteolysis, shown in Table 5.2, is followed by SDS-PAGE and densitometric scanning.

The mechanism of selective cleavage at Asp-Pro bond in acidic conditions has been discussed by Piszkiewicz.[167] It is proposed to involve intramolecular catalysis by aspartyl carboxylate anion displacement of the protonated nitrogen of the peptide bond. The properties of proline are clearly instrumental in enhancing the lability of the bond linking an aspartyl residue to a prolyl residue. Piszkiewicz suggested that in peptide linkages, there is a greater basicity of the prolyl nitrogen compared with those of other amino acid residues. This greater basicity would, by analogy with the hydrolysis of anilides,[168] enhance the rate of hydrolysis of the peptide bond by increasing protonation of the leaving group.

17 kDa band: TSKVPVFRLD.....

19 kDa band: PLLRLRYFPE.....

**Figure 5.6** N-terminal sequence for the two fragments of acid-cleaved REXH

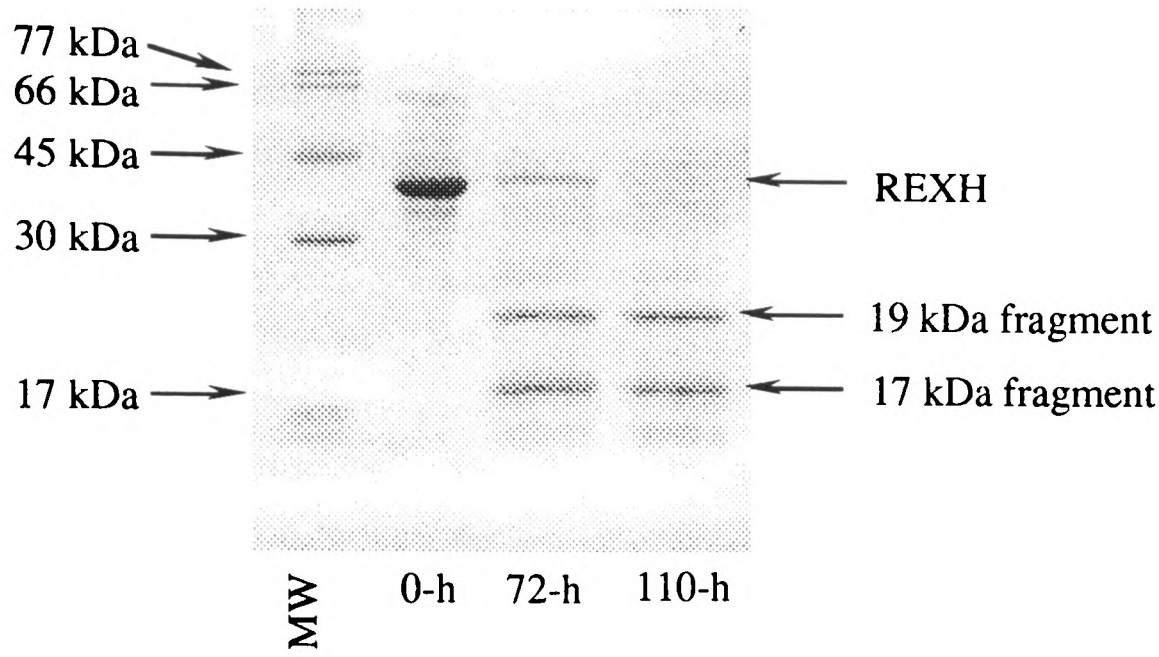



**Figure 5.7** SDS-PAGE gels showing progress of acid-catalysed specific cleavage of REXH

| Time (hours) | Extent of cleavage (%) |
|---|---|
| 0 | 0 |
| 72 | 80 |
| 96 | 90 |
| 120 | 95 |

**Table 5.2** Progress of acid-catalysed cleavage of REXH at Asp-Pro bond

Based on the sequence in Figure 5.5, the 17 kDa fragment would contain 29 acidic residues (Glu and Asp) and 17 basic residues (Arg and Lys) while the corresponding numbers are 18 and 21 respectively for the 19 kDa fragment. It seems feasible then to exploit the charge difference to separate the two fragments by ion-exchange chromatography if reversed-phase HPLC is unworkable. The 17 kDa fragment, which would bear the N-terminal block in native REXH, could thus be isolated and digested further to give smaller manageable fragments. Unfortunately, this line of research was discontinued before this point was reached owing to competing research priorities.

### 5.2.3 α-KG analogues as competitive inhibitors of REXH

The working hypothesis for the mechanism of REXH as outlined in Chapter 4 was modelled after the detailed stereochemical model proposed by Günzler[58] for the action of prolyl 4-hydroxylase. The latter model has presently become a prototype for mechanistic studies into other α-KG dependent dioxygenases. The Günzler mechanism emphasises the availability of a binding domain for α-KG within the ferrous active site of the enzyme. As depicted in Figure 4.1,

the ferrous ion is assumed to form a 5-membered chelate with the planar, bidentate oxo carboxylate segment of α-KG while a lipophilic spacer presents the distal carboxylate to an unspecified anion binding site on the enzyme.

It was believed that compounds which could mimic this binding behaviour of α-KG but could not undergo the catalytically essential oxidative decarboxylation would be inhibitors of prolyl 4-hydroxylase.[58] This idea led to the successful prediction, by Hanauske-Abel and Günzler, of the potent inhibitory properties of the pyridine dicarboxylates, the benzene dicarboxylates and the dihydroxybenzoates.[169-171] This was quickly followed by the discovery of similar patterns of inhibition for two related enzymes also involved in collagen biosynthesis, prolyl 3-hydroxylase and lysyl hydroxylase.[172] Kinetic studies revealed competitive inhibition with respect to α-KG, thus proving that these compounds do indeed replace α-KG as iron-chelators at the binding site of all three enzymes.

Subsequently, inhibition studies using similar α-KG analogues became quite a popular component in the characterization of other α-KG dependent dioxygenases. The salient results obtained with some α-KG dependent dioxygenases, listing only some of the strongest inhibitors and succinate, are summarised in Table 5.3.

| Enzyme | Ref. | $K_m$ for α-KG (μM) | $K_i$ (μM) | | | | |
|---|---|---|---|---|---|---|---|
| | | | (91) | (92) | (93) | (94) | (95) |
| Prolyl 4-hydroxylase | [169, 170] | 20 | 2 | 5 | 9 | 0.8 | 400 |
| Prolyl 3-hydroxylase | [172] | 5 | 3 | - | - | 15 | 800 |
| Lysyl hydroxylase | [172] | 100 | 50 | - | - | 150 | 10 000 |
| Flavanone 3β-hydroxylase | [64] | 20 | 1.2 | - | - | 40 | - |
| Flavone Synthase I | [65] | 16 | 1.8 | - | - | - | - |
| Hyoscyamine 6β-hydroxylase | [62] | 43 | 9 | 90 | - | - | - |
| γ-butyrobetaine hydroxylase | [150] | 18 | 0.2 | 0.6 | 13 | 14 | 70 |
| Proline 4-hydroxylase | [173] | 32 | 5.4 | 32 | - | 49 | 120 |

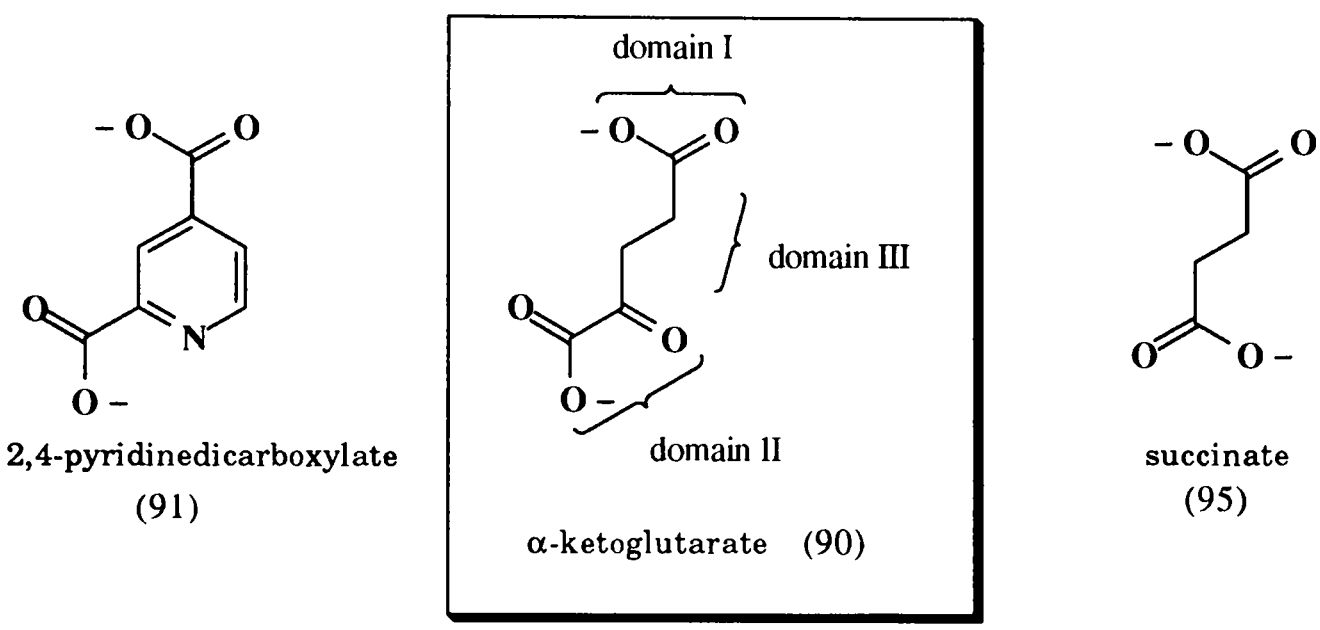
**Table 5.3** α-KG analogues as competitive inhibitors of α-KG dependent dioxygenases

It is noteworthy that compounds (91) to (94) were consistently found to be among the most potent inhibitors for all the enzymes studied. Even more remarkable are the following observations:

(1) For each enzyme, the $K_i$ for the strongest inhibitor is always lower than the $K_m$ for α-KG. This implies a higher enzyme affinity for the inhibitor than for α-KG, an indication of just how tightly the enzyme binds to the inhibitor.

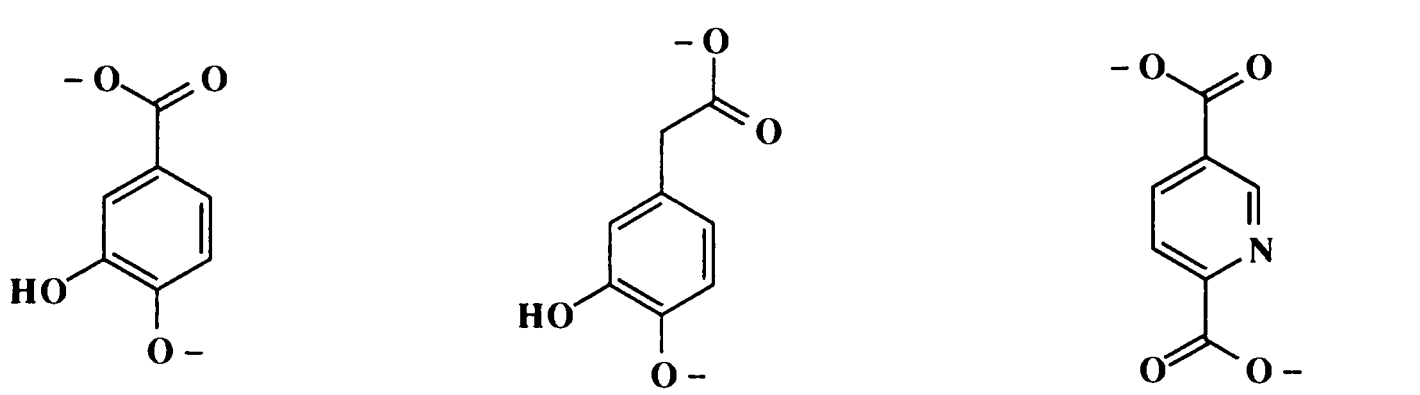
(2) With the single exception of prolyl 4-hydroxylase, the order of inhibition for the compounds listed is identical among all the enzymes studied, decreasing from (91) to (95) in terms of inhibitory activity. Prolyl 4-hydroxylase is apparently unique in being most strongly inhibited by (94).

In the course of this project, REXH was examined with several of the well studied α-KG analogues (Figure 5.8), and the results are given in Table 5.4.




2,4-pyridinedicarboxylate (91)

succinate (95)



3,4-dihydroxybenzoate (92)

3,4-dihydroxyphenylacetate (93)

2,5-pyridinedicarboxylate (94)

**Figure 5.8** Structural analogues of α-KG evaluated as inhibitors for REXH

| Compound | (91) | (92) | (93) | (94) | (95) |
|---|---|---|---|---|---|
| $IC_{50}$ (μM) | 1.2 | 3.6 | 21 | 220 | 600 |

**Table 5.4** α-KG analogues as competitive inhibitors of REXH

Although $IC_{50}$ values are not the same as $K_i$ values, their relationship can normally be simplified into a simple linear form,[174] such that the same interpretation used in comparing relative $K_i$ values can be applied to the comparison of relative $IC_{50}$ values. Thus, the order of

inhibition for REXH follows the main trend observed in Table 5.3. Collectively, all the inhibition studies carried out on α-KG dependent dioxygenases, including the results presented in Tables 5.3. and 5.4, point to the following conclusions:

(1) Such studies have contributed towards the understanding of the structure of this class of enzymes as these inhibitors, and other tested but non-inhibitory compounds, serve as reporter molecules in mapping the α-KG binding site. It is now widely accepted[150] that three subsites are present within the latter (Figure 5.8) : subsite I (a positively charged region) for anchoring the C-5 carboxyl group of α-KG, subsite II (*cis*-coordination site at iron) is used for chelation with C1-C2 segment of α-KG and subsite III (a hydrophobic region) serves to house a non-polar linker for domains I and II. Any compound capable of furnishing groups for recognition at these three subsites can potentially be inhibitory to the enzyme.

(2) All the enzymes studied, except prolyl 4-hydroxylase, share a very similar topology in their α-KG binding site. In particular, due to the highest potency exhibited by (91), α-KG is thought to bind in the staggered conformation as depicted in Figure 5.8. Prolyl 4-hydroxylase could be unique in having α-KG binding in a carboxylate-aligned manner akin to structure (94), its most potent inhibitor.

In summary, this work suggests that REXH possesses an α-KG binding site that is topologically similar to those in other α-KG dependent dioxygenases.

### 5.2.4 Iron binding to REXH

The preponderance of iron at the active centres of metalloproteins is perhaps a direct consequence of the highest abundance of this metal on planet Earth, as living systems evolve by maximising economy with the resources provided by Nature. Iron-containing proteins carry out a number of important functions in life, the most prominent of which include enzymic catalysis, dioxygen transport and electron transfer.[175] The iron-containing enzymes can be broadly classified into two groups based on whether the iron is present as part of a haem prosthetic group in the enzyme.

The α-KG dependent dioxygenases belong to the category of non-haem iron-dependent enzymes, which also include pteridine-dependent hydroxylases and catechol dioxygenases.[43] While iron seems to be an integral part of haem enzymes, such as cytochromes P-450,[93] there is apparently only a loose association between the metal and members of the α-KG dependent family.[66] The evidence described below for the latter enzyme class points to an equilibrium between enzyme-bound and free ferrous ions in solution.

(1) By all accounts, exogenous iron has to be added to highly purified enzyme preparations to bring about full activity.[60, 61, 176] The occasional low level of activity observed has been attributed to contaminating iron in the purification buffers.

(2) Strong iron-chelators such as EDTA, 2,2'-bipyridyl, *o*-phenanthroline and bathophenanthroline have been reported to inhibit some α-KG enzymes.[26, 62] These compounds presumably formed very strong complexes with iron in solution, thus depleting free iron in solution and hence reducing the amount of enzyme-bound iron.

(3) A number of divalent cations, including $Zn^{2+}$, $Mn^{2+}$, $Co^{2+}$, $Cu^{2+}$, $Ni^{2+}$, $Ca^{2+}$, $Mg^{2+}$, $Cd^{2+}$ and $Hg^{2+}$, have been reported to inhibit the enzymes competitively with respect to $Fe^{2+}$.[60-62, 64, 65] These alternative metal ions presumably displace iron from the active site but are not capable of the chemistry that results in dioxygen activation.

In view of the above, it is of interest to evaluate the tightness of iron binding to REXH. In particular, simple experiments would be conducted to assay the amount of iron bound to highly purified samples of REXH. There is a special significance about quantitating the amount of iron associated with REXH because of the bifunctionality of the enzyme. For instance, a result that show two equivalents of iron per mole of enzyme would suggest strongly the presence of separate active sites for the expandase and hydroxylase activities.

During the project, the Fish method[177] of iron determination was employed for the purpose of accurate quantitation of iron in purified samples of native and soluble recombinant REXH. The advantages of this method over other colorimetric assays are:

(1) The sensitivity of this method, owing to the high molar absorptivity (27 900) of the specific complex formed between the chelator and ferrous ion. This chelator, Ferrozine, suffers from little interference from other metals except copper (II), which is effectively removed by neocuproine, a potent copper (II) chelator.

(2) The novel digestion procedure for the protein, employing acid-permanganate treatment, proves to be highly reproducible and thorough, in contrast to the more conventional method of ashing.

(3) The method is simple to carry out.

The Fish assay indicated that there was 0.42 mole of iron bound per mole of native REXH and 0.44 mole of iron bound per mole of soluble recombinant REXH, both enzymes belonging to the purest fraction purified by methods outlined in Chapter 3. This implies that ferrous ions are not bound firmly to the protein molecules, a situation similar to the other α-KG dependent enzymes as described above. It is not known if the ferrous ions bind specifically at the active site or non-specifically on the protein surface. The loose association between the protein and metal ions is consistent with the results obtained from ESMS, where ferrous ion is deemed to be readily lost from denatured protein. Based on these results, it is impossible to infer anything about the number of active sites in REXH. Nevertheless, this experiment is a

further demonstration of the similar active site shared by REXH and others from its family, adding to the results obtained from the screening of α-KG analogues.

## 5.3 Experimental

### 5.3.1 General techniques

Amino acid sequencing was performed by Mr Antony Willis of the MRC Immunochemistry Unit (Department of Biochemistry, Oxford University) on the Applied Biosystems automated amino acid sequencer model 470A, fitted with an on-line model 120A for detection of phenylthiohydantoin-amino acids. Methods for HPLC, UV-visible spectrophotometry, SDS-PAGE, enzyme incubation, agar plate bioassay and freeze-drying were similar to those described in Chapters 3 and 4.

### 5.3.2 Electrospray mass spectrometry of proteins

The electrospray mass spectrometer used was a VG Bio-Q Quadrupole instrument, fitted with an atmospheric pressure ionisation source, operated by Dr Robin Aplin of the Dyson Perrins Laboratory. Electrophoretically homogeneous REXH, as obtained in Chapter 2, was buffer-exchanged into 5 mM Tris-HCl, pH 7.5 at RT, containing 1 mM DTT, using PD-10 (Pharmacia) gel filtration columns. The protein concentration was adjusted to 2 mg/ml (*ca.* 60 pmole/μl) either by dilution into excess buffer, or ultra-filtration using a PM-10 (Amicon) membrane. This was then mixed with methanol and formic acid (the proton source) to give a final concentration of *ca.* 30 pmole/μl enzyme, 50% (v/v) methanol and 1% (v/v) formic acid. The sample solution (20 μl) was injected into the electrospray system *via* a loop injector (Rheodyne 5717) at a typical flow rate of 5 μl/min (Applied Biosystems model 140A dual syringe pump). The mass spectrometer was scanned over a m/z range of 700 to 1500 Da/e. The instrument was calibrated with myoglobin (20 pmol/μl, MW = 16950.5 Da).

### 5.3.3 Peptide cleavage at Asp-Pro bond

Soluble recombinant REXH (0.5 mg, purified according to Chapter 3) was freeze-dried in a 1.5 ml eppendorf. 0.25 ml of cleavage solution [30% (v/v) formic acid in 7 M guanidinium-HCl] was added to the freeze-dried protein, followed by vortex stirring with warming at 60°C to promote the dissolution. The solubilised protein was left to stand at 37°C for 5 days, during which 10 μl aliquots were removed every 24 hr for SDS-PAGE analysis. For the electrophoretic analysis, the aliquot was diluted 10 times with Milli-Q water. Trichloroacetic acid (TCA) was added to a final concentration of 10% (v/v) and stirred on ice for 20 min. The TCA-precipitated protein was collected by centrifuging at 28000 x *g* for 10 min. The precipitate was then analysed on 15% polyacrylamide gels.

### 5.3.4 $IC_{50}$ determination for α-KG analogues as competitive inhibitors of soluble recombinant REXH

Each α-KG analogue under investigation was assessed at 5 different concentrations: 10 mM, 1 mM, 0.1 mM, 0.01 mM and 0 mM. Reaction mixtures contained the analogue (at each of the above concentrations), DTT (1 mM), ascorbate (1 mM), α-KG (1 mM), ferrous sulfate (50 μM), ammonium sulfate (50 mM), penicillin N (0.28 mM) and enzyme (*ca.* 0.3 mU) in 50 mM Tris-HCl, pH 7.5 at RT, in a total volume of 100 μl. Each concentration of analogue was done in triplicates. Mixtures were incubated at 28°C in an orbital shaker operating at 200 rpm. After 10 min, the reaction was quenched with 100 μl of cold methanol and the precipitated protein removed by microfuge centrifugation. The supernatant (100 μl) was introduced into a 8-mm well in an agar plate seeded with *E. coli* X580 and penicillinase. The plates were incubated at 37°C for 12 hr, after which the diameter of the inhibition zone was measured. The amount of cephalosporins (DAOC and DAC) was estimated by reference to a calibration graph obtained with standards of cephalosporin C. The reaction rate at each concentration of analogue was computed from the amount of product (cephalosporins) formed and the time of incubation. $IC_{50}$ values were determined from a plot of reciprocal of reaction rate versus logarithm of inhibitor concentration.

### 5.3.5 Iron assay

The method of iron assay for REXH was a modification of the Fish method.[177]

| Reagent A | Reagent B |
|---|---|
| 0.6 M HCl | 80 mg Ferrozine (Sigma) |
| 2.25% (w/v) aqueous $KMnO_4$ | 80 mg neocuproine (Sigma) |
| | 8.8 g ascorbic acid |
| | 9.7 g ammonium acetate, all dissolved in water to 25 ml |

Reagent A was prepared immediately before use by mixing equal volumes of 1.2 M HCl and 4.45% (w/v) aqueous $KMnO_4$. This reagent releases chlorine gas and must be prepared in a fume hood. Reagent B was prepared by first dissolving ammonium acetate and ascorbic acid in minimal water, then adding the Ferrozine and neocuproine, and finally filling up to the final volume of water. It could be stored at -20°C for up to 6 weeks in an amber bottle. For analysis, 0.5 ml of Reagent A was mixed with 1 ml of enzyme solution and incubated for 2 hr at 60°C, followed by cooling on ice. Reagent B (100 μl) was then added and the absorbance at 562 nm measured after 10 min. A calibration curve was obtained by parallel treatment of several standards of ferrous sulfate over a concentration range of 0 to 200 μM.

# CHAPTER 6

# INTRODUCTION TO CLAVAMS

## 6.1 Occurrence and biological activity

Since the first clinical use of penicillin about half a century ago, the family of β-lactam antibiotics has grown to include the numerous structural types described in Section 1.6. Among these structural types, the clavam nucleus (13) differs from the penam nucleus (7) simply in the substitution of sulfur with oxygen, its isoelectronic congener, at the 1-position. Because of this structural resemblance and their reported co-occurrence in several species of *Streptomyces* bacteria,[178] it was of interest to find out if the two bicyclic nuclei are formed by chemically analogous pathways. However, the clavams are different from the rest of the β-lactam family in two intriguing aspects.

(a) Occurrence

Whereas other classes of β-lactams, *viz.* penams, cephems, carbapenems and monocyclic β-lactams, have been found to be produced by more than one genus of micro-organism, the clavams have so far been isolated exclusively from *Streptomyces* bacteria (see Table 1.2). This, however, should not be taken to imply that *Streptomyces* species are the sole natural producers of clavams, as our present knowledge simply reflects the current state of effort involved in the screening programs. A more systematic approach is required to address the issue of whether the production of clavams is unique to the *Streptomyces* bacteria. To date, only fifteen naturally occurring clavams have been reported in the literature (Figure 6.1) and these are described below in a historical context.

The isolation and structural elucidation of clavulanic acid (31) from *S. clavuligerus*, reported by Beecham scientists in 1976,[179] represents the first example of the microbial production of the novel clavam nucleus. This was followed by the discovery, in Glaxo laboratories,[180] of three more clavams from culture fluids of the same organism. The three new clavams, whose structures as shown in Figure 6.1 have all been verified by total syntheses, were clavam-2-carboxylate (96),[181] 2-hydroxymethylclavam (97)[182, 183] and 2-formyloxymethyl-clavam (98).[183] Clavalanine (99), patented by Roche scientists in 1980[184] and synthesized in 1985[185], was the next clavam to be isolated from *S. clavuligerus*. The last two clavam metabolites detected in culture broths of this organism were clavaminic acid (36)[35] and dihydroclavaminic acid (35).[37, 38] These were shown, through collaborative efforts between Beecham scientists and our group at Oxford, to be mid-stage biosynthetic intermediates leading to clavulanic acid (31).

In collaboration with Ciba-Geigy, Prof. Hans Zähner at the University of Tübingen (Germany) identified two new clavams produced by *S. antibioticus* ssp *antibioticus* Tü 1718. The isolation of 2-(2-hydroxyethyl)clavam (100) and valclavam (37) were reported in 1981[186, 187]

and 1985[44, 188] respectively, with the total synthesis of (100) being reported in 1987.[189] As yet, there have been no report of a total synthesis or X-ray structure of valclavam (37).

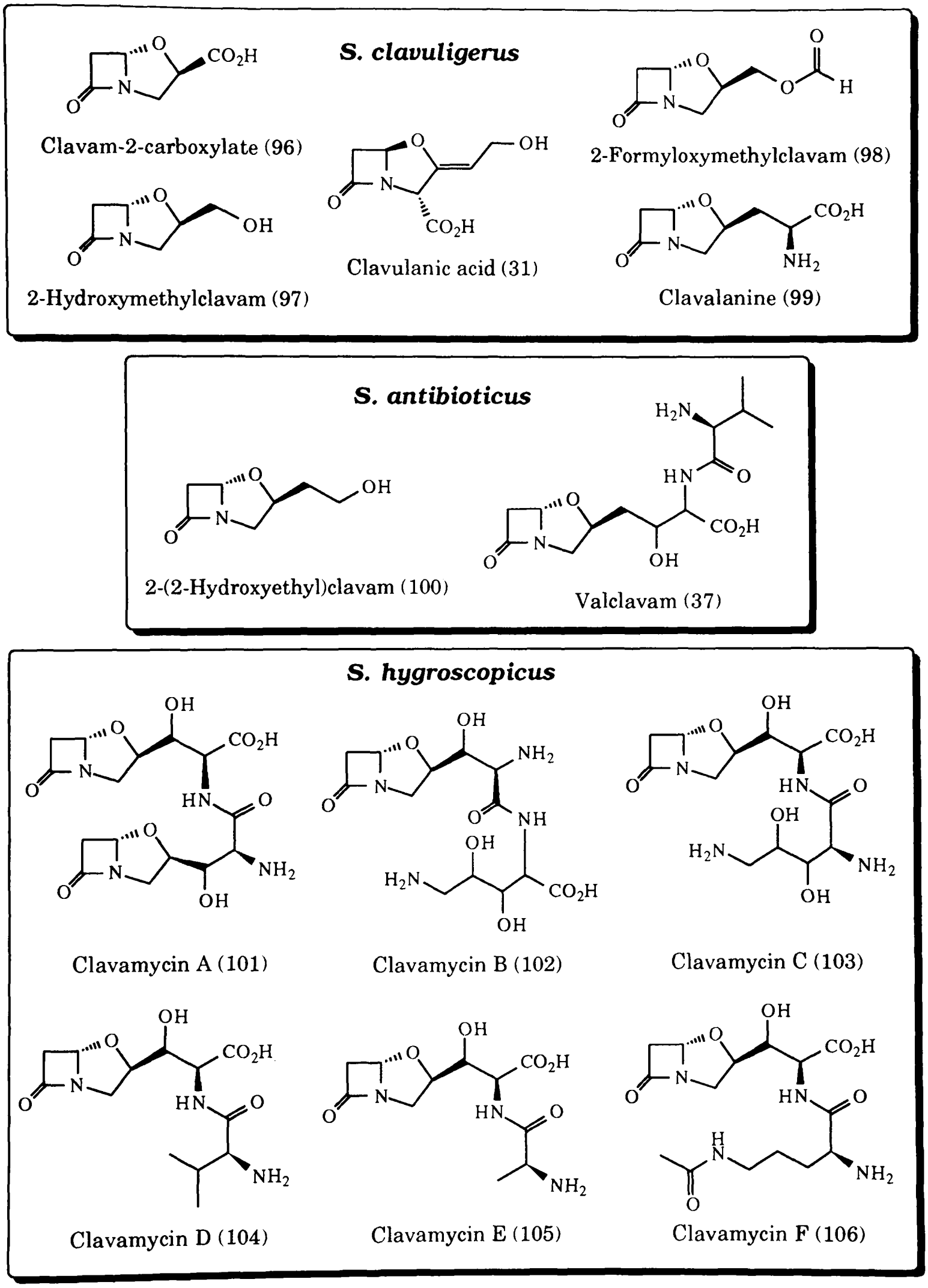



**Figure 6.1** Clavams from *Streptomyces* spp.

The last six clavams to be discovered, reported by Sandoz scientists in 1986,[190, 191] were extracted from fermentation broths of two strains of *S. hygroscopicus*. Clavamycins A, B and C were produced by *S. hygroscopicus* NRRL 15846 and clavamycins D, E and F were produced by *S. hygroscopicus* NRRL 15879. No synthesis of the clavamycins has been reported.

Following their initial discovery, some clavams were found to be produced by other *Streptomyces* spp. For instance, *S. jumonjinensis*,[192] *S. katsurahamanus*[193] and *S. lipmanii*[35] have all been shown to be producers of clavulanic acid (31). 2-Hydroxymethylclavam (97) was also isolated from *S. hygroscopicus* NRRL 15846.[190] Clavamycin A (101) was patented as clavam antibiotic CA-31 when isolated from *S. platensis*.[194] Clavamycins D (104) and E (105) were patented as clavam antibiotics CA-146A and CA-146B respectively when isolated from *S. lavendulae*.[195] From the above examples, it is conceivable that other species of *Streptomyces* not aforementioned are producers of some compounds in Figure 6.1 and also that more clavam structures remain to be discovered.

Just as clavaminic acid (36) and dihydroclavaminic acid (35) are biogenetic precursors of clavulanic acid (31), it is highly probable that some of the structures in Figure 6.1 are biological derivatives of one another, either anabolically or catabolically. There is also the possibility that some of these structures are artefacts arising from the isolation procedures. For instance, Clavamycin C (103) could have easily arisen from Clavamycin A (101) through the loss of the β-lactam segment of one clavam nucleus.

(b) Biological activity

The biological activity of the clavams differs from that of the penicillins and cephalosporins in two interesting ways: (i) its spectrum of susceptible organisms and (ii) its mode of action. Indeed, none of these clavams were discovered through the type of screening programs used to detect penicillins and cephalosporins. Instead, they were the 'hits' obtained from screening programs designed to identify antifungal agents, or other biologically active substances. Thus, clavulanic acid (31) was the product of extensive screening for novel β-lactamase inhibitors.[33] Although possessing weak antibacterial activity, it is a potent irreversible inhibitor of Class A β-lactamases[196] and is currently used synergistically with β-lactamase-sensitive penicillins in several antibiotic formulations (refer to entry 'clavulanate' in the British National Formulary). A number of synthetic β-lactamase inhibitors based on the clavam skeleton in clavulanic acid (31) were subsequently prepared by Beecham scientists.[183]

All the other clavams depicted in Figure 6.1 resulted from screening protocols that revealed their potent antifungal activity against a number of animal and plant pathogens. In addition, valclavam (37), clavalanine (99) and 2-(2-hydroxyethyl)clavam (100) were found to inhibit numerous Gram-positive and Gram-negative bacteria grown in a chemically defined medium.[44, 186, 197] The antibacterial activity of these compounds was completely reversed by methionine and its biosynthetic precursors cystathionine and homocysteine. Reversal of the activity of clavalanine (99) by *O*-succinylhomoserine but not by homoserine suggested that the enzyme homoserine *O*-succinyltransferase is inhibited (Figure 6.2).[197]

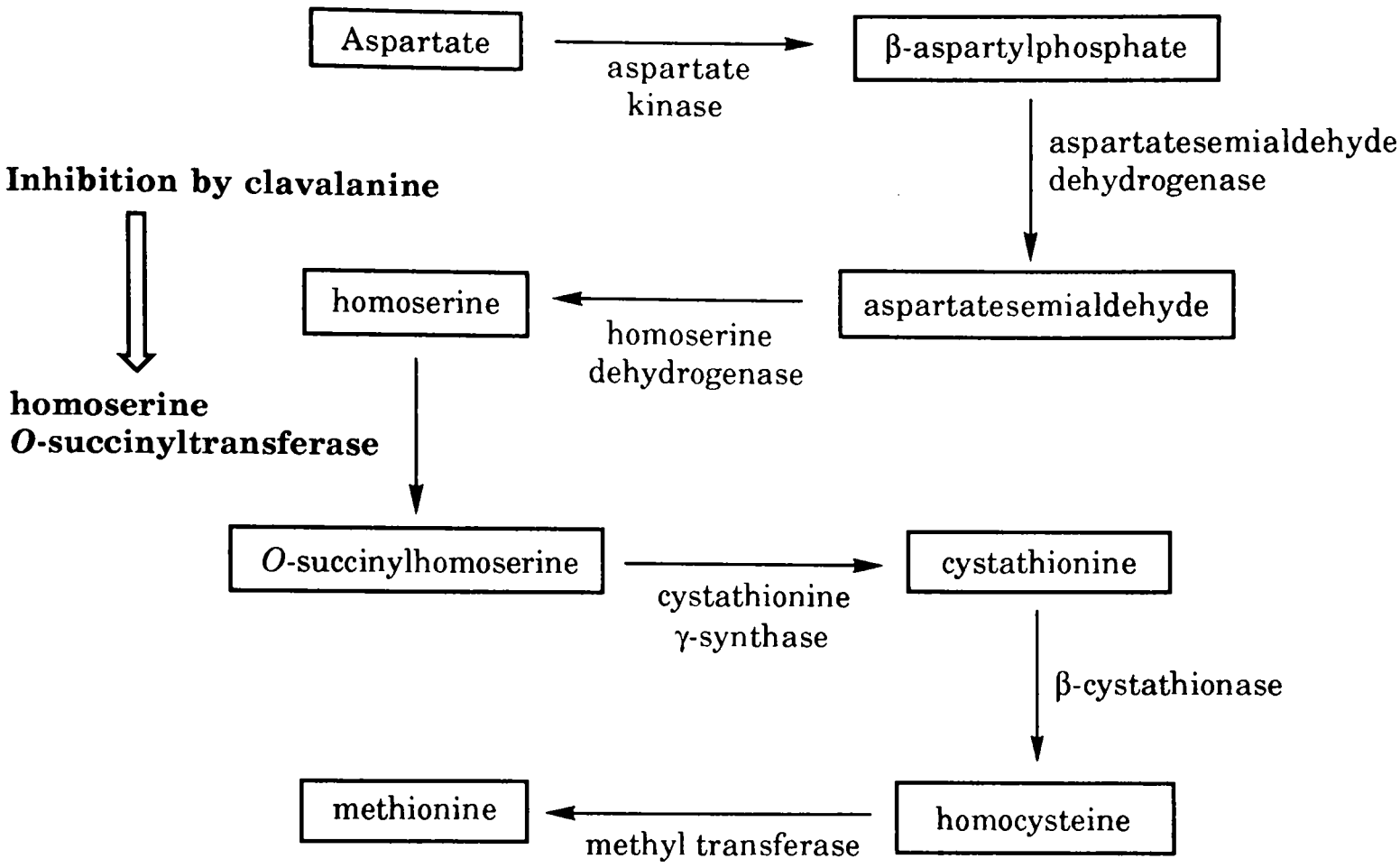



**Figure 6.2** Methionine biosynthesis in *E. coli*

Zähner *et al.*[198] subsequently demonstrated that the molecular basis for the bacteriostatic action against *E. coli* exhibited by valclavam (37) and 2-(2-hydroxyethyl)clavam (100) was a non-competitive inhibition of homoserine *O*-succinyltransferase, thus blocking methionine biosynthesis. The two substances interfered with the growth of eukaryotes such as *Saccharomyces cerevisiae* via the inhibition of RNA formation *in vivo*, although RNA polymerases are not inhibited *in vitro*. In stark contrast to the β-lactams in clinical usage, these clavams as well as clavalanine (99) are not recognised by penicillin-binding proteins and hence do not inhibit cell wall biosynthesis. They also did not inhibit β-lactamases, nor were they susceptible to the action of these enzymes.

The clavamycins share with the above three clavams their potent antifungal properties and lack of interaction with β-lactamases, but differ in having no antibacterial activity.[190] Thus, it would seem that the features common to all known naturally occurring clavams, except clavulanic acid (31), are (i) predominantly antifungal properties and (ii) non-interaction with β-lactamases.

The striking observation that clavulanic acid (31) and clinical β-lactams interact with penicillin-recognising enzymes (β-lactamases and *DD*-peptidases), whereas the other naturally occurring clavams do not, has been postulated to be a direct consequence of the stereochemical difference at the bridgehead carbon (C-5).[44, 186, 190, 197] Whilst clavulanic acid (31) shares the 5*R*-stereochemistry with clinically used β-lactams, all the other clavams in Figure 6.1 possess

the *S*-configuration at the corresponding chiral centre. Some workers go further to advocate that this fundamental structural difference is also responsible for the contrasting antimicrobial spectra (antibacterial vs antifungal) of the two groups of compounds.[185] In this regard, it would be interesting to speculate on the activity spectra for the 5-epimers of the antifungal clavams and also 5-epimers of the antibacterial clinical β-lactams. In conclusion, the intriguing biological activity of the 5*S*-clavams promises to add a new and enriching dimension to the vast field of β-lactam research.

## 6.2 Biosynthesis of clavulanic acid

Until now, the major efforts in elucidating the biosynthesis of clavams have been restricted to studies on clavulanic acid (31). The sharp focus on this compound can be attributed to its important clinical application as well as its pioneer status in the chronology of clavam discovery. The biosynthetic pathway to clavulanic acid (31), which has been mentioned briefly in Chapter 1, will be treated in greater detail here for two reasons. Firstly, the brisk pace of research in this field resulted in substantial new information which was not covered in Chapter 1. Secondly, being the only known clavam pathway, it could serve as a basis for formulating biosynthetic hypotheses for other clavams such as those studied in this project. All the presently known steps leading to clavulanic acid (31) are shown in Figure 6.3.

The penams and cephems are derived from a tripeptide precursor and incorporate non-amino acyl units (such as a carbamoyl group, an acetoxy group or a phenoxy group) only during the later stages of their biosynthetic pathways (Figure 1.1). In contrast, clavulanic acid (31) is apparently a mixed-precursor metabolite right from the first committed step of its biosynthesis. In retro-biosynthetic terms, the structure of clavulanic acid (31) can be decomposed into two segments: (i) a $C_3$ unit forming the carbon skeleton of the β-lactam ring and (ii) a $C_5$ unit incorporating the rest of the oxazolidine ring and the whole of the hydroxyethylidene side arm.

The $C_3$ unit has been shown, by whole-cell feeding experiments with *S. clavuligerus*, to be derived from any of several $C_3$ compounds which are interconvertable through primary metabolism. These compounds include glycerol,[199, 200] pyruvate,[201] *D*-glycerate,[201] *D*-lactate[202] and β-hydroxypropionate.[203] Due to the different parameters employed by different workers during the feeding experiments, it is difficult to make a rigorous comparison of their reported levels of incorporation in an attempt to pin-point the particular $C_3$ compound that is most directly channelled into the clavulanic acid (31) pathway. There is as yet no consensus on this issue. Incorporation of glycerol was shown to proceed with absolute stereospecificity,[204] the 5-H in clavulanic acid (31) being derived from the *pro*-(R) hydrogen in the *pro*-(R) hydroxymethylene group of glycerol (Figure 6.4)

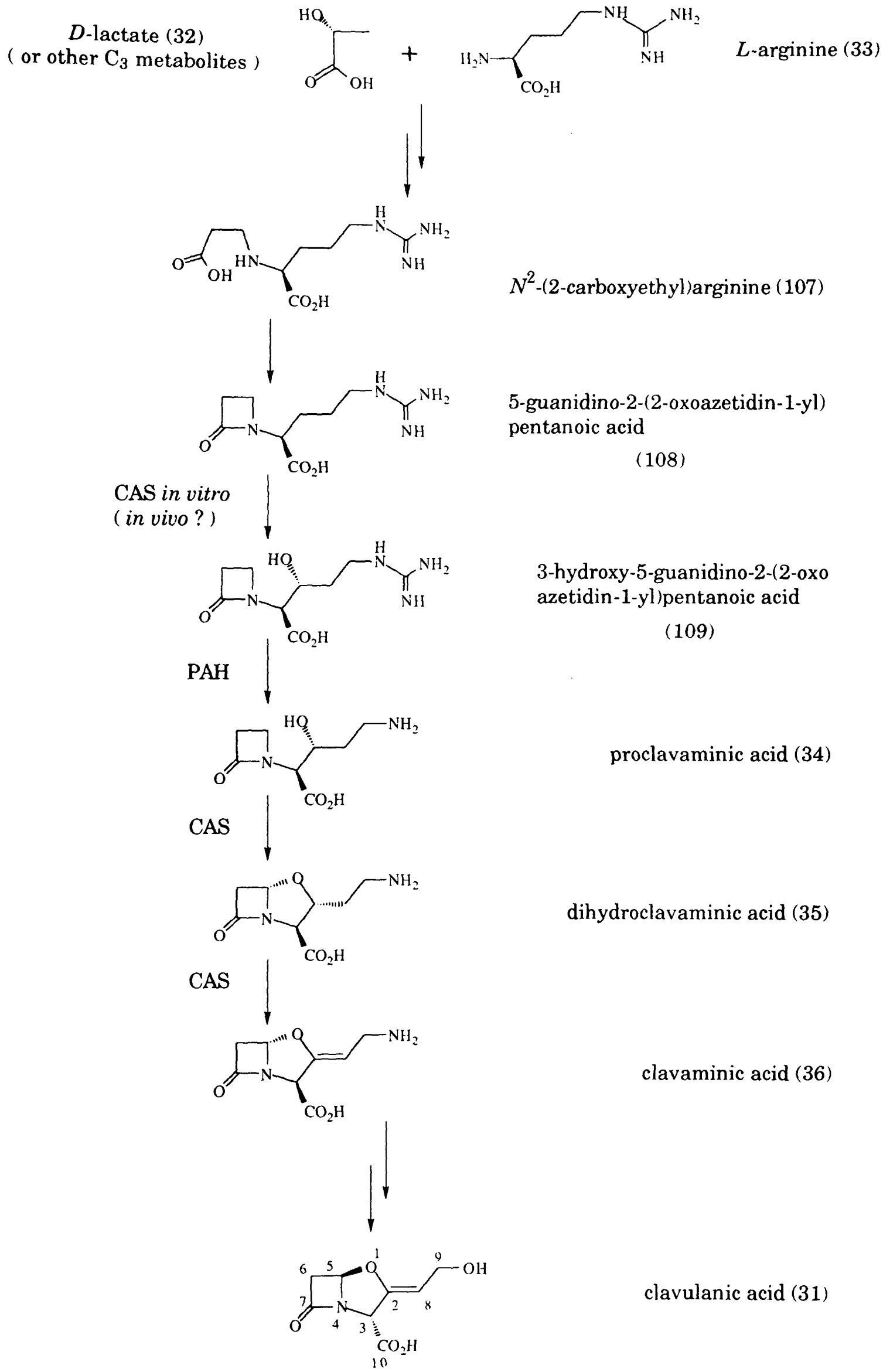



**Figure 6.3** Biosynthesis of clavulanic acid (31)

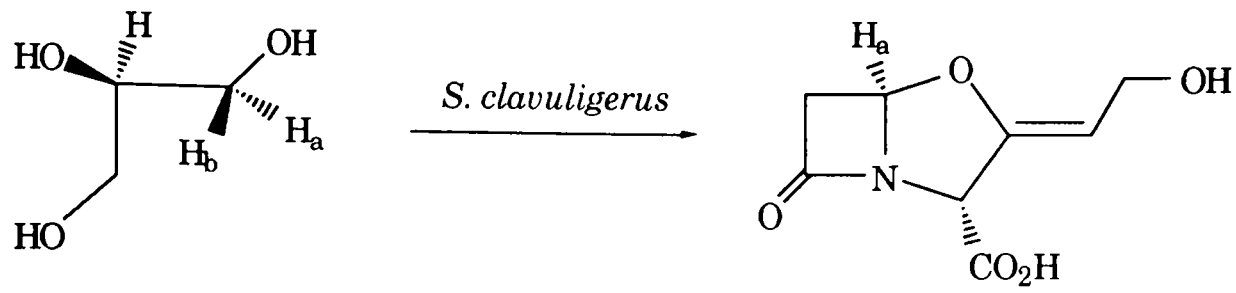



**Figure 6.4** Stereochemical outcome of glycerol incorporation into clavulanic acid (31)

Early feeding experiments with glutamate,[205] ornithine[206-208] and arginine,[207] involving wild-type *S. clavuligerus*, have shown all three compounds to be precursors for the $C_5$ unit of clavulanic acid (31). Again, the metabolic interconversions among these amino acids (Figure 6.5) and the variations in the design of the feeding experiments make it impossible to deduce which of these amino acids is utilised in the committed step to clavulanic acid (31).

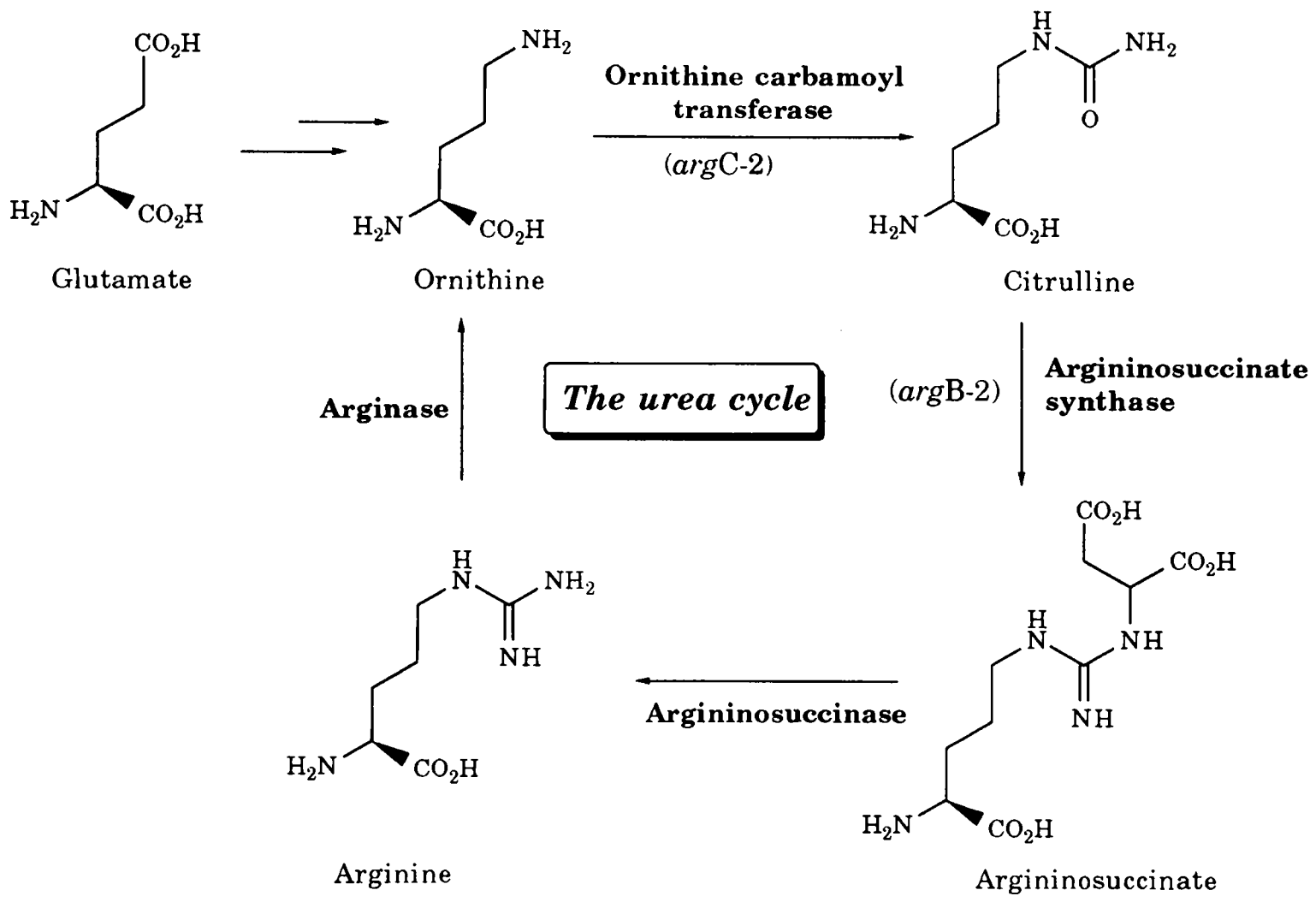



**Figure 6.5** Metabolic relationship among glutamate, ornithine and arginine

This long-standing ambiguity was settled by a very recent set of feeding experiments using mutants of *S. clavuligerus* auxotrophic for arginine.[34] The two mutants, designated *arg*C-2 and *arg*B-2, are deficient in the enzyme ornithine carbamoyl transferase and argininosuccinate synthase respectively (Figure 6.5). The observation that both mutants incorporate arginine but not ornithine into clavulanic acid (31) provided conclusive evidence for the former as the entry point to clavulanic acid (31) pathway. This conclusion was further verified by the discovery, described below, of early biosynthetic intermediates containing the arginino residue.

Thus, feeding experiments performed by Beecham scientists[209] demonstrated the

intermediacy of three arginine derivatives among the earliest stages of the clavulanic acid (31) pathway. The three compounds, in their postulated biosynthetic sequence, are $N^2$-(2-carboxyethyl)arginine (107), 5-guanidino-2-(2-oxoazetidin-1-yl)pentanoic acid (108) and 3-hydroxy-5-guanidino-2-(2-oxoazetidin-1-yl)pentanoic acid (109). The compound (107) has also been isolated from culture broths of *S. clavuligerus* dclH 65, a mutant blocked in clavulanic acid (31) production.[210] The formation of amide link as the ring closure step in β-lactam formation is conceptually the opposite of β-lactamase action. Though unprecedented in the other known β-lactam pathways[32], it is speculated to be involved in carbapenem biosynthesis.[211] This difference provides yet another distinctive feature in the biochemistry of clavams. The enzyme mediating this step has not been isolated yet.

The next step, the hydroxylation of (108) to (109), has been shown to be an *in vitro* activity of CAS[132] though more experiments are required to prove the same role for CAS *in vivo*. The Beecham group also succeeded in isolating a new enzyme,[209] proclavaminic acid amidino hydrolase (PAH), that converts (109) into proclavaminic acid (34). Only the second enzyme so far discovered in the pathway, PAH shows significant sequence homology[212] to arginase (of the urea cycle) which also catalyses the conversion of a guanidino to an amino group. Indeed, PAH was activated by $Mn^{2+}$ ions, a phenomenon reported for arginases from *Bacillus anthracis*[213] and staphylococcal sources.[213, 214]

The identification of the CAS-mediated transformation of proclavaminic acid (34) to clavaminic acid (36) was the first major breakthrough in biosynthetic studies on clavulanic acid (31).[35, 36] This reaction was subsequently shown to occur in two stages via the saturated clavam dihydroclavaminic acid (35),[37, 38, 215] indicating that an oxidative cyclisation preceded desaturation. CAS, an α-KG dependent dioxygenase, therefore resembles REXH in being capable of various oxidative modes of catalysis (hydroxylation, cyclisation and desaturation). Further to this similarity, CAS was found to exist as two isozymes in *S. clavuligerus*,[39, 216] just as the ring expandase and 3'-hydroxylase are separate enzymes in the same organism. Though the two isozymes were found to exhibit comparable kinetic parameters for both cyclisation and desaturation, they could represent the beginning of evolutionary specialisation for each enzymic activity. Both CAS isozymes have been isolated,[39] cloned[216] and overexpressed.[202] Mechanistic studies on CAS shall be described in Chapter 9.

The steps preceding $N^2$-(2-carboxyethyl)arginine (107) and those occurring after clavaminic acid (36) are not known in any detail. Townsend[217] has shown that the allylic hydroxy group in clavulanic acid (31) is derived from molecular oxygen, indicating the involvement of an oxygenase during the oxidative deamination from clavaminic to clavulanic acid (31). Townsend[218] further demonstrated that the 5-*pro*-(R) hydrogen in ornithine is transformed into the 9-*pro*-(S) hydrogen of clavulanic acid (31), with concomitant loss of the 5-*pro*-(S) hydrogen in ornithine (Figure 6.6). The simultaneous configurational inversion and proton substitution was interpreted as strong evidence for the intermediacy of an aldehyde species (110 in Figure 6.6) between clavaminic and clavulanic acids.

The biosynthetic genes for clavulanic acid (31) and cephamycin C (28) have been shown to occur together as a 'super-cluster' in three *Streptomyces* species, suggesting that genetic regulation of these two pathways are closely intertwined.[178]

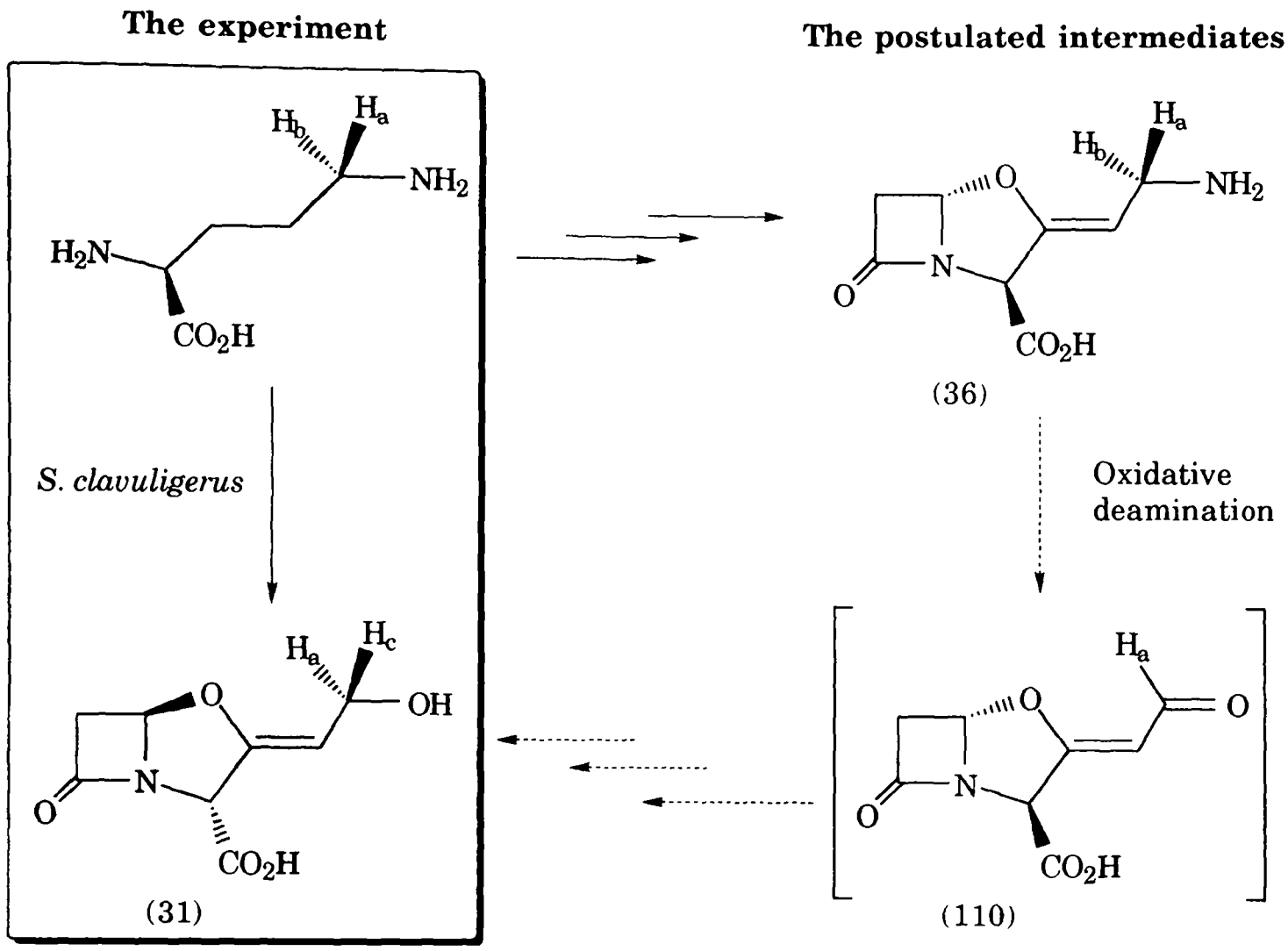



**Figure 6.6** Incorporation of stereospecifically labelled ornithine into clavulanic acid (31)

# CHAPTER 7

# FERMENTATION, BIOASSAY, ISOLATION & CHARACTERISATION OF VALCLAVAM

## 7.1 Introduction

The initiation of biosynthetic studies on valclavam (37) represents an extension of our group's long-standing interest[32] in the bioorganic chemistry of β-lactams. When the idea of biosynthetic studies into a second clavam was first conceived, the sole precedent being clavulanic acid (31), the strategy was to identify a compound (from Figure 6.1) whose structure suggested a plausible biosynthetic hypothesis that could be used for preliminary experimental design. Clavalanine (99), produced by *S. clavuligerus*, was the first to be selected since its structure can be retro-biosynthetically divided into a lactate-derived $C_3$ unit (the carbon skeleton of the β-lactam ring) and an ornithine-derived $C_5$ unit by analogy to the biosynthesis of clavulanic acid (31). More interestingly, this hypothesis requires the ornithine to be incorporated in the opposite regiochemical sense to that established for clavulanic acid (31), *i.e.* such that the β-lactam nitrogen in clavalanine (99) originates from the δ-amino but not the α-amino group of ornithine (Figure 7.1).

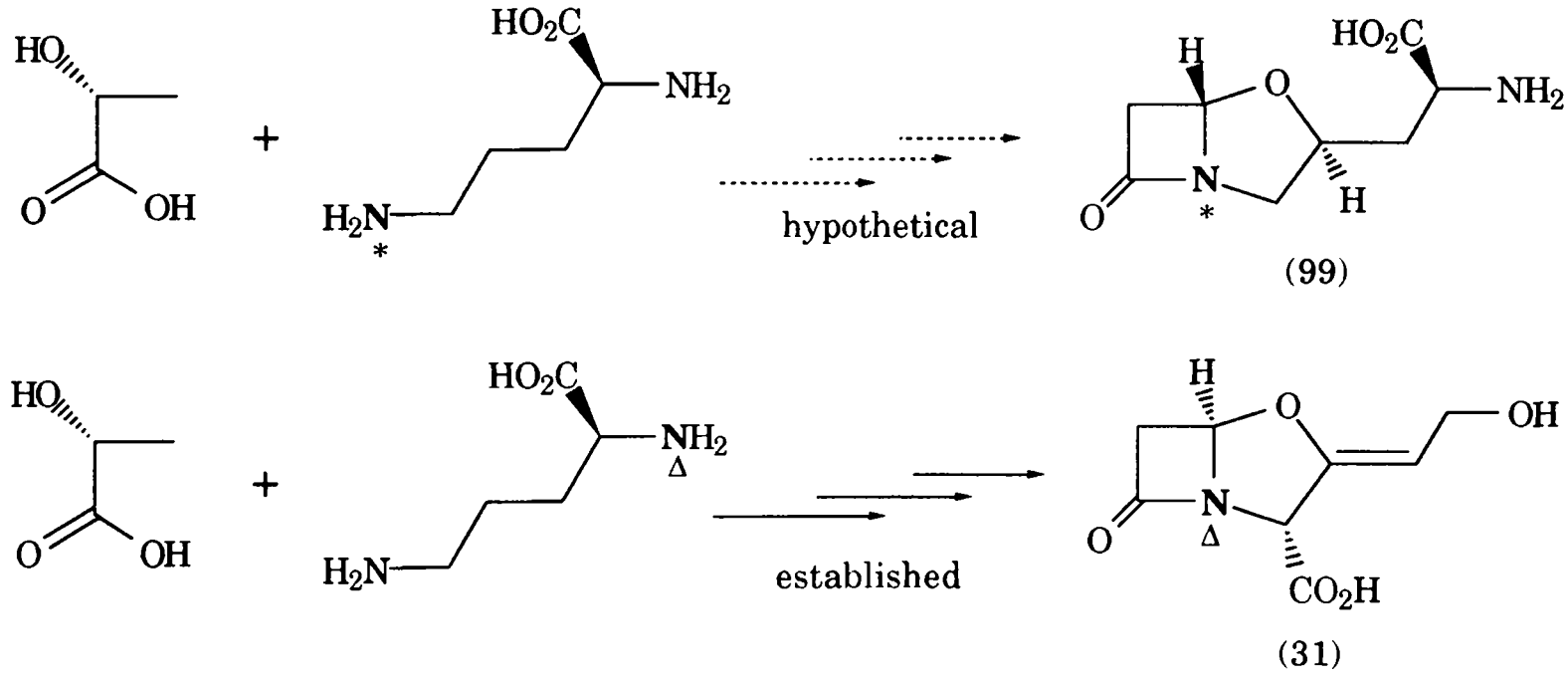



**Figure 7.1** Biogenetic hypothesis for clavalanine (99), in comparison with established route for clavulanic acid (31)

The clavalanine-producing strain of *S. clavuligerus* ATCC 27064, termed 1715B, was a gift from Hoffmann-La Roche. The decision to start our biosynthetic studies with clavalanine (99) was also prompted by the experience of our laboratory in handling other strains of this *Streptomyces* species, which were used in our enzymological studies[32] on CAS, ACVS, DAOCS and DACS.

At the commencement of work on clavalanine (99), we were also interested in other clavams that are structurally similar to this compound but are produced by other *Streptomyces* spp. In particular, we were looking for clavam metabolites that might be biosynthetically related to clavalanine (99), with a view to conducting comparative studies of the biosynthetic pathways or even switching the target of our studies if necessary. In the event, 2-(2-hydroxyethyl)clavam (100) was thought to be a suitable candidate. The producer organism, *S. antibioticus* ssp *antibioticus* Tü 1718, was given to us by Prof Dr Hans Zähner (University of Tübingen) who suggested[219] instead that valclavam (37) would be more amenable to our purposes by virtue of its higher titre from this organism and its greater stability compared to 2-(2-hydroxyethyl)clavam (100).

## 7.2 Results and Discussion

### 7.2.1 Development of bioassay for clavalanine (99)

One of the first priorities when initiating a biosynthetic study of a natural product is to devise a reliable means of monitoring its production from natural sources. Reliability in this context encompasses accuracy, precision, sensitivity and selectivity. The bacterium *Bacillus sp.* ATCC 27860 was the organism found to be most sensitive to clavalanine (99) under a wide-ranging screening program conducted by Roche.[197] A slope of this organism provided by Roche was examined for our bioassay development.

*S. clavuligerus* ATCC 27064 is known to be a prolific producer of β-lactam antibiotics. Penicillin N (24),[220] DAOC (25),[221] cephamycin C (27)[220], clavulanic acid (31)[179] and three other clavams[180] had all been isolated from fermentation broths prior to the discovery of clavalanine (99). A mutant of *S. clavuligerus*, designated IT1, has also been found to produce non-β-lactam antibiotics, *viz.* holomycin and a tunicamycin-related complex of antibiotics.[222] Roche scientists have demonstrated the frequent co-production of other β-lactam compounds with clavalanine (99).[197]

| Compound | Load (μg) | Kill zone (mm) |
|---|---|---|
| Clavalanine | 5 | 58 |
| Clavulanic acid | 5 | 0 |
| Penicillin N | 5 | 28 |
| DAOC | 40 | 27 |
| DAC | 40 | 32 |
| Cephamycin C | 40 | 19 |
| Holomycin | 50 | 20 |

**Table 7.1** Growth inhibition of *Bacillus* sp. ATCC 27860 by antibiotics from *S. clavuligerus*

A semi-quantitative comparison of the inhibitory activity of some of the antibiotics produced by *S. clavuligerus* (Table 7.1) showed that clavalanine (99) is most potent against *Bacillus* sp. ATCC 27860 - a desirable situation in the context of our work. Nevertheless, it would be ideal to devise a selective bioassay for clavalanine (99) with the ability to exclude the biological activity of all other antibiotics in the culture filtrate.

The activity of clavalanine (99) against *Bacillus* has been shown to be bacteriostatic and reversible by *L*-methionine and several of its biosynthetic precursors.[197] It is most probably an inhibitor of homoserine *O*-succinyltransferase, and therefore of methionine biosynthesis, as had been unequivocally demonstrated for two other clavams[198] [valclavam (37) and 2-(2-hydroxyethyl)clavam (100)]. As such, the use of two sets of bioassay plates (with and without methionine) was considered potentially useful in the indication of clavams. Interference from penicillin N (24) and other penicillins was effectively removed by penicillinase (DIFCO Labs, Bacto Penase Concentrate), whilst cephalosporins were readily degraded by P-99 β-lactamase from *Enterobacter cloacae*. A mixture of all the compounds in Table 7.1, including a gift of authentic clavalanine from Roche, gave a clear inhibition zone in the methionine-negative plate and no such zone in the methionine-positive plates. Consequently, this design of the bioassay was deemed to be sensitive and selective for clavalanine (99) detection in the culture filtrate of *S. clavuligerus* ATCC 27064. A logarithmic plot of the clavalanine concentration vs. the diameter of the inhibition zone gave an approximately linear relationship (Figure 7.2), suggesting the utility of the bioassay in estimating antibiotic titre in crude broth.

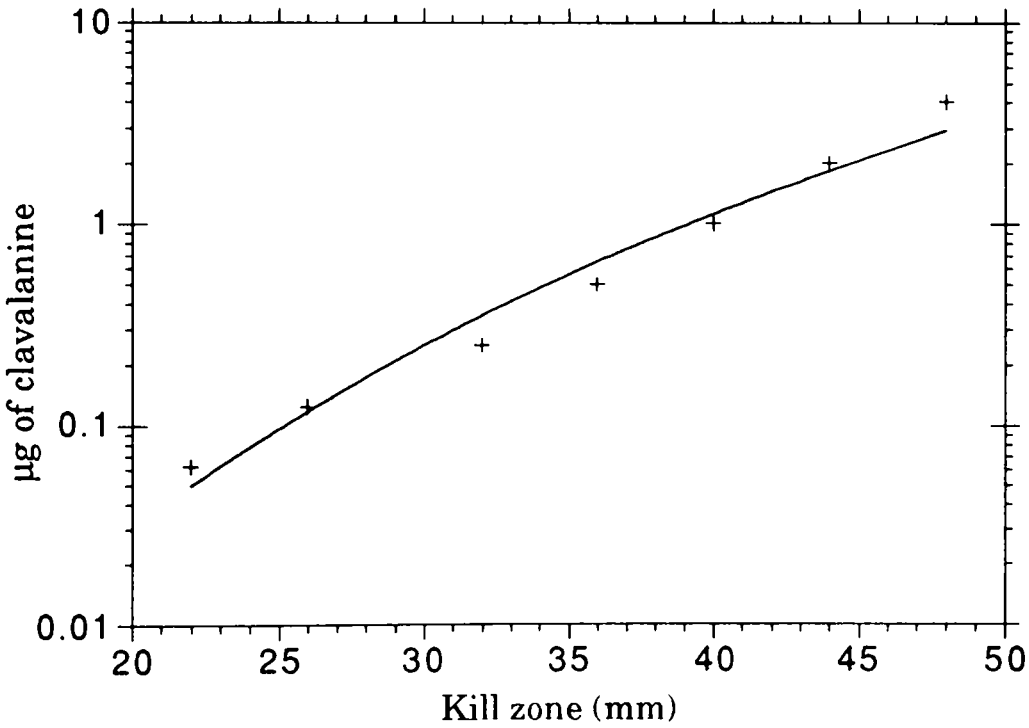



**Figure 7.2** Plot of clavalanine concentration vs. kill zone in *Bacillus* bioassay plate

## 7.2.2 Fermentation of *S. clavuligerus* ATCC 27064 for clavalanine

Despite repeated attempts over a period of five months, there was no success in producing clavalanine (99) from the fermentations although the organism was a gift of Roche who were the first to isolate the antibiotic.[184] Three more strains of *S. clavuligerus* given to us

by Roche also failed to produce clavalanine (99). Instead, the culture filtrates consistently demonstrated anti-*Bacillus* activity that was not reversible by methionine. The HPLC fraction corresponding to the retention time of an authentic sample of clavalanine (99) was not bioactive, whereas some other fractions were bioactive in plates with or without methionine. The bioactive components in these non-clavalanine fractions were not susceptible to β-lactamases in the bioassay plates and must therefore be non-β-lactam in nature. Overall, the absence of clavalanine (99) was as unequivocal as the presence of non-β-lactam antibiotics. No effort was made to investigate the non-clavam activity in order not to lose sight of the original project direction. In a final communication[223] with Roche, it was revealed that the exact clavalanine-producing strain might not, in fact, have been given to us due to logistical problems in a 1985 reorganisation of their culture collection (*i.e.* the strain was lost).

It is well known among microbiologists that the *Streptomyces* bacteria exhibit a high degree of genetic heterogeneity and fluidity even at the species level.[83] The heterogeneity manifests itself in the different biochemical characteristics exhibited by strains that appear morphologically identical. Genetic 'fluidity' implies that each strain can 'obtain' new genes or 'discard' old ones readily under the appropriate growth conditions. Strain selection programs are therefore commonly practised in industry for the purpose of re-isolating the genes encoding the production of a particular metabolite, so that its yield can be improved. In anticipation of the futility of further efforts to produce clavalanine (99), valclavam (37) became the next candidate for our studies.

### 7.2.3 Development of an assay for valclavam (37)

(a) Imidazole derivatisation assay

The imidazole derivatisation assay is a sensitive method employed in the detection and quantitation of clavulanic acid (31).[76, 224] It involves the quantitative reaction of clavulanic acid (31) with imidazole and the subsequent HPLC determination of the adduct (see Figure 7.3) which absorbs strongly at 312 nm ($\varepsilon$ = 26900 $M^{-1}cm^{-1}$).

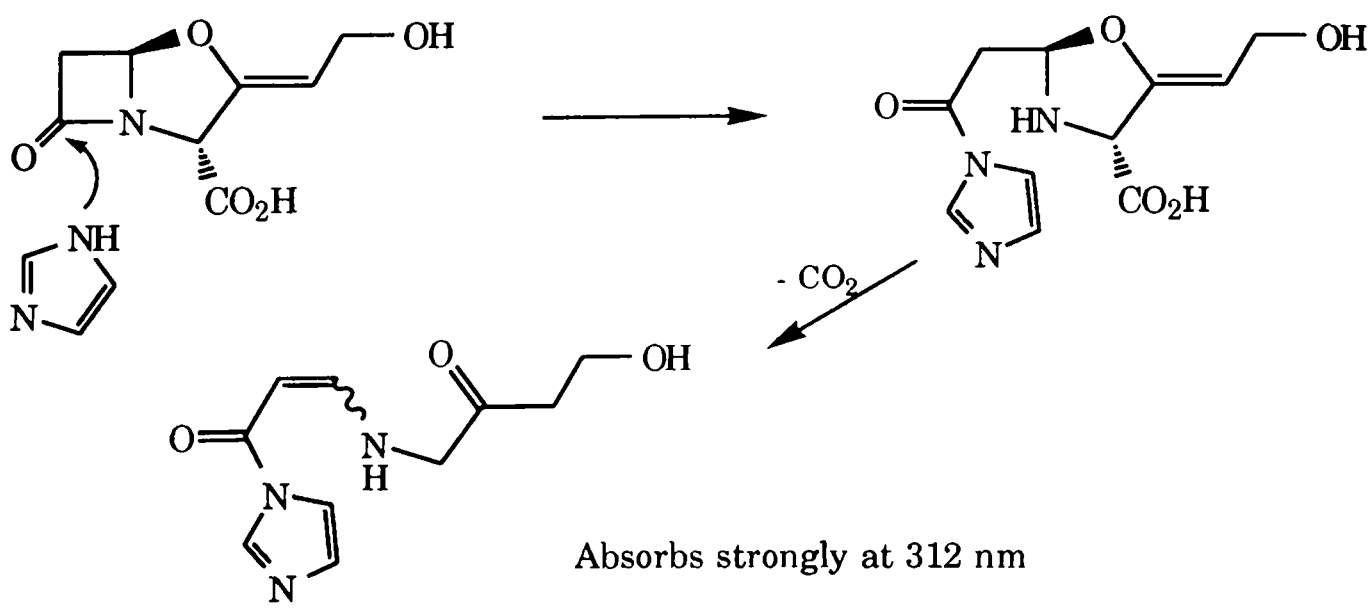



**Figure 7.3** Reaction of clavulanic acid (31) with imidazole

The first attempt to develop an assay for valclavam consisted of adapting the above method of imidazole derivatisation. It was found that valclavam (37) was not quantitatively derivatised by imidazole, since a sample treated according to the clavulanate-imidazole protocol gave a HPLC-detectable valclavam peak which was biologically active. Unreacted valclavam (37) continued to be present even after a five-fold increase in imidazole concentration. The diminished reactivity of valclavam (37) towards imidazole is thought to be a consequence of its lack of the exocyclic double bond at C-2. Therefore, the imidazole assay was deemed to be unsuitable for the determination of valclavam (37).

(b) Bioassay

Like clavalanine (99), valclavam (37) is a 5*S*-configured clavam exhibiting mainly antifungal activity and a restricted antibacterial spectrum. Their similar biological activity suggests that the basic bioassay protocol designed for clavalanine (99) could be similarly applied to valclavam. *Bacillus* sp. ATCC 27860, the test organism for clavalanine, was found to be only mildly inhibited by valclavam (37), giving only a 23 mm kill zone with a 50 μg loading (*cf.* Table 7.1). It was therefore decided to use *E. coli* ATCC 12435 as the test organism for three reasons: (i) it had proved to be highly sensitive to valclavam in a screening done by Zähner *et al.*[198], (ii) it could be procured from a public culture collection and (iii) the bacterium *E. coli* has been used extensively in our group for bioassay purposes, *e.g.* the enzyme bioassay for REXH. Although no β-lactamase-sensitive antibiotics have been reported to be produced by *S. antibioticus* ssp *antibioticus* Tü 1718, β-lactamases were added to the bioassay as a precautionary measure. Reversal of inhibition by methionine was again a useful criterion for clavam identification. Indeed, a mixture of valclavam (37) with the β-lactams in Table 7.1 demonstrated the selectivity of the bioassay. Again, the linear relationship between the logarithm of valclavam concentration and kill zone diameter was useful for estimating valclavam titre in crude broth (Figure 7.4).

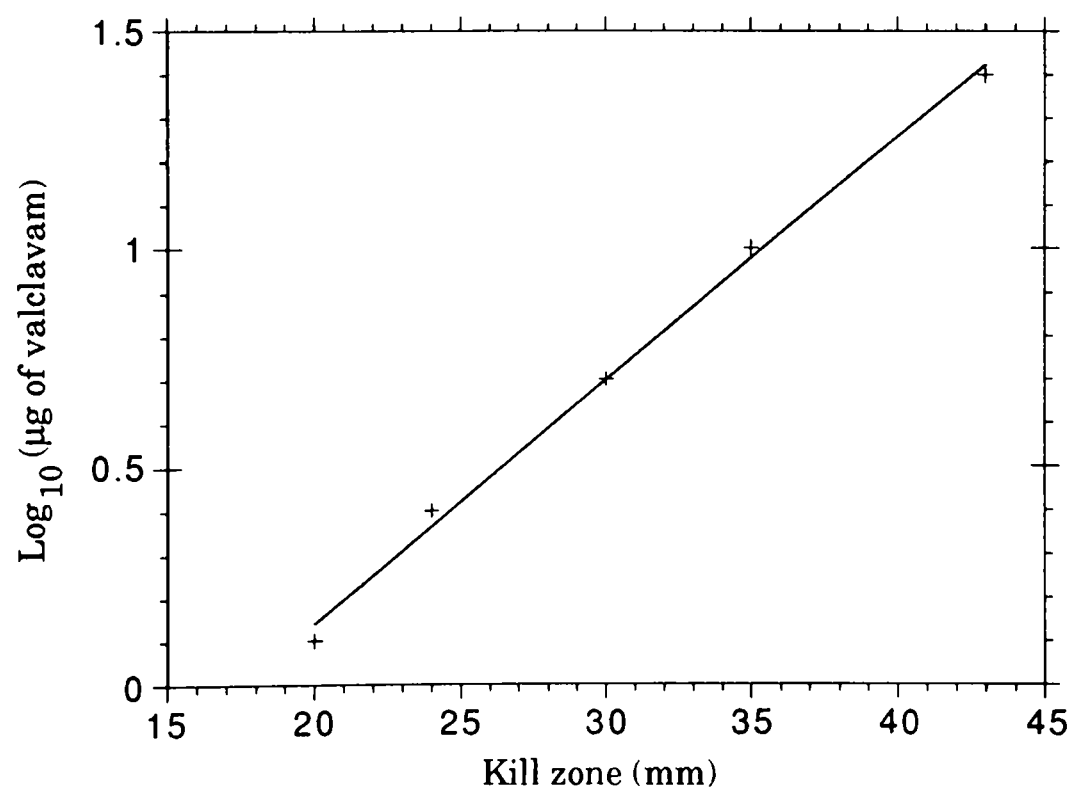



**Figure 7.4** Plot of logarithm of valclavam concentration vs. kill zone in *E. coli* bioassay plate

### 7.2.4 Optimisation of fermentation parameters for valclavam production from *S. antibioticus* ssp *antibioticus* Tü 1718

The organism was a gift of Prof Hans Zähner (University of Tübingen, Germany) who isolated the organism and discovered its production of valclavam (37).[44] The production of valclavam (37) reported in the literature was carried out in fermenters of greater than 5-litre capacity. For our biosynthetic investigations, smaller shake-flask growths were deemed to be appropriate. Hence, parameters used for the fermenters had to be adapted for our shake-flask cultures using 100 ml of medium in 500-ml Erlenmeyer flasks.

In the entire course of this project, a total of 23 batches of fermentations were performed, out of which valclavam was produced in 16 runs. The failures occurred mainly in the earlier runs. Antibiotic production became more predictable in the later runs as conditions were controlled more strictly. It was found that stringent requirements exist for the production of valclavam (37), failing which the biosynthetic machinery for this antibiotic is not operational. The important parameters found to affect the reproducibility of valclavam production are discussed below.

Firstly, as is often the case with microbial production of secondary metabolites, the precise composition of the production medium was found to be fundamental in modulating the biochemical characteristics of the valclavam producer.[83] In this case, even different proprietary brands of soybean meal were discriminated by the organism. Thus, no valclavam (37) was produced when Soyalose 105 (Central Soya) and Arkasoy 50 soybean meals were used. Soybean meal from BDH Biochemicals was the only variety which tested positive for valclavam production. We thought that this difference may have arisen from the variation in fat content of the soybean meals from different manufacturers, being high in the BDH variety and low in the other two brands.

Secondly, the inclusion of glycerol in the culture medium was found to be useful in ensuring consistent valclavam production. Glycerol was not a component of the fermentation method described in the literature.[44] Based on our biosynthetic hypothesis (discussed in Chapter 8) that the β-lactam carbons in valclavam (37) could be assembled from glycerol or closely related $C_3$ primary metabolites, we were hoping to exploit 'the precursor effect'[83] to encourage antibiotic production from *S. antibioticus*. That this hope was fulfilled indirectly contributes to the validation of our hypothesis. In retrospect, it is probable that the high-fat soybean meal promotes antibiotic production through the enhanced supply of glycerol from triglyceride metabolism.

Thirdly, the precise history of the seed culture used as inoculum for the final production flasks was crucial in ensuring reproducible results in production level. Initially, seed cultures were prepared in two steps, involving an intermediate agar slant grown from the glycerol freezes. Both the agar slant medium and the seed medium were based on popular formulations that encourage vigorous propagation of *Streptomyces* bacteria and so differed considerably from

the valclavam production medium. It was found that the transfer of this organism through various media can sometimes deselect the valclavam producing strain from the apparently heterogeneous population, resulting in successful production only about half of the time. Again, this behaviour demonstrates the genetic heterogeneity and fluidity of *Streptomyces* species, in this case *S. antibioticus*. To reduce the probability of medium selection against those cells carrying the valclavam biosynthetic genes, it was decided to minimise the stages prior to the production flasks and to avoid other culture media if possible. Therefore, the seed inoculum was prepared directly from glycerol freezes and the seed medium was made identical to the production medium. Since the inoculation from glycerol freeze to seed flask is much smaller (1% v/v in contrast to 5% v/v in production medium), a slightly longer incubation period was needed to bring about a heavy seed culture. This one-step method of preparing the seed inoculum led to consistent valclavam production.

Fourthly, the pH of the production medium, though always adjusted to the optimal value of 7.2 at the start of fermentation, was found to be liable to erratic fluctuations. In several instances of unsuccessful bacterial growth (and hence antibiotic production), the pH of the medium had become very low (< 5.0). The reasons for this drastic drop in pH are not known but it is apparently the cause of this very low growth rate. One possible way of mitigating this problem would be to include in the production medium an acid-quenching agent (such as calcium carbonate) or a non-metabolisable buffer.

Fifthly, the degree of aeration in the shake flask, as dictated by the shaking speed, has to be maintained at an optimal level by shaking at 250 rpm. Deviations in either direction from this optimal shaking speed led to a gradual reduction in growth rate (data not shown).

Lastly, temperature control was a crucial factor in maintaining antibiotic production. The optimum temperature reported in the literature is 26 - 27°C. During one of the summer experiments to determine the time profile of antibiotic production, the rising ambient temperature inadvertently caused the incubator temperature to rise above 30°C for 12 hr. The valclavam titre stagnated at 6 hr after the onset of this 'heat shock' and virtually disappeared at the end of the 12th hour. Thereafter, all fermentations for valclavam production were carried out in an incubator shaker containing both heating and cooling units, instead of only a heater fan, and temperature control (to ± 1°C) became possible. Such a critical temperature dependence has also been observed for the production of clavulanic acid (31) by *S. clavuligerus*.[225]

When the effect of the above fermentation parameters had been investigated, some of them by systematic study whilst others through serendipity, it was possible to obtain a time profile for the antibiotic production in conjunction with growth rate. A typical time profile for the production of valclavam, under the fermentation conditions described in Section 7.3.5, is shown in Figure 7.5. The growth profile is typical of bacterial cultures in having 3 phases: a lag phase (~ 0th to 60th hour), a log phase (trophophase, ~ 60th to 150th hour) and a stationary phase (~ 150th hour onwards). The antibiotic profile consisted of a lag phase (~ 0th to 80th hour), a production phase (idiophase, ~ 80th to 150th hour) and a degradation phase (~ 150th hour

onwards). The overlap of the trophophase and idiophase in *S. antibioticus* in Figure 7.5 is common among the filamentous microorganisms (the actinomycetes and fungi), in contrast to unicellular bacterial cultures with well defined trophophase and idiophase.[83] The particular nature of the antibiotic production profile impinges directly on the time-frame of the whole-cell feeding experiments carried out in Chapter 8.

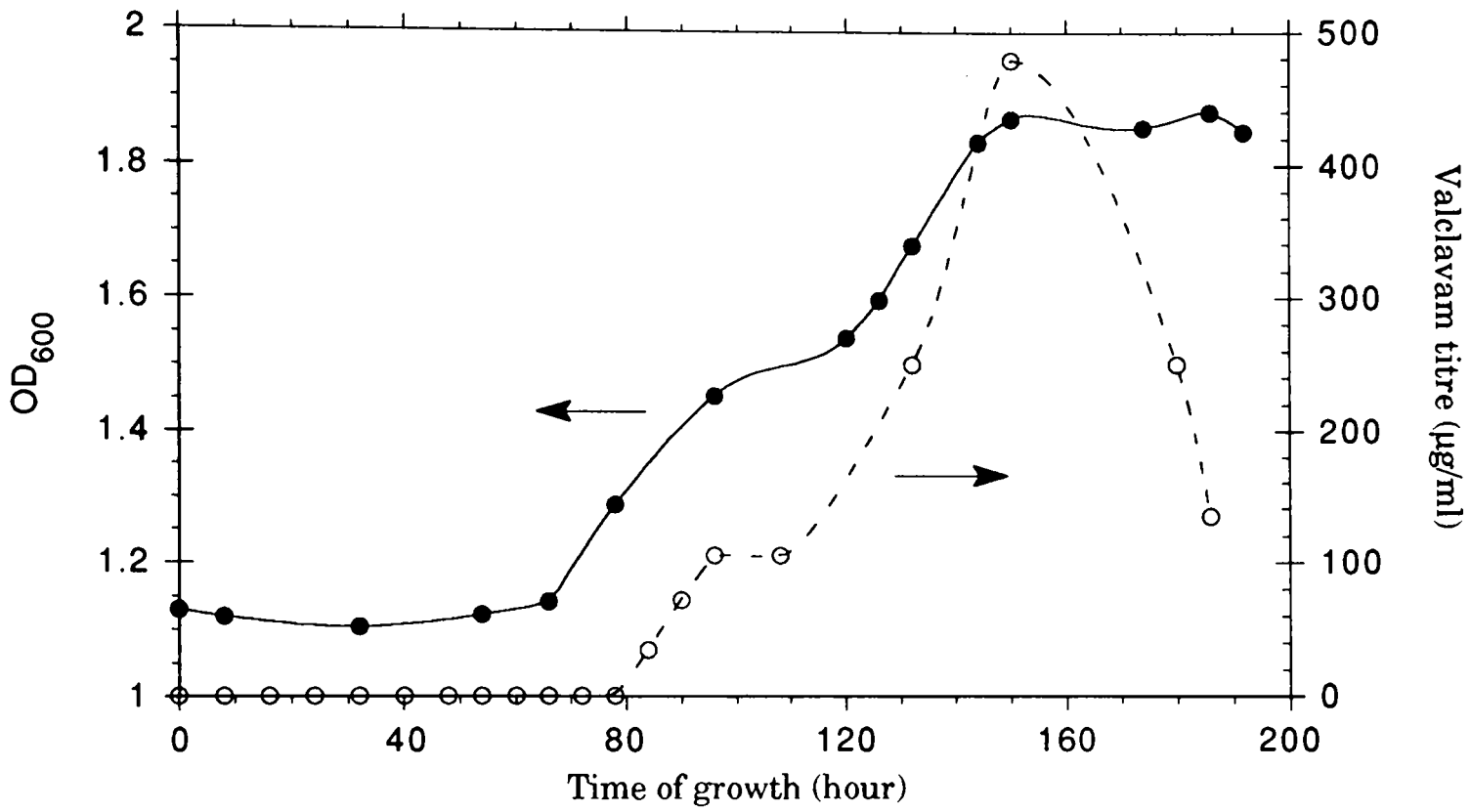



**Figure 7.5** A typical time profile for growth rate and valclavam production of *S. antibioticus* ssp *antibioticus* Tü 1718

## 7.2.5 Isolation and characterisation of valclavam (37) and its dipeptide fragment - Structural revision of valclavam and Tü 1718B

Based on the time course of valclavam production shown in Fig. 7.5, cell harvesting and recovery of the antibiotic was routinely carried out on the 6th day (between 140th and 150th hr) after inoculation. The literature method for the isolation of valclavam is shown in Fig. 7.6.

In essence, the scheme involves 3 high-capacity, low-resolution chromatographic steps (organic adsorption with Amberlite XAD-16 resin, ion-exchange with Amberlite IRA 401-$Cl^-$ resin and gel filtration with Trisacryl GF resin) before a final high-resolution, low-capacity HPLC step. Such a scheme is conventional for processing large quantities (over 10 litres) of culture filtrate, as typically encountered in industry. The number of steps involved and the nature of each step determine directly the time-frame of the entire purification process.

In the context of this project, it was deemed desirable to simplify the above scheme to suit the smaller scale of operation and to shorten the time-frame of the antibiotic isolation procedure. Hence, the first 3 chromatography steps in Figure 7.6 were omitted since they conferred low resolution and were mainly useful for large-scale work. The modified scheme devised for the routine purification of valclavam (37) is shown in Figure 7.7.

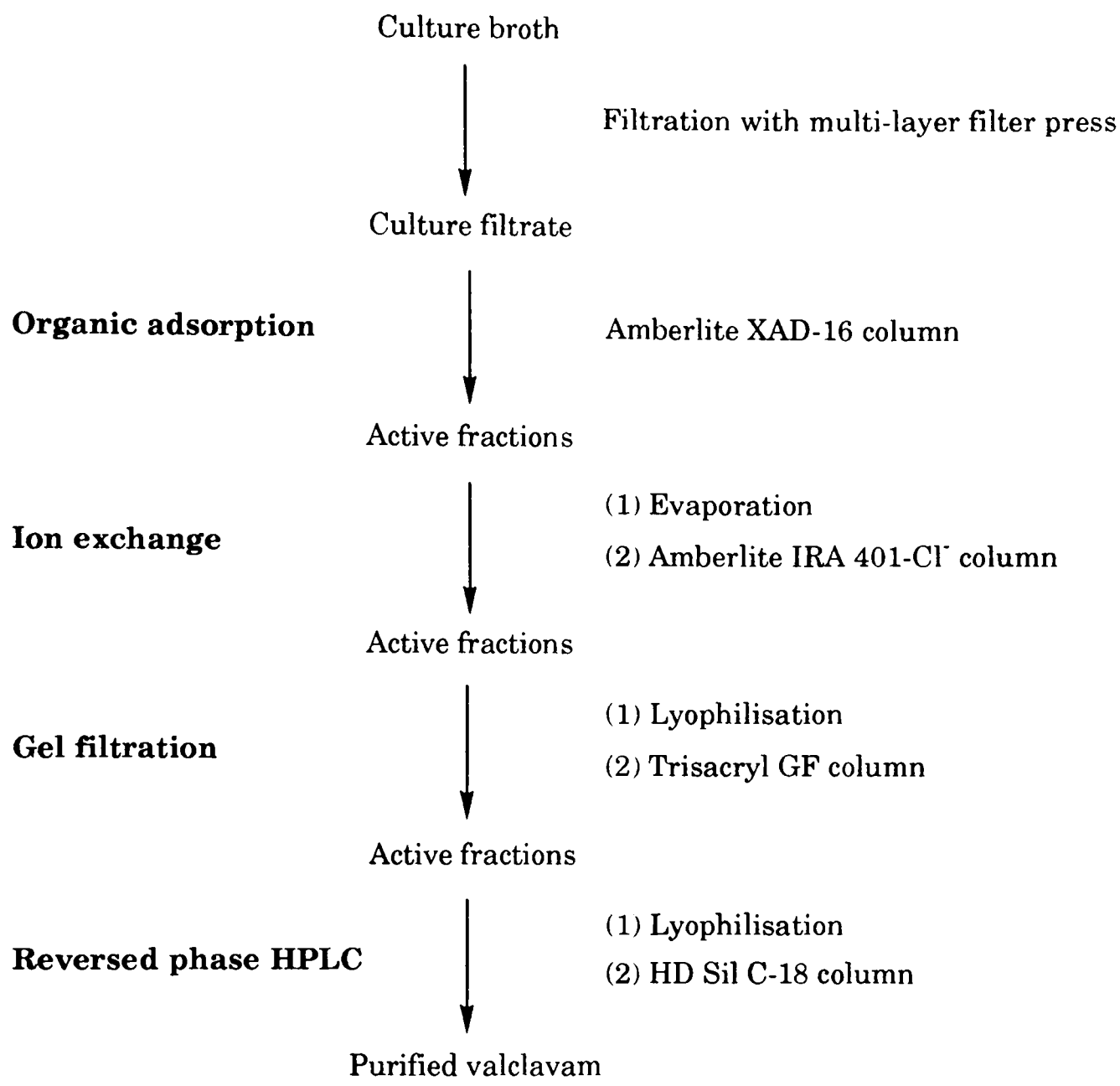



**Figure 7.6** Isolation of valclavam (37) from culture broth of *S. antibioticus*[226]

Thus, in the routine purification of valclavam (37), lyophilisate from the crude culture filtrate was subjected directly to reversed phase HPLC utilising an acetonitrile-phosphate gradient (see Section 7.3) which is adapted from a methanol-based analytical method in the literature.[226] A typical chromatogram from such a HPLC run is given in Figure 7.8, showing that valclavam eluted sharply at about 11.2 min under these conditions. Since the sample loaded onto the column was a very crude heterogeneous mixture, the resolution provided by the HPLC column was lost gradually. After about 20 runs, the resolution had to be restored by cleaning-in-place (CIP) procedures described in the commercial literature, or by repacking the guard column.

The second HPLC procedure employed an identical gradient as the first HPLC procedure but utilised methanol as the organic modifier. The valclavam ($R_t$ = 13.2 min) eluted from this column gave a $^1H$ NMR (500 MHz) spectrum as shown in Figure 7.9. This spectrum is similar to that acquired during the original structural elucidation carried out at Ciba-Geigy.[227]

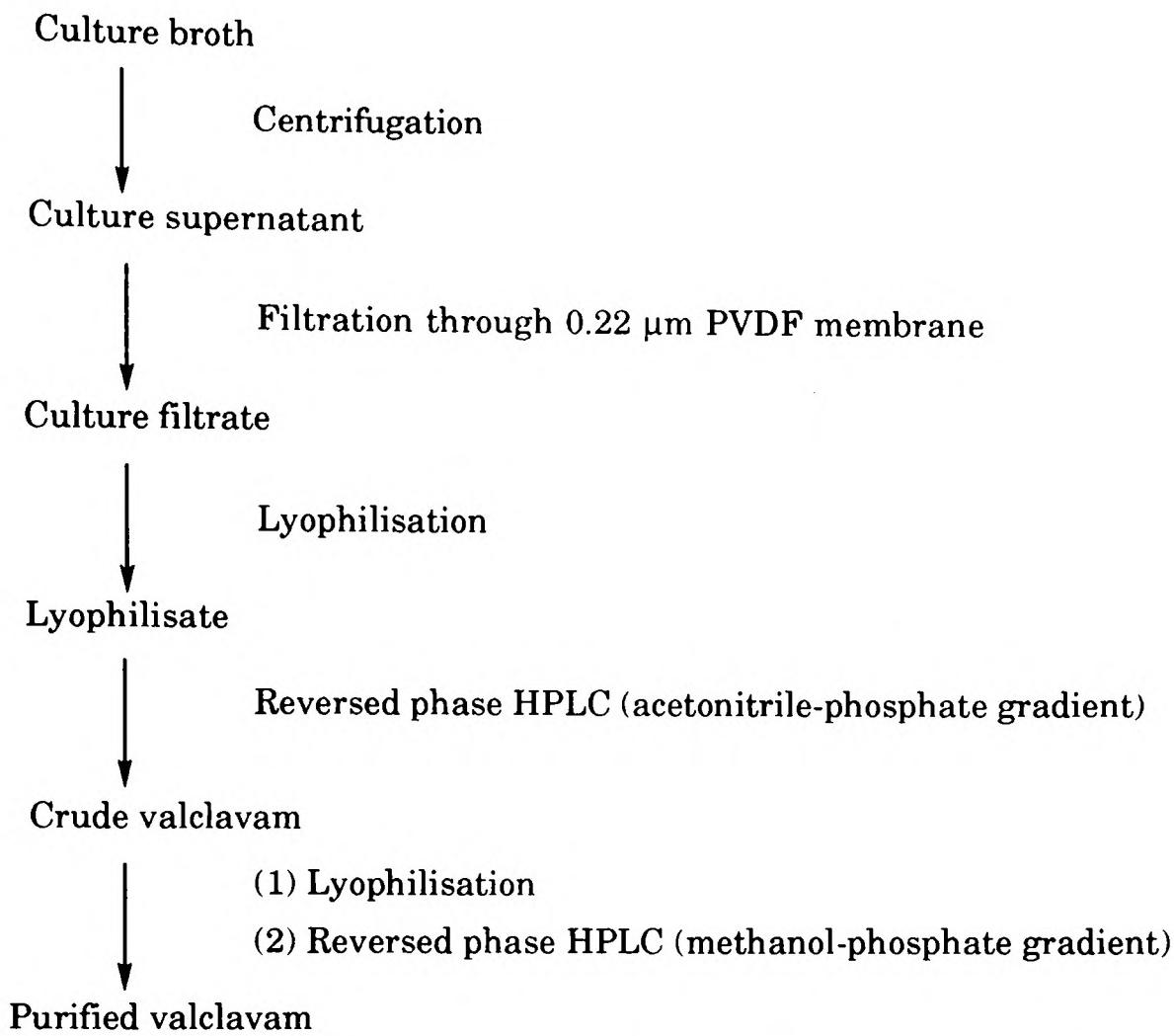



**Figure 7.7** Modified scheme for routine small scale isolation of valclavam

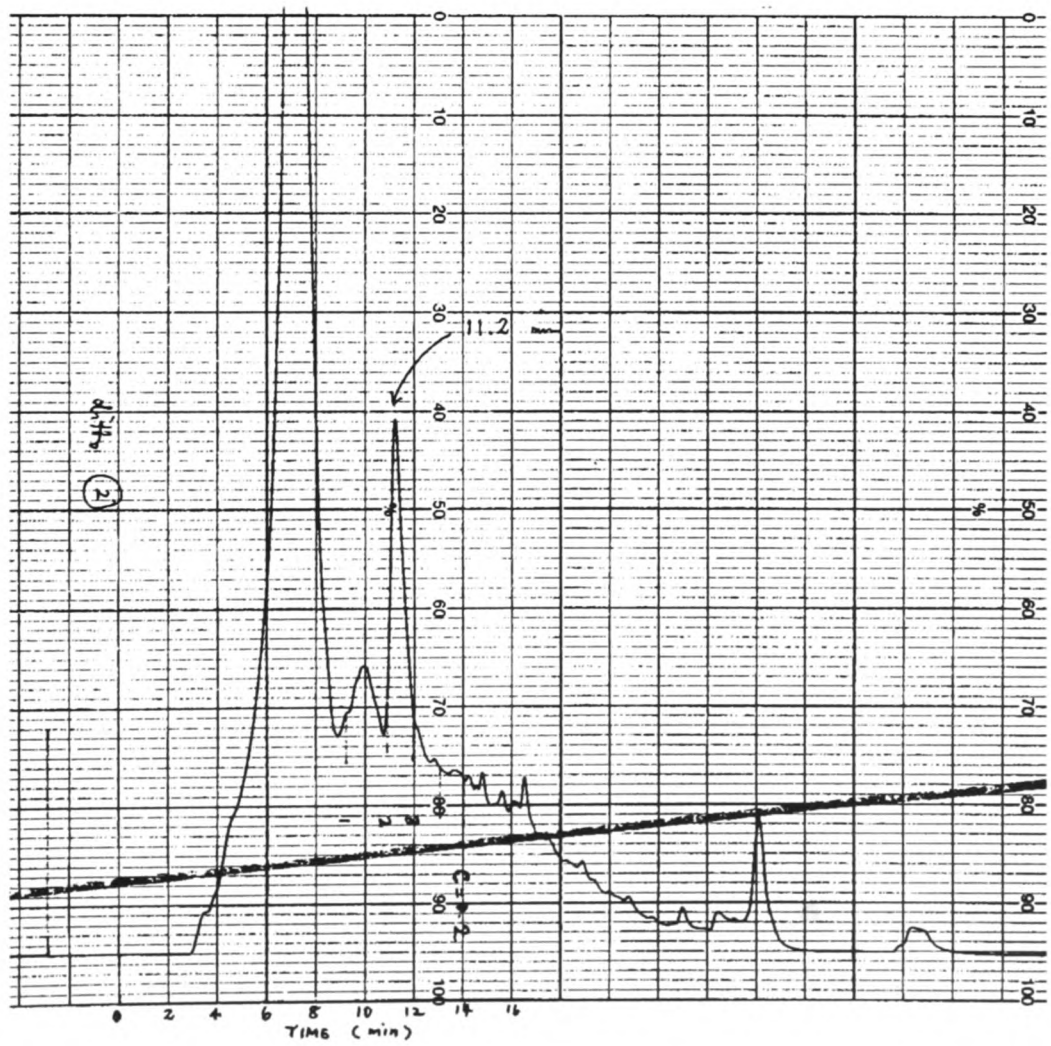
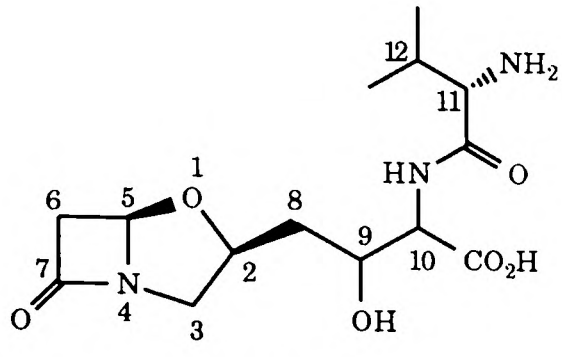
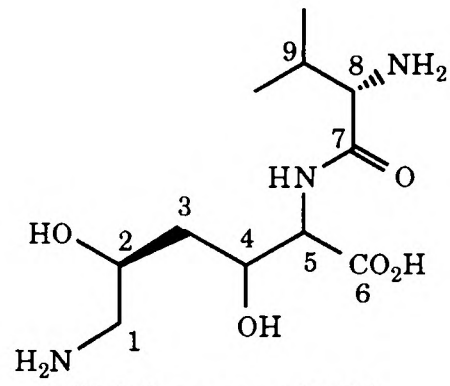



**Figure 7.8** HPLC chromatogram for the culture filtrate of *S. antibioticus* ssp *antibioticus* Tü 1718, with valclavam eluted at 11.2 min




Valclavam (37)




Val-fragment (111)

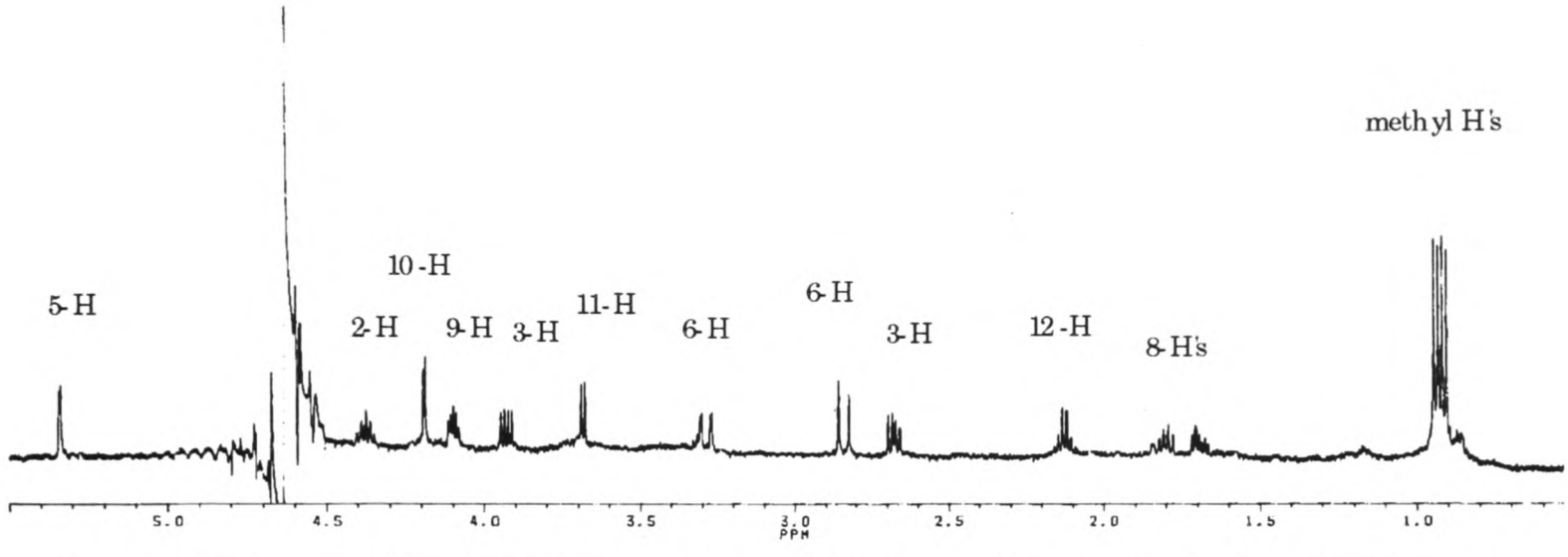



**Figure 7.9** 500 MHz $^1$H NMR spectrum of valclavam (37) obtained after 2 HPLC runs

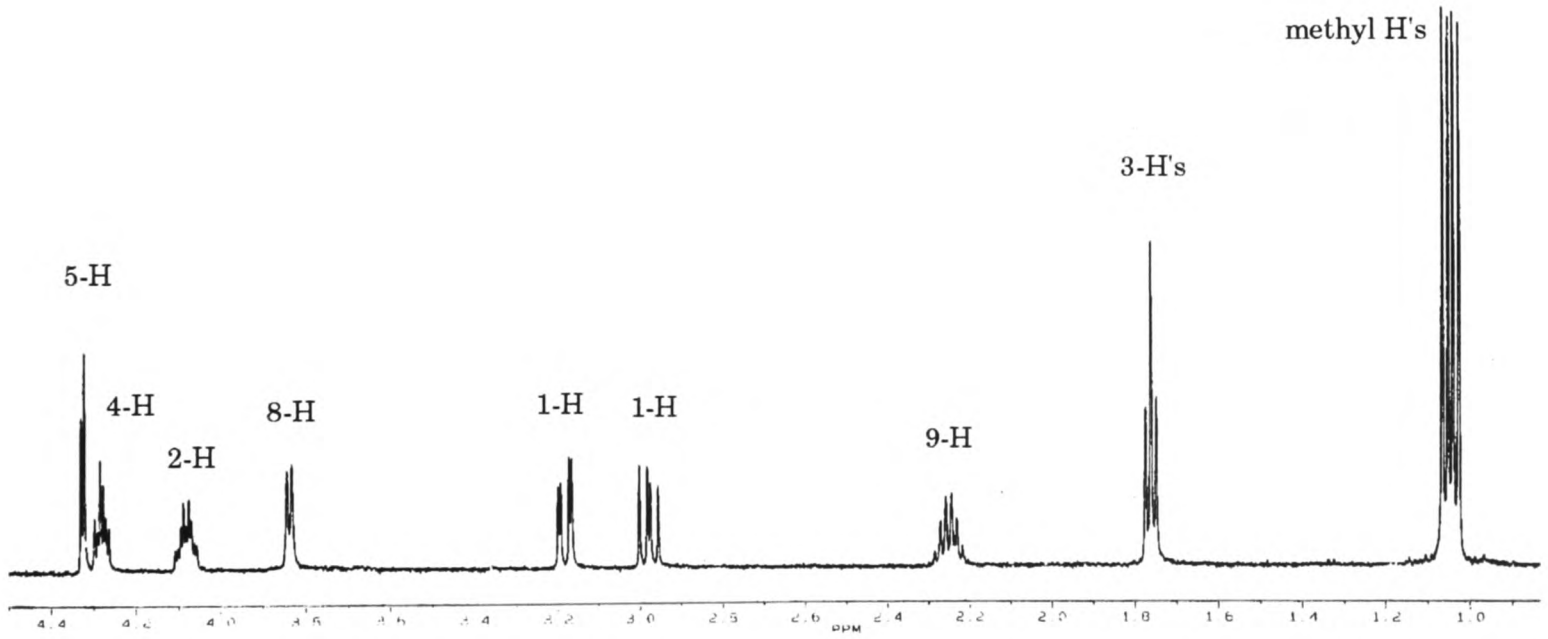



**Figure 7.10** 500 MHz $^1$H NMR spectrum of val-fragment (111) after HPLC in $NH_4HCO_3$

Upon standing in aqueous solution, valclavam was found to degrade to a stable fragment, herein called val-fragment (111), which could be subsequently isolated by HPLC in 10 mM ammonium bicarbonate buffer, pH 8.0 ($C_{18}$ 250 x 10 mm, flow rate = 4 ml/min, detection at 210 nm, $R_t$ = 6.5 min). Even at the most stable pH value of 7, valclavam has a half-life of about 2 days at 20$^o$C. From its 500 MHz $^1$H NMR spectrum (Figure 7.10) and a m/z value of 278 for the protonated molecular ion in ESMS, the degradation product was tentatively assigned the structure (111) which could have arisen from valclavam through the pathway depicted in Figure 7.11. An ESMS spectrum for a degrading sample of valclavam gave molecular ions for valclavam ($MH^+$ = 330), val-fragment ($MH^+$ = 278) and the intermediate imine (structure 112, $MH^+$ = 348).

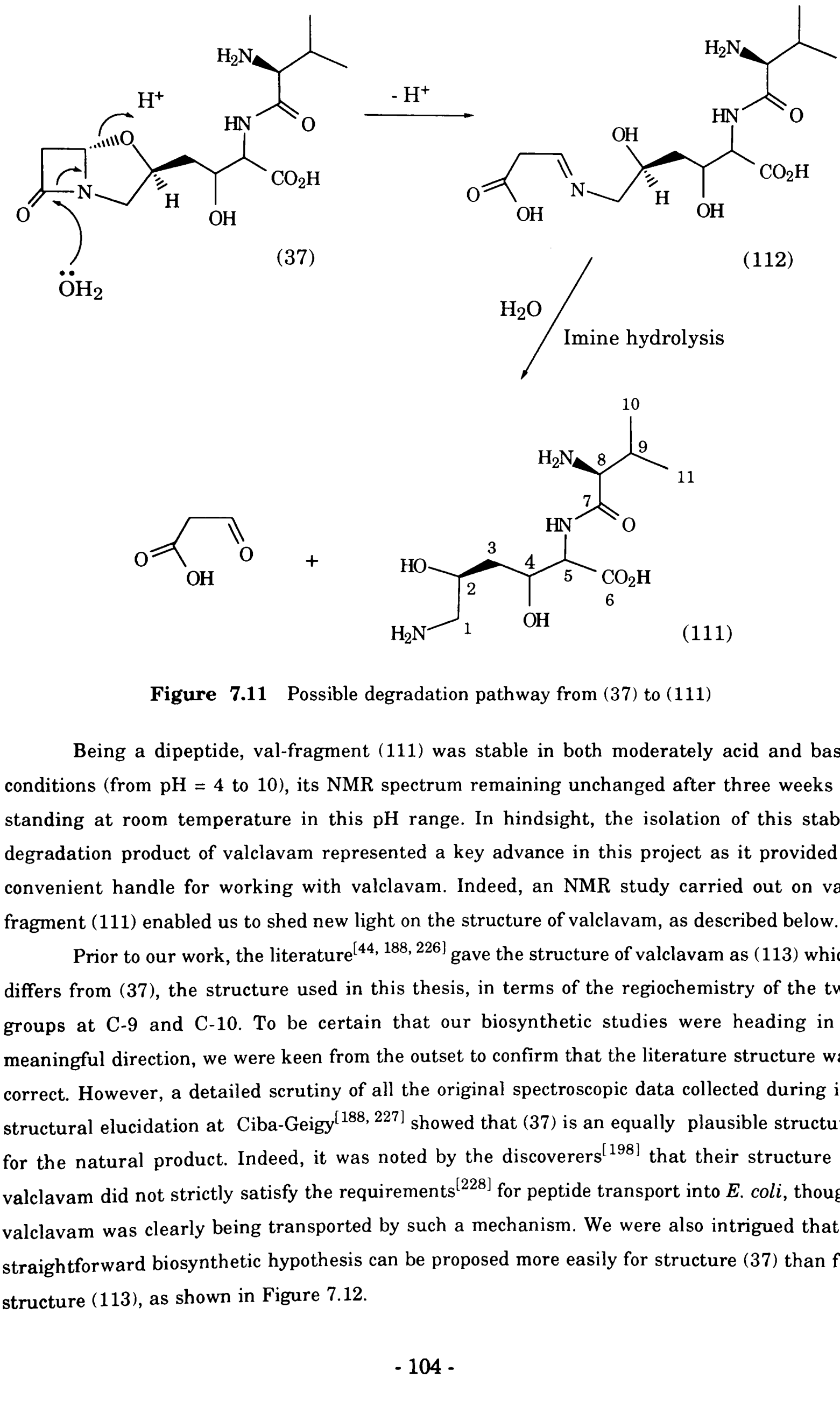



**Figure 7.11** Possible degradation pathway from (37) to (111)

Being a dipeptide, val-fragment (111) was stable in both moderately acid and basic conditions (from pH = 4 to 10), its NMR spectrum remaining unchanged after three weeks of standing at room temperature in this pH range. In hindsight, the isolation of this stable degradation product of valclavam represented a key advance in this project as it provided a convenient handle for working with valclavam. Indeed, an NMR study carried out on val-fragment (111) enabled us to shed new light on the structure of valclavam, as described below.

Prior to our work, the literature[44, 188, 226] gave the structure of valclavam as (113) which differs from (37), the structure used in this thesis, in terms of the regiochemistry of the two groups at C-9 and C-10. To be certain that our biosynthetic studies were heading in a meaningful direction, we were keen from the outset to confirm that the literature structure was correct. However, a detailed scrutiny of all the original spectroscopic data collected during its structural elucidation at Ciba-Geigy[188, 227] showed that (37) is an equally plausible structure for the natural product. Indeed, it was noted by the discoverers[198] that their structure of valclavam did not strictly satisfy the requirements[228] for peptide transport into *E. coli*, though valclavam was clearly being transported by such a mechanism. We were also intrigued that a straightforward biosynthetic hypothesis can be proposed more easily for structure (37) than for structure (113), as shown in Figure 7.12.

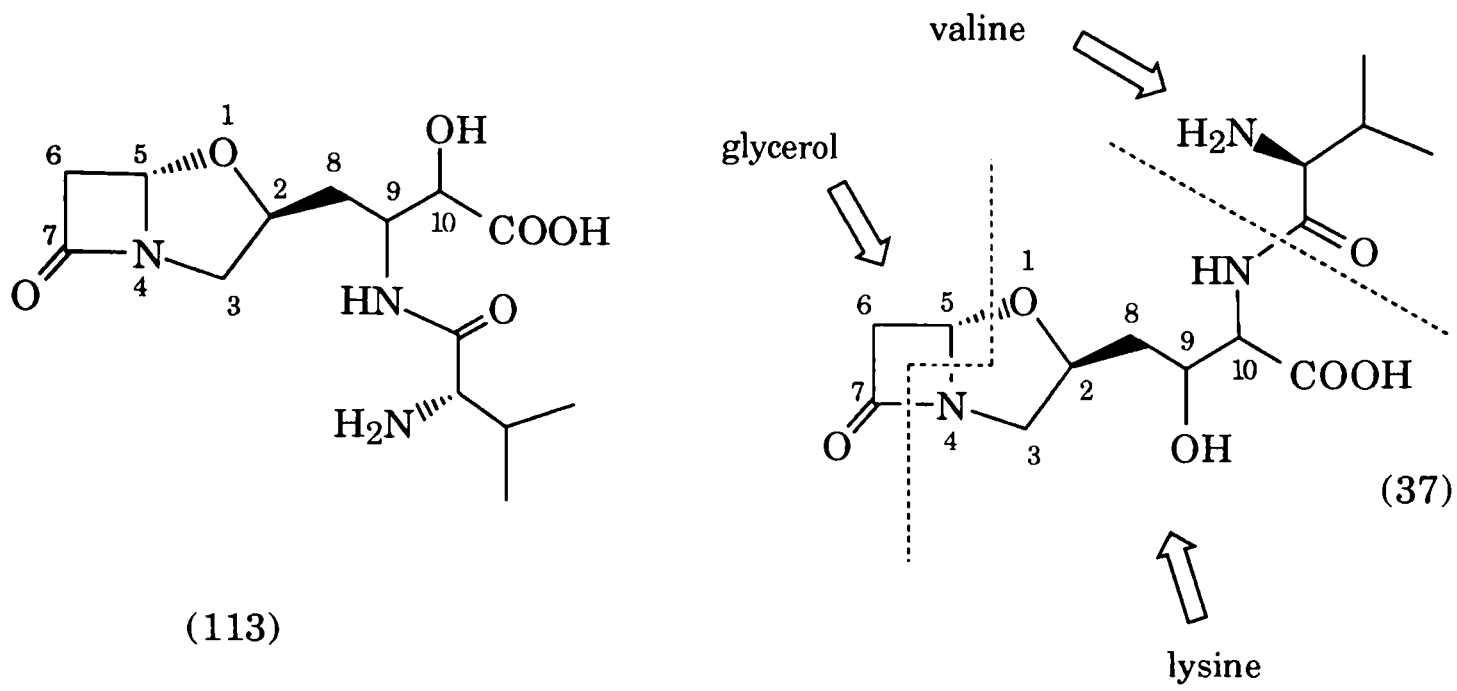



**Figure 7.12** Two possible structures for valclavam based on Ciba-Geigy data[188, 227]

We anticipated that the question of whether (113) or (37) represents the structure of valclavam correctly could be answered by looking at the corresponding regiochemistry in val-fragment (111). Thus, a new priority emerged at this stage of our project, *viz.* to determine whether val-fragment possesses structure (111) or (114). The latter structure can be derived from the literature structure (113) in an analogous fashion to the transformation from (37) to (111) (Figure 7.13).

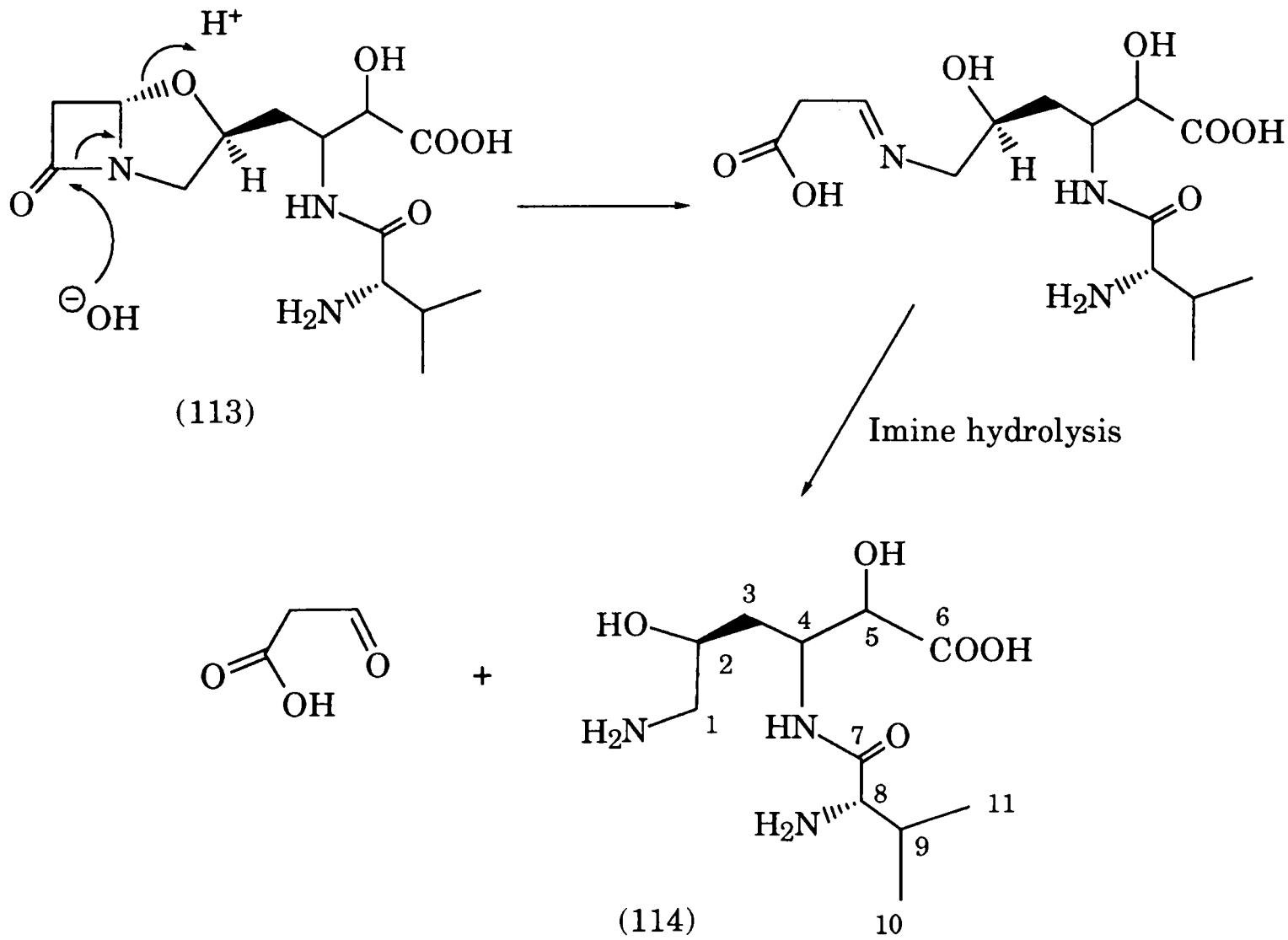



**Figure 7.13** Possible degradation pathway from (113) to (114)

A 500 MHz $^{1}H$-$^{1}H$ COSY spectrum of val-fragment was consistent with the connectivity as indicated in either structure (111) or (114). Signals for all 11 carbon atoms of val-fragment could be observed after a 12-hr 1D acquisition. Assignment of all the protonated carbon resonances was achieved via a HMQC shift correlation experiment.[229] Long-range (two- and three-bond) $^{1}H$-$^{13}C$ correlations were obtained from HMBC experiments.[230] These were recorded on a basic sample (pD 7.5 - 8.5, from the residual ammonium bicarbonate) and on the same sample after acidification (pD < 1.0). This was necessary because at the basic pD many proton resonances were broadened, most notably 8-H, and failed to show correlations. These resonances were readily sharpened on acidification. Correlations were recognised directly from the 2D plot and from columns of the 2D spectrum corresponding to the $^{13}C$ dimension. Chemical shift assignments of val-fragment are given in Table 7.2, and observed long-range correlations summarised in Table 7.3. All long-range correlations were consistent with the dipeptide moieties identified in COSY spectra and were sufficient to assign the two carbonyl resonances unambiguously. The regiochemistry of val-fragment was established by the observation of long-range correlations of protons 4 and 5. Thus, only 5-H showed correlations with both carbonyl resonances, whereas 4-H correlated with C-6 only (Figure 7.14). Such observations would be consistent with structure (111) only. An attempt to use fragmentation patterns in Fast Atom Bombardment-Mass Spectrometry to corroborate this regiochemical assignment failed to give any conclusive results.

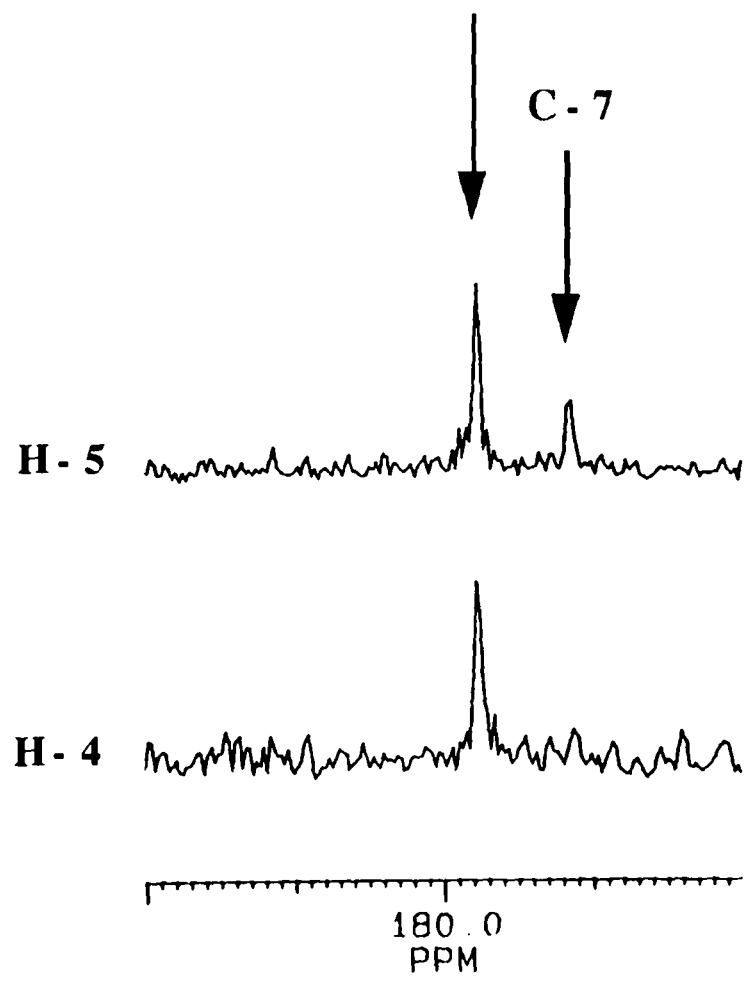



**Figure 7.14** Sections of $^{13}C$ columns of a HMBC experiment recorded for the basic sample of val-fragment. Long-range correlations to carbonyl groups are shown for protons 4 and 5.

| Carbon atom | Acidic pD | | Basic pD | |
|---|---|---|---|---|
| | $^{1}H$ | $^{13}C$ | $^{1}H$ | $^{13}C$ |
| 1 | 2.97, 3.18 | 47.3 | 2.99, 3.20 | 47.2 |
| 2 | 4.05 | 68.4 | 4.08 | 68.4 |
| 3 | 1.78 | 40.5 | 1.75 | 40.9 |
| 4 | 4.47 | 71.0 | 4.28 | 72.0 |
| 5 | 4.68 | 59.6 | 4.33 | 61.6 |
| 6 | - | 175.8 | - | 178.9 |
| 7 | - | 172.8 | - | 175.9 |
| 8 | 4.00 | 61.3 | 3.84 | 62.8 |
| 9 | 2.30 | 33.0 | 2.26 | 33.9 |
| 10 | 1.05 | 19.6 | 1.00 | 19.8 |
| 11 | 1.08 | 20.6 | 1.03 | 21.2 |

**Table 7.2** Proton and carbon assignments of val-fragment at pD < 1.0 and *ca.* 8.0, and at 311 K. Shifts are relative to TSP at 0.00 ppm. Proton shifts were measured in a 1D spectrum whereas carbon shifts were taken from columns of the 2D HMBC experiment.

| Proton | Carbon-13 (Basic) | Carbon-13 (Acidic) |
|---|---|---|
| 1[a] | 2 | 2, 3 |
| 2[b] | - | - |
| 3 | 1, 2, 4, 5 | 1, 2, 4, 5 |
| 4 | 2, 3, 6 | 3, 6 |
| 5 | 3, 4, 6, 7 | 6, 7 |
| 8[c] | - | 7, 9, 10, 11 |
| 9 | 7, 8, 10, 11 | 10, 11 |
| 10 | 8, 9, 11 | 8, 9, 11 |
| 11 | 8, 9, 10 | 8, 9, 10 |

[a] Correlations were observed for the high-field proton only.
[b] No correlations were observed for proton 2, presumably because of the high multiplicity and low intensity of the proton resonance.
[c] In basic solution, no correlations were observed for proton 8 as the resonance was very broad. This resonance was a sharp doublet in acidic solution and gave rise to the listed correlations.

**Table 7.3** Summary of the long-range $^{1}H$-$^{13}C$ correlations observed in HMBC experiments performed on val-fragment

Thus, the range of NMR experiments that have been performed on val-fragment (111) included 1D $^{1}H$ NMR, $^{1}H$-$^{1}H$ COSY, 1D $^{13}C$ NMR, HMQC and HMBC. The HMBC experiment proved to be crucial in the assignment of val-fragment as structure (111) rather than its regio-isomer (114). A corollary of this result would be that the structure of valclavam hitherto reported as (113) was wrong. Therefore, the structure of valclavam should be revised to (37). Having settled the issue of regiochemistry, synthetic efforts are underway[231] to clarify the remaining

stereochemical questions in order to establish the complete structure of these compounds.

Val-fragment (111) provided another surprise when it was discovered that it possessed its own antibiotic activity independent of valclavam (37). It was observed that a fresh sample of valclavam inhibited *E. coli* ATCC 12435 only in the absence of methionine, whereas an old sample produced inhibition whether or not methionine was present. The degree of inhibition in the methionine-containing plate increased with the age of valclavam sample, suggesting that a substance derived from valclavam was responsible for this activity. The use of purified val-fragment (111) confirmed it as the bioactive species. Whilst the mode of action of val-fragment (111) is not known, it is definitely not due to the inhibition of methionine biosynthesis as in the case of valclavam. Due to the dipeptide nature of val-fragment (111), *i.e. L*-valyl-β,δ-dihydroxylysine, it is probable that the compound is an inhibitor of some peptide receptors in *E. coli*, such as peptide-transport systems or essential proteases. It is remarkable that the degradation product of an antibiotic possesses a distinct antibiotic activity. A plot of the logarithm of val-fragment (111) titre vs. kill zone in the *E. coli* bioassay plate exhibits linearity as shown in Figure 7.15. It is not clear from these preliminary results if the smaller extent of inhibition in plates containing methionine implies any interference with its biosynthesis.

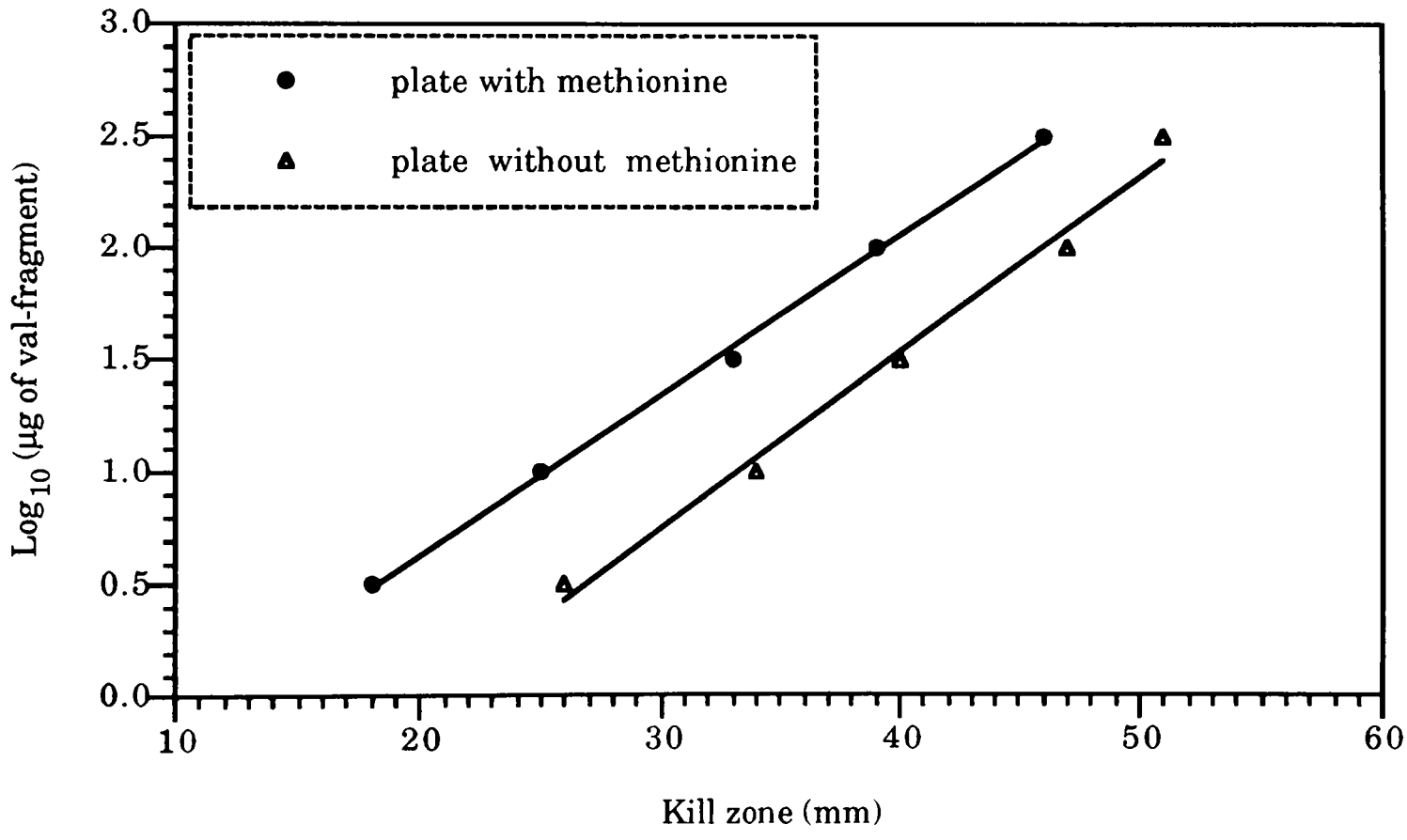
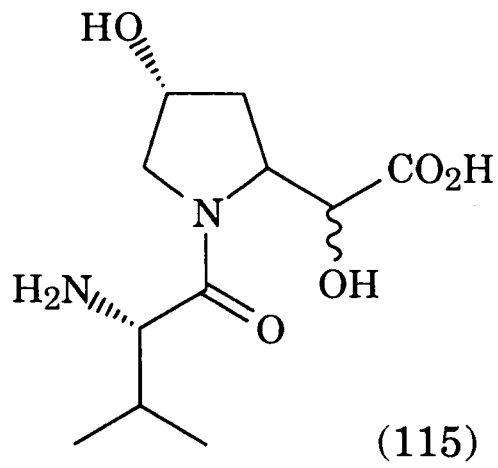



**Figure 7.15** Inhibition of *E. coli* ATCC 12435 by val-fragment (111)

After the revision of the structure of valclavam, our attention was attracted by a report[232] on Tü 1718B, a pseudopeptide antibiotic also isolated from *S. antibioticus* ssp *antibioticus* Tü 1718 and assigned the structure (115). A recent synthesis[233] of the two possible epimers of structure (115) revealed both to be lacking in the antibiotic properties exhibited by the natural product. Furthermore, their $^{1}H$ and $^{13}C$ NMR spectra clearly differed

from the corresponding spectra of the natural product, indicating that structure (115) cannot be the correct representation of Tü 1718B.




(115)

We were then surprised to find that the $^1H$ NMR signals reported for Tü 1718B are essentially similar to our corresponding data for val-fragment (111) (Table 7.4), suggesting the same identity for the two compounds.

| Hydrogen atom(s) | δ (ppm) and multiplicity | |
|---|---|---|
| | **Tü 1718B (pD = 8.4)** | **Val-fragment (pD ~ 8.0)** |
| 2 x 1-H | 2.95 (q), 3.15 (q) | 2.99 (q), 3.20 (q) |
| 2-H | 4.05 (m) | 4.08 (m) |
| 2 x 3-H | 1.75 (t) | 1.75 (t) |
| 4-H | 4.25 (m) | 4.28 (m) |
| 5-H | 4.30 (d) | 4.33 (d) |
| 8-H | 3.90 (d) | 3.84 (d) |
| 9-H | 2.30 (m) | 2.26 (m) |
| 3 x 10-H | 1.03 (d) | 1.00 (d) |
| 3 x 11-H | 1.05 (d) | 1.03 (d) |

**Table 7.4** Comparison of $^1H$ NMR signals for Tü 1718B and val-fragment (111)

The two largest FD-MS ions for Tü 1718B were reported to give m/z values of 278 and 260. The 260 peak was assigned by the authors as the $M^+$ ion, whereas the 278 peak was unaccounted for. We now believe that their interpretation was incorrect. Since the MS was performed in the field desorption (FD) mode on a Varian MAT 711 instrument, the parent molecular ion should be protonated and register a m/z value of 261 for structure (115). In addition, an ammonium or water adduct of the parent ion is not normally encountered in the FD mode and the unaccounted 278 peak becomes totally inexplicable. Therefore, we concluded that the 278 peak belonged to the $MH^+$ ion and the 260 peak arised from the loss of a water molecule, in agreement with the ESMS spectrum we obtained for val-fragment (111) (Figure 7.16).

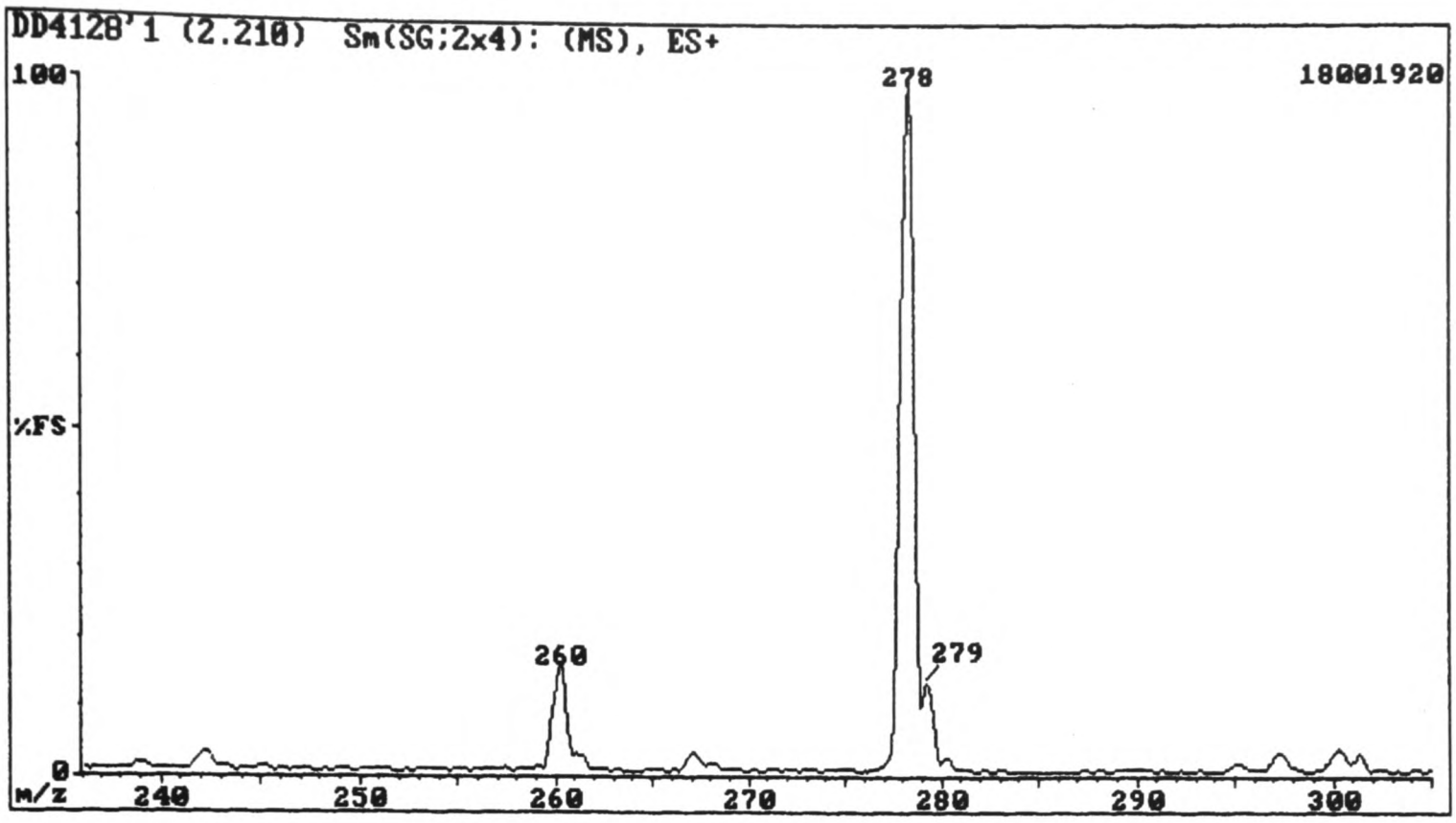



**Figure 7.16** ESMS spectrum of val-fragment (111)

Tü 1718B is also reported to inhibit the growth of *E. coli*, just as val-fragment (111) was shown earlier to be active against *E. coli*. The optical properties of val-fragment (111) were subsequently found to coincide with those of Tü 1718B: a UV spectrum showing a maximum at 210 nm and a CD spectrum (Figure 7.17) showing a negative Cotton effect at 217 nm with a positive band below 200 nm.

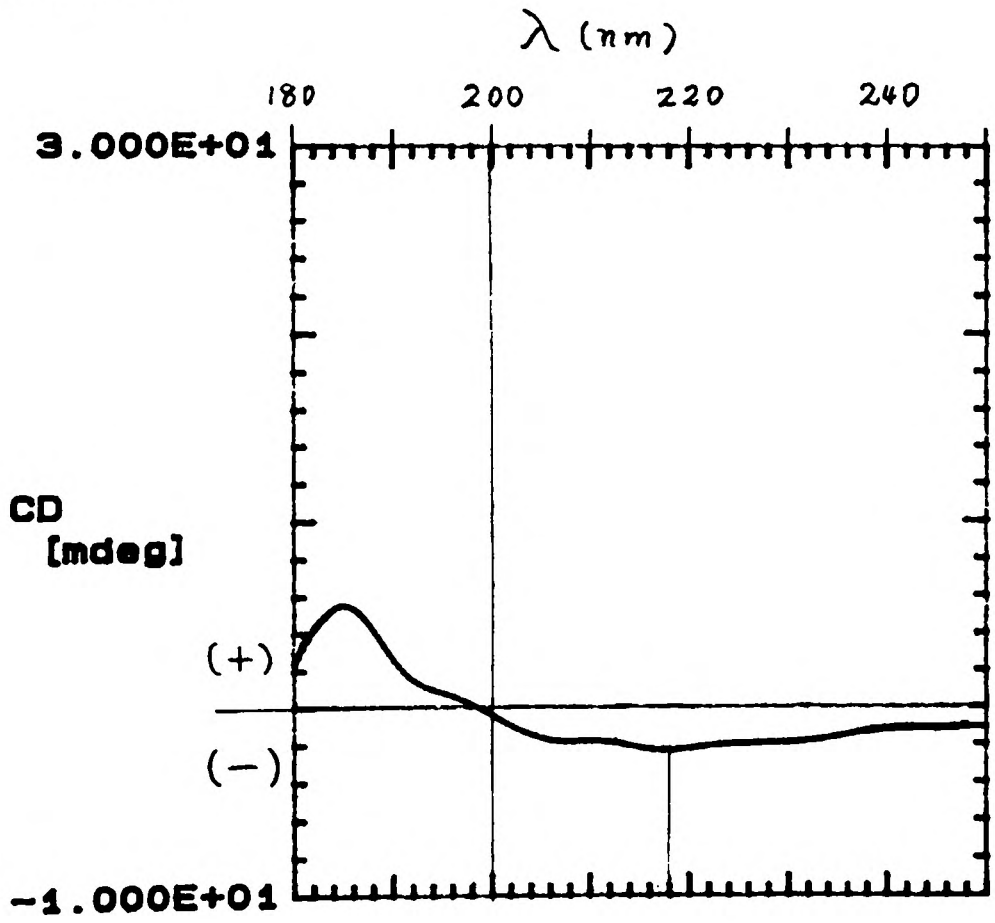



**Figure 7.17** CD spectrum of val-fragment (111)

We were, however, unable to repeat the observation that *L*-4-hydroxyproline could be obtained from the oxidative degradation of Tü 1718B with periodate according to the method of Yanuka.[234] Structure (115) could give rise to *L*-4-hydroxyproline via 2 steps: (i) an amide cleavage under the strongly acidic condition of the periodate oxidation and (ii) degradation of the α-hydroxyacid moiety to a carboxylic group. This was cited as the strongest evidence for the presence of a cyclic moiety as in structure (115). We thought it possible, although unlikely, that under the conditions of periodate oxidation, val-fragment (111) could yield *L*-4-hydroxyproline via the scheme in Figure 7.18. Our failure to detect *L*-4-hydroxyproline from such a reaction remains a discrepancy which cannot be readily explained.

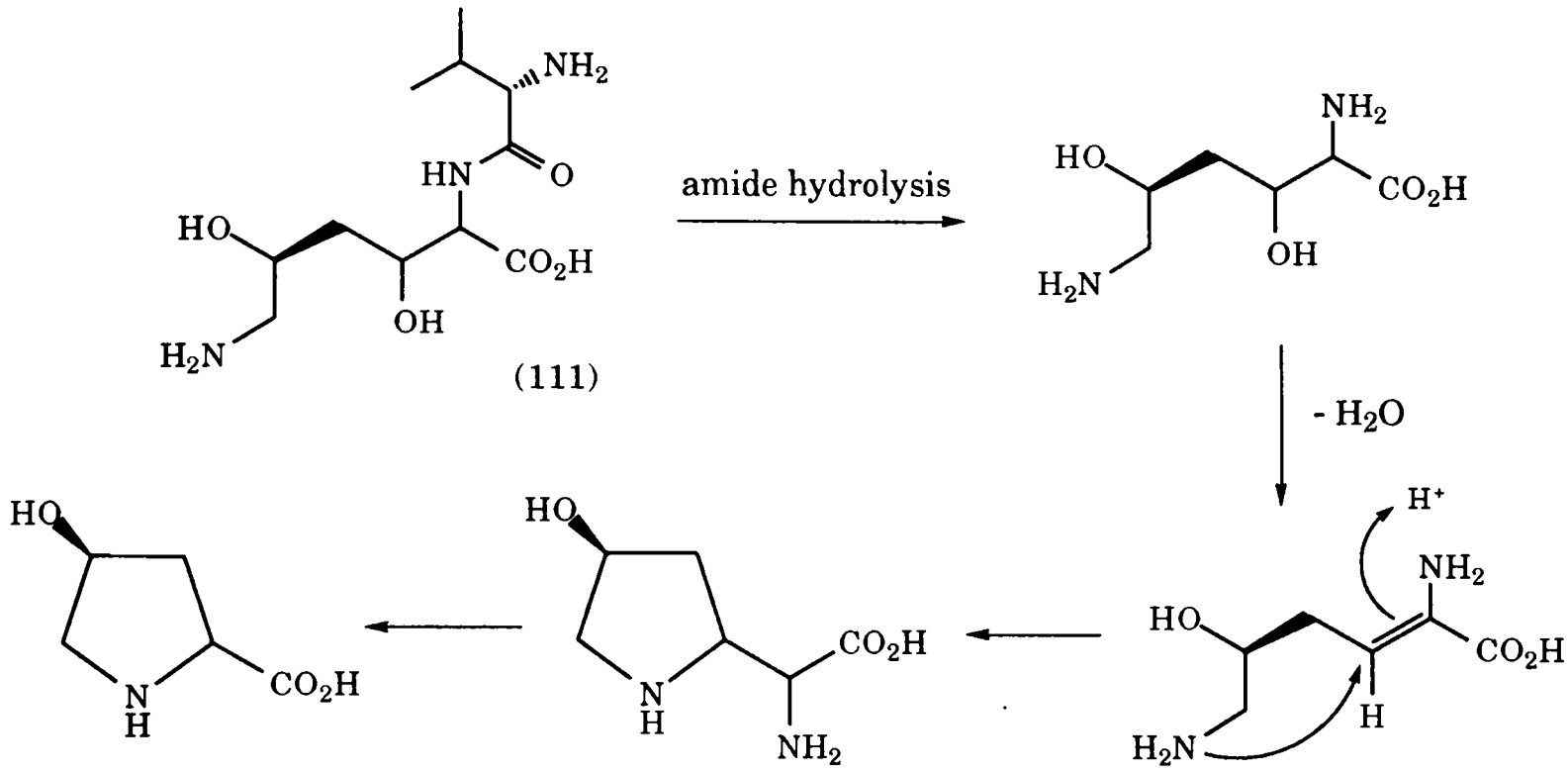



**Figure 7.18** Hypothetical formation of 4-hydroxyproline from val-fragment (111) under conditions of periodate oxidation[234]

Nevertheless, our spectroscopic (NMR, UV, CD, MS) and biological data were sufficient to convince us that the correct structure for Tü 1718B is (111) and not (115). In addition, we believe that Tü 1718B is a degradation fragment of valclavam (37) and probably not a biosynthetic end-product in its own right.

## 7.3 Experimental

### 7.3.1 General techniques

Amino acid analysis was performed by Mr Anthony Willis (MRC Immunochemistry Unit, Oxford University) on an Applied Biosystems automated amino acid sequencer model 470A, fitted with an on-line model 120A for detection of phenylthiohydantoin-amino acids. UV and CD spectra were acquired by Dr Alison Rodger using a JASCO-720 spectropolarimeter under nitrogen-purged conditions. Other facilities, *e.g.* for HPLC, ESMS, bioassay incubation and freeze-drying were similar to those described in Chapters 3, 4 and 5.

### 7.3.2 Bioassay for clavalanine (99)

The holed-plate assay method, using *Bacillus sp.* ATCC 27860 as the test organism, was employed for the detection of clavalanine (99) in culture filtrates. Bioassay plates were prepared by Mr John Keeping in an aseptic manner as follows :

Cultures of *Bacillus sp.* ATCC 27860 were maintained as glycerol freezes stored at -80°C. Inoculum for bioassay plates was grown up overnight, at 37°C, as 5 ml agar slants on Oxoid Nutrient Broth (CM 1). The inoculum (4 ml) was added to 400 ml of Davis minimal agar, maintained at 45°C. [Composition of medium (g/l) : $K_2HPO_4$, 7.0 ; $KH_2PO_4$, 3.0 ; $Na_3$-citrate, 0.5 ; $MgSO_4.7H_2O$, 0.1 ; $(NH_4)_2SO_4$, 1.0 ; glucose (autoclaved separately), 2.0 ; Oxoid No.1 Agar, 10.0 ; deionised water added to 1 litre, pH 7.0]

Penicillinase (0.4 ml) (DIFCO Laboratories, Bacto Penase Concentrate, 10 000 000 IU/ml) was added to the diluted bacterial suspension, followed by 2.0 ml of cephalosporinase (P-99 β-lactamase from *Enterobacter cloacae*, 2 mg/ml). *DL*-Methionine (4 ml) (BDH, 40 mg/ml in water) was then added to give the methionine-positive plates. The bacterial agar suspension, with the above additives, was poured into 8-cm Petri dishes (18 ml per dish) and allowed to solidify. When stored at 4°C, these plates were usable for up to 2 weeks. The methionine-negative plates were prepared in the same manner as above except that no methionine was added. For detecting clavalanine (99), 100 µl of the sample solution was loaded into 8-mm wells (made with a cork-borer in both sets of plates. The loaded plates were then incubated at 37°C for 12 hr.

### 7.3.3 Fermentation of *S. clavuligerus* ATCC 27064 for clavalanine production and HPLC analysis of the culture filtrate

Cultures of *S. clavuligerus* were maintained as glycerol freezes stored at -80°C. The glycerol cultures were thawed and streaked onto slopes of SC agar and incubated at 28°C for up to 14 days. When well sporulated, the slopes could be stored at 4°C for several weeks. {Composition of SC agar (g/l) : Dextrin, 10.0 ; $KH_2PO_4$, 1.0 ; $MgSO_4.7H_2O$, 1.0 ; $(NH_4)_2SO_4$, 1.0 ; NaCl, 1.0 ; $CaCO_3$, 4.0 ; Oxoid No.1 Agar, 10.0 ; Trace Element Solution, 1 ml ; deionised water added to 1 litre, pH 7.0 , [Composition of Trace Element Solution : $FeSO_4.7H_2O$, 0.1 g ; $MnSO_4.H_2O$, 0.1 g ; $ZnSO_4.7H_2O$, 0.1 g ; deionised water added to 100 ml]}

Sterile deionised water (5 ml) was added to a well .sporulated slope and the surface growth carefully scraped off with a sterile needle to produce a suspension. The suspension was then used to inoculate 100 ml of seed medium in a 500-ml Erlenmeyer flask. [Composition of seed medium (g/l) : Tryptone Soya Broth (Oxoid CM 129), 30.0 ; glycerol, 10.0 ; deionised water added to 1 litre, pH 7.0]

The flasks were incubated at 28°C and 250 rpm on a gyratory shaker for about 64 hr. The seed culture was used to inoculate the production medium (5 ml inoculum per 100 ml

growth medium). [Composition of production medium (g/l) : glycerol, 20.0 ; Soyalose 105 (Central Soya), 15.0 ; $K_2HPO_4$, 1.0 ; $CoCl_2.6H_2O$, 0.01 ; deionised water added to 1 litre, pH 7.0]

The flasks were incubated at 27°C and 250 rpm on a gyratory shaker for 4 days. Mycelia were separated from the culture fluid by centrifugation at 9000 rpm (14400 *g*) for 40 min using a Beckman J2-21 centrifuge with a JA-10 rotor.

The culture filtrate was analysed on a Hypersil $C_{18}$ column, 250 x 10 mm, eluted isocratically with Milli-Q water at a flow rate of 2 ml/min. Eluate was monitored at 230 nm and peaks were collected as separate fractions. Each fraction (100 µl) was introduced into 8-mm wells made in both sets of *Bacillus* bioassay plates. The plates were incubated at 37°C for 12 hr before the diameters of inhibition zones were measured.

### 7.3.4 Bioassay for valclavam (37)

The holed-plate assay method, using *E. coli* ATCC 12435 as the test organism, was employed for the detection of valclavam (37) in culture filtrates. Bioassay plates were prepared by Mr John Keeping in an aseptic manner as follows :

Cultures of *E. coli* ATCC 12435 were maintained as glycerol freezes stored at -80°C. Inoculum for the bioassay plates was grown up overnight at 37°C as 5 ml agar slants on Oxoid Nutrient Broth (CM 1). The inoculum (4 ml) was diluted into 400 ml of a modified minimal agar, maintained at 45°C. [Composition of medium (g/l) : $K_2HPO_4$, 7.0 ; $KH_2PO_4$, 3.0 ; NaCl, 0.1 ; $MgSO_4.7H_2O$, 0.1 ; $(NH_4)_2SO_4$, 1.0 ; glucose (autoclaved separately), 4.0 ; Oxoid No.1 Agar, 10.0 ; deionised water added to 1 litre, pH 7.2]

Penicillinase (0.4 ml) (DIFCO Laboratories, Bacto Penase Concentrate, 10 000 000 IU/ml) was added to the diluted bacterial suspension, followed by 2.0 ml of cephalosporinase (P-99 β-lactamase from *Enterobacter cloacae*, 2 mg/ml). *DL*-Methionine (4 ml) (BDH, 40 mg/ml in water) was then added to give the methionine-positive plates. The bacterial agar suspension, with all the additives, was poured into 8-cm Petri dishes (18 ml per dish) and allowed to solidify. When stored at 4°C, the plates were usable for up to 2 weeks. The methionine-negative plates were prepared in the same manner as above except that no methionine was added. For detecting valclavam (37), 100 µl of the sample solution was loaded into 8-mm wells (made with a cork-borer) in both sets of plates. The loaded plates were then incubated at 37°C for 12 hr.

### 7.3.5 Fermentation of *S. antibioticus* ssp *antibioticus* Tü 1718 for the production of valclavam

Cultures of *S. antibioticus* were maintained as glycerol freezes stored at -80°C. The seed flasks were inoculated directly from the glycerol freezes, using about 1 ml for every 100 ml of seed medium in a 500-ml Erlenmeyer flask. [Composition of seed medium (g/l) : *D*-mannitol, 30.0 ; soybean meal (BDH), 20.0 ; glycerol 20.0 ; deionised water added to 1 litre, pH 7.2]

These flasks were incubated at 27°C and 250 rpm on a gyratory shaker for 3 to 4 days.

When a heavy growth was obtained, as judged by microscopic examination of the mycelia, 5 ml of the seed culture was used to inoculate each 100 ml aliquot of production medium in 500-ml Erlenmeyer flasks. The production medium was identical to the seed medium.

The flasks were incubated at 27°C and 250 rpm on a gyratory shaker for 5 to 6 days. Mycelia were separated from culture fluid by centrifugation at 9000 rpm (14400 *g*) for 40 min using a Beckman J2-21 centrifuge with a JA-10 rotor.

### 7.3.6 Isolation of valclavam (37) and Tü 1718B (111)

A 6-day old culture broth of *S. antibioticus* ssp *antibioticus* Tü 1718 was pooled from 100-ml cultures grown in separate 500-ml Erlenmeyer flasks. The mycelia was removed by centrifugation at 28000 *g* for 1 hr. The culture supernatant was clarified by filtration through a 0.22 µm PVDF membrane. The culture filtrate was then lyophilised to yield a brown flaky residue which was suitable for long-term storage. The first chromatographic step utilised a preparative $C_{18}$ column (Rainin, 250 x 21.4 mm), eluting with the acetonitrile-phosphate gradient listed in Table 7.5.

Gradient :

| Time | Flow (ml/min) | %A | %B |
|---|---|---|---|
| Initial | 12 | 100 | 0 |
| 36 | 12 | 80 | 20 |
| 43 | 12 | 0 | 100 |
| 50 | 12 | 0 | 100 |
| 58 | 12 | 100 | 0 |
| 75 | 12 | 100 | 0 |

Solvent A : 0.205 mM $Na_2HPO_4$, 0.140 mM $KH_2PO_4$, pH 7.0 Solvent B: acetonitrile
Column : Dynamax-300A $C_{18}$ 250 x 21.4 mm, 5 µm UV detection : 210 nm
Sample load : ~ 50 mg in 1 ml solvent A System configuration: As in Section 4.6.1

**Table 7.5** HPLC conditions for crude lyophilisate of culture filtrate of *S. antibioticus*

The valclavam peak ($R_t$ = 11.2 min) was collected and lyophilised to a light brownish powder. Further purification was achieved with a second reversed phase HPLC, employing identical conditions as in Table 7.5 except that methanol was used as the organic modifier instead of acetonitrile. The valclavam peak was again collected as a single fraction ($R_t$ = 13.2 min). The identity of the collected fraction was confirmed by bioassay and NMR.

When isolation of Tü 1718B was desired, the valclavam fraction from the first HPLC was adjusted to pH 8 - 9 and left overnight (*i.e.* > 12 hr) at 37°C. More than 80% of valclavam would have degraded to Tü 1718B by then. The sample was lyophilised to a powdery residue for storage. The lyophilisate was dissolved in 0.5 ml of 10 mM ammonium bicarbonate, pH 8, and used for reversed phase HPLC [Hypersil $C_{18}$ 250 x 10 mm, 10 mM ammonium bicarbonate (pH 8), Flow = 4 ml/ min, Detect at 210 nm] where Tü 1718B eluted at 6.5 min.

### 7.3.7 HMQC and HMBC experiments

The experiments were performed by Dr Timothy Claridge on a Bruker AM500 spectrometer operating at 500.13 and 125.77 MHz for proton and carbon respectively. Samples were estimated to be < 1 mg in quantity, and were prepared in 0.4 ml of 99.96% $D_2O$ with TSP added as internal reference. The probe temperature was held at 311 K and all 2D $^1H$-$^{13}C$ correlation experiments were performed on a non-spinning sample. Experiments were performed on both basic (from HPLC purification) and acidified samples.

Heteronuclear Multiple-Quantum Coherence (HMQC) shift correlation : HMQC spectra were acquired according to the method of Bax and Subramanian,[229] in which a BIRD sequence acts to presaturate protons not coupled to $^{13}C$, prior to the HMQC sequence. The $1/2J$ delays for the evolution of one-bond H-C couplings were set to 3.5 ms ($J$ = 143 Hz). The relaxation delay, and the period between the BIRD sequence and the HMQC sequence were 300 ms. The acquisition time $t_2$ was 205 ms (2K data points) and $t_1$ was 25.6 ms (256 increments of 128 scans and 2 dummy scans). Spectra were acquired in phase sensitive mode with low power presaturation of residual HOD and without broad-band $^{13}C$ decoupling during acquisition.

Heteronuclear Multiple-Bond Connectivity (HMBC) experiments : HMBC spectra were acquired according to the method of Bax and Summers,[230] with a low-pass $J$-filter to suppress one-bond correlations. The delay in the low-pass filter was set to 3.4 ms ($^1J$ = 147 Hz) with the delay for the evolution of long-range couplings set to 60 ms ($^nJ$ = 8.3 Hz). The acquisition time $t_2$ was 350 ms (2K data points) and $t_1$ was 34.2 ms (1024 increments of 64 scans and 2 dummy scans, 1.5 s relaxation delay between scans). Spectra were acquired in the absolute value mode.

Processing of HMQC and HMBC spectra made use of a phase-shifted sine-bell window for $t_2$ data, with a single zero filling and no window for $t_1$ data.

### 7.3.8 Periodate oxidation of Tü 1718B

To a solution of Tü 1718B (0.5 mg) in 2 ml of a 2:2:1 mixture of water:acetone:acetic acid, 2 mg of sodium periodate was added and stirred at 40 - 50 °C for 24 hr. The end product was partitioned into ethyl acetate and aqueous extracts and evaporated separately. Both residues were submitted for amino acid analysis to detect *L*-4-hydroxyproline.

### 7.3.9 Imidazole derivatisation assay for valclavam (37)

Our method was adapted from that reported by Bird[224] for assaying clavulanates. Valclavam (1.25 mg in 1 ml $H_2O$) was added with 40 µl of 2 M imidazole-HCl, pH 6.8 and incubated at 37°C for 15 min. The reaction mixture was analysed by injecting 25 µl into a $C_{18}$ column (250 x 4.6 mm) eluting isocratically with 0.1 M $NaH_2PO_4$ in 6% (v/v) methanol, pH 3.2 at 2 ml/ min. The eluate was monitored at 312 nm. Clavulanic acid (31) was used as a control for the feasibility of the reaction.

# CHAPTER 8

# BIOSYNTHETIC PRECURSORS FOR VALCLAVAM

## 8.1 Introduction

The conventional research methodology employed for a biosynthetic project was outlined in Section 1.4. Once an acceptable level of the secondary metabolite can be obtained reproducibly from a chosen producer, speculations about its biosynthesis are made which collectively serve as the working hypothesis for designing further experiments. The preliminary studies carried out in Chapter 7, together with knowledge of secondary metabolism in *Streptomyces* bacteria, provided a basis for formulating biosynthetic hypotheses for valclavam (37) and laid the groundwork for studies described in this chapter. Our working hypotheses for the biosynthesis of valclavam (37) are given below.

(i) Hypothesis A

As mentioned in Section 7.2.5, the structure of valclavam (37) can be divided into three biosynthetic units (see Figure 7.12). These are (a) a $C_3$ unit, comprising the β-lactam carbons, which probably arised from a $C_3$ intermediate in the primary metabolic pool, (b) a $C_5$ unit, comprising the N-terminal *L*-valyl residue, most probably derived intact from *L*-valine, and (c) a $C_6$ unit between the $C_3$ and $C_5$ units, possibly derived from lysine. The functionalisation at C-2 and C-9 of the $C_6$ unit could conceivably result from the action of hydroxylases, consistent with the wide occurrence of hydroxylated amino acyl moieties in *Streptomyces* metabolites.[16, 235] The analogy might be stretched further by postulating that these hydroxylases are ferrous-dependent dioxygenases, consistent with the occurrence of such enzymes in other β-lactam pathways[236] (see Sections 1.8 and 2.2).

(ii) Hypothesis B

Townsend *et al*[237] recently demonstrated that *L*-ornithine is a precursor of clavam-2-carboxylate (96), being incorporated into the clavam in the same regiochemical sense as clavulanic acid (31) (Figure 8.1). Moreover, the relatively advanced biosynthetic intermediate to clavulanic acid (31), proclavaminic acid (34), was incorporated into clavam-2-carboxylate (96) with equal efficiency .

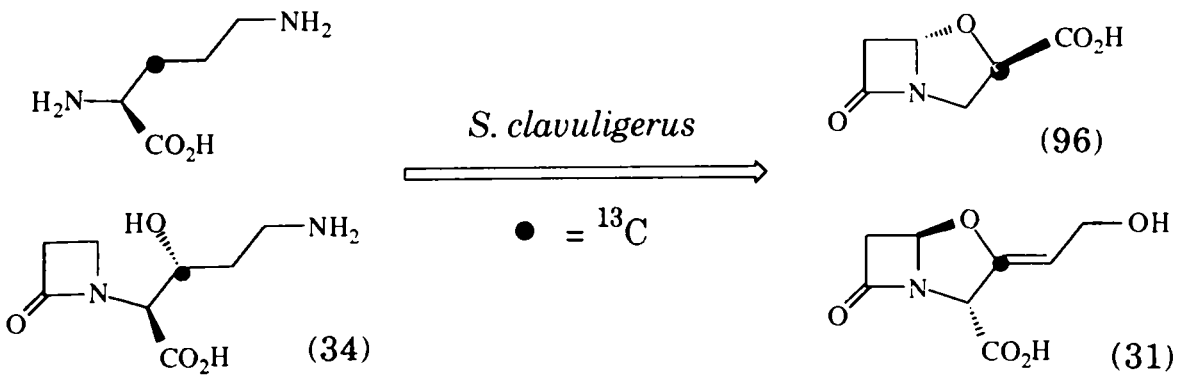



**Figure 8.1** Townsend's incorporation experiments with clavam-2-carboxylate (96)

It was concluded that the two clavams, despite their biologically significant stereochemical difference at C-5, share the early part of their biosynthetic pathways at least up to proclavaminic acid (34). The authors went further to suggest that all the other clavams in Figure 6.1, including valclavam (37), may be similarly derived from *L*-ornithine. Taking into consideration the latest evidence that *L*-arginine (33) (rather than *L*-ornithine) is a more direct precursor to clavulanic acid (31),[34] we have modified and expanded Townsend's proposal to the hypothesis depicted in Figure 8.2.

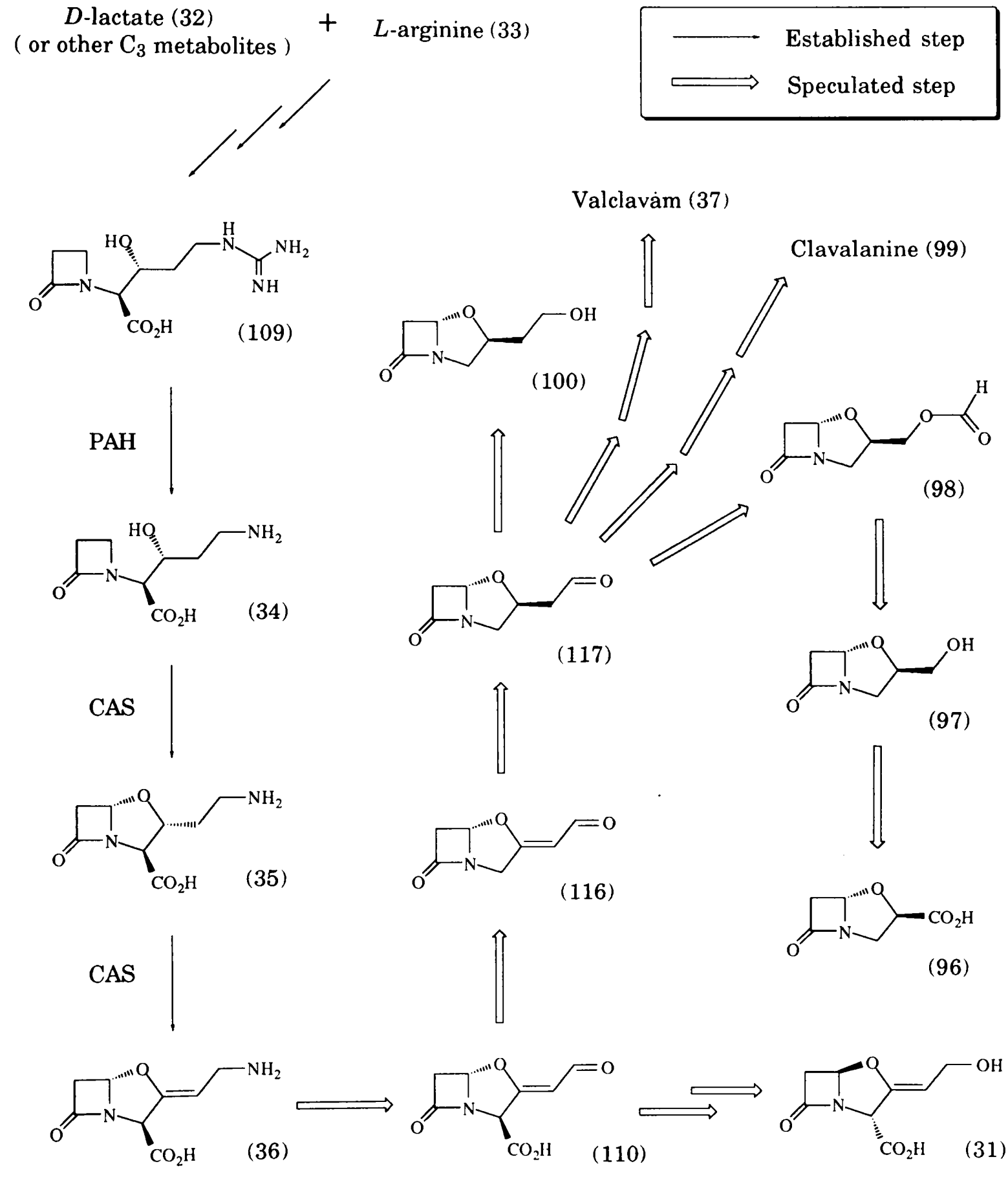



**Figure 8.2** Hypothesis B for the biosynthesis of clavams in *Streptomyces* bacteria

Thus, we proposed that the 5*S*-clavams share an even larger part of their biosynthetic pathways with clavulanic acid (31), up to and including the hypothetical α,β-unsaturated aldehyde (110). The latter compound has previously been proposed to be an intermediate in the remarkable 'double inversion' (at C-3 and C-5) of clavaminic acid (36) to clavulanic acid (31) (Figure 6.6).[217] It can conceivably arise from clavaminic acid (36) through oxidative deamination, a common reaction in amino acid metabolism. The α,β-unsaturated carbonyl group can then serve as an electron sink in the decarboxylation of (110) as shown in Figure 8.3.

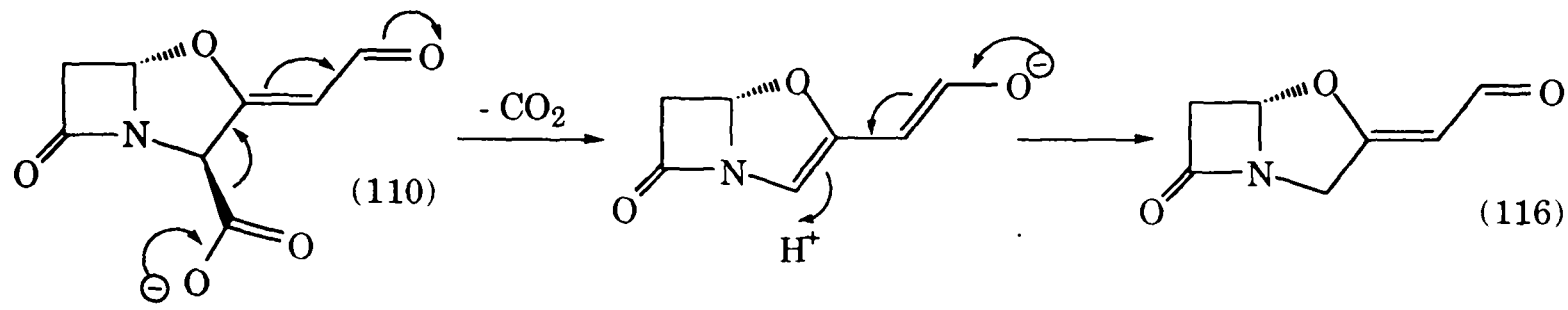
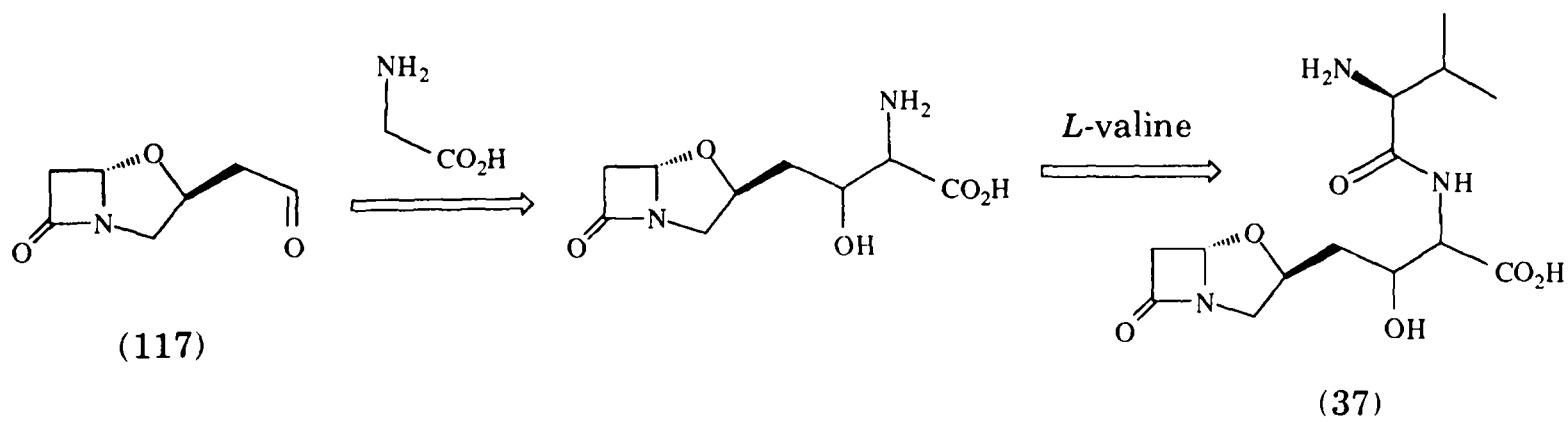



**Figure 8.3** Decarboxylation of (110) facilitated by its α,β-unsaturated carbonyl group

The reduction of (116) at the double bond would generate the saturated aldehyde (117) which may be the point of divergence to all the 5*S*-clavams in *S. clavuligerus* and *S. antibioticus*. A further reduction of (117), now at the aldehyde group itself, immediately produces 2-(2-hydroxyethyl)clavam (100). A Baeyer-Villiger oxidation, a reaction encountered in biological systems,[238, 239] would yield 2-formyloxymethylclavam (98). Ester hydrolysis of (98) would give 2-hydroxymethylclavam (97) which can be ultimately oxidised to clavam-2-carboxylate (96). All the chemical transformations alluded to have well established precedents in biological systems.[1]

Valclavam (37) contains two additional carbons in its $C_6$ unit compared with (117). We hypothesised that these extra carbons originate from an aldol-type condensation with glycine (Figure 8.4). Hypothesis A can be simply tested by whole-cell feeding experiments with potential precursors, whereas the more elaborate Hypothesis B offers scope for other investigations such as enzyme detection and whole-cell feeding experiments with potential mid-stage intermediates. The following section describes the outcome of our feeding experiments.




(117) (37)

**Figure 8.4** Hypothetical route from (117) to valclavam (37)

## 8.2 Results and Discussion

The elucidation of a biosynthetic pathway often begins with whole-cell feeding experiments to identify the earliest precursors to a natural product. The producer organism is fed, in separate experiments, with an array of specifically labelled compounds which are hypothetical precursors of the natural product. The classical radiolabelling method (using $^{14}C$ and $^{3}H$) is still used extensively but labelling methods employing stable isotopes (*e.g.* $^{13}C$, $^{2}H$, $^{18}O$, $^{15}N$) are becoming increasingly popular. Stable isotope incorporation can be monitored with spectroscopic techniques such as NMR and mass spectrometry. The extent and regiochemistry of incorporation of these exogenous substances are then used to draw conclusions about their roles in the biosynthesis of the natural product. In certain cases, the feeding of stereospecifically labelled precursors can reveal the mechanism of the biosynthetic enzymes even before they are isolated.

Since secondary metabolites are produced during a specific period of the growth of an organism,[3, 83] the timing of feeding and harvesting is important to maximise the incorporation for a true precursor. If a true precursor is fed well ahead of the production phase, most of it will be channelled into other metabolic pathways (*e.g.* amino acids for protein synthesis), leading to sub-optimal levels of incorporation. On the other hand, late feeding can also result in diminished incorporation because (i) the early steps of the pathway might have lost their full efficiency and (ii) most unlabelled product formation has already occurred. Another potential cause for low incorporation of a true precursor is poor cellular uptake, a problem usually difficult to overcome. In harvesting, it is most important to avoid undue delay which might give rise to product degradation. Thus, based on the time course of valclavam (37) production obtained in Chapter 7, feeding of labelled compounds was routinely performed between the 60th and 72nd hour after inoculation, whilst the culture filtrate was routinely collected between the 5th and 6th day. Two different fermentation media were used for the feeding experiments - a glycerol-based medium as described in Section 7.3.5 and the other without glycerol. It has been observed,[225] during biosynthetic studies on clavulanic acid (31), that a glycerol-based medium can dramatically decrease the amount of $C_3$ precursor incorporated into clavulanic acid (31). This is due, at least in part, to the rapid flux through and equilibrium among related primary metabolites. To prevent this problem with our studies on valclavam (37), we omitted glycerol from feeding experiments with $C_3$ compounds, despite our findings (see Chapter 7) that the addition of glycerol led to consistent and enhanced production of valclavam (37). Valclavam (37) titre was assayed on a daily basis during the feeding experiments to ensure that an acceptable level of production was obtained before harvesting. The entire data accrued in the project is compiled into Table 8.1.

Since val-fragment (111) is much more stable than valclavam (37), it was the preferred choice for isolation and subsequent analysis for isotope incorporation. In cases where incorporation into val-fragment (111) was observed, the labelled positions were further localised

to the $C_6$ or $C_5$ unit. However, where the $C_3$ unit was suspected to be the site of incorporation, valclavam (37) had to be isolated intact.

| Compound fed[(a)] | No. of expt | Analysed as valclavam (V) or val-fragment (F) | Average level of incorporation (%) |
|---|---|---|---|
| [U-$^{14}$C]-*L*-valine | 2 | F | 6.75 |
| [U-$^{14}$C]-*L*-ornithine | 2 | F | 2.87 |
| [U-$^{14}$C]-*L*-arginine[(b)] | 2 | F | 18.63 |
| [$^{14}$COOH]-*L*-arginine | 1 | F | 0.67 |
| [$^{14}CH_3$]-*L*-methionine | 2 | F | 5.79 |
| [U-$^{14}$C]-*L*-lysine | 2 | F | 0.88 |
| [U-$^{14}$C]-glycine | 2 | F | 0.93 |
| [U-$^{14}$C]-*L*-proline | 1 | F | 0.75 |
| [U-$^{14}$C]-*L*-glutamate[(b)] | 1 | F | 0.60 |
| [U-$^{14}$C]-*L*-aspartate | 1 | F | 0.73 |
| [U-$^{14}$C]-glycerol | 2 | V | 5.43 |
| [U-$^{14}$C]-glycerol | 1 | F | 0.66 |
| [U-$^{14}$C]-*L*-lactate | 1 | V | 5.15 |
| [U-$^{14}$C]-*D*-lactate | 1 | V | 4.47 |
| [1-$^{14}$C]-pyruvate | 1 | V | 1.96 |
| [1,2,2'-$^{13}C_3$]-proclavaminate | 1 | F | 0 |

(a) All radiolabelled compounds were fed at 0.1 mmole per 100-ml fermentation at a specific activity of 100 μCi/mmole. $^{13}C_3$-PCA was fed at 0.0625 mmole per 100-ml fermentation.

(b) The specific activity of val-fragment (111) was multiplied by five-quarters to take into account the loss of 1-$^{14}$COOH. The guanidino group of *L*-arginine was not labelled.

**Table 8.1** Summary of whole-cell feeding experiments with *S. antibioticus* ssp *antibioticus* Tü 1718

The following conclusions can be inferred from Table 8.1:

### (i) Baseline level of incorporation

The range of values obtained for incorporation of radiolabelled compounds suggests that the baseline for these experiments varies between 0.6 and 0.9 %. This fluctuation is probably attributable to background radiation in the measurements as well as low non-specific incorporation into the metabolite. Assuming competent cellular uptake of common amino acids, which is justifiable for a bacterium like *S. antibioticus*,[228] it is concluded that *L*-lysine, glycine, *L*-proline, *L*-glutamate and *L*-aspartate are not precursors for valclavam (37). The failure to

observe *L*-lysine incorporation immediately implies that Hypothesis A is untenable, *i.e.* the $C_6$ unit of valclavam (37) is not derived directly from this amino acid.

### (ii) Origin of the $C_5$ unit

The significant level (6.75%) of incorporation of *L*-valine into val-fragment (111) suggests its direct utilisation as a building block of valclavam (37). This is not surprising in view of the intact *L*-valyl residue in the $C_5$ unit. Interestingly, *L*-valine is also the progenitor of penicillins, cephalosporins and cephamycins (see Figure 1.1) but its fate in these β-lactams is far from intact.

### (iii) Origin of the $C_3$ unit

By analogy to the biosynthesis of clavulanic acid (31), it is speculated that the β-lactam carbons of valclavam (37) were derived from any of several $C_3$ compounds which are interconvertable through primary metabolism. These compounds include glycerol,[199, 200] pyruvate,[201] *D*-glycerate,[201] *D*-lactate[202] and β-hydroxypropionate.[203] The figures in Table 8.1 bore out our predictions for glycerol, *L*-lactate, *D*-lactate and pyruvate (*D*-glycerate and β-hydroxypropionate were not tested). The results for glycerol, where both valclavam (37) and val-fragment (111) have been analysed for radiolabel incorporation, unambiguously localise the site of labelling at the β-lactam ring. Since these feeding experiments were carried out in parallel under identical conditions, the data are perhaps a more accurate indication of the relative efficiency of utilisation compared with values in the literature. It is therefore probable that, relative to pyruvate (1.97%), glycerol (5.43%) and the lactates (4.47%, 5.15%) are metabolically closer to a putative $C_3$ compound that is directly committed into the valclavam (37) pathway. This was the first evidence of a partially shared pathway between valclavam (37) and clavulanic acid (31). However, the identity of the direct precursor for the $C_3$ unit remains an elusive problem.

### (iv) Origin of the $C_6$ unit

With the elimination of *L*-lysine as a possible precursor for the $C_6$ unit, we were left with only one working hypothesis. It was therefore encouraging to observe the exceptionally high incorporation of [U-$^{14}$C]-*L*-arginine into val-fragment (111). Moreover, the lack of radiolabel incorporation from [$^{14}$COOH]-*L*-arginine argues for a regiochemistry of incorporation that is identical to that in clavulanic acid (31) biosynthesis, followed by the decarboxylation of (110) to (116) as speculated in Hypothesis B. The lower level of incorporation of [U-$^{14}$C]-*L*-ornithine is consistent with the fact that it has to be converted, via the urea cycle, into *L*-arginine which is the point of entry into the clavam pathway. However, the wide disparity in the incorporation levels of these amino acids was never observed in studies on the clavulanic acid (31) pathway (Table 8.2). It is possible that this discrepancy could be attributed to the difference in producer organisms. Perhaps, the urea cycle (or the cellular uptake of *L*-ornithine) is much less efficient in *S. antibioticus* compared to the corresponding process(es) in *S. clavuligerus*.

| Compound fed | % level of incorporation into | | |
|---|---|---|---|
| | clavulanic acid (31) | | valclavam (37) |
| | Townsend[206] | Martin[207] | From Table 8.1 |
| [U-$^{14}$C]-*L*-proline | 0.11 | - | 0.75 |
| [U-$^{14}$C]-*L*-glutamate | 0.24 | 0.37 | 0.60 |
| [U-$^{14}$C]-*L*-ornithine | 2.11 | 3.61 | 2.87 |
| [U-$^{14}$C]-*L*-arginine | 1.61 | 6.32. | 18.63 |

**Table 8.2** Comparison of results of feeding experiments by different groups

The lack of incorporation of [U-$^{14}$C]-glycine was disappointing. We decided to experiment with simpler and commoner carbon donors and, to our surprise, found a significant incorporation of [$^{14}CH_3$]-*L*-methionine into val-fragment (111). This suggests that the two additional carbons in the $C_6$ unit originate from the methyl groups of *L*-methionine, presumably via the one-carbon donor *S*-adenosylmethionine. Though the chemistry from two methyl groups to the α-amino acid group in valclavam (37) seems complicated and biologically excessive, it is somewhat analogous to the findings for the biosynthesis of thienamycin (12), where the hydroxyethyl side chain was also derived from the methyl groups of *L*-methionine[240, 241] (Figure 8.5).

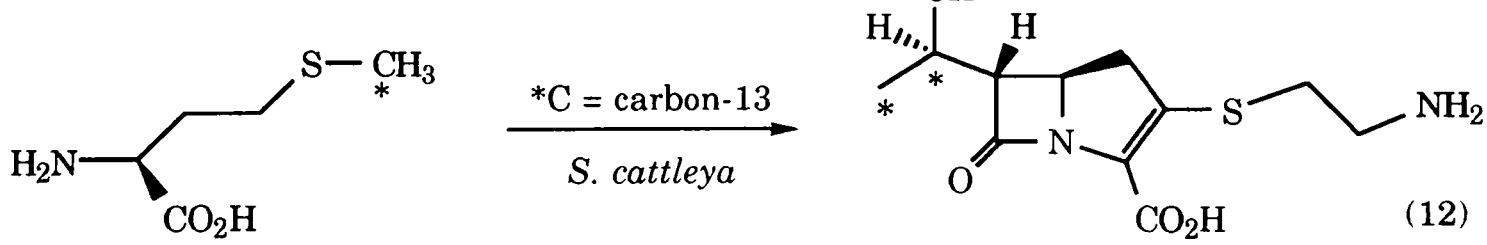
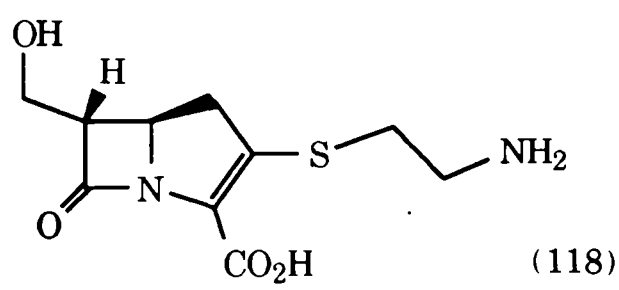
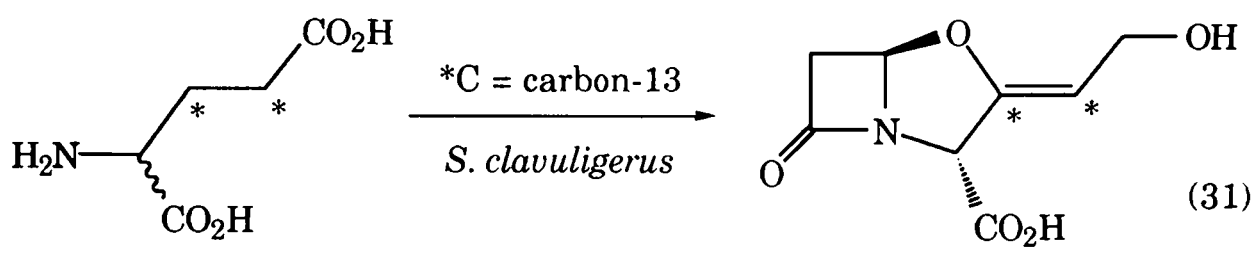



**Figure 8.5** Incorporation of methionine into thienamycin (12)

It would be ideal to demonstrate unequivocally by stable isotope NMR (after feeding with [$^{13}CH_3$]-*L*-methionine) that the atoms C-5 and C-6 of val-fragment (111) are indeed the positions labelled by the methyl groups. This line of investigation is currently being pursued in a continuing project. It is not known if each carbon atom is added individually to the putative $C_3$-$C_4$ intermediate [equivalent to the hypothetical α,β-unsaturated aldehyde (117)] or if a $C_2$ molecule (derived from methyl groups of *L*-methionine) is utilised. In the case of thienamycin (12), the carbon atoms are probably added sequentially to a parent carbapenem nucleus, an idea supported by the isolation of northienamycin (118) from culture broths of *S. cattleya*.[242]



At this stage, the entire carbon skeleton of valclavam (37) has been tentatively assigned to various precursors, but the result for [U-$^{14}$C]-*L*-glutamate warrants further comment. *L*-Glutamate is metabolically linked to the amino acids of the urea cycle as shown in Figure 6.5. Consequently, this amino acid has the potential to be incorporated into clavulanic acid (31) and valclavam (37). Indeed, Elson *et al.*[205] observed the specific incorporation of [3,4-$^{13}C_2$]-*DL*-glutamate into clavulanic acid (31) (Figure 8.6), though later work by the groups of Martin[207] and Townsend[206] independently found [U-$^{14}$C]-*L*-glutamate to be incorporated much less efficiently than [U-$^{14}$C]-*L*-ornithine and [U-$^{14}$C]-*L*-arginine (see Table 8.2).




(31)

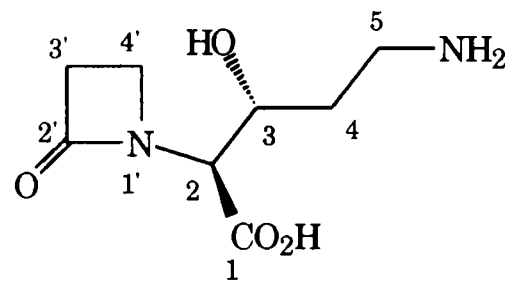
**Figure 8.6** Incorporation of [3,4-$^{13}C_2$]-*DL*-glutamate into clavulanic acid (31)

Our own results show only background incorporation of [U-$^{14}$C]-*L*-glutamate. This may be partly due to the low efficiency of the urea cycle in *S. antibioticus*, as suggested earlier to account for the wide disparity in the incorporation of *L*-ornithine vs. *L*-arginine. In summary, our results show that the $C_6$ unit of valclavam (37) is derived basically from *L*-arginine and the methyl groups of *L*-methionine.

### (v) Feeding of [1,2,2'-$^{13}C_3$]-proclavaminic acid (119)

1 = 2 = 99% $^{13}$C

2' = 92% $^{13}$C




(119)

Our tentative conclusions concerning the identity of the precursors for the $C_3$ and $C_6$ units of valclavam (37) suggest a high probability that the valclavam (37) pathway in *S. antibioticus* shares some of the early steps with the clavulanic acid (31) pathway in *S. clavuligerus*. Hypothesis B can now be tested further by whole-cell feeding experiments using known intermediates in the clavulanic acid (31) pathway. A comprehensive set of feeding experiments might even establish the exact point of divergence of the pathways in the two *Streptomyces* species.

It has been shown that PCA (34), an established intermediate to clavulanic acid (31) (see Figure 6.3), is poorly taken up by the mycelia of *S. clavuligerus*, resulting in only 3.5% incorporation which is deemed very low for such a mid-stage intermediate.[36] However, the authors claimed their calculation revealed that virtually all the PCA transported into the cells

was utilised for clavulanic acid (31) biosynthesis. Subsequently, Townsend[237] observed an even lower level of incorporation (1%) of PCA into clavulanic acid (31), with an identical level of incorporation into clavam-2-carboxylate (Figure 8.1).

We undertook the feeding of [1,2,2'-$^{13}C_3$]-PCA (119), a gift of SmithKline Beecham Pharmaceuticals, to fermentation of *S. antibioticus* and detected no $^{13}C$ enrichment in the isolated val-fragment (111) either by NMR or mass spectrometry. Poor cellular uptake of (119) could have been responsible for the lack of incorporation even if the compound is a true intermediate to valclavam (37). Indeed, the structure of PCA still retains a large part of the structure of ornithine, a molecule which we have postulated to be transported into *S. antibioticus* less efficiently than into *S. clavuligerus*. Hence, at this stage, it is not possible to draw any conclusion as to the extent of overlap between the pathways for valclavam (37) and clavulanic acid (31).

### (vi) Summary of precursor incorporation

The early precursors for various portions of valclavam (37) are depicted in Figure 8.7:

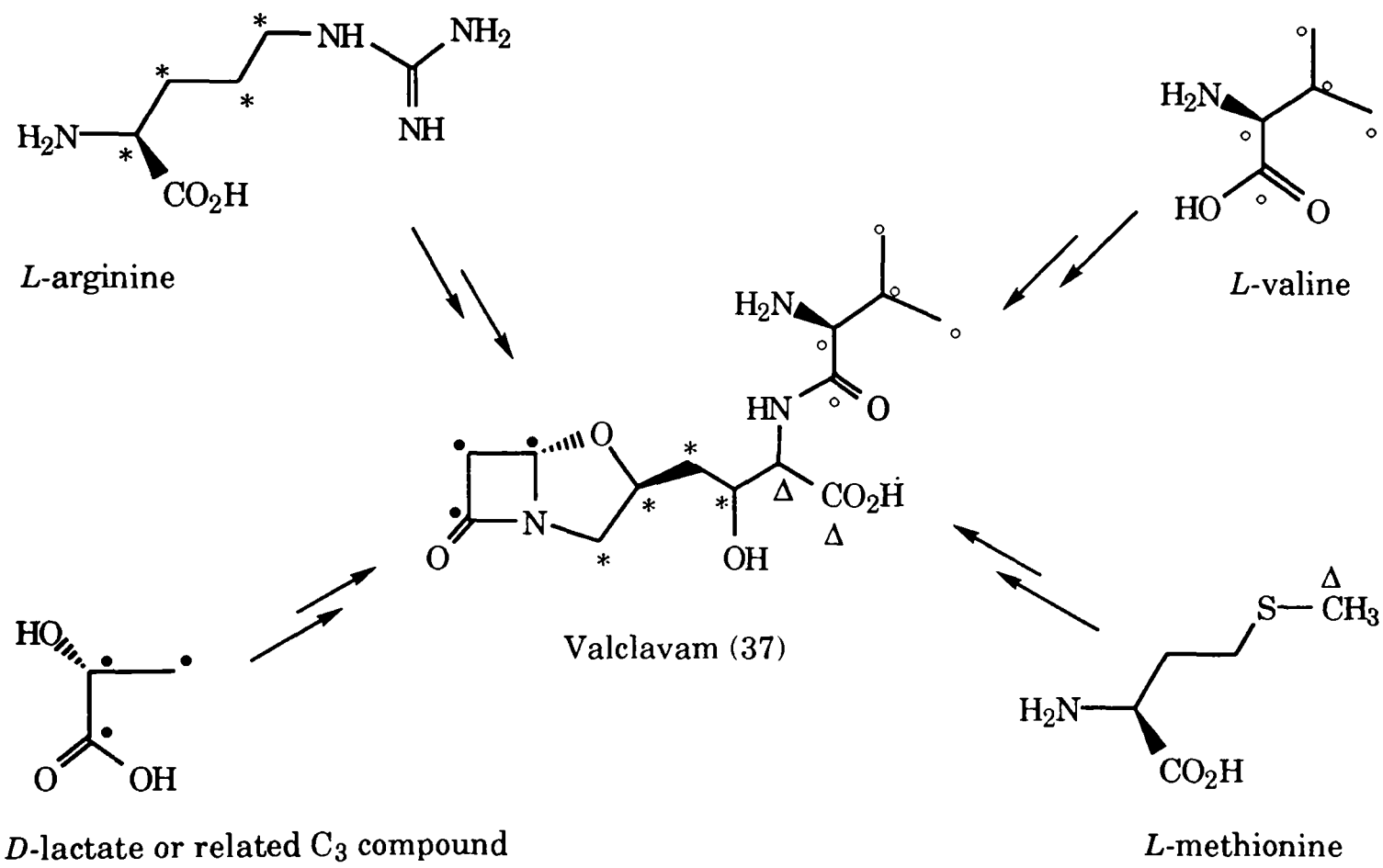



**Figure 8.7** The building blocks of valclavam (37)

## 8.3 Experimental

### 8.3.1 Feeding experiments for *S. antibioticus* ssp *antibioticus* Tü 1718

Fermentations were carried essentially according to Section 7.3.5, except that glycerol was omitted from the production medium used in feeding experiments with labelled $C_3$ compounds.

Radiochemicals were purchased from the following suppliers: Amersham International ([U-$^{14}$C]-*L*-valine, [U-$^{14}$C]-*L*-ornithine, [U-$^{14}$C]-*L*-arginine, [U-$^{14}$C]-*L*-lysine, [U-$^{14}$C]-glycerol), ICN Radiochemicals ([U-$^{14}$C]-*L*-lactate, [U-$^{14}$C]-*D*-lactate), New England Nuclear Products ([$^{14}CH_3$]-*L*-methionine, [U-$^{14}$C]-glycine, [U-$^{14}$C]-*L*-aspartate, [1-$^{14}$C]-pyruvate), Sigma ([U-$^{14}$C]-*L*-proline, [U-$^{14}$C]-*L*-glutamate) and American Radiolabelled Chemicals Inc. ([$^{14}$COOH]-*L*-arginine). These radiochemicals were diluted with a 'cold' sample to obtain a specific activity of 100 μCi/mmole (concentration = 0.1 mmole/ml). At the time of feeding (between the 60th and 72nd hour after inoculation), 1 ml of the diluted radiochemical was filtered through a sterile 0.22 μm filter and administered to 100-ml culture in an aseptic environment. A triplicate set of feedings was done for each labelled compound. Bioassay for valclavam (37) titre (throughout the feeding experiment) and product isolation (between the 5th and 6th day after inoculation) were performed as in Sections 7.3.4 and 7.3.6.

## 8.3.2 Determination of $^{14}$C incorporation

The HPLC-isolated valclavam (37) or val-fragment (111) was quantified by 500 MHz $^{1}$H NMR, based on the integral for the 6 methyl protons vs. the integral for the TSP signal at 0 ppm. Thus,

$$\mu\text{mole of product isolated} = \frac{9 \times \mu\text{mole of TSP added} \times \text{integral for 6 methyl protons}}{6 \times \text{integral for TSP at 0 ppm}}$$

Samples which have been quantified were mixed with Optiphase 'Safe' liquid scintillation cocktail (LKB Scintillation Products) and the radioactivity counted on a LKB-Wallac 1215 RackBeta Liquid Scintillation Counter operating in the $^{14}$C mode. The cpm values recorded were converted to nanoCurie (*n*Ci) by reference to a calibration plot for each labelled compound. The specific activity of the isolated product and its % incorporation of radiolabelled compound were calculated as follows:

$$\text{Specific activity of product isolated (μCi/mmole)} = \frac{n\text{Ci of radioactivity present (by LSC)}}{\mu\text{mole of product isolated (by NMR)}}$$

$$\%\ \text{Level of incorporation} = \frac{\text{Specific activity of product isolated (μCi/mmole)}}{\text{Specific activity of labelled compound (μCi/mmole)}} \times 100\ \%$$

Samples were purified by HPLC until the valclavam (37) or val-fragment (111) peak exhibited constant specific radioactivity.

### 8.3.3 Determination of $^{13}C$ incorporation

A significant incorporation of [1,2,2'-$^{13}C_3$]-PCA (119) into valclavam (37) would result in enrichment of the $^{13}C$ content at C-1 of val-fragment (111). This was investigated by the examination of the $^{13}C$ satellites for the 1-H signal of val-fragment (111) in a 1D proton spectrum. $^{1}H$ NMR was performed by Dr Timothy Claridge on a Bruker AMX 500 spectrometer running at 499.98 MHz. ESMS, as described in Chapter 4, was also used to determine possible $^{13}C$ enrichment in val-fragment (111). An authentic sample of val-fragment (111) from a normal fermentation, not fed with (119), was used as the control.

# CHAPTER 9

# BIOSYNTHETIC ENZYMES FOR VALCLAVAM

## 9.1 Enzymes in the clavulanic acid pathway

The results of Chapter 8 provided a strong impetus for further attempts towards the testing of Hypothesis B. As shown in Figure 8.2, it is speculated that two enzymes (*i.e.* CAS and PAH) in the clavulanic acid (31) pathway also play a direct role in the biosynthesis of valclavam (37). It would seem logical, therefore, that efforts should be made to look for similar enzymatic activities in *S. antibioticus* ssp *antibioticus* Tü 1718. CAS and PAH, the only enzymes to be discovered so far in the clavulanic acid (31) pathway, have been described briefly in Section 6.2. Whilst PAH has not been characterized in much detail due to its relative novelty,[209] several mechanistic studies have been performed on CAS and are described below.

Townsend *et al.*[243, 244] demonstrated the retention of deuterium at C-2 and C-3' of labelled PCA's (120/122) during their oxidative cyclisation/desaturation to the corresponding labelled clavaminic acids (121/123), indicating the non-participation of these sub-structures in the chemistry of CAS action (Figure 9.1). Further incubation experiments[245] with PCAs stereospecifically deuteriated at the C-4' position showed the loss of 4'(*S*)-atom only (Figure 9.2). Such a retention of configuration during an oxidative cyclisation process has precedents in the actions of IPNS and REXH.[32]

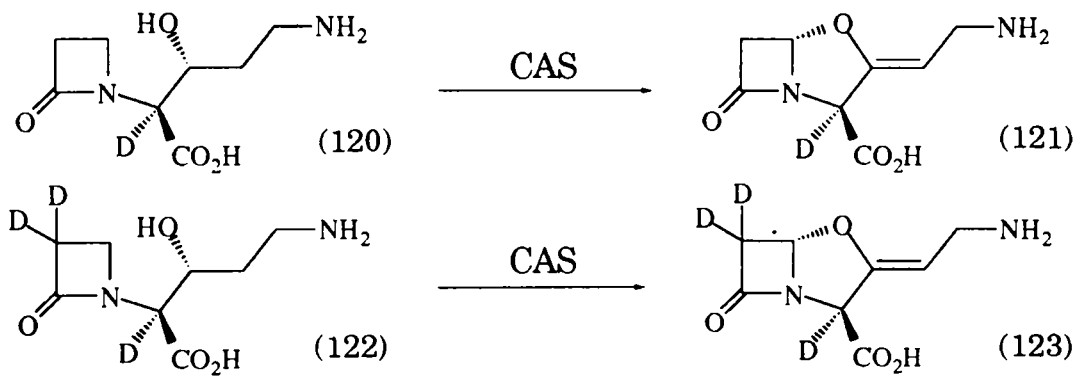



**Figure 9.1** Incubation of CAS with deuteriated proclavaminic acids (120/122)[243, 244]

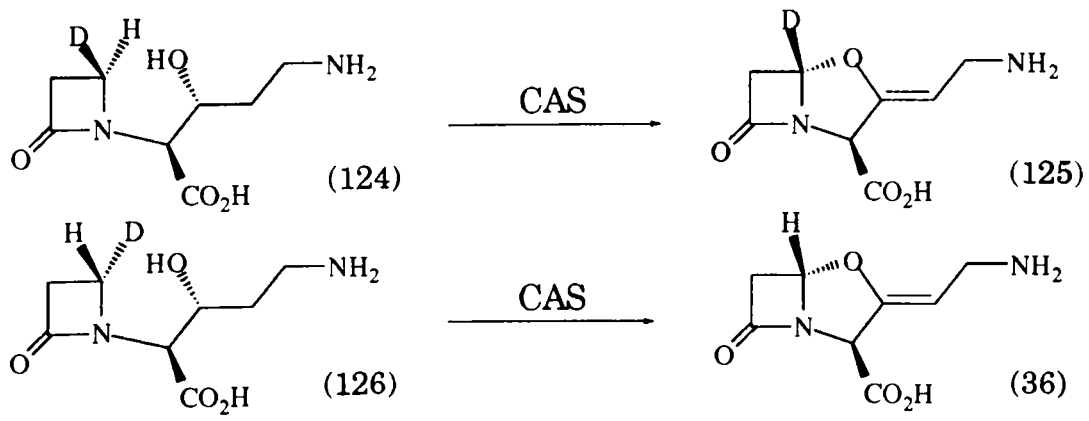



**Figure 9.2** Stereochemical retention during oxidative cyclisation by CAS[245]

As mentioned in Section 6.2, the hydroxylation of (108) to (109) in the clavulanic acid (31) pathway is an *in vitro*, and probably *in vivo*, activity of CAS. This reaction was subsequently shown to proceed with a high (> 95%) degree of retention of configuration at C-3 (Figure 9.3), consistent with the findings for other α-KG dependent dioxygenases.[246]

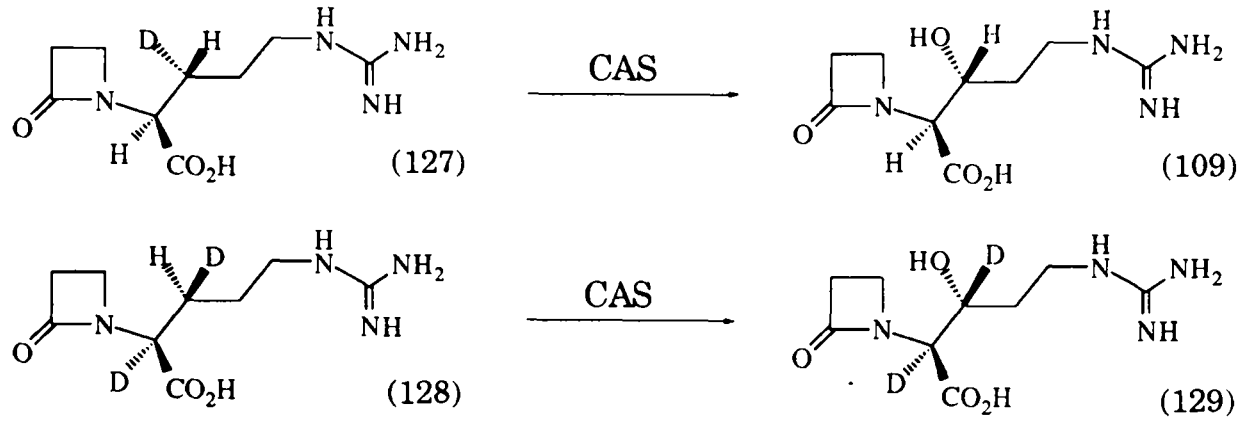



**Figure 9.3** Stereochemical retention during hydroxylation by CAS[246]

The above experiments demonstrate obvious mechanistic parallels between CAS and the other oxygenases (IPNS and REXH) in the wider context of β-lactam biosynthesis. Preliminary studies[247, 248] have shown that CAS might exhibit a narrower substrate specificity compared with IPNS[108] or REXH (see Chapter 4), being able to act on just a small number of the wide range of substrate analogues tested so far. This contrast in the degree of substrate specificity between CAS and IPNS/REXH may reflect differences in the extent of evolutionary progress made by these proteins. As shown in Figure 9.4, the alternative substrates of CAS included $N_\alpha$-acetyl-*L*-arginine (130),[76] $N_\alpha$-acetyl-*L*-ornithine (132),[76] dehydroxy-PCA (135)[132] and the γ-lactam analogue (137) of PCA (34).[144]

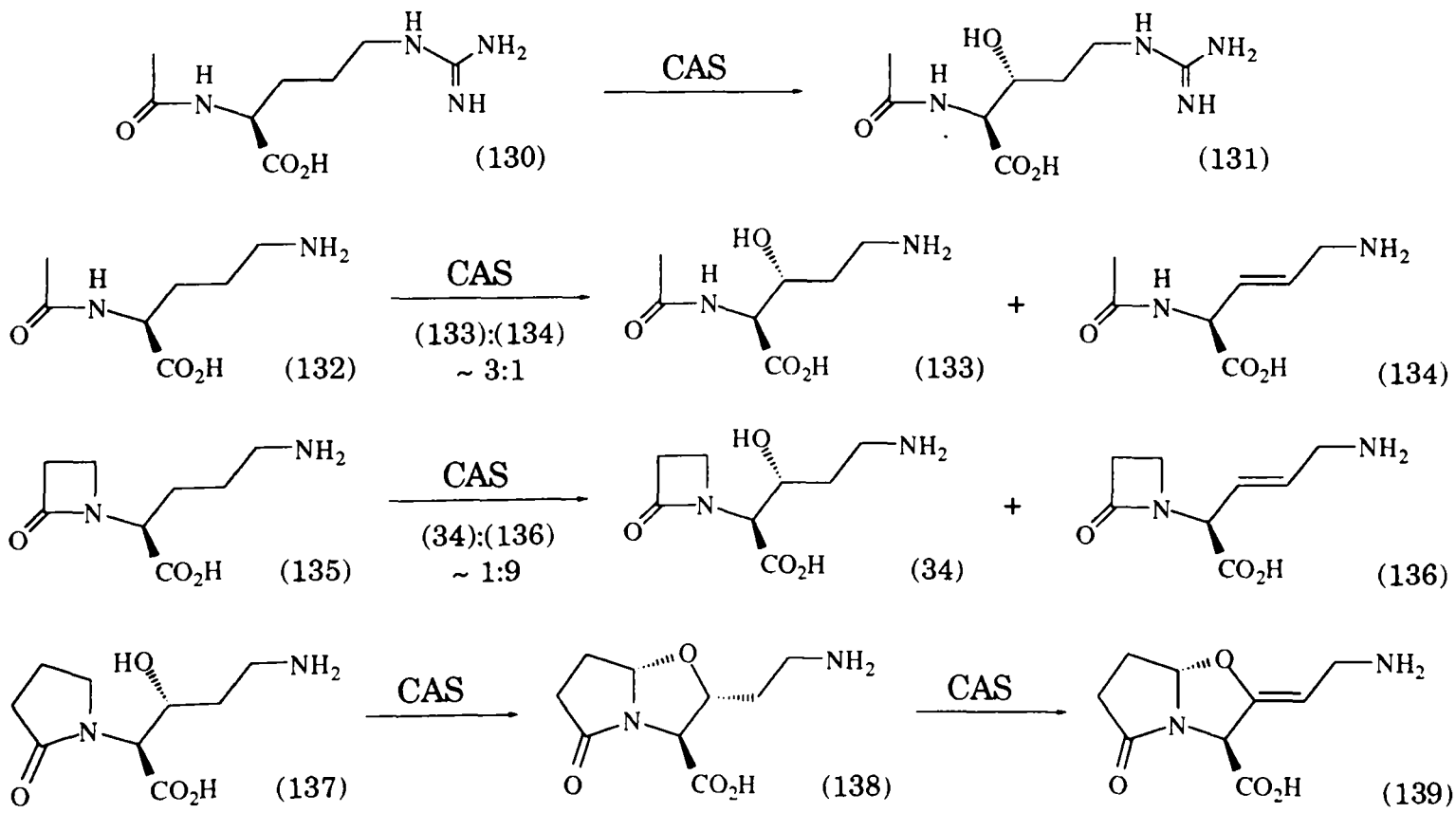



**Figure 9.4** Alternative substrates for CAS

The next section describes our attempts to seek out the activities of CAS and PAH in *S. antibioticus* ssp *antibioticus* Tü 1718. The work in this chapter was a collaboration with Dr Yoshiyuki Fujishima of the Oxford Centre for Molecular Sciences, University of Oxford.

## 9.2 Results and Discussion

The four methods listed below have been developed for the detection of CAS activity[76]:

(a) Direct HPLC determination of clavaminic acid (36) production from PCA (34)

(b) Imidazole derivatisation of clavaminic acid (36) and HPLC determination

(c) Hydroxylation of $N_\alpha$-acetyl-*L*-arginine (130) to (131) [see Figure 9.4] and HPLC

(d) Decarboxylation of [1-$^{14}$C]-α-KG and radioactivity counting of $^{14}CO_2$ production

Method (d) was unsuitable for our purpose due mainly to its low specificity in the context of our experiments with crude cell-free extract (CFE). Method (c) was adopted as the routine assay for CAS due to its following advantages over methods (a) and (b):

(i) a ready commercial supply of $N_\alpha$-acetyl-*L*-arginine compared with the limited supply of PCA (34) from total synthesis

(ii) relative procedural simplicity leading to faster throughput in sample processing

(iii) the better resolution between substrate and product peaks in the HPLC analysis, conferring higher reliability for product identification

The search for CAS activity was carried out on *S. antibioticus* cells harvested daily over a 6-day fermentation. This procedure was prompted by the common observation[3, 83] that antibiotic synthesizing-enzymes are expressed only during the antibiotic production phase of a microbial growth profile, presumably to maximise cellular economy. With CFEs from the 2nd, 3rd and 4th-day harvest, we were thus able to detect significant turnover of (130) into a new product that co-eluted with (131) on HPLC and also gave NMR signals identical to those of (131). No such conversion was observed with CFEs from the 1st, 5th and 6th-day harvest. Enzyme activity (1 IU = 1 μmole of product formed per minute) and valclavam (37) titre was plotted against time of fermentation, as shown in Figure 9.5.

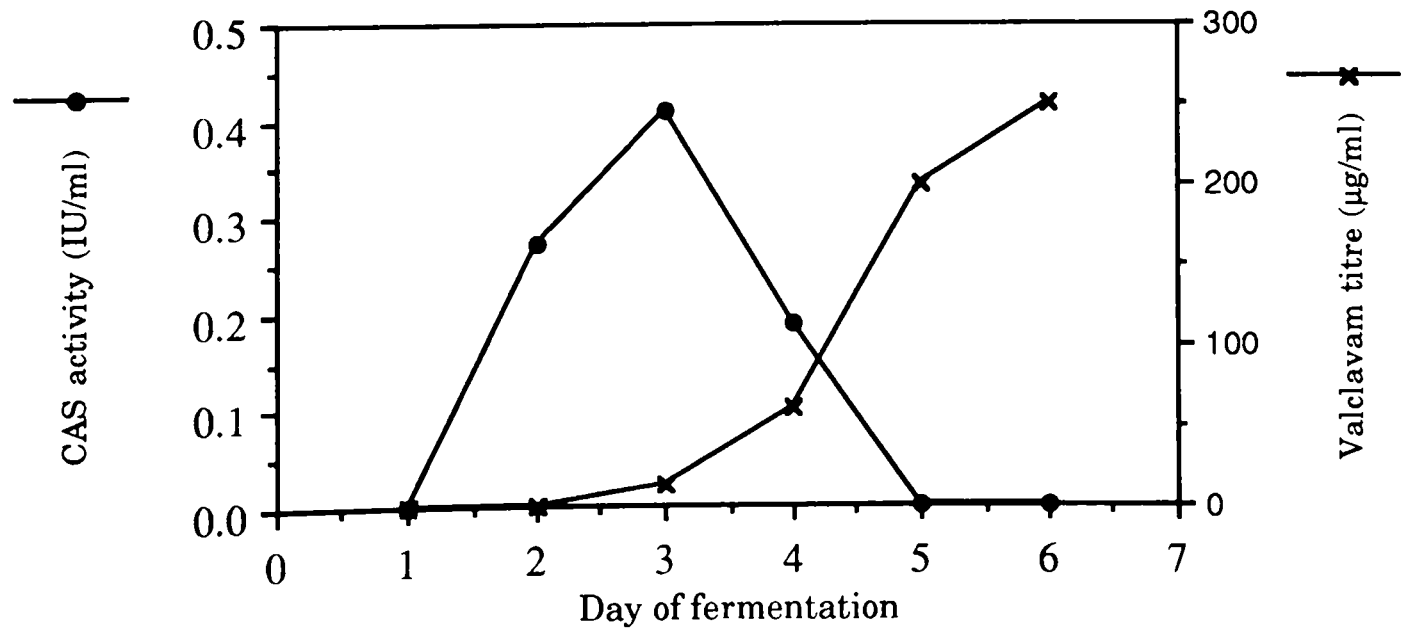



**Figure 9.5** CAS activity and valclavam (37) production during *S. antibioticus* fermentation

The antibiotic profile was similar to that obtained in Figure 7.5, with valclavam (37) appearing on the 3rd day and rising gradually to its peak on the 6th day. The CAS activity profile was interesting, reminiscent of the time course of IPNS production during *C. acremonium* fermentations.[26] In both cases, the enzyme activities were observed to appear on the 2nd day, peaked on the 3rd day, dropped by ~ 50% on the 4th day and disappeared on the 5th day. This suggested that the enzymes have served the full term of their biological duty by the 5th day. With the mirroring of valclavam (37) production in the production of cephalosporin C by *C. acremonium*, the comparison of the two pathways become even more intriguing. It is tempting to believe that the similar time course of CAS production (in *S. antibioticus*) and IPNS production (in *C. acremonium*) is not merely fortuitous, but reflects the equivalent positions of these enzymes within the time-scale of their respective biosynthetic pathways. By now, IPNS is well established as the second enzyme in the six steps from three precursor amino acids to cephalosporin C (see Figure 1.1). Though only two enzymes have been discovered in the clavulanic acid (31) pathway, it is likely that CAS mediates the steps in the early part of the valclavam (37) pathway, as speculated in Figure 8.2. One wonders if it is a sheer coincidence that both enzymes carry out oxidative cyclisation to give analogous β-lactam structures.

Whilst the presence of $N_\alpha$-acetyl-*L*-arginine hydroxylation activity was unequivocal, it was not rigorous proof of the presence of CAS. This is due to many literature reports[249-251] of the natural occurrence of various hydroxy-*L*-arginines, all pointing to the existence of enzymes capable of hydroxylating *L*-arginine or its simple derivatives. In the event, the presence of CAS in *S. antibioticus* was verified by the conversion of its natural substrate, PCA (34), into a mixture of dihydroclavaminic and clavaminic acids (HPLC and NMR data not shown) by a 2nd-day CFE. We then set out to attempt an enzyme isolation with the 3rd-day CFE, with a view to obtaining sufficient quantities of protein for N-terminal sequencing. CAS activity was monitored by $N_\alpha$-acetyl-*L*-arginine hydroxylation during the purification. Table 9.1 summarizes the efforts of this preliminary purification.

| Step[a] | Total protein (mg) | Total activity | | Specific activity (IU/mg) | Purification (fold) |
|---|---|---|---|---|---|
| | | IU | Recovery | | |
| Crude sonicate | 120 | 3.27 | 100 | 0.028 | 1 |
| Q-Sepharose FF pool | 7.5 | 1.13 | 34 | 0.151 | 5.4 |
| Superdex 75 pool | 0.71 | 0.087 | 2.7 | 0.123 | 4.4 |
| Mono Q pool | 0.14 | 0.088 | 2.7 | 0.629 | 22.5 |

(a) Not all the enzyme eluted from a step is pooled for the next step.

**Table 9.1** Purification of CAS from *S. antibioticus* ssp *antibioticus* Tü 1718

The first FPLC step, a strong anion exchange on the Q-Sepharose Fast Flow column, gave four active fractions as shown in Figure 9.6. The SDS-PAGE gel is shown in Figure 9.7.

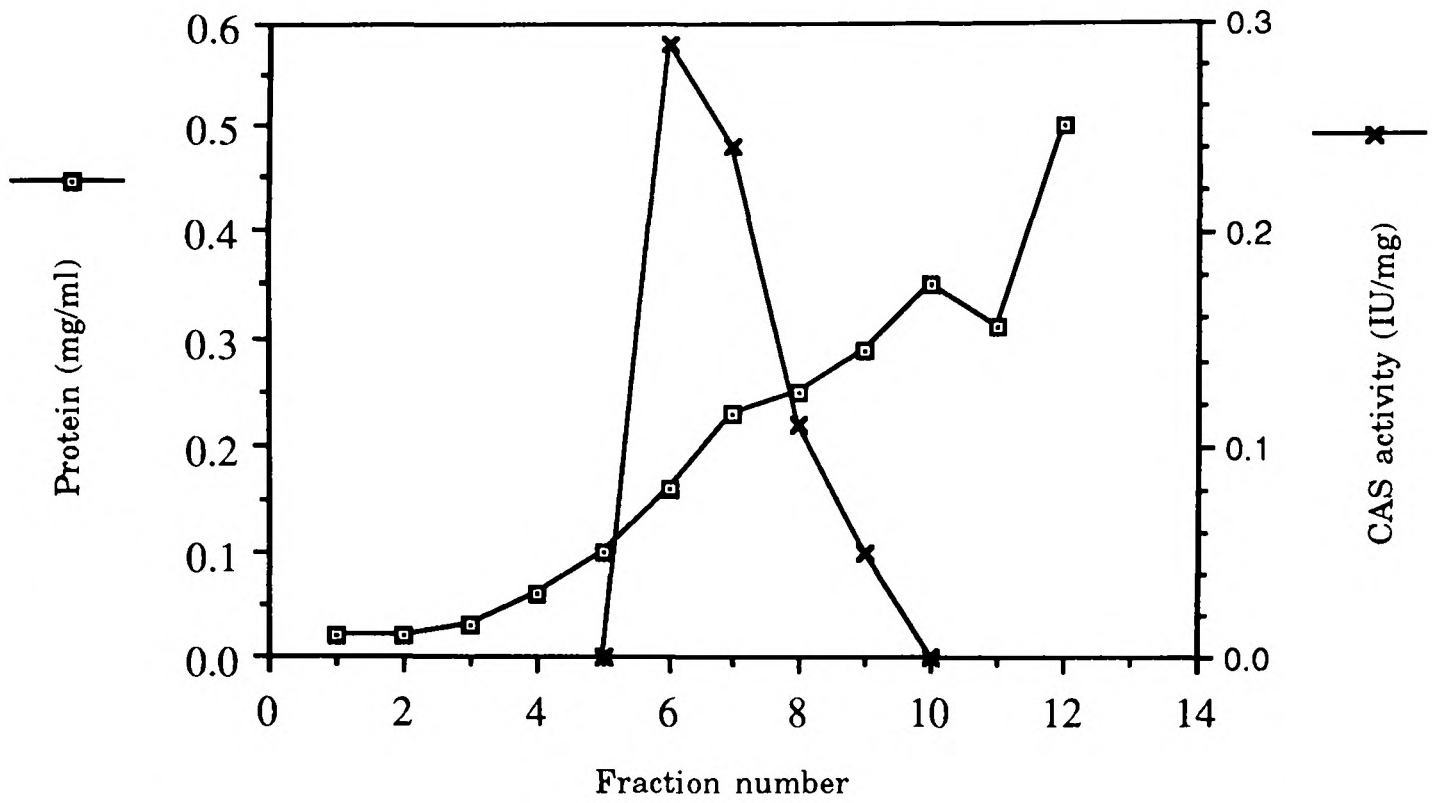



**Figure 9.6** Protein elution profile from a Q-Sepharose FF run on CAS

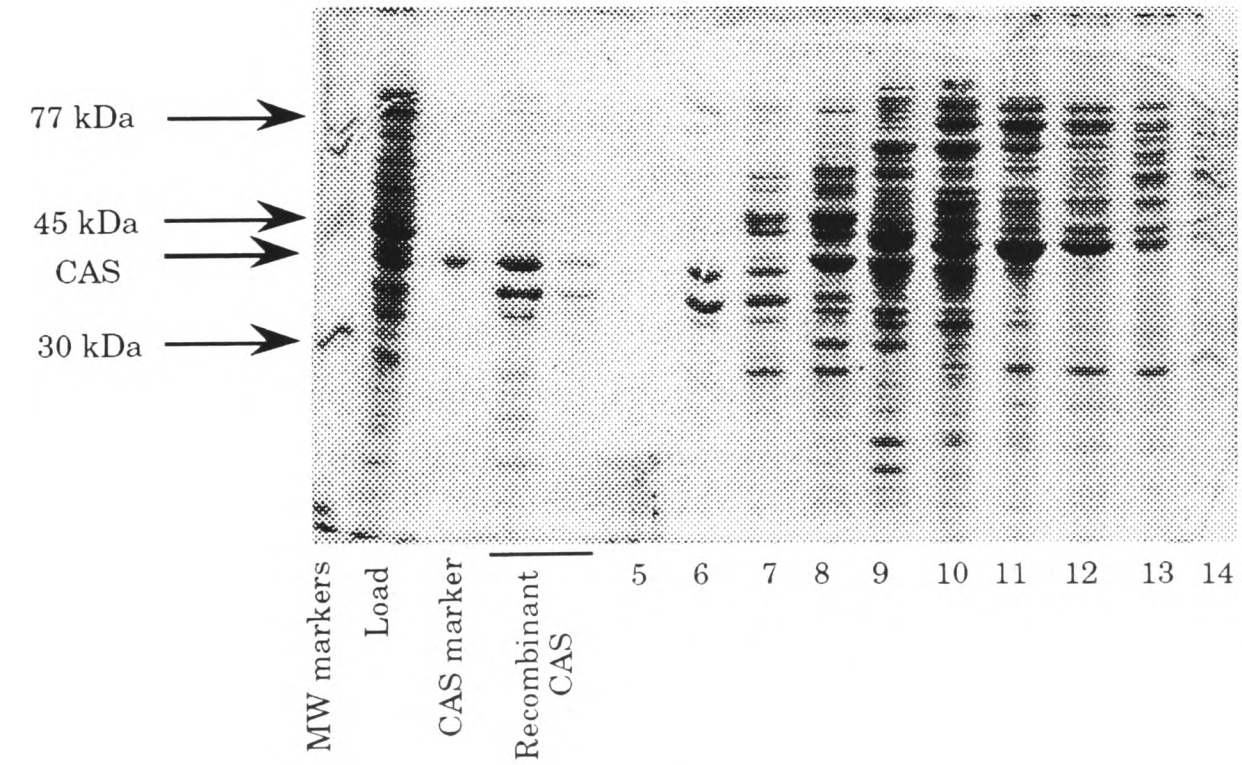



**Figure 9.7** SDS-PAGE gel of some fractions from Q-Sepharose FF run on CAS

The second FPLC step, gel filtration on the Superdex 75 column resulted in a substantial loss of activity which was probably attributable to the concentration step prior to FPLC. The protein elution profile and SDS-PAGE gel are shown in Figures 9.8 and 9.9 respectively.

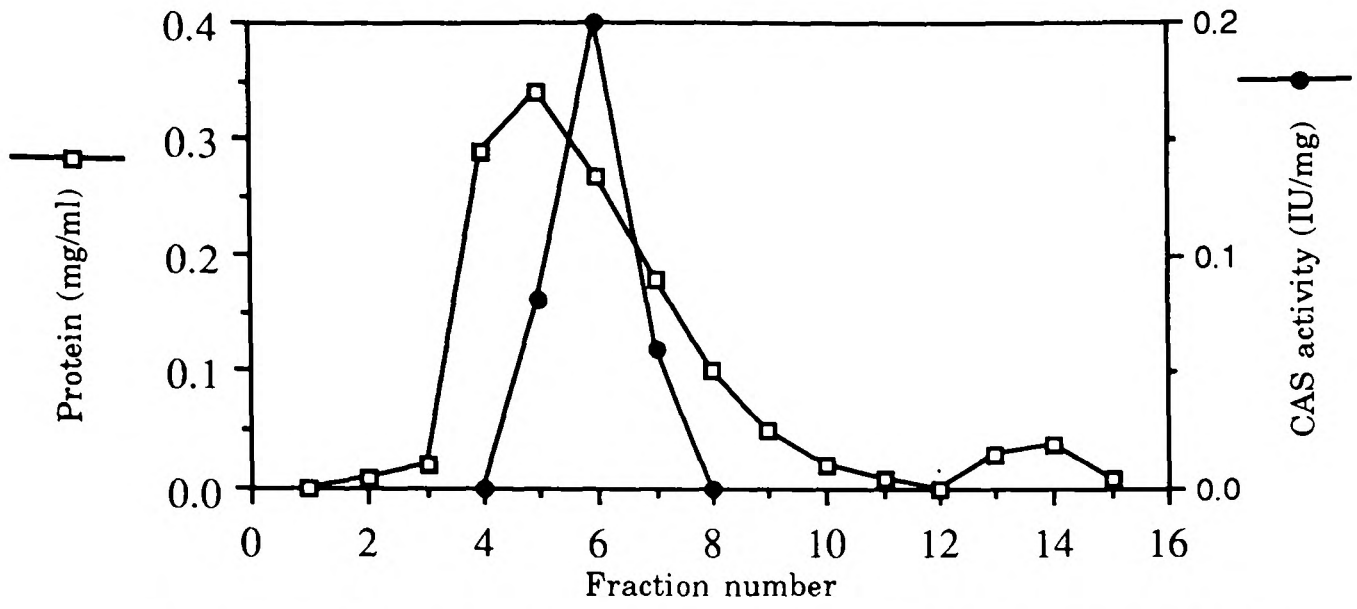



**Figure 9.8** Protein elution profile from a Superdex 75 run on CAS

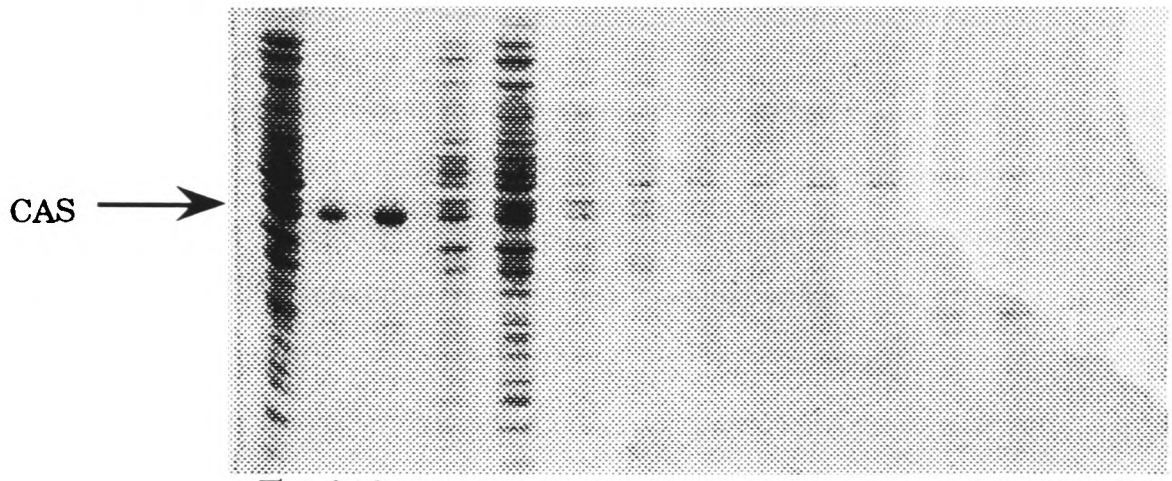



**Figure 9.9** SDS-PAGE gel of some fractions from Superdex 75 run on CAS

The final step, another strong anion-exchange chromatography but with the Mono Q resin, gave an active fraction (fraction 9 in Figures 9.10 and 9.11) that contained a major band (~ 40% of total fraction protein) found to co-migrate electrophoretically with the 36 kDa CAS from *S. clavuligerus*.

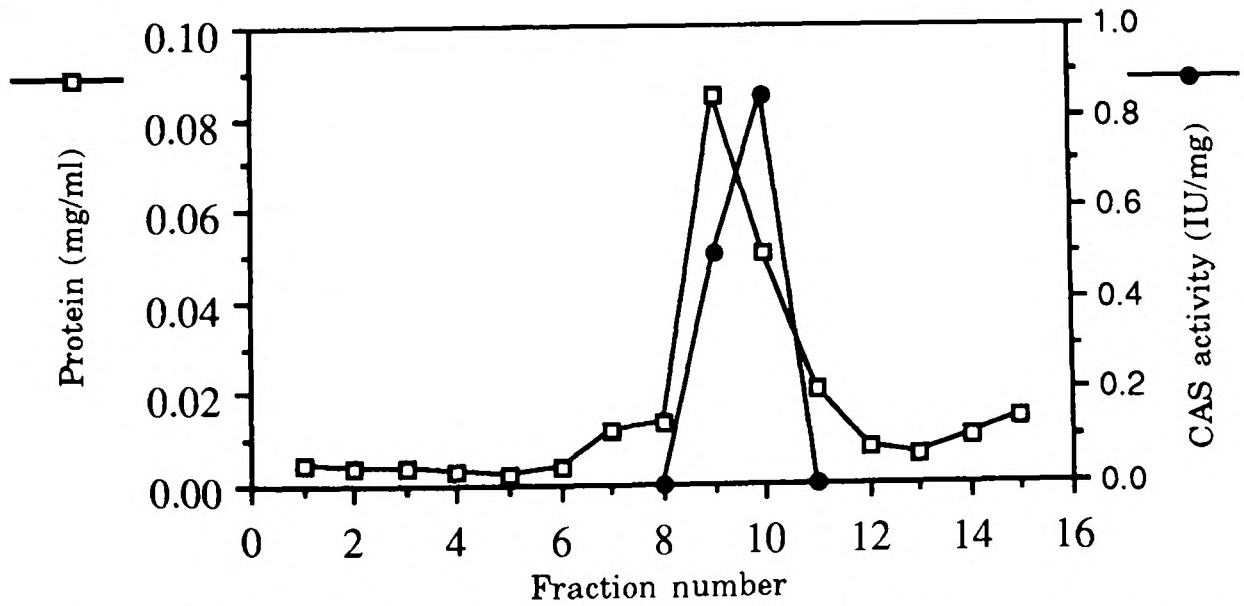



**Figure 9.10** Protein elution profile from a Mono Q run on CAS

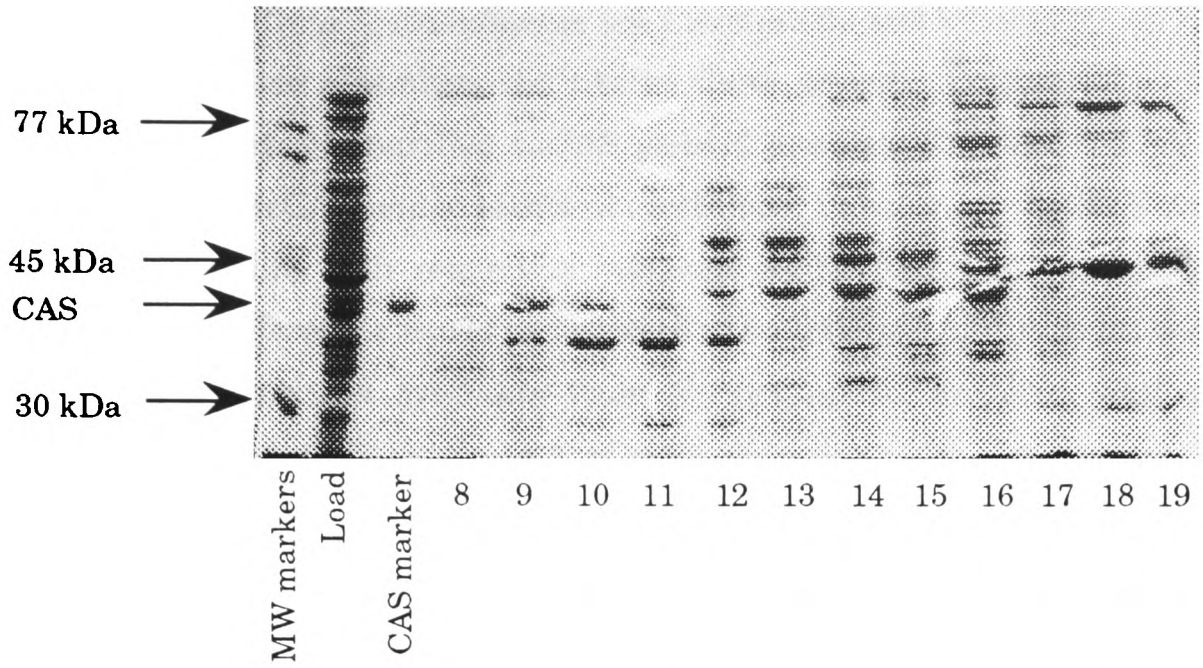



**Figure 9.11** SDS-PAGE gel of some fractions from Mono Q run on CAS

The major band from fraction 9 of the last FPLC step was blotted from the SDS-PAGE gel and subjected to N-terminal analysis. The 33-residue N-terminal sequence obtained, when aligned with N-terminal sequences[216] of the CAS isozymes from *S. clavuligerus*, convincingly identified this band as a CAS isozyme from *S. antibioticus* (Figure 9.12).

| | | |
|---|---|---|
| $sc_1$CAS | TSVDCTAYGPELRALAARLPRTPRADLYAFLDA | (published) |
| saCAS | TVVDCSEYSADLLALASRLPRIPRQDLYGFLDA | (our data) |
| $sc_2$CAS | PIVDCTPYRDELLALASELPEVPRADLHGFLDE | (published) |

(Legend : Single underline = conservative change, Double underline = identity
sa = *S. antibioticus*, $sc_1$ = *S. clavuligerus* isozyme 1, $sc_2$ = *S. clavuligerus* isozyme 2)

**Figure 9.12** Comparison of 33-residue N-terminal sequences from different CAS enzymes

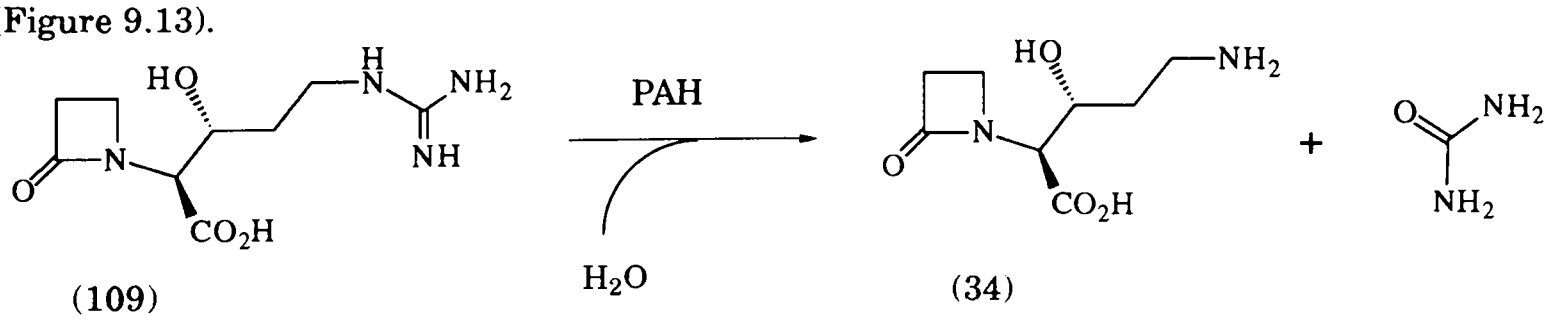
It is interesting to examine the sequence homology reflected in Figure 9.12 vis-à-vis the homologies previously discussed for REXH, DAOCS and DACS. As mentioned in Section 2.2, in terms of entire amino acid sequence, REXH is 57% identical to DAOCS, 54% identical to DACS, with 59% identity between DAOCS and DACS. On the other hand, in terms of the 33-residue N-terminal sequence, saCAS is 67% identical to $sc_1$CAS and 61% identical to $sc_2$CAS, with 61% identity between $sc_1$CAS and $sc_2$CAS. However, the complete sequences for the two isozymes of *S. clavuligerus* show 82% identity due to greater consensus away from the N- and C-termini.[216] If this homology pattern can be extrapolated to include the saCAS, we should expect the corresponding homology to be ~ 20% higher than the above. Nevertheless, the CAS enzymes bear a closer resemblance to one another than what is observed for REXH, DAOCS and DACS. In retrospect, this discovery is not surprising in the light of two factors:

(a) The CAS isozymes from *S. clavuligerus* are functionally and kinetically very similar, and saCAS probably also resembles the two in these aspects. Preliminary incubations of PCA (34) with partially purified saCAS yielded products (35) and (36) in ratios often encountered with scCAS under identical conditions. In contrast, DAOCS and DACS are largely specialised in their respective reactions whilst REXH is competent in both.[53] Such functional differences would be expected to be reflected in structural details.

(b) All three CAS enzymes originate from *Streptomyces* bacteria, and it is open to speculation whether CAS exists as 2 or more isozymes in *S. antibioticus*. In contrast, REXH is of fungal origin (*C. acremonium*) whilst DAOCS and DACS are of a bacterial origin (*S. clavuligerus*).

Given their co-occurrence in *S. clavuligerus* and their mechanistic similarities, these α-KG dependent dioxygenases (*i.e.* $sc_1$CAS, $sc_2$CAS, DAOCS and DACS) represent the closest chemical parallels currently identifiable between the two β-lactam biosynthetic pathways (clavams vs. cephems). If an evolutionary relation existed between these two pathways, one would expect it to be most evident in these four enzymes. However, comparison of the DNA sequences for all four proteins provided no convincing evidence of a common evolutionary origin for these two pathways.[216] This comparison and a growing body of chemical evidence lend some credence to the notion of several separately evolved solutions to the task of β-lactam biosynthesis.[236] Perhaps, the competitive advantage conferred on microorganisms by the ability to make β-lactam antibiotics is sufficiently great that it has been achieved several times over by quite independent means.

Having established the existence of CAS in *S. antibioticus*, the presence of PAH becomes highly probable since the latter is an 'upstream' enzyme in the pathway. Our assay for PAH was based on a spectrophotometric determination of urea produced by the PAH-catalysed reaction (Figure 9.13).




(109) (34)

**Figure 9.13** Production of urea from PAH catalysis

PAH activity was monitored alongside the production of CAS and valclavam (37), again using CFE prepared from mycelia harvested over a 6-day fermentation of *S. antibioticus*. The CAS and valclavam (37) levels obtained were essentially similar to Figure 9.5 and PAH activity was indeed present in the CFEs. We were surprised, however, to find that PAH activity did not decrease after the 3rd day (Figure 9.14). A rationale for this observation would be that the same PAH continues to be functioning in other metabolic pathways, probably serving a similar 'arginase'-type role for other simple guanidino compounds.

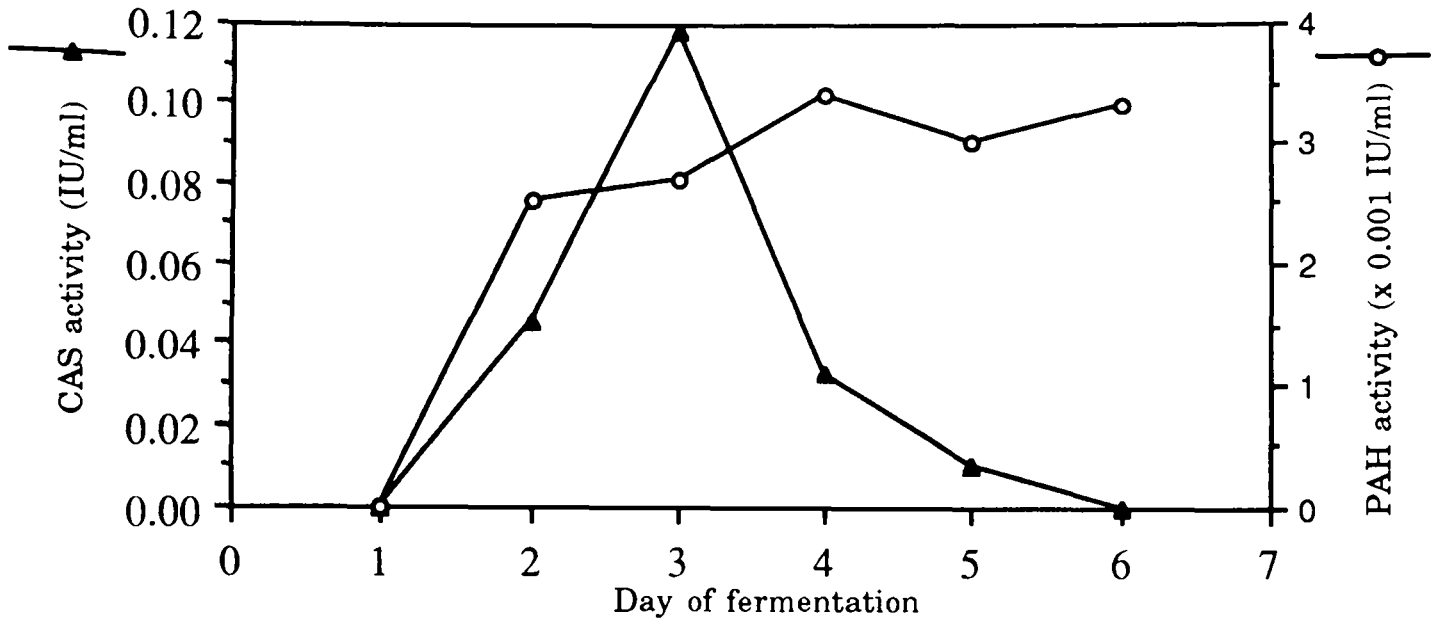



**Figure 9.14** PAH activity during the fermentation of *S. antibioticus*

In view of the fact that CAS activity has been reported in other clavulanic acid-producing organisms, *e.g. S. jumonjinensis*[192] and *S. katsurahamanus*,[193] we monitored for clavulanic acid (31) production in the last two fermentations. The failure to detect this substance, or any other clavams, allowed us to propose the involvement of PAH and CAS in valclavam (37) biosynthesis.

In conclusion, the results of Chapters 7, 8 and 9 provided the first definitive evidence for the overlap of the clavulanic acid (31) pathway in *S. clavuligerus* and the valclavam (37) pathway in *S. antibioticus*. This may not come as a surprise since clavulanic acid (31) and valclavam (37) are structurally homologous molecules produced by two closely related species. The case for a common origin of all clavams is thus reinforced. In the event of a verification, such a sharing of pathways will provide an interesting parallel to the biosynthesis of penicillins, cephalosporins and cephamycins (Figure 1.1). If indeed these two extended biosynthetic schemes (*i.e.* sulfur-based β-lactams vs. oxygen-based β-lactams) do not have a common ancestral origin, then they must represent a most convincing demonstration of the power and beauty of convergent evolution.

## 9.3 Experimental

### 9.3.1 General techniques

Techniques and instrumentation for protein purification and analysis (FPLC, Bradford assay, SDS-PAGE, N-terminal amino acid sequencing) were essentially identical to those in Chapter 3. High-field NMR, ESMS and HPLC were also carried out according to procedures described in Chapter 4. Fermentation of *S. antibioticus* ssp *antibioticus* Tü 1718, bioassay for valclavam (37) and imidazole derivatisation assay [for clavulanic acid (31)] were performed as described in Chapter 7. "Tris buffer" refers to 50 mM Tris adjusted with HCl to pH 7.5 at room temperature.

### 9.3.2 Preparation of cell-free extract of *S. antibioticus* ssp *antibioticus* Tü 1718

Cell-free extracts of *S. antibioticus* ssp *antibioticus* Tü 1718 were prepared according to a modification of Townsend's protocol[39]:

Three 100-ml cultures were harvested each day for 6 days and stored at -80°C. Frozen cell paste (5 g) was resuspended in 15 ml of Tris buffer containing 10 mM KCl, 10 mM $MgCl_2$, 10% (v/v) glycerol, 5 mg/l soybean trypsin inhibitor, 0.5 mM PMSF and 0.5 mM benzamidine. The suspension was given 4 bursts of sonication at 10 sec each, using a 380-W sonicator at power setting 5. Cell debris was removed by centrifugation at 28 000 *g* for 20 min, and the supernatant treated with one-tenth its volume of 10% aqueous streptomycin sulfate. After 10 min of equilibration, the precipitate was removed by centrifugation at 28 000 *g* for 30 min. The supernatant was filtered through 0.45 μm sterile filter and exchanged into the desired buffer (appropriate for each experiment) using PD-10 columns (Pharmacia).

### 9.3.3 Assay for CAS activity

Unless otherwise specified, the assay for CAS activity was carried out using a modification of the protocol developed by Lloyd[76]:

Incubation mixtures contained, in a total volume of 100 μl, α-KG (10 mM), $FeSO_4$ (1 mM), $N_\alpha$-acetyl-*L*-arginine (130) (10 mM) and enzyme (70 μl added) in Tris buffer containing 10 mM KCl and 10 mM $MgCl_2$. Mixtures were incubated at ambient temperature (22 - 25°C) in an orbital shaker operating at 200 rpm. After 5 or 10 min, the reaction was quenched by flash-heating for 15 sec in a 60-W microwave oven at power setting 4. The precipitated protein was removed by microfuge centrifugation. The supernatant (25 μl) was introduced into a $C_{18}$ column (250 x 4.6 mm) eluting isocratically with Milli Q water at 1 ml/min. The unreacted substrate (130) [$R_t$ = 6.8 min] and hydroxylated product (131) [$R_t$ = 4.6 min] were monitored at 214 nm. The amount of product formed was quantified by cutting out the HPLC peak and comparing with a calibration curve.

### 9.3.4 Assay for PAH activity

The assay for PAH activity was adapted from an unpublished protocol developed by SmithKline Beecham[252]:

CFE's were prepared from frozen cell pastes collected daily over a 6-day fermentation. The CFE's were exchanged into 50 mM Tris-HCl, pH 8.5 at RT. The incubation mixture contained, in a total volume of 0.5 ml, 0.1 ml of buffer-exchanged CFE, 1 mM of acetohydroxamic acid, 0.1 mM of $MnCl_2$ and 1 mM of 3-hydroxy-5-guanidino-2-(2-oxoazetidin-1-yl)pentanoic acid (109) in 50 mM Tris-HCl, pH 8.5 at RT. The mixtures were incubated at 28°C without shaking. A 50 μl aliquot of each incubation mixture (for 1-day to 6-day CFE) was collected at fixed time

intervals (0, 0.25, 0.5, 1, 2 and 3-hour) and frozen on dry ice. Each aliquot was then mixed with 0.5 ml of BUN colour reagent and 0.75 ml of BUN acid reagent (both are part of the Sigma Urea Assay Kit, cat. no 535-B) and heated at 100°C for 10 min. The mixtures were cooled rapidly in an ice-bath and its absorbance measured at 525 nm. The amount of urea produced was obtained from a calibration chart using urea standards (2, 4, 6, 8 and 10 μg/ml) prepared from a stock solution in the Sigma Kit. The experimental control consisted of an enzyme-free reaction mixture treated in an otherwise identical manner. The production of urea against time was plotted for each CFE and the initial reaction velocity estimated from the graph.

### 9.3.5 Incubation with PCA (34)

The 2nd-day CFE was prepared according to Section 9.3.2 and exchanged into a buffer containing 10 mM $NH_4HCO_3$, 10 mM KCl, 10 mM $MgCl_2$ and 1 mM DTT. The buffer-exchanged CFE (1.4 ml) was mixed with 0.2 ml of $FeSO_4$ (10 mM), 0.2 ml of α-KG (100 mM) and 0.2 ml of PCA (100 mM). The 2-ml mixture was incubated at 27°C for 60 min in an orbital shaker operating at 200 rpm. The reaction was quenched with 5 ml of acetone and centrifuged to remove precipitated protein. Acetone was removed by rotary evaporation and the aqueous portion freeze-dried for NMR analysis.

### 9.3.6 Preliminary purification of CAS from *S. antibioticus* ssp *antibioticus* Tü 1718

The preliminary enzyme purification was performed with a 3rd-day CFE, prepared according to Section 9.3.2. The CFE (30 ml) was filtered through a 0.22 μm filter and loaded onto a 15-ml Q-Sepharose Fast Flow column (XK 16/20 housing) previously equilibrated with buffer A (50 mM Tris, pH 7.5, 10 mM $MgCl_2$, 10 mM KCl, 1 mM DTT) and buffer B (buffer A with 1 M NaCl). Running at 2 ml/min, a linear gradient of 0% to 5% buffer B was applied for the first 15 ml, followed by 5% to 25% buffer B over the next 75 ml. Fractions of 7.5 ml were collected, assayed for protein content and CAS activity and analysed by SDS-PAGE. The active and reasonably pure fractions were pooled together and concentrated by ultrafiltration. The concentrate was directly loaded onto a Superdex 75 10/30 column, previously equilibrated with buffer A. Fractions of 1 ml were collected at 0.25 ml/min. The fractions were assayed for protein content and CAS activity and analysed by SDS-PAGE. The active and reasonably pure fractions were pooled together and loaded onto a Mono Q 5/5 column, previously equilibrated with buffers A and B. Running at 0.5 ml/min, a linear gradient of 0% to 5% buffer B was applied for the first 1 ml, followed by 5% to 25% buffer B over the next 10 ml. Fractions of 1 ml were collected, assayed for protein content and CAS activity and analysed by SDS-PAGE.

# REFERENCES


1. J. Mann, *Secondary metabolism* , 2nd Edn, pp.1-374 (Oxford University Press, Oxford, 1987)

2. A. L. Demain, in *Secondary metabolites: their function and evolution* : (ed. D. J. Chadwick and J. Whelan) pp.3-23 (John Wiley & Sons, Chichester, 1992)

3. J. F. Martin and A. L. Demain, in *The filamentous fungi* : (ed. J. E. Smith and D. R. Berry) pp.426-450 (Edward Arnold, London, 1978)

4. D. H. Williams and R. A. Maplestone, in *Secondary metabolites: their function and evolution* : (ed. D. J. Chadwick and J. Whelan) pp.45-63 (John Wiley & Sons, Chichester, 1992)

5. J. Davies, in *Secondary metabolites: their function and evolution* : (ed. D. J. Chadwick and J. Whelan) pp.1-2 (John Wiley & Sons, Chichester, 1992)

6. Y. Fujishima, *D.Phil. Thesis,* University of Oxford (1992)

7. L. J. Nisbet, in *Secondary metabolites: their function and evolution* : (ed. D. J. Chadwick and J. Whelan) pp.215-235 (John Wiley & Sons, Chichester, 1992)

8. R. Thiericke and J. Rohr, *Nat. Prod. Rep.*, 265-289 (1993)

9. K. C. Nicolaou, C. Riemer, M. A. Kerr, D. Rideout and W. Wrasidlo, *Nature*, **364**, 464-466 (1993)

10. G. G. Yarbrough, D. P. Taylor, R. T. Rowlands, M. S. Crawford and L. L. Lasure, *J. Antibiot.*, **46**, 535-544 (1993)

11. L. Crombie, *Nat. Prod. Rep.*, **9**, 193-198 (1992)

12. A. M. Bridges, P. F. Leadlay, W. P. Revill and J. Staunton, *J. Chem. Soc., Chem. Commun.*, 776-777 (1991)

13. A. M. Bridges, P. F. Leadlay, W. P. Revill and J. Staunton, *J. Chem. Soc., Chem. Commun.*, 778-779 (1991)

14. D. E. Cane, *Tetrahedron*, **47**, xiii-xiv (1991)

15. T. Cavalier-Smith, in *Secondary metabolites: their function and evolution* : (ed. D. J. Chadwick and J. Whelan) pp.64-87 (John Wiley & Sons, Chichester, 1992)

16. H. Kleinkauf and H. von Döhren, in *Biochemistry of peptide antibiotics: recent advances in the biotechnology of β-lactams and microbial bioactive peptides* : (ed. H. Kleinkauf and H. von Döhren) pp.1-32 (Walter de Gruyter, Berlin, 1990)

17. M. I. Page, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.129-147 (Chapman & Hall, London, 1992)

18. H. C. Neu, *Science*, **257**, 1064-1073 (1992)

19. J. M. Frère, M. Nguyen-Distèche, J. Coyette and B. Joris, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.148-197 (Chapman & Hall, London, 1992)

20. S. G. Waley, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.198-228 (Chapman & Hall, London, 1992)

21. R. D. G. Cooper, *Bioorg. Med. Chem.*, **1**, 1-17 (1993)

22. Y. Aharonowitz, G. Cohen and J. F. Martin, *Ann. Rev. Microbiol.*, **46**, 461-495 (1992)

23. S. Usui and C.-A. Yu, *Biochim. Biophy. Acta*, **999**, 78-85 (1989)

24. L. Láiz, P. Liras, J. M. Castro and J. F. Martín, *J. Gen. Microbiol.*, **136**, 663-671 (1990)

25. S. E. Jensen, D. W. S. Westlake and S. Wolfe, *J. Antibiot.*, **38**, 263-265 (1985)

26. J. E. Baldwin, R. M. Adlington, J. B. Coates, M. J. C. Crabbe, N. P. Crouch, J. W. Keeping, G. C. Knight, C. J. Schofield, H.-H. Ting, C. A. Vallejo, M. Thorniley and E. P. Abraham, *Biochem. J.*, **245**, 831-841 (1987)

27. J. E. Dotzlaf and W.-K. Yeh, *J. Bacteriol.*, **169**, 1611-1618 (1987)

28. A. Scheidegger, M. T. Küenzi and J. Nüesch, *J. Antibiot.*, **37**, 522-531 (1984)

29. A. Matsuda and K. Matsuyama. *An acetyltransferase gene-containing DNA*, European Patent EP 0 450 758 A1 (1991)

30. Y. Fujisawa and T. Kanzaki, *Agric. Biol. Chem.*, **39**, 2043-2048 (1975)

31. A. Scheidegger, A. Gützwiller, M. T. Küenzi, A. Fiechter and J. Nüesch, *J. Biotech.*, **3**, 109-117 (1985)

32. J. E. Baldwin and C. J. Schofield, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.1-78 (Chapman & Hall, London, 1992)

33. C. Reading and M. Cole, *Antimicrob. Agents Chemother.*, **11**, 852-857 (1977)

34. B. P. Valentine, C. R. Bailey, A. Doherty, J. Morris, S. W. Elson, K. H. Baggaley and N. H. Nicholson, *J. Chem. Soc., Chem. Commun.*, 1210-1211 (1993)

35. S. W. Elson, K. H. Baggaley, J. Gillett, S. Holland, N. Nicholson, J. T. Sime and S. R. Woroniecki, *J. Chem. Soc., Chem. Commun.*, 1736-1738 (1987)

36. S. W. Elson, K. H. Baggaley, J. Gillett, S. Holland, N. Nicholson, J. T. Sime and S. R. Woroniecki, *J. Chem. Soc., Chem. Commun.*, 1739-1740 (1987)

37. J. E. Baldwin, R. M. Adlington, J. S. Bryans, A. O. Bringhen, J. B. Coates, N. P. Crouch, M. D. Lloyd, C. J. Schofield, S. W. Elson, K. H. Baggaley, R. Cassels and N. Nicholson, *J. Chem. Soc., Chem. Commun.*, 617-619 (1990)

38. J. E. Baldwin, R. M. Adlington, J. S. Bryans, A. O. Bringhen, J. B. Coates, N. P. Crouch, M. D. Lloyd, C. J. Schofield, S. W. Elson, K. H. Baggaley, R. Cassels and N. Nicholson, *Tetrahedron*, **47**, 4089-4100 (1991)

39. S. P. Salowe, E. N. Marsh and C. A. Townsend, *Biochemistry*, **29**, 6499-6508 (1990)

40. B. W. Bycroft, C. Maslen, S. J. Box, A. Brown and J. W. Tyler, *J. Antibiot.*, **41**, 1231-1242 (1988)

41. J. O' Sullivan, A. M. Gillum, C. A. Aklonis, M. L. Souser and R. B. Sykes, *Antimicrob. Agents Chemother.*, **21**, 558-564 (1982)

42. J. O' Sullivan, M. L. Souser, C. C. Kao and C. Aklonis, *Antimicrob. Agents Chemother.*, **23**, 598-602 (1983)

43. O. Hayaishi (ed.), *Molecular mechanisms of oxygen activation* , pp.1-651 (Academic Press, London, 1974)

44. H. Peter, J. Rabenhorst, F. Röhl and H. Zähner, in *Recent advances in chemotherapy* : (ed. J. Ishigami) pp.237-238 (University of Tokyo Press, Tokyo, 1985)

45. M. Kohsaka and A. L. Demain, *Biochem. Biophy. Res. Commun.*, **70**, 465-473 (1976)

46. M. Yoshida, T. Konomi, M. Kohsaka, J. E. Baldwin, S. Herchen, P. D. Singh, N. A. Hunt and A. L. Demain, *Proc. Natl. Acad. Sci. USA*, **75**, 6253-6257 (1978)

47. M. K. Turner, J. E. Farthing and S. J. Brewer, *Biochem. J.*, **173**, 839-850 (1978)

48. J. Kupka, Y.-Q. Shen, S. Wolfe and A. Demain, *FEMS Microbiol. Lett.*, **16**, 1-6 (1983)

49. D. J. Hook, L. T. Chang, R. P. Elander and R. B. Morin, *Biochem. Biophy. Res. Commun.*, **87**, 258-265 (1979)

50. C. M. Stevens, E. P. Abraham, F. C. Huang and C. J. Sih, *FASEB*, **34**, 625 (1975)

51. S. M. Samson, J. E. Dotzlaf, M. L. Slisz, G. W. Becker, R. M. V. Frank, L. E. Veal, W.-K. Yeh, J. R. Miller, S. W. Queener and T. D. Ingolia, *Bio/Tech.*, **5**, 1207-1214 (1987)

52. J. E. Baldwin, J. M. Blackburn, R. J. Heath and J. D. Sutherland, *Bioorg. & Med. Chem. Lett.*, **2**, 663-668 (1992)

53. W.-K. Yeh, J. E. Dotzlaf and G. W. Huffman, in *50 years of penicillin application - A symposium in honour of Sir Edward P. Abraham - 1990* : (ed. H. Kleinkauf and H. von Döhren) pp.180-197 (W. de Gruyter, Berlin, *In press*)

54. S. M. Samson, J. L. Chapman, R. Belagaje, S. W. Queener and T. D. Ingolia, *Proc. Natl. Acad. Sci. USA*, **84**, 5705-5709 (1987)

55. M. T. Abbott and S. Udenfriend, in *Molecular mechanisms of oxygen activation* : (ed. O. Hayaishi) pp.167-214 (Academic Press Inc., New York, 1974)

56. J. Matsuda, S. Okabe, T. Hashimoto and Y. Yamada, *J. Biol. Chem.*, **266**, 9460-9464 (1991)

57. R. Myllylä, V. Günzler, K. I. Kivirikko and D. D. Kaska, *Biochem. J.*, **286**, 923-927 (1992)

58. H. M. Hanauske-Abel and V. Günzler, *J. Theor. Biol.*, **94**, 421-455 (1982)

59. R. Myllylä, L. Tuderman and K. I. Kivirikko, *Eur. J. Biochem.*, **80**, 349-357 (1977)

60. E. De Carolis and V. De Luca, *J. Biol. Chem.*, **268**, 5504-5511 (1993)

61. L. Tuderman, R. Myllylä and K. I. Kivirikko, *Eur. J. Biochem.*, **80**, 341-348 (1977)

62. T. Hashimoto and Y. Yamada, *Eur. J. Biochem.*, **164**, (1987)

63. S. Udenfriend and G. Cardinale, in *Oxygenases and oxygen metabolism* : (ed. M. Nozaki) pp.99-110 (Academic Press, New York, 1982)

64. L. Britsch and H. Grisebach, *Eur. J. Biochem.*, **156**, 569-577 (1986)

65. L. Britsch, *Arch. Biochem. Biophy.*, **282**, 152-160 (1990)

66. B. Siegel, *Bioorg. Chem.*, **8**, 219-226 (1979)

67. H. A. Albrecht, G. Beskid, J. G. Christenson, J. W. Durkin, V. Fallat, N. H. Georgopapadakou, D. D. Keith, F. M. Konzelmann, E. R. Lipschitz, D. H. McGarry, J. Siebelist, C. C. Wei, M. Weigele and R. Yang, *J. Med. Chem.*, **34**, 669-675 (1991)

68. W.-K. Yeh and S. W. Queener, *Ann. N.Y. Acad. Sci.*, **613**, 128-141 (1990)

69. P. L. Skatrud, A. J. Tietz, T. D. Ingolia, C. A. Cantwell, D. L. Fisher, J. L. Chapman and S. W. Queener, *Bio/Tech.*, **7**, 477-485 (1989)

70. W.-K. Yeh, S. K. Ghag and S. W. Queener, *Ann. N.Y. Acad. Sci.*, **672**, 396-408 (1992)

71. A. Kornberg, *Methods in Enzymology*, **182**, 1-5 (1990)

72. J. Kupka, Y.-Q. Shen, S. Wolfe and A. L. Demain, *Can. J. Microbiol.*, **29**, 488-496 (1983)

73. M. M. Bradford, *Anal. Biochem.*, **72**, 248-254 (1976)

74. U. K. Laemmli, *Nature*, **227**, 680-685 (1970)

75. R. J. Heath, *D.Phil. Thesis,* University of Oxford (1992)

76. M. D. Lloyd, *D.Phil. Thesis,* University of Oxford (1992)

77. F. A. O. Marston, *Biochem. J.*, **240**, 1-12 (1986)

78. C. H. Schein, *Bio/Tech.*, **7**, 1141-1149 (1989)

79. T. Kohno, D. F. Carmichael, A. Sommer and R. C. Thompson, *Methods in Enzymology*, **185**, 187-195 (1990)

80. C. B. Anfinsen, E. Haber, M. Sela and F. H. White, *Proc. Natl. Acad. Sci. USA*, **47**, 1309-1314 (1961)

81. P. S. Kim and R. L. Baldwin, *Ann. Rev. Biochem.*, **59**, 631-660 (1990)

82. V. V. Mozhaev and K. Martinek, *Enzyme Microb. Technol.*, **4**, 299-309 (1982)

83. J. F. Martin and A. L. Demain, *Microbiol. Rev.*, **44**, 230-251 (1980)

84. A. Hershko and A. Ciechanover, *Ann. Rev. Biochem.*, **61**, 761-807 (1992)

85. G. E. Means and R. E. Feeney, *Chemical modification of proteins* , 1st Edn, (Holden-Day,

San Francisco, 1971)

86. P. V. Dandaram and F. Eckstein (ed.), *Theory and practice in affinity techniques* , (Academic Press, London, 1978)

87. A. R. Fersht, *Enzyme structure and mechanism* , 2nd Edn, pp.1-475 (W.H. Freeman, New York, 1985)

88. R. T. Aplin, J. E. Baldwin, C. J. Schofield and S. G. Waley, *FEBS Lett.*, **277**, 212-214 (1990)

89. Y. Fujita, A. Gottlieb, B. Peterkofsky, S. Udenfriend and B. Witkop, *J. Am. Chem. Soc.*, **86**, 4709-4716 (1964)

90. B. Lindblad, G. Lindstedt, M. Tofft and S. Lindstedt, *J. Am. Chem. Soc.*, **91**, 4604-4606 (1969)

91. P. Hobza, J. Hurych and R. Zahradnik, *Biochim. Biophy. Acta*, **304**, 466-472 (1973)

92. G. A. Hamilton, in *Progress in Bioorganic Chemistry* : (ed. E. T. Kaiser and F. J. Kézdy) pp.83-157 (Wiley-Interscience, New York, 1971)

93. P. R. Ortiz de Montellano (ed.), *Cytochrome P-450: Structure, Mechanism and Biochemistry* , (Plenum Press, New York, 1986)

94. A. G. Prescott, *J. Exptl. Bot.*, **44**, 849-861 (1993)

95. Y. Fujishima, *unpublished results*

96. F. P. Guengerich and T. L. MacDonald, *Acc. Chem Res.*, **17**, 9-16 (1984)

97. J. E. Baldwin, *J. Heterocyclic Chem.*, **27**, 71-78 (1990)

98. R. D. Miller, L. L. Huckstep, J. P. McDermott, S. W. Queener, S. Kukolja, D. O. Spry, T. K. Elzey, S. M. Lawrence and N. Neuss, *J. Antibiot.*, **34**, 984-993 (1981)

99. J. E. Baldwin, R. M. Adlington, R. T. Aplin, N. P. Crouch, G. Knight and C. J. Schofield, *J. Chem. Soc., Chem. Commun.*, 1651-1654 (1987)

100. J. E. Baldwin, R. M. Adlington, N. P. Crouch, C. J. Schofield, N. J. Turner and R. T. Aplin, *Tetrahedron*, **47**, 9881-9900 (1991)

101. J. E. Baldwin, R. M. Adlington, R. T. Aplin, N. P. Crouch, C. J. Schofield and H.-H. Ting, *J. Chem. Soc., Chem. Commun.*, 1654-1656 (1987)

102. J. E. Baldwin, R. M. Adlington, N. P. Crouch, N. J. Turner and C. J. Schofield, *J. Chem. Soc., Chem. Commun.*, 1141-1143 (1989)

103. C. A. Townsend, A. B. Theis, A. S. Neese, E. B. Barrabee and D. Poland, *J. Am. Chem. Soc.*, **107**, 4760-4767 (1985)

104. E. P. Abraham, C.-P. Pang, R. L. White, D. H. G. Crout, M. Lutstorf, P. J. Morgan and A. E. Derome, *J. Chem. Soc., Chem. Commun.*, 723-724 (1983)

105. J. E. Baldwin, R. M. Adlington, T. W. Kang, E. Lee and C. J. Schofield, *J. Chem. Soc., Chem. Commun.*, 104-106 (1987)

106. J. E. Baldwin, R. M. Adlington, R. T. Aplin, N. P. Crouch and R. Wilkinson, *Tetrahedron*, **48**, 6853-6862 (1992)

107. C. A. Townsend and E. B. Barrabee, *J. Chem. Soc., Chem. Commun.*, 1586-1588 (1984)

108. J. E. Baldwin and M. Bradley, *Chem. Rev.*, **90**, 1079-1088 (1990)

109. J. E. Baldwin, R. M. Adlington, M. J. C. Crabbe, G. C. Knight, T. Nomoto, C. J. Schofield and H.-H. Ting, *Tetrahedron*, **43**, 3009-3014 (1987)

110. S. W. Queener and N. Neuss, in *Chemistry and Biology of β-lactam Antibiotics* : (ed. R. B. Morin and M. Gorman) pp.45 (Academic Press, New York, 1982)

111. C. J. Schofield and S. Sukumaran, *unpublished results*

112. M. E. Wood, *D.Phil. Thesis,* University of Oxford (1990)

113. M. E. Wood, *unpublished results*

114. M. E. Wood, *personal communication*

115. S. W. Leppard, *B.A. Part II Thesis,* University of Oxford (1990)

116. D. Miller, *B.A. Part II Thesis,* University of Oxford (1993)

117. J. E. Baldwin, R. M. Adlington, N. P. Crouch, J. B. Coates, J. W. Keeping, C. J. Schofield, W. A. Shuttleworth and J. D. Sutherland, *J. Antibiot.*, **41**, 1694-1695 (1988)

118. R. M. Adlington, J. E. Baldwin, B. Chakravarti, M. Jung, S. E. Moroney, J. A. Murphy, P. D. Singh, J. J. Usher and C. Vallejo, *J. Chem. Soc., Chem. Commun.*, 153-154 (1983)

119. J. E. Baldwin, R. M. Adlington, N. P. Crouch, C. J. Schofield and H.-H. Ting, *J. Chem.*

*Soc., Chem. Commun.*, 1556-1558 (1987)

120. J. E. Baldwin, R. M. Adlington, N. P. Crouch, R. J. Heath, I. A. C. Pereira and J. D. Sutherland, *Bioorg. & Med. Chem. Lett.*, **2**, 669-672 (1992)

121. J. E. Baldwin, R. M. Adlington, N. P. Crouch and I. A. C. Pereira, *J. Chem. Soc., Chem. Commun.*, 1448-1451 (1992)

122. J. E. Baldwin, R. M. Adlington, N. P. Crouch and I. A. C. Pereira, *Tetrahedron*, **49**, 4907-4922 (1993)

123. L. D. Thornburg and J. Stubbe, *J. Am. Chem. Soc.*, **111**, 7632-7633 (1989)

124. A. J. Castellino and T. C. Bruice, *J. Am. Chem. Soc.*, **110**, 7512-7519 (1988)

125. J. E. Baldwin and E. P. Abraham, *Nat. Prod. Rep.*, 129-145 (1988)

126. J. E. Baldwin, R. M. Adlington, N. P. Crouch, J. W. Keeping, S. W. Leppard, J. Pitlik, C. J. Schofield, W. J. Sobey and M. E. Wood, *J. Chem. Soc., Chem. Commun.*, 768-770 (1991)

127. C.-K. Liu, C.-A. Hsu and M. T. Abbott, *Arch. Biochem. Biophy.*, **159**, 180-187 (1973)

128. J. E. Baldwin, R. M. Adlington, C. J. Schofield, W. J. Sobey and M. E. Wood, *J. Chem. Soc., Chem. Commun.*, 1012-1015 (1989)

129. V. J. Chen, A. M. Orville, M. R. Harpel, C. A. Frolik, K. K. Surerus, E. Münck and J. D. Lipscomb, *J. Biol. Chem.*, **264**, 21677-21681 (1989)

130. V. J. Chen, C. A. Frolik, S. Samson, P. D. Gesellchen, A. M. Orville, M. R. Harpel, J. D. Lipscomb, K. K. Surerus and E. Münck, *J. Inorg. Biochem.*, **36**, H052 (1989)

131. L.-J. Ming, L. Que, A. Kriauciunas, C. A. Frolik and V. J. Chen, *Inorg. Chem.*, **29**, 1111-1112 (1990)

132. J. E. Baldwin, M. D. Lloyd, B.-W. Son, C. J. Schofield, T. Sewell, S. W. Elson, K. H. Baggaley and N. H. Nicholson, *J. Chem. Soc., Chem. Commun.*, 500-502 (1993)

133. J. E. Baldwin, R. M. Adlington, N. P. Crouch, I. A. C. Pereira, R. T. Aplin and C. Robinson, *J. Chem. Soc., Chem. Commun.*, 105-108 (1993)

134. J. E. Baldwin, R. M. Adlington, N. P. Crouch and C. J. Schofield, *Tetrahedron*, **44**, 643-650 (1988)

135. M. Liersch, J. Nüesch and H. J. Treichler, in *Proceedings of the Second International Symposium on the Genetics of Industrial Microorganisms* : (ed. K. D. McDonald) pp.179-195 (Academic Press, London, 1976)

136. Y. Fujisawa, H. Shirafuji, M. Kida, K. Nara, M. Yoneda and T. Kanzaki, *Agric. Biol. Chem.*, **39**, 1295 (1975)

137. Y. Fujisawa and T. Kanzaki, *J. Antibiot.*, **28**, 372-378 (1975)

138. S. W. Queener, O. K. Sebek and C. Vézina, in *Annual Review of Microbiology* : pp.593-636 (Annual Reviews Inc., 1978)

139. L. N. Jungheim and R. J. Ternansky, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.306-324 (Chapman & Hall, London, 1992)

140. J. E. Baldwin, G. P. Lynch and J. Pitlik, *J. Antibiot.*, **44**, 1-24 (1991)

141. R. A. Pascal, M. A. Oliver and Y. C. J. Chen, *Biochemistry*, **24**, 3158-3165 (1985)

142. S. Udenfriend, C. T. Clark, J. Axelrod and B. B. Brodie, *J. Biol. Chem.*, **208**, 731-739 (1954)

143. B. B. Brodie, J. Axelrod, P. A. Shore and S. Udenfriend, *J. Biol. Chem.*, **208**, 741-750 (1954)

144. J. E. Baldwin, R. M. Adlington, J. S. Bryans, M. D. Lloyd, T. J. Sewell, C. J. Schofield, K. H. Baggaley and R. Cassels, *J. Chem. Soc., Chem. Commun.*, 877-879 (1992)

145. R. Eisenthal and A. Cornish-Bowden, *Biochem. J.*, **139**, 715-720 (1974)

146. A. Cornish-Bowden and R. Eisenthal, *Biochem. J.*, **139**, 721-730 (1974)

147. J. E. Baldwin, J. B. Coates, M. G. Moloney, A. J. Pratt and A. C. Willis, *Biochem. J.*, **266**, 561-567 (1990)

148. L.-J. Ming, L. Que, A. Kriauciunas, C. A. Frolik and V. J. Chen, *Biochemistry*, **30**, 11653-11659 (1991)

149. C. R. Randall, Y. Zang, A. E. True, L. Que, J. M. Charnock, C. D. Garner, Y. Fujishima, C. J. Schofield and J. E. Baldwin, *Biochemistry*, **32**, 6664-6673 (1993)

150. S.-F. Ng, H. M. Hanauske-Abel and S. Englard, *J. Biol. Chem.*, **266**, 1526-1533 (1991)

151. B. T. Chait and S. B. H. Kent, *Science*, **257**, 1885-1894 (1992)

152. I. Jardine, *Methods in Enzymology*, **193**, 441-455 (1990)

153. T. R. Covey, R. F. Bonner, B. I. Shushan and J. Henion, *Rapid Commun. Mass Spectrom.*, **2**, 249-256 (1988)

154. H. E. Witkowska, B. N. Green and S. Smith, *J. Biol. Chem.*, **265**, 5662-5665 (1990)

155. P. Caffrey, B. Green, L. Packman, B. Rawlings, J. Staunton and P. Leadlay, *Eur. J. Biochem.*, **195**, 823-830 (1991)

156. Y.-T. Li, Y.-L. Hsieh, J. D. Henion and B. Ganem, *J. Am. Soc. Mass Spectrom.*, **4**, 631-637 (1993)

157. E. C. Huang, B. N. Pramanik, A. Tsarbopoulos, P. Reichert, A. K. Ganguly, P. P. Trotta and T. L. Nagabhushan, *J. Am. Soc. Mass Spectrom.*, **4**, 624-630 (1993)

158. S. M. Arfin and R. A. Bradshaw, *Biochemistry*, **27**, 7979-7984 (1988)

159. C. Flinta, B. Persson, H. Jörnvall and G. von Heijne, *Eur. J. Biochem.*, **154**, 193-196 (1986)

160. R. L. Kendall and R. A. Bradshaw, *J. Biol. Chem.*, **267**, 20667-20673 (1992)

161. A. Yoshida and M. Lin, *J. Biol. Chem.*, **247**, 952-957 (1972)

162. S. Tsunasawa and F. Sakiyama, *Methods in Enzymology*, **106**, 165-170 (1984)

163. F. Wold, *Ann. Rev. Biochem.*, **50**, 783-814 (1981)

164. P. Matsudaira, *Methods in Enzymology*, **182**, 602-613 (1990)

165. M. Landon, *Methods in Enzymology*, **47**, 145-149 (1976)

166. J. Jauregui-Adell and J. Marti, *Anal. Biochem.*, **69**, 468-473 (1975)

167. D. Piszkiewicz, M. Landon and E. L. Smith, *Biochem. Biophy. Res. Commun.*, **40**, 1173-1178 (1970)

168. M. L. Bender and R. J. Thomas, *J. Am. Chem. Soc.*, **83**, 4183-4189 (1961)

169. K. Majamaa, H. M. Hanauske-Abel, V. Günzler and K. Kivirikko, *Eur. J. Biochem.*, **138**, 239-245 (1984)

170. K. Majamaa, V. Günzler, H. M. Hanauske-Abel, R. Myllylä and K. Kivirikko, *J. Biol. Chem.*, **261**, 7819-7823 (1986)

171. G. Tschank, M. Raghunath, V. Günzler and H. M. Hanauske-Abel, *Biochem. J.*, **248**, 625-633 (1987)

172. K. Majamaa, T. M. Turpeenniemi-Hujanen, P. Latipää, V. Günzler, H. M. Hanauske-Abel, I. E. Hassinen and K. Kivirikko, *Biochem. J.*, **229**, 127-133 (1985)

173. J. E. Baldwin, C. C. Lawrence and C. J. Schofield, *unpublished results*

174. Y.-C. Cheng and W. H. Prusoff, *Biochem. Pharmacol.*, **22**, 3099-3108 (1973)

175. K. D. Karlin, *Science*, **261**, 701-708 (1993)

176. M. Pänkäläinen and K. I. Kivirikko, *Biochim. Biophy. Acta*, **229**, 504-508 (1971)

177. W. W. Fish, *Methods in Enzymology*, **158**, 357-364 (1988)

178. J. M. Ward and J. E. Hodgson, *FEMS Microbiol. Lett.*, **110**, 239-242 (1993)

179. T. T. Howarth, A. G. Brown and T. J. King, *J. Chem. Soc., Chem. Commun.*, 266-267 (1976)

180. D. Brown, J. R. Evans and R. A. Fletton, *J. Chem. Soc., Chem. Commun.*, 282-283 (1979)

181. E. Arribas, C. Carreiro and A. M. Valdeolmillos, *Tetrahedron Lett.*, **29**, 1609-1612 (1988)

182. T. Konosu and S. Oida, *Chem. Pharm. Bull.*, **39**, 2212-2215 (1991)

183. P. H. Bentley and E. Hunt, *J.C.S. Perkin I*, 2222-2227 (1980)

184. M. Kellett, D. Pruess and J. P. Scannell. *2-(3-Alanyl)clavam antibiotic*, US Patent 4 202 819 (1980)

185. S. De Bernardo, J. P. Tengi, G. J. Sasso and M. Weigele, *J. Org. Chem.*, **50**, 3457-3462 (1985)

186. M. Wanning, H. Zähner, B. Krone and A. Zeeck, *Tetrahedron Lett.*, **22**, 2539-2540 (1981)

187. H. Zähner, M. Wanning, B. Krone, A. Zeeck and H. Peter. *Clavam derivatives*, European Patent EP 0 057 664 A2 (1982)

188. H. Peter and H. Zähner. *Clavam compounds*, German Patent DE 3427651 (1985)

189. D. Hoppe and T. Hilpert, *Tetrahedron*, **43**, 2467-2474 (1987)

190. H. D. King, J. Langhärig and J. J. Sanglier, *J. Antibiot.*, **39**, 510-515 (1986)

191. H. U. Naegeli, H.-R. Loosli and A. Nussbaumer, *J. Antibiot.*, **39**, 516-524 (1986)

192. C. M. Vidal. *Fermentative manufacture of clavulanic acid, using Streptomyces jumonjinensis*, Spanish Patent ES 550549 (1987)

193. K. Kitano, K. Kintaka and K. Katamoto. *Clavulanic acid production by Streptomyces katsurahamanus*, Japanese Patent JP 53-104769 (1978)

194. Y. Kawamura, E. Kondo, K. Matsumoto, N. Tsuji and K. Tawara. *Clavam antibiotic CA-31*, Japanese Patent JP 62-10089 (1987)

195. Y. Kawamura, E. Kondo, K. Matsumoto, N. Tsuji and K. Tawara. *Clavam antibiotics CA-146A and CA-146B*, Japanese Patent JP 62-10088 (1987)

196. R. F. Pratt, in *The chemistry of β-lactams* : (ed. M. I. Page) pp.229-271 (Chapman & Hall, London, 1992)

197. D. L. Pruess and M. Kellett, *J. Antibiot.*, **36**, 208-212 (1983)

198. F. Röhl, J. Rabenhorst and H. Zähner, *Arch. Microbiol.*, **147**, 315-320 (1987)

199. S. W. Elson and R. S. Oliver, *J. Antibiot.*, **31**, 586-592 (1978)

200. I. Stirling and S. W. Elson, *J. Antibiot.*, **32**, 1125-1129 (1979)

201. C. A. Townsend and M.-F. Ho, *J. Am. Chem. Soc.*, **107**, (1985)

202. S. W. Elson *et al, unpublished results*

203. A. L. Gutman, V. Ribon and A. Boltanski, *J. Chem. Soc., Chem. Commun.*, 1627-1629 (1985)

204. C. A. Townsend and S.-S. Mao, *J. Chem. Soc., Chem. Commun.*, 86-89 (1987)

205. S. W. Elson and R. S. Oliver, *J. Antibiot.*, **35**, 81-86 (1982)

206. C. A. Townsend and M.-F. Ho, *J. Am. Chem. Soc.*, **107**, 1065-1066 (1985)

207. J. Romero, P. Liras and J. F. Martin, *Appl. Environ. Microbiol.*, **52**, 892-897 (1986)

208. B. W. Bycroft, A. Penrose, J. Gillett and S. W. Elson, *J. Chem. Soc., Chem. Commun.*, 980-981 (1988)

209. S. W. Elson, K. H. Baggaley, M. Davison, M. Fulston, N. H. Nicholson, G. D. Risbridger and J. W. Tyler, *J. Chem. Soc., Chem. Commun.*, 1212-1214 (1993)

210. S. W. Elson, K. H. Baggaley, M. Fulston, N. H. Nicholson, J. W. Tyler, J. Edwards, H. Holms, I. Hamilton and D. M. Mousdale, *J. Chem. Soc., Chem. Commun.*, 1211-1212 (1993)

211. J. M. Williamson, *CRC Critical Reviews in Biotechnology*, **4**, 111-128 (1986)

212. A. Earl *et al*, *unpublished results*

213. E. Soru, *J. Chromatogr.*, **20**, 325-333 (1965)

214. E. Soru and O. Zaharia, *Rev. Roum. Biochim.*, **13**, 49-60 (1976)

215. S. P. Salowe, W. J. Krol, D. Iwata-Reuyl and C. A. Townsend, *Biochemistry*, **30**, 2281-2292 (1991)

216. E. N. Marsh, M. D.-T. Chang and C. A. Townsend, *Biochemistry*, **31**, 12648-12657 (1992)

217. C. A. Townsend and W. J. Krol, *J. Chem. Soc., Chem. Commun.*, 1234-1236 (1988)

218. C. A. Townsend, M.-F. Ho and S.-S. Mao, *J. Chem. Soc., Chem. Commun.*, 638-639 (1986)

219. H. Zähner, *personal communication to Dr C. J. Schofield*

220. R. Nagarajan, L. D. Boeck, M. Gorman, R. L. Hamill, C. E. Higgens, M. M. Hoehn, W. M. Stark and J. G. Whitney, *J. Am. Chem. Soc.*, **93**, 2308-2310 (1971)

221. C. E. Higgens, R. L. Hamill, T. H. Sands, M. M. Hoehn, N. E. Davis, R. Nagarajan and L. D. Boeck, *J. Antibiot.*, **27**, 298-300 (1974)

222. M. Kenig and C. Reading, *J. Antibiot.*, **32**, 549-554 (1979)

223. D. L. Pruess, *personal communication to Prof J. E. Baldwin*

224. A. E. Bird, J. M. Bellis and B. C. Gasson, *Analyst*, **107**, 1241-1245 (1982)

225. S.-S. Mao, *Ph.D. Thesis,* The Johns Hopkins University (1987)

226. J. Rabenhorst, *Ph.D. Thesis,* University of Tübingen (1986)

227. H. Peter, *personal communication to Dr C. J. Schofield*

228. J. W. Payne and C. Gilvarg, in *Bacterial transport* : (ed. B. P. Rosen) pp.325-383 (Marcel Dekker, New York, 1978)

229. A. Bax and S. Subramanian, *J. Magn. Reson.*, **67**, 565-569 (1986)

230. A. Bax and M. F. Summers, *J. Am. Chem. Soc.*, **108**, 2093-2094 (1986)

231. V. Lee and C. J. Schofield, *personal communication*

232. A. Kern, G. Bovermann, G. Jung, M. Wanning and H. Zähner, *Liebigs Ann. Chem.*, 361-365 (1989)

233. H.-T. Postels and W. A. König, *Liebigs Ann. Chem.*, 1281-1287 (1992)

234. Y. Yanuka, R. Katz and S. Sarel, *Tetrahedron Lett.*, **14**, 1725-1728 (1968)

235. H. Kleinkauf and H. von Döhren, in *Annual Review of Microbiology* : pp.259-289 (Annual Reviews Inc., 1987)

236. C. A. Townsend, *Biochem. Soc. Trans.*, **21**, 208-213 (1993)

237. D. Iwata-Reuyl and C. A. Townsend, *J. Am. Chem. Soc.*, **114**, 2762-2763 (1992)

238. F. Petit and R. Furstoss, *Tetrahedron: Asymmetry*, **4**, 1341-1352 (1993)

239. M. J. Taschner, D. J. Black and Q.-Z. Chen, *Tetrahedron: Asymmetry*, **4**, 1387-1390 (1993)

240. J. M. Williamson, E. Inamine, K. E. Wilson, A. W. Douglas, J. M. Liesch and G. Albers-Schönberg, *J. Biol. Chem.*, **260**, 4637-4647 (1985)

241. D. R. Houck, K. Kobayashi, J. M. Williamson and H. G. Floss, *J. Am. Chem. Soc.*, **108**, 5365-5366 (1986)

242. K. E. Wilson, A. J. Kempf, J. M. Liesch and B. H. Arison, *J. Antibiot.*, **36**, 1109-1117 (1983)

243. W. J. Krol, A. Basak, S. P. Salowe and C. A. Townsend, *J. Am. Chem. Soc.*, **111**, 7625-7626 (1989)

244. C. A. Townsend and A. Basak, *Tetrahedron*, **47**, 2591-2602 (1991)

245. A. Basak, S. P. Salowe and C. A. Townsend, *J. Am. Chem. Soc.*, **112**, 1654-1656 (1990)

246. J. E. Baldwin, K. D. Merritt, C. J. Schofield, S. W. Elson and K. H. Baggaley, *J. Chem.*

*Soc., Chem. Commun.*, 1301-1302 (1993)

247. S. W. Elson, K. H. Baggaley, S. Holland, N. H. Nicholson, J. T. Sime and S. R. Woroniecki, *Bioorg. & Med. Chem. Lett.*, **2**, 1503-1508 (1992)

248. D. Iwata-Reuyl, A. Basak, L. S. Silverman, C. A. Engle and C. A. Townsend, *J. Nat. Prod.*, **56**, 1373-1396 (1993)

249. M. Ito-Kagawa, Y. Koyama and S. Kondo, *J. Antibiot.*, **37**, 487-493 (1984)

250. Y. Fujita, *Bull. Chem. Soc. Jpn.*, **33**, 1379-1381 (1960)

251. E. A. Bell and A. S. L. Tirimanna, *Nature*, **197**, 901-902 (1963)

252. K. H. Baggaley, *personal communication to Dr C. J. Schofield*

## PUBLICATIONS ARISING FROM THIS THESIS

1. 'The Role of Sulphur in Cephalosporin Biosynthesis'
   J. E. Baldwin, K.-C. Goh, M. E. Wood, C. J. Schofield, R. D. G. Cooper and G. W. Huffman, *Bioorg. & Med. Chem. Lett.*, **1**, 421-424 (1991)

2. 'Oxidation of Deacetylcephalosporin C by DAOC/DAC Synthase'
   J. E. Baldwin, K.-C. Goh and C. J. Schofield, *J. Antibiot.*, **45**, 1378-1381 (1992)

3. 'Revised Structures for Tü 1718B and Valclavam'
   J. E. Baldwin, T. D. W. Claridge, K.-C. Goh, J. W. Keeping and C. J. Schofield, *Tetrahedron Lett.*, **34**, 5645-5648 (1993)