Studies on Enzymes Mechanism and Selectivity Using Synthetic Substrate Analogues

A thesis submitted to the Board of the Faculty of Physical Sciences of the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Luc Henry

Lincoln College, Oxford

Trinity Term 2012

STUDIES ON ENZYMES MECHANISM AND SELECTIVITY USING SYNTHETIC SUBSTRATE ANALOGUES

Abstract

Organic chemistry is a valuable tool for studying enzyme mechanisms. Upon incubation with a specific enzyme, synthetic substrate analogues labeled with heavy atoms or carrying extra functional groups can provide mechanistic insights. In the present work, new compounds were synthesised in order to study the mechanism and substrate selectivity of two enzymes: human γ-butyrobetaine hydroxylase and bacterial carboxymethylproline synthase.

γ-Butyrobetaine hydroxylase (BBOX) is an Fe(II) and 2-oxoglutarate (2OG)-dependent oxygenase that catalyses the stereospecific hydroxylation of γ-butyrobetaine, the final step of L-carnitine (L-Car) biosynthesis in mammals. Substrate analogues were synthesised to probe BBOX specificity *in vitro*. Some of those unnatural substrates were oxidised by BBOX and the products identified using a range of analytical techniques. 3-(2,2,2-Trimethylhydrazinium)propionate (THP) is a clinically used BBOX inhibitor. Under standard assay conditions, THP was oxidised by BBOX. NMR studies have identified the products of this reaction to be malonic acid semialdehyde, formaldehyde, dimethylamine and 3-amino-4-(methylamino)butanoic acid. The formation of 3-amino-4-(methylamino)butanoic acid suggests that BBOX can catalyse a Stevens type rearrangement involving N-N bond cleavage and C-C bond formation. The proposed structures and mechanisms were confirmed by mass spectrometric and NMR analyses using [^{13}C]-labeled THP as well as synthetic standards of both enantiomers of 3-amino-4-(methylamino)butanoic acid. Although the structure of the rearrangement product was confirmed, the stereochemistry remains unknown. Altogether, these studies revealed the unprecedented nature of a BBOX-catalysed C–C bond formation reaction upon THP oxidation and may inspire the design of improved inhibitors for BBOX and other 2OG oxygenases.

Pectobacterium carotovorum CarB and *Streptomyces cattleya* ThnE are two carboxymethylproline synthases (CMPS) that catalyse an early step in carbapenem antibiotics biosynthesis. CMPS produces (2*S*,5*S*)-carboxymethylproline (*t*-CMP) from malonyl-CoA and L-glutamate semi-aldehyde. L-Glutamate semi-aldehyde exists in equilibrium with L-5-hydroxyproline and L-pyrroline-5-carboxylate in solution (collectively abbreviated L-GHP). Because of the high stereoselectivity of *t*-CMP formation and the growing interest in novel carbapenem antibiotics, CMPS is potentially an interesting biocatalyst. A series of L-GHP analogues were synthesised and tested as CMPS substrates in an attempt to produce unnatural *t*-CMP derivatives enzymatically. Methyl-substituted L-GHP analogues were accepted by CMPS and the *t*-CMP products could be further carried through to the corresponding bicyclic carbapenams using CarA, a β-lactam synthetase. These results demonstrate the versatility of the early carbapenem biosynthetic pathway and the possibility of introducing structural diversity using synthetic substrate analogues. A crystal structure of *S. cattleya* ThnE was obtained in complex with L-proline and coenzyme A, giving the first insight into substrate binding. This structural information will potentially allow further rational mutagenesis studies aiming to broaden the range of unnatural L-GHP analogues accepted by CMPS.

À mes parents, Janine et Jacques-André Henry, à qui je dois absolument tout.

Acknowledgements

The work presented in this thesis was made possible by a Lord Florey fellowship from the Berrow Foundation (Lincoln College, Oxford). I would like to thank the trustees of the Berrow Foundation for this generous award that gave me the opportunity to come to Oxford and has supported me during the first three years of my DPhil studies.

In addition, the last months of my stay in Oxford have been founded by a generous scholarship from the Vice Chancellors' Fund at the University of Oxford. Therefore I would like to extend my sincere gratitude to the trustees of this fund.

***

I would like to thank my supervisor, Prof Christopher J. Schofield, for giving me the opportunity to spend over three years in his laboratory, for his support, precious advice and for the possibility he gave me to freely choose the direction of my research.

I would also like to extend my thanks to all members of the Schofield group, from October 2008 till March 2012, and in particular to the following people in the Chemistry Research Laboratory: J. Ruben Gomez-Castellanos for being a highly productive and motivated coworker, in addition to being a good friend, Dr Refaat B. Hamed for giving me the opportunity to work on the biocatalysis project which has been an extremely productive collaboration, for his advice, skills and experience with these difficult CMPS assays, Ivanhoe K. H. Leung for inviting me into the BBOX project and for his motivation in taking it further, as well as his expertise in NMR techniques.

The excellent and crucial technical support was provided by Wendy Sobey and Herminio Mansojubier within the Schofield group. Practical support was supplied by Dr James McCullagh, Lingzhi Gong, Colin Sparrow, James Wickens and Robin Procter of the CRL Mass Spectrometry service as well as Dr Barbara Odell and Dr Tim Claridge of the CRL NMR service.

The following people spent a great deal of their precious time proof-reading this thesis: Ivanhoe K. H. Leung, Dr Refaat B. H. Hamed, Dr Richard J. Hopkinson, Dr Adam P. Hardy, Dr Michael A. McDonough, Dr Emily Flashman and Dr Sander S. Van Berkel.

***

In no particular order, I would like to add the people I have met outside the laboratory and who made my limited free time absolutely priceless: Tom, Serge, Josephine, Dorothea, Bonnie, Georges, Sarah, Gabrielle, Dante, Guy, Xavier, Justyna, Alain, Philippe, Jérémie, Cecilia, Anne-Claire and Tonya.

Finally, I would like to thank Marc and Gaetan amongst all the people I left behind in Switzerland, one more time, or simply neglected to visit elsewhere, for still today being the friends I knew before embarking on this scientific quest. I would especially like to thank all the members of my family for their continuous support during the past years.

Tout le reste n'est que garniture.

Collaborations

I have to give credit to the following people for their significant contribution to the work presented in this thesis. In addition, details for their involvement in the obtention of the results described later are given in the beginning of each chapter.

The results presented in Chapter 1 were obtained in collaboration with Dr Grazyna Kochan (human BBOX expression and purification), Dr Tobias J. Krojer (human BBOX crystallization and structure elucidation), Ivanhoe K. H. Leung (NMR assays and kinetic analysis, as well as NMR characterization of the THP reaction products), Dr Filippo Sladojevich (chiral normal phase HPLC), Nikita Loic (advice for the mass spectrometry analyses and HPLC separation of the 3-amino-4-(methylamino)butanoic acid enantiomers), Anna Rydzik (helpful discussion and follow up of the project), Dr Michael McDonough (analysis of the crystallographic data), Dr Timothy D. W. Claridge, Dr Frank von Delft, Prof Udo Oppermann and Prof Christopher J. Schofield (data analysis and supervision).

The results presented in Chapters 3 to 6 were obtained in collaboration with Dr Christian Ducho, Dr John L. Sorensen, Dr Jasmin Mecinović (preliminary synthetic studies on substituted L-GHP analogues), Dr Refaat B. Hamed (CMPS and β-LS assays preparation, products isolation and characterisation), J. Ruben Gomez-Castellanos, Daniel Harding (preparation and recombinant expression of CMPS variants as well as MatB protein), Dr Michael McDonough (collection and preliminary analysis of crystallographic data), Dr Timothy D. W. Claridge and Prof Christopher J. Schofield (data analysis and supervision). In addition, I would like to thank Prof. Manuela Tosin (formerly at the University of Cambridge, UK, now at the University of Warwick, UK) for her generous gift of the expression vectors containing the *coaA*, *coaD* and *coaE* genes.

The work presented in Chapter 7 was made possible thanks to the initial cloning work for the *thnN* and *thnO* genes from *S. cattleya* genomic DNA done by Dr Refaat B. Hamed. The LC-MS analyses of ThnN and ThnO purified proteins was done by Dr Adam Hardy. The tryptic digest experiment on ThnN was performed with the help of Dr Alexander Wolf and further MS-MS analysis was carried out by Dr Mukram M. Mackeen at the Henry Wellcome Building for Molecular Physiology, Department of Clinical Medicine, University of Oxford. In addition, Ivanhoe K. H. Leung (advice on NMR experiments) and Prof Christopher J. Schofield (data analysis and supervision) have contributed to this work.

I certify that the work described in this thesis is entirely my own, except where I have either acknowledge help from a named person or given reference to a previously published source. Text taken from another source is enclosed in quotation marks and a reference is given.

Oxford, United Kingdom, July 2012

Contents

List of Abbreviations, Acronyms and Symbols

β-LS	β-Lactam synthetase
2OG	2-Oxoglutarate
ACC	Acetyl-CoA carboxylase
ACP	Acyl-carrier protein
ACN	Acetonitrile
ACV	δ-(L-α-Aminoadipoyl)-L-cysteinyl-D-valine
ADP	Adenosine-5′-diphosphate
AMP	Adenosine-5′-monophosphate
amu	Atomic mass unit
app. quant.	Apparent quantitative
APS	Ammonium persulfate
ATP	Adenosine-5′-triphosphate
γ-BB	γ-Butyrobetaine
BBOX	γ-Butyrobetaine hydroxylase
BEVS	Baculovirus expression vector system
BHL	*N*-Butanoyl-L-homoserine lactone
Calc.	Calculated
CAR	Carboxylic acid reductase
CAS	Clavaminic acid synthase
Ccr	Crotonyl-CoA carboxylase reductase
CD	Circular dichroism
CDI	*N*,*N*'-Carbonyldiimidazole
CMP	Carboxymethylproline
CMPS	Carboxymethylproline synthase
CoASH	Coenzyme A
CS	Crotonase superfamily
CTP	Cytidine-5′-triphosphate
D_2O	Deuterium oxide
Da	Dalton = atomic mass unit
DCM	Dichloromethane
DIBAL	Diisobutylaluminium hydride
DIPEA	*N*,*N*-Diisopropylethylamine
DMP	Dess-Martin periodinane
DMSO	Dimethylsulfoxide
DMSO-d_6	Deuterated dimethylsulfoxide
DNA	Deoxyribonucleic acid
d.r.	Diastereomeric ratio
DSBH	Double-stranded beta-helix
DTT	Dithiothreitol
ECC	Enzyme cascade catalysis
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
e.g.	Exempli gratia (for example)
ELSD	Evaporative light scattering detector
ESI	Electrospray ionisation
Et_2O	Diethylether

Et_3N	Triethylamine
EtOAc	Ethyl Acetate
EtOH	Ethanol
ESI	Electrospray ionization
EWG	Electron withdrawing group
FA	Fatty acid
FAS	Fatty acid synthase
FCC	9-Fluorenone-4-carbonyl chloride
FDAA	N_α-(2,4-Dinitro-5-fluorophenyl)-L-alaninamide
FDVA	N_α-(2,4-Dinitro-5-fluorophenyl)-L-valinamide
FLEC	(+)-1-(9-Fluorenyl)ethyl chloroformate
Fmoc	Fluorenylmethyloxycarbonyl
FPLC	Fast protein liquid chromatography
GABA	γ-Aminobutyric acid
L-GHP	Equilibrium mixture of L-GSA, L-5HP and L-P5C
L-GSA	L-Glutamate semialdehyde
GTP	Guanosine-5′-triphosphate
HBTU	*O*-Benzotriazole-*N*,*N*,*N*',*N*'-tetramethyl-uronium-hexafluoro-phosphate
HCA	Hexachloroacetone
Hex	Hexane
HHL	*N*-Hexanoyl-L-homoserine lactone
HMBC	Heteronuclear multiple bond correlation
L-5HP	5-Hydroxy-L-proline
HOBt	Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HR	High resolution
HTML	3-Hydroxy-N^ϵ-trimethyllysine
HTMLA	3-Hydroxy-N^ϵ-trimethyllysine aldolase
HWE	Horner-Wadsworth-Emmons
*i*Bu	Isobutyl
Incub.	Incubation
*i*Pr	Isopropyl
IPNS	Isopenicillin N synthase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
K_D	Dissociation constant
LAT	Lysine aminotransferase
LC	Liquid chromatography
LDA	Lithium diisopropylamide
LR	Low resolution
m/z	Mass over charge ratio
MAHT	Malonic acid half thioester
MeOH	Methanol
mp	Melting point
MS	Mass spectrometry
MTS	Mitochondrial targeting sequence
MW	Microwave
n/a	Non applicable
NAD(P)	Nicotinamide adenine dinucleotide (phosphate)
NBS	*N*-Bromosuccinimide

NCBI	National center for biotechnology information
N.D.	Not determined
n.i.	Not isolated
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NOG	*N*-Oxalylglycine
NMR	Nuclear magnetic resonance
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OAT	Ornithine aminotransferase
OD	Optical Density
OHHL	*N*-(3-Oxohexanoyl)-L-homoserine lactone
L-P5C	L-Pyrroline-5-carboxylate
P5P	Pyridoxal-5′-phosphate
PBP	Penicillin-binding protein
PDB	Protein data bank
PDCA	Pyridine dicarboxylate
PHD	Plant homeodomain
PHD2	Prolyl hydroxylase 2
PK	Polyketide
PKS	Polyketide synthase
PLE	Porcine liver esterase
PPARα	Peroxisome proliferator-activated receptors α
PPh_3	Triphenylphosphine
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
R_f	Retention factor
rt	Room temperature
SAM	*S*-Adenosylmethionine
SAR	Structure-activity relationship
SDS	Sodium monododecyl sulfate
SNAc	*N*-Acetylcysteamine
t-Bu	*tert*-Butyl
TBAF	Tetra-*n*-butylammonium fluoride
TBDMS	*tert*-Butyldimethylsilyl
TCA	Tricarboxylic acid
TEMED	*N*,*N*,*N*',*N*'-Tetramethyl-ethylenediamine
Temp.	Temperature
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
THP	3-(2,2,2-Trimethylhydrazinium)propionate
TLC	Thin layer chromatography
TMABA	N^{δ}-Trimetylaminobutyraldehyde
TMABADH	N^{δ}-Trimetylaminobutyraldehyde dehydrogenase
TML	N^{ϵ}-Trimethyllysine
TMLH	N^{ϵ}-Trimethyllysine hydroxylase
TOCSY	Total correlation spectroscopy
TOF	Time of flight
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
p-Ts	*p*-Toluenesulfonyl
UTP	Uridine-5′-triphosphate

SI units, standard 1- or 3-letters amino acids abbreviations and standard DNA abbreviations are used throughout.

NMR multiplicities are as follow: *s*: singlet; *bs*: broad singlet; *d*: doublet; *t*: triplet; *q*: quadruplet; *quint*: quintuplet; *spt*: septuplet; *m*: multiplet; *app*: apparent.

Chapter 1

Biochemical Studies on Human γ-Butyrobetaine Hydroxylase

1.1 General Introduction

1.1.1 L-Carnitine: Biological Roles and Distribution

L-Carnitine **1** ((3*R*)-3-hydroxy-4-(trimethylammonio)butanoate or L-Car, Figure 1.1) is a low molecular weight zwitterionic quaternary amine that is found in all forms of life, spanning from the simplest bacteria to mammals. Investigations of the biological role of L-Car **1** only started in the 1950s, years after it was first isolated from muscle extracts in 1905 and its structure was determined in 1927 [1, 2].

L-Car belongs to a family of neutral chemical compound called betaines. Betaines are zwitterions and therefore contain a positively charged cationic functional group which bears no hydrogen atom, in this case a constitutively charges ammonium group, and a negatively charged functional group. Betaines are found widely in nature where they carry out crucial biological roles. In addition to betaines, other constitutively charges quaternary ammonium ions are crucial in biology. Choline **2** (2-hydroxy-*N*,*N*,*N*-trimethylethanaminium), for example, is the precursor for important components of the cell membrane and a precursor of the neurotransmitter acetylcholine **3** (2-acetoxy-*N*,*N*,*N*-trimethylethanaminium). Amongst betaines, muscarine **4** is a neurotoxic compound produced by certain mushrooms, which mimics acetylcholine binding to muscarinic acetylcholine receptors,

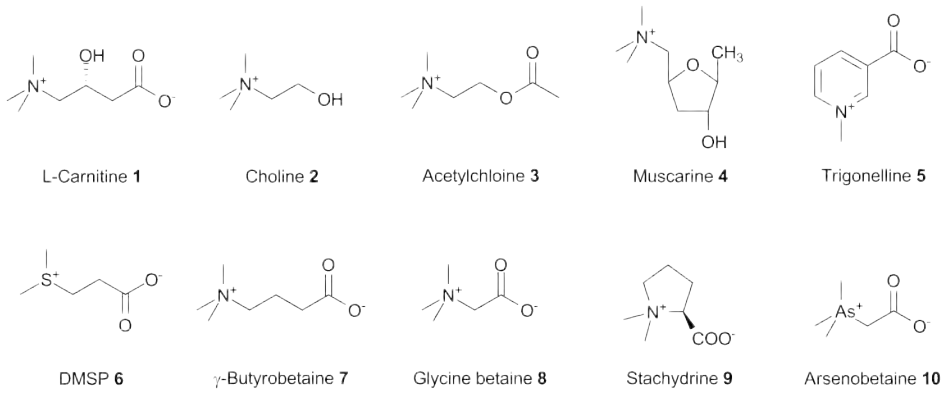


Figure 1.1: Structure of L-carnitine and other natural betaines. L-Carnitine **1** and γ-butyrobetaine **7** are the two betaines relevant to my studies. The structures of other naturally occurring betaines are given for comparison: choline **2**, acetylcholine **3**, muscarine **4**, trigonelline **5**, dimethylsulfoniopropionate **6** (DMSP), glycine betaine **8**, stachidrine **9** and arsenobetaine **10**.

trigonelline **5** is a metabolite of niacin and dimethylsulfoniopropionate **6** (DMSP) is an intermediate in the carbon and sulfur fixation by marine phytoplankton [3] (Figure 1.1). The occurrence and roles of naturally occurring quaternary ammonium ions are highly diverse and have been reviewed by Anthoni et al. [4].

Biological Roles of L-Carnitine

In prokaryotes, L-Car **1** is proposed to serve as an osmoprotectant, in a similar manner to other quaternary ammonium ions, *e.g.* 4-(trimethylammonio)butanoate **7** (γ-butyrobetaine or γ-BB), glycine betaine **8**, proline betaine **9** (stachydrine), and arsenobetaine **10** (Figure 1.1) [4]. L-Car **1** can be accumulated by bacterial cells and actively exported across the membrane when the intracellular salt concentration is significantly above that of the environment. This mechanism allows survival in rapidly changing environments without large variations in the intracellular concentrations of functionally important ions, such as potassium, sodium and calcium. In the enterobacteria *Escherichia coli* for example, L-Car **1** is specifically and actively transported across the membrane through by CaiT, a member of the betaine/choline/carnitine transporter (BCCT) family [5].

In eukaryotes, L-Car **1** is essential for energy homeostasis and participates in the excretion of toxic metabolites. The fundamental metabolic role of L-Car in energy homeostasis is mediated through reversible esterification of its C-3 hydroxyl group [6]. L-Car *O*-acylation allows the active

and specific transport of long-chain fatty acids (12 carbons and more) as fatty acyl-L-Car species across the mitochondrial membrane (Figure 1.2). Mitochondrial β-oxidation of long-chain fatty acids is a key process in energy metabolism and is ubiquitously and strictly dependent on L-Car **1** [7]. Fatty acids are activated in the cytosol by acyl-CoA synthetases. The mitochondrial inner membrane is not permeable to the resulting fatty acyl-CoA species and an active transport involving three distinct proteins is required. Carnitine palmitoyl transferase I (CPT-I) catalyses the cytosolic transesterification of acyl-CoA into acyl-carnitine which is then transported against free carnitine or acetyl-carnitine across the membrane by a family of organic cation transporters (OCTN1-3 in humans) [8]. Carnitine palmitoyl transferase II (CPT-II) then catalyses transesterification to give fatty acyl-CoA species that enter the β-oxidation cycle in the mitochondrial matrix (Figure 1.2) [9]. An osmolyte role has also been proposed for L-Car **1** in eukaryotes, and in particular in the context of the human eye lens [10].

In humans, L-Car **1** is both obtained from the diet (75%) and produced endogenously (25%) [12]. Because of its absolute requirement for fatty acid metabolism, L-Car is known as a 'conditionally essential nutrient' [13]. The tissue distribution of L-Car **1** [14], and the defects in endogenous production or intake from external sources have been extensively studied and associated with several pathological conditions [6], including diabetes [15].

1.1.2 L-Carnitine Biosynthesis

In eukaryotes, the biosynthesis of L-Car **1** proceeds in four enzyme-catalysed steps, starting from N^{ϵ}-trimethyllysine **11** (TML) (Scheme 1.1). The literature related to the L-Car **1** biosynthesis pathway has been reviewed at several occasions, including recently by Vaz and coworkers [11, 16].

Feeding experiments using radiolabeled L-lysine in the filamentus fungus *Neurospora crassa* [17, 18, 19] and in rats [20] showed low but significant incorporation into L-Car **1** and its precursor γ-BB **7**. Broquist and coworkers first identified TML **11**, instead of L-lysine, as a precursor of L-Car biosynthesis in 1973 [18, 21]. Although labeling studies suggest that the pool of free TML **11** comes from lysosomal or proteasomal degradation of proteins containing post-translationally modified trimethyllysine residues (*e.g.* methylated histone tails [22, 23]), the identification of an

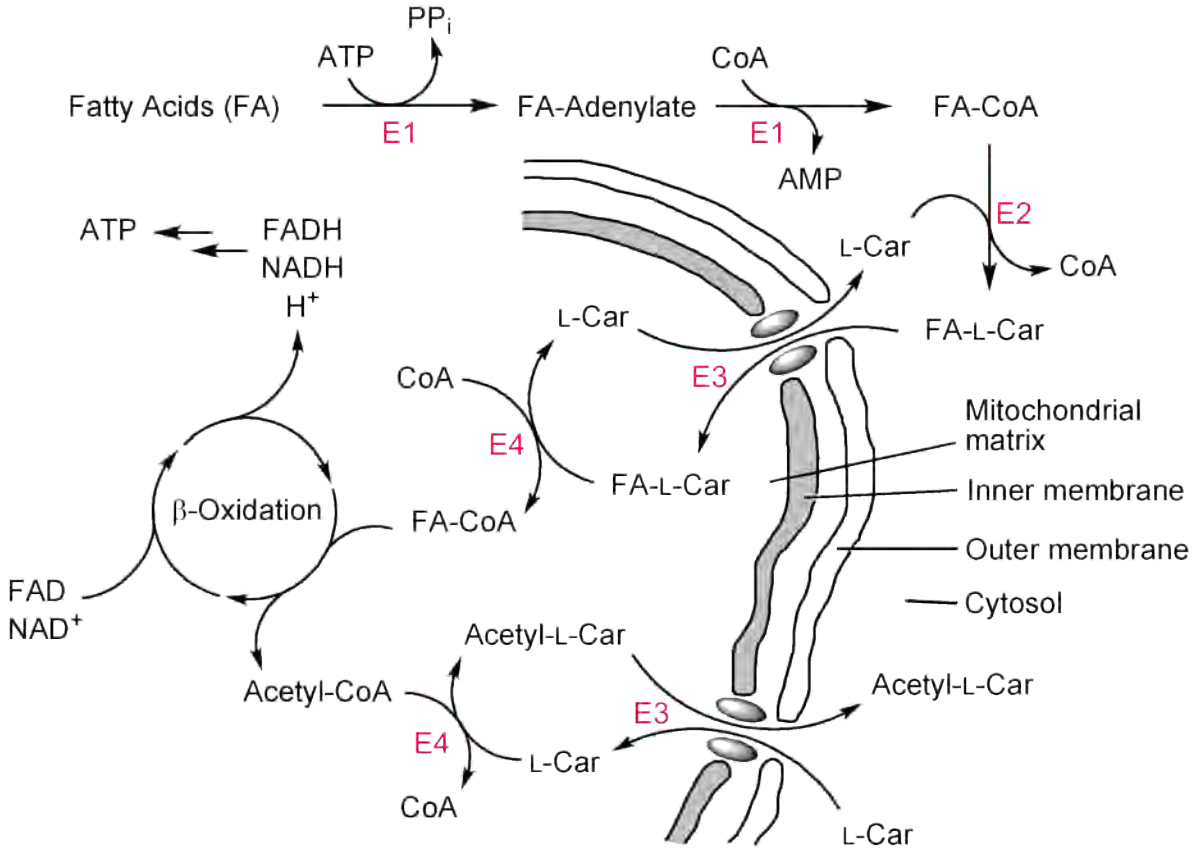


Figure 1.2: Main biological role of L-carnitine in eukaryotes. In eukaryotes, L-Car **1** is required for the active transport of fatty acids across the mitochondrial inner membrane. Fatty acids are transfered from coenzyme A onto L-Car **1** by cytosolic CPT-I and the resulting fatty acyl-carnitine (FA-L-Car) is transported across the mitochondrial membrane by OCTN3. When they reach the mitochondrial lumen, they are transformed back into fatty acyl-CoA species which enter the β-oxidation pathway to produce ATP [6, 7]. Enzymes mentioned in the text are highlighted in red. E1, acyl-CoA synthetase; E2, carnitine palmitoyl transferase I (CPT-I); E3, organic cation transporter 3 (OCTN-3); E4, carnitine palmitoyl transferase I (CPT-II). Figure modified from reference [11].

S-adenosylmethionine-dependent L-lysine methyltransferase in *N. crassa* [24, 25] suggests that there could be a L-Car biosynthetic pathway starting from free L-lysine. Some preliminary studies suggest that such a pathway exists in humans [26, 27]. If this hypothesis is demonstrated to be wrong, the capacity to produce free TML **11** by methylation of free lysine might have been lost in mammals thought reductive evolution, and more generally in species where the diet provides L-Car **1** in sufficient amounts.

A complete L-Car **1** biosynthesis pathway proposed to involve four enzymes was recently described in the yeast *Candida albicans* [28]. These four enzymes are N^{ϵ}-trimethyllysine hydroxylase (TMLH), 3-hydroxy-N^{ϵ}-trimethyllysine aldolase (HTMLA), N^{δ}-trimethylaminobutyraldehyde dehydrogenase (TMABADH) and γ-butyrobetaine hydroxylase (BBOX) (Scheme 1.1). Gene deletion strains were constructed and characterised for each of the four enzymes, showing a strict dependence and therefore non-promiscuous activity for each of them. This is in good agreement with experimental evidence showing that *C. albicans* is able to grow on minimal fatty acid medium without L-Car supplementation and suggesting that *C. albicans* has its own carnitine biosynthesis machinery. In contrast, no homologous genes were identified in the yeast *Saccharomyces cerevisiae*, in good agreement with the observation that *S. cerevisiae* is strictly dependent on L-Car supplementation for growth on fatty acids, ethanol, or acetate [28].

N^{ϵ}-Trimethyllysine Hydroxylase

N^{ϵ}-Trimethyllysine hydroxylase (TMLH, EC 1.14.11.8) is the first enzyme of the canonical L-Car biosynthesis pathway and belongs to the superfamily of 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenases [29]. TMLH introduces an hydroxyl group at the C-3 position of free TML **11** to give 3-hydroxy-N^{ϵ}-trimethyllysine **12** (HTML) (Scheme 1.1). The stereochemical course of the reaction catalysed by TMLH is unknown, although a study suggested that the product might be (3*S*)-HTML by comparison with the ^{1}H NMR coupling pattern of the two diastereomers of non-methylated 3-hydroxylysine [30]. TMLH activity was first described in rat liver mitochondrial extracts [31] and the gene associated with the enzyme function later identified and characterised in *N. crassa* [32], rats [33], and humans [34]. In contrast to the other enzymes in the L-Car biosynthesis pathway which

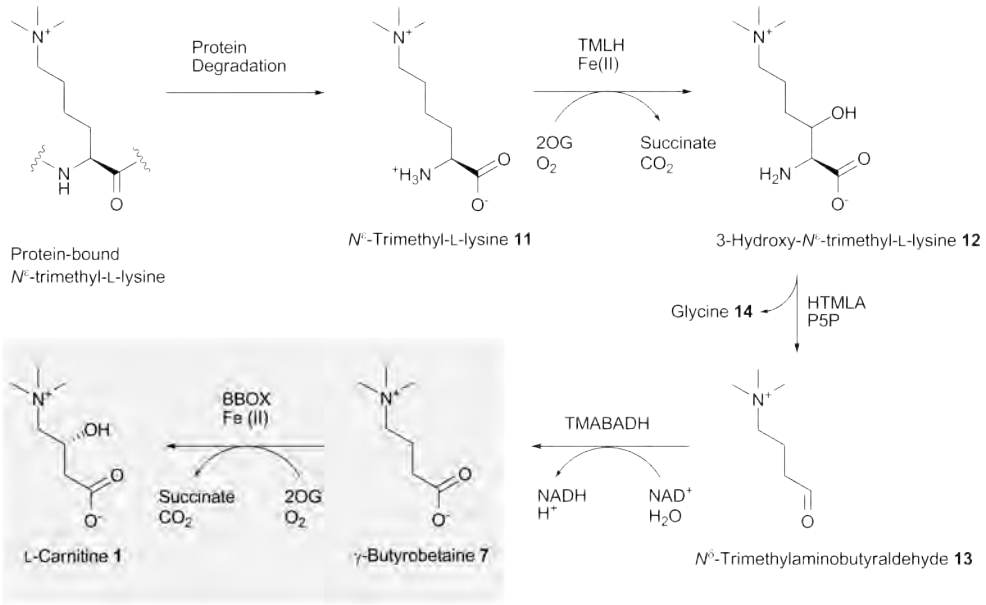
are present in the cytosol, TMLH activity was found to be associated with mitochondrial fractions in mammals [31]. The identification of the gene coding for the human TMLH enzyme enabled amino acids sequence analyses that revealed the presence of a 40 amino acids *N*-terminal mitochondrial targeting sequence (MTS) in both isoforms of the human protein [34]. The MTS is highlighted with red asterisks in Table 1.1. The presence of a MTS in the rat TMLH sequence confirmed the observation of an association with mitochondrial membranes [35]. A MTS is characterised by an alternating pattern of hydrophobic and positively charged amino acids (amphipathic helix) at the *N*-terminus of proteins directed to the mitochondrial matrix [36].

3-Hydroxy-N^{ϵ}-trimethyllysine Aldolase

The second enzyme in the L-Car biosynthesis pathway is 3-hydroxy-N^{ϵ}-trimethyllysine aldolase (HTMLA, EC 4.1.2.'X'). HTMLA activity was described in the yeast *C. albicans* [28], however it has remained elusive in mammals. No gene has been associated with the HTMLA activity found in rat and human tissues so far, although it is proposed to be a pyridoxal 5′-phosphate (P5P)-dependent aldolase. Labeling studies suggest that HTMLA catalyses the aldolytic cleavage (or retro-aldol reaction) of HTML **12** between carbons 2 and 3, generating N^{δ}-trimethylaminobutyraldehyde **13** (TMABA) and glycine **14** (Scheme 1.1) [38]. Further support for this mechanism comes from a study in which perfused rat livers were injected with proteins containing [^{3}H]-labeled trimethyllysine residues. The tissues showed accumulation of labeled 3-hydroxy-6-*N*-trimethyllysine after inhibiting P5P-dependent enzymes with 1-amino-D-proline [39]. 1-Amino-D-proline inhibits P5P-dependent enzymes by reacting with the aldehyde carbonyl group of the coenzyme to form a stable hydrazone. Interestingly, the aldolytic cleavage of HTML **12** into TMABA **13** and glycine **14** that is proposed to be catalysed by HTMLA in L-Car **1** biosynthesis was found to be performed by the enzyme glycine hydroxymethyltransferase (EC 2.1.2.1). Glycine hydroxymethyltransferase is a P5P-dependent aldolase that has been used by Henderson and coworkers in a coupled assay to quantify TMLH activity [31].

```
hBBOX1  ----------------------------------------MACTIQKAEALDGAHLMQILWYDEEESLYPAVWLRDNC  38
hTMLH   MWYHRLSHLHSRLQDLLKGGVIYPALPQPNFKSLLPLAVHWHHTASKSLTCAWQQHEDHFELKYANTVMRFDYVWLRDHC  80
        **************************************** :.: . : * :  .    * * :    :  *****:*

hBBOX1  PCSDCYLDSAKARKLLVEALDVNIGIKGLIFDRKKVYITWPDEHYSEFQADWLKKRCFSKQARAKLQRELFFPECQYWGS  118
hTMLH   RSASCYNSKTHQRSLDTASVDLCIKPKTIRLDETTLFFTWPDGHVTKYDLNWLVKNSYEGQKQKVIQPRILWNAEIYQ--  160
         .:.** ..:: *.* . ::*: *  * : :*...::::*** * :::: :** *..:. * :  :* .:::    *

hBBOX1  ELQLPTLDFEDVLRYDEHAYKWLSTLKKVGIVRLTGASDKPGEVSKLGKRMGFLYLTFYGHTWQVQDKIDANNVAYTTGK  198
hTMLH   QAQVPSVDCQSFLETNEGLKKFLQNFLLYGIAFVENVPPTQEHTEKLAERISLIRETIYGRMWYFTSDFSRGDTAYTKLA  240
        : *:*::* :..*. :*   *:*..:   **. : ... .  ...**.:*:.::  *:**: * . ..:. .:.***.

hBBOX1  LSFHTDYPALHHPPGVQLLHCIKQTVTGGDSEIVDGFNVCQKLKKNNPQAFQILSSTFVDFTDIG-VDYCDFSVQSKHKI  178
hTMLH   LDRHTDTTYFQEPCGIQVFHCLKHEGTGGRTLLVDGFYAAEQVLQKAPEEFELLSKVPLKHEYIEDVGECHNHMIGIGPV  320
        *. *** . ::.* *:*::**:*:  *** : :**** ..::: :: *: *:::*.. :..  *  *. *.  : .   :

hBBOX1  IELDDK-GQVVRINFNNATRDTIFDVPVERVQPFYAALKEFVDLMNSKESKFTFKMNPGDVITFDNWRLLHGRRSYEAGT  238
hTMLH   LNIYPWNKELYLIRYNNYDRAVINTVPYDVVHRWYTAHRTLTIELRRPENEFWVKLKPGRVLFIDNWRVLHGRECFTG--  400
        :::     ::  *.:**  * .*  ** : *: :*:* : :.  :.  *.:* .*:** *: :****:****..: .

hBBOX1  EISRHLEGAYADWDVVMSRLRILRQRVENGN                                                  387
hTMLH   --YRQLCGCYLTRDDVLNTARLLGLQA----                                                  421
        *:* *.*   * *:.  *:*  :.
```

Table 1.1: Sequence alignment for *Homo sapiens* BBOX (gi|4502369) and TMLH (gi|8922625) generated using Clustal W [37]. Residues chelating either the structural zinc (Cys-38, Cys-40, Cys-43 and His-82 in BBOX) or the catalytic iron (His-202, Asp-204 and His-347 in BBOX) are highlighted in green; Residue forming the hydrophobic cage of the active site (Tyr-177, Trp-181 and Tyr-194 in BBOX) are highlighted in blue, together with the two residues interacting with the substrate terminal carboxylate (Asn-191 and Asn-292 in BBOX); the lysine residue (Arg-360) interacting with the terminal carboxylate of 2OG/NOG is highlighted in red. The predicted mitochondrial targeting sequence (MTS) of hTMLH is indicated with red asterisks.



Scheme 1.1: L-Carnitine biosynthesis pathway. L-Carnitine biosynthesis in mammals proceeds in four steps from free N^{ϵ}-trimethyllysine **11**. TML, N^{ϵ}-trimethyllysine **11**; TMLH, N^{ϵ}-trimethyllysine hydroxylase; HTML, 3-hydroxy-N^{ϵ}-trimethyllysine **12**; HTMLA, 3-hydroxy-N^{ϵ}-trimethyllysine aldolase; TMABA, N^{δ}-trimethylaminobutyraldehyde **13**; TMABADH, N^{δ}-trimethylaminobutyraldehyde dehydrogenase; BBOX, γ-butyrobetaine hydroxylase; P5P, pyridoxal-5′-phosphate.

N^{δ}-Trimethylaminobutyraldehyde Dehydrogenase

The third enzyme in the L-Car biosynthesis pathway has N^{δ}-trimethylaminobutyraldehyde dehydrogenase (TMABADH) activity but, as for HTMLA, no gene has been specifically associated with TMABADH activity in mammals. In humans, the conversion of TMABA **13** into γ-BB **7** (Scheme 1.1) is proposed to be catalysed by a protein encoded by the *ALDH9* gene [40]. Aldehyde dehydrogenase 9 (EC 1.2.1.19) is a nicotinamide adenine dinucleotide (NAD^+)-dependent oxidoreductase and was shown to have broad substrate specificity, including towards aminobutyraldehyde and its trimethylated form TMABA **13** [41]. This gene is of interest because it also acts as a γ-aminobutyraldehyde dehydrogenase, forming 4-aminobutanoic acid **15** (γ-aminobutyric acid or GABA), an important neurotransmitter. Similarly, three different aldehyde dehydrogenase enzymes could be purified from bovine liver tissues, one of them having a preference for TMABA **13** while the two others showed a broader substrate specificity [42].

γ-Butyrobetaine Hydroxylase

The fourth and last enzyme in the L-Car biosynthesis pathway is γ-butyrobetaine hydroxylase (BBOX, EC 1.14.11.1). BBOX, like TMLH, is a member of the Fe(II) and 2OG-dependent oxygenase superfamily of enzymes and will be described in details in Sections 1.1.5 and 1.1.6. BBOX catalyses the hydroxylation of γ-BB **7** at the C-3 position to give L-carnitine **1** in a stereoselective fashion. The high sequence similarity between BBOX and TMLH (identities = 101/354 (29%), positives = 163/354 (46%), gaps = 16/354 (5%), Table 1.1) suggests a common evolutionary ancestor. As described below, the activity and specificity of BBOX has so far only been studied using partially purified prokaryotic and animal enzyme preparations (Section 1.2). The present study describes work on the *in vitro* characterisation of pure recombinant human BBOX.

The absence of specific genes responsible for HTMLA and TMABADH activity in mammals suggests that the mammalian L-Car biosynthesis pathway might have only two specific enzymes, TMLH and BBOX, which evolved from a common ancestor, and that the two intermediate steps might be catalysed by promiscuous enzymes. Alternatively, the enzymes responsible for specific HTMLA and TMABADH activity have yet to be identified. In the mouse, the genes coding for the known L-Car biosynthesis enzymes, as well as the genes coding for OCTN2 and OCTN3, have been shown to be regulated by the peroxisome proliferator-activated receptor α (PPARα) [43]. PPARα is a ligand activated transcription factor which has an important role in lipid metabolism and energy homeostasis [44]. Looking for genes regulated by the same transcriptional mechanism might lead to the identification of potential HTMLA and TMABADH candidates.

1.1.3 L-Carnitine Metabolism

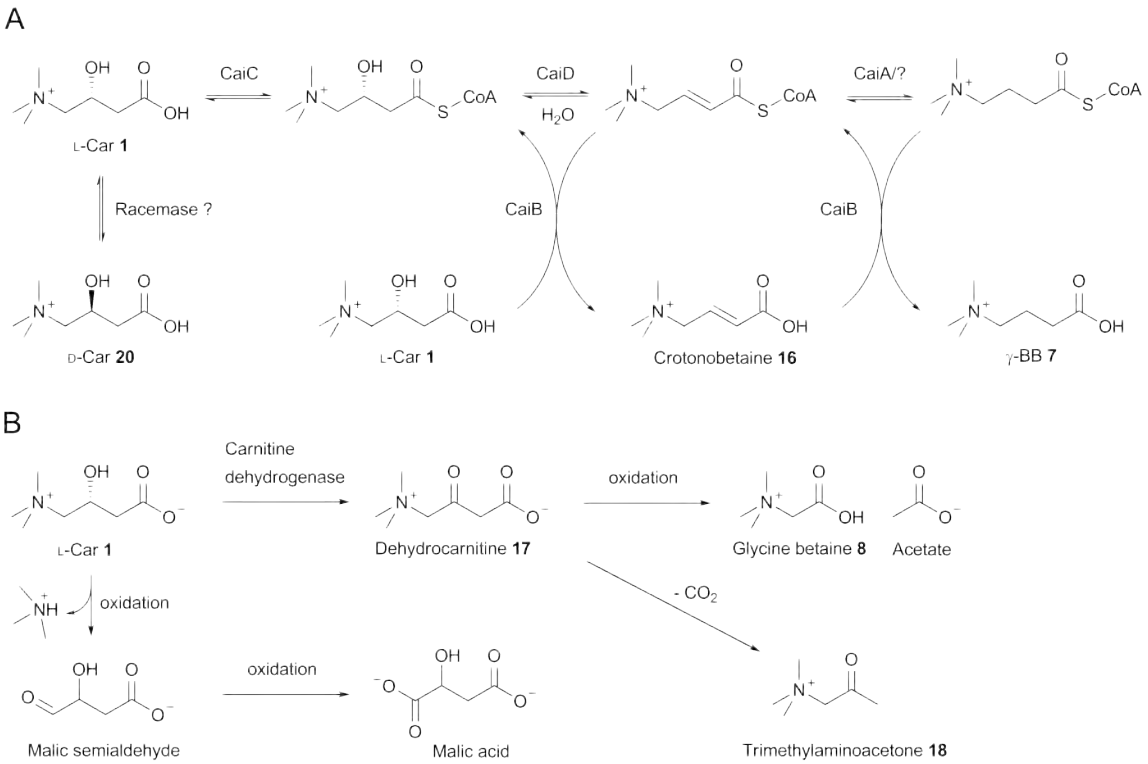
L-Car **1** is known to be metabolised by microorganisms but not by mammals [45]. Accordingly, the pharmacokinetic behaviour of L-Car in humans is dependent both on urinary excretion and on enterobacterial degradation in the gastrointestinal tract [46]. The prokaryotic degradation pathways of L-Car **1** vary depending on the oxygen availability and have been reviewed in detail by Rebouche and Seim [45] as well as by Lohninger, Pittner and Pittner [47].

Under anaerobic conditions, L-Car **1** is not assimilated as a source of carbon and nitrogen, but rather metabolised to γ-BB **7**, *via* crotonobetaine **16** ((2*E*)-4-(trimethylammonio)but-2-enoate) in a fully reversible manner (Scheme 1.2 A). The first enzyme that was identified as a member of the anaerobic L-Car **1** degradation pathway in *E. coli* is a crotonobetainyl-CoA/carnitine CoA-transferase that has been later crystallised in the presence of CoA [48]. This transferase was demonstrated to be encoded by the *caiB* gene [49] situated in an operon which contains six open reading frames (*caiTABCDE*) [5]. CaiT is a L-Car/γ-BB antiporter [50] of which the structure was recently solved [51, 52]. CaiA is proposed to be one of the two components of a flavin adenine dinucleotide (FAD)-dependent crotonobetaine reductase pair of enzymes [53]. CaiC is an ATP-dependent betaine/CoA ligase which catalyses the formation of the acyl-CoA substrates used by the other enzymes in the pathway [54]. CaiD is a carnitine-CoA dehydratase catalysing the reversible conversion of carnitinyl-CoA to crotonobetainyl-CoA [55], and CaiE has not been characterised and shows no homology to proteins of known function. CaiF, an additional gene found in the 3′ region adjacent to the cluster and transcribed in the opposite direction to the other *cai* genes was demonstrated to encode an activator of the *cai* operon-encoded carnitine metabolism [56].

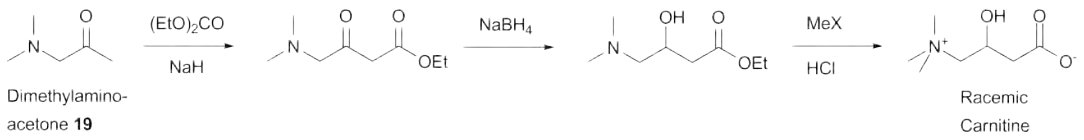
Under aerobic conditions, microorganisms can utilise L-Car **1** as the sole carbon, nitrogen, and energy source through at least two degradation pathways (Scheme 1.2 B). The C-3 oxidation of L-Car **1** produces dehydrocarnitine **17** (3-carboxy-*N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium), which can be metabolised into glycine betaine **8** and, presumably, acetate. The conversion of dehydrocarnitine **17** into glycine betaine **8** is in competition with the non-enzymatic decarboxylation of dehydrocarnitine **17** that yields *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18**. Alternatively, cleavage of the C-4–N bond leads to the formation of trimethylamine and malic acid. The further utilisation of both glycine betaine **8** and trimethylamine by bacteria has been well documented and leads to the formation of methane, carbon dioxide and ammonia [57].

1.1.4 L-Carnitine Supplementation

L-Car **1** supplementation has been advertised as an aid to weight loss, to enhance the sense of well-being, and improve exercise performance. It is commonly used by athletes despite the absence of consensus on the efficiency of oral uptake, even when using high doses [58, 59]. In addition, L-



Scheme 1.2: L-Carnitine degradation pathways in prokaryotes. (A) Under anaerobic conditions, a major portion of L-Car **1** is converted to γ-BB **7** *via* crotonobetaine **16**. This pathway is fully reversible and is exploited for L-Car **1** production (Section 1.1.4) [45]; (B) Under aerobic conditions L-Car can be oxidised into dehydrocarnitine **17** and glycine betaine **8**, or to malic acid and trimethylamine.



Scheme 1.3: Synthesis of racemic carnitine. The synthetic route proposed by Comber and Brouillette for the production of racemic Car from 1-dimethylamino-2-propanone **19**.

Car **1** supplementation is reported to be beneficial under many pathological conditions, including hypertension-related organ damage [60] and noninsulin-dependent diabetes mellitus [15]. Beneficial effects for L-Car **1** supplementation have also been proposed in combination with cancer treatments such as ifosfamide and cisplatin, which cause urinary loss of L-Car [61, 62], but, so far, the only direct link between BBOX and cancer has been a study integrating micro array data from published literature and showing that the gene for BBOX1 is differentially expressed in cancer tissues [63].

Because of the importance of L-Car **1** in energy metabolism and health, there is a commercial interest in its production. Several methodologies have been described and patented, but so far the commercially viable chemical syntheses, such as the route starting from 1-dimethylamino-2-propanone **19** proposed by Comber and Brouillette, have only yielded the carnitine racemate (Scheme 1.3) [64].

There are few reports of the impact of the unnatural enantiomer D-carnitine **20** ((3*S*)-3-hydroxy-4-(trimethylammonio)butanoate or D-Car) on metabolism and health, however it is believed to inhibit L-carnitine-mediated pathways, and it is banned from food supplements in Europe and in the United States. Commercial asymmetric chemical synthesis of L-Car **1** has been considered but found to be non-viable, and so was the recycling of the unnatural enantiomer from a synthetic racemic mixture through a bioprocess after fractionated crystallisation. Biological racemic resolution procedures yielding enantiopure L-Car **1** have been described, *e.g.* using an acetylcholinesterase that hydrolyses the D- but not the L-isomer of *O*-acetylcarnitine [65]. Other chemical and biological methods converting the D-Car **20** unnatural enantiomer into L-Car **1** exist but they are not economically feasible on a large scale either. The industrial preparation of L-Car **1** is therefore achieved by fermentation [66, 67, 68]. Because *de novo* biosynthesis has not been observed in prokaryotes and eukaryotic cells are difficult to grow on production scale, the fine chemical supplier Lonza developed an industrial process that starts with the achiral substrate γ-BB **7** and exploits the anaerobic pathway (Scheme

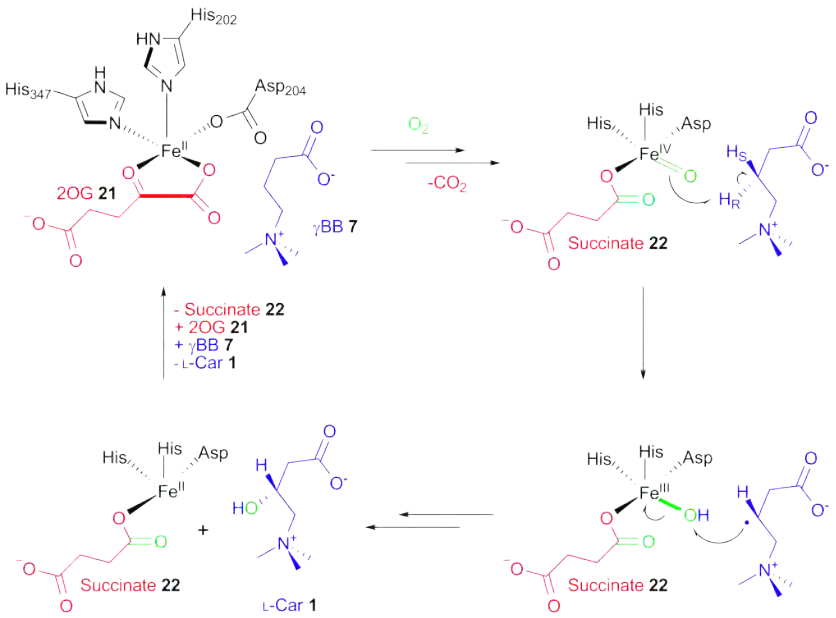
1.2 A). A classical strain development program yielded an *Achromobacter xylosoxydans* production strain with high productivity, a high precursor uptake rate, and a high product tolerance [66].

In contrast with supplementation, lowering L-Car **1** levels in tissues *via* the inhibition of its biosynthetic pathway has drawn limited attention. BBOX is the only enzyme in the L-Car biosynthetic pathway that has been considered as a potential target (Section 1.1.5). A decrease in L-Car availability is believed to stimulate glucose utilisation, while decreasing fatty acid oxidation, a metabolic switch that may be beneficial under certain physiological or pathological conditions.

1.1.5 BBOX Mechanism and Inhibition

γ-Butyrobetaine hydroxylase (BBOX) is the only enzyme in the L-Car biosynthesis pathway that has been extensively characterised. Some medicinal chemistry studies have led to the identification of potential BBOX inhibitors (Section 1.2.1). Members of the Fe(II) and 2-oxoglutarate (2-oxopentanedioic acid or 2OG)-dependent oxygenase superfamily of enzymes have long been therapeutic targets and their mechanism, including that of BBOX, is well studied (Scheme 1.4) [29]. The catalytically important metal ion is coordinated in a tridentate manner by the side chain of active site amino acid residues, two histidines and a carboxylate, either aspartate or glutamate [69]. In the case of BBOX, these residues are His-202, His-347 and Asp-204. In a resting state, the three remaining coordination are proposed to be occupied by water molecules. In the presence of 2OG **21** (shown in red in Scheme 1.4), two of the remaining coordination sites are occupied by the α-keto carboxylate moiety of the 2OG cosubstrate and the last site is occupied by a water molecule (not shown). Molecular oxygen (shown in green in Scheme 1.4) then enters the active site and reacts to form a reactive Fe(IV)-oxo intermediate through oxidative decarboxylation of 2OG **21** into succinate **22** [70]. In the presence of the substrate, the Fe(IV)-oxo species abstracts an hydrogen through a homolytic C-H bond cleavage, yielding an Fe(III)-hydroxyl and a free carbon radical on the substrate. This radical can react with the hydroxyl group to form the hydroxylated product and a reduced Fe(II) metal center. In the absence of substrate, uncoupled turnover of 2OG **21** can occur, leading to an Fe(IV)-oxo species which will most likely slowly inactivate the enzyme through self-oxidation.

Based on this understanding of the BBOX mechanism, γ-BB analogues containing a 'radical



Scheme 1.4: Proposed BBOX catalytic cycle. Many Fe(II) and 2OG-dependent enzymes have been studied in detail and a common mechanism has been proposed. 2OG **21** (shown in red) and Fe(II) are oxidised by molecular oxygen (in green) to give succinate **22** and an Fe(IV)-oxo species which abstract an hydrogen atom from the substrate γ-BB (shown in blue) in a homolytic fashion. The hydroxyl group now sitting on the Fe(III) centre can react with the free radical formed on the substrate to form the hydroxylated product.

clock' cyclopropane group were designed and tested as competitive substrates and inhibitors (*e.g.* compounds **23**, **24**, **25** and **26**, Figure 1.7) [71, 72]. These compounds were synthesised primarily for mechanistic studies but also based on the idea that they could trap the radical intermediate and inhibit the enzyme by forming a covalent adduct. This covalent approach to enzyme inhibition has been used for other radical generating enzymes [73], including Fe(II) and 2OG-dependent enzymes such as clavaminate synthase [74]. Ziering and Pascal proposed that 3-(2,2-dimethylcyclopropyl)propanoic acid **26** does indeed inhibit BBOX through this mechanism [72]. On the other hand, the three compounds synthesised (compounds **23**, **24** and **25**, Figure 1.7) and tested by Petter *et al.* showed very poor inhibition [71]. These authors explained a lack of inhibitory properties due to the bulky cyclopropane group preventing the compounds from having sufficient binding affinities.

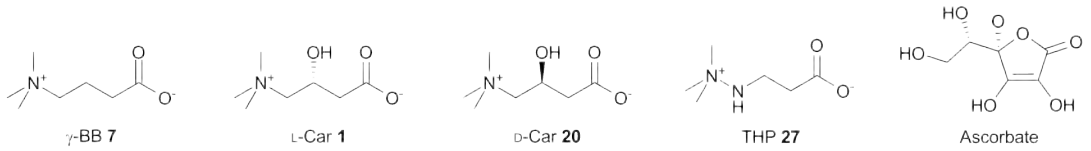


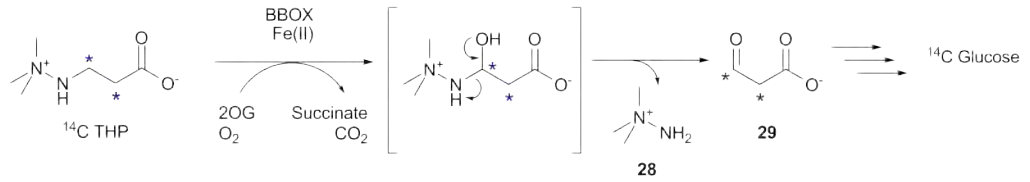
Figure 1.3: Structures of γ-butyrobetaine, L-carnitine, D-carnitine, THP and ascorbate. The structural similarities between the BBOX substrate γ-BB **7**, product L-Car **1**, unnatural enantiomer D-Car **20** and clinically used inhibitor THP **27** suggest a competitive mechanism of inhibition. The structure of the reducing agent ascorbate is given for reference.

THP: A Clinically Used BBOX Inhibitor

3-(2,2,2-Trimethyldiazan-2-ium-1-yl)propanoate

27 commonly known as 3-(2,2,2-trimethylhydrazinium)propionate (THP, Figure 1.3), also named MET-88, meldonium or quaterine, and marketed under the brand name Mildronate, was identified as a human BBOX inhibitor in the late 1980s [75]. Several studies have been conducted *in vivo* using THP **27** as a canonical BBOX inhibitor and some reports confirmed the non-competitive mode of action with respect to the natural substrate γ-BB **7** [76] while others suggested that THP could be acting as a BBOX competitive substrate [77, 78]. Beside the action of THP **27** as BBOX inhibitor, it has been proposed to disrupt renal carnitine reabsorption by OCTN1-3 transporters *via* a competitive mechanism [78]. Beneficial effects have been found to be associated with THP administration, through modulation of energy homeostasis, but also through its action on targets other than BBOX, such as carnitine acetyltransferase [79]. THP **27** administration to rats has been shown to have no effect on *TMLH* and *BBOX1* at a transcriptional level [80], however effects on glucose and fatty acid metabolism-related gene transcription and expression levels have been described [81, 82, 83].

Despite THP **27** being clinically used in some countries, its exact mode of action remains unclear. Some studies have investigated the action of the drug on metabolic processes [84], but the molecular mechanism is not fully understood. The pharmacokinetics and metabolic fate of THP **27** have been studied in perfused rat hearts using [^{14}C]-labeled THP and major radioactive metabolites were identified as carbon dioxide, glucose, succinic acid and 3-hydroxypropionic acid (Scheme 1.5) [77]. This observation led to the hypothesis that THP **27** is indeed hydroxylated by BBOX at the C-3 position, giving an unstable hemihydrazinal species that spontaneously fragments into trimethylhydrazine **28** and labeled malonate semialdehyde **29** (Scheme 1.5). Radioactive malonate semialdehyde **29** can then be metabolised to propionic acid or acetyl-CoA, participating



Scheme 1.5: Previously proposed metabolic fate of THP. The mechanism proposed by Yohisue et al. is based on observations made for the metabolic fate of [^{14}C]-labeled THP in perfused rat hearts [77]. The absence of label on the trimethylhydrazinium moiety did not allow the authors to describe any fate for this part of the drug and the formation of trimethylhydrazine **28** was speculated. [^{14}C]-labels are indicated by asterisks.

in glyconeogenesis and then be excreted as [^{14}C]-labeled CO_2 through the glycolytic pathway and tricarboxylic acid (TCA) cycle. Although these results point towards THP **27** being a BBOX substrate, other more recent reports state that THP **27** is a substrate analogue which cannot be hydroxylated [85].

The results described in Section 1.3 suggest an alternate mechanism which explains the absence of trimethylhydrazine **28** formation upon BBOX-catalysed THP oxidation.

1.1.6 Human BBOX Crystal Structure and Mechanistic Implications

Until recently, none of the enzymes involved in the biosynthesis of L-Car **1** had been structurally characterised. In 2010, an apo crystal structure of human BBOX was reported at 2.0 Å resolution (PDB ID: 3N6W) [85]. The authors of this report obtained pure recombinant protein from an *E. coli* expression system using a codon optimised gene, albeit in low yield. This result was achieved using a plasmid carrying the GroEL and GroES bacterial chaperones, in addition to the *BBOX1* gene cloned in a second plasmid, suggesting that proper folding of BBOX is crucial for expression and activity.

After several unsuccessful attempts to obtain pure and active recombinant human BBOX using *E. coli* expression systems in our laboratory, collaborators at the Structural Genomic Consortium site in Oxford, UK, found that satisfactory expression of recombinant BBOX could be achieved using a baculovirus expression vector system (BEVS) and insect cell culture. Cultures of insect cells infected with BEVS allow the expression and purification of recombinant human proteins after complex post-translational modifications (*e.g.* proteolysis, *N*- and *O*-glycosylation, acylation, amidation, carboxymethylation, phosphorylation, and prenylation) that absent in bacteria [86, 87]. Insect cells also have robust machinery for the correct folding of mammalian proteins. Because

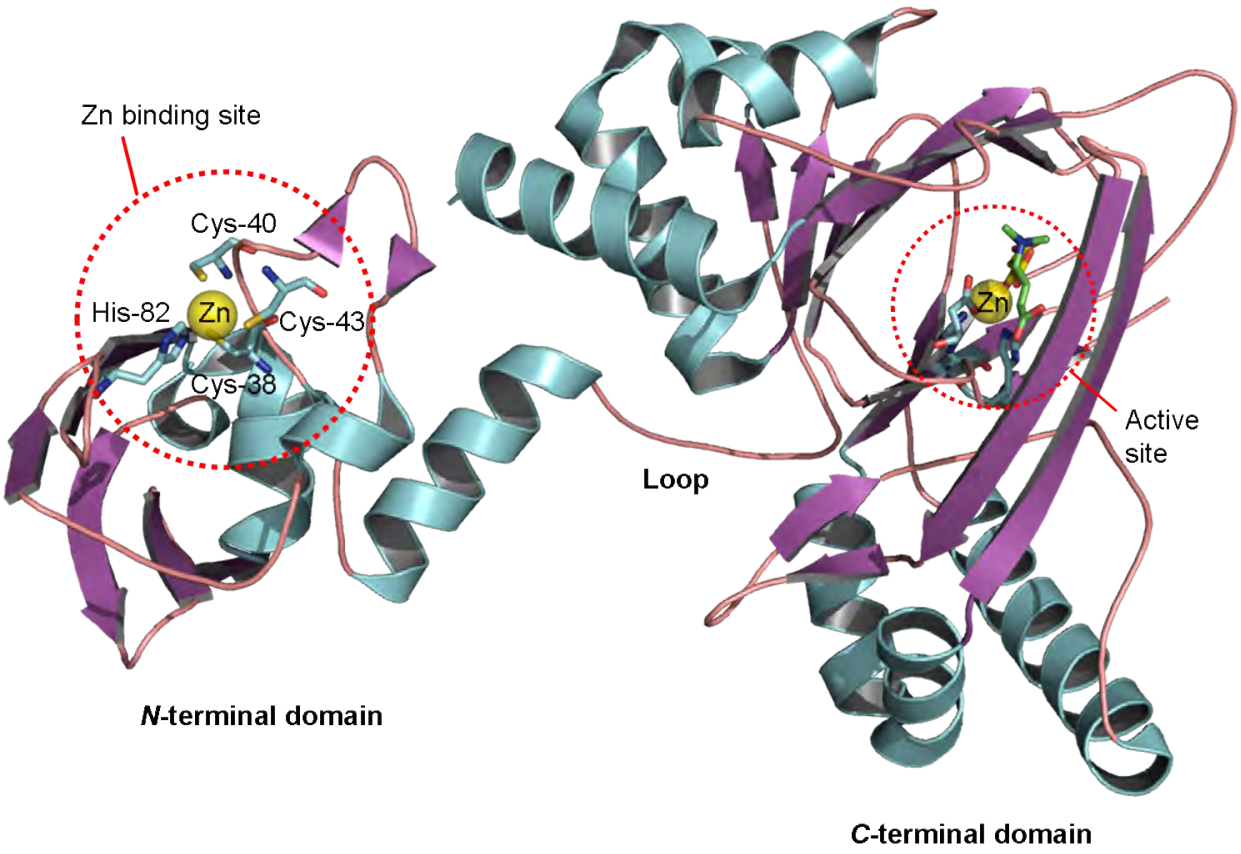


Figure 1.4: View from a crystal structure of human γ-butyrobetaine hydroxylase. The *N*-terminal zinc binding domain contains the first 110 amino acids. This domain is connected to the catalytic *C*-terminal domain by a loop composed of residues 111–123. The *C*-terminal domain shows a DSBH fold common to Fe(II) and 2OG-dependent oxygenases. This figure was generated using PyMOL [91].

BBOX has never been described to carry any post-translational modification, proper folding of the polypeptide is probably the key factor for explaining the success of the BEVS approach. Highly pure BBOX was obtained using this methodology.

The same collaborators crystallised BBOX in the presence of Zn(II) and *N*-oxalylglycine **30** (NOG) as substitutes for Fe(II) and 2OG **21**. Fe(II) and 2OG-dependent oxygenases are known to be inhibited by divalent metals, such as Zn(II), Mn(II) or Cu(II) [88], and metal chelating species such as NOG **30** [89]. The replacement of Fe(II) and 2OG **21** by non-reactive analogues allows the formation of stable crystallographic substrate complexes. The crystal structures for the enzyme in complex with natural substrate γ-BB **7** and the clinically used inhibitor THP **27** (Figure 1.3) were obtained at 1.78 Å and 1.82 Å resolution, respectively (PDB ID: 3O2G and 3MS5) [90].

The Overall Structure of BBOX

The overall BBOX structure shows a two-domain architecture (Figure 1.4). The first 110 *N*-terminal residues form a zinc binding domain involved in crystallographic dimerisation. This interaction is probably relevant in solution as BBOX was previously shown to be an homodimer by polyacrylamide gel electrophoresis [76, 92]. The residues involved in the binding of the zinc cation are Cys-38, Cys-40, Cys-43 and His-82. This domain is only partly conserved in TMLH where two of the three cysteines, Cys-80 and Cys-85 (the second one being replaced by Ser-82), and a histidine, His-124, are present (Table 1.1). The *C*-terminal domain of BBOX contains the catalytic activity and shows a typical double-stranded beta-helix (DSBH) fold, common to Fe(II) and 2OG-dependent oxygenases [93]. The DSBH structural fold contains the conserved HXD...H motif (His-202, Asp-204, and His-347 in BBOX) which coordinates the catalytic metal (Figure 1.5). The two domains are linked by a loop composed of residues 111–123.

The Active Site of BBOX

The active site of BBOX is buried in the core of the *C*-terminal domain with only a narrow opening to the solvent. The size of this tunnel might enable diffusion of molecular oxygen and release of carbon dioxide, but the differences between the substrate bound and the apo structures suggest that a conformational change occurs in order to enable the substrates to access the metal centre and the products to be released after hydroxylation. In the case of the apo BBOX structure, the loops surrounding the active site are not clearly defined, indicating their conformational flexibility in the absence of substrates. These regions are visible in the substrate and inhibitor structures, due to lower flexibility, and partly cover the active site, seemingly blocking the access of the substrate to the metal centre.

Several specific interactions can explain the binding of substrates and the stereochemical course of the reaction. The terminal carboxylate of NOG forms hydrogen bond interactions with the side chain of Arg-360 while the two other carbonyl oxygen atoms coordinate the metal. The structures solved in the presence of the substrate γ-BB **7** or the inhibitor THP **27** indicate that thse two compounds adopt

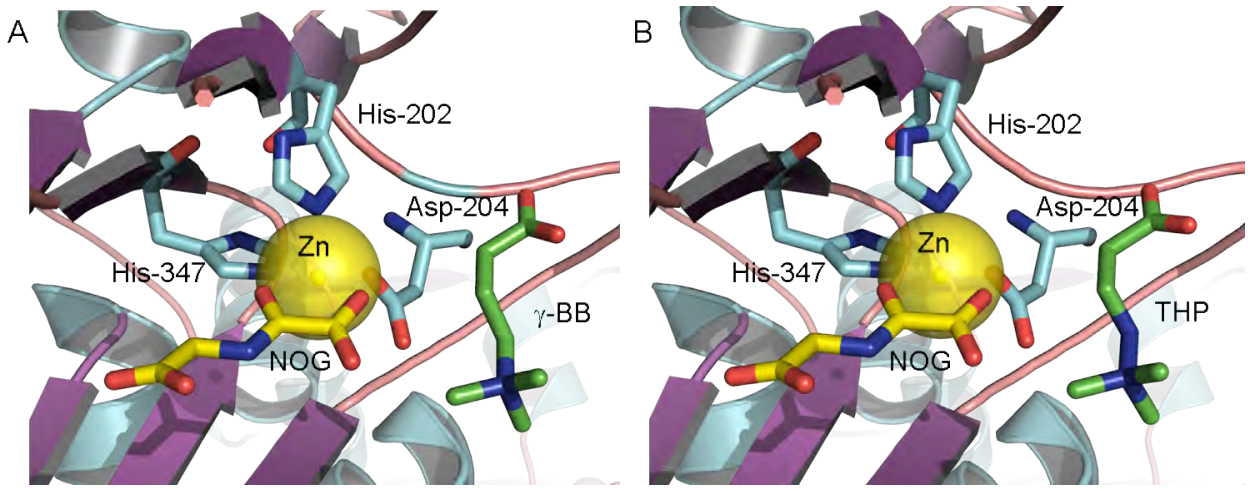


Figure 1.5: Views from the active site of BBOX containing γBB or THP. The active site geometry is near identical in the structure with (A) the natural substrate γBB **7** and (B) the inhibitor THP **27**, both displayed as green sticks. The crystallographic Zn(II) replacing the catalytic Fe(II) is shown as a yellow sphere. It is coordinated by His-202, Asp-204 and His-347, displayed as blue sticks. NOG **30**, an unreactive 2OG analogue, is displayed as yellow sticks. This figure was generated using PyMOL [91].

a near identical conformation in the active site (Figure 1.5). This is not surprising considering the structural similarities between γ-BB **7** and THP **27**, and it suggests that indeed BBOX inhibition is due to THP being a competitive binder, and possibly a competitive substrate. The terminal carboxylates of γ-BB **7** and THP **27** are positioned to form hydrogen bond interactions with two asparagine residues (Asn-191 and Asn-292) while the trimethylammonium or trimethylhydrazinium function sits in an electron rich aromatic cage formed by by the side chains of three aromatic residues: Tyr-177, Trp-181 and Tyr-194. This binding mode is similar to the way the *N*-trimethylammonium group of posttranslationally modified lysines in histone tails bind to the plant homeo domain (PHD) and tudor domain (TD)-containing proteins [94, 95]. The binding mode of the substrate can also explain the retention of configuration after hydrogen abstraction [46]. Based on the C-3 atom position, the *pro-R* hydrogen must be pointing towards the catalytic metal, while the *pro-S* hydrogen points away.

Altogether, these crystallographic results allow the rationalisation of past observations, such as dimerisation, as well as results presented below on the basis of structural features.

1.1.7 Objectives

The objectives of the work described herein were to characterise the purified recombinant human BBOX *in vitro*, and in particular its cofactor dependence, substrate specificity, and the mechanism of inhibition by the clinically used drug THP **27**. THP **27** was of particular interest because it represents a rare example of an Fe(II) and 2OG-dependent oxygenase inhibitor which is not an analogue of the 2OG cofactor, but rather a substrate analogue. The second part of this chapter describes the observation that THP **27** not only inhibits BBOX, but is also a substrate undergoing oxidative reaction to give unanticipated products (Section 1.3). The identity of a product which was formed through an unprecedented rearrangement reaction was confirmed by synthesis and analytical techniques.

The results presented in this chapter were obtained in collaboration with Dr Grazyna Kochan (human BBOX expression and purification), Dr Tobias J. Krojer (human BBOX crystallisation and structure elucidation), Ivanhoe K. H. Leung (NMR assays and kinetic analysis, as well as NMR characterisation of the THP reaction products), Nikita Loik (advice for the mass spectrometry analyses and HPLC separation of the 3-amino-4-(methylamino)butanoic acid enantiomers), Anna Rydzik (helpful discussion and follow up of the project), Dr Michael McDonough (analysis of the crystallographic data), Dr Timothy D. W. Claridge, Dr Frank von Delft, Prof Udo Oppermann and Prof Christopher J. Schofield (data analysis and supervision).

Part of the results presented in this chapter are published and described in references [90] and [96].

1.2 BBOX Substrate Analogues and Inhibition Studies

1.2.1 Introduction

From the 1970s, several mechanistic studies were conducted on BBOX, either using crude cell extracts, or on partially purified fractions of mammalian or bacterial protein. These studies resulted in the determination of kinetic parameters and isotope effects [97], stereoselectivity [98], substrate specificity [99, 100, 101], cofactor requirement [102, 103] and mechanism of inhibition by small molecules, including substrate [71] and 2OG analogues [104]. In most cases, [^{14}C] or [^{3}H]-radiolabeled substrates were used to quantify BBOX activity. Examples include the quantification of [^{14}C]-CO_2 liberated from the coupled decarboxylation of [^{14}C]-2OG, or the [^{3}H] released into the aqueous medium from the [2,3-^{3}H]-γ-BB substrate **31** (Figure 1.7 B) [101, 71]. Other quantification methodologies were based on chromatographic separation and scintillation measurement for [^{3}H] or [^{14}C]-radiolabeled starting material and products, such as [methyl-^{3}H]-labeled γ-BB, [methyl-^{14}C]-labeled γ-BB, [1,2,3,4-^{14}C]-γ-BB and [methyl-^{14}C]-labeled N^{ϵ}-trimethyllysine [102, 31, 24].

The 2OG analogues reported to exhibit good BBOX inhibition are generic Fe(II) and 2OG-dependent enzyme inhibitors, such as NOG **30**, pyridine 2,4-dicarboxylate **32** (2,4-PDCA) and 3,4-dihydroxybenzoate **33** (Figure 1.6) [104]. These compounds compete with the 2OG **21** cosubstrate for iron chelation and therefore may be considered to be unlikely to be selective towards BBOX. From a pharmacological perspective, because the Fe(II) and 2OG-dependent oxygenase family of enzyme is large and its members carry diverse biological functions, unspecific inhibition is likely to have undesirable side effects. Accordingly, there is a strong interest in the development of inhibitors specific for some pharmacologically relevant members, including BBOX [105].

It was important to replicate to some extent, and expand the previously published results on BBOX using highly purified recombinant human protein for two reasons. First, all the previous data were obtained either from tissue homogenates, cell extracts or only partially purified enzyme preparations, and none of them from human sources. The presence of many unknown components in the described assays likely added complexity to the interpretation of the results. Also, the kinetic parameters and specificity of the human enzyme might differ from bacterial or other mammalian enzymes. Finally, most assays were conducted using radiolabeled substrates and cosubstrate. Higher accuracy and

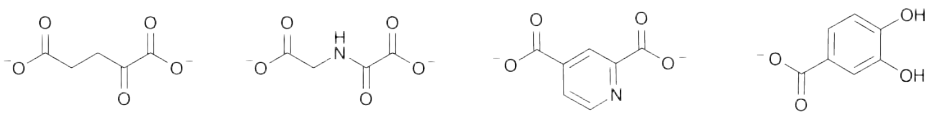


Figure 1.6: Structure of 2-oxoglutarate and related general oxygenase inhibitors. *N*-Oxalyl glycine **30** (NOG), 2,4-pyrridine dicarboxylate **32** and 3,4-dihydroxybenzoate **33** chelate the catalytically important metal in a competitive manner with regard to 2OG **21** and are generic inhibitors of Fe(II) and 2OG-dependent oxygenases.

sensitivity can now be obtained using pure recombinant enzyme and modern spectroscopic and spectrometric techniques such as high field nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). These techniques do not require the use of radioactive substances. A set of compounds previously tested as BBOX substrates or inhibitors was compiled from the literature (Figure 1.7) and a selection was tested with recombinant BBOX *in vitro* to compare our results with the observations made using other techniques [101, 71].

The identification of the molecular mechanism for the inhibition of human BBOX by the clinically used compound THP **27** was of particular interest. The structural similarities between THP **27** and BBOX natural substrate γ-BB **7** suggested a competitive inhibition mechanism, possibly through oxidation of THP **27** to yield unknown metabolites. The identification of these products is described in Section 1.3.

1.2.2 BBOX Activity Assay Development

Various assays have been described for the characterisation of Fe(II) and 2OG-dependent oxygenases (reviewed recently in Rose et al. [105]). As described above, the characterisation of BBOX activity has so far been mostly based on the use of radiolabeled substrates. A standard assay for this family of enzyme that has been used for BBOX characterisation allows the quantification of enzyme activity by measuring the [^{14}C]-labeled carbon dioxide released from labeled 2OG [78]. Additionally, direct detection of [^{14}C] and [^{3}H]-labeled substrates and products [106, 107] as well as a coupled assay in the presence of carnitine acetyltransferase and [^{14}C]-labeled acetate have been extensively used [108]. Because radiolabeled chemicals are dangerous to handle, an alternative method was investigated in order to assay BBOX activity.

Both the natural substrate and product of BBOX lack a chromophore that can be detected using

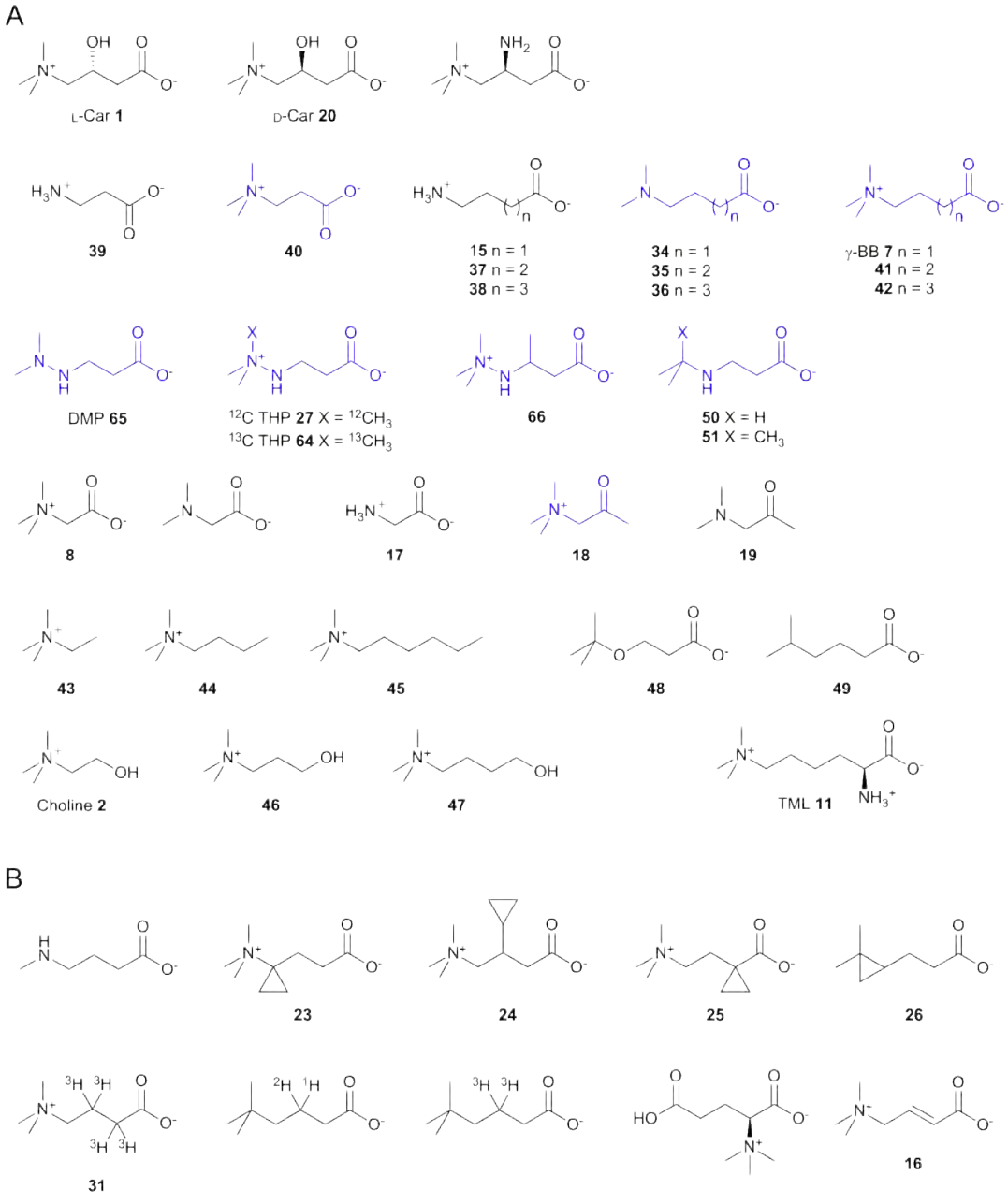


Figure 1.7: Structures of BBOX substrates and inhibitors. (A) Structures of the compounds that were tested in this study. The compounds for which the synthesis is described in this work are highlighted in blue; (B) Structures of the compounds that were tested in previous studies and that were not tested in this study.

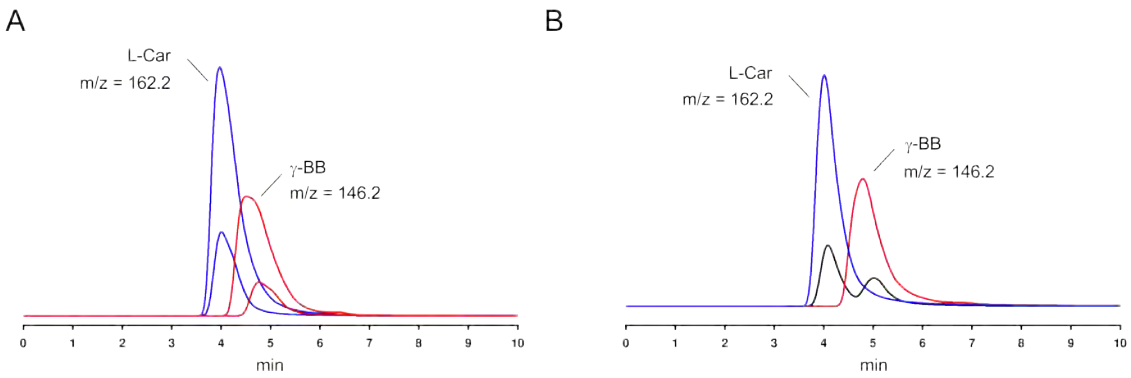


Figure 1.8: Investigations into an LC-MS based assay for measuring BBOX activity. Two different elution profiles for pure standard of the starting material γ-BB **7** (m/z = 146.2, shown in red) and the product L-Car **1** (m/z = 162.2, shown in blue) on a Synergi Polar-RP column (250 mm x 4.6 mm, 4 μ particle size). The traces are extracted ion chromatograms. The flow rate was set at 1 mL per min and the mobile phases were A: 0.1% formic acid in water and B: 0.1% formic acid in acetonitrile. (A) 0.5 nmol (top traces) and 0.1 nmol (bottom trace) of compounds eluted using 2% B in A for 5 min followed by 2–98% B in A over 30 min, L-Car **1** (R_t = 4.1 min) γ-BB **7** (R_t = 4.6 min); (B) 0.5 nmol of compounds eluted using 2% B in A for 5 min followed by 2–20% B in A over 30 min, L-Car **1** (R_t = 4.1 min) γ-BB **7** (R_t = 5.1 min). A mixture of 0.1 nmol of each γ-BB **7** and L-Car **1** can be separated using these conditions (black trace).

standard UV absorbance coupled to high performance liquid chromatography (HPLC) separation. An initial attempt to solve this issue involved a LC-MS assay exploiting the difference in mass and elution times between γ-BB **7** and L-Car **1**. Because of the hydrophilicity and highly similar structural features of the two compounds, the separation on available reverse phase HPLC columns was found to be limited (Figure 1.8). L-Car **1** (m/z = 162.2, blue trace, Figure 1.8 A) eluted within the void volume of the column (R_t = 4.1 min), while γ-BB **7** (m/z = 146.2, red trace, Figure 1.8 A) was only retained for a short time (R_t = 4.6 min). An attempt to optimise the chromatography condition by decreasing the slope of the gradient only moderately improved the separation as L-Car **1** (m/z = 162.2, blue trace, Figure 1.8 B) still eluted within the void volume (R_t = 4.1 min), and γ-BB **7** (m/z = 146.2, red trace, Figure 1.8 B) was only retained for a bit longer (R_t = 5.1 min). The sensitivity of mass spectrometry allowed the obtention of a baseline separation when injecting 0.1 nmol of each compound (black trace, Figure 1.8 B). However, even with these parameters in hand, it would probably make the detection and quantification of small differences in turnover or inhibition relatively difficult.

As an alternative, a ^{1}H NMR spectroscopy-based assay was developed by Ivanhoe K. H. Leung to measure the activity of recombinant human BBOX in real time. The course of the hydroxylation

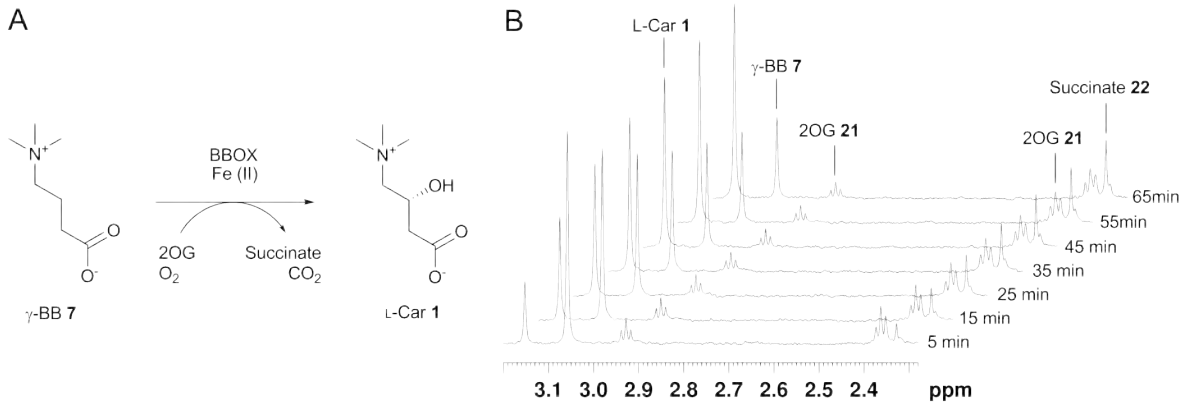


Figure 1.9: Time course ^{1}H NMR monitoring of BBOX activity. (A) The reaction catalysed by BBOX; (B) Monitoring of BBOX-catalysed conversion of γ-BB **7** into L-Car **1** was followed using ^{1}H NMR spectroscopy at 700 MHz and by measuring the conversion of the methyl signals for the starting material (δ = 3.06 ppm) and the product (δ = 3.16 ppm). Spectra were recorded at 10 min intervals with the first spectrum being recorded 5 min after sample mixing.

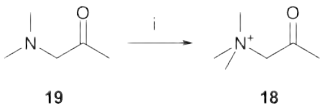
reaction could be followed based on changes in intensity of the signals corresponding to the *N*-methyl proton resonances of the trimethylammonium moiety of the substrate **7** (δ = 3.06 ppm) and the product **1** (δ = 3.16 ppm) (Figure 1.9). The identity of the signals was initially confirmed using standard samples of L-Car **1** and γ-BB **7**. The inhibition assays were performed following that same protocol to observe the effect of various compounds on the conversion of γ-BB **7** into L-Car **1**. Studies on the identification of alternate substrates was performed in the absence of the natural substrate γ-BB **7** by looking for the appearance of new species.

An initial set of experiments were performed to confirm the requirements for BBOX enzymatic activity. BBOX was found to be highly active (k_{cat} = 5.13 s^{-1} with respect to 2OG **21**, k_{cat} = 5.3 s^{-1} with respect to γ-BB **7**) [90]. The activity of BBOX was found to be dependent on the presence of Fe(II). The hydroxylation of the natural substrate γ-BB **7** was closely coupled to oxidation of the 2OG **21** as previously reported [109]. BBOX activity was enhanced in the presence of reducing agents, such as ascorbate (Figure 1.3), but the *in vivo* relevance of this observation is unclear [110, 103, 102, 111, 112]. The addition of potassium ions had stimulatory effect [113, 100] and substrate inhibition could be observed at high concentration of γ-BB **7** as reported [114, 100].

1.2.3 D-Carnitine is a BBOX Substrate

There was an interest in testing the unnatural enantiomer D-Car **20** as a potential substrate for BBOX. D-Car **20** is present in food supplement preparations in amounts varying from 0.5 to 6% of total carnitine content [115] and could potentially have harmful effects on health. A limited number of studies have investigated the physiological relevance of D-Car and even though it is described as physiologically inactive, some side effects are observed when carnitine is used as a racemate [66]. For example, when D-Car **20** was injected directly into the rat blood stream, a depletion of the natural L-isomer in heart and skeletal muscle tissues was observed [116]. The unnatural enantiomer was described to be responsible for this depletion and a competitive transport mechanism was suggested. This hypothesis was supported by the observation that D-Car **20** inhibits the transport of L-Car **1** by organic anion transporters from the blood stream into heart and skeletal muscle tissues [117].

To gain molecular insights into the action of BBOX on D-Car **20**, Nordin and coworkers tested the radiolabeled compound as a substrate for purified human and pseudomonas BBOX [118]. They found that D-Car promotes enzymatic 2OG **21** decarboxylation but no product could be identified. 2OG turnover was therefore thought to be due to uncoupled oxidation [109]. Only [^{14}C]-D-Car itself was recovered from the assay. At the time, the authors wrote that "the reaction catalysed by γ-butyrobetaine hydroxylase is stereospecific ... which means that further hydroxylation of the product, L-carnitine, cannot be expected. Hydroxylation of the product enantiomer was considered as a possibility in analogy with the hydroxylation of D-octopamin by dopamine β-hydroxylase (EC 1.14.17.1) [119] ... The absence of any product from D-carnitine and the fact that also L-carnitine catalyzed the partial reaction rule out hydroxylation of D-carnitine as a cause of the observed decarboxylation of 2-oxoglutarate" [118]. When D-Car **20** was incubated with BBOX in a ^{1}H NMR spectroscopy-based assay as described above (Section 1.2.2), two additional signals corresponding to the methyl groups of trimethylammonium species could be observed, indicating that D-Car is a BBOX substrate and that at least two different products are formed. Beside the singlet corresponding to D-Car itself (δ = 3.13 ppm), another singlet (δ = 3.21 ppm) rapidly appeared and then slowly decreased with time, while a third singlet (δ = 3.19 ppm) appeared in a manner proportional to the disappearance of the second signal (Figure 1.10 B) [90]. The same two species were observed when L-Car **1** was incubated in the presence of BBOX but only to a very low extent, suggesting that L-Car



Scheme 1.6: The synthesis of *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** triflate salt from 1-dimethylamino-2-propanone **19**. (i) 1.1 eq. Methyl triflate, CH_2Cl_2, rt, 30 min, 85% [120].

can further react in the presence of BBOX. This observation could also be due to low amounts of D-Car **20** present in the L-Car **1** sample from commercial sources.

As mentioned by Nordin and coworkers [118], there is precedent for the hydroxylation of a substrate at a position already carrying a hydroxyl group and subsequent elimination of water to form the corresponding ketone; in the context of oxygen sensing, the human prolyl hydroxylase 2 (PHD2) hydroxylates the prolyl residue of its substrate HIF-1α (Scheme 1.7 A) [121]. The recombinant PHD2 catalytic domain ($PHD2_{181-426}$) was shown to introduce a 4-oxoprolyl functionality in a truncated substrate when the prolyl residue of the HIF-$1\alpha_{556-574}$ is replaced by a *cis*-4-prolyl residue (Scheme 1.7 B) [122]. In a similar manner, although it is not an Fe(II) and 2OG-dependent enzyme, dopamine β-monooxygenase, which normally hydroxylates dopamine to form noradrenaline (Scheme 1.7 C), can turn over the substrate analogue *S*-octopamine to form (4-hydroxyphenyl)(oxo)methanaminium (Scheme 1.7 D) [119]. Having these examples in mind, we hypothesised that the transient species formed upon incubation of D-Car **20** in the presence of BBOX was dehydrocarnitine **17** and would spontaneously and non-enzymatically decarboxylate to form *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** (Scheme 1.7 E).

The BBOX-catalysed reaction with D-Car **20** was followed by mass spectrometry and both the formation of the labile intermediate **17** and *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** could be observed in real time in the crude assay (Figure 1.10 A), further confirming the ^{1}H NMR-based observation. *N*,*N*,*N*-Trimethyl-2-oxopropan-1-aminium **18** was unequivocally identified as the final product of the reaction by comparing the ^{1}H NMR chemical shifts and heteronuclear correlations of the product in the crude assay with a synthetic standard. The triflate salt of *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** was obtained in one step as previously described by Biscoe and Breslow *via* methylation of commercially available 1-dimethylamino-2-propanone **19** using methyl

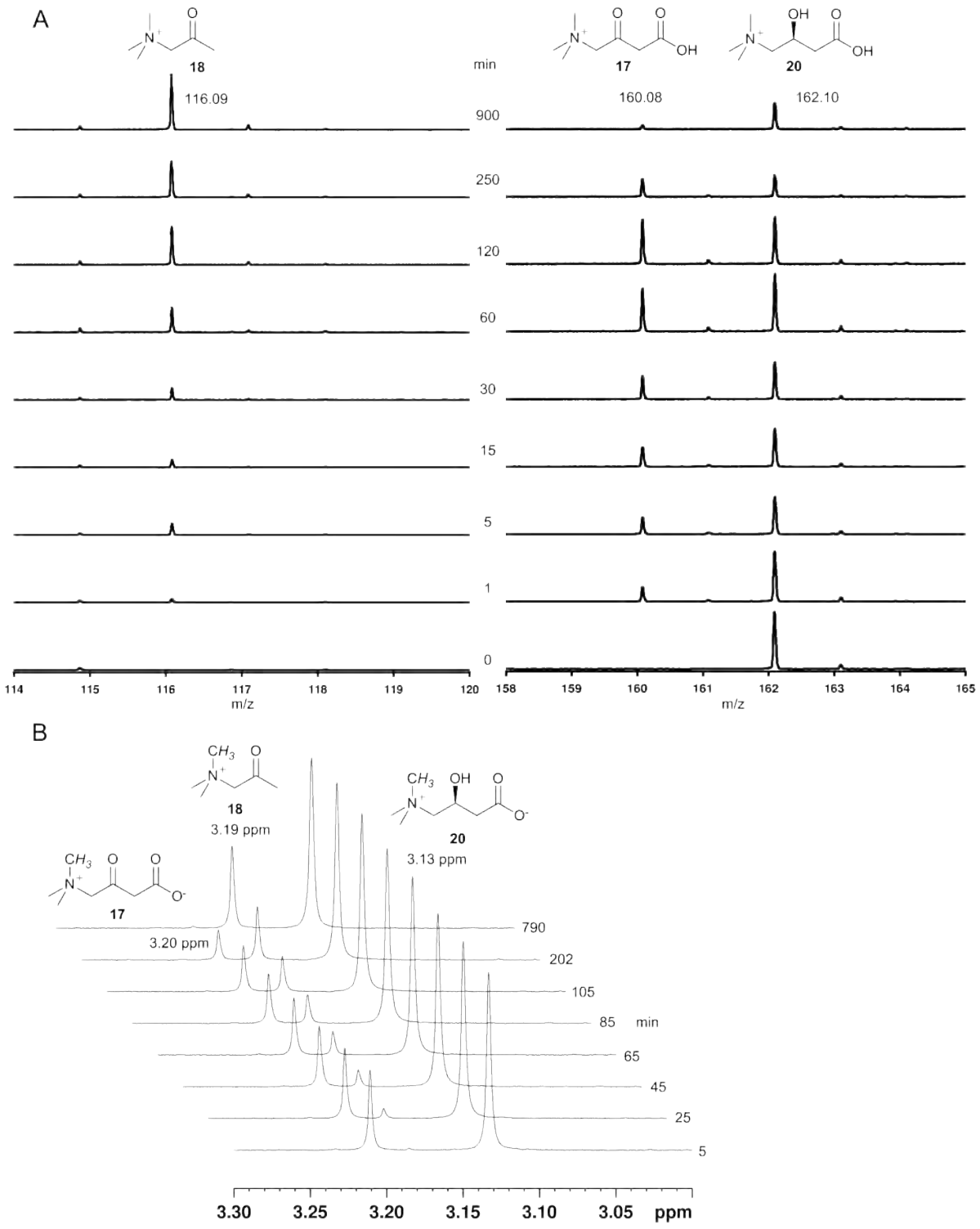


Figure 1.10: The unnatural D-carnitine enantiomer is a substrate for BBOX *in vitro*. MS and ^{1}H NMR monitoring of BBOX-catalysed conversion of D-Car **20** into *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18**. (A) The Reaction of D-Car **20** with BBOX was followed by mass spectrometry. The D-Car **20** signal (m/z = 162.10) was found to decrease while a transient species dehydrocarnitine **17** (m/z = 160.08) is converted non-enzymatically to *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** (m/z = 116.09); (B) ^{1}H NMR time course experiment showing the conversion of D-Car **20** into *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18**. L-Car **1** is also converted to *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** *via* dehydrocarnitine **17** albeit much slower (data not shown). The assay mixture contained 1 μM BBOX, 25 μM Fe(II), 1 mM D-Car, 1.5 mM 2OG, 5 mM sodium ascorbate, 80 mM KCl, and 50 mM Tris-D11 pH 7.5 in 95% H_2O and 5% D_2O.

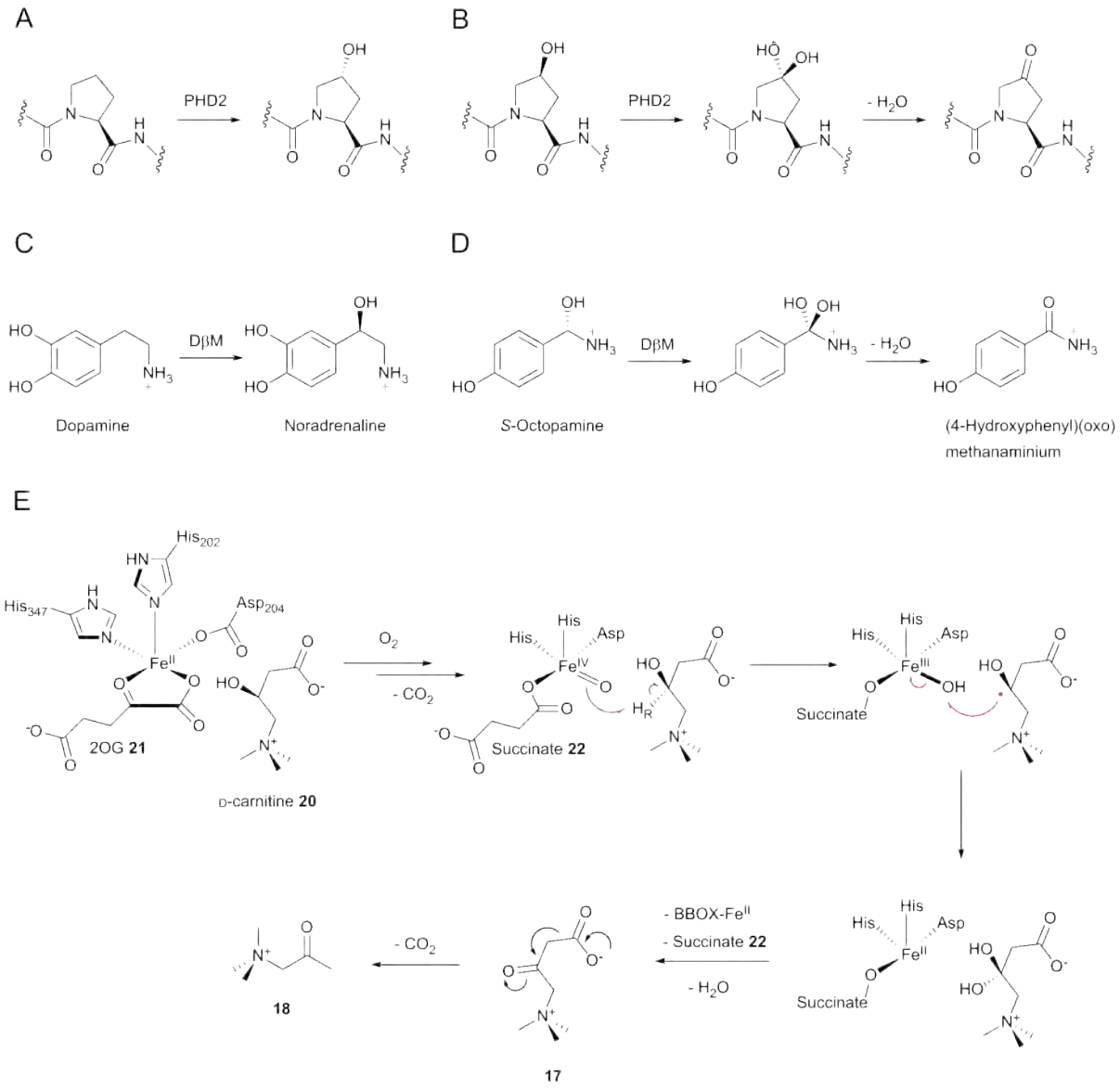
trifluoromethanesulfonate (Scheme 1.6) [120, 123] and further purified by preparative HPLC prior to NMR and MS characterisation. This purified compound was used to confirm the identity of the product in the crude assay.

From a biological perspective, the observation that the unnatural enantiomer of carnitine is converted into *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** by BBOX is interesting because high levels of compound **18** were observed in urine after D-Car **20** administration in mouse and rat [1]. However, no mechanism was suggested for the faster metabolism of D-Car **20** to *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** compared to L-Car **1** in animal tissues. Our results may explain this observation.

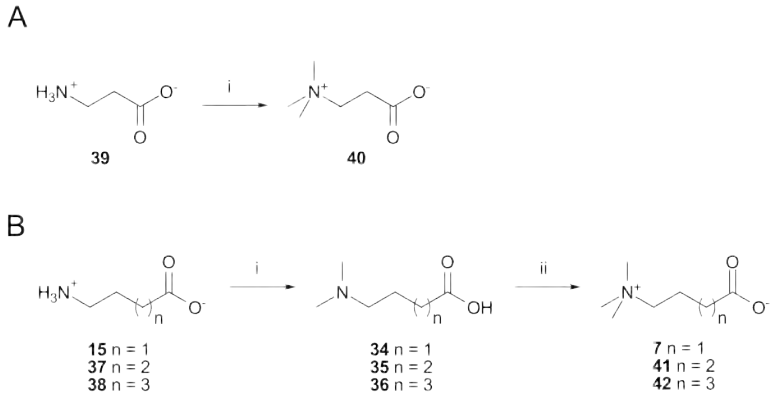
1.2.4 Synthesis of BBOX substrate analogues

In order to perform a structure-activity relationship (SAR) study on BBOX, a series of γ-BB analogues were tested as substrates, or for their ability to compete with γ-BB **7** and inhibit BBOX activity. Several commercially available compounds were tested as obtained from suppliers without further purification. Others were obtained by chemical synthesis (Scheme 1.8, highlighted in blue in Figure 1.7). Many of the compounds tested were selected for their structural similarities with the natural substrate **7**. Three main structural features were investigated: 1 - The tolerance for substrates with a carbon-chain longer or shorter than γ-BB **7**; 2 - The importance of the carboxylate functionality; 3 - The importance of the quaternary ammonium ion. The results for the activity and inhibition assays are given in Section 1.2.5.

A series of dimethylated tertiary amines were synthesised from the corresponding amino acids to test whether the requirement for all three methyl groups of the quaternary ammonium functionality was strict. Dimethylamines 4-(dimethylamino)butanoic acid **34**, 5-(dimethylamino)pentanoic acid **35** and 6-(dimethylamino)hexanoic acid **36** were obtained in apparent quantitative yields using a Eschweiler-Clarke modification of the Leuckardt reaction from the corresponding amino acids **15**, **37** and **38**, respectively (Scheme 1.8 B; for a review see Moore [124]). The Eschweiler-Clarke reaction is useful for the formation of tertiary amines, and also for preventing the overalkylation to the quaternary ammonium species that occurs when using an alkylation reagent such as methyl iodide. In a first



Scheme 1.7: Oxygenases can catalyse the ketonisation of unnatural substrates. (A) Hydroxylation reaction catalysed by PHD2 on the HIF-1α substrate [121]; (B) The HIF-1$\alpha_{556-574}$ peptide with a mutation to a *cis*-4-hydroxyprolyl residue at position Pro-564 is a substrate giving the 4-oxoprolyl containing product upon PHD2-catalysed oxidation [122]; (C) Dopamine hydroxylation reaction catalysed by dopamine β-monooxygenase; (D) Ketonisation of (*S*)-octopamine into (4-hydroxyphenyl)(oxo)methanaminium by dopamine β-monooxygenase [119]; (E) Proposed mechanism for the reaction of BBOX with the unnatural D-Car **20** enantiomer. Enzymatic hydroxylation is followed by spontaneous ketonisation and decarboxylation to give *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18**.

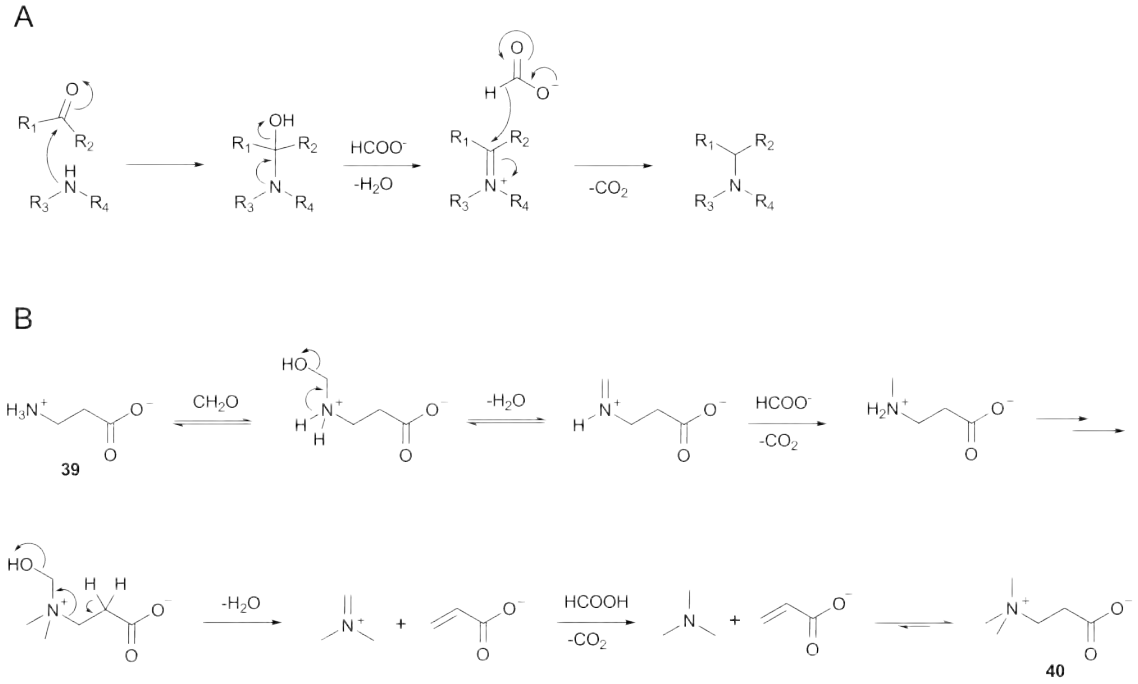


Scheme 1.8: Synthesis of BBOX substrate analogues. (A) (i) 37% aq. Formaldehyde sol., 97% formic acid, reflux, 3 d, 78%; (B) (i) 37% aq. Formaldehyde sol., 2 eq. $NaBH_3(CN)$, 2 d, app. quant.; (ii) 5 eq. MeI, $KHCO_3$, MeOH, 0 °C–rt, 12-48 h, 89–91%.

step, an imine is formed in equilibrium with formaldehyde and a primary amine (Scheme 1.9 A). Under reflux conditions, formate acts as a hydride donor and reduces the imine to the corresponding monomethyl amine. This reaction can be repeated a second time to give the dimethylated tertiary amine. A formaldehyde adduct of the tertiary amine can still form, but as dehydration cannot give the corresponding imine, the reaction stops at this stage (Scheme 1.9 A). Removal of the excess of reagents is then performed by coevaporation with water. It is important to evaporate the excess of formaldehyde solution when the mixture is still warm as formaldehyde tends to polymerise upon cooling.

As it has been previously described, subjecting β-alanine **39** to the same reductive amination conditions gave 3-(trimethylammonio)propanoate **40** (β-alanine betaine) *via* an anomalous Eschweiler-Clarke reaction (Scheme 1.8 A) [125]. The driving force for this reaction is the formation of a conjugated π-system in an acrylate intermediate after elimination of a trimethyliminium cation. This imine will be reduced by formate to give trimethylamine, which can then add into the acrylate double bond to form the quaternary ammonium of β-alanine betaine **40** (Scheme 1.9 B).

The iodide salts of 4-(trimethylammonio)butanoate **7**, 5-(trimethylammonio)pentanoate **41** and 6-(trimethylammonio)hexanoate **42** were synthesised from the corresponding dimethylamines **34**, **35** and **36**, using methyl iodide as the alkylating agent as previously described (Scheme 1.8 B) [99]. Stirring the iodide salts formed upon methylation allowed the oxidation of iodide in water, and further extraction of iodine with diethylether. The quaternary ammonium salts were found to be



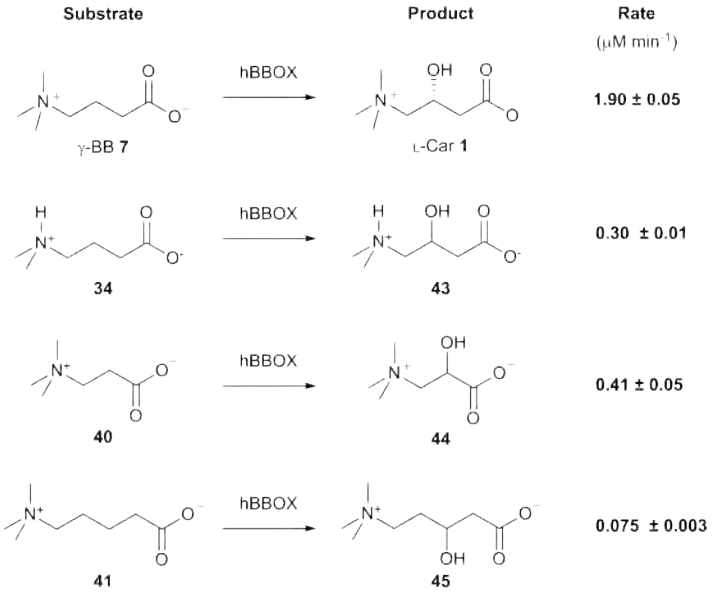
Scheme 1.9: Proposed mechanism for the Leuckardt reaction. (A) The reductive amination of formaldehyde using a secondary amine and formic acid as an hydride donor; (B) Proposed mechanism for the anomalous Eschweiler-Clarke reaction observed when β-alanine is used as a substrate [125].

white crystalline solids with good water solubility.

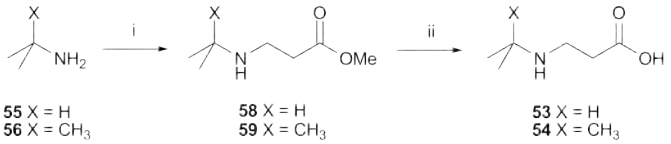
1.2.5 BBOX Catalyses the Hydroxylation of Substrate Analogues

Amongst the compounds tested as potential BBOX substrates, some γ-BB analogues were found to be hydroxylated by BBOX under standard assay conditions, although to a much lower extent than the natural substrate γ-BB **7** (Scheme 1.10) [90]. *N*,*N*-Dimethylaminobutyric acid **34** and 3-(trimethylammonio)propanoate **40** were found to be hydroxylated by BBOX to give 4-(dimethylamino)-3-hydroxybutanoic acid **43** and 2-hydroxy-3-(trimethylammonio)propanoate **44**, respectively, as described [99]. In addition, the γ-BB analogues with an extra methylene group, 5-(trimethylammonio)pentanoate **41** was also found to be hydroxylated at the C-3 position to yield 3-hydroxy-5-(trimethylammonio)pentanoate **45**. 5-(Trimethylammonio)pentanoate **41** was reported not to induce 2OG **21** turnover and was therefore not previously considered a substrate [101].

In addition to the compounds for which the synthesis is described in Section 1.2.4, several commercially available compounds were tested as BBOX substrates or inhibitors. The structures of all the compounds tested are displayed in Figure 1.7. A range of compounds



Scheme 1.10: γ-BB analogues are BBOX substrates. Some of the γ-BB analogues tested were hydroxylated by BBOX. The products were identified by MS and NMR analyses. The position of the hydroxyl group was identified using 2D NMR experiments. 4-(dimethylamino)-3-hydroxybutanoic acid **43**, MS: m/z = 148.1; ^{1}H NMR (700 MHz, D_2O): δ = 4.29 (*m*), 3.14 (*m*), 2.87 (*s*), 2.37 (*m*); ^{13}C NMR (176 MHz, D_2O): δ = 63.0, 61.5, 44.5, 42.0; 2-hydroxy-3-(trimethylammonio)propanoate **44**, MS: m/z = 148.0987; ^{1}H NMR (700 MHz, D_2O): δ = 4.47 (*m*), 3.65 (*m*), 3.17 (*s*); ^{13}C NMR (176 MHz, D_2O): δ = 69.5, 68.5, 54.0; 3-hydroxy-5-(trimethylammonio)pentanoate **45**, MS: m/z = 176.1297; ^{1}H NMR (700 MHz, D_2O): δ = 3.99 (*m*), 3.41 (*m*), 3.07 (*s*), 2.38 (*m*), 1.92 (*m*). The rate of product formation is given in μM min^{-1}. Assay mixtures contained 12.5 nM BBOX, 50 μM Fe(II), 50 μM substrate, 1,5 mM 2OG, 5 mM sodium ascorbate, 80 mM KCl and 50 mM Tris-D_{11} pH 7.5 in D_2O. The standard error is given for at least three independent measurements [90].



Scheme 1.11: Synthesis of 3-(isopropylamino)propanoic acid (X = H) and 3-(*tert*-butylamino)propanoic acid (X = CH_3). (i) 0.75 eq. Methyl acrylate, 0 °C–rt, 12–48 h, 93%; (ii) 1 eq. NaOH THF/H_2O, 16 h, app. quant.

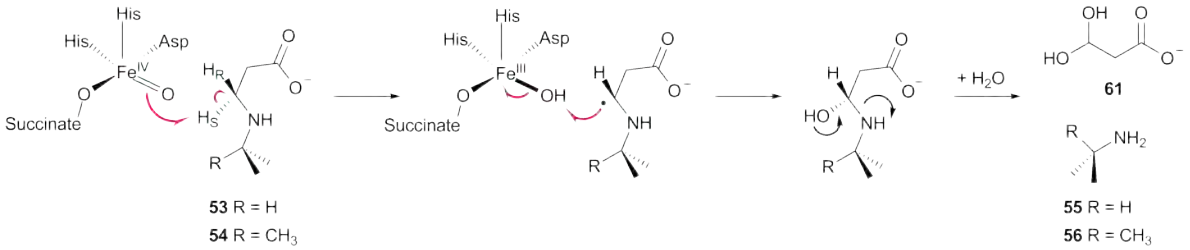
lacking the terminal carboxylate functionality, such as aliphatic trimethylammonium species *N*,*N*,*N*-trimethylethanaminium **46**, *N*,*N*,*N*-trimethylbutan-1-aminium **47** and *N*,*N*,*N*-trimethylhexan-1-aminium **48**, or choline **2** and related species 3-hydroxy-*N*,*N*,*N*-trimethylpropan-1-aminium **49** and 4-hydroxy-*N*,*N*,*N*-trimethylbutan-1-aminium **50** containing a terminal hydroxyl group were tested. None of these compounds led to detectable formation of products, confirming the importance of the carboxylate for substrate binding.

Several compounds from which the terminal trimethylammonium group was absent were also tested. In 3-*tert*-butoxypropanoic acid **51**, the trimethylammonium group is replaced by a *tert*-butyl group and the C-4 methylene by an oxygen atom. In 5-methylhexanoic acid **52**, the trimethylammonium group is replaced by an isobutyl group. None of these compounds led to detectable formation of products, suggesting that the charged terminal ammonium group is required for BBOX activity.

1.2.6 3-Alkylaminopropanoic Acids are BBOX Substrates

Two THP analogues, *N*-isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54**, in which the terminal nitrogen atom is replaced by a carbon atom, were of particular interest as probes for the hydroxylation followed by elimination mechanism initially proposed for THP **27** (Scheme 1.5) [77].

N-Isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54**, were synthesised by adapting a procedure reported for the synthesis of THP **27** (Section 1.3.1) [78], using propan-2-amine **55** (isopropylamine) and 2-methylpropan-2-amine **56** (*tert*-butylamine) instead of dimethylhydrazine **57** (Scheme 1.11). Methyl *N*-isopropyl-β-alaninate **58** and methyl *N*-*tert*-butyl-β-alaninate **59** were obtained in near quantitative yields by addition of methyl acrylate to neat isopropylamine **55** and *tert*-butylamine **56**. Evaporation of the excess of volatile amines gave the pure products without the



Scheme 1.12: 3-(Isopropylamino)propanoic acid and 3-(*tert*-butylamino)propanoic acid are BBOX substrates. The proposed mechanism is based on ^{1}H NMR observation of isopropylamine **55** and *tert*-butylamine **56** formation when BBOX activity was assayed in the presence of *N*-isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54**, respectively.

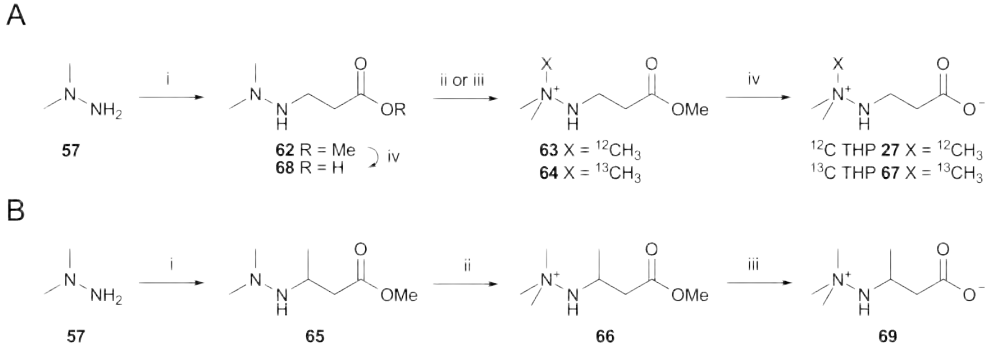
need for any purification. The methyl esters were hydrolysed under alkaline conditions to afford the sodium salts of *N*-isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54** in apparent quantitative yields.

Both *N*-isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54** were oxidised by BBOX in the presence of 2OG **21**, Fe(II) and ascorbate, as judged from the set of new signals that could be observed by ^{1}H NMR spectroscopy. The products formed were postulated to be isopropylamine **55** and *tert*-butylamine **56**, respectively, together with malonic acid semialdehyde **60** in an hydrated form **61**. The identities of isopropylamine **55** and *tert*-butylamine **56** were confirmed in NMR studies involving the addition of pure standards for these two compounds to the assay mixture. Based on these observations, a mechanism was proposed in which an Fe(IV)-oxo mediated proton abstraction occurs, followed by reaction between the resulting radical with the hydroxyl coordinating the metal. This hydroxylated species can eliminate water to form an imine which can then be hydrolysed to give malonic acid semialdehyde and the corresponding amine (Scheme 1.12). These observations support the previously proposed THP metabolic fate (Section 1.1.5, Scheme 1.5) [77].

1.3 A Rearrangement Reaction Catalysed by BBOX

1.3.1 Synthesis of 2-(2-Carboxyethyl)-1,1,1-trimethylhydrazinium, commonly known as 3-(2,2,2-Trimethylhydrazinium)propionate (THP), and Analogues

In order to investigate the mechanism of inhibition of BBOX by THP **27**, this compound and some analogues were synthesised by adapting a reported procedure [78]. Dimethyl hydrazine **57** was added to methyl acrylate to form methyl 3-(2,2-dimethylhydrazino)propanoate **62** (Scheme 1.13 A). 2-(3-Methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium **63** and the [^{13}C]-labeled derivative **64** were obtained *via* alkylation of methyl 3-(2,2-dimethylhydrazino)propanoate **62** using methyl iodide or [^{13}C]-methyl iodide, respectively. The addition reaction was also performed using methyl crotonate in place of methyl acrylate, yielding methyl 3-(2,2-dimethylhydrazino)butanoate **65**. The dimethylhydrazine group of methyl 3-(2,2-dimethylhydrazino)butanoate **65** was alkylated using methyl iodide to give 2-(4-methoxy-4-oxobutan-2-yl)-1,1,1-trimethylhydrazinium **66** (Scheme 1.13 B). Alkaline hydrolysis of the methyl esters in a mixture of water and tetrahydrofuran gave the sodium salts of 3-(2,2,2-trimethyldiazan-2-ium-1-yl)propanoate **27**, [^{13}C]-labeled 3-(2,2,2-trimethyldiazan-2-ium-1-yl)propanoate **67**, 3-(2,2-dimethylhydrazino)propanoic acid **68** (DHP) and 3-(2,2,2-trimethyldiazan-2-ium-1-yl)butanoate **69**. A small amount of these salts were subjected to HPLC purification before enzyme inhibition assays.

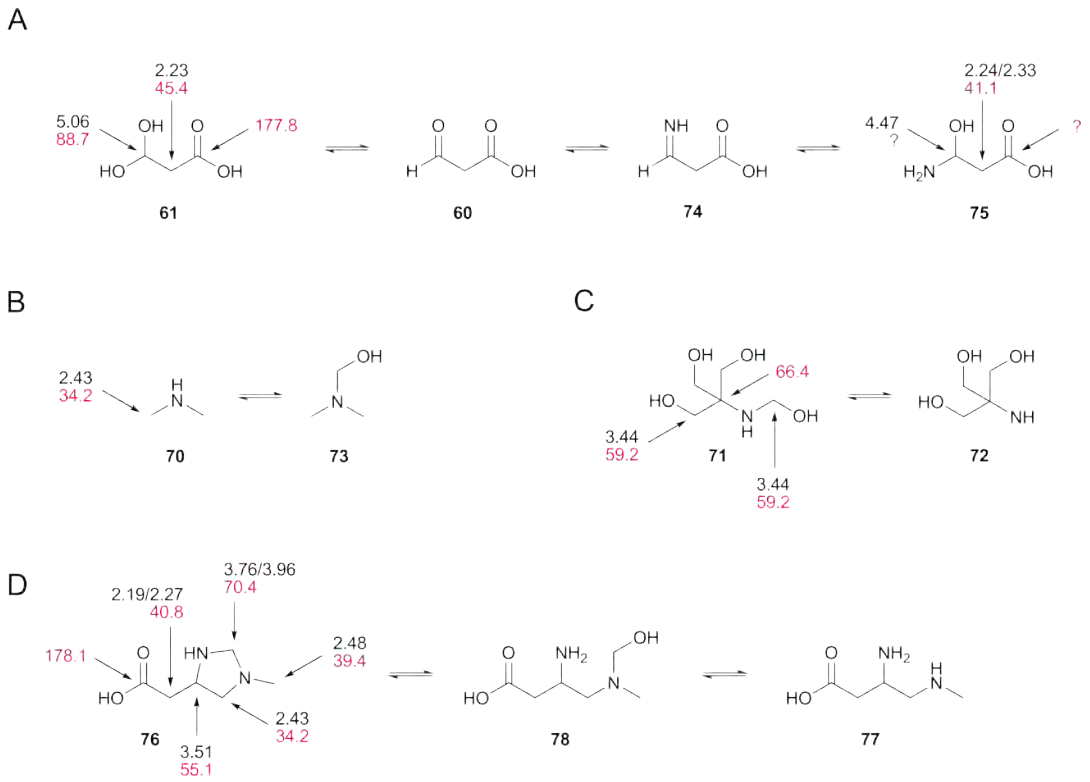


Scheme 1.13: Synthesis of THP and analogues. (A) (i) 0.75 eq. Methyl acrylate, 0 °C–r.t., 24 h, 93%; (ii) 1.6 eq. MeI (X = ^{12}C), 1:9 EtOH/Cyclohexane, 0 °C–r.t., 16 h, 85%; (iii) 1.6 eq. [^{13}C]-MeI (X = ^{13}C), 1:9 EtOH/Cyclohexane, 0 °C–rt, 16 h, 89%; (iv) 1 eq. NaOH, 1:1 THF/H_2O, 16 h, app. quant.; (B) (i) 0.5 eq. Methyl crotonate, 0 °C–r.t., 24 h, 93%; (ii) 1.6 eq. MeI, 1:9 EtOH/cyclohexane, 0 °C–r.t., 16 h, 85%; (iii) 1 eq. NaOH, 1:1 THF/H_2O, 16 h, app. quant.

1.3.2 THP as a BBOX Substrate

The inhibition of human BBOX by THP **27** was studied using the ^{1}H NMR assay described in Section 1.2.2 [90]. THP was found to be an inhibitor with an IC_{50} value of 66 μM. This is in good agreement with the IC_{50} value of 62 μM described previously for purified recombinant human enzyme [85]. However, the ^{1}H NMR analysis suggested that THP **27** is not a reversible competitive inhibitor. Upon incubation of THP **27** under standard BBOX assay conditions, but in the absence of the γ-BB substrate, THP was shown to be converted into several unanticipated products (Scheme 1.14). These products were characterised using a combination of 1D TOCSY, COSY, ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC experiments and were identified as dimethylamine **70**, hydrated malonate semialdehyde **61**, formaldehyde, which forms a characteristic adduct **71** with the amino group of Tris buffer **72**, and an initially unidentified product. It is worth pointing out that no signals corresponding to the methyl groups of trimethylhydrazine or trimethylamine could be observed, ruling out the formation of these compounds. Although it was not observed, the hemiaminal **73** could be formed by the reaction of dimethylamine with formaldehyde. The identity of the peaks corresponding to dimethylamine and formaldehyde were confirmed by addition of pure standards to the reaction mixture. The complexity of the reaction mixture came from the presence of multiple species. For example, hydrated malonate semialdehyde was postulated to stem from imine **74** *via* the hydrated form 3-amino-3-hydroxypropanoic acid **75** (Scheme 1.14 A). In addition, the initially unidentified product was assigned to the cyclic aminal **76** formed by reaction of 3-amino-4-(methylamino)butanoic acid **77** with formaldehyde, likely *via* the hemiaminal **78** (Scheme 1.14 D) [96]. The characterisation of compound **76** is detailed below and a possible mechanism for the BBOX-catalysed formation of 3-amino-4-(methylamino)butanoic acid **77** is described in Section 1.3.3.

The THP analogue lacking the trimethylhydrazinium group, 3-(2,2-dimethylhydrazino)propanoic acid **68** (DHP), was tested in the BBOX assay; Formation of the Tris-formaldehyde adduct **71** was observed, suggesting a similar process to THP oxidation, likely involving a demethylation step. However, the low levels of product formation in the case of DHP **68** did not allow any conclusive product identification to be made (data not shown). Racemic 3-(2,2,2-trimethyldiazan-2-ium-1-yl)butanoate **69** was also tested in the BBOX assay but no product was observed by NMR spectroscopy. Compound **69** was of interest because a quaternary carbon centre may be formed if it



Scheme 1.14: THP is a BBOX substrate and is converted into various products under standard assay conditions. Observed ^{1}H (in black) and ^{13}C chemical shift values. (A) Malonic acid semialdehyde **60**, present in the hydrated form **61**, and the hydrated form of the imine **74**, 3-amino-3-hydroxypropanoic acid **75**; (B) Dimethyl amine **70** likely in equilibrium with the formaldehyde adduct **73**; (C) Tris buffer **72** and the corresponding formaldehyde adduct **71**; (D) initially uncharacterised rearrangement product **77**. As shown in Figure 1.18, product **77** is likely present in the assay as the cyclic formaldehyde adduct **76** *via* the hemiaminal **78**.

followed the same fate as proposed for THP **27** (Section 1.3.3, Scheme 1.16).

NMR Identification of a Rearrangement Product

The structure of the unanticipated rearrangement product originating from THP oxidation was assigned to 3-amino-4-(methylamino)butanoic acid **77** by analysis of the signals observed in various NMR spectroscopy experiments [96]. 1D TOCSY (Figure 1.15) and 2D COSY (Figure 1.16) experiments pointed towards a structure with a 4 carbon atom chain. This implies a rearrangement reaction involving a C–C bond formation as THP **27** only has a 3 carbon chain. ^{1}H-^{13}C HMBC correlation between protons H-a/H-a′ and a low chemical shift carbon resonance (178.1 ppm) suggested the presence of a terminal carboxylate. The ^{13}C chemical shift of carbon C-b and C-c indicated that they were likely to be connected to an heteroatom, probably a nitrogen. ^{1}H-^{13}C HMBC

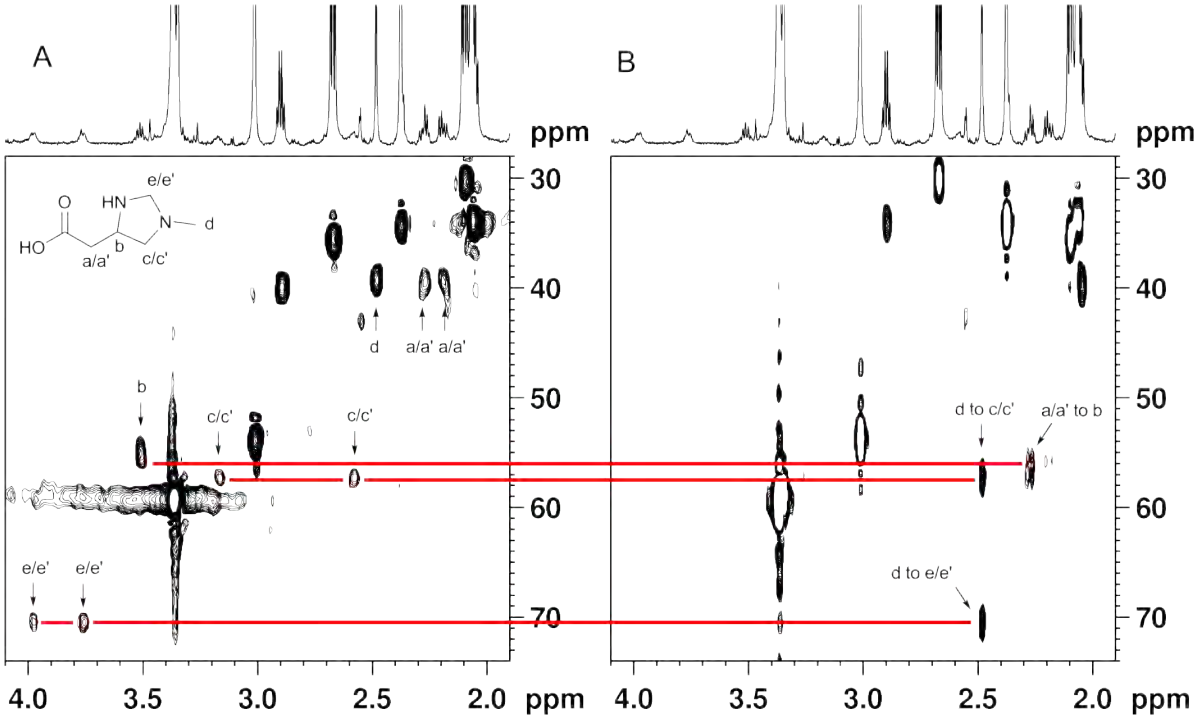


Figure 1.11: Initial NMR characterisation of the cyclic formaldehyde adduct **76** of 3-amino-4-(methylamino)butanoic acid **77** based on 2D ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC experiments. (A) 2D ^{1}H-^{13}C HSQC NMR spectrum for compound **76** in the assay mixture; (B) 2D ^{1}H-^{13}C HMBC NMR spectrum for compound **76** in the assay mixture.

experiments also showed correlations between the C-d methyl protons and two carbon atoms, C-c and C-e (Figure 1.11).

To exclude the possibility that the additional carbon atom found in the 4 carbon chain of 3-amino-4-(methylamino)butanoic acid **77** could have originated from 2OG **21**, the reaction was performed using [1,2,3,4-^{13}C]-2OG as a cosubstrate; however the labels were only found to be retained in succinate (data not shown).

Confirmation of the Rearrangement Product using MS Analyses

The preliminary structural assignment of the rearrangement product **77** based on NMR analyses was uncertain because of the absence of information regarding the nature of the heteroatoms present. Mass spectrometry (MS) analyses were therefore performed to test the assignment of the unexpected product as 3-amino-4-(methylamino)butanoic acid **77**. Despite numerous attempts, and using different instruments, no signal corresponding to the molecular ion of compound **77** could be observed, neither in positive nor in negative ionisation modes. Because 3-amino-4-

(methylamino)butanoic acid **77** could not be directly observed by MS, a derivatisation methodology using a commercially available reagent was investigated. 6-Aminoquinolyl-*N*-hydroxysuccinimidyl carbamate **79** (1-[(quinolin-6-ylcarbamoyl)oxy]pyrrolidine-2,5-dione commonly known as AccQ-Tag, Scheme 1.22 B) is selective for amines in aqueous media over a broad range of pH values [126]. The rate of reaction with primary or secondary amines was demonstrated to be much faster than the rate of hydrolysis. This reagent was originally described for the detection of very low abundance of amines, down to femtomol quantities [126]. The aminoquinoline moiety has a strong absorption and fluorescent properties with emission maximum at λ = 395 nm when excited at λ = 250 nm and is therefore especially well suited for HPLC-MS analysis when both the fluorescence and mass can be detected. AccQ-Tag **79** is now commercially available.

In order to optimise the derivatisation and HPLC conditions, the methodology was first tested on the α-amino acid L-ornithine (Figure 1.12). L-Ornithine is an isostere of 3-amino-4-(methylamino)butanoic acid **77** and also carries two potentially nucleophilic amino functionalities. It should therefore show a similar mass spectrum for its mono and bis-derivatised products. The presence of two amino groups in 3-amino-4-(methylamino)butanoic acid **77** implied that formation of two different mono-derivatised products, as well as the bis-derivatisation could occur, potentially giving a complicated elution pattern. An excess of reagent was therefore used in order to maximise the formation of the bis-derivatised product. L-Ornithine was derivatised following the supplier instructions and the fluorescence and mass spectra were analysed on a UPLC-MS spectrometer after separation on a C18 reverse phase column (Figure 1.12 A). One of the mono-derivatised L-ornithine adducts was observed in low amounts (Figure 1.12 C, m/z = 303.10, R_t = 1.18 min), followed by the bis-derivatised L-ornithine product (Figure 1.12 D, m/z = 473.11, R_t = 6.12 min). Even though the hydrolysis of AccQ-Tag **79** was described to be much slower than its reactivity towards amines, the product of hydrolysis quinolin-6-amine **80** could be observed after it reacted with another derivatisation reagent molecule to form compound **81** (m/z = 315.07, R_t = 6.31 min) (Scheme 1.22 B, and Figure 1.12 B). The relative intensities of the isotopic peaks for each species reflects the natural abundance of the ^{13}C element in this organic compound containing five carbon atoms (m/z = 100%, (m+1)/z = 5.5%, (m+2)/z = 0.3%). The ^{1}H NMR assay mixture containing the rearrangement product 3-amino-4-(methylamino)butanoic acid **77** was subjected to the same derivatisation procedure. An

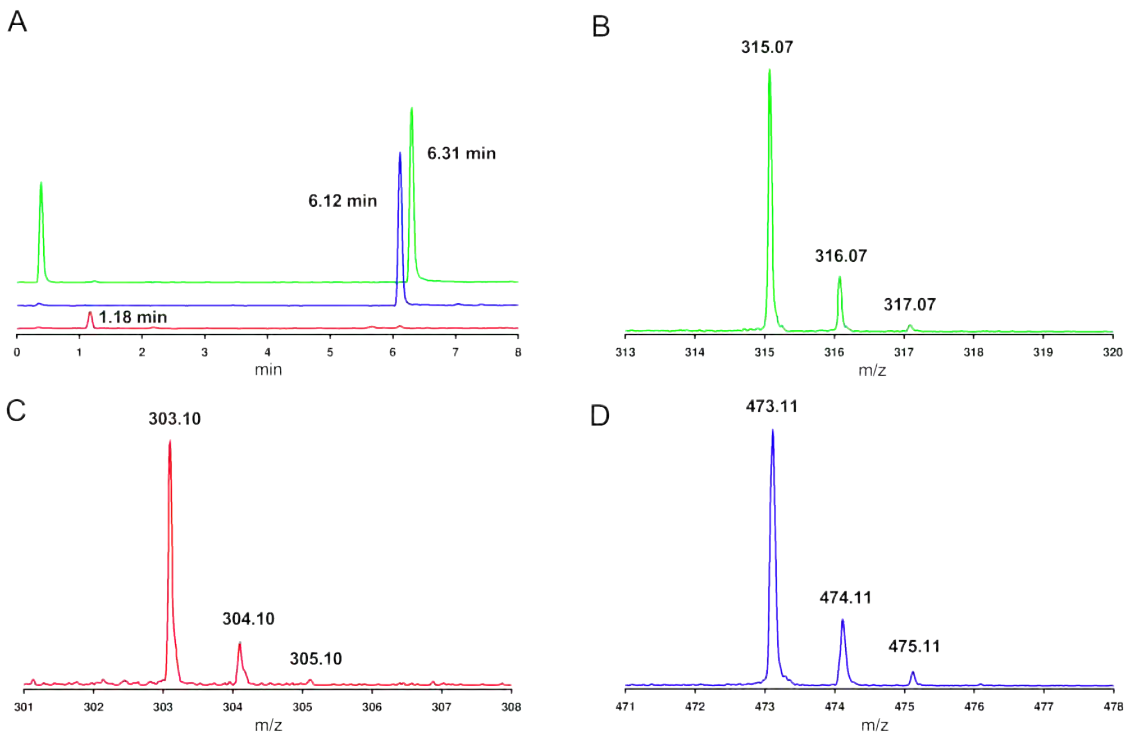
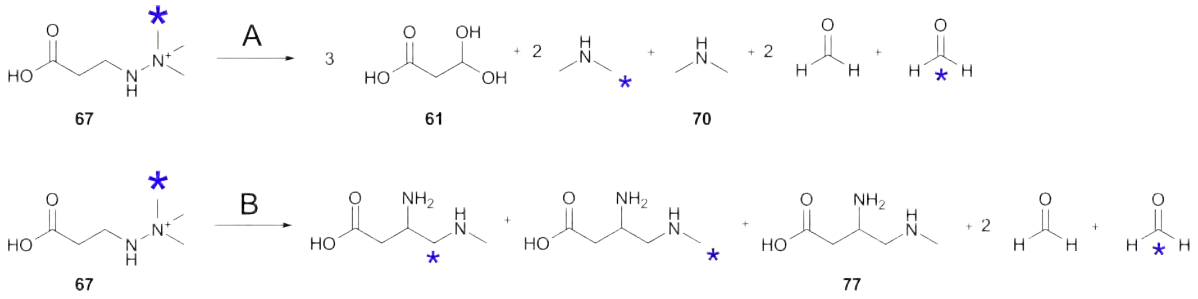


Figure 1.12: Ornithine derivatisation with AccQ-Tag. (A) Single ion chromatograms for the derivatisation reaction of L-ornithine using AccQ-Tag. Monoderivatised L-ornithine (m/z = 303.10) is shown in red (R_t = 1.18 min), bis-derivatised L-ornithine (m/z = 473.11) is shown in blue (R_t = 6.12 min), the excess AccQ-Rag reagent reacts with its own amine hydrolysis product to give a urea (m/z = 315.07) shown in green (R_t = 6.31 min) (Scheme 1.22 B); (B) Mass spectrum for compound **81**; (C) Mass spectrum for monoderivatised L-ornithine; (D) Mass spectrum for bis-derivatised L-ornithine.

excess of AccQ-Tag **79** was used not only to minimise mono-derivatised product formation, but also to overcome the competition between the compound of interest and the amino group of the Tris buffer **72** present in large excess. The resulting mixture was analysed by LC-MS as for L-ornithine. The mono-derivatised 3-amino-4-(methylamino)butanoic acid compound was observed (Figure 1.13 C, m/z = 303.10, R_t = 1.14 min) as well as the bis-derivatised 3-amino-4-(methylamino)butanoic acid compound (Figure 1.13 E, m/z = 473.11, R_t = 5.91 min) and mono-derivatised dimethyl amine (Figure 1.13 G, m/z = 216.09, R_t = 4.16 min). This last observation shows the utility of the AccQ-Tag reagent for the detection of low amounts of volatile amines lacking a chromophore.

Synthesis of [^{13}C]-Labeled THP and BBOX Assays

In order to gain more confidence about the origin of the extra carbon atom present in the 4-carbon chain of the 3-amino-4-(methylamino)butanoic acid rearrangement product **77**, the synthesis of THP



Scheme 1.15: Possible fates of the [^{13}C] label when [^{13}C]-THP **67** is subjected to BBOX reaction. THP labeled with a ^{13}C atom at one of its methyl group is a substrate for BBOX. According to the proposed mechanism (Scheme 1.16), the [^{13}C] label has a 2/3 chance of being found in dimethylamine **70**, and a 1/3 chance of being found in formaldehyde when following path A. When following path B, the [^{13}C] label can have three different fates, two of which retain the label in the rearrangement product at different positions, with the remaining one yielding [^{13}C]-labeled formaldehyde.

with a [^{13}C]-label at one of the methyl group of the trimethylhydrazinium moiety was carried out using commercially available [^{13}C]-methyl iodide (Section 1.3.1, Scheme 1.13 A).

The NMR assay was then performed in the presence of [^{13}C]-THP **67** and the products containing the label were confirmed by ^{13}C NMR to be the Tris-[^{13}C]-formaldehyde adduct, 3-amino-4-(methylamino)butanoic acid **77** with a label at either the C-d methyl group or the inserted C-c methylene, as well as dimethylamine (data not shown). This ^{1}H NMR assay in which [^{13}C]-THP **67** was used as a substrate was further analysed by MS after derivatisation with AccQ-Tag **79**. Considering that only one of the three methyl group of THP was labeled, and that, based on the proposed mechanism, it can follow three different fates, the predicted probability for the label to be retained in compound **77** is 2/3 (Figure 1.15 B). This is in good agreement with the MS results where the observed isotope ratio is 1:2 (Figure 1.13 D and F). The label was also found in dimethylamine **70** where the observed isotope ratio is again 1:2 (Figure 1.13 H). This confirms a fragmentation route where the label has 2/3 chance of being retained in dimethylamine and 1/3 chance of being converted to formaldehyde (Scheme 1.15 A).

1.3.3 Mechanism of the BBOX-Catalysed Rearrangement

As described above, the careful analysis of the ^{1}H-^{13}C NMR correlation spectroscopy data allowed the determination of the connectivity in the unexpected 3-amino-4-(methylamino)butanoic acid **77** product. The mass spectrometry analysis confirmed the structural hypothesis and both MS and ^{13}C

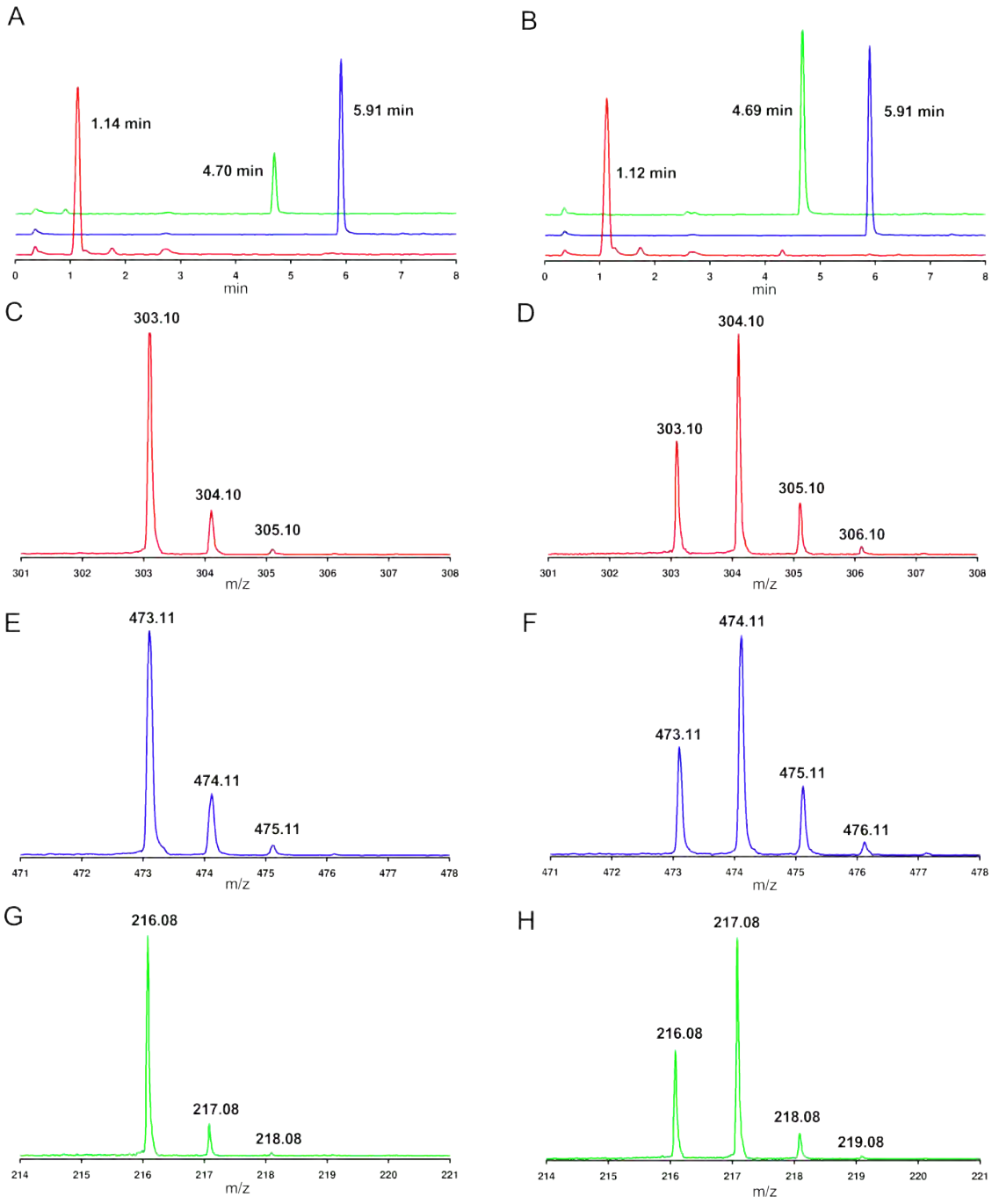
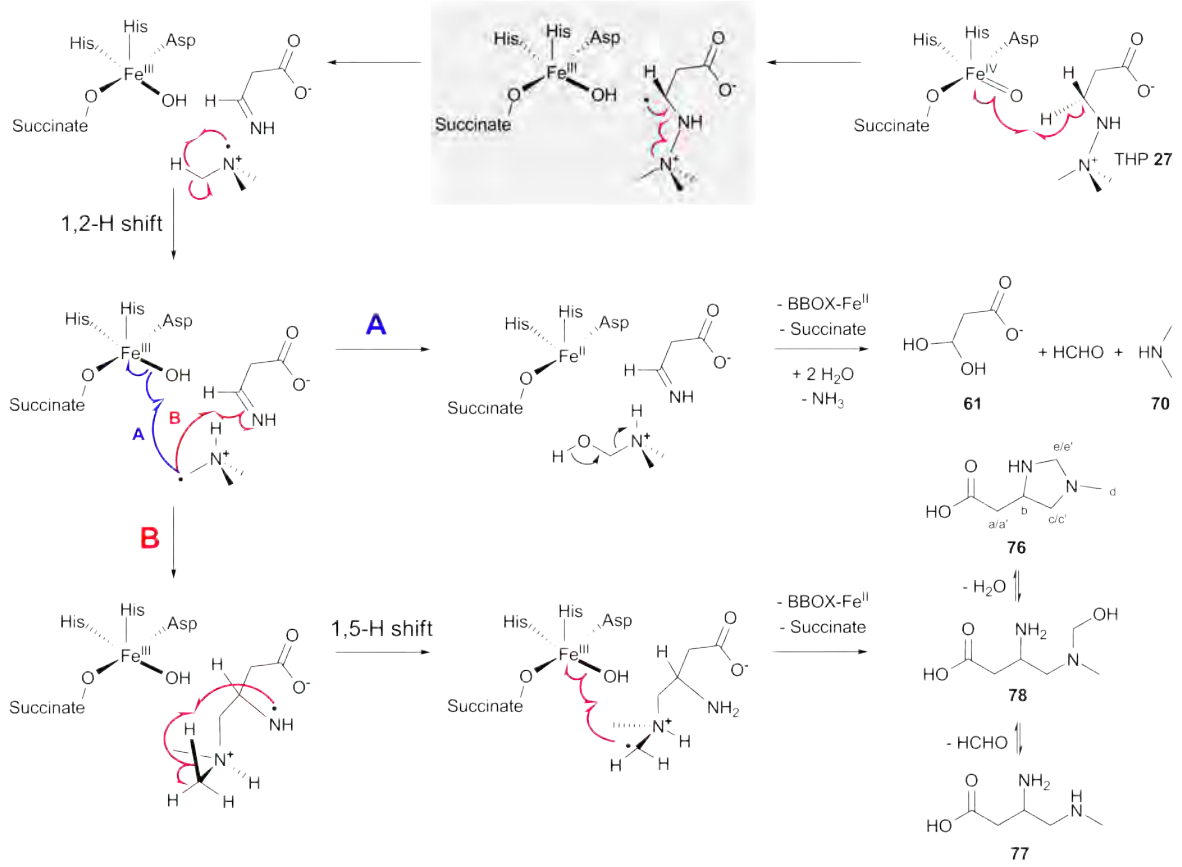


Figure 1.13: MS identification of 3-amino-4-(methylamino)butanoic acid and dimethylamine confirms the rearrangement reaction mechanism. (A) Single ion chromatograms for assay mixture using THP **27**. Monoderivatised 3-amino-4-(methylamino)butanoic acid **77** (m/z = 303.10) is shown in red (R_t = 1.14 min), bis-derivatised **77** (m/z = 473.11) is shown in blue (R_t = 5.91 min), derivatised dimethylamine (m/z = 216.09) is shown in green (R_t = 4.70 min); (B) Single ion chromatograms for the assay mixture using [^{13}C]-THP with BBOX. Monoderivatised **77** (m/z = 304.10) is shown in red (R_t = 1.12 min), bis-derivatised **77** (m/z = 474.11) is shown in blue (R_t = 5.91 min) and derivatised dimethylamine (m/z = 216.09) is shown in green (R_t = 4.69 min); (C) Mass spectrum for mono-derivatised **77** from the THP assay; (D) Mass spectrum for mono-derivatised **77** from the [^{13}C]-THP assay; (E) Mass spectrum for bis-derivatised **77** from the THP assay; (F) Mass spectrum for bis-derivatised **77** from the [^{13}C]-THP assay; (G) Mass spectrum for derivatised dimethylamine **70** from the THP assay; (H) Mass spectrum for derivatised dimethylamine **70** from the [^{13}C]-THP assay.

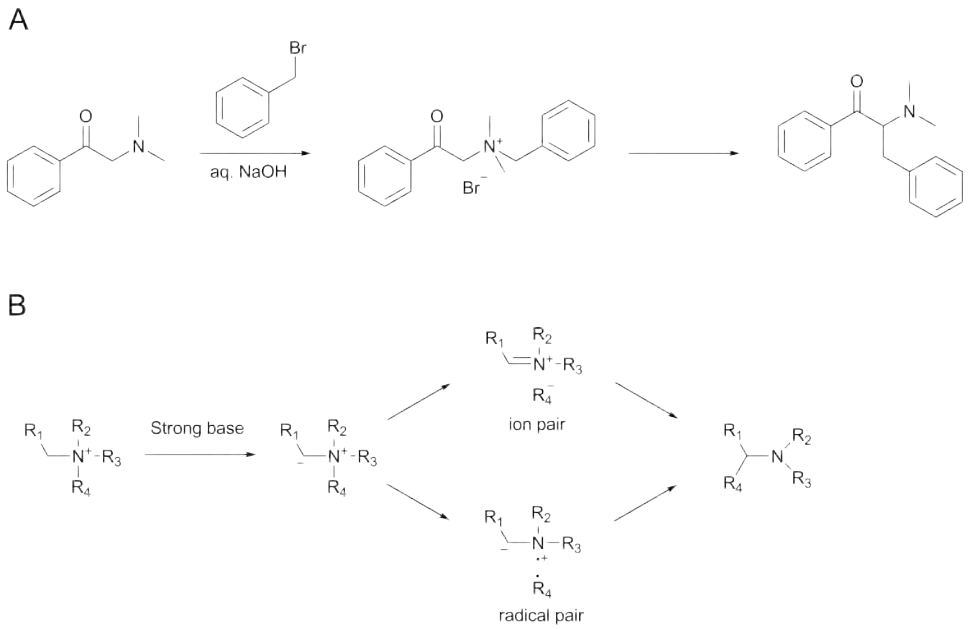
NMR techniques assigned the inserted carbon atom as originating from one of the methyl groups of THP, as shown by the insertion of a [^{13}C]-label at that position when [^{13}C]-THP was used as a substrate. Altogether these results allow a plausible mechanism to be drawn for the reaction of THP **27** with BBOX *in vitro* (Scheme 1.16).

Starting from hydrogen abstraction from THP **27** by the Fe(IV)-oxo species, the radical species formed (highlighted in a gray rectangle in Scheme 1.16) does not react with the hydroxyl group residing on the Fe(III) metal centre as in the case of the natural substrate (Scheme 1.4), but undergoes a homolytic N–N bond cleavage to give a trimethylammonium radical cation. The *N*-radical might be expected to yield trimethylamine or the *N*-hydroxytrimethylamine by reacting with the solvent or the Fe(III)-hydroxyl group. Instead, dimethylamine **70** was observed suggesting a *N*-demethylation step. This implies that a 1,2-hydrogen shift occurs, yielding the methyl radical which can then follow two paths. This methyl radical can react with the Fe(III)-bound hydroxyl group to form a hydroxymethylamine species (Scheme 1.16, Path A) which can spontaneously demethylate to form dimethylamine and formaldehyde. Alternatively, the methyl radical can add into the imine form of malonate semialdehyde, forming a new C–C bond and an amine radical (Scheme 1.16, Path B). This can then undergo a 1,5-hydrogen shift to give again a methyl radical that can react with the hydroxyl group on the metal. The hydroxymethyl species can again spontaneously lose formaldehyde, or form a cyclic adduct **76** with the second amino group. Under the assay conditions, 3-amino-4-(methylamino)butanoic acid **77** seems to participate to an equilibrium shifted towards a cyclic formaldehyde adduct (Scheme 1.16). The rearrangement product is therefore represented as the cyclic structure **76** in the following NMR figures [96]. Based on ^{1}H NMR observation of the approximate relative abundance of these species, path A leading to dimethylamine is thought to be the major reaction path. Path B probably accounts for less than 10% of the total reaction, yielding very low levels of 3-amino-4-(methylamino)butanoic acid **77**.

The closest parallel to the BBOX-catalysed production of 3-amino-4-(methylamino)butanoic acid **77** from THP **27** in synthetic organic chemistry is the Stevens rearrangement. The first example of this rearrangement reaction was reported by Stevens and coworkers and involves the 1,2-shift of a phenyl group from a ammonium salt, to form the corresponding tertiary amine (Scheme 1.17 A)



Scheme 1.16: Proposed mechanism for the reaction between THP and BBOX *in vitro* [96]. After hydrogen abstraction from THP **27** and radical formation (intermediate boxed in gray), a homolytic cleavage of the N–N bond gives a trimethylammonium radical cation which undergoes a 1,2-H shift. The methyl radical can then follow two paths: (A) it can react with the hydroxyl sitting on the metal to form a hydroxymethylamine species which can eliminate formaldehyde to give the observed malonate semialdehyde **61** and dimethylamine **70**; (B) The radical can also attack the C of the imine to give a terminal amine radical which undergoes a 1,5-H shift. The resulting methyl radical can then react with the hydroxyl sitting on the metal to form the hydroxymethylamine species. Elimination of formaldehyde at this stage will give 3-amino-4-(methylamino)butanoic acid **77**, while cyclisation via elimination of water gives the observed cyclic adduct **76**.



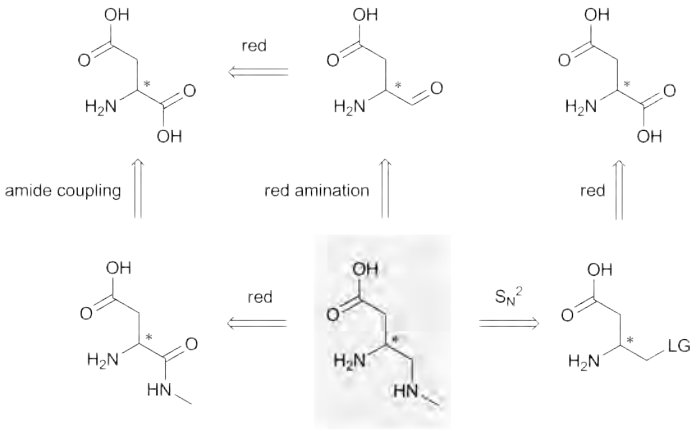
Scheme 1.17: Outline mechanism for the Stevens rearrangement. (A) First reported example of Stevens rearrangement [127]; (B) The proposed mechanism goes *via* an ion pair or a radical pair intermediate [129].

[127]. Thomson and Stevens reported a second example where a phenylacyl group migrates in the same way [128]. The mechanism proposed for this reaction involves an initial deprotonation under harsh conditions to form an ylide and the 1,2-shift can then go through an ion-pair or a radical-pair mechanism (Scheme 1.17 B). Various examples of this reaction are found in the literature and have been recently reviewed in Vanecko et al. [129].

1.3.4 Synthesis of the Rearrangement Product

As mentioned in Section 1.3.3, the minor product of the reaction between BBOX and THP **27** is probably present as an equilibrium between a free 3-amino-4-(methylamino)butanoic acid **77** species and a cyclic formaldehyde adduct **76**. To confirm the identity of 3-amino-4-(methylamino)butanoic acid **77** as the rearrangement product observed when THP **27** is incubated with BBOX, the synthesis of a standard of compound **77** was investigated in order to perform formaldehyde titration and reproduce this equilibrium. After retrosynthetic analysis, several routes were envisaged, starting from aspartic acid (Scheme 1.18).

The key features of compound **77** are the β-amino acid functionality and the *N*-methyl secondary amino group. Monomethylated secondary amines are difficult to access from the corresponding



Scheme 1.18: Retrosynthetic analysis for 3-amino-4-(methylamino)butanoic acid. Various routes were envisaged to access racemic or enantiomerically pure (*S*) and (*R*)-3-amino-4-(methylamino)butanoic acid, 3-amino-4-(methylamino)butanoic acid.

primary amines due to their increased reactivity after a first methylation step, usually leading to overalkylated amines and eventually the quaternary ammonium species. Alkylation using methyl halide or methyl triflate would give the quaternary ammonium ion and reductive amination using formaldehyde would yield the tertiary amine. Both procedures would be difficult to stop at the monomethylated stage. To circumvent the problem of overalkylation, synthetic routes where the monomethylamine was installed as such were investigated.

Attempted Synthesis of 3-Amino-4-(methylamino)butanoic Acid *via* Amide Reduction

The formation of *tert*-butyl N^2-(*tert*-butoxycarbonyl)-*N*-methyl-L-α-asparaginate **82** from the corresponding protected L-aspartic acid derivative **83** and an excess of monomethylamine under peptide coupling conditions, followed by reduction of the amide to the secondary amine *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** was initially investigated (Scheme 1.19 A). Monomethylamine was reacted with protected L-aspartic acid **83** in the presence of peptide coupling reagents 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride and hydroxybenzotriazole. The coupling proceeded in good yield and the resulting product **82** was easily purified. However, the attempted reduction step was found to yield no desired product **84** (Scheme 1.19 A), probably due to the low chemoselectivity of this methodology for the reduction of amides in the presence of esters. When *tert*-butyl N^2-(*tert*-butoxycarbonyl)-*N*-methyl-L-α-asparaginate **82** was reacted in the presence of borane dimethyl sulfide complex, the *tert*-butyl

ester was found to be reduced first to give the *tert*-butyl ether N^2-(*tert*-butoxycarbonyl)-*O*-*tert*-butyl-*N*-methyl-L-homoserinamide **85** and traces of the secondary amine *tert*-butyl [(2*S*)-4-*tert*-butoxy-1-(methylamino)butan-2-yl]carbamate **86** where both the ester and the amide were reduced. The temperature and reaction time were varied in an attempt to optimise product formation, but only the unwanted products were formed, the only difference being the amount of recovered starting material **82**.

Attempted Synthesis of 3-Amino-4-(methylamino)butanoic Acid *via* Nucleophilic Substitution

Two routes involving the nucleophilic substitution by monomethyl amine as a key step were also investigated. One exploited *p*-toluenesulfonyl (*p*-tosyl), the other iodide, as leaving groups, respectively (Scheme 1.19 B and C).

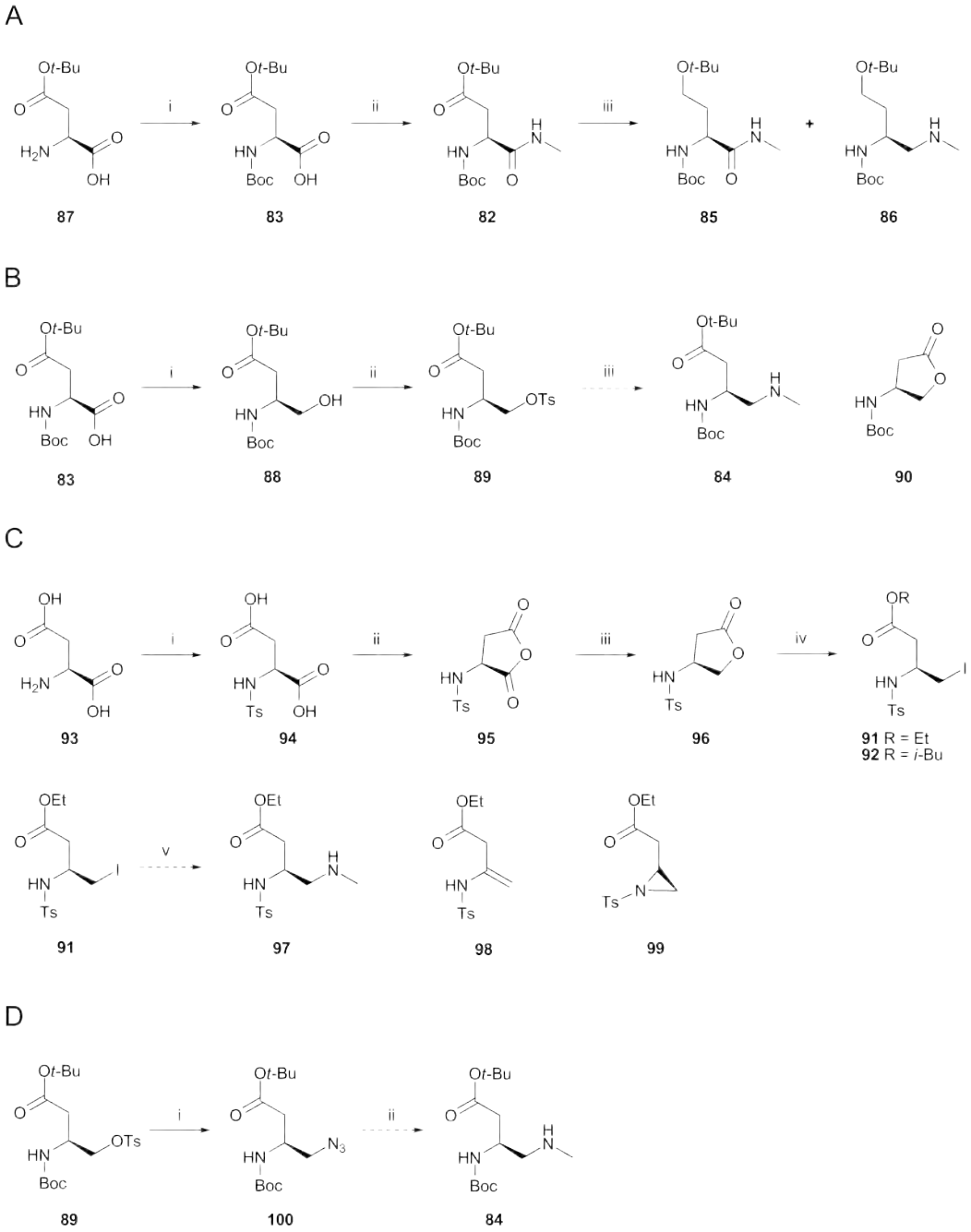
L-Aspartic acid protected as a γ-*tert*-butyl ester **87** was commercially available and was further protected with a *N*-Boc functionality to give urethane **83** (Scheme 1.19 B). This combination of protecting groups was chosen to allow convenient removal under acidic conditions as a last step in the synthesis of 3-amino-4-(methylamino)butanoic acid. Compound **83** was converted to the protected (3*S*)-homoserine alcohol **88** using isobutyl chloroformate to form an anhydride intermediate which was then reduced using sodium borohydride according to reported procedures [130, 131]. The hydroxyl group of *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **88** was tosylated as described [132], to give *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-[(4-methylphenyl)sulfonyl]oxybutanoate **89** which could be subjected to nucleophilic substitution reactions. Direct substitution using a methyl amine solution in tetrahydrofuran was performed at 100 °C in a sealed vial under microwave irradiation. The only product that could be isolated after reaction work-up was *tert*-butyl [(3*S*)-5-oxotetrahydrofuran-3-yl]carbamate **90**, corresponding to the intramolecular cyclisation reaction. Under these reaction conditions where the *tert*-butyl ester may be cleaved to some extent by traces of *p*-toluenesulfonic acid, and an intramolecular process is probably favoured; The carboxylic acid oxygen can act as a nucleophile to substitute the *p*-toluenesufonyl group and form the 5-membered lactone **90**. The lactone (*S*)-*tert*-butyl 5-oxotetrahydrofuran-3-ylcarbamate **90** may be further opened by monomethylamine under the reaction conditions, however the resulting

free carboxylic acid was probably lost during work-up or purification.

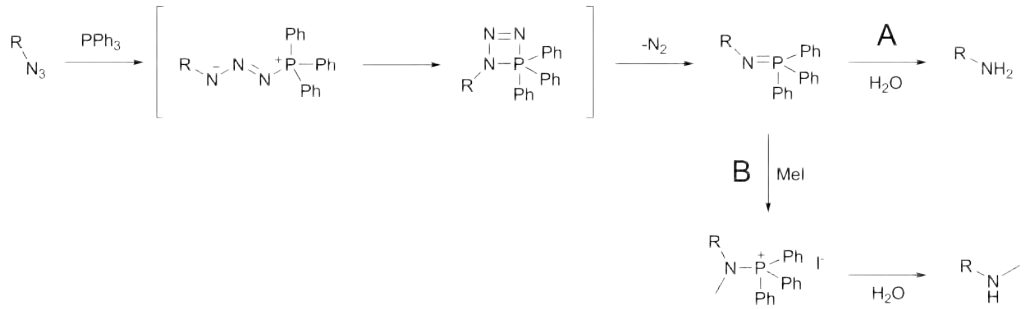
Ethyl and isobutyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate, **91** and **92** respectively, were obtained by adapting reported procedures [133, 134] (Scheme 1.19 C). The amino group of L-aspartic acid **93** was protected as a *N*-tosyl functionality and the resulting diacid **94** was cyclised to give *N*-[(3*S*)-2,5-dioxotetrahydrofuran-3-yl]-4-methylbenzenesulfonamide **95**. Selective reduction of **95** with sodium borohydride gave 4-methyl-*N*-[(3*S*)-5-oxotetrahydrofuran-3-yl]benzenesulfonamide **96** which could be opened using trimethylsilyl iodide in ethanol or isopropanol under mild conditions to afford deoxy-iodo-β-homoserine ethyl or isopropyl ester, **91** and **92**, respectively. When ethyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **91** was subjected to nucleophilic substitution in the presence of a large excess of a monomethylamine solution in tetrahydrofuran, only traces of the desired product ethyl (3*S*)-4-(methylamino)-3-[(4-methylphenyl)sulfonyl]aminobutanoate **97** could be isolated, together with low amounts of the elimination product ethyl 3-[(4-methylphenyl)sulfonyl]aminobut-3-enoate **98** and the aziridine ethyl (2*S*)-1-[(4-methylphenyl)sulfonyl]aziridin-2-ylacetate **99** (Scheme 1.19 C), as judged by MS and ^{1}H NMR analysis of the crude reaction mixture. These products were not present in sufficient amount to be fully characterised. The same reaction was observed in the case of isobutyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **92**.

Attempted Synthesis of 3-Amino-4-(methylamino)butanoic Acid *via* a Staudinger Reaction

A strategy to access monomethylated secondary amines from azides exploits the Staudinger reaction (Scheme 1.20) [135]. The Staudinger reaction involves the treatment of an alkyl azide with triphenylphosphine to produce a phosphoazide which eliminates nitrogen gas and yields an iminophosphorane. The corresponding primary amine can then be accessed by hydrolysis of this iminophosphorane (Scheme 1.20, Path A). Following a similar methodology, the formation of a secondary amine from the primary azide can be achieved through *in situ* 'protection' of a primary amine as an iminophosphorane using triphenylphosphine. The iminophosphorane can be alkylated using an alkyl halide to give the corresponding disubstituted aminophosphonium which cannot get further alkylated. Hydrolysis of that intermediate then yields the monoalkylated secondary



Scheme 1.19: Unsuccessful syntheses of 3-amino-4-(methylamino)butanoic acid. (A) (i) $(Boc)_2O$, 5:3 THF/H_2O, Na_2CO_3, 24 h, 97%; (ii) EDC HCl, HOBt, DIPEA, NH_2Me, THF, 0 °C–rt, 6 h, 84%; (iii) $BH_3 \bullet Me_2S$, THF, 0 °C–rt, 16 h, no product isolated; (B) (i) *N*-Methylmorpholine, isobutyl chloroformate, -30 °C, $NaBH_4$, THF, 30 min, 91%; (ii) pyridine, *p*-TsCl, ACN, 2 d., 38%; (iii) 2 M $MeNH_2$ in THF, MW, 100 °C, 1 h, 48%; (C) (i) 10% aq. NaOH, Me_2CO, *p*-TsCl, iPr_2EtN, 0 °C–rt, 18 h, 93%; (ii) Ac_2O, rt, 12 h, 75%; (iii) $NaBH_4$, THF, 0 °C–rt, 2 h. 45%; (iv) $(CH_3)_3SiI$, CH_2Cl_2, EtOH, rt, 24 h, 67%; (v) 2 M $MeNH_2$ in THF, MW, 100 °C, 1 h, n.i.; (D) (i) NaN_3, DMF, 60 °C, 67%; (ii) Solid supported $(Ph)_3P$, MeI, n.i.



Scheme 1.20: The Staudinger reaction and monomethylation of amines. Hydrolysis of the iminophosphorane intermediate give the primary amine. Alkylation of the iminophosphorane gives the corresponding aminophosphonium species which cannot undergo overalkylation. Hydrolysis of this intermediate gives the secondary amine.

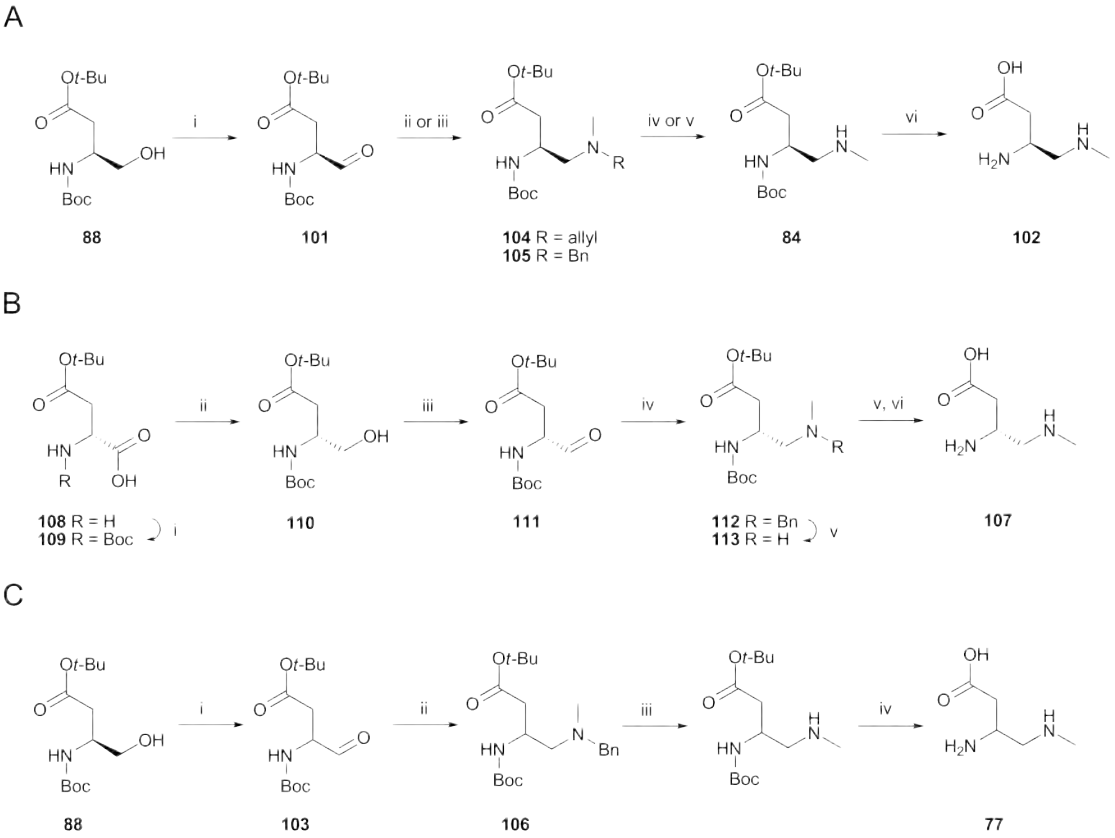
amine (Scheme 1.20, Path B). Classon and coworkers have described a procedure using solid-supported triphenylphosphine for a convenient removal of the reagent by filtration, after alkylation and hydrolysis [136]. In an attempt to install the monomethylated secondary amino group of (3*S*)-3-amino-4-(methylamino)butanoic acid **77** *via* this strategy, *tert*-butyl (3*S*)-4-azido-3-[(*tert*-butoxycarbonyl)amino]butanoate **100** was obtained in moderate yield by substitution of the tosyl group of compound **89** with sodium azide (Scheme 1.19 D). Unfortunately, when tested on compound **100**, the Staudinger reaction in the presence of either triphenylphosphine or the polymer-bound version, followed by alkylation with iodomethane or methyl triflate did not yield any desired product **84**.

Successful Synthesis of 3-Amino-4-(methylamino)butanoic Acid *via* Reductive Amination

The successful syntheses of both enantiomers of 3-amino-4-(methylamino)butanoic acid were achieved following a strategy that involved the reductive amination of an α-amino acid aldehyde with a monomethyl amine synthon (Scheme 1.21). Chiral α-amino acid aldehydes are useful building blocks that can typically be accessed using reductive or oxidative methods that have been reviewed in great detail [137]. Reductive methods include reduction of the corresponding α-amino esters, Weinreb amides, morpholine amides [138] or related activated forms of carboxylic acids derivatives [139]. Examples of oxidative methods include Swern [140], Dess-Martin [141] and manganese (IV) oxide (MnO_2) [142] procedure for the oxidation of β-amino alcohols.

A concern about the use of a reductive amination strategy was a possible epimerisation of the α-amino aldehyde intermediate. The relative stability of α-amino aldehydes obtained from protected serine and threonine amino acids has been studied in in the context of the so-called 'Garner's aldehyde' originally described by Garner and Park [143] and reviewed in Liang et al. [144]. In contrast, the relative instability of other α-amino aldehydes is often avoided by rapidly carrying the next synthetic step on the crude aldehyde. The synthesis of the (3*S*)-enantiomer of *tert*-butyl 3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **101** has been described from *N*-Boc-L-aspartic acid (*O*-*t*-Bu)-OH **83** using a Swern oxidation procedure in two different reports [130, 131]. These authors did not report and optical purity or specific rotation for this compound; however, they mentioned the instability and rapid racemisation of the α-amino aldehyde and therefore did not purify it prior to the Wittig reaction that followed in their synthetic route [131].

In the synthesis described herein, the (3*S*)-enantiomer **102** of 3-amino-4-(methylamino)butanoic acid was successfully obtained starting from the amino acid chiral pool (Scheme 1.21 A). Protected L-aspartic acid **83** was reduced to the protected homoserine **88** as above [130, 131]. the Dess-Martin oxidation procedure gave enantiomerically enriched product **101** after chromatography, as judged from the specific rotation measurement. The specific rotation of α-amino aldehyde **101** was measured again after six months at -20 °C and the result suggested that the compound probably racemised to a large extent, although not completely. In contrast, the Swern oxidation methodology afforded racemic *tert*-butyl 3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **103** after chromatography, as judged by the absence of specific rotation, possibly because the work-up conditions were not as mild as for the Dess-Martin procedure. Reductive amination using a solution of monomethylamine in tetrahydrofuran and sodium triacetoxyborohydride as the reducing agent did not yield any of the expected product *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84**. Two reasons can explain this observation. Monomethylamine is a gas commercially available as aqueous or tetrahydrofuran solutions. The low concentrations that can be achieved in the reaction mixture may not favour the formation of the imine in equilibrium with the aldehyde in the presence of an excess of water. Also, *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** is a polar compound and purification using silica chromatography was found to be difficult. Elution with a relatively polar mobile phase was necessary to recover low amounts of product and it is possible that



Scheme 1.21: Successful syntheses of enantioenriched and racemic 3-amino-4-(methylamino)butanoic acid. (A) (i) Dess-Martin periodinane, CH_2Cl_2, rt, 45 min, 69%; (ii) *N*-Allylmethylamine, $NaBH(OAc)_3$, rt, 16 h, 86%; (iii) *N*-Benzylmethylamine, $NaBH(OAc)_3$, rt, 16 h, 61%; (iv) 5% mol/mol $Pd(PPh_3)_4$, *N*,*N*′-dimethylbarbituric acid, MeOH, rt, 3h, 29%; (v) H_2, Pd C, THF, 3 h, app. quant. (vi) 2M HCl in Et_2O, rt, 2 h, app. quant.; (B) (i) $(Boc)_2O$, K_2CO_3, THF/H2O, rt, 24 h, 97%; (ii) *N*-Methylmorpholine, *i*Bu-chloroformate, $NaBH_4$, THF, -30 °C–rt, 2h, 91%; (iii) Dess-Martin periodinane, CH_2Cl_2, rt, 45 min, 69%; (iv) *N*-Benzylmethylamine, $NaBH(OAc)_3$, rt, 16 h, 81%; (v) H_2, palladium on activated carbon, THF, 3 h, app. quant.; (vi) 2M HCl in Et_2O, rt, 2 h, app. quant.; (C) (i) Oxalyl chloride, DMSO, CH_2Cl_2, -78 °C, 60 min, -78 °C–rt, 30 min, 62%; (ii) *N*-Benzylmethylamine, $NaBH(OAc)_3$, rt, 16 h, 81%; (iii) H_2, palladium on activated carbon, THF, 3 h, app. quant.; (iv) 2M HCl in Et_2O, rt, 2 h, app. quant.

most of the product could have been retained on the silica stationary phase.

To address these problems, the solution of monomethylamine in tetrahydrofuran was replaced by neat allylmethylamine or benzylmethylamine, both non-volatile liquids, and the reductive amination to the corresponding products *tert*-butyl (3*S*)-4-[allyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **104** and *tert*-butyl (3*S*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **105** using sodium triacetoxyborohydride was found to proceed in good yield (Scheme 1.21 A). This route required one additional deprotection step to remove the *N*-allyl or *N*-benzyl groups but allowed the convenient purification of the reductive amination products. The *N*-allyl deprotection using tetrakis(triphenylphosphine)palladium(0) was found to go to completion but removal of the catalyst required the difficult chromatography of *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84**. On the other hand, removal of the benzyl group under hydrogenation conditions using palladium(0) on activated carbon yielded the clean product *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** simply by filtration. The racemate was also carried further through reductive amination with benzylmethylamine and sodium triacetoxyborohydride to give racemic *tert*-butyl 4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **106**, and then through the two successive deprotection steps, to afford the racemic 3-amino-4-(methylamino)butanoic acid **77** (Scheme 1.21 C).

The (3*R*)-enantiomer **107** of 3-amino-4-(methylamino)butanoic acid was synthesised using an identical route to the one used for the (3*S*)-enantiomer, starting from protected D-aspartic acid **108** (Scheme 1.21 B). Protection of the amino group of **108** with di-*tert*-butyl-dicarbonate yielded (2*R*)-4-*tert*-butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **109**. The carboxylic acid function of **109** was reduced as for **83** to yield *tert*-butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **110** [130, 131]. The Dess-Martin oxidation of the protected homoserine **110** gave the enantiomerically enriched aldehyde **111** which was subjected to reductive amination to yield *tert*-butyl (3*R*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **112**. Reductive amination was expected to yield stable chiral products and assuming that the two next sequential deprotection steps, i.e. the hydrogenation of the *N*-benzyl protecting group and the removal of

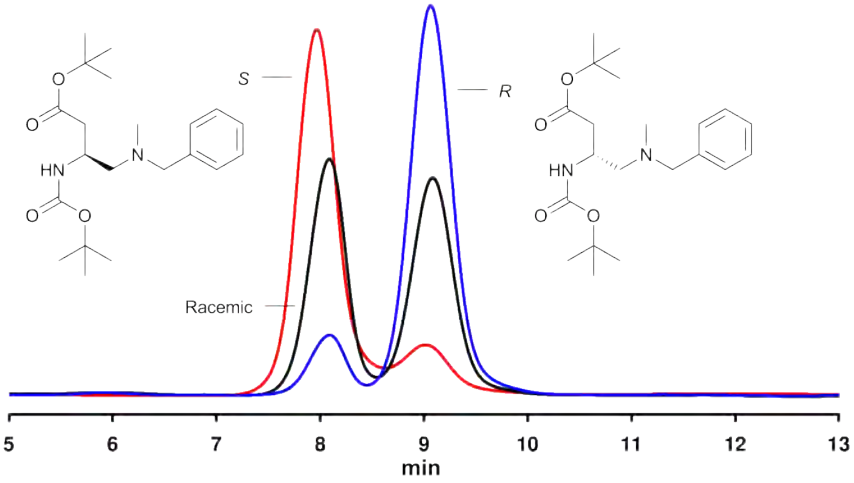


Figure 1.14: Determination of the enantiomeric excess (ee) for (3*S*) and (3*R*)-*tert*-butyl 4-(benzyl(methyl)amino)-3-(*tert*-butoxycarbonylamino)butanoate using normal phase chiral HPLC. Approximately 20 μg of samples in isopropanol were injected on a Chiracel AD column (250 mm x 4.6 mm x 5 μ) and eluted using 1 mL min^{-1} of 5% isopropanol in *n*-hexane. The signal represents the absorbance at 220 nm. The elution profile for the racemic product obtained by reductive amination of the racemic aldehyde with benzylmethylamine is shown in black. The elution profile of the (*S*)-enantiomer obtained starting from L-aspartic acid is shown in red. The elution profile of the (*R*)-enantiomer obtained starting from D-aspartic acid is shown in blue. The full HPLC traces and the integration values are given in Appendix A.1 (3*S*), A.2 (3*R*) and A.3 (racemate).

tert-butyl esters and *N*-Boc groups, would not induce racemisation, the enantiomeric excess of both the *tert*-butyl (3*S*)- and (3*R*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate enantiomers **105** and **112** was assessed by chiral HPLC (Figure 1.14). Racemic *tert*-butyl 4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **106** was used as a standard to develop the normal phase chiral HPLC separation conditions. As expected from the values of the specific rotation obtained for the two products **105** and **112**, the chiral HPLC analysis shows enrichment for the (*S*)- and (*R*)-enantiomer, respectively (**105**: ee = 77%; **112**: ee = 71%). Compounds **105** and **112** were deprotected under hydrogenation conditions to give *tert*-butyl (3*S*)- and (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate, respectively compounds **84** and **113**, which were deprotected under acidic conditions to give the free β-amino acids (3*S*)-3-amino-4-(methylamino)butanoic acid **102** and (3*R*)-3-amino-4-(methylamino)butanoic acid **107**, respectively.

1.3.5 Structural Confirmation for the Rearrangement Product of the Reaction Between BBOX and THP

The mass spectrometry (MS) and nuclear magnetic resonance (NMR) characteristics of a synthetic standard of the racemic 3-amino-4-(methylamino)butanoic acid **77** were recorded and compared with the characteristics of the rearrangement product present in the BBOX assay. However, similar to the rearrangement product present in the assay, the molecular ion for racemic 3-amino-4-(methylamino)butanoic acid **77** or the enantiomerically enriched compounds **102** and **107** could not be observed by MS. Derivatisation with AccQ-Tag **79** was performed as above (Section 1.3.2) and gave results identical to these obtained from the assay mixture (data not shown). As expected, the ^{1}H and ^{13}C NMR NMR data were in accordance with the proposed structure; However the chemical shift values were different from the ones observed in the assay, probably due to the absence of formaldehyde in the sample (Figure 1.17).

To verify the hypothesis that 3-amino-4-(methylamino)butanoic acid forms a cyclic formaldehyde adduct **76** in the assay mixture (Figure 1.14), the racemic 3-amino-4-(methylamino)butanoic acid **77** synthetic standard was titrated with [^{12}C]-formaldehyde under buffered conditions similar to the assay. The ^{1}H NMR spectrum was recorded (Figure 1.15 A) and 1D TOCSY NMR analysis with irradiation at 3.58 ppm reproduced the chemical shift and coupling patterns observed in the reaction mixture (Figure 1.15 B and C). The addition of formaldehyde induced a large difference in both the ^{1}H and the ^{13}C NMR chemical shifts, as judged by HSQC (Figure 1.17), reproducing the pattern observed in the assay (Figure 1.11). It is proposed that this large chemical shift difference between H-c and H-c′ are due to the rearrangement product **77** being present mainly as the cyclic formaldehyde adduct **76**. To further confirm the presence of the cyclic formaldehyde adduct **76**, a solution of synthetic 3-amino-4-(methylamino)butanoic acid **77** was subjected to titration with [^{13}C]-formaldehyde. The enrichment in [^{13}C]-labeled compound allowed to observe HMBC cross-peaks between the bridging carbon atom C-e originating from the [^{13}C]-formaldehyde and the protons H-b, H-c/c′ and H-d found 3 bonds away (Figure 1.18) [96].

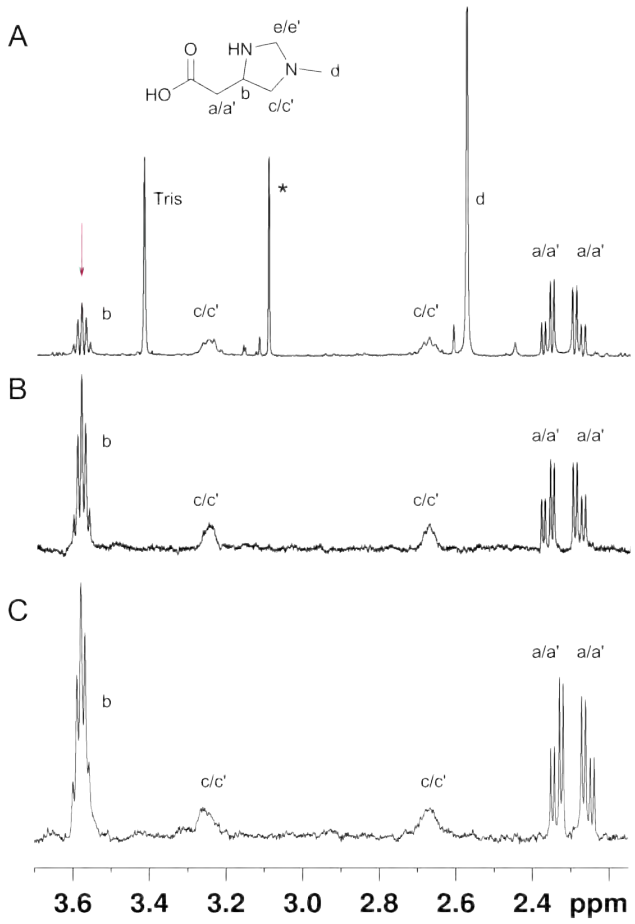


Figure 1.15: ^{1}H and 1D TOCSY NMR analyses for the rearrangement product and the 3-amino-4-(methylamino)butanoic acid synthetic standard. Spectra were recorded at 700 MHz in solutions buffered in Tris/Tris-D_{11} (pH 7.5) in 95% H_2O and 5% D_2O at 280 K. When a formaldehyde solution in the same buffer was titrated to the synthetic standard of 3-amino-4-(methylamino)butanoic acid, the cyclic adduct was formed. For 1D TOCSY analysis, specific irradiation was applied at 3.58 ppm corresponding to the signal for H-b. (A) ^{1}H NMR for synthetic 3-amino-4-(methylamino)butanoic acid in the presence of [^{12}C]-formaldehyde. An impurity from the formaldehyde solution is indicated with a star; (B) 1D TOCSY NMR for synthetic 3-amino-4-(methylamino)butanoic acid in the presence of [^{12}C]-formaldehyde; (C) 1D TOCSY NMR of the rearrangement product in the assay mixture [90].

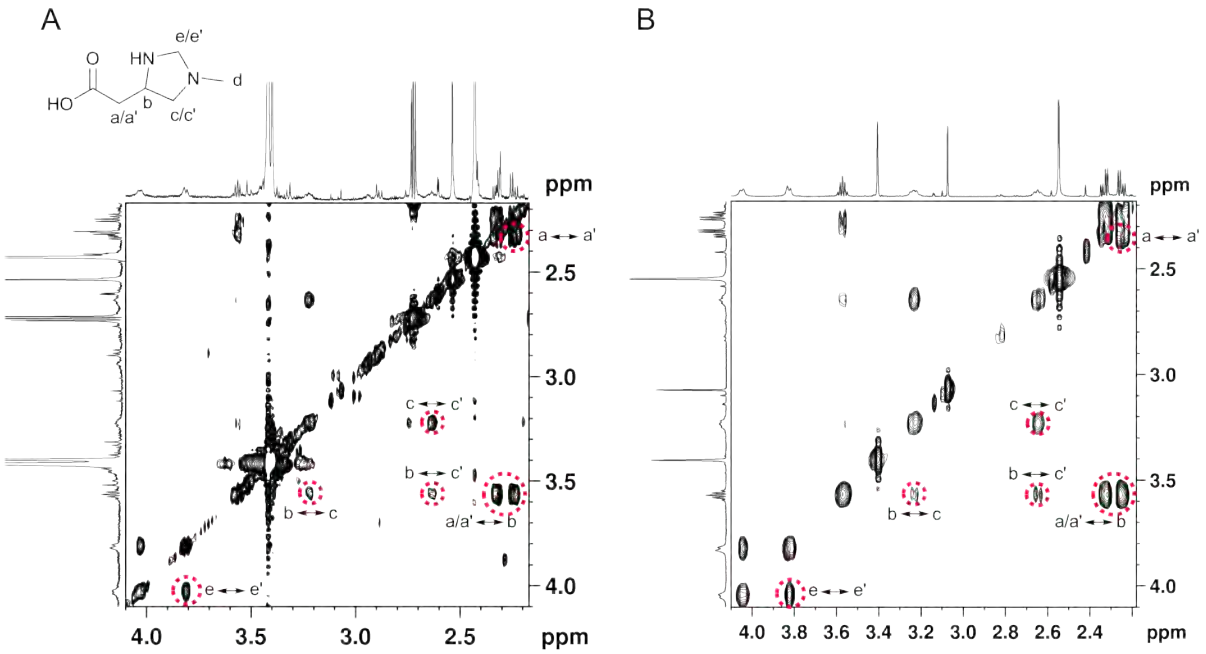


Figure 1.16: 2D COSY analysis for rearrangement product and synthetic standard. Titration of [^{12}C]-formaldehyde to synthetic 3-amino-4-(methylamino)butanoic acid reproduces the ^{1}H-^{1}H NMR correlation observed in the assay. A formaldehyde solution was titrated with the racemic standard of 3-amino-4-(methylamino)butanoic acid **77** to obtain the cyclic adduct **76**. (A) 2D COSY NMR spectrum for compound **76** formed in the assay; (B) 2D COSY NMR spectrum for synthetic standard of compound **77** in the presence of formaldehyde [90].

1.3.6 Attempted Stereochemical Assignment

We were interested in the identification of the C-3 stereochemistry of the rearrangement product 3-amino-4-(methylamino)butanoic acid product **77** formed upon reaction of BBOX with the clinically used inhibitor THP **27**. The mechanism suggesting how BBOX can turn THP **27** into 3-amino-4-(methylamino)butanoic acid **77** implies a series of highly reactive radical species (Scheme 1.16). These radical species would probably react with water molecules if they were exposed to the solvent. The absence of water molecules near the reaction centre can be partly explained by the hydrophobicity of the trimethylamonium binding pocket and the closed active site (Section 1.1.6). Since the protein active site environment is chiral, some stereoselectivity could be envisaged in the formation of the rearrangement product through C–C bond formation. Having the two enantiomerically enriched amino acids **102** and **107** in hand, as well as the racemate **77** (Section 1.3.4), it was desirable to develop an analytical method for their separation and then compare the result with the characteristics of the rearrangement product present in the assay mixture.

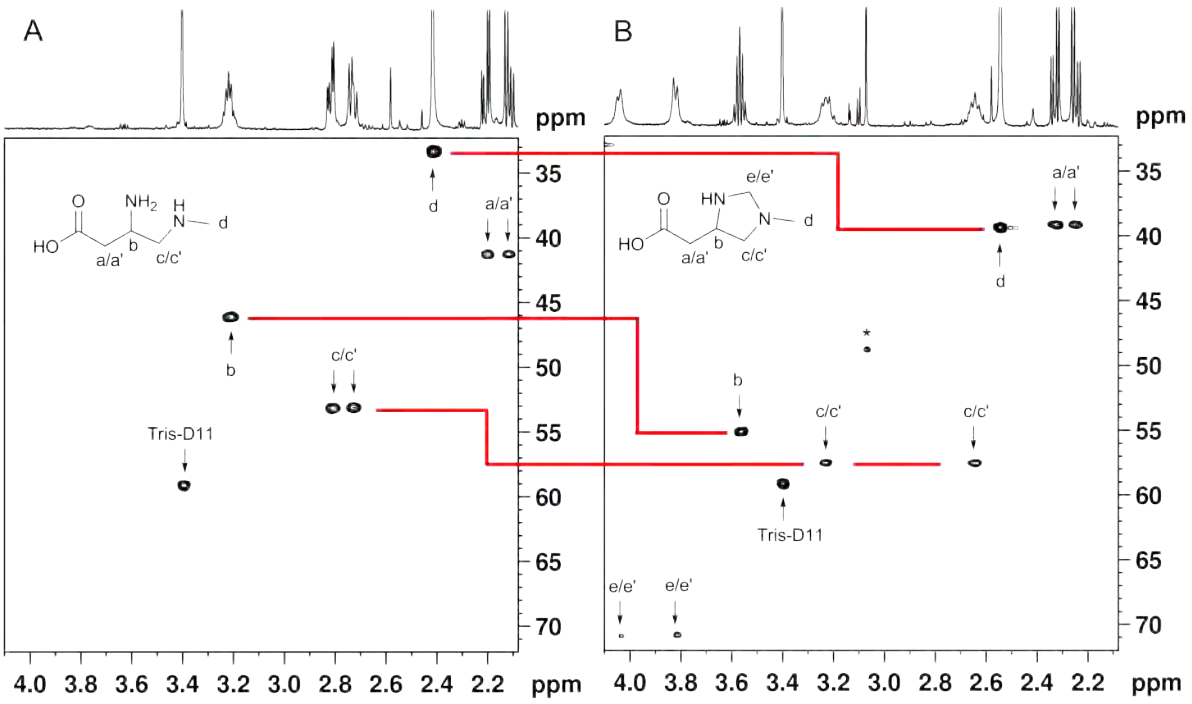


Figure 1.17: 2D HSQC NMR analysis for synthetic standard with and without [^{12}C]-formaldehyde. The addition of [^{12}C]-formaldehyde to racemic 3-amino-4-(methylamino)butanoic acid **77** induces large chemical shift differences consistent with a cyclic formaldehyde adduct **76**. (A) 2D ^{1}H-^{13}C HSQC NMR for compound **77** in the absence of [^{12}C]-formaldehyde; (B) 2D ^{1}H-^{13}C HSQC NMR for compound **76** in the presence of [^{12}C]-formaldehyde [96].

Because of the extremely low levels of 3-amino-4-(methylamino)butanoic acid **77** in the assay mixture, and the presence of numerous other components, including a large excess of Tris buffer, salts, cofactors and protein, a sensitive and reliable analytical technique was needed. As described in Section 1.3.2, the product itself could not be observed directly by mass spectrometry, therefore two different techniques to determine the stereochemistry of the rearrangement product could be envisaged. First, a selective and efficient derivatisation *in situ* followed by chromatography separation, and second, the use of a ^{1}H NMR chiral shift reagent. Only the chromatography option was further investigated because is was anticipated to be easier to implement and the NMR method might be complicated by the multiple species present.

Towards the Chromatographic Resolution of the 3-Amino-4-(methylamino)butanoic Acid Enantiomers

Several approaches for the chromatographic separation of α-amino acid enantiomers have been

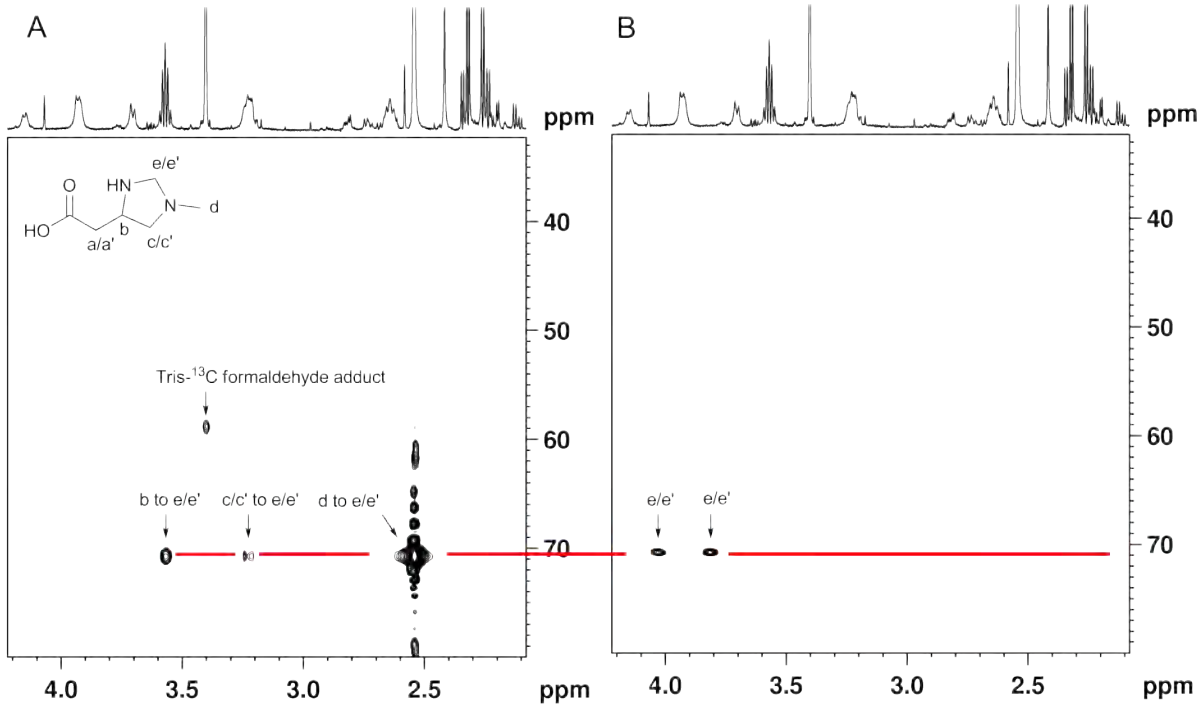
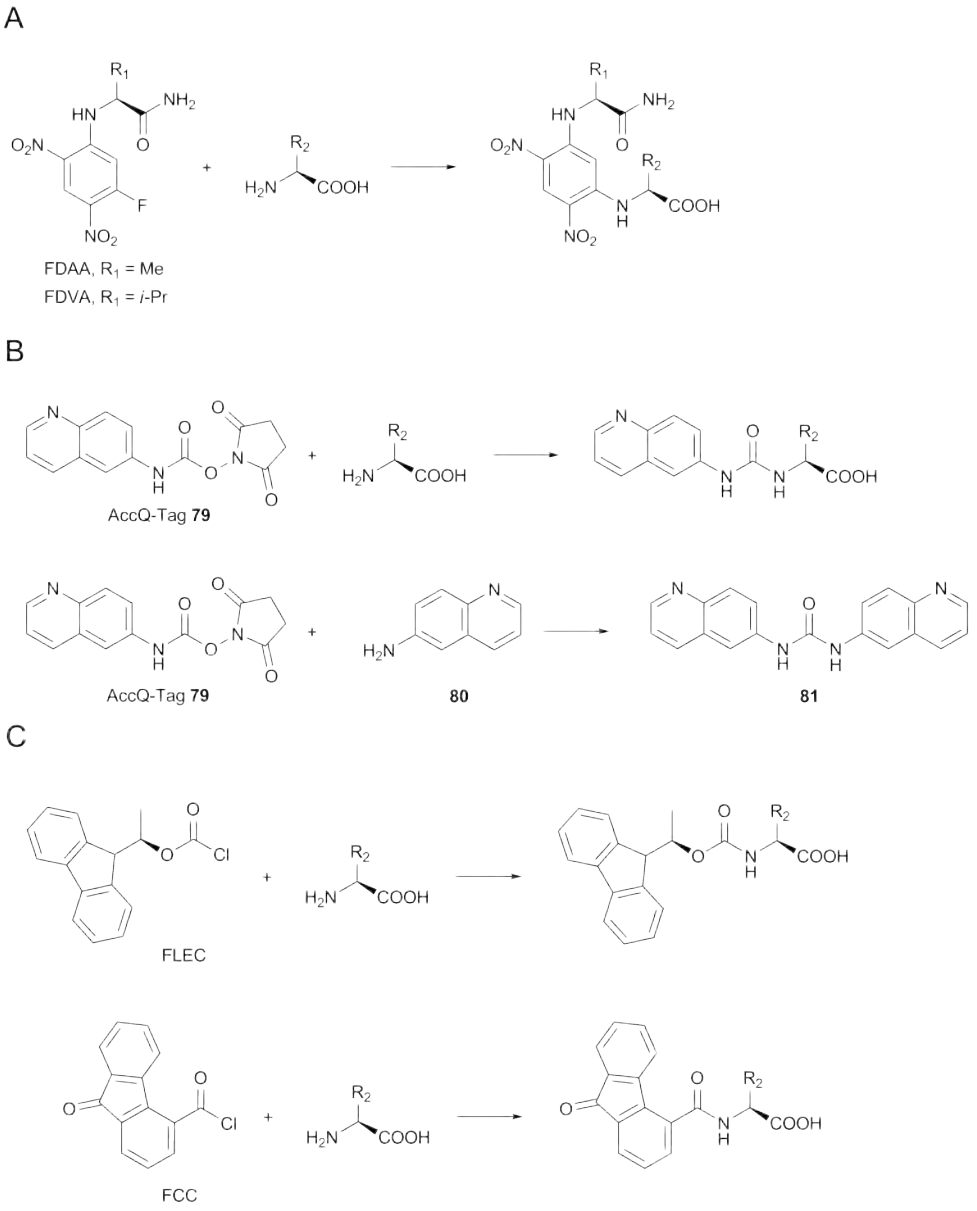


Figure 1.18: 2D HMBC NMR analysis for synthetic standard in the presence of [^{13}C]-formaldehyde. Addition of [^{13}C]-formaldehyde to (3*S*)-3-amino-4-(methylamino)butanoic acid **102** reveals long-range correlations consistent with a cyclic aminal formaldehyde adduct. (A) 2D ^{1}H-^{13}C HMBC NMR for 3-amino-4-(methylamino)butanoic acid cyclic adduct **76** in the presence of [^{13}C]-formaldehyde. The 3-bond correlations between the added [^{13}C]-formaldehyde and protons H-b, H-c/c′ and H-d are clearly present; (B) 2D ^{1}H-^{13}C HSQC NMR for 3-amino-4-(methylamino)butanoic acid cyclic adduct **76** in the presence of [^{13}C]-formaldehyde. The value obtained by single quantum correlation for the carbon chemical shift (70.1 ppm) confirms that the cross-peaks observed in the HMBC spectrum correspond to correlations with a C atom originating from [^{13}C]-formaldehyde [96].

explored [145]. In contrast, few examples of methodologies for the separation and identification of β-amino acid enantiomers are described in the literature. The existing examples apply to the separation of β-amino acid with hydrophobic side chains, reporting difficult and poor separation when using direct or indirect separation on chiral columns [146]. The relative hydrophilicity of 3-amino-4-(methylamino)butanoic acid **77** was expected to lead to even poorer separation. In order to improve the retention of 3-amino-4-(methylamino)butanoic acid on reverse phase HPLC columns, an indirect separation was tested, i.e. after derivatisation of the analyte using various reagents known to react with amine functionalities (Scheme 1.22). This indirect approach can be further divided in two categories. In the first category, the pure enantiomers are derivatised with a chiral auxiliary to generate diastereomers which can be separated on an achiral stationary phase. In the second, the enantiomers are derivatised with an achiral functionality and separated with the help of a chiral stationary phase.

Derivatisation of 3-Amino-4-(methylamino)butanoic Acid Using a Chiral Reagent

Initially, an indirect separation strategy using a conventional C18 reverse phase column was investigated in an attempt to separate the two enantiomers after derivatisation with a chiral auxiliary. Marfey's reagent is a chiral variant of the highly reactive amine derivatisation reagent 2,4-dinitrofluorobenzene, or Sanger reagent, originally developed to identify free amino groups in proteins [147]. Marfey's reagent carries a fluorescent 2,4-dinitrobenzyl group, an α-amino acid based chiral centre and a reactive fluorine leaving group. When reacted with a primary or secondary amino group, it undergoes substitution at the fluorine position, releasing fluoride and forming the aromatic amine covalent adduct. Two versions of chiral Marfey's reagent are commercially available: N_α-(2,4-dinitro-5-fluorophenyl)-L-alaninamide (FDAA) and N_α-(2,4-dinitro-5-fluorophenyl)-L-valinamide (FDVA) (Scheme 1.22 A). Marfey's reagent has proven very useful to separate enantiomers of various natural and unnatural α-amino acids, as well as other chiral amines, and was therefore a reagent of choice to separate the enantiomers of 3-amino-4-(methylamino)butanoic acid **77** [148]. However, any attempt to separate the two enantiomers of 3-amino-4-(methylamino)butanoic acid (*S*)-**102** and (*R*)-**107** after derivatisation with FDVA failed, despite an optimisation of the gradient and promising results for ornithine and lysine, two amino acids also containing two amine functionalities (data not shown).



Scheme 1.22: Examples of derivatisation agents for amino groups. (A) FDAA (R_1 = Me) and FDVA (R_1 = *i*-Pr) react with primary and secondary chiral amines to give diastereomeric products through a S_N2 mechanism; (B) AccQ-Tag^{TM} reacts with with primary and secondary amines to give the corresponding urea; (C) FLEC, a chiral Fmoc derivative, and FCC react with primary and secondary amines to give the corresponding amides.

The derivatisation of 3-amino-4-(methylamino)butanoic acid enantiomers with (+)-1-(9-fluorenyl)ethyl chloroformate (FLEC), a chiral analogue of the fluorenylmethyloxycarbonyl (Fmoc) protecting group was envisaged but not tested (Scheme 1.22 C). This chloroformate has been shown to react with α-amino acids to form the corresponding diastereomeric carbamates which can be separated [149].

Attempted Separation on a Chiral Stationary Phase

Separation of the two enantiomers of 3-amino-4-(methylamino)butanoic acid using a strategy based on chiral HPLC columns was investigated. Three different columns were tested. Because 3-amino-4-(methylamino)butanoic acid enantiomers are highly polar compounds, the underivatised (3*S*)-3-amino-4-(methylamino)butanoic acid **102** and (3*R*)-amino-4-(methylamino)butanoic acid **107** eluted with the void volume on all available chiral columns, using any of the tested elution systems and conditions (data not shown). Derivatisation of the amino groups of 3-amino-4-(methylamino)butanoic acid **77** was therefore performed prior to chiral HPLC analyses.

The first column tested was a LiChroCART® ChiraSpher® NT column (250 mm x 4.0 mm, 5 μm particle size), which has a stationary phase based on silica particles coated with poly-(*N*-acryloyl-(*S*)-phenylalanine ethyl ester. The bis-derivatisation with AccQ-Tag **79** which led to the production of bulky aromatic ureas was likely to increase the hydrophobicity of 3-amino-4-(methylamino)butanoic acid to a great extent, therefore increasing the retention time. However, the bis-derivatised adducts of (3*S*)-3-amino-4-(methylamino)butanoic acid **102** (Figure 1.19 A–D, red trace) and (3*R*)-3-amino-4-(methylamino)butanoic acid **107** (Figure 1.19 A–D, green trace) eluted with the same retention time, as shown by overlapping traces for the enriched enantiomers together with a racemic mixture (Figure 1.19 A–D, black trace). L- and D-proline were also subjected to derivatisation with AccQ-Tag **79**; The resulting products were tested as controls, however no separation could be achieved, suggesting an unsuited stationary phase (Figure 1.19 E and F).

Two chiral columns specifically designed for the separation of polar analytes were then tested. ChiroBiotic V™ and ChiroBiotic T™ columns are packed with stationary phases based on silica particles coated with cyclic peptides. The ChiroBiotic V separation is based on a stationary phase

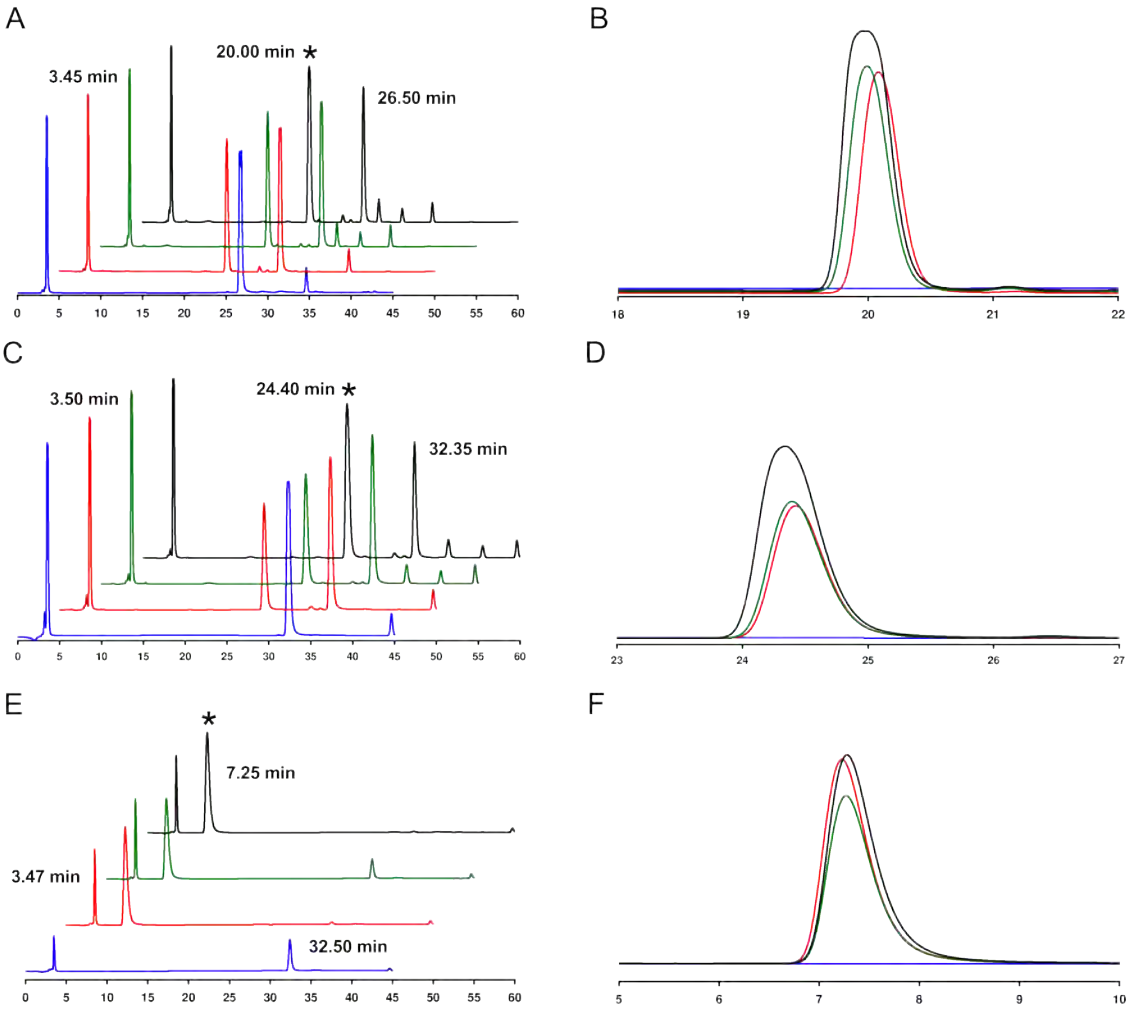


Figure 1.19: HPLC traces for standards derivatised with AccQ-Tag using a ChiraSphere® NT Column. Attempts to separate compounds (3*S*)-3-amino-4-(methylamino)butanoic acid **102** and (3*R*)-3-amino-4-(methylamino)butanoic acid **107** on a ChiraSphere NT column (250 mm x 4 mm, particle size 5 μm) after derivatisation with AccQ-Tag **79**. The UV absorbance was monitored at 245 nm. L-proline and D-proline derivatised with AccQ-Tag could not be separated on this column either. (A) Full chromatograms for control with no analyte (shown in blue), pure **102** (shown in red), pure **107** (shown in green) and a equimolar mixture of **102** and **107** (shown in black). The elution gradient was as follow: 2% B in A for 5 min, 2–98% B in A over 40 min. A = 20 mM NH_4OOCH, B = ACN; (B) Expansion of chromatograms for control (shown in blue), pure **102** (shown in red), pure **107** (shown in green) and a equimolar mixture of **102** and **107** (shown in black); (C) Full chromatograms for control (shown in blue), pure **102** (shown in red), pure **107** (shown in green) and a equimolar mixture of **102** and **107** (shown in black). The elution gradient was as follow: 2% B in A for 5 min, 2–98% B in A over 40 min. A = 20 mmol NH_4OOCH, B = MeOH; (D) Expansion of chromatograms for control (shown in blue), pure **102** (shown in red), pure **107** (shown in green) and a equimolar mixture of **102** and **107** (shown in black); (E) Full chromatograms for control (shown in blue), pure L-proline (shown in red), pure D-proline (shown in green) and a equimolar mixture of L-proline and D-proline (shown in black). The elution gradient was as follow: 2% B in A for 5 min, 2–98% B in A over 40 min. A = 20 mM NH_4OOCH, B = ACN; (F) Expansion of chromatograms for control (shown in blue), pure L-proline (shown in red), pure D-proline (shown in green) and a equimolar mixture of L-proline and D-proline (shown in black).

functionalised with vancomycin, a glycopeptide natural product which contains 18 chiral centres. The ChiroBiotic T separation is based on teicoplanin, a natural product amphoteric glycopeptide which contains 23 chiral centres. Using the ChiroBiotic V and ChiroBiotic T columns, samples containing the L- and D-enantiomers of proline derivatised with AccQ-Tag could be easily separated. When eluted from the ChiroBiotic V column using an isocratic concentration of methanol in ammonium formate buffer (pH 6.3), the D-proline 6-aminoquinolyl-*N*-hydroxysuccinimidyl urea was found to elute first (R_t = 11.55 min, Figure 1.20 A and B, green trace), followed by the L-proline urea (R_t = 13.4 min, Figure 1.20 A and B, red trace). A baseline separation could be obtained for a equimolar mixture of D- and L-proline derivatives (Figure 1.20 A and B, black trace). Similar results were obtained using the ChiroBiotic T column (data not shown). Unfortunately, when the two 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate bis-derivatised enantiomers of 3-amino-4-(methylamino)butanoic acid were eluted under the same conditions, no separation was obtained despite a much longer retention time (R_t = 25.2 min, data not shown). This was confirmed by the observation of a single peak for the racemic mixture (R_t = 25.2 min, data not shown). Varying the elution conditions, for example using a 2–20% or 2–98% gradient of methanol in buffer, or using more acidic (pH 5.0) or basic (pH 7.5) buffers, did not improve the separation, only altering the retention time to some extent.

Bis-acetylation of the amino groups of 3-amino-4-(methylamino)butanoic acid **77** was investigated (Figure 1.21) hoping to achieve a 'trade-off' between the poor retention of the free β-amino acid and the introduction of the bulky 6-aminoquinolyl group that seemed to hinder separation. Bis-acetylation of 3-amino-4-(methylamino)butanoic acid was performed using an excess of acetic anhydride in a mixture of water and tetrahydrofuran. Because the resulting bis-acetylated β-amino acids do not carry any functionality having a good UV absorption, the samples were analysed using the HPLC ChiroBiotic V and ChiroBiotic T columns coupled to a quadrupole mass spectrometer analyser (Figure 1.21 A–C). A racemic mixture of lysine and ornithine α-amino acid samples were subjected to the same acetylation procedure and run under the same conditions as controls. The bis-acetylated lysine and ornithine derivatives were expected to show similar chromatographic behaviour to bis-acetylated 3-amino-4-(methylamino)butanoic acid under identical chromatography conditions. The good separation for the two bis-acetylated enantiomers of lysine (Figure 1.21 A–C, black traces) and ornithine (Figure 1.21 A–C, green traces) demonstrates the ability of the ChiroBiotic V and

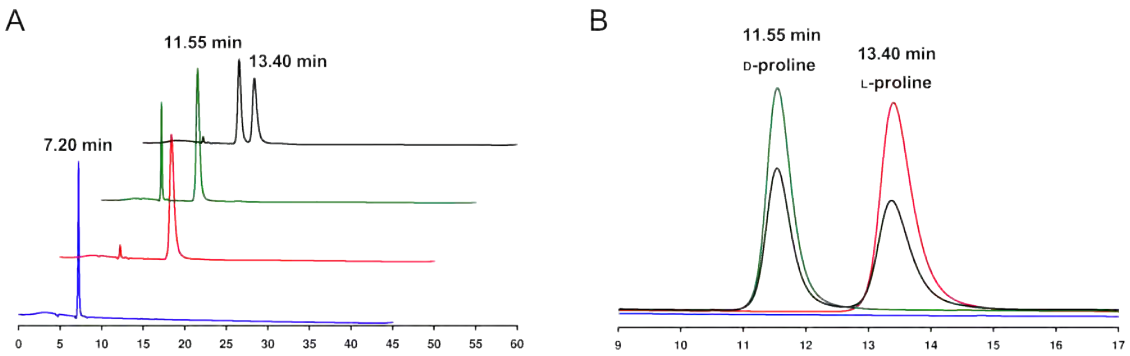


Figure 1.20: HPLC traces for proline enantiomers derivatised with AccQ-Tag using ChiroBiotic V column. The two enantiomers of proline were derivatised with AccQ-Tag and eluted from a vancomycin-functionalised stationary phase ChiroBiotic V column (250 mm x 4 mm, particle size 5 μm) column using 1 mL min^{-1} isocratic 20% MeOH in 20 mM NH_4OOCH (pH 6.3) as a mobile phase. Similar results were obtained using a ChiroBiotic T column. The UV absorbance was monitored at 254 nm. (A) Full HPLC chromatogram for control (shown in blue), derivatised L-proline (shown in red), derivatised D-proline (shown in green) and a equimolar mixture of the two derivatised enantiomers of proline (shown in black); (B) Expansion of HPLC chromatograms for control (shown in blue), derivatised L-proline (R_t = 13.40 min, shown in red), derivatised D-proline (R_t = 11.55 min, shown in green) and a equimolar mixture of the two derivatised enantiomers of proline (shown in black).

ChiroBiotic T columns to discriminate between α-amino acid enantiomers. In comparison, the two bis-acetylated enantiomers of 3-amino-4-(methylamino)butanoic acid showed poor (Figure 1.21 B and C, red and blue traces) or no separation at all (Figure 1.21 A, red and blue traces) depending on the gradient. Replacing the methanol mobile phase by acetonitrile and the formic acid mobile phase by acetic acid, ammonium formate or ammonium acetate at various pH did not improve the separation. The near-baseline separation observed when a shallow gradient was used (Figure 1.21 C, red and blue traces) was promising, but when a concentration of Tris buffer comparable to the assay mixture was added to the sample prior to acetylation, the two enantiomers co-eluted again. This is probably due to a large excess of Tris interfering with the interactions between the stationary phase and the 3-amino-4-(methylamino)butanoic acid analytes.

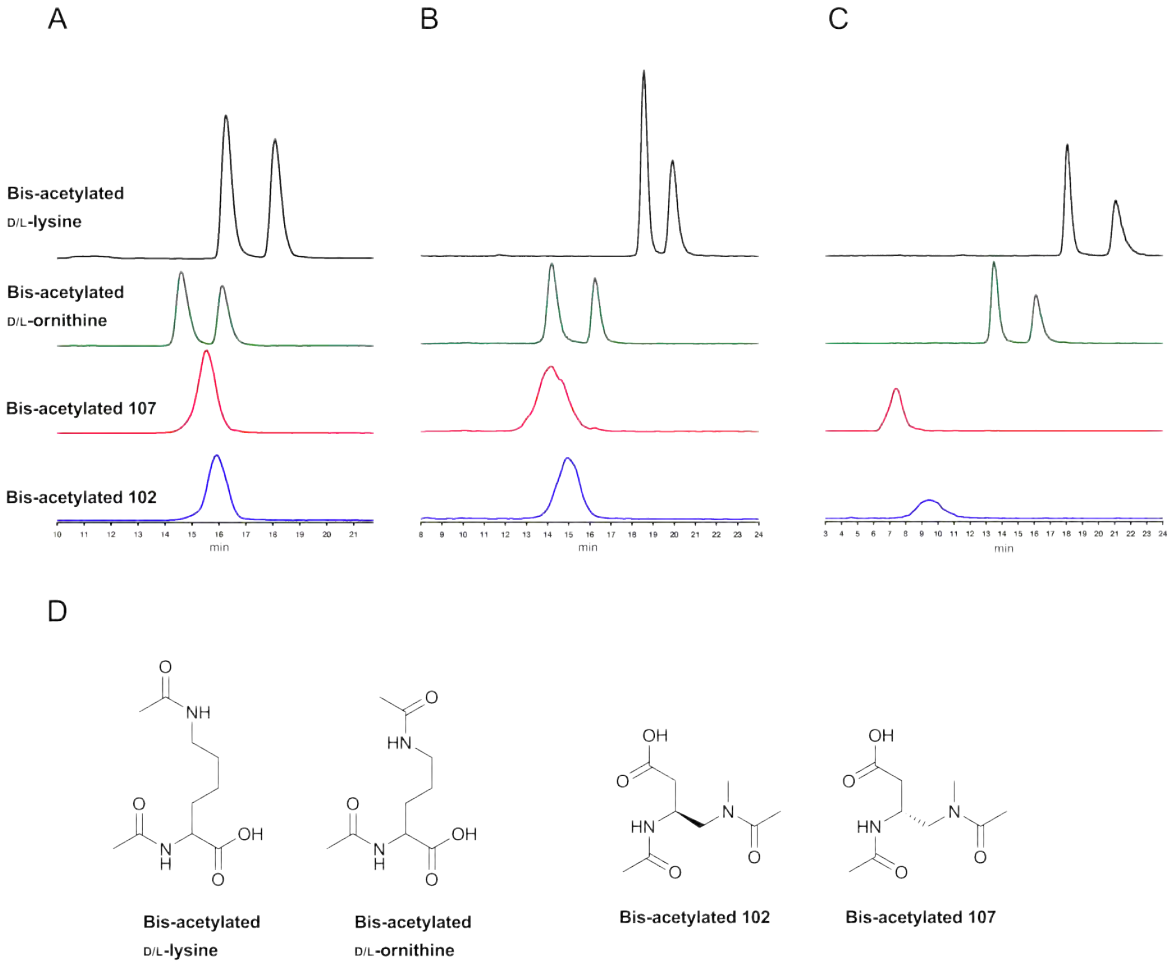


Figure 1.21: Attempted separation of acetylated 3-amino-4-(methylamino)butanoic acid standards. Attempts to separate (3*S*)-3-amino-4-(methylamino)butanoic acid **102** and (3*R*)-3-amino-4-(methylamino)butanoic acid **107** after derivatisation with acetic anhydride and eluted on a vancomycin-functionalised stationary phase ChiroBiotic V column (250 mm x 4 mm, particle size 5 μm) using various gradients of B (MeOH) in A (0.05% HCOOH in water pH 3.5). Extracted ion chromatograms are shown for bis-acetylated lysine (m/z = 231.3, black trace), bis-acetylated ornithine (m/z = 217.2, green trace), bis-acetylated (3*R*)-3-amino-4-(methylamino)butanoic acid (m/z = 217.2, red trace) and (3*S*)-3-amino-4-(methylamino)butanoic acid (m/z = 217.2, blue trace). (A) 2% B in A for 5 min, 2–98% B in A over 40 min; (B) 2% B in A for 5 min, 2–50% B in A over 40 min; (C) 2% B in A for 5 min, 2–20% B in A over 40 min; (D) Structures of the bis-acetylated derivatives of amino acids D/L-lysine, D/L-ornithine, (3*S*)-3-amino-4-(methylamino)butanoic acid **102** and (3*R*)-3-amino-4-(methylamino)butanoic acid **107**.

1.4 Summary for Chapter 1

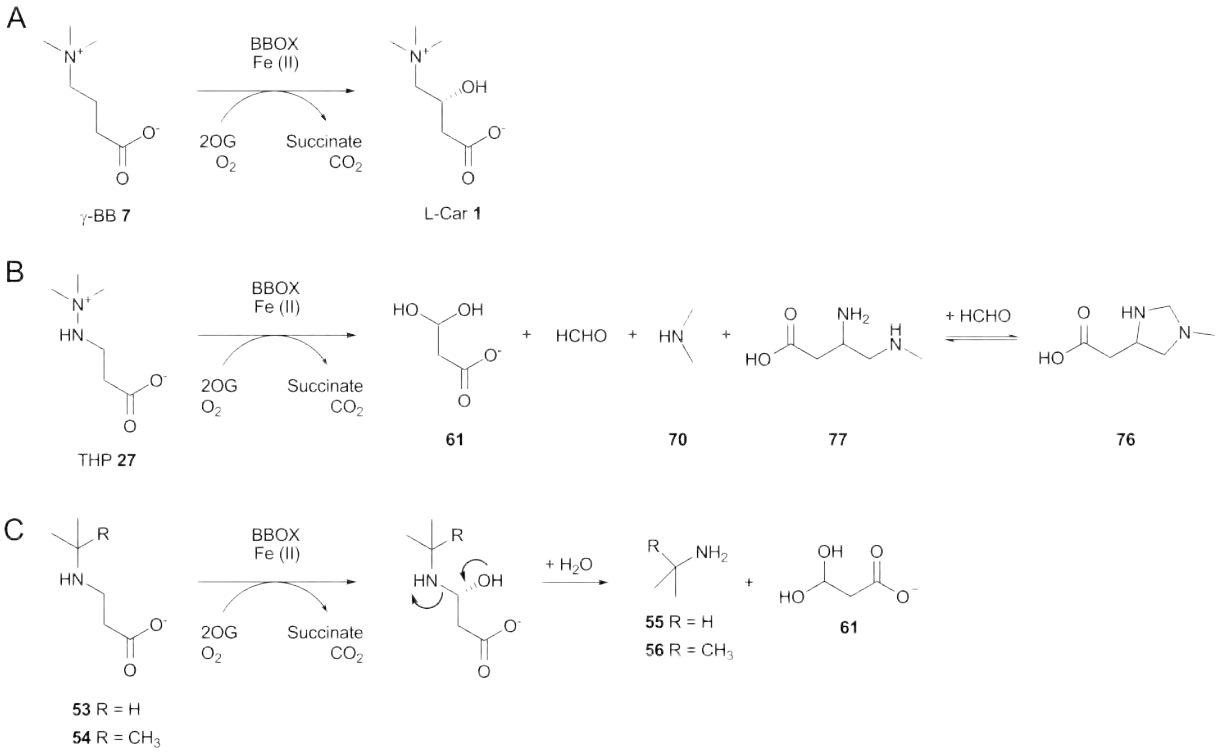
1.4.1 BBOX Activity and Substrate Specificity

γ-Butyrobetaine hydroxylase (BBOX) catalyses the Fe(II) and 2OG-dependent hydroxylation of γ-BB **7** at the C-3 position to give L-Car **1** (Scheme 1.23 A). In Section 1.2.2, the activity of purified recombinant human BBOX was measured *in vitro* and in real time in collaboration with Ivanhoe K. H. Leung using a ^{1}H NMR spectroscopy-based assay. The cofactor requirement of the enzyme was determined and the kinetic parameters measured for the 2OG cofactor and γ-BB substrate **7**.

Under the same assay conditions, the unnatural D-enantiomer of carnitine was found to be a BBOX substrate (Section 1.2.3). When D-Car **20** was subjected to BBOX-catalysed oxidation, 3-dehydrocarnitine **17** (3-carboxy-*N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium) was identified as a transient species that spontaneously decarboxylated to give *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18**. This observation suggests a ketonisation mechanism previously proposed for another Fe(II) and 2OG-dependent enzyme, human prolyl hydroxylase 2 (PHD2) [122], and for the copper-dependent dopamine β-monooxygenase [119]. Since industrial preparations of L-Car **1** were found to contain detectable levels of the unnatural D-enantiomer, the biological relevance of *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** which is produced *via* BBOX-catalysed oxidation of D-Car **20** *in vitro* should be investigated.

In order to test them as potential substrates or inhibitors, various γ-BB analogues were obtained *via* chemical synthesis or from chemical suppliers when commercially available (Section 1.2.1, Figure 1.7). Dimethylamino analogues of γ-BB **7** were obtained in apparent quantitative yields from the corresponding amino acids using formaldehyde under Leuckardt reductive amination conditions. Methylation of the tertiary amine products was performed using methyl iodide as an alkylating agent. The dimethyl and trimethylamino compounds were tested as BBOX substrates or inhibitors in a ^{1}H NMR spectroscopy-based assay either in the absence, or in the presence of the natural substrate γ-BB **7** (Section 1.2.2).

A series of conclusions can be drawn from the SAR studies described in Section 1.2.5 regarding



Scheme 1.23: BBOX was found to catalyse the oxidation of unnatural substrates. (A) The role of BBOX in the biosynthesis of L-Car **1**; (B) The clinically used BBOX inhibitor THP **27** was oxidised to yield several species, including 3-amino-4-(methylamino)butanoic acid **77**; (C) *N*-Isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54** are BBOX substrates.

the substrate specificity of BBOX: 1 - The quaternary ammonium ion is probably important for binding. However, the observation that dimethylaminobutanoic acid **34** and the THP analogues, *N*-isopropyl-β-alanine acid **53** and *N*-*tert*-butyl-β-alanine **54**, are BBOX substrates suggests that the constitutively charged trimethylammonium group is not strictly required. Despite the possible role of the quaternary ammonium positive charge in an electrostatic interaction with an electron-rich aromatic pocket, the binding of the substrate is probably achieved mainly through a set of hydrophobic interactions between the three methyl groups of the trimethylammonium moiety and this hydrophobic aromatic cage [90]. When one methyl group is missing, *e.g.* in compound **34**, the amino group is probably protonated but no longer a good substrate. This supports the importance of all three methyl groups for optimal hydrophobic interactions. 2 - The carboxylate group seems to be crucial for binding since analogues bearing a hydroxyl, *e.g.* choline **2**, 3-hydroxy-*N*,*N*,*N*-trimethylpropan-1-aminium **49** and 4-hydroxy-*N*,*N*,*N*-trimethylbutan-1-aminium **50** (Section 1.2.1, Figure 1.7) or simply an alkyl chain, *e.g.* *N*,*N*,*N*-trimethylethanaminium **46**, *N*,*N*,*N*-trimethylbutan-1-aminium **47** and *N*,*N*,*N*-trimethylhexan-1-aminium **48**, (Section 1.2.1, Figure 1.7) were found to be neither BBOX substrates, nor inhibitors. This observation is supported by the BBOX crystallographic data. The carboxylate group of the substrate γ-BB **7** and the inhibitor THP **27** is positioned to form hydrogen bond interactions with the side chains of Asn-191 (2.9 Å), Asn-292 (2.9 Å) and the backbone amide of Tyr-205 (2.9 Å) [90]. 3 - Shorter and longer substrates than γ-BB **7** are accepted by BBOX when the *N*-trimethylammonium and carboxylate functionalities are retained. 3-(Trimethylammonio)propanoate **40** with one less methylene than γ-BB **7** and 5-(trimethylammonio)pentanoate **41** with one additional methylene group were found to be hydroxylated by BBOX, although to a lesser extent than the natural substrate γ-BB **7** (Scheme 1.10).

1.4.2 The Clinically Used Drug THP Inhibits BBOX Likely Through a Competitive Mechanism

THP **27** is a clinically used BBOX inhibitor for which the mode of action remains unclear. Using a ^{1}H NMR spectroscopy-based assay, we demonstrated for the first time that THP **27** inhibits the BBOX-catalysed hydroxylation of γ-BB **7** likely *via* a competitive mechanism (Section 1.3.2), in contradiction with the non-competitive mode of action proposed by others [75]. In the absence of the natural substrate γ-BB **7**, THP **27** was found to be oxidised by BBOX and several

products were observed (Scheme 1.23 B). Analysis of the reaction mixture using MS and NMR techniques identified dimethylamine **70**, hydrated malonic acid semialdehyde **61**, formaldehyde and 3-amino-4-(methylamino)butanoic acid **77** (Scheme 1.23). These products were unexpected and suggest that the BBOX-catalysed oxidation of THP occurs *via* a mechanism different from the predicted hydroxylation, followed by fragmentation that would yield malonate semialdehyde and *N*,*N*,*N*-trimethylhydrazine (Scheme 1.5) [77]. When *N*-isopropyl-β-alanine **53** and *N*-*tert*-butyl-β-alanine **54** were subjected to BBOX-catalysed oxidation, the products of hydroxylation followed by fragmentation were observed (Scheme 1.23 C).

1.4.3 3-Amino-4-(methylamino)butanoic Acid is a Rearrangement Product of the BBOX-Catalysed oxidation of THP

A plausible mechanism for the BBOX-catalysed oxidation of THP **27** was proposed, based on the products observed in the assay mixture using NMR and MS analyses (Scheme 1.16). The formation of the unexpected product 3-amino-4-(methylamino)butanoic acid **77** was confirmed using two approaches; the chemical synthesis of [^{13}C]-labeled THP **67** with a label at one of *N*-methyl group was performed. [^{13}C]-Labeled THP **67** was incubated with BBOX and the fates of the label were identified. ^{13}C-^{1}H NMR correlation spectroscopy and ^{13}C chemical shift comparison with authentic standards showed that, besides dimethylamine and formaldehyde, the label is retained as a *N*-methyl group as well as inserted into a new C–C bond in 3-amino-4-(methylamino)butanoic acid **77**. The MS analyses confirmed the predicted isotopic distribution based on the proposed mechanism; the label has 2/3 chance to be retained in the 3-amino-4-(methylamino)butanoic acid **77** and dimethylamine products, and 1/3 chance to be found in the formaldehyde produced by demethylation (Scheme 1.15). These results confirmed the proposed mechanism for the BBOX-catalysed oxidation of THP (Scheme 1.16). Comparison of the relative abundance of the products suggests that after proton abstraction by the Fe(IV)-oxo species, the reaction follows preferentially the path leading to the formation of dimethylamine **70** (Scheme 1.16, Path A), 3-amino-4-(methylamino)butanoic acid **77** is formed in low amounts *via* a second, less favoured reaction (Scheme 1.16, Path B).

The syntheses of racemic 3-amino-4-(methylamino)butanoic acid **77** as well as enantioenriched (3*S*)-isomer **102** and (3*R*)-isomer **107** were performed successfully from the amino acid chiral pool

(Section 1.3.4). 3-Amino-4-(methylamino)butanoic acid **77** and the two enantiomers **102** and **107** were obtained in six steps from L-aspartic acid 5-*tert*-butyl ester (Scheme 1.21). Having access to pure standard enabled the confirmation of the structure of the rearrangement product **77**. In addition, formaldehyde titration experiments, including [^{13}C]-formaldehyde (data not shown) [150], demonstrated that under the assay conditions, compound **77** forms the cyclic formaldehyde adduct **76** in which the carbon atom originating from formaldehyde is bridging the two amino groups (Scheme 1.14 D). Despite repeated efforts, no robust analytical methodology could be developed to separate the two 3-amino-4-(methylamino)butanoic acid enantiomers (Section 1.3.6). A chiral HPLC method allowing separation of the (3*S*)-isomer **102** and (3*R*)-isomer **107** could have led to investigations on the stereospecificity of the C–C bond forming reaction. This could be the subject of future investigations.

References

[1] Bremer, J. (1983) Carnitine–Metabolism and functions. *Physiol Rev*, **63**, 1420–1480.

[2] Fraenkel, G. and Friedman, S. (1957) Carnitine. *Vitam Horm*, **15**, 73–118.

[3] Reisch, C. R., Stoudemayer, M. J., Varaljay, V. A., Amster, I. J., Moran, M. A., and Whitman, W. B. (2011) Novel pathway for assimilation of dimethylsulphoniopropionate widespread in marine bacteria. *Nature*, **473**, 208–211.

[4] Anthoni, U., Christophersen, C., Hougaard, L., and Nielsen, P. (1991) Quaternary ammonium compounds in the biosphere - An example of a versatile adaptive strategy. *Comparative Biochemistry and Physiology B - Comparative Biochemistry*, **99**, 1–18.

[5] Eichler, K., Bourgis, F., Buchet, A., Kleber, H. P., and Mandrand-Berthelot, M. A. (1994) Molecular characterization of the *cai* operon necessary for carnitine metabolism in *Escherichia coli*. *Mol Microbiol*, **13**, 775–786.

[6] Zammit, V. A., Ramsay, R. R., Bonomini, M., and Arduini, A. (2009) Carnitine, mitochondrial function and therapy. *Adv Drug Deliv Rev*, **61**, 1353–1362.

[7] Houten, S. M. and Wanders, R. J. A. (2010) A general introduction to the biochemistry of mitochondrial fatty acid β-oxidation. *J Inherit Metab Dis*, **33**, 469–477.

[8] Lahjouji, K., Mitchell, G. A., and Qureshi, I. A. (2001) Carnitine transport by organic cation transporters and systemic carnitine deficiency. *Mol Genet Metab*, **73**, 287–297.

[9] Kerner, J. and Hoppel, C. (2000) Fatty acid import into mitochondria. *Biochim Biophys Acta*, **1486**, 1–17.

[10] Peluso, G., Barbarisi, A., Savica, V., Reda, E., Nicolai, R., Benatti, P., and Calvani, M. (2000) Carnitine: An osmolyte that plays a metabolic role. *J Cell Biochem*, **80**, 1–10.

[11] Vaz, F. M. and Wanders, R. J. A. (2002) Carnitine biosynthesis in mammals. *Biochem J*, **361**, 417–429.

[12] Steiber, A., Kerner, J., and Hoppel, C. L. (2004) Carnitine: A nutritional, biosynthetic, and functional perspective. *Mol Aspects Med*, **25**, 455–473.

[13] Rebouche, C. J. (1992) Carnitine function and requirements during the life cycle. *FASEB J*, **6**, 3379–3386.

[14] Foster, D. W. (2004) The role of the carnitine system in human metabolism. *Ann N Y Acad Sci*, **1033**, 1–16.

[15] Mingrone, G. (2004) Carnitine in type 2 diabetes. *Ann N Y Acad Sci*, **1033**, 99–107.

[16] Strijbis, K., Vaz, F. M., and Distel, B. (2010) Enzymology of the carnitine biosynthesis pathway. *IUBMB Life*, **62**, 357–362.

[17] Horne, D. W., Tanphaichitr, V., and Broquist, H. P. (1971) Role of lysine in carnitine biosynthesis in *Neurospora crassa*. *J Biol Chem*, **246**, 4373–4375.

[18] Horne, D. W. and Broquist, H. P. (1973) Role of lysine and ϵ-*N*-trimethyllysine in carnitine biosynthesis. I. Studies in *Neurospora crassa*. *J Biol Chem*, **248**, 2170–2175.

[19] Sachan, D. S. and Broquist, H. P. (1980) Synthesis of carnitine from ϵ-*N*-trimethyllysine in post mitochondrial fractions of *Neurospora crassa*. *Biochem Biophys Res Commun*, **96**, 870–875.

[20] Hoppel, C. L., Cox, R. A., and Novak, R. F. (1980) N^6-Trimethyllysine metabolism. 3-Hydroxy-N^6-trimethyllysine and carnitine biosynthesis. *Biochem J*, **188**, 509–519.

[21] Tanphaichitr, V. and Broquist, H. P. (1973) Role of lysine and ϵ-*N*-trimethyllysine in carnitine biosynthesis. II. Studies in the rat. *J Biol Chem*, **248**, 2176–2181.

[22] Paik, W. K., Nochumson, S., and Kim, S. (1977) Carnitine biosynthesis via protein methylation. *Trends Biochem Sci*, **2**, 159–161.

[23] Dunn, W. A. and Englard, S. (1981) Carnitine biosynthesis by the perfused rat liver from exogenous protein-bound trimethyllysine. Metabolism of methylated lysine derivatives arising from the degradation of 6-*N*-[methyl-^{3}H]lysine-labeled glycoproteins. *J Biol Chem*, **256**, 12437–12444.

[24] Rebouche, C. J. and Broquist, H. P. (1976) Carnitine biosynthesis in *Neurospora crassa*: Enzymatic conversion of lysine to ϵ-*N*-trimethyllysine. *J Bacteriol*, **126**, 1207–1214.

[25] Borum, P. R. and Broquist, H. P. (1977) Purification of *S*-adenosylmethionine: ϵ-*N*-L-lysine methyltransferase. The first enzyme in carnitine biosynthesis. *J Biol Chem*, **252**, 5651–5655.

[26] Khan-Siddiqui, L. and Bamji, M. S. (1983) Lysine-carnitine conversion in normal and undernourished adult men – Suggestion of a nonpeptidyl pathway. *Am J Clin Nutr*, **37**, 93–98.

[27] Vijayasarathy, C., Khan-Siddiqui, L., Murthy, S. N., and Bamji, M. S. (1987) Rise in plasma trimethyllysine levels in humans after oral lysine load. *Am J Clin Nutr*, **46**, 772–777.

[28] Strijbis, K., van Roermund, C. W. T., Hardy, G. P., van den Burg, J., Bloem, K., de Haan, J., van Vlies, N., Wanders, R. J. A., Vaz, F. M., and Distel, B. (2009) Identification and characterization of a complete carnitine biosynthesis pathway in *Candida albicans*. *FASEB J*, **23**, 2349–2359.

[29] Hausinger, R. P. (2004) Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol Biol*, **39**, 21–68.

[30] Novak, R. F., Swift, T. J., and Hoppel, C. L. (1980) N^6-Trimethyl-lysine metabolism. Structural identification of the metabolite 3-hydroxy-N^6-trimethyl-lysine. *Biochem J*, **188**, 521–527.

[31] Hulse, J. D., Ellis, S. R., and Henderson, L. M. (1978) Carnitine biosynthesis. β-Hydroxylation of trimethyllysine by an α-ketoglutarate-dependent mitochondrial dioxygenase. *J Biol Chem*, **253**, 1654–1659.

[32] Swiegers, J. H., Vaz, F. M., Pretorius, I. S., Wanders, R. J. A., and Bauer, F. F. (2002) Carnitine biosynthesis in *Neurospora crassa*: Identification of a cDNA coding for ϵ-*N*-trimethyllysine hydroxylase and its functional expression in *Saccharomyces cerevisiae*. *FEMS Microbiol Lett*, **210**, 19–23.

[33] Vaz, F. M., Ofman, R., Westinga, K., Back, J. W., and Wanders, R. J. (2001) Molecular and biochemical characterization of rat ϵ-*N*-trimethyllysine hydroxylase, the first enzyme of carnitine biosynthesis. *J Biol Chem*, **276**, 33512–33517.

[34] Monfregola, J., Cevenini, A., Terracciano, A., van Vlies, N., Arbucci, S., Wanders, R. J. A., D'Urso, M., Vaz, F. M., and Ursini, M. V. (2005) Functional analysis of TMLH variants and definition of domains required for catalytic activity and mitochondrial targeting. *J Cell Physiol*, **204**, 839–847.

[35] van Vlies, N., Ofman, R., Wanders, R. J. A., and Vaz, F. M. (2007) Submitochondrial localization of 6-*N*-trimethyllysine dioxygenase – Implications for carnitine biosynthesis. *FEBS J*, **274**, 5845–5851.

[36] Omura, T. (1998) Mitochondria-targeting sequence, a multi-role sorting sequence recognized at all steps of protein import into mitochondria. *J Biochem*, **123**, 1010–1016.

[37] Larkin, M. A., et al. (2007) Clustal W and clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

[38] Hochalter, J. B. and Henderson, L. M. (1976) Carnitine biosynthesis: the formation of glycine from carbons 1 and 2 of 6-*N*-trimethyl-L-lysine. *Biochem Biophys Res Commun*, **70**, 364–368.

[39] Dunn, W. A., Aronson, N. N., and Englard, S. (1982) The effects of 1-amino-D-proline on the production of carnitine from exogenous protein-bound trimethyllysine by the perfused rat liver. *J Biol Chem*, **257**, 7948–7951.

[40] Vaz, F. M., Fouchier, S. W., Ofman, R., Sommer, M., and Wanders, R. J. (2000) Molecular and biochemical characterization of rat γ-trimethylaminobutyraldehyde dehydrogenase and evidence for the involvement of human aldehyde dehydrogenase 9 in carnitine biosynthesis. *J Biol Chem*, **275**, 7390–7394.

[41] Ambroziak, W. and Pietruszko, R. (1991) Human aldehyde dehydrogenase. Activity with aldehyde metabolites of monoamines, diamines, and polyamines. *J Biol Chem*, **266**, 13011–13018.

[42] Hulse, J. D. and Henderson, L. M. (1980) Carnitine biosynthesis. Purification of 4-*N*'-trimethylaminobutyraldehyde dehydrogenase from beef liver. *J Biol Chem*, **255**, 1146–1151.

[43] Koch, A., Knig, B., Stangl, G. I., and Eder, K. (2008) PPARα mediates transcriptional upregulation of novel organic cation transporters-2 and -3 and enzymes involved in hepatic carnitine synthesis. *Exp Biol Med (Maywood)*, **233**, 356–365.

[44] Schoonjans, K., Staels, B., and Auwerx, J. (1996) Role of the peroxisome proliferator-activated receptor (PPAR) in mediating the effects of fibrates and fatty acids on gene expression. *J Lipid Res*, **37**, 907–925.

[45] Rebouche, C. J. and Seim, H. (1998) Carnitine metabolism and its regulation in microorganisms and mammals. *Annu Rev Nutr*, **18**, 39–61.

[46] Evans, A. M. and Fornasini, G. (2003) Pharmacokinetics of L-carnitine. *Clin Pharmacokinet*, **42**, 941–967.

[47] Lohninger, A., Pittner, G., and Pittner, F. (2005) L-Carnitine: New aspects of a known compound – A brief survey. *Monatshefte fur Chemie / Chemical Monthly*, **136**, 1255–1268.

[48] Rangarajan, E. S., Li, Y., Iannuzzi, P., Cygler, M., and Matte, A. (2005) Crystal structure of *Escherichia coli* crotonobetainyl-CoA:Carnitine CoA-transferase (CaiB) and its complexes with CoA and carnitinyl-CoA. *Biochemistry*, **44**, 5728–5738.

[49] Eichler, K., Schunck, W. H., Kleber, H. P., and Mandrand-Berthelot, M. A. (1994) Cloning, nucleotide sequence, and expression of the *Escherichia coli* gene encoding carnitine dehydratase. *J Bacteriol*, **176**, 2970–2975.

[50] Jung, H., Buchholz, M., Clausen, J., Nietschke, M., Revermann, A., Schmid, R., and Jung, K. (2002) CaiT of *Escherichia coli*, a new transporter catalyzing L-carnitine/γ-butyrobetaine exchange. *J Biol Chem*, **277**, 39251–39258.

[51] Schulze, S., Kster, S., Geldmacher, U., van Scheltinga, A. C. T., and Khlbrandt, W. (2010) Structural basis of Na^+-independent and cooperative substrate/product antiport in CaiT. *Nature*, **467**, 233–236.

[52] Tang, L., Bai, L., hua Wang, W., and Jiang, T. (2010) Crystal structure of the carnitine transporter and insights into the antiport mechanism. *Nat Struct Mol Biol*, **17**, 492–496.

[53] Preusser, A., Wagner, U., Elssner, T., and Kleber, H. P. (1999) Crotonobetaine reductase from *Escherichia coli* consists of two proteins. *Biochim Biophys Acta*, **1431**, 166–178.

[54] Bernal, V., Arense, P., Blatz, V., Mandrand-Berthelot, M. A., Cánovas, M., and Iborra, J. L. (2008) Role of betaine:CoA ligase (CaiC) in the activation of betaines and the transfer of coenzyme A in *Escherichia coli*. *J Appl Microbiol*, **105**, 42–50.

[55] Elssner, T., Engemann, C., Baumgart, K., and Kleber, H. P. (2001) Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA hydratase and a CoA-transferase, in the hydration of crotonobetaine to L-carnitine by *Escherichia coli*. *Biochemistry*, **40**, 11140–11148.

[56] Eichler, K., Buchet, A., Lemke, R., Kleber, H. P., and Mandrand-Berthelot, M. A. (1996) Identification and characterization of the *caiF* gene encoding a potential transcriptional activator of carnitine metabolism in *Escherichia coli*. *J Bacteriol*, **178**, 1248–1257.

[57] Oren, A. (1990) Formation and breakdown of glycine betaine and trimethylamine in hypersaline environments. *Antonie Van Leeuwenhoek*, **58**, 291–298.

[58] Karlic, H. and Lohninger, A. (2004) Supplementation of L-carnitine in athletes: Does it make sense? *Nutrition*, **20**, 709–715.

[59] Brass, E. P. (2004) Carnitine and sports medicine: Use or abuse? *Ann N Y Acad Sci*, **1033**, 67–78.

[60] Mate, A., Miguel-Carrasco, J. L., and Vzquez, C. M. (2010) The therapeutic prospects of using L-carnitine to manage hypertension-related organ damage. *Drug Discov Today*, **15**, 484–492.

[61] Graziano, F., et al. (2002) Potential role of levocarnitine supplementation for the treatment of chemotherapy-induced fatigue in non-anaemic cancer patients. *Br J Cancer*, **86**, 1854–1857.

[62] Block, K. I., Koch, A. C., Mead, M. N., Tothy, P. K., Newman, R. A., and Gyllenhaal, C. (2008) Impact of antioxidant supplementation on chemotherapeutic toxicity: A systematic review of the evidence from randomized controlled trials. *Int J Cancer*, **123**, 1227–1239.

[63] Dawany, N. B., Dampier, W. N., and Tozeren, A. (2011) Large-scale integration of microarray data reveals genes and pathways common to multiple cancer types. *Int J Cancer*, **128**, 2881–2891.

[64] Comber, R. N. and Brouillette, W. J. (1985) An efficient synthesis of D,L-carnitine hydrochloride *via* the acylation of an α-aminoketone. *Organic Preparations and Procedures International*, **17**, 175–181.

[65] Dropsy, E. P. and Klibanov, A. M. (1984) Cholinesterase-catalyzed resolution of D,L-carnitine. *Biotechnol Bioeng*, **26**, 911–915.

[66] Meyer, H.-P. and Robins, K. T. (2005) Large scale bioprocess for the production of optically pure L-carnitine. *Monatshefte fur Chemie / Chemical Monthly*, **136**, 1269–1277.

[67] Naidu, G. S. N., Lee, I. Y., Lee, E. G., Kang, G. H., and Park, Y. H. (2000) Microbial and enzymatic production of L-carnitine. *Bioprocess Biosyst Eng*, **23**, 627–635.

[68] Bernal, V., Sevilla, A., Cnovas, M., and Iborra, J. L. (2007) Production of L-carnitine by secondary metabolism of bacteria. *Microb Cell Fact*, **6**, 31.

[69] Bruijnincx, P. C. A., van Koten, G., and Gebbink, R. J. M. K. (2008) Mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad: Recent developments in enzymology and modeling studies. *Chem Soc Rev*, **37**, 2716–2744.

[70] Krebs, C., Fujimori, D. G., Walsh, C. T., and Bollinger, J. M. (2007) Non-heme Fe(IV)-oxo intermediates. *Acc Chem Res*, **40**, 484–492.

[71] Petter, R. C., Banerjee, S., and Englard, S. (1990) Inhibition of γ-butyrobetaine hydroxylase by cyclopropyl-substituted γ-butyrobetaines. *J Org Chem*, **55**, 3088–3097.

[72] Ziering, D. L. and Pascal, R. A. (1990) Mechanism-based inhibition of bacterial γ-butyrobetaine hydroxylase. *J Am Chem Soc*, **112**, 834–838.

[73] Suckling, C. J. (1988) The cyclopropyl group in studies of enzyme mechanism and inhibition. *Angew Chem Int Ed Engl*, **27**, 537–552.

[74] Iwata-Reuyl, D., Basak, A., Silverman, L. S., Engle, C. A., and Townsend, C. A. (1993) Synthesis and reaction of potential alternate substrates and mechanism-based inhibitors of clavaminate synthase. *J Nat Prod*, **56**, 1373–1396.

[75] Simkhovich, B. Z., Shutenko, Z. V., Meirena, D. V., Khagi, K. B., Mezapuke, R. J., Molodchina, T. N., Kalvins, I. J., and Lukevics, E. (1988) 3-(2,2,2-Trimethylhydrazinium)propionate (THP) – A novel γ-butyrobetaine hydroxylase inhibitor with cardioprotective properties. *Biochem Pharmacol*, **37**, 195–202.

[76] Galland, S., Borgne, F. L., Guyonnet, D., Clouet, P., and Demarquoy, J. (1998) Purification and characterization of the rat liver γ-butyrobetaine hydroxylase. *Mol Cell Biochem*, **178**, 163–168.

[77] Yoshisue, K., Yamamoto, Y., Yoshida, K., Saeki, M., Minami, Y., Esumi, Y., and Kawaguchi, Y. (2000) Pharmacokinetics and biological fate of 3-(2,2,2-trimethylhydrazinium)propionate dihydrate (MET-88), a novel cardioprotective agent, in rats. *Drug Metab Dispos*, **28**, 687–694.

[78] Spaniol, M., Brooks, H., Auer, L., Zimmermann, A., Solioz, M., Stieger, B., and Krahenbuhl, S. (2001) Development and characterization of an animal model of carnitine deficiency. *Eur J Biochem*, **268**, 1876–1887.

[79] Jaudzems, K., Kuka, J., Gutsaits, A., Zinovjevs, K., Kalvinsh, I., Liepinsh, E., Liepinsh, E., and Dambrova, M. (2009) Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism. *J Enzyme Inhib Med Chem*, **24**, 1269–1275.

[80] Davis, A. T., Monroe, T. J., Ingersoll, K. E., VanWyk, L., and Wojtas, M. (2007) Dietary mildronate supplementation has no effect on carnitine biosynthetic enzyme mRNA expression in rat. *Nutrition Research*, **27**, 225–229.

[81] Liepinsh, E., Vilskersts, R., Skapare, E., Svalbe, B., Kuka, J., Cirule, H., Pugovics, O., Kalvinsh, I., and Dambrova, M. (2008) Mildronate decreases carnitine availability and up-regulates glucose uptake and related gene expression in the mouse heart. *Life Sci*, **83**, 613–619.

[82] Degrace, P., Demizieux, L., Gresti, J., Tsoko, M., André, A., Demaison, L., and Clouet, P. (2004) Fatty acid oxidation and related gene expression in heart depleted of carnitine by mildronate treatment in the rat. *Mol Cell Biochem*, **258**, 171–182.

[83] Sokolovska, J., Isajevs, S., Sugoka, O., Sharipova, J., Lauberte, L., Svirina, D., Rostoka, E., Sjakste, T., Kalvinsh, I., and Sjakste, N. (2011) Correction of glycaemia and GLUT1 level by mildronate in rat streptozotocin diabetes mellitus model. *Cell Biochem Funct*, **29**, 55–63.

[84] Shutenko, Z., Meirena, D., Kagan, T., Syakste, N., and Kalvinsh, I. (1995) Mildronate: Mechanisms of action and prospects for correction of pathologic states. *Pharmaceutical Chemistry Journal*, **29**, 309–312.

[85] Tars, K., et al. (2010) Crystal structure of human γ-butyrobetaine hydroxylase. *Biochem Biophys Res Commun*, **398**, 634–639.

[86] Kost, T. A., Condreay, J. P., and Jarvis, D. L. (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol*, **23**, 567–575.

[87] Jarvis, D. L. (2009) Baculovirus-insect cell expression systems. *Methods Enzymol*, **463**, 191–222.

[88] Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J Biol Chem*, **272**, 23031–23036.

[89] Cunliffe, C. J., Franklin, T. J., Hales, N. J., and Hill, G. B. (1992) Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analogue *N*-oxaloglycine and its derivatives. *J Med Chem*, **35**, 2652–2658.

[90] Leung, I. K. H., Krojer, T. J., Kochan, G. T., Henry, L., von Delft, F., Claridge, T. D. W., Oppermann, U., McDonough, M. A., and Schofield, C. J. (2010) Structural and mechanistic studies on γ-butyrobetaine hydroxylase. *Chem Biol*, **17**, 1316–1324.

[91] Schrödinger, LLC (2010) The PyMOL molecular graphics system, version 1.3r1.

[92] Lindstedt, S. and Nordin, I. (1984) Multiple forms of γ-butyrobetaine hydroxylase (EC 1.14.11.1). *Biochem J*, **223**, 119–127.

[93] McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J., and Schofield, C. J. (2010) Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr Opin Struct Biol*, **20**, 659–672.

[94] Horton, J. R., Upadhyay, A. K., Qi, H. H., Zhang, X., Shi, Y., and Cheng, X. (2010) Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat Struct Mol Biol*, **17**, 38–43.

[95] Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R.-M. (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science*, **312**, 748–751.

[96] Henry, L., Leung, I. K., Claridge, T. D., and Schofield, C. J. (2012) γ-Butyrobetaine hydroxylase catalyses a stevens type rearrangement. *Bioorg Med Chem Lett*, pp. –.

[97] Blanchard, J. S. and Englard, S. (1983) γ-Butyrobetaine hydroxylase: Primary and secondary tritium kinetic isotope effects. *Biochemistry*, **22**, 5922–5929.

[98] Englard, S., Blanchard, J. S., and Midelfort, C. F. (1985) γ-Butyrobetaine hydroxylase: Stereochemical course of the hydroxylation reaction. *Biochemistry*, **24**, 1110–1116.

[99] Lindstedt, G. (1967) Hydroxylation of γ-butyrobetaine to carnitine in rat liver. *Biochemistry*, **6**, 1271–1282.

[100] Wehbie, R. S., Punekar, N. S., and Lardy, H. A. (1988) Rat liver γ-butyrobetaine hydroxylase catalyzed reaction: Influence of potassium, substrates, and substrate analogues on hydroxylation and decarboxylation. *Biochemistry*, **27**, 2222–2228.

[101] Lindstedt, G., Lindstedt, S., and Tofft, M. (1970) γ-Butyrobetaine hydroxylase from *Pseudomonas* sp AK1. *Biochemistry*, **9**, 4336–4342.

[102] Dunn, W. A., Rettura, G., Seifter, E., and Englard, S. (1984) Carnitine biosynthesis from γ-butyrobetaine and from exogenous protein-bound 6-*N*-trimethyl-L-lysine by the perfused guinea pig liver. Effect of ascorbate deficiency on the in situ activity of γ-butyrobetaine hydroxylase. *J Biol Chem*, **259**, 10764–10770.

[103] Rebouche, C. J. (1991) Ascorbic acid and carnitine biosynthesis. *Am J Clin Nutr*, **54**, 1147S–1152S.

[104] Ng, S. F., Hanauske-Abel, H. M., and Englard, S. (1991) Cosubstrate binding site of *Pseudomonas* sp. AK1 γ-butyrobetaine hydroxylase. Interactions with structural analogs of α-ketoglutarate. *J Biol Chem*, **266**, 1526–1533.

[105] Rose, N. R., McDonough, M. A., King, O. N. F., Kawamura, A., and Schofield, C. J. (2011) Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev*, **40**, 4364–4397.

[106] Carter, A. L. and Stratman, F. W. (1980) A rapid and sensitive assay of γ-butyrobetaine hydroxylase. *FEBS Lett*, **111**, 112–114.

[107] Englard, S., Horwitz, L. J., and Mills, J. T. (1978) A simplified method for the measurement of γ-butyrobetaine hydroxylase activity. *J Lipid Res*, **19**, 1057–1063.

[108] McGarry, J. D. and Foster, D. W. (1976) An improved and simplified radioisotopic assay for the determination of free and esterified carnitine. *J Lipid Res*, **17**, 277–281.

[109] Holme, E., Lindstedt, S., and Nordin, I. (1984) Uncoupling and isotope effects in γ-butyrobetaine hydroxylation. *Biosci Rep*, **4**, 433–440.

[110] Furusawa, H., et al. (2008) Vitamin C is not essential for carnitine biosynthesis in vivo: Verification in vitamin C-depleted senescence marker protein-30/gluconolactonase knockout mice. *Biol Pharm Bull*, **31**, 1673–1679.

[111] Nelson, P. J., Pruitt, R. E., Henderson, L. L., Jenness, R., and Henderson, L. M. (1981) Effect of ascorbic acid deficiency on the *in vivo* synthesis of carnitine. *Biochim Biophys Acta*, **672**, 123–127.

[112] Thoma, W. J. and Henderson, L. M. (1984) Effect of vitamin C deficiency on hydroxylation of trimethylaminobutyrate to carnitine in the guinea pig. *Biochim Biophys Acta*, **797**, 136–139.

[113] Punekar, N. S., Wehbie, R. S., and Lardy, H. A. (1987) γ-Butyrobetaine hydroxylase and the protective role of glutathione peroxidase. *J Biol Chem*, **262**, 6720–6724.

[114] Lindstedt, G., Lindstedt, S., and Nordin, I. (1977) Purification and properties of γ-butyrobetaine hydroxylase from *Pseudomonas* sp AK1. *Biochemistry*, **16**, 2181–2188.

[115] Sánchez-Hernández, L., Castro-Puyana, M., García-Ruiz, C., Crego, A. L., and Marina, M. L. (2010) Determination of L- and D-carnitine in dietary food supplements using capillary electrophoresis-tandem mass spectrometry. *Food Chemistry*, **120**, 921–928.

[116] Paulson, D. J. and Shug, A. L. (1981) Tissue specific depletion of L-carnitine in rat heart and skeletal muscle by D-carnitine. *Life Sci*, **28**, 2931–2938.

[117] Tsuchida, H., et al. (2010) Identification of a novel organic anion transporter mediating carnitine transport in mouse liver and kidney. *Cell Physiol Biochem*, **25**, 511–522.

[118] Holme, E., Linstedt, S., and Nordin, I. (1982) Uncoupling in the γ-butyrobetaine hydroxylase reaction by D- and L-carnitine. *Biochem Biophys Res Commun*, **107**, 518–524.

[119] May, S. W., Phillips, R. S., Mueller, P. W., and Herman, H. H. (1981) Dopamine β-hydroxylase. Demonstration of enzymatic ketonization of the product enantiomer *S*-octopamine. *J Biol Chem*, **256**, 2258–2261.

[120] Biscoe, M. R. and Breslow, R. (2003) Hydrophobically directed selective reduction of ketones. *J Am Chem Soc*, **125**, 12718–12719.

[121] Schofield, C. J. and Ratcliffe, P. J. (2004) Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol*, **5**, 343–354.

[122] Loenarz, C., Mecinović, J., Chowdhury, R., McNeill, L. A., Flashman, E., and Schofield, C. J. (2009) Evidence for a stereoelectronic effect in human oxygen sensing. *Angew Chem Int Ed Engl*, **48**, 1784–1787.

[123] Biscoe, M. R., Uyeda, C., and Breslow, R. (2004) Requirements for selective hydrophobic acceleration in the reduction of ketones. *Org Lett*, **6**, 4331–4334.

[124] Moore, M. L. (2004) *The Leuckart Reaction*. John Wiley & Sons, Inc.

[125] Rahal, S. and Badache, L. (1991) An anomalous Eschweiler-Clarke reaction. *Tetrahedron Lett*, **32**, 3847–3848.

[126] Cohen, S. A. and Michaud, D. P. (1993) Synthesis of a fluorescent derivatizing reagent, 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate, and its application for the analysis of hydrolysate amino acids via high-performance liquid chromatography. *Anal Biochem*, **211**, 279–287.

[127] Stevens, T. S., Creighton, E. M., Gordon, A. B., and MacNicol, M. (1928) Degradation of quaternary ammonium salts. Part I. *J Chem Soc*, **1928**, 3193–3197.

[128] Thomson, T. and Stevens, T. S. (1932) Degradation of quaternary ammonium salts. Part VII. New cases of radical migration. *J Chem Soc*, **1932**, 1932–1940.

[129] Vanecko, J. A., Wan, H., and West, F. G. (2006) Recent advances in the Stevens rearrangement of ammonium ylides. Application to the synthesis of alkaloid natural products. *Tetrahedron*, **62**, 1043–1062.

[130] Bowers, A. A., Greshock, T. J., West, N., Estiu, G., Schreiber, S. L., Wiest, O., Williams, R. M., and Bradner, J. E. (2009) Synthesis and conformation-activity relationships of the peptide isosteres of FK228 and largazole. *J Am Chem Soc*, **131**, 2900–2905.

[131] Ewing, W. R., et al. (1999) Design and structure-activity relationships of potent and selective inhibitors of blood coagulation factor Xa. *J Med Chem*, **42**, 3557–3571.

[132] Bergeron, R. J., Wiegand, J., Weimar, W. R., Vinson, J. R., Bussenius, J., Yao, G. W., and McManis, J. S. (1999) Desazadesmethyldesferrithiocin analogues as orally effective iron chelators. *J Med Chem*, **42**, 95–108.

[133] Jefford, C. W. and McNulty, J. (1994) A practical synthesis of (2*S*,3*R*)-3-amino-2-methylpentanoic acid from L-aspartic acid. *Helv Chim Acta*, **77**, 2142–2146.

[134] Jefford, C. W., McNulty, J., Lu, Z.-H., and Wang, J. B. (1996) The enantioselective synthesis of β-amino acids, their α-hydroxy derivatives, and the *N*-terminal components of bestatin and microginin. *Helv Chim Acta*, **79**, 1203–1216.

[135] Cristau, H. J., Garcia, C., Kadoura, J., and Torreilles, E. (1990) *Phosphorous, Sulfur Silicon Relat Elem*, **49/50**, 151.

[136] Ayesa, S., Samuelsson, B., and Classon, B. (2008) A one-pot, solid-phase synthesis of secondary amines from reactive alkyl halides and an alkyl azide. *Synlett*, **2008**, 97–99.

[137] Gryko, D., Chalko, J., and Jurczak, J. (2003) Synthesis and reactivity of *N*-protected-α-amino aldehydes. *Chirality*, **15**, 514–541.

[138] Douat, C., Heitz, A., Martinez, J., and Fehrentz, J.-A. (2000) Synthesis of *N*-protected α-amino aldehydes from their morpholine amide derivatives. *Tetrahedron Lett*, **41**, 37–40.

[139] Pavol and Zlatoidsky (1995) Preparation of N^2-protected amino acid aldehydes via reduction of corresponding acid halides with lithium tris-(tert-butoxy)-aluminium hydride. *Tetrahedron Lett*, **36**, 7281–7284.

[140] Omura, K. and Swern, D. (1978) Oxidation of alcohols by 'activated' dimethyl sulfoxide. A preparative, steric and mechanistic study. *Tetrahedron*, **34**, 1651–1660.

[141] Dess, D. B. and Martin, J. C. (1983) Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones. *J Org Chem*, **48**, 4155–4156.

[142] Sergeev, M. E., Pronin, V. B., and Voyushina, T. L. (2005) Procedure for the oxidation of β-amino alcohols to α-amino aldehydes. *Synlett*, **2005**, 2802–2804.

[143] Garner, P. and Park, J. M. (1987) The synthesis and configurational stability of differentially protected β-hydroxy-α-amino aldehydes. *J Org Chem*, **52**, 2361–2364.

[144] Liang, X., Andersch, J., and Bols, M. (2001) Garner's aldehyde. *J Chem Soc, Perkin Trans 1*, **2001**, 2136–2157.

[145] Ilisz, I., Berkecz, R., and Péter, A. (2008) Application of chiral derivatizing agents in the high-performance liquid chromatographic separation of amino acid enantiomers: A review. *J Pharm Biomed Anal*, **47**, 1–15.

[146] Péter, A., Arki, A., Vékes, E., Tourwé, D., Lázár, L., Fülöp, F., and Armstrong, D. W. (2004) Direct and indirect high-performance liquid chromatographic enantioseparation of β-amino acids. *J Chromatogr A*, **1031**, 171–178.

[147] Sanger, F. (1945) The free amino groups of insulin. *Biochem J*, **39**, 507–515.

[148] Bhushan, R. and Brckner, H. (2004) Marfey's reagent for chiral amino acid analysis: A review. *Amino Acids*, **27**, 231–247.

[149] Einarsson, S., Josefsson, B., Moller, P., and Sanchez, D. (1987) Separation of amino acid enantiomers and chiral amines using precolumn derivatization with (+)-1-(9-fluorenyl)ethyl chloroformate and reversed-phase liquid chromatography. *Anal Chem*, **59**, 1191–1195.

[150] Henry, L., Leung, I. K. H., Claridge, T. D. W., and Schofield, C. J. (2012) γ-butyrobetaine hydroxylase catalyses a stevens type rearrangement. *Manuscript submitted*.

[151] Gottlieb, H. E., Kotlyar, V., and Nudelman, A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *J Org Chem*, **62**, 7512–7515.

[152] Gasparini, G., Prins, J., Leonard, and Scrimin, P. (2008) Exploiting neighboring-group interactions for the self-selection of a catalytic unit. *Angew Chem Int Ed Engl*, **47**, 2475–2479.

[153] Dokichev, T., Latypova, D., Shakirov, R., Biglova, R., and Talipov, R. (2010) Catalyzed synthesis of β-amino acids esters. *Russ J Org Chem*, **46**, 755–757.

[154] Bergmeier, S. C., Lee, W. K., and Rapoport, H. (1993) Chirospecific synthesis of (1*S*,3*R*)-1-amino-3-(hydroxymethyl)cyclopentane, a precursor for carbocyclic nucleoside synthesis. Intramolecular aziridine cyclizations. *J Org Chem*, **58**, 5019–5022.

[155] Ensch, C. and Hesse, M. (2002) Total syntheses of the spermine alkaloids (–)-(*R*,*R*)-hopromine and (±)-homaline. *Helv Chim Acta*, **85**, 1659–1677.

[156] Calvisi, G., Dell'Uomo, N., De Angelis, F., Dejas, R., Giannessi, F., and Tinti, M. (2003) A practical and stereoconservative synthesis of (*R*)-3-amino-4-(trimethylammonio)butanoate ((*R*)-aminocarnitine), and its trimethylphosphonium and simple ammonium analogues starting from D-aspartic acid. *Eur J Org Chem*, **2003**, 4501–4505.

[157] Torne, C. W., Sanderson, S. J., Mottram, J. C., Coombs, G. H., and Meldal, M. (2004) Combinatorial library of peptidotriazoles: Identification of [1,2,3]-triazole inhibitors against a recombinant *Leishmania mexicana* cysteine protease. *J Comb Chem*, **6**, 312–324.

1.5 Materials and Methods

1.5.1 General Material and Methods

Unless otherwise stated, chemicals were from Alfa Aesar (Heysham, UK), Aldrich (Dorset, UK), Acros chemicals (Loughborough, UK), Bachem (St. Helens Merseyside, UK) or Fisher Scientific (Loughborough, UK) and used without further purification. Solvents for chemical transformations, work-up and chromatography were from Rathburn (Walkerburn, UK) or Aldrich (Dorset, UK) at HPLC grade, and used without further distillation. Dried solvents were from Aldrich (Dorset, UK) or by filtration through columns containing activated aluminium oxide under argon. Silica gel 60 F_{254} analytical thin layer chromatography (TLC) plates were from Merck (Darmstadt, Germany) and visualised under UV light, or with potassium permanganate stain. Chromatographic purifications were performed using prepacked SNAP columns on a Biotage SP1 Purification system (Uppsala, Sweden). Yields refer to purified, dried, and spectroscopically pure compounds. NMR tubes were obtained from Aldrich (5 mm) or from Bruker (1 and 2 mm). Deuterated solvents were obtained from Aldrich and Apollo Scientific Ltd. All ^{1}H and ^{13}C NMR Spectra were recorded using either Bruker AV 500 MHz with ^{13}C cryoprobe and variable temperature setup or Bruker DQX 400 MHz. All chemical shifts are given in ppm relative to the solvent peak [151]. Coupling constants *J* are given in Hz to the nearest 0.5 Hz value. Low Resolution (LR) Electrospray Ionisation (ESI) Mass Spectrometry (MS) data (m/z) were obtained on a Fision VG Platform using quadrupole analyser. High Resolution (HR) mass spectrometry data (m/z) data were obtained from an Bruker MicroTOF instrument using ESI source and Time of Flight (TOF) analyser. Infrared (IR) spectra were recorded on a Bruker Tensor 27 instrument.

Preparative HPLC

Preparative HPLC was performed using a UltiMate® 3000 system (Dionex), a Vydac 218TP C18 column (250 mm x 25 mm 10–15μ, Grace Davison Discovery Sciences) and the following mobile phase: A = 0.1% formic acid in water, B = 0.1% formic acid in acetonitrile. Compounds with no detectable chromophore were purified by splitting a small portion of the flow into a PL-ELS 2100 Ice™ (Varian Inc.) evaporative light scattering detector (ELSD). Fractions containing the compounds of interest were collected in Falcon plastic tubes and freeze-dried.

Analytical HPLC

Analytical reverse phase HPLC was performed using a 1290 Infinity™ system (Agilent Technologies) and various mobile phase compositions, as described in the text. The following columns were used in this study: Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μ particle size) (Phenomenex), ChiroBiotic V (250 mm x 4.6 mm, 5 μ particle size) (Advanced Separation Technologies Inc.), ChiroBiotic T (250 mm x 4.6 mm, 5 μ particle size) (Advanced Separation Technologies Inc.), ChiraSpher NT column (250 mm x 4.0 mm, 5 μm particle size) (LiChroCART).

Normal phase chiral HPLC for enantiomeric excess determination was performed using a 1290 Infinity system (Agilent Technologies). Approximately 20 μg of sample in isopropanol was injected on a Chiracel AD column (250 mm x 4.6 mm x 5 μ) and eluted using 1 mL min^{-1} of 5% isopropanol in *n*-hexane and the product identified by following the absorbance at 220 nm.

1.5.2 BBOX Assay Monitoring Using NMR and Mass Spectrometry Analyses

All the work on NMR and MS monitoring of the BBOX assays and inhibition using the substrate analogues synthesised in this study was performed by Ivanhoe K. H. Leung as described recently [90]. The extensive NMR analysis leading to the identification of 3-amino-4-(methylamino)butanoic acid as the rearrangement product was also performed by Ivanhoe K. H. Leung [96].

NMR Assays

NMR assays for BBOX activity and product identification experiments were conducted at 700 MHz using a Bruker Avance III spectrometer equipped with an inverse TCI cryoprobe optimised for ^{1}H observation and 3mm diameter Bruker MATCH micro tubes (Hilgenberg) containing 160 μL final sample volume. Solutions were buffered using Tris or Tris-$^{2}H_{11}$ (pH 7.5) dissolved in 100% D_2O or 95% H_2O/5% D_2O. Pulse tip-angle calibration using the single-pulse nutation method was undertaken for each condition using a blank assay sample. Reaction products were identified and assigned by a combination of multiplicity edited HSQC, HMBC, 1D-TOCSY, and COSY where applicable.

Standard kinetic assays were conducted at 298 K in solutions containing 12.5 nM BBOX, 50 μM Fe(II), 50 μM substrate, 1.5 mM 2OG, 5 mM ascorbate, 50 mM Tris-$^{2}H_{11}$ (pH 7.5), and 80 mM KCl. Reactions were started by the addition of BBOX. Kinetic time courses were followed by using standard ^{1}H acquisition (ten transients with a relaxation delay of 1.2 s). The total time lapse between addition of the enzyme and the end of the first acquisition was 4 min. Time courses were followed for at least 15 min, and each point was carried out at least in triplicate.

Inhibition assays were conducted at 298 K in solutions containing 12.5 nM BBOX, 50 μM Fe(II), 50 μM γ-BB, 1.5 mM 2OG, 5 mM ascorbate, 50 mM Tris-$^{2}H_{11}$ (pH 7.5), 80 mM KCl, and varying concentrations of inhibitor. Reactions were initiated by BBOX addition. The extent of turnover of γ-BB to give L-Car was measured 3 min 45 s after mixing by using standard ^{1}H acquisition (ten transients with a relaxation delay of 1.2 s). Response curves were fitted using OriginPro 8.0 (OriginLab, Northampton, MA, USA). Each point was carried out at least in triplicate.

MS Assays

MS experiments for substrate analogue hydroxylation were conducted using a Waters LCT Premier XE (Waters, Milford, MA, USA) and/or Bruker MicrOTOF. Assay conditions were as for the NMR work, except that the solution was buffered in 100% H_2O Tris-$^{2}H_{11}$ (pH 7.5); quenching was by addition of 80% MeCN.

1.5.3 Synthesis of Trimethyl-(2-oxo-2-methyl-ethyl)ammonium Triflate

***N*,*N*,*N*-Trimethyl-2-oxopropan-1-aminium triflate 18** [120, 123]

The triflate salt of *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** was prepared as previously reported [120, 123]. 1-Dimethylamino-2-propanone **19** (200 mg, 2.0 mmol) was dissolved in anhydrous dichloromethane (40 mL) and methyl triflate (380 mg, 1.1 eq.) was added dropwise at room temperature. The reaction was allowed to stir under nitrogen atmosphere for 30 min and an orange oil was then precipitated upon addition of *n*-hexane (200 mL). After decantation of the hexane layer, the oil was dried under high vacuum and the triflate salt of *N*,*N*,*N*-trimethyl-2-oxopropan-1-aminium **18** recovered as an orange solid (450 mg, 85%). For assay purpose, a small fraction (20 mg) of the product was desalted using preparative reverse phase HPLC purification.

mp 55–57 °C; IR (neat) ν/cm^{-1}: 1737 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 4.48 (*s*, 2H), 3.23 (*s*, 9H), 2.19 (*s*, 3H); ^{13}C NMR (400 MHz, D_2O): δ = 202.2, 69.8, 54.1, 28.1; LR ESI-MS: 116.09 ($[M]^+$).

1.5.4 Synthesis of BBOX Substrate Analogues

3-(Trimethylammonio)propanoate 40 [125]

3-(Trimethylammonio)propanoate **40** (commonly known as β-alanine betaine) chloride salt was prepared as reported [125]. β-Alanine (2.00 g, 22.4 mmol) was dissolved into an aqueous formaldehyde solution (37%, 6.5 mL) and aqueous formic acid (90%, 7.5 mL). The reaction mixture was stirred under reflux for 24 h. Concentrated hydrochloric acid (3 mL) was added and the solvent was evaporated to dryness. The hydrochloride salt of compound **40** was recovered as a white solid, washed with *n*-hexane and recrystallised from ethanol/acetone mixture (2.92 g, 78%).

mp 198–200 °C (lit. 119–120 °C for the free base [125]); IR (neat) ν/cm^{-1}: 3439 (OH), 1736 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.62 (*t*, *J* = 7.5, 2H), 3.10 (*s*, 9H), 2.92 (*t*, *J* = 7.5, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 173.6, 61.9, 63.4, 28.4; HR ESI-MS: 132.1012 ($[M]^+$, $C_6H_{14}NO_2^+$, calc. 132.1019).

4-(Dimethylamino)butanoic acid 34

4-(Dimethylamino)butanoic acid **34** is commercially available. 4-Aminobutanoic acid **15** (515 mg, 5 mmol) was treated as **40** above and the reaction mixture was stirred under reflux for 40 h. Concentrated hydrochloric acid (3 mL) was added and the solvent was evaporated to dryness. The hydrochloride salt of 4-(dimethylamino)butanoic acid **34** was recovered as a white solid, washed with *n*-hexane and recrystallised from ethanol/acetone mixture (655 mg, app. quant.).

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3412 (OH), 1736 (OC=O); ^{1}H NMR (400 MHz, MeOD): δ = ; ^{13}C NMR (100 MHz, MeOD): δ = ; LR ESI-MS: 132.10 ($[M+H]^+$).

4-(Trimethylammonio)butanoate 7 (γ-BB) [152]

Several procedures have been reported for the synthesis of 4-(trimethylammonio)butanoate **7** including from ethyl 4-chlorobutyrate and trimethylamine followed by hydrolysis [152]. 4-(Trimethylammonio)butanoate **7** is commercially available too. 4-(Dimethylamino)butanoic acid **34** (131 mg, 1.0 mmol) and potassium hydrogen carbonate (450 mg, 4.5 eq.) were combined in methanol (10 mL) under a N_2 atmosphere. The resulting suspension was stirred for 15 min at 0 °C prior to addition of iodomethane (0.3 mL, 5 eq.). The reaction was allowed to slowly reach room temperature, then stirred for 3 days. The precipitate that formed was washed three times with dichloromethane (5 mL) followed by warm ethyl acetate to give 4-(trimethylammonio)butanoate **7** potassium iodide salt as a white solid (139 mg, 95%). For assay purpose, a small fraction (20 mg) of the product was desalted using preparative reverse phase HPLC.

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3389 (b, OH), 1738 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = ; ^{13}C NMR (100 MHz, D_2O): δ = ; HR ESI-MS: 146.1170 ($[M]^+$, $C_7H_{16}NO_2^+$, calc. 146.1176).

5-(Dimethylamino)pentanoic acid 35

5-Aminopentanoic acid **37** (5-aminovaleric acid) (0.65 g, 5.5 mmol) was treated as **40** above and the reaction mixture was stirred under reflux for 40 h. Concentrated hydrochloric acid (3 mL) was added and the solvent was evaporated to dryness. The hydrochloride salt of 5-

(dimethylamino)pentanoic acid **35** was recovered as a white solid, washed with *n*-hexane and recrystallised from ethanol/acetone mixture (1.0 g, app. quant.).

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3462 (b, OH), 1757 (OC=O); ^{1}H NMR (400 MHz, MeOD): δ = 3.00-2.96 (*m*, 2H), 2.75 (*s*, 6H), 2.31 (*t*, *J* = 7.0, 2H), 1.76-1.68 (*m*, 2H), 1.65-1.58 (*m*, 2H); ^{13}C NMR (100 MHz, MeOD): δ = 168.0, 57.6, 42.8, 34.2, 24.1, 22.2; HR ESI-MS: ([M+Na]$^{+}$, $C_7H_{15}NNaO_2^{+}$, calc. 168.1000).

5-(Trimethylammonio)pentanoate 41

5-Aminopentanoic acid **37** (118 mg, 1.0 mmol) and potassium hydrogen carbonate (450 mg, 4.5 eq.) were combined in methanol (10 mL) under a N_2 atmosphere and the resultant suspension was stirred for 15 min at 0 °C prior to addition of iodomethane (0.3 mL, 5 eq.). The reaction was allowed to slowly reach room temperature, then stirred for 3 days. The precipitate formed was washed three times with dichloromethane (5 mL) followed by warm ethyl acetate to give 5-(trimethylammonio)pentanoate **41** potassium iodide salt as a white solid (292 mg, 89%). For assay purpose, a small fraction (20 mg) of the product was desalted using preparative reverse phase HPLC.

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3424 (b, OH), 1755 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.29-3.25 (*m*, 2H), 3.05 (*s*, 9H), 2.29 (*t*, *J* = 7.5, 2H), 1.79-1.71 (*m*, 2H), 1.61-1.53 (*m*, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 180.9, 170.6, 66.6, 53.2, 35.4, 22.3; HR ESI-MS: 160.1332 ([M]$^{+}$, $C_8H_{18}NO_2^{+}$, calc. 160.1332).

6-(Dimethylamino)hexanoic acid 36

6-Aminocaproic acid (1.31 g, 10 mmol) was dissolved into a 37% formaldehyde aqueous solution (10 mL) and 90% formic acid (6 mL) and the reaction mixture was stirred under reflux conditions for 40 h. Concentrated hydrochloric acid (3 mL) was added and the solvent was evaporated to dryness. The hydrochloride salt of 6-(dimethylamino)hexanoic acid **36** was recovered as a white solid, washed with *n*-hexane and recrystallised from ethanol/acetone mixture (2.0 g, app. quant.).

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3487 (b, OH), 1746 (OC=O); ^{1}H NMR (400 MHz, MeOD): δ = 2.97-2.93 (*m*, 2H), 2.73 (*s*, 6H), 2.26 (*t*, *J* = 7.0, 2H), 1.73-1.57 (*m*, 4H), 1.40.1.32 (*m*, 2H); ^{13}C NMR (100 MHz, MeOD): δ = 168.1, 57.7, 42.7, 34.4, 26.2, 24.6, 24.1; HR ESI-MS: ([M+Na]$^{+}$, $C_8H_{17}NNaO_2^{+}$, calc. 182.1157).

6-(Trimethylammonio)hexanoate 42

6-(Trimethylammonio)hexanoate **42** potassium iodide salt (314 mg, 91%) was obtained from 6-aminohexanoic acid **38** (131 mg, 1.0 mmol) following the same procedure as used for the preparation of 5-(trimethylammonio)pentanoate **41**. For assay purpose, a small fraction (20 mg) of the product was desalted using preparative reverse phase HPLC.

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3423 (b, OH), 1754 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.28-3.23 (*m*, 2H), 3.04 (*s*, 9H), 2.26 (*t*, *J* = 7.5, 2H), 1.79-1.71 (*m*, 2H), 1.62-1.55 (*m*, 2H), 1.36-1.29 (*m*, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 181.1, 170.5, 66.8, 53.1, 35.5, 24.7, 22.4; HR ESI-MS: 174.1493 ([M]$^{+}$, $C_9H_{20}NO_2^{+}$, calc. 174.1489).

1.5.5 Synthesis of THP and Analogues

General procedure A for the synthesis of the methyl esters: [78]

The following protocol was adapted from a reported procedure [78]. Methyl acrylate (1.8 mL, 20.0

mmol) or methyl crotonate (2.1 mL, 20.0 mmol) were added over 3 h to ice cold dimethylhydrazine **57**, 2-methylpropan-2-amine **56**, or propan-2-amine **55** (30.0 mmol in each case) under a nitrogen atmosphere with protection from light. The reaction was slowly allowed to reach room temperature, then stirred for 12–48 h. The excess reagent was carefully removed at atmospheric pressure to give the products as colourless liquids. Due to their low boiling points and because ^{1}H NMR analyses implied a purity higher than 95%, the crude compounds were used for the subsequent step without further purification.

Methyl 3-(2,2-dimethylhydrazino)propanoate 62 [78]

The general procedure A using dimethylhydrazine **57** (2.3 mL, 1.5 eq.) and methyl acrylate (2.7 mL, 20.0 mmol) (reaction time: 48 h) yielded methyl 3-(2,2-dimethylhydrazino)propanoate **62** as a pale yellow oil (2.9 g, 99%).

TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.30; IR (neat) ν/cm^{-1}: 3452 (NH), 1738 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.62 (*s*, 3H), 2.98 (*t*, *J* = 6.5, 2H), 2.47 (*t,* *J* = 6.5, 2H), 2.36 (*s*, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.3, 51.4, 47.5, 44.0, 33.4; HR ESI-MS: 147.1122 ([M+H]$^+$, $C_6H_{15}N_2O_2^+$, calc. 147.1122).

Methyl 3-(2,2-dimethylhydrazino)butanoate 65

The general procedure A using dimethylhydrazine **57** (2.3 mL, 1.5 eq.) and methyl crotonate (3.2 mL, 20 mmol) (reaction time: 24 h) yielded methyl 3-(2,2-dimethylhydrazino)butanoate **65** as a pale yellow liquid (3.0 g, 94%).

TLC: 19:1 CH_2Cl_2/MeOH, R_f = 0.15; IR (neat) ν/cm^{-1}: 3448 (NH), 1724 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.63 (*s*, 3H), 3.31-3.26 (*m*, 1H), 2.49 (*dd*, *J* = 15.5, 7.0, 1H), 2.37 (*s*, 6H), 2.24 (*dd*, *J* = 15.5, 6.0, 1H), 1.03 (*d*, *J* = 6.0, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.9, 52.1, 46.4, 43.2, 33.2, 26.2; HR ESI-MS: 161.1286 ([M+H]$^+$, $C_7H_{17}N_2O_2^+$, calc. 161.1285).

Methyl *N*-*tert*-butyl-β-alaninate 59

The general procedure A using *tert*-butylamine (3.1 mL, 1.5 eq.) and methyl acrylate (reaction time: 12 h) yielded methyl *N*-*tert*-butyl-β-alaninate **59** as a colourless oil (3.0 g, 95%).

TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.20; IR (neat) ν/cm^{-1}: 3323 (NH), 1739 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.63 (*s*, 3H), 2.77 (*t*, *J* = 6.5, 2H), 2.44 (*t*, *J* = 6.5, 2H), 1.05 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.3, 51.4, 50.2, 37.9, 35.4, 28.9; HR ESI-MS: 160.1334 ([M+H]$^+$, $C_8H_{18}NO_2^+$, calc. 160.1332).

Methyl *N*-isopropyl-β-alaninate 58 [153]

The general procedure A using isopropylamine (2.6 mL, 1.5 eq.) and methyl acrylate (reaction time: 16 h) yielded methyl *N*-isopropyl-β-alaninate **58** as a colourless oil (2.75 g, 93%).

TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.10; IR (neat) ν/cm^{-1}: 3289 (NH), 1727 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.63 (*s*, 3H), 2.82 (*t*, *J* = 6.5, 2H), 2.76 (*sept*, *J* = 6.5, 1H), 2.46 (*t*, *J* = 6.5, 2H), 1.01 (*d*, *J* = 6.5, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.2, 51.4, 48.3, 42.4, 34.7, 22.8; HR ESI-MS: 146.1176 ([M+H]+, $C_7H_{16}NO_2^+$, calc. 146.1176).

2-(3-Methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium 63 [78]

Iodomethane (0.125 mL, 1.6 eq.) was added to an ice cold solution of methyl 3-(2,2-dimethylhydrazino)propanoate **62** (200 mg, 1.3 mmol) in a 9:1 mixture of cyclohexane/ethanol (1

mL) under a nitrogen atmosphere and protection from the light. The reaction was slowly allowed to reach room temperature and left stirring for 16 h. An orange oil precipitated, which was dissolved in warm EtOH (1 mL) after removing the solvent by decantation. A pale yellow precipitate formed upon addition of cyclohexane (20 mL). The iodide salt of 2-(3-methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium **63** was collected by decantation of the hexane layer and dried under high vacuum (296 mg, 79%).

mp 105–107 °C; IR (neat) ν/cm^{-1}: 3451 (NH), 1734 (OC=O); ^{1}H NMR (400 MHz, MeOD): δ = 3.71 (*s*, 3H), 3.35 (*s*, 9H), 3.29 (*t*, *J* = 6.5, 2H), 2.56 (*t*, *J* = 6.5, 2H); ^{13}C NMR (100 MHz, MeOD): δ = 172.0, 55.3, 54.7, 39.0, 31.9; LR ESI-MS: 162.12 ([M]+).

[^{13}C]-Labeled 2-(3-methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium 64

The same procedure as above using methyl 3-(2,2-dimethylhydrazino)propanoate **62** (200 mg, 1.3 mmol) and [^{13}C]-MeI (0.125 mL, 1.6 eq.) gave the labeled 2-(3-methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium **64** as a pale yellow solid (285 mg, 76%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.69 (*s*, 3H), 3.31 (*d*, *J* = 3.0, 6H), 3.31 (*d*, *J* = 145.0, 3H), 3.27 (*t*, *J* = 6.5, 2H), 2.59 (*t*, *J* = 6.5, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 54.7; HR ESI-MS: 162.1321 ([M]+, $C_6{}^{13}CH_{17}N_2O_2{}^+$, calc. 162.1318).

General procedure B for the hydrolysis of the methyl ester of THP and its analogues:

The methyl esters (1 mmol) were hydrolysed in a 1:1 THF/water mixture (2 mL) using one equivalent of sodium hydroxide (40 mg, 1 mmol) over 16 h at room temperature. Evaporation of the solvent gave the products as sodium salts in an apparently quantitative yield (purity $\geq$ 95% as judged by ^{1}H/^{13}C NMR). For assay purposes, a small fraction (20 mg) of the products was desalted using reverse phase HPLC.

3-(2,2-Dimethylhydrazino)propanoic acid 68 (DHP)

Following the general procedure B, hydrolysis of methyl 3-(2,2-dimethylhydrazino)propanoate **62** (180 mg, 1.2 mmol) gave the sodium salt of 3-(2,2-dimethylhydrazino)propanoic acid **68** as a white solid (app. quant.).

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 1623 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.11 (*t*, *J* = 6.5, 2H), 2.80 (*s*, 6H), 2.40 (*t*, *J* = 6.5, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.1, 44.4, 42.8, 35.1; HR ESI-MS: 155.0792 ([M+H]$^+$, $C_5H_{13}N_2O_2{}^+$, calc. 155.0796).

***N-tert*-Butyl-β-alanine 54**

Following the general procedure B, hydrolysis of methyl *N-tert*-butyl-β-alaninate **59** (160 mg, 1 mmol) gave the sodium salt of *N-tert*-butyl-β-alanine **54** as a white solid (165 mg, app. quant.).

mp 180–184 °C; IR (neat) ν/cm^{-1}: 3423 (NH), 1619 (OC=O); 1H NMR (400 MHz, D_2O): δ = 3.03 (*t*, *J* = 7.0, 2H), 2.43 (*t*, *J* = 7.0, 2H), 1.25 (*s*, 9H); ^{13}C NMR (100 MHz, D_2O): δ = 179.2, 55.5, 38.8, 34.4, 25.8; HR ESI-MS: 146.1177 ([M+H]$^+$, $C_7H_{16}NO_2{}^+$, calc. 146.1176).

***N*-Isopropyl-β-alanine 53**

Following the general procedure B, hydrolysis of methyl *N*-isopropyl-β-alaninate **58** (145 mg, 1 mmol) gave the sodium salt of *N*-isopropyl-β-alanine **53** as a pale orange solid (150 mg, app. quant.).

mp 135–136 °C; IR (neat) ν/cm^{-1}: 3416 (NH), 1749 (OC=O); 1H NMR (400 MHz, D_2O): δ = 3.41 (*spt*, *J* = 6.5, 1H), 3.27 (*t*, *J* = 6.5 Hz, 2H), 2.78 (*t*, *J* = 6.5 Hz, 2H), 1.28 (*d*, *J* = 6.5 Hz, 6H);

^{13}C NMR (100 MHz, D_2O): δ = 174.5, 51.5, 40.5, 30.7, 18.5; LR ESI-MS: 132.09 ($[M+H]^+$), 154.08 ($[M+Na]^+$).

3-(2,2,2-Trimethyldiazan-2-ium-1-yl)propanoate 27 (THP) [78]

Hydrolysis of 2-(3-methoxy-3-oxopropyl)-1,1,1-trimethylhydrazinium **63** (288 mg, 1 mmol) following the general procedure B yielded the sodium iodide salt of 3-(2,2,2-trimethyldiazan-2-ium-1-yl)propanoate **27** as a pale yellow solid (293 mg, app. quant.).

mp $\geq$ 300 °C.; IR (neat) ν/cm^{-1}: 3443 (NH), 1602 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.26 (*s*, 9H), 3.15 (*t*, *J* = 6.5, 2H), 2.31 (*t*, *J* = 6.5, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 179.0, 54.0, 39.9, 34.2; LR ESI-MS: 147.11 ($[M]^+$), 169.10 ($[M-H+Na]^+$).

[^{13}C]-Labeled 3-(2,2,2-Trimethylhydrazinium)propionate 67 ([^{13}C]-THP)

Following the general procedure B, the sodium iodide salt of [^{13}C]-labeled methyl-3-(2,2,2-trimethylhydrazinium)propionate **67** was obtained as a pale yellow solid (295 mg, app. quant.).

mp $\geq$ 300 °C; IR (neat) ν/cm^{-1}: 3442 (NH), 1604 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.26 (*d*, *J* = 3.0, 6H), 3.26 (*d*, *J* = 145.0, 3H), 3.15 (*t*, *J* = 6.5, 2H), 2.31 (*t*, *J* = 6.5, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 179.0, 54.0, 39.9, 34.2; HR ESI MS: 170.0986 ($[M+H]^+$, $C_5{}^{13}CH_{14}N_2NaO_2{}^+$, calc. 170.0981).

1.5.6 Synthesis of the Proposed Rearrangement Product

(2*S*)-4-*tert*-Butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid 83

L-Aspartic acid 4-*tert*-butyl ester **87** (1.89 g, 10 mmol) was dissolved in a mixture of THF (20 mL) and H_2O (12 mL) at 0 °C. After stirring for 5 min, di-*tert*-butyl dicarbonate (4.35 g, 2 eq.) was added in small portions followed by sodium carbonate (4.20 g, 4 eq.). The reaction was stirred at room temperature for 24 h, then the solvent was removed under reduced pressure. The residue was dissolved in 10% sodium hydrogen carbonate aqueous solution (50 mL), washed with diethylether (3 x 20 mL) and citric acid was added to the aqueous layer until pH 4.0 was reached. The product was extracted with ethyl acetate (3 x 100 mL). The combined organic phases were dried over magnesium sulfate, filtered and evaporated to give pure (2*S*)-4-*tert*-butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **83** as a white foam (2.82 g, 97%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.45; IR (neat) ν/cm^{-1}: 1717 (C=O); $[\alpha]_D^{25}$ = + 26.7° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.64 (*app bs*, 1H), 5.53 (*d*, *J* = 8.5, 1H), 4.57 (*dt*, *J* = 8.5, 4.5, 1H), 2.94 (*dd*, *J* = 17.0, 5.0, 1H), 2.74 (*dd*, *J* = 17.0, 5.0, 1H), 1.44 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.2, 170.3, 155.6, 82.0, 80.3, 49.9, 37.7, 28.2, 28.0; HR ESI-MS: 312.1411 ($[M+Na]^+$, $C_{13}H_{23}NNaO_6{}^+$, calc. 312.1418).

(2*R*)-4-*tert*-Butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid 109

D-Aspartic acid 4-*tert*-butyl ester **108** (757 mg, 4 mmol) was treated as compound **87** above to give (2*R*)-4-*tert*-butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **109** as a colourless oil (1.15 g, 99%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.45; IR (neat) ν/cm^{-1}: 1717 (C=O); $[\alpha]_D^{25}$ = - 27.3° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.18 (*app bs*, 1H), 5.53 (*d*, *J* = 8.5, 1H, NH), 4.57 (*dt*, *J* = 8.5, 4.5, 1H), 2.92 (*dd*, *J* = 17.0, 5.0, 1H), 2.74 (*dd*, *J* = 17.0, 5.0, 1H), 1.44 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.2, 170.3, 155.6, 82.0, 80.3, 49.9, 37.7, 28.2, 28.0; HR

ESI-MS: 312.1415 ([M+Na]$^+$, $C_{13}H_{23}NNaO_6{}^+$, calc. 312.1418).

tert-Butyl N^2-(*tert*-butoxycarbonyl)-*N*-methyl-L-α-asparaginate 82

(2*S*)-4-*tert*-Butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **83** (1.89 g, 10 mmol) was dissolved in dried tetrahydrofuran (50 mL). The solution was cooled to 0 °C prior to addition of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (3.83 g, 2 eq.) and hydroxybenzotriazole (3.38 g, 2 eq.). Diisopropylethylamine (3.5 mL, 2 eq.) was added followed by a 2 M monomethylamine solution in tetrahydrofuran (10 mL, 2 eq.). The reaction was allowed to reach room temperature and stirred under N_2 atmosphere for 16 h. Solvent was then evaporated, the residue was redissolved in 10% w/v aqueous ammonium chloride solution (50 mL) and the solution extracted with ethyl acetate (3 x 50 mL). Combined organic phases were dried over magnesium sulfate, filtered and the filtrate evaporated to give the crude product as a yellow solid. The residue was purified by chromatography (SNAP-100 column, gradient elution using 12–100% ethyl acetate in *n*-hexane). Purified *tert*-butyl N^2-(*tert*-butoxycarbonyl)-*N*-methyl-L-α-asparaginate **82** was recovered as a low melting point white solid (2.54 g, 84%).

mp $\geq$ 300 °C; TLC: 1:1 hex/EtOAc, R_f = 0.10; IR (neat) ν/cm^{-1}: 3324 (NH), 1699 (C=O), 1661 (C=O); $[\alpha]_D^{25}$ = - 27.2° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.63 (*app bs*, 1H), 5.70 (*app bs*, 1H), 4.39 (*app bs*, 1H), 2.74 (*m*, 4H), 2.57 (*dd*, *J* = 16.5, 6.0, 1H), 1.37 (*s*, 18H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.3, 170.8, 155.4, 81.3, 80.0, 50.7, 37.5, 28.1, 27.8, 26.1; HR ESI-MS: 325.1737 ([M+Na]$^+$, $C_{14}H_{26}N_2NaO_5{}^+$, calc. 325.1734).

N^2-(*tert*-Butoxycarbonyl)-*O*-*tert*-butyl-*N*-methyl-L-homoserinamide 85

tert-Butyl N^2-(*tert*-butoxycarbonyl)-*N*-methyl-L-α-asparaginate **82** (673 mg, 2.2 mmol) was dissolved in tetrahydrofuran (3 mL) and the solution was cooled to 0 °C. A 2 M solution of borane dimethyl sulfide complex in tetrahydrofuran (2.2 mL, 2 eq.) was added and the mixture stirred at room temperature for 16 h. Gas evolution was observed when the reaction was quenched by slow addition of methanol (10 mL) over 1 h. The solvent was then evaporated and the residue purified by chromatography (SNAP-100 column, gradient elution using 10–100% ethyl acetate in *n*-hexane). Some traces of starting material was recovered followed by the undesired product **85** (66 mg, 10%).

TLC: 1:1 hex/EtOAc, R_f = 0.10; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.74 (*app bs*, 1H), 5.93 (*d*, *J* = 6.5, 1H), 4.16-4.12 (*m*, 1H), 3.52-3.46 (*m*, 1H), 3.44-3.40 (*m*, 1H), 2.76 (*d*, *J* = 5.0, 3H), 1.99-1.88 (*m*, 2H), 1.39 (*s*, 9H), 1.14 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 174.8, 172.5, 79.6, 73.3, 59.3, 53.8, 32.5, 28.3, 27.3, 26.0; HR ESI-MS: 311.1942 ([M+Na]$^+$, $C_{14}H_{28}N_2NaO_4{}^+$, calc. 311.1947).

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate 88 [130, 131]

The synthesis of compound **88** was performed following a reported procedure [130, 131]. (2*S*)-4-*tert*-Butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **83** (2.14 g, 7.4 mmol) was dissolved in anhydrous tetrahydrofuran (50 mL). The solution was cooled to -30 °C before addition of *N*-methylmorpholine (0.82 mL, 1 eq.) and stirring for 10 min. Isobutyl chloroformate (0.98 mL, 1 eq.) was then added dropwise and the reaction warmed to -10 °C until the formation of the corresponding anhydride was found complete, as shown by TLC. The solution was cooled again to -30 °C prior to addition of sodium borohydride (945 mg, 3 eq.) followed by methanol (5 mL) dropwise over 10 min. After 1 h, the temperature was raised to 20 °C and the reaction left stirring for another hour before quenching with saturated ammonium chloride aqueous solution (50 mL). The tetrahydrofuran was evaporated under reduced pressure and the aqueous layer was extracted with diethylether (3 x 100 mL). The combined organic phases were dried over magnesium sulfate, filtered and the filtrate evaporated. The residue was purified by chromatography (SNAP-100 column, gradient

elution using 10–100% ethyl acetate in *n*-hexane). *tert*-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **88** was isolated as a low melting point white solid (1.85 g, 91%).

TLC: 1:1 hex/EtOAc, R_f = 0.30; IR (neat) ν/cm^{-1}: 3365 (O-H), 1700 (NC=O), 1698 (OC=O); $[\alpha]_D^{25}$ = + 3.8° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.32 (*app bs*, 1H), 3.92 (*app bs*, 1H), 3.60 (*app bs*, 2H), 3.35 (*app bs*, 1H), 2.52-2.41 (*m*, 2H), 1.40 (*s*, 9H), 1.39 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.0, 155.7, 81.0, 79.5, 64.4, 49.5, 37.3, 28.3, 27.9; HR ESI-MS: 298.1618 ([M+Na]$^+$, $C_{13}H_{25}NNaO_5^+$, calc. 298.1625).

tert-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate 110

The synthesis of *tert*-butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **110** was performed following the same procedure as for the (*S*)-enantiomer **88** [130, 131]. (*R*)-4-*tert*-butoxy-2-(*tert*-butoxycarbonylamino)-4-oxobutanoic acid **109** (900 mg, 3.1 mmol) was treated as **88** above to give compound **110** as a low melting point white solid (750 mg, 88%).

TLC: 1:1 hex/EtOAc, R_f = 0.30; IR (neat) ν/cm^{-1}: 3365 (O-H), 1714 (NC=O), 1697 (OC=O); $[\alpha]_D^{25}$ = - 4.0° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.27 (*app bs*, 1H), 3.95 (*app bs*, 1H), 3.65 (*app bs*, 2H), 3.01 (*app bs*, 1H), 2.56-2.45 (*m*, 2H), 1.44 (*s*, 9H), 1.42 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.1, 155.8, 81.2, 79.6, 64.7, 49.6, 37.4, 28.3, 28.0; HR ESI-MS: 298.1618 ([M+Na]$^+$, $C_{13}H_{25}NNaO_5^+$, calc. 298.1625).

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-[(4-methylphenyl)sulfonyl]oxybutanoate 89 [132]

The following procedure was adapted from a previously reported synthesis [132]. *tert*-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **88** (1.70 g, 6.2 mmol) was dissolved in dried acetonitrile (50 mL) together with pyridine (0.97 mL, 1.95 eq.) and *p*-toluenesulfonyl chloride (1.9 g, 1.6 eq.). After stirring under N_2 atmosphere at room temperature for 2 days, the reaction was concentrated under reduced pressure and the residue purified by chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). *tert*-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-[(4-methylphenyl)sulfonyl]oxybutanoate **89** was recovered as a white solid (1.01 g, 38%).

mp 45–48 °C; TLC: 8:2 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 3407 (N-H), 1751 (NC=O), 1720 (OC=C); $[\alpha]_D^{25}$ = - 16.2° (c 1.3, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.73 (*d*, *J* = 8.0, 2H), 7.30 (*d*, *J* = 8.0, 2H), 5.09 (*app bs*, 1H), 4.12 (*app bs*, 1H), 4.02 (*app bs*, 2H), 2.44 (*d*, *J* = 6.0, 2H), 2.39 (*s*, 3H), 1.37 (*s*, 9H), 1.36 (*J*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.7, 154.7, 144.9, 132.4, 129.8, 127.8, 81.3, 79.6, 70.3, 46.4, 36.3, 28.1, 27.8, 21.5; HR ESI-MS: 452.1698 ([M+Na]$^+$, $C_{20}H_{31}NNaO_7S^+$, calc. 452.1713).

(*S*)-*tert*-Butyl 5-oxotetrahydrofuran-3-ylcarbamate 90

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-[(4-methylphenyl)sulfonyl]oxybutanoate **89** (100 mg, 0.23 mmol) was dissolved in 2 M monomethylamine solution in THF (2.0 mL, 2.0 mmol, 9 eq.) in a sealed vial and the reaction was stirred under microwave irradiation at 100 °C (250 MW) for 1 h. The reaction was evaporated and the residue purified by chromatography (SNAP-10 column, gradient elution using 0–10% methanol in ethyl acetate, both containing 0.1% triethylamine) to yield the lactone **90** as a colourless oil (31 mg, 48%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.97 (*app bs*, 1H), 4.54 (*t*, *J* = 8.5, 1H), 4.21-4.14 (*m*, 1H), 4.05 (*dd*, *J* = 9.0, 6.0, 1H), 2.62-2.49 (*m*, 2H), 1.45 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.6, 159.1, 82.0, 69.4, 48.9, 40.7, 28.0; LR ESI-MS: 289.21 ([M+H]$^+$).

***N*-[(4-Methylphenyl)sulfonyl]-L-aspartic acid 94** [134]

N-[(4-Methylphenyl)sulfonyl]-L-aspartic acid **94** was synthesised following a reported procedure [134]. L-Aspartic acid (1.33 g, 10 mmol) was dissolved in a 10% w/w sodium hydroxide aqueous solution (8.0 mL, 2 eq.) and the solution cooled to 0 °C before addition of *p*-toluenesulfonyl chloride (2.1 g, 1.1 eq.) followed by *N*,*N*-diisopropylethylamine (1.9 mL, 1.1 eq.) and acetone (15 mL). The clear solution was stirred at room temperature for 18 h and diluted with 1 N sodium hydroxide solution (50 mL). The aqueous phase was washed with diethylether (2 x 20 mL), cooled to 0 °C and the pH was carefully adjusted to 1 using concentrated hydrochloric acid. The product was then extracted with diethylether (4 x 20 mL) and the combined organic layers dried over magnesium sulfate. The solvent was evaporated to yield *N*-[(4-methylphenyl)sulfonyl]-L-aspartic acid **94** as a white solid (2.66 g, 93%).

mp 107–109 °C (lit. 114–116 °C [134]); $[\alpha]_D^{25}$ = + 2.8° (c 1.0, CH_2Cl_2) (lit. + 1.6° (c 1.0, AcOH)); ^{1}H NMR (400 MHz, MeOD): δ = 7.75 (*d*, *J* = 8.0, 2H), 7.35 (*d*, *J* = 8.0, 2H), 4.18 (*t*, *J* = 6.0, 1H), 2.70 (*m*, 2H), 2.41 (*s*, 3H); ^{13}C NMR (100 MHz, MeOD): δ = 173.6, 173.5, 144.9, 139.3, 130.7, 128.4, 53.8, 38.9, 21.6; LR ESI-MS: 286.05 ([M-H]$^-$).

***N*-[(3*S*)-2,5-Dioxotetrahydrofuran-3-yl]-4-methylbenzenesulfonamide 95** [133, 154]

Compound **95** was synthesised following a reported procedure [133]. *N*-[(4-Mmethylphenyl)sulfonyl]-L-aspartic acid **94** (2.3 g, 8 mmol) was dissolved in neat acetic anhydride (20 mL) for 18 h and the volatiles were removed under vacuum at room temperature. The residue was triturated in cold diethylether and the resulting white solid dried after decanting the solvent to yield *N*-[(3*S*)-2,5-dioxotetrahydrofuran-3-yl]-4-methylbenzenesulfonamide **95** (1.62 g, 75%).

mp 138-142 °C (lit. 137-138 °C [133], 145-148 °C [154]); $[\alpha]_D^{25}$ = + 6.8° (c 1.0, CH_2Cl_2) (lit. + 7.2° (c 1.0, $CHCl_3$)); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.78 (*d*, *J* = 8.0, 2H), 7.36 (*d*, *J* = 8.0, 2H), 5.66-5.61 (*m*, 1H), 4.54-4.48 (*m*, 1H), 3.36-3.29 (*m*, 1H), 3.13-3.06 (*m*, 1H), 2.46 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.9, 168.1, 138.2, 129.9, 126.8, 72.7, 50.3, 34.6, 21.2; LR ESI-MS: 268.06 ([M-H]$^-$).

4-Methyl-*N*-[(3*S*)-5-oxotetrahydrofuran-3-yl]benzenesulfonamide 96 [154, 133, 155]

Compound **96** was synthesised following a reported procedure [133]. A solution of *N*-[(3*S*)-2,5-dioxotetrahydrofuran-3-yl]-4-methylbenzenesulfonamide **95** (1.60 g, 6 mmol) in dried tetrahydrofuran (20 mL) was added slowly at 0 °C to a suspension of sodium borohydride (238 mg, 1.05 eq.) in dried tetrahydrofuran (10 mL). After stirring for 1 h at 0 °C and 1 h at room temperature, the reaction mixture was diluted with ethanol (50 mL) and 1 M hydrochloric acid solution (20 mL). The mixture was stirred under reflux conditions for 2 h and the volatiles were removed under vacuum. The aqueous layer was extracted with ethyl acetate (3 x 20 mL) and the combined organic phases were dried over magnesium sulfate and evaporated. The residue was purified by chromatography (SNAP-100, gradient elution using 0–80% ethyl acetate in *n*-hexane) and 4-methyl-*N*-[(3*S*)-5-oxotetrahydrofuran-3-yl]benzenesulfonamide **96** was recovered as a white solid (690 mg, 45%).

mp 108-112 °C (lit. 89-91 °C [154], 112-113 °C [133], 109-112 °C [155]); $[\alpha]_D^{25}$ = - 9.8° (c 1.0, CH_2Cl_2) (lit. - 15.7° (c = 1.0, EtOH)); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.74 (*d*, *J* = 8.0, 2H), 7.33 (*d*, *J* = 8.0, 2H), 5.90 (*d*, *J* = 6.0, 1H), 4.41-4.35 (*m*, 1H), 4.16-4.10 (m, 2H), 2.64 (*dd*, *J* = 18.0, 7.5, 1H), 2.44 (*s*, 3H), 2.37 (*dd*, *J* = 18.0, 5.0, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 174.9, 144.3, 136.6, 130.1, 127.0, 73.0, 49.7, 34.7, 21.5; HR ESI-MS: 254.0491 ([M-H]$^-$, $C_{11}H_{12}NO_4S^-$, calc. 254.0493).

Ethyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate 91 [134, 156]

Ethyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **91** was synthesised by adapting a reported procedure [156]. Iodotrimethylsilane (0.5 mL, 3 eq.) was added at 0 °C to a solution of 4-methyl-*N*-[(3*S*)-5-oxotetrahydrofuran-3-yl]benzenesulfonamide **96** (300 mg, 1.2 mmol) in anhydrous dichloromethane (20 mL) and ethanol (1 mL, 10 eq.). The reaction was stirred at room temperature for 24 h and water (10 mL) was added before stirring for 5 min. The phases were separated and the organic layer was washed with a 5% sodium sulfite aqueous solution and water, dried over magnesium sulfate and evaporated. The residue was purified by silica chromatography (SNAP-25, gradient elution using 0–100% ethyl acetate in *n*-hexane) and ethyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **91** was recovered as a colourless oil (330 mg, 67%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.77 (*d*, *J* = 8.0, 2H), 7.32 (*d*, *J* = 8.0, 2H), 5.36 (*d*, *J* = 9.0, 1H), 4.11-4.02 (*m*, 2H), 3.60-3.52 (*m*, 1H), 3.33 (*dd*, *J* = 10.5, 4.5, 1H), 3.25 (*dd*, *J* = 10.5, 6.5, 1H), 2.69 (*dd*, *J* = 16.5, 5.5, 1H), 2.55 (*dd*, *J* = 16.5, 6.5, 1H), 2.44 (s, 3H), 1.22 (*t*, *J* = 7.0, 3H); LR ESI-MS: 434.02 ([M+Na]$^+$).

Isobutyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate 92 [156]

Isobutyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **92** was synthesised following a reported procedure [156]. 4-Methyl-*N*-[(3*S*)-5-oxotetrahydrofuran-3-yl]benzenesulfonamide **96** (300 mg, 1.2 mmol) was treated as above using isobutanol (1.1 mL) instead of ethanol and silica chromatography (SNAP-25 column, gradient elution using 0–100% ethyl acetate in *n*-hexane) yielded isobutyl (3*S*)-4-iodo-3-[(4-methylphenyl)sulfonyl]aminobutanoate **92** as a colourless oil (332 mg, 63%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.74 (*d*, *J* = 8.0, 2H), 7.32 (*d*, *J* = 8.0, 2H), 5.32 (*d*, *J* = 9.0, 1H), 3.93-3.87 (*m*, 2H), 3.52-3.48 (*m*, 1H), 3.53-3.49 (*m*, 1H), 3.30 (*dd*, *J* = 10.5, 4.5, 1H), 3.20 (*dd*, *J* = 10.5, 6.5, 1H), 2.71 (*dd*, *J* = 16.5, 5.0, 1H), 2.52 (*dd*, *J* = 16.5, 6.5, 1H), 2.40 (*s*, 3H), 1.90-1.88 (*m*, 1H), 0.92 (*d*, *J* = 6.5, 6H); LR ESI-MS: 462.03 ([M+Na]$^+$).

***tert*-Butyl (3*S*)-4-azido-3-[(*tert*-butoxycarbonyl)amino]butanoate 100** [157]

The synthesis of compound **100** was reported previously using a different procedure [157]. *tert*-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-[(4-methylphenyl)sulfonyl]oxybutanoate **89** (710 mg, 1.65 mmol) was dissolved in dried dimethylformamide (5 mL) together with sodium azide (170 mg, 1.6 eq.). The reaction was stirred at 60 °C and monitored by TLC. Upon completion, the mixture was diluted with water (50 mL) and extracted with ethyl acetate (3 x 50 mL). The combined organic phases were dried over magnesium sulfate, filtered and the filtrate was evaporated to give the crude which was purified by chromatography (SNAP-50 column, gradient elution using 0–60% ethyl acetate in *n*-hexane). Purified *tert*-butyl (3*S*)-4-azido-3-[(*tert*-butoxycarbonyl)amino]butanoate **100** was recovered as a colourless oil (330 mg, 67%).

TLC: 7:3 hex/EtOAc, R_f = 0.55; IR (neat) ν/cm^{-1}: 2113 (N_3), 1722 (C=O); $[\alpha]_D^{25}$ = - 9.6° (c 1.0, CH_2Cl_2) (lit. - 10.0° (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.16 (*app bs*, 1H), 4.04 (*app bs*, 1H), 3.49-3.38 (*m*, 2H), 2.45 (*d*, *J* = 6.5, 2H), 1.43 (*s*, 9H), 1.42 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.2, 154.9, 81.4, 79.7, 53.7, 47.4, 37.5, 28.2, 27.9; HR ESI-MS: 323.1689 ([M+Na]$^+$, $C_{13}H_{24}N_4NaO_4^+$, calc. 323.1690).

***tert*-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate 101**

The aldehyde **101** was obtained from *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-

hydroxybutanoate **88** using the Dess-Martin oxidation procedure. A solution of Dess-Martin periodinane (1.2 g, 1.15 eq.) in CH_2Cl_2 (25 mL) was added to a solution of *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **88** (680 mg, 2.47 mmol) in CH_2Cl_2 (25 mL) and the reaction was monitored by TLC while stirring at room temperature. After 2 h, dichloromethane (25 mL) was added and the excess reagent quenched with a saturated sodium hydrogen carbonate aqueous solution (50 mL) containing sodium thiosulfate (100 mg). The organic layer was separated and washed with saturated sodium hydrogen carbonate aqueous solution (50 mL), water (50 mL), dried over magnesium sulfate, filtered and evaporated. The residue was purified by silica chromatography (SNAP-50 column, gradient elution using 0–70% ethyl acetate in *n*-hexane)and compound **101** was recovered as a colourless oil (470 mg, 69%). This product was used immediately in the reductive amination step to prevent racemisation.

TLC: 1:1 hex/EtOAc, R_f = 0.55; IR (neat) ν/cm^{-1}: 3366 (NH), 1717 (C=O); $[\alpha]_D^{25}$ = - 22.6° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.56 (*s*, 1H), 5.63 (*d*, *J* = 7.5, 1H), 4.28-4.23 (*m*, 2H), 2.81 (*dd*, *J* = 17.0, 5.0, 1H), 2.67 (*dd*, *J* = 17.0, 5.0, 1H), 1.39 (*s*, 9H), 1.36 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 199.3, 170.1, 155.4, 81.7, 80.1, 56.0, 35.6, 28.1, 27.8; HR ESI-MS: 274.1651 ([M+H]$^+$, $C_{13}H_{24}NO_5^+$, calc. 274.1649).

tert-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate 111

tert-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **111** was obtained using the same Dess-Martin procedure as for the (*S*)-enantiomer **101**. *tert*-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **110** (680 mg, 2.5 mmol) was subjected to Dess-Martin oxidation as above to give the aldehyde **111** as a colourless oil (456 mg, 67%). This compound was used immediately in the reductive amination step to prevent racemisation.

TLC: 7:3 hex/EtOAc, R_f = 0.15; IR (neat) ν/cm^{-1}: 3362 (NH), 1717 (C=O); $[\alpha]_D^{25}$ = + 22.0° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.64 (*s*, 1H), 5.61 (*d*, *J* = 7.5, 1H), 4.32-4.26 (*m*, 2H), 2.88 (*dd*, *J* = 17.0, 5.0, 1H), 2.71 (*dd*, *J* = 17.0, 5.0, 1H), 1.46 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 199.3, 170.1, 155.4, 81.7, 80.1, 56.0, 35.6, 28.1, 27.8; HR ESI-MS: 274.1653 ([M+H]$^+$, $C_{13}H_{24}NO_5^+$, calc. 274.1649).

tert-Butyl 3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate 103 [130, 131]

The racemic aldehyde was synthesised *via* the Swern oxidation of *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-hydroxybutanoate **88** by adapting previously reported procedures [130, 131]. (*S*)-*tert*-Butyl 3-(*tert*-butoxycarbonylamino)-4-hydroxybutanoate 2 (3.05 g, 11.1 mmol) was dissolved in dried dichloromethane (20 mL) and the solution added dropwise to a premixed solution of oxalyl chloride (1.2 mL, 1,25 eq.) and dimethylsulfoxide (2.0 mL, 2.5 eq.) in anhydrous dichloromethane (100 mL) at -78 °C. The reaction was stirred for 30 min at -78 °C and Et_3N (8.3 mL, 5.4 eq.) was added dropwise. After another 30 min, water (50 mL) was added and the reaction allowed to reach room temperature over 30 min. The phases were separated, the aqueous layer extracted with dichloromethane (2 x 50 mL) and the organic phases combined, dried over magnesium sulfate, filtered and evaporated to give the crude product as a yellow oil. This residue was purified by chromatography (SNAP-100 column, gradient elution using 0–70% ethyl acetate in *n*-hexane) and *tert*-butyl 3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **103** was eluted first (1.87 g, 62%) followed by the unreacted starting material **88** (0.68 g, 22% recovery). All spectroscopic data were identical to the pure enantiomers **101** and **111** above, except for the optical rotation $[\alpha]_D^{25}$ = 0.0° (c 1.0, CH_2Cl_2).

tert-Butyl (3*S*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate 105

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **101** (430 mg, 1.57 mmol) was

dissolved in dried tetrahydrofuran (10 mL) together with *N*-benzylmethylamine (2.0 mL, 10 eq.) and the mixture was stirred at room temperature under N_2 atmosphere for 30 min. Sodium triacetoxyborohydride (530 mg, 1.5 eq.) was added in small portions. The reaction was stirred for 16 h. *N*-Benzylmethylamine (1.0 mL, 5 eq.) was added followed by sodium triacetoxyborohydride (530 mg, 1.5 eq.) after stirring for 30 min. After 6 h, the reaction was diluted with ethyl acetate (50 mL) and the organic phase washed with saturated sodium hydrogen carbonate solution (20 mL), brine (20 mL), dried over sodium hydrogen carbonate, filtered and evaporated to give the crude product as an oil containing a large excess of *N*-benzylmethylamine. Purification by chromatography (SNAP-50 column, gradient elution using 0–60% ethyl acetate in *n*-hexane) yielded the pure product *tert*-butyl (3*S*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **105** as a colourless oil (365mg, 61%). The ee was 77% as determined by chiral HPLC analysis (Appendix A.2).

TLC: 7:3 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 3367 (NH), 1713 (C=O); $[\alpha]_D^{25}$ = + 5.0° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.31-7.21 (*m*, 5H), 5.09 (*app bs*, 1H), 4.04 (*app bs*, 1H), 3.51 (*m*, 2H), 2.58 (*m*, 1H), 2.48-2.44 (*m*, 3H), 2.21 (*s*, 3H), 1.45 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.1, 155.4, 139.0, 129.0, 128.1, 126.9, 80.6, 62.5, 60.3, 46.0, 42.3, 38.4, 28.4, 28.0; HR ESI-MS: 379.2578 ([M+H]$^+$, $C_{21}H_{35}N_2O_4^+$, calc. 379.2591).

tert-Butyl (3*R*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate 112

The (3*R*) aldehyde **111** (261 mg, 0.96 mmol) was subjected to reductive amination with benzylmethylamine as for compound **105** to give *tert*-butyl (3*R*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **112** as a colourless oil (291 mg, 81%). The ee was 71% as determined by chiral HPLC analysis (Appendix A.1).

TLC: 7:3 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 3375 (NH), 1719 (C=O); $[\alpha]_D^{25}$ = - 4.0° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.31-7.21 (*m*, 5H), 5.10 (*app bs*, 1H), 4.04 (*app bs*, 1H), 3.51 (*m*, 2H), 2.58 (*m*, 1H), 2.48-2.44 (*m*, 3H), 2.21 (*s*, 3H), 1.45 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): 171.2, 155.5, 139.0, 128.9, 128.1, 127.0, 80.7, 62.5, 60.3, 46.0, 42.4, 38.4, 28.4, 28.1; LR ESI-MS: 379.22 ([M+H]$^+$), 401.23 ([M+Na]$^+$).

tert-Butyl 4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate 106

The racemic aldehyde *tert*-butyl 3-[(*tert*-butoxycarbonyl)amino]-4-oxobutanoate **103** (184 mg, 0.67 mmol) was subjected to reductive amination with benzylmethylamine as for compound **105** to give racemic *tert*-butyl 4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **106** as a colourless oil (226 mg, 89%). The spectroscopic data for the racemic product was identical than for both enantiomers, except for the optical rotation $[\alpha]_D^{25}$ = 0.0° (c 1.0, CH_2Cl_2). There was no ee as determined by chiral HPLC analysis (Appendix A.3).

tert-Butyl (3*S*)-4-[allyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate 104

Aldehyde **101** (450 mg, 1.6 mmol) was converted to *tert*-butyl (3*S*)-4-[allyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **104** following the same reductive amination procedure as for **105** using *N*-allylmethylamine (1.35 mL, 7.5 eq.) instead. Compound **104** was obtained as a colourless oil (464 mg, 86%).

TLC: 7:3 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 3367 (NH), 1715 (C=O); $[\alpha]_D^{25}$ = + 4.5° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.79 (*ddt*, *J* = 17.1, 10.4, 6.4, 1H), 5,13 (*m*, 2H), 3.94 (*m*, 1H), 2.99 (*dd*, *J* = 5.8, 4.8, 2H), 2.58 (*m*, 1H), 2.41 (*m*, 3H), 2.22 (*s*, 3H), 1.44 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.2, 155.4, 135.7, 117.4, 80.6, 79.1, 67.2, 59.6, 46.1, 42.4, 38.3, 28.4, 28.1; HR ESI-MS: 379.2578 ([M+H]$^+$, $C_{21}H_{35}N_2O_4^+$, calc. 379.2591).

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate 84

tert-Butyl (3*S*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **105** (243 mg, 0.64 mmol) was dissolved in dried tetrahydrofuran (10 mL) and catalytic palladium on activated carbon was added (10% w/w, 25 mg). The black suspension was placed under H_2 atmosphere (1 bar) and the reaction stirred vigorously for 12 h. The catalyst was removed by filtration through a pad of Celite® and *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** was obtained as a colourless oil after evaporation of the filtrate (178 mg, 96%).

Alternatively, **84** was obtained from the deprotection of allyl amine **104** (120 mg, 0.37 mmol) using 5% mol/mol tetrakis(triphenylphosphane)palladium(0) (20 mg) in methanol (5 mL) and *N*,*N*′-dimethylbarbituric acid as a scavanger (115 mg, 2 eq.) over 3 h at room temperature. Solvent was then evaporated and the crude purified by chromatography (SNAP-10) column using a 0–10% methanol in ethyl acetate gradient. Purified *tert*-butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** was recovered as a colourless oil (30 mg, 29%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.10; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.13 (*app bs*, 1H), 4.99 (*app bs*, 1H), 2.70 (*dd*, *J* = 12.0, 6.5, 1H), 2.62 (*dd*, *J* = 12.0, 5.5, 1H), 2.50 (*dd*, *J* = 15.0, 5.5, 1H), 2.42 (*s*, 3H), 2.41 (*dd*, J = 15.0, 6.5, 1H), 1.44 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.9, 155.4, 80.9, 79.2, 54.8, 47.4, 38.6, 36.4, 28.4, 28.0; HR ESI-MS: 289.2112 ($[M+H]^+$, $C_{14}H_{29}N_2O_4^+$, calc. 289.2122).

tert-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate 113

tert-Butyl (3*R*)-4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **112** (233 mg, 0.61 mmol) was treated as **105** and the pure product (*R*)-*tert*-butyl 3-(*tert*-butoxycarbonylamino)-4-(methylamino)butanoate was obtained as a colourless oil after evaporation of the filtrate (166 mg, 94%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.10; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.39 (*app bs*, 1H), 4.07 (*app bs*, 1H), 2.79 (*dd*, *J* = 12.0, 7.0, 1H), 2.62 (*dd*, *J* = 12.0, 5.5, 1H), 2.49 (*m*, 2H + 3H), 1.43 (*s*, 9H), 1.42 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.6, 155.5, 81.1, 79.5, 54.0, 46.7, 38.6, 35.5, 28.3, 28.0; HR ESI-MS: 289.2118 ($[M+H]^+$, $C_{14}H_{29}N_2O_4^+$, calc. 289.2122).

(3*S*)-3-Amino-4-(methylamino)butanoic acid 102

tert-Butyl (3*S*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **84** (120 mg, 0.42 mmol) was deprotected using a 2.0 M solution of hydrochloric acid in diethylether over 2 h at room temperature. The solvent was then decanted and the hydrochloride salt hydrate of the β-amino acid **102** recovered as a white foamy solid (77 mg, app. quant.). This product was purified using preparative HPLC for further NMR and MS analysis.

^{1}H NMR (400 MHz, D_2O): δ = 4.02 (*m*, 1H), 3.44 (*d*, *J* = 6.5, 2H), 2.95 (*dd*, *J* = 18.0, 5.5, 1H), 2.91 (*dd*, *J* = 18.0, 6.5, 1H), 2.79 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 173.1, 50.2, 45.2, 34.8, 34.1.

(3*R*)-3-Amino-4-(methylamino)butanoic acid 107

tert-Butyl (3*R*)-3-[(*tert*-butoxycarbonyl)amino]-4-(methylamino)butanoate **113** (165 mg, 0.57 mmol) was treated with a 2.0 M solution of hydrochloric acid in diethylether as compound **84** and the hydrochloride salt hydrate of (*R*)-3-amino-4-(methylamino)butanoic acid **107** was recovered as a white foamy solid (102 mg app. quant.). This product was purified using preparative HPLC prior to further NMR and MS analysis.

^{1}H NMR (400 MHz, D_2O): δ = 4.02 (*m*, 1H), 3.44 (*d*, *J* = 6.5, 2H), 2.95 (*dd*, *J* = 18.0, 5.5, 1H),

2.91 (*dd*, *J* = 18.0, 6.5, 1H), 2.79 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 173.1, 50.2, 45.2, 34.8, 34.1.

3-Amino-4-(methylamino)butanoic acid 77

Racemic 3-amino-4-(methylamino)butanoic acid **77** was obtained from *tert*-butyl 4-[benzyl(methyl)amino]-3-[(*tert*-butoxycarbonyl)amino]butanoate **106** in two steps as described for the (3*S*) and (3*R*)-enantiomers. The NMR spectroscopic data for compound **77** was identical to these for compounds **102** and **107**.

1.5.7 Derivatisation Reactions

Amino acids stock solutions

α-Amino acids stock solutions (38 mM) in water were prepared from commercially available L- and D-proline, L- and D-lysine, L- and D-ornithine. Stock solutions of (*S*)-**102**, (*R*)-**107** and racemic 3-amino-4-(methylamino)butanoic acid **77** β-amino acids were performed in water at a 38 mM concentration and kept at -20 °C until used. Dilutions to 38 μM were made directly prior to derivatisation reactions and chromatographic analysis.

Marfey's Reagent Derivatisation Reaction

Marfey's reagent N_α-(5-fluoro-2,4-dinitrophenyl)-L-valinamide (FDVA) **114** was purchased from Sigma-Aldrich as a solid. A 1% FDVA stock solution in acetone was prepared immediately prior to derivatisation assays.

All samples contained 170 μL final volume. The blank sample was composed of 20 μL $NaHCO_3$ in 150 μL water. The control reaction sample contained 20 μL $NaHCO_3$, 50 μL water and 100 μL reagent solution. The amino acid samples contained 20 μL $NaHCO_3$, 50 μL amino acid solution (38 μM) and 100 μL reagent solution. After adding each component, the samples containing 11 μM final amino acid concentration were mixed for 1 min and then heated at 40 °C for 1 h. The samples were then quenched with 830 μL 0.1 N aqueous HCl and injected onto the HPLC-MS instrument as such.

6-Aminoquinolyl-*N*-hydroxysuccinimidyl Carbamate Derivatisation Reaction

The 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate (AccQ-Tag) derivatisation kit was used following the procedure described by the supplier (Waters). The reagent solution was prepared by adding dried acetonitrile (1 mL) to the provided amount of AccQTag reagent. Mixing was performed by vortex briefly and the heating for 10 min at 55 °C. The borate buffer solution was used as provided.

The assay samples were as described for the NMR assay in a 160 μL final volume. Samples contained 12.5 nM BBOX, 50 μM Fe(II), 50 μM THP or [^{13}C]-labeled THP, 1.5 mM 2OG, 5 mM ascorbate, 80 mM KCl, all in 50 mM Tris H_2O (pH 7.5). The BBOX assays carried in the presence of THP or [^{13}C]-labeled THP were quenched by addition of 10% v/v acetonitrile and the samples were freeze-dried. The resulting solid was redissolved in the borate buffer supplied with the derivatisation reagent before adding a solution of 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate in dried acetonitrile. The derivtization reaction was carried at 60 °C for 15 min. A portion of the derivatised samples (2 μL) was directly subjected to MS analysis.

Standard amino acid samples derivatisation reaction were carried out in 100 μL final volume. The

blank sample was composed of borate buffer only. The control reaction sample contained 90 μL buffer and 10 μL derivatisation reagent. The amino acid samples contained 89 μL of borate buffer, 1 μL of amino acid stock solution (38 μM) and 10 μL of reagent solution. After adding each component, the samples containing 380 nM final amino acid concentration were vortexed for 1 min and then heated at 60 °C for 15 min. A portion of the derivatised samples (2 μL) was directly subjected to MS analysis on a LCT Premier XE (Waters) after elution through a C18 column (100 mm x 2.1 mm, 2μ particles size, Waters). The MS data were analysed using the MassLynx^{TM} v4.0 software (Waters).

Amino acid Acetylation Reaction

All samples for acetylation had a 170 μL final volume. A blank sample was composed of 20 μL $NaHCO_3$ in 150 μL water. The control reaction sample contained 20 μL $NaHCO_3$, 50 μL water and 100 μL reagent solution. The amino acid samples contained 20 μL $NaHCO_3$, 50 μL amino acid solution (38 μM) and 100 μL reagent solution. After adding each component, the samples containing 11 μM final amino acid concentration were mixed for 1 min and then heated at 40 °C for 1 h. The samples were then quenched with 830 μL aqueous HCl (0.1 M) and 10 μL injected onto the HPLC-MS instrument as such.

1.5.8 Enantiomeric Separation Using Analytical HPLC

HPLC Analysis of samples derivatised with Marfey's Reagent

HPLC-MS analysis of Marfey's reagent derivatised compounds was performed on a LCT Premier XE (Waters) after elution through a C18 column (100 mm x 2.1 mm, 2μ particles size, Waters). The MS data was analysed using the MassLynx^{TM} v4.0 software (Waters).

HPLC Analysis of samples derivatised with 6-Aminoquinolyl-*N*-hydroxysuccinimidyl Carbamate

The HPLC-UV analysis of AccQ-Tag derivatised compounds was performed using a ChiraSpher NT column, as well as ChiroBiotic V or ChiroBiotic T columns, on an Agilent HPLC binary pump system. The UV absorbance was monitored at 265 nm. The samples (10 μL) were eluted using a flow rate of 1 mL min^{-1} with the following gradients: isocratic elution using A: 20 mM NH_4OOCH or NH_4OOCCH_3 pH 6.5 and B: methanol; 20% B in A for 40 min. Shallow gradient using A: 20mM NH_4OOCH or NH_4OOCCH_3 pH 6.5 and B: methanol or acetonitrile; 2% B in A for 5 min followed by 2–20% B in A over 40 min. Steep gradient using A: 20 mM NH_4OOCH or NH_4OOCCH_3 pH 6.5, B: methanol or acetonitrile; 2% B in A for 5 min followed by 2–20% B in A over 20 min. Shallow acidic gradient using A: 0.1% v/v HCOOH or CH_3COOH in H_2O pH 2.5 and B: methanol or acetonitrile; 2% B in A for 5 min followed by 2–20% B in A over 60 min. After the end of the gradient, the mobile phase was set to reach 98% B over 5 min. The column was washed with 98% B for 5 min and re-equilibrated at 2% B for 5 min before the next sample was injected.

LC-MS Analysis of Acetylated Samples

LC-MS analysis of AccQ-Tag derivatised and acetylated compounds was performed using ChiroBiotic V and ChiroBiotic T columns (250 mm x 4.6 mm, 5 μm pore size) on a Waters 1525m Binary HPLC Pump system with a Waters 2777 Sample Manager coupled to a Micromass® Quattro micro^{TM} API mass spectrometer using the positive electrospray ionisation mode. Eluent A was 0.1% v/v HCOOH in water pH 2.5, eluent B was 0.1% v/v HCOOH in acetonitrile. The samples (10 μL) were eluted using a flow rate of 1 mL min^{-1} with 2% eluent over 5 min, followed by a gradient of

2–98%, 2–50% or 2–20% B over 40 min, and then to 98% B over 5 min. The column was washed with 98% B for 5 min and the column was re-equilibrated at 2% B for 5 min before the next sample was injected.

Chapter 2

Carbapenem Antibiotic Biosynthesis

2.1 The Carbapenems - An Important Class of β-Lactam Antibiotics

2.1.1 General Introduction to β-Lactam Antibacterial Agents

One of the most significant improvements in the health of the world's population has been the development of antibacterial drugs and in particular that of β-lactam antibiotics [1, 2]. The β-lactam antibiotics can be divided into several main classes, depending on the structure of their nuclei. There are two main families of monocyclic β-lactams, the *N*-sulfonated monobactams [3] and the *N*-alkylated nocardicins, chlorocardicins and formadicins (Figure 2.1 A). Bicyclic β-lactams can contain either a 5-membered (Figure 2.1 B) or a 6-membered ring (Figure 2.1 C) that is fused to their bicyclic β-lactam ring. The penams, penems and cephems contain a sulphur atom in their bicyclic nucleus, while the clavams, oxapenems and oxacephems contain an oxygen atom (Figure 2.1).

β-Lactam antibiotics remain amongst the most potent and useful treatments against bacterial infections. The bactericidal activity of penicillins was first observed in the late 1920s [4] and the chemical structure of the penam core was elucidated in the early 1940s. Although the total chemical synthesis of penicillin G was described some years later [5], production by fermentation remains the most cost-effective method for producing clinically used penicillins. However, because of the emergence and fast spread of resistance mechanisms amongst pathogenic bacteria, there has been a constant need for novel antibiotics. While finding a cure for new infectious agents was an initial driver

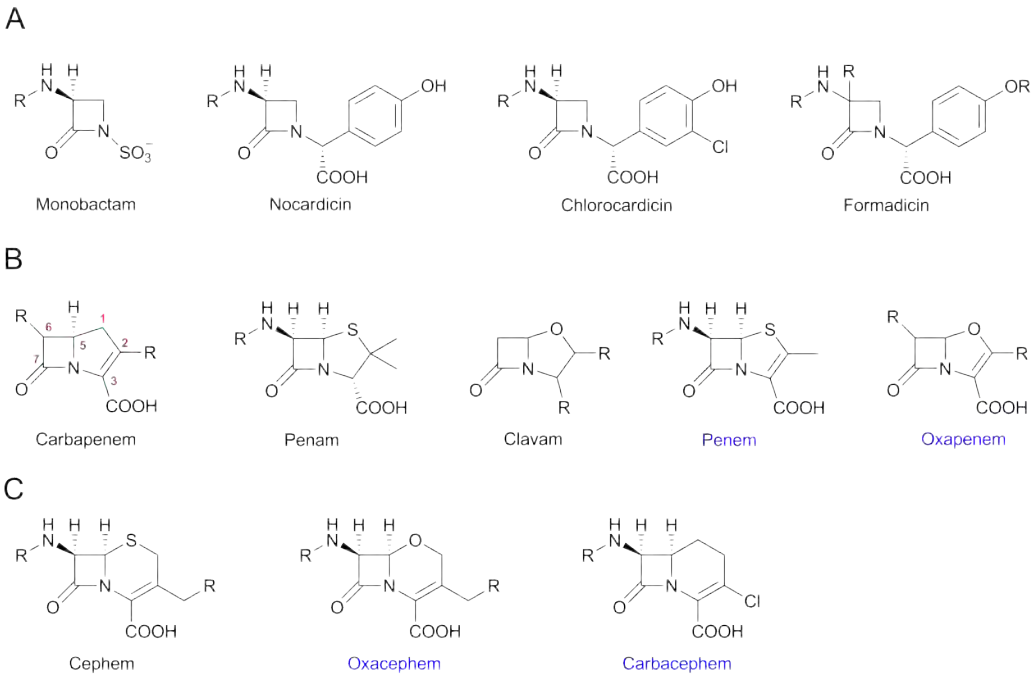


Figure 2.1: Structures for the main classes of β-Lactam antibiotics. The β-lactam nuclei found in clinically used natural and synthetic antibiotics. Carbon atom numbering is shown in red for the carbapenem nucleus based on the assumption that carbapenems are dethia-1-carbapen-2-em derivatives of penicillin. Another nomenclature is based on the IUPAC recommendations which identify carbapenems as 7-oxo-1-azabicyclo[3.2.0]hept-2-ene-2-carboxylate derivatives. Classes of β-lactams which do not have naturally occurring members are shown in blue. (A) Examples of monocyclic β-lactam nuclei; (B) 5-Membered bicyclic β-lactam nuclei; (C) 6-Membered β-lactam bicyclic nuclei.

for antibacterial research, much current effort is now directed towards keeping up with increasingly resistant known pathogens [6].

The consensus antibacterial mode of action of β-lactam antibiotics, including that of the carbapenems, is based on their ability to inhibit the bacterial cell wall biosynthesis and repair machinery. More specifically, most β-lactam antibacterial agents inhibit the transpeptidase activity of penicillin binding proteins (PBP) responsible for cross-linking bacterial cell wall components. The hypothesis that β-lactam antibiotics are PBP substrate analogues is based on a structural similarity between the β-lactam ring of the antibiotics and the D-alanine-D-alanine fragments of the transpeptidase substrate [7]. By preventing growth and repair of the cell wall peptidoglycan, a fundamental structure only present in prokaryotes, β-lactam antibiotics have the advantage of having a specific activity against bacterial species.

Pathogenic bacteria have acquired or developed resistance mechanisms for all known antibacterial agents. At least three mechanisms contribute to bacterial resistance to β-lactam antibiotics, i.e. active drug efflux, target modification, and degradation. The β-lactam antibiotics are inactivated by β-lactamase enzymes, a resistance mechanism that is increasingly threatening their antibacterial efficacy [8]. Two different families of β-lactamases have evolved from two distinct ancestors: the serine-β-lactamases and the metallo-β-lactamases. Although they are called "carbapenemases", many of the β-lactamases responsible for the resistance to carbapenems inactivate most of the hydrolysable β-lactam antibacterial agents. In order to retain antibacterial efficacy, β-lactam antibiotics are often used in combination with β-lactamase inhibitors such as sulbactam, tazobactam and the natural product clavulanic acid.

2.1.2 The Discovery and Development of Carbapenems

Carbapenems belong to a sub-family of β-lactam antibiotics that were discovered half a century after penicillins, as part of a quest for β-lactams with improved antibacterial activity or β-lactamase inhibitory properties. The carbapenems differ from the penicillins and cephalosporins by the absence of a heteroatom in their "carbon only" 5-membered ring fused to the β-lactam ring. The carbapenem carbon atom numbering that is used throughout the work described in the thesis is based

on the assumption that carbapenems are dethia-1-carbapen-2-em derivatives of penicillin (Figure 2.1 B). Carbapenems are also desaturated between C-2 and C-3. Most natural β-lactams have been isolated from fermentation broths of actinomycetes. More specifically, the majority of the carbapenems identified so far have been isolated from *Streptomyceteae* ssp. with two exceptions from other prokaryotic species. *Pectobacterium carotovorum* is the producer of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** (Figure 2.4 and Figure 2.2 A) and *Kitasatosporia papulosa* of AB-110-D, the only carbapenem with a *Z*-geometry at its C-2 side chain (Figure 2.3 A).

Thienamycin **116** (Figure 2.4 and Figure 2.2 D), named after its β-thioenamine functionality, was isolated and characterised in 1976 and was described as one of the most potent broad-spectrum natural antibiotics [9]. Thienamycin **116** was isolated from cultures of soil bacteria *Streptomyces cattleya* and *Streptomyces penemifaciens* together with *N*-acetylthienamycin **117** and *N*-acetyldehydrothienamycin **118** (Figure 2.2 C). The total synthesis and the determination of the absolute configuration of thienamycin **116** were published in the following years [10, 11]. In the same period, the structurally related olivanic acids (*e.g.* MM4550, Figure 2.3 A) were isolated from a culture of *Streptomyces olivaceus* [12, 13, 14]. Since the discovery of thienamycins and olivanic acids, the carbapenem family of natural secondary metabolites was found to contain more than 40 different structures, including the simplest possible carbapenem (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**, carpetimycins, asparenomycins, pluracidomycins and the PS series of carbapenems. A pharmacological survey of the carbapenems has been compiled by Wise [15].

Despite the exceptional potency and antibacterial spectrum of thienamycin **116**, there were two problems that ruled out its use in the clinic. First, thienamycin **116** was found to have a low chemical stability in solution: it has been proposed that this concentration-dependent instability is due to intramolecular opening of the β-lactam by the amine group of the C-2 side-chain [16]. This issue was solved through the *N*-formimidoyl capping of the cysteamine side chain to yield clinically used carbapenems *e.g.* imipenem **119** [17]. Second, the β-lactam ring of both thienamycin **116** and imipenem **119** was found to be hydrolysed *in vivo* by a human renal enzyme, dehydropeptidase-1 (DHP-1) [18]. In order to prevent this metabolic degradation, imipenem has since been co-administered with cilastatin, an inhibitor of DHP-1. The development of the imipenem-cilastatin

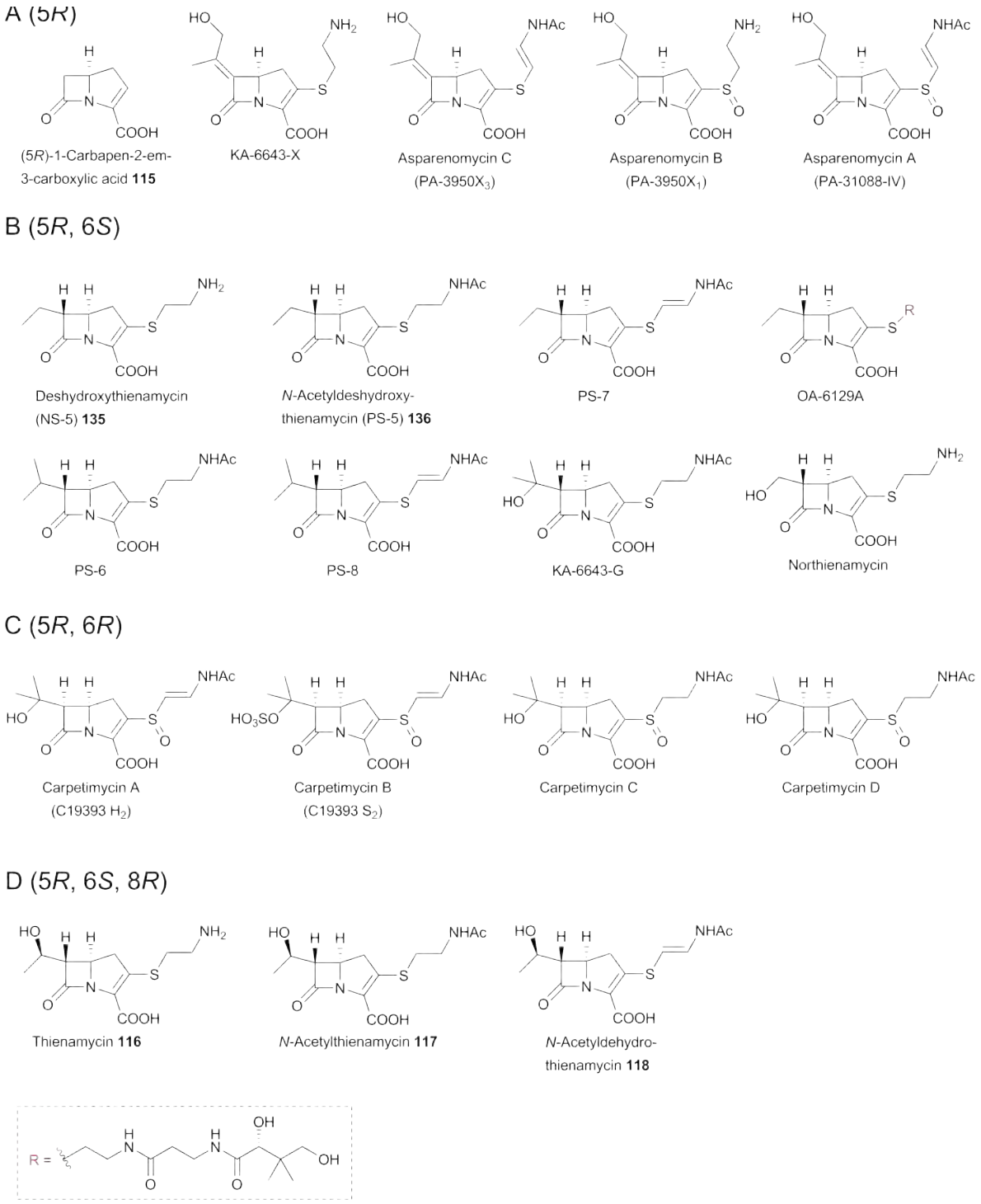


Figure 2.2: Structures of naturally occurring carbapenems. The C-6 pantethenate side chain found in the OA-6129 carbapenem family is highlighted in a dotted box. (A) All the carbapenems, including compounds with only one chiral center, share the (5*R*) sterochemistry. (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** is produced by P. carotovorum, while the C-5–C-6 unsaturated asparenomycins are produced by *S. tokunonesis* and *S. argenteolus*; (B) Carbapenems having a (5*R*,6*S*) stereochemistry; (C) Carpetimycins produced by *Streptomyces* sp. KC-6643 all share the (5*R*,6*R*) stereochemistry; (D) The first isolated carbapenem, thienamycin **116**, as well as the *N*-acetyl derivative **117** and the *N*-acetyldehydrothienamycin **118** (PS-5) all share the (5*R*,6*S*,8*R*) stereochemistry.

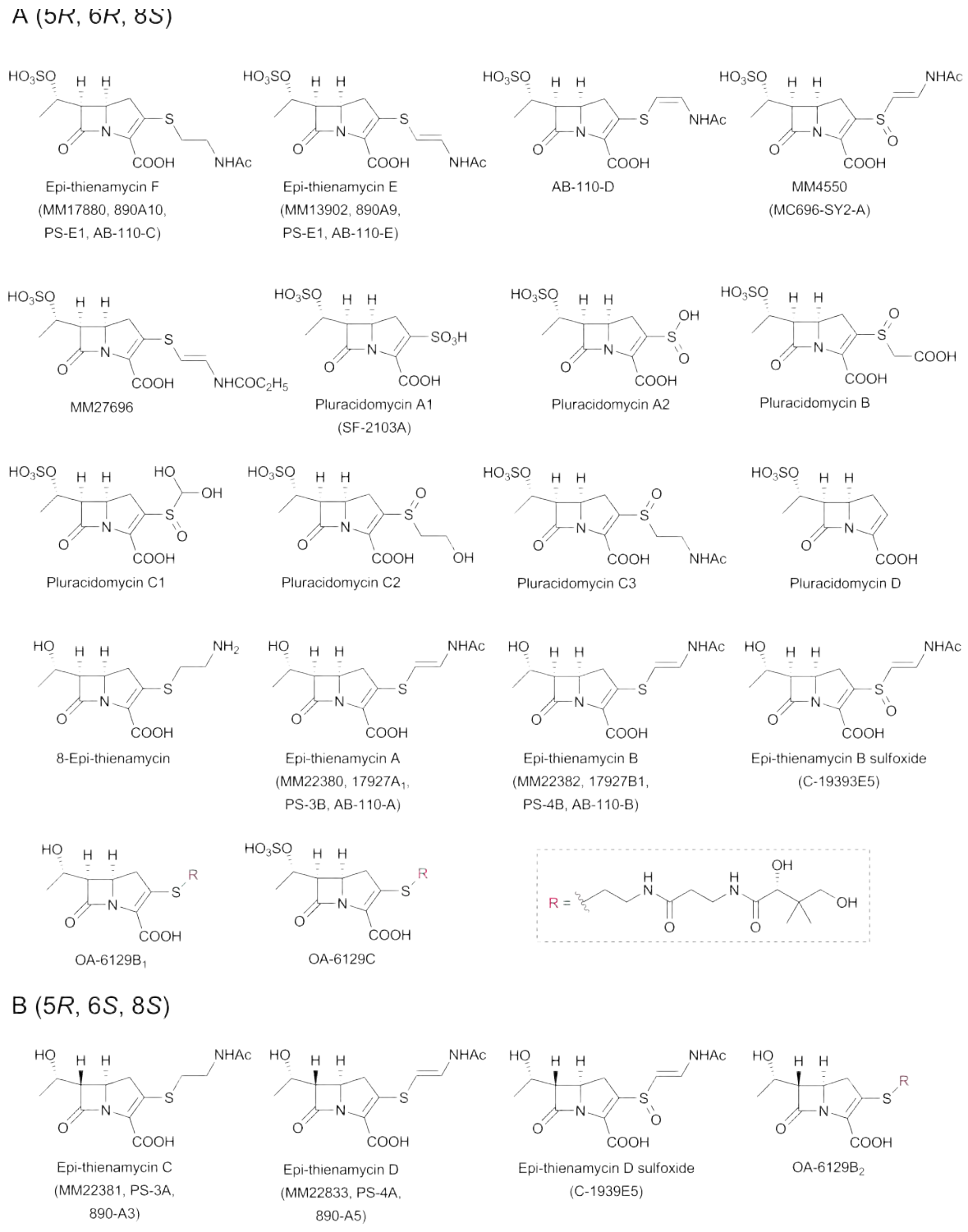


Figure 2.3: Structures of naturally occurring carbapenems (continued). The C-6 pantethenate side chain found in the OA-6129 carbapenem family is highlighted in a dotted box. (A) The largest sub-family of carbapenems share the (5*R*,6*R*,8*S*) stereochemistry; (B) Carbapenems sharing the (5*R*,6*S*,8*S*) stereochemistry.

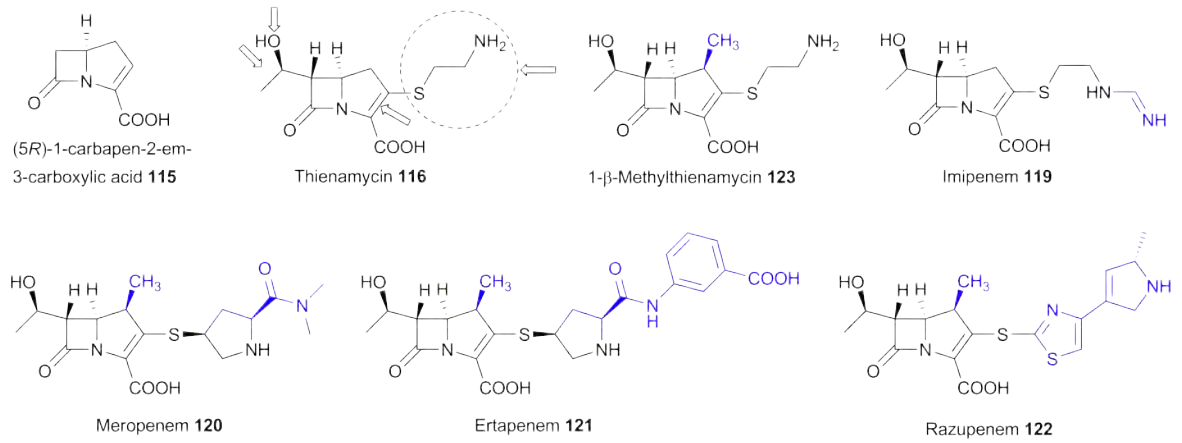


Figure 2.4: Structures of natural and synthetic carbapenems. The chemical structures of natural products (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** and thienamycin **116** together with that of some synthetic analogues, 1-β-methyl thienamycin **123**, imipenem **119**, meropenem **120**, ertapenem **121** and razupenem **122**. The functionally important features of carbapenems that are described in the text are highlighted with arrows or dotted circles. The unnatural functionalities added in synthetic carbapenems are highlighted in blue.

combination has been reviewed in Kahan *et al.* [19].

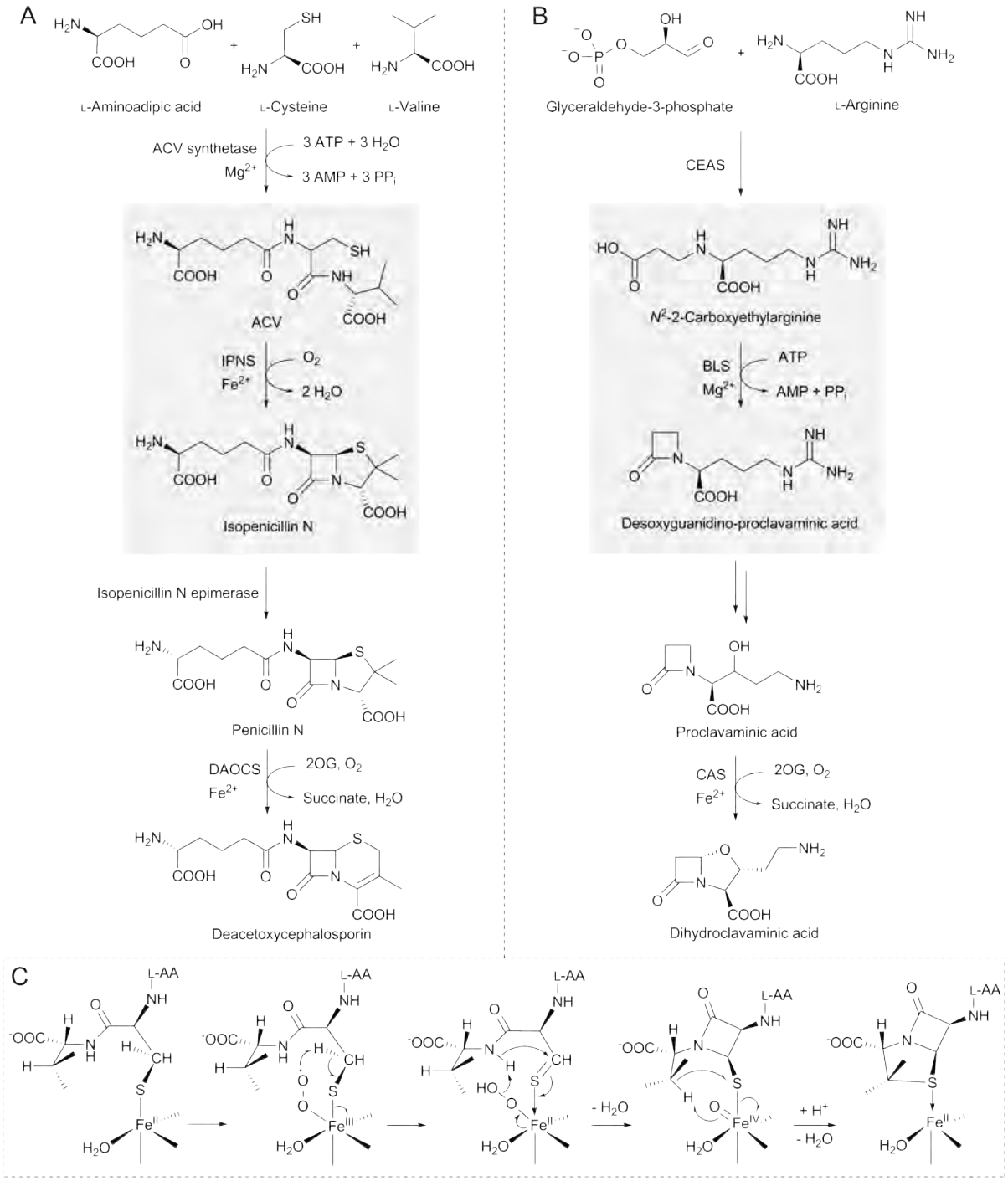
As an alternative to the imipenem-cilastatin combination, several generations of thienamycin derivatives have been produced, *e.g.* meropenem **120**, ertapenem **121** and razupenem **122** (Figure 2.4). The structure of these synthetic thienamycin derivatives is derived from the observation that a β-methyl substituent at the position C-1 of the carbapenem core of 1-β-methylthienamycin **123** provides resistance to DHP-1 degradation and superior antibacterial activity [20, 21, 22]. The stereochemistry at this center was shown to be important and α-methylcarbapenem isomers retain the reduced DHP-1 susceptibility, but not the antibacterial activity [20]. Thienamycin **116** displays lowered susceptibility to most serine β-lactamases when compared to penicillins. Synthetic carbapenem antibiotics derived from thienamycin are therefore often used to treat serious infections including those involving multi-drug resistant bacteria [23]. However β-lactamases, and in particular the metallo β-lactamases, that can catalyse carbapenem hydrolysis have been described [24]. In contrast to many β-lactam antibiotics, including the penicillins and cephalosporins which are obtained by fermentation or from fermented precursors, clinically useful carbapenems are solely produced by chemical synthesis. There is therefore an interest in understanding their biosynthesis in order to further develop biocatalytic or fermentation routes and produce useful antibiotics or intermediates in a sustainable manner.

All carbapenems share the (5*R*)-stereochemistry, including the simplest carbapenem (5*R*)-1-

carbapen-2-em-3-carboxylic acid **115** (Figure 2.4 and Figure 2.2 A), but differ in the way they are substituted at their C-2 and C-6 positions. Carbapenems also display various stereochemistries at C6 and C8. The C-6 side chain can be methyl, ethyl, or isopropyl, that can be saturated, desaturated, hydroxylated, or sulphated (Figure 2.2 and Figure 2.3). The thioether-linked C-2 side chain displays large structural variations. Collectively, the natural carbapenems isolated from various organisms demonstrate the power of biosynthesis to generate a broad range of compounds with a common template.

2.1.3 β-Lactam Antibiotics Biosynthesis - General Considerations

Similar to most prokaryotic secondary metabolites, the genes encoding for the biosynthesis of β-lactam antibiotics are found clustered in the genome of the producers. Two of such clusters have been associated with the biosynthesis of a carbapenem antibiotic and have been sequenced (Figure 2.5). Bioinformatic analysis and biochemical characterisation of the polypeptides encoded by the genes present in these clusters showed limited homology with other β-lactam antibiotics biosynthetic enzymes, suggesting that carbapenem natural products are produced using a unique strategy that differs significantly from sulphur-containing bicyclic β-lactams such as penams and cephems, or oxygen-containing clavams [26]. A central step in penam and cephem biosynthesis is the oxidative formation of the penam bicyclic system in a single step from a linear δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine (ACV) precursor by isopenicillin N synthase (IPNS, Scheme 2.1 A) [27]. An outline mechanism for IPNS is shown in Scheme 2.1 C. In contrast, the biosynthesis of the clavam bicyclic nucleus found in clavulanic acid requires two enzymatic steps [28]. In clavam biosynthesis, the β-lactam ring is formed by the action of BLS, a β-lactam synthetase (β-LS) enzyme, using an ATP-dependent strategy (Scheme 2.1 B). The fused five-membered ring is then formed *via* an enzymatic oxidative C-O bond formation catalysed by clavaminic acid synthase (CAS). The bicyclic carbapenem nucleus is formed following yet another strategy: the 5-membered ring is constructed first by the carboxymethylproline synthase (CMPS) enzyme (Scheme 2.2, Section 2.2) [29]. The β-lactam is then formed by CarA, a β-LS enzyme, using an ATP-dependent strategy, analogous to that of BLS in clavulanic acid biosynthesis [30]. In addition to the β-lactam ring formation, other steps in the biosynthesis of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** and thienamycin **116** biosynthesis have been studied at a molecular level and the results of these studies are summarised in Section 2.2



Scheme 2.1: Biosynthetic strategies for the production of β-lactam secondary metabolites [25]. The formation of the β-lactam ring in penicillin and calvulanic acid biosynthesis (boxed in gray) proceeds in two distinct mechanisms. (A) The biosynthesis of penicillins and cephalosporins starts from the ACV tripeptide precursor. IPNS is a non-heme Fe(II)-dependent enzyme that catalyses the oxidative formation of the penicillin bicyclic nucleus in one step. Following epimerisation of isopenicillin N, deacetoxycephalosporin synthase (DAOCS) catalyses the oxidative ring expansion to yield the cephalosporin nucleus. ACV, δ-(L-α-aminoadipoyl)-L-cysteinyl-D-valine; IPNS, isopenicillin N synthase; DAOCS, deacetoxycephalosporin C synthase; (B) Schematic representation of key steps in the clavulanic acid biosynthetic pathway. CEAS, carboxyethylarginine synthase; BLS, β-lactam synthetase; CAS, clavaminic acid synthase; (C) Outline mechanism for the penicillin nucleus formation catalysed by IPNS. This figure was modified from reference [25].

and Section 2.3.

2.2 (5*R*)-1-Carbapen-2-em-3-carboxylic Acid Biosynthesis

(5*R*)-1-Carbapen-2-em-3-carboxylic acid **115** (Scheme 2.2) is the simplest member of the carbapenem family of β-lactam antibiotics; it has been isolated from the plant pathogens *P. carotovorum* (formerly *Erwinia carotovora*) and *Serratia marcescens* [31]. The biosynthetic pathway leading to (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** has been studied in an effort to unveil the biochemical events necessary for the production of the carbapenem nucleus. These results made a crucial step towards the understanding of the early biosynthesis pathway of C-2 and C-6 substituted carbapenems biosynthesis, *e.g.* thienamycin. Aspects of the biosynthesis of carbapenems have been reviewed before by Williamson in 1986 [32] as well as from a genetic perspective by Coulthurst, Barnard and Salmond in 2005 [33].

2.2.1 Labelling Studies

Pioneering labelling studies suggested that L-glutamate **124** and acetate are precursors for the biosynthesis of (5*S*)-1-carbapen-2-em-3-carboxylic acid **115**. Bycroft and Maslen used L-[^{14}C]-glutamate to feed *Serratia sp.* (ATCC 39006) cultures and showed substantial incorporation into (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** [34]. In addition, the incorporation of two carbon atoms derived from an intact [^{13}C]-acetate unit into positions 6 and 7 of the β-lactam ring of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** was observed [35]. The same authors also identified (3*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** and (3*S*,5*R*)-1-carbapenam-3-carboxylic acid **126** as possible biosynthetic intermediates, although they initially assigned the incorrect (3*R*,5*S*) stereochemistry to **126** [35]. Subsequent work by Bycroft and coworkers rectified this assignment [36], and Townsend and coworkers unambiguously confirmed the corrected absolute configurations of the natural carbapenams **125** and **126**, as well as the carbapenem **115** using synthetic standards [37]. Several reports describe the synthesis of (3*S*,5*S*)-1-carbapenam-3-carboxylic acid **125**, (3*S*,5*R*)-1-carbapenam-3-carboxylic acid **126** or the unnatural (3*R*)-isomers [37, 38, 39, 40].

2.2.2 (5*R*)-1-Carbapen-2-em-3-carboxylic Acid Biosynthetic Genes

The activities of only three enzymes are strictly required for the biosynthesis of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** from primary metabolites: CarB, CarA and CarC (Scheme 2.2) [42]. In

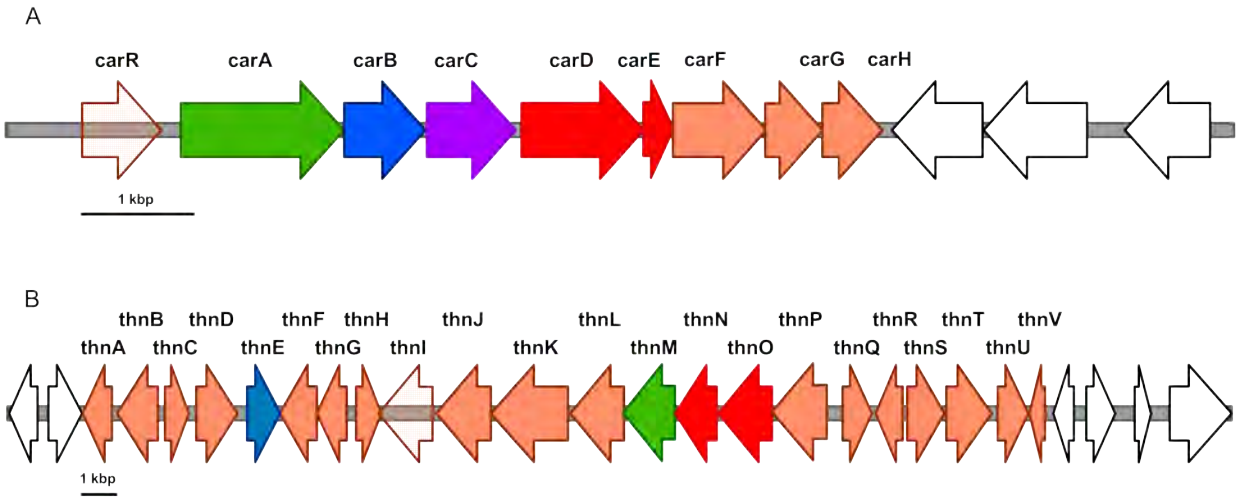


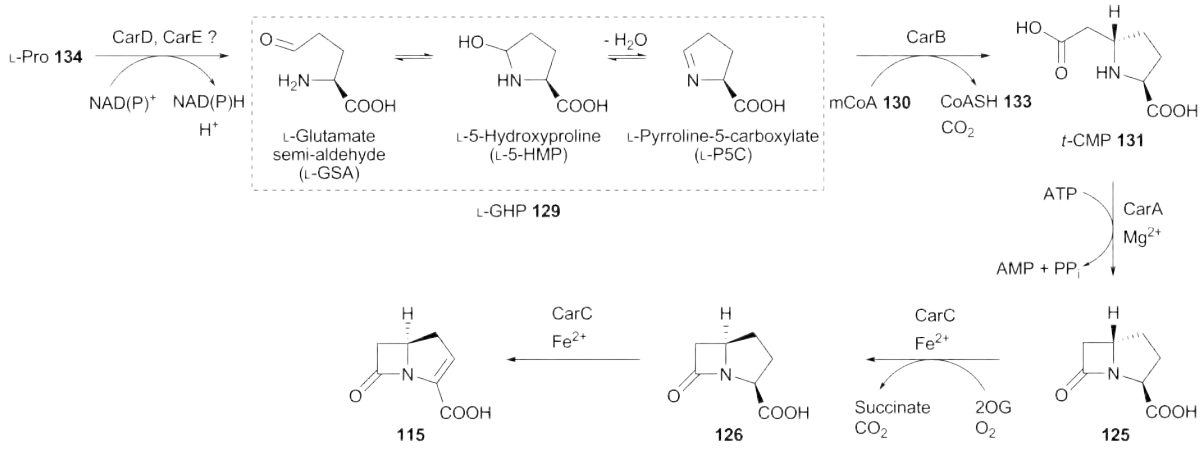
Figure 2.5: Schematic organisation of the *P. carotovorum* (5*S*)1-carbapen-2-em-3-carboxylic acid and *S. cattleya* thienamycin biosynthesis gene clusters [41, 26]. Coding sequences are indicated by arrows; the CMPS genes *carB* and *thnE* are indicated in blue; the β-LS genes *carA* and *thnM* are in red; the Fe(II) and 2OG-dependent epimerase/desaturase *carC* gene is in purple and has no clear homologue in the thienamycin cluster; the pairs of genes proposed to code for the L-GHP production are in red; regulatory genes are in transparent beige; coding sequences outside the proposed boundaries of the clusters are in white. (A) The deduced biosynthetic gene cluster for (5*S*)-1-carbapen-2-em-3-carboxylic acid **115** in *P. carotovorum*; (B) The deduced gene cluster for thienamycin in *S. cattleya*. This figure was generated using Vector NTI software.

P. carotovorum, these three proteins were found within a cluster of nine genes (*carRABCDEFGH*, Figure 2.5 A) which was identified by mutational analysis [41]. Bioinformatic analysis of the coding sequences provided preliminary functions for the products of these genes (Table 2.1) [43]. Deletion mutants of the first five genes resulted in abolished (*carABC*) or reduced (*carDE*) production of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**, while deletion of the last three (*carFGH*) led to suicidal lethality, pointing towards their involvement in a resistance mechanism [41, 43]. The proposed role of each of the corresponding proteins is described in the following paragraphs. Using the sequence of the *P. carotovorum* cluster, searches in all available bacterial genomes identified homologous clusters. For example, the cluster found in the insect pathogen *Photorhabdus luminescens* (formerly *Xenorhabdus luminescens*) [44] contains eight genes (*cpmABCDEFGH*), each one of them having homology to one *P. carotovorum* gene.

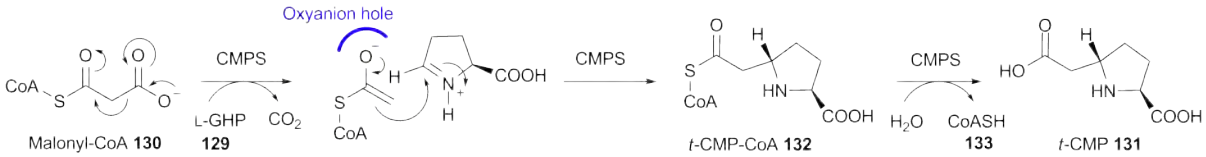
CarB

Because experiments using radiolabelled substrates suggested that acetyl-CoA **127** and L-

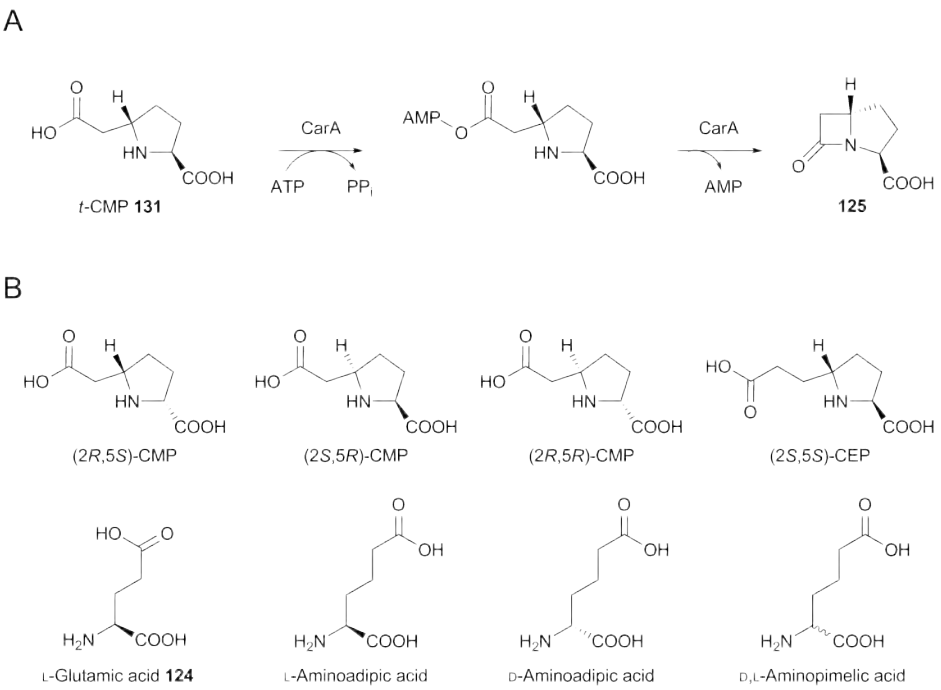
glutamate **124** were the precursors for the carbapenem nucleus, activated forms of L-glutamate, *e.g.* glutamate-5-phosphate **128** or L-GHP **129**, were investigated as substrates for the CarB-catalysed reaction [45]. L-GHP **129** is the acronym used throughout this work to designate the equilibrium mixture of L-glutamate semialdehyde (L-GSA), 5-hydroxy-L-proline (L-5HP) and L-pyrroline-5-carboxylate (L-P5C) (shown in a dotted box in Scheme 2.2) [29, 46]. When acetyl-CoA **127** was tested as a CarB substrate in the presence of L-GHP, no product was observed. In contrast, the incubation of malonyl-CoA **130** with CarB in the presence of L-GHP **129** resulted in the production of *t*-CMP **131** [29]. The stereochemistry of the product was assigned based on comparison of the circular dichroism (CD) spectrum with that of synthetic standards [29]. In agreement with the labelling studies showing incorporation of the acetyl moiety of acetyl-CoA **127** [35], the observation that malonyl-CoA **130**, and not acetyl-CoA **127**, is the CarB substrate could be explained by the *in vivo* activity of the biotin-dependent acetyl-CoA carboxylase (ACC) (EC 6.4.1.2) that produces malonyl-CoA **130** from acetyl-CoA **127**. The formation of *t*-CMP **131** from L-GHP **129** and malonyl-CoA **130** implies that CarB can catalyse three distinct steps [29]. Decarboxylative formation of an acetyl-CoA enolate from malonyl-CoA **130** is followed by C–C bond formation between the resulting enolate and L-GHP **129** to yield an acyl-CoA intermediate **132**. Finally, enzyme-catalysed thioester hydrolysis yields coenzyme A **133** (CoASH) and the (2*S*,5*S*)-isomer of *t*-CMP **131** (Scheme 2.3). Because CMPS is the main focus of the following chapters, its function and mechanism are described in detail in Section 2.4.



Scheme 2.2: General scheme for the biosynthesis of (5*R*)-1-carbapen-2-em-3-carboxylic acid. Three enzymes only are strictly required for the formation of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** from primary metabolites: CarB, CarA and CarC. mCoA, malonyl-coenzyme A.



Scheme 2.3: CMPS enzymes catalyse three successive reactions. Carboxymethylproline synthase first promotes malonyl-CoA **130** decarboxylation and stabilises the intermediate enolate in an oxyanion hole (represented in blue). In the presence of the second substrate, L-GHP **129**, CMPS catalyses the formation of a C–C bond to give an acyl-CoA intermediate **132**. Finally the enzyme hydrolyses the thioester bond to release CoASH **133** and *t*-CMP **131**. In the absence of L-GHP **129**, the enolate gets protonated to give acetyl-CoA **127** which ultimately gets hydrolysed to CoASH **133** and acetate.

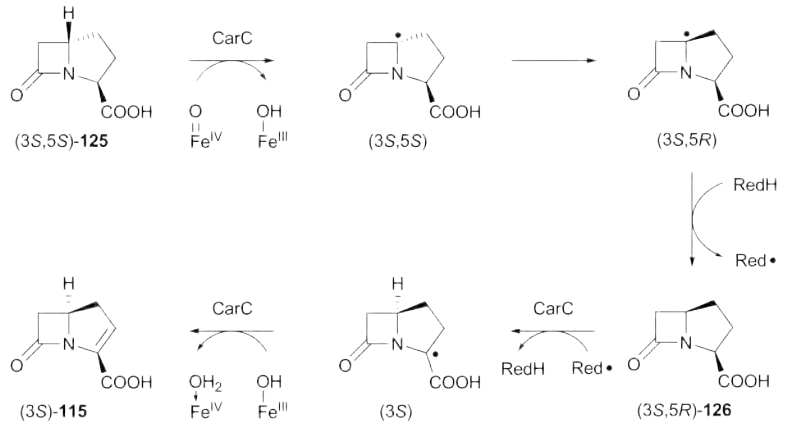


Scheme 2.4: CarA-catalysed β-lactam formation. (A) CarA catalyses the Mg(II) and ATP-dependent cyclisation of *t*-CMP **131** into (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** *via* an adenylate intermediate; (B) In addition to *t*-CMP **131**, several substrate analogues have been shown to be CarA substrates, demonstrating the promiscuity of this β-LS enzyme.

CarA

CarA is a β-lactam synthetase (β-LS) that catalyses the Mg(II) and ATP-dependent cyclisation of (2*S*,5*S*)-5-carboxymethylproline **131** (*t*-CMP) into (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** (Scheme 2.4 A) [47, 30]. CarA is evolutionary related to BLS, an enzyme involved in the formation of the monocyclic β-lactam of clavulanic acid (Scheme 2.1 B) [47]. Both CarA and BLS share homology with asparagine synthetase, another Mg(II) and ATP-dependent enzyme [48]. A crystal structure of CarA has been reported by Miller *et al.* [49] and comparison with the BLS structure further support mechanistic similarities [50, 51].

Townsend and coworkers have demonstrated that CarA is a promiscuous enzyme, showing very little decrease in turnover number when all four stereoisomers of *t*-CMP **131** were used [30]. In addition, other alternate substrates were accepted by CarA, *e.g.* incubation in the presence of L-glutamic acid yielded the five-membered lactam L-pyroglutamic acid (Scheme 2.4 B). Attempts to exploit this substrate promiscuity for biocatalytic purposes are described in Chapter 4.



Scheme 2.5: Epimerisation and desaturation reactions catalysed by CarC. The proposed mechanism involves the formation of an Fe(IV)-oxo species through an oxidative decarboxylation of the cosubstrate 2OG.

CarC

CarC is an Fe(II) and 2-oxoglutarate (2OG)-dependent oxygenase that catalyses the epimerisation and desaturation of (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** to give first (2*S*,5*R*)-1-carbapenam-3-carboxylic acid **126** and then (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** (Scheme 2.5) [52]. Both the epimerisation and desaturation steps are required for antibacterial activity. In a similar way to CarB, the observation that CarC can catalyse two different chemical transformations in a stepwise manner may explain the requirement for only three enzymes in the process of building an active biomolecule from primary metabolites [37].

Sequence analysis indicated that CarC shows homology to clavaminic acid synthase (CAS) (Scheme 2.1 B) [53]. Although the CarC-catalysed reaction is reminiscent of the oxidation and desaturation steps catalysed by CAS, CarC catalysis only involves one single two-electron oxidation, while CAS catalyses two two-electron oxidations. Thus, in theory a single catalytic cycle is needed to carry out the two CarC-catalysed transformations, consuming only one equivalent of 2OG cofactor. Various possible mechanisms for the reactions catalysed by CarC have been proposed [52]. Although no consensus understanding of the CarC mechanism at an atomic level has been reached, one proposition was brought forward based on labelling and computational studies (Scheme 2.5) [54, 55], together with structural information obtained from a CarC crystal structure in complex with the Fe(II) and 2OG **21** cofactors, as well as the substrate analogue L-*N*-acetylproline [52]. Most Fe(II) and

2OG-dependent oxygenases share a common mechanism involving the initial activation of the metal center with molecular oxygen. This mechanism is described in more detail in Chapter 1, for another Fe(II) and 2OG-dependent oxygenase: human γ-butyrobetaine hydroxylase (hBBOX) (Section 1.1.5, Scheme 1.4). In brief, the oxidative decarboxylation of the 2OG cosubstrate generates a reactive Fe(IV)-oxo species which can then abstract a hydrogen radical from the substrate [56]. In the case of CarC, the C-5 bridgehead hydrogen of the substrate is probably abstracted by the Fe(IV)-oxo species, consistent with loss of label at that position (Scheme 2.5) [37]. Thermodynamically favorable epimerisation at this stage, followed by oxidation across C-2/C-3 involving a hydrogen radical donor (RedH in Scheme 2.5, possibly a tyrosine residue), could yield the carbapenem **115** [55].

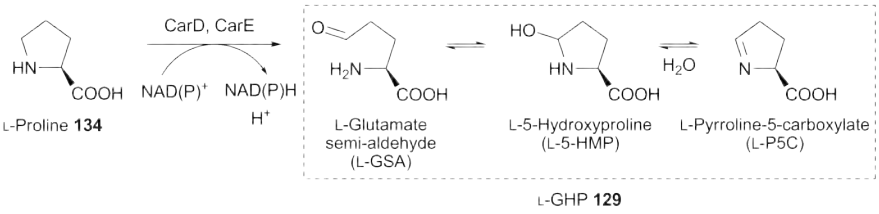
The chemical synthesis of the three unnatural diastereomers of the CarA product (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** enabled testing test them as CarC substrates [38]. As for CarA, CarC was found to be a promiscuous enzyme accepting at least three of the four isomers, resulting in the formation of one single product (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**. This observation suggests that the epimerisation step is not crucial for the desaturation step.

CarD and CarE

The functions of CarD and CarE has not been studied experimentally. CarD shows sequence similarity to proline dehydrogenase enzymes, while CarE shows sequence similarities to ferrodoxins, a family of enzymes that contain an iron-sulfur cluster allowing them to mediate electron transfer [41]. CarD and CarE are proposed to work in tandem in the biosynthesis of the CMPS substrate L-GHP **129** from L-proline **134** (Scheme 2.6 A) [41]. Because L-P5C is a primary metabolite, CarD and CarE are probably not strictly required for (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** biosynthesis, in agreement with mutagenesis studies [43].

CarR

In *P. carotovorum*, the transcription of the (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** biosynthetic genes is positively regulated by the LuxR-type activator protein, CarR, which is a *N*-acyl homoserine lactone receptor [57]. Activation of CarR is dependent upon *N*-(3-oxohexanoyl)-L-



Scheme 2.6: Proposed role for CarD and CarE. Based on bioinformatic analyses, CarD and CarE are proposed to produce L-GHP from L-proline

homoserine lactone (OHHL, Figure 2.6) synthesis by the LuxI homologue CarI. CarI is encoded by the *carI* gene which is situated outside the boundaries of the (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** biosynthesis gene cluster [58].

In *Serratia* sp. the transcriptional activation of the (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** biosynthetic genes is dependent on the synthesis of *N*-butanoyl-L-homoserine lactone (BHL, Figure 2.6) and *N*-hexanoyl-L-homoserine lactone (HHL, Figure 2.6) by SmaI. In comparison, no CarR homologue could be found in the *P. luminescens* cluster [44], suggesting a different regulation mechanism that is still to be identified. The transcriptional regulation of carbapenem biosynthesis has been reviewed in McGowan *et al.* [59, 33].

CarF, CarG and CarH

The products of the *carF* and *carG* genes have *N*-terminal signal peptides that are proposed to mediate the extracellular export of the polypeptides [43]. CarF and CarG are potentially translationally coupled and are proposed to be involved in an unprecedented β-lactam resistance mechanism, based on the absence of homologues in the NCBI Protein database and the observation that their deletion from the *P. carotovorum* chromosome results in a carbapenem-sensitive phenotype and suicidal lethality [43]. In addition, when a carbapenem-sensitive *E. coli* strain was transformed with a plasmid containing the *carRA-I* genes, it not only gained the ability to produce (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125**, but also became resistant to (2*S*,5*S*)-1-carbapenam-3-carboxylic acid **125** and not to other carbapenems [43].

The function of CarH is unknown and deletion of the corresponding gene from the *P. carotovorum* genome showed no effect on the production of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**, nor

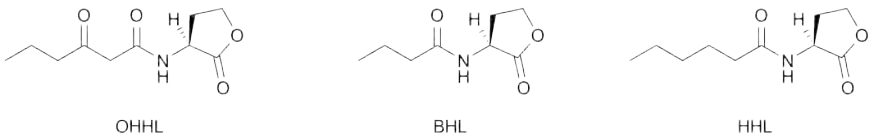


Figure 2.6: Structures of L-homoserine lactones. OHHL, *N*-(3-oxohexanoyl)-L-homoserine lactone; BHL, *N*-butanoyl-L-homoserine lactone; HHL, *N*-hexanoyl-L-homoserine lactone.

on *P. carotovorum* resistance to compound **115**. CarH has no homologue in the NCBI Protein database but two observations seem to indicate that it may be involved in the production of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**. First, the position of the *carH* gene suggests that it might be both co-transcribed and translationally coupled with other genes in the cluster. Second, the *carH* gene was identified within a near identical cluster in the (5*R*)-1-carbapen-2-em-3-carboxylic acid producer *Serratia marcesens* while the DNA sequence outside the cluster boundaries was different. These observations suggest that the carbapenem biosynthesis cluster was probably acquired in its entirety by *P. carotovorum* and *S. marcesens* and that a functional copy of *carH* was maintained during the process strongly suggesting its relevance [43]. It is notable that CarF-H homologues are found clustered together in *Pectobacterium wasabiae* (Table 2.1), a species for which the genome has been sequenced (NCBI accession number CP001790) but which does not contain any CarA-E homologues. This observation suggests that this non-producing species might have acquired the carbapenem resistance mechanism, without producing carbapenems.

Protein	Size (aa)	NCBI Accession Number	Structure PDB ID	Proposed Function [Ref]	Closest Homologue (species)	Identity/Similarity (%)
CarR	244	AAC45995	–	Transcriptional regulator [57]	ExpR (*Erwinia chrysanthemi*)	41/63
CarA	503	AAD38229	1Q15, 1Q19 [49]	β-Lactam synthetase [46, 30]	Asparagine synthase (*Dickeya zeae*)	55/71
CarB	250	AAD38230	2A7K, 2A81 [60]	Carboxymethylproline synthase [29, 46, 61]	Enoyl-CoA hydratase/isomerase (*Pantoea sp.* At-9b)	65/80
CarC	273	AAD3823	1NX8 [52]	Epimerase/desaturase [37]	CAS-like (*Dickeya zeae*)	82/91
CarD	376	AAD38232	–	Proline dehydrogenase [41]	Proline dehydrogenase (*Pantoea sp.* At-9b)	53/73
CarE	92	AAD38233	–	Ferredoxin; 2Fe-2S subgroup [41]	ferredoxin (*Dickeya zeae*)	56/79
CarF	288	AAD38234	–	Resistance mechanism [43]	Hypothetical protein (*Pectobacterium wasabiae*)	78/84
CarG	177	AAD38235	–	Resistance mechanism [43]	Hypothetical protein (*Pectobacterium wasabiae*)	84/93
CarH	184	AAD38235	–	Unknown [43]	Hypothetical protein (*Pectobacterium wasabiae*)	85/02

Table 2.1: Proposed functions of deduced *car* gene products in *Pectobacterium carotovorum*.

2.3 Biosynthesis of C-2 and C-6 Substituted Carbapenems

2.3.1 Labelling and Mutagenesis Studies

An initial proposal for the biosynthetic pathway of thienamycin **116** was based on the observation that acetate and glutamate are incorporated when added to fermentation broths of *S. cattleya* [62], linking the thienamycin biosynthesis with that of (5*R*)-1-carbapen-2-em-3-carboxylic acid (Section 2.2). Although the sequence of events leading to the introduction of the additional modifications of the carbapenem nucleus remain elusive, labelling studies have identified precursors for the C-2 cysteaminyl and C-6 hydroxyethyl side chains present in thienamycin **116**.

The C-2 cysteaminyl side chain has initially been postulated to originate from L-cysteine, rather than CoASH **133**. The observation that both fully labelled [^{14}C]- or [^{35}S]-cystine were incorporated in the cysteaminyl side chain of thienamycin **116** but not [^{35}S]-pantethine supported this proposal [62]. In addition, the single radiolabel of 1-[^{14}C]-cystine was not significantly incorporated into thienamycin **116**, in agreement with the possibility that the C-1 atom is released as CO_2 when cysteamine is enzymatically converted into L-cysteine through α-decarboxylation. A series of mutagenesis experiments coupled to LC-MS analysis of the intermediates present in the culture broths of the corresponding mutants suggested that carbapenams and carbapenems with a cysteinyl side chains were produced [63]. An observation that argues against the hypothesis that the cysteaminyl side chain comes from cysteine is the presence of pantothenate side chains in carbapenems of the OA-6129 series (Figures 2.2 and 2.3) [64]. Mutants of *S. fulvoviridis*, a species that usually produces the PS-5 carbapenems carrying a C-2 *N*-acetylcysteaminyl side chain, were shown to lack the acylase activity needed to hydrolyse the pantothenate moiety and therefore produced carbapenems with a C-2 pantetheinyl side chain instead [65]. Based on this observation, the C-2 side chain was also proposed to originate from pantothenate, although radiolabeled β-alanine was shown to be incorporated efficiently but not pantothenate itself [66]. In order to determine the sequence of biosynthetic events that link the OA-6129 metabolites together, various blocked mutants of *S. fulvoviridis* have been constructed. Isolation of the carbapenem intermediates produced by three of the mutants suggests that OA-6129 A is hydroxylated at C-8 to form OA-6129 B2, isomerisation of OA-6129 B2 at C-6 gives OA-6129 Bl and finally sulfation at the C-8 hydroxyl yields OA-6129C [67].

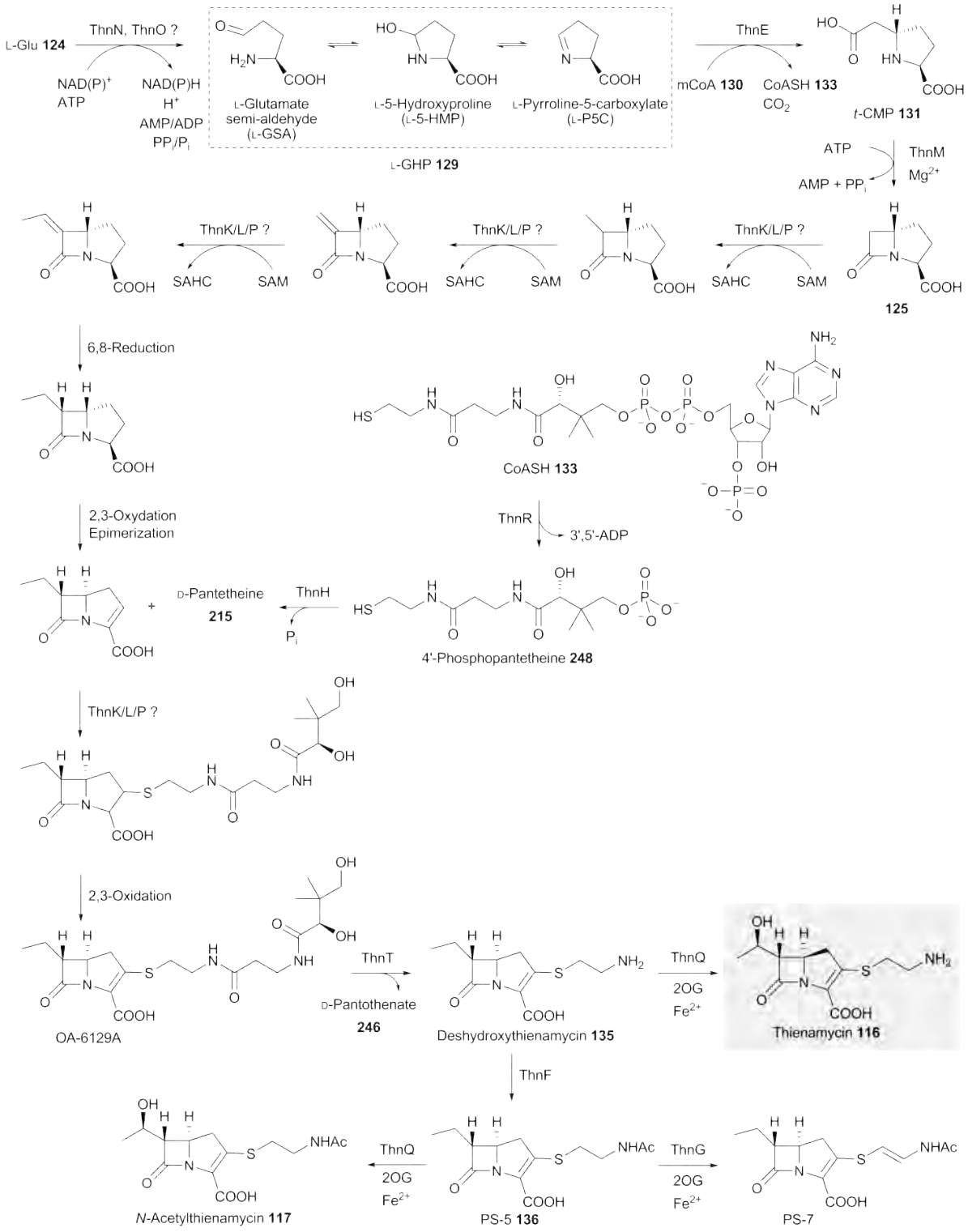
In addition, the isolation of the carbapenams OA-6129 D and OA-6129 E [68], which are structurally related to the OA-6129 carbapenems, but which lack the unsaturation between C-2 and C-3, suggest that there is an enzymatic oxidation step involved. Altogether, the results described in this paragraph point out the absence of a consensus regarding the origin and the mechanism of introduction of the cysteaminyl side chain of C-2 substituted carbapenems.

The C-6 1-hydroxyethyl side chain of the carbapenems has been proposed to originate from two sequential methylation steps by *S*-adenosylmethionine (SAM)-dependent enzymes, followed by hydroxylation. Radiolabeling experiments using L-[methyl-^{3}H]-methionine and L-[methyl-^{14}C]-methionine showed incorporation of the labels in thienamycin produced by *S. cattleya* cultures [62, 69]. Further support for this hypothesis comes from an experiment showing that SAM-dependent enzyme inhibition by cycloleucine abolished thienamycin production by S. cattleya [62]. The enzymatic introduction of the hydroxyl group was proposed based on the report that a mutant strain of *S. cattleya* produces deshydroxythienamycin **135** (NS-5) and *N*-acetyldeshydroxythienamycin **136** (PS-5) but not thienamycin **116**, although no molecular characterisation of the mutated gene was reported at the time [70].

2.3.2 Thienamycin Biosynthetic Enzymes

The cluster of genes encoding for the enzymes involved in the biosynthesis of thienamycin has been identified in *S. cattleya* [26]. This cluster contains 22 potential genes spanning across about 26 kbp (Figure 2.5) [26]. Only two of the predicted genes have clear homologs in the gene cluster encoding for (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** biosynthesis i.e. *thnE* and *thnM* (Figure 2.5). ThnE and ThnM show 32% and 24% sequence identity to CarB and CarA, respectively, the first two enzymes in the carbapenem pathway; no close identifiable CarC homologue was found [26]. In light of the information obtained from the predicted *thn* gene products (Tables 2.2 and 2.3), a plausible pathway for the biosynthesis of thienamycin has been proposed (Scheme 2.7) [26]. Further bioinformatic analysis enabled the discovery of a highly similar cluster in *S. flavogriseus* and a recent report suggests that it is cryptic under laboratory conditions [71].

The complete genome of *S. cattleya* NRRL 8057 was sequenced and annotated by the Genoscope - Centre National de Sequencage, in France, and released during the course of the present study [72]. This genome is composed of one linear chromosome of 6.28 Mbp and one linear megaplasmid of 1.81



Scheme 2.7: Overview of the proposed biosynthetic steps yielding thienamycin **116** and *N*-acetylthienamycin **117**. The exact sequence of the biosynthetic events presented here is speculative. Insertional inactivation experiments allowed the identification of several intermediates that are not shown in this scheme [63]. Figure modified from reference [26]

Mbp, both of them showing an average GC content of about 73%. 5779 protein-coding genes were annotated on the chromosome and 1713 on the plasmid. *S. flavogriseus* ATCC 33331 has a 7.34 Mbp main chromosome and two plasmids, pSFLA01 and pSFLA02, of 0.19 and 0.13 Mbp, respectively. As in *S. cattleya*, the *S. flavogriseus* cryptic thienamycin biosynthetic gene cluster is located on the *S. flavogriseus* choromosome [71].

Protein	Size (aa)	NCBI Accession Number	Proposed Function [Ref]	Closest Homologue (Species)	Identity/Similarity (%)
ThnA	259	CCB71846	NAD(P) dependent oxidoreductase	2-hydroxycyclohexanecarboxyl-CoA dehydrogenase (*Streptomyces* sp. AA4)	65/76
ThnB	257	CCB71847	Transcriptional regulator	TetR family transcriptional regulator (*Streptosporangium roseum*)	66/74
ThnC	209/212	CCB71848	Lactone efflux transmembrane protein	Lysine exporter protein LysE/YggA (*Catenulispora acidiphila* DSM 44928)	58/71
ThnD	363	CCB71849	Oxidoreductase	Alcohol dehydrogenase (*Vulcanisaeta distributa* DSM 14429)	31/50
ThnE	271/294	CCB71850	Carboxymethylproline synthase [73, 74]	Enoyl-CoA hydratase/isomerase (*Streptomyces flavogriseus* ATCC 33331)	76/86
ThnF	327	CCB71851	*N*-Acetyltransferase [75]	GCN5-related *N*-acetyltransferase (*Streptomyces flavogriseus* ATCC 33331)	54/62
ThnG	263	CCB71852	Fe(II) and 2OG-dependent oxygenase [76]	Phytanoyl-CoA dioxygenase (*Streptomyces flavogriseus* ATCC 33331)	67/81
ThnH	224	CCB71853	Hydrolase [75] [75]	HAD-superfamily hydrolase (*Streptomyces flavogriseus* ATCC 33331)	60/70
ThnI	476	CCB71854	Transcriptional activator [77]	Transcriptional activator ClaR (*Streptomyces clavuligerus* ATCC 27064)	33/46
ThnJ	483	CCB71855	Putative transport protein	Major facilitator superfamily transporter (*Streptomyces flavogriseus* ATCC 33331)	59/69
ThnK	681	CCB71856	Methyltransferase	Radical SAM superfamily member (*Streptomyces flavogriseus* ATCC 33331)	78/86
ThnL	474	CCB71857	Methyltransferase	Radical SAM superfamily member (*Streptomyces flavogriseus* ATCC 33331)	81/87

The proposed functions are based on BLAST [78] searches of the NCBI Protein database or biochemical data. In the case of ThnE, ThnF, ThnG, ThnH, and ThnI, the proposed function has been verified by biochemical studies and references are given.

Table 2.2: Proposed function of the *thn* gene products in *Streptomyces cattleya*.

Protein	Size (aa)	NCBI Accession Number	Proposed Function [Ref]	Closest Homologue (Species)	Identity/Similarity (%)
ThnM	458	CCB71858	β-Lactam synthetase [74]	Asparagine synthase (*Streptomyces flavogriseus* ATCC 33331)	64/73
ThnN	367	CCB71859	Carboxylate reductase component	AMP-dependent synthetase/ligase (*Streptomyces flavogriseus* ATCC 33331)	86/92
ThnO	472	CCB71860	Carboxylate reductase component	Hypothetical protein (*Streptomyces flavogriseus* ATCC 33331)	77/84
ThnP	484	CCB71861	Methyltransferase	Radical SAM domain protein (*Streptomyces flavogriseus* ATCC 33331)	83/91
ThnQ	259	CCB71862	Fe(II) and 2OG-dependent oxygenase [76]	Phytanoyl-CoA dioxygenase (*Streptomyces flavogriseus* ATCC 33331)	71/80
ThnR	196/240	CCB71863	Hydrolase [75]	NUDIX hydrolase (*Streptomyces flavogriseus* ATCC 33331)	61/73
ThnS	329	CCB71864	β-Lactamase	Zn-dependent hydrolase (*Saccharomonospora viridis* DSM 43017)	50/62
ThnT	382/399	CCB71865	Hydrolase [75]	Peptidase S58 DmpA (*Streptomyces flavogriseus* ATCC 33331)	81/89
ThnU	268	CCB71866	Transcriptional regulator	SARP family transcriptional regulator (*Streptomyces flavogriseus* ATCC 33331)	61/73
ThnV	137	CCB71867	Cysteine transferase	Cysteine transferase (*Streptomyces viridochromogenes* DSM 40736)	81/89

The proposed functions are based on BLAST [78] searches of the NCBI Protein database or biochemical data. In the case of ThnM, ThnQ, ThnR and ThnT, the proposed function has been verified by biochemical studies and references are given.

Table 2.3: Proposed function of the *thn* gene products in *Streptomyces cattleya* (continued).

2.4 Carboxymethylproline Synthase - An Unusual Member of the Crotonase Superfamily of Enzymes

CMPS belongs to the crotonase superfamily (CS) of enzymes (also called the enoyl-CoA hydratase superfamily), whose members share a common structural fold and similar catalytic mechanism [79, 80]. The common CS fold is composed of a repeated crotonase motif that consists of a $\beta\beta\alpha$ sequence. This fold assembles into two approximately perpendicular β-sheets surrounded by α-helices that enables the binding of the substrate, usually a coenzyme A thioester derivative (Figure 2.7 and Section 6.1.1, Figure 6.1). Catalytic activity is promoted by the positioning of the acyl-CoA thioester carbonyl group into a conserved oxyanion hole formed of two backbone amide hydrogens.

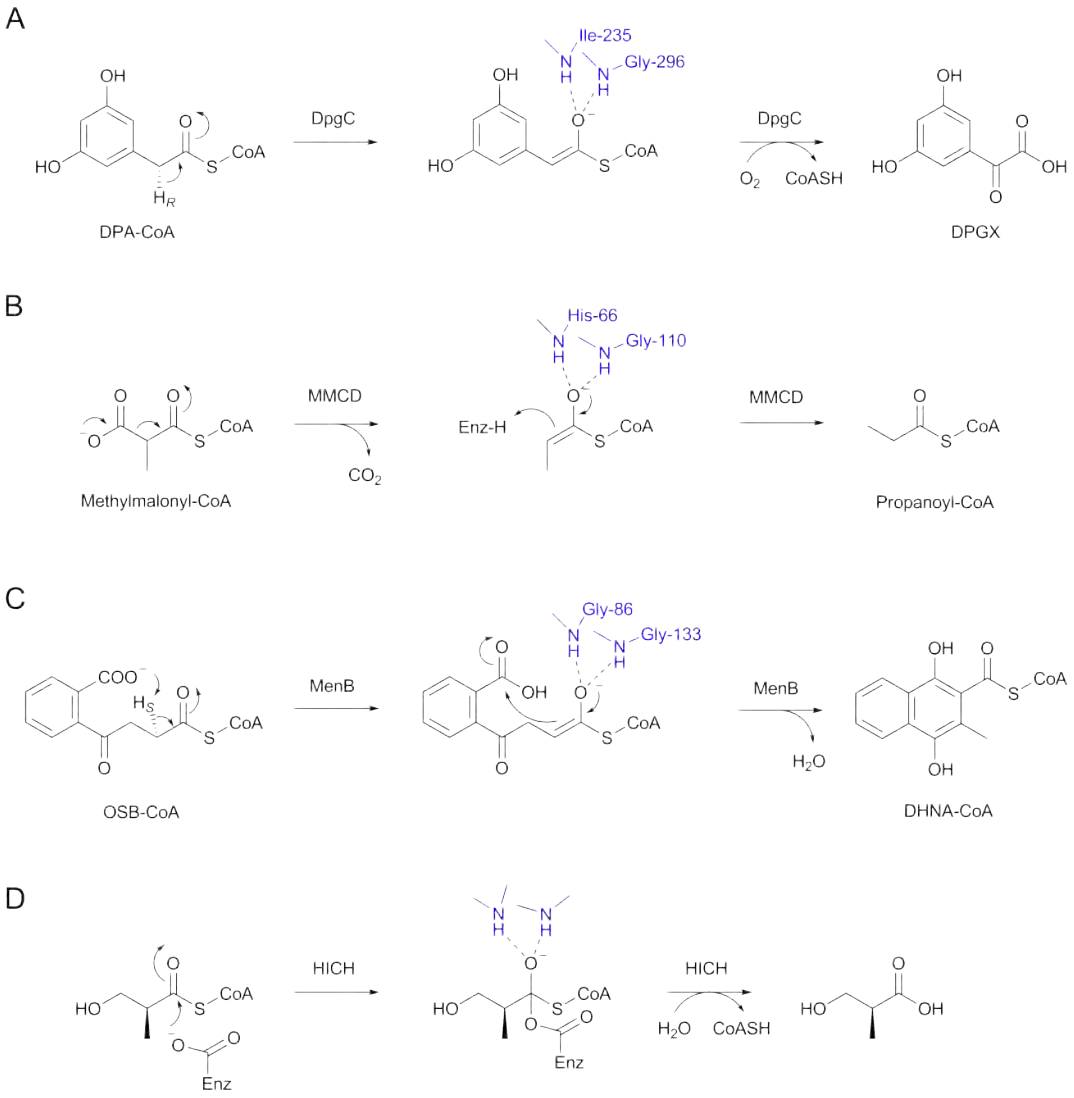
Multiple reactions are catalysed by members of the CS, *e.g.* cofactor independent oxidation [81], decarboxylation [82], C–C bond formation [83] and thioester hydrolysis [84] (Scheme 2.8); all these are proposed to involve an enolate intermediate. This common mechanistic feature and the conserved oxyanion hole suggests that the CS members have evolved from a common ancestor [79].

Because they catalyse a reaction similar to that of the CMPS, and because the crystal structures for three of them have been reported in complex with stable substrate analogues, the four members of the CS described below are of particular interest.

2.4.1 Examples of Enzymes of the CS Superfamily

3,5-Dihydroxyphenylglyoxylate Synthase

3,5-Dihydroxyphenylglyoxylate synthase (DpgC, EC 1.2.3.X) catalyses the cofactor-independent oxidation of 3,5-dihydroxyphenylacetyl-CoA (DPA-CoA) into 3,5-dihydroxyphenylglyoxylate (DPGX) (Scheme 2.8 A) [81]. This reaction is a key step in the biosynthesis of the non-proteinogenic amino acid 3,5-dihydroxyphenylglycine (DPG). DPG is then used by non-ribosomal peptide synthetases (NRPS) for the biosynthesis of cyclic peptide antibiotics of the vancomycin and teicoplanin family. A structure of DpgC from *Streptomyces toyocaensis* A47934, a vancomycin producer, was solved in complex with a stable analogue of the acyl-CoA substrate together with molecular oxygen (Figure 2.7 A, Section 5.1.5 Figure 5.2 A and B, PDB ID: 2NP9) [81]. This crystal structure allowed the proposal of a plausible mechanism for DpgC catalysis. An O_2 molecule



Scheme 2.8: Members of the CS catalyse diverse chemical reactions. Outlined mechanisms for four members of the CS showing the formation of an oxyanion intermediate during catalysis. For each enzyme, the backbone of the residues stabilising this intermediate (oxyanion hole) is highlighted in blue. (A) Cofactor independent dioxygenation reaction catalysed by DpgC [81]; (B) Decarboxylation of methylmalonyl-CoA by MMCD [82]; (C) C–C bond formation catalysed by MenB [83]; (D) Thioester hydrolysis by HICH.

is present in the active site hydrophobic pocket and oriented by a hydrogen bond with a backbone NH. That allows oxygen to be positioned to react with the enolate formed through substrate thioester pro-(*R*)-hydrogen α-deprotonation. A peroxide intermediate then displaces the CoA thiolate *via* an addition-elimination mechanism involving the oxianion hole to give a 1,2-dioxetanone and CoASH. A second pro-(*R*)-hydrogen deprotonation and ring opening produces the α-ketocarboxylate DPGX [81].

Methylmalonyl-CoA Decarboxylase

Methylmalonyl-CoA decarboxylase (MMCD, EC 4.1.1.41) catalyses the decarboxylation of methylmalonyl-CoA to yield propionyl-CoA (Scheme 2.8 B). A crystal structure of MMCD has been solved in complex with a thioether malonyl-CoA analogue (Figure 2.7 B, Figure 5.2 C and D) [82]. Modeling studies suggest that the backbone amide hydrogens of His-66 and Gly-110 form an oxyanion hole stabilising the thioester carboxylate, while a tyrosine residue orients the terminal methylmalonyl-CoA carboxylate so that it is orthogonal to the plane of the thioester carbonyl group. A similar orthogonal orientation of the malonyl-CoA terminal carboxylate is probably required for the decarboxylation step catalysed by CMPS [82].

1,4-Dihydroxy-2-naphthoyl-CoA Synthase

1,4-Dihydroxy-2-naphthoyl-CoA synthase (MenB, EC 4.1.3.36) catalyses the conversion of *O*-succinylbenzoyl-CoA (OSB-CoA) into 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) *via* an intramolecular Claisen condensation reaction (C–C bond formation) (Scheme 2.8 C) [83]. MenB is involved in the menaquinone biosynthesis pathway. Menaquinone is a redox active polyisoprenylated naphtoquinone used by most Gram-positive and some Gram-negative bacteria in the electron transport chain. MenB from *Mycobacterium tuberculosis* and *E. coli* were recently crystallised in complex with a stable substrate analogue (Figure 2.7 C) [83]. These structures are described in Section 5.1.2 (Figure 5.2 E and F, PDB ID: 3T88) and based on this structural information, the MenB mechanism was proposed to require two sets of active site residues crucial for C–C bond formation [83]. The oxyanion hole formed by the amide backbone of two glycine residues (Gly-86 and Gly-133 in

E. coli MenB) stabilises the thioenolate formed *via* an intramolecular α-proton abstraction. This thioenolate can then react with a free carboxylate in an intramolecular Claisen condensation fashion *via* a tetrahedral hemiacetal intermediate (Scheme 2.8 C). Two tyrosine residues are proposed to activate the free carboxylate for nucleophilic attack, and then stabilise the tetrahedral intermediate. Elimination of water yields the β-ketothioester and subsequent tautomerisation to the enol is driven by the thermodynamically favoured formation of the aromatic ring system.

3-Hydroxyisobutyryl-CoA Hydrolase

3-Hydroxyisobutyryl-CoA hydrolase (HICH, EC 3.1.2.4) is a CS member that catalyses the hydrolysis of an acyl-CoA thioester (Scheme 2.8 D). 3-Hydroxyisobutyryl-CoA is an intermediate in the metabolism of L-valine into propionyl-CoA [84, 85]. The structure of human HICH has been solved (PDB ID: 3BPT) but no description has been published and the detailed mechanism for this enzyme is not understood. Because the structure was solved in the absence of any substrate or analogue, the oxyanion hole residues have not been identified so far.

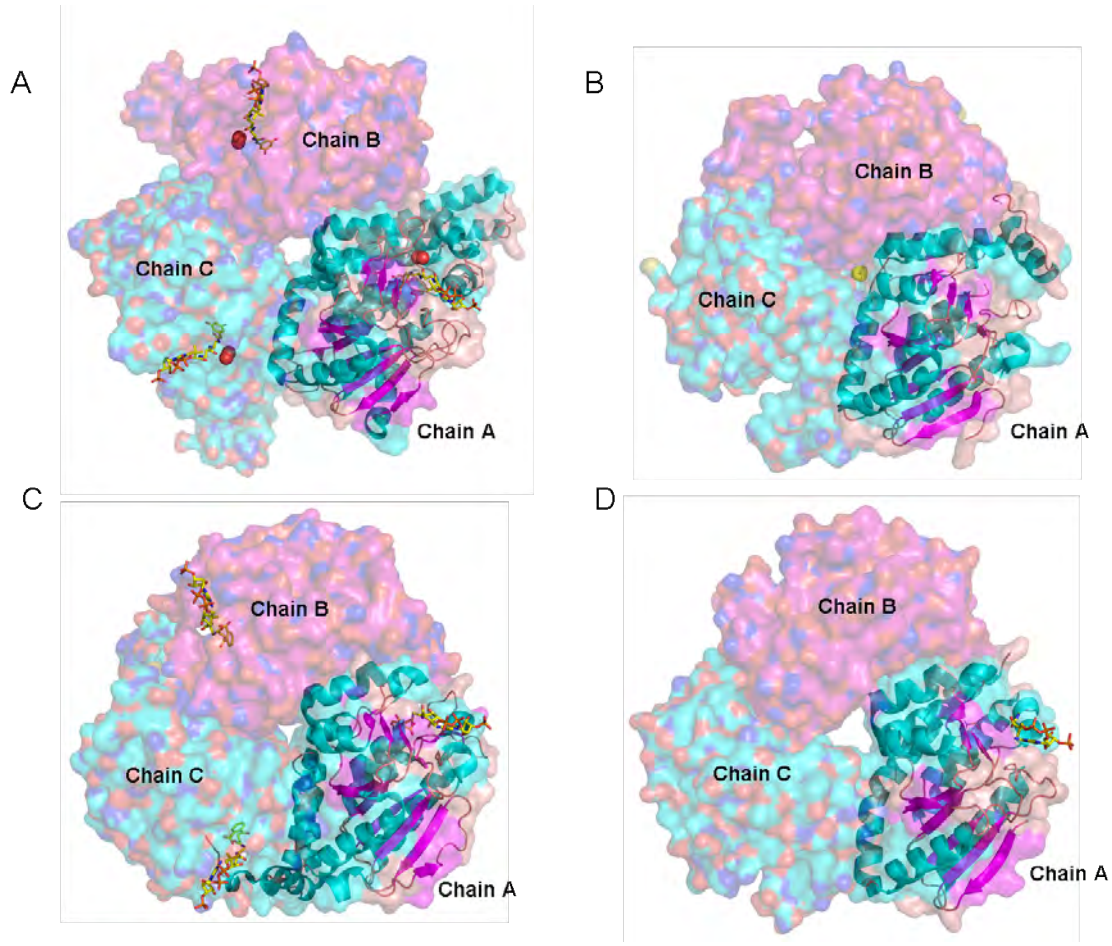


Figure 2.7: View of the crystal structures of four CS members show the conserved trimeric arrangement of CS enzymes. DpgC, MMCD, MenB and CarB crystallographic trimers. Cartoon representation of chain A shows α-helices in turquoise, β-sheets in purple and loops in pink. The surface is displayed for chain B (blue) and chain C (magenta). A view from the active site of DpgC, MMCD and MenB is shown in Figure 5.2. (A) *S. toyocaensis* DpgC trimer (PDB ID: 2NP9) showing the presence of one oxygen molecule (displayed as red spheres) and one stable 3,5-dihydroxyphenylacetyl-aza(dethia)CoA ligand per monomer; (B) *E. coli* MMCD trimer (PDB ID: 1EF8) with a nickel cation (shown as a yellow sphere) situated on the 3-fold rotational axis and coordinated by one histidine residue of each monomer (His-220); (C) *M. tuberculosis* MenB trimer (PDB ID: 3T8A); (D) *P. carotovorum* CarB trimer (PDB ID: 2A81). This figure was generated using PyMOL [86].

2.5 CMPS Mechanistic Investigations

Two genes encoding for the enzymes that catalyse the formation of *t*-CMP **131** in the context of the biosynthesis of a carbapenem natural product have been expressed in *E. coli* and the recombinant proteins characterised biochemically: CarB from *P. carotovorum* and ThnE from *S. cattleya*. Except for the first 46 amino acids, *S. cattleya* ThnE displays a high sequence identity with CarB (Appendix B.1, Table B.1). Because they share most of their characteristics, both ThnE and CarB enzymes are sometimes referred to as carboxymethylproline synthases (CMPS) throughout the text.

Searching for CMPS homologues in the NCBI Protein database using the BLAST algorithm [78], a series of sequences showing similarity to ThnE and CarB can be found in the context of carbapenem gene clusters. These genes are located within short genetic distance to genes showing homology to the *carA*/*thnM* β-LS genes. Table 2.4 shows the degree of conservation between the characterised CMPS enzymes ThnE (Entry I) and CarB (Entry V), together with ThnE* (Entry II) and other predicted CMPS. As judged from the other genes in the predicted clusters, the first two species are anticipated to produce thienamycin **116** or related C-2 and C-6 substituted carbapenems, while the last five species probably produce the unsubstituted (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**. A sequence alignment for *S. cattleya* ThnE, *S. flavogriseus* ThnE*, *P. carotovorum* CarB and *Photobacterium luminescens* CpmB is shown in Appendix B.1 (Table B.1).

The mechanism of CMPS catalysis has been investigated using a range of mechanistic probes and by comparing the stereochemistry of the resulting products with synthetic standards. CarB and ThnE have been demonstrated to catalyse the formation of *t*-CMP **131** in three distinct catalytic steps i.e. decarboxylation, C–C bond formation and thioester hydrolysis, employing malonyl-CoA **130** and L-GHP **129** as substrates (Scheme 2.3) [29, 73]. Multiple mechanisms were initially proposed for each step of the CMPS-catalysed formation of *t*-CMP **131**. Five different paths have been suggested for C–C bond formation (Scheme 2.9 paths A-E) and three paths for the thioester hydrolysis of the *t*-CMP-CoA intermediate **132** (Scheme 2.10 paths A-C). The sequence of binding events for the substrates to the active site is still not fully understood but a plausible mechanism has been proposed after careful elimination of selected possibilities based on experimental data. Of particular interest was the synthesis of a set of deuterium labelled L- and D-GHP analogues which were tested for CarB-

Entry	Species	Size (aa)	Identities (%)	Similarities (%)	Gaps (%)	GenBank Accession Number
I	*Streptomyces cattleya* NRRL 8057	272	272/272 (100)	272/272 (100)	0/272 (0)	CCB71850
II	*Streptomyces flavogriseus* ATCC 33331	263	191/251 (76)	217/251 (86)	0/251 (0)	ADW01632
III	*Rhodococcus fascians* D188 [87]	274	106/240 (44)	146/240 (61)	3/240 (1)	CAC43338
IV	*Dickeya zeae* Ech1591	250	92/241 (38)	141/241 (59)	0/241 (0)	ACT05105
V	*Pectobacterium carotovorum*	250	93/242 (38)	138/242 (57)	0/242 (0)	AAD38230
VI	*Photorhabdus luminescens* [44]	249	81/229 (35)	135/229 (59)	1/229 (0)	CAE12478
VII	*Pantoea sp.* At-9b	253	81/239 (34)	131/239 (55)	0/239 (0)	ADU71473

List of predicted CMPS compiled from BLAST search using the ThnE amino acid sequence. CMPS sequences were retrieved from the NCBI protein database searches for ThnE sequence similarities using the BLAST tool [78]. The degree of conservation with ThnE is given for each example. The genes coding for putative CMPS shown in this table were found within the context of a carbapenem biosynthesis gene cluster, as judged by neighbouring genes containing either a *carA* or a *thnM* β-LS homologue. As judged from the neighboring genes, the first two species are expected to produce a C-2 and C-6 substituted carbapenem, such as thienamycin **116**, while the last five probably produce (5*R*)-1-carbapen-2-em-3-carboxylic acid **115**.

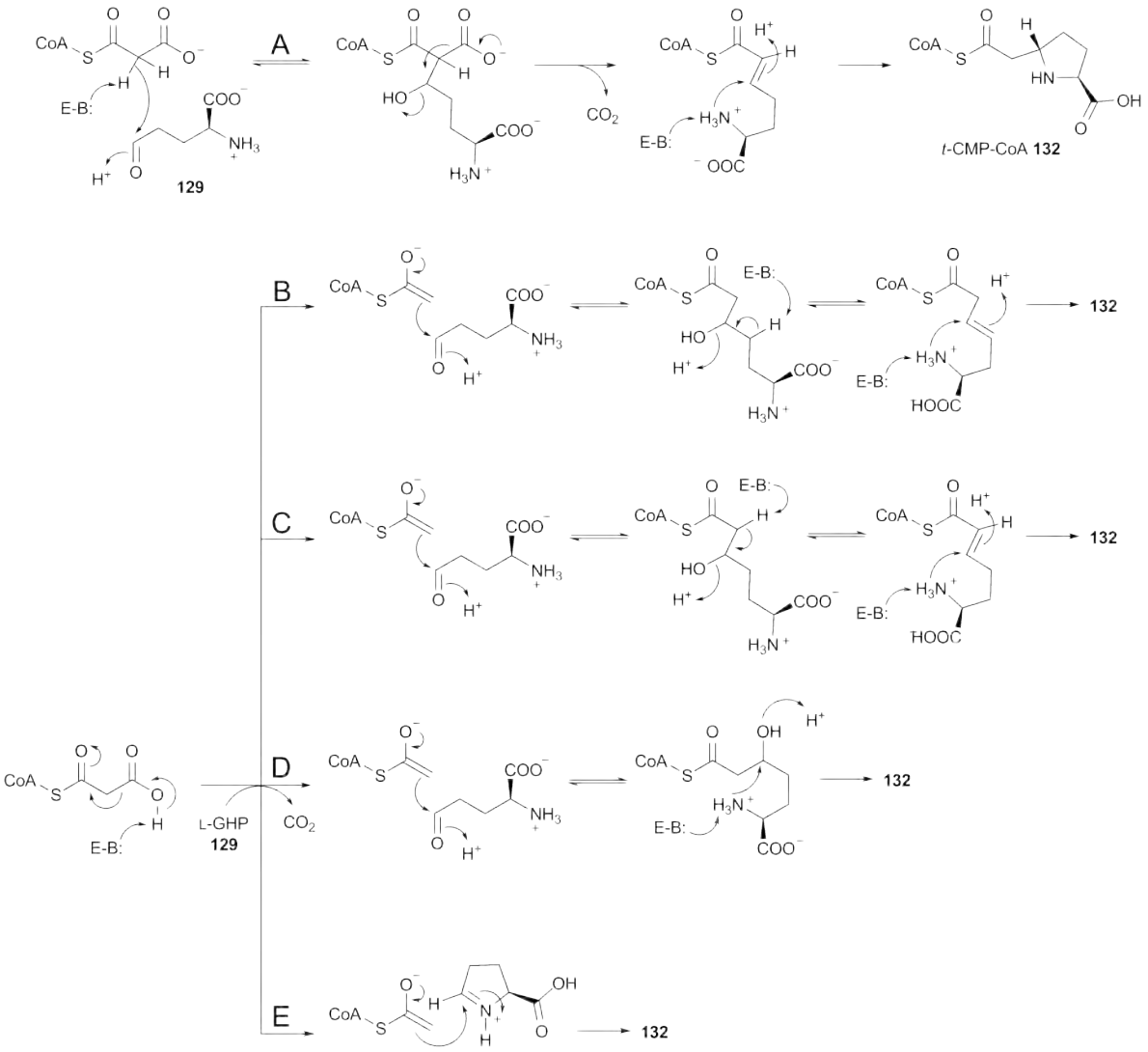
Table 2.4: Predicted carboxymethylproline synthase homologues showing homology to ThnE.

and ThnE-catalysed formation of *t*-CMP **131** [88, 89]. An important result obtained in these studies was the demonstration that L-GHP **129** is a substrate for the CMPS-catalysed formation of *t*-CMP **131**, while its D-enantiomer is not accepted by the enzyme [88]. This is in agreement with the results of an experiment in which the amount of CoASH **133** formed from excess malonyl-CoA **130** was found to be half of the amount of racemic DL-GHP in the CarB assay, as judged by colorimetric quantification of free thiol production coupled to the CMPS activity [46]. The observation that L-GHP **129**, but not D-GHP, is a CMPS substrate also has implications for the interpretation of the results described in Chapter 4. A summary of investigations on the CMPS mechanism is given in the next paragraphs.

2.5.1 Mechanism of CMPS-Catalysed Decarboxylation

Four of the five mechanisms proposed for CMPS-catalysed formation of the *t*-CMP-CoA intermediate **132** involve initial debarboxylation of malonyl-CoA to form an enolate that can then attack either the aldehyde, hemiaminal or imine form of L-GHP **129** (Scheme 2.9 paths B-E). In these mechanisms, the enolate is proposed to be stabilised in an oxyanion hole formed by two backbone amide hydrogen atoms (Gly-62 and Met-108 in CarB, Gly-107 and Val-153 in ThnE) [60]. The decarboxylation reaction of malonic acid half thioesters (MAHT) is probably slow in the absence of catalyst and under buffered physiological conditions. The rapid decarboxylation of malonyl-CoA **130** observed in the presence of CMPS is catalysed by the enzyme and is probably driven by active site residues promoting a stereoelectronically favoured decarboxylation of **130**. Decarboxylation of malonyl-CoA **130** is stereoelectronically favoured if the terminal malonyl-CoA carboxylate oxygen atoms are in a plane orthogonal to the oxyanion hole-stabilised thioester enolate. In addition, the polarisation of the thioester carbonyl group by the oxyanion hole residues likely assists in the catalysis of the decarboxylation step.

In the absence of L-GHP substrate, CMPS was shown to catalyse decarboxylation of malonyl-CoA and alkylmalonyl-CoA analogues, such as methyl and dimethylmalonyl-CoA **137**, to yield the corresponding acetyl, propionyl and isobutyryl-CoA species [46, 61]. This observation of an uncoupled decarboxylation reaction suggests that decarboxylation and the C–C bond formation in the presence of L-GHP **129** does not happen in a fully concerted process (Scheme 2.9, path A).



Scheme 2.9: Possible mechanisms for the CMPS-catalysed C–C bond formation [61]. (A) Aldol reaction prior to decarboxylation. This path is the only one not involving decarboxylation as a first step; (B) Aldol reaction followed by elimination of water *via* thioester α-proton abstraction and 5-*endo*-trig ring closure; (C) Aldol reaction followed by elimination of water *via* thioester γ-proton abstraction and 5-*exo*-trig ring closure; (D) Aldol reaction followed by direct substitution; (E) Mannich reaction with the imine form of L-GHP.

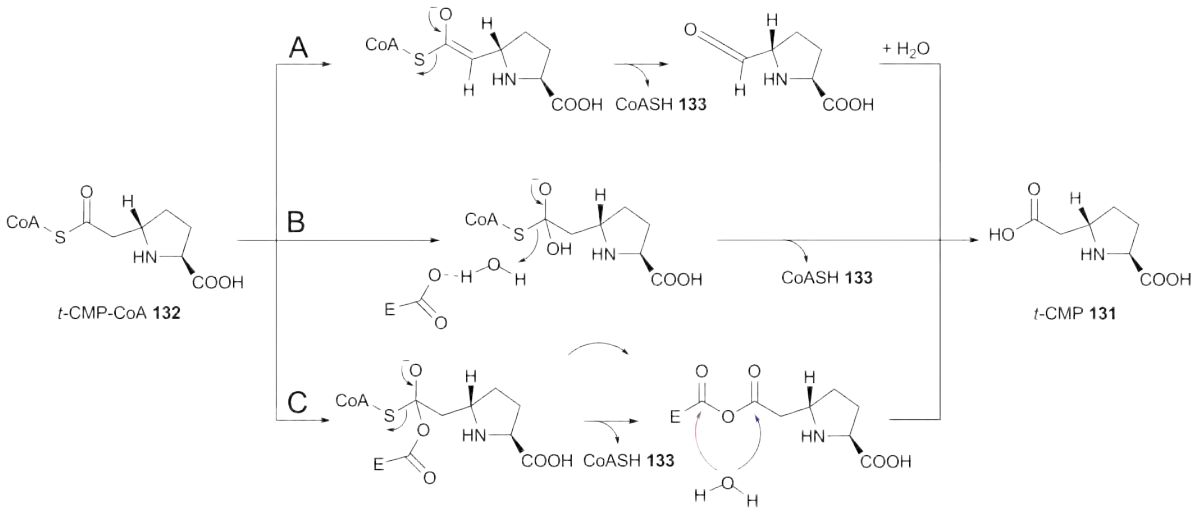
2.5.2 Mechanism of CMPS-Catalysed C–C Bond Formation

Two possible paths for ring formation involve a favoured 5-*exo*-trig ring closure (Scheme 2.9 paths A and C) while another one involves unfavored 5-*endo*-trig ring closure (Scheme 2.9 path B) [90]. The path that involves C–C bond formation prior to decarboxylation (Scheme 2.9 path A) is unlikely, based on the observation that uncoupled decarboxylation occurs, as described above, and impossible in the case of dimethylmalonyl-CoA **137** which is a substrate for the CMPS-catalysed formation of 6,6-dimethyl-*t*-CMP **138** [73, 61]. The proton abstraction required at the malonyl-CoA C-2 position cannot occur in the case of dimethylmalonyl-CoA **137**. The same observation applies to the path involving first C–C bond formation, followed by the 5-*exo*-trig ring closure which would require proton abstraction to form the α-unsaturated thioester intermediate (Scheme 2.9 path C). In a similar way, the observation that both deuterium labels are retained when the reaction is carried in aqueous buffer but in the presence of [$^2H_2$]-malonyl-CoA also excludes paths A and C. When 4,4-[$^2H_2$]-L-GHP was incubated with CMPS in the presence of malonyl-CoA, both the labels were retained at the corresponding C-4 position of the *t*-CMP product [89]. The impossibility of a mechanism involving proton abstraction at this position, followed by 5-*endo*-trig ring closure (Scheme 2.9 path B) was further confirmed with the synthesis and the incubation of 4,4-dimethyl-L-GHP **139** (Scheme 4.3, Section 4.3.1). Overall, path E was identified as the most likely mechanism for C–C bond formation [61].

2.5.3 Mechanism of CMPS-Catalysed Thioester Hydrolysis

The observation by Townsend and coworkers that *t*-CMP-CoA **132** can be isolated from CMPS incubations and characterised by MS [46] implies that the last step of the CMPS-catalysed *t*-CMP **131** formation is the hydrolysis of the thioester bond of this intermediate. The hypothesis that CMPS can catalyse the hydrolysis of *t*-CMP-CoA **132** has been confirmed further by the incubation of synthetic **132** in the presence of CarB, leading to the production of CoASH **133** and *t*-CMP **131** [46]. The decarboxylation products of the uncoupled reaction mentioned above, i.e. acetyl, propionyl and isobutyryl-CoA, were also found to be hydrolysed to yield CoASH **133** and the corresponding carboxylate [46, 61].

An elimination-addition mechanism has been suggested for thioester hydrolysis [91]. The elimination-addition mechanism involving a ketene intermediate (Scheme 2.10 path A) can be ruled



Scheme 2.10: Possible mechanisms for CMPS-catalysed thioester hydrolysis [61]. (A) Mechanism involving a ketene intermediate; (B) Thioester hydrolysis mediated by a water molecule activated by the active site glutamate residue; (C) Formation of an glutamate anhydride, followed by hydrolysis.

out in the case of CMPS catalysis, again based on the observation that dimethylmalonyl-CoA **137** is a substrate [61]. A path involving an activated water molecule (Scheme 2.10 path B) or a carboxylate anhydride formed by an active site glutamate residue (Scheme 2.10 path C) have been envisaged, in a way similar to the mechanisms proposed for 3-hydroxyisobutyryl-CoA hydrolase (HICH) [92]. This glutamate residue is Glu-131 in CarB, and Glu-176 in ThnE. Although this possibility cannot be ruled out completely, the glutamate-derived anhydride that is proposed for HICH is unlikely to be involved in CMPS-catalysed thioester hydrolysis. Experiments using multiple turnover of CMPS in $H_2{}^{18}O$ showed incorporation of one ^{18}O atom in the *t*-CMP product but none in the CMPS glutamate residue [60, 61]. The CarB variants in which the glutamate residue Glu-131 was replaced by alanine or glutamine showed complete loss of CMPS activity, while the aspartate mutant CarB E131D retained some activity, suggesting a catalytic role for the carboxylate function of the residue at this position [61].

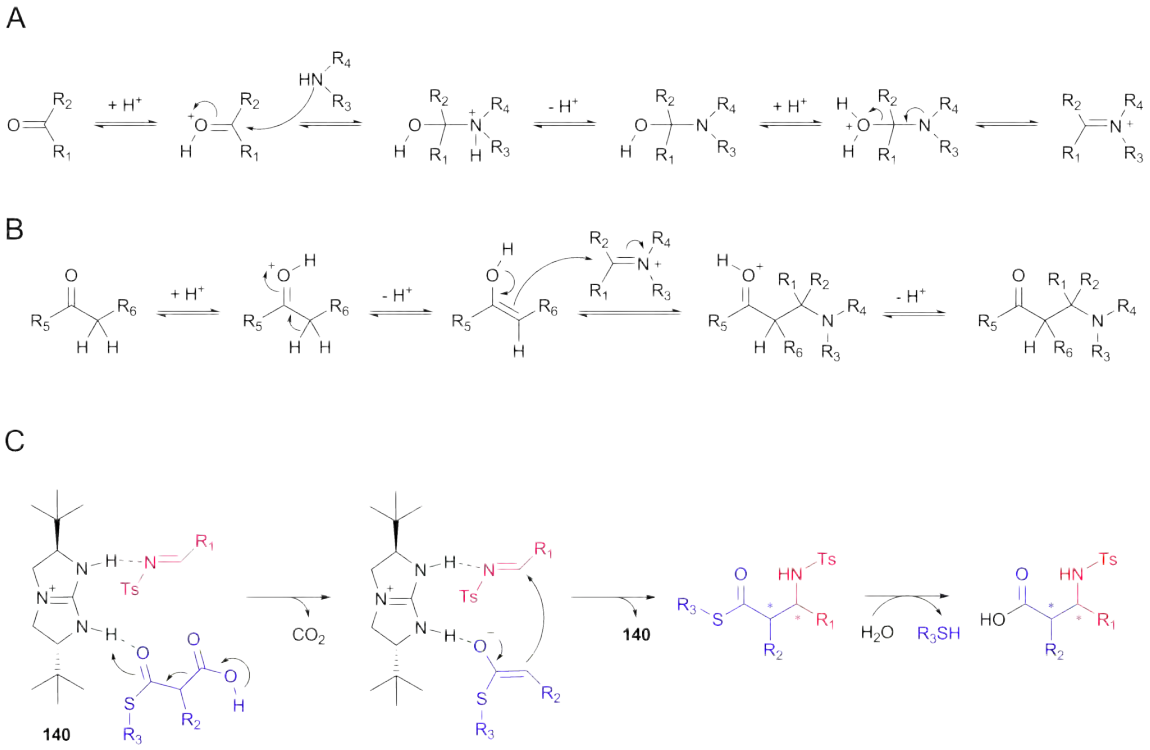
2.6 Objectives

The objectives of the studies described in the following chapters were to further characterise the early biosynthesis pathway of carbapenem antibiotics and investigate the possibility of engineering CMPS enzymes for the production of *t*-CMP analogues.

Stereoselective reactions are important transformations in organic chemistry. Except for the use of 6-oxocamphor hydrolase in β-diketone desymmetrisation [93], there are few reports of biocatalytic studies on CS enzymes. There is a general interest in developing stereoselective biocatalysts [94]. Our interest in CMPS biocatalysis emerges from the promising results previously obtained for these enzymes, and the chemical efforts that have been spent in generating biomimetic enantioselective addition reactions using artificial oxyanion holes [95]. Inspired by the mechanism of polyketide synthase (PKS) enzymes, organocatalysts have been developed for enantioselective 1,4-addition reactions of malonic acid half thioester (MAHT) to nitroolefins [96]. Similar to CMPS catalysis, other synthetic routes have exploited the decarboxylation of a MAHT to generate an enolate intermediate stabilised by a chiral bicyclic guanidine **140** (Scheme 2.11) [95]. The Mannich reaction of this enolate with a secondary imine was demonstrated to proceed with good stereocontrol. The ease of thioester hydrolysis allows the convenient two-step generation of the corresponding α,β-substituted β-amino acids. Considering the technically challenging and costly methodologies that synthetic chemists use for making a chiral molecule, CMPS is a potentially interesting target for biocatalysis studies.

Accordingly, the substrate tolerances of two characterised CMPS enzymes, CarB from *P. carotovorum* and ThnE from *S. cattleya* as well as some of their variants, were explored in order to expend their biocatalytic scope. Two specific questions were asked: (i) Is it possible to broaden the substrate tolerance of CMPS by using methyl-substituted L-GHP analogues in order to introduce additional functionalities into the *t*-CMP product? (ii) Is it possible to synthesise stable analogues of the CMPS substrates in order to get a CMPS crystal structure containing more detailed information than the structures available?

Investigations on the CMPS substrate tolerance are described in Chapter 4 using L-GHP analogues synthesised in Chapter 3. A synthetic route was investigated for each L-GHP analogue carrying a



Scheme 2.11: Proposed mechanism for the Mannich reaction. (A) First half-reaction generates an iminium cation through condensation of an aldehyde and an amine; (B) Second half-reaction involves the acid-catalysed formation of an enolate and C–C bond formation; (C) Representative example of enantioselective decarboxylative Mannich reaction used in synthetic chemistry. Starting from a MAHT and a pre-formed imine, branched β-amino acids were synthesised using a chiral oxyanion hole mimic catalyst such as a chiral bicyclic guanidine [95].

methyl group at position C-2, C-3, C-4 or C-5. After synthesis and isolation in a protected form, these substrate analogues were deprotected and tested as CMPS substrates for the formation of carboxymethylproline derivatives. Access to synthetic L-GHP analogues with a methyl group at position C-2, C-3, C-4 and C-5, or a cysteaminyl side chain at C-3 were expected to give further insight into the CMPS mechanism and provide a biocatalytic solution to access functionalised proline analogues with a high degree of stereoselectivity and possibly new chiral tertiary and quaternary centers.

Chapter 6 describes an attempt to obtain a CMPS crystal structure in complex with stable CoA analogues described in Chapter 5. The structural information obtained for a series of CS members in the presence of non-hydrolysable substrate analogues, as described in Section 5.1.2, strongly encouraged us to test a similar approach in the context of CMPS crystallography. So far, CMPS crystal structures have only been reported in the absence of ligand, or in complex with acetyl-CoA **127** but without visible electron density for the portion of the ligand binding to the active site [60]. These results are therefore lacking information about the binding mode of the substrates. A crystal structure of the *S. cattleya* ThnE protein has not been reported and an insight into the residues responsible for substrate positioning and specificity would allow better understanding of the catalytic mechanism. In addition, further CMPS mutagenesis to generate useful biocatalysts would benefit from a more accurate picture.

References

[1] Wennergren, G. and Lagercrantz, H. (2007) One sometimes finds what one is not looking for (Sir Alexander Fleming): The most important medical discovery of the 20th century. *Acta Paediatr*, **96**, 141–144.

[2] Kardos, N. and Demain, A. L. (2011) Penicillin: The medicine with the greatest impact on therapeutic outcomes. *Appl Microbiol Biotechnol*, **92**, 677–687.

[3] Sykes, R. B. and Bonner, D. P. (1985) Discovery and development of the monobactams. *Rev Infect Dis*, **7**, S579–S593.

[4] Fleming, A. (1929) On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the insolation of *B. influenzae*. *Br J Exp Pathol*, **10**, 226–236.

[5] Sheehan, J. C., Buhle, E. L., Corey, E. J., Laubach, G. D., and Ryan, J. J. (1950) The total synthesis of a 5-phenyl penicillin: Methyl 5-phenyl-(2-carboxymethyl)-penicillinate. *J Am Chem Soc*, **72**, 3828–3829.

[6] Theuretzbacher, U. (2011) Resistance drives antibacterial drug development. *Curr Opin Pharmacol*, **11**, 433–438.

[7] Tipper, D. J. and Strominger, J. L. (1965) Mechanism of action of penicillins: A proposal based on their structural similarity to acyl-D-alanyl-D-alanine. *Proc Natl Acad Sci U.S.A.*, **54**, 1133–1141.

[8] Cornaglia, G., Giamarellou, H., and Rossolini, G. M. (2011) Metallo-β-lactamases: A last frontier for β-lactams? *Lancet Infect Dis*, **11**, 381–393.

[9] Demain, A. L. and Elander, R. P. (1999) The β-lactam antibiotics: Past, present, and future. *Antonie Van Leeuwenhoek*, **75**, 5–19.

[10] Johnston, D. B. R., Schmitt, S. M., Bouffard, F. A., and Christensen, B. G. (1978) Total synthesis of ($\pm$)-thienamycin. *J Am Chem Soc*, **100**, 313–315.

[11] Albers-Schonberg, G., et al. (1978) Structure and absolute configuration of thienamycin. *J Am Chem Soc*, **100**, 6491–6499.

[12] Corbett, D. F., Eglington, A. J., and Howarth, T. T. (1977) Structure elucidation of MM 17880, a new fused β-lactam antibiotic isolated from *Streptomyces olivaceus*; A mild β-lactam degradation reaction. *J Chem Soc, Chem Commun*, **1977**, 953–954.

[13] Butterworth, D., Cole, M., Hanscomb, G., and Rolinson, G. N. (1979) Olivanic acids, a family of β-lactam antibiotics with β-lactamase inhibitory properties produced by *Streptomyces* species. I. Detection, properties and fermentation studies. *J Antibiot (Tokyo)*, **32**, 287–294.

[14] Brown, A. G., Corbett, D. F., Eglington, A. J., and Howarth, T. T. (1979) Structures of olivanic acid derivatives MM 22380, MM 22381, MM 22382 and MM 22383; four new antibiotics isolated from *Streptomyces olivaceus*. *J Antibiot (Tokyo)*, **32**, 961–963.

[15] Wise, R. (1986) *In vitro* and pharmacokinetic properties of the carbapenems. *Antimicrob Agents Chemother*, **30**, 343–349.

[16] Kahan, J. S., et al. (1979) Thienamycin, a new β-lactam antibiotic. I. Discovery, taxonomy, isolation and physical properties. *J Antibiot (Tokyo)*, **32**, 1–12.

[17] Leanza, W. J., Wildonger, K. J., Miller, T. W., and Christensen, B. G. (1979) *N*-Acetimidoyl- and *N*-formimidoylthienamycin derivatives: Antipseudomonal β-lactam antibiotics. *J Med Chem*, **22**, 1435–1436.

[18] Kropp, H., Sundelof, J. G., Hajdu, R., and Kahan, F. M. (1982) Metabolism of thienamycin and related carbapenem antibiotics by the renal dipeptidase, dehydropeptidase. *Antimicrob Agents Chemother*, **22**, 62–70.

[19] Kahan, F. M., Kropp, H., Sundelof, J. G., and Birnbaum, J. (1983) Thienamycin: Development of imipenen-cilastatin. *J Antimicrob Chemother*, **12 Suppl D**, 1–35.

[20] Shih, D. H., Cama, L., and Christensen, B. G. (1985) Synthetic carbapenem antibiotics III. 1-Methyl thienamycin. *Tetrahedron Lett*, **26**, 587–590.

[21] Kanazawa, K., Nouda, H., and Sunagawa, M. (1997) Structure-activity relationships of carbapenem compounds to anti-*Haemophilus influenzae* activity and affinity for penicillin-binding proteins. Effect of 1 β-methyl group and C-2 side chain. *J Antibiot (Tokyo)*, **50**, 162–168.

[22] Guthikonda, R. N., Cama, L. D., Quesada, M., Woods, M. F., Salzmann, T. N., and Christensen, B. G. (1987) Structure-activity relationships in the 2-arylcarbapenem series: Synthesis of 1-methyl-2-arylcarbapenems. *J Med Chem*, **30**, 871–880.

[23] Llarrull, L. I., Testero, S. A., Fisher, J. F., and Mobashery, S. (2010) The future of the β-lactams. *Curr Opin Microbiol*, **13**, 551–557.

[24] Frase, H., Shi, Q., Testero, S. A., Mobashery, S., and Vakulenko, S. B. (2009) Mechanistic basis for the emergence of catalytic competence against carbapenem antibiotics by the GES family of β-lactamases. *J Biol Chem*, **284**, 29509–29513.

[25] Kershaw, N. J., Caines, M. E. C., Sleeman, M. C., and Schofield, C. J. (2005) The enzymology of clavam and carbapenem biosynthesis. *Chem Commun*, **34**, 4251–4263.

[26] Núnez, L. E., Méndez, C., Brana, A. F., Blanco, G., and Salas, J. A. (2003) The biosynthetic gene cluster for the β-lactam carbapenem thienamycin in *Streptomyces cattleya*. *Chem Biol*, **10**, 301–311.

[27] Cohen, G., Shiffman, D., Mevarech, M., and Aharonowitz, Y. (1990) Microbial isopenicillin n synthase genes: structure, function, diversity and evolution. *Trends Biotechnol*, **8**, 105–111.

[28] Townsend, C. A. (2002) New reactions in clavulanic acid biosynthesis. *Curr Opin Chem Biol*, **6**, 583–589.

[29] Sleeman, M. C. and Schofield, C. J. (2004) Carboxymethylproline synthase (CarB), an unusual carbon-carbon bond-forming enzyme of the crotonase superfamily involved in carbapenem biosynthesis. *J Biol Chem*, **279**, 6730–6736.

[30] Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. *Biochemistry*, **42**, 7836–7847.

[31] Parker, W. L., Rathnum, M. L., Wells, J. S., Trejo, W. H., Principe, P. A., and Sykes, R. B. (1982) SQ 27,860, a simple carbapenem produced by species of *Serratia* and *Erwinia*. *J Antibiot (Tokyo)*, **35**, 653–660.

[32] Williamson, J. M. (1986) The biosynthesis of thienamycin and related carbapenems. *Crit Rev Biotech*, **4**, 111–131.

[33] Coulthurst, S. J., Barnard, A. M. L., and Salmond, G. P. C. (2005) Regulation and biosynthesis of carbapenem antibiotics in bacteria. *Nat Rev Microbiol*, **3**, 295–306.

[34] Bycroft, B. W., Maslen, C., Box, S. J., Brown, A. G., and Tyler, J. W. (1987) The isolation and characterisation of (3*R*,5*R*)- and (3*S*,5*R*)-carbapenam-3-carboxylic acid from *Serratia* and *Erwinia* species and their putative biosynthetic role. *J Chem Soc, Chem Commun*, **1987**, 1623–1625.

[35] Bycroft, B. W., Maslen, C., Box, S. J., Brown, A., and Tyler, J. W. (1988) The biosynthetic implications of acetate and glutamate incorporation into (3*R*,5*R*)-carbapenam-3-carboxylic acid and (5*R*)-carbapen-2-em-3-carboxylic acid by *Serratia sp. J Antibiot (Tokyo)*, **41**, 1231–1242.

[36] Bycroft, B. W. and Chhabra, S. R. (1989) A chiral synthesis of trans-carbapenam-3-carboxylic acid and the assignment of (3*S*,5*S*) configuration to the corresponding product from *Serratia* and *Erwinia* species. *J Chem Soc, Chem Commun*, **1989**, 423–425.

[37] Stapon, A., Li, R., and Townsend, C. A. (2003) Carbapenem biosynthesis: Confirmation of stereochemical assignments and the role of CarC in the ring stereoinversion process from L-proline. *J Am Chem Soc*, **125**, 8486–8493.

[38] Sleeman, M. C., Smith, P., Kellam, B., Chhabra, S. R., Bycroft, B. W., and Schofield, C. J. (2004) Biosynthesis of carbapenem antibiotics: New carbapenam substrates for carbapenem synthase (CarC). *ChemBioChem*, **5**, 879–882.

[39] Stapon, A., Li, R., and Townsend, C. A. (2003) Synthesis of (3*S*,5*R*)-carbapenam-3-carboxylic acid and its role in carbapenem biosynthesis and the stereoinversion problem. *J Am Chem Soc*, **125**, 15746–15747.

[40] Bycroft, B. W., Chhabra, S. R., Kellam, B., and Smith, P. (2003) Convenient syntheses of (3*S*,5*S*)-carbapenam-3-carboxylates and their biosynthetic relevance. *Tetrahedron Lett*, **44**, 973–976.

[41] McGowan, S. J., Sebaihia, M., Porter, L. E., Stewart, G. S., Williams, P., Bycroft, B. W., and Salmond, G. P. (1996) Analysis of bacterial carbapenem antibiotic production genes reveals a novel β-lactam biosynthesis pathway. *Mol Microbiol*, **22**, 415–426.

[42] Li, R., Stapon, A., Blanchfield, J. T., and Townsend, C. A. (2000) Three unusual reactions mediate carbapenem and carbapenam biosynthesis. *J Am Chem Soc*, **122**, 9296–9297.

[43] McGowan, S. J., Sebaihia, M., O'Leary, S., Hardie, K. R., Williams, P., Stewart, G. S., Bycroft, B. W., and Salmond, G. P. (1997) Analysis of the carbapenem gene cluster of *Erwinia carotovora*: Definition of the antibiotic biosynthetic genes and evidence for a novel β-lactam resistance mechanism. *Mol Microbiol*, **26**, 545–556.

[44] Derzelle, S., Duchaud, E., Kunst, F., Danchin, A., and Bertin, P. (2002) Identification, characterization, and regulation of a cluster of genes involved in carbapenem biosynthesis in *Photorhabdus luminescens*. *Appl Environ Microbiol*, **68**, 3780–3789.

[45] McGowan, S. J., Bycroft, B. W., and Salmond, G. P. (1998) Bacterial production of carbapenems and clavams: Evolution of β-lactam antibiotic pathways. *Trends Microbiol*, **6**, 203–208.

[46] Gerratana, B., Arnett, S. O., Stapon, A., and Townsend, C. A. (2004) Carboxymethylproline synthase from *Pectobacterium carotorova*: A multifaceted member of the crotonase superfamily. *Biochemistry*, **43**, 15936–15945.

[47] Bachmann, B. O., Li, R., and Townsend, C. A. (1998) β-lactam synthetase: A new biosynthetic enzyme. *Proc Natl Acad Sci U.S.A.*, **95**, 9082–9086.

[48] Raber, M. L., Arnett, S. O., and Townsend, C. A. (2009) A conserved tyrosyl-glutamyl catalytic dyad in evolutionarily linked enzymes: Carbapenam synthetase and β-lactam synthetase. *Biochemistry*, **48**, 4959–4971.

[49] Miller, M. T., Gerratana, B., Stapon, A., Townsend, C. A., and Rosenzweig, A. C. (2003) Crystal structure of carbapenam synthetase (CarA). *J Biol Chem*, **278**, 40996–41002.

[50] Miller, M. T., Bachmann, B. O., Townsend, C. A., and Rosenzweig, A. C. (2002) The catalytic cycle of β-lactam synthetase observed by X-ray crystallographic snapshots. *Proc Natl Acad Sci U.S.A.*, **99**, 14752–14757.

[51] Miller, M. T., Bachmann, B. O., Townsend, C. A., and Rosenzweig, A. C. (2001) Structure of β-lactam synthetase reveals how to synthesize antibiotics instead of asparagine. *Nat Struct Biol*, **8**, 684–689.

[52] Clifton, I. J., Doan, L. X., Sleeman, M. C., Topf, M., Suzuki, H., Wilmouth, R. C., and Schofield, C. J. (2003) Crystal structure of carbapenem synthase (CarC). *J Biol Chem*, **278**, 20843–20850.

[53] Zhou, J., Kelly, W. L., Bachmann, B. O., Gunsior, M., Townsend, C. A., and Solomon, E. I. (2001) Spectroscopic studies of substrate interactions with clavaminate synthase 2, a multifunctional α-KG-dependent non-heme iron enzyme: Correlation with mechanisms and reactivities. *J Am Chem Soc*, **123**, 7388–7398.

[54] Topf, M., Sandala, G. M., Smith, D. M., Schofield, C. J., Easton, C. J., and Radom, L. (2004) The unusual bifunctional catalysis of epimerization and desaturation by carbapenem synthase. *J Am Chem Soc*, **126**, 9932–9933.

[55] Borowski, T., Broclawik, E., Schofield, C. J., and Siegbahn, P. E. M. (2006) Epimerization and desaturation by carbapenem synthase (CarC). A hybrid DFT study. *J Comput Chem*, **27**, 740–748.

[56] Krebs, C., Fujimori, D. G., Walsh, C. T., and Bollinger, J. M. (2007) Non-heme Fe(IV)-oxo intermediates. *Acc Chem Res*, **40**, 484–492.

[57] Welch, M., Todd, D. E., Whitehead, N. A., McGowan, S. J., Bycroft, B. W., and Salmond, G. P. (2000) *N*-Acyl homoserine lactone binding to the CarR receptor determines quorum-sensing specificity in *Erwinia*. *EMBO J*, **19**, 631–641.

[58] McGowan, S., Sebaihia, M., Jones, S., Yu, B., Bainton, N., Chan, P. F., Bycroft, B., Stewart, G. S., Williams, P., and Salmond, G. P. (1995) Carbapenem antibiotic production in *Erwinia carotovora* is regulated by CarR, a homologue of the LuxR transcriptional activator. *Microbiology*, **141**, 541–550.

[59] McGowan, S. J., Holden, M. T., Bycroft, B. W., and Salmond, G. P. (1999) Molecular genetics of carbapenem antibiotic biosynthesis. *Antonie Van Leeuwenhoek*, **75**, 135–141.

[60] Sleeman, M. C., Sorensen, J. L., Batchelar, E. T., McDonough, M. A., and Schofield, C. J. (2005) Structural and mechanistic studies on carboxymethylproline synthase (CarB), a unique member of the crotonase superfamily catalyzing the first step in carbapenem biosynthesis. *J Biol Chem*, **280**, 34956–34965.

[61] Batchelar, E. T., Hamed, R. B., Ducho, C., Claridge, T. D. W., Edelmann, M. J., Kessler, B., and Schofield, C. J. (2008) Thioester hydrolysis and C–C bond formation by carboxymethylproline synthase from the crotonase superfamily. *Angew Chem Int Ed Engl*, **47**, 9322–9325.

[62] Williamson, J. M., Inamine, E., Wilson, K. E., Douglas, A. W., Liesch, J. M., and Albers-Schnberg, G. (1985) Biosynthesis of the β-lactam antibiotic, thienamycin, by *Streptomyces cattleya*. *J Biol Chem*, **260**, 4637–4647.

[63] Rodríguez, M., Núnez, L. E., Brana, A. F., Méndez, C., Salas, J. A., and Blanco, G. (2011) Mutational analysis of the thienamycin biosynthetic gene cluster from *Streptomyces cattleya*. *Antimicrob Agents Chemother*, **55**, 1638–1649.

[64] Fukagawa, Y., Okabe, M., Azuma, S., Kojima, I., Ishikura, T., and Kubo, K. (1984) Studies on the biosynthesis of carbapenem antibiotics. I. Biosynthetic significance of the OA-6129 group of carbapenem compounds as the direct precursors for PS-5, epithienamycins A and C and MM 17880. *J Antibiot (Tokyo)*, **37**, 1388–1393.

[65] Kubo, K., Ishikura, T., and Fukagawa, Y. (1984) Studies on the biosynthesis of carbapenem antibiotics. II. Isolation and functions of a specific acylase involved in the depantothenylation of the OA-6129 compounds. *J Antibiot (Tokyo)*, **37**, 1394–1402.

[66] Kubo, K., Ishikura, T., and Fukagawa, Y. (1985) Studies on the biosynthesis of carbapenem antibiotics. IV. β-Alanine in the C-3 pantetheinyl side chain of the OA-6129 group of carbapenems. *J Antibiot (Tokyo)*, **38**, 904–911.

[67] Kojima, I., Fukagawa, Y., Okabe, M., Ishikura, T., and Shibamoto, N. (1988) Mutagenesis of OA-6129 carbapenem-producing blocked mutants and the biosynthesis of carbapenems. *J Antibiot (Tokyo)*, **41**, 899–907.

[68] Yoshioka, T., Watanabe, A., Kojima, I., Shimauchi, Y., Okabe, M., Fukagawa, Y., and Ishikura, T. (1984) Structures of OA-6129D and E, new carbapenam antibiotics. *J Antibiot (Tokyo)*, **37**, 211–217.

[69] Houck, D. R., Kobayashi, K., Williamson, J. M., and Floss, H. G. (1986) Stereochemistry of methylation in thienamycin biosynthesis: Example of a methyl transfer from methionine with retention of configuration. *J Am Chem Soc*, **108**, 5365–5366.

[70] Rosi, D., Drozd, M. L., Kuhrt, M. F., Terminello, L., Came, P. E., and Daum, S. J. (1981) Mutants of *Streptomyces cattleya* producing *N*-acetyl and deshydroxy carbapenems related to thienamycin. *J Antibiot (Tokyo)*, **34**, 341–343.

[71] Blanco, G. (2011) Comparative analysis of a cryptic thienamycin-like gene cluster identified in *Streptomyces flavogriseus* by genome mining. *Arch Microbiol*.

[72] Barbe, V., Bouzon, M., Mangenot, S., Badet, B., Poulain, J., Segurens, B., Vallenet, D., Marlire, P., and Weissenbach, J. (2011) Complete genome sequence of *Streptomyces cattleya* NRRL 8057, a producer of antibiotics and fluorometabolites. *J Bacteriol*, **193**, 5055–5056.

[73] Hamed, R. B., Batchelar, E. T., Mecinović, J., Claridge, T. D. W., and Schofield, C. J. (2008) Evidence that thienamycin biosynthesis proceeds via C-5 epimerization: ThnE catalyzes the formation of (2*S*,5*S*)-*trans*-carboxymethylproline. *ChemBioChem*, **10**, 246–250.

[74] Bodner, M. J., Li, R., Phelan, R. M., Freeman, M. F., Moshos, K. A., Lloyd, E. P., and Townsend, C. A. (2011) Definition of the common and divergent steps in carbapenem β-lactam antibiotic biosynthesis. *ChemBioChem*, **12**, 2159–2165.

[75] Freeman, M. F., Moshos, K. A., Bodner, M. J., Li, R., and Townsend, C. A. (2008) Four enzymes define the incorporation of coenzyme A in thienamycin biosynthesis. *Proc Natl Acad Sci U.S.A.*, **105**, 11128–11133.

[76] Bodner, M. J., Phelan, R. M., Freeman, M. F., Li, R., and Townsend, C. A. (2010) Non-heme iron oxygenases generate natural structural diversity in carbapenem antibiotics. *J Am Chem Soc*, **132**, 12–13.

[77] Rodríguez, M., Núnez, L. E., Brana, A. F., Méndez, C., Salas, J. A., and Blanco, G. (2008) Identification of transcriptional activators for thienamycin and cephamycin C biosynthetic genes within the thienamycin gene cluster from *Streptomyces cattleya*. *Mol Microbiol*, **69**, 633–645.

[78] Altschul, S. F., Madden, T. L., Schffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389–3402.

[79] Hamed, R. B., Batchelar, E. T., Clifton, I. J., and Schofield, C. J. (2008) Mechanisms and structures of crotonase superfamily enzymes – How nature controls enolate and oxyanion reactivity. *Cell Mol Life Sci*, **65**, 2507–2527.

[80] Holden, H. M., Benning, M. M., Haller, T., and Gerlt, J. A. (2001) The crotonase superfamily: Divergently related enzymes that catalyze different reactions involving acyl coenzyme A thioesters. *Acc Chem Res*, **34**, 145–157.

[81] Widboom, P. F., Fielding, E. N., Liu, Y., and Bruner, S. D. (2007) Structural basis for cofactor-independent dioxygenation in vancomycin biosynthesis. *Nature*, **447**, 342–345.

[82] Benning, M. M., Haller, T., Gerlt, J. A., and Holden, H. M. (2000) New reactions in the crotonase superfamily: Structure of methylmalonyl CoA decarboxylase from *Escherichia coli*. *Biochemistry*, **39**, 4630–4639.

[83] Li, H.-J., Li, X., Liu, N., Zhang, H., Truglio, J. J., Mishra, S., Kisker, C., Garcia-Diaz, M., and Tonge, P. J. (2011) Mechanism of the intramolecular Claisen condensation reaction catalyzed by MenB, a crotonase superfamily member. *Biochemistry*, **50**, 9532–9544.

[84] Hawes, J. W., Jaskiewicz, J., Shimomura, Y., Huang, B., Bunting, J., Harper, E. T., and Harris, R. A. (1996) Primary structure and tissue-specific expression of human β-hydroxyisobutyryl-coenzyme A hydrolase. *J Biol Chem*, **271**, 26430–26434.

[85] Shimomura, Y., Murakami, T., Nakai, N., Huang, B., Hawes, J. W., and Harris, R. A. (2000) 3-hydroxyisobutyryl-coa hydrolase. *Methods Enzymol*, **324**, 229–240.

[86] Schrödinger, LLC (2010) The PyMOL molecular graphics system, version 1.3r1.

[87] Maes, T., Vereecke, D., Ritsema, T., Cornelis, K., Thu, H. N., Montagu, M. V., Holsters, M., and Goethals, K. (2001) The att locus of *Rhodococcus fascians* strain D188 is essential for full virulence on tobacco through the production of an autoregulatory compound. *Mol Microbiol*, **42**, 13–28.

[88] Sorensen, J. L., Sleeman, M. C., and Schofield, C. J. (2005) Synthesis of deuterium labelled L- and D-glutamate semialdehydes and their evaluation as substrates for carboxymethylproline synthase (CarB)–implications for carbapenem biosynthesis. *Chem Commun*, **9**, 1155–1157.

[89] Ducho, C., Hamed, R. B., Batchelar, E. T., Sorensen, J. L., Odell, B., and Schofield, C. J. (2009) Synthesis of regio- and stereoselectively deuterium-labelled derivatives of L-glutamate semialdehyde for studies on carbapenem biosynthesis. *Org Biomol Chem*, **7**, 2770–2779.

[90] Baldwin, J. E. (1976) Rules for ring closure. *J Chem Soc, Chem Commun*, **1976**, 734–736.

[91] Douglas, K. T. (1986) Elimination-addition pathways for thiol esters. *Acc Chem Res*, **19**, 186–192.

[92] Wong, B. J. and Gerlt, J. A. (2003) Divergent function in the crotonase superfamily: An anhydride intermediate in the reaction catalyzed by 3-hydroxyisobutyryl-CoA hydrolase. *J Am Chem Soc*, **125**, 12076–12077.

[93] Grogan, G., Roberts, G. A., Bougioukou, D., Turner, N. J., and Flitsch, S. L. (2001) The desymmetrization of bicyclic β-diketones by an enzymatic retro-Claisen reaction. A new reaction of the crotonase superfamily. *J Biol Chem*, **276**, 12565–12572.

[94] Wohlgemuth, R. (2010) Asymmetric biocatalysis with microbial enzymes and cells. *Curr Opin Microbiol*, **13**, 283–292.

[95] Pan, Y., Kee, C. W., Jiang, Z., Ma, T., Zhao, Y., Yang, Y., Xue, H., and Tan, C.-H. (2011) Expanding the utility of Bronsted base catalysis: Biomimetic enantioselective decarboxylative reactions. *Chem Eur J*, **17**, 8363–8370.

[96] Lubkoll, J. and Wennemers, H. (2007) Mimicry of polyketide synthases – Enantioselective 1,4-addition reactions of malonic acid half-thioesters to nitroolefins. *Angew Chem Int Ed Engl*, **46**, 6841–6844.

Chapter 3

Synthesis of L-GHP Analogues.

3.1 Introduction

Structural analogues of the CMPS substrate L-GHP **129** are of interest both from a biocatalysis and a biosynthesis perspective. This chapter describes the synthesis of a series of L-GHP analogues that enabled further investigations into the CMPS substrate tolerance. All the compounds described below were tested as CMPS substrates, as described in Chapter 4.

From a biocatalysis perspective, CarB, ThnE and their variants could be valuable tools for the stereoselective production of substituted proline derivatives. Functionalised saturated *N*-heterocycles in general, and proline derivatives in particular, are widely found in pharmaceutically active and clinically used molecules, such as saxagliptin (BMS-477118) [1] and captopril (SQ 14225) [2] (Figure 3.1 A), as well as in biologically active natural products [3], such as the antibiotic siderochelin A [4], cyclic peptide antibiotic bottromycin A2 [5, 6], nostopeptolides A1-3, polyoxypeptin A, neurotoxin kainic acid [7] and proteasome inhibitor lactacystin [8, 9] (Figure 3.1 B). In chemical synthesis, organocatalysis using proline-based ligands has proven a very powerful tool for stereoselective chemical transformations. CMPS biocatalysis could provide a rapid and convenient access to a catalogue of substituted carboxymethylproline chiral compounds which could be useful as building blocks for synthesis.

From a biosynthesis perspective, L-GHP analogues could help clarify the sequence of events taking place in the multi-enzyme pathway producing thienamycin **116**, olivanic acid, and related C-2 and C-6 substituted carbapenems. The syntheses of two C-3 cysteaminyl-substituted L-GHP analogues were carried out in order to test the possibility that ThnE could accept a substrate with a

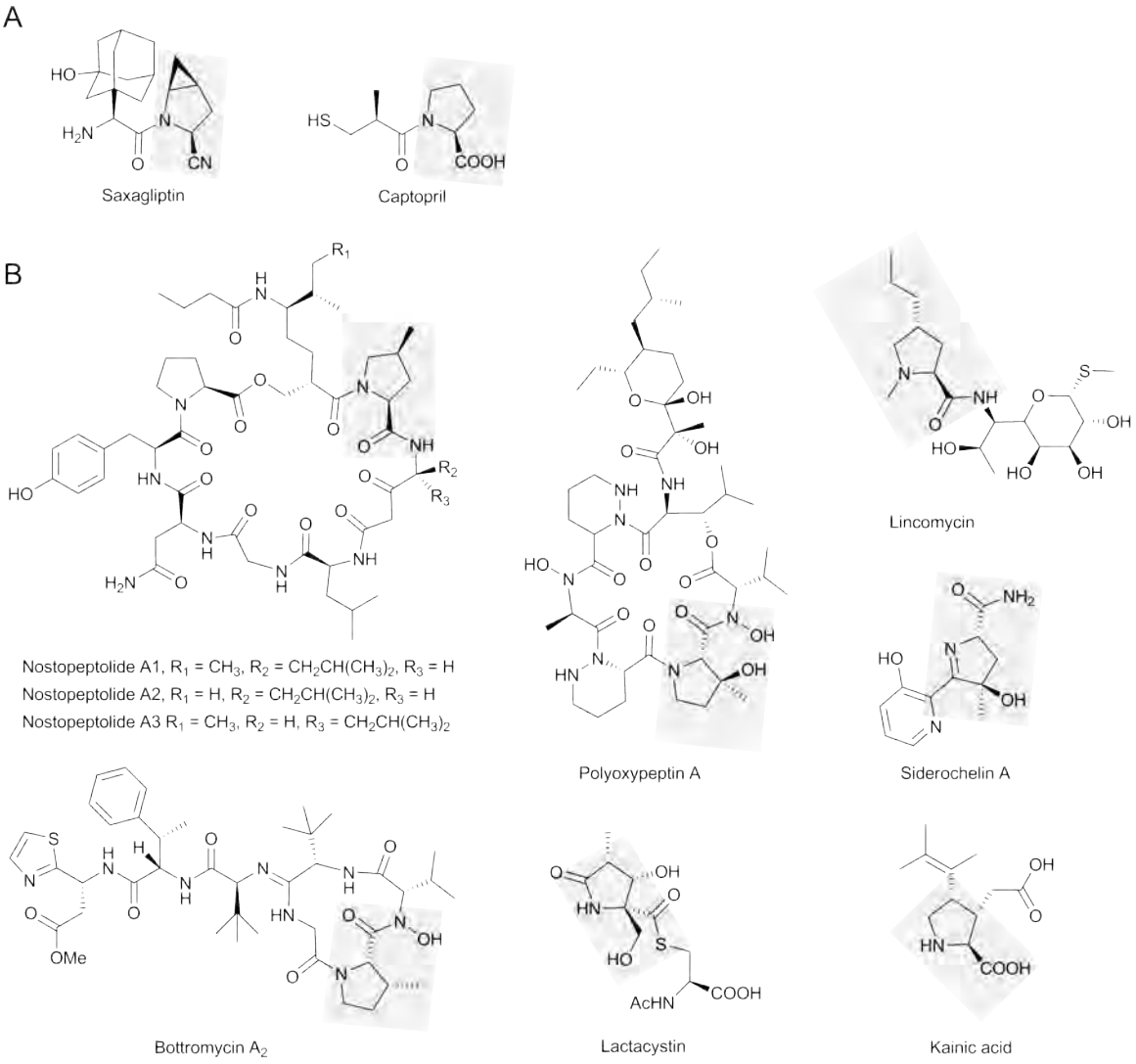


Figure 3.1: Proline derivatives in (A) clinically used drugs and (B) natural products.

cysteaminyl functionality. In addition, some of these L-GHP analogues did provide further evidence for the proposed C–C bond formation mechanism, as described in Chapter 4.

The results presented in this chapter were obtained in collaboration with Dr Christian Ducho, Dr John L. Sorensen and Dr Jasmin Mecinović who performed preliminary synthetic studies on L-GHP analogues. Part of the results described in this chapter are in reference [10].

3.2 Synthesis of Methyl-Substituted GHP Analogues

In order to test the ability of CMPS enzymes or their variants to accept methyl-substituted GHP analogues, derivatives carrying a methyl group at position C-2, C-3, C-4 or C-5 of the amino acid aldehyde substrate were synthesised in their protected form (Figure 3.2). The protecting groups were chosen to allow convenient deprotection for the CMPS assay. The *tert*-butyloxycarbonyl (Boc) and *tert*-butyl (*t*-Bu) groups were the protecting groups of choice, respectively of the amine and the carboxylate function. The deprotection conditions required to remove Boc and *t*-Bu groups have proven to be compatible with the CMPS assay, without further purification of the unstable free L-GHP derivatives [11, 12]. All of the final protected semialdehyde compounds described in this section were found to be difficult to characterise; In addition to showing very poor ionisation under MS analysis conditions, the ^{1}H and ^{13}C NMR spectra for the final protected products were found to be complicated due to some isomerism, as previously observed in the case of 4*S*-ethyl-5-hydroxy-L-proline [13]. The results for the CMPS-catalysed formation of substituted *t*-CMP products using the L-GHP analogues described below will be discussed in Chapter 4.

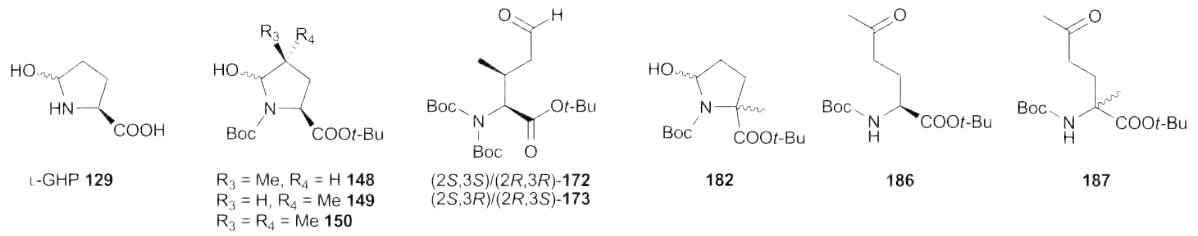
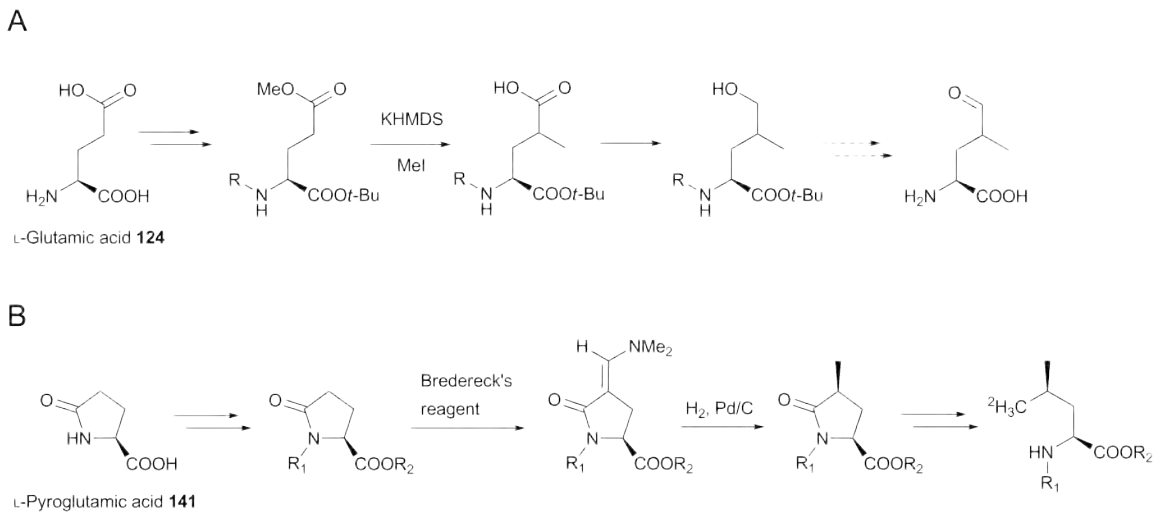


Figure 3.2: Structures of methyl-L-GHP analogues synthesised in this work. L-GHP derivatives carrying a methyl group at position C-2, C-3, C-4 or C-5 were synthesised to be tested as CMPS substrates.

3.2.1 Synthesis of C-4 Methyl-Substituted L-GHP

Various synthetic routes yielding 4-alkyl substituted derivatives of glutamic acid, pyroglutamic acid or proline have been described, either starting from L-glutamic acid **124** [14, 15] or from 5-oxo-L-proline **141**, commonly known as L-pyroglutamic acid [16, 17, 18].

Koskinen and Rapoport have reported the synthesis of 4-substituted prolines from L-glutamic acid **124** using an alkylation step involving a strong base and an alkyl halide [14]. The alkylation

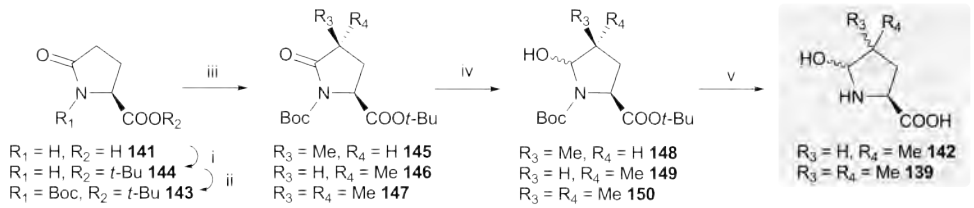


Scheme 3.1: Previously reported synthesis towards C-4 methyl-substituted L-glutamate derivatives. (A) Strategy described by Moore *et al.* [20]. KHMDS, potassium bis(trimethylsilyl)amide; MeI, methyl iodide; (B) Strategy described by August *et al.* [16]. Bredereck's reagent, *tert*-butoxy-bis(dimethyamino)methane; H_2, hydrogen gas; Pd/C, palladium(0) on activated charcoal.

of protected glutamate derivatives was systematically studied by Hanessian and Margarita [15] and the methodology successfully used by others, for example for the synthesis of labeled chloroleucine in the context of barbamide biosynthesis studies [19]. In the context of 4-methylproline biosynthesis investigations, Moore and coworkers obtained (2*S*)-5-hydroxyleucine using alkylation of the protected L-glutamate γ-methyl ester, followed by reduction (Scheme 3.1 A) [20]. A possibility that has not been tested experimentally would be to oxidise the protected 5-hydroxyleucine to the corresponding aldehyde which upon deprotection would yield an epimeric mixture of 4-methyl-L-GHP **142** (Scheme 3.1 A).

Young and coworkers have used 4-methyl-L-pyroglutamic acid as an intermediate in the stereospecific synthesis of labeled leucine [16]. Starting from 5-oxo-L-proline **141**, their synthesis of 4-methyl-L-pyroglutamic acid involved the formation of an enaminone using Bredereck's reagent, followed by hydrogenation (Scheme 3.1 B).

In the present work, 4-methyl-L-GHP **142** and 4,4-dimethyl-L-GHP **139** were obtained from 5-oxo-L-proline **141** in four steps by adapting a route described by Gu and Li (Scheme 3.2) [18]. This strategy involved the base-promoted alkylation of di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** with methyl iodide. Alternatively, methyl triflate could be used as an alkylating agent [21].



Scheme 3.2: Synthesis of C-4 methyl-substituted L-GHP analogues from L-pyroglutamate. (i) Isobutylene, H_2SO_4, 1,4-dioxane, -78 °C to rt, 6 d., 51%; (ii) $(Boc)_2O$, DMAP, Et_3N, CH_2Cl_2, 3 days, rt, 85%; (iii) LiHMDS, methyl triflate, toluene/THF 2:1, -78 °C, 6 h., 21-62%; (iv) Lithium triethylborohydride, THF, -78 °C, 84-87%; (v) 10% aq. HCOOH, 60 °C, 1 h, n.i.

The carboxylate functionality of 5-oxo-L-proline **141** was protected as a *tert*-butyl ester using isobutylene and sulphuric acid [22] to yield *tert*-butyl 5-oxo-L-prolinate **144**. Further protection of the lactam nitrogen with di-*tert*-butyl dicarbonate gave di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143**. The alkylation of di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** yielding a mixture of the C-4 monomethylated products *cis*- and *trans*-di-*tert*-butyl (2*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate, **145** and **146**, was found to be difficult to control. Although the optimised conditions described by Young and coworkers for this transformation were used [23], some overalkylation leading to di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate **147** was observed. Overalkylation likely occurs because the C-4 proton of the monoalkylated products **145** and **146** is more acidic than the C-4 protons of the starting material **143**. The stereochemical course of this alkylation step was not important in our case, as the final acidic deprotection step of either pure (4*S*)- or (4*R*)-isomer of di-*tert*-butyl (2*S*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate, **148** and **149**, respectively, would lead to C-4 epimerisation of 4-methyl-L-GHP **142** in aqueous medium. Nonetheless, the two diastereomers *cis*- and *trans*-di-*tert*-butyl (2*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate, **145** and **146**, were separated by chromatography and their stereochemistry assigned by comparison with a previous characterisation [23]. The results show a 4.5:1 preference for the *cis* product **145**, in a ratio similar to that reported in Charrier *et al.* [23]. When using an excess of base and an excess of methyl iodide, di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate **147** was obtained in good yield.

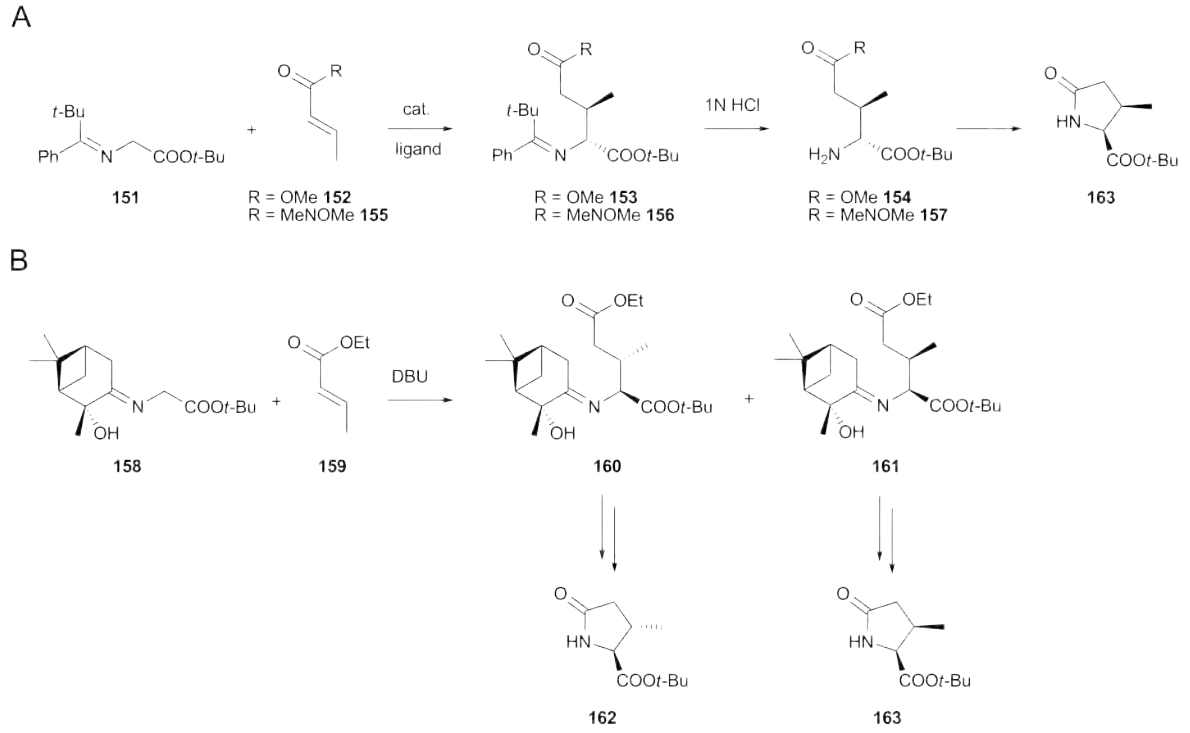
The reduction of the 5-membered lactams *cis*-di-*tert*-butyl (2*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **145**, *trans*-di-*tert*-butyl (2*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **146** and di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate **147** with lithium triethylborohydride

gave the corresponding hemiaminals di-*tert*-butyl (2*S*,4*S*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate **148**, di-*tert*-butyl (2*S*,4*R*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate **149** and di-*tert*-butyl (2*S*)-5-hydroxy-4,4-dimethylpyrrolidine-1,2-dicarboxylate **150** (Scheme 3.2). After purification by chromatography, these hemiaminals were stored at -20 °C until deprotection immediately prior to the CMPS assays described in Chapter 4.

3.2.2 Synthesis of C-3 Methyl-Substituted GHP

There are examples of synthetic routes towards 3-methylglutamate derivatives in the literature. Kobayashi and coworkers described a synthesis of enantiomerically enriched 3-methyl-D-glutamic acid from *N*-(*tert*-butylphenylmethylidene)glycine *tert*-butyl ester **151**, a glycine Schiff base, and methyl crotonate **152** using a chiral calcium complex as a catalyst [24, 25] (Scheme 3.3 A). The 1,4-addition reaction product **153** was deprotected to the free amine **154** which was found to spontaneously cyclise to give the corresponding pyroglutamate derivative. When methyl crotonate **152** was replaced by the Weinreb crotonamide **155**, the 1,4-addition reaction product **156** was also obtained in good yield. Acidic deprotection of the Schiff base to give the free amine **157**, followed by protection as a *N*-*tert*-butyloxycarbonyl might allow the conversion of the Weinreb amide into the corresponding aldehyde using Schwartz reagent (Cp_2ZrHCl) as reported for the synthesis of labelled GHP analogues [12]. In another example, Martinez and coworkers reported that the Michael addition of *tert*-butyl glycinate Schiff base **158**, carrying a *N*-chiral 2-hydroxy-3-pinanone auxiliary [26], to ethyl crotonate **159** afforded one set of diastereomer **160** and **161** [27]. Compounds **160** and **161** were separated and were further carried through to the enantiomerically pure (2*S*,3*S*)- and (2*S*,3*R*)-3-methyl glutamic acids *via* the corresponding pyroglutamate derivatives **162** and **163** (Scheme 3.3 B). After *N*-Boc protection, compounds **162** and **163** could be reduced with lithium triethylborohydride to the corresponding hemiaminals, in a similar way to the 4-methyl pyroglutamate derivatives described above (Section 3.2.1, Scheme 3.2).

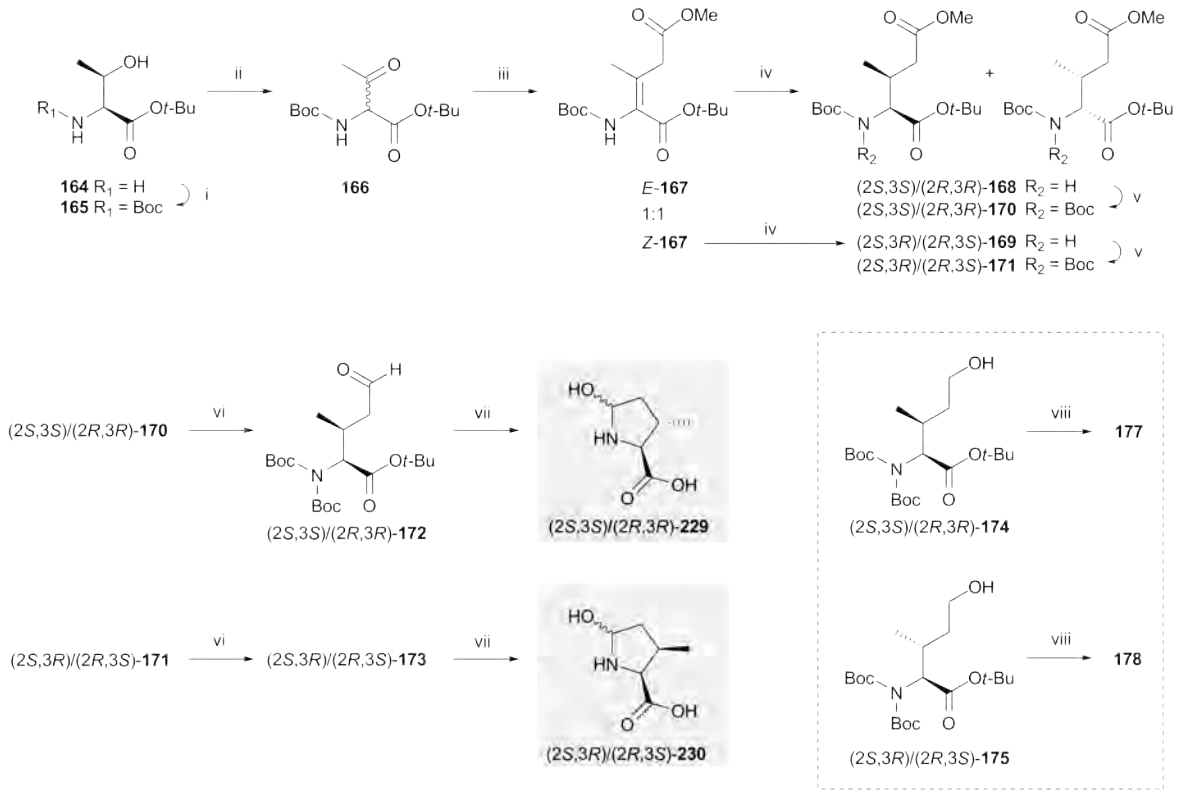
In the present work, two diastereomeric sets of 3-methyl-GHP enantiomers were obtained from commercially available *tert*-butyl L-threoninate **164** in seven steps, using a procedure involving a Horner-Wadsworth-Emmons (HWE) reaction (Scheme 3.4). A *tert*-butyloxycarbonyl group was installed on the amine group of *tert*-butyl L-threoninate **164** and oxidation of



Scheme 3.3: Previously reported synthesis towards C-3 methyl-substituted L-glutamate derivatives. (A) Strategy described by Kobayashi *et al.* [25]. cat, catalyst; HCl, hydrochloric acid; (B) Strategy described by Wehbe *et al.* [27]. DBU, 1,8-diazabicyclo-[5.4.0]undec-7-ene.

the protected amino acid **165** with Dess-Martin periodinane yielded the corresponding ketone *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-3-oxobutanoate **166**. Compound **166** was reacted under HWE conditions in the presence of methyl(triphenylphosphoranylidene)acetate to give a 1:1 mixture of chromatographically separable *E*- and *Z*-isomers of 1-*tert*-butyl 5-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-methylpent-2-enedioate **167**. These two isomers were carefully characterised by NMR analysis, including 1D NOE experiments, in order to assign the *E* or *Z* configuration (Figure 3.3). Key NOE correlation signals were observed with the N*H* proton. In the case of *Z*-**167**, irradiation of the CH_2 methylene protons at 3.26 ppm gave a strong NOE with the N*H* proton, as well as weaker signals for the methyl ester CH_3 and the methyl CH_3 protons. For the same compound, irradiation of the C-3 methyl protons gave no signal for the N*H* proton (Figure 3.3 A). The opposite observation was made for *E*-**167** (Figure 3.3 B).

Z-**167** and *E*-**167** were separately subjected to hydrogenation to give two diasteromeric sets of non-separable protected 3-methylglutamic acid isomers (2*S*,3*S*)/(2*R*,3*R*)-**168** and (2*S*,3*R*)/(2*R*,3*S*)-**169**. As anticipated, the optical rotation for both mixtures of (2*S*,3*S*)/(2*R*,3*R*)-**168** and (2*S*,3*R*)/(2*R*,3*S*)-**169**



Scheme 3.4: Synthesis of C-3 methyl-substituted L-GHP analogues. (i) $(Boc)_2O$, Et_3N, CH_2Cl_2, rt, 97%; (ii) Dess-Martin periodinane, CH_2Cl_2, rt, 96%; (iii) Methyl(triphenylphosphoranylidene)acetate, toluene, reflux, 62%; (iv) H_2 (g), palladium on activated charcoal, EtOAc, rt, 98%; (v) $(Boc)_2O$, DMAP, Et_3N,, CH_2Cl_2, 85%; (vi) DIBAL-H, Et_2O, -78 °C, 70%; (vii) 10% aq. HCOOH, 60 °C, 1 h, n.i.; (viii) Dess-Martin periodinane, CH_2Cl_2, 90 min, 88%.

was found to be zero. Further protection of the amino group with a second *N*-*tert*-butyloxycarbonyl group allowed the reduction of the methyl ester of (2*S*,3*S*)/(2*R*,3*R*)-1-*tert*-butyl 5-methyl *N*,*N*-bis(*tert*-butoxycarbonyl)-3-methylglutamate **170** and (2*S*,3*R*)/(2*R*,3*S*)-1-*tert*-butyl 5-methyl *N*,*N*-bis(*tert*-butoxycarbonyl)-3-methylglutamate **171** to the corresponding aldehydes (2*S*,3*S*)/(2*R*,3*R*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **172** and (2*S*,3*R*)/(2*R*,3*S*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **173**. The alcohol by-products (2*S*,3*S*)/(2*R*,3*R*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-hydroxyisoleucinate **174** and (2*S*,3*R*)/(2*R*,3*S*)-**175** could be isolated and converted to the desired aldehydes **172** and **173** using Dess-Martin periodinane oxidation (dotted frame, Scheme 3.4). Aldehydes **172** and **173** were stored at -20 °C until further use.

3.2.3 Synthesis of C-2 Methyl-Substituted GHP

The synthesis of 2-methyl-L-GHP was investigated in order to test the possibility of introducing steric bulk and a pre-installed quaternary center into the *t*-CMP product *via* CMPS catalysis. The presence of a quaternary center at C-2 would also confirm that the CMPS mechanism cannot involve proton abstraction at this position.

A procedure reported by Seebach and coworkers that was proposed to allow 'self-reproduction' of chirality was investigated in an initial attempt to obtain enantioenriched (2*S*)-5-hydroxy-2-methylpyrrolidine-2-carboxylic acid **176** [28]. The authors showed that protection of L-proline **134** as the bicyclic product (3*R*,5*S*)-3-*tert*-butyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **177** using pivalaldehyde allows a two-step deprotonation followed by alkylation procedure without losing the stereochemistry at C-5 (Scheme 3.5 A) [28]. Deprotonation of (3*R*,5*S*)-3-*tert*-butyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **177** at C-5 to yield the non-racemic lithium enolate (Scheme 3.5 A) was carried out with lithium diisopropylamide. The alkylation of this stabilised anion with methyl iodide yielded the bicyclic compound **178** containing a chiral C-5 quaternary center.

A variant of the self-reproduction of chirality methodology was described by Amedjkouh and Ahlberg, using trichloroacetaldehyde for the protection of L-proline **134**, rather than pivalaldehyde (Scheme 3.5 B) [29]. These authors also reported the protection of 5-oxo-L-proline **141** using trichloroacetaldehyde to form the corresponding (2*R*,5*S*)-2-trichloromethyl-1-aza-3-oxabicyclo-[3.3.0]-octan-4,8-dione **179**. Self-reproduction of chirality has been applied to the synthesis of various

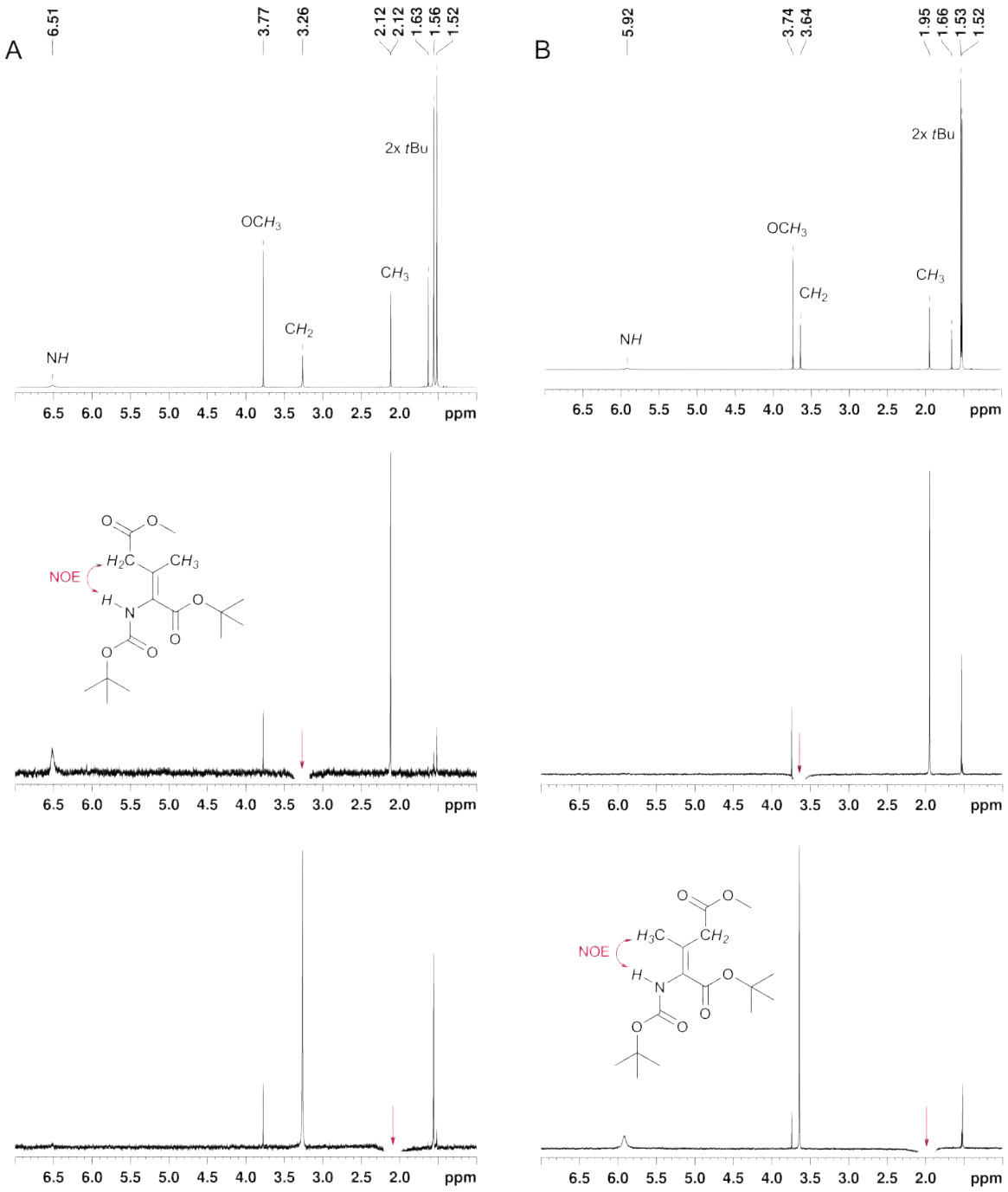
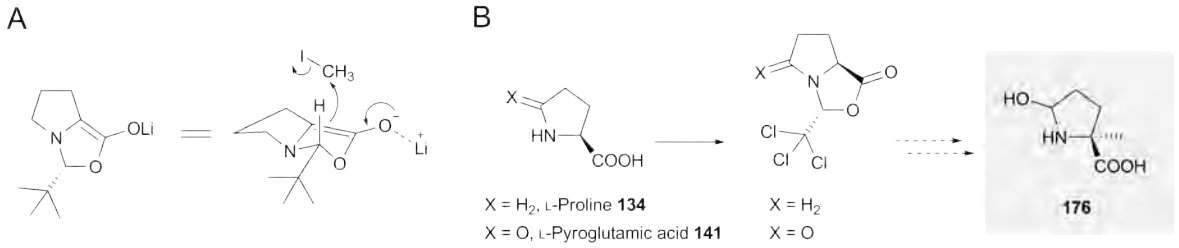
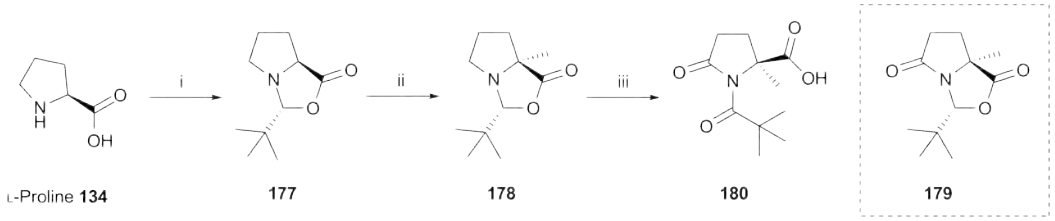


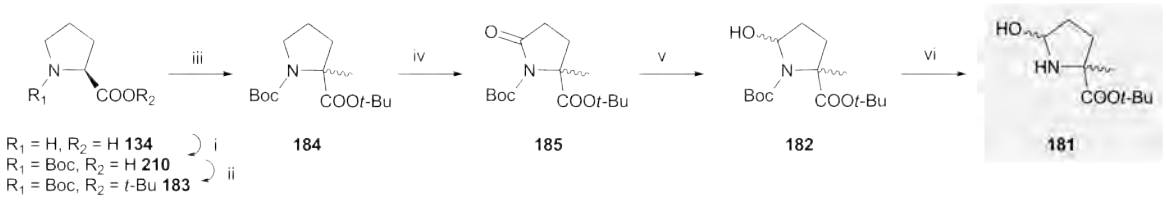
Figure 3.3: NOE analysis for the two isomers of 1-*tert*-butyl 5-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-methylpent-2-enedioate. The double bond geometry of *Z*-**167** and *E*-**167** was determined using 1D NOE NMR experiments recorded at 500 MHz in $CDCl_3$. (A) NOE analysis for compound *Z*-**167**. From top to bottom: ^{1}H NMR spectrum; irradiation at 3.26 ppm (CH_2) shows NOE with the N*H* signal (6.51 ppm); irradiation at 2.12 ppm (CH_3) does not show any NOE with the N*H* signal (6.51 ppm); (B) NOE analysis for compound *E*-**167**. From top to bottom: ^{1}H NMR spectrum; irradiation at 3.64 ppm (CH_2) does not show any NOE with the N*H* signal (5.92 ppm); irradiation at 1.95 ppm (CH_3) shows NOE with the N*H* signal (5.92 ppm).



Scheme 3.5: Proposed mechanism for alkylation with "self-reproduction" of chirality. (A) Lithium enolate conformation proposed by Seebach to explain the "self-reproduction" of chirality observed under alkylation conditions [28]; (B) Trichloroacetaldehyde protection of proline and pyroglutamic acid described by Amedjkouh and Ahlberg [29].



Scheme 3.6: Unsuccessful route towards enantiomerically enriched (2*S*)-2-GHP. Self-reproduction of chirality allowed to access chiral (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-*c*][1,3]oxazol-1-one **178** but not (7a*S*)-3-*tert*-butyl-7a-methyldihydro-1*H*-pyrrolo[1,2-*c*][1,3]oxazole-1,5(6*H*)-dione **179**. (A) (i) Trimethylacetaldehyde, cat. CF_3COOH, toluene, Dean-Stark reflux, 7 d; (ii) Lithium diisopropylamide, MeI, THF, -78 °C–r.t., 6 h, 23%; (iii) Ruthenium tetroxide *in situ* from ruthenium(IV) oxide and sodium periodate, 2:3 ethyl acetate/water, rt, 16 h, 85%.

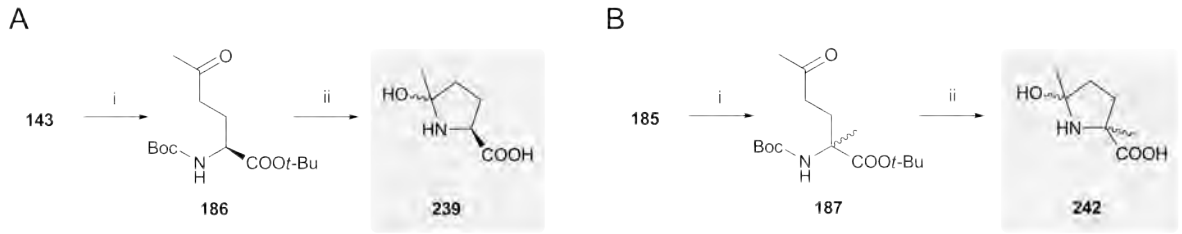


Scheme 3.7: Synthesis of racemic C-2-methyl substituted L-GHP analogue. (i) $(Boc)_2O$, Et_3N, 1,4-dioxane, rt; (ii) EDC HCl, DMAP, 1:1 *t*-BuOH/CH_2Cl_2, rt, 78% over 2 steps; (iii) *n*-BuLi, DIPEA, MeI, THF, -78 °C, 65%; (iv) Ruthenium tetroxide *in situ* from ruthenium(IV) oxide and sodium periodate, 1:4 EtOAc/water, rt, 16 h., 85%; (v) Lithium triethylborohydride, THF, -78 °C, 68%; (vi) 10% aq. HCOOH, 60 °C, 1 h, n.i.

chiral proline analogues, for asymmetric organocatalysis [30] and in the context of conformationally restricted natural product analogues [31, 32].

In an attempt to obtain enantioenriched C-2 methyl-L-GHP **176**, the Seebach procedure was reproduced. L-Proline **134** was reacted with pivalaldehyde under a nitrogen atmosphere using a Dean-Stark apparatus, in order to trap the water produced by the condensation reaction and keep the system under strictly anhydrous conditions. The extremely moisture sensitive bicyclic product **177** was found to undergo rapid hydrolysis back to L-proline **134** when in contact with air. Accordingly, only partial characterisation of **177** was carried out, and this compound was immediately subjected to α-alkylation as described [28]. After alkylation using lithium diisopropylamide as a base, and methyl iodide as an alkylating agent, the product containing a chiral C-5 quaternary center **178** was found to be more stable towards hydrolysis than its precursor **177**. Indeed, (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **178** was purified by distillation and fully characterised. The oxidation procedure using ruthenium tetroxide that has been described for various proline derivatives [33] did not yield any of the expected bicyclic product **179** (dotted frame, Scheme 3.6), but yielded instead the bis-oxidised product 1-(2,2-dimethylpropanoyl)-2-methyl-5-oxo-L-proline **180**.

A procedure yielding racemic 2-methyl-GHP **181** was investigated instead. Di-*tert*-butyl 5-hydroxy-2-methylpyrrolidine-1,2-dicarboxylate **182** was obtained from L-proline **134** in six steps (Scheme 3.7). The key transformation was the alkylation at the C-2 position of protected L-proline **183** through the generation of a tertiary carbanion using butyllithium as a base, and therefore loss of stereochemistry at this center, as reported [31]. Alkylation of this carbanion was carried



Scheme 3.8: Synthesis of C-5 methyl-substituted L-GHP analogues. (A) (i) MeMgBr, THF, 87%; (ii) 10% aq. HCOOH, 60 °C, 1 h, n.i.; (B) (i) MeMgBr, THF, 84%; (ii) 10% aq. HCOOH, 60 °C, 1 h, n.i.

out using methyl iodide to give protected racemic 2-methylproline **184**. Oxidation using catalytic ruthenium (VIII) tetroxide formed *in situ* from ruthenium (IV) dioxide and an excess of aqueous sodium periodate in a biphasic reaction mixture, gave the racemic pyroglutamate product di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185**. This oxidation procedure was described in Yoshifuji *et al.* [33] and was initially adapted from the work of Sheehan and Tulis [34]. This first practical example of reaction converting L-proline into L-pyroglutamic acid was demonstrated to be compatible with a variety of *N*-protecting groups [33]. Reduction of the di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185** with lithium triethylborohydride [35] yielded di-*tert*-butyl 5-hydroxy-2-methylpyrrolidine-1,2-dicarboxylate **182** which was stored at -20 °C until deprotection for CMPS assays, as described in Chapter 4.

3.2.4 Synthesis of C-5 Methyl-Substituted L-GHP Analogues

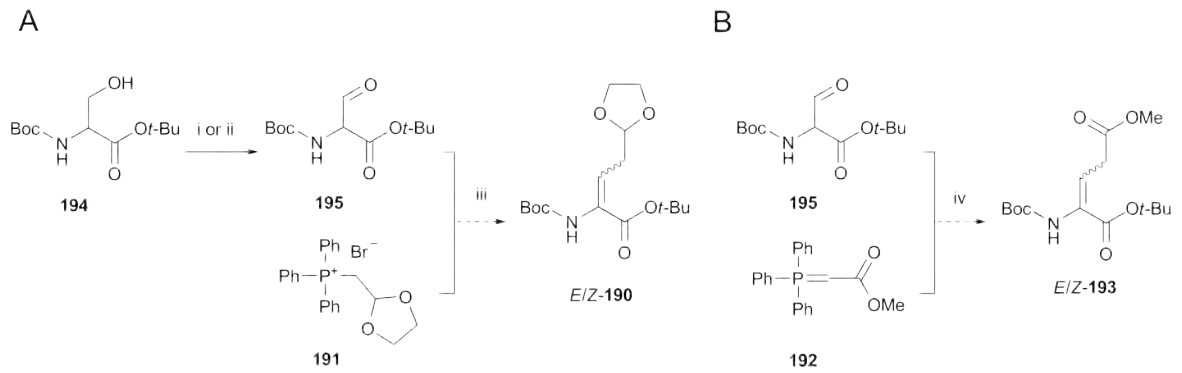
The synthesis of 5-methyl-L-GHP was required to complete a systematic investigation of methyl-substituted L-GHP CMPS substrates. It would also enable testing the possibility of CMPS being able to install a new quaternary center *via* C–C bond formation. *tert*-Butyl *N*-(*tert*-butoxycarbonyl)-5-oxo-L-norleucinate **186** was synthesised by Grignard addition of methyl magnesium bromide to di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** followed by deprotection (Scheme 3.8 A). *tert*-Butyl *N*-(*tert*-butoxycarbonyl)-2-methyl-5-oxonorleucinate **187** was synthesised using the same reaction conditions, starting from di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185**, as described in the context of racemic 2-methyl-L-GHP **181** synthesis (Scheme 3.8 B).

3.3 3-Cysteaminyl Substituted GHP Analogues

An interest in exploring the tolerance of CMPS toward substituants at the C-3 position of the L-GHP substrate emerges from the presence of a cysteamine side chain at the equivalent position in thienamycin **116** (Chapter 2, Figure 2.4). So far, the enzyme encoding for the introduction of this side chain in the thienamycin biosynthesis has not been identified. Three enzymes, ThnR, ThnH and ThnT, are proposed to be involved in the stepwise truncation of coenzyme A **133** to cysteamine (Chapter 2, Scheme 2.7) [36]. However, these results do not exclude the possibility that the cysteaminyl side chain is introduced before the CMPS-catalysed step in the thienamycin biosynthesis pathway. In order to investigate this possibility, a synthetic route towards two C-3 cysteaminyl-substituted L-GHP analogues was investigated.

A key intermediate envisaged for the synthesis of 3-cysteaminyl-L-GHP **188** and 3-(*N*-acetyl)-cysteaminyl-L-GHP **189** was the protected α,β-dehydroglutamic acid aldehyde **190** (Scheme 3.9 and 3.10). Because of its conjugated unsaturated system, *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** was anticipated to be a good substrate for the nucleophilic conjugate addition of thiols. The double bond formed during the Wittig reaction yielding 1-*tert*-butyl 5-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-methylpent-2-enedioate **167** described in the context of 3-methyl-GHP synthesis was found to spontaneously isomerise from the 3,4- to the 2,3-position (Section 3.2.2, Scheme 3.4). This observation suggested that *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** could be accessed from protected serine using a similar oxidation followed by Wittig coupling strategy (Scheme 3.9). This route was considered to be attractive because both the phosphonium ylides **191** and **192** were commercially available, yielding **190** and **193**, respectively.

Unfortunately, the oxidation of commercially available *tert*-butyl *N*-(*tert*-butoxycarbonyl)-L-serinate **194** using Swern [37] or Dess-Martin [38] conditions failed to produce the desired amino acid aldehyde *tert*-butyl *N*-(*tert*-butoxycarbonyl)-3-oxo-L-alaninate **195** in suitably pure form. Despite the presence of a signal characteristic for an aldehyde proton (δ ^{1}H = 9.70 ppm) in the ^{1}H NMR spectrum of the crude reaction mixture, no pure product could be isolated after chromatography, possibly due to instability of the aldehyde on silica. Repeated attempts to carry out the Wittig reaction on the crude aldehyde **195** after minimal aqueous work-up were unsuccessful. In contrast, the same oxidation

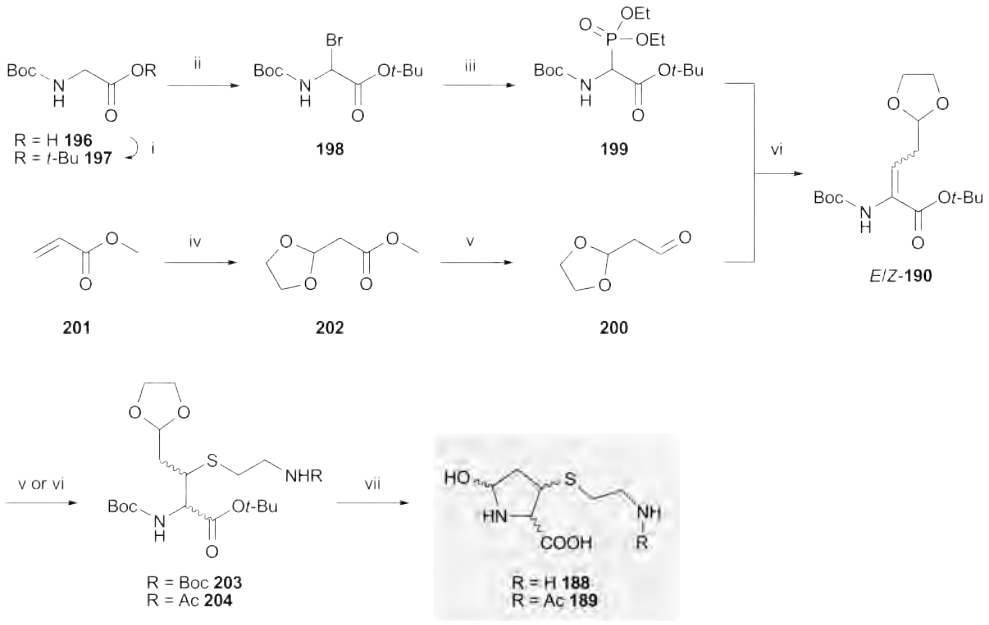


Scheme 3.9: Unsuccessful syntheses of *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** and methyl ester **193**. (i) Dess-Martin periodinane, CH_2Cl_2, rt, N.D.; (ii) $SOCl_2$, DMSO, Et_3N, CH_2Cl_2, -78 °C, 1 h, N.D.; (iii) n-BuLi, 0 °C–65 °C, 4 h, n.i.; (iv) 125 °C, 4 h, n.i.

conditions were successfully applied to the related protected L-threonine **165** (Section 3.2.2).

Thus, instead of starting from serine, compound **190** was synthesised starting from [(*tert*-butoxycarbonyl)amino]acetic acid **196** (*N*-Boc glycine, Scheme 3.10 A), following a route described in a study where the *Z*- and *E*-isomers of *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** were used as precursors for deuterated L-GHP analogues [39]. Protection of the carboxylate of *N*-Boc glycine **196** using isobutylene and sulphuric acid [22], or *tert*-butanol and *N*,*N*-DMF-di(neopentyl)acetal [40], gave *tert*-butyl [(*tert*-butoxycarbonyl)amino]acetate **197** in moderate yield. The latter method, initially described by Eschenmoser and coworkers for benzyl esters formation [41], and applied in the context of *tert*-butyl ester formation in Ducho *et al.* [39] proved to be the most efficient. Radical bromination of **197** using strictly one equivalent of *N*-bromosuccinimide (NBS) and catalytic benzoylperoxide yielded the protected α-bromoglycine **198**. This unstable α-bromoamino acid was used for the next step without purification. The formation of *tert*-butyl [(*tert*-butoxycarbonyl)amino](diethoxyphosphoryl)acetate **199** using triethylphosphite and *tert*-butyl bromo[(*tert*-butoxycarbonyl)amino]acetate **198** under Michaelis-Arbuzov reaction conditions was found to proceed in a near quantitative yield from *tert*-butyl [(*tert*-butoxycarbonyl)amino]acetate **197**.

The 1,3-dioxolane monoprotected dialdehyde **200** was obtained in two steps from methyl acrylate **201** (Scheme 3.10). The starting material alkene was converted to the acetal form of malonate semialdehyde methyl ester **202** according to a previously reported procedure [42]. Reduction of the methyl ester using a slight excess of diisobutylaluminium hydride gave near quantitative conversion to the corresponding aldehyde **200** as determined by ^{1}H NMR analysis of the crude. The aldehyde **200**



Scheme 3.10: Synthesis of C-3 cysteaminyl substituted L-GHP Analogues. (i) *N*,*N*-DMF-di(neopenthyl)acetal, *t*-BuOH, toluene, reflux, 5 h, 60%; (ii) Benzoyl peroxide, NBS, CCl_4, reflux, 4 h, n.i.; (iii) $P(OEt)_3$, dry CH_2Cl_2, 93% over 2 steps; (iv) Methyl acrylate, ethylene glycol, DME, cat. $PdCl_2$, $CuCl_2$, O_2, 50 °C, 24 h, 74%; (v) 1M DIBAL in THF, dry CH_2Cl_2, -78 °C, 6 h, 72%; (vi) KO*t*-Bu, dry THF, -78 °C to rt, 18 h, 91%; (v) SNAc, K_2CO_3, CH_2Cl_2, rt, 3 h, 65%; (vi) *N*-Boc cysteamine, K_2CO_3, CH_2Cl_2, rt, 3 h, 62%; (vii) 10% HCOOH, 60 °C, 1h, n.i.

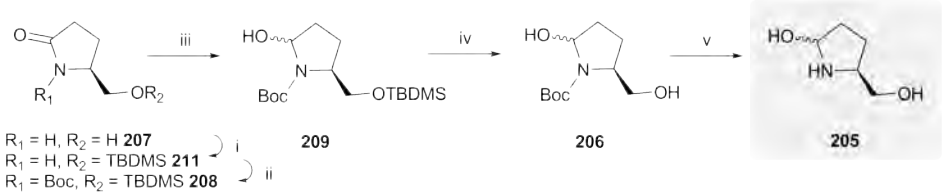
was found to be unstable and therefore directly used in the subsequent Horner-Wadsworth-Emmons (HWE) reaction without further purification.

When aldehyde **200** and *tert*-butyl [(*tert*-butoxycarbonyl)amino](diethoxyphosphoryl)acetate **199** were reacted in the presence of potassium *tert*-butoxide as a base, the HWE reaction gave a 1:15 mixture of *E*- and *Z*-isomers of *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** with an overall yield of 91%. The two isomers could be separated by silica chromatography; *E*-**190** was found to elute prior to *Z*-**190**. The *E*- and *Z*-configurations of the two isomers of *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate **190** were assigned based on the observation of NOE signals in ^{1}H NMR spectroscopy as reported [39]. The nucleophilic conjugate addition of *N*-Boc cysteamine or *N*-acetylcysteamine (SNAc) was carried out on *tert*-butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate *Z*-**190** under mild basic conditions. This reaction probably yields a mixture of four non-separable stereoisomers, giving complicated ^{1}H and ^{13}C NMR spectra. Compounds **203** and **204** were stored at -20 °C until deprotection to yield 3-cysteaminyl-L-GHP **188** and 3-(*N*-acetyl)cysteaminyl-L-GHP **189** which were tested in the CMPS assay without purification.

3.4 Synthesis of an L-GHP Analogue Lacking the Carboxylate Function

The α-carboxylate moiety of L-GHP is probably important for substrate positioning in the CMPS active site considering that the D-GHP enantiomer is not a substrate for CMPS enzymes [12]. In addition, the information obtained from a CMPS crystal structure in complex with L-proline suggests that the carboxylate is forming hydrogen-bonding interactions (Section 6.3.2, Figure 6.5 B). In order to test if an hydrogen bond interaction is sufficient for positioning the pyrroline substrate for productive CMPS catalysis, the synthesis of (5*S*)-5-(hydroxymethyl)pyrrolidin-2-ol **205**, a L-GHP analogue carrying an hydroxymethyl moiety in place of the carboxylate group, was performed.

tert-Butyl (5*S*)-2-hydroxy-5-(hydroxymethyl)pyrrolidine-1-carboxylate **206** was synthesised from the commercially available (*S*)-5-(hydroxymethyl)pyrrolidin-2-one **207** in four steps and an overall yield of 32% (Scheme 3.11). The two initial protection steps have been described [43, 44]. The reduction of the lactam **208** to the hemiaminal **209** was performed using lithium triethylborohydride as described above for the synthesis of C-4 and C-2 methyl-substituted GHP analogues (Sections 3.2.1 and 3.2.3). Compound **209** was found to show complicated ^{1}H NMR patterns and low intensity ^{13}C NMR signals, probably due to the presence of conformational isomers, as often observed in compounds containing *N*-Boc groups. The *tert*-butyldimethylsilyl ether protecting group of compound **209** was removed and the *N*-Boc protected L-GHP analogue **206** was stored at -20 °C until deprotection for the CMPS assay.



Scheme 3.11: Synthesis of (5*S*)-5-(hydroxymethyl)pyrrolidin-2-ol. (i) TBDMSCl, rt, 24 h, 99%; (ii) $(Boc)_2O$, DMAP, Et_3N, rt, 24 h, 70%; (iii) $LiEt_3BH$, -78 °C, 6 h, 57%; (iv) TBAF, 0 °C, 2h, 81%; (v) 10% HCOOH, 60 °C, 1h, n.i.

3.5 Summary for Chapter 3

3.5.1 Methyl-Substituted L-GHP Analogues Can be Synthesised from L-Proline, L-Tyrosine or L-Pyroglutamic Acid

The syntheses of protected methyl-substituted amino acid semialdehydes from L-proline, L-tyrosine or L-pyroglutamic acid was successfully carried out and is described in Section 3.2 (Scheme 3.12). These CMPS substrate analogues with one or two methyl groups at every position of the proline ring added diversity to the pool of CMPS substrates. The protected amino acid semialdehyde species synthesised were found to be difficult to characterise, showing low ionisation properties under MS analysis conditions, and complicated ^{1}H NMR patterns as well as ^{13}C NMR signals of low intensity. Hence, they were deprotected and the free amino acid semialdehydes were used immediately in the CMPS assay without characterisation.

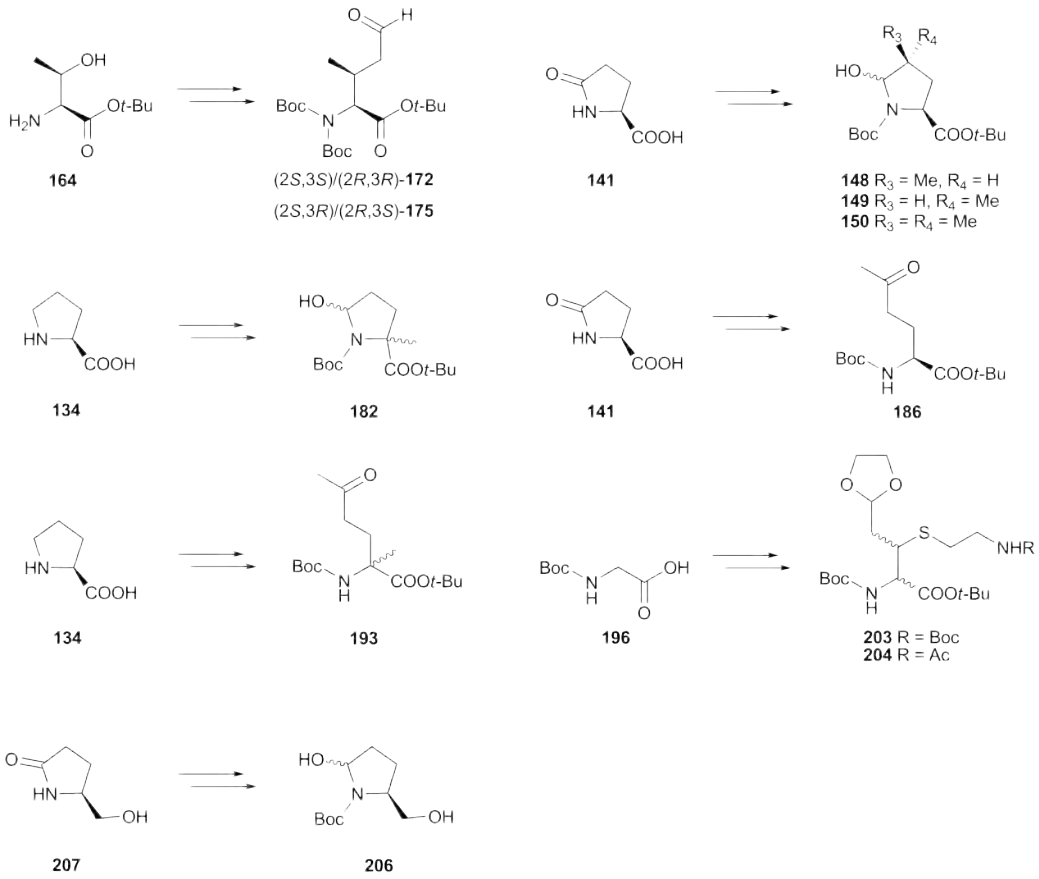
3.5.2 Perspectives

Several other amino acid semialdehyde substrates could be of interest to further expend the utility of CMPS as biocatalysts and possibly produce chemical compounds with interesting properties. The synthesis of 3,3-dimethyl-L-GHP would be particularly relevant to expend investigations into methyl-substituted GHP analogues. Access to synthetic 3,3-dimethyl-L-GHP would enable testing of the ability of CMPS enzymes to produce the corresponding 3,3-dimethyl-*t*-CMP substrate for CarA β-lactam formation (Scheme 3.13 A). 2,2-Dimethylcarbapenam would be a first step towards the generation of analogues of penicillins, e.g. penicillin G (Scheme 3.13 B).

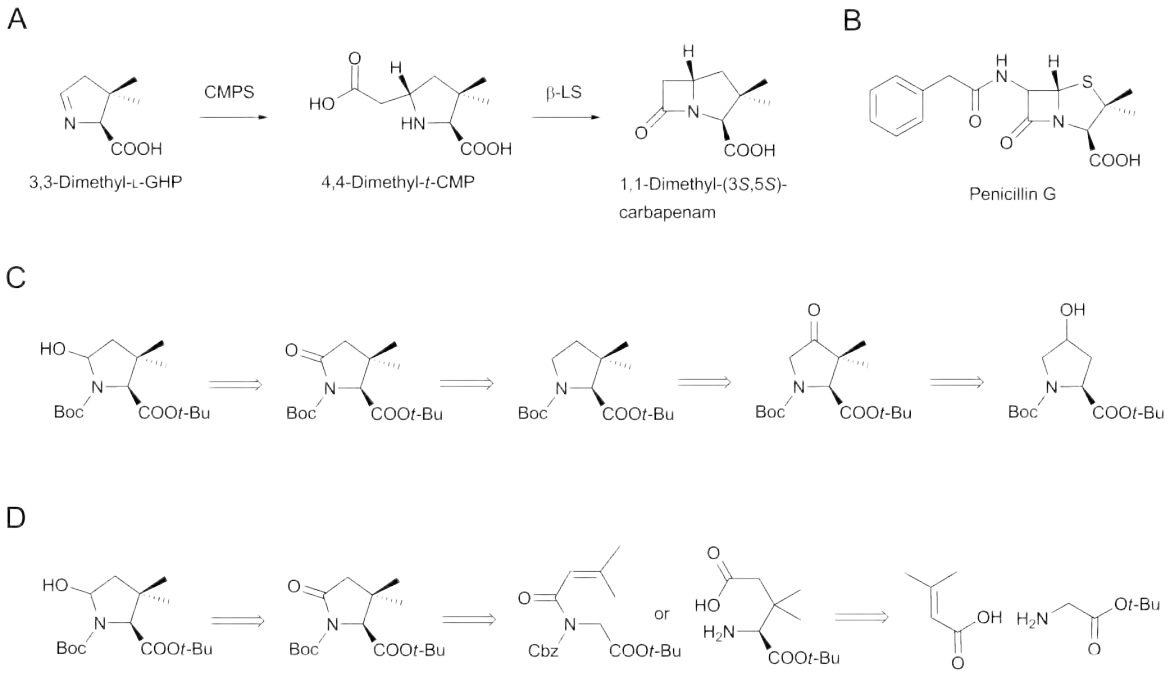
Two different synthetic routes could be envisaged for 3,3-dimethyl-L-GHP. One is the synthesis of 3,3-dimethyl-L-GHP *via* 3,3-dimethylproline [45], starting from 4-hydroxyproline (Scheme 3.13 C). Oxidation of 4-hydroxy-L-proline to the keto product would allow alkylation at position C-3, although this step might have low regioselectivity and yield the C-2 alkylated product as a side-product. Reduction to the alcohol and elimination of the hydroxyl group after activation would give 3,3-dimethylproline as described by Sharma and Lubell [45]. Oxidation at position C-5 this time, using the ruthenium tetroxide procedure described in Section 2.2.2 for the protected racemic 2-methylproline **184** would yield the corresponding pyroglutamate analogue that can be reduced to

the protected semialdehyde of interest.

Alternatively, the 1,4-addition of a Schiff base to a 3-methylbut-2-enoic acid derivative using a strategy similar to the procedure described by Kobayashi and coworkers for the synthesis of 3-methyl glutamic acid derivatives might yield the protected 3,3-dimethyl glutamic acid [25, 27]. The same approach could also be attempted after forming an amide bond, yielding the 3,3-dimethylpyroglutamate derivative *via* intramolecular 1,4-addition (Scheme 3.13 D).



Scheme 3.12: Protected methyl-substituted amino acid semialdehydes synthesised in this work



Scheme 3.13: Synthesis of 3,3-dimethyl-L-GHP could allow the CMPS catalysed generation of the 3,3-dimethyl-*t*-CMP CarA substrate. (A) Possible biocatalytic route towards 2,2-dimethyl carbapenam; (B) Structure of penicillin G; (C) Retrosynthetic analysis for 3,3-dimethyl-L-GHP.

References

[1] Augeri, D. J., et al. (2005) Discovery and preclinical profile of saxagliptin (BMS-477118): a highly potent, long-acting, orally active dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem*, **48**, 5025–5037.

[2] Gavras, H., Brunner, H. R., Turini, G. A., Kershaw, G. R., Tifft, C. P., Cuttelod, S., Gavras, I., Vukovich, R. A., and McKinstry, D. N. (1978) Antihypertensive effect of the oral angiotensin converting-enzyme inhibitor SQ 14225 in man. *N Engl J Med*, **298**, 991–995.

[3] Mauger, A. B. (1996) Naturally occurring proline analogues. *J Nat Prod*, **59**, 1205–1211.

[4] Liu, W. C., Fisher, S. M., Wells, J. S., Ricca, C. S., Principe, P. A., Trejo, W. H., Bonner, D. P., Gougoutos, J. Z., Toeplitz, B. K., and Sykes, R. B. (1981) Siderochelin, a new ferrous-ion chelating agent produced by *Nocardia*. *J Antibiot (Tokyo)*, **34**, 791–799.

[5] Nakamura, S., Tanaka, N., and Umezawa, H. (1966) Bottromycin A1, A2 and their structures. *J Antibiot (Tokyo)*, **19**, 10–12.

[6] Shimamura, H., Gouda, H., Nagai, K., Hirose, T., Ichioka, M., Furuya, Y., Kobayashi, Y., Hirono, S., Sunazuka, T., and Omura, S. (2009) Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew Chem Int Ed Engl*, **48**, 914–917.

[7] Nitta, I., Watase, H., and Tomiie, Y. (1958) Structure of kainic acid and its isomer, allokainic acid. *Nature*, **181**, 761–762.

[8] Omura, S., Fujimoto, T., Otoguro, K., Matsuzaki, K., Moriguchi, R., Tanaka, H., and Sasaki, Y. (1991) Lactacystin, a novel microbial metabolite, induces neuritogenesis of neuroblastoma cells. *J Antibiot (Tokyo)*, **44**, 113–116.

[9] Omura, S., Matsuzaki, K., Fujimoto, T., Kosuge, K., Furuya, T., Fujita, S., and Nakagawa, A. (1991) Structure of lactacystin, a new microbial metabolite which induces differentiation of neuroblastoma cells. *J Antibiot (Tokyo)*, **44**, 117–118.

[10] Hamed, R. B., Henry, L., Gomez-Castellanos, J. R., Mecinović, J., Ducho, C., Sorensen, J. L., Claridge, T. D. W., and Schofield, C. J. (2012) Crotonase catalysis enables flexible production of functionalized prolines and carbapenams. *J Am Chem Soc*, **134**, 471–479.

[11] Sleeman, M. C., Sorensen, J. L., Batchelar, E. T., McDonough, M. A., and Schofield, C. J. (2005) Structural and mechanistic studies on carboxymethylproline synthase (CarB), a unique member of the crotonase superfamily catalyzing the first step in carbapenem biosynthesis. *J Biol Chem*, **280**, 34956–34965.

[12] Sorensen, J. L., Sleeman, M. C., and Schofield, C. J. (2005) Synthesis of deuterium labelled L- and D-glutamate semialdehydes and their evaluation as substrates for carboxymethylproline synthase (CarB)–implications for carbapenem biosynthesis. *Chem Commun*, **9**, 1155–1157.

[13] Moody, C. M. and Young, D. W. (1997) Stereospecific synthesis of naturally-occurring 4-alkylideneglutamic acids, 4-alkylglutamates and 4-alkylprolines. *J Chem Soc, Perkin Trans 1*, **1997**, 3519–3530.

[14] Koskinen, A. M. P. and Rapoport, H. (1989) Synthesis of 4-substituted prolines as conformationally constrained amino acid analogs. *J Org Chem*, **54**, 1859–1866.

[15] Hanessian, S. and Margarita, R. (1998) 1,3-asymmetric induction in dianionic allylation reactions of amino acid derivatives - synthesis of functionally useful enantiopure glutamates, pipecolates and pyroglutamates. *Tetrahedron Lett*, **39**, 5887–5890.

[16] August, R. A., Khan, J. A., Moody, C. M., and Young, D. W. (1996) Stereospecific synthesis of (2*S*,4*R*)-[5,5,5-$^{2}H_3$]-leucine. *J Chem Soc, Perkin Trans 1*, **1996**, 507–514.

[17] Katoh, M., Mizutani, H., and Honda, T. (2005) Stereoselective synthesis of Nuphar quinolizidine alkaloid, (-)-deoxynupharidine. *Tetrahedron Lett*, **46**, 5161–5163.

[18] Gu, Z.-Q. and Li, M. (2003) A concise synthesis of (2*S*,4*R*)- and (2*S*,4*S*)-4-methylglutamic acid. *Tetrahedron Lett*, **44**, 3203–3205.

[19] Gerwick, W. H., Leslie, P., Long, G., Marquez, B. L., and Willis, C. L. (2003) [6-^{13}C]-(2*S*,4*S*)-5-Chloroleucine: Synthesis and incubation studies with cultures of the cyanobacterium, *Lyngbya majuscula*. *Tetrahedron Lett*, **44**, 285–288.

[20] Luesch, H., Hoffmann, D., Hevel, J. M., Becker, J. E., Golakoti, T., and Moore, R. E. (2003) Biosynthesis of 4-methylproline in cyanobacteria: Cloning of *nosE* and *nosF* genes and biochemical characterization of the encoded dehydrogenase and reductase activities. *J Org Chem*, **68**, 83–91.

[21] Tarver, J. E., Terranova, K. M., and Joullié, M. M. (2004) Hetero-Diels-Alder and pyroglutamate approaches to (2*S*,4*R*)-2-methylamino-5-hydroxy-4-methylpentanoic acid. *Tetrahedron*, **60**, 10277–10284.

[22] Roeske, R. (1963) Preparation of *tert*-butyl esters of free amino acids. *J Org Chem*, **28**, 1251–1253.

[23] Charrier, J.-D., Duffy, J. E. S., Hitchcock, P. B., and Young, D. W. (2001) Unusual stereoselectivity in the alkylation of pyroglutamate ester urethanes. *J Chem Soc, Perkin Trans 1*, **2001**, 2367–2371.

[24] Tsubogo, T., Saito, S., Seki, K., Yamashita, Y., and Kobayashi, S. (2008) Development of catalytic asymmetric 1,4-addition and [3 + 2] cycloaddition reactions using chiral calcium complexes. *J Am Chem Soc*, **130**, 13321–13332.

[25] Kobayashi, S., Tsubogo, T., Saito, S., and Yamashita, Y. (2008) Calcium-catalyzed diastereo- and enantioselective 1,4-addition of glycine derivatives to α,β-unsaturated esters. *Org Lett*, **10**, 807–809.

[26] Achqar, A. E., Boumzebra, M., Roumestant, M.-L., and Viallefont, P. (1988) 2-hydroxy-3-pinamone as chiral auxiliary in the asymmetric synthesis of β-amino acids. *Tetrahedron*, **44**, 5319–5332.

[27] Wehbe, J., Rolland, V., Roumestant, M., and Martinez, J. (2003) Glutamate transporter blockers: Enantiomerically pure (2*S*,3*S*)- and (2*S*,3*R*)-3-methyl glutamic acids. *Tetrahedron: Asymmetry*, **14**, 1123–1126.

[28] Seebach, D., Boes, M., Naef, R., and Schweizer, W. B. (1983) Alkylation of amino acids without loss of the optical activity: Preparation of α-substituted proline derivatives. A case of self-reproduction of chirality. *J Am Chem Soc*, **105**, 5390–5398.

[29] Amedjkouh, M. and Ahlberg, P. (2002) Synthesis of chiral diamines using novel 2-trichloromethyloxazolidin-4-one precursors derived from 5-oxo-proline and proline. *Tetrahedron: Asymmetry*, **13**, 2229–2234.

[30] Tong, S.-T. A., Harris, P. W. R., Barker, D., and Brimble, M. A. (2008) Use of (*S*)-5-(2-methylpyrrolidin-2-yl)-1H-tetrazole as a novel and enantioselective organocatalyst for the aldol reaction. *Eur J Org Chem*, **2008**, 164–170.

[31] Clive, D. L. J., Li, Z., and Yu, M. (2007) Intramolecular conjugate displacement: A general route to hexahydroquinolizines, hexahydroindolizines, and related [m,n,0]-bicyclic structures with nitrogen at a bridgehead. *J Org Chem*, **72**, 5608–5617.

[32] Harris, P. W., Brimble, M. A., Muir, V. J., Lai, M. Y., Trotter, N. S., and Callis, D. J. (2005) Synthesis of proline-modified analogues of the neuroprotective agent glycyl-L-prolyl-glutamic acid (GPE). *Tetrahedron*, **61**, 10018–10035.

[33] Yoshifuji, S., Tanaka, K.-I., Kawai, T., and Nitta, Y. (1986) A novel synthesis of L-pyroglutamic acid derivatives from L-proline: Utility of *N*-protecting groups for ruthenium tetroxide oxidation of cyclic α-amino acids. *Chem Pharm Bull*, **34**, 3873–3878.

[34] Sheehan, J. C. and Tulis, R. W. (1974) Oxidation of cyclic amines with ruthenium tetroxide. *J Org Chem*, **39**, 2264–2267.

[35] Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. *Biochemistry*, **42**, 7836–7847.

[36] Freeman, M. F., Moshos, K. A., Bodner, M. J., Li, R., and Townsend, C. A. (2008) Four enzymes define the incorporation of coenzyme A in thienamycin biosynthesis. *Proc Natl Acad Sci U.S.A.*, **105**, 11128–11133.

[37] Omura, K. and Swern, D. (1978) Oxidation of alcohols by 'activated' dimethyl sulfoxide. A preparative, steric and mechanistic study. *Tetrahedron*, **34**, 1651–1660.

[38] Dess, D. B. and Martin, J. C. (1983) Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones. *J Org Chem*, **48**, 4155–4156.

[39] Ducho, C., Hamed, R. B., Batchelar, E. T., Sorensen, J. L., Odell, B., and Schofield, C. J. (2009) Synthesis of regio- and stereoselectively deuterium-labelled derivatives of L-glutamate semialdehyde for studies on carbapenem biosynthesis. *Org Biomol Chem*, **7**, 2770–2779.

[40] Buchi, H., Steen, K., and Eschenmoser, A. (1964) *N*,*N*-dimethylformamide dineopentylacetal: A reagent for esterifying carboxylic acids with benzyl alcohols. *Angew Chem Int Ed Engl*, **3**, 62–63.

[41] Brechbühler, H., Büchi, H., Hatz, E., Schreiber, J., and Eschenmoser, A. (1965) Die Reaktion von Carbonsauren mit Acetalen des *N*, *N*-Dimethylformamids: Eine Veresterungsmethode. *Helv Chim Acta*, **48**, 1746–1771.

[42] Hosokawa, T., Ohta, T., Kanayama, S., and Murahashi, S. (1987) Palladium(II)-catalyzed acetalization of terminal olefins bearing electron-withdrawing substituents with optically active diols. *J Org Chem*, **52**, 1758–1764.

[43] Ackermann, J., Matthes, M., and Tamm, C. (1990) Approaches to the synthesis of cytochalasans part 9. A versatile concept leading to all structural types of cytochalasans. *Helv Chim Acta*, **73**, 122–132.

[44] Fiaux, H., Kuntz, D. A., Hoffman, D., Janzer, R. C., Gerber-Lemaire, S., Rose, D. R., and Juillerat-Jeanneret, L. (2008) Functionalized pyrrolidine inhibitors of human type II α-mannosidases as anti-cancer agents: Optimizing the fit to the active site. *Bioorg Med Chem*, **16**, 7337–7346.

[45] Sharma, R. and Lubell, W. D. (1996) Regioselective enolization and alkylation of 4-oxo-*N*-(9-phenylfluoren-9-yl)proline: Synthesis of enantiopure proline, valine and hydroxyproline, valine chimeras. *J Org Chem*, **61**, 202–209.

[46] Besson, M., Delbecq, F., Gallezot, P., Neto, S., and Pinel, C. (2000) Asymmetric synthesis of 2-methyl cyclohexane carboxylic acids by heterogeneous catalysis: Mechanistic aspects. *Chem Eur J*, **6**, 949–958.

[47] Winterfeld, G. A., Ito, Y., Ogawa, T., and Schmidt, R. R. (1999) A novel and efficient route towards α-GalNAc-Ser and α-GalNAc-Thr building blocks for glycopeptide synthesis. *Eur J Org Chem*, **1999**, 1167–1171.

[48] Kelly, G. T., Sharma, V., and Watanabe, C. M. H. (2008) An improved method for culturing *Streptomyces sahachiroi*: Biosynthetic origin of the enol fragment of azinomycin B. *Bioorg Chem*, **36**, 4–15.

[49] You, S.-L. and Kelly, J. W. (2003) Total synthesis of dendroamide A: Oxazole and thiazole construction using an oxodiphosphonium salt. *J Org Chem*, **68**, 9506–9509.

[50] Huy, P. and Schmalz, H.-G. (2011) Practical one-pot double functionalizations of proline. *Synthesis*, **2011**, 954–960.

3.6 Materials and Methods for Chapter 3

3.6.1 Synthesis of 4-Mono and 4,4-Dimethylglutamate Semialdehyde from L-Pyroglutamic Acid

***tert*-Butyl 5-oxo-L-prolinate 144** [46]

The synthesis of compound **144** was adapted from a reported procedure [46] for the *tert*-butyl protection of carboxylic acids using isobutylene under acidic conditions [22]. 5-Oxo-L-proline **141** (10.0 g, 77.5 mmol) was dissolved in dioxane (100 mL) and concentrated sulphuric acid was added to a final concentration of 2% v/v (2 mL). The solution was cooled to -78 °C and isobutylene was condensed (approximately 20 mL) by bubbling the gas through the solvent. The sealed flask was allowed to reach room temperature again and the reaction was kept under pressure and stirring conditions for 6 days. The pressure was carefully released before evaporation of the solvent. The residue was dissolved in diethyl ether (200 mL) and the solution was washed twice with a saturated sodium hydrogen carbonate aqueous solution (100 mL), dried over magnesium sulphate, filtered and evaporated to give *tert*-butyl 5-oxo-L-prolinate **144** as a white solid (7.3 g, 51%).

mp 98–100 °C (lit. 102 °C [46]); TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.40; IR (KBr) ν/cm^{-1}: 3267 (NH), 1737 (OC=O), 1704 (NC=O), 1680 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.11 (*app bs*, 1H), 4.14 (*dd*, *J* = 8.0, 5.5, 1H), 2.49-2.29 (*m*, 3H), 2.22-2.14 (*m*, 1H), 1.48 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 177.7, 171.0, 82.4, 56.0, 29.3, 27.9, 24.8; HR ESI-MS : 208.0941 ([M+Na]$^+$, $C_9H_{15}NNaO_3^+$, calc. 208.0944).

Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate 143 [16]

The synthesis of di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** was adapted from a reported procedure [16]. *tert*-Butyl 5-oxo-L-prolinate **144** (6.0 g, 32.4 mmol) was dissolved in anhydrous dichloromethane (250 mL). Triethylamine (6.8 mL, 1.5 eq.) was added dropwise followed by 4-dimethylaminopyridine (5.93 g, 1 eq.). After stirring for 10 min at room temperature, di-*tert*-butyl dicarbonate (10.6 g, 1.5 eq.) was added in small portions. The reaction was left under a nitrogen atmosphere for 3 days before quenching with water (50 mL). The phases were separated and the organic layer was washed with a saturated sodium hydrogen carbonate solution (100mL), water (100 mL), dried over magnesium sulphate, filtered and evaporated to give an orange oil. This crude product was purified using silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** was isolated as a white solid (7.85 g, 85%).

mp 53–54 °C (lit. 54–56 °C [16]); TLC: 1:1 hex/EtOAc, R_f = 0.35; IR (KBr) ν/cm^{-1}: 1786 (OC=O), 1783 (NC=O), 1723 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.47 (*dd*, *J* = 9.5, 2.5, 1H), 2.64-2.55 (*m*, 1H), 2.49-2.42 (*m*, 1H), 2.33-2.22 (*m*, 1H), 2.02-1.94 (*m*, 1H), 1.49 (*s*, 9H), 1.47 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 173.5, 170.3, 149.2, 83.2, 82.2, 59.5, 31.1, 27.9, 21.6; HR ESI-MS : 308.1464 ([M+Na]$^+$, $C_{14}H_{23}NNaO_5^+$, calc. 308.1468).

***trans*- And *cis*-di-*tert*-butyl (2*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate 146 and 145** [23]

The synthesis of compounds **146** and **145** was adapted from a reported procedure [23]. Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** (1.20 g, 4.2 mmol) was dissolved in anhydrous toluene (10 mL) and the solution was cooled to -78 °C prior to addition of a lithium bis(trimethylsilyl)amide solution in tetrahydrofuran (1 M, 5 mL, 1.2 eq.). After stirring under a nitrogen atmosphere for 1 h, methyl triflate (0.57 mL, 1.2 eq.) was added dropwise *via* syringe

and the reaction mixture was stirred under these conditions for 2 h before addition of lithium bis(trimethylsilyl)amide solution in tetrahydrofuran (1 M, 5 mL, 1.2 eq.). After 4 h, the reaction was quenched with a saturated ammonium chloride aqueous solution (20 mL) and allowed to reach room temperature. The phases were separated and the aqueous layer was extracted with ethyl acetate. The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give the crude product which was purified using silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). The *trans*-isomer **146** eluted first (165 mg, 13%), followed by the *cis*-isomer **145** (730 mg, 58%) and the unreacted starting material **143** (84 mg, 7% recovery).

Di-*tert*-butyl (2*S*,4*R*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **146**: mp 55–58 °C (lit. 62–64 °C [23]); TLC: 7:3 hex/EtOAc, R_f = 0.35; IR (KBr) ν/cm^{-1}: 1777 (OC=O), 1743 (NC=O), 1735 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.42 (*dd*, *J* = 9.5, 1.0, 1H), 2.71-2.60 (*m*, 1H), 2.24 (*ddd*, *J* = 13.0, 8.5, 1.0, 1H), 1.89 (*ddd*, *J* = 13.0, 12.0, 9.5, 1H), 1.52 (*s*, 9H), 1.48 (*s*, 9H), 1.21 (*d*, *J* = 7.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 175.8, 170.3, 149.5, 83.1, 82.2, 57.7, 36.5, 30.7, 27.9, 15.2; HR ESI-MS : 322.1639 ([M+Na]$^+$, $C_{15}H_{25}NNaO_5^+$, calc. 322.1625).

Di-*tert*-butyl (2*S*,4*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **145**: mp 64–67 °C (lit. 68–69 °C [23]); TLC: 7:3 hex/EtOAc, R_f = 0.25; IR (KBr) ν/cm^{-1}: 1758 (OC=O), 1745 (NC=O), 1713 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.42 (*dd*, *J* = 9.0, 6.0, 1H), 2.65-2.51 (*m*, 2H), 1.61-1.50 (*m*, 1H), 1.51 (*s*, 9H), 1.48 (*s*, 9H), 1.27 (*d*, *J* = 7.0, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.0, 170.6, 149.6, 83.3, 82.1, 58.1, 37.4, 29.6, 27.9, 16.5; HR ESI-MS : 322.1635 ([M+Na]$^+$, $C_{15}H_{25}NNaO_5^+$, calc. 322.1625).

Di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate 147

Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** (1.0 g, 3.5 mmol) was dissolved in anhydrous toluene (10 mL) under a nitrogen atmosphere. The solution was cooled to -78 °C prior to addition of a lithium bis(trimethylsilyl)amide solution in tetrahydrofuran (1 M, 5 mL, 1.2 eq.). After stirring under a nitrogen atmosphere for 1 h, methyl triflate (0.44 mL, 1.1 eq.) was added dropwise *via* syringe and the reaction mixture was stirred under these conditions for 2 h before adding once more a lithium bis(trimethylsilyl)amide solution in tetrahydrofuran (1 M, 5.0 mL, 1.2 eq.) followed by methyl triflate (0.44 mL, 1.1 eq.). After 4 h, the reaction was quenched with a saturated ammonium chloride aqueous solution (20 mL) and allowed to reach room temperature. The phases were separated and the aqueous layer was extracted three times with ethyl acetate. The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give the crude product which was purified using silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). The main product eluted first and found to be di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate **147** (650 mg, 59%), followed by the *trans*-isomer **146** and *cis*-isomer **145** (210 mg, 20% cumulated yields) and finally some unreacted starting material **143** (25 mg, 2.5% recovery).

Dimethylated product **147**: mp 95–97 °C (lit. 100–102 °C [23]); TLC: 7:3 hex/EtOAc, R_f = 0.40; IR (KBr) ν/cm^{-1}: 1785 (OC=O), 1730 (NC=O); $[\alpha]_D^{25}$ = - 37.9° (c 1.5, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.41 (*dd*, *J* = 9.5, 4.5, 1H), 2.19 (*dd*, *J* = 13.0, 9.5, 1H), 1.89 (*dd*, *J* = 13.0, 4.5, 1H), 1.52 (*s*, 9H), 1.48 (*s*, 9H), 1.22 (*s*, 3H), 1.21 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 178.3, 170.7, 149.8, 83.2, 82.1, 56.5, 41.6, 36.7, 27.8, 25.8, 25.2; HR ESI-MS : 336.1778 ([M+Na]$^+$, $C_{15}H_{25}NNaO_5^+$, calc. 336.1781).

Di-*tert*-butyl (2*S*,4*R*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate 149

Di-*tert*-butyl (2*S*,4*R*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **146** (60 mg, 0.20 mmol) was dissolved in dried tetrahydrofuran (3 mL) under a nitrogen atmosphere. The solution was then cooled to -78 °C, prior to dropwise addition of lithium triethylborohydride solution in tetrahydrofuran (1 M, 0.2 mL, 1.0 eq.). The reaction was stirred under these conditions for 3 h, when lithium

triethylborohydride solution was added again (1 M, 0.2 mL, 1.0 eq.) After 3 h at -78 °C, the reaction was quenched with a saturated sodium hydrogenocarbonate aqueous solution (1 mL), followed by a 40% aqueous hydrogen peroxide solution (0.1 mL). The temperature was raised to 0 °C and the mixture was stirred for 30 min. The tetrahydrofuran was evaporated and the resulting aqueous layer extracted with ethyl acetate. The combined organic phases were washed with brine (10 mL), dried over magnesium sulphate and evaporated to give the crude product which was purified using silica chromatography (SNAP-10 column, gradient elution using 7–60% ethyl acetate in *n*-hexane). The product was isolated as a colourless oil (51 mg, 84%) and was found to be an equilibrating mixture of di-*tert*-butyl (2*S*,4*R*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate **149** and the corresponding hydrated aldehyde (3:7 in $CDCl_3$).

TLC: 8:2 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 1742 (OC=O), 1698 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.33, 6.21 (2x *app bs*, 0.3H), 5.37, 5.08 (2x *app bs*, 0.7H), 4.53-4.42 (*m*, 0.3H), 4.31-4.07 (*m*, 0.7H), 3.01-2.87 (*m*, 0.3H), 2.49-2.41 (*m*, 0.3H), 2.26-1.97 (*m*, 2H), 1.94-1.83 (*m*, 0.7H), 1.68 (*app bs*, 1H), 1.48 (s, 9H), 1.45 (s, 9H), 1.09-1.01 (*m*, 2H); HR ESI-MS: 324.1782 ($[M+Na]^+$, $C_{15}H_{27}NNaO_5^+$, calc. 324.1781).

Di-*tert*-butyl (2*S*,4*S*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate 148

Di-*tert*-butyl (2*S*,4*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **145** (30 mg, 0.10 mmol) was dissolved in dried tetrahydrofuran (2 mL) under a nitrogen atmosphere and treated as **146** above. The residue was purified by silica chromatography (SNAP-10 column, gradient elution using 7–60% ethyl acetate in *n*-hexane) and the product was isolated as a colourless oil (26 mg, 87%). The purified product gave a complicated NMR pattern probably due to an mixture of di-*tert*-butyl (2*S*,4*S*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate **148** in equilibrium with other conformational isomers.

The same reaction was performed on a larger scale using di-*tert*-butyl (2*S*,4*S*)-4-methyl-5-oxopyrrolidine-1,2-dicarboxylate **145** (240 mg, 0.80 mmol). Two sequential additions of substoichiometric amount of lithium triethylborohydride (1 M, 0.72 mL, 0.9 eq.) at -78 °C followed by the same work-up and purification procedures gave di-*tert*-butyl (2*S*,4*S*)-5-hydroxy-4-methylpyrrolidine-1,2-dicarboxylate **148** as a low melting point white solid (225 mg, 93%).

mp $\leq$ 35 °C; TLC: 8:2 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 1744 (OC=O), 1704 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.25-7.16 (*m*, 0.3H, NH), 6.33, 6.20 (2x *app bs*, 0.7H), 5.43-5.35 (*m*, 0.15H), 5.22-5.18 (*m*, 0.15H), 4.53-4.41 (*m*, 0.7H), 4.29-4.07 (*m*, 0.3H), 3.00-2.87 (*m*, 0.7H), 2.61-2.56 (*m*, 0.7H), 2.47-2.42 (*m*, 1H), 2.35-2.29 (*m*, 0.3H), 2.19-2.09 (*m*, 0.3H), 1.68 (*app. s*, 2H), 1.48 (*s*, 9H), 1.44 (*s*, 9H), 1.08 (*app. d*, 1H); HR ESI-MS: 324.1782 ($[M+Na]^+$, $C_{15}H_{27}NNaO_5^+$, calc. 324.1781).

Di-*tert*-butyl (2*S*)-5-hydroxy-4,4-dimethylpyrrolidine-1,2-dicarboxylate 150

Di-*tert*-butyl (2*S*)-4,4-dimethyl-5-oxopyrrolidine-1,2-dicarboxylate **147** (200 mg, 0.64 mmol) was dissolved in dried tetrahydrofuran (5 mL) under a nitrogen atmosphere and treated as **146** above. The residue was purified by silica chromatography (SNAP-25 column, gradient elution using 7–60% ethyl acetate in *n*-hexane) and di-*tert*-butyl (2*S*)-5-hydroxy-4,4-dimethylpyrrolidine-1,2-dicarboxylate **150** was isolated as a colourless oil (170 mg, 85%). The purified product gave a complicated NMR pattern, probably due to the presence of several conformational isomers.

TLC: 8:2 hex/EtOAc, R_f = 0.30; IR (KBr) ν/cm^{-1}: 1743 (OC=O), 1704 (NC=O) 1681 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.06, 4.98, 4.93, 4.82 (4x *s*, 1H), 4.20-4.16 (*m*, 0.4H), 4.14-4.07 (*m*, 0.6H), 2.29-2.19 (*m*, 0.30H), 2.02-1.78 (*m*, 1.6H), 1.71-1.67 (*m*, 0.25H), 1.62-1.59 (*m*, 0.25H), 1.46 (*s*, 9H), 1.43 (*s*, 9H), 1.09, 1.05, 0.96 (3x *s*, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.0, 172.4, 171.8, 155.0, 154.2, 153.7, 88.7, 88.6, 88.4, 81.6, 81.2, 81.0, 80.7, 80.6, 59.0, 58.8, 58.4, 43.2, 42.0,

41.3, 41.0, 40.8, 39.8, 28.3, 28.2, 28.1, 28.0, 27.9, 27.8, 25.9, 25.5, 22.5, 21.9, 21.7; HR ESI-MS: 338.1939 ($[M+Na]^+$, $C_{16}H_{29}NNaO_5{}^+$, calc. 338.1938).

3.6.2 Synthesis of 3-Methylglutamate Semialdehyde from L-Threonine-O*t*-Bu

***tert*-Butyl *N*-(*tert*-butoxycarbonyl)-L-threoninate 165** [47]

Triethylamine (4.0 mL, 2.5 eq.) was added to a solution of *tert*-butyl L-threoninate **164** (2.0 g, 11.4 mmol) and di-*tert*-butyl dicarbonate (3.0 g, 1.2 eq.) in anhydrous dichloromethane (100 mL). The reaction mixture was stirred under a nitrogen atmosphere at room temperature for 16 h before adding water (30 mL). The phases were separated and the organic phase was washed with a saturated sodium hydrogen carbonate solution (30 mL) and brine (30 mL), dried over magnesium sulphate, filtered and evaporated to give an orange oil, further purified using silica chromatography (SNAP-100 column, gradient elution with 0–100% ethyl acetate in *n*-hexane). *tert*-Butyl *N*-(*tert*-butoxycarbonyl)-L-threoninate **165** was obtained as a white solid (3.04 g, 97%). Spectroscopic data were in good agreement with these reported [47].

mp 68–70 °C (lit. 70 °C [47]); TLC: 7:3 hex/EtOAc, R_f = 0.25; IR (KBr) ν/cm^{-1}: 3437 (OH), 1742 (OC=O), 1689 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.30 (*app bs*, 1H), 4.23 (*app bs*, 1H), 4.12 (*app bs*, 1H), 2.15 (*app bs*, 1H), 1.48 (*s*, 9H), 1.45 (*s*, 9H), 1.24 (*d*, *J* = 2.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 170.5, 156.1, 82.3, 79.8, 68.5, 59.1, 28.3, 28.0, 19.9; HR ESI-MS: 298.1623 ($[M+Na]^+$, $C_{13}H_{25}NNaO_5{}^+$, calc. 298.1625).

textittert-Butyl 2-[(*tert*-butoxycarbonyl)amino]-3-oxobutanoate 166 [48, 49]

The oxidation of *tert*-Butyl *N*-(*tert*-butoxycarbonyl)-L-threoninate **165** was adapted from reported procedures [48, 49]. A solution of the Dess-Martin periodinane (4.62 g, 1.2 eq.) in dichloromethane (50 mL) was added at room temperature under a nitrogen atmosphere to a solution of the protected amino acid **165** in dichloromethane (50 mL). The reaction mixture was stirred under these conditions for 90 min and then diluted with dichloromethane (50 mL) before quenching with a saturated sodium hydrogen carbonate solution containing 0.5 M sodium sulphite. The phases were separated and the organic layer was washed with a saturated sodium hydrogen carbonate (50 mL), water (50 mL) dried over magnesium sulphate, filtered and evaporated to give the crude which was eluted from a pad of silica using a 7:3 *n*-hexane/ethyl acetate mixture. The solvent was evaporated and *tert*-butyl 2-(*tert*-butoxycarbonyl-amino)-3-oxobutanoate **166** was obtained as a pale yellow oil (1.77g, 71%).

TLC: 7:3 hex/EtOAc, R_f = 0.40; IR (KBr) ν/cm^{-1}: 1750 (OC=O), 1719 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.66 (*app bs*, 1H), 4.90 (*d*, *J* = 7.0, 1H), 2.34 (*app bs*, 3H), 1.47 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 199.5, 165.6, 154.8, 83.6, 80.2, 64.9, 28.2, 27.7; LR ESI-MS: 272.16 ($[M-H]^-$).

1-*tert*-Butyl 5-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-methylpent-2-enedioate 167

tert-Butyl 2-(*tert*-butoxycarbonyl-amino)-3-oxobutanoate **166** (1.75 g, 6.4 mmol) was dissolved in toluene (70 mL) and methyl(triphenylphosphoranylidene)acetate (2.57 g, 1.2 eq.) was added. The reaction mixture was placed under a nitrogen atmosphere and refluxed until reaction completion, as judged by TLC analysis. After cooling the mixture to room temperature, the solvent was evaporated to give the crude as a dark red oil. This residue was purified using silica chromatography (SNAP-100 column, gradient elution with 0–50% ethyl acetate in *n*-hexane). The *Z*-isomer of 1-*tert*-butyl 5-methyl 2-[(*tert*-butoxycarbonyl)amino]-3-methylpent-2-enedioate **167** eluted first (720 mg, 31%)

followed by the *E*-isomer (725 mg, 31%) and were both obtained as white solids (62% cummulated yield). Assignments of *Z*-**167** and *E*-**167** configurations were based on 1D NOESY NMR analyses (Figure 3.3).

Z-**167**: mp 86–87 °C; TLC: 8:2 hex/EtOAc, R_f = 0.25; IR (KBr) ν/cm^{-1}: 3369 (NH), 1746 (OC=O), 1698 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.49 (*app bs*, 1H), 3.71 (*s*, 3H), 3.20 (*s*, 2H), 2.05 (*s*, 3H), 1.49 (*s*, 9H), 1.45 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 163.7, 153.8, 127.2, 81.4, 52.3, 40.6, 28.2, 28.0, 19.6; HR ESI-MS: 352.1726 ([M+Na]$^+$, $C_{16}H_{27}NNaO_6^+$, calc. 352.1731).

E-**167**: mp 115–117 °C; TLC: 8:2 hex/EtOAc, R_f = 0.20; IR (KBr) ν/cm^{-1}: 3306 (NH), 1726 (OC=O), 1688 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.91 (*app bs*, 1H), 3.67 (*s*, 3H), 3.57 (*s*, 2H), 1.88 (*s*, 3H), 1.46 (*s*, 18H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 171.0, 163.7, 126.1, 81.8, 51.9, 39.2, 28.2, 28.0, 21.0; HR ESI-MS: 352.1726 ([M+Na]$^+$, $C_{16}H_{27}NNaO_6^+$, calc. 352.1731).

(2*S*,3*R*)/(2*R*,3*S*)-1-*tert*-Butyl 5-methyl *N*-(*tert*-butoxycarbonyl)-3-methylglutamate 169

Z-**167** (50 mg, 0.15 mmol) was dissolved in a 1:1 mixture of tetrahydrofuran and methanol (5 mL), and 10% mol/mol palladium on activated carbon (5 mg, 10% w/w) was added. The reaction was stirred at room temperature under a hydrogen atmosphere for 24 h. After the reaction was completed, as judged by TLC, the mixture was filtered through a pad of Celite® 545 and the clear resulting solution was evaporated to give an enantiomeric mixture of (2*S*,3*R*)-**169** and (2*R*,3*S*)-**169** as a colourless oil (48 mg, 98%).

TLC: 7:3 hex/EtOAc, R_f = 0.40; IR (neat) ν/cm^{-1}: 3371 (NH), 1735 (OC=O), 1716 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.09 (*d*, *J* = 8.3, 1H), 4.32 (*dd*, *J* = 8.5, 2.0, 1H), 3.69 (*s*, 3H), 2.58 (*app bs*, 1H), 2.41 (*d*, *J* = 15.5, 5.5, 1H), 2.14 (*dd*, *J* = 15.5, 8.5, 1H), 1.48 (*s*, 9H), 1.44 (*s*, 9H), 0.88 (*d*, *J* = 7.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 172.9, 170.9, 155.7, 82.2, 79.7, 57.1, 51.6, 37.9, 33.1, 28.3, 28.0, 14.4; HR ESI-MS: 354.1883 ([M+Na]$^+$, $C_{16}H_{29}NNaO_6^+$, calc. 354.1887).

(2*S*,3*S*)/(2*R*,3*R*)-1-*tert*-Butyl 5-methyl *N*-(*tert*-butoxycarbonyl)-3-methylglutamate 168

E-**167** (300 mg, 0.91 mmol) was dissolved in ethyl acetate (15 mL) and subjected to hydrogenation as for *Z*-**167** to give an enantiomeric mixture of (2*S*,3*S*)-**168** and (2*R*,3*R*)-**168** as a colourless oil (300 mg, 99%).

TLC: 7:3 hex/EtOAc, R_f = 0.40; IR (neat) ν/cm^{-1}: 3365 (NH), 1741 (OC=O), 1721 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.12 (*d*, *J* = 8.0, 1H), 4.14 (*dd*, *J* = 8.0, 4.5, 1H), 3.67 (*s*, 1H), 2.46-2.39 (*m*, 2H), 2.18 (*dd*, *J* = 15.5, 8.5, 1H), 1.47 (*s*, 9H), 1.44 (*s*, 9H), 1.00 (*d*, *J* = 6.5, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 172.8, 170.8, 155.4, 82.3, 79.7, 58.0, 51.7, 37.4, 33.7, 28.3, 28.0, 16.4; HR ESI-MS: 354.1892 ([M+Na]$^+$, $C_{16}H_{29}NNaO_6^+$, calc. 354.1887).

(2*S*,3*R*)/(2*R*,3*S*)-1-*tert*-Butyl 5-methyl *N*,*N*-bis(*tert*-butoxycarbonyl)-3-methylglutamate 171

(2*S*,3*R*)/(2*R*,3*S*)-**169** (45 mg, 0.14 mmol) was dissolved in dried tetrahydrofuran (2 mL) together with di-*tert*-butyl dicarbonate (100 mg, 6 eq.) and 4-dimethylaminopyridine (10 mg, 1.5 eq.). The reaction was stirred at 60 °C under a nitrogen atmosphere for 18 h. Di-*tert*-butyl dicarbonate (100 mg, 6 eq.) was added again and the reaction stirred under the same conditions for 6 h. The solvent was evaporated and the residue purified using silica chromatography (SNAP-10 column, gradient elution with 0–30% ethyl acetate in *n*-hexane) to give an enantiomeric mixture of (2*S*,3*R*)-**171** and (2*R*,3*S*)-**171** as a white solid (50 mg, 86%).

mp 58–62 °C; TLC: 7:3 hex/EtOAc, R_f = 0.50; IR (KBr) ν/cm^{-1}: 1734 (OC=O), 1670 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.60 (*d*, *J* = 9.0, 1H), 3.68 (*s*, 3H), 2.89-2.81 (*m*, 2H), 2.27 (*dd*, *J* = 16.0, 9.5, 1H), 1.51 (*s*, 18H), 1.45 (*s*, 9H), 0.91 (*d*, *J* = 7.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$):

δ = 172.9, 169.1, 152.5, 82.9, 81.3, 61.4, 51.5, 40.0, 30.2, 28.0, 27.9, 16.3; HR ESI-MS: 454.2412 ([M+Na]$^+$, $C_{21}H_{38}NNaO_8^+$, calc. 454.2411).

(2*S*,3*S*)/(2*R*,3*R*)-1-*tert*-Butyl 5-methyl *N*,*N*-bis(*tert*-butoxycarbonyl)-3-methylglutamate 170

(2*S*,3*S*)/(2*R*,3*R*)-**170** (300 mg, 0.91 mmol) was dissolved in tetrahydrofuran (10 mL) together with di-*tert*-butyl dicarbonate (1.25 g, 6 eq.) and 4-dimethylaminopyridine (30 mg, 0.25 eq.). The reaction was stirred at room temperature under a nitrogen atmosphere for 3 days. Di-*tert*-butyl dicarbonate (0.5 g, 2.5 eq.) and 4-dimethylaminopyridine (30 mg, 0.25 eq.) were added again and the reaction was kept under the same conditions for 24 h. The solvent was evaporated and replaced by ethyl acetate (50 mL) and the organic layer was washed with 5% potassium sulphate aqueous solution (50 mL), 5% sodium hydrogen carbonate aqueous solution (50 mL) and water (50 mL), dried over magnesium sulphate, filtered and evaporate. The residue was purified using silica chromatography (SNAP-100 column, gradient elution using 0–50% ethyl acetate in *n*-hexane) to give an enantiomeric mixture of (2*S*,3*S*)-**170** and (2*R*,3*R*)-**170** as a white solid (330 mg, 85%).

mp 57–59 °C; TLC: 7:3 hex/EtOAc, R_f = 0.45; IR (KBr) ν/cm^{-1}: 1732 (OC=O), 1699 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.51 (*d*, *J* = 9.0, 1H), 3.67 (*s*, 3H), 2.85-2.76 (*m*, 1H), 2.43 (*dd*, *J* = 15.0, 4.0, 1H), 2.08 (*dd*, *J* = 15.0, 10.0, 1H), 1.51 (*s*, 18H), 1.45 (*s*, 9H), 1.16 (*d*, *J* = 6.5, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 173.0, 168.6, 152.4, 83.0, 81.3, 62.8, 51.5, 38.1, 30.7, 27.9, 19.1; HR ESI-MS: 454.2418 ([M+Na]$^+$, $C_{21}H_{38}NNaO_8^+$, calc. 454.2411).

(2*S*,3*R*)/(2*R*,3*S*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate 173

(2*S*,3*R*)/(2*R*,3*S*)-1-*tert*-Butyl 5-methyl *N*,*N*-bis(*tert*-butoxycarbonyl)-3-methylglutamate **171** (30 mg, 0.07 mmol) was dissolved in dried diethyl ether (2 mL). The solution was cooled to -78 °C under a nitrogen atmosphere before addition of a 1 M diisobutyl aluminium hydride solution in hexanes (0.21 mL, 3.0 eq.). After stirring for 90 min, the reaction was quenched with water (0.1 mL) and allowed to reach room temperature over 30 min. The aluminium hydroxide precipitate was filtered off through a pad of Celite® 545, washed with diethyl ether and the filtrate was evaporated to give a colourless oil which was purified using silica chromatography (SNAP-10 column, gradient elution using 0–50% ethyl acetate in *n*-hexane). An enantiomeric mixture of (2*S*,3*R*)- and (2*R*,3*S*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **173** eluted first (23 mg, 82%), followed by an enantiomeric mixture of the undesired alcohols (2*S*,3*R*)- and (2*R*,3*S*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-hydroxyisoleucinate **175** (6 mg, 5%). Both compounds were isolated as colourless oils.

(2*S*,3*R*)/(2*R*,3*S*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **173**: TLC: 8:2 hex/EtOAc, R_f = 0.40; IR (neat) ν/cm^{-1}: 1743 (OC=O), 1701 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.77 (*d*, *J* = 2.0, 1H), 4.54 (*d*, *J* = 9.5, 1H), 2.99-2.94 (*m*, 2H), 2.47 (*ddd*, *J* = 17.5, 9.5, 2.5, 1H), 1.52 (*s*, 18H), 1.44 (*s*, 9H), 0.91 (*d*, *J* = 7.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ 201.6, 169.1, 152.6, 83.0, 81.6, 61.8, 50.0, 28.0, 27.9, 16.7; HR ESI-MS: 424.2306 ([M+Na]$^+$, $C_{20}H_{35}NNaO_7^+$, calc. 424.2306).

(2*S*,3*R*)/(2*R*,3*S*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-hydroxyisoleucinate **175**: TLC: 8:2 hex/EtOAc, R_f = 0.10; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.49 (*dd*, *J* = 9.5, 1H), 3.71 (*app bs*, 1H), 2.56-2.49 (*m*, 1H), 2.24 (*app bs*, 1H), 1.89-1.81 (*m*, 1H), 1.59 (*app bs*, 1H), 1.51 (*s*, 18H), 1.45 (*s*, 9H), 0.88 (*d*, *J* = 7.0, 3H); HR ESI-MS: 426.2463 ([M+Na]$^+$, $C_{20}H_{37}NNaO_7^+$, calc. 426.2462).

(2*S*,3*S*)/(2*R*,3*R*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate 172

(2*S*,3*S*)/(2*R*,3*R*)-**170** (50mg, 0.11 mmol) was dissolved in dried diethyl ether (3 mL) and the solution was cooled to -78 °C under a nitrogen atmosphere before addition of a 1 M diisobutyl

aluminium hydride solution in tetrahydrofuran (0.09 mL, 0.8 eq.) and stirring for 90 min. Another 0.8 eq. of DIBAL-H solution was added and after 90 min under these conditions, the reaction was quenched with water (0.05 mL) and allowed to reach room temperature over 30 min. The emulsion was filtered through a pad of Celite® 545, washed with diethyl ether and evaporated to give a colourless oil which was purified using silica chromatography (SNAP-10 column, gradient elution using 0–100% ethyl acetate in *n*-hexane) to yield a mixture two non-differenciable isomers (2*S*,3*S*) and (2*R*,3*R*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **172** (16 mg, 34%) and a mixture two non-differenciable isomers of the undesired alcohol (2*S*,3*S*)and (2*R*,3*R*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-hydroxyisoleucinate (2*S*,3*S*)/(2*R*,3*R*)-**174** (32 mg, 64%). This side product could be reoxidised in dichloromethane (5 mL) by adding a solution of Dess-Martin periodinane (44 mg, 1.5 eq,) in dichloromethane (5 mL) at room temperature under a nitrogen atmosphere and stirring for 90 min. The reaction was quenched by addition of a saturated solution of sodium hydrogen carbonate (10 mL) containing sodium sulphite (10 mg). The phases were separated and the organic layer were washed with a saturated sodium hydrogen carbonate solution (5 mL), water (5 mL), dried over magnesium sulphate, filtered and evaporated to give a colourless oil further purified using silica chromatography (SNAP-10 column, gradient elution using 0–50% ethyl acetate in *n*-hexane) to yield the aldehyde (2*S*,3*S*)/(2*R*,3*R*)-*tert*-butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **172** as a pale yellow oil (28 mg, 88%).

(2*S*,3*S*)/(2*R*,3*R*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-oxoisoleucinate **172**: TLC: 1:1 hex/EtOAc, R_f = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.71-9.69 (*m*, 1H), 4.52 (*dd*, *J* = 9.0, 3.0, 1H), 3.00-2.90 (*m*, 1H), 2.51-2.45 (*m*, 1H), 2.25-2.17 (*m*, 1H), 1.51 (*s*, 18H), 1.45 (*s*, 9H), 1.16 (*dd*, *J* = 6.5, 3.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 201.9, 168.6, 152.7, 83.2, 81.5, 62.7, 47.4, 28.6, 27.9 (2x), 19.6; LR ESI-MS: 424.23 ([M+Na]$^+$), 825.46 ([2M+Na]$^+$).

(2*S*,3*S*)/(2*R*,3*R*)-*tert*-Butyl *N*,*N*-bis(*tert*-butoxycarbonyl)-5-hydroxyisoleucinate **174**: mp 77–79 °C; TLC: 1:1 hex/EtOAc, R_f = 0.45; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.43 (*dd*, *J* = 9.0, 3.0, 1H), 3.76-3.69 (*m*, 1H), 3.66.3.59 (*m*, 1H), 2.49-2.41 (*m*, 1H), 1.70.1.61 (*m*, 1H), 1.59 (*app bs*, 1H), 1.50 (*s*, 18H), 1.43 (*s*, 9H), 1.38-1.26 (*m*, 1H), 1.11 (*dd*, *J* = 6.5, 3.0, 3H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 169.2, 152.8, 82.9, 81.1, 63.5, 60.6, 35.4, 29.6, 27.9, 18.8; LR ESI-MS: 426.23 ([M+Na]$^+$), 829.42 ([2M+Na]$^+$).

3.6.3 Synthesis of 2-Methylglutamate Semialdehyde from L-Proline

(3*R*,5*S*)-3-*tert*-Butyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one 177 [28]

The synthesis of (3*R*,5*S*)-3-*tert*-butyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **177** has been reported [28]. To a suspension of L-proline **134** (0.67 g, 5.8 mmol) in pentane (25 mL) was added trimethylacetaldehyde (3.15 mL, 5 eq.) and catalytic trifluoroacetic acid (20 μL). The reaction was stirred for 4 days under reflux and vigorous stirring with azeotropic removal of the water formed in the condensation reaction (Dean-Stark apparatus). At the end of the reaction, the solvent was evaporated under reduced pressure and the crude product immediately placed under nitrogen atmosphere before carrying out the next step. Because of the extremely moisture sensitive character of the product, the NMR analysis was carried out in a sealed tube filled with nitrogen gas.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.50 (*s*, 1H), 3.79 (*dd*, *J* = 9.0, 4.5, 1H), 3.19 (*dt*, *J* = 10.5, 6.5, 1H), 2.79 (*dt*, *J* = 10.5, 6.5, 1H), 2.18-2.09 (*m*, 1H), 2.06-1.98 (*m*, 1H), 1.84-1.75 (*m*, 1H), 1.73-1.62 (*m*, 1H), 0.91 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 178.0, 108.1, 62.7, 58.3, 37.4, 29.6, 25.1, 24.1.

(3*R*,5*S*)-3-*tert*-Butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one 178 [28]

The synthesis of (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **178** has been reported [28]. A solution of lithium diisopropylamide in tetrahydrofuran (2 M, 12 mL, 1.1 eq.) was added at -78 °C to a solution of (3*R*,5*S*)-3-*tert*-butyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **177** under nitrogen atmosphere. The reaction was stirred for 1 h before methyl iodide was added dropwise and the resulting suspension allowed to reach 0 °C. After 16 h, the mixture was poured into a 1:3 mixture of water and dichloromethane (200 mL). The phases were separated and the organic layer was dried over magnesium sulphate. After evaporation of the solvent, the residue was purified by Kugelrohr distillation and (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **178** was obtained as a colourless oil that crashed out in cold diethyl ether as a low melting point white solid (1.0 g, 23%).

mp $\leq$ 35 °C; bp 104–106 °C, 0.1 mm Hg (lit. 85 °C, 0.05 mm Hg [28]); $[\alpha]_D^{25}$ = - 32.6° (c 1.0, CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.25 (*s*, 1H), 3.14-3.07 (*m*, 1H), 2.84-2.79 (*m*, 1H), 2.19-2.13 (*m*, 1H), 1.85-1,78 (*m*, 1H), 1.77-1.64 (*m*, 1H), 1.37 (*s*, 3H), 0.90 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 179.0, 105.8, 68.7, 57.8, 38.6, 36.3, 25.4, 25.3, 24.0; HR ESI-MS : 198.1491 ([M+H]$^+$, $C_{11}H_{20}NO_2^+$, calc. 198.1489).

1-(2,2-Dimethylpropanoyl)-2-methyl-5-oxo-L-proline 180

1-(2,2-Dimethylpropanoyl)-2-methyl-5-oxo-L-proline **180** was obtained as the main product in an attempt to access (7a*S*)-3-*tert*-butyl-7a-methyldihydro-1*H*-pyrrolo[1,2-*c*][1,3]oxazole-1,5(6*H*)-dione **179** from (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **178** following a reported procedure for the oxidation of prolines [33]. A solution of (3*R*,5*S*)-3-*tert*-butyl-7a-methyltetrahydro-1*H*-pyrrolo[1,2-c][1,3]oxazol-1-one **178** (1.0 g, 5.0 mmol) in ethyl acetate (10 mL) was added to a suspension of catalytic ruthenium (IV) oxide hydrate (100 mg) in 10% sodium periodate aqueous solution (15 mL) and the emulsion was vigorously stirred for 40 h. After the reaction was complete as judged by TLC, the phases were separated and the aqueous layer was extracted three times with ethyl acetate. The combined organic phases were treated with isopropanol (3 mL) for 2 h in order to decompose unreacted ruthenium (VIII) tetroxide and later washed with 1 N hydrochloric acid (50 mL), dried over magnesium sulphate and evaporated to give a dark yellow oil. The crude was purified by silica chromatography (SNAP-25 column, gradient elution using 7–60% ethyl acetate in *n*-hexane) and 1-(2,2-dimethylpropanoyl)-2-methyl-5-oxo-L-proline **180** was obtained as a colourless oil (633 mg, 56%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.09 (*app bs*, 1H), 2.76-2.67 (*m*, 1H), 2.64-2.56 (*m*, 1H), 2.37-2.30 (*m*, 1H), 2.06-1.99 (*m*, 1H), 1.59 (*s*, 3H), 1.33 (*s*, 9H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 181.9, 178.5, 173.7, 66.7, 42.1, 31.4, 30.8, 26.2, 22.1; HR ESI-MS : 226.1092 ([M+H]$^+$, $C_{11}H_{20}NO_2^+$, calc. 226.1085).

(2*S*)-Di-*tert*-butyl pyrrolidine-1,2-dicarboxylate 183 [50]

L-Proline **134** (2.0 g, 17.37 mmol) was dissolved in 1,4-dioxane (50 mL) under a nitrogen atmosphere. Di-*tert*-butyl dicarbonate (7.64 g, 2 eq.) was added followed by triethylamine (9.6 mL, 4 eq.). The reaction mixture was stirred at room temperature for 16 h and the solvent was evaporated under reduced pressure to give the crude 1-(*tert*-butoxycarbonyl)-L-proline **210** which was immediately dissolved in a 1:1 mixture of *tert*-butanol and dichloromethane (25 mL each). 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide (2.35 g, 1 eq.) and 4-dimethylaminopyridine (2.12 g, 1 eq.) were subsequently added. The reaction mixture was stirred at room temperature for 18 h before diluting with ethyl acetate (50 mL) and quenching with a saturated sodium hydrogen carbonate solution (50 mL). The phases were separated and the aqueous layer was extracted twice with ethyl acetate (50 mL). The combined organic phases were washed with brine (50 mL), water (50 mL), dried over magnesium sulphate and evaporated to give the crude product as a yellow oil which was purified

using silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). (2*S*)-Di-*tert*-butyl pyrrolidine-1,2-dicarboxylate **183** was isolated as a colourless oil (3.65 g, 77.5% from L-proline **134**). Spectroscopic analyses show the presence of two conformational isomers, in good agreement with reported data [50].

TLC: 1:1 hex/EtOAc, R_f = 0.65; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.15 (*dd*, *J* = 8.5, 3.0, 0.3H), 4.07 (*dd*, *J* = 8.5, 3.0, 0.7H), 3.53-3.30 (*m*, 2H), 2.21-2.08 (*m*, 1H), 1.98-1.75 (*m*, 3H), 1.43 (*s*, 12.5H), 1.40 (*s*, 5.5H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 172.3, 172.2, 154.3, 153.9, 80.7, 79.5, 79.3, 59.6, 59.4, 46.4, 46.2, 30.8, 29.8, 28.3 (2x), 27.9 (2x), 24.1, 23.3; HR ESI-MS : 294.1677 ($[M+Na]^+$, $C_{14}H_{25}NNaO_4^+$, calc. 294.1676).

Di-*tert*-butyl 2-methylpyrrolidine-1,2-dicarboxylate 184

The synthesis of racemic di-*tert*-butyl 2-methylpyrrolidine-1,2-dicarboxylate **184** was adapted from a reported procedure for the C-2 alkylation of *N*-Boc-L-proline methyl ester [31]. A solution of *n*-butyl lithium in hexanes (2.5 M, 1.04 mL, 1.3 eq.) was added at -78 °C to a stirred solution of diisopropylamine (0.365 mL, 1.3 eq.) in dried tetrahydrofuran (10 mL) and the mixture was kept under a nitrogen atmosphere for 20 min prior to dropwise addition of (2*S*)-di-*tert*-butyl pyrrolidine-1,2-dicarboxylate **183** (543 mg, 2 mmol) dissolved in dried tetrahydrofuran (10 mL). After stirring at -78 °C for 20 min, methyl iodide (0.25 mL, 2.0 eq.) was added, the temperature was raised to 0 °C for 1 h and then to room temperature. When the reaction was complete, as judged by TLC analysis, it was quenched with a saturated ammonium chloride solution (20 mL), tetrahydrofuran was evaporated and the aqueous residue was extracted with diethyl ether. Combined organic phases were dried over magnesium sulphate, filtered and evaporated to give the crude as a pale orange oil. Di-*tert*-butyl 2-methylpyrrolidine-1,2-dicarboxylate **184** was obtained as a colourless oil (369 mg, 65%) after silica chromatography purification (SNAP-100 column, gradient elution with 0–50% ethyl acetate in *n*-hexane). NMR analysis shows two sets of signals corresponding to two conformational isomers.

TLC: 7:3 hex/EtOAc, R_f = 0.50; IR (neat) ν/cm^{-1}: 1735 (OC=O), 1699 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 3.59-3.48 (*m*, 1.3H), 3.44 (*t*, *J* = 6.7, 0.7H), 2.17-2.05 (*m*, 1H), 1.93-1.76 (*m*, 3H), 1.50 (*s*, 0.9H), 1.48 (*s*, 2.1H), 1,45-1.43 (*m*, 18H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 173.8, 173.7, 153.8, 80.6, 80.3, 79.7, 79.0, 65.6, 65.0, 48.1, 48.0, 40.7, 39.1, 28.5, 28.4, 27.9, 27.8, 23.4, 23.2, 22.5, 22.2; HR ESI-MS : 308.1832 ($[M+Na]^+$, $C_{15}H_{27}NNaO_4^+$, calc. 308.1832).

Di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate 185 [33]

The synthesis of di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185** was adapted from a reported procedure for the oxidation of prolines [33]. A solution of di-*tert*-butyl 2-methylpyrrolidine-1,2-dicarboxylate **184** (360 mg, 1.26 mmol) in ethyl acetate (5 mL) was added to a suspension of catalytic ruthenium (IV) oxide hydrate (50 mg) in 10% sodium periodate aqueous solution (10 mL) and the emulsion was vigorously stirred for 40 h. After the reaction was complete as judged by TLC, the phases were separated and the aqueous layer was extracted three times with ethyl acetate. The combined organic phases were treated with isopropanol (2 mL) for 2 h in order to decompose unreacted ruthenium (VIII) tetroxide and later washed with water (50 mL), dried over magnesium sulphate and evaporated to give a dark yellow oil. The crude was purified by silica chromatography (SNAP-25 column, gradient elution using 5–40% ethyl acetate in *n*-hexane) and di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185** was obtained as a colourless oil (320 mg, 85%).

TLC: 8:2 hex/EtOAc, R_f = 0.10; IR (neat) ν/cm^{-1}: 1741 (OC=O), 1706 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 2.68-2.43 (*m*, 2H), 2.16 (*ddd*, *J* = 13.0, 9.5, 6.0, 1H), 1.90 (*ddd*, *J* = 13.0, 9.5, 8.0,1H), 1.62 (*s*, 3H), 1.52 (*s*, 9H), 1.46 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.9, 171.6, 149.5, 83.4, 81.9, 65.8, 30.8, 30.6, 27.9, 27.8, 23.0; HR ESI-MS : 322.1624 ($[M+Na]^+$,

$C_{15}H_{25}NNaO_5^+$, calc. 322.1625).

Di-*tert*-butyl 5-hydroxy-2-methylpyrrolidine-1,2-dicarboxylate 182

Di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185** (320 mg, 1.07 mmol) was dissolved in dried tetrahydrofuran (5 mL) under a nitrogen atmosphere and the solution was cooled to -78 °C prior to dropwise addition of lithium triethylborohydride solution in tetrahydrofuran (1 M, 1 mL, 0.9 eq.). The reaction was stirred under these conditions for 3 h and lithium triethylborohydride solution was added again (1 M, 1 mL, 0.9 eq.). After another 3 h at -78 °C, the reaction was quenched with a saturated sodium hydrogen carbonate aqueous solution (5 mL) followed by a 40% aqueous hydrogen peroxide solution (0.1 mL). The temperature was raised to 0 °C and the mixture was stirred for 30 min. Tetrahydrofuran was evaporated and the resulting aqueous layer extracted three times with ethyl acetate (20 mL). The combined organic phases were washed with brine (20 mL), dried over magnesium sulphate, filtered and evaporated to give the crude product which was purified by silica chromatography (SNAP-25 column, gradient elution using 5–60% ethyl acetate in *n*-hexane). Di-*tert*-butyl 5-hydroxy-2-methylpyrrolidine-1,2-dicarboxylate **182** was obtained as a colourless oil (215 mg, 67%). Spectroscopic analyses show the presence of two conformational isomers.

TLC: 8:2 hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 1735 (OC=O), 1703 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.57 (*app bs*, 0.7H), 5.42 (*t*, *J* = 4.0, 0.3H), 2.45-2.37 (*m*, 1H), 1.99-1.83 (*m*, 3H), 1.62-1.54 (*m*, 3H), 1.48-1.40 (*m*, 18H); ^{13}C NMR (400 MHz, $CDCl_3$): δ = 173.3, 153.9, 83.3, 82.8, 81.2, 81.1, 80.9, 80.5, 66.3, 66.0, 37.2, 35.5, 32.0, 30.5, 28.4, 28.3, 27.8, 27.7, 22.1, 21.1; HR ESI-MS : 324.1781 ($[M+Na]^+$, $C_{15}H_{27}NNaO_5^+$, calc. 324.1781).

3.6.4 Synthesis of 5-Methyl-L-glutamate Semialdehyde and 2,5-Dimethylglutamate Semialdehyde

***tert*-Butyl *N*-(*tert*-butoxycarbonyl)-5-oxo-L-norleucinate 186**

Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** (0.5 g, 1.75 mmol) was dissolved in dried tetrahydrofuran (10 mL) and the solution cooled to -40 °C. At this temperature, a 3.0 M solution of methyl magnesium bromide in diethyl ether (0.70 mL, 2.1 mmol) was added and the reaction mixture stirred at -40 °C for 2 h. The reaction was quenched with a saturated ammonium chloride aqueous solution (5 mL). The reaction mixture was then allowed to reach to room temperature, diluted with diethyl ether (5 mL) and water (5 mL), and the phases were separated. The organic phase was washed with water (20 mL), dried over sodium sulphate and evaporated under reduced pressure. The crude residue was purified by silica chromatography (SNAP-25 column, gradient elution using 0–100% ethyl acetate in *n*-hexane) to yield *tert*-butyl *N*-(*tert*-butoxycarbonyl)-5-oxo-L-norleucinate **186** as a colourless oil (457 mg, 87%).

TLC: 1:1 *n*-hex/EtOAc, R_f = 0.60; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.08 (*d*, *J* = 7.0, 1H), 4.14-4.09 (*m*, 1H), 2.60-2.42 (*m*, 2H), 2.13 (*s*, 3H), 2.11-2.04 (*m*, 1H), 1.87-1.77 (*m*, 1H), 1.45 (*s*, 9H), 1.42 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 207.5, 171.4, 155.4, 82.0, 79.7, 53.4, 39.4, 29.9, 28.3, 27.9, 26.8; HR ESI-MS: 324.1777 ($[M+Na]^+$, $C_{15}H_{27}NNaO_5^+$, calc. 324.1781).

***tert*-Butyl *N*-(*tert*-butoxycarbonyl)-2-methyl-5-oxonorleucinate 187**

A solution of di-*tert*-butyl 2-methyl-5-oxopyrrolidine-1,2-dicarboxylate **185** (100 mg, 0.33 mmol) in dried tetrahydrofuran (5 mL) at -40 °C was treated with 3.0 M solution of methyl magnesium bromide in diethyl ether (0.13 mL, 1.2 eq.) as for compound **143** and purified to yield *tert*-butyl *N*-(*tert*-butoxycarbonyl)-2-methyl-5-oxonorleucinate **187** as a colourless oil (88 mg, 84%).

TLC: 1:1 *n*-hex/EtOAc, R_f = 0.70; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.36 (*app bs*, 1H), 2.49-2.42 (*m*, 1H), 2.34-2.23 (*m*, 1H), 2.13 (*s*, 3H), 2.08-1.93 (*m*, 1H), 1.86-1.82 (*m*, 1H), 1.66 (*s*, 3H), 1.46 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.1, 173.0, 92.2, 81.9, 80.7, 66.2, 59.1, 37.2, 30.7, 28.3, 27.8, 25.6; HR ESI-MS: 338.1932 ($[M+Na]^+$, $C_{16}H_{29}NNaO_5{}^+$, calc. 338.1938).

3.6.5 Synthesis of 3-Cysteamine-GHP Derivatives from *N*-Boc-Glycine

***tert*-Butyl [(*tert*-butoxycarbonyl)amino]acetate 197** [39]

tert-Butyl [(*tert*-butoxycarbonyl)amino]acetate **197** was prepared as described [39]. *N*,*N*-Dimethylformamide dineopentylacetal (15.0 mL, 3 eq.) was added dropwise over 30 min to a refluxed solution of [(*tert*-butoxycarbonyl)amino]acetic acid **196** (3.0 g, 17.1 mmol) in a 3:1 toluene/*tert*-butanol mixture (75 mL). The reaction was kept under reflux for 6 h until reaction completion, then allowed to cool down to room temperature before washing twice with a saturated sodium hydrogen carbonate solution (50 mL) and water (50 mL). The combined aqueous layers were reexctracted with toluene (20 mL) and organic layers were dried over magnesium sulphate, filtered and evaporated to give the crude product as a yellow oil, further purified by silica chromatography (SNAP-100 column, gradient elution with 0–20% ethyl acetate in *n*-hexane). The *tert*-butyl [(*tert*-butoxycarbonyl)amino]acetate **197** was obtained as a white solid (2.35 g, 60%).

mp 54–55 °C; TLC: 9:1 hex/EtOAc, R_f = 0.30; IR (KBr) ν/cm^{-1}: 3391 (NH), 1743 (OC=O), 1712 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.03 (*app bs*, 1H), 3.77 (*s*, 2H), 1.43 (*s*, 9H), 1.42 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 169.4, 155.6, 81.8, 79.6, 43.0, 28.2, 28.0; HR ESI-MS: 254.1359 ($[M+Na]^+$, $C_{11}H_{21}NNaO_4{}^+$, calc. 254.1363).

***tert*-Butyl [(*tert*-butoxycarbonyl)amino](diethoxyphosphoryl)acetate 199** [39]

To a solution of *tert*-butyl [(*tert*-butoxycarbonyl)amino]acetate **197** (2.0 g, 8.6 mmol) in degassed methyl tetrachloride (40 mL) was added a catalytic amount of radical reaction initiator benzoyl peroxide (5% mol/mol, 110 mg) followed by *N*-bromosuccinimide (1.55 g, 1 eq.). The reaction mixture was stirred under reflux conditions for 4 h, cooled to 0 °C, filtered to remove the precipitate and the filtrate was evaporated after washing the solid phase with chloroform to give the crude *tert*-butyl bromo[(*tert*-butoxycarbonyl)amino]acetate **198** as a pale yellow oil. This compound (TLC: 1:1 hex/EtOAc, R_f = 0.35) was found to be unstable and was therefore stored under a nitrogen atmosphere at -20 °C and used for the next step without further purification or characterisation. A solution of triethylphosphite (0.85 mL, 1.1 eq.) in anhydrous dichloromethane was added dropwise under a nitrogen atmosphere and at room temperature to a solution of the crude *tert*-butyl bromo[(*tert*-butoxycarbonyl)amino]acetate **198** in anhydrous dichloromethane (20 mL) and the reaction was stirred for 3 h before solvent evaporation. The resulting residue was purified using silica chromatography (SNAP-100 column, gradient elution with 0–20% ethyl acetate in *n*-hexane) and the product **199** was obtained as a low melting point white solid (2.18 g, 70% over 2 steps).

mp $\leq$ 35 °C; TLC: 1:1 hex/EtOAc, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.33 (*d*, *J* = 9.0, 1H), 4.69 (*dd*, *J* = 22.0, 9.0, 1H), 4.20-4.11 (*m*, 4H), 1.47 (*s*, 9H), 1.42 (*s*, 9H), 1.31 (*t*, *J* = 7.0, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 165.9, 154.9 (*d*, $^2J_{PC}$ = 8.0), 129.8, 128.1, 83.2, 80.4, 63.4 (*d*, $^2J_{PC}$ = 6.5), 63.2 (*d*, $^2J_{PC}$ = 6.5), 52.7 (*d*, $^1J_{PC}$ = 147.0), 28.1, 27.8, 16.3 (*d*, $^3J_{PC}$ = 5.5), 16.2 (*d*, $^3J_{PC}$ = 5.5); HR ESI-MS: 390.1654 ($[M+Na]^+$, $C_{15}H_{30}NNaO_7{}^+$, calc. 390.1652).

Methyl (1,3-dioxolan-2-yl)-acetate 202 [42]

Compound **202** was synthesised as described [42]. Palladium (II) chloride (612 mg, 10% mol/mol)

and copper (II) chloride (3.45 g, 1 eq.) were placed in a round-bottom flask under an oxygen atmosphere. A solution of methyl acrylate **201** (3.14 mL, 34.8 mmol) and ethylene glycol (1.95 mL, 1 eq.) in dimethyl ether ethylene glycol (40 mL) was added and the resulting suspention was stirred at 50 °C for 20 h. The reaction mixture was then diluted with diethyl ether (150 mL), filtered through a pad of Celite® 545 and evaporated under reduced pressure to give the crude product as a brown oil which was purified using silica chromatography (SNAP-100 column, gradient elution with 0–60% ethyl acetate in *n*-hexane). Methyl (1,3-dioxolan-2-yl)-acetate **202** was obtained as a yellow liquid (3.78 g, 75%).

TLC: 1:1 hex/EtOAc, R_f = 0.55; IR (neat) ν/cm^{-1}: 1742 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.95 (*t*, *J* = 5.3, 1H), 3.72-3.64 (*m*, 2H), 3.62-3.54 (*m*, 2H), 3.40 (*s*, 3H), 2.37 (*d*, *J* = 5.3, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 168.9, 100.2, 64.2, 50.8, 38.8; HR ESI-MS : 169.0469 ([M+Na]$^+$, $C_6H_{10}NaO_4^+$, calc. 169.0471).

2-(1,3-Dioxolan-2-yl)-ethanal 200 [39]

2-(1,3-Dioxolan-2-yl)-ethanal **200** was prepared as described [39]. To a solution of the methyl ester **202** (1.75 g, 12.0 mmol) in anhydrous dichloromethane (75 mL) cooled to -78 °C and under a nitrogen atmosphere, was added a 1 M solution of diisobutyl aluminium hydride in tetrahydrofuran (13.2 mL, 1.1 eq.) and the reaction was stirred under these conditions for 2 h. The addition of diisobutyl aluminium hydride (13.2 mL, 1.1 eq.) was repeated twice at an interval of 2 h for a total reaction time of 6 h before quenching with methanol (12 mL) followed by water (6 mL). The mixture was allowed to slowly reach room temperature and was then filtered through a pad of silica which was washed with diethyl ether and the filtrate was then evaporated to give the aldehyde **200** as a colourless oil (1.0 g, 72%). This unstable compound was stored at -20 °C and used for the next step without further purification or characterisation.

***tert*-Butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate 190** [39]

The synthesis of compound **190** has been described [39]. To a solution of potassium *tert*-butoxide (673 mg, 1.1 eq.) in dried tetrahydrofuran (40 mL) cooled to -78 °C and under a nitrogen atmosphere, was added a solution of *tert*-butyl [(*tert*-butoxycarbonyl)amino](diethoxyphosphoryl)acetate **199** (2.0 g, 5.4 mmol) in dried tetrahydrofuran (20 mL). After 5 min, a solution of crude aldehyde **200** (1.0 g, 1.6 eq.) in dried tetrahydrofuran (20 mL) was added dropwise over 10 min under the same conditions and the reaction mixture was slowly allowed to reach room temperature over 20 h before quenching with methanol (20 mL). The solvent was evaporated under reduced pressure and the residue redissolved in ethyl acetate (150 mL) and water (150 mL). The phases were separated and the organic layer washed 3 more times with water, dried over magnesium sulphate, filtered and evaporated, giving the crude product as a yellow oil. This residue was purified using silica chromatography (SNAP-100 column, gradient elution with 0–80% diethyl ether in petroleum ether 30-40) eluting *E*-**190** (114 mg, 6%) first, and followed by *Z*-**190** (1.71 g, 85%), both obtained as low melting point solids (91% overall yield).

Minor isomer *E*-**190**: mp $\leq$ 35 °C; TLC: 6:4 pet 30-40/Et_2O, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.74 (*app bs*, 1H + NH), 4.98 (*t*, *J* = 5.0, 1H), 4.03-3.95 (*m*, 2H), 3.90-3.82 (*m*, 2H), 2.90 (*dd*, *J* = 7.5, 5.0, 2H), 1.52 (*s*, 9H), 1.44 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 163.2, 152.9, 127.7, 103.7, 83.0, 80.0, 65.0, 53.5, 33.3, 28.2, 28.1; HR ESI-MS: 352.1732 ([M+Na]$^+$, $C_{16}H_{27}NNaO_6^+$, calc. 352.1731).

Major isomer *Z*-**190**: mp $\leq$ 35 °C; TLC: 6:4 pet 30-40/Et_2O, R_f = 0.15; IR (KBr) ν/cm^{-1}: 3371 (NH), 1714 (OC=O), 1695 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.37 (*app. t*, *J* = 7.5, 1H + NH), 4.98 (*t*, *J* = 4.5, 1H), 4.00-3.92 (*m*, 2H), 3.89-3.80 (*m*, 2H), 2.56 (*dd*, *J* = 7.5, 4.5, 2H), 1.46 (*s*, 9H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 163.6, 153.1, 125.9, 103.1, 81.5, 80.2, 65.7,

65.0, 32.9, 28.1, 27.9; HR ESI-MS: 352.1732 ([M+Na]$^+$, $C_{16}H_{27}NNaO_6^+$, calc. 352.1731).

3-(*N*-Acetylcysteamine-*tert*-butyl 2-*tert*-butyloxycarbonylamino)-4-(1,3-dioxolan-2-yl)-butanoate 204

tert-Butyl 2-[(*tert*-butoxycarbonyl)amino]-4-(1,3-dioxolan-2-yl)but-2-enoate *Z*-**190** (320 mg, 0.97 mmol) was dissolved in anhydrous dichloromethane (5 mL), potassium carbonate (829 mg, 6 eq.) was added and the resulting orange suspension was placed under a nitrogen atmosphere prior to addition of acetyl cysteamine (0.1 mL, 1 eq.) at room temperature. The reaction was found to be completed after 2 h as judged by TLC, and the solvent was evaporated. The residue was redissolved in diethyl ether (20 mL) and saturated sodium chloride aqueous solution (20 mL), the phases were separated and the aqueous layer was extracted with diethyl ether. The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give a yellow oil. This crude was purified using silica chromatography (SNAP-25 column, gradient elution with 5–100% ethyl acetate in *n*-hexane) to give two non-separable pairs of diastereomers of the protected 3-*N*-acetylcysteaminyl-GHP **204** as a colourless oil (282 mg, 65%).

TLC: 9:1 DCM/MeOH, R_f = 0.40; IR (neat) ν/cm^{-1}: 3318 (NH), 1716 (OC=O), 1661 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.55 (*app bs*, 1H), 6.44 (*d*, *J* = 6.5, 0.35H), 6.06 (*d*, *J* = 6.5, 0.65H), 5.06 (*m*, 1H), 4.28 (*dd*, *J* = 9.0, 4.5, 0.65H), 4.23 (*dd*, *J* = 9.0, 5.5, 0.35H), 3.90 (*m*, 2H), 3.81 (*m*, 2H), 3.23 (*m*, 2.65H), 3.15 (*dd*, *J* = 12.5, 6.5, 0.35H), 2.64 (*m*, 2H), 1.94 (*m*, 1H), 1.86 (*m*, 1H), 1.81 (*s*, 3H), 1.45 (*s*, 9H), 1.42 (*s*, 9H); HR ESI-MS : 471.2130 ([M+Na]$^+$, $C_{20}H_{36}N_2NaO_7S^+$, calc. 471.2135).

3-(*N*-*tert*-Butyloxycarbonylcysteamine-*tert*-butyl 2-*tert*-butyloxycarbonylamino)-4-(1,3-dioxolan-2-yl)-butanoate 203

Z-**190** (280 mg, 0.85 mmol) was dissolved in anhydrous dichloromethane (5 mL) at room temperature and powdered potassium carbonate (829 mg, 7 eq.) was added before placing the resulting orange suspension under a nitrogen atmosphere. *N*-Boc cysteamine was then added and the reaction was stirred under these conditions for 3 h. After reaction completion, as judged by TLC, solvent was evaporated and the residue was redissolved in diethyl ether (20 mL) and saturated sodium chloride aqueous solution (20 mL). The phases were separated and the aqueous layer was extracted twice with diethyl ether. The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give a pale yellow oil. This crude was purified using silica chromatography (SNAP-25 column, gradient elution with 10–100% diethyl ether in petroleum ether 30-40) to give the protected 3-cysteaminyl-GHP **203** as a low melting point white solid containing two non-separable pairs of diastereomers (314 mg, 62%).

mp $\leq$ 35 °C; TLC: 1:1 hex/EtOAc, R_f = 0.40; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.44-5.37 (*m*, 0.3H) 5.24-5.11 (*m*, 2.4H), 5.00 (*m*, 0.2H), 4.48 (*dd*, *J* = 9.5, 2.5, 0.8H), 4.44 (*dd*, *J* = 8.5, 4.0, 0.2H), 4-02-3.91 (*m*, 2H), 3.89-3.81 (*m*, 2H), 3.43-3.39 (*m*, 0.8H), 3.31 (*app bs*, 2H), 3.22-3.18 (*m*, 0.2H), 2.79-2.61 (*m*, 2H), 2.03-1.95 (*m*, 1H), 1.91-1.84 (*m*, 1H), 1.47-1.42 (*m*, 27H); HR ESI-MS : 529.2554 ([M+Na]$^+$, $C_{23}H_{42}N_2NaO_8S^+$, calc. 529.2554).

3.6.6 Synthesis of (5*S*)-5-(Hydroxymethyl)pyrrolidin-2-ol

(*S*)-5-([*tert*-Butyl(dimethyl)silyl]oxymethyl)pyrrolidin-2-one 211 [43, 44]

(*S*)-5-([*tert*-Butyl(dimethyl)silyl]oxymethyl)pyrrolidin-2-one **211** was prepared as described [43, 44]. (*S*)-5-(Hydroxymethyl)pyrrolidin-2-one **207** (0.5 g, 4.34 mmol) and imidazole (0.74 g, 2.5 eq.)

were dissolved in dried dimethylformamide (5 mL) and *tert*-butyldimethylsilyl chloride (0.79 g, 1.2 eq.) was added. After 24 h at room temperature, the reaction was diluted with diethyl ether (50 mL) and the organic layer was washed with water (20 mL) and brine (20 mL). The organic phase was dried over magnesium sulphate, filtered and the solvent was evaporated to give the product **211** as a colourless oil (0.99 g, 99%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.45 (*app bs*, 1H), 3.72-3.68 (*m*, 1H), 3.61-3.55 (*m*, 1H), 3.46-3.42 (*m*, 1H), 2.37-2.23 (*m*, 2H), 2.19-2.09 (*m*, 1H), 1.78-1.70 (*m*, 1H), 0.86 (*s*, 9H), 0.03 (*s*, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 178.3, 66.6, 55.7, 29.8, 25.7, 25.6, 22.7, 19.9, 18.1, -3.7, -5.6; LR ESI-MS: 230.15 ($[M+H]^+$).

***tert*-Butyl (*S*)-2-([*tert*-butyl(dimethyl)silyl]oxymethyl)5-oxopyrrolidin-1-carboxylate 208** [43, 44]

tert-
Butyl (*S*)-2-([*tert*-butyl(dimethyl)silyl]oxymethyl)5-oxopyrrolidin-1-carboxylate **208** was prepared as described [43, 44]. To a solution of (*S*)-5-([*tert*-butyl(dimethyl)silyl]oxymethyl)pyrrolidin-2-one **211** (0.95 g, 4.14 mmol) in anhydrous dichloromethane (20 mL) was added di-*tert*-butyl dicarbonate (1.9 g, 2 eq.), and 4-dimethylaminopyridine (0.55 g, 1 eq.), followed by triethylamine (0.9 mL, 1.5 eq.) under a nitrogen atmosphere. The reaction was stirred at room temperature for 24 h and then diluted with diethyl ether (30 mL). The organic layer was washed with water (50 mL) followed by a 0.1 N hydrochloric acid solution (50 mL), dried over magnesium sulphate, filtered and the solvent evaporated. The residue was purified by silica chromatography (SNAP-50 column, gradient elution using 0–100% ethyl acetate in *n*-hexane) and the product **208** was obtained as a colourless oil (0.95 g, 70%).

TLC: 7:3 *n*-hex/EtOAc, R_f = 0.40; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.15-4.12 (*m*, 1H), 3.90-3.86 (*m*, 1H), 3.66-3.63 (*m*, 1H), 2.71-2.62 (*m*, 1H), 2.37-2.29 (*m*, 1H), 2.12-1.94 (*m*, 2H), 1.49 (*s*, 9H), 0.84 (*s*, 9H), 0.01 (*s*, 3H), -0.01 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 174.8, 149.9, 82.5, 64.2, 58.8, 32.2, 28.0, 25.7, 21.0, 18.1, -5.7 (2x); LR ESI-MS: 330.20 ($[M+H]^+$).

***tert*-Butyl (2*S*)-2-([*tert*-butyl(dimethyl)silyl]oxymethyl)-5-hydroxypyrrolidine-1-carboxylate 209**

tert-Butyl (*S*)-2-([*tert*-butyl(dimethyl)silyl]oxymethyl)5-oxopyrrolidin-1-carboxylate **208** (0.9 g, 2.73 mmol) was dissolved into dried tetrahydrofuran (30 mL) under a nitrogen atmosphere and the solution was cooled to -78 °C prior to dropwise addition of a solution of lithium triethylborohydride in tetrahydrofuran (1 M, 2.5 mL, 0.9 eq.). The reaction was stirred at -78 °C for 3 h and a solution of lithium triethylborohydride in tetrahydrofuran (1 M, 2.5 mL, 0.9 eq.) was added again. After 3 h at -78 °C, the reaction was quenched by slow addition of a saturated sodium hydrogen carbonate aqueous solution (15 mL) followed by a 40% aqueous hydrogen peroxide solution (1 mL). The temperature was raised to 0 °C and the mixture stirred for 30 min. The organic solvent was evaporated and the aqueous layer extracted with diethyl ether. The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified by silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane). The product **209** was obtained as a colourless oil (0.52 g, 57%).

TLC: 8:2 *n*-hex/EtOAc, R_f = 0.30; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.49-5.24 (*m*, 1H), 4.00-3.47 (*m*, 3H), 2.22-1.78 (*m*, 4H), 1.48 (*s*, 9H), 0.90 (*s*, 9H), 0.08 (*s*, 3H), 0.03 (*s*, 3H); LR ESI-MS: 240.13 ($[M+Na]^+$).

***tert*-Butyl (5*S*)-2-hydroxy-5-(hydroxymethyl)pyrrolidine-1-carboxylate 206**

tert-Butyl (2*S*)-2-([*tert*-butyl(dimethyl)silyl]oxymethyl)-5-hydroxypyrrolidine-1-carboxylate **209** (100 mg, 0.3 mmol) was dissolved in dried tetrahydrofuran (2.5 mL) and the solution was cooled to 0 °C prior to addition of a 1 M tetra-*n*-butylammonium fluoride solution in tetrahydrofuran (0.35 mL, 1.1 eq.). The reaction was found to be completed after 2 h, as judged by TLC, and was quenched with water (5 mL). Dichloromethane (20 mL) was added and the phases were separated. The organic layer was dried over magnesium sulphate, filtered and evaporated. The residue was purified by silica chromatography (SNAP-25 column, gradient elution using 12–100% ethyl acetate in *n*-hexane). The product **206** was obtained as a colourless oil (53 mg, 81%).

TLC: 1:1 *n*-hex/EtOAc, $R_f = 0.10$; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.68-4.84 (*m*, 1H), 4.04-3.52 (*m*, 3H), 2.47-1.63 (*m*, 4H), 1.48 (*s*, 9H).

Chapter 4

CMPS Assays Using L-GHP Analogues

4.1 Introduction

4.1.1 Substrate Analogues Previously Tested for CMPS Catalysis

Prior to the start of my work, studies on the CMPS-catalysed production of carboxymethylproline derivatives had shown that wildtype CMPS, together with a series of CMPS variants, can accept a range of substrate analogues (*vide infra*, Figure 4.2).

Alkylmalonyl-CoA Derivatives are CMPS Substrates

Studies on wildtype CarB have shown that methylmalonyl-CoA **212** was accepted as a CMPS substrate to yield 6-methyl-*t*-CMP. However, the C-6 stereochemistry of the *t*-CMP product was not determined [1]. In addition, dimethylmalonyl-CoA **137** and ethylmalonyl-CoA **213** were accepted by CMPS, while isopropylmalonyl-CoA was not a CMPS substrate [2, 3, 4]. The substrate tolerances of wildtype CarB and ThnE were also found to be different, *e.g.* ThnE did not accept ethylmalonyl-CoA **213** as a substrate while CarB did [4]. The screening of a series of CarB and ThnE active site variants showed that some had improved C-6 stereoselectivity compared to wildtype enzymes with respect to the production of the *t*-CMP analogues alkylated at the C-6 position (Scheme 4.2 A) [5]. CarB variants could reach d.r. values as high as 95:5 in favor of the (6*R*)-6-methyl-*t*-CMP product (CarB M108V) or 10:90 in favor of the (6*S*)-6-methyl-*t*-CMP product (CarB W79F/M108A), a significant improvement compared to the wildtype CarB which showed no preference (d.r. = 50:50) [4]. Other

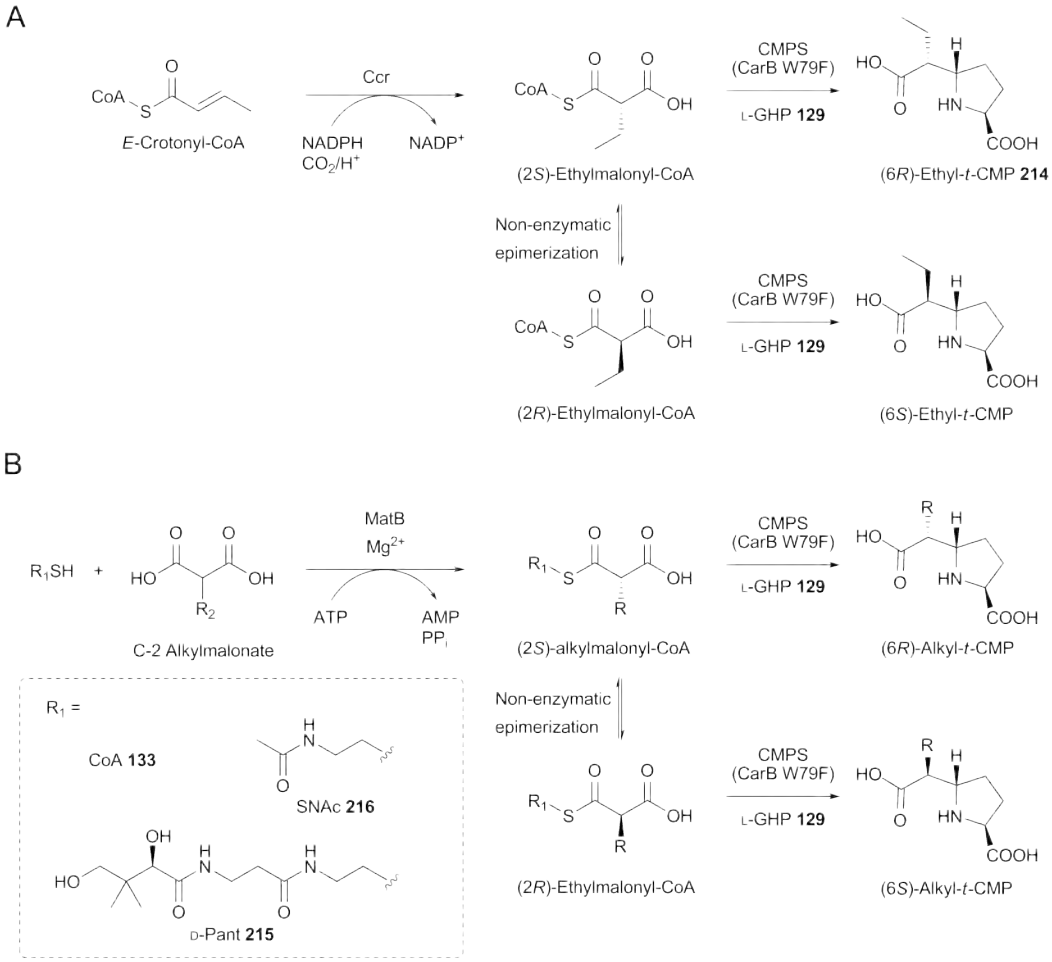
CarB variants showed stereoselectivity when incubated in the presence of epimeric ethylmalonyl-CoA **213** and L-GHP **129**. Stereoselective formation of both the (6*R*)-6-ethyl-*t*-CMP product (CarB M108L, d.r. = 72:28) and the (6*S*)-6-ethyl-*t*-CMP product (CarB W79A, d.r. = 12:88) could be achieved [5]. Wildtype ThnE shows a natural preference for the (6*R*)-6-methyl-*t*-CMP product (d.r. = 80:20). ThnE variants with decreased or inverted stereoselectivity were characterised (respectively ThnE W124F, d.r. = 55:45 and ThnE V153A, d.r. = 9:91) [5].

Enzymatic Production of Alkylmalonyl-CoA Derivatives

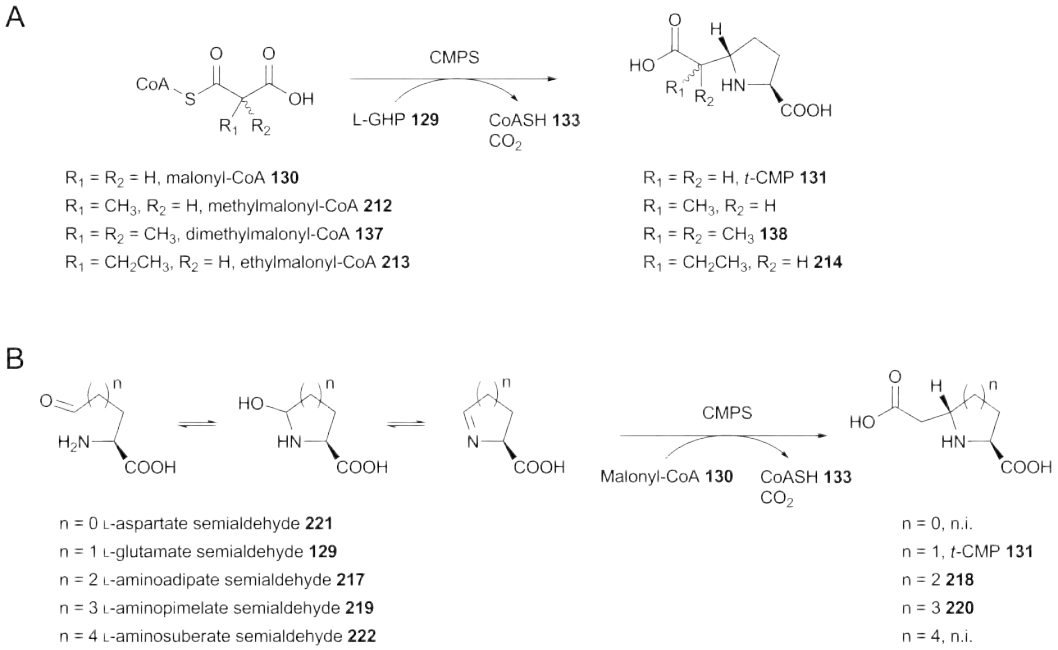
The use of recombinant crotonyl-CoA carboxylase reductase (CCR) from *Rhodobacter spheroides* [6] to produce enantiopure C-2 substituted malonyl-CoA substrates for CMPS assays was reported by Hamed *et al.* (Scheme 4.1 A) [5]. A Ccr/CMPS coupled assay allowed the generation of (6*R*)-ethyl-*t*-CMP **214** with excellent stereocontrol. However, this methodology is limited to the production of ethylmalonyl-CoA **213**, and is also hampered by the requirement for stoichiometric amount of crotonyl-CoA and NADPH. Biocatalytic processes greatly benefit from methodologies exploiting the regeneration of cofactors and substrates from cheap starting materials [7].

Enzymatic regeneration of malonyl-CoA derivatives and truncated versions thereof has been described [8]. One enzyme that could potentially produce C-2 alkylated malonyl-CoA substrates for the CMPS assay is the bacterial malonyl-CoA synthetase (MatB). The broad substrate specificity of *Rhizobium trifolii* MatB was first observed and reported by Kohsla and coworkers, and a dozen malonyl-CoA derivatives could be produced in detectable yield [9]. Together with these results, an application in engineered polyketide biosynthesis has been proposed [9]. Dordick and coworkers demonstrated the ability of MatB to regenerate malonyl-CoA from catalytic amounts of CoA in a polyketide synthase (PKS) coupled assay, but the authors did not try to expand the malonate extender unit structural diversity beyond the natural MatB substrate [8].

Two recent reports took the prospect of using MatB as a biocatalyst one step further; Hughes and Keatinge-Clay described the crystal structure of *Streptomyces coelicolor* MatB and the ability of this enzyme to use not only CoASH **133**, but also D-pantetheine **215** and *N*-acetylcysteamine **216** (SNAc) as substrates to form the corresponding malonic acid half thioester (MAHT, Scheme 4.1



Scheme 4.1: Towards C-6 alkylated *t*-CMP diversity and stereocontrol. Two examples of CMPS coupled assays that yield C-6 substituted *t*-CMP derivatives. (A) Ccr was successfully used in a CMPS coupled assay to increase the stereocontrol of product formation beyond that when using CMPS alone. When Ccr produces the CMPS substrate in a stereoselective manner, only one enolate product is formed upon decarboxylation, giving one single *t*-CMP product. A kinetic resolution process occurs to give near complete stereoselectivity, likely because the non-enzymatic epimerisation of the C-2 alkylmalonyl-CoA is relatively slow compared to the CMPS catalysed formation of 6-alkyl-*t*-CMP; (B) The stereoselective formation of C-2-akylated malonyl-CoA derivatives by MatB could potentially allow a similar kinetic resolution process when coupled to CMPS-catalysed formation of C-6 alkyl-*t*-CMP products.



Scheme 4.2: CMPS accepts a range of malonyl-CoA and amino acid aldehyde analogues. (A) CMPS enzymes catalyse the reaction between C-2 alkylated malonyl-CoA analogues and L-GHP to give the corresponding C-6 substituted *t*-CMP products; (B) Besides the natural substrate where n = 1, CMPS enzymes also accept L-GHP analogues with n = 2 and 3 to give the corresponding 6- and 7-membered heterocycles. CMPS assays using amino acid semialdehyde substrates with n = 0 and 4 did not yield any detectable product formation.

B) [10]. Inspired by this structural work on *S. coelicolor* MatB, Koryakina and Williams improved the enzyme ability to take C-2 substituted malonate substrates by applying a saturation mutagenesis methodology to two active site residues in *R. trifolii* MatB [11].

The possibility of using a bacterial malonyl-CoA synthetase (MatB) to generate MAHT from catalytic CoASH **133** or SNAc **216**, and C-2 substituted malonic acid derivatives, is currently being investigated in our laboratory (Scheme 4.1 B).

L-GHP Analogues are CMPS Substrates

CMPS biocatalysis has proven useful in the production of functionalised *N*-heterocycles of different ring size when used in combination with amino acid semialdehyde unnatural substrate analogues [12]. Based on the CarB crystal structures and biocatalytic studies, amino acid substitution in the CarB active site has been promising for the identification of CMPS variants with expanded substrate tolerance for bulkier L-GHP analogues, i.e. L-amino acid aldehydes with longer chains than L-GHP **129** (Scheme 4.2 B) [12]. Although the yield of the reaction was lower than for

L-GHP **129** itself, L-aminoadipate semialdehyde **217** was found to be a wildtype CarB substrate, yielding the (2*S*,6*S*)-6-(carboxymethyl)-pipecolic acid **218** [12]. A CarB variant with an active site histidine residue substituted for an alanine (CarB H229A) was shown to convert L-aminopimelate semialdehyde **219** into the 7-membered *N*-heterocylic product (2*S*,7*S*)-7-(carboxymethyl)-azepane-carboxylic acid **220**. The anticipated product of CMPS catalysis was not observed when L-aspartate semialdehyde **221** and L-aminosuberate semialdehyde **222** were used as substrates, showing the limits of CMPS tolerance to L-GHP analogues.

This chapter is a summary of investigations into the CMPS substrate selectivity carried out in order to probe the potential of ThnE and CarB variants as biocatalysts. The results below were obtained in collaboration with Dr Refaat B. Hamed and J. Ruben Gomez-Castellanos. Active site variants of CarB and ThnE were screened to select the biocatalysts giving high yields of the desired *t*-CMP product, and high stereoselectivity when applicable. The results for the CMPS and β-LS assays described in Section 4.3 were collected and analysed by J. Ruben Gomez-Castellanos and Dr Refaat B. Hamed, using CMPS variants prepared and purified by Daniel Harding, J. Ruben Gomez-Castellanos and Dr Refaat B. Hamed. The *t*-CMP and carbapenam products were purified and characterised by Dr Refaat B. Hamed. Part of the results described below have been published in reference [13].

4.2 Recombinant Expression and Purification of ThnE

For assay and crystallography purposes, ThnE was purified after recombinant expression in *E. coli*. The gene encoding for wildtype ThnE had previously been cloned from *S. cattleya* genomic DNA in the Schofield group [4]. When purified from *E. coli*, full length wildtype ThnE was found to be active, however, ESI-MS and Edman degradation analyses suggested that *N*- and *C*-terminal proteolysis had occurred. When the sequences of CarB and ThnE were compared, it was apparent that CarB lacked the first 46 amino acids of ThnE (Appendix Table B.1). It was therefore postulated that truncation of this ThnE *N*-terminal region would not interfere with catalysis. Indeed, recombinant ThnE Δ2–46 was found to give identical results to the full-length enzyme and showed higher levels of expression and purity [4]. In addition to this truncation, a silent mutation at position V-153 of wildtype ThnE lead to a significant increase in expression levels, possibly by decreasing the codon bias [5].

The gene encoding for ThnE V153V Δ2–46 had previously been cloned in a pET24-a(+) vector. After expression in *E. coli*, the recombinant protein was purified using a reported two-step procedure (Figure 4.1 A and B, fraction 25) [4]. In order to obtain a ThnE V153V Δ2–46 protein sample pure enough for the crystallisation experiments described in Chapter 6, an additional size exclusion chromatography step was performed. The fraction containing ThnE V153V Δ2–46 with the highest purity based on SDS-PAGE analysis was concentrated to 22 mg mL^{-1} (Figure 4.1 C, fraction 25). The other fractions were pooled together, concentrated to 18 mg mL^{-1} and kept for assay purposes.

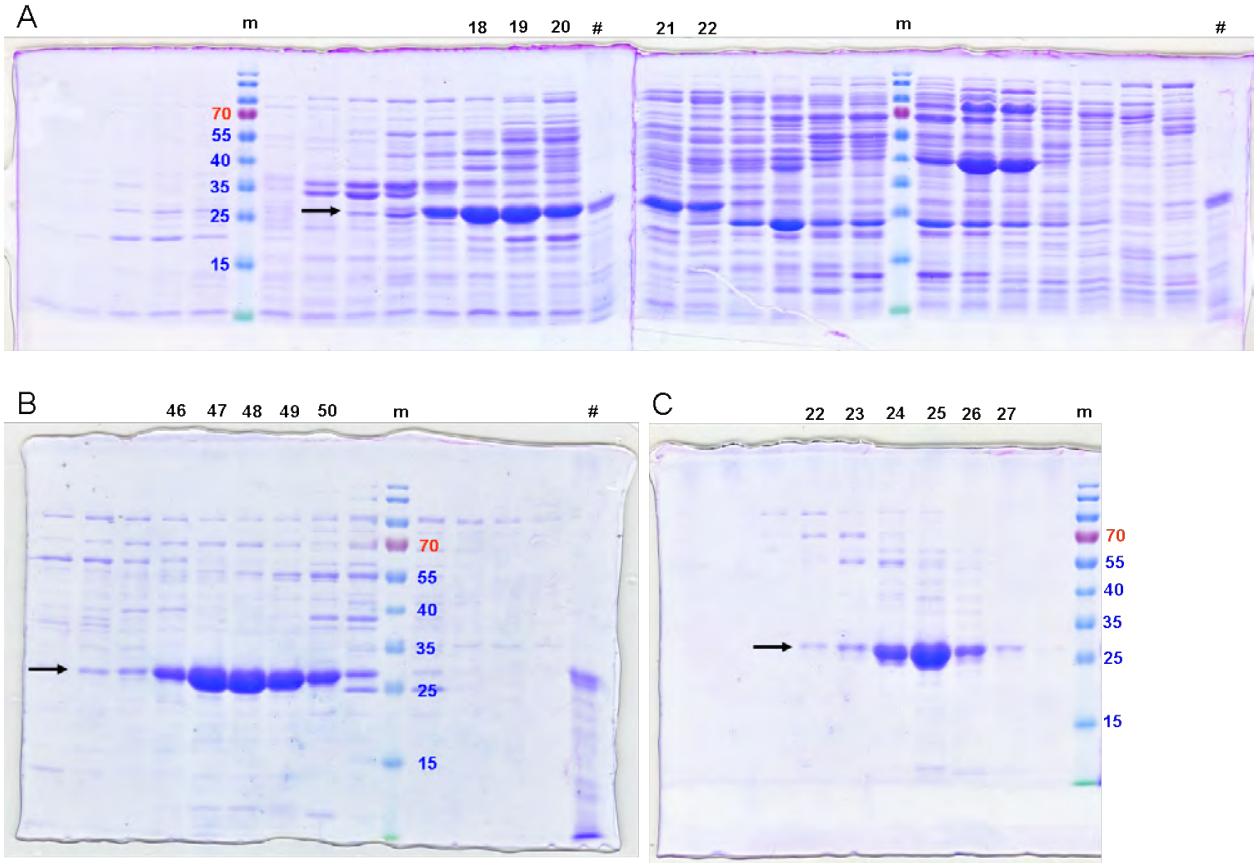


Figure 4.1: Purification of ThnE recombinant protein. SDS-PAGE analysis of ThnE purification after recombinant expression in *E. coli*. All samples were run on 12.5% acrylamide gels. The fractions containting the protein of interest (indicated by an arrow) are labelled. The molecular weights of the markers are given in kDa. The hash symbol indicate a previously purified ThnE Δ2–46 sample used as a reference. ThnE expected size was 27.4 kDa. (A) Fractions of the anion exchange chromatography purification step; (B) Fractions of the hydrophobic interactions chromatography; (C) Fractions of the size exclusion chromatography step. Fraction 25 was concentrated separately for attempted crystallisation.

4.3 Enzymatic Production of Substituted *t*-CMPs and Carbapenams from L-GHP Analogues

L-GHP **129** and L-GHP substrate analogues for CMPS assays were obtained by deprotection of the synthetic amino acid aldehyde precursors described in Chapter 3 using 10% formic acid in water as reported [12, 13]. The *t*-CMP products were detected by LC-MS analysis after performing the reported CMPS assay on an analytical scale [13]. Scale-up of the assay allowed isolation of the *t*-CMP products after LC-MS guided purification. The stereochemistry of all *t*-CMP products was assigned based on NMR data, including coupling constant (*J*) values and 2D NOESY analyses, assuming that the GHP analogues only react in their (2*S*)-stereochemistry (Section 2.5), and retain their stereochemistry at C-2 during the reaction. A summary of the results obtained for the CMPS catalysed formation of substituted *t*-CMP from GHP analogues carrying a methyl group at position C-2, C-3, C-4 and C-5 is given in Table 4.1. The CMPS variants giving the highest yield and/or diastereomeric ratio (d.r.) are shown in comparison with wildtype CarB and ThnE. The d.r. for the 4-methyl-*t*-CMP products was determined by ^{1}H NMR spectroscopy analysis of the mixture of products obtained under standard assay conditions. The isolated yield refers to cumulated yields for multiple steps, including deprotection of the amino acid aldehydes, CMPS enzymatic reaction, LC-MS guided purification and lyophilisation. The products were quantified by ^{1}H NMR spectroscopy using $[^{2}H]_4$-trimethylsilylpropionate as an external standard as reported [12].

4.3.1 Preparation of C-4 Methylated *t*-CMP and Carbapenam Derivatives

The deprotection of both synthetic enantioenriched (4*S*)-4-methyl-*N*-Boc-L-GHP *tert*-butyl ester **148** and (4*R*)-4-methyl-*N*-Boc-L-GHP *tert*-butyl ester **149** in aqueous formic acid was expected to give an epimeric mixture of 4-methyl-L-GHP **142** (Scheme 4.3 A), based on the observation that a deuterium atom can be incorporated at position C-4 when the deprotection reaction was carried out with deuterated formic acid in deuterium oxide (data not shown). The incorporation of two deuterium atoms was observed when L-GHP **129** was obtained under similar deprotection conditions [14].

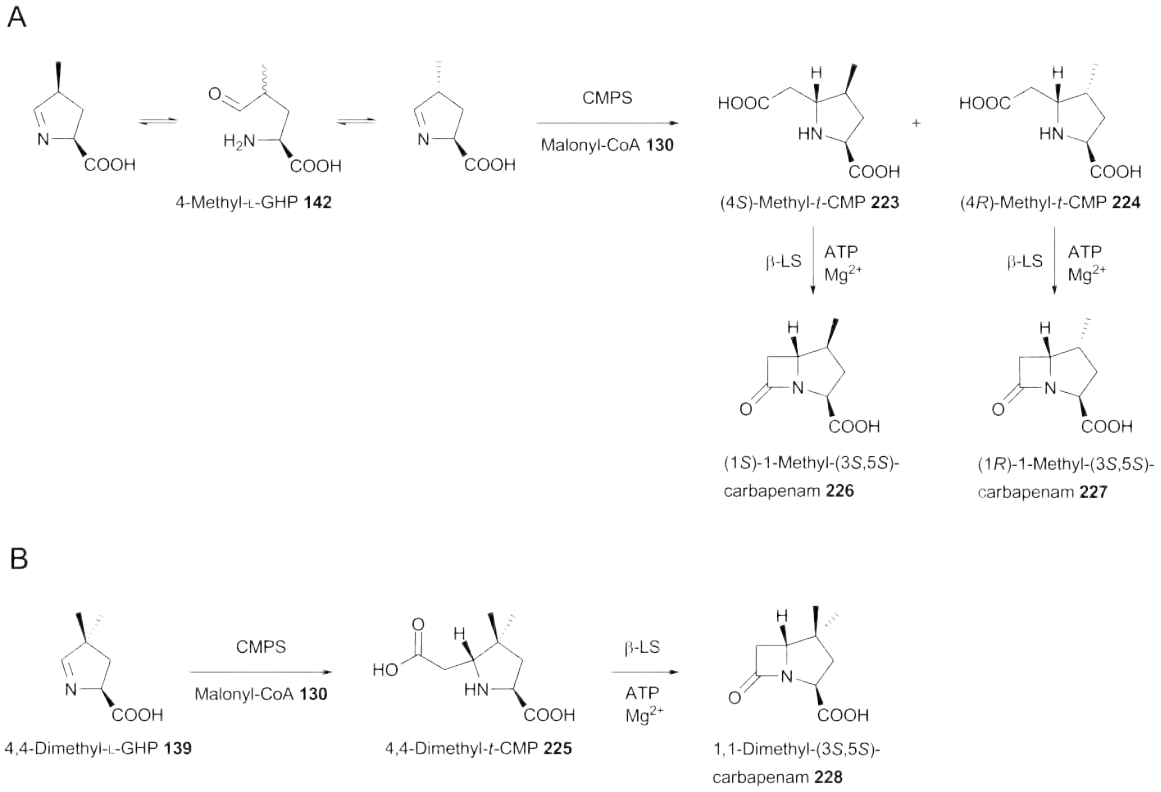
Incubation of 4-methyl-L-GHP **142** with wildtype CarB and malonyl-CoA **130** gave a $\cong$ 1:1 mixture of the C-4 epimeric products (4*S*)-4-methyl-*t*-CMP **223** and (4*R*)-4-methyl-*t*-CMP **224** as observed by analytical LC-MS (Figure 4.2 B). Given that previous studies have shown the ability

of CMPS variants to select from equilibrating mixtures of alkylmalonyl-CoA derivatives [5], the possibility of using CMPS variants in order to produce 4-methyl-*t*-CMP in a stereoselective manner from epimeric mixture of 4-methyl-L-GHP **142** was investigated.

4-Methyl-L-GHP **142** was incubated in the presence of malonyl-CoA **130** and all available CarB variants, prepared as described [5], and the ratio of products determined by analytical LC-MS (Figure 4.2 B) as well as NMR. All the CarB variants tested produced a mixture of both (4*S*)-4-methyl-*t*-CMP **223** and (4*R*)-4-methyl-*t*-CMP **224** (Scheme 4.3 A) but with varying yields and 4*S*:4*R* d.r. as follow: CarB M108A 54:46, CarB M108L 56:44, CarB M108V 93:7, CarB M108I 90:10, CarB Q111N 70:30, CarB W79F 61:39, CarB W79F/M108A 62:38, CarB W79A 64:36 and CarB H229A 65:35. Some of these variants represent a significant improvement in stereoselectivity compared to the wildtype CarB enzyme that shows no marked preference (Table 4.1, Entry I).

Screening of wildtype and active site variants of ThnE revealed that they can catalyse the formation of the two 4-methyl-*t*-CMP C-4 epimers **223** and **224**, with a general preference for the (4*S*)-epimer **223**. Wildtype ThnE shows a d.r. of 92:8 in favor of the (4*S*)-epimer **223** (Table 4.1, Entry II) which could be improved to 99:1 with a single point mutation in ThnE H274A (Table 4.1, Entry II). Scale-up allowed isolation of the (4*S*)-4-methyl-*t*-CMP **223** and (4*R*)-4-methyl-*t*-CMP **224** products and the d.r. confirmed by ^{1}H NMR analyses [13]. These results demonstrate the ability of some CMPS variants to stereoselectively produce (4*S*)-4-methyl-*t*-CMP **223** from an equilibrating mixture of the corresponding GHP substrate analogues, likely through a dynamic kinetic resolution process. Notably, (4*S*)-4-methyl-*t*-CMP **223** has the same C-4 stereochemistry as the 1-β-methylcarbapenems showing decreased susceptibility to DHP-1 hydrolysis (Chapter 2, Figure 2.4).

Under CMPS assay conditions, 4,4-dimethyl-L-GHP **139** was converted to a single product (Scheme 4.3 B, Table 4.1, Entry IV) assigned as 4,4-dimethyl-*t*-CMP **225** on the basis of LC-MS (Figure 4.2 B) and NMR analyses [13]. This result further confirms the proposed CMPS-catalysed C–C bond formation mechanism. As previously demonstrated with a L-GHP substrate carrying two deuterium atoms at position C-4 [14], the observation that 4,4-dimethyl-L-GHP **139** is a CMPS substrate rules out the possibility that the C–C bond formation proceeds *via* a mechanism that would



Scheme 4.3: The formation of 4-methyl and 4,4-dimethyl-*t*-CMP by CMPS. (A) An equilibrating mixture of 4-methyl-L-GHP **142** was found to be a substrate for CMPS catalysis yielding (4*S*)-methyl-*t*-CMP **223** and (4*R*)-methyl-*t*-CMP **224** in different ratios, depending on the CMPS variant (Table 4.1, entries I–III) [13]. The (4*S*)- and (4*R*)-methyl-*t*-CMP products **223** and **224** could be separated by semi-preparative HPLC purification and further converted into the corresponding carbapenams using CarA; (B) 4,4-Dimethyl-L-GHP **139** is also a CMPS substrate (Table 4.1, entries IV–V). The CarA assay in the presence of 4,4-dimethyl-*t*-CMP **225** yielded the corresponding 1,1-dimethyl-(3*S*,5*S*)-carbapenam [13].

involve C-4 deprotonation and a C-4/C-5 unsaturated intermediate (Scheme 2.9, Path B).

Upon incubation with the *P. carotovorum* β-LS (CarA), (4*S*)-4-methyl-*t*-CMP **223** and (4*R*)-4-methyl-*t*-CMP **224** and 4,4-dimethyl-*t*-CMP **225** were converted to the corresponding bicyclic compounds (1*S*)-1-methyl-(3*S*,5*S*)-carbapenam **226**, (1*R*)-1-methyl-(3S,5S)-carbapenam **227** and 1,1-dimethyl-(3*S*,5*S*)-carbapenam **228** (Scheme 4.3) [13]. In the case of the CarA-catalysed reaction of a 1:1 mixture of C-4 epimers **223** and **224**, the β-LS exhibited a diastereoselective bias towards the (4*S*)-epimer with a (4*R*)/(4*S*) ratio of $\sim$1:3. Notably, the resulting major product (1*S*)-1-methyl-(3*S*,5*S*)-carbapenam **226** has the same stereochemistry at C-4 as the commercially available carbapenems with improved stability towards DHP-1 (Figure 2.4).

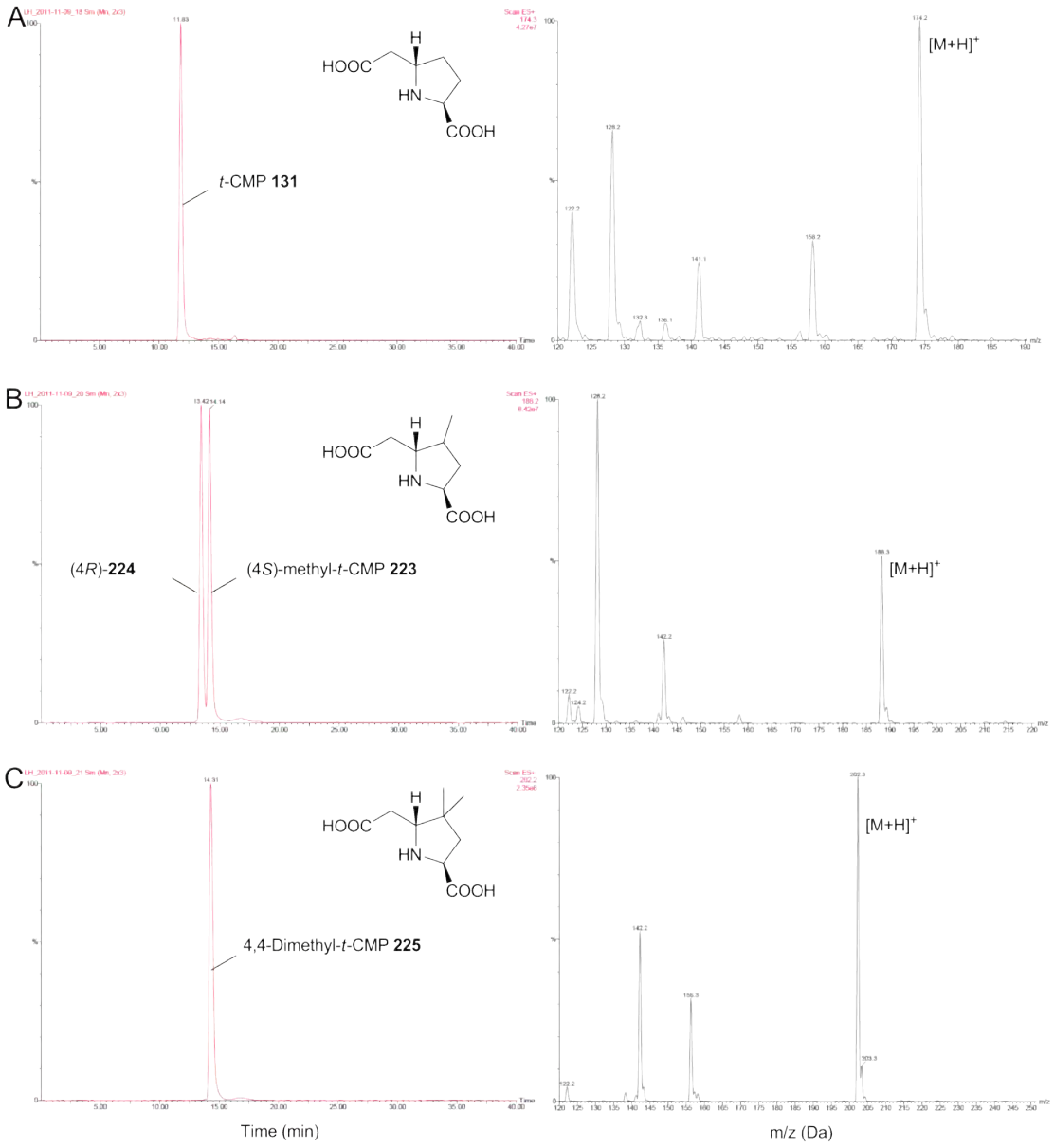


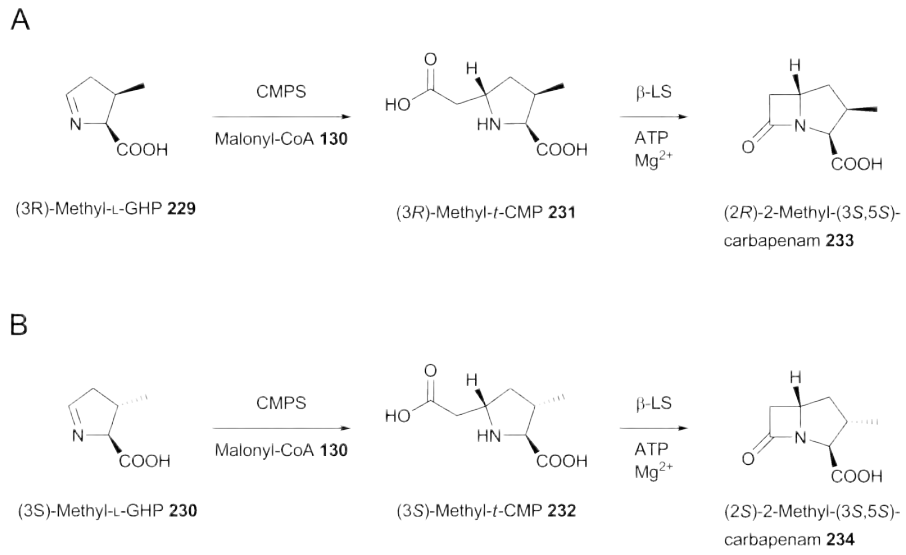
Figure 4.2: Analytical LC-MS detection of *t*-CMP products after CMPS catalysis. Extracted ion chromatogram (left) and mass spectrum (right) for *t*-CMP products. The relative yields for CMPS variants were determined by integrating the MS signal and using *p*-aminosalicylic acid as an internal standard (not shown). (A) Upon incubation of CMPS with L-GHP **129**, a single *t*-CMP product **131** is observed in positive ionisation mode (R_t = 11.8 min, m/z = 174.2); (B) When wildtype CarB is incubated with an epimeric mixture of 4-methyl-L-GHP, two isomers of 4-methyl-*t*-CMP (m/z = 188.2 in positive ionisation mode) can be separated. (4*R*)-methyl-*t*-CMP **224** elutes first (R_t = 13.4), followed by the (4*S*)-isomer **223** (R_t = 14.1 min) based on NMR characterisation [13]; (C) A single product was observed for 4,4-dimethyl-*t*-CMP in positive ionisation mode (R_t = 14.3 min, m/z = 202.2).

4.3.2 CMPS Catalyses the Formation of 3-Methyl-*t*-CMP but not of 3-Cysteaminyl-*t*-CMP

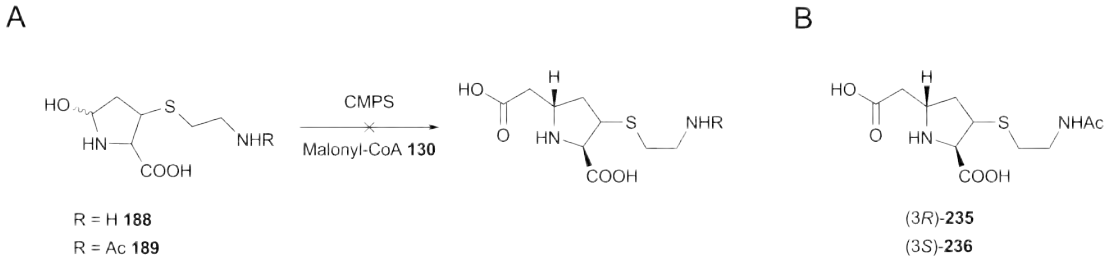
Our interest in testing GHP analogues substituted at C-3 arises from the presence of a cysteamine substituent at that position of clinically used carbapenem antibiotics (Figure 2.4). The possibility that a C-3 functionality could be accepted by an engineered CMPS might allow the introduction of a handle that could for example be chemically modified.

Deprotection of 3-methyl-GHP derivatives **172** and **173** under mild acidic conditions yielded (2*S*,3*S*)/(2*R*,3*R*)-3-methyl-GHP **229** and (2*S*,3*R*)/(2*R*,3*S*)-3-methyl-GHP **230**, respectively. These substrates were used in the CMPS assay without further purification (Scheme 4.4). Incubation of either **230** or **229** separately with various CMPS variants in the presence of malonyl-CoA **130** led to the identification of wildtype CarB as the best catalyst for the formation of both (3*R*)-methyl-*t*-CMP **231** (Scheme 4.4 A) and (3*S*)-methyl-*t*-CMP **232** (Scheme 4.4 B). Scale-up allowed the isolation and characterisation of a single major product for the incubation of (3*S*)- or (3*R*)-methyl-L-GHP separately, demonstrating that CMPS can catalyse the formation of C-3-substituted *t*-CMP derivatives without loss of stereochemistry at that position. The respective yields indicated that (2*S*,3*S*)-3-methyl-GHP **229** (Table 4.1, Entry VI) is a better substrate than the (2*S*,3*R*)-methyl-GHP analogue **230** (Table 4.1, Entry VIII). Wild type ThnE gave similar results, although in lower isolated yields (Table 4.1, Entries VII and IX). The bicyclic product of β-LS catalysis (2*R*)-2-methyl-(3*S*,5*S*)-carbapenam **233** and (2*S*)-2-methyl-(3*S*,5*S*)-carbapenam **234** were obtained upon incubation of β-LS CarA with either (3*R*)-methyl-*t*-CMP **231** or (3*S*)-methyl-*t*-CMP **232**.

In order to investigate the possibility of introducing a cysteamine side-chain prior to the CMPS-catalysed step, 3-cysteaminyl-GHP **188** and 3-(*N*-acetyl)-cysteaminyl-GHP **189** were obtained via acidic deprotection of the precursors **204** and **203** described in Section 3.3. Upon incubation of **188** or **189** in the presence of malonyl-CoA **130** and CMPS, no formation of the expected C-3 cysteaminyl-substituted *t*-CMP product was observed by LC-MS analysis (Scheme 4.5 A). The hypothesis of the cysteamine side chain being introduced before the ThnE catalysed step in thienamycin biosynthesis is likely to be invalid in light of a recent report that the two enantiomers **235** and **236** of the *t*-CMP substituted at the C-3 position with a *N*-acetylcysteamine group were not accepted as substrate for the



Scheme 4.4: CMPS catalyses the formation of 3-methyl-*t*-CMP. Each 3-methyl-*t*-CMP product was further turned into the corresponding carbapenams using CarA. (A) Incubation of (3*R*)-3-methyl-L-GHP **230** with malonyl-CoA and CMPS yielded (3*R*)-3-methyl-*t*-CMP **231** (Table 4.1, Entries VIII and IX) [13]; (B) Incubation of (3*S*)-3-methyl-L-GHP **229** with malonyl-CoA and CMPS yielded (3*S*)-3-methyl-*t*-CMP **232** (Table 4.1, Entries VI and VII) [13].



Scheme 4.5: 3-Cysteaminyl-GHP analogues are not CMPS substrates. (A) 3-cysteaminyl-GHP **188** and 3-(*N*-acetyl)-cysteaminyl-GHP **189** are not CMPS substrates; (B) The two 3-cysteaminyl-*t*-CMP compounds tested by Townsend and coworkers were not accepted by *S. cattleya* β-LS ThnM.

β-lactam synthetase ThnM (Scheme 4.5 B) [15].

4.3.3 CMPS Can Form *t*-CMP Products with Quaternary Carbon Centers

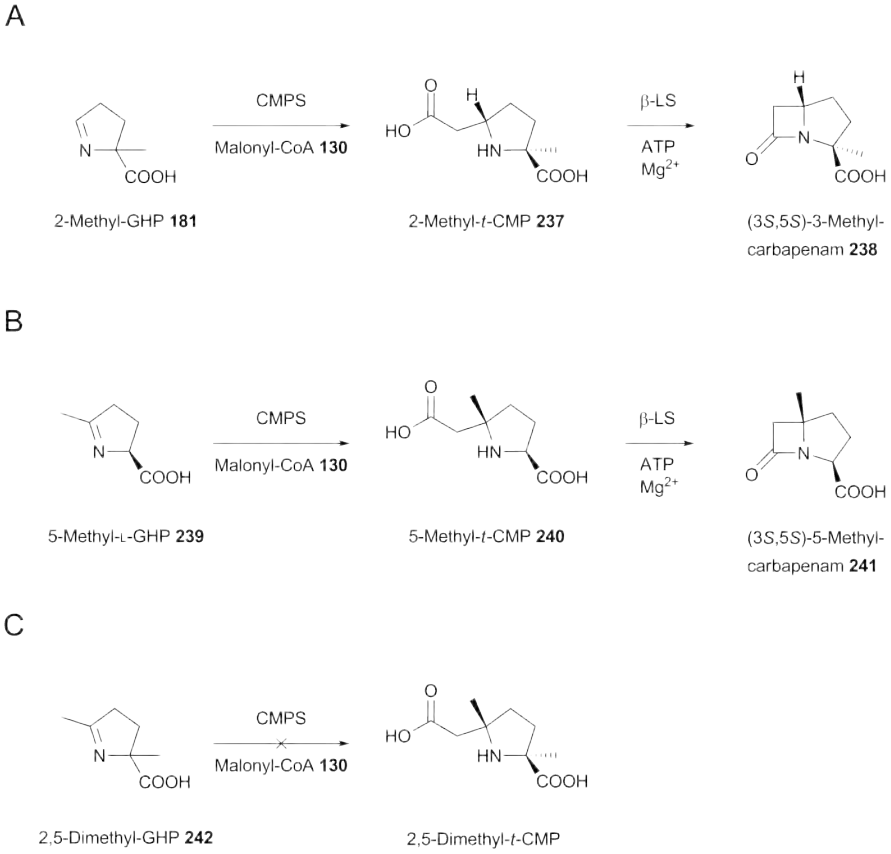
Generating quaternary carbon centers under mild conditions and in a stereoselective manner remains a challenge in organic chemistry. By testing C-2 and C-5 methyl-substituted GHP as substrates for CMPS-catalysed formation of the corresponding *t*-CMP products, it was hoped that CMPS could form a new quaternary center in a stereoselective manner. The production of bicyclic β-lactams with quaternary centers on the nucleus is of particular interest because such compounds have been chemically synthesised and can show improved stability to hydrolysis [16, 17]. There is precedent for

the biocatalytic production of bicyclic β-lactams with a quaternary center including on isopenicillin N synthase (IPNS) [17]. During penicillin and cephalosporin biosynthesis, IPNS catalyses the oxidative formation of the bicyclic core of isopenicillin N from an ACV tripeptide precursor (Section 2.1.3, Scheme 2.1). Substitution of the L-cysteine residue of ACV by (3*R*)-methyl-L-cysteine allowed the IPNS-catalysed formation of a 5α-methylpenicillin. This product showed no antibacterial activity but was more stable than penicillin G to β-lactamase 1 from *Bacillus cereus* [17]. Another report described the chemical synthesis of racemic 3-methylpenicillin G and 3,5-dimethylpenicillin G but these compounds did not show any antibacterial activity either [18].

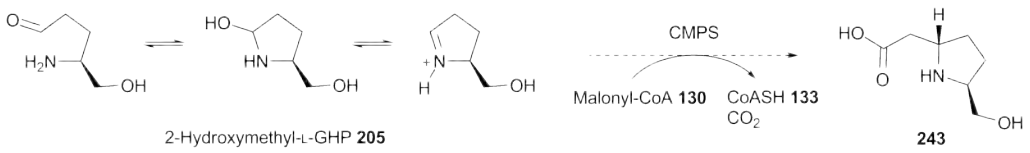
The CarB H229A variant was previously found to accept bulkier amino acid aldehyde substrates (Section 4.1.1) [12]. In the presence of racemic 2-methyl-GHP **181** and malonyl-CoA **130**, CarB H229A produced a single product having the expected mass for 2-methyl-*t*-CMP **237** (Scheme 4.6 A). CarB H229A gave the highest yield compared to other CMPS variants tested (Table 4.1 entries X–XII) [13]. Scale-up followed by LC-MS-guided preparative purification and NMR NOE analyses revealed a *trans*-relationship between the H-5 and the C-2 methyl group and the product was thus assigned (2*S*,5*S*)-2-methyl-*t*-CMP **237**. Upon incubation with CarA, (2*S*,5*S*)-2-methyl-*t*-CMP **237** was transformed into the bicyclic β-lactam (3*S*,5*S*)-3-methyl-carbapenam **238**.

Perhaps the most interesting result came from the observation that 5-methyl-L-GHP **239** was a CMPS substrate, yielding a single *t*-CMP derivative (Scheme 4.6 B). The ThnE V153A variant produced detectable levels of the corresponding 5-methyl-*t*-CMP **240** in a stereoselective manner (Table 4.1, Entry XV). This observation suggests that engineered CMPS biocatalysts can form a new quaternary carbon center *via* a sterically demanding C–C bond formation reaction between the enolate formed upon malonyl-CoA decarboxylation and 5-methyl-L-GHP **239**. This 5-methyl-*t*-CMP **240** was incubated with CarA in the presence of ATP to yield the 5-methylcarbapenam **241**.

In contrast, no product was observed when 2,5-dimethyl-GHP **242** was incubated with the available CMPS variants (Scheme 4.6 C). This is not surprising considering that two different active site substitutions were necessary for the production of the two monomethylated products **237** and **240** and that these compounds were not produced by any of the other CMPS variants. A ThnE variant with both V153A and H274A amino acid substitutions could accommodate 2,5-dimethyl-GHP **242** and should therefore be investigated.



Scheme 4.6: CMPS catalyses the formation of *t*-CMP products containing a quaternary center. (A) 2-Methyl-L-GHP **181** was accepted as a CMPS (CarB H229A) substrate. The resulting 2-methyl-*t*-CMP **237** could be further converted to the corresponding 3-methyl-(3*S*,5*S*)-carbapenam **238** by the β-LS CarA; (B) 5-Methyl-L-GHP **239** was accepted as a CMPS (ThnE V153A) substrate. The resulting 5-methyl-*t*-CMP **240** was also converted to the corresponding 5-methyl-(3*S*,5*S*)-carbapenam **241** by the β-LS CarA; (C) Incubation of 2,5-dimethyl-L-GHP in the presence of the CMPS variants available did not give any detectable amount of product.



Scheme 4.7: (5*S*)-5-(Hydroxymethyl)pyrrolidin-2-ol is not a CMPS Substrate. The L-GHP analogue lacking the 2-carboxylate functionality was not accepted as a CMPS substrate.

4.3.4 (5*S*)-5-(Hydroxymethyl)pyrrolidin-2-ol is not a CMPS Substrate

(5*S*)-5-(Hydroxymethyl)pyrrolidin-2-ol **205** was tested as a CMPS substrate after acidic deprotection of the *N*-Boc protected precursor **206**. When incubated in the presence of (5*S*)-5-(hydroxymethyl)pyrrolidin-2-ol **205**, none of the tested CMPS variants produced detectable amounts of *t*-CMP product **243** (Scheme 4.7). This observation suggests that the C-2 carboxylate moiety in L-GHP is important for productive catalysis.

4.4 Summary for Chapter 4

4.4.1 CMPS Enzymes Accept Methyl-Substituted L-GHP Substrates

The results presented in Section 4.3 demonstrate the utility of CMPS enzymes for the preparation of *t*-CMP derivatives with a methyl-substituant at all four carbon positions of the proline ring. Wildtype CarB and ThnE, as well as a series of active site variants, were systematically screened by analytical LC-MS for the formation of *t*-CMP products when incubated in the presence of malonyl-CoA **130** and GHP substrate analogues [13]. The results are summarised in Scheme 4.7 and in Table 4.1. The observation that CMPS is a stereoselective, robust and versatile biocatalyst is encouraging because it may allow the production of unnatural carbapenem by fermentation. L-GHP analogues carrying other substituent than methyl groups have to be tested in the future in order to further broaden the scope of CMPS biocatalysis.

When an equilibrating mixture of 4-methyl-L-GHP **142** was incubated in the presence of malonyl-CoA **130**, different CMPS variants gave different ratio of (4*S*)- to (4*R*)-methyl-*t*-CMP products. This observation suggest that these variants can have different preferences for either the (4*S*)- or (4*R*)-epimer of the substrate (Scheme 4.3) [13]. While wildtype CarB displayed no stereochemical preference, exploiting this dynamic equilibrium with stereoselective CMPS variants enabled the production of enantioenriched (4*S*)-methyl-*t*-CMP **223** with up to 98% ee (Table 4.1, Entry III). This result is probably achieved *via* a dynamic kinetic resolution process which is taking place in the assay. The preferred substrate is depleted by a given CMPS variant and the C-4 epimerisation equilibrium therefore continuously displaced. (4*S*)-Methyl-*t*-CMP **223** was of particular interest because it could potentially be a direct precursor for the clinically used carbapenem antibiotics carrying the corresponding C-2 β-methyl group (Figure 2.4).

CMPS variants also accepted C-2 racemates of (3*S*)- and (3*R*)-methyl-L-GHP and gave the corresponding enantiopure (2*S*,3*S*)- and (2*S*,3*R*)-methyl-*t*-CMP products (Table 4.1, Entries VI–IX). The relative stereochemistry at C-2, C-3 and C-5 for these compounds was verified by NOESY NMR spectroscopy [13].

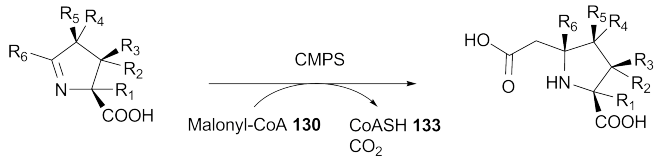
Under standard *in vitro* assay conditions, CMPS variants were found to accept 2-methyl-L-GHP **181** with a quaternary center at C-2, yielding an enantiopure *t*-CMP product **237** from a racemic synthetic substrate analogue. A CarB variant with an increased space in the active site (CarB H229A) was found to yield a 5–15 fold increase in 2-methyl-*t*-CMP **237** production compared to wildtype CarB or ThnE (Table 4.1, Entries X–XII). This is in good agreement with the same variant giving the best yields for products with a larger ring size than *t*-CMP (Scheme 4.2) [12]. This result is important for two reasons. First the ability of CMPS to discriminate between the two C-2 enantiomers allows to stereoselectively produce a compound with a chiral quaternary center. Second, the absence of proton at C-2 rules out the possibility of a CMPS mechanism proceeding *via* C-2 proton abstraction.

Finally, compound **239**, the L-GHP analogue carrying a methyl group at position C-5, was also found to be a CMPS substrate. When 5-methyl-L-GHP **239** was incubated under standard assay conditions in the presence of malonyl-CoA **130**, a single 5-methyl-*t*-CMP diastereomer (2*S*,5*S*)-5-methyl-CMP **240** was observed. This result is of importance from a biocatalytic perspective. Although quaternary carbons are a common feature of terpenoid natural products, they are formed mainly *via* enzyme-catalysed polyene cyclisations and cationic rearrangements in which fused rings are formed *via* cascade cyclisations. The synthesis of isolated and contiguous all-carbon quaternary stereocenters is a particularly challenging task, and biocatalysis may be a promising solution.

These results suggest that functionalities absent from the natural CMPS substrates could be introduced at an early stage in a reprogrammed carbapenems biosynthesis pathway. The β-methyl substituent at the C-1 position of the clinically used carbapenems which confers resistance to DHP-1 (Section 2.1.2, Figure 2.4) could be installed by an organism containing a CMPS variant using 4-methyl-L-GHP as a substrate and was therefore of particular interest.

4.4.2 Methyl-Substituted *t*-CMP Derivatives are β-LS Substrates

In agreement with the previously reported β-LS promiscuity [19], CarA was found to accept the methyl-substituted *t*-CMP derivatives obtained from CMPS-catalysed condensation of a methyl-L-GHP substrates with malonyl-CoA. The resulting carbapenam CarA products were isolated and the presence of methyl group was found to yield improved β-lactam stability towards hydrolysis [13]. However, none of the carbapenam products had antibacterial activity [13]. This is in agreement with

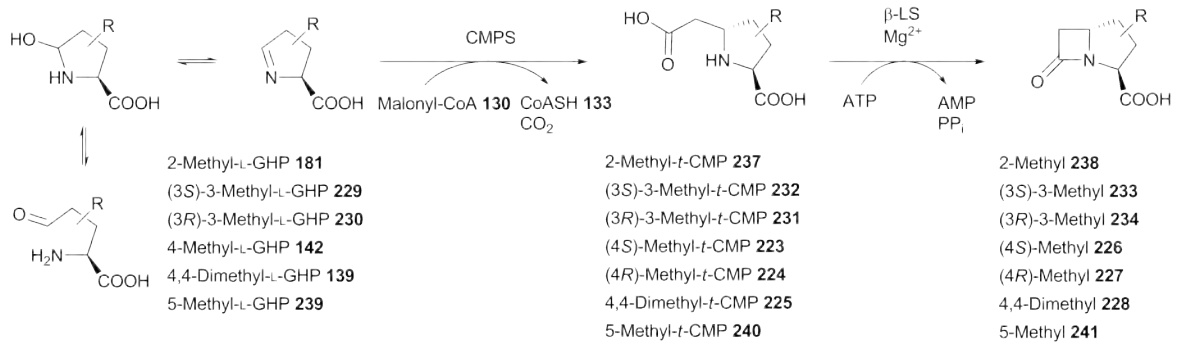


Entry	GHP	R^a	CMPS variant[b]	CMP Product	R^a	d.r.[†] (%)	Yield[‡] (%)
I	**142**	R_4 or $R_5 = CH_3$	CarB	**223/224**	$R_5 = CH_3$ (**223**)	51:49	31
II	**142**	R_4 or $R_5 = CH_3$	ThnE	**223/224**	$R_5 = CH_3$ (**223**)	92:8	44
III	**142**	R_4 or $R_5 = CH_3$	ThnE H274A	**223/224**	$R_5 = CH_3$ (**223**)	99:1	39
IV	**139**	$R_4 = R_5 = CH_3$	CarB	**225**	$R_4 = R_5 = CH_3$	n/a	48
V	**139**	$R_4 = R_5 = CH_3$	ThnE	**225**	$R_4 = R_5 = CH_3$	n/a	32
VI	**229**	$R_2 = CH_3$	CarB	**232**	$R_2 = CH_3$	n/a	16 (32)
VII	**229**	$R_2 = CH_3$	ThnE	**232**	$R_2 = CH_3$	n/a	9 (18)
VIII	**230**	$R_3 = CH_3$	CarB	**231**	$R_3 = CH_3$	n/a	12 (24)
IX	**230**	$R_3 = CH_3$	ThnE	**231**	$R_3 = CH_3$	n/a	2 (4)
X	**181**	$R_1 = CH_3$	CarB	**237**	$R_1 = CH_3$	n/a	3 (6)
XI	**181**	$R_1 = CH_3$	ThnE	**237**	$R_1 = CH_3$	n/a	$\geq$ 1 (2)
XII	**181**	$R_1 = CH_3$	CarB H229A	**237**	$R_1 = CH_3$	n/a	16 (32)
XIII	**239**	$R_6 = CH_3$	CarB	**240**	$R_6 = CH_3$	n/a	-
XIV	**239**	$R_6 = CH_3$	ThnE	**240**	$R_6 = CH_3$	n/a	1
XV	**239**	$R_6 = CH_3$	ThnE V153A	**240**	$R_6 = CH_3$	n/a	10

Summary of the results for the chemoenzymatic synthesis of functionalised *t*-CMP from malonyl-CoA **130** and methylated L-GHP derivatives using CMPSs. [a]R = H where unspecified; [b]The CMPS variant giving the highest yield and/or diastereomeric ratio are shown in comparison with wildtype CarB and ThnE; [†]d.r.: diastereomeric ratio (*S*:*R*) as determined by ^{1}H NMR spectroscopy for standard assay conditions. [‡]The yield (isolated) refers to cumulated yields for the following steps: deprotection of amino acid aldehydes, enzymatic reaction, LC-MS purification and lyophilisation. Products were quantified by ^{1}H NMR spectroscopy using $[^2H]_4$-trimethylsilylpropionate as an external standard as reported [12].

Table 4.1: CMPS-catalysed formation of methyl-substituted *t*-CMP products.

the requirement for the CarC-catalysed epimerisation and desaturation reactions required for (5*S*)-1-carbapen-2-em-3-carboxylic acid **115** antibacterial activity (Section 2.1.2).



Scheme 4.8: Methyl-substituted carbapenams can be accessed using CMPS and β-LS biocatalysis. The L-GHP analogues carrying a methyl group at position C-2, C-3, C-4 or C-5 of the amino acid semialdehyde were found to be substrates for CMPS enzymes. Further incubation of these *t*-CMP derivatives with β-LS and ATP yielded the corresponding methyl-substituted carbapenams.

References

[1] Gerratana, B., Arnett, S. O., Stapon, A., and Townsend, C. A. (2004) Carboxymethylproline synthase from *Pectobacterium carotorova*: A multifaceted member of the crotonase superfamily. *Biochemistry*, **43**, 15936–15945.

[2] Sleeman, M. C. and Schofield, C. J. (2004) Carboxymethylproline synthase (CarB), an unusual carbon-carbon bond-forming enzyme of the crotonase superfamily involved in carbapenem biosynthesis. *J Biol Chem*, **279**, 6730–6736.

[3] Batchelar, E. T., Hamed, R. B., Ducho, C., Claridge, T. D. W., Edelmann, M. J., Kessler, B., and Schofield, C. J. (2008) Thioester hydrolysis and C–C bond formation by carboxymethylproline synthase from the crotonase superfamily. *Angew Chem Int Ed Engl*, **47**, 9322–9325.

[4] Hamed, R. B., Batchelar, E. T., Mecinović, J., Claridge, T. D. W., and Schofield, C. J. (2008) Evidence that thienamycin biosynthesis proceeds via C-5 epimerization: ThnE catalyzes the formation of (2*S*,5*S*)-*trans*-carboxymethylproline. *ChemBioChem*, **10**, 246–250.

[5] Hamed, R. B., Gomez-Castellanos, J. R., Thalhammer, A., Harding, D., Ducho, C., Claridge, T. D. W., and Schofield, C. J. (2011) Stereoselective C-C bond formation catalysed by engineered carboxymethylproline synthases. *Nat Chem*, **3**, 365–371.

[6] Erb, T. J., Brecht, V., Fuchs, G., Mller, M., and Alber, B. E. (2009) Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc Natl Acad Sci U S A*, **106**, 8871–8876.

[7] Zhao, H. and van der Donk, W. A. (2003) Regeneration of cofactors for use in biocatalysis. *Curr Opin Biotechnol*, **14**, 583–589.

[8] Kim, M. I., Kwon, S. J., and Dordick, J. S. (2009) *In vitro* precursor-directed synthesis of polyketide analogues with coenzyme A regeneration for the development of antiangiogenic agents. *Org Lett*, **11**, 3806–3809.

[9] Pohl, N. L., Hans, M., Lee, H. Y., Kim, Y. S., Cane, D. E., and Khosla, C. (2001) Remarkably broad substrate tolerance of malonyl-CoA synthetase, an enzyme capable of intracellular synthesis of polyketide precursors. *J Am Chem Soc*, **123**, 5822–5823.

[10] Hughes, A. J. and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for *in vitro* polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem Biol*, **18**, 165–176.

[11] Raber, M. L., Castillo, A., Greer, A., and Townsend, C. A. (2009) A conserved lysine in β-lactam synthetase assists ring cyclization: Implications for clavam and carbapenem biosynthesis. *ChemBioChem*, **10**, 2904–2912.

[12] Hamed, R. B., Mecinović, J., Ducho, C., Claridge, T. D. W., and Schofield, C. J. (2010) Carboxymethylproline synthase catalysed syntheses of functionalised *N*-heterocycles. *Chem Commun*, **46**, 1413–1415.

[13] Hamed, R. B., Henry, L., Gomez-Castellanos, J. R., Mecinović, J., Ducho, C., Sorensen, J. L., Claridge, T. D. W., and Schofield, C. J. (2012) Crotonase catalysis enables flexible production of functionalized prolines and carbapenams. *J Am Chem Soc*, **134**, 471–479.

[14] Ducho, C., Hamed, R. B., Batchelar, E. T., Sorensen, J. L., Odell, B., and Schofield, C. J. (2009) Synthesis of regio- and stereoselectively deuterium-labelled derivatives of L-glutamate semialdehyde for studies on carbapenem biosynthesis. *Org Biomol Chem*, **7**, 2770–2779.

[15] Bodner, M. J., Li, R., Phelan, R. M., Freeman, M. F., Moshos, K. A., Lloyd, E. P., and Townsend, C. A. (2011) Definition of the common and divergent steps in carbapenem β-lactam antibiotic biosynthesis. *ChemBioChem*, **12**, 2159–2165.

[16] Onoue, H. and Narukawa, Y. (1989) Synthesis and antibacterial activity of 5-methylcarbapenems. *J Antibiot (Tokyo)*, **42**, 1100–1113.

[17] Baldwin, J. E., Adlington, R. M., Moss, N., and Robinson, N. G. (1987) Penicillin biosynthesis: Active site mapping with L-α-aminoadipoyl-(C-methyl-L-cysteinyl)-D-valine variants. *J Chem Soc, Chem Commun*, **1987**, 1664–1667.

[18] Meyer, R., Schöllkopf, U., Madawinata, K., and Stafforst, D. (1978) Totalsynthese von ($\pm$)-2,2-diethyl-3-methyl- und ($\pm$)-2,2-diethyl-3,6-dimethyl-6-(phenoxyacetamido)penam-3-carbonsäure. *Justus Liebigs Annalen der Chemie*, **1978**, 1982–1989.

[19] Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. *Biochemistry*, **42**, 7836–7847.

[20] Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbour Laboratory Press, 2nd edn.

[21] Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, **182**, 319–326.

[22] Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

4.5 Materials and Methods for Chapter 4

4.5.1 General Material and Methods for Molecular Biology

All the molecular biology protocols were carried out according to standard procedures in place in the group of Prof Christopher J. Schofield at the University of Oxford.

IPTG was from Melford Laboratories Ltd., electrophoresis grade agarose was from Bioline, TEMED, SDS and APS were from Sigma-Aldrich, a 30% acrylamide/bisacrylamide mixture (37.5:1) was from Roche Diagnostics Ltd. Molecular weight marker P7708S for SDS-PAGE was from New England Biolabs. Bacto Tryptone, Yeast Extract and Bacto Agar for use in culture media were from Oxoid and Difco. Unless otherwise stated, plasmids and restriction enzymes were from Promega, Novagen, New England BioLabs, and Stratagene. Molecular weight marker for SDS-PAGE (prestained protein marker) was from Invitrogen. 1 kb DNA ladder for DNA electrophoresis was from New England Biolabs. Custom DNA oligonucleotides were form Sigma. Other molecular biology materials were from QiaGen and Roche. Competent *E. coli* cells were from Stratagene. Fast protein liquid chromatography (FPLC) columns and equipment, and small-scale gel filtration columns (PD-10) were from Amersham Biosciences. Spin concentrators for protein concentration were from Amicon. Water was purified by a Millipore Milli-Q™ system fitted with a 0.22 μm filter at the outlet. All standard solutions used in molecular biology and microbiology were prepared according to standard procedures using Milli-Q™ water and autoclaved, or sterilised by filtration, as required. Optical density (OD) of cell cultures was measured by monitoring the absorbance at 600 nm on a Cary® 4000 spectrometer (Varian Inc.).

pH measurements

The pH values for solutions and buffers were measured using a Mettler Delta 340 pH meter, with a Mettler Toledo 410 electrode. Calibration was carried out at pH 4.0, 7.0 and 10.0 immediately prior to use with standards suppled by Fisher Scientific. The electrode was stored in saturated potassium chloride aqueous solution (4 M).

Centrifugation

Samples up to 1.5 mL were centrifuged at room temperature using an AccuSpin™ Micro (Fisher Scientific) or at 4 °C using a Microfuge® 22R (Beckman-Coulter). Larger samples (15–800 mL) were centrifuged at 4 °C using either an Allegra™ 21R or Avanti™ J-25 (Beckman-Coulter). Typically, *E. coli* cell cultures were centrifugated on an Avanti using a JA-10 rotor at 8 krpm. Cell lysate centrifugation was carried out on the Avanti using a JA-25.50 rotor at 18 krpm, and proteins were concentrated by centrifugation on the Allegra using a SX4250 swing bucket rotor at 4 krpm.

Oligonucleotide concentration measurements

The concentration of DNA oligonucleotides was determined using a NanoDrop® ND-1000 spectrometer. DNA can be quantified by exploiting the absorbance maximum of the nucleobases at λ = 260 nm. The purity of the DNA preparation were estimated by calculating the ratio of measured absorbances at 260 and 280 nm, where protein and single-stranded DNA/RNA contaminants have a stronger absorbance. A 260 nm/280 nm ratio of 1.8 or higher is usually acceptable for molecular biology applications.

Polymerase chain reaction

Polymerase chain reactions (PCR) were carried out using a Techne® Endurance TC-312 thermal cycler. Custom oligonucleotides were obtained as lyophilised pellets and were resuspended in sterile water to a stock concentration of 100 ng mL^{-1}. Fresh dilution to 10 ng mL^{-1} were made immediately prior to PCR protocol. The gene fragment amplification PCR experiments were carried out on reaction mixtures as described in Table 4.2 using the PCR protocol described in Table 4.3.

Reagent	Concentration	Volume (μL)
DNA template	25 ng mL^{-1}	1.0/3.0/5.0
Forward primer	10 μM	1.0
Reverse primer	10 μM	1.0
dNTP mix	10 mM	1.0
$MgCl_2$	50 mM	1.5
DMSO		10.0
Pfu buffer	10x	5.0
DNA polymerase	2.5 U μL^{-1}	1.0
Water		28.5/26.5/24.5

Pfu Ultra® DNA polymerase and the corresponding buffer was obtained from Stratagene. dNTP mix was purchased as a 100 mM stock solution from Stratagene and diluted to 10 mM immediately prior to use.

Table 4.2: PCR mix for gene amplification.

Step	Time	Temperature (°C)
1. Initial hold	2 min	98
2. Denaturing	30 sec	98
3. Annealing	30 sec	72
4. Extension	1 min kbp^{-1}	72
5. Denaturing	30 sec	95
6. Annealing	30 sec	68
7. Extension	1 min kbp^{-1}	72
8. Final extension	10 min	72

Steps 2–4 were repeated for 5 cycles, then steps 5–7 were repeated for 30 cycles.

Table 4.3: PCR protocol for gene amplification.

Agarose gel electrophoresis

DNA samples were analysed or purified by horizontal agarose gel electrophoresis. Agarose gels were typically made by melting 1–1.5% agarose in TAE buffer (1x) under microwave irradiation. SYBR Safe DNA staining agent stock solution in DMSO (Invitrogen) was added to a final 1/10000 v/v concentration just before casting the gel into an H2-SET gel tray (Anachem) to a depth of 0.7 cm. The gel was allowed to solidify at ambient temperature after insertion of a comb at one end of the tray to allow for sample loading directly into the gel. Samples were loaded into the wells created by the comb with 6x loading buffer (5:1 dilution). The well size was varied from 5 μL for analytical electrophoresis, to 25 μL for preparative electrophoresis. The gels were run in TAE buffer (1x) at

a constant voltage (80 V) at room temperature using a Bio-Rad microprocessor Model 200/2.0 until suitable fragment separation was achieved. Analysis of the DNA fragment size was by reference to a GeneRuler® (Fermentas) 1 kb DNA ladder. DNA was visualised using UV irradiation in a Gel Logic 200 imaging system (Kodak) and 535 nm emission filter. For purification purpose, the bands corresponding to the fragment of interest were cut out using a sharp blade.

6x DNA loading buffer	
Bromophenol blue	0.25% (w/v)
SDS	1% (w/v)
Glycerol	30% (v/v)

50x TAE buffer (The pH was adjusted to pH 8.3 using conc. HCl or 5 M NaOH)	
Tris	2 M
Acetic acid (glacial)	1 M
EDTA	50 mM

PCR product purification

PCR mixture were purified using the following protocol: the PCR reaction mixtures were separated by electrophoresis on a 1.2% Agarose in TAE buffer and the bands corresponding to the fragment of desired length, as judged by UV irradiation, were cut out using a sharp blade. The DNA fragments were then extracted from the gel and purified using a gel extraction kit (QIAgen) according to the supplier instructions.

DNA restriction

The restriction enzymes employed in this work were available commercially and used according to the supplier instructions. They were used either to prepare vectors and PCR-amplified DNA for ligation, or to check for the presence of the DNA insert by performing an analytical restriction digest. If identical restriction sites are used at either end of a DNA sequence to be inserted into a vector, ligation may occur in two orientations, only one of which produces the construct of interest. To avoid this possibility, two different restriction sites were used at the two ends of the inserts.

Gene ligation

Doubly-digested vectors and PCR amplified inserts were mixed in a pre-chilled 50 μL Eppendorf tube in quantities according to the following formula adapted from manufacturer's documentation for LigaFast™ Rapid DNA Ligation System (Promega):

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng of insert}$$

Typically, 50 ng vector were used with a three-fold molar excess of insert over vector. The vector-insert mixture was incubated in a water bath at 45 °C for 5 min, mixed again and then chilled on ice for 5 min. The reaction was diluted with sterile water before adding T4 DNA ligase and 10x reaction buffer (New England Biolabs) as described in table 4.4. The reaction mixture was mixed again and left in an open ice bucket overnight. 1 μL of each ligation reaction was transformed into 50 μL *E. coli* XL10-Gold® cells (Stratagene) on the following day.

Reagent	Quantity
Vector	50 ng
Insert	see formula above
T4 ligase	200 U
10x reaction buffer	1 μL
Milli-Q H_2O	to complete to 10 μL

Table 4.4: Ligation mix.

Plasmid transformation

Plasmid DNA from miniprep purification was transformed into *E. coli* XL1 Blue (plasmid production), BL21 DE3, BL21-CodonPlus (DE3)-RP or Rosetta 2 (protein expression) competent cells using a heat-shock protocol. 2 μL of 100 ng μL plasmid solution from miniprep was added to 20 μL of *E. coli* XL1 Blue cells thawed on ice and the sample left on ice for 20 min. The sample was incubated at 42 °C for 45 sec, left on ice again for 2 min, before adding 0.7 mL of preheated (37 °C) SOC medium. After 30 min incubation at 37 °C, 200 μL of the sample was spread on an agar plate containing 35 μg mL^{-1} of the appropriate antibiotic.

Ligation products were transformed into ultracompetent *E. coli* XL10 Gold® cells using the following protocol: 2 μL of ligation mixture was added to 20 μL of *E. coli* XL10 Gold® ultracompetent cells, followed by 1 μL of β-mercaptoethanol, and the sample left on ice for 20 min. The sample was incubated at 42 °C for 45 sec, left on ice again for 2 min, before adding 0.7 mL of preheated (37 °C) SOC medium. After 30 min incubation at 37 °C, 200 μL of the sample was spread on an agar plate containing 35 μg mL^{-1} of the appropriate antibiotic.

Mutagenesis PCR products were transformed directly into ultracompetent *E. coli* XL10 Gold® cells using the following protocol: the PCR mix was subjected to digestion of methylated template DNA for 90 min at 37 °C using 1 μL of DpnI enzyme (Stratagene). 2 μL of this mixture was then added to 20 μL of *E. coli* XL10 Gold ultracompetent cells, followed the transformation protocol described above for ligation products.

Purification of plasmid DNA

Plasmid DNA was typically purified from 5 mL of bacterial cultures grown overnight using a QIAprep Spin Miniprep Kit (QIAGEN) according to the manufacturer's instructions. Small quantities of plasmid DNA for cloning and sequencing may be isolated from *E. coli* cells using the alkaline lysis method as described by Sambrook *et al.* [20], which exploits the different sizes and topologies of genomic and plasmid DNA. Under mild alkaline conditions, the hydrogen bonds between the two DNA strands are broken, thereby separating the two strands. In the case of genomic DNA, this leads to precipitation of the DNA in complex with cellular proteins. The two strands of comparatively small plasmid DNA, however, stay in solution in an interlinked manner due to the closed, circular nature of the plasmid and reassociate on being transferred to a solution at neutral pH. Lysing bacterial cells in an alkaline medium is therefore an easy and efficient method to purify plasmid DNA.

Sequencing

Gene sequencing was outsourced to the SourceBioscience sequencing facility (Oxford, UK). All constructs in pET24-a(+) and pET28-a(+) plasmids were sequenced using primers complementary to

the T7 promoter and T7 terminator sequences.
The T7 promoter primer sequence (20 mer) was as follow : `TAATACGACTCACTATAGGG`.
The T7 terminator primer sequence (20 mer) was as follow: `CTAGTTATTGCTCAGCGGTG`.

Growth media

All growth media were prepared according to Sambrook *et al.* [20] using Milli-Q water and sterilised by autoclaving at 121 °C for 20 min (TouchClave® II autoclave; LTE Scientific Ltd.). The pH of all growth media was adjusted to 7.0 using conc. HCl or 5.0 M NaOH prior to sterilisation.

2TY medium (1 L)

Tryptone	16 g
Yeast extract	10 g
Sodium chloride	5 g

Luria-Bertani (LB) medium (1 L)

Tryptone	10 g
Yeast extract	5 g
Sodium chloride	10 g

SOC medium (0.5 L)

Tryptone	10 g
Yeast extract	2.5 g
Sodium chloride	5 mL of 1 M solution (10 mM final conc.)
Potassium chloride	1.25 mL of a 1 M solution (2 mM final conc.)
Magnesium chloride	5 mL of a 1 M solution (5 mM final conc.)

The pH was adjusted to 7.0 and the solution autoclaved as described above.
10 mL of a sterile 1 M glucose solution were added before use.

Agar plates

For the preparation of agar plates, Bacto™ Agar (Becton Dickinson & Co.) was added to the medium to a final concentration of 15 g L^{-1} before autoclaving. Sterilised medium was allowed to cool to 55 °C, mixed with the appropriate antibiotic stock solution and poured into Petri dishes (Sarstedt Ltd.) in a laminar flow microbiological Safety Cabinet Class II (Envair Ltd.). Plates were stored at 4 °C for a maximum of one month prior to use.

4.5.2 Protein Expression and Purification

Protein expression Trials

Expression trials were typically carried out in 600 mL conical flasks containing 100 mL of sterile 2TY growth medium. The antibiotic corresponding to the resistance gene present in the expression plasmid was added to a final concentration of 3.0–5.0 mg L^{-1} from a 3.0–5.0 mg mL^{-1} sterile stock in water (ampicillin) or ethanol (kanamycin). The flasks were inoculated using either individual colonies from Agar plates of fresh transformation after overnight incubation at 37 °C or directly from glycerol stocks and incubated at 37 °C in a shaker (180 rpm) until the end of the linear growth curve was reached, as judged by optical density ($\cong$ 0.6). The protein expression was induced by addition of

various IPTG concentration (0 mM, 0.2 mM, 0.5 mM, 1.0 mM) from a sterile 1M stock in water. The culture were then incubated at various temperature (15 °C, 28 °C and 37 °C) for various time period (4 h, 8 h, 16 h) and the cells pelleted by centrifugation. cell pellets were resuspended in 5 mL lysis buffer in 50 mL polycarbonate centrifuge tubes and lysed on ice by sonication (6 x 30 seconds sonication with 30 seconds rest in between) using an MSE Soniprep 150 sonicator (Sanyo) at 10 μm amplitude. After centrifugation in an Avanti^{TM} J-25 centrifuge (Beckman Coulter, Inc.; JA-25.50 rotor, 22000 rpm, 25 min, 4 °C), the resulting supernatant was treated as 'soluble' protein extract. The resulting pellet was resuspended in 5 mL lysis buffer as before and treated as 'insoluble' protein fraction. Total protein concentrations were measured as described below using a NanoDrop ND-1000 spectrometer and fractions containing equal quantities of total protein analysed by SDS-PAGE.

Large Scale Protein Expression

Large scale protein expression cultures were carried out in 2 L conical flasks containing 600 mL of 2TY growth medium. The antibiotic corresponding to the resistance gene present in the expression plasmid was added to a final concentration of 3.0–5.0 mg L^{-1} from a 3.0–5.0 mg mL^{-1} sterile stock in water (ampicillin) or ethanol (kanamycin). The flasks were inoculated with 6 mL of overnight culture and incubated at 37 °C in a shaker (180 rpm) until the end of the linear growth curve was reached, as judged by optical density ($\cong$ 0.6). The protein expression was induced by addition of IPTG to the optimal final concentration from a sterile 1M stock in water. After the optimal expression period at the optimal temperature, as judged from expression trials, the cultures were pelleted by centrifugation on an Avanti instrument using a JA-10 rotor at 8 krpm for 5 min and the pellets stored at -80 °C until lysis and protein purification. Typically, each flask would produce 4–6 g of bacterial pellet (6–10 g L^{-1}).

The pET-24a(+) plasmid containing the gene encoding for ThnE V153V Δ2–46 between NdeI and EcoRI restriction sites (generously provided by Dr Refaat B. Hamed) was transformed into *E. coli* BL21-CodonPlus (DE3)-RP competent cells (Stratagene). Colonies were grown in 5 mL 2TY medium and the overnight cultures used to inoculate 2TY medium containing 50 μg L^{-1} kanamycin and 50 μg L^{-1} chloramphenicol. Cells were grown in 600 mL 2TY medium in shaken flasks at 37 °C until the media had an OD $\cong$ 0.6 and the temperature was reduced to 15 °C. IPTG (0.5 mM final concentration) was added and cells were allowed to grow for a further 16 h before harvesting by centrifugation.

Bacterial Cell Lysis

The pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5) and protease inhibition cocktail (Roche) and DNAse (Roche) were added. Bacterial cell lysis was performed by sonication (3 sec pulses every 3 sec, for 5–10 min) at 0 °C and the cell membrane debris were removed by centrifugation on an Avanti instrument using a JA-25.50 rotor at 18 krpm for 45 min at 4 °C. The clear lysate was sterile-filtered using a 0.2 μm pore size filter and loaded onto the chromatography column.

Protein Purification

Proteins were purified using chromatographic methods, including affinity, ion exchange, hydrophobic interaction and size exclusion (gel filtration) chromatography as described below. All protein purification protocols were carried out at 4 °C using Äkta FPLC^{TM} systems (GE Healthcare). Buffers were freshly prepared using Milli-Q water and filtered through a 0.2 μm polyamide filter (Millipore U.K. Ltd.) before use. Samples for size exclusion chromatography up to 2 mL were

loaded using a 2 mL loop, larger samples for other methodologies were loaded using the FPLC pump. Purification procedures were followed by monitoring absorbance at λ = 280 nm with the UPC-900 monitor on the FPLC and by SDS-PAGE. All columns were cleaned after each use according to the manufacturers' instructions. All other columns were stored in 20% ethanol.

Anion exchange chromatography

The cell lysate of ThnE V153V Δ2–46 recombinant expression cultures was applied on a 55 mL Q-Sepharose FF column previously washed and equilibrated in 50 mM Tris-HCl (pH 7.5). The filtered cell lysate was then loaded onto the column though the pumps. The column initially was washed with 5 column volumes of 50 mM Tris-HCl (pH 7.5) to remove the unbound protein. A 0–0.6 M NaCl in 50 mM Tris-HCl (pH 7.5) gradient was then run over 10 column volumes. ThnE V153V Δ2–46 protein eluted at $\approx$ 0.30 M NaCl. The presence of the protein of interest was determined by SDS-PAGE analysis and the protein concentration was measured after pooling the various fractions.

Hydrophobic interaction chromatography

The fractions containing ThnE V153V Δ2–46 were pooled and mixed with an equal volume of 2 M ammonium sulphate in 50 mM Tris-HCl (pH 7.5) before being loaded onto a 75 mL Phenyl-Sepharose HP column that had been washed and pre-equilibrated with 1 M ammonium sulphate in 50 mM Tris-HCl (pH 7.5). Unbound protein was then removed by washing the column with another 5 volumes of ammonium sulphate in 50 mM Tris-HCl (pH 7.5) before a gradient was run from 1 M ammonium sulphate in 50 mM Tris-HCl (pH 7.5) to 50 mM Tris-HCl (pH 7.5) over 10 column volumes. ThnE V153V Δ2–46 eluted at $\approx$ 0.3 M ammonium sulphate. The fractions containing ThnE V153V Δ2–46 and with purity $\geq$ 95%, as judged by SDS-PAGE analysis, were pooled and concentrated ($\leq$ 2 mL) as described below.

Size exclusion chromatography

Size exclusion chromatography, also commonly called gel filtration chromatography, was used to further purify the proteins after previous chromatography steps. For crystallography purposes, ThnE V153V Δ2–46 was purified using size exclusion chromatography.

The protein sample ($\leq$ 2 mL) was applied on a 300 mL Superdex^{TM} S200 column, equilibrated with 100 mM Tris-HCl (pH 7.5) containing 100 mM NaCl. The proteins were eluted using 1 column volume of 100 mM Tris-HCl (pH 7.5) containing 100 mM NaCl. The fractions containing the proteins of interest and with purity $\geq$ 95%, as judged by SDS-PAGE analysis, were pooled and concentrated as described below.

Concentration of protein solutions

Protein solutions were concentrated using Amicon® Ultra-4 or Ultra-15 Ultracel® PL ultrafiltration devices (Millipore U.K. Ltd.) with 5000, 10000 or 30000 Da nominal molecular weight limit-membranes in an Allegra^{TM} 21R centrifuge (Beckman Coulter, Inc.; S4180 rotor, 4000 rpm, 4 °C).

Protein concentration measurements

The concentration of protein solutions was determined using a NanoDrop® ND-1000 spectrometer using the following considerations. The molar absorption coefficient of a peptide or protein is related to its tryptophane (W), tyrosine (Y) and cysteine (C) amino acid composition. At

280 nm, this value is approximated by the weighted sum of the 280 nm molar absorption coefficients of these three constituent amino acids [21], as described in the following equation ϵ = (n(W) $\times$ 5690) + (n(Y) $\times$ 1280) + (n(C) $\times$ 120) where n is the number of each residue and the stated values are the amino acid molar absorption at 280 nm.

SDS-PAGE

The SDS-PAGE protocol was first described by Laemmli [22]. Before being loaded onto the gel, protein samples were boiled at 95 °C for 5 min in SDS-PAGE Sample loading buffer. Polyacrylamide gels (5% stacking gel, 12% resolving gel) were prepared using the Mini Protean II system (Bio-Rad Laboratories Ltd.) and electrophoresis was typically conducted at 170 V. After the end of the run, the gels were stained in a Coomassie solution for 15 min, followed by the time necessary for removal of the stain from the gel in a destain solution. For the composition of these solutions, see below.

Reagent	3% Stacking gel (mL)	12.5% Resolving gel (mL)
H_2O	3.1	1.9
30% acrylamide mix	0.5	3.8
1.5 M Tris, pH 8.8	-	3.0
1.0 M Tris, pH 6.8	1.25	-
10% SDS	0.05	0.12
10% APS	0.05	0.2
TEMED	0.005	0.02

Polyacrylamide gel compositions used in this work. Volumes of reagents given are those required for two gels.

Table 4.5: SDS PAGE Polyacrylamide gel composition

3x SDS-PAGE Sample loading buffer	
β-mercaptoethanol	16%
Bromophenol blue	0.06%
Glycerol	30%
SDS	20%
Tris, pH 6.8	0.24 M

10x SDS-PAGE Running buffer	
Glycine	1.92 M
Tris	0.25 M
SDS	1%

SDS-PAGE Staining solution	
Coomassie® Brilliant Blue R	0.1%
Acetic acid	10%
Ethanol	20%

SDS-PAGE Destaining solution	
Acetic acid	10%
Isopropanol	20%

4.5.3 CMPS Assays Using L-GHP and Analogues

CMPS assays were carried out using ThnE prepared as described above and ThnE variants, CarB wildtype and CarB variants prepared by J. R. Castellanos-Gomez as described previously [5]. All the stock solution of CMPS enzymes were stored in 50 mM Tris-HCl pH 7.5 at a concentration of about 20 mg mL^{-1}.

Deprotection of amino acid aldehydes for CMPS assay

Solutions of CMPS substrate and substrate analogues were obtained by deprotection of the appropriate synthetic amino acid semialdehydes **149**, **244**, **150**, **245**, **172**, **173**, **182**, **186**, **187**, **203**, **204** and **206**. The protected L-GHP or analogue (4–5 mg) was dissolved in 10% aqueous formic acid (0.5 mL) to give a 15 mM final concentration solution, and the reaction mixture stirred at 60 °C for 1h as previously described for other CMPS substrates [14]. After cooling to room temperature, these 15 mM solutions of substrate were used directly for CMPS assay purpose without purification or stored frozen at -80 °C.

Standard CMPS Assay

The analytical CMPS assays were carried out at pH 7.9 in a final reaction volume of 70 μL in a 0.5 ml Eppendorf tube by sequential addition of the components as follow: the amino acid aldehyde (75 nmol, 5 μL of 15 mM solution) was mixed in 600 mM Tris-HCl pH 9.0 (35 μL) together with malonyl-CoA (80 nM, 8 μL of 10 mM solution) and the mixture diluted with water (18.5 μL) before CMPS was added (0.3 nmol, 3 μL of 1.5 mM solution). The assay mixture was kept at 37 °C for 5–30 min depending on the experiment, then the reaction quenched with a cold 250 μM *p*-aminosalicylic acid solution in acetonitrile (70 μL). The samples were kept on ice for 15 min, and then centrifuged at 13 krpm for 15 min at 4 °C to remove the precipitated protein. The 100 μL supernatant was transfered in a vial for analysis.

The preparative CMPS assays and spectroscopic identification of the product, the assay way scaled up 10 times, to a final volume of 700 μL. The products were purified and analysed by Dr Refaat B. Hamed as described in Hamed *et al.* [13].

CMPS Assay LC-MS Analysis

Analytical analysis of the formation of *t*-CMP analogues was monitored by LC-MS using a Primesep 100 column (Sielc, 250 mm x 4.6 mm, 10 μm particle size) on a Waters 1525m Binary HPLC Pump system with a Waters 2777 Sample Manager coupled to a Micromass® Quattro micro™ API mass spectrometer using the positive electrospray ionisation mode. Eluent A was 0.1% formic acid in water (v/v), eluent B was 0.1% formic acid in acetonitrile (v/v). The samples (100 μL) were eluted using a flow rate of 1 mL min^{-1} with 5% eluent over 10 min, followed by a gradient of 5–70% B over 18 min, and then 70–95% B over 5 min The column was washed with 95% B for 10 min and the column was re-equilibrated at 5% B for 20 min before the next sample was injected.

The resulting chromatograms were analysed using the MassLynx™ v4.0 software (Waters).

Chapter 5

Stable Coenzyme A Analogues

5.1 Introduction

5.1.1 Coenzyme A - Biological Roles and Biosynthesis

Coenzyme A **133** (CoASH, Figure 5.1 A) is ubiquitous in living organisms. CoASH **133** is composed of a 3′,5′-diphosphoadenosine group connected through a pyrophosphate linkage to a phosphopantetheine moiety. CoASH **133** plays a central role in metabolism and is thought to be a cofactor in over 100 metabolic processes involving 4% of all proteins found in any given organism [1].

Coenzyme A Diverse Biological Roles

Many of the biological roles of CoASH **133** are mediated by the formation of a thioester linkage with metabolites and xenobiotics bearing a carboxylic acid function. Specific or promiscuous ligases and transferases catalyse the reversible formation of acyl-CoA species, *e.g.* malonyl-CoA **130**, acetyl-CoA **127** or fatty acyl-CoA species with various chain lengths and degrees of desaturation (Figure 5.1 A). These CoASH conjugates can then be used by other enzymes as substrates or cofactors, or be transported into specific organelles. Because CoASH **133** and the corresponding thioester derivatives are crucial components of the cell machinery, enzymes that produce acyl-CoA derivatives, transport them across membranes or use them as substrate and cofactors are of particular interest in biology and pharmacology. These enzymes participate in energy homeostasis when involved in metabolic processes such as the tricarboxylic acid (TCA) cycle and fatty acid (FA) synthesis and β-oxidation.

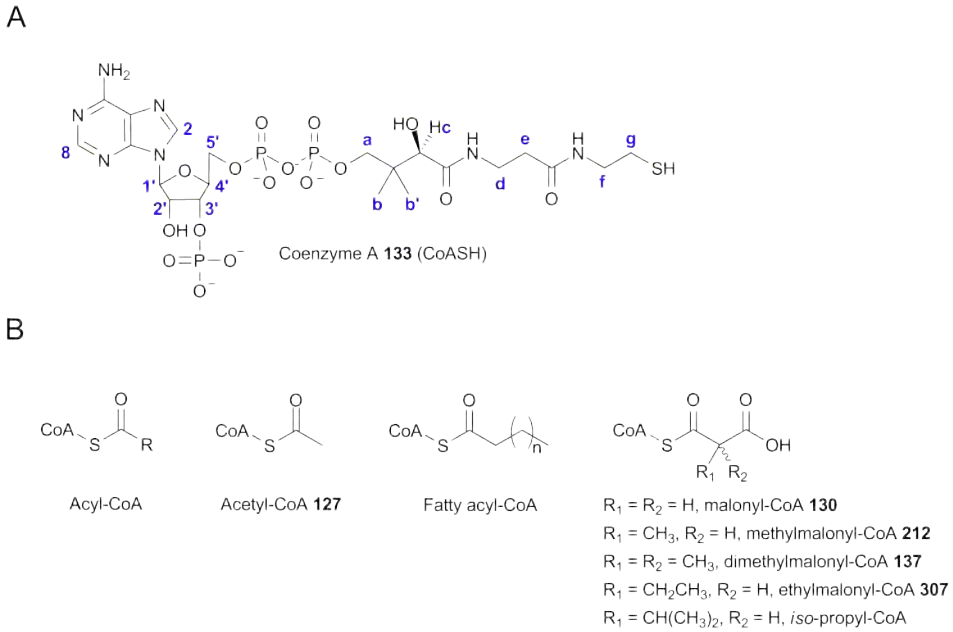
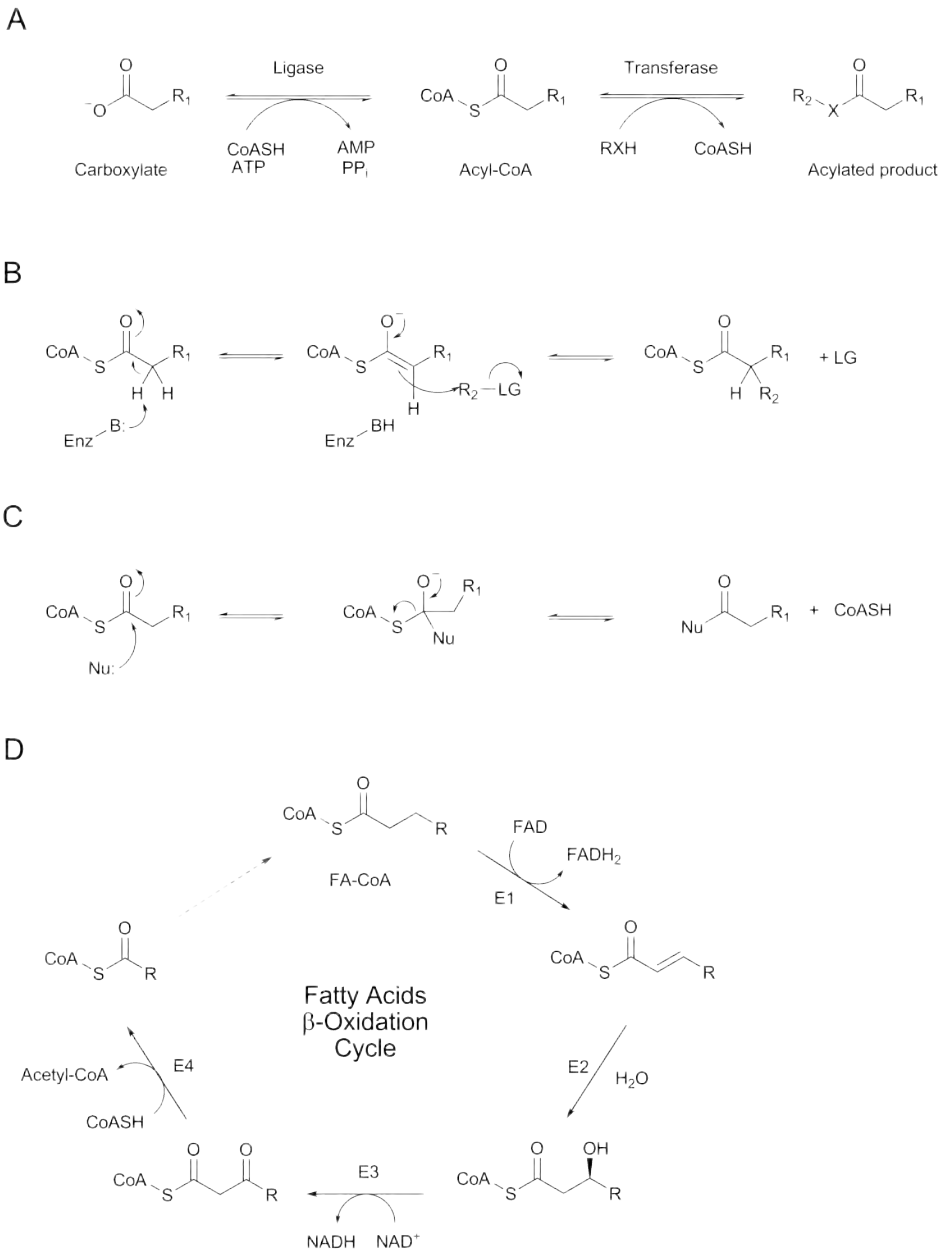


Figure 5.1: Structures of coenzyme A and acyl-coenzyme A derivatives. (A) Chemical structure of CoASH **133**. Proton labeling was used by Patel and Walt [3]; (B) Representative examples of acyl-CoA derivatives.

Another example is the regulation of gene expression *via* acetylation of histone tails [2].

Reactivity of acyl-CoA derivatives is based on the properties of the thioester functionality. In a thioester bond, lone pairs on the sulphur share much less electron density with the carbonyl group than the oxygen in the corresponding ester or carboxylate functionalities. This is due to difference in overlap between the sp^3 orbitals of the oxygen and sulphur lone pair with the carbonyl π^* orbitals. Because the sulphur lone pair electrons are not shared with the carbonyl, deprotonation of the carbon adjacent to the carbonyl enables the formation of a carbanion stabilised by conjugation. For this reason, the protons at the α-position of the carbonyl group are more acidic in the case of the thioester than the ester or carboxylate. Enzyme-catalysed deprotonation of an acyl-CoA substrate by a basic amino acid residue can occur at physiological pH to form the corresponding enolate allowing, for example, aldol reactions to occur. This enolate can react with an electrophile, typically an aldehyde, ketone or another thioester (Scheme 5.1 B). The thioester bond is also reactive towards nucleophilic attack, for example by an alcohol or amine, followed by CoASH release (Scheme 5.1 C). In conclusion, acyl-CoA derivatives containing at least one proton adjacent to the carbonyl group of the thioester can act both as nucleophiles and electrophiles.



Scheme 5.1: Examples of reactions carried on acyl-coenzyme A substrates. (A) Specific or promiscuous ligases or transferases catalyse the two step transfer of a free carboxylate onto a nucleophilic substrate *via* the acyl-CoA intermediate; (B) Deprotonation at the α position of an acyl-CoA cosubstrate yields the corresponding enolate which can then react with an electrophile; Enz, enzyme; LG, leaving group; (C) Nucleophilic attack on the carbonyl group results in the formation of a tetrahedral transition state, followed by elimination of CoASH; Nu, nucleophile; (D) Schematic representation of the fatty acyl-CoA β-oxidation cycle. E1, acyl CoA dehydrogenase; E2, enoyl CoA hydratase; E3, L-β-hydroxyacyl CoA dehydrogenase; E4, β-ketothiolase.

The activation of a carboxylate as a thioester can also be achieved following the enzymatic transfer of the phosphopantotheine moiety onto a protein serine residue. This protein modification is catalysed by phosphopantotheine transferases (PPTase) that recognise specific amino acid sequences of apoacyl-carrier protein (ACP) domains, yielding a holoACP. The phosphopantotheinylation of ACP was first reported by Vagelos and coworkers [4]. An example of such PPTase mediated phosphopantotheinylation is described for the carboxylic acid reductase (CAR) in Section 7.1.3 (Scheme 7.4). However, other examples include ACP domains found in the assembly line of a major class of metabolites, the polyketides (PK) and fatty acids (FA) on the one hand, and in the peptidyl-carrier protein (PCP) domains found in the assembly line of the nonribosomal peptides (NRP), on the other.

CoASH **133** can also be used as a carrier to increase the affinity or specificity of a protein for its ligand. Organelle-specific acyltransferase activity, coupled to active transport of acyl-CoA species across membranes enables control of metabolite concentrations in sub-cellular compartments. As described in Chapter 1, fatty acyl-CoA species play a crucial role in energy homeostasis; the transport of fatty acids across the mitochondrial membrane is mediated by fatty acyl-carnitine transporters. A fatty acyl-CoA β-oxidation cycle occurs in the mitochondrion, where fatty acid substrates undergo four consecutive enzymatic steps (Scheme 5.1 D) [5]. The reaction of acyl-CoA dehydrogenase yields the α,β-unsaturated CoA ester. Addition of water across the double bond is catalysed by crotonase (or enoyl-CoA hydratase) yielding a β-hydroxyester. Even though it is unmodified in these two steps, the thioester is believed to facilitate both the dehydrogenase and crotonase reactions by stabilising the partial negative charge on the α-carbon atom. The third step does not involve the thioester, but the CoA moiety is crucial for recognition. Nucleophilic attack by a second CoASH molecule and elimination of acetyl-CoA **127** completes the cycle.

Overall, the significance of CoASH **133** explains the extensive interest in developing an understanding of the enzymes that use CoA as a cofactor or a substrate.

Coenzyme A Biosynthesis

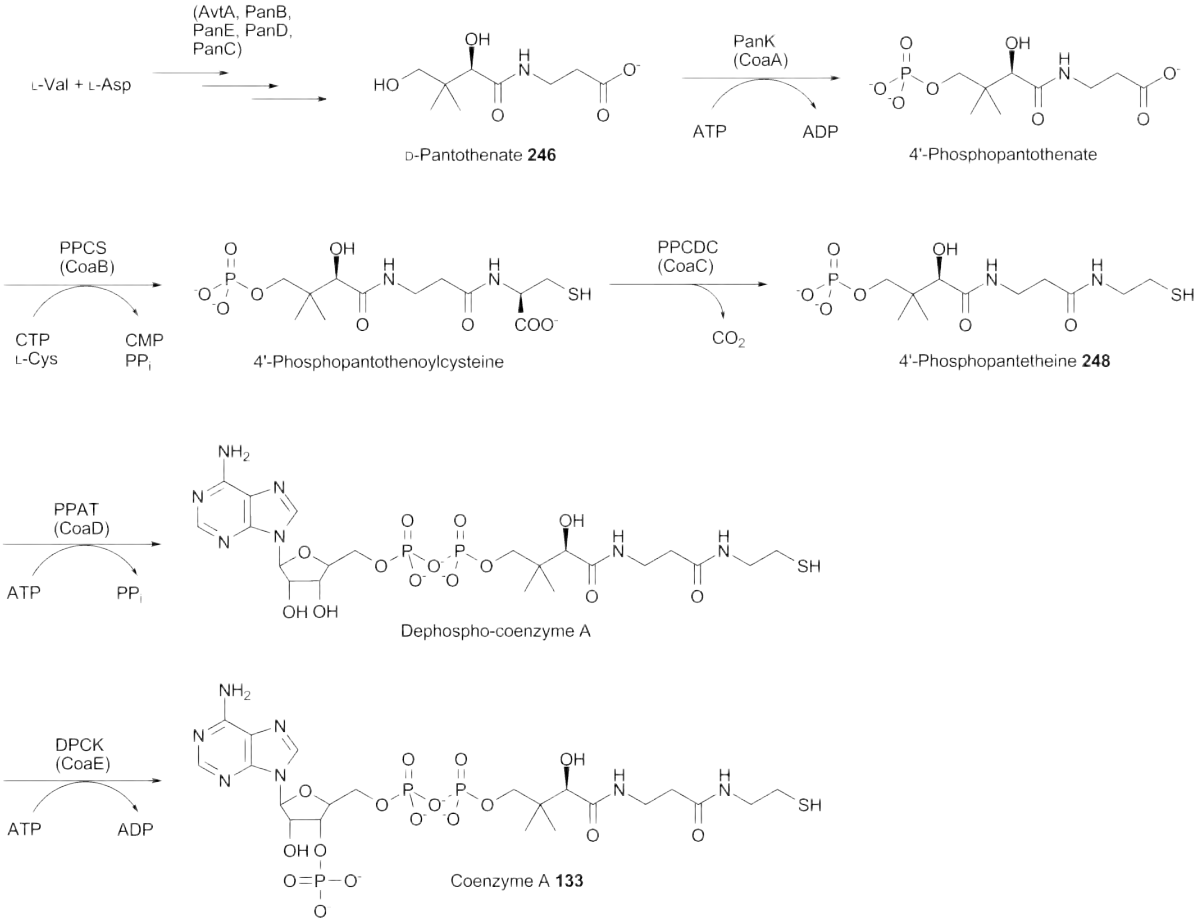
The biosynthesis of CoASH **133** is conserved across all living organisms and proceeds in five enzymatic steps, starting from D-pantothenate **246** (*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanine also commonly called vitamin B5), cysteine and ATP (Scheme 5.2) [6]. D-Pantothenate **246** itself is synthesised *de novo* from L-valine and L-aspartate in bacteria and plants [7, 8], but is formally a vitamin in mammals because they lack the D-pantothenate **246** biosynthetic enzymes. Even though the individual enzymatic steps have been known since the 1950s [9, 10], the enzymes responsible for the biosynthesis of CoASH **133** and their respective genes were not identified until much later. One reason is the absence of sequence homology between eukaryotic and prokaryotic enzymes, suggesting a convergent evolution of the CoASH biosynthetic pathways. In addition, the bacterial genes involved in CoASH biosynthesis are not organised in a cluster, unlike other biosynthetic pathways such as those coding for carbapenem β-lactam antibiotics (Sections 2.1.2 and 2.1.3). The identification of genes responsible for the biosynthesis of CoASH **133** and a detailed understanding of the biochemical transformations catalysed by the corresponding enzymes led to a renewed interest in CoASH **133** biosynthesis, *e.g.* as an antimicrobial target [11], and for its role in heritable neurodegenerative disorders [12].

5.1.2 Chemical Synthesis of Coenzyme A and Coenzyme A Analogues

Access to pure CoASH **133** and acyl-CoA derivatives is crucial for the characterisation of CoA-utilising enzymes, in the context of mechanistic and structural investigations *in vitro*, but also for the elucidation of their function *in vivo*. Acyl-CoA and analogues can be divided in two groups; those which can be accessed from CoASH **133** itself, and those which cannot. The application of CoA derivatives as mechanistic probes has been reviewed by Mishra and Drueckhammer in 2000 [13], and recent examples are summarised in this section with a specific focus on crystallographic studies where protein structures were obtained in complex with unnatural acyl-CoA analogues.

Analogues Accessible From Coenzyme A

Analogues which can be obtained from CoASH **133** are typically thioesters or thioethers, with



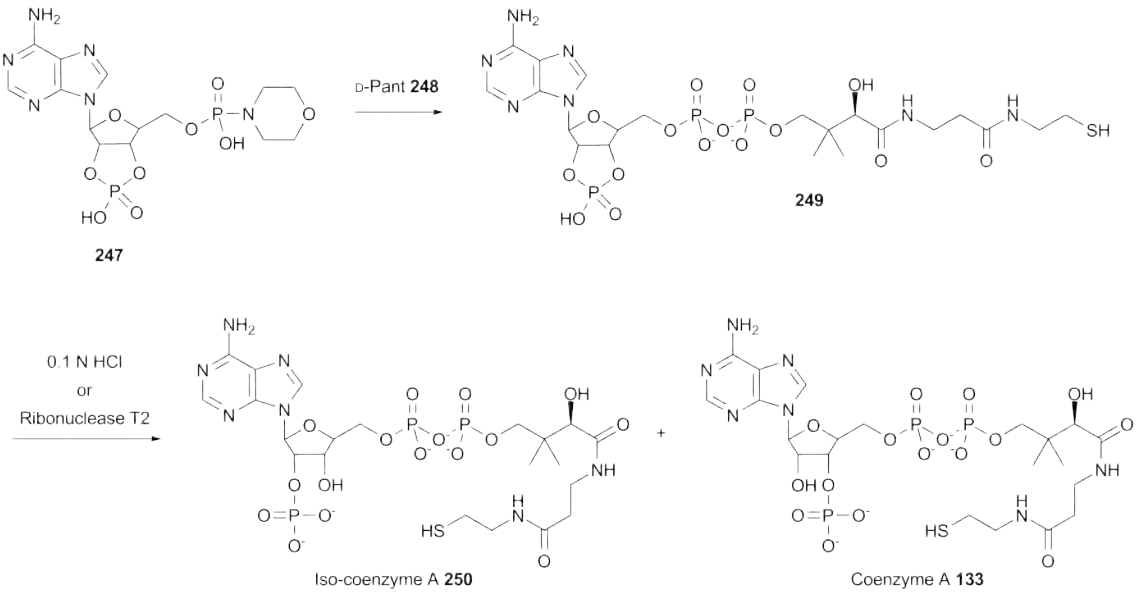
Scheme 5.2: Coenzyme A biosynthetic pathway. The genes coding for the first five steps are only found in bacteria and plants, while the last five steps are biochemically conserved across all living organisms (prokaryotic enzymes are given in brackets). Mammals obtain D-pantothenate exclusively from the diet. Not shown in this pathway is the 4′-phosphorylation of D-pantetheine **215** by CoaA, allowing bypass of the first biosynthetic steps. AvtA, valine–pyruvate aminotransferase; PanB, 3-methyl-2-oxobutanoate hydroxymethyltransferase; PanE, 2-dehydropantoate reductase; PanD, aspartate α-decarboxylase; PanC, pantoate-β-alanine ligase; PanK (CoaA), pantothenate kinase; PPCS (CoaB), phosphopantothenoylcysteine synthetase; PPCDC (CoaC), phosphopantothenoylcysteine decarboxylase; PPAT (CoaD), phosphopantetheine adenylyltransferase; DPCK (CoaE), dephospho-CoA kinase.

the exception of desulpho-coenzyme A where the thiol group has been reduced to the corresponding methyl group [14]. Commercial preparations of CoASH **133** are obtained from bacterial sources. The terminal thiol group of CoASH **133** is nucleophilic and thioesters can therefore be accessed by chemical or enzymatic acylation methodologies using natural or unnatural carboxylic acids [13]. Thioether derivatives of CoASH can be obtained *via* nucleophilic substitution to an alkyl halide, yielding acyl-CoA analogues with an inserted methylene or ketomethylene moiety (Scheme 5.4 B). Many examples of such thioester and thioether analogues have been reported [13]. Selected examples of such analogues used in the context of crystallographic studies are described in Section 5.1.5.

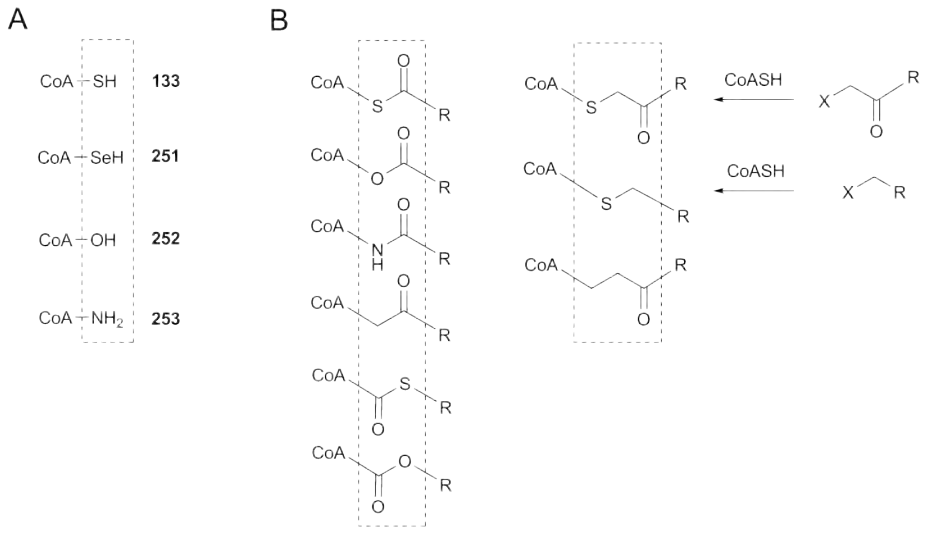
Total Synthesis of Coenzyme A and Analogues

The first successful total chemical synthesis of CoASH **133** was described by Moffatt and Khorana [15] and their procedure has since been improved by others; in particular by Michelson who introduced an enzymatic transformation as the final step [16]. The final two steps of the original chemical synthesis involved reaction of an adenosine 2′,3′-cyclic phosphate-5′-phosphoromorpholidate **247** with 4′-phospho-D-pantetheine **248** to give the intermediate **249**, followed by acidic treatment that yielded a 1:1 mixture of 2′ and 3′ ribose phosphate (iso-CoASH **250** and CoASH **133**, respectively, Scheme 5.3) [15]. Michelson observed that the opening of the 2′,3′-cyclic phosphate intermediate **249** by ribonuclease T2 (or ribonucleate 3′-adenylohydrolase) gave the 3′-phosphate exclusively, doubling the yield of CoA analogue preparations (Scheme 5.3) [16].

The total chemical synthesis and purification of CoASH **133** and its analogues remains a challenging task [17]. Nevertheless, this synthetic methodology has been used to generate a wide range of acyl-CoA analogues that cannot be accessed from CoASH **133** directly, and where the labile thioester was replaced by a more stable ester, amide or methylene linker (Figure 5.7 C). Several biochemical investigations have also been performed using synthetic CoA analogues produced by total synthesis, *e.g.* seleno(dethia)CoA **251** [18], hydroxy(dethia)CoA **252** [19], amino(dethia)CoA **253**, and a range of carba(dethia)CoA derivatives [20].



Scheme 5.3: The two last steps of CoASH total synthesis described by Moffatt and Khorana [15]. 4′-Phospho-D-pantetheine **248** was prepared in one step form D-pantetheine **248**, while adenosine 2′,3′-cyclic phosphate-5′-phosphoromorpholidate **247** was prepared from adenosine in three steps. Coupling followed by acidic hydrolysis of the 2′,3′-cyclic phosphate gave a 1:1 mixture of CoASH **133** and isoCoASH **250**. When ribonuclease T2 activity is exploited, the 2′,3′-cyclic phosphate can be hydrolysed specifically at the 2′ position to produce CoASH **133** regiospecifically.



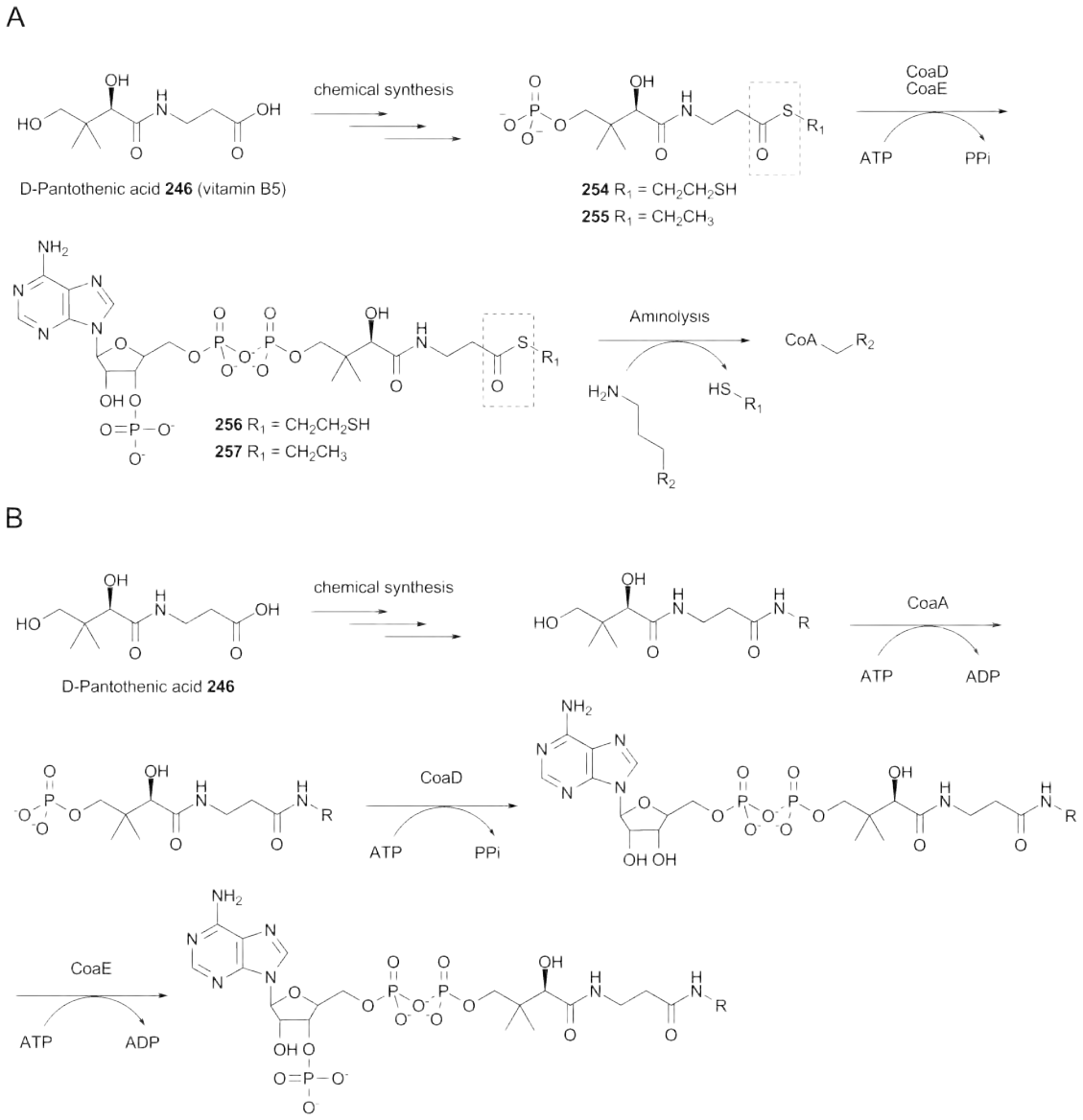
Scheme 5.4: Examples of CoA analogues. (A) Structure of CoASH **127**, CoASeH **251**, CoAOH **252** and $CoANH_2$ **253**; (B) Structure of the acyl-CoA thioester functionality together with oxa(dethia), aza(dethia) and carba(dethia) analogues; (C) Other examples of stable acyl-CoA analogues. The two thioether derivatives can be accessed from CoASH and the corresponding halide (X = Cl, Br).

5.1.3 Chemoenzymatic Synthesis of Coenzyme A and Analogues

Pioneering work by Novelli and coworkers in the late 1940s and early 1950s led to the isolation of CoASH **133** [21] and the determination of its structure [22]. These findings were followed by the identification of pantetheine as a CoASH biosynthetic precursor, and the isolation of the enzymes responsible for CoASH biosynthesis from pantetheine in pigeon liver extracts [9, 10].

Using the protocol described by Novelli and coworkers for the purification of the biosynthetic enzymes from liver extracts, Stewart and Ball reported the first chemoenzymatic synthesis of a CoA analogue, showing that *in vitro* conversion of 4′-phospho-D-hydroxy(dethia)pantetheine to hydroxy(dethia)coenzyme A **252** (CoAOH) is possible in two enzymatic steps [23]. Hydroxy(dethia)CoA **252** had been previously obtained by total synthesis by Stewart and coworkers [19]. Several analogues, including desulfo-coenzyme A [24], were subsequently obtained following this procedure using partially purified enzymes from tissue extracts. Another elegant approach reported by Martin and Drueckhammer [25] enabled the generation of CoA analogues and was inspired by the work of Stewart and Ball [23]; This methodology involved the chemoenzymatic generation of CoA analogues, carrying a thioester bond in place of the C-8 amide bond (dotted box, Scheme 5.5 A). Two 4′-phosphopantetheine analogues **254** and **255** were used as substrates in an enzyme cascade catalysis (ECC) assay with an enzyme preparation from *Brevibacterium ammoniagenes*. Two enzymes present in the assay, CoaD and CoaE, enabled the construction of the 3′,5′-diphosphoadenosine moiety of the CoA analogues **256** and **257**. Aminolysis of the C-8 thioester bond yielded the corresponding amides *in vitro* (Scheme 5.5 A); this technique has been used to generate a variety of probes. Schwartz and Drueckhammer described an analogue of the acetyltransferases tetrahedral hemiacetal transition state where the thioester group was replaced with a secondary alcohol, and this compound was tested for the inhibition of various acetyltransferases [26]. Of particular interest in the context of CMPS catalysis is another example where the aminolysis strategy was used to access the acetyl-CoA enolate analogues, as described in Section 5.2.5 (Scheme 5.17) [27].

More recently, an ECC approach to generate unnatural CoA derivatives has been described by Wright and coworkers (Scheme 5.5 B) [28]. The observation that bacterial pantothenate kinase (CoaA) was promiscuous enough to take D-pantetheine **215** as a substrate [29] in place



Scheme 5.5: Two different strategies have been described for the chemoenzymatic production of CoA analogues. Only three enzymes are required to produce CoA analogues from pantotheine analogues: CoaA, CoaD and CoaE. The reaction can be carried out in a 'one pot' fashion by addition of an excess of ATP and the three purified enzymes. (A) Strategy described by Martin and Drueckhammer [25]. Chemoenzymatic production of pantetheine thioester analogue followed by aminolysis enables the generation of CoA analogues; (B) Strategy described by Wright and coworkers [28].

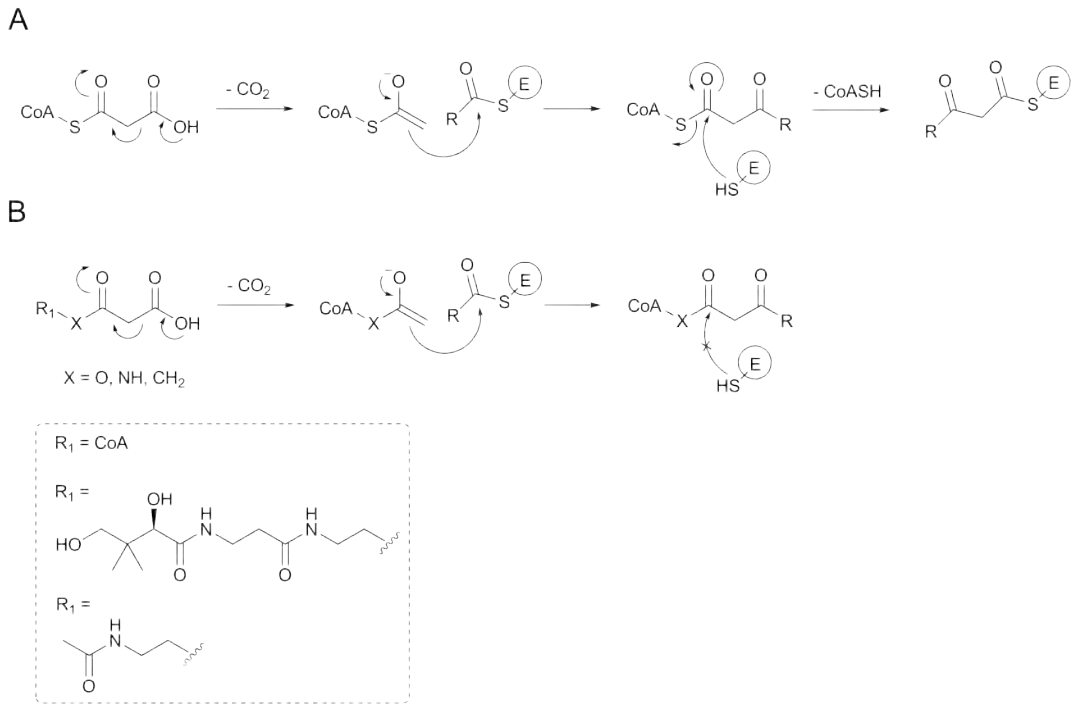
of D-pantothenate **246** suggested that analogues where the terminal thiol was replaced by other functionalities could be accepted as well. The promiscuity of the last two enzymes in the CoA biosynthetic pathway, phosphopantetheine adenylyltransferase (CoaD) and dephospho-CoA kinase (CoaE), enabled the generation of CoA analogues in a 'one-pot' fashion. A wide range of pantetheine analogue precursors were prepared synthetically and successfully tested *in vitro* as substrates for the three biosynthetic enzymes purified from recombinant expression in *E. coli* (Scheme 5.5 B). This ECC approach was applied by others to generate CoA analogues carrying reactive and bulky functionalities in a straight forward manner [30] yielding highly informative results in recent years, some of which are described in the next sections.

5.1.4 Coenzyme A Analogues as Probes

Hydroxy(dethia)CoA **252** [19], the first CoA analogue to be synthesised, has been recognised as a potential inhibitor of CoA-utilising enzymes [31]. Several other unnatural CoA analogues have since been synthesised using the different methodologies described above and tested *in vivo* or *in vitro* as non-natural substrates or inhibitors. Some examples of CoA analogues are of interest in the context of the present work and are described below.

Non-Hydrolysable Malonyl-CoA Analogues

Polyketides are produced in bacteria, fungi, plants and protists by polyketide synthases (PKSs) through repetitive condensation steps that attach malonyl-CoA or C-2 substituted malonyl-CoA units to a starter unit (Scheme 5.6 A) [32]. The diversity of natural extender units has been reviewed in Chan *et al.* [33]. The identification of the intermediates of this iterative process has proven difficult. Spencer and coworkers reported malonyl-CoA thioether analogues as the first trapping agent for the identification of intermediates in Type III PKS catalysis [34]. The ability of malonyl-CoA analogues to undergo decarboxylation and condensation was exploited to generate non-hydrolysable CoA derivatives that could be isolated and characterised. The ECC strategy described in Section 5.1.3 was used to generate malonyl-aza(dethia)CoA **258** and methylmalonyl-aza(dethia)CoA **259** from 3-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-3-oxopropanoic acid **260** (malonyl-aza(dethia)pantetheine) and 3-[2-(*N*-



Scheme 5.6: Non-hydrolysable malonyl-CoA analogues as trapping agents. (A) Polyketide formation by decarboxylative malonate condensation reaction. (B) In the presence of stable malonyl-CoA analogues (or truncated analogues), the intermediate formed cannot be loaded back onto the enzyme, enabling its identification [35]. The same strategy was applied to investigations into the mechanism of type I modular PKS [36]. E, Enzyme (ketosynthase domain of a type III PKS).

[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-2-methyl-3-oxopropanoic acid **261** (methylmalonyl-aza(dethia)pantetheine), as well as malonyl-oxo(dethia)CoA **262** and methylmalonyl-oxo(dethia)CoA **263** from 3-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethoxy]-3-oxopropanoic acid **264** (malonyl-oxa(dethia)pantetheine) and 3-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethoxy]-2-methyl-3-oxopropanoic acid **265** (methylmalonyl-oxo(dethia)pantetheine), respectively. These compounds were tested for PKS-catalysed formation of stable intermediates (Scheme 5.6 B) [35, 36].

Truncated Analogues

Truncation of the 3′,5′-diphosphoadenosine functional group of CoA enables the generation of CoA analogues requiring simplified synthetic and purification methodologies. In order to simplify the synthesis of probes for the study of CoA and acyl-CoA utilising enzymes, 4′-phosphopantetheine,

pantetheine and SNAc derivatives have been synthesised and their biological relevance investigated. Jencks and coworkers have used various truncated CoA analogues, such as pantetheine, pantothenol, *N*-acetylaletheine and SNAc in order to determine the contribution of each CoA moiety to substrate binding energies, *e.g.* for succinyl-CoA:3-oxoacid CoA transferase (EC 2.8.3.5) [37, 38]. Reports also showed that truncated acyl-CoA analogues can be accepted as substrates by FAS [39] and PKS enzymes [40].

CoA derivatives have demonstrated good inhibition of CoA-utilising enzymes *in vitro*, but their inability to cross cell membranes prevented their use *in vivo*, *e.g.* as antibacterial agents. Neutral pantetheine analogues, on the other hand, can be transported across bacterial membranes by pantothenate permease (PanF) and therefore reach their intracellular targets [41]. Strauss and Begley first demonstrated that some cell-permeable truncated CoA analogues, such as pantetheine derivatives, could be processed *in vivo* by the CoA biosynthetic enzymes described above (Section 5.1.1, Scheme 5.2) and yield the corresponding CoA derivatives [29]. Based on this observation, the toxicity of *N*-pentylpantothenamide was proposed to be due to its *in vivo* conversion to ethyl(dethia)CoA, which then acts as an inhibitor of CoA- and acetyl-CoA-utilising enzymes. This knowledge was recently applied in the precursor-mediated inhibition of aminoglycoside *N*-6′-acetyltransferase (AAC), an enzyme responsible for aminoglycoside resistance through the acetylation of kanamycin, rendering the antibiotic inactive [42]. After detailed medicinal chemical investigations, Auclair and coworkers used this *in vivo* activation of pantetheine analogue precursors to generate aminoglycoside-CoA bisubstrate inhibitors of AAC with antibacterial activity [43]. Burkart and coworkers have used cell permeable acyl-pantetheine analogues to perform *in vivo* labelling of phosphopantetheinylated proteins, including fluorescent derivatives [30, 44, 45], as well as alkynyl or aza derivatives that can be substrates for click chemistry [46].

5.1.5 Coenzyme A Analogues in Structural Biology

In their review of the literature published prior to 2000, Mishra and Drueckhammer reported that the crystal structures of about 20 CoA ester-utilising enzymes had been solved, most of them in complex with their CoA thioester substrate or an analogue thereof [13]. Wierenga and coworkers have made significant contributions to the understanding of the interactions and structural folds required for

CoA binding in various families of enzymes [47, 48]. There are now more than 550 structures of proteins using CoASH or its derivatives as a cofactor, but only 50 crystal structures (10%) have been determined in complex with a ligand. The development of synthetic and chemoenzymatic methodologies to access stable acyl-CoA analogues allowed to obtain very informative structures within the past twelve years.

Numerous protein crystal structures were solved in complex with truncated analogues based on the pantetheine scaffold because these compounds are relatively easily accessible synthetically, *e.g.* fluoroacetyl-CoA thioesterase from *S. cattleya* in complex with fluoroacetyl-carba(dethia)pantetheine (PDB ID: 3KVZ) [49]. However more structural information can be extracted from the crystal structures described in the next paragraphs; these crystal structures were solved in complex with CoA analogues carrying the 3′,5′-diphosphoadenosine group.

CoA Thioether Analogues

Most of the enzyme crystal structures solved in complex with CoA analogues were co-crystallised with compounds containing a stable thioether (Table 5.1, Entries I-XVI, Scheme 5.4 C). The crystal structures of acetyltransferases were solved in complex with bi-substrate transition-state mimics in which the amino group of the acceptor substrate was linked to acetyl-CoA. Examples include the transcriptional co-activator p300/CBP (CREBBP) co-crystallised with a lysine-CoA bi-substrate inhibitor (Table 5.1, Entry I) [54], GCN5 histone lysine acetyltransferase (HAT) with a 20-mer peptide-CoA bi-substrate analogue [55], and serotonin *N*-acetyltransferase with an *S*-acetyltryptamine-CoA bi-substrate analogue [56]. The crystal structure of the *E. coli* methylmalonyl-CoA decarboxylase (Section 2.4.1) was determined in complex with 2*S*-carboxypropyl-CoA (Figure 5.2 A and B, Table 5.2, Entry IX, PDB ID: 1EF9) [50].

Carba(dethia)CoA Analogues

Various examples of protein crystal structures were reported in complex with carba(dethia)CoA analogues where the sulphur atom is replaced by a methylene group (Scheme 5.4 B), thus keeping the CoA chain length near constant and placing the carbonyl group at the same position as in the

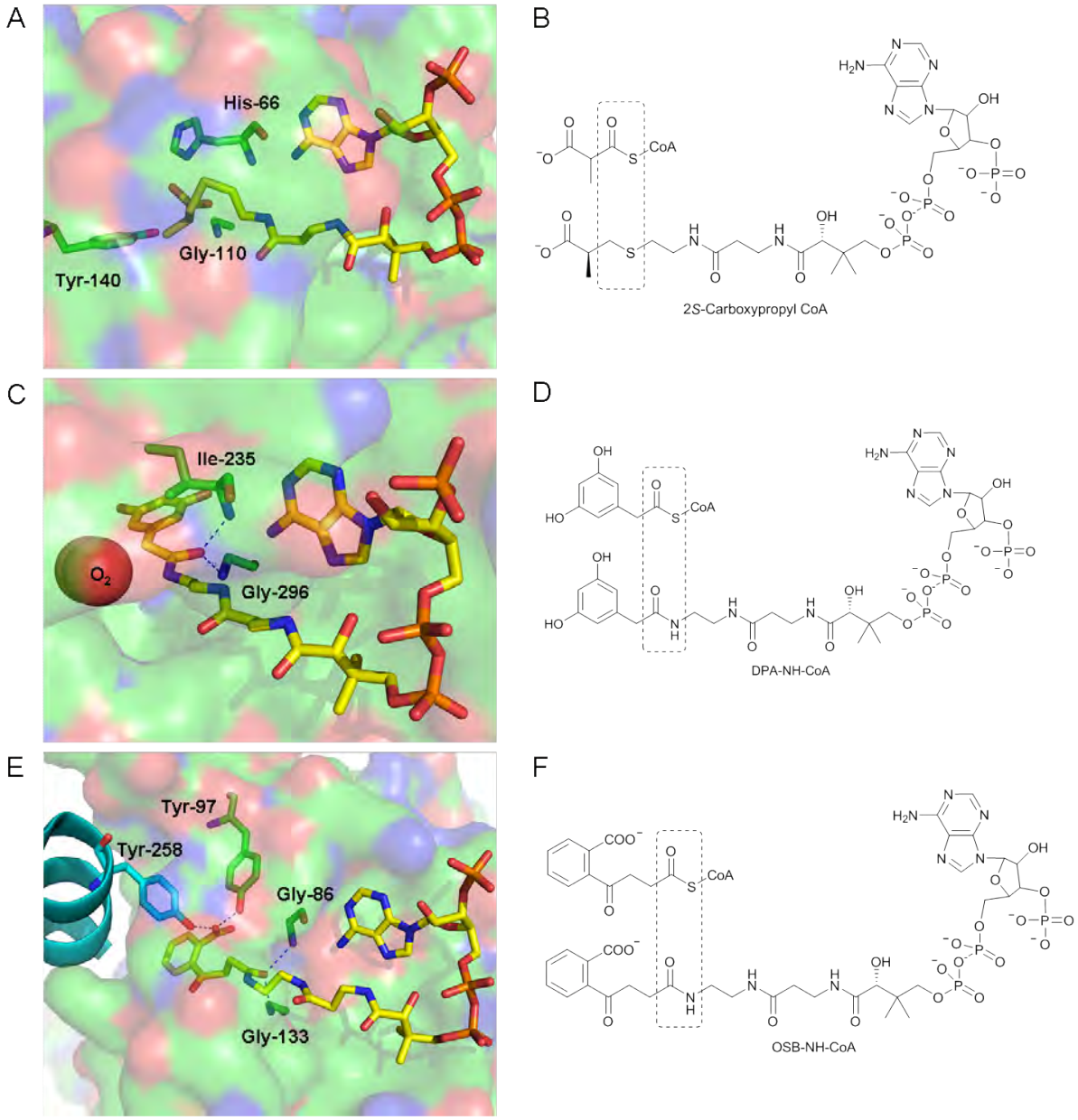


Figure 5.2: View from the active site of CS enzymes in complex with stable CoA substrate analogues. Three members of the crotonase superfamily, DpgC, MMCD and MenB, were crystallised in complex with stable acyl-CoA substrate analogues (displayed in yellow sticks). The outline mechanism for these enzymes is given in Figure 2.8. (A) MMCD in complex with 2*S*-carboxypropyl-CoA (PDB ID: 1EF9) [50]; (B) Chemical structure of 2*S*-carboxypropyl-CoA showing the stable thioether in place of the thioester bond (dotted frame); (C) DpgC in complex with 3,5-dihydroxyphenylacetyl-aza(dethia)CoA (DPA-NH-CoA) and molecular oxygen (red spheres) (PDB ID: 2NP9) [51]. The oxyanion hole residues Ile-235 and Gly-296 are displayed in green sticks; (D) Chemical structure of DPA-NH-CoA showing the amide bond in place of the thioester (dotted frame); (E) *Mycobacterium tuberculosis* MenB in complex with *O*-succinylbenzoyl-aza(dethia)CoA (OSB-NH-CoA) (PDB ID: 3T8A) [52]; (F) Chemical structure of OSB-NH-CoA showing the amide bond in place of the thioester (dotted frame). This figure was generated using PyMOL [53].

thioester functionality (Table 5.2, Entries XVII-XXIII). Examples of structures solved in complex with carba(dethia)CoA derivatives include citrate synthase [57] and methylmalonyl-CoA mutase [58].

Aza(dethia)CoA Analogues

Three protein crystal structures have been solved in complex with acyl-CoA analogues in which the sulphur atom was replaced by a nitrogen atom, resulting in an amide functionality, which are much more stable towards hydrolysis than the corresponding thioesters.

Two members of the CS family of enzymes described in Section 2.4.1 have been co-crystallised with aza(dethia)CoA substrate analogues; Bruner and coworkers reported the structure of the dioxygenase DpgC from *Streptomyces toyocaensis* in complex with the substrate analogue 2-(3,5-dihydroxyphenyl)acetyl-aza(dethia)CoA (DPA-NH-CoA, Figure 5.2 C and D, Table 5.2, Entry XXIV, PDB ID: 2NP9) [51]. Crystal structures of *Mycobacterium tuberculosis* and *E. coli* MenB have been solved by Tonge and coworkers in complex with the substrate analogues *O*-succinylbenzoyl-aza(dethia)CoA (OSB-NH-CoA) (Figure 5.2 E and F, Table 5.2, Entries XXV and XXVI, PDB ID: 3T8A and 3T88, respectively) [52]. In addition, tetrahydrodipicolinate *N*-succinyltransferase from *Mycobacterium bovis* was co-crystallised with succinyl-aza(dethia)CoA (Table 5.2, Entry XXVII, PDB ID: 1KGQ) [59]. It is interesting to note that another tetrahydrodipicolinate *N*-succinyltransferase crystal structure was determined in complex with the natural substrate succinyl-CoA. A comparison between the two structures show that the carbonyl group of the amide and that of the thioester functionality are pointing in opposite directions, making hydrogen-bond interactions with different residues or water molecules in the active site. This observation suggest that the amido functionality does not perfectly match the binding properties of the thioester.

Entry	Protein	Species	Ligand	PDB ID	Ref.
			Thioethers		
I	CREBBP histone acetyltransferase	*Homo sapiens*	Lysine-CoA	3BIY	[54]
II	GCN5 histone acetyltransferase	*Tetrahymena thermophila*	N″-(2-CoA)-Propanoyl-lysine	1M1D	[55]
III	Serotonin *N*-acetyltransferase	*Ovis aries*	CoA-*S*-Acetyl tryptamine	1CJW	[56]
IV	Serotonin *N*-acetyltransferase	*Ovis aries*	CoA-*S*-Acetyl 5-bromotryptamine	1KUV	[60]
V	Choline *O*-acetyltransferase	*Homo sapiens*	*S*-(2-Oxopropyl)-CoA	2FY5	[61]
VI	Polysialic acid *O*-acetyltransferase	*Neisseria meningitidis*	*S*-(2-Oxopropyl)-CoA	2WLG	[62]
VII	Spermidine/spermine *N*1-acetyltransferase	*Homo sapiens*	*N*1-Acetylspermine-*S*-CoA	2JEV	[63]
VIII	4-Chlorobenzoate:CoA ligase	*Alcaligenes* sp.	4-Chlorophenacyl-CoA	3CW9	[64]
IX	Methylmalonyl-CoA decarboxylase	*Escherichia coli*	2-Carboxypropyl-CoA	1EF9	[50]
X	Methylmalonyl-CoA mutase	*Propionibacterium freudenreichii*	2-Carboxypropyl-CoA	7REQ	[65]
XI	Methylmalonyl-CoA mutase	*Propionibacterium freudenreichii*	3-Carboxypropyl-CoA	6REQ	[65]
XII	4-Hydroxybenzoyl-CoA thioesterase	*Pseudomonas* sp. CBS3	4-Hydroxybenzyl-CoA	1LO8	[66]
XIII	4-Hydroxybenzoyl-CoA thioesterase	*Arthrobacter* sp.	4-Hydroxybenzyl-CoA	1Q4U	[66]
XIV	4-Hydroxybenzoyl-CoA thioesterase	*Pseudomonas* sp. CBS3	4-Hydroxyphenacyl-CoA	1LO7	[66]
XV	4-Hydroxybenzoyl-CoA thioesterase	*Arthrobacter* sp.	4-Hydroxyphenacyl-CoA	1Q4T	[66]
XVI	Acyl-CoA Binding Domain 4	*Homo sapiens*	Stearoyl-CoA	2WH5	

Table 5.1: Protein structures in complex with stable coenzyme A analogues.

Entry	Protein	Species	Ligand	PDB ID	Ref.
			Carba(dethia)CoA		
XVII	Citrate synthase	*Gallus gallus*	Trifluoroacetonyl-CoA	6CSC	[57]
XVIII	Citrate synthase	*Gallus gallus*	α-Fluoro-amidocarboxymethyl-(dethia)CoA	1CSR	[57]
XIX	Citrate synthase	*Gallus gallus*	α-Fluoro-carboxymethyl(dethia)CoA	1CSS	[57]
XX	Citrate synthase	*Gallus gallus*	Nitromethyl-(dethia)CoA	1AMZ	
XXI	Citrate synthase	*Thermoplasma acidophilum*	*S*-Citryl-carba(dethia)CoA	2R9E	
XXII	Methylmalonyl-CoA Mutase	*Propionibacterium freudenreichii*	Succinyl-carba(dethia)CoA	5REQ	[58]
XXIII	Methylmalonyl-CoA Mutase	*Propionibacterium freudenreichii*	Methylmalonyl-carba(dethia)CoA	5REQ	[58]
			Aza(dethia)CoA		
XXIV	3,5-Dihydroxyphenylglyoxylate synthase	*Streptomyces toyocaensis*	2-(3,5-Dihydroxyphenyl)acetyl-aza(dethia)CoA	2NP9	[51]
XXV	1,4-Dihydroxy-2-naphthoyl-CoA synthase	*Mycobacterium tuberculosis*	*O*-Succinylbenzoyl-aza(dethia)CoA	3T8A	[52]
XXVI	1,4-Dihydroxy-2-naphthoyl-CoA synthase	*Escherichia coli*	*O*-Succinylbenzoyl-aza(dethia)CoA	3T88	[52]
XXVII	Tetrahydrodipicolinate *N*-succinyltransferase	*Mycobacterium bovis*	Succinyl-aza(dethia)CoA	1KGQ	[59]

Table 5.2: Protein structures in complex with coenzyme A analogues (continued).

5.2 CoA Analogues as CMPS Substrate Analogues

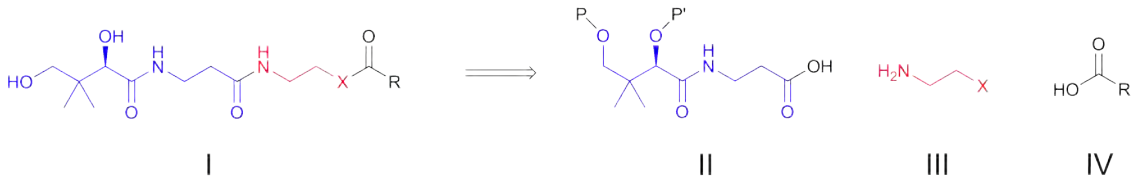
5.2.1 Introduction

The examples described in Section 5.1.4 and Section 5.1.5 demonstrate the utility of coenzyme A analogues, or truncated versions thereof, in the study of CoA-utilising enzymes *in vivo* and *in vitro*. The malonyl-CoA analogues with an ester functionality which mimics the thioester linkage have been previously described and were found to be good substrates of the malonyl-CoA utilising PKS subunits [35]. It was therefore hypothesised that malonyl and methylmalonyl-CoA analogues may also be accepted as CMPS substrates. The replacement of the thioester by an ester, or an amide functionality may allow decarboxylation and C-C bond formation to occur and could slow down or prevent subsequent hydrolysis of the resulting intermediate. Accordingly, this strategy may allow the isolation and characterisation of the corresponding stable *t*-CMP-CoA intermediates. In this section, the synthesis of stable CoA analogues for CMPS biochemical and crystallographic studies is described.

Chemoenzymatic Synthesis of Stable Acyl-CoA Derivatives

The chemoenzymatic synthesis of stable analogues of both the malonyl-CoA **130** substrate and the thioester intermediate of catalysis, *t*-CMP-CoA **132**, can be described as a two step process (Scheme 5.5 B). First, the chemical synthesis of a range of (dethia)pantetheine and acyl-(dethia)pantetheine analogues was carried out from the hemi-calcium salt of D-pantothenic acid **246**. Second, these pantetheine and acyl-pantetheine analogues were subjected to the one-pot enzymatic cascade catalysis (ECC) assay on a preparative scale yielding the corresponding (dethia)CoA and acyl-(dethia)CoA derivatives.

The chemical synthesis of the pantetheine or acyl-pantetheine moiety was designed based on the atom substituting for the sulphur present in the natural thioesters. Three different sets of pantetheine analogues were synthesised. One set was with a nitrogen atom in place of the sulphur (aza(dethia)pantetheine), one with an oxygen (oxa(dethia)pantetheine) and one where the sulphur atom was replaced by a methylene moiety (carba(dethia)pantetheine). In the case of the aza and oxa



Scheme 5.7: Retrosynthetic analysis of the pantetheine precursor for the CoA analogues chemoenzymatic synthesis. The pantetheine precursors **I** that were used for the chemoenzymatic synthesis of the CoA analogues described in this work can be split into three fragments; the pantothenate synthon **II**, the ethylene linker **III** and the terminal acyl functionality **IV**. X = O, NH, CH_2

analogues, retrosynthetic analysis of the pantetheine fragment (Scheme 5.7, Compound **I**) suggests three synthons. The first synthon (Scheme 5.7, Synthon **II**) can be obtained from commercially available D-pantothenic acid **246** through protection of the two hydroxyl groups enabling further functionalisation of the terminal carboxylate. The second synthon (Scheme 5.7, Synthon **III**) is an ethylene linker that can be derived from ethane-1,2-diamine **266** or 2-aminoethanol **267**. The third synthon (Scheme 5.7, Synthon **IV**) was either malonate half-esters or a *t*-CMP moiety. In the case of the carba analogues, synthons **III** and **IV** have to be synthesised as a single entity.

The CoA biosynthesis enzymes used in the ECC assay (i.e. CoaA, CoaD and CoaE) were recombinantly expressed in *E. coli* and purified using a one step affinity purification procedure (Section 5.2.2, Figure 5.3) [35]. Various pantetheine and acyl-pantetheine analogues were converted to the corresponding acyl-CoA analogues using these purified enzymes. The use of the acyl-CoA analogues for CMPS assays and crystallographic studies is described in Section 5.3 and Chapter 6, respectively.

5.2.2 Expression and Purification of Coenzyme A Biosynthetic Enzymes

Genes coding for *E. coli* CoA biosythesis enzymes cloned into expression vectors were a generous gift from Dr Manuela Tosin (University of Cambridge, now at the University of Warwick, U.K.). The expression and purification procedure for CoaA, CoaD and CoaE was adapted from described methodologies [28, 35]. The genes coding for the pantothenate kinase (CoaA) and the 4′-phosphopantetheine adenosine transferase (CoaD) were supplied cloned into the pET-29b(+) vector, between NcoI and HindIII restriction sites. Therefore the genes are under the control of a IPTG-

Protein	Function	Tag	Size (aa)	Molecular Weight (Da)	Extinction Coefficient ϵ
CoaA	Pantothenate kinase	*N*-term His_6	343	39312	43810
CoaD	Phosphopantetheine adenylyltransferase	*N*-term His_6	182	20289	8250
CoaE	Dephospho-CoA kinase	*C*-term His_6	238	26070	23470

The *coaA* and *coaD* genes were cloned from *E. coli* BL21 (DE3) genomic DNA into the pET-29b(+) vector, between the NcoI and HindIII restriction sites. The *coaE* gene was cloned from *E. coli* BL21 (DE3) genomic DNA into the pET-20b(+) vector, between the NdeI and XhoI restriction sites. The size is given for the expected products of the corresponding expression constructs (Appendix D.4). The extinction coefficient ϵ was calculated using a reported equation [68].

Table 5.3: The coenzyme A biosynthetic enzymes used in this work.

inducible T7 promoter and the corresponding proteins carry an *N*-terminal hexahistidine tag. The dephosphocoenzyme A kinase (CoaE) was supplied cloned into a pET-20b(+) vector, between NdeI and XhoI restriction sites. Therefore this gene is also under the control of a T7 promoter, however the corresponding protein carries a *C*-terminal hexahistidine tag. Before expression trials were conducted, the DNA sequence of these constructs was verified by sequencing. All three sequences were searched against the NCBI Nucleotide database using the blastn algorithm [67] and found to be identical to the *E. coli* BL21 (DE3) genes.

A single step affinity purification protocol utilising the hexahistidine tag was used to isolate the proteins of interest from a cell lysate. The elution fractions containing the protein of interest were pooled and concentrated. The final concentrations for each protein were measured based on the calculated absorption coefficients (Table 5.3) and were found to be 8 mg mL^{-1} (CoaA), 9 mg mL^{-1} (CoaD) and 18 mg mL^{-1} (CoaE); the corresponding yields were 150 mg L^{-1} (CoaA), 200 mg L^{-1} (CoaD) and 225 mg L^{-1} (CoaE), respectively. The yields refer to isolated pure protein per liter of growth medium. The protein solution was aliquoted and frozen at -80 °C until used in the ECC assays.

5.2.3 Chemoenzymatic Synthesis of Malonyl-CoA Analogues for Mechanistic and Crystallographic Studies

Synthesis of Synthon II

Synthon II was obtained *via* a one-step protection of the commercially available D-pantothenic

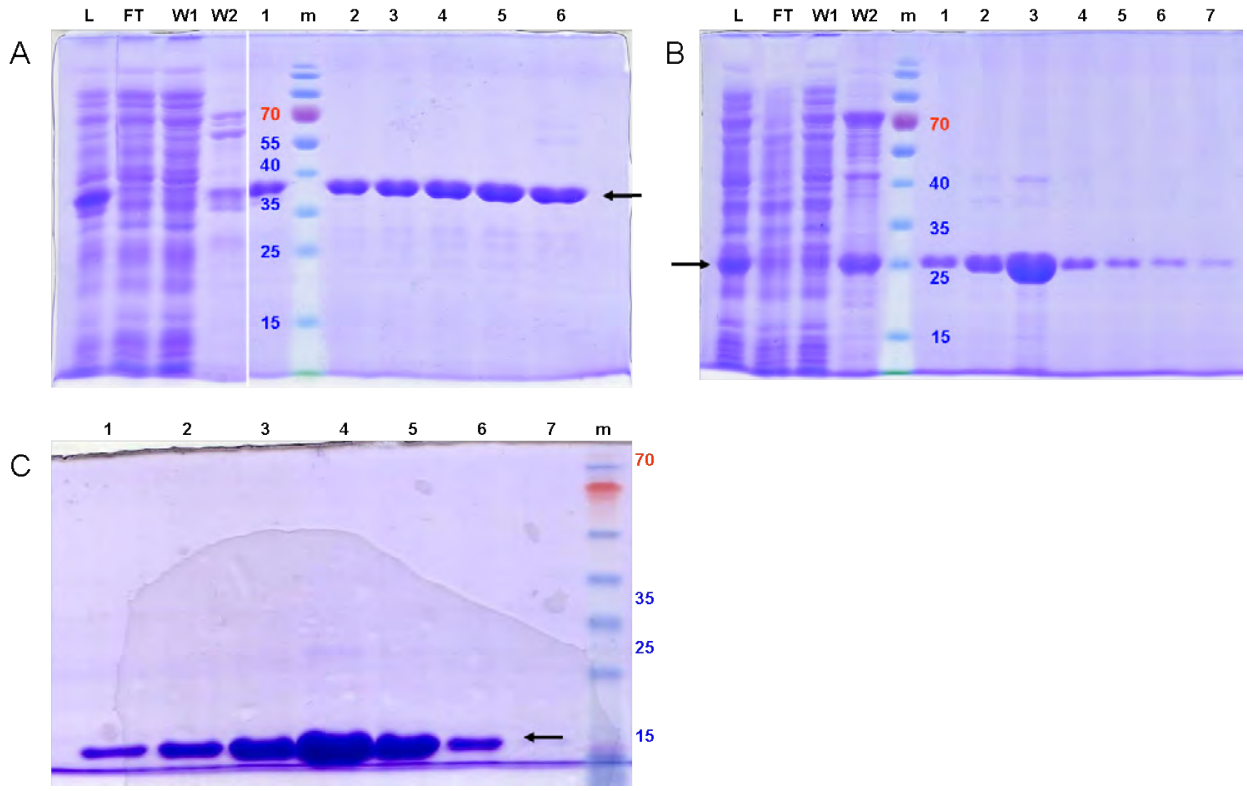


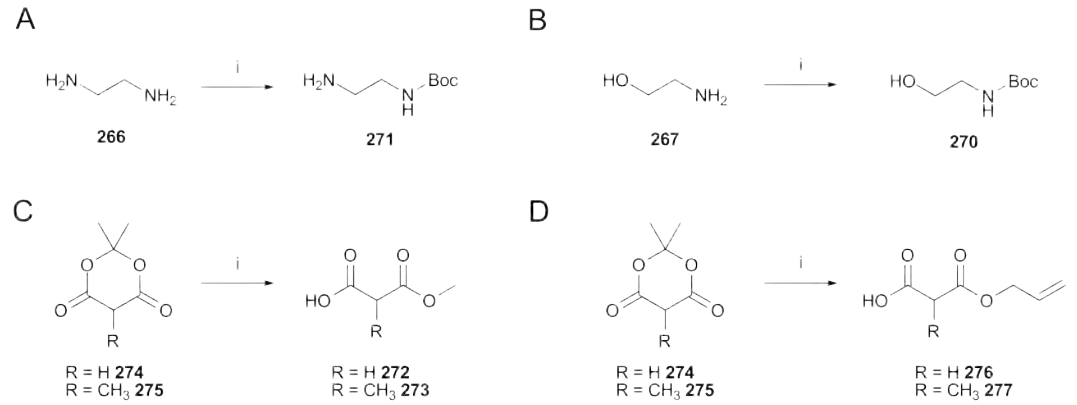
Figure 5.3: SDS-PAGE analysis of CoA biosynthesis enzymes purification steps. Purification was performed after recombinant expression in *E. coli* of CoaA, CoaD and CoaE proteins carrying an hexahistidine tag. All samples were run on 12% acrylamide gels. L, total cell lysate; FT, flow through; W, wash fractions; E, elution fractions; m, prestained protein ladder. The molecular weights of the markers are given in kDa. (A) CoaA calculated molecular weight: 39.3 kDa; (B) CoaD calculated molecular weight: 20.2 kDa; (C) CoaE calculated molecular weight: 26.0 kDa.

acid **246** hemi-calcium salt. The diol protecting group strategy was designed to allow deprotection under either acidic or alkaline conditions as a last step in the synthesis of synthon **I**. For acidic deprotection, strategies have been reported using 2,2-dimethoxypropane [69] or acetone [30] to form *N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **268** (Scheme 5.10). Reaction in the presence of anisaldehyde was reported to yield *N*-[(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **269** (Scheme 5.14) [70, 71]. The hydroxyl groups could also be protected individually *via* acetylation (Scheme 5.9) [35] or trifluoroacetylation [25]. As an alternative, both alcohols have been protected as triethylsilyl esters in the context of crotonyl-oxa(dethia)CoA chemoenzymatic synthesis [72].

N-[(4*R*)-2,2,5,5-Tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **268** was obtained in good yield from D-pantothenic acid **246** following a reported procedure (Scheme 5.10) [35]. D-Pantothenic acid **246** was also reacted with anisaldehyde in the presence of *p*-toluene sulfonic acid to yield *N*-[(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **269**, albeit in a poorer yield (Scheme 5.14 B). The introduction of the 4-methoxyphenyl functionality in synthon **II** was investigated because it gives a greater hydrophobic character to the product than the isopropylidene group allowing a more straight forward purification of the subsequent products from polar reagents and starting materials. It also introduced a UV active group that allows convenient detection using UV absorbance or fluorescence techniques. However, the isopropylidene group was eventually preferred due to the difficulty of removing anisaldehyde by evaporation after deprotection, while acetone was readily removed. Anisaldehyde was also found to interfere with the CoA biosynthesis enzyme assay, even when low amounts were still present. Therefore, HPLC purification of the pantetheine analogues was required prior to the assay. As an alternative to acidic deprotection, the two hydroxyl groups were protected with acetic anhydride in order to allow alkaline deprotection of the corresponding esters, as reported [35].

Synthesis of Synthon III

Four different ethylene linkers were required in order to perform the orthogonal connection of fragments II and III or III and IV, depending on the synthetic sequence. Ethane-1,2-diamine **266** and

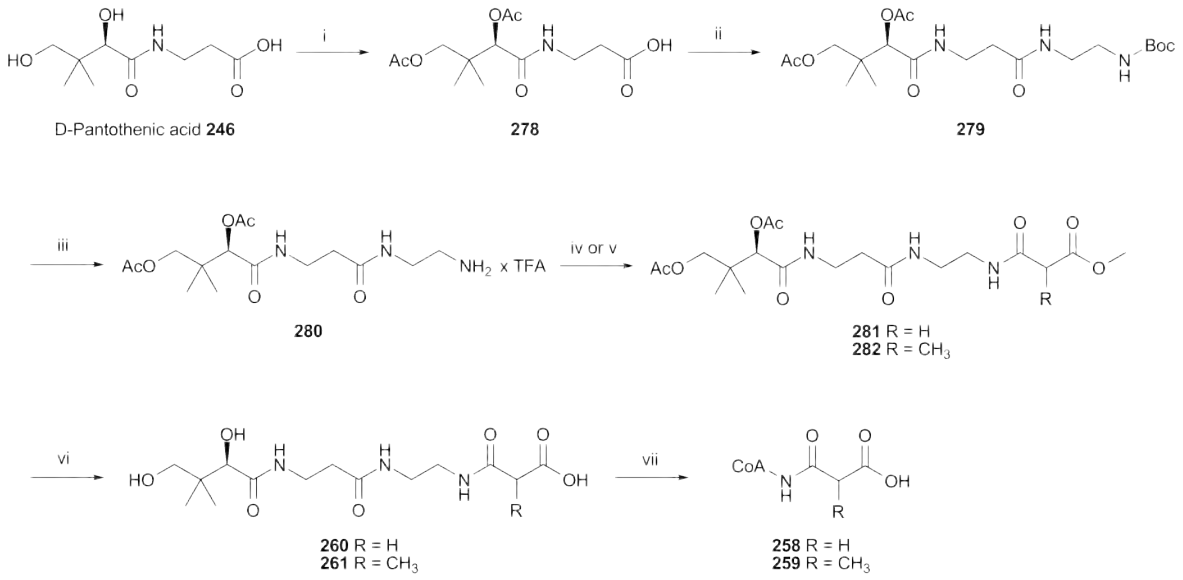


Scheme 5.8: Synthesis of pantetheine synthons III and IV. (A) (i) $(Boc)_2O$, 1,4-dioxane, rt, 4 h, 90%; (B) (i) $(Boc)_2O$, 1,4-dioxane, rt, 4 h, 90%; (C) (i) MeOH, reflux, 4 d, 88%; (D) (i) Allyl alcohol, dry ACN, reflux, 16 h, 99%.

2-aminoethanol **267** were commercially available and were protected to enable orthogonal coupling of either the amino or hydroxyl functionality. The amino group of 2-aminoethanol was protected using one equivalent of di-*tert*-butyl dicarbonate following a reported procedure [73] to give *tert*-butyl (2-hydroxyethyl)carbamate **270** in good yield. The mono-protection of ethane-1,2-diamine **266** was successfully achieved by adapting reported procedures [74, 75] using an excess of the volatile amine. Removal of the excess reagent by evaporation gave the non-volatile *tert*-butyl (2-aminoethyl)carbamate **271** in good yield.

Synthesis of Synthon IV

A series of acyl-pantetheine analogues were synthesised by coupling the protected pantetheine analogues with malonate and 2-methylmalonate derivatives. The half-methyl esters of malonate and methylmalonate, 3-methoxy-3-oxopropanoic acid **272** and 3-methoxy-2-methyl-3-oxopropanoic acid **273**, have been previously accessed by monohydrolysis of dimethylmalonates [76]. In this work, the half-methyl esters **272** and **273** were obtained by methanolysis of 2,2-dimethyl-dioxane-4,6-dione **274** (Meldrum's acid) and 2,2,5-trimethyl-dioxane-4,6-dione **275**, respectively (Scheme 5.8 C) [77]. Similarly, the monoallyl malonate derivatives 3-(allyloxy)-3-oxopropanoic acid **276** and 3-(allyloxy)-2-methyl-3-oxopropanoic acid **277** were obtained from 2,2-dimethyl-dioxane-4,6-dione **274** (Meldrum's acid) and 2,2,5-trimethyl-dioxane-4,6-dione **275**, respectively, following a procedure described by Barrett and coworkers [78] (Scheme 5.8 D). In both cases, the malonate



Scheme 5.9: Chemoenzymatic production of aza(dethia)CoA derivatives analogues starting from D-pantothenate. (i) Ac_2O, cat. I_2, 0 °C–rt, 48 h., 97%; (ii) **271**, Et_3N, HOBt, EDC HCl, dry THF, 0 °C–rt, 32 h., 59%; (iii) 10% TFA in CH_2Cl_2, 4 h, rt, app. quant.; (iv) **272**, DIPEA, HOBt, EDC HCl, 0 °C–rt, 24 h., 22%; (v) **273**, DIPEA, HOBt, EDC HCl, 0 °C–rt, 24 h., 45%; (vi) K_2CO_3, MeOH, rt, 16 h, app. quant.; (vii) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt.

and methylmalonate half esters were obtained in good yield and evaporation of the reaction solvent gave the products without any purification step being required.

Synthesis of Malonyl and Methylmalonyl-aza(dethia)CoA

The synthesis of methylmalonyl-aza(dethia)pantetheine **261** was reported during the course of this study [36]. However that of malonyl-aza(dethia)pantetheine **260** has not been reported. Malonyl and methylmalonyl-aza(dethia)CoA **258** and **259** were obtained from the corresponding aza(dethia)pantetheine analogues described by Tosin *et al.* [36], followed by an ECC conversion to the CoA analogues using CoaA, CoaD and CoaE *in vitro* (Scheme 5.9).

N-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-
β-alanine **278** was chosen for alkaline deprotection of the acetyl functionality as a last step. The reaction of *N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanine **278** with synthon III *tert*-butyl (2-aminoethyl)carbamate **271** using 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride and hydroxybenzotriazole as coupling reagents yielded (14*R*)-14-acetoxy-2,2,15,15-tetramethyl-

4,9,13-trioxo-3-oxa-5,8,12-triazahexadecan-16-yl acetate **279** in moderate yield. The apparent quantitative removal of the *N*-Boc group of compound **279** using trifluoroacetic acid, followed by coupling of the resulting trifluoroacetate salt of (2*R*)-4-acetoxy-1-(3-[(2-aminoethyl)amino]-3-oxopropylamino)-3,3-dimethyl-1-oxobutan-2-yl acetate **280** with either malonate half-methyl ester **272** or methylmalonate half-methyl ester **273** as synthon **IV** yielded the protected pantetheine analogues **281** and **282**. One-step deprotection of the malonate methyl ester and the acetyl protecting groups of **281** and **282** using potassium carbonate in cold methanol gave the fully deprotected 3-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-3-oxopropanoic acid **260** and 3-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-2-methyl-3-oxopropanoic acid **261**, respectively. These two precursors were purified by preparative HPLC prior to ECC conversion to the CoA derivative. After the assay, the two aza(dethia)CoA derivatives **258** and **259** were purified by preparative HPLC and isolated as white solids after lyophilisation.

Synthesis of Malonyl and Methylmalonyl-oxa(dethia)CoA

Similarly to the synthesis of malonyl and methylmalonyl-aza(dethia)CoA, **258** and **259**, malonyl-oxa(dethia)CoA **262** and methylmalonyl-oxa(dethia)CoA **263** were obtained *via* the corresponding malonyl-oxa(dethia)pantetheine and methylmalonyl-oxa(dethia)pantetheine precursors (Scheme 5.10). The synthesis of malonyl-oxa(dethia)pantetheine **264** has been reported [35] and the strategy described by Tosin *et al.* was used as a guideline to access methylmalonyl-oxa(dethia)pantetheine **265** [36]. The synthesis of compound **265** was reported during the course of this study [79].

The protected oxa(dethia)pantetheine precursor (4*R*)-*N*-3-[(2-hydroxyethyl)amino]-3-oxopropyl-2,2,5,5-tetramethyl-
1,3-dioxane-4-carboxamide **283** was obtained by coupling *N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **268** to 2-aminoethanol **267** using ethylchloroformate (Scheme 5.10). 3-(Allyloxy)-3-oxopropanoic acid **276** and 3-(allyloxy)-2-methyl-3-oxopropanoic acid **277** were used as synthon IV and the malonate ester was formed using HBTU as the coupling reagent to give the fully protected malonyl- and methylmalonyl-oxa(dethia)pantetheine compounds, allyl 2-[(*N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethyl malonate **284** and allyl 2-[(*N*-

[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethyl methylmalonate **285**. The allyl-ester deprotection step was carried out under mild conditions using a catalytic amount of tetrakis(triphenylphosphine)palladium to give 3-oxo-3-2-[(*N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethoxypropanoic acid **286** and 2-methyl-3-oxo-3-2-[(*N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethoxypropanoic acid **287**. The isopropylidene group was removed under mild acidic conditions using DOWEX 50WX8-400 (H^+) resin [35]. The two resulting malonyl and methylmalonyl-oxa(dethia)pantetheine analogues, **264** and **265**, were subjected to preparative HPLC purification prior to the one-pot ECC conversion to the CoA derivative. Malonyl-oxa(dethia)CoA **262** and methylmalonyl-oxa(dethia)CoA **263** were purified by preparative HPLC and isolated as white solids after lyophilisation. As expected, the malonate protons were found to be much more acidic than for the aza(dethia)CoA analogues **258** and **259**, as observed by rapid deuterium exchange when ^{1}H NMR was recorded in D_2O. In the case of malonyl-oxa(dethia)CoA **262**, the malonate singlet was visible at 3.45 ppm when the ^{1}H NMR was recorded in 90% H_2O (Figure 5.4, Left), However rapid exchange suppressed the observation of the singlet after 1 h in D_2O (Figure 5.4, Right). When recorded immediately after sample preparation, the ^{1}H NMR spectrum for methylmalonyl-oxa(dethia)CoA **263** shows the anticipated methylmalonate C*H* quartet at 3.59 ppm overlaping with one of the CH_a signals. The corresponding methyl group signal at 1.32 ppm has a coupling constant of J = 7.5 Hz (Appendix D.24).

Chemoenzymatic Synthesis of Malonyl-carba(dethia)CoA

The chemoenzymatic synthesis of malonyl-carba(dethia)CoA methyl ester **288** was performed as reported by Tosin *et al.* with minor modifications (Scheme 5.11) [35]. The protected pantothenate derivative **278** was used as synthon II and coupled to *tert*-butyl γ-aminobutyrate to yield *tert*-butyl 4-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)butanoate **289**. The *tert*-butyl ester was removed with trifluoroacetic acid and the free carboxylic acid of 4-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)butanoic acid **290** was directly reacted with 2,2-dimethyl-1,3-dioxane-4,6-dione **274** to yield β-keto esters using Meldrum's acid [80]. The resulting cyclic adduct **291** was subjected to decarboxylative methanolysis to give methyl 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-3-oxohexanoate **292**. Removal of the acetyl groups

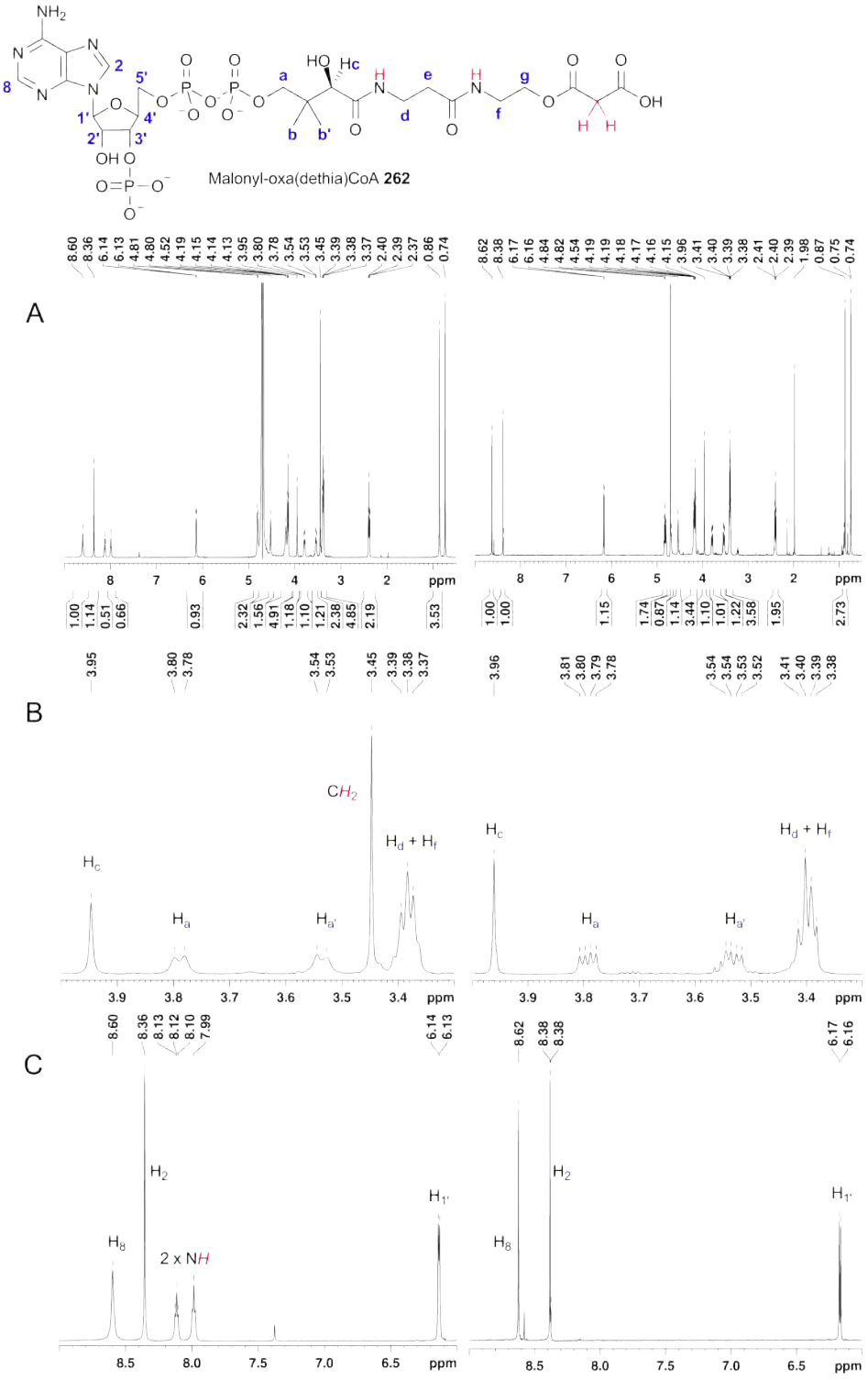
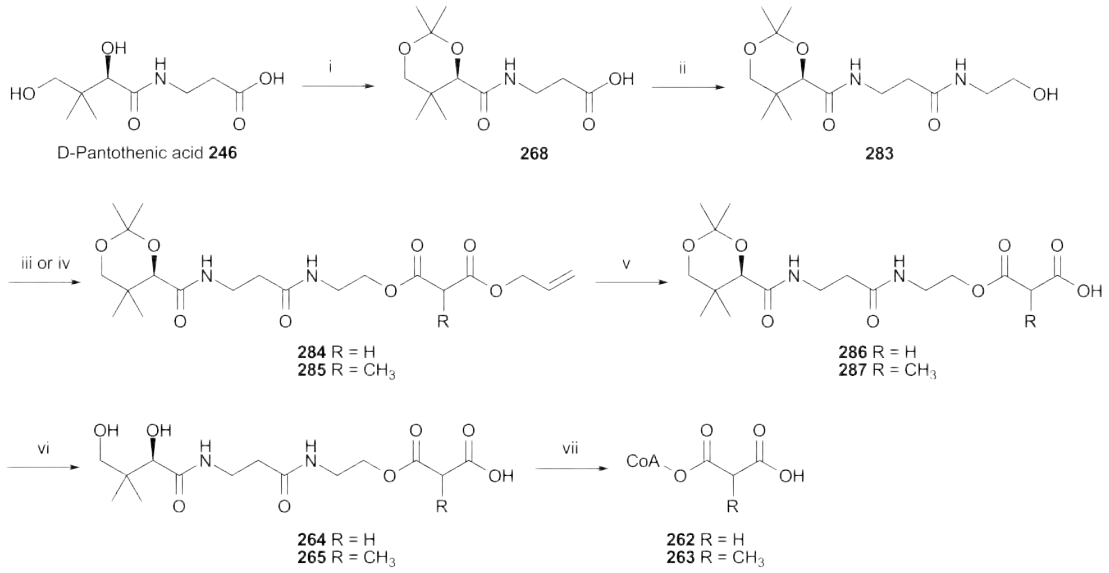


Figure 5.4: Deuterium exchange in malonyl-oxa(dethia)CoA. Exchangeable protons that are visible in 90% H_2O but absent in D_2O are highlighted in red in the structure of malonyl-oxa(dethia)CoA **262**. The nomenclature used to describe the CoA protons was used previsouly by Patel and Walt [3]. (A) Full ^{1}H NMR spectrum for **262** recorded in 90% H_2O at 500MHz (left) and in D_2O at 500MHz (right); (B) Expansion for the 4.0–3.3 ppm region. The malonate protons give a singlet at 3.45 ppm in 90% H_2O (left) and are exchanged for deuterium in D_2O (right); (C) Expansion for the 9.0–6.0 ppm region. Amide protons give two triplets at 8.12 and 7.99 ppm in 90% H_2O (left) and are exchanged for deuterium in D_2O (right).



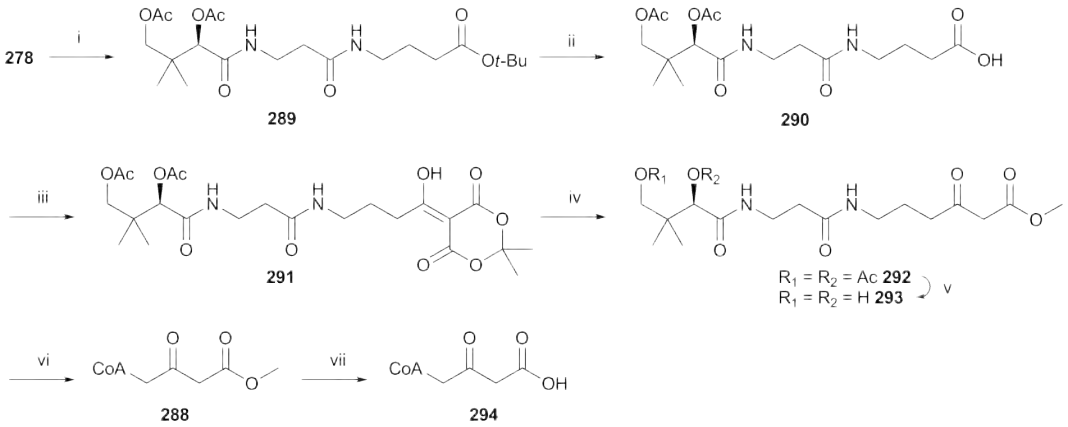
Scheme 5.10: Chemoenzymatic production of oxa(dethia)CoA derivatives starting from D-pantothenate. (i) DMP, *p*-TsOH, rt, 18 h, 81%; (ii) Et_3N, ethylchloroformate, 2-aminoethanol, dry CH_2Cl_2, 0 °C–rt, 3 h, 55%; (iii) **276**, DIPEA, HBTU, dry THF, 0 °C–r.t., 16 h, 30%; (iv) **277**, DIPEA, HBTU, dry THF, 0 °C–rt, 16 h, 70%; (v) $Pd(Ph_3P)_4$, pyrrolidine, ACN, 0 °C, 3 h, 93%; (vi) DOWEX 50WX8-400 (H^+), H_2O, rt, 16 h, app. quant.; (vii) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 43-52%.

under mild alkaline conditions afforded methyl 6-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)-3-oxohexanoate **293**, the methyl ester of malonyl-carba(dethia)pantetheine .

Malonyl-carba(dethia)pantetheine methyl ester **293** was subjected to the ECC reaction in the presence of 10 equivalents of ATP. After overnight incubation, the product was purified by preparative HPLC (Appendix D.25 and D.26). In an attempt to isolate the free acid **294**, the methyl ester **288** was incubated with pork liver esterase (PLE) overnight at room temperature, however the yield of methyl ester hydrolysis by the enzyme was found to be relatively low, as judged by HPLC (Appendix D.27), and the product too unstable to be observed by mass spectrometry analysis, probably due to rapid decarboxylation. For this reason, a PLE/CMPS coupled assay was used in an attempt to generate the non-hydrolysable *t*-CMP-CoA analogue *t*-CMP-carba(dethia)CoA **295** (Section 5.3.1).

5.2.4 Chemoenzymatic Synthesis of *t*-CMP-CoA Analogues

Two possible routes toward the amide and ester analogues of *t*-CMP-CoA **132** were investigated in parallel. The first route involved the ECC synthesis of CoAOH **252** or $CoANH_2$ **253**, followed by chemical coupling to a protected *t*-CMP moiety and a final deprotection step (Scheme 5.15 A). The second route involved the chemical synthesis of *t*-CMP-pantetheine analogues followed

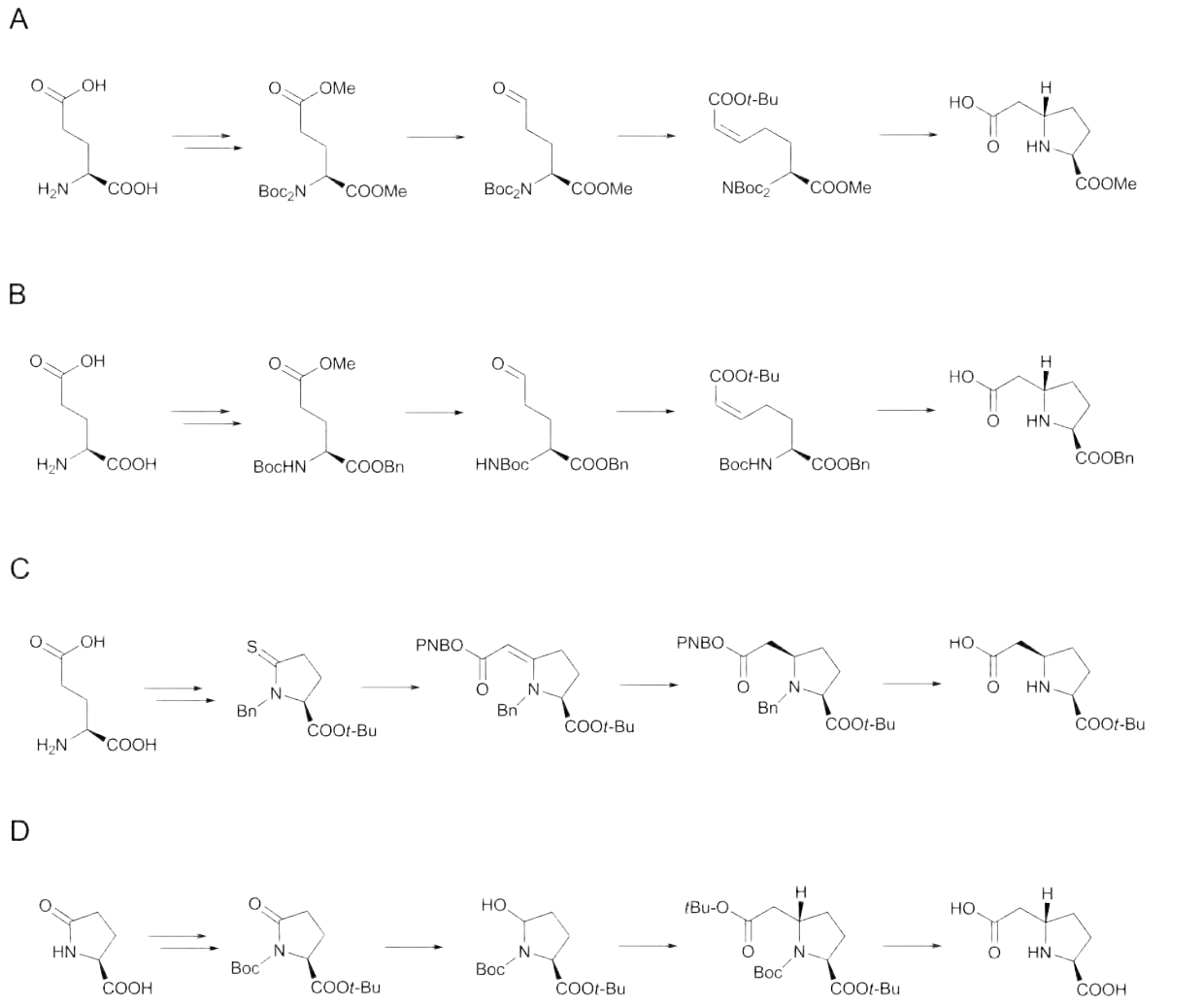


Scheme 5.11: Chemoenzymatic production of malonyl-carba(dethia)CoA analogue **294**. (i) GABA-*O*-*t*Bu, Et_3N, HOBt, EDC HCl, CH_2Cl_2, 0 °C–rt, 18 h, 93%; (ii) 10% TFA in CH_2Cl_2, 3 h, rt, app. quant.; (iii) Meldrum's acid, DMAP, EDC HCl, CH_2Cl_2, 0 °C–rt, 18 h; (iv) MeOH, reflux, 18 h, 77% over two steps; (v) K_2CO_3, MeOH, 0 °C, 5 h, 55%; (vi) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 43-52%; (vii) PLE, 50 mM HEPES pH 8.0, 25 mM KCl, rt, 18 h, 28%.

by the ECC conversion to stable *t*-CMP-CoA analogues (Scheme 5.16). The two routes were successfully applied by others to the synthesis of non-hydrolysable acyl-CoA analogues for crystallographic studies (Section 5.1.5). Bruner and coworkers used the coupling strategy to obtain 3,5-dihydroxyphenylacetyl-aza(dethia)CoA (DPA-NH-CoA) from $CoANH_2$ **253** as ligand for DpgC (Figure 5.2 D) [51], while Tonge and coworkers used the enzymatic one-pot reaction as a last step in the synthesis of the MenB ligand *O*-succinylbenzoyl-aza(dethia)CoA (OSB-NH-CoA) (Figure 5.2 F) [52].

Chemical Synthesis of *t*-CMP Synthon IV

A synthesis of *t*-CMP using a protecting group strategy compatible with later coupling to pantetheine or CoA analogues was required. Various synthetic approaches to produce the CMP fragment have been published (Scheme 5.12). The synthesis of *trans*-(3*S*,5*S*)-CMP and its enantiomer *trans*-(3*R*,5*R*)-CMP was reported by Bycroft *et al.* starting from protected L-glutamate and D-glutamate, respectively (Scheme 5.12 A and B for the *trans*-(3*S*,5*S*)-CMP enantiomer) [81]. Both the methyl ester (Scheme 5.12 A) and the benzyl ester (Scheme 5.12 B) of the C-2 carboxylate were described. In these two routes, a Wittig olefination of the γ-glutamate aldehyde using *tert*-butyl(triphenylphosphoranylidene)acetate followed by a thermodynamically controlled



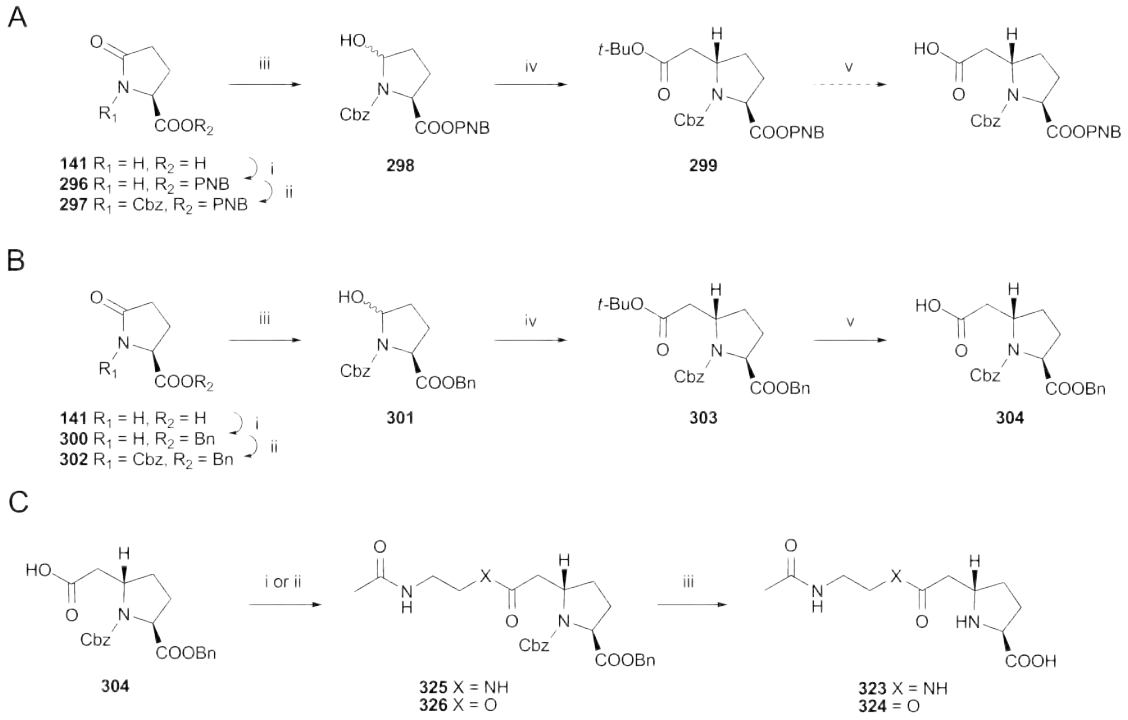
Scheme 5.12: Published syntheses of carboxymethylproline. (A) Route described in Bycroft *et al.* starting from L-glutamic acid and yielding the C-2 carboxylate methyl ester of *t*-CMP [81]; (B) Route described in Bycroft *et al.* yielding the C-2 carboxylate benzyl ester of *t*-CMP [81]; (C) Route described in Gerratana *et al.* starting from L-glutamic acid and yielding the C-2 carboxylate *tert*-butyl ester of *cis*-CMP [84]; (D) Route described in Gerratana *et al.* starting from L-pyroglutamic acid and yielding the benzyl ester of *t*-CMP [84].

intramolecular Michael reaction step allowed the construction of the pyrrolidine ring and was optimised for the preparation of the *trans*-stereoisomers with a *trans*/*cis* ratio of about 11:1. Gerratana *et al.* preferentially obtained the *cis*-CMP diastereomer using a strategy starting from L-glutamic acid and involving an Eschenmoser coupling [82] previously applied to the synthesis of proline derivatives by Rapoport and coworkers [83] (Scheme 5.12 C). This *cis*-CMP compound was tested as a substrate for CarA β-LS activity (Figure 2.4, Section 2.1.2) [84].

In the present case, the *t*-CMP synthon **IV** was synthesised from L-pyroglutamic acid **141** using a tandem Horner-Wadsworth-Emmons (HWE) and Michael reaction previously used for the synthesis

of 5-substituted proline derivatives [85] and used by Townsend and coworkers in the context of carbapenam synthase CarA investigations. Gerratana *et al.* described the synthesis of the free carboxylate of *t*-CMP [84] as well as the benzyl ester (Scheme 5.12 D) [86, 84]. Di-*tert*-butyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **143** used to generate L-GHP analogues methylated at the C-4 position (Section 2.2.2) could not be exploited as synthon IV due to protecting group orthogonality reasons. The presence of two carboxylate functions on *t*-CMP requires orthogonal protecting groups to allow selective amide, ester or thioester formation at the C-6 carboxylate. Because either the nitrogen or C-2 carboxylate of *t*-CMP carried a protecting group that was meant to be removed with the C-6 carboxylate protecting group, none of the *t*-CMP derivatives described previously could be used without modifications. The most convenient HWE-Michael procedure using commercially available *tert*-butyl phosphonoacetate would install an acid labile *tert*-butyl protected C-6 carboxylate (Scheme 5.12 A and B), this orthogonal protecting groups would then be removed under alkaline or hydrogenation conditions.

An initial attempt to obtain *t*-CMP as a C-2 4-nitrobenzyl ester was carried out starting from L-pyroglutamic acid **141**; However, only a low yield of the desired product was obtained (Scheme 5.13 A). After protection of the C-2 carboxylate of L-pyroglutamic acid **141** with 4-nitrobenzyl bromide to give 4-nitrobenzyl 5-oxo-L-prolinate **296**, the lactam nitrogen was protected using di-*tert*-butyl dicarbonate. The reduction of 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **297** yielded 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-hydroxypyrrolidine-1,2-dicarboxylate **298**. When the hemiaminal **298** was subjected to the HWE-Michael conditions, only small amounts of the desired product 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **299** could be isolated after chromatography. The 4-nitrobenzyl group of 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **297** was replaced by a benzyl group which can also be removed by hydrogenation, although less efficiently (Scheme 5.13 B). Several procedures for the protection of the L-pyroglutamate carboxylate group as a benzyl ester have been reported [87, 88, 89, 90]. Benzyl 5-oxo-L-prolinate **300**, was obtained *via* the alkylation of the carboxylate with benzyl bromide by adapting these procedures. The lactam nitrogen was then protected with a carboxybenzyl group using a previously reported procedure [91]. Dibenzyl (2*S*)-5-hydroxypyrrolidine-1,2-dicarboxylate **301** was synthesised by reduction



Scheme 5.13: Synthesis of protected *t*-CMP as synthon IV and truncated *t*-CMP-CoA analogues. (A) (i) PNBBr, K_2CO_3, cat. I_2, dry DMF, 48 h, 85%; (ii) CbzCl, *n*-BuLi, THF, -78 °C–0 °C, 3 h, 67%; (iii) ; (iv); (B) (i) BnBr, K_2CO_3, $Na2SO_4$, cat. I_2, DMF, 48 h, 80%; (ii) CbzCl, *n*-BuLi, THF, -78 °C–0 °C, 3 h, 54%; (iii) $LiEt_3BH$, THF, -78 °C, 6 h, 72%; (iv) *tert*-Butyldiethylphosphonoacetate, KH, DMF, rt, 24h, 69%; (v) CF_3COOH, rt, 6h, app. quant.; (C) (i) *N*-(2-aminoethyl)acetamide, DIPEA, HOBt, EDC HCl, dry CH_2Cl_2, 0 °C–rt, 24 h, 61%; (ii) *N*-(2-hydroxyethyl)acetamide, DIPEA, HOBt, EDC HCl, dry CH_2Cl_2, 0 °C–rt, 24 h, 61%; (iii) H_2, Pd/C, EtOH, rt, 5 h, 98%.

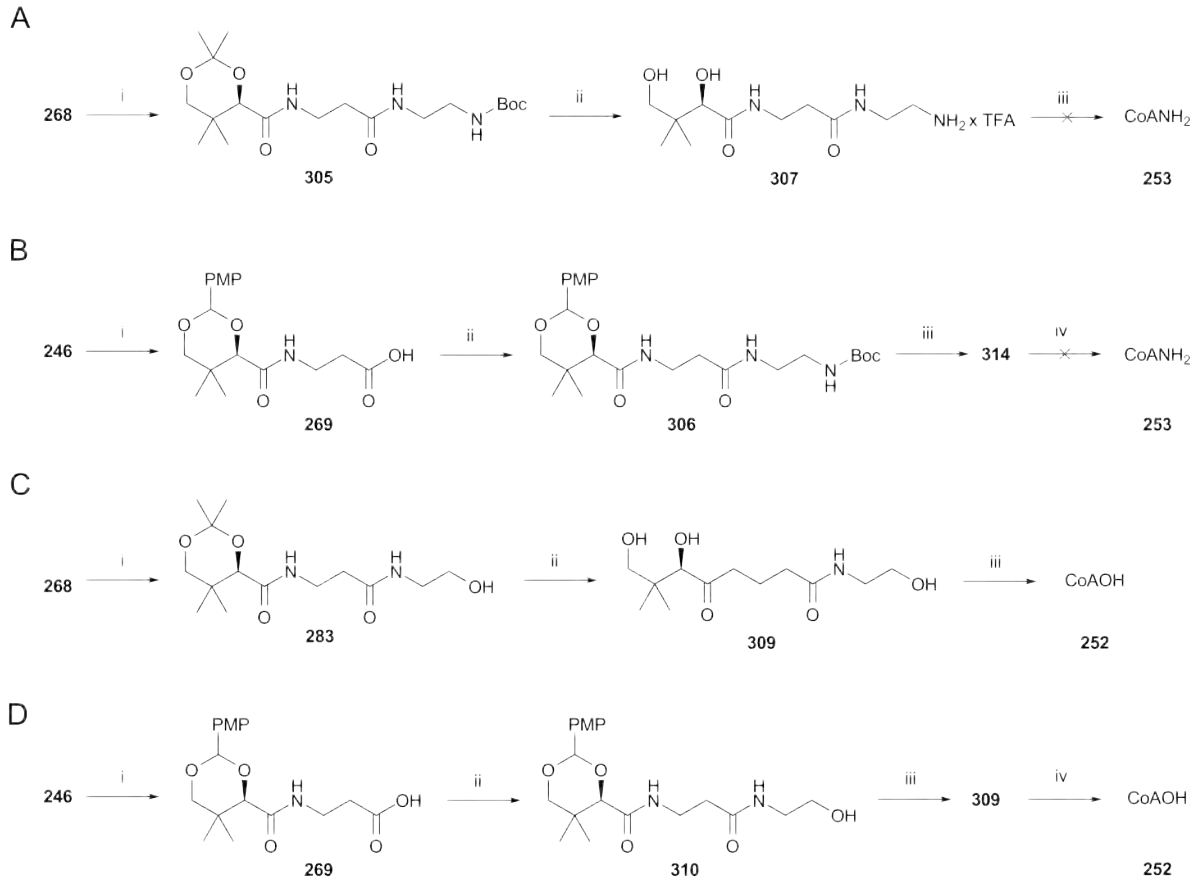
of dibenzyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **302** and was further subjected to the HWE-Michael reaction conditions in the presence of *tert*-butyldiethylphosphonoacetate. The corresponding product, dibenzyl (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **303**, was obtained in moderate yield as a 7:3 mixture of *trans*- and *cis*-isomer as judged by ^{1}H NMR analysis (not shown). This ratio is different from the one reported for this reaction on different substrates [85, 86] and the difference may be attributed to the replacement of the *N*-Boc by an *N*-Cbz protecting group for the hemiaminal **301**. The stereochemistry at C-5 was assigned based on the analysis of 2D NOE experiments. Acidic deprotection of the *tert*-butyl ester with trifluoroacetic acid gave (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** in apparent quantitative yield. This compound was used in the subsequent coupling step with pantetheine and CoA analogues without purification.

Attempted Chemoenzymatic Synthesis of Amino(dethia)CoA

The chemoenzymatic synthesis of amino(dethia)CoA **253** has been reported [92, 93]. In the present work, *N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **268** and *N*-[(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **269** were used as synthon **II** and were coupled to synthon **III**, *tert*-butyl (2-aminoethyl)carbamate **271**, using coupling reagents (Scheme 5.14 A and B). The products **305** and **306** were obtained in good yield after purification by chromatography. In both cases, acidic deprotection yielded the trifluoroacetate salt of (2*R*)-*N*-3-[(2-aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide **307**. (2*R*)-*N*-3-[(2-Aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide **307** was purified by HPLC to remove the trifluoroacetic acid present before using this precursor in the CoA biosynthesis assay; trifluoroacetic acid was found to interfere with the ECC conversion to the CoA derivative. Purified (2*R*)-*N*-3-[(2-aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide **307** was then used in the ECC assay in the presence of an excess of ATP. Unfortunately, at the optimal pH, the product of CoaA and CoaD catalysis, dephospho-amino(dethia)CoA **308**, was observed, but the final CoaE-catalysed step did not proceed, yielding no detectable amount of amino(dethia)CoA **253**. The absence of conversion of dephospho-amino(dethia)CoA **308** into amino(dethia)CoA **253** had been observed previously by Liu and Bruner and these authors reported that increasing the pH of the assay from 8.0 to 9.0 prior to adding CoaE solved the problem. However, in my hands, the ECC did not yield sufficient amounts of amino(dethia)CoA **253** for characterisation. Instead, dephospho-amino(dethia)CoA **308** was isolated in good yield and characterised after preparative HPLC purification.

Chemoenzymatic Synthesis of Hydroxy(dethia)CoA

Hydroxy(dethia)CoA **252** (CoAOH) was the first CoA analogue produced using a chemoenzymatic approach reported by Stewart and Ball (Section 5.1.3) [23]. However, CoAOH **252** was not obtained using the purified *E. coli* proteins CoaA, CoaD and CoaE. Instead the authors used a crude enzyme system isolated from pork liver [23]. In order to perform the ECC assay, the hydroxy(dethia)pantetheine precursor (2*R*)-2,4-dihydroxy-*N*-3-[(2-hydroxyethyl)amino]-



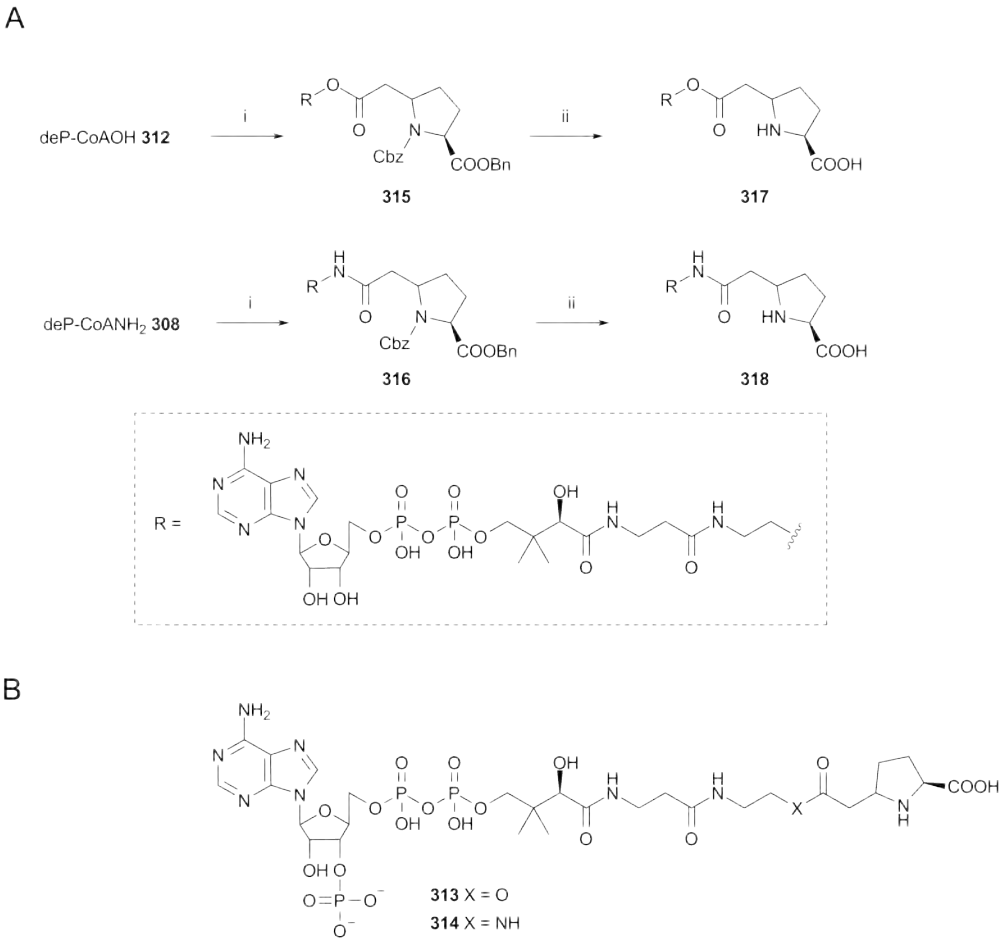
Scheme 5.14: Chemoenzymatic synthesis of amino(dethia)CoA **253** and hydroxy(dethia)CoA **252**. (A) (i) **271**, DIPEA, HOBt, EDC HCl, dry THF, 0 °C–rt, 18 h, 96%; (ii) 10% TFA in CH_2Cl_2, 3 h, rt, app. quant.; (iii) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, n.i.; (B) (i) *p*-TsOH, anisaldehyde, THF, rt, 5 h, 46%; (ii) **271**, DIPEA, HOBt, EDC HCl, dry THF, 0 °C–rt, 18 h., 88%; (iii) 10% TFA in CH_2Cl_2, 3 h, rt, app. quant.; (iv) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, n.i.; (C) (i) Et_3N, ethylchloroformate, 2-aminoethanol, dry CH_2Cl_2, 0 °C–rt, 3 h, 55%; (ii) DOWEX 50WX8-400 (H^+), H_2O, rt, 16 h, app. quant.; (iii) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 15%; (D) (i) *p*-TsOH, anisaldehyde, THF, rt, 5 h, 46%; (ii) Et_3N, EtOCOCl, 2-aminoethanol, dry CH_2Cl_2, 0 °C–rt, 16 h, 96%; (iii) 10% TFA in CH_2Cl_2, 3 h, rt, app. quant.; (iv) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 15%.

3-oxopropyl-3,3-dimethylbutanamide **309** could be obtained *via* the deprotection of either (4*R*)-*N*-3-[(2-hydroxyethyl)amino]-3-oxopropyl-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** or (4*R*)-*N*-3-[(2-hydroxyethyl)amino]-3-oxopropyl-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide **310** under acidic conditions (Scheme 5.14 C and D), or *via* alkaline deprotection of (2*R*)-4-acetoxy-1-(3-[(2-hydroxyethyl)amino]-3-oxopropylamino)-3,3-dimethyl-1-oxobutan-2-yl acetate **311**. Acidic deprotection of compounds **283** and **310** gave (2*R*)-2,4-dihydroxy-*N*-3-[(2-hydroxyethyl)amino]-3-oxopropyl-3,3-dimethylbutanamide **309** (hydroxy(dethia)pantetheine) in apparent quantitative yield. Compound **309** was purified by preparative HPLC prior to enzymatic conversion to CoAOH **252** to remove any impurity that could inhibit the 'one-pot' ECC reaction. The purified hydroxy(dethia)pantetheine **309** was subjected to ECC in the presence of an excess of ATP, and CoAOH **252** was fully characterised after preparative HPLC purification. Additionally, the intermediate of catalysis, dephospho-hydroxy(dethia)CoA **312**, could be isolated and characterised.

Coupling of Dephospho-amino(dethia)CoA and Dephospho-hydroxy(dethia)CoA with *t*-CMP

In an initial attempt to obtain non-hydrolysable *t*-CMP-CoA analogues, a route was investigated involving the chemoenzymatic synthesis of amino(dethia)CoA **253** and hydroxy(dethia)CoA **252** followed by coupling to a *t*-CMP moiety. This route is similar to that described for the synthesis of 3,5-dihydroxyphenylacetyl-aza(dethia)CoA (DPA-NH-CoA) by Bruner and coworkers (Figure 5.2 A and B) [51]. By adapting the procedure reported by Bruner and coworkers, *t*-CMP-aza(dethia)CoA **313** and *t*-CMP-oxa(dethia)CoA **314** were anticipated to be accessible from the protected *t*-CMP synthon **IV** (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** using a coupling strategy (Scheme 5.15 A). The synthesis of *t*-CMP-aza(dethia)CoA **313** and *t*-CMP-oxa(dethia)CoA **314** differs from that of DPA-NH-CoA because protection of the C-2 *t*-CMP carboxylate for orthogonal coupling with the C-6 carboxymethyl moiety is required. A benzyl group was used in order to allow deprotection under hydrogenation conditions as a last step.

Dephospho-hydroxy(dethia)CoA **312** and dephospho-amino(dethia)CoA **308** were used in test reactions in an attempt to obtain the protected *t*-CMP-dephospho-oxa(dethia)CoA **315** and protected



Scheme 5.15: Attempted synthesis of stable *t*-CMP-CoA analogues. This strategy was adapted from a procedure reported by Bruner and coworkers [51]. (A) (i) **304**, HBTU, HOBt, DIPEA, 0 °C–rt, 4 h, 41–42%; (ii) H_2, Pd/C, 1:1 THF/H_2O, rt, 5 h, 67–78%; (B) Structure of *t*-CMP-oxa(dethia)CoA **314** (X = O) and *t*-CMP-aza(dethia)CoA **313** (X = N).

t-CMP-dephospho-aza(dethia)CoA **316**. Dephospho-hydroxy(dethia)CoA **312** was coupled to (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** to yield protected *t*-CMP-dephospho-oxa(dethia)CoA **315** (Scheme 5.15 A). Compound **315** was purified by preparative HPLC and was obtained in moderate yield as a white solid after lyophilisation. Dephospho-aza(dethia)CoA **316** was coupled to (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** using the same procedure as for **315** (Scheme 5.15 A). *t*-CMP-dephospho-aza(dethia)CoA **316** was obtained in moderate yield as a white solid after preparative HPLC purification. Both the protected *t*-CMP-dephospho-oxa(dethia)CoA **315** and protected *t*-CMP-dephospho-aza(dethia)CoA **316** were subjected to catalytic hydrogenation conditions in order to remove the *t*-CMP benzyl and carboxybenzyl groups. *t*-CMP-dephospho-oxa(dethia)CoA **317** and *t*-CMP-dephospho-aza(dethia)CoA **318** were isolated and characterised after removal of the palladium catalyst by filtration.

These results suggest that the 3′-phosphate of CoA is not labile under hydrogenation conditions. This strategy could therefore be used to generate *t*-CMP-aza(dethia)CoA **313** and *t*-CMP-oxa(dethia)CoA **314** from amino(dethia)CoA **253** or hydroxy(dethia)CoA **252** and a *t*-CMP fragment *via* a coupling step followed by deprotection. However, another strategy may be preferred due to the following limitations; the low yield of the coupling step that requires preparative HPLC purification prior to the final deprotection step hampered synthesis on a acceptable scale. Although hydroxy(dethia)CoA **252** could be produced in significant amounts from the corresponding hydroxy(dethia)pantetheine precursor, amino(dethia)CoA **253** could not be obtained from the amino(dethia)pantetheine (Scheme 5.14).

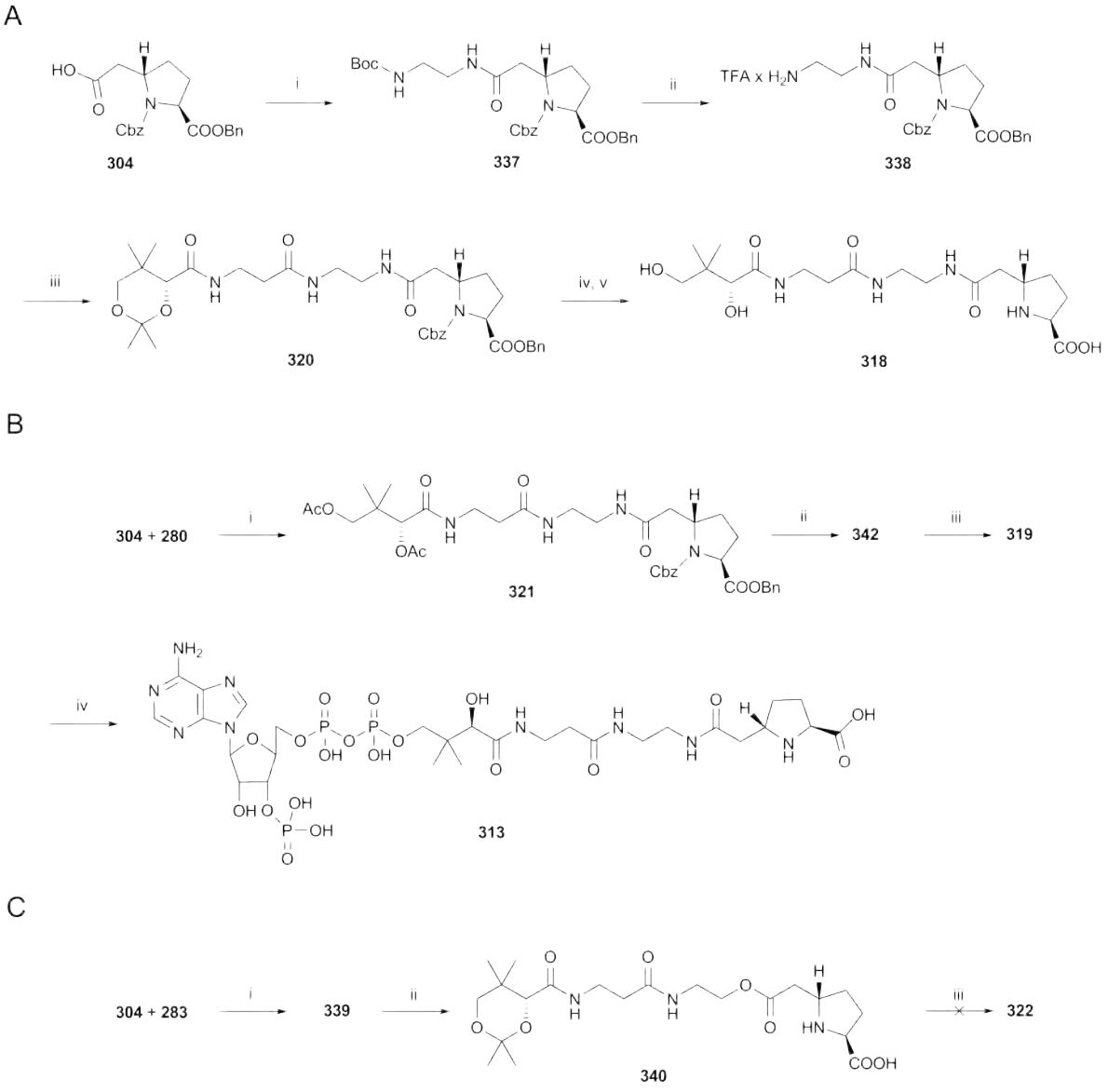
Coupling of Amino and Hydroxy(dethia)pantetheine with *t*-CMP followed by ECC

A second strategy to access non-hydrolysable *t*-CMP-CoA analogues was based on the work of Tonge and coworkers who used the ECC reaction as a last step in the synthesis of the MenB ligand *O*-succinylbenzoyl-aza(dethia)CoA (OSB-NH-CoA) (Figure 5.2 E and F) [52]. In the present work, the *t*-CMP was coupled to the amino(dethia)pantetheine or hydroxy(dethia)pantetheine and the precursor was fully deprotected before the ECC reaction (Scheme 5.16). This strategy was anticipated to overcome the limitations of the strategy described in the previous paragraph.

Two synthetic routes were investigated for the synthesis of the *t*-CMP-aza(dethia)pantetheine precursor **319** (5-(2-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-2-oxoethyl)-L-proline, Scheme 5.16 A and B). One involved the *O*,*O*′-isopropylidene-protected intermediate **320** (Scheme 5.16 A) and the other the *O*,*O*′-diacetyl-protected intermediate **321** (Scheme 5.16 B). Acidic or alkaline deprotection, followed by removal of the benzyl and carboxybenzyl *t*-CMP protecting groups under hydrogenation conditions yielded the *t*-CMP-aza(dethia)pantetheine precursor **319**. In a first set of ECC experiments, the *t*-CMP-aza(dethia)pantetheine precursor **319** was found not to be accepted by the CoA biosynthetic enzymes. This could be due to the substrate being contaminated with trifluoroacetic acid or other chemicals, which interfered with the enzymatic steps in the aqueous assay. Purification of 5-(2-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethyl]amino-2-oxoethyl)-L-proline **319** by preparative HPLC after deprotection allowed the successful completion of the ECC assay. *t*-CMP-aza(dethia)CoA **313** was obtained in moderate yield from compound **319** after overnight incubation with CoaA, CoaD and CoaE in the presence of 10 equivalents of ATP, followed by preparative HPLC purification. In contrast, *t*-CMP-oxa(dethia)CoA **314** could not be obtained using this strategy as neither the acidic or alkaline deprotection conditions left the ester linkage of *t*-CMP-oxa(dethia)pantetheine **322** (5-2-[2-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)ethoxy]-2-oxoethyl-L-proline) intact (Scheme 5.16 C).

Synthesis of Truncated *t*-CMP-CoA Analogues

Two truncated *t*-CMP-CoA analogues, 5-2-[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324**, were synthesised to be tested as CMPS ligands for crystallography. Coupling of *N*-(2-aminoethyl)-acetamide or *N*-(2-hydroxyethyl)acetamide to the protected *t*-CMP derivative **304** gave respectively dibenzyl (2*S*)-5-2-[(2-acetamidoethyl)amino]-2-oxoethylpyrrolidine-1,2-dicarboxylate **325** and dibenzyl (2*S*)-5-[2-(2-acetamidoethoxy)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate **326** in moderate yields (Scheme 5.13 C). Removal of the carboxybenzyl and benzyl protective groups under catalytic hydrogenation conditions using palladium on activated carbon gave the free compounds 5-2-[(2-acetamidoethyl)amino]-



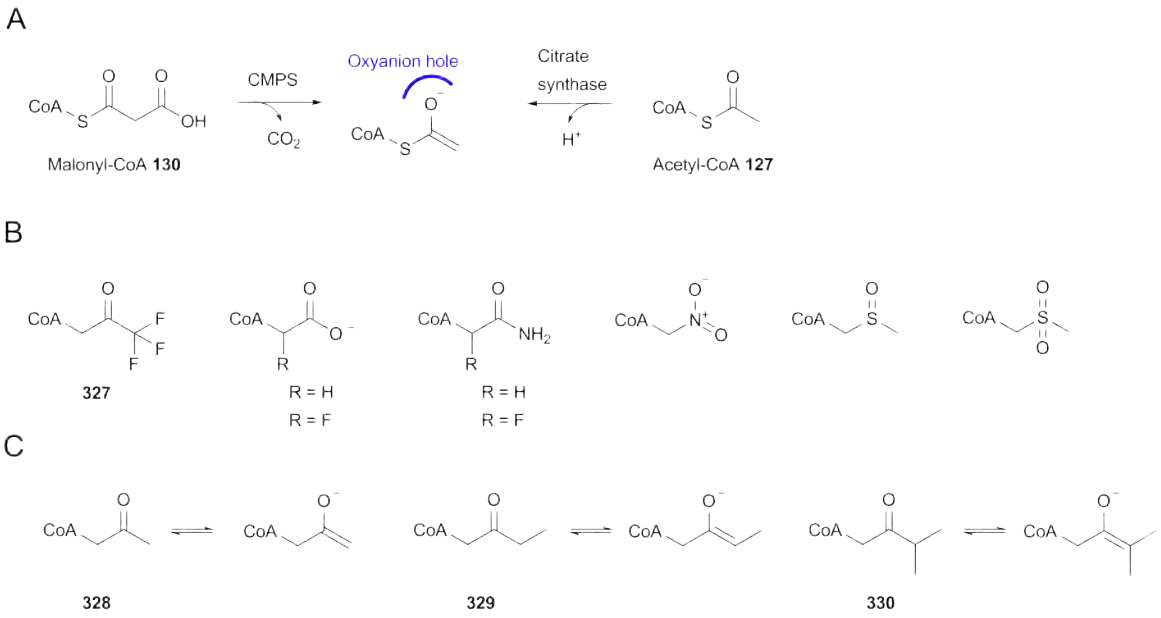
Scheme 5.16: Chemoenzymatic production of *t*-CMP-CoA analogues. (A) (i) **271**, EDC HCl, HOBt, DIPEA, rt, 24 h, 83%; (ii) CF_3COOH, rt, 6 h, app. quant.; (iii) PyBOP, Et_3N, 16 h, 58%; (iv) H_2, Pd/C, rt, 4 h, app. quant.; (v) DOWEX 50WX8-400 (H^+) resin, rt, 16 h, 99%; (B) (i) EDC HCl, HOBt, Et_3N, 0 °C–rt, 16 h, 96%; (ii) H_2, Pd/C, rt, 4 h, app. quant.; (iii) K_2CO_3, 0 °C, 5 h, 59%; (iv) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 52%; (C) (i) HBTU, DIPEA, 16 h, 41%; (ii) H_2, Pd/C, rt, 4 h, app. quant.; (iii) DOWEX 50WX8-400 (H^+) resin, rt, 12 h, n.i.

2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324** in apparent quantitative yields. A small amount of 5-2-[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324** was purified by preparative HPLC prior to use in ThnE crystallisation trials (Section 6.4).

5.2.5 Chemoenzymatic Synthesis of Enolate Analogues for Crystallographic Studies

Several stable enolate analogues have been synthesised previously, using the aminolysis strategy for biochemical studies (Section 5.1.3, Scheme 5.5 A), *e.g.* the trifluoroacetyl-CoA analogue trifluoroacetyl-carba(dethia)CoA **327** (Scheme 5.17 B) and the acetyl-CoA analogue acetyl-carba(dethia)CoA **328** (Scheme 5.17 C) [25]. Some of these compounds have been tested as citrate synthase and chloramphenicol acetyltransferase inhibitors [27]. The mechanism of citrate synthase catalysis involves deprotonation of acetyl-CoA **127** to give the corresponding enolate (Scheme 5.17 A). This enolate is identical to the oxyanion-hole stabilised CMPS catalysis intermediate formed *via* decarboxylation of malonyl-CoA (Scheme 5.17 A). The observation that methylmalonyl-CoA **212** and dimethylmalonyl-CoA **137** are CMPS substrates suggests that CarB is capable of catalysing their decarboxylation to yield the corresponding enolates [94]. Propionyl-carba(dethia)CoA **329** and isobutyryl-carba(dethia)CoA **330** are carba(dethia)CoA analogues of these enolates (Scheme 5.17 C). They contain a keto function in equilibrium with the corresponding enol tautomer. The equilibrium for propionyl-carba(dethia)CoA **329** and isobutyryl-carba(dethia)CoA **330** can be expected to be shifted more towards the enol form than for acetonyl-carba(dethia)CoA **328** due to higher substitution of the resulting double bond (Scheme 5.17 C). The equilibrium position was determined by NMR for isobutyryl-carba(dethia)CoA **330** (Figure 5.5). In the context of crystallographic studies on CMPS enzymes, it was envisaged that the propionyl-carba(dethia)CoA **329** and isobutyryl-carba(dethia)CoA **330** enol forms could be anchored into the oxyanion hole, improving the binding of the CoA cofactor, thus allowing co-crystallisation. The synthesis of propionyl-carba(dethia)CoA **329** and isobutyryl-carba(dethia)CoA **330** as mimic of the intermediates of CMPS catalysis was therefore investigated.

A route to access propionyl-carba(dethia)CoA **329** and isobutyryl-carba(dethia)CoA **330** was designed, starting from the malonyl-carba(dethia)pantetheine methyl ester **292**, as described above



Scheme 5.17: Analogues of acetyl-CoA enolate. (A) Decarboxylation of malonyl-CoA **130** in CMPS catalysis and deprotonation of acetyl-CoA **127** in citrate synthase give the same enolate intermediate; (B) Stable analogues of the enolate intermediate that have been reported and have been tested as citrate synthase and chloramphenicol acetyltransferase inhibitors. These compounds were obtained using the aminolysis strategy described by Martin and Drueckhammer [25, 27]; (C) Keto-enol tautomerism in carba(dethia) analogues of acyl-CoA derivatives.

(Scheme 5.11, Section 5.2.3). Using an alkylation step described by Leadlay and coworkers [36], the extent of alkylation could be controlled by the amounts of alkylating agent added (Scheme 5.18). In the presence of potassium carbonate, a solution of methyl 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-3-oxohexanoate **292** in acetone was treated with an excess of methyl iodide to give the dialkylated product, methyl 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-2,2-dimethyl-3-oxohexanoate **331**, in moderate yield with traces of monoalkylated product methyl 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-2-methyl-3-oxohexanoate **332**. Under the same conditions, reducing the amount of methyl iodide to 1.5 equivalents yielded traces of dialkylated product **331** with the desired monoalkylated product **332** produced in moderate yield.

Optimised conditions for the mild alkaline hydrolysis of the hydroxyl protecting groups of **332** and **331** have been described by Leadlay and coworkers, yielding methyl 6-(N-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)-2-methyl-3-oxohexanoate **333** and methyl 6-(*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanylamino)-2,2-dimethyl-3-oxohexanoate **334** (Scheme 5.18, dashed box) [36]. Raising the temperature for the hydrolysis reaction to

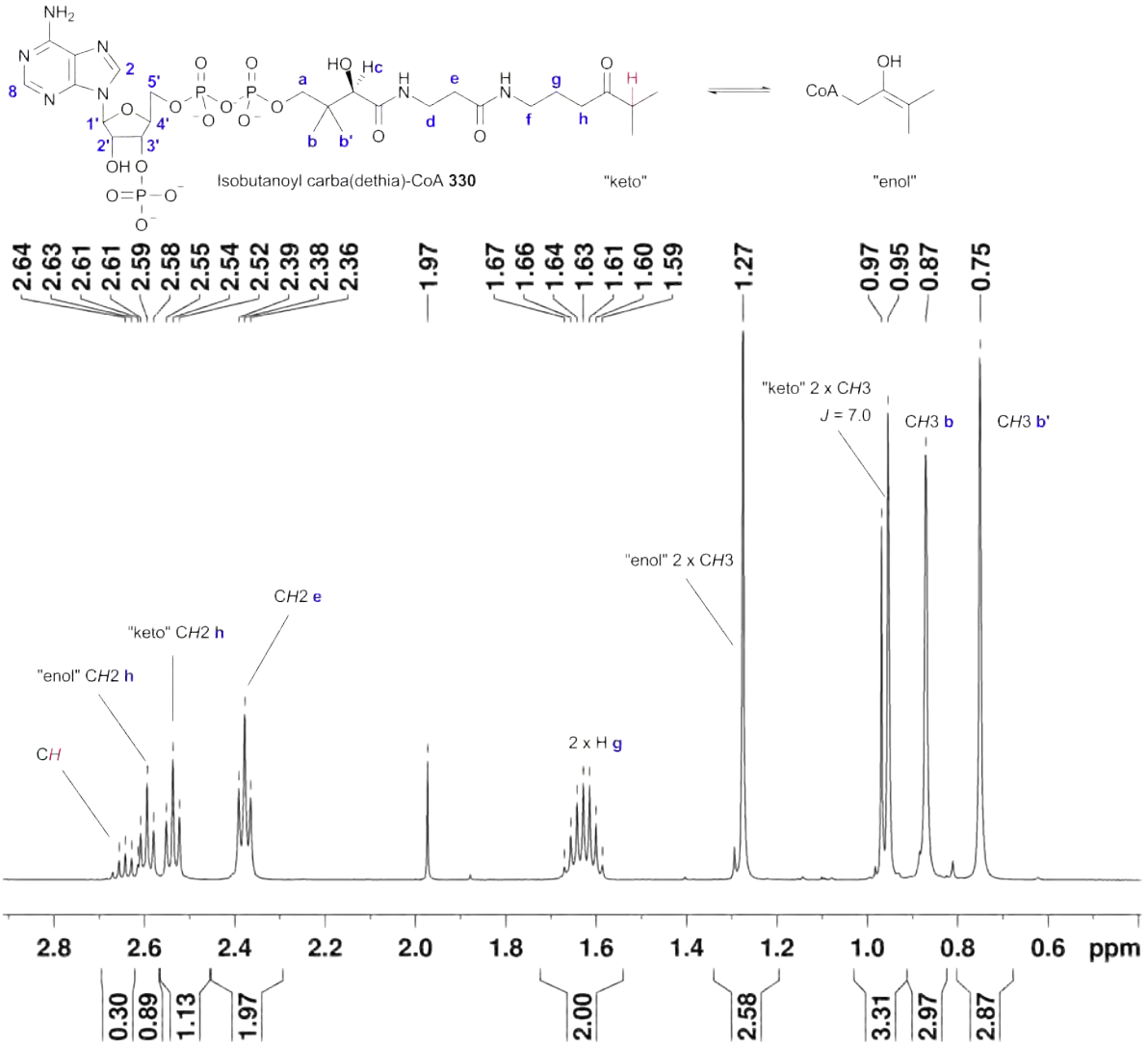
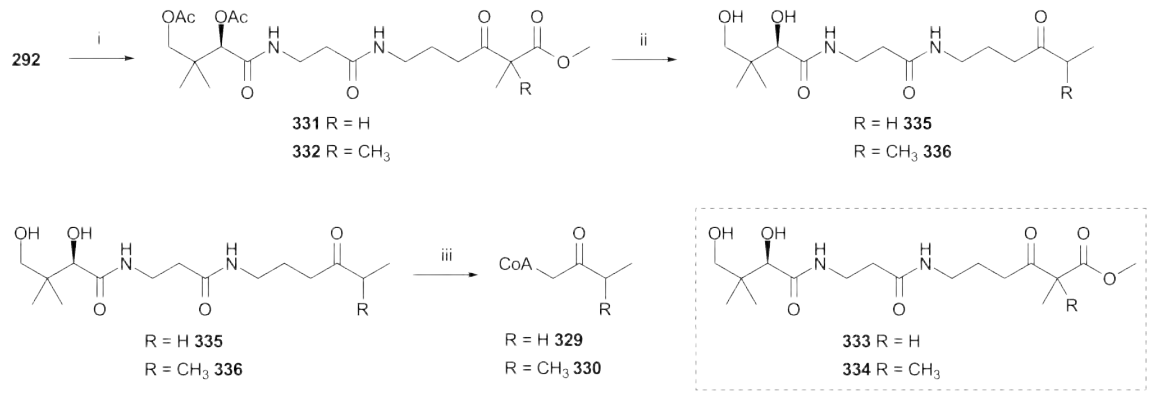


Figure 5.5: ^{1}H NMR showing keto-enol tautomerism for isobutyryl-carba(dethia)CoA. Isobutyryl-carba(dethia)CoA **330** is present as a mixture of keto and enol form in aqueous solution at a ratio of about 1.1:0.9, according to integration values for H-h protons. The spectrum was recorded at 500 MHz in D_2O. Deuterium exchange explains the low integration value for the keto C*H* signal (δ = 2.62 ppm, integration value ≅ 0.30, should be ≅ 0.45). This C*H* signal disappeared completely after 30 min at 20 °C.



Scheme 5.18: Chemoenzymatic synthesis of CoA enolate analogues. (i) MeI, K_2CO_3, $(CH_3)_2CO$, 0 °C–rt, 18 h, 34–53%; (ii) K_2CO_3, MeOH, rt, 16 h, 82–85%; (iii) CoaA, CoaD, CoaE, 50 mM Tris, 20 mM KCl, 10 mM $MgCl_2$, ATP, pH 8.0, 16 h, rt, 43–52%.

25 °C yielded the decarboxylated compound (2*R*)-2,4-dihydroxy-3,3-dimethyl-*N*-3-oxo-3-[(4-oxohexyl)amino]propylbutanamide **335** and (2*R*)-2,4-dihydroxy-*N*-3-[(4-hydroxy-5-methylhex-4-en-1-yl)amino]-3-oxopropyl-3,3-dimethylbutanamide **336**. The ketones **335** and **336** were subjected to the ECC assay in the presence of an excess of ATP to yield pure propanoyl carba(dethia)-CoA **329** and isobutyryl carba(dethia)-CoA **330** after preparative HPLC purification.

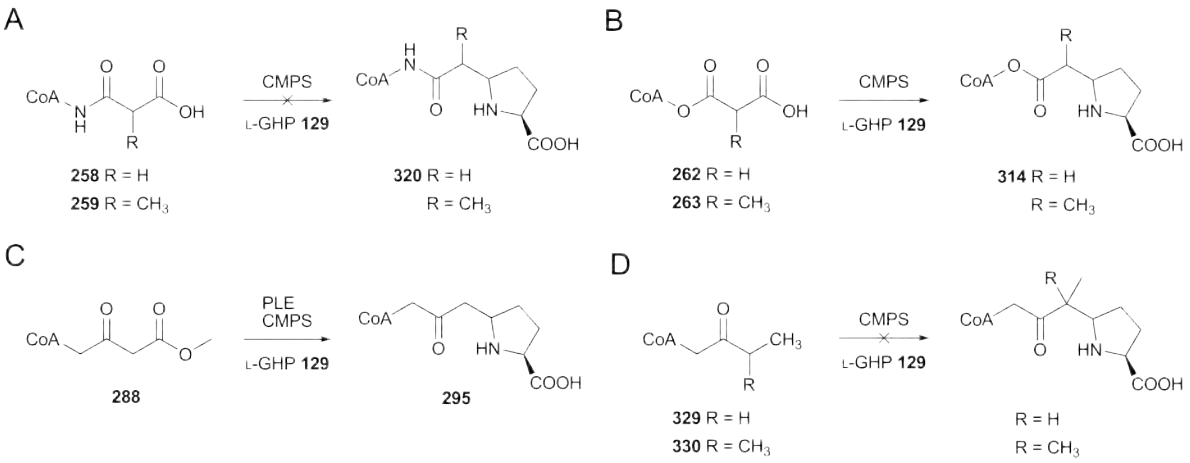
5.3 CMPS Assays using CoA Analogues

Although the existence of a *t*-CMP-CoA intermediate **132** in the CMPS reaction has been suggested based on MS analysis, no NMR data has been reported to date [95]. Several attempts to reproduce these results described in Gerratana *et al.* in our laboratory have failed, possibly because different HPLC conditions were used. Other parameters might also account for the absence of observable *t*-CMP-CoA **132**, such as the use of a different buffer for the CMPS assay. Tris buffer gave improved CMPS catalysis when compared to the phosphate buffer used by Townsend and coworkers, possibly by increasing the rate of thioester hydrolysis. It was hypothesised that malonyl-aza(dethia)CoA **258**, malonyl-oxa(dethia)CoA **262** or their C-2 methyl derivatives **259** and **263**, as well as malonyl-carba(dethia)CoA **294** would yield non-hydrolysable *t*-CMP-CoA analogues when incubated with CMPS in the presence of L-GHP **129**, enabling the characterisation of the intermediate of CMPS catalysis.

5.3.1 CMPS Accepts Malonyl-oxa(dethia)CoA as a Substrates

Seven malonyl-CoA analogues (compounds **258**, **259**, **262**, **263**, **288**, **329** and **330**) were tested in a standard CMPS assay and in the absence of malonyl-CoA **130** (Scheme 5.19). The formation of the corresponding stable *t*-CMP-CoA intermediate analogues was monitored by LC-MS analysis.

A stable *t*-CMP-CoA intermediate analogue was formed when using one of the malonyl-CoA analogues; Incubation of malonyl-oxa(dethia)CoA **262** with CarB or ThnE in the presence of L-GHP **129** yielded *t*-CMP-oxa(dethia)CoA **314**, CoAOH **252** and *t*-CMP **131**, as observed by LC-MS analysis (Figures 5.6 and 5.7). The yield for *t*-CMP **131** was found to be lower when using malonyl-oxa(dethia)CoA **262** compared to malonyl-CoA **130** (Figure 5.6 B). The difference in *t*-CMP production may be due to a slower decarboxylation rate for malonyl-oxa(dethia)CoA **262** than for malonyl-CoA **130**, and/or due to slower hydrolysis of the intermediate. Relatively higher stability of the ester towards hydrolysis may explain why the *t*-CMP-oxa(dethia)CoA **314** can be observed using LC-MS analysis while the more labile *t*-CMP-CoA **132** cannot. To allow full characterisation, the separation of *t*-CMP-oxa(dethia)CoA **314** from other hydroxy(dethia)CoA derivatives present in the assay would require further optimisation of the HPLC conditions; however, MS analysis provided evidence for the production of compound **314** (Figure 5.7).



Scheme 5.19: Seven substrate analogues were tested for the CMPS-catalysed formation of the corresponding *t*-CMP-CoA analogues. (A) Neither malonyl-aza(dethia)CoA **258** nor methylmalonyl-aza(dethia)CoA **259** were CMPS substrates, as judged by the absence of *t*-CMP-CoA formation; (B) Malonyl-oxa(dethia)CoA **262**, but not methylmalonyl-oxa(dethia)CoA **263**, was accepted as a CMPS substrate and yielded *t*-CMP-oxa(dethia)CoA **314**, CoAOH **252** and *t*-CMP **131**; (C) When malonyl-carba(dethia)CoA methyl ester **288** was coincubated with porcine liver esterase (PLE) and CarB, only low amounts of *t*-CMP-carba(dethia)CoA **295** were observed by LC-MS and the product could not be characterised; (D) Neither propanoyl carba(dethia)CoA **329** or isobutyryl carba(dethia)CoA **330** are CMPS substrates.

Malonyl-carba(dethia)CoA methyl ester **288** was hydrolysed using porcine liver esterase (PLE) to yield malonyl-carba(dethia)CoA **294**. This ester hydrolysis reaction was coupled to a CMPS assay in the presence of L-GHP **129** (Scheme 5.19 C); However, there was no compelling evidence that malonyl-carba(dethia)CoA **294** was accepted by CMPS as a substrate. *t*-CMP-carba(dethia)CoA **295** was not produced in sufficient amounts for characterisation; The low efficiency of the PLE-catalysed hydrolysis produced low quantities of malonyl-carba(dethia)CoA **294**. When malonyl-carba(dethia)CoA methyl ester **288** was incubated with PLE only, the yield of the ester hydrolysis reaction was found to be low, even after an extended period of time, as judged by HPLC (Appendix D.27).

No product formation was observed when CMPS was incubated with malonyl aza(dethia)CoA **258**, methylmalonyl-aza(dethia)CoA **259** and methylmalonyl-oxa(dethia)CoA **263** in the presence of L-GHP **129** (Scheme 5.19 A and B). Propanoyl carba(dethia)CoA **329** and isobutyryl carba(dethia)CoA **330**, two carba(dethia)-analogues of the enolate intermediate of CMPS catalysis, were also tested as CMPS substrates. They did not form any detectable amount of the expected

t-CMP-carba(dethia)CoA derivatives (Scheme 5.19 D).

5.3.2 CMPS Assays in the Presence of *t*-CMP-aza(dethia)CoA

When CMPS was incubated with *t*-CMP-aza(dethia)CoA **313** alone, no *t*-CMP formation was observed, even after 5 hours, an incubation 10 times longer than the standard CMPS assay. This observation suggests that either *t*-CMP-aza(dethia)CoA **313** cannot bind to the CMPS active site, or that the amide functionality cannot be hydrolysed by CMPS.

To test whether *t*-CMP-aza(dethia)CoA **313** can compete with malonyl-CoA binding in the CMPS active site, inhibition of the CarB-catalysed reaction was monitored. The CMPS assay was performed using malonyl-CoA **130** and 4-methyl-L-GHP **142** as substrates in the presence of increasing concentrations of the non-hydrolysable *t*-CMP-CoA analogue **313**. Using identical conditions, but shorter reaction times compared to standard assay, the formation of 4-methyl-*t*-CMP was followed by LC-MS analysis using *p*-aminosalicylic acid as an internal standard. The CMPS-catalysed formation of both (4*S*)-4-methyl-*t*-CMP **223** and (4*R*)-4-methyl-*t*-CMP **224** was found to be unaltered by addition of up to 10 equivalents of *t*-CMP-aza(dethia)CoA **313** to the natural substrate malonyl-CoA **130** (Figure 5.8). Even though the experiments carried out cannot exclude that *t*-CMP-aza(dethia)CoA **313** binds to CMPS at all, the results in hand suggest that it probably does not have a high binding affinity for ThnE.

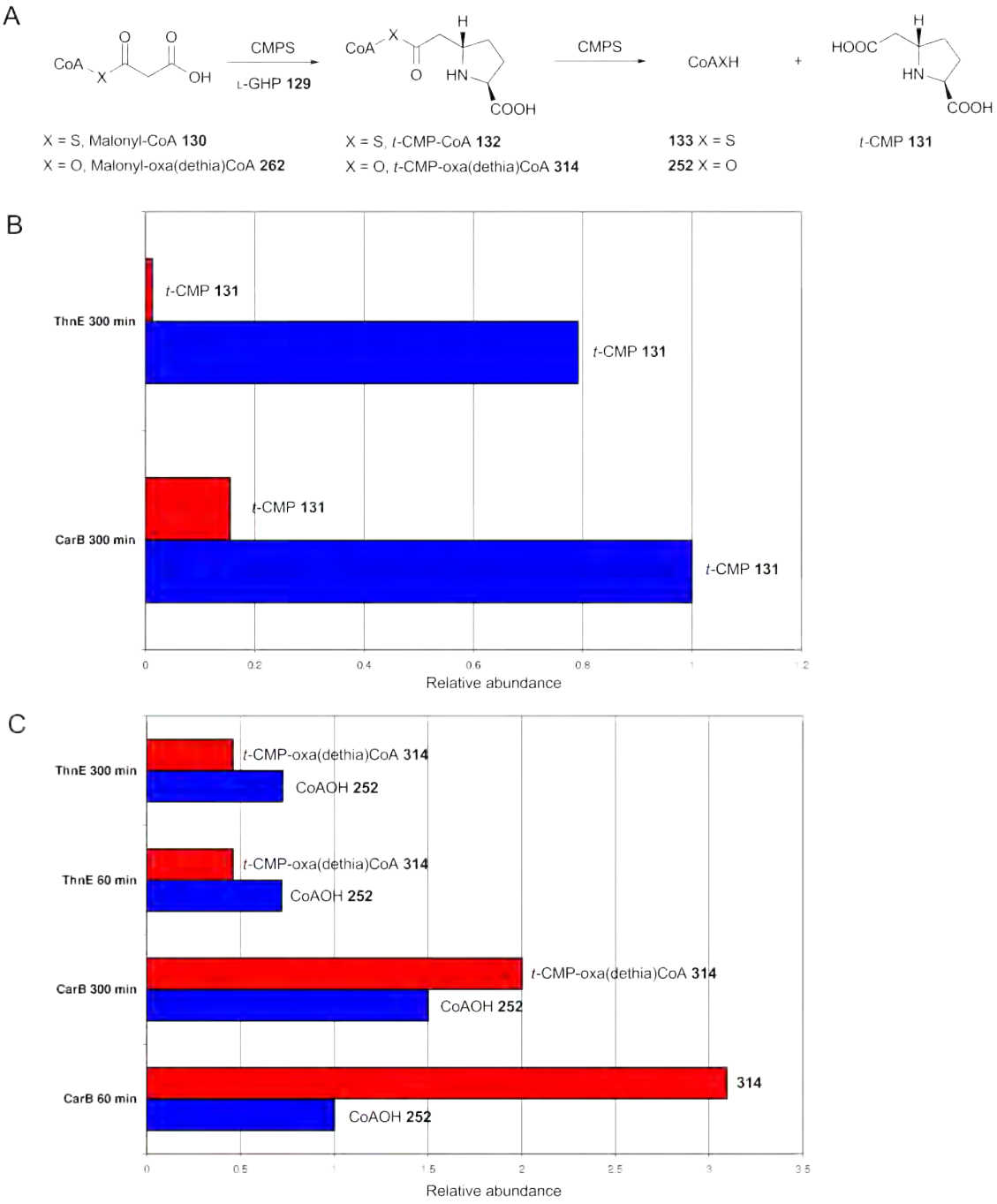


Figure 5.6: Comparison between malonyl-CoA **130** and malonyl-oxa(dethia)CoA **262** as substrates for the CMPS-catalysed production of *t*-CMP. The standard CMPS assay was performed in the presence of L-GHP **129**. Formation of *t*-CMP **131** and CoA derivatives was monitored using LC-MS analysis. (A) General scheme for the CMPS-catalysed reaction in the presence of L-GHP **129** and malonyl-CoA **130** or malonyl-oxa(dethia)CoA **262**. (B) Quantification of *t*-CMP **131** formation for the CarB and ThnE reactions using L-GHP and malonyl-CoA **130** (blue bars) or malonyl-oxa(dethia)CoA **262** (red bars) after 5 h incubation. (C) Comparison between CarB and ThnE-catalysed formation of *t*-CMP-oxa(dethia)CoA **314** (red bars) and further hydrolysis into hydroxy(dethia)CoA **252** (CoAOH) (blue bars) after 1 h and 5 h incubation. The relative abundance of the two species should not be compared as their ionisation efficiency is probably different.

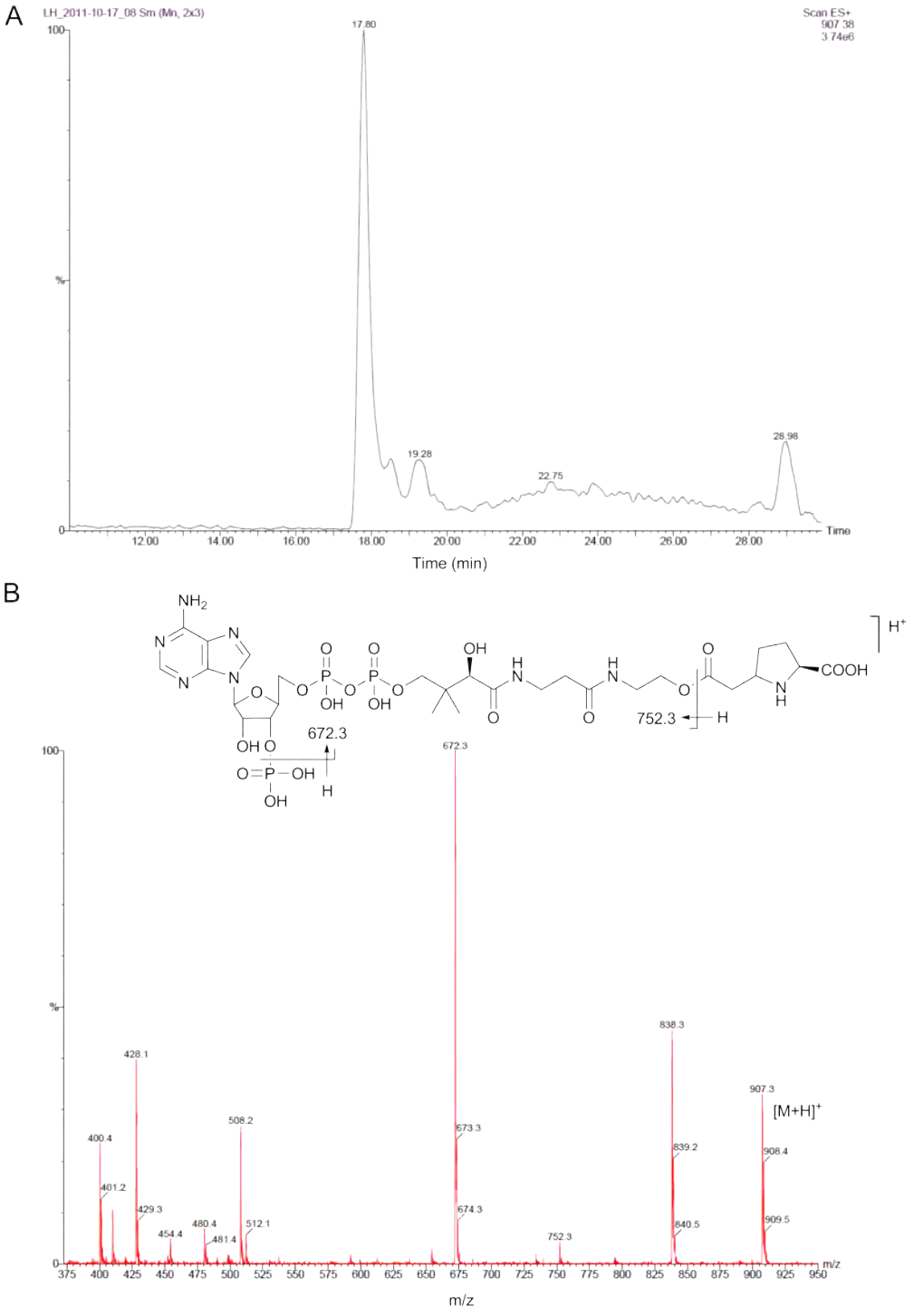


Figure 5.7: CMPS accepts malonyl-oxa(dethia)CoA **262** as substrate and yields *t*-CMP-oxa(dethia)CoA. The formation of *t*-CMP-oxa(dethia)CoA **314** was monitored using LC-MS analysis. (A) LC-MS extracted ion chromatogram shows low but significant formation of *t*-CMP-oxa(dethia)CoA **314** (R_t = 17.8 min, m/z = 907.3); (B) Mass spectrum for the *t*-CMP-oxa(dethia)CoA **314** chromatogram peak (R_t = 17.8 min). The molecular ion (m/z = 907.3) is visible as well as the fragments corresponding to the loss of the *t*-CMP moiety (m/z = 752.3) and further loss of the 3′-phosphate (m/z = 672.3).

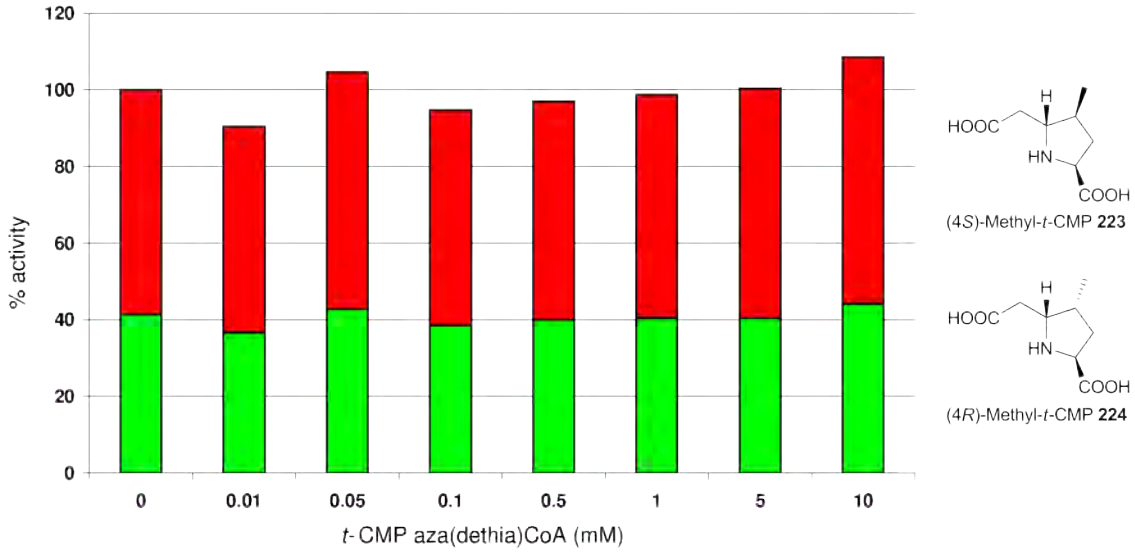


Figure 5.8: *t*-CMP-aza(dethia)CoA does not inhibit CMPS formation of 4-methyl-*t*-CMP at low concentrations. The amount of (4*S*)-4-methyl-*t*-CMP **223** (red bars) and (4*R*)-4-methyl-*t*-CMP **224** (green bars) was constant over a range of *t*-CMP-aza(dethia)CoA **313** concentrations. The CMPS inhibition assays were carried out for 5 min in 450 mM Tris (pH 7.9) with increasing concentrations of *t*-CMP-aza(dethia)CoA **313** in the presence of 2 mM 4-methyl-L-GHP and 1 mM malonyl-CoA **130**. *p*-Aminosalicylic acid was added as an internal standard for quantification when quenching the assay.

5.4 Summary for Chapter 5

5.4.1 Stable Acyl-CoA Derivatives can be Accessed using a Chemoenzymatic Strategy

A series of stable acyl-CoA analogues were obtained as described in Section 5.2, *via* an enzyme cascade catalysis (ECC) strategy, using three enzymes from the CoA biosynthesis pathway. CoaA, CoaD and CoaE from *E. coli* had been previously shown to demonstrate substrate promiscuity and had therefore be used to generate a diversity of acyl-CoA derivatives. CoaA, CoaD and CoaE were obtained separately *via* recombinant expression in *E. coli* and purified using a one-step affinity protocol. The pantetheine analogue precursors were synthesised from D-pantothenate and subjected to the ECC reaction in the presence of an excess of ATP. Preparative HPLC purification yielded the CoA analogues which were characterised by mass spectrometry and NMR spectroscopy. Two malonyl-CoA and two methylmalonyl-CoA analogues of which the sulphur atom was replaced by with a nitrogen or an oxygen atom were obtained using this procedure. Two routes towards analogues of the proposed *t*-CMP-CoA **132** intermediate of CMPA catalysis were investigated. While the *t*-CMP-aza(dethia)CoA **313** was obtained *via* the synthesis of the corresponding *t*-CMP-aza(dethia)pantetheine precursor followed by enzymatic conversion to the final CoA analogue, attempts to isolate *t*-CMP-oxa(dethia)CoA **314** using the same strategy was unsuccessful. In addition, two additional stable analogues, propanoyl carba(dethia)-CoA **329** and isobutyryl carba(dethia)-CoA **330**, as well as two truncated *t*-CMP-CoA analogues, 5-2-[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324**, were synthesised to be tested in co-crystallisation experiments with *S. cattleya* ThnE.

5.4.2 From all Malonyl-CoA Analogues Tested, only Malonyl-oxa(dethia)CoA is a CMPS Substrate

The malonyl-CoA analogues described in Section 5.2 were tested as CMPS substrates. When malonyl-aza(dethia)CoA **258**, malonyl-oxa(dethia)CoA **262**, or the C-2 methyl derivatives **259** and **263** were tested as CMPS substrates, only malonyl-oxa(dethia)CoA **262** was found to react to give the corresponding *t*-CMP-oxa(dethia)CoA **314** as well as the hydrolysis products CoAOH **252** and *t*-CMP **131**. The formation of these three products supports a stepwise mechanism involving C–C

bond formation followed by hydrolysis. Malonyl-carba(dethia)CoA **294** could be generated *in situ* using PLE-catalysed hydrolysis of malonyl-carba(dethia)CoA methyl ester **288** (Scheme 5.19 C). Malonyl-carba(dethia)CoA **294** was found to yield low amounts of *t*-CMP-carba(dethia)CoA **295** as judged by LC-MS analysis. It is somehow surprising that compound **294** is not a better substrate, but the low amounts of product formed may be attributed to the low efficiency of the PLE-catalysed ester hydrolysis. However, the malonyl-CoA analogues described in Section 5.2 could not be tested as CMPS inhibitors due to the limitations of the synthetic methodology. Larger amounts would be needed in order to perform inhibition studies. Measuring the CMPS activity in the presence of these analogues could provide an understanding of their binding properties and explain their inability to co-crystallise with ThnE.

t-CMP-aza(dethia)CoA is Neither a CMPS Substrate Nor an Inhibitor

When *t*-CMP-aza(dethia)CoA **314** was incubated in the presence of CarB or ThnE, *t*-CMP formation was not observed, suggesting that these enzymes cannot hydrolyse the amide bond *t*-CMP-aza(dethia)CoA **314**. Alternatively, *t*-CMP-aza(dethia)CoA **314** might not bind for conformational reasons. The inability for an analogue of the intermediate of catalysis to bind to the CMPS active site would explain the absence of CMPS inhibition by compound **314** when it was added to a standard CMPS assay, as well as the absence of CMPS crystals in complex with compound **314** when co-crystallisation experiments were carried out (Chapter 6).

References

[1] Begley, T. P., Kinsland, C., and Strauss, E. (2001) The biosynthesis of coenzyme A in bacteria. *Vitam Horm*, **61**, 157–171.

[2] Cebrat, M., Kim, C. M., Thompson, P. R., Daugherty, M., and Cole, P. A. (2003) Synthesis and analysis of potential prodrugs of coenzyme A analogues for the inhibition of the histone acetyltransferase p300. *Bioorg Med Chem*, **11**, 3307–3313.

[3] Patel, S. S. and Walt, D. R. (1987) Substrate specificity of acetyl coenzyme A synthetase. *J Biol Chem*, **262**, 7132–7134.

[4] Majerus, P. W., Alberts, A. W., and Vagelos, P. R. (1965) Acyl carrier protein. IV. The identification of 4'-phosphopantethein as the prosthetic group of the acyl carrier protein. *Proc Natl Acad Sci U S A*, **53**, 410–417.

[5] Voet, D. and Voet, J. G. (2008) *Biochemistry*. John Wiley & Sons.

[6] Leonardi, R., Zhang, Y.-M., Rock, C. O., and Jackowski, S. (2005) Coenzyme A: Back in action. *Prog Lipid Res*, **44**, 125–153.

[7] Webb, M. E., Smith, A. G., and Abell, C. (2004) Biosynthesis of pantothenate. *Nat Prod Rep*, **21**, 695–721.

[8] Coxon, K. M., Chakauya, E., Ottenhof, H. H., Whitney, H. M., Blundell, T. L., Abell, C., and Smith, A. G. (2005) Pantothenate biosynthesis in higher plants. *Biochem Soc Trans*, **33**, 743–746.

[9] Hoagland, M. B. and Novelli, G. D. (1954) Biosynthesis of coenzyme A from phosphopantetheine and of pantetheine from pantothenate. *J Biol Chem*, **207**, 767–773.

[10] Levintow, L. and Novelli, G. (1954) The synthesis of coenzyme A from panthetheine: preparation and properties of panthetheine kinase. *J Biol Chem*, **207**, 761–765.

[11] Spry, C., Kirk, K., and Saliba, K. J. (2008) Coenzyme A biosynthesis: An antimicrobial drug target. *FEMS Microbiol Rev*, **32**, 56–106.

[12] Gordon, N. (2002) Pantothenate kinase-associated neurodegeneration (Hallervorden-Spatz syndrome). *Eur J Paediatric Neurology*, **6**, 243–247.

[13] Mishra, P. K. and Drueckhammer, D. G. (2000) Coenzyme A analogues and derivatives: Synthesis and applications as mechanistic probes of coenzyme A ester-utilizing enzymes. *Chem Rev*, **100**, 3283–3310.

[14] Chase, J. F., Middleton, B., and Tubbs, P. K. (1966) A coenzyme A analogue, desulpho-CoA; Preparation and effects on various enzymes. *Biochem Biophys Res Commun*, **23**, 208–213.

[15] Moffatt, J. G. and Khorana, H. G. (1961) Nucleoside polyphosphates. XII. The total synthesis of coenzyme A. *J Am Chem Soc*, **83**, 663–675.

[16] Michelson, A. (1964) Synthesis of coenzyme A. *Biochim Biophys Acta*, **93**, 71–77.

[17] Mautner, H. G. (1970) Synthesis of coenzyme A analogs. *Methods Enzymol*, **18**, 338–350.

[18] Günther, W. H. H. and Mautner, H. G. (1965) The synthesis of selenocoenzyme A. *J Am Chem Soc*, **87**, 2708–2716.

[19] Miller, T. L., Rowley, G. L., and Stewart, C. J. (1966) Coenzyme A analogs. Synthesis of D-oxypantetheine-4' phosphate and oxy-coenzyme A. *J Am Chem Soc*, **88**, 2299–2304.

[20] Zhao, Y., Michenfelder, M., and Rétey, J. (1994) Synthesis, characterization, and enzymic conversion of nonhydrolysable analogues of propionylcoenzyme A. *Can J Chem*, **72**, 164–169.

[21] Lipmann, F., Kaplan, N. O., Novelli, G. D., Tuttle, L. C., and Guirard, B. M. (1950) Isolation of coenzyme A. *J Biol Chem*, **186**, 235–243.

[22] Baddiley, J., Thain, E. M., Novelli, G. D., and Lipmann, F. (1953) Structure of coenzyme A. *Nature*, **171**, 76.

[23] Stewart, C. J. and Ball, W. J. (1966) Coenzyme A analogs. II. enzymatic conversion of D-oxypantetheine 4'-phosphate to oxy-coenzyme A. *Biochemistry*, **5**, 3883–3886.

[24] Stewart, C. J., Thomas, J. O., Ball, W. J., and Aguirre, A. R. (1968) Coenzyme A analogs. III. Chemical synthesis of desulfopantetheine 4'-phosphate and its enzymic conversion to desulfo-coenzyme A. *J Am Chem Soc*, **90**, 5000–5004.

[25] Martin, D. P. and Drueckhammer, D. G. (1992) Combined chemical and enzymic synthesis of coenzyme A analogs. *J Am Chem Soc*, **114**, 7287–7288.

[26] Schwartz, B. and Drueckhammer, D. G. (1996) A stereochemical probe of the tetrahedral intermediate in the reactions of acetyl-coenzyme A dependent acetyltransferases. *J Am Chem Soc*, **118**, 9826–9830.

[27] Martin, D. P., Bibart, R. T., and Drueckhammer, D. G. (1994) Synthesis of novel analogs of acetyl coenzyme A: Mimics of enzyme reaction intermediates. *J Am Chem Soc*, **116**, 4660–4668.

[28] Nazi, I., Koteva, K. P., and Wright, G. D. (2004) One-pot chemoenzymatic preparation of coenzyme A analogues. *Anal Biochem*, **324**, 100–105.

[29] Strauss, E. and Begley, T. P. (2002) The antibiotic activity of *N*-pentylpantothenamide results from its conversion to ethyldethia-coenzyme A, a coenzyme A antimetabolite. *J Biol Chem*, **277**, 48205–48209.

[30] Worthington, A. S. and Burkart, M. D. (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org Biomol Chem*, **4**, 44–46.

[31] Stewart, C. J. and Miller, T. L. (1965) Oxy-coenzyme A: A competitive inhibitor of coenzyme A in the phosphotransacetylase reaction. *Biochem Biophys Res Comm*, **20**, 433–438.

[32] Fischbach, M. A. and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem Rev*, **106**, 3468–3496.

[33] Chan, Y. A., Podevels, A. M., Kevany, B. M., and Thomas, M. G. (2009) Biosynthesis of polyketide synthase extender units. *Nat Prod Rep*, **26**, 90–114.

[34] Spiteller, D., Waterman, C. L., and Spencer, J. B. (2005) A method for trapping intermediates of polyketide biosynthesis with a nonhydrolyzable malonyl-coenzyme A analogue. *Angew Chem Int Ed Engl*, **44**, 7079–7082.

[35] Tosin, M., Spiteller, D., and Spencer, J. B. (2009) Malonyl carba(dethia)- and malonyl oxa(dethia)-coenzyme A as tools for trapping polyketide intermediates. *ChemBioChem*, **10**, 1714–1723.

[36] Tosin, M., Betancor, L., Stephens, E., Li, W. M. A., Spencer, J. B., and Leadlay, P. F. (2010) Synthetic chain terminators off-load intermediates from a type I polyketide synthase. *ChemBioChem*, **11**, 539–546.

[37] Whitty, A., Fierke, C. A., and Jencks, W. P. (1995) Role of binding energy with coenzyme A in catalysis by 3-oxoacid coenzyme A transferase. *Biochemistry*, **34**, 11678–11689.

[38] Fierke, C. A. and Jencks, W. P. (1986) Two functional domains of coenzyme A activate catalysis by coenzyme A transferase. Pantetheine and adenosine 3'-phosphate 5'-diphosphate. *J Biol Chem*, **261**, 7603–7606.

[39] Sumper, M. and Lynen, F. (1972) Substrate specificity of fatty acid synthetase from yeast. *FEBS Lett*, **28**, 142–144.

[40] Pohl, N. L., Gokhale, R. S., Cane, D. E., and Khosla, C. (1998) Synthesis and incorporation of an *N*-acetylcysteamine analogue of methylmalonyl-CoA by a modular polyketide synthase. *J Am Chem Soc*, **120**, 11206–11207.

[41] Jackowski, S. and Alix, J. H. (1990) Cloning, sequence, and expression of the pantothenate permease (panF) gene of *Escherichia coli*. *J Bacteriol*, **172**, 3842–3848.

[42] Davies, J. and Wright, G. D. (1997) Bacterial resistance to aminoglycoside antibiotics. *Trends Microbiol*, **5**, 234–240.

[43] Vong, K., Tam, I. S., Yan, X., and Auclair, K. (2012) Inhibitors of aminoglycoside resistance activated in cells. *ACS Chem Biol*.

[44] Meier, J. L., Mercer, A. C., Rivera, H., and Burkart, M. D. (2006) Synthesis and evaluation of bioorthogonal pantetheine analogues for *in vivo* protein modification. *J Am Chem Soc*, **128**, 12174–12184.

[45] Clarke, K. M., Mercer, A. C., Clair, J. J. L., and Burkart, M. D. (2005) *In vivo* reporter labeling of proteins via metabolic delivery of coenzyme A analogues. *J Am Chem Soc*, **127**, 11234–11235.

[46] Mercer, A. C., Meier, J. L., Hur, G. H., Smith, A. R., and Burkart, M. D. (2008) Antibiotic evaluation and *in vivo* analysis of alkynyl coenzyme A antimetabolites in *Escherichia coli*. *Bioorg Med Chem Lett*, **18**, 5991–5994.

[47] Modis, Y. and Wierenga, R. (1998) Two crystal structures of *N*-acetyltransferases reveal a new fold for CoA-dependent enzymes. *Structure*, **6**, 1345–1350.

[48] Benning, M. M., Wesenberg, G., Liu, R., Taylor, K. L., Dunaway-Mariano, D., and Holden, H. M. (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. strain CBS-3. *J Biol Chem*, **273**, 33572–33579.

[49] Dias, M. V. B., Huang, F., Chirgadze, D. Y., Tosin, M., Spiteller, D., Dry, E. F. V., Leadlay, P. F., Spencer, J. B., and Blundell, T. L. (2010) Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK. *J Biol Chem*, **285**, 22495–22504.

[50] Benning, M. M., Haller, T., Gerlt, J. A., and Holden, H. M. (2000) New reactions in the crotonase superfamily: Structure of methylmalonyl CoA decarboxylase from *Escherichia coli*. *Biochemistry*, **39**, 4630–4639.

[51] Widboom, P. F., Fielding, E. N., Liu, Y., and Bruner, S. D. (2007) Structural basis for cofactor-independent dioxygenation in vancomycin biosynthesis. *Nature*, **447**, 342–345.

[52] Li, H.-J., Li, X., Liu, N., Zhang, H., Truglio, J. J., Mishra, S., Kisker, C., Garcia-Diaz, M., and Tonge, P. J. (2011) Mechanism of the intramolecular Claisen condensation reaction catalyzed by MenB, a crotonase superfamily member. *Biochemistry*, **50**, 9532–9544.

[53] Schrödinger, LLC (2010) The PyMOL molecular graphics system, version 1.3r1.

[54] Liu, X., Wang, L., Zhao, K., Thompson, P. R., Hwang, Y., Marmorstein, R., and Cole, P. A. (2008) The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature*, **451**, 846–850.

[55] Poux, A. N., Cebrat, M., Kim, C. M., Cole, P. A., and Marmorstein, R. (2002) Structure of the GCN5 histone acetyltransferase bound to a bisubstrate inhibitor. *Proc Natl Acad Sci U S A*, **99**, 14065–14070.

[56] Hickman, A. B., Namboodiri, M. A., Klein, D. C., and Dyda, F. (1999) The structural basis of ordered substrate binding by serotonin *N*-acetyltransferase: enzyme complex at 1.8 $\mathring{a}$ resolution with a bisubstrate analog. *Cell*, **97**, 361–369.

[57] Schwartz, B., Drueckhammer, D. G., Usher, K. C., and Remington, S. J. (1995) α-fluoro acid and α-fluoro amide analogs of acetyl-CoA as inhibitors of citrate synthase: Effect of pKa matching on binding affinity and hydrogen bond length. *Biochemistry*, **34**, 15459–15466, pMID: 7492547.

[58] Thomä, N. H., Meier, T. W., Evans, P. R., and Leadlay, P. F. (1998) Stabilization of radical intermediates by an active-site tyrosine residue in methylmalonyl-CoA mutase. *Biochemistry*, **37**, 14386–14393.

[59] Beaman, T. W., Vogel, K. W., Drueckhammer, D. G., Blanchard, J. S., and Roderick, S. L. (2002) Acyl group specificity at the active site of tetrahydridipicolinate n-succinyltransferase. *Protein Sci*, **11**, 974–979.

[60] Scheibner, K. A., Angelis, J. D., Burley, S. K., and Cole, P. A. (2002) Investigation of the roles of catalytic residues in serotonin *N*-acetyltransferase. *J Biol Chem*, **277**, 18118–18126.

[61] Kim, A.-R., Rylett, R. J., and Shilton, B. H. (2006) Substrate binding and catalytic mechanism of human choline acetyltransferase. *Biochemistry*, **45**, 14621–14631.

[62] Lee, H. J., Raki, B., Gilbert, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. J. (2009) Structural and kinetic characterizations of the polysialic acid *O*-acetyltransferase OatWY from *Neisseria meningitidis*. *J Biol Chem*, **284**, 24501–24511.

[63] Hegde, S. S., Chandler, J., Vetting, M. W., Yu, M., and Blanchard, J. S. (2007) Mechanistic and structural analysis of human spermidine/spermine N^1-acetyltransferase. *Biochemistry*, **46**, 7187–7195.

[64] Reger, A. S., Wu, R., Dunaway-Mariano, D., and Gulick, A. M. (2008) Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry*, **47**, 8016–8025.

[65] Mancia, F., Smith, G. A., and Evans, P. R. (1999) Crystal structure of substrate complexes of methylmalonyl-CoA mutase. *Biochemistry*, **38**, 7999–8005.

[66] Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H. M. (2003) The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU. *J Biol Chem*, **278**, 43709–43716.

[67] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403–410.

[68] Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, **182**, 319–326.

[69] Kurihara, T., Matsubara, Y., Harusawa, S., and Yoneda, R. (1991) Chemistry of ethyl 2-[*t*-butoxycarbonyl(methyl)amino]methyl-3-hydroxy-3-phenyl (or 3-vinyl)propionate: mechanistic considerations in the formation of tetrahydro-1,3-oxazin-2-ones. *J Chem Soc, Perkin Trans 1*, **1991**, 3177–3183.

[70] Mandel, A. L., La Clair, J. J., and Burkart, M. D. (2004) Modular synthesis of pantetheine and phosphopantetheine. *Org Lett*, **6**, 4801–4803.

[71] Patrone, J. D., Yao, J., Scott, N. E., and Dotson, G. D. (2009) Selective inhibitors of bacterial phosphopantothenoylcysteine synthetase. *J Am Chem Soc*, **131**, 16340–16341.

[72] Dai, M., Feng, Y., and Tonge, P. J. (2001) Synthesis of crotonyl-oxyCoA: A mechanistic probe of the reaction catalyzed by enoyl-CoA hydratase. *J Am Chem Soc*, **123**, 506–507.

[73] Callahan, J. F., Ashton-Shue, D., Bryan, H. G., Bryan, W. M., Heckman, G. D., Kinter, L. B., McDonald, J. E., Moore, M. L., and Schmidt, D. B. a. (1989) Structure-activity relationships of novel vasopressin antagonists containing C-terminal diaminoalkanes and (aminoalkyl)guanidines. *J Med Chem*, **32**, 391–396.

[74] Lowe, A. J., Dyson, G. A., and Pfeffer, F. M. (2008) Factors influencing anion binding stoichiometry: The subtle influence of electronic effects. *Eur J Org Chem*, **2008**, 1559–1567.

[75] Krapcho, A. P. and Kuell, C. S. (1990) Mono-protected diamines. *N-tert*-Butoxycarbonyl-α,ω-alkanediamines from α,ω-alkanediamines. *Synth Comm*, **20**, 2559–2564.

[76] Niwayama, S., Cho, H., and Lin, C. (2008) Highly efficient selective monohydrolysis of dialkyl malonates and their derivatives. *Tetrahedron Lett*, **49**, 4434–4436.

[77] Page, P. C. B., Moore, J. P., Mansfield, I., McKenzie, M. J., Bowler, W. B., and Gallagher, J. A. (2001) Synthesis of bone-targeted oestrogenic compounds for the inhibition of bone resorption. *Tetrahedron*, **57**, 1837–1847.

[78] Navarro, I., Basset, J.-F., Hebbe, S., Major, S. M., Werner, T., Howsham, C., Brackow, J., and Barrett, A. G. M. (2008) Biomimetic synthesis of resorcylate natural products utilizing late stage aromatization: Concise total syntheses of the marine antifungal agents 15G256ι and 15G256β. *J Am Chem Soc*, **130**, 10293–10298.

[79] Tran, L., Broadhurst, R. W., Tosin, M., Cavalli, A., and Weissman, K. J. (2010) Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. *Chem Biol*, **17**, 705–716.

[80] Oikawa, Y., Sugano, K., and Yonemitsu, O. (1978) Meldrum's acid in organic synthesis. 2. A general and versatile synthesis of β-keto esters. *J Org Chem*, **43**, 2087–2088.

[81] Bycroft, B. W., Chhabra, S. R., Kellam, B., and Smith, P. (2003) Convenient syntheses of (3*S*,5*S*)-carbapenam-3-carboxylates and their biosynthetic relevance. *Tetrahedron Lett*, **44**, 973–976.

[82] Roth, M., Dubs, P., Gotschi, E., and Eschenmoser, A. (1971) Sulfidkontraktion via alkylative Kupplung: Eine Methode zur Darstellung von β-Dicarbonylderivaten. Uber synthetische Methoden, 1. Mitteilung. *Helv Chim Acta*, **54**, 710–734.

[83] Petersen, J. S., Fels, G., and Rapoport, H. (1984) Chirospecific syntheses of (+)- and (-)-anatoxin A. *J Am Chem Soc*, **106**, 4539–4547.

[84] Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Inhibition and alternate substrate studies on the mechanism of carbapenam synthetase from *Erwinia carotovora*. *Biochemistry*, **42**, 7836–7847.

[85] Collado, I., Ezquerra, J., Vaquero, J. J., and Pedregal, C. (1994) Diastereoselective functionalization of 5-hydroxy prolinates by tandem Horner-Emmons-Michael reaction. *Tetrahedron Lett*, **35**, 8037–8040.

[86] Stapon, A., Li, R., and Townsend, C. A. (2003) Carbapenem biosynthesis: Confirmation of stereochemical assignments and the role of CarC in the ring stereoinversion process from L-proline. *J Am Chem Soc*, **125**, 8486–8493.

[87] August, R. A., Khan, J. A., Moody, C. M., and Young, D. W. (1996) Stereospecific synthesis of (2*S*,4*R*)-[5,5,5-$^2H_3$]-leucine. *J Chem Soc, Perkin Trans 1*, **1996**, 507–514.

[88] Suaifan, G. A. R. Y., Arafat, T., and Threadgill, M. D. (2007) Synthetic approaches to peptides containing the L-gln- L-val- D(*S*)-Dmt motif. *Bioorg Med Chem*, **15**, 3474–3488.

[89] Bernardi, F., Garavelli, M., Scatizzi, M., Tomasini, C., Trigari, V., Crisma, M., Formaggio, F., Peggion, C., and Toniolo, C. (2002) Pseudopeptide foldamers: The homo-oligomers of pyroglutamic acid. *Chem Eur J*, **8**, 2516–2525.

[90] Tarver, J. E., Terranova, K. M., and Joullié, M. M. (2004) Hetero-Diels-Alder and pyroglutamate approaches to (2*S*,4*R*)-2-methylamino-5-hydroxy-4-methylpentanoic acid. *Tetrahedron*, **60**, 10277–10284.

[91] Hungate, R. W., Chen, J. L., Starbuck, K. E., Macaluso, S. A., and Rubino, R. S. (1996) New cytochalasins: Synthetic studies of a novel HIV-1 protease inhibitor. *Tetrahedron Lett*, **37**, 4113–4116.

[92] Strieker, M., Nolan, E. M., Walsh, C. T., and Marahiel, M. A. (2009) Stereospecific synthesis of *threo*- and *erythro*-β-hydroxyglutamic acid during kutzneride biosynthesis. *J Am Chem Soc*, **131**, 13523–13530.

[93] Liu, Y. and Bruner, S. D. (2007) Rational manipulation of carrier-domain geometry in nonribosomal peptide synthetases. *ChemBioChem*, **8**, 617–621.

[94] Batchelar, E. T., Hamed, R. B., Ducho, C., Claridge, T. D. W., Edelmann, M. J., Kessler, B., and Schofield, C. J. (2008) Thioester hydrolysis and C–C bond formation by carboxymethylproline synthase from the crotonase superfamily. *Angew Chem Int Ed Engl*, **47**, 9322–9325.

[95] Gerratana, B., Arnett, S. O., Stapon, A., and Townsend, C. A. (2004) Carboxymethylproline synthase from *Pectobacterium carotorova*: A multifaceted member of the crotonase superfamily. *Biochemistry*, **43**, 15936–15945.

[96] Gottlieb, H. E., Kotlyar, V., and Nudelman, A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *J Org Chem*, **62**, 7512–7515.

[97] Crawford, J. M., Dancy, B. C. R., Hill, E. A., Udwary, D. W., and Townsend, C. A. (2006) Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc Natl Acad Sci U S A*, **103**, 16728–16733.

[98] Virga, K. G., Zhang, Y.-M., Leonardi, R., Ivey, R. A., Hevener, K., Park, H.-W., Jackowski, S., Rock, C. O., and Lee, R. E. (2006) Structure-activity relationships and enzyme inhibition of pantothenamide-type pantothenate kinase inhibitors. *Bioorg Med Chem*, **14**, 1007–1020.

5.5 Materials and Methods for Chapter 5

5.5.1 General Material and Methods for Synthesis and Assays

Unless otherwise stated, chemicals were from Alfa Aesar, Aldrich, Acros chemicals, Bachem or Fisher Scientific and used without further purification. Solvents for chemical transformations, work-up and chromatography were from Rathburn or Aldrich at HPLC grade, and used without further distillation. Anhydrous solvents were from Aldrich (Dorset, U.K.) or dried by filtration through columns containing activated aluminium oxide under argon. Silica gel 60 F_{254} analytical thin layer chromatography (TLC) plates were obtained from Merck and visualised under UV light, or with potassium permanganate stain. Chromatographic purifications were performed using prepacked SNAP columns on a Biotage SP1 Purification system (Uppsala, Sweden). Yields refer to purified, dried, and spectroscopically pure compounds. NMR tubes (5 mm) were obtained from Aldrich or from Bruker (1 and 2 mm). Deuterated solvents were obtained from Aldrich and Apollo Scientific Ltd. All ^{1}H and ^{13}C NMR Spectra were recorded using either Bruker AV 500 MHz with ^{13}C cryoprobe and variable temperature setup or Bruker AV 400 MHz spectrometer. All chemical shifts are given in ppm relative to the solvent peak [96]. Coupling constants *J* are given in Hz to the nearest 0.5 Hz value. Low Resolution (LR) Electrospray Ionisation (ESI) Mass Spectrometry (MS) data (m/z) were obtained on a Fision VG Platform using quadrupole analyser. High Resolution mass spectrometry (HR-MS) data were obtained from a MicroTOF® instrument (Bruker) using ESI source and Time of Flight (TOF) analyser. Matrix-assisted laser desorption/ionisation (MALDI) MS was performed using a Micro MX time of flight (TOF) instrument (Waters). Infrared (IR) spectra were recorded on a Tensor® 27 instrument (Bruker). Preparative HPLC was performed on a UltiMate® 3000 system (Dionex) using a Vydac 218TP C18 column (250 mm x 25 mm 10-15μ, Grace Davison Discovery Sciences) and the following mobile phase: A = 0.1% formic acid in water, B = 0.1% formic acid in acetonitrile. The detection of compounds with no chromophore was achieved using a PL-ELS 2100 Ice™ (Varian Inc.) evaporative light scattering detector (ELSD). Analytic HPLC was performed on a 1290 Infinity™ system (Agilent Technologies) using various columns and mobile phases as described in the text.

Lyophilisation

According to the volume of the sample to be lyophilised, a 15 mL or 50 mL falcon tube, or 1.5 mL Eppendorf tube was used as a container. Lyophilisation was performed by flash freezing the samples in liquid nitrogen before applying a vacuum of $\leq$ 10 mbar on a Edwards Freeze Drier Micro Modulyo until the samples went dry.

5.5.2 General Material and Methods for Molecular Biology

All the molecular biology protocols were carried out according to standard procedures in place in the group of Prof Christopher J. Schofield at the University of Oxford, and were essentially the same as in Section 4.4.1.

5.5.3 Expression and Purification of the CoA Biosynthesis Enzymes

Protein expression Trials

Expression trials were performed as described in Section 4.4.2.

Large Scale Protein Expression

The genes coding for *E. coli* coenzyme A biosythesis enzymes were cloned into expression vectors and were a generous gift of Dr. Manuela Tosin (University of Cambridge, now at the University of Warwick, U.K.). The expression and purification procedure for all three enzymes was adapted from previously described methodologies [28, 35]. Before expression trials were conducted, the DNA sequence of these constructs were verified by sequencing. All three sequences were searched in the NCBI Nucleotide database using the blastn algorithm [67] and found to be identical to the *E. coli* BL21 (DE3) genes. Because these proteins were expressed in their original producer species, the yields were expected and found to be high. The genes coding for the *E. coli* pantothenate kinase and the phosphopantothenate adenosyl transferase (CoaA and CoaD) had been cloned into the pET-29b(+) vector, between NcoI and HindIII restriction sites. The genes are regulated by a T7 promoter and the corresponding proteins carried a *N*-terminal hexahistidine tag. The dephosphocoenzyme A kinase (CoaE) had been cloned into a pET-20b(+) vector, between NdeI and XhoI restriction sites. This gene is also regulated by an T7 promoter, however the corresponding protein carries a *C*-terminal hexahistidine tag. The plasmids pET-29b(+) and pET-20b(+) contain a kanamycin and ampicillin resistance gene, respectively. The three constructs were independently transformed into *E. coli* BL21 (DE3) cells for protein expression.

Nickel-affinity chromatography

The one-step purification of the recombinant proteins was carried out following a standard protocol with little modification from the supplier instructions. For each CoA biosynthesis protein, the filtered cell lysate was loaded at a 1 mL min^{-1} flow rate using the FPLC pump on a 5 mL HisTrap^{TM}FF column (GE healthcare) previously charged with with nickel (II) ions and, after washing the unbound proteins with buffer, the tagged protein was eluted with a 60–600 mM imidazole gradient over 10 column volumes. Analysis of the flow through and wash fractions indicated a fully efficient binding of the target protein to the resin. The fractions containing eluted proteins, as judged by monitoring the UV absorbance, were screened by SDS-PAGE electrophoresis (Figure 5.3). Fractions containing the proteins of interest were pooled, mixed with 20 mM Tris-HCl pH 7.5 containing 300 mM NaCl and 100 mM EDTA to reach a final concentration of 1 mM EDTA and incubated at 4 °C for 1 h to chelate nickel ions leached off the column. The EDTA and the imidazole were removed from the protein samples by overnight dialysis as below. The following buffers were used:

Ni Lysis buffer	
Tris	50 mM
Imidazole	10 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Wash buffer	
Tris	50 mM
Imidazole	40 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Elution buffer	
Tris	50 mM
Imidazole	500 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Strip buffer	
Tris	20 mM
Sodium chloride	100 mM
EDTA	100 mM

Dialysis

After nickel-affinity chromatography, the imidazole containing buffer of CoaA, CoaD and CoaE protein samples was exchanged for 50 mM HEPES buffer pH 8.0 containing 250 mM NaCl and 2 mM $MgCl_2$ by dialysis over night at 4 °C using a 10 kDa cut-off membrane.

5.5.4 Synthesis of Pantetheine and Acyl-pantetheine Analogues

Benzyl 5-oxo-L-prolinate 300 [87]

L-Pyroglutamic acid **141** (5.0 g, 38.7 mmol) was dissolved in dried dimethylformamide (150 mL) together with benzyl bromide (4.58 mL, 0.99 eq.) and potassium hydrogen carbonate (3.85 g, 1.1 eq.). Catalytic sodium sulphate (550 mg, 10% mol/mol) and sodium iodide (29 mg, 0.5% mol/mol) were added and the reaction was stirred at room temperature for 48 h. The solvent was removed under vacuum and the residue partitioned between ethyl acetate and a saturated sodium hydrogen carbonate aqueous solution. The organic phase was washed with brine and water, dried over magnesium sulphate and evaporated to give benzyl 5-oxo-L-prolinate **300** as a pale pink oil (6.7 g, 80%). This oil was used in the next step without further purification as the NMR chemical shifts were in accordance with previously reported data [87].

TLC: 1:1 *n*-hex/EtOAc, R_f = 0.40; IR (neat) ν/cm^{-1}: 3091, 3065, 3034 (CH_{aryl}), 1749 (OC=O), 1708 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.38-7.30 (*m*, 5H), 7.06 (*app bs*, 1H), 5.17 (*s*, 2H), 4.26 (*dd*, J = 8.5, 5.0, 1H), 2.48-2.38 (*m*, 1H), 2.36-2.25 (*m*, 2H), 2.22-2.13 (*m*, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 178.1, 171.9, 135.0, 128.5, 128.2, 67.1, 55.4, 29.1, 24.6; LR ESI-MS : 242.10 ([M+Na]$^+$), 461.20 ([2M+Na]$^+$).

Dibenzyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate 302

Benzyl 5-oxo-L-prolinate **300** (5.3 g, 24.1 mmol) was dissolved in dried tetrahydrofuran (100 mL) and the solution was cooled to -78 °C under a nitrogen atmosphere. A 2.5 M solution of *n*-butyl lithium in hexanes (10.6 mL, 1.1 eq.) was added dropwise and the reaction turned into a suspension which was stirred for 30 min. A 2.9 M solution of benzyl chloroformate in toluene (4.1 mL, 1.2 eq.) was added dropwise and the mixture was stirred at -78 °C for 2 h, then allowed to reach 0 °C over 1 h. The reaction was quenched at this temperature by addition of saturated sodium hydrogen carbonate aqueous solution (50 mL) and allowed to reach room temperature. The organic solvent were removed under vacuum and the resulting aqueous layer extracted three times with dichloromethane. The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified by silica chromatography (SNAP-100 column, gradient elution using 0–100% ethyl acetate in *n*-hexane) and the product recrystallised from a 1:1 ethyl acetate/*n*-hexane mixture to yield dibenzyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **302** as a white solid (4.6 g, 54%).

mp 105–108 °C; TLC: 1:1 *n*-hexane/EtOAc, R_f = 0.30; IR (neat) ν/cm^{-1}: 3065, 3044, 3031 (CH_{aryl}), 1792 (OC=O), 1737 (NC=O), 1709 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.37-7.32 (*m*, 8H), 7.30-7.27 (*m*, 2H), 5.23 (*s*, 2H), 5.13 (*s*, 2H), 4.73 (*dd*, *J* = 9.5, 2.5, 1H), 2.69-2.59 (*m*, 1H), 2.54-2.47 (*m*, 1H), 2.40-2.24 (*m*, 1H), 2.10-2.03 (*m*, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.8, 170.8, 150.8, 134.9, 128.7, 128.6, 128.6, 128.4, 128.3, 128.1, 68.3, 67.4, 58.7, 31.0, 21.7; HR ESI-MS : 376.1154 ([M+Na]$^+$, $C_{20}H_{19}NNaO_5{}^+$, calc. 376.1155).

Dibenzyl (2*S*)-5-hydroxypyrrolidine-1,2-dicarboxylate 301

A solution of dibenzyl (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **302** (3.0 g, 8.49 mmol) in dried tetrahydrofuran (50 mL) was cooled to -78 °C under a nitrogen atmosphere prior to dropwise addition of a 1 M solution of lithium triethylborohydride in tetrahydrofuran (8.0 mL, 0.95 eq.). The reaction was stirred under these conditions for 3 h and 1.0 M solution of lithium triethylborohydride in tetrahydrofuran (8.0 mL, 0.95 eq.) was added again. After another 3 h, the reaction was quenched by slow addition of saturated sodium hydrogen carbonate solution (25 mL) and a 40% aqueous hydrogen peroxide solution (2 mL). The temperature was raised to 0 °C and the mixture stirred for 30 min. The mixture was allowed to slowly reach room temperature and the tetrahydrofuran was remove under reduced pressure. The aqueous layer was extracted three times with diethyl ether, and the combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified using silica chromatography (SNAP-100 column, gradient elution with 0–80% ethyl acetate in *n*-hexane) and dibenzyl (2*S*)-5-hydroxypyrrolidine-1,2-dicarboxylate **301** was obtained as a colourless oil (2.17 g, 72%).

TLC: 1:1 hex/EtOAc, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.37-7.21 (*m*, 10H), 5.97-5.59 (*m*, 1H), 5.28-4.97 (*m*, 4H), 4.58-4.38 (*m*, 1H), 3.62 (*app bs*, 1H), 2.74-2.53 (*m*, 1H), 2.50-2.25 (*m*, 1H), 2.23-1.96 (*m*, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.3, 172.0, 171.8, 171.7, 171.1, 154.6, 135.9 (x2), 135.4, 135.3, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 127.8, 127.6, 82.9, 82.7, 82.1, 81.8, 81.6, 68.1, 67.9, 67.5, 67.2 (x2), 66.9, 66.8, 66.7, 60.3, 59.5, 59.4, 59.1, 59.0, 33.7, 33.5, 32.4, 32.2, 31.9, 31.0, 28.0, 27.9, 27.1, 26.8, 25.8, 22.9, 21.0, 14.1, 13.9; HR ESI-MS : 378.1306 ($[M+Na]^+$, $C_{20}H_{21}NNaO_5^+$, calc. 378.1312).

Dibenzyl (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate 303

tert-Butyldiethylphosphonoacetate (1.40 mL, 1.02 eq.) was added dropwise to a suspension of 30% potassium hydride on mineral oil (800 mg, 1.02 eq.) in dried dimethylformamide (25 mL). The reaction was stirred at room temperature under a nitrogen atmosphere for 1 h prior to dropwise addition of dibenzyl (2*S*)-5-hydroxypyrrolidine-1,2-dicarboxylate **301** (2.10 g, 5.91 mmol) in dried dimethylformamide (25 mL). The mixture was allowed to stir at room temperature for 24 h before quenching with a saturated ammonium chloride solution (100 mL) and dilution with diethyl ether (100 mL). The phases were separated and the aqueous layer was extracted two times with diethyl ether (100 mL). The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified using silica chromatography (SNAP-100 column, gradient elution with 0–40% ethyl acetate in hexane) and dibenzyl (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **303** was obtained as a colourless oil (1.85 g, 69%).

TLC: 8:2 *n*-hexane/EtOAc, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): 7.37-7.22 (*m*, 10H), 5.25-5.09 (*m*, 2H), 5.06-4.96 (*m*, 2H), 4.49-4.39 (*m*, 1.7H), 4.35-4.27 (*m*, 0.3H), 3.20 (*dd*, *J* = 16.0, 3.5, 0.2H), 2.96 (*dd*, *J* = 15.5, 3.0, 0.5H), 2.75 (*dd*, *J* = 15.0, 3.0, 0.3H), 2.67 (*dd*, *J* = 15.5, 5.0, 0.1H), 2.45-2.33 (*m*, 0.3H), 2.30-2.12 (*m*, 2.7H), 1.97 (*dd*, *J* = 12.0, 6.0, 1H), 1.83 (*dd*, *J* = 12.0, 6.0, 1H), 1.46-1.43 (m, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.6, 172.3, 172.0, 170.7, 170.4, 170.3, 154.5, 153.9, 153.6, 136.4, 136.3, 135.5, 135.3, 128.5, 128.4, 128.3, 128.2 (x2), 128.0 (x2), 127.9, 127.8 (x2), 127.7, 127.5., 127.4, 80.7, 80.5, 80.4, 67.0, 66.8 (x2), 66.7, 59.8, 59.7, 59.5, 55.6, 55.3, 54.7, 40.2, 39.2, 39.0, 28.8, 28.1, 28.0 (x2); HR ESI-MS : 476.2049 ($[M+Na]^+$, $C_{26}H_{31}NNaO_6^+$, calc. 476.2044).

2-((5*S*)-1,5-Bis(benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid 304

Dibenzyl (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **303** (1.80 g, 4.0 mmol) was deprotected in dichloromethane (20 mL) containing 10% vol/vol trifluoroacetic acid over 6 h.

The solvent was evaporated and the excess trifluoroacetic acid coevaporated with toluene, followed by dichloromethane. The free acid was obtained as a pale orange low melting point solid (1.57 g, app. quant.).

TLC: 1:1 hex/EtOAc, R_f = 0.10; IR (neat) ν/cm^{-1}: 3091, 3066, 3035 (CH_{aryl}), 1742 (OC=O), 1708 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.36-7.22 (*m*, 10H), 5.24-5.11 (m, 2H), 5.08-4.96 (*m*, 2H), 4.51-4.46 (*m*, 1H), 4.43-4.41 (*m*, 0.7H), 4.38-4.32 (*m*, 0.3H), 3.28 (*dd*, *J* = 16.5, 4.0, 0.2H), 3.06 (*dd*, *J* = 16.0, 3.5, 0.5H), 3.01 (*dd*, *J* = 16.0, 4.0, 0.2H), 2.86 (*dd*, *J* = 15.5, 3.5, 0.3H), 2.61-2.49 (*m*, 0.3H), 2.44-2.37 (*m*, 0.7H), 2.31-2.15 (*m*, 2H), 2.07-1.95 (*m*, 1H), 1-87-1.79 (*m*, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.3, 172.2, 154.2, 136.1, 135.3, 128.6, 128.5, 128.4 (x2), 128.3, 128.1 (x2), 128.0, 127.9, 127.8, 67.4, 67.2, 67.0, 66.9, 60.0, 59.7, 55.1, 54.4, 38.8, 38.1, 30.1, 29.2, 28.8, 28.5, 28.3; HR ESI-MS : 420.1412 ([M+Na]$^+$, $C_{22}H_{23}NNaO_6^+$, calc. 420.1418).

4-Nitrobenzyl 5-oxo-L-prolinate 296

5-Oxo-L-proline **141** (5.0 g, 38.7 mmol) was dissolved in dried dimethylformamide (150 mL) together with 4-nitrobenzyl bromide (8.30 g, 0.99 eq.) and potassium hydrogen carbonate (3.85 g, 1.1 eq.). Catalytic amount of sodium sulphate (550 mg, 10% mol/mol) and sodium iodide (29 mg, 0.5% mol/mol) was added and the reaction was stirred at room temperature for 48 h. The solvent was removed under vacuum and the residue partitioned between ethyl acetate and a saturated sodium hydrogen carbonate aqueous solution. The organic phase was washed with brine and water, dried over magnesium sulphate and evaporated to give 4-nitrobenzyl 5-oxo-L-prolinate **296** as a low melting point white solid (8.5 g, 85%).

TLC: 1:1 *n*-hex/EtOAc, R_f = 0.40; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.26-8.23 (*m*, 2H), 7.56-7.52 (*m*, 2H), 6.59 (*app bs*, 1H), 5.29 (*s*, 2H), 4.35 (*ddd*, *J* = 9.0, 5.0, 0.5, 1H), 2.58-2.49 (*m*, 1H), 2.47-2.33 (*m*, 2H), 2.30-2.21 (*m*, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 177.8, 171.6, 147.9, 142.1, 128.6, 123.9, 65.7, 55.3, 29.1, 24.8; HR ESI-MS : ([M+Na]$^+$, $C_{12}H_{12}N_2NaO_5^+$, calc. 287.0644).

1-Benzyl 2-(4-nitrobenzyl) (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate 297

4-Nitrobenzyl 5-oxo-L-prolinate **296** (5.0 g, 18.9 mmol) was dissolved in dried tetrahydrofuran (100 mL) and the solution was cooled to -78 °C under a nitrogen atmosphere. A 2.5 M solution of *n*-butyl lithium in hexanes (8.3 mL, 1.1 eq.) was added dropwise and the reaction turned into a deep red suspension which was stirred for 30 min. A 2.9 M solution of benzyl chloroformate in toluene (4.1 mL, 1.2 eq.) was added dropwise and the mixture was stirred at -78 °C for 1 h, then allowed to reach 0 °C over 1 h. The reaction was quenched at this temperature by addition of saturated sodium hydrogen carbonate aqueous solution (30 mL) and allowed to reach room temperature. The organic solvent were removed under vacuum and the resulting aqueous layer extracted three times with dichloromethane. The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified by chromatography (SNAP-100 column, gradient elution with 0–100% ethyl acetate in *n*-hexane) and the product recrystallised from a 1:9 ethyl acetate/*n*-hexane mixture to yield 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **297** as a low melting point white solid (5.0 g, 67%).

mp 102–106 °C; TLC: 1:1 *n*-hex/EtOAc, R_f = 0.25; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.15-8.11 (*m*, 2H), 7.39 (*app. d*, *J* = 8.5, 2H), 7.34-7.27 (*m*, 5H), 5.27-5.13 (*m*, 4H), 4.76 (*dd*, *J* = 9.5, 3.0, 1H), 2.68-2.48 (*m*, 2H), 2.44-2.34 (*m*, 1H), 2.11-2.03 (*m*, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 175.8, 172.6, 170.6, 150.7, 147.6, 141.9, 134.7, 128.4, 128.3, 128.2, 127.8, 123.6, 68.2, 65.6, 58.5, 30.8, 21.5; HR ESI-MS : 421.0996 ([M+Na]$^+$, $C_{20}H_{18}N_2NaO_7^+$, calc. 421.1006).

(2*S*)-1-Benzyl 2-(4-nitrobenzyl) 5-hydroxypyrrolidine-1,2-dicarboxylate 298

1-Benzyl 2-(4-nitrobenzyl) (2*S*)-5-oxopyrrolidine-1,2-dicarboxylate **297** (3.0 g, 7.53 mmol) was dissolved into dried tetrahydrofuran (50 mL) and the solution was cooled to -78 °C under a nitrogen atmosphere. A 2 M lithium triethylborohydride solution in tetrahydrofuran (8.0 mL, 1.06 eq.) was added dropwise, the reaction was stirred at -78 °C under a nitrogen atmosphere for 8 h and quenched by slow addition of a saturated sodium hydrogen carbonate aqueous solution (50 mL) followed by methanol (5 mL). The mixture was allowed to reach room temperature and stirred for 1 h more. The volatile solvents were remove under reduced pressure and the resulting aqueous layer was extracted with diethyl ether. The combined organic layers were dried over magnesium sulphate and evaporated. The residue was purified by chromatography (SNAP-100 column, gradient elution with 0–100% ethyl acetate in *n*-hexane) to yield (2*S*)-1-benzyl 2-(4-nitrobenzyl) 5-hydroxypyrrolidine-1,2-dicarboxylate **298** as a colourless oil (2.48 g, 83%). The NMR analyses show a complicated pattern probably due to conformational isomers.

TLC: 1:1 *n*-hex/EtOAc, $R_f = 0.20$; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.22-8.10 (*m*, 2H), 7.52-7.21 (*m*, 7H), 5.74-5.63 (*m*, 1H), 5.31-5.02 (*m*, 4H), 4.57 (*dd*, *J* = 16.0, 9.0, 0.65H), 4.49 (*t*, *J* = 8.0, 0.15H), 4.43 (*t*, *J* = 7.5, 0.2H), 3.74-3.71 (*m*, 0.45H), 3.62-3.60 (*m*, 0.25H), 3.05-3.01 (*m*, 0.3H), 2.62-2.49 (*m*, 1H), 2.36-2.13 (*m*, 1H), 2.11-1.91 (*m*, 2H); LR ESI-MS : 397.10 ($[M-H]^-$).

1-Benzyl 2-(4-nitrobenzyl) (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate 299

tert-Butyl diethylphosphonoacetate (0.21 mL, 1 eq.) was added dropwise to a stirring suspension of 60% sodium hydride on mineral oil (35 mg, 1 eq.) in dried dimethylformamide (10 mL). The reaction was stirred under a nitrogen atmosphere for 1 h at room temperature and a solution of (2*S*)-1-benzyl 2-(4-nitrobenzyl) 5-hydroxypyrrolidine-1,2-dicarboxylate **298** (350 mg, 0.87 mmol) in dried dimethylformamide (10 mL) was added dropwise. The reaction was stirred at room temperature overnight before quenching with a saturated ammonium chloride aqueous solution (50 mL). This aqueous solution was extracted with diethyl ether and the combined organic phases were dried over magnesium sulphate and evaporated. The residue was purified by chromatography (SNAP-50 column, gradient elution with 0–100% ethyl acetate in *n*-hexane) to yield 1-benzyl 2-(4-nitrobenzyl) (2*S*)-5-(2-*tert*-butoxy-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **299** as a yellow oil (144 mg, 33%).

TLC: 8:2 *n*-hex/EtOAc, $R_f = 0.40$; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.22-8.08 (*m*, 2H), 7.54-7.23 (*m*, 7H), 5.28-4.95 (*m*, 4H), 4.49-4.40 (*m*, 1H), 4.21-4.10 (*m*, 1H), 3.18 (*d*, *J* = 16.0, 3.5, 0.2H), 2.97 (*d*, *J* = 15.0, 3.5, 0.5H), 2.76 (*d*, *J* = 16.0, 3.0, 0.3H), 2.38-2.12 (*m*, 2H), 2.01-1.96 (*m*, 1H), 1.89-1.83 (*m*, 1H), 1.65 (*app bs*, 1H), 1.45 (*s*, 9H); LR ESI-MS : 399.10 ($[M-H]^-$).

***tert*-Butyl (2-aminoethyl)carbamate 271** [74, 75]

tert-Butyl (2-aminoethyl)carbamate **271** was synthesised by adapting reported procedures [74, 75]. A solution of di-*tert*-butyl dicarbonate (8.73 g, 40 mmol) in 1,4-dioxane (10 mL) was added dropwise over 30 min to a solution of ethane-1,2-diamine **266** (8.1 mL, 3 eq.) in 1,4-dioxane (40 mL). A white precipitate was formed slowly as the reaction was left under stirring conditions at room temperature. After 4 h, the mixture was filtered and the clear solution was concentrated to yield *tert*-butyl (2-aminoethyl)carbamate **271** as a white low melting point solid (5.79 g, 90%).

mp $\leq$ 35 °C; TLC: 3:1 CH_2Cl_2/MeOH, $R_f = 0.50$; IR (neat) ν/cm^{-1}: 3443 (NH), 3364 (NH), 1704 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): = 5.11 (*app bs*, 1H), 3.11 (*app. q.*, *J* = 5.5, 2H), 2.74 (*app. t*, *J* = 6.0, 2H), 1.39 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 156.14, 79.0, 43.3, 41.7, 28.3; HR ESI-MS : 183.1104 ($[M+Na]^+$, $C_7H_{16}N_2NaO_2^+$, calc. 183.1104).

***tert*-Butyl (2-hydroxyethyl)carbamate 270** [73]

The synthesis of *tert*-butyl (2-hydroxyethyl)carbamate **270** has been described using a different procedure [73]. A solution of di-*tert*-butyl dicarbonate (10.92 g, 50 mmol) in 1,4-dioxane (20 mL) was added dropwise over 30 min to a solution of 2-aminoethanol **267** (3.02 mL, 1 eq.) in 1,4-dioxane (30 mL). After 4 h at room temperature, the reaction was completed, as judged by TLC, and a saturated sodium hydrogen carbonate aqueous solution was added (50 mL). The 1,4-dioxane was removed by evaporation and the aqueous phase was extracted with ethyl acetate. The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give *tert*-butyl (2-hydroxyethyl)carbamate **270** as a pale yellow oil (7.30 g, 90%).

TLC: 3:1 *n*-hex/EtOAc, R_f = 0.20; IR (neat) ν/cm^{-1}: 3319 (OH), 1683 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.12 (*app bs*, 1H), 3.69-3.65 (*m*, 2H), 3.28-3.25 (*m*, 2H), 3.05 (*app bs*, 1H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 156.8, 79.6, 62.4, 43.1, 28.3; HR ESI-MS : 184.0942 ($[M+Na]^+$, $C_7H_{15}NNaO_3{}^+$, calc. 184.0944).

Dibenzyl (2*S*)-5-[2-((2-[(*tert*-butoxycarbonyl)amino]ethyl)amino)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate 337

To a solution of 2-((5*S*)-1,5-bis(Benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid **304** (440 mg, 0.97 mmol) in dried tetrahydrofuran (5 mL) at 0 °C was added 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (382 mg, 2 eq.), hydroxybenzotriazole (338 mg, 2.5 eq.) and *tert*-butyl (2-aminoethyl)carbamate **271** (240 mg, 1.5 eq.). After stirring at 0 °C for 15 min, diisopropylethylamine (0.87 mL, 5 eq.) was added dropwise and the reaction was stirred at room temperature for 24 h, quenched with a 20% ammonium chloride aqueous solution (20 mL) before dilution with diethyl ether (100 mL). Phases were separated and the aqueous layer was extracted once more with diethyl ether. Combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified using silica chromatography (SNAP-50 column, gradient elution with 0–70% (10% methanol in ethyl acetate) in *n*-hexane) and dibenzyl (2*S*)-5-[2-((2-[(*tert*-butoxycarbonyl)amino]ethyl)amino)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate **337** was obtained as a white solid (435 mg, 83%).

mp 120–123 °C; TLC: 19:1 DCM/MeOH R_f = 0.35; IR (KBr) ν/cm^{-1}: 3364 (NH), 3329 (NH) 3063, 3032 (CH_{aryl}), 1740 (OC=O), 1690 (NC=O), 1641 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.39-7.22 (*m*, 10H), 5.28-4.94 (*m*, 4H), 4.47-4.35 (*m*, 2H), 3.38-3.17 (*m*, 5H), 2.72 (*dd*, *J* = 14.0, 3.0, 0.65H), 2.59 (*dd*, *J* = 14.0, 2.5, 0.35H), 2.34-1.94 (*m*, 5H), 1.43 (*s*, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.2, 171.3, 154.4, 136.2, 135.4, 128.7, 128.6, 128.4, 128.3, 128.2, 128.1, 127.9, 67.3, 67.1, 59.7, 56.4, 28.4; HR ESI-MS : 562.2513 ($[M+Na]^+$, $C_{29}H_{37}N_3NaO_7{}^+$, calc. 562.2524).

Dibenzyl (2*S*)-5-(2-[(2-aminoethyl)amino]-2-oxoethyl)pyrrolidine-1,2-dicarboxylate 338

Dibenzyl (2*S*)-5-[2-((2-[(*tert*-butoxycarbonyl)amino]ethyl)amino)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate **337** (400 mg, 0.74 mmol) was deprotected in a mixture of 10% trifluoroacetic acid in dichloromethane over 6 h. Coevaporation of the solvent with toluene followed by dichloromethane yielded the trifluoroacetate salt of dibenzyl (2*S*)-5-(2-[(2-aminoethyl)amino]-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **338** as a colourless oil (400 mg, app. quant.). This compound was used for further steps without purification.

IR (neat) ν/cm^{-1}: 3300 (NH), 3091, 3069, 3037 (CH_{aryl}), 1745 (OC=O), 1680 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.13-7.81 (*m*, 3H), 7.35-7.19 (*m*, 10H), 5.20-4.92 (*m*, 4H), 4.50-4.42 (*m*, 2H), 3.53 (*app bs*, 1H), 3.41 (*app bs*, 1H), 3.12-3.08 (*m*, 2H), 2.67-1.67 (*m*, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.6, 172.1, 161.3, 161.0, 154.9, 135.8, 135.2, 128.6-127.3 (multiple peaks), 67.2, 66.8, 59.5, 56.2, 41.03, 39.8, 37.3, 28.9, 28.1; LR ESI-MS : 440.17 ($[M+H]^+$).

Dibenzyl (2*S*)-5-(2-[(2-acetamidoethyl)amino]-2-oxoethyl)pyrrolidine-1,2-dicarboxylate 325

2-((5*S*)-1,5-Bis(benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid **304** (250 mg, 0.63 mmol) was dissolved in dried tetrahydrofuran (5 mL) under a nitrogen atmosphere together with 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (241 mg, 2 eq.), hydroxybenzotriazole (212 mg, 2.5 eq.) and *N*-(2-aminoethyl)acetamide (120 μL, 2.0 eq.) at 0 °C. Diisopropylethylamine (550 μL, 5 eq.) was then added dropwise to the mixture and the reaction was stirred at room temperature for 24 h, quenched with a 20% solution of ammonium chloride (20 mL) before dilution with diethyl ether (100 mL). The phases were separated and the aqueous layer was extracted once more with diethyl ether. The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified using silica chromatography (SNAP-25 column, gradient elution with 0–10% methanol in ethyl acetate). Dibenzyl (2*S*)-5-(2-[(2-acetamidoethyl)amino]-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **325** was obtained as a low melting point white solid (186 mg, 61%).

mp $\leq$ 35 °C; TLC: 9:1 DCM/MeOH, R_f = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.35-7.20 (*m*, 10H), 7.11 (*app bs*, 0.2H), 6.82 (*app bs*, 0.8H), 6.76 (*app bs*, 0.5H), 6.56 (*app bs*, 0.2H), 6.46 (*app bs*, 0.3H), 5.24-4.94 (*m*, 4H), 4.49-4.36 (*m*, 2H), 3.33-3.09 (*m*, 4H), 2.73 (*dd*, *J* = 14.0, 4.0, 0.7H), 2.60 (*dd*, *J* = 14.0, 2.5, 0.3H), 2.34-2.03 (*m*, 3H), 1.99-1.84 (*m*, 5H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.1, 171.9, 171.7, 171.3, 171.0, 154.6, 154.3, 136.3, 136.1, 135.5, 135.2, 135.0, 128.6-127.6 (10 peaks), 67.2, 67.0, 66.8, 66.7, 59.8, 59.5, 56.2, 55.3, 41.0, 40.8, 40.0, 39.9, 39.5, 39.3, 28.9, 28.8, 28.5, 28.2, 27.2, 23.0; HR ESI-MS : 504.2107 ([M+Na]$^+$, $C_{26}H_{31}N_3NaO_6^+$, calc. 504.2105).

Dibenzyl (2*S*)-5-[2-(2-acetamidoethoxy)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate 326

2-((5*S*)-1,5-Bis(benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid **304** (250 mg, 0.63 mmol) was dissolved in dried tetrahydrofuran (5 mL) under a nitrogen atmosphere together with 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (241 mg, 2 eq.), hydroxybenzotriazole (212 mg, 2.5 eq.) and *N*-(2-hydroxyethyl)acetamide (116 μL, 1.5 eq.) at 0 °C. Diisopropylethylamine (550 μL, 5 eq.) was then added dropwise to the mixture and the reaction was stirred at room temperature for 24 h, quenched with a 20% ammonium chloride aqueous solution (20 mL) before dilution with diethyl ether (100 mL). The phases were separated and the aqueous layer was extracted once more with diethyl ether. The combined organic phases were dried over magnesium sulphate, filtered and evaporated. The residue was purified using silica chromatography (SNAP-25 column, gradient elution with 0–10% methanol in ethyl acetate). Dibenzyl (2*S*)-5-[2-(2-acetamidoethoxy)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate **326** was obtained as a colourless oil (141 mg, 46%).

TLC: 9:1 DCM/MeOH, R_f = 0.35; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.35-7.22 (*m*, 10H), 6.62 (*app bs*, 0.6H), 6.39 (*app bs*, 0.1H), 6.03 (*app bs*, 0.3H), 5.23-4.94 (*m*, 4H), 4.55-4.30 (*m*, 4H), 4.26-4.20 (*m*, 0.7H), 4.16-4.00 (*m*, 1.3H), 3.68-3.60 (*m*, 0.7H), 3.52-3.46 (*m*, 1.3H), 3.38-3.19 (*m*, 0.7H), 3.05-2.91 (*m*, 0.3H), 2.76 (*dd*, J = 15.0, 6.0, 0.7H), 2.68 (*dd*, J = 15.0, 3.0, 0.3H), 2.59-2.46 (*m*, 0.3H), 2.43-2.12 (*m*, 1.7H), 2.07-1.94 (*m*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.1, 171.0, 170.6, 154.4, 136.0, 135.2, 128.5-127.7 (10 peaks), 67.3, 67.1, 67.0, 66.8, 63.7, 63.5, 63.2, 59.9, 59.6, 55.5, 54.6, 51.6, 39.6, 39.2, 38.7, 38.5, 29.3, 29.1, 28.3, 28.2, 23.0; HR ESI-MS : 505.1937 ([M+Na]$^+$, $C_{26}H_{30}N_2NaO_7^+$, calc. 505.1945).

5-(2-[(2-Acetamidoethyl)amino]-2-oxoethyl)-L-proline 323

Dibenzyl (2*S*)-5-(2-[(2-acetamidoethyl)amino]-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **325** (100 mg, 0.21 mmol) was dissolved in ethanol (2 mL) and palladium(0) on activated carbon was added (10 mg, 10% w/w). The black suspension was stirred under a hydrogen atmosphere for 5 h until reaction completion, as judged by TLC, and then filtered through a pad of Celite® 545. The filtrate was evaporated and the residue purified by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 15 min) to yield 5-(2-[(2-acetamidoethyl)amino]-2-oxoethyl)-L-proline **323** as

a white solid (R_t = 9.5 min, 34 mg, 64%). ^{1}H NMR analysis shows a 7:3 *trans*/*cis* mixture of C-5 diastereomers.

^{1}H NMR (500 MHz, D_2O): δ = 4.17-4.12 (*m*, 1H), 3.99-3.89 (*m*, 1H), 3.32-3.26 (*m*, 4H), 2.75-2.67 (*m*, 2H), 2.46-2.39 (*m*, 0.7H), 2.34-2.27 (*m*, 0.3H), 2.24-2.12 (*m*, 1.4H), 2.07-1.96 (*m*, 0.6H), 1.94 (*s*, 3H), 1.81-1.65 (*m*, 1H); ^{13}C NMR (125 MHz, D_2O): δ = 174.5, 174.4, 174.3, 171.9, 171.8, 61.2, 60.9, 57.4, 57.2, 38.6, 38.5, 36.5, 36.3, 29.7, 28.8, 28.2, 28.1, 21.6; HR ESI-MS : 280.1269 ($[M+Na]^+$, $C_{11}H_{19}N_3NaO_4^+$, calc. 280.1268).

5-[2-(2-Acetamidoethoxy)-2-oxoethyl]-L-proline 324

Dibenzyl (2*S*)-5-[2-(2-acetamidoethoxy)-2-oxoethyl]pyrrolidine-1,2-dicarboxylate **326** (250 mg, 0.21 mmol) was dissolved in ethanol (2 mL) and palladium(0) on activated carbon was added (10 mg, 10% w/w). The black suspension was stirred under a hydrogen atmosphere for 5 h until reaction completion, as judged by TLC, and then filtered through a pad of Celite® 545. The filtrate was evaporated and the residue purified by preparative HPLC to yield 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324** as a white solid (R_t = 10.5 min, 67 mg, 50%). ^{1}H NMR analysis shows a 7:3 *trans*/*cis* mixture of C-5 diastereomers.

^{1}H NMR (500 MHz, D_2O): δ = 4.13-4.08 (*m*, 1H), 3.95-3.82 (*m*, 1H), 3.26-3.21 (*m*, 4H), 2.70-2.66 (*m*, 2H), 2.42-2.34 (*m*, 0.7H), 2.30-2.23 (*m*, 0.3H), 2.20-2.07 (*m*, 1H), 2.01-1.93 (*m*, 1H), 1.89 (s, 3H), 1.72-1.59 (*m*, 1H); ^{13}C NMR (125 MHz, D_2O): δ = 174.4, 174.2, 174.1, 172.2, 172.1, 61.0, 60.9, 54.3, 54.2, 38.9, 38.7, 35.9, 35.6, 28.7, 28.6, 28.1, 28.0, 21.5; HR ESI-MS : 281.1116 ($[M+Na]^+$, $C_{11}H_{18}N_2NaO_5^+$, calc. 281.08).

***N*-([(4*R*)-2,2,5,5-Tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine 268** [35, 97]

N-([(4*R*)-2,2,5,5-Tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **268** was obtained by adapting a previously reported procedure [69]. The hemicalcium salt of D-pantothenic acid **246** ((*R*)-3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanoic acid, 11.91 g, 50 mmol) was suspended in 2,2-dimethylpropane (100 mL) and *p*-toluene sulfonic acid (9.51 g, 1 eq.) was added in one portion. The suspension was stirred at room temperature for 16 h and the white insoluble solid was filtered off and washed with acetone. The filtrate was evaporated and the resulting white solid was washed with warm *n*-hexane to yield the hemi-calcium salt of *N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanine **268** (10.50 g, 81%).

mp 90–92 °C. (lit. 88–90 °C [97]); TLC: 4:1 EtOAc/MeOH, R_f = 0.40; IR (KBr) /cm-1: 3450 (COOH), 3390 (NH), 1727 (OC=O), 1630 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.05 (*app bs*, 1H), 4.11 (*s*, 1H), 3.70 (*d*, *J* = 11.5, 1H), 3.64-3.56 (*m*, 1H), 3.53-3.45 (*m*, 1H), 3.29 (*d*, *J* = 11.5, 1H), 2.62 (*dt*, *J* = 6.0, 2.0, 2H), 1.46 (*s*, 3H), 1.43 (k, 3H) 1.04 (*s*, 3H), 0.98 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 170.2, 99.1, 77.1, 71.4, 34.1, 33.9, 33.0, 30.9, 29.4, 22.0, 18.8, 18.7; HR ESI-MS : 282.1322 ($[M+Na]^+$, $C_{12}H_{21}NNaO_5^+$, calc. 282.1312).

(4*R*)-*N*-(3-[(2-Hydroxyethyl)amino]-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide 283 [35]

N-([(4*R*)-2,2,5,5-Tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **268** hemi-calcium salt (2.59 g, 10 mmol) was dissolved together with triethylamine (3.0 mL, 2 eq.) in anhydrous dichloromethane (50 mL) at 0 °C. Ethylchloroformate (1.05 mL, 1.1 eq.) was added dropwise and the corresponding anhydride was formed over 30 min before adding 2-aminoethanol (1.20 mL, 2 eq.). The reaction was stirred at 0 °C for 30 min followed by 3 h at room temperature. Evaporation of the solvent gave the crude product which was purified by chromatography on a SNAP-100 column using a gradient of 0–10% methanol in ethyl acetate. Fractions containing the product were

pooled and evaporated to give (4*R*)-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** as a white solid (1.65g, 55%).

mp 172–175 °C; TLC: 9:1 EtOAc/MeOH, R_f = 0.25; IR (neat) ν/cm^{-1}: 3303 (NH, OH), 1739 (NC=O), 1638 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.06 (*t*, *J* = 6.0, 1H), 6.68 (*app bs*, 1H), 4.07 (*s*, 1H), 3.72 (*d*, *J* = 5.0, 1H), 3.69 (*d*, *J* = 11.5, 1H), 3.58 (*dt*, *J* = 11.5, 6.0, 1H), 3.43 (*ddt*, 1H), 3.29 (*d*, *J* = 11.5, 1H), 2.97 (*t*, *J* = 5.5, 1H), 2.49 (*t*, *J* = 6.0, 2H), 1.47 (*s*, 3H), 1.43 (*s*, 3H), 1.04 (*s*, 3H), 0.98 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.9, 170.6, 99.1, 77.1, 71.3, 62.0, 42.5, 36.3, 35.0, 32.9, 29.4, 22.1, 18.8, 18.7; HR ESI-MS: 325.1734 ([M+Na]$^+$, $C_{14}H_{26}N_2NaO_5^+$, calc. 325.1726).

3-(Allyloxy)-3-oxopropanoic acid 276 [78]

3-(Allyloxy)-3-oxopropanoic acid **276** was obtained by adapting a previously reported procedure [78]. 2,2-Dimethyl-dioxane-4,6-dione **274** (1.44 g, 10 mmol) and allyl alcohol (0.68 mL, 1 eq.) were stirred under reflux conditions in dried acetonitrile (5 mL) for 24 h under a nitrogen atmosphere. The solvent was then evaporated to give 3-(allyloxy)-3-oxopropanoic acid **276** as a colourless liquid (1.44 g, app. quant.).

IR (neat) ν/cm^{-1}: 1714 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.23 (*app bs*, 1H), 5.96-5.86 (*m*, 1H), 5.37-5.25 (*m*, 2H), 4.66 (*app. d*, *J* = 6.0, 2H), 3.47 (*s*, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.9, 166.2, 131.2, 119.0, 66.33, 40.86; HR ESI-MS: 167.0319 ([M+Na]$^+$, $C_6H_8NaO_4^+$, calc. 167.0315).

3-(Allyloxy)-2-methyl-3-oxopropanoic acid 277

3-(Allyloxy)-2-methyl-3-oxopropanoic acid **277** was synthesised following the same procedure as for compound **276** [78]. 2,2,5-Trimethyl-dioxane-4,6-dione **275** (1.58g, 10 mmol) was dissolved in dried acetonitrile (5 mL) and allyl alcohol (5 mL) and the reaction was stirred under reflux conditions for 24 h under a nitrogen atmosphere. The solvent was then evaporated to give 3-(allyloxy)-2-methyl-3-oxopropanoic acid **277** as a colourless oil (1.56 g, app. quant.).

IR (neat) ν/cm^{-1}: 1722 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.48 (*app bs*, 1H), 5.96-5.86 (*m*, 1H), 5.34 (*dd*, *J* = 17.5, 1.5, 1H), 5.25 (*dd*, *J* = 10.5, 1.5, 1H), 4.66 (*d*, *J* = 6.0, 2H), 3.53 (*q*, *J* = 7.5, 1H), 1.47 (*d*, *J* = 7.5, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 175.9, 169.4, 131.4, 118.7, 66.1, 45.9, 13.5; HR ESI-MS: 181.0473 ([M+Na]$^+$, $C_7H_{10}NaO_4^+$, calc. 181.0471).

Allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl malonate 284 [35]

Malonic acid mono allyl ester **276** (175 mg, 1.2 eq.) and *N*,*N*-diisopropylethylamine (0.26 mL, 1.5 eq.) were added to a solution of (4*R*)-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** (305 mg, 1.0 mmol) in dried tetrahydrofuran (20 mL). The solution was cooled to 0 °C, *O*-Benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluorophosphate (HBTU, 455 mg, 1.2 eq.) was added and the reaction stirred at room temperature for 16 h. The solvent was evaporated and replaced by ethyl acetate (50 mL), the organic phase was washed twice with saturated sodium hydrogen carbonate solution (2 x 20 mL) dried over magnesium sulphate, filtered and the filtrate evaporated to give allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl malonate **284** as a colourless oil (300 mg, 70%).

TLC: 19:1 EtOAc/MeOH, R_f = 0.55; IR (neat) ν/cm^{-1}: 3413 (NH), 3308 (NH), 1734 (C=O), 1658 (C=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.05 (*t*, *J* = 6.0, 1H), 6.53 (*t*, *J* = 5.5, 1H), 5.90 (*ddt*, *J* = 17.0, 11.0, 5.5, 1H), 5.33 (*dd*, *J* = 17.0, 1.0, 1H), 5.26 (*dd*, *J* = 11.0, 1.0, 1H), 4.64 (*dt*, *J* = 5.5, 1.0, 2H), 4.23 (*m*, 2H), 4.05 (*s*, 1H), 3.66 (*d*, *J* = 11.5, 1H), 3.58-3.44 (*m*, 4H), 3.43 (*s*, 2H), 3.26 (*d*,

J = 11.5, 1H), 2.43 (*t*, *J* = 6.0, 2H), 1.44 (*s*, 3H), 1.40 (*s*, 3H), 1.01 (*s*, 3H), 0.95 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.3, 170.1, 166.7, 166.3, 131.2, 119.0, 99.0, 71.3, 66.3, 64.1, 41.2, 38.5, 38.1, 35.8, 34.6, 32.8, 29.4, 22.0, 18.8, 18.6; HR ESI-MS: 451.2058 ($[M+Na]^+$, $C_{20}H_{32}N_2NaO_8^+$, calc. 451.2051).

3-Oxo-3-2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethoxypropanoic acid 286 [35]

Allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl malonate **284** (295 mg, 0.69 mmol) was dissolved in acetonitrile (10 mL) and the solution was cooled to 0 °C, prior to addition of pyrrolidine (25 μL, 1 eq.) and tetrakis (triphenylphosphine)-palladium (33 mg, 10% mol/mol). The reaction was found to be complete after 2 h, as judged by TLC, and the mixture filtered through a pad of Celite® 545. The filtrate was evaporated, the residue dissolved in ethyl acetate and extracted twice with water. The combined aqueous phases were freeze-dried to give compound **286** as a white solid (247 mg, 78%). For characterisation purposes, a small portion was purified by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 15 min, R_t = 17.0 min). The remaining crude product was used for the next step without further purification.

^{1}H NMR (400 MHz, D_2O): δ = 4.21 (*t*, *J* = 5.5, 2H), 3.94 (*s*, 1H), 3.52-3.40 (*m*, 5H), 3.34 (*d*, *J* = 11.0, 1H), 2.46 (*t*, *J* = 6.5, 2H), 2.18 (*s*, 6H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 175.5, 174.6, 174.5, 169.5, 76.1, 68.7, 64.5, 38.9, 38.5, 35.8, 35.6, 30.6, 20.8, 19.4; HR ESI-MS: 411.1730 ($[M+Na]^+$, $C_{17}H_{28}N_2NaO_8^+$, calc. 411.1738).

3-[2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-3-oxopropanoic acid 264 [35]

3-Oxo-3-2-[(*N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethoxypropanoic acid **286** (88 mg, 0.23 mmol) was dissolved in water (10 mL) and DOWEX 50WX8-400 (H^+) resin (100 mg), previously washed with 1 M hydrochloric acid and then with water to pH 7.0. The resulting suspension was stirred at room temperature for 12 h and then filtered. The filtrate was evaporated and the residue was purified by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 15 min) to give 3-[2-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-3-oxopropanoic acid **264** (R_t = 14.0 min, 65 mg, 82%).

^{1}H NMR (400 MHz, D_2O): δ = 4.21 (*t*, *J* = 5.5, 2H), 3.93 (*s*, 1H), 3.52-3.40 (*m*, 5H), 3.34 (*d*, *J* = 11.0, 1H), 2.45 (*t*, *J* = 6.5, 2H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 175.5, 174.5, 171.2, 169.4, 76.1, 68.7, 64.6, 38.9, 38.5, 35.8, 35.6, 20.8, 19.4; HR ESI-MS: 371.1422 ($[M+Na]^+$, $C_{14}H_{24}N_2NaO_8^+$, calc. 371.1425).

Allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl methylmalonate 285 [36]

(4*R*)-*N*-(3-[(2-Hydroxyethyl)amino]-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** (305 mg, 1.0 mmol) was dissolved in dried tetrahydrofuran (20 mL) at room temperature together with methyl malonic acid mono allyl ester **277** (190 mg, 1.2 eq.) and *N*,*N*-diisopropylethylamine (0.26 mL, 1.5 eq.). The solution was cooled to 0 °C, *O*-benzotriazole-*N*,*N*,*N*′,*N*′tetramethyluroniumhexafluoro-phosphate (455 mg, 1.2 eq.) was added and the reaction stirred at room temperature for 16 h. The solvent was evaporated and replaced by ethyl acetate (50 mL), the organic phase was washed with saturated sodium hydrogen carbonate solution (20 mL) and water (20 mL), dried over magnesium sulphate, filtered and the filtrate evaporated to give allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl methylmalonate **285** as a colourless oil (313 mg, 70%).

TLC: 19:1 EtOAc/MeOH, R_f = 0.65; IR (neat) ν/cm^{-1}: 3415 (NH), 3310 (NH), 1732 (C=O), 1659 (C=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.05 (*t*, *J* = 6.0, 1H), 6.39 (*t*, *J* = 5.0, 1H), 5.95-5.85 (*m*, 1H), 5.33 (*dd*, *J* = 17.0, 1.5, 1H), 5.26 (*dd*, *J* = 10.0, 1.0, 1H), 4.64 (*dt*, *J* = 6.0, 1.5, 2H), 4.35-4.28 (*m*, 1H), 4.18-4.10 (*m*, 1H), 4.06 (*d*, *J* = 1.5, 1H), 3.67 (*d*, *J* = 11.5, 1H), 3.59-3.46 (*m*, 5H), 3.26 (*d*, *J* = 11.5, 1H), 2.42 (*t*, *J* = 6.0, 2H), 1.45 (*s*, 3H), 1.44 (*d*, *J* = 7.5, 3H), 1.41 (*s*, 3H), 0.98 (*s*, 3H), 0.92 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.3, 170.1, 170.0, 169.7, 131.4, 118.7, 98.9, 71.3, 66.0, 63.8, 45.9, 38.5, 38.1, 35.7, 34.6, 32.8, 29.3, 22.0, 18.7, 18.6, 13.4; HR ESI-MS: 465.2207 ([M+Na]$^+$, $C_{21}H_{34}N_2NaO_8^+$, calc. 465.2192).

2-Methyl-3-oxo-3-2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethoxypropanoic acid 287 [36]

Allyl 2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl methylmalonate **285** (295 mg, 0.69 mmol) was dissolved in acetonitrile (10 mL) and the solution was cooled to 0 °C, prior to addition of pyrrolidine (25 μL, 1 eq.) and tetrakis (triphenylphosphine)-palladium (33 mg, 10% mol/mol). The reaction was found to be complete after 2 h, as judged by TLC, and the mixture filtered through a pad of Celite® 545. The filtrate was evaporated, the residue dissolved in ethyl acetate and extracted twice with water. The combined aqueous phase was freeze-dried to give compound **287** as a white solid (247 mg). For characterisation purposes, a small portion was purified by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 15 min, R_t = 17.5 min).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.26-4.17 (*m*, 1H), 3.94 (*s*, 1H), 3.59 (*q*, *J* = 7.0, 1H), 3.52-3.39 (*m*, 5H), 3.34 (*d*, *J* = 11.0, 1H), 2.45 (*t*, *J* = 6.5, 2H), 2.18 (*s*, 6H), 1.35 (*d*, *J* = 7.0, 3H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 175.5, 174.8, 174.5, 172.8, 76.1, 68.7, 64.5, 46.7, 38.9, 38.5, 35.8, 35.6, 30.8, 20.8, 19.4, 13.4, 13.3; HR ESI-MS: 425.1894 ([M+Na]$^+$, $C_{18}H_{30}N_2NaO_8^+$, calc. 425.1894).

3-[2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-2-methyl-3-oxopropanoic acid 265 [36]

2-Methyl-3-oxo-3-2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethoxypropanoic acid **287** (26.6 mg, 56 μmol) was dissolved in water (10 mL) and DOWEX 50WX8-400 (H^+) resin (100 mg), previously washed with 1N HCl and then water to pH 7.0. The resulting suspension was stirred at room temperature for 12 h and then filtered. The filtrate was evaporated and the residue was purified by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 15 min) to give 3-[2-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-2-methyl-3-oxopropanoic acid **265** as a colourless oil (R_t = 14.5 min, 17.2 mg, 85%).

^{1}H NMR (400 MHz, D_2O): δ = 4.25-4.18 (*m*, 1H), 3.93 (*s*, 1H), 3.60 (*q*, *J* = 7.0, 1H), 3.52-3.39 (*m*, 5H), 3.34 (*d*, *J* = 11.0, 1H), 2.45 (*t*, *J* = 6.5, 2H), 1.35 (*d*, *J* = 7.0, 3H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 175.5, 174.6, 174.5, 172.8, 76.1, 68.7, 64.5, 46.6, 38.9, 38.5, 35.8, 35.6, 20.8, 19.4, 13.3; HR ESI-MS: 385.1575 ([M+Na]$^+$, $C_{15}H_{26}N_2NaO_8^+$, calc. 385.1581).

Dibenzyl (2*S*)-5-(2-oxo-2-(2-[(*N*-[(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl-β-alanyl)amino]ethoxy)ethyl)pyrrolidine-1,2-dicarboxylate 339

(4*R*)-*N*-(3-[(2-Hydroxyethyl)amino]-3-oxopropyl)-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** (15.3 mg, 50 μmol) was dissolved in dried tetrahydrofuran together with (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** (14 mg, 35 μmol) and diisopropylethylamine (17 μL, 3.0 eq.). The solution was cooled to 0 °C and 2-(1H-benzotriazole-1-yl)-1,1,3,3-

tetramethyluronium hexafluorophosphate (19 mg, 50 mol) was added before the reaction was allowed to reach room temperature. After stirring for 16 h, the solvent was evaporated and the residue purified by silica chromatography (SNAP-25 column, gradient 0–10% methanol in dichloromethane) to yield dibenzyl (2*S*)-5-(2-oxo-2-2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethoxyethyl)pyrrolidine-1,2-dicarboxylate **339** as a colourless oil (9.5 mg, 41%). The purified product gives a complicated NMR pattern due to the presence of two diastereomers and multiple conformational isomers. The signals are reported as observed.

TLC: 9:1 EtOAc/MeOH, R_f = 0.75; IR (neat) ν/cm^{-1}: 3334 (NH), 3067, 3034 (CH$_{aryl}$), 1742 (OC=O), 1706 (NC=O), 1669 (NC=O); ^{1}H NMR (400 MHz, D$_2$O): δ = 7.41-7.26 (*m*, 10H), 7.15 (*app bs*, 0.7H), 7.09 (*app bs*, 0.3H), 6.71 (*app bs*, 0.7H), 6.21 (*app bs*, 0.3H), 5.27-5.18 (*m*, 1H), 5.16-5.00 (*m*, 3H), 4.60-4.39 (*m*, 2H), 4.31-4.26 (*m*, 0.7H), 4.20-4.07 (*m*, 1.3H), 3.72 (*d*, *J* = 11.5, 1H), 3.68-3.49 (*m*, 2H), 3.33 (*d*, *J* = 11.5, 1H), 3.10-3.02 (*m*, 0.3H), 2.86-2.76 (*m*, 1H), 2.64-2.58 (*m*, 0.3H), 2.53 (*dd*, *J* = 15.0, 5.0, 0.7H), 2.48-2.43 (*m*, 2H), 2.37-2.19 (*m*, 3H), 2.14-1.96 (*m*, 2H), 1.89-1.97 (*m*, 2H), 1.52 (*s*, 3H), 1.46 (*s*, 3H), 1.08 (*s*, 3H), 1.03 (*s*, 3H); ^{13}C NMR (100 MHz, D$_2$O): δ = 172.1, 171.9, 171.4, 171.2, 171.0, 170.9, 170.2, 170.0, 154.4, 136.3, 136.0, 135.3, 128.6-127.6 (18 signals), 99.0, 77.1, 71.5, 71.4, 67.4, 67.2, 66.9, 66.8, 63.6, 63.5, 63.4, 60.0, 59.9, 59.7, 55.9, 55.5, 55.3, 54.6, 51.7, 39.6, 38.9, 38.5, 38.4, 35.9, 35.7, 34.8, 34.7, 33.0, 30.6, 29.4, 29.2, 28.8, 28.3, 22.1, 18.9, 18.7; HR ESI-MS: 704.3158 ([M+Na]$^+$, C$_{36}$H$_{47}$N$_3$NaO$_{10}$$^+$, calc. 704.3154).

5-(2-Oxo-2-(2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethoxy)ethyl)-L-proline 340

Compound **339** (9.5 mg, 20 μmol) was dissolved in tetrahydrofuran (2 mL), and water (1 mL) was added, followed by palladium on activated carbon (2 mg, 20% w/w), and the reaction was placed under a hydrogen gas atmosphere for 4 h. After reaction completion, as judged by TLC, the mixture was filtered through a pad of Celite® 545 and the clear resulting solution was evaporated to give compound **340** as a colourless oil (6.0 mg, 93%).

^{1}H NMR (400 MHz, D$_2$O): δ = 4.20 (*t*, *J* = 5.5, 2H), 4.13 (*s*, 1H), 3.76-3.71 (*m*, 2H), 3.50-3.43 (*m*, 3H), 3.26 (*d*, *J* = 11.5, 1H), 2.92-2.89 (*m*, 1H), 2.45 (*t*, *J* = 6.5, 2H), 2.33-2.23 (*m*, 2H), 1.89-1.86 (*m*, 2H), 1.64-1.58 (*m*, 2H), 1.45 (*s*, 3H), 1.44 (*s*, 3H), 0.99 (*s*, 3H), 0.97 (*s*, 3H); HR ESI-MS: 480.2310 ([M+Na]$^+$, C$_{21}$H$_{35}$N$_3$NaO$_8$$^+$, calc. 480.2316).

Dibenzyl (2*S*)-5-[2-oxo-2-((2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)amino)ethyl]pyrrolidine-1,2-dicarboxylate 320

N-([(4*R*)-2,2,5,5-Tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **268** (72 mg, 1.2 eq.) was dissolved in anhydrous dichloromethane (2 mL) and triethylamine (0.12 mL, 3.6 eq.) was added followed by benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (175 mg, 1.5 eq.); the solution was then stirred at room temperature for 30 min. A solution of dibenzyl (2*S*)-5-2-[(2-aminoethyl)amino]-2-oxoethylpyrrolidine-1,2-dicarboxylate **338** (100 mg, 0.23 mmol) in anhydrous dichloromethane (2 mL) was then added dropwise and the reaction was stirred at room temperature under a nitrogen atmosphere for 16 h. The mixture was then diluted with dichloromethane (20 mL) and washed with saturated sodium hydrogen carbonate aqueous solution (20 mL) followed by water (20 mL). The combined aqueous layers were extracted with dichloromethane (20 mL) and the combined organic phases were dried over magnesium sulphate, filtered and evaporated to give the crude which was purified using silica chromatography (SNAP-25 column, gradient elution with 0–100% (10% methanol in ethyl acetate) in *n*-hexane) yielding dibenzyl (2*S*)-5-[2-oxo-2-((2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)amino)ethyl]pyrrolidine-1,2-dicarboxylate **320** as a white solid (110 mg, 58%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.30; IR (KBr) ν/cm^{-1}: 3342 (NH), 3055, 3028, 3012 (CH$_{aryl}$),

1753 (OC=O), 1713 (NC=O), 1689 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.36-7.20 (*m*, 10H), 5.22-4.93 (*m*, 4H), 4.42-4.38 (*m*, 2H), 4.06 (*s*, 1H), 3.66 (*d*, *J* = 11.5, 1H), 3.60-3.42 (*m*, 2H), 3.31 (*app bs*, 2H), 3.25 (*d*, *J* = 11.5, 1H), 3.15 (*m*, 1H), 2.68 (*dd*, *J* = 14.0, 4.5, 0.65H), 2.61 (*dd*, *J* = 14.0, 3.0, 0.35H), 2.42-2.39 (*m*, 2H), 2.34-2.21 (*m*, 2H), 2.13-2.08 (*m*, 1H), 1.98-1.94 (*dd*, *J* = 12.5, 6.5, 1H), 1,89-1.94 (*m*, 1H), 1.45 (*s*, 3H), 1.40 (*s*, 3H), 0.99 (*s*, 3H), 0.95 (*s*, 3H); HR ESI-MS : 703.3302 ($[M+Na]^+$, $C_{14}H_{16}N_2NaO_5^+$, calc. 703.3314).

_N_-[(2_R_)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanine 278 [35]

The hemicalcium salt of D-pantothenic acid **246** (4.80 g, 20.0 mmol) was suspended in acetic anhydride (100 mL) with a catalytic amount of iodine (0.30 mg, 1.2 mmol). The reaction mixture was stirred at 0 °C for 2 h followed by 48 h at room temperature. The solvent was evaporated and the residue was redissolved in dichloromethane (200 mL) before washing with a 1M sodium thiosulphate aqueous solution (100 mL). The phases were separated and the organic layer was dried over magnesium sulphate, filtered and evaporated giving the corresponding anhydride as pale yellow oil. This crude material was dissolved in a mixture of tetrahydrofuran and water (2:1, 50 mL) and stirred vigorously at room temperature for 16 h. After removal of the solvent, *N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanine **278** was obtained as a white solid (5.85 g, 97%).

mp 82–84 °C; IR (KBr) ν/cm^{-1}: 3367 (COOH), 1735 (OC=O), 1690 (OC=O), 1652 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.74 (*app bs*, 1H), 6.78 (*app bs*, 1H), 5.29 (*d*, *J* = 4.0, 1H), 4.96 (*d*, *J* = 4.0, 1H), 4.02 (*dd*, *J* = 11.0, 4.0, 1H), 3.82 (*dd*, *J* = 11.0, 4.0, 1H), 3.60-3.53 (*m*, 1H), 3.50-3.42 (*m*, 1H), 2.57 (*dd*, *J* = 9.5, 5.5, 2H), 2.13 (*d*, *J* = 4.0, 3H), 2.06 (*d*, *J* = 4.0, 3H), 1.05 (*d*, *J* = 4.0, 3H), 1.02 (*d*, *J* = 4.0, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.2, 171.1, 169.9, 168.3, 69.2, 53.4, 37.1, 34.5, 33.3, 21.3, 20.8, 20.7, 20.6; HR ESI-MS: 326.1209 ($[M+Na]^+$, $C_{13}H_{21}NNaO_7^+$, calc. 326.1210).

***tert*-Butyl 4-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)butanoate 289** [35]

N-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanine **278** (4.65 g, 15.3 mmol), *tert*-butyl-δ-aminobutanoate hydrochloride (3.3 mg, 1.1 eq.) and 1-hydroxybenzotriazole hydrate (2.27 g, 1.1 eq.) were dissolved in anhydrous dichloromethane (20 mL) before adding triethylamine (4.3 mL, 2 eq.) under a nitrogen atmosphere. The reaction mixture was cooled down to 0 °C and *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride (2.85 g, 1.2 eq.) was added. After stirring for 30 min on ice and then at room temperature for 18 h, the mixture was diluted with dichloromethane (50 mL) and washed with a saturated sodium hydrogen carbonate aqueous solution (50 mL) followed by water (50 mL). The organic layer was dried over magnesium sulphate, filtered and evaporated to give *tert*-butyl 4-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)butanoate **289** as a white solid (6.33 g, 93%).

mp 50–55 °C; TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.50; IR (KBr) ν/cm^{-1}: 3261 (NH), 3093 (NH), 1736 (NC=O), 1727 (NC=O), 1672 (OC=O), 1643 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.06 (*t*, *J* = 5.6, 1H), 6.44 (*t*, *J* = 5.0, 1H), 4.86 (*s*, 1H), 3.96 (*d*, *J* = 11.0, 1H), 3.78 (*d*, *J* = 11.0, 1H), 3.56-3.40 (*m*, 2H), 3.25-3.15 (*m*, 2H), 2.31 (*t*, *J* = 6.0, 2H), 2.21 (*t*, *J* = 7.0, 2H), 2.09 (*s*, 3H), 2.00 (*s*, 3H), 1.80-1.70 (*m*, 2H), 1.38 (*s*, 9H), 1.00 (*s*, 3H), 0.96 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.5, 171.5, 170.8, 169.8, 167.9, 80.4, 76.8, 69.2, 38.8, 36.9, 35.1, 34.8, 32.8, 27.9, 24.5, 21.2, 20.7, 20.6, 20.5; HR ESI-MS: 467.2350 ($[M+Na]^+$, $C_{21}H_{36}N_2NaO_8^+$, calc. 467.2364).

4-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)butanoic acid 290 [35]

tert-Butyl 4-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)butanoate **289** (1.5 g, 3.37 mmol) was dissolved in anhydrous dichloromethane (18 mL) and trifluoroacetic acid (2 mL)

was added dropwise. After stirring at room temperature for 6 h, the excess of trifluoroacetic acid was co-evaporated several times with toluene and then with dichloromethane to yield the free acid **290** as a pale orange oil (1.26 g, 96%) which requiered no further purification.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.92 (*app bs*, 1H), 7.31 (*app bs*, 1H), 7.05 (*app bs*, 1H), 4.87 (*s*, 1H), 4.01 (*d*, *J* = 11.0, 1H), 3.84 (*d*, *J* = 11.0, 1H), 3.52-3.51 (*m*, 2H), 3.35-3.21 (*m*, 2H), 2.48 (*t*, *J* = 6.0, 2H), 2.39 (*t*, *J* = 7.0, 2H), 2.13 (*s*, 3H), 2.06 (*s*, 3H), 1.89-1.78 (*m*, 2H), 1.04 (*s*, 3H), 1.02 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.7, 172.8, 171.1, 170.3, 169.3, 76.7, 69.2, 39.1, 37.1, 35.7, 35.0, 31.3, 24.1, 21.2, 20.7, 20.6, 20.5; HR ESI-MS: 411.1736 ($[M+Na]^+$, $C_{17}H_{28}N_2NaO_8^+$, calc. 411.1738).

(2*R*)-4-Acetoxy-1-[(3-([4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl]amino)-3-oxopropyl)amino]-3,3-dimethyl-1-oxobutan-2-yl acetate 291 [35]

4-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)butanoic acid **290** (1.20 g, 3.09 mmol) was dissolved into anhydrous dichloromethane (30 mL) together with 2,2-dimethyl-1,3-dioxane-4,6-dione **274** (450 mg, 1.01 eq.) and 4-dimethylaminopyridine (450 mg, 1.2 eq.). The mixture was stirred at 0 °C for 10 min and *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide hydrochloride was added (650 mg, 1.1 eq.). The temperature was slowly raised to room temperature and after 16 h, the reaction was diluted with dichloromethane (50 mL), washed with 0.1 M hydrochloric acid aqueous solution (20 mL), dried over magnesium sulphate, filtered and evaporated to yield compound **291** as a yellow oil (1.55 g) which was subjected to methanolysis without further purification.

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.08 (*t*, *J* = 6.0, 1H), 6.29 (*t*, *J* = 5.5, 1H), 4.87 (*s*, 1H), 4.02 (*d*, J = 11.0, 1H), 3.84 (*d*, *J* = 11.0, 1H), 3.52 (*apt q.*, *J* = 5.5, 2H), 3.44-3.36 (*m*, 2H), 3.29-3.21 (*m*, 2H), 3.16-3.00 (m, 2H), 2.46-2.34 (*m*, 2H), 2.14 (*s*, 3H), 2.05 (*s*, 3H), 1.98-1.85 (*m*, 2H), 1.75 (*s*, 3H), 1.74 (*s*, 3H), 1.06 (*s*, 3H), 1.03 (*s*, 3H); HR ESI-MS: 513.2092 ($[M-H]^-$, $C_{23}H_{33}N_2O_{11}^-$, calc. 513.2090).

Methyl 6-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate 292 [35]

(2*R*)-4-Acetoxy-1-[(3-([4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-ylidene)-4-hydroxybutyl]amino)-3-oxopropyl)amino]-3,3-dimethyl-1-oxobutan-2-yl acetate **291** (780 mg, 1.5 mmol) was dissolved in dried methanol (20 mL) and the solution was stirred under reflux conditions for 18 h. The solvent was then evaporated to give the crude methyl ester as a yellow oil. The residue was purified by silica chromatography (SNAP-50 column, gradient elution using 0–10% methanol in dichloromethane). Methyl 6-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate **292** was obtained as a colourless oil (520 mg, 77% over 2 steps).

TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.50; IR (KBr) ν/cm^{-1}: 1735 (NC=O), 1669 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.49 (*t*, *J* = 5.5, 1H), 6.07 (*t*, *J* = 5.5, 1H), 4.90 (*s*, 1H), 4.03 (*d*, *J* = 11.0, 1H), 3.84 (*d*, *J* = 11.0, 1H), 3.74 (*s*, 3H), 3.52-3.40 (*m*, 4H), 3.25-3.12 (*m*, 2H), 2.62 (*t*, *J* = 6.5, 2H), 2.36 (*t*, *J* = 5.5, 2H), 2.15 (*s*, 3H), 2.06 (*s*, 3H), 1.80-1.73 (*m*, 2H), 1.06 (*s*, 3H), 1.03 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 202.6, 171.8, 171.0, 170.0, 168.1, 167.8, 77.1, 69.3, 52.4, 48.9, 40.3, 38.7, 37.1, 35.2, 35.1, 23.0, 21.3, 20.9, 20.8, 20.7; HR ESI-MS : 467.2001 ($[M+Na]^+$, $C_{20}H_{32}N_2NaO_9^+$, calc. 467.2000).

Methyl 6-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate 293 [35]

Methyl 6-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate **292** (520 mg, 1.2 mmol) was dissolved in methanol (20 mL) on ice and potassium carbonate (250 mg,

1.5 eq.) was added prior to stirring for 5 h at 0 °C. The solvent was evaporated and the yellow residue dissolved in water (3 mL) and extracted with ethyl acetate (3 x 10 mL). The combined organic phases were dried over magnesium sulphate, filtered and evaporated to give methyl 6-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate **293** as a pale yellow oil (230 mg, 55%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.94 (*t*, *J* = 5.5, 1H), 6.70 (*t*, *J* = 5.5, 1H), 4.00 (*s*, 1H), 3.73 (*s*, 3H), 3.60-3.51 (*m*, 2H), 3.49 (*s*, 2H), 3.29-3.17 (*m*, 2H), 2.61 (*t*, *J* = 7.0, 2H), 2.43 (*t*, *J* = 6.0, 2H), 1.83-1.76 (*m*, 2H), 0.98 (*s*, 3H), 0.91 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 202.8, 173.9, 171.7, 168.0, 77.4, 70.7, 52.5, 48.9, 40.3, 39.3, 38.7, 35.8, 35.2, 23.1, 21.4, 20.4; HR ESI-MS : 383.1785 ($[M+Na]^+$, $C_{16}H_{28}N_2NaO_7^+$, calc. 383.1789).

6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2-methyl-3-oxohexanoate 332 [36]

Methyl 6-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate **292** (635 mg, 1.42 mmol) was dissolved in acetone (10 mL) at 0 °C, potassium carbonate (197 mg, 1 eq.) was added and the mixture was stirred for 5 min. Methyl iodide (130 μL, 1.5 eq.) was then added dropwise and the reaction mixture allowed to reach room temperature and kept under stirring conditions for 18 h. The solvent was evaporated and replaced by ethyl acetate (50 mL) and water (10 mL). The phases were separated and the organic layer was washed with brine (10 mL), dried over magnesium sulphate, filtered and evaporated to give the crude which was purified by chromatography (SNAP-100 column, gradient elution with 0–10% methanol in ethyl acetate). The dimethylated compound eluted first 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-2,2-dimethyl-3-oxohexanoate **331** (13 mg), followed by the monomethylated product **332** which was isolated as a colourless oil (223 mg, 34%).

TLC: 19:1 EtOAc/MeOH, R_f = 0.40; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.07 (*t*, *J* = 5.5, 0.25H), 7.01 (*app bs*, 0.75H), 6.35 (*app bs*, 1H), 4.86 (*s*, 1H), 3.98 (*d*, *J* = 11.0, 1H), 3.79 (*d*, *J* = 11.0, 1H), 3.68 (*s*, 2.2H), 3.64 (*s*, 0.8H), 3.54-3.42 (*m*, 3H), 3.19-3.14 (*m*, 2H), 2.65-2.45 (*m*, 2H), 2.34-2.30 (*m*, 2H), 2.10 (*s*, 3H), 2.01 (*s*, 3H), 1.78-1.69 (*m*, 2H), 1.29 (*d*, *J* = 7.5, 3H), 1.01 (*s*, 3H), 0.97 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 205.4, 171.6, 171.0, 170.8, 169.8, 168.0, 76.9, 69.2, 52.4, 52.3, 38.6, 38.2, 36.9, 35.2, 34.9, 23.0, 21.2, 20.7, 20.6 (2x), 12.7; HR ESI-MS : 481.2152 ($[M+Na]^+$, $C_{21}H_{34}N_2NaO_9^+$, calc. 481.2157).

(2*R*)-2,4-Dihydroxy-3,3-dimethyl-*N*-(3-oxo-3-[(4-oxohexyl)amino]propyl)butanamide 335

6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2-methyl-3-oxohexanoate **332** (22 mg, 48 μmol) was dissolved in methanol (2 mL) and the solution cooled to 0 °C. Potassium carbonate (250 mg, 1.5 eq.) was added prior to stirring for 16 h at room temperature. The solvent was evaporated and the residue subjected to HPLC purification (2% B in A over 5 min followed by 2–98% B in A over 15 min) to give the decarboxylated compound **335** (R_t = 15.5 min, 12.9 mg, 85%) as a mixture of equilibrating keto-enol tautomers, followed by methyl 6-((N-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2-methyl-3-oxohexanoate **333** (R_t = 17.5 min). When the same reaction was carried out at 0 °C over 5 h, compound **333** was found to be the only product.

(2*R*)-2,4-Dihydroxy-3,3-dimethyl-*N*-(3-oxo-3-[(4-oxohexyl)amino]propyl)butanamide **335**: ^{1}H NMR (400 MHz, D_2O): δ = 3.93 (*s*, 1H), 3.52-3.40 (*m*, 3H), 3.34 (*d*, *J* = 11.0), 3.12 (*t*, *J* = 7.0, 2H), 2.69 (*q*, *J* = 7.0, 0.4H), 2.55-2.48 (*m*, 3.2H), 2.43 (*t*, *J* = 6.5, 2H), 1.73-1.65 (*m*, 3.2H), 0.96 (*t*, *J* = 7.5, 1.8H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = = 218.7, 175.5, 174.2, 76.1, 68.7, 39.4, 39.1, 38.9, 36.1, 35.8, 35.7, 23.1, 20.8, 19.4, 7.6; HR ESI-MS : 339.1885 ($[M+Na]^+$, $C_{15}H_{28}N_2NaO_5^+$, calc. 339.1890).

Methyl 6-((N-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2-methyl-3-oxohexanoate **333**: ^{1}H NMR (400 MHz, D_2O): δ = 3.94 (*s*, 1H), 3.81 (*q*, *J* = 7.0, 0.3H), 3.53-3.40 (*m*, 3H), 3.34 (*d*, *J* = 11.0, 1H), 3.31 (*s*, 0.6H), 3.20 (*s*, 0.6H), 2.71-2.61 (*m*, 2H), 2.44 (*t*, *J* = 6.5, 2H), 1.75-1.68 (*m*, 2H), 1.35 (*s*, 1H), 1.28 (*d*, *J* = 7.0, 2H), 0.87 (*s*, 3H), 0.84 (*s*, 3H); HR ESI-MS : 397.1946 ($[M+Na]^+$, $C_{17}H_{30}N_2NaO_7^+$, calc. 397.1945).

6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2,2-dimethyl-3-oxohexanoate 331

Methyl 6-(*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanylamino)-3-oxohexanoate **292** (140 mg, 0.32 mmol) and powdered potassium carbonate (138 mg, 3 eq.) were mixed in dried dimethylformamide (10 mL) and the suspension cooled to 0 °C. Methyl iodide (95 μL, 5 eq.) was added and the reaction left at room temperature under stirring conditions for 18 h. Water (20 mL) and ethyl acetate (50 mL) were added and the phases separated. The organic layer was washed with brine (20 mL), dried over magnesium sulphate, filtered and evaporated to give the crude mixture, further purified by chromatography (SNAP-25 column, gradient elution using 0–10% methanol in ethyl acetate). 6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2,2-dimethyl-3-oxohexanoate **331** was isolated as a colourless oil (78 mg, 53%).

TLC: 19:1 EtOAc/MeOH, R_f = 0.45; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.00 (*app bs*, 1H), 6.24 (*app bs*, 1H), 4.88 (*s*, 1H), 4.00 (*d*, *J* = 11.0, 1H), 3.81 (*d*, *J* = 11.0, 1H), 3.70 (*s*, 3H), 3.56-3.40 (*m*, 2H), 3.25-3.12 (*m*, 2H), 2.49 (*t*, *J* = 6.5, 2H), 2.33 (*t*, *J* = 5.5, 2H), 2.12 (*s*, 3H), 2.03 (*s*, 3H), 1.80-1.73 (*m*, 2H), 1.34 (*s*, 6H), 1.03 (*s*, 3H), 0.99 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 207.7, 174.2, 171.6, 170.9, 169.9, 168.0, 76.9, 69.2, 55.4, 52.5, 38.7, 37.0, 35.3, 35.2, 35.0, 23.4, 21.9, 21.3, 20.8, 20.7, 20.6; HR ESI-MS : 495.2309 ($[M+Na]^+$, $C_{20}H_{32}N_2NaO_9^+$, calc. 495.2313).

Methyl 6-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2,2-dimethyl-3-oxohexanoate 334

6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2,2-dimethyl-3-oxohexanoate **331** (20 mg, 42 μmol) was dissolved in methanol (5 mL) at 0 °C and powdered potassium carbonate (18 mg, 3 eq.) was added. The reaction was stirred for 5 h at 0 °C, the solvent was evaporated and the residue subjected to HPLC purification (2% B in A over 5 min followed by 2–98% B in A over 15 min) to give compound **334** as a colourless oil (R_t = 17.5 min).

^{1}H NMR (500 MHz, D_2O): δ = 3.98 (*s*, 1H), 3.75 (*s*, 3H), 3.57-3.44 (*m*, 3H), 3.39 (*d*, *J* = 11.5, 1H), 3.15 (*t*, *J* = 7.0, 2H), 2.67 (*t*, *J* = 7.0, 2H), 2.48 (*t*, *J* = 6.5, 2H), 1.74 (*quint*, *J* = 7.0, 2H), 1.39 (*s*, 6H), 0.92 (*s*, 3H), 0.88 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 213.5, 175.5, 174.2, 76.2, 68.8, 56.4, 53.5, 39.0, 35.9, 35.7, 35.8, 23.2, 21.6, 20.9, 19.5; HR ESI-MS: 411.2102 ($[M+Na]^+$, $C_{18}H_{32}N_2NaO_7^+$, calc. 411.2102).

(2*R*)-2,4-Dihydroxy-*N*-(3-[(4-hydroxy-5-methylhex-4-en-1-yl)amino]-3-oxopropyl)-3,3-dimethylbutanamide 336

6-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-2,2-dimethyl-3-oxohexanoate **331** (20 mg, 42 μmol) was dissolved in methanol and powdered potassium carbonate (18 mg, 3 eq.) was added. The reaction was stirred for 16 h at room temperature, the solvent was evaporated and the residue subjected to HPLC purification (2% B in A over 5 min followed by 2–98% B in A over 30 min) to give the decarboxylated compound **336** as a $\cong$ 2:1 mixture of keto-enol tautomers (R_t = 19.0 min, 11.3 mg, 82%).

^{1}H NMR (500 MHz, D_2O): δ = 3.93 (*s*, 1H), 3.52-3.39 (*m*, 3H), 3.34 (*d*, *J* = 11.0, 1H), 3.11 (*t*, *J* = 7.0, 2H), 2.76-2.64 (*m*, 1.5H), 2.60 (*t*, *J* = 7.0, 1.5H), 2.43 (*t*, *J* = 6.5, 2H), 1.75-1.64 (*m*, 2H),

1.34 (*s*, 2H), 1.02 (*d*, *J* = 7.0, 4H), 0.87 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 175.6, 175.5, 174.2, 174.1, 76.1, 68.8, 41.2, 39.2, 39.1, 38.9, 37.7, 36.0, 35.8, 35.7 (2x), 23.0, 20.9, 19.4, 17.9, 17.8; HR ESI-MS: 397.1932 ($[M+Na]^+$, $C_{17}H_{30}N_2NaO_7^+$, calc. 397.1945).

***tert*-Butyl (2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate 305**

(*R*)-3-(2,2,5,5-Tetramethyl-1,3-dioxane-4-carboxamido)propanoic acid **268** (520 mg, 2.0 mmol) was dissolved in dried tetrahydrofuran (10 mL) together with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (621 mg, 2 eq.), hydroxybenzotriazole (770 mg, 2.5 eq.) and *tert*-butyl (2-aminoethyl)carbamate **271** (330 mg, 1.05 eq.) and the solution was cooled to 0 °C. *N*,*N*-Diisopropylethylamine (1.74 mL, 5 eq.) was the added dropwise and the mixture was stirred at room temperature for 18 h. The reaction was quenched with a 25% w/v ammonium chloride aqueous solution (20 mL) and after evaporation of tetrahydrofuran, the aqueous residue was extracted with diethyl ether (4 x 30 mL). Evaporation of the solvent gave an oily residue which was dissolved in ethyl acetate (50 mL) and washed with a saturated sodium hydrogen carbonate aqueous solution (20 mL). The organic phase was dried over magnesium sulphate, filtered and evaporated to give *tert*-butyl (2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate **305** as a low melting point white solid (773 mg, 96%).

TLC: 19:1 EtOAc/MeOH, R_f = 0.50; IR (neat) ν/cm^{-1}: 3325 (NH), 1697 (NC=O), 1655 (NC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.06 (*t*, *J* = 6.0, 1H), 6.81 (*app bs*, 1H), 5.22 (*app bs*, 1H), 4.05 (*s*, 1H), 3.66 (*d*, *J* = 11.5, 1H), 3.59-3.45 (*m*, 2H), 3.36-3.28 (*m*, 2H), 3.26-3.20 (*m*, 3H), 2.42 (*t*, *J* = 6.0, 2H), 1.44-1.40 (*m*, 15H), 1.00 (*s*, 3H), 0.94 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.6, 170.2, 156.7, 99.0, 79.5, 70.1, 71.3, 40.4, 40.2, 35.9, 34.8, 32.9, 29.4 (28.3), 22.1, 18.8, 18.6; HR ESI-MS: 424.2404 ($[M+Na]^+$, $C_{19}H_{35}N_3NaO_6^+$, calc. 424.2418).

***N*-([(4*R*)-2-(4-Methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine 269**

The hemicalcium salt of D-pantothenic acid **246** (4.80 g, 20.0 mmol) was suspended in tetrahydrofuran (150 mL) and *p*-toluene sulfonic acid (7.99 g, 1 eq.) was added in one portion, followed by anisaldehyde (5.1 mL, 1 eq.). The reaction mixture was stirred at room temperature for 10 h. After reaction completion, as judged by TLC, the white suspension was filtered on Celite® 545 and the solid washed several times with acetone. The filtrate was evaporated under reduced pressure and the resulting solid was washed several times with a warm 1:1 mixture of *n*-hexane and ethyl acetate to give *N*-([(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **269** as a white solid (6.50 g, 46%).

mp 82–85 °C; IR (neat) ν/cm^{-1}: 3383 (NH), 1724 (NC=O), 1618 (OC=O); TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.35; ^{1}H NMR (400 MHz, d_6-DMSO): δ = 7.44 (*d*, *J* = 8.5, 2H), 6.94 (*d*, *J* = 8.5, 2H), 5.53 (*s*, 1H), 4.10 (*s*, 1H), 3.77 (*s*, 3H), 3.64 (*dd*, *J* = 17.0, 11.5, 2H), 3.32-3.22 (*m*, 2H), 2.41 (*t*, *J* = 7.0, 2H), 1.01 (*s*, 3H), 0.95 (*s*, 3H); ^{13}C NMR (100 MHz, d_6-DMSO): δ = 173.2, 168.3, 159.7, 130.5, 127.8, 113.4, 100.5, 83.2, 77.4, 55.2, 34.3, 33.8, 32.6, 21.6, 19.1; HR ESI-MS: 360.1410 ($[M+Na]^+$, $C_{17}H_{23}NNaO_6^+$, calc. 360.1418).

***tert*-Butyl (2-[(*N*-([(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate 306**

N-([(4*R*)-2-(4-Methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **269** (2.0 g, 5.9 mmol) was dissolved in dry tetrahydrofuran (50 mL) together with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (2.3 g, 2 eq.) hydroxybenzotriazole (2.0 g, 2.5 eq.) and *tert*-butyl (2-aminoethyl)carbamate **271** (1.91 g, 1.5 eq.) and the solution was placed under

a nitrogen atmosphere. Diisopropylethylamine (5.2 mL, 5 eq.) was added dropwise and the reaction stirred for 16 h at room temperature before quenching with 25% ammonium chloride aqueous solution (100 mL), diluting with diethyl ether (200 mL) and extracting once more with diethyl ether (200 mL) after phase separation. The combined organic phases were washed with brine (100 mL) and water (100 mL), dried over magnesium sulphate, filtered and evaporated to give the crude which was purified by chromatography (SNAP-100) using ethyl acetate elution to yield *tert*-butyl (2-[(*N*-([(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate **306** as a low melting point white solid (2.48 g, 88%).

TLC: 19:1 CH_2Cl_2/MeOH, R_f = 0.50; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.44 (*d*, *J* = 8.5, 2H), 7.04 (*t*, *J* = 6.0, 1H), 6.92 (*d*, *J* = 8.5, 2H), 6.61 (*t*, *J* = 5.0, 1H), 5.46 (*s*, 1H), 4.98 (*app bs*, 1H), 4.09 (*s*, 1H), 3.81 (*s*, 3H), 3.73-3.64 (*m*, 2H), 3.55 (*q*, *J* = 6.0, 2H), 3.35-3.27 (*m*, 2H), 3.22-3.18 (*m*, 2H), 2.41 (*t*, *J* = 6.5, 2H), 1.43 (*s*, 9H), 1.10 (*s*, 3H), 1.09 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.4, 169.5, 160.2, 156.7, 130.1, 127.5, 113.7, 101.3, 83.8, 79.6, 78.4, 55.3, 40.5, 40.1, 36.0, 34.9, 33.0, 28.3, 21.8, 19.1; HR ESI-MS: 502.2526 ([M+Na]$^+$, $C_{24}H_{37}N_3NaO_7^+$, calc. 502.2524).

(2*R*)-*N*-3-[(2-Aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide 307

tert-Butyl (2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate **305** (200 mg, 0.5 mmol) was dissolved in dichloromethane (18 mL) and trifluoroacetic acid was added dropwise at room temperature. TLC monitoring showed reaction completion after 4 h and the solvent was evaporated under vacuum. The excess acid was removed by coevaporation with water followed by methanol to give the trifluoroacetate salt of (2*R*)-*N*-3-[(2-aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide **307** as a white solid (187 mg, app. quant.). For assay purposes, a small portion of the title amine **307** was purified by preparative HPLC (2% B in A over 5 min followed by 2–20% B in A over 15 min and 20–98% B in A over 5 min) to give (2*R*)-*N*-3-[(2-aminoethyl)amino]-3-oxopropyl-2,4-dihydroxy-3,3-dimethylbutanamide **307** as a colourless oil (R_t = 9.0 min)). Alternatively, compound **307** was obtained by deprotection of *tert*-butyl (2-[(*N*-([(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate **306** using the same acidic conditions described for *tert*-butyl (2-[(*N*-([(4*R*)-2,2,5,5-tetramethyl-1,3-dioxan-4-yl]carbonyl)-β-alanyl)amino]ethyl)carbamate **305**.

^{1}H NMR (400 MHz, D_2O): δ = 4.37 (*s*, 1H), 4.13 (*d*, *J* = 9.0, 1H), 4.08 (*d*, *J* = 9.0, 1H), 3.52 (*t*, *J* = 6.0, 2H), 3.27 (*t*, *J* = 6.5, 2H), 3.15 (*t*, *J* = 6.0, 2H), 2.72 (*t*, *J* = 6.5, 2H), 1.17 (*s*, 3H), 0.99 (*s*, 3H); ^{13}C NMR (100 MHz, D_2O): δ = 180.3, 173.6, 77.4, 75.9, 41.0, 39.7, 37.2, 36.0, 32.3, 21.9, 18.6; HR ESI-MS: 284.1582 ([M+Na]$^+$, $C_11H_23N_3NaO_4^+$, calc. 284.1581).

(14*R*)-14-Acetoxy-2,2,15,15-tetramethyl-4,9,13-trioxo-3-oxa-5,8,12-triazahexadecan-16-yl acetate 279

N-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanine **278** (1.30 g, 5.0 mmol) was dissolved in anhydrous dichloromethane (25 mL) followed by *tert*-butyl 2-aminoethylcarbamate (881 mg, 1.1 eq.), triethylamine (1.4 mL, 2 eq.) and hydroxybenzotriazole (745 mg, 1.1 eq.). The solution was cooled to 0 °C and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (1.15 g, 1.2 eq.) was added. The reaction was stirred at 0 °C for 30 min and at room temperature for 18 h. The mixture was then diluted with dichloromethane (50 mL) and the organic phase washed with 0.1 M aqueous hydrochloric acid (30 mL) followed by a sodium hydrogenocarbonate solution (30 mL), dried over magnesium sulphate, filtered and evaporated to give (14*R*)-14-acetoxy-2,2,15,15-tetramethyl-4,9,13-trioxo-3-oxa-5,8,12-triazahexadecan-16-yl acetate **279** as a white solid (1.93 g, 86%).

mp $\leq$ 35 °C; TLC: 19:1 EtOAc/MeOH, R_f = 0.50; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.19 (*app bs*, 1H), 6.71 (*app bs*, 1H), 5.33 (*app bs*, 1H), 4.83 (*s*, 1H), 4.02 (*d*, *J* = 11.0, 1H), 3.82 (*d*, *J* = 11.0,

1H), 3.56-3.48 (*m*, 1H), 3.44-3.40 (*m*, 2H), 3.30-3.18 (*m*, 3H), 2.43-2.29 (*m*, 2H), 2.13 (*s*, 3H), 2.04 (*s*, 3H), 1.40 (*s*, 9H), 1.04 (*s*, 3H), 1.01 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.1, 170.9, 170.4, 168.14, 156.8, 79.5, 77.2, 69.2, 40.3, 36.9, 35.5, 35.2, 28.3, 21.3, 20.8, 20.7, 20.4; HR ESI-MS: 468.2313 ($[M+Na]^+$, $C_{20}H_{35}N_3NaO_8{}^+$, calc. 468.2316).

(2*R*)-4-Acetoxy-1-((3-[(2-aminoethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate 280

The synthesis of compound **280** was previously reported using another procedure [36]. (*R*)-2,2,15,15-Tetramethyl-4,9,13-trioxo-3-oxa-5,8,12-triazahexadecane-14,16-diyl diacetate (700 mg, 1.73 mmol) was dissolved in dichloromethane (18 mL) and trifluoroacetic acid (2 mL) was added dropwise. The reaction was left under vigorous stirring at room temperature and the reaction was found to be completed in 4 h, as judged by TLC. The solvent was evaporated and excess trifluoroacetic acid was removed by coevaporation with water. The trifluoroacetate salt of the title amine **280** was obtained as a colourless oil (790 mg, app. quant.).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 10.69 (*app bs*, 1H), 7.88 (*app bs*, 2H), 7.48 (*app bs*, 1H), 4.74 (*s*, 1H), 3.96 (*d*, *J* = 11.0, 1H), 3.86 (*d*, *J* = 11.0, 1H), 3.49 (*app bs*, 2H), 3.18 (*app bs*, 1H), 2.43 (*app bs*, 1H), 2.09 (*s*, 3H), 2.03 (*s*, 3H), 1.00 (*s*, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 173.9, 171.4, 170.9, 169.5, 161.2, 117.2, 77.6, 69.1, 40.2, 37.3, 37.0, 35.7, 35.3, 20.9, 20.6, 20.4; HR ESI-MS: 346.1968 ($[M+Na]^+$, $C_{15}H_{28}N_3O_6{}^+$, calc. 346.1973).

3-Methoxy-3-oxopropanoic acid 272 [77]

The synthesis of 3-methoxy-3-oxopropanoic acid **272** was adapted from a previously described procedure [77]. 2,2-Dimethyl-1,3-dioxane-4,6-dione **274** (720 mg, 5.0 mmol) was dissolved in methanol (15 mL) and the reaction was stirred under reflux conditions for 4 d. The solvent was evaporated under reduced pressure to yield 3-methoxy-3-oxopropanoic acid **272** as a colourless oil (521 mg, 88% yield).

IR (neat) ν/cm^{-1}: 1718 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.71 (s, 3H), 3.29 (s, 2H); ^{13}C NMR (100 MHz, D_2O): δ = 174.6, 172.3, 52.9, 44.8; LR ESI-MS: 117.02 $[M-H]^-$.

3-Methoxy-2-methyl-3-oxopropanoic acid 273 [76]

3-Methoxy-2-methyl-3-oxopropanoic acid **273** was previously described using an alternative procedure [76] and was synthesised using the same procedure as for compound **272** [77]. 2,2,5-Trimethyl-dioxane-4,6-dione **275** (790 mg, 5.0 mmol) was dissolved in MeOH (15 mL) and the reaction was stirred under reflux conditions for 4 d. The solvent was evaporated under reduced pressure to yield 3-methoxy-2-methyl-3-oxopropanoic acid **273** as a colourless oil (554 mg, 84% yield).

IR (neat) ν/cm^{-1}: 1710 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 11.26 (*app bs*, 1H), 3.76 (*s*, 3H), 3.50 (*q*, *J* = 7.5, 1H), 1.45 (*d*, *J* = 7.5, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 175.9, 170.2, 52.7, 45.7, 13.5; LR ESI-MS: 131.03 $[M-H]^-$.

Methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoate 281

(2*R*)-4-Acetoxy-1-((3-[(2-aminoethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate **280** trifluoroacetate salt (300 mg, 0.65 mmol) was dissolved in dried tetrahydrofuran (10 mL) together with 3-methoxy-3-oxopropanoic acid **272** (130mg, 1.4 eq.), diisopropylethylamine (0.225 mL, 2.0 eq.) and hydroxybenzotriazole (123 mg, 1.4 eq.). The solution was cooled to

0 °C and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (190 mg, 1.5 eq.) was added. After stirring for 16 h under these conditions, the solvent was evaporated and replaced by dichloromethane (30 mL). The organic phase was washed two times with saturated sodium hydrogen carbonate aqueous solution (2 x 20 mL), dried over magnesium sulphate, filtered and evaporated to yield methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoate **281** as a colourless oil (62 mg, 22%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.20; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.42 (*app bs*, 1H), 7.09 (*app bs*, 1H), 6.70 (*app bs*, 1H), 4.82 (*s*, 1H), 4.01 (*d*, *J* = 11.0, 1H), 3.84 (*d*, *J* = 11.0, 1H), 3.73 (*s*, 3H), 3.54-3.34 (*m*, 4H), 3.31 (*s*, 2H), 2.43-2.30 (*m*, 2H), 2.14 (*s*, 3H), 2.05 (*s*, 3H), 1,05 (*s*, 3H), 1.03 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.4, 171.0, 170.6, 169.5, 168.2, 166.3, 77.4, 69.2, 53.5, 41.5, 39.7, 39.4, 36.9, 35.6, 35.4, 20.3, 20.8, 20.7, 20.6; HR ESI-MS: 468.1942 ($[M+Na]^+$, $C_{19}H_{31}N_3NaO_9^+$, calc. 468.1953).

3-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoic acid 260

Methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoate **281** (60 mg, 13 μmol) was dissolved in methanol (5 mL) and potassium carbonate (75 mg, 4.0 eq.) was added. The reaction was stirred at room temperature for 18 h, the solvent was evaporated and the residue dissolved in water. The pH was adjusted to 7.0 using 1 M hydrochloric acid solution and the aqueous phase freeze-dried to yield compound **260** as a colourless oil. This residue was desalted by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 30 min) to yield 3-([2-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoic acid **260** as a colourless oil (R_t = 15.5 min, 28 mg, 62%).

^{1}H NMR (500 MHz, D_2O): δ = 3.94 (*s*, 1H), 3.50-3.42 (*m*, 3H), 3.38-3.27 (*m*, 7H), 2.45 (*t*, *J* = 6.5, 2H), 0.88 (*s*, 3H), 0.84 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 175.5, 174.6, 172.2, 169.6, 76.1, 68.7, 42.5, 39.1, 38.9 (2x), 35.8, 35.7, 20.8, 19.4; HR ESI-MS: 346.1626 ($[M-H]^-$, $C_{14}H_{24}N_3O_7^-$, calc. 346.1620).

Methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoate 282 [36]

(2*R*)-4-Acetoxy-1-((3-[(2-aminoethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate **280** trifluoroacetate salt (325 mg, 0.71 mmol) was treated as above using 3-methoxy-2-methyl-3-oxopropanoic acid **273** (132 mg, 1.4 eq.). Reaction work-up gave a diastereometic mixture of methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoate **282** as a colourless oil (145 mg, 45%).

TLC: 9:1 EtOAc/MeOH, R_f = 0.20; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.12 (*app bs*, 2H), 6.74 (*app bs*, 1H), 4.83 (*d*, *J* = 2.0, 1H), 4.01 (*d*, *J* = 11.0, 1H), 3.84 (*d*, *J* = 11.0, 1H), 3.72 (*s*, 3H), 3.55-3.22 (*m*, 7H), 2.43-2.29 (*m*, 2H), 2.13 (*s*, 3H), 2.04 (*s*, 3H), 1.39 (*dd*, *J* = 7.5, 1.0, 3H), 1.05 (*s*, 3H), 1.02 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.4, 172.3, 172.2, 170.1, 170.5 (2x), 170.2 (2x), 168.2, 77.4, 77.3, 69.2, 52.5, 46.8, 46.7, 39.9, 39.3 (2x), 36.9 (2x), 35.6 (2x), 35.4 (2x), 21.3 (2x), 20.7 (2x), 20.5 (2x), 14.5, 14.3; HR ESI-MS: 482.2109 ($[M+Na]^+$, $C_{20}H_{33}N_3NaO_9^+$, calc. 482.2109).

3-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoic acid 261 [36]

Methyl 3-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoate **282** (40 mg, 87 μmol) was treated in the

same way as compound **281** above and the residue was desalted by preparative HPLC (2% B in A over 5 min followed by 2–98% B in A over 30 min) to yield 3-([2-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoic acid **261** as a colourless oil (R_t = 15.5 min, 22 mg, 70%).

^{1}H NMR (500 MHz, D_2O): δ = 3.94 (*s*, 1H), 3.50-3.42 (*m*, 4H), 3.36-3.28 (*m*, 5H), 2.4 (*t*, *J* = 6.5, 2H), 1.30 (*d*, *J* = 7.0, 3H), 0.88 (*s*, 3H), 0.84 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 175.5, 175.0, 174.5, 173.6, 76.1, 68.7, 47.2, 39.1, 38.9 (2x), 35.8, 35.7, 20.8, 19.4, 13.7; HR ESI-MS: 360.1779 ([M-H]$^-$, $C_{15}H_{26}N_3O_7^-$, calc. 360.1779).

Dibenzyl (2*S*)-5-(2-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)pyrrolidine-1,2-dicarboxylate 321

(2*R*)-4-Acetoxy-1-((3-[(2-aminoethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate **280** trifluoroacetate salt (280 mg, 1.8 eq.) was dissolved in anhydrous dichloromethane (10 mL) together with (5*S*)-1,5-bis[(benzyloxy)carbonyl]pyrrolidin-2-ylacetic acid **304** (139 mg, 0.35 mmol), triethylamine (0.35 mL, 7.0 eq.) and hydroxybenzotriazole (150 mg, 2.8 eq.). The solution was cooled to 0 °C and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (150 mg, 2.8 eq.) was added. After stirring for 2 h on ice and 16 h at room temperature, the solvent was evaporated and the residue was purified by silica chromatography (SNAP-100 column, gradient elution with a 0–10% methanol in ethyl acetate) to yield dibenzyl (2*S*)-5-(2-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **321** as a colourless oil (245 mg, 96%).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.37-7.20 (*m*, 10H), 7.17-7.15 (*m*, 0.6H), 7.07 (*app bs*, 0.4H), 6.96-6.78 (*m*, 1.7H), 6.59-6.52 (*m*, 0.3H), 5.29-4.78 (*m*, 5H), 4.48-4.35 (*m*, 2H), 4.05-4.01 (*m*, 1H), 3.86-3.83 (*m*, 1H), 3.56-3.40 (*m*, 3H), 3.36-3.30 (*m*, 2H), 3.27-3.12 (*m*, 1H), 2.70 (*dd*, *J* = 13.5, 4.0, 0.7H), 2.58 (*dd*, *J* = 13.5, 3.0, 0.3H), 2.43-2.20 (*m*, 5H), 2.14 (*s*, 3H), 2.05 (*s*, 3H), 2.01-1.88 (*m*, 2H), 1.06 (*s*, 3H), 1.04 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 172.2, 172.1, 171.6, 171.0, 170.5, 168.2, 154.5, 136.0, 135.3, 128.7, 128.6, 128.5, 128.4, 128.1, 128.0, 127.9, 127.6, 77.2, 69.3, 67.1, 66.8, 59.6, 56.4, 41.1, 39.7, 39.5, 36.7, 35.5, 35.2, 28.6, 28.3, 21.3, 20.8 (2x), 20.6; HR ESI-MS: 747.3216 ([M+Na]$^+$, $C_{37}H_{48}N_4NaO_{11}^+$, calc. 747.3212).

5-(2-([2-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)-L-proline 341

Dibenzyl (2*S*)-5-(2-([2-((*N*-[(2*R*)-2,4-diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)pyrrolidine-1,2-dicarboxylate **321** (69 mg, 95 mol) was dissolved in tetrahydrofuran (5 mL) and water (5 mL) was added, followed by palladium on activated carbon (7 mg, 10% w/w). The reaction was placed under an hydrogen atmosphere for 4 h. After reaction completion, as judged by TLC, the mixture was filtered through a pad of Celite® 545 and the resulting clear solution was evaporated to give compound **341** as a white foam (47 mg, app. quant.).

^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.81 (*s*, 1H), 4.11-4.04 (*m*, 1.7H), 4.00-3.89 (*m*, 1H), 3.85 (*d*, *J* = 11.0, 1H), 3.81-3.75 (*m*, 0.3H), 3.61-3.52 (*m*, 0.7H), 3.50-3.40 (*m*, 2H), 3.34-3.26 (*m*, 2H), 3.24-3.18 (*m*, 0.3H), 2.78-2.66 (*m*, 1.7H), 2.43-2.39 (*m*, 2H), 2.36-2.17 (*m*, 2.3H), 2.13 (*s*, 3H), 2.05 (*s*, 3H), 1.86-1.58 (*m*, 2H), 1.06 (*s*, 3H), 1.03 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 174.5, 174.2, 172.7, 172.4, 172.1, 170.9, 78.4, 70.5, 62.9, 59.0, 40.2, 40.1, 38.2, 38.1, 37.0, 36.6, 30.7, 29.8, 22.1, 21.1, 20.9, 20.8; HR ESI-MS: 523.2385 ([M+Na]$^+$, $C_{22}H_{36}N_4NaO_9^+$, calc. 523.2374).

5-(2-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)-L-proline 319

5-(2-([2-((*N*-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)-L-proline **341** (35 mg, 70 μmol) was dissolved in methanol (10 mL) and the solution cooled to 0 °C prior to addition of potassium carbonate (35 mg, 3.5 eq.). The reaction was stirred at 0 °C for 5 h and the solvent was evaporated. The residue was purified by preparative HPLC (2% B in A over 5 min followed by 2–20% B in A over 15 min) to give the product as a white solid (R_t = 16.8 min, 17.1 mg, 59%).

^{1}H NMR (400 MHz, D_2O): δ = 4.19-4.13 (*m*, 1H), 4.01-3.90 (*m*, 2H), 3.53-3.41 (*m*, 3H), 3.36 (*d*, *J* = 11.5, 1H), 3.32-3.26 (*m*, 4H), 2.77-2.72 (*dd*, *J* = 11.5, 7.0, 2H), 2.47-2.39 (*m*, 2.7H), 2.35-2.27 (*m*, 0.3H), 2.26-2.13 (*m*, 1.3H), 2.06-1.97 (*m*, 0.7H), 1.80-1.65 (*m*, 1H), 0.88 (*s*, 3H), 0.85 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 175.5, 174.6, 172.3, 172.2, 76.1, 68.7, 61.7, 61.3, 57.8, 57.6, 39.1, 39.0, 36.9, 36.7, 35.8 (2x), 30.2, 29.2, 28.6 (2x), 20.9, 19.4; HR ESI-MS: 439.2170 ([M+Na]$^+$, $C_{18}H_{32}N_4NaO_7^+$, calc. 439.2163).

(4*R*)-*N*-(3-[(2-Hydroxyethyl)amino]-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide 310

Ethylchloroformate (0.2 mL, 1 eq.) was added dropwise to a cold solution of *N*-([(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxan-4-yl]carbonyl)-β-alanine **269** (675 mg, 2.0 mmol) and triethylamine (0.7 mL, 2 eq.) in dichloromethane (10 mL). The reaction was stirred at 0 °C for 30 min. 2-Aminoethanol **267** (0.25 mL, 2 eq.) was then added, the reaction kept on ice for 30 min and then left at room temperature for 16 h. The solvent was then evaporated and replaced by ethyl acetate (10 mL). The insoluble white solid was filtered off, washed with ethyl acetate and the filtrate was evaporated to give the crude product. This residue was purified by silica chromatography (SNAP-25 column, gradient elution with 2–20% methanol in dichloromethane) to yield (4*R*)-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide **310** as a white solid (730 mg, 96%).

mp 119–122 °C; TLC: 9:1 CH_2Cl_2/MeOH, R_f = 0.35; IR (neat) ν/cm^{-1}: 3421 (OH), 1664 (OC=O); ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.44 (*d*, *J* = 8.5, 2H), 7.05 (*t*, *J* = 6.0, 1H), 6.93 (*d*, *J* = 8.5, 2H), 6.58 (*app bs*, 1H), 5.46 (*s*, 1H), 4.08 (*s*, 1H), 3.82 (*s*, 3H), 3.72 (*d*, *J* = 11.5, 1H), 3.67-3.63 (*m*, 3H), 3.55 (*q*, *J* = 6.5, 2H), 3.43-3.31 (*m*, 2H), 3.09 (*app bs*, 1H), 2.45 (*t*, *J* = 6.5, 2H), 1.11 (*s*, 3H), 1.09 (*s*, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 171.8, 169.9, 169.3, 130.1, 127.5, 113.8, 101.4, 83.8, 78.4, 62.0, 55.3, 42.4, 36.3, 35.1, 33.1, 21.8, 19.1; HR ESI-MS: 403.1839 ([M+Na]$^+$, $C_{19}H_{28}N_2NaO_6^+$, calc. 403.1840).

(2*R*)-4-Acetoxy-1-((3-[(2-hydroxyethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate 311

N-[(2*R*)-2,4-Diacetoxy-3,3-dimethylbutanoyl]-β-alanine **278** (305 mg, 1.0 mmol) was dissolved in dichloromethane (5 mL) and triethylamine (0.14 mL, 1 eq.) was added. The solution was cooled to 0 °C and ethylchloroformate (0.1 mL, 1 eq.) was added dropwise. The reaction was stirred for 10 min and 2-aminoethanol (0.12 mL, 2 eq.) was added. The reaction was kept at 0 °C for 30 min and at room temperature overnight. The reaction was quenched with 1 M aqueous hydrochloric acid solution (2 mL) and the phases were separated. The organic phase was washed once more with 1 M aqueous hydrochloric acid solution (2 mL), dried over magnesium sulphate and evaporated to yield (2*R*)-4-acetoxy-1-((3-[(2-hydroxyethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate **311** as a low melting point white solid (339 mg, 98%).

mp $\leq$ 35 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.30 (*t*, *J* = 6.0, 1H), 7.18 (*t*, *J* = 5.5, 1H), 4.79 (*s*, 1H), 3.94 (*d*, *J* = 11.0, 1H), 3.78 (*d*, *J* = 11.0, 1H), 3.46-3.40 (*m*, 2H), 3.35-3.20 (*m*, 3H), 2.39-2.36 (*m*, 2H), 2.08 (*s*, 3H), 1.99 (*s*, 3H), 1.18-1.14 (*m*, 2H), 0.98 (*s*, 3H), 0.95 (*s*, 3H); HR ESI-MS: 369.1630 ([M+Na]$^+$, $C_{15}H_{26}N_2NaO_7^+$, calc. 369.1632).

(2*R*)-2,4-Dihydroxy-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-3,3-dimethylbutanamide 309 [98]

(2*R*)-2,4-Dihydroxy-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-3,3-dimethylbutanamide **309** was previously synthesised using a different procedure [98]. (4*R*)-*N*-3-[(2-Hydroxyethyl)amino]-3-oxopropyl-2,2,5,5-tetramethyl-1,3-dioxane-4-carboxamide **283** (302 mg, 1.0 mmol) was dissolved in tetrahydrofuran (25 mL) and water (10 mL) was added followed by DOWEX 50WX8-400 ($^+$) resin (500 mg), previously washed with 1 M hydrochloric acid solution and then water to pH 7.0. The resulting suspension was stirred at room temperature for 16 h and then filtered. The resin was washed with water (5 mL) and methanol (5 mL) and the filtrate was evaporated to give (2*R*)-2,4-dihydroxy-*N*-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-3,3-dimethylbutanamide **309** as a colourless oil (260 mg, 99%).

Alternatively, (2*R*)-4-acetoxy-1-((3-[(2-hydroxyethyl)amino]-3-oxopropyl)amino)-3,3-dimethyl-1-oxobutan-2-yl acetate **311** (173 mg, 0.5 mmol) was dissolved in methanol (5 mL) at 0 °C and the mixture stirred for 4 h in the presence of potassium carbonate (210 mg, 3 eq.). The solvent was evaporated and the residue used for the ECC assay described below without further purification.

IR (neat) ν/cm^{-1}: 3298 (OH), 1637 (OC=O); ^{1}H NMR (400 MHz, D_2O): δ = 3.81 (*s*, 1H), 3.51 (*t*, *J* = 5.5, 2H), 3.44-3.36 (*m*, 3H), 3.30 (*d*, *J* = 11.0, 1H), 3.21 (*t*, *J* = 5.5, 2H), 2.37 (*t*, *J* = 6.5, 2H), 0.83 (*s*, 6H); ^{13}C NMR (100 MHz, D_2O): δ = 175.1, 173.1, 76.3, 69.4, 60.6, 42.0, 39.4, 35.5 (2x), 21.5, 21.0; HR ESI-MS: 285.1421 ([M+Na]$^+$, $C_{11}H_{22}N_2NaO_5^+$, calc. 285.1421).

5.5.5 Enzymatic Conversion of Pantetheine Analogues into CoA Analogues

General Procedure

The enzymatic synthesis of CoA analogues from the corresponding pantetheine analogues was performed by adapting the 'one-pot' enzyme cascade catalysis (ECC) procedure from Tosin *et al.* [35] and initially described by Wright and coworkers [28]. Pantethenate substrates (10-50 μmol) were dissolved in reaction buffer (50 mM HEPES pH 9.0, 25 mM KCl, 10 mM $MgCl_2$) together with ATP disodium salt hydrate (8-12 eq.) and the pH was adjusted to 8.0 using 10% aqueous sodium hydroxide solution. Three enzymes (CoaA, CoaD, CoaE, 900 μg each) were added at 30 min intervals and the reaction was left at room temperature for 16 h. One assay volume of cold acetonitrile was added and, after vortexing for 1 min, the mixture was left at 0 °C for 15 min before centrifugation (45 min, 13 krpm) to remove the precipitated proteins. The supernatant was freeze-dried, redissolved in 0.1% aqueous formic acid (2.0–8.0 mL) and subjected to preparative HPLC.

Malonyl-carba(dethia)CoA methyl ester 288 [35]

Methyl 6-((*N*-[(2*R*)-2,4-dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)-3-oxohexanoate **293** (10 mg, 27.7 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 5 min followed by 2–20% B in A over 25 min) gave malonyl-carba(dethia)-CoA methyl ester **288** as a white foam (R_t = 17.1 min, 10.6 mg, 45%).

^{1}H NMR (500 MHz, D_2O): δ = 8.66 (*s*, 1H), 8.42 (*s*, 1H), 6.21 (*d*, *J* = 5.5, 1H), 4.86 (*app bs*, 2H), 4.58 (*app bs*, 1H), 4.24 (*app bs*, 2H), 4.00 (*s*, 1H), 3.84 (*dd*, *J* = 10.0, 5.0, 1H), 3.70 (*s*, 3H), 3.58 (*dd*, *J* = 10.0, 5.0, 1H), 3.44 (*t*, *J* = 6.5, 2H), 3.11 (*t*, *J* = 7.0, 2H), 2.62 (*t*, *J* = 7.0, 2H), 2.42 (*t*, *J* = 6.5, 2H), 1.70 (*quin*, *J* = 7.0, 2H), 0.91 (*s*, 3H), 0.79 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 207.6, 174.8, 173.9, 170.1, 150.0, 148.6, 144.7, 142.6, 118.6, 87.5, 83.6, 74.2, 71.9, 65.1, 52.8, 40.0, 38.4,

35.5, 22.3, 20.8, 18.2; ^{31}P NMR (202 MHz, D_2O): δ = 0.9 (*s*, 1P), -9.7 (*app bs*, 1P), -10.3 (*app bs*, 1P); HR ESI-MS: 870.1457 ([M+Na-2H]$^-$, $C_{26}H_{40}N_7NaO_{19}P_3^-$, calc. 870.1457).

Malonyl-carba(dethia)CoA 294 [35]

Malonyl carba(dethia)-coenzyme A methyl ester **288** (15.0 mg, 17.7 μmol) was dissolved in reaction buffer (500 μL, 50 mM HEPES pH 8.0, 25 mM KCl, 10 mM $MgCl_2$) and a suspension of porcine liver esterase (PLE) in 3.2 M ammonium sulphate was added (25 μL, 130 U). The reaction was left at 25 °C for 18 h and the protein was then removed by centrifugation through a 3kDa cutoff membrane. The flow through was directly subjected to preparative HPLC (2% B in A over 10 min followed by 2–20% B in A over 10 min and 20–98% B in A over 5 min) to give malonyl carba(dethia)-CoA **294** (R_t = 21.0 min, 4.2 mg, 28%) and the unreacted methyl ester (R_t = 22.2 min).

^{1}H NMR (500 MHz, D_2O): δ = 8.63 (*s*, 1H), 8.39 (*s*, 1H), 6.18 (*d*, *J* = 6.0, 1H), 4.87-4.83 (*m*, 2H), 4.56-4.54 (*m*, 1H), 4.22-4.18 (*app bs*, 2H), 3.97 (*s*, 1H), 3.80 (*dd*, *J* = 10.0, 4.5, 1H), 3.70 (*s*, 3H), 3.53 (*dd*, *J* = 9.5, 4.5, 1H), 3.41 (*t*, *J* = 7.0, 2H), 3.07 (*t*, *J* = 7.0, 2H), 2.60 (*t*, *J* = 7.0, 1H), 2.51 (*t*, *J* = 7.0, 1H), 2.39 (*t*, *J* = 6.5, 2H), 1.65 (*quin*, *J* = 7.0, 2H), 0.88 (*s*, 3H), 0.75 (*s*, 3H).

Propanoyl-carba(dethia)CoA 329

(2*R*)-2,4-Dihydroxy-3,3-dimethyl-*N*-(3-oxo-3-[(4-oxohexyl)amino]propyl)butanamide **335** (8.0 mg, 25 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min followed by 2–20% B in A over 15 min) gave propanoyl-carba(dethia)-CoA **329** as a white foam (R_t = 21.5 min, 10 mg, 50%).

^{1}H NMR (500 MHz, D_2O): δ = 8.59 (*s*, 1H), 8.36 (*s*, 1H), 6.14 (*d*, *J* = 6.0, 1H), 4.84-4.78 (*m*, 2H), 4.56 (*app bs*, 1H), 4.23-4.17 (*m*, 2H), 3.97 (*s*, 1H), 3.80 (*d*, *J* = 8.8, 1H), 3.56 (*d*, *J* = 9.5, 1H), 3.39-3.36 (*m*, 2H), 3.03 (*t*, *J* = 6.8, 2H), 2.46-2.41 (*m*, 2H), 2.36 (*t*, *J* = 6.5, 1H), 1.65-1.58 (*m*, 2H), 0.89-0.86 (*m*, 6H), 0.74 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 5.3 (s, 1P), - 9.5 (*d*, *J* = 20.0, 1P), -10.1 (*d*, *J* = 20.0, 1P); HR ESI-MS: 848.1420 ([M-3H+2Na]$^-$, $C_{25}H_{39}N_7Na_2O_{17}P_3^-$, calc. 848.1416).

Isobutyryl-carba(dethia)CoA 330

(2*R*)-2,4-Dihydroxy-*N*-(3-[(4-hydroxy-5-methylhex-4-en-1-yl)amino]-3-oxopropyl)-3,3-dimethylbutanamide **336** (12.5 mg, 40 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 20 min and 20–98% B in A over 2 min) gave isobutyryl-carba(dethia)-CoA **330** as a white foam (R_t = 24.0 min, 13.4 mg, 42%). The ^{1}H NMR data suggest that the product is present as a keto-enol equilibrium mixture.

^{1}H NMR (500 MHz, D_2O): δ = 8.63 (*s*, 1H), 8.41 (*s*, 1H), 6.19 (*d*, *J* = 5.5, 1H), 4.88-4.87 (*m*, 1H), 4.58 (*app bs*, 1H), 4.25 (*m*, 2H), 3.99 (*s*, 1H), 3.85 (*dd*, *J* = 9.5, 4.5, 1H), 3.60 (*dd*, *J* = 9.5, 4.5, 1H), 3.44 (*m*, 2H), 3.09 (*t*, *J* = 7.0, 2H), 2.69 (*spt*, *J* = 7.0, 0.4H), 2.64 (*t*, *J* = 7.0, 2H), 2.58 (*t*, *J* = 7.0, 2H), 2.42 (*t*, *J* = 6.5, 2H), 1.67 (*quint*, *J* = 7.0, 2H), 1.32 (*s*, 2.6H), 1.01 (*d*, *J* = 7.0, 3.4H), 0.91 (*s*, 3H), 0.79 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.7 (s, 1P), - 9.8 (bs, 1P), - 10.3 (bs, 1P); HR ESI-MS: 862.1578 ([M-3H+2Na]$^-$, $C_{26}H_{43}N_7Na_2O_{17}P_3^-$, calc. 862.1573).

Dephospho-amino(dethia)CoA 308

(2*R*)-*N*-(3-[(2-Aminoethyl)amino]-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide **307** (26.6 mg, 100 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 10 min and

20–98% B in A over 5 min) gave the intermediate of catalysis dephospho-CoANH_2 **308** (R_t = 15.6 min, 44.7 mg, 65%).

^{1}H NMR (500 MHz, D_2O): δ = 8.60 (*s*, 1H), 8.37 (*s*, 1H), 6.14 (*d*, *J* = 5.5, 1H), 4.51 (*dd*, *J* = 5.0, 4.0, 1H), 4.38-4.35 (*m*, 1H), 4.25-4.16 (*m*, 2H), 3.99 (*s*, 1H), 3.80 (*dd*, *J* = 10.0, 5.0, 1H), 3.60 (*dd*, J = 10.0, 5.0, 1H), 3.47-3.42 (*m*, 4H), 3.09 (*t*, *J* = 6.0, 2H), 2.46 (*t*, *J* = 6.5, 2H), 0.90 (*s*, 3H), 0.83 (*s*, 3H); ^{13}C NMR (125 MHz, D_2O): δ = 175.6, 175.2, 150.5, 148.7, 145.4, 142.7, 118.9, 88.3, 84.5 (*d*, *J* = 9.0), 75.0, 74.9, 72.1 (*d*, *J* = 6.0), 70.6, 65.4 (*d*, *J* = 5.0), 39.6, 38.8 (*d*, *J* = 8.0), 37.2, 35.9, 35.8, 21.1, 19.2; ^{31}P NMR (202 MHz, D_2O): δ = - 9.6 (*app bs*, 1P), - 10.1 (*d*, *J* = 20.0, 1P); HR ESI-MS: 693.1759 ([M+Na]$^+$, $C_{21}H_{36}N_8NaO_{13}P_2^+$, calc. 693.1769).

Malonyl aza(dethia)CoA 258

3-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-3-oxopropanoic acid **260** (12.6 mg, 38.0 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 10 min and 20–98% B in A over 5 min) gave malonyl-aza(dethia)CoA **258** as a white foam (R_t = 19.6 min, 13.4 mg, 42%).

^{1}H NMR (500 MHz, D_2O): δ = 8.62 (*s*, 1H), 8.41 (*s*, 1H), 6.15 (*d*, *J* = 5.5, 1H), 4.89-4.84 (*app bs*, 1H), 4.73 (*app bs*, 1H), 4.58 (*app bs*, 1H), 4.38 (*m*, 1H), 4.28-4.20 (*m*, 2H), 4.00 (*s*, 1H), 3.85 (*dd*, *J* = 10.0, 5.0, 1H), 3.59 (*dd*, *J* = 10.0, 5.0, 1H), 3.44 (*t*, *J* = 6.5, 2H), 3.33-3.26 (*m*, 4H), 2.43 (*t*, *J* = 6.5, 2H), 0.92 (*s*, 3H), 0.81 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.9 (*s*, 1P), - 9.7 (*d*, *J* = 21.0, 1P), - 10.2 (*d*, *J* = 21.0, 1P); HR ESI-MS: 859.1448 ([M+Na]$^+$, $C_{24}H_{39}N_8NaO_{19}P_3^+$, calc. 859.1437).

Methylmalonyl-aza(dethia)CoA 259

3-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-methyl-3-oxopropanoic acid **261** (12.6 mg, 34.9 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 10 min and 20–98% B in A over 5 min) gave methylmalonyl-aza(dethia)CoA **259** as a white foam (R_t = 19.4 min, 8.6 mg, 29%) and some intermediate of catalysis **342** (see below).

^{1}H NMR (500 MHz, D_2O): δ = 8.61 (*s*, 1H), 8.40 (*s*, 1H), 6.18 (*d*, *J* = 5.5, 1H), 4.87 (*app bs*, 1H), 4.57 (*app bs*, 1H), 4.25 (*app bs*, 1H), 4.28-4.20 (*m*, 2H), 3.98 (*s*, 1H), 3.84 (*dd*, J = 9.5, 4.5, 1H), 3.61-3.59 (*m*, 1H), 3.43 (*q*, *J* = 7.5, 1H), 3.32-3.24 (*m*, 4H), 2.43 (*t*, *J* = 6.5, 2H), 1.28 (*d*, *J* = 7.5, 3H), 0.91 (*s*, 3H), 0.79 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.7 (*s*, 1P), - 9.9 (*app bs*, 1P), - 10.3 (*app bs*, 1P); HR ESI-MS: ([M+Na]$^+$, $C_{25}H_{41}N_8NaO_19P_3^+$, calc. 873.1599).

Methylmalonyl-dephospho-aza(dethia)CoA 342

Methylmalonyl-aza(dethia)(dephospho)CoA **342** was separated from compound **259** (see above) by preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 20 min) and obtained as a white foam (R_t = 21.7 min, 5.8 mg, 21%).

^{1}H NMR (500 MHz, D_2O): δ = 8.65 (*s*, 1H), 8.42 (*s*, 1H), 6.17 (*d*, *J* = 5.5, 1H), 4.88-4.85 (*m*, 1H), 4.57 (*t*, *J* = 4.5, 1H), 4.38 (*app bs*, 1H), 4.22 (*app bs*, 1H), 4.00 (*s*, 1H), 3.84 (*dd*, J = 9.5, 4.5, 1H), 3.58 (*dd*, J = 9.5, 4.5, 1H), 3.45-3.43 (*m*, 1H), 3.31-3.26 (*m*, 4H), 2.43 (*t*, *J* = 6.5, 2H), 1.29 (*d*, *J* = 7.5, 3H), 0.92 (*s*, 3H), 0.80 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = - 9.7 (*d*, *J* = 21.0, 1P), - 10.2 (*d*, *J* = 21.0, 1P); HR ESI-MS: 779.1770 ([M+Na]$^+$, $C_{24}H_{38}N_8NaO_{16}P_2^+$, calc. 779.1773).

Hydroxy(dethia)CoA 252

(2*R*)-2,4-Dihydroxy-
N-(3-[(2-hydroxyethyl)amino]-3-oxopropyl)-3,3-dimethylbutanamide **309** (5.2 mg, 19.8 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 5 min, followed by 2–20% B in A over 15 min and 20–98% B in A over 5 min) gave hydroxy(dethia)CoA **252** as a white foam (R_t = 14.9 min, 6.2 mg, 41%) and some of the intermediate of catalysis dephospho-hydroxy(dethia)CoA **312** (R_t = 16.2 min, 2.0 mg, 15%).

^{1}H NMR (500 MHz, D_2O): δ = 8.60 (*s*, 1H), 8.36 (*s*, 1H), 6.16 (*d*, *J* = 5.5, 1H), 4.83-4.78 (*m*, 1H), 4.76 (*app bs*, 1H), 4.54 (*app bs*, 1H), 4.22-4.16 (*m*, 2H), 3.95 (*s*, 1H), 3.78 (*dd*, *J* = 9.5, 4.5, 1H), 3.56-3.52 (*m*, 3H), 3.41 (*t*, *J* = 6.5, 2H), 3.22 (*t*, *J* = 5.5, 2H), 2.40 (*t*, *J* = 6.5, 2H), 0.86 (*s*, 3H), 0.74 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.9 (*s*, 1P), - 9.7 (*d*, *J* = 20.5, 1P), - 10.4 (*d*, *J* = 20.5, 1P); HR ESI-MS: 752.1451 ([M+Na]$^+$, $C_{21}H_{37}N_7NaO_{17}P_3^+$, calc. 752.1453).

Malonyl-oxa(dethia)CoA 262 [35]

3-[2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-3-
oxopropanoic acid **264** (9.5 mg, 27.3 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 5 min, followed by 2–20% B in A over 15 min and 20–98% B in A over 5 min) gave malonyl-oxo(dethia)CoA **263** as a white foam (R_t = 19.6 min, 9.7 mg, 43%).

^{1}H NMR (500 MHz, D_2O): δ = 8.67 (*s*, 1H), 8.43 (*s*, 1H), 6.21 (*d*, *J* = 5.5, 1H), 4.88-4.85 (*m*, 1H), 4.74 (*s*, 1H), 4.25 (*d*, *J* = 7.5, 1H), 4.21 (*t*, *J* = 5.5, 2H), 4.01 (*s*, 1H), 3.84 (*dd*, *J* = 9.5, 4.5, 1H), 3.58 (*dd*, *J* = 9.5, 4.5, 1H), 3.44-3.41 (*m*, 4H), 2.45 (*t*, *J* = 6.5, 2H), 0.92 (*s*, 3H), 0.79 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.7 (*s*, 1P), - 9.8 (*app bs*, 1P), - 10.3 (*app bs*, 1P); HR ESI-MS: 858.1129([M-2H+Na]$^-$, $C_{24}H_{36}N_7NaO_{20}P_3^-$, calc. 858.1131).

Methylmalonyl-oxa(dethia)CoA 263

3-[2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethoxy]-2-methyl-3-
oxopropanoic acid **265** (11.5 mg, 31.7 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 10 min and 20–98% B in A over 2 min) gave methylmalonyl-oxo(dethia)CoA **263** as a white foam (R_t = 20.2 min, 14.2 mg, 52%).

^{1}H NMR (400 MHz, D_2O): δ = 8.61 (*s*, 1H), 8.40 (*s*, 1H), 6.18 (*d*, *J* = 5.5, 1H), 4.88-4.84 (*m*, 1H), 4.57 (*app bs*, 1H), 4.25 (*app bs*, 1H), 4.22-4.18 (*m*, 2H), 3.98 (*s*, 1H), 3.85-3.84 (*m*, 1H), 3.61-3.56 (*m*, 2H), 3.47-3.39 (*m*, 4H), 2.43 (*t*, *J* = 6.5, 2H), 1.32 (*d*, *J* = 7.5, 3H), 0.91 (*s*, 3H), 0.79 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.7 (*s*, 1P), - 9.9 (*app bs*, 1P), - 10.3 (*app bs*, 1P); HR ESI-MS: 872.1293 ([M-2H+Na]$^-$, $C_{25}H_{38}N_7NaO_{20}P_3^-$, calc. 872.1288).

Protected (2*S*)-carboxymethylprolyl-dephospho-aza(dethia)CoA 316

2-((5*S*)-1,5-Bis(benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid **304** (4 mg, 10 μmol) was dissolved in dried dimethylformamide (200 μL) and *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluoro-phosphate (HBTU, 4 mg, 8 eq.) was added followed by a solution of hydroxybenzotriazole (HOBt, 4 mg, 8 eq.) in dried dimethylformamide (200 μL). The mixture was cooled to 0 ° before addition of diisopropylethylamine (17 μL). After stirring under these conditions for 10 min, a solution of dephospho-amino(dethia)CoA **308** (2.5 mg, 3.7 μmol) in a 1:1 mixture of DMF and water (200 μL) was added and the reaction stirred at room temperature for 4 h. The reaction mixture was diluted with 0.1% aqueous trifluoroacetic acid and subjected directly to preparative HPLC (2% B in A over 5 min, followed by 2–98% B in A over 30 min). The fractions containing the product were lyophilised and protected (2*S*)-carboxymethylprolyl-dephospho-aza(dethia)CoA **316**

obtained as a white foam (R_t = 14.2 min, 1.7 mg, 42%).

^{1}H NMR (500 MHz, D_2O): δ = 8.62 (*s*, 1H), 8.35 (*s*, 1H), 7.37-7.30 (*m*, 8H), 7.22-7.20 (*m*, 1H), 7.17-7.15 (*m*, 1H), 6.13 (*d*, *J* = 5.5, 1H), 5.19-4.89 (*m*, 4H), 4.74 (*m*, 1H), 4.52 (*m*, 1H), 4.44-4.30 (*m*, 2H), 4.22 (*app bs*, 1H), 4.00 (*s*, 1H), 3.84 (*dd*, *J* = 9.5, 4.5, 1H), 3.73-3.65 (*m*, 1H), 3.57 (*dd*, *J* = 9.5, 3.0, 1H), 3.43-3.39 (*m*, 2H), 3.25-3.06 (*m*, 4H), 2.58-2.52 (*m*, 1H), 2.42-2.28 (*m*, 1H), 2.09-1.99 (*m*, 2H), 1.74 (*dd*, *J* = 12.0, 6.5, 1H), 1.32-1.25 (*m*, 4H), 0.92 (*s*, 3H), 0.79 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = - 9.59 (*d*, *J* = 20.5, 1P), - 10.2 (*d*, *J* = 20.5, 1P); HR ESI-MS: 1050.3376 ($[M+H]^+$, $C_{43}H_{58}N_9O_{18}P_2^+$, calc. 1050.3370).

(2*S*)-Carboxymethylprolyl-dephospho-aza(dethia)CoA 318

Protected (2*S*)-carboxymethylprolyl-dephospho-aza(dethia)CoA **316** (1.65 mg, 1.6 μmol) was dissolved in a 1:1 water/THF mixture (1 mL) and palladium on activated carbon was added (1 mg, 70% w/w). The reaction vessel was flushed several times with nitrogen and vacuum cycles before replacing the atmosphere with hydrogen gas. The reaction was left at room temperature for 5 h, the catalyst filtered off on a pad of Celite and solvent evaporated to give (2*S*)-carboxymethylprolyl-dephospho-aza(dethia)CoA **318** as a colorless oil (1.17 mg, 78%).

^{1}H NMR (500 MHz, D_2O): δ = 8.47 (*s*, 1H), 8.23 (*s*, 1H), 6.1 (*d*, *J* = 6.0, 1H), 4.74 (*m*, 1H), 4.51-4.49 (*m*, 1H), 4.36-4.35 (*m*, 1H), 4.19 (*dd*, *J* = 5.0, 3.0, 2H), 3.97 (*s*, 1H), 3.80 (*dd*, *J* = 10.0, 5.0, 2H), 3.65 (*t*, *J* = 8.0, 1H), 3.61-3.53 (*m*, 2H), 3.41 (*d*, *J* = 6.5, 2H), 3.27-3.23 (*m*, 4H), 2.50-2.36 (*m*, 3H), 2.25-2.11 (*m*, 1H), 1.96-1.90 (*m*, 2H), 1.79-1.71 (*m*, 1H), 1.57-1.54 (*m*, 1H), 1.44-1.40 (*m*, 1H), 0.86 (*s*, 3H), 0.75 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = - 9.63 (*d*, *J* = 20.5, 1P), - 9.63 (*d*, *J* = 20.5, 1P), - 10.16 (*d*, *J* = 20.5); HR ESI-MS: 824.2399 ($[M-H]^-$, $C_{28}H_{44}N_9O_{16}P_2^-$, calc. 824.2387).

(2*S*)-Carboxymethylprolyl-aza(dethia)CoA 313

5-(2-([2-((*N*-[(2*R*)-2,4-Dihydroxy-3,3-dimethylbutanoyl]-β-alanyl)amino)ethyl]amino)-2-oxoethyl)-L-proline **319** (10.9 mg, 26 μmol) was incubated under ECC conditions in the presence of CoaA, CoaD and CoaE as described above and preparative HPLC (2% B in A over 10 min, followed by 2–20% B in A over 10 min and 20–98% B in A over 2 min) yielded (2*S*)-carboxymethylprolyl aza(dethia)CoA **313** as a white foam (R_t = 20.2 min, 14.2 mg, 52%).

^{1}H NMR (500 MHz, D_2O): δ = 8.65 (*s*, 1H), 8.42 (*s*, 1H), 6.20 (*d*, *J* = 5.5, 1H), 4.87-4.83 (*m*, 1H), 4.58 (*app bs*, 1H), 4.40-4.35 (*m*, 1H), 4.27-4.20 (*m*, 2H), 4.03-3.93 (*m*, 2H), 3.82 (*dd*, *J* = 9.5, 4.5, 1H), 3.60 (*dd*, *J* = 9.5, 4.5, 1H), 3.45 (*t*, *J* = 6.5, 2H), 2.79-2.73 (*m*, 2H), 2.51-2.43 (*m*, 3H), 2.28-2.20 (*m*, 1H), 2.13-2.06 (*m*, 1H), 0.91 (*s*, 3H), 0.83 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = 0.92 (*s*, 1P), - 9.5 (*app bs*, 1P), - 10.1 (*app bs*, 1P); HR ESI-MS: 926.1859 ($[M-2H+Na]^-$, $C_{28}H_{44}N_9NaO_{19}P_3^-$, calc. 926.1859).

Protected (2*S*)-carboxymethylprolyl-dephospho-oxa(dethia)CoA 315

2-((5*S*)-1,5-Bis(benzyloxycarbonyl)pyrrolidin-2-yl)acetic acid **304** (4 mg, 10 μmol) was dissolved in dried dimethylformamide (200 μL) and *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluoro-phosphate (HBTU, 4 mg, 8 eq.) was added followed by a solution of hydroxybenzotriazole (HOBt, 4 mg, 8 eq.) in dried dimethylformamide (200 μL). The mixture was cooled to 0 °C before addition of diisopropylethylamine (17 μL). After stirring under these conditions for 10 min, a solution of hydroxy(dethia)CoA **252** (2.5 mg, 3.7 μmol) in a 1:1 mixture of DMF and water (200 μL) was added and the reaction stirred at room temperature for 4 h. The reaction mixture was diluted with 0.1% aqueous trifluoroacetic acid and subjected directly to preparative HPLC (2% B in A over 5 min, followed by 2–98% B in A over 30 min). The fractions containing the product

were lyophilised and protected (2*S*)-carboxymethylprolyl-dephospho-oxa(dethia)CoA **315** obtained as a white foam (R_t = 15.2 min, 1.6 mg, 41%).

^{1}H NMR (500 MHz, D_2O): δ = 8.64-8.04 (*m*, 2H), 7.37-7.02 (*m*, 10H), 6.26-5.96 (*m*, 1H), 5.16-4.96 (*m*, 4H), 4.55-4.15 (*m*, 4H), 4.01 (*s*, 1H), 3.84-3.81 (*m*, 1H), 3.60-3.53 (*m*, 2H), 3.43 (*t*, *J* = 6.5, 2H), 3.25 (*t*, *J* = 5.5, 2H), 3.19-3.15 (*m*, 1H), 2.96-2.58 (*m*, 3H), 2.44 (*t*, *J* = 6.5, 2H), 2.24-2.13 (*m*, 1H), 2.11-2.03 (*m*, 1H), 1.92-1.85 (*m*, 1H), 1.77-1.73 (*m*, 1H), 0.91 (*s*, 3H), 0.78 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = - 9.6 (*app bs*, 1P), - 10.3 (*app bs*, 1P); HR ESI-MS: 1051.3203 ($[M+H]^+$, $C_{43}H_{57}N_8O_{19}P_2^+$, calc. 1051.3210).

(2*S*)-Carboxymethylprolyl-dephospho-oxa(dethia)CoA 317

Protected (2*S*)-carboxymethylprolyl-dephospho-oxa(dethia)CoA **315** (1.60 mg, 1.5 μmol) was dissolved in a 1:1 water/THF mixture (1 mL) and palladium on activated carbon was added (1 mg, 70% w/w). The reaction vessel was flushed several times with nitrogen and vacuum cycles before replacing the atmosphere with hydrogen gas. The reaction was left at room temperature for 5 h, the catalyst filtered off on a pad of Celite® 545 and solvent evaporated to give (2*S*)-carboxymethylprolyl-dephospho-oxo(dethia)CoA **317** as a colorless oil (0.95 mg, 67%).

^{1}H NMR (400 MHz, D_2O): δ = 8.47 (*s*, 1H), 8.23 (*s*, 1H), 6.10 (*d*, *J* = 5.5, 1H), 4.75-4.73 (*m*, 1H), 4.5 (*d*, *J* = 5.0, 3.8, 1H), 4.37-4.34 (*m*, 1H), 4.19 (*dd*, *J* = 5.0, 3.0, 2H), 4.17-4.09 (*m*, 1H), 3.97 (*s*, 1H), 3.93-3.87 (*m*, 1H), 3.80 (*dd*, *J* = 10.0, 5.0, 1H), 3.59 (*t*, *J* = 5.5, 2H), 3.54 (*dd*, *J* = 10.0, 4.5, 1H), 3.42 (*t*, *J* = 6.5, 2H), 3.25 (*t*, *J* = 5.5, 2H), 2.65-2.56 (*m*, 2H), 2.46-2.39 (*m*, 2H), 2.22-2.13 (*m*, 1H), 2.03-1.97 (m, 1H), 1.77-1.69 (*m*, 1H), 1.29-1.23 (*m*, 1H), 0.86 (s, 3H), 0.74 (*s*, 3H); ^{31}P NMR (202 MHz, D_2O): δ = - 9.6 (*d*, *J* = 21.0, 1P), - 10.2 (*d*, *J* = 21.0, 1P); HR ESI-MS: 825.2239 ($[M-H]^-$, $C_{28}H_{43}N_8O_{17}P_2^-$, calc. 825.2227).

5.5.6 CMPS Assays using Malonyl-CoA Analogues

The malonyl and methylmalonyl-CoA analogues, where the sulphur atom was replaced by a nitrogen or an oxygen atom, were tested as CMPS substrates using the following procedure.

CMPS Assay

The deprotection of amino acid aldehydes substrates was performed as described in Chapter 4 (Section 4.5.3), and the CMPS assay was carried at pH 7.9 at a final volume of 70 μL as in Section 4.5.3 but in the absence of malonyl-CoA **130**. A 15 mM solution of amino L-GHP or 4-methyl-L-GHP in 10% aqueous formic acid (5 μL, 75 nmol) was mixed in 900 mM Tris pH 9.0 (35 μL) together with a 20 mM solution of malonyl-oxa(dethia)CoA **262**, methylmalonyl-oxa(dethia)CoA **263**, malonyl aza(dethia)CoA **258** or methylmalonyl aza(dethia)CoA **259** (8 μL, 160 nmol) and the mixture diluted with water (20 μL) before CMPS was added (0.3 nmol, 3 μL of 1.5 mM solution).

The assay mixture was kept at 37 °C for 1–5 h and quenched with a 250 μM *p*-aminosalicylic acid solution in acetonitrile (70 μL). The samples were kept on ice for 15 min, and then centrifuged at 13000 rpm for 15 min at 4 °C to remove the precipitated protein. The supernatant was transfered in a vial for analysis.

CMPS Assay LC-MS Analysis

The formation of *t*-CMP **131** or 6-methyl-*t*-CMP **343** was monitored as described in Section 4.5.3. The formation of *t*-CMP-CoA derivatives was monitored by analytical reverse phase LC-MS using a Phenomonex Luna C18 column (150 mm x 4.6 mm, 5 μm particle size). The resulting chromatograms

were analysed using the MassLynx™ v4.0 software.

5.5.7 CMPS Assays in the Presence of *t*-CMP-aza(dethia)CoA

CMPS Assay using *t*-CMP-aza(dethia)CoA as a potential CMPS substrate

t-CMP-aza(dethia)CoA **313** was tested as a CMPS substrate using the following procedure. The deprotection of amino acid aldehydes was performed as described in Chapter 4 (Section 4.5.3), and the CMPS assay was carried out at pH 7.9 at a final volume of 70 μL as described in Section 4.5.3 but in the absence of malonyl-CoA **130**. A 15 mM solution of amino L-GHP or 4-methyl-L-GHP in 10% aqueous formic acid (5 μL, 75 nmol) was mixed in 900 mM Tris pH 9.0 (35 μL) together with a 20 mM solution of *t*-CMP aza(dethia)CoA **313** (8 μL, 160 nmol) and the mixture diluted with water (20 μL) before CMPS was added (0.3 nmol, 3 μL of 1.5 mM solution).

The assay mixture was kept at 37 °C for 4 h and quenched with a 250 μM *p*-aminosalicylic acid solution in acetonitrile (70 μL). The samples were kept on ice for 15 min, and then centrifuged at 13000 rpm for 15 min at 4 °C to remove the precipitated protein. The supernatant was transfered in a vial for analysis.

The formation of *t*-CMP **131** was monitored as described in Chapter 4 (Section 4.5.3). The resulting chromatograms were analysed using the MassLynx™ v4.0 software.

CMPS Assay using *t*-CMP-aza(dethia)CoA as a potential CMPS inhibitor

t-CMP aza(dethia)CoA **313** was tested for CMPS inhibition using the following procedure. The deprotection of amino acid aldehydes was performed as described in Chapter 4 (Section 4.5.3), and the CMPS assay was carried out at pH 7.9 at a final volume of 70 μL. A 15 mM solution of L-GHP or 4-methyl-L-GHP in 10% aqueous formic acid (5 μL, 75 nmol) was mixed in 900 mM Tris pH 9.0 (35 μL) together with a 10 mM solution of malonyl-CoA **130** (8 μL, 80 nmol). A solution of *t*-CMP-aza(dethia)CoA **313** was added to various final concentrations ranging from 0.01 mM to 10 mM and the mixture diluted with water (20 μL) before CMPS was added (0.3 nmol, 3 μL of 1.5 mM solution). The control experiment was carried in the same way but in the absence of *t*-CMP-aza(dethia)CoA **313**.

The assay mixture was kept at 37 °C for 5 min and quenched with a 250 μM *p*-aminosalicylic acid solution in acetonitrile (70 μL). The samples were kept on ice for 15 min, and then centrifuged at 13000 rpm for 15 min at 4 °C to remove the precipitated protein. The supernatant was transfered in a vial for analysis.

The formation of *t*-CMP **131** was monitored as in Chapter 4 (Section 4.5.3). The resulting chromatograms were analysed using the MassLynx™ v4.0 software.

Chapter 6

Crystallographic Studies on ThnE

6.1 Introduction

6.1.1 Crystallographic Studies on CarB

Crystal structures of CarB from of *P. carotovorum* have been determined in an apo form to 2.24 Å resolution, as well as in complex with acetyl-CoA **127** and a bicine molecule to 3.15 Å resolution. The structural features revealed in these structures have been discussed by Sleeman *et al.* [1]. The structural information were exploited for mutagenesis work, as described in Section 2.4.1 [2]; Analysis of the crystal packing in the apo-CarB structure revealed that the uncomplexed CarB protein crystallises as a homotrimer in the orthorhombic spacegroup $P2_12_12_1$. When complexed with acetyl-CoA **127** and bicine, CarB also crystallises as a homotrimer in the hexagonal spacegroup $P6_3$. In the $P6_3$ crystal form, only one of the monomers (chain A) contains bound ligands. This observation was proposed to be due to a crystal packing 'artifact' and the hypothesis is supported by the results described below for ThnE (Section 2.5.2).

The main difference between the monomer containing an acetyl-CoA ligand (chain A) and the other two monomers (chain B and C) is the position of the α-2 and α-10 helices [1]. The α-2 helix folds onto the active site to interact with the adeninyl group of CoA. The phosphate groups of CoA are directed away from the active site and are sandwiched between chain A residues and a symmetry-related trimer chain B residues. This 3′,5′-diphosphoadenosine group is probably exposed to the solvent in solution. The α-10 helix is bent towards the active site in the bicine-bound monomer.

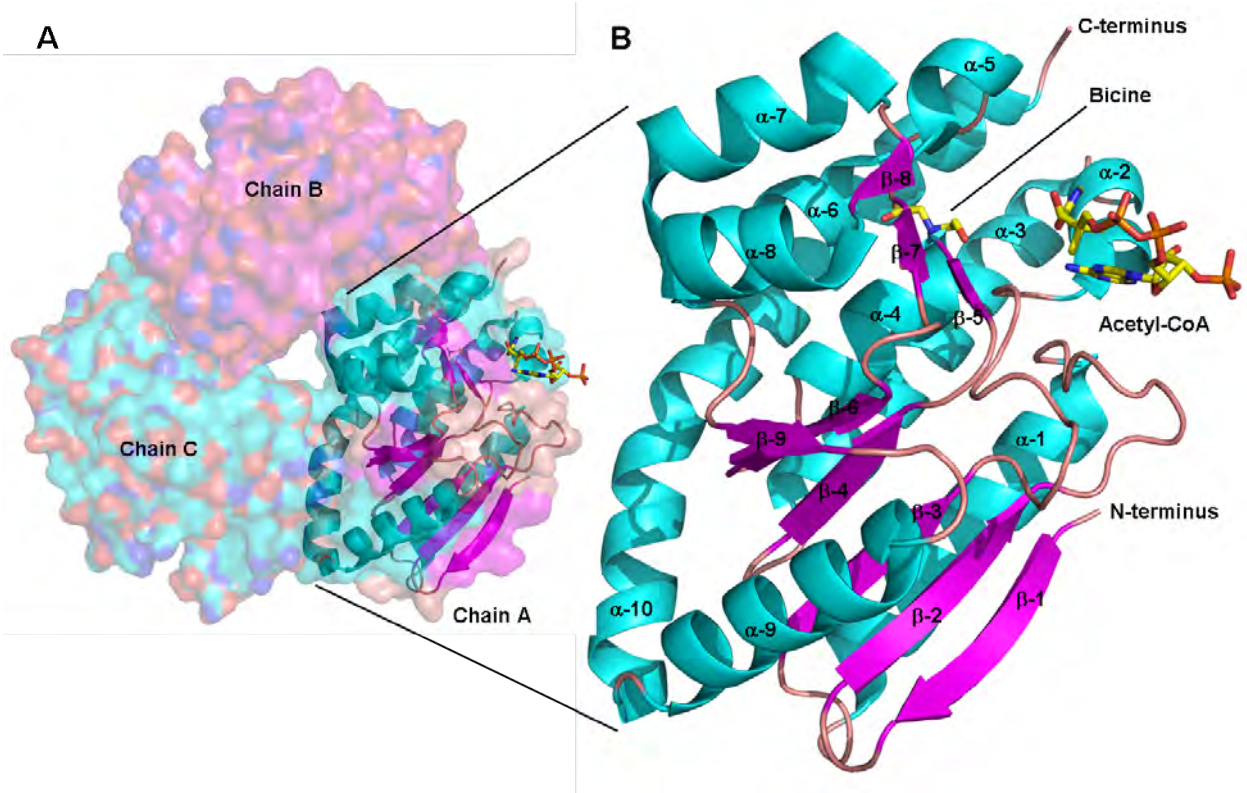


Figure 6.1: Views from CarB crystal structure (PDB ID: 2A81). The presence of only one ligand per trimer is probably a crystal packing 'artifact'. The ribbon representation of chain A shows α-helices in turquoise, β-sheets in purple and loops in pink. The surface is displayed for chain B (blue) and chain C (magenta). (A) CarB crystallographic trimer probably reflects the oligomeric state observed in solution; (B) CarB monomer with secondary structure elements labeled. The bicine and acetyl-CoA ligands bound to the active site are displayed as yellow sticks. This figure was generated using PyMOL [3].

Instead of the linear α-10 helix observed in the apo-CarB structure, the bent helix α-10 observed in the complex apparently enables a hydrogen-bond between His-229 and the bicine carboxylate (Figure 6.2 B). This interaction may mimic L-GHP binding and substrate positioning in the active site.

The crystal structure of CarB in complex with acetyl-CoA and bicine does not contain visible electron density for all atoms of the CoA ligand (Figure 6.2 A). The 3′,5′-diphosphoadenosine group as well as part of the pantetheine moiety are well defined but the acetyl group and the sulphur atom are not visible, potentially indicating the flexibility of this region of the ligand, or partial hydrolysis of the thioester leading to low occupancy for the acetyl function, as suggested by the results described below in the case of ThnE (Section 2.5.2). Comparisons with structures of other CS members suggest that the residues forming the oxyanion hole in CarB are Gly-62 and Met-108 (Figure 6.2). However, the absence of electron density for the ligand in the oxyanion hole does not allow conclusive determination of their involvement in the binding of acyl-CoA. On the other hand, the bicine ligand is well defined and binds in the active site, probably in a similar overall position to L-GHP, through a series of hydrogen bonds between its carboxylate and CarB active site residues, including the side chain imidazole of His-229 (2.77 Å) and the side chain nitrogen of Gln-111 (2.82 Å).

6.1.2 Objectives

The objectives of the crystallographic work presented herein were to obtain a structure of ThnE from *S. cattleya*, a homologous CMPS enzyme, for comparison with the CarB structure described [1]. A ThnE structure may help to explain the differences in stereoselectivity observed between CarB and ThnE-catalysed reactions in the presence of alkylmalonyl-CoA and methyl-substituted L-GHP substrate analogues (Section 4.3). It may also provide more insights into the CMPS mechanism, in particular if a substrate/intermediate complex could be obtained. Indeed, there is a substantial gap in knowledge due to the lack of a CMPS structure in complex with L-GHP.

All of the substrate analogues and non-hydrolysable analogues of the intermediate of catalysis described in Chapter 5 were tested in ThnE co-crystallisation experiments. These analogues include malonyl-oxa(dethia)CoA **262**, malonyl-aza(dethia)CoA **258**, their C-2 methyl derivatives **259** and **263**, malonyl-carba(dethia)CoA methyl ester **288**, the two truncated *t*-CMP-CoA analogues 5-2-

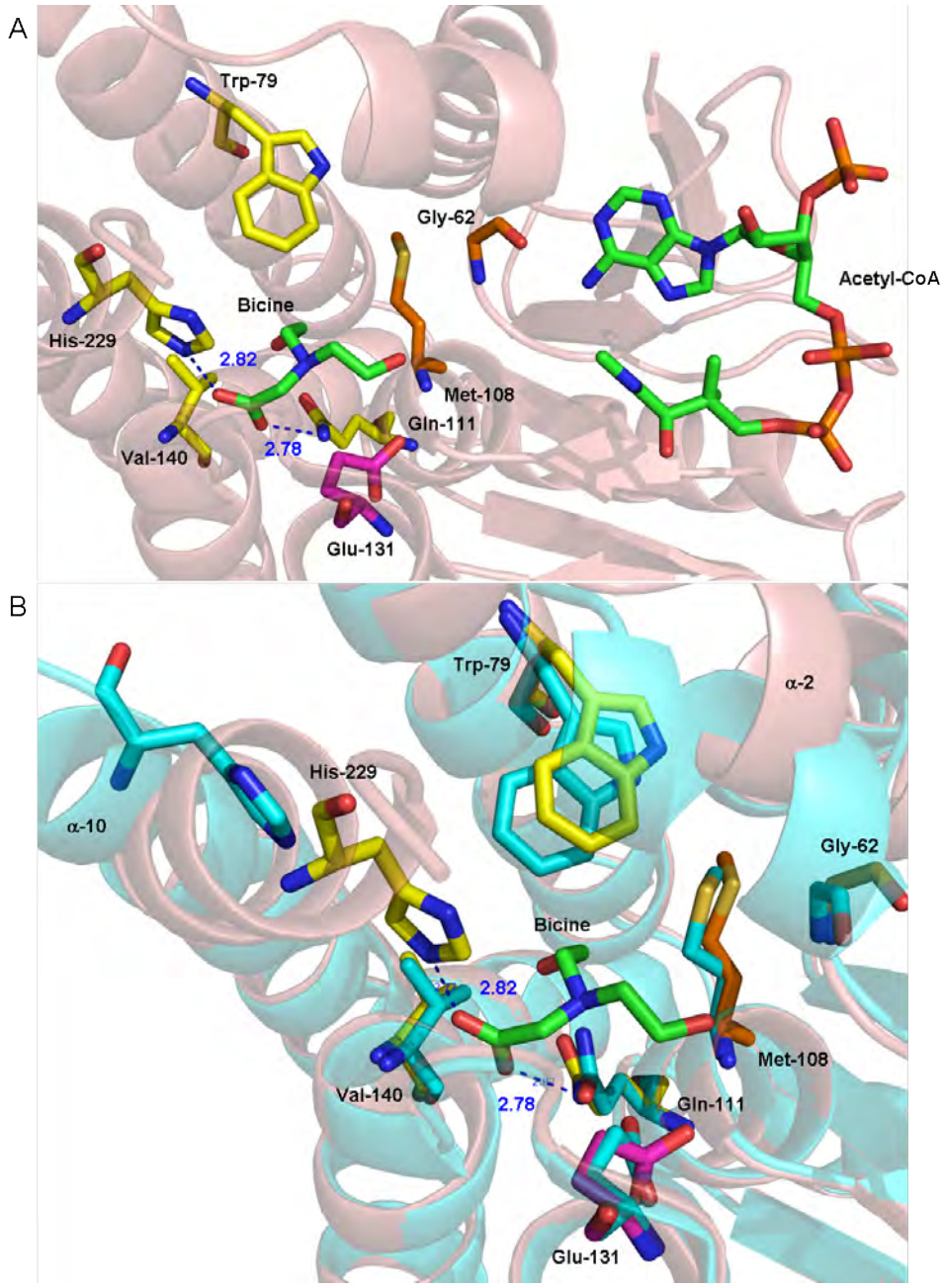


Figure 6.2: View from the active site of CarB (PDB ID: 2A81). Residues thought to be important for catalysis are highlighted in sticks representation. The secondary structure elements are represented as ribbons. Some residues were removed for clarity. (A) The active site of CarB in complex with acetyl-CoA **127** and bicine. The two oxyanion hole-forming residues, Gly-62 and Met-108 are represented as dark orange sticks. Glu-131 proposed to participate in thioester hydrolysis is represented as purple sticks. Some residues substituted for biocatalysis purposes are represented as yellow sticks; (B) Comparison between the secondary structure elements in apo-CarB (in blue) and CarB in complex with acetyl-CoA **127** and bicine (in pink), as well as some active site residues. The α-10 helix is bending in the latter structure, enabling a hydrogen-bond between His-229 (represented in yellow sticks) and the bicine ligand (green sticks). This figure was generated using PyMOL [3].

[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324**, as well as *t*-CMP-aza(dethia)CoA **313** (Figure 6.7). The purification of ThnE is described in Chapter 4.

X-ray diffraction data collection was performed with the help of Dr Michael A. McDonough and Dr Rasheduzzaman Chowdhury on an in-house X-ray source at the Chemistry Research Laboratory in Oxford, U.K., or at the Diamond Light Source, Harwell, U.K. (beamlines I03-1 and I04-1). X-ray diffraction data processing was performed by Dr Michael A. McDonough at the Chemistry Research Laboratory in Oxford, U.K.

6.2 Apo-ThnE Protein Crystallography

6.2.1 Crystallisation trials

In order to test for conditions where ThnE would reproducibly crystallise, initial experiments were set up both at 4 °C and 20 °C using five different commercial crystallisation screens, each containing 96 unique conditions. These screens provide starting points to find hit conditions and were performed using a high throughput sitting drop vapor diffusion methodology in Intelliplate three-drop 96-well crystallisation plates. A 25 mg mL^{-1} ThnE solution in 100 mM Tris-HCl (pH 7.5) was diluted to 20 mg mL^{-1} with water. Three different protein dilutions were used for each condition, including 2:1, 1:1 and 1:2 as ratios of protein to well solution.

The JCSG-plus screen is described as "the most complete sparse matrix screen available today" by its supplier Molecular Dimensions, and was developed based on work carried out at the Joint Centre for Structural Genomics. The JCSG screen provides a wide range of conditions that were shown to promote crystallisation for a diverse set of proteins. The Crystal Screen is also a sparse matrix screen provided by Hampton Research. The SaltRX screen is a salt versus pH matrix crystallisation screen provided by Hampton Research. The PEG/ion screen is a crystallisation screen designed to evaluate high purity polyethylene glycol 3350 and 48 unique salts representing a very complete range of anions and cations frequently used in the crystallisation of biological macromolecules. The Index screen samples a series of diverse and specially formulated reagent zones to identify reagent classes and pH values that are effective in producing crystals.

Apo-ThnE crystals were found to grow under various conditions, indicating a relatively large tolerance of the enzyme for salt and precipitant concentration (Table 6.1). ThnE protein was found to precipitate under acidic conditions (pH $\leq$ 7.0) and therefore, crystals only grew under mildly basic conditions (7.5 $\leq$ pH $\leq$ 9.1). The apo-ThnE protein was found to crystallise best at 20 °C and crystals started growing within 4–5 h after setting up the plates. CarB had been shown to crystallise over a similar period of time (personal communication by Dr Michael A. McDonough). Although in the case of CarB, diffraction quality crystals were harvested after 3–5 days [1]. Diffraction quality crystals for apo-ThnE were obtained after 18 h. The best initial hits were obtained using the JCSG screen, well

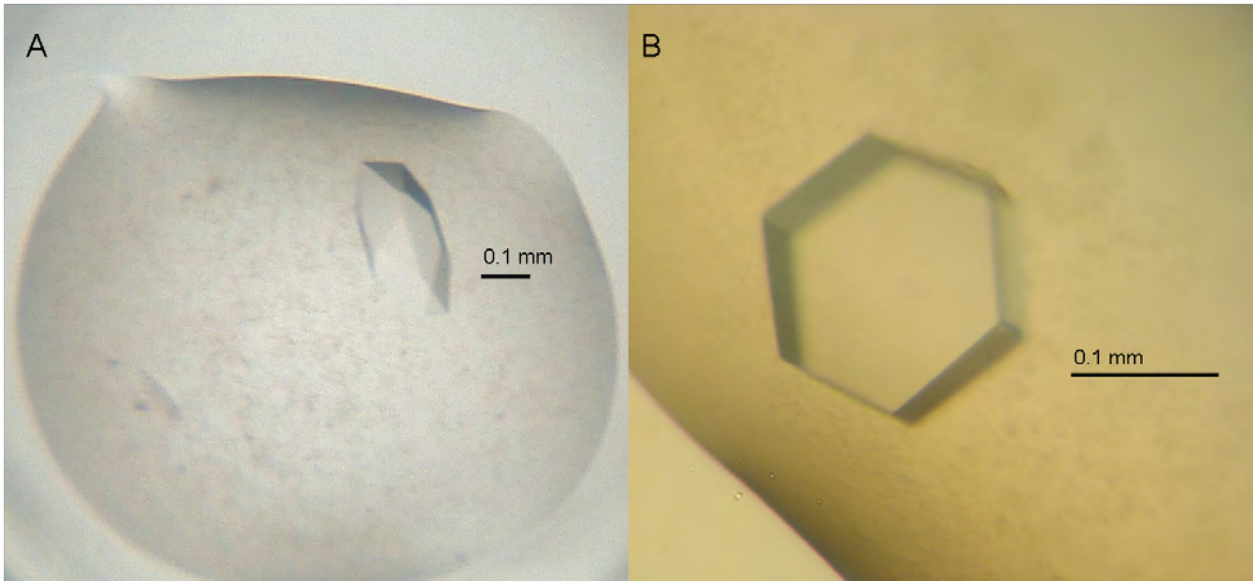


Figure 6.3: Crystal of apo-ThnE and ThnE in complex with acetyl-CoA and L-proline. A 0.1 mm scale bar is given for comparison. (A) Typical rhombohedral morphology for apo-ThnE crystal grown in reservoir solution containing 0.67 M trisodium citrate and 0.12 M tris (pH 8.5). The approximate size of this crystal was 350 x 160 μm; (B) Typical hexagonal plate morphology for ThnE crystals in complex with a CoA derivative. This crystal was grown in reservoir solution containing 0.89 M trisodium citrate (pH 9.0) and 0.08 M tris (pH 8.5). The approximate size of this crystal was 120 x 120 x 40 μm and contained acetyl-CoA **127** and L-proline **134**.

A8, containing 0.2 M ammonium formate and 20% PEG 3350, and with a 1:2 ratio of protein to well solution. The crystals obtained from the screens which were tested for diffraction are detailed in Table 6.1.

6.2.2 Results

All of the apo-ThnE crystals for which X-ray diffraction data were collected showed the same morphology and crystallographic unit cell constants (Figure 6.3 A, space group $P2_12_12_1$; a, b, c = 58.3 Å, 112.1 Å, 113.8 Å; α, β, γ = 90°, 90°, 90°).

The structure of apo-ThnE was determined to 1.85 Å resolution and shows that the crystal packing blocks the active site to an extent that is predicted to prevent the binding of the CoA ligand. The *C*-terminal amino acids (residues 277–293) forming helix α-11 in the ThnE structure in complex with CoA are not visible in the apo-ThnE structure. Instead, a second crystallographic trimer molecule occupies the position of the adenosyl moiety of CoA observed in the complexed structure. This observation suggests that soaking experiments with this crystal form and CoA or stable analogues

might be unsuccessful. The blocked CoA binding site in apo-ThnE is similar to that observed in apo-CarB [1]. The absence of any ligand in the active site of ThnE, even when the crystallisation condition contained a large excess of malonate, acetate, citrate and related neutralised organic acids, may reflect the synergy required between the binding of the CoA moiety and the access for the L-GHP substrate, as suggested by the presence of a putative acetate ion in the active site of a ThnE structure in complex with CoA (Figure 6.5 A).

In the CarB structures, the observed movement of residue His-229 located on helix α-10 upon binding of the CoA ligand may explain the inability of L-GHP **129** analogues, *e.g.* L-proline, to bind in the active site in the absence of CoA. However, it is notable that the CarB variant where His-229 was substituted to an alanine was found to retain the wildtype activity towards L-GHP and analogues, or even showed increased turnover for non-natural substrates such as 2-methyl-L-GHP **181**, probably for steric reasons (Table 4.1). There is a conserved histidine residue in ThnE (His-274) but its position does not change to such a large extent comparing the apo-ThnE structure to ThnE in complex with L-proline **134** and acetyl-CoA **130** (Section 2.5.2, Figure 6.6). Similarly to CarB, the ThnE variant where His-274 was substituted for an alanine retained the activity of the wildtype (Table 4.1)

The ThnE V153A variant was crystallised under the same conditions and the structure of the apo-protein was determined to 1.90 Å resolution (Table 6.2, Entry II). The electron density for the substituted amino acid is visible and confirms that the substitution is present and increases the size of the active site (data now shown). This additional space potentially explains the tolerance of this ThnE variant for bulkier substrates compared to L-GHP (Table 4.1, Entry XV). The structure of the ThnE V153A variant has yet to be refined further, but it may provide useful information for the rationalisation of its substrate tolerance. ThnE V153A was the best variant for the production of the *t*-CMP product involving the introduction of a quaternary carbon center at C-5 (Table 4.1, Entry XV).

Entry	Screen	Buffer, Ions (pH)	PEG (%)	Temperature (°C)	Time (d)
I	JCSG	0.2 M Ammonium formate	PEG 3350 (20)	20	1
II	JCSG	1.6 M Trisodium citrate (6.5)		20	7
III	JCSG	2.1 M Malic acid (7.0)		4	7
IV	PEG-Ion	0.2 M Diammonium hydrogen sulfate		20	3
V	PEG-Ion	0.1 M Sodium malonate (6.0)	PEG 3350 (12)	20	3
VI	PEG-Ion	0.1 M Sodium malonate (7.0)	PEG 3350 (12)	20	3
VII	PEG-Ion	0.1 M Malic acid (7.0)	PEG 3350 (12)	20	1
VIII	PEG-Ion	0.1 M Diammonium tartrate (7.0)	PEG 3350 (12)	20	5
IX	PEG-Ion	0.1M Tris (8.5), 2% v/v Tacsimate† (8.0)	PEG 3350 (16)	20	7
X	INDEX	0.1 M HEPES (7.0), 15% v/v Tacsimate† (7.0)	PEG 3350 (2)	20	1

Composition of the crystallisation conditions that gave apo-ThnE crystals. Three of the commercial screens produced hits. The time necessary for harvestable crystals to grow is indicated in days. † Tacsimate™ (Hampton Research) is a mixture of organic acid salts, including 1.83 M malonic acid, 0.25 M ammonium citrate tribasic, 0.12 M succinic acid, 0.3 M D,L-malic acid, 0.4 M sodium acetate trihydrate, 0.5 M sodium formate and 0.16 M ammonium tartrate dibasic (titrated to the required final pH using sodium hydroxide).

Table 6.1: Crystallography screen conditions that gave apo-ThnE crystals.

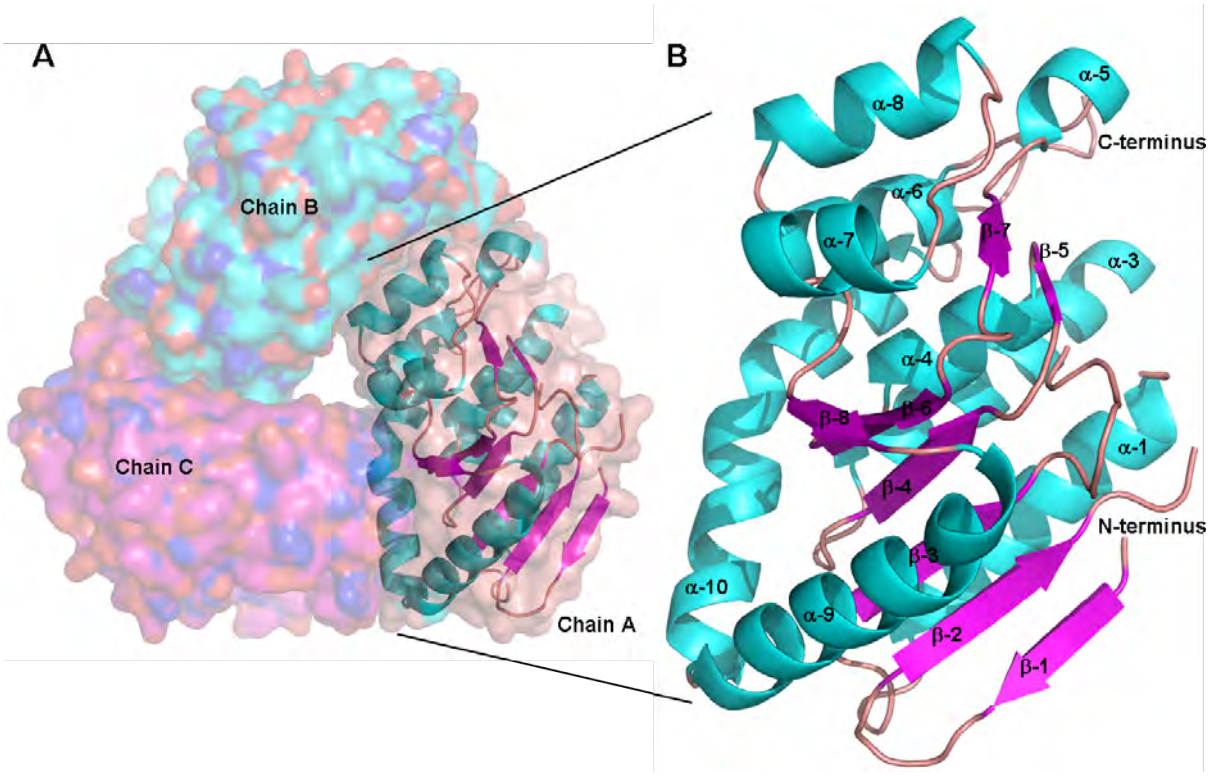


Figure 6.4: Views from a crystal structure of apo-ThnE. No ligand was present in this structure. Helices α-2 and α-11 are likely flexible and therefore no electron density was visible in the apo-protein structure. The ribbon representation of chain A shows α-helices in turquoise, β-sheets in purple and loops in pink. The surface is displayed for chain B (blue) and chain C (magenta). (A) ThnE crystallographic trimer probably reflects the oligomeric state observed in solution [4]; (B) ThnE monomer with secondary structure elements labeled. This figure was generated using PyMOL [3].

6.3 ThnE Crystal Structure in Complex with Acetyl-CoA, CoASH and Proline

6.3.1 Crystallisation trials

An optimisation screen in the presence of a series of commercially available CoA ligands was set up based on the results obtained for apo-ThnE using the methodology described above (Section 2.5.1). A 25 mg mL^{-1} ThnE solution in 100 mM Tris-HCl (pH 7.5) was diluted to 20 mg mL^{-1} with 100 mM ligand stock solution (pH 7.5) prior to crystallisation trials. The final concentration of all ligands in the protein solution was 20 mM. The best conditions for apo-ThnE crystallography were 0.2 M ammonium formate and 20% PEG 3350 (Table 6.1, Entry I) and a two-dimensional optimisation screen varying the ammonium formate and the PEG concentrations was designed. ThnE was mixed with malonyl-CoA **130**, acetyl-CoA **127** and a mixture of acetyl-CoA **127** or CoASH **133** with L-proline **134**. Crystals grew in a similar form compared to the apo-ThnE crystals and the electron density maps obtained from X-ray diffraction data revealed no ligand bound to the enzyme. This observation was attributed to the inability of ThnE to bind to the ligand under conditions favorable for the formation of apo-protein crystals, due to crystal packing interfering with the CoA binding site.

The five commercial 'coarse' screens described above were tested again, this time in the presence of a 25-fold excess of malonyl-CoA **130**. Hexagonal crystals (100–150 μm) appeared under some conditions after 3–4 days and were stable for weeks (Figure 6.3 B). The best conditions were from the SaltRX screen containing 1.2 M trisodium citrate (pH 9.1) and 0.1 M Tris HCl (pH 8.5), or 1.4 M trisodium citrate (pH 9.1) and 0.1 M HEPES pH (7.5). No hexagonal crystals grew in the JCSG screen but instead, crystals with a morphology similar to the apo-ThnE form were observed. The apo-ThnE orthorhombic spacegroup $P2_12_12_1$ was confirmed by indexing images collected for these crystals.

Two optimisation grid screens were set up, one at pH 8.0 and the other one at pH 9.0. Both varied reservoir solution containing 0.6–1.4 M trisodium citrate and 0.06–0.2 M Tris HCl. The ThnE protein was found to crystallise in the presence of malonyl-CoA **130** in a highly reproducible manner, giving hexagonal plate crystals under conditions contained within 0.89–1.32 M trisodium citrate and

0.06–0.16 M Tris HCl.

6.3.2 Results

X-ray diffraction data were collected for ThnE in complex with three CoA thioester ligand combinations. One for ThnE co-crystallised with malonyl-CoA (2.00 Å resolution, Table 6.2, Entry III), one for ThnE co-crystallised with CoASH and L-proline (1.66 Å resolution, Table 6.2, Entry IV) and one for ThnE co-crystallised with acetyl-CoA and L-proline (1.65 Å resolution, Table 6.2, Entry V). The ThnE crystals co-crystallised with these combination of ligands showed the same hexagonal plate morphology and similar crystallographic unit cell constants (Figure 6.3 B, space group $P6_322$; a, b, c = 110.0 Å, 110.0 Å, 110.5 Å; α, β, γ = 90°, 90°, 120°).

The main differences between these structures and the apo-ThnE structure described above are the presence of additional density for helices α-2 (residues 105–118) and α-11 (residues 277–293) as well as for a loop connecting sheet β-2 and helix α-1 (residues 63–68) (shown in pink in Figure 6.6). These regions of the proteins are probably stabilised by the presence of CoA derivatives, helices α-2 and α-11 forming a lid above the active site, and preventing a second crystallographic trimer to fill this space in a way similar to the apo-ThnE crystal form (Section 2.5.1).

Unlike CarB, in which only one chain of the crystallographic trimer contains a CoA ligand, all three monomers of the ThnE crystallographic trimer contain density corresponding to a CoA ligand in the active site (not shown). This observation suggests that either CarB and ThnE trimers have different ligand binding properties, or that the presence of only one CoA ligand by CarB trimer is an 'artifact' from the crystallisation conditions. When ThnE was co-crystallised in the presence of malonyl-CoA **130** alone, the cosubstrate was observed to undergo apparent decarboxylation and thioester hydrolysis to give CoASH. This proposal is based on the observation of an apparent acetate molecule binding in the active site, and given that there is no acetate in the crystallisation solution (Figure 6.5 A). This uncoupled malonyl-CoA turnover has been observed for CarB in solution [5].

The carboxylate of the apparent acetate molecule in the ThnE structure co-crystallised with malonyl-CoA (Table 6.2, Entry III) forms the same hydrogen-bond interactions as the L-proline

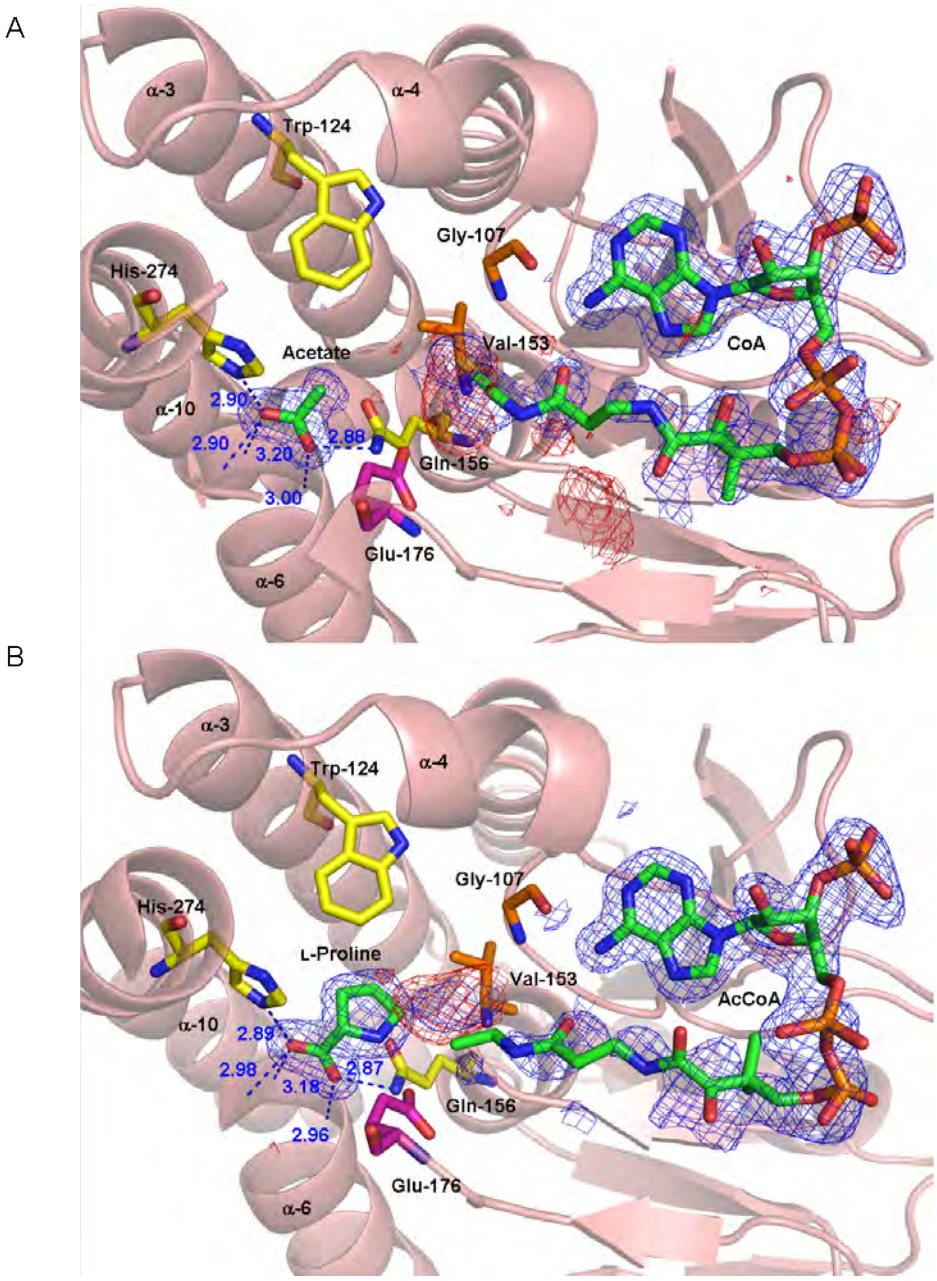


Figure 6.5: Views from the active site of ThnE. Residues thought to be important for catalysis are highlighted in sticks representation. The ligands are displayed as green sticks. The electron density maps is shown for the ligands; The blue map is contoured at 1σ and calculated with coefficients of the form ($2F_0$ - F_c), where F_0 was the native structure factor amplitude and F_c the calculated structure factor amplitude. The red map was calculated with coefficients of the form (F_0 - F_c). The secondary structure elements are represented as cartoons. The two oxyanion hole-forming residues, Gly-107 and Val-153 are represented as dark orange sticks. Glu-176 proposed to participate to thioester hydrolysis is represented as purple sticks. Other active site residues are represented as yellow sticks. Ligand interactions are shown with blue dotted lines and indicated in Å. Some residues, including helices α-2 and α-11, were removed for clarity. (A) The active site of ThnE in complex with CoASH **133** and acetate; (B) The active site of ThnE in complex with acetyl-CoA **127** and L-proline **134**. This figure was generated using PyMOL [3].

carboxylate in the other two complexes (Table 6.2, Entries IV and V). The residues involved in the binding of the L-proline carboxylate are the side-chain of residues His-274 and Gln-156 as well as amide hydrogens (Asn-184, Phe-185 and Gly-186) which are located in the *N*-terminus of helix α-6 (Figure 6.5 B). In the case of L-proline, the amine of the ligand is positioned to form an additional hydrogen-bonding interaction with the carboxylate of Glu-176 (2.61 Å, not shown in Figure 6.5 for clarity), the residue proposed to participate in thioester hydrolysis by comparison with CarB (Section 2.5.3, Scheme 2.10) [5]. Unfortunately, none of the structures in complex with CoA derivatives contain visible electron density for the entire CoA ligand. There is well-defined electron density for the adenosyl moiety in every case. However, only part of the pantetheinyl chain is visible, with the density being increasingly diffuse when further away from the adenosyl group (Figure 6.5 A and B). There is strong electron density (shown in red mesh in Figure 6.5 A and B) near the residues forming the oxyanion hole but it is not defined well enough to model a thiol or thioester functionality. This might be due to alternative conformations or partial occupancy for acetyl-CoA and CoASH, as well as protein in which the pantetheine moiety is in a conformation pointing out of the active site.

Taken together, the results obtained for the crystallisation of ThnE the presence of malonyl-CoA **130** or acetyl-CoA **127** and L-proline **134** were promising; Decarboxylation of malonyl-CoA **130** to yield acetyl-CoA **127**, which was further hydrolysed to CoASH and acetate within the crystallisation time frame was not unexpected, given the likely instability of malonyl-CoA **130**. Importantly, the presence of visible electron density for apparent acetate and L-proline **134** was considered to give an indication on how the L-GHP substrate may bind in the active site. These structural insights should allow further mutagenesis work in order to modulate the CMPS substrate tolerance without impairing the substrate binding.

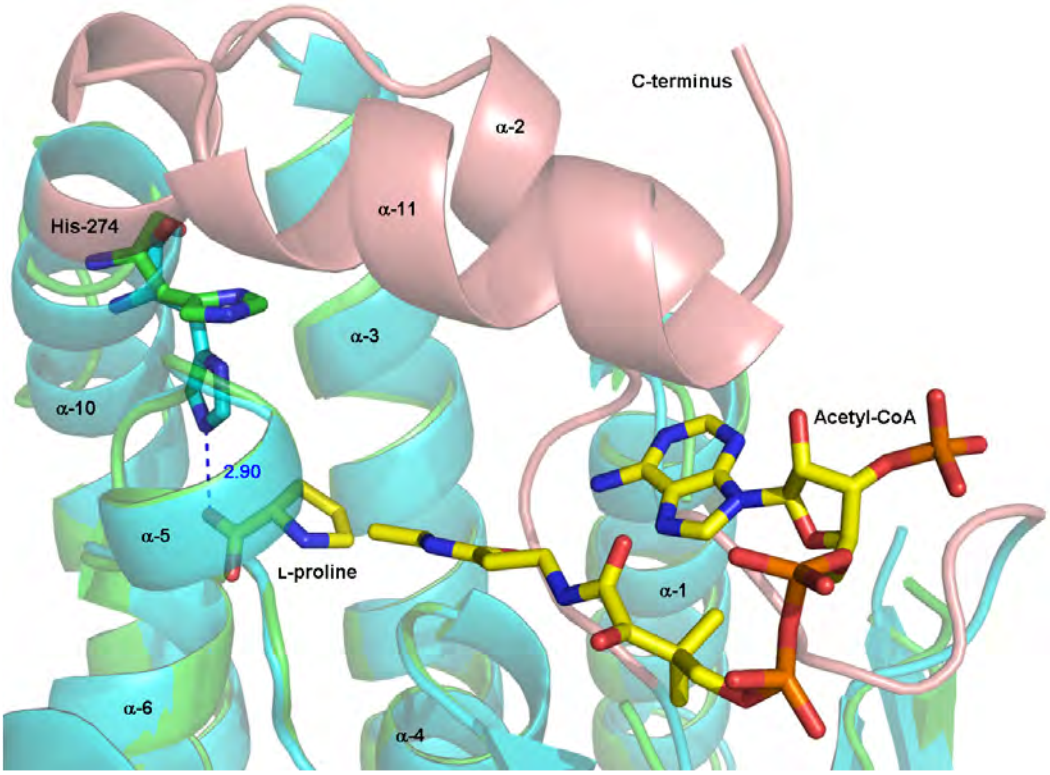


Figure 6.6: Comparison between apo-ThnE and ThnE in complex with L-proline. Superimposition of the apo-ThnE structure (in green) with the ThnE structure in complex with with acetyl-CoA **127** and L-proline **134** (in yellow). Helices α-2 and α-11 are likely disordered in the apo-ThnE structure and therefore only visible in the structure in complex with the ligands and are highlighted in pink. The secondary structure elements are represented as ribbons and labelled if necessary. The His-274 is highlighted with sticks representation. Ligand interactions are shown with blue dotted lines, with distances indicated in Å. Some residues were removed for clarity. This figure was generated using PyMOL [3].

6.4 ThnE Crystallisation Trials with Stable Acyl-CoA Analogues

6.4.1 Crystallisation trials

The stable acyl-CoA analogues described in Chapter 5 were used in co-crystallisation experiments in an attempt to obtain a ThnE structure in complex with a ligand displaying electron density near the oxyanion hole. The stable nature of the ligands tested would prevent the decarboxylation and hydrolysis reactions observed for malonyl-CoA **130** and which led to the difficult interpretation of the observed electron density. The ligands were added to the protein sample in an identical manner to malonyl-CoA **130** or acetyl-CoA **127** prior to setting up the crystallization plates. The same optimisation screen as described in Section 6.2.1 was set up with a series of analogues, including malonyl-oxa(dethia)CoA **262**, malonyl-aza(dethia)CoA **258**, their C-2 methyl derivatives **259** and **263**, malonyl-carba(dethia)CoA methyl ester **288**, the two truncated *t*-CMP-CoA analogues 5-2-[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324**, as well as *t*-CMP-aza(dethia)CoA **313** (Figure 6.7). The protein (20 mg mL^{-1}) and ligand (20 mM) concentrations were kept constant.

6.4.2 Results

The co-crystallisation screens in the presence of non-natural analogues of the CMPS substrate or intermediate of catalysis yielded no crystals under most conditions. Some conditions allowed ThnE crystals to grow, but X-ray diffraction analysis showed that they were apo-ThnE protein crystals.

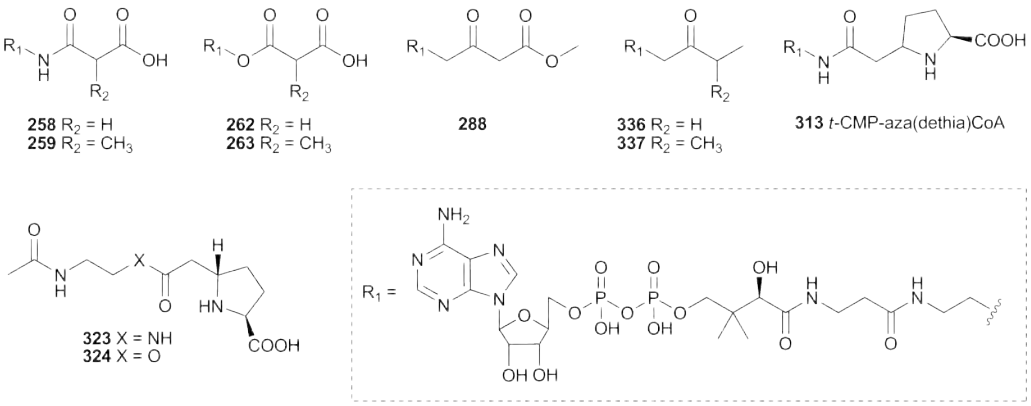


Figure 6.7: Structures of the stable substrate and intermediate analogues tested as ligands for ThnE co-crystallisation. The synthesis of these compounds is described in Chapter 5.

Since *t*-CMP-aza(dethia)CoA **313** was neither a ThnE substrate nor an inhibitor (Section 5.3.2), it might not be expected not to bind. The stable acyl-CoA analogues that have been co-crystallised with other CS enzymes were all substrate analogues, and not analogues of intermediates. In the case of 3,5-dihydroxyphenylglyoxylate synthase (DpgC) for example, the non-hydrolysable substrate analogue DPA-NH-CoA was demonstrated to bind with an affinity similar to that of the natural substrate DPA-CoA, consistent with the analogue acting as a substrate mimic (Section 5.1.5) [6].

In contrast, it is perhaps more surprising that no ThnE crystals could be obtained in complex with malonyl-aza(dethia)CoA **258**, malonyl-oxa(dethia)CoA **262** or their C-2 methyl derivatives **259** and **263**. Malonyl-oxa(dethia)CoA **262** was demonstrated to be a CMPS substrate (Section 5.3.1) and the sulphur substitution for oxygen was not expected to influence the substrate binding to a great extent. No crystals could be obtained in the presence of stable thioester enolate analogues. Under all conditions, the crystallisation trials in the presence of propanoyl-carba(dethia)CoA **329**, isobutanoyl-carba(dethia)CoA **330**, 5-2-[(2-acetamidoethyl)amino]-2-oxoethyl-L-proline **323** and 5-[2-(2-acetamidoethoxy)-2-oxoethyl]-L-proline **324** or *t*-CMP-aza(dethia)CoA **313** were found to contain precipitate only.

Entry	I	II	III	IV	V
Data set	RG000398 A8 3	RG000562 B11 2	RG000703 B10 2	RG000922 A6 3	RG000974 B5 2
Protein	ThnE WT	ThnE V153A	ThnE WT	ThnE WT	ThnE WT
Ligands	Apo	Apo	Malonyl-CoA	CoASH, L-proline	Acetyl-CoA, L-proline
Space group	$P2_12_12_1$	$P2_12_12_1$	$P6_322$	$P6_322$	$P6_322$
Wavelength (Å)	1.5412	1.5412	0.9163	0.9796	0.9794
Unit cell (a, b, c) (Å)	58.3, 110.9, 113.8	58.3, 112.1, 113.8	109.7, 109.7, 110.1	110.0, 110.0, 110.6	109.6, 109.6, 110.5
Angles (deg)	90, 90, 90	90, 90, 90	90, 90, 120	90, 90, 120	90, 90, 120
Resolution range† (Å)	30–1.85 (1.92–1.85)	40–1.90 (1.97–1.90)	50–2.00 (2.07–2.00)	95–1.66 (1.70–1.66)	95–1.65 (1.69–1.65)
Completeness† (%)	98.6 (97.3)	96.1 (94.1)	99.7 (100)	99.7 (99.0)	99.4 (94.2)
R_{work}‡ (%)	27	N.D.	15	N.D.	15
R_{free}‡ (%)	30	N.D.	16	N.D.	16

† Numbers in brackets are for the inner shell.

‡ R_{work} and R_{free} are given only for the structures which were refined for generating figures, based on 10% of the total reflections.

Table 6.2: Summary of ThnE Crystallography results.

6.5 Summary for Chapter 6

6.5.1 A CMPS Crystal Structure in Complex with L-Proline and CoA Gives Insights into Substrate Binding

Initial crystallisation screens showed that a variety of conditions result in the crystallisation of ThnE protein in an apo form. X-ray diffraction analyses showed that in all these crystallisation conditions, the apo-ThnE crystals formed had the same morphology, spacegroup and cell constants. No electron density was visible in the active site even in the presence of high concentrations of titrated organic acids such as malonate and acetate. Optimisation screens allowed the reproduction of crystal formation in a robust manner, but addition of various ligands, including all of the CoA analogues and truncated *t*-CMP-CoA analogues described in Section 5.2 did not produce any ThnE crystals in complex with ligand.

A second series of ThnE crystallography screens was set up in the presence of malonyl-CoA, in the hope that the protein in complex with a ligand would crystallise under conditions different from that of apo-ThnE. A number of hits were obtained and optimisation screens enabled the formation of ThnE crystals in complex with malonyl-CoA **130**, acetyl-CoA **127** or CoASH **133**. When L-proline **134** was added to acetyl-CoA **127** or CoASH **133**, electron density was visible for this second ligand as well. All the crystals containing a CoA ligand showed the same morphology, different from the apo-enzyme, having an hexagonal plate shape (Figure 6.3 B). A complete dataset was collected for three of these crystals and two structures refined. When ThnE was co-crystallised with malonyl-CoA **130**, the ligand was found to undergo decarboxylation and thioester hydrolysis to yield CoASH **133** and acetate, both binding in the active site. Visible electron density is present for both ligands, although the CoASH map is not clearly defined towards the end of the pantetheine moiety. There is a visible electron density near the oxyanion hole, suggesting the presence of a sulphur atom, but since there is no continuous density for the entire CoASH ligand, the refinement process was found to be difficult for that region. Reproducing the crystallisation conditions in the presence of acyl-CoA ligands but using acidic conditions might be a way to prevent, or slow down, the decarboxylation and hydrolysis reactions observed in the case of malonyl-CoA. However, the recombinant ThnE protein was found to precipitate under acidic conditions (pH $\leq$ 7.0) suggesting that using a lower pH for

crystallography might not be possible.

6.5.2 None of the Stable CoA Analogues Co-Crystallise with CMPS

The absence of ThnE crystal formation with any of the synthetic analogues was disappointing. The absence of binding properties for the non-hydrolysable and truncated *t*-CMP-CoA analogues could be explained by a low affinity of the intermediate of catalysis *t*-CMP-CoA **132** for the enzyme, in accordance with the observation that in the CMPS assay, *t*-CMP-CoA **132** and *t*-CMP-oxa(dethia)CoA **314** can escape the enzyme-catalysed hydrolysis step to a certain extent (Figure 5.7 A). However, there is no clear explanation why malonyl-aza(dethia)CoA **258** and malonyl-oxo(dethia)CoA **262** could not be co-crystallised with ThnE under conditions similar to malonyl-CoA **130** co-crystallisation.

The presence of L-proline **134** and acetate in two of the ThnE structures enables speculation on the position of the CMPS substrate, L-GHP **129**, in the active site. Most probably, the carboxylate function of L-GHP **129** binds in a similar way to that of L-proline and acetate. This is a good indication for further mutagenesis studies. Broadening the substrate tolerance of CarB and ThnE would allow to generate more product diversity, eventually moving away from *t*-CMP derivatives. Potentially, the structural information contained in the ThnE structure in complex with L-proline could allow mutagenesis work towards enzyme repurposing. An interesting target would for example be CMPS variants that could accept D-GHP.

References

[1] Sleeman, M. C., Sorensen, J. L., Batchelar, E. T., McDonough, M. A., and Schofield, C. J. (2005) Structural and mechanistic studies on carboxymethylproline synthase (CarB), a unique member of the crotonase superfamily catalyzing the first step in carbapenem biosynthesis. *J Biol Chem*, **280**, 34956–34965.

[2] Hamed, R. B., Mecinović, J., Ducho, C., Claridge, T. D. W., and Schofield, C. J. (2010) Carboxymethylproline synthase catalysed syntheses of functionalised *N*-heterocycles. *Chem Commun*, **46**, 1413–1415.

[3] Schrödinger, LLC (2010) The PyMOL molecular graphics system, version 1.3r1.

[4] Hamed, R. B., Batchelar, E. T., Mecinović, J., Claridge, T. D. W., and Schofield, C. J. (2008) Evidence that thienamycin biosynthesis proceeds via C-5 epimerization: ThnE catalyzes the formation of (2*S*,5*S*)-*trans*-carboxymethylproline. *ChemBioChem*, **10**, 246–250.

[5] Batchelar, E. T., Hamed, R. B., Ducho, C., Claridge, T. D. W., Edelmann, M. J., Kessler, B., and Schofield, C. J. (2008) Thioester hydrolysis and C–C bond formation by carboxymethylproline synthase from the crotonase superfamily. *Angew Chem Int Ed Engl*, **47**, 9322–9325.

[6] Widboom, P. F., Fielding, E. N., Liu, Y., and Bruner, S. D. (2007) Structural basis for cofactor-independent dioxygenation in vancomycin biosynthesis. *Nature*, **447**, 342–345.

[7] Otwinowski, Z. and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. Charles W. Carter, J. (ed.), *Macromolecular Crystallography Part A*, vol. 276 of *Methods in Enzymology*, pp. 307 – 326, Academic Press.

[8] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40**, 658–674.

[9] Adams, P. D., et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213–221.

[10] Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr*, **65**, 1074–1080.

[11] Weininger, D. (1990) SMILES. 3. DEPICT. Graphical depiction of chemical structures. *J Chem Info Comp Sci*, **30**, 237–243.

[12] Emsley, P. and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, **60**, 2126–2132.

6.6 Materials and Methods for Chapter 6

6.6.1 CMPS Crystallography

Crystallisation screens

Commercial 'coarse' screens tested for the determination of hit conditions were JCSG-plus (Molecular Dimensions), PEG/Ion (Hampton Research), Index (Hampton Research), Crystal and SaltRX (Hampton Research). Optimisation screens were set up in 96-well blocks (2 mL in each well) using a Janus^{TM} automated workstation (Perkin-Elmer). Screen components were commercially available from Hampton Research.

Apo-ThnE V153V and ThnE V153A crystallisation trials

Crystals of apo-ThnE V153V Δ2–46 were grown using a high throughput sitting drop vapor diffusion methodology in Intelli-plate^{TM} three-drop 96-well crystallisation plates (Art Robbins Instruments) and incubated at 4 °C or 20 °C in a Minstrel III^{TM} automated workstation (Rigaku). Three different ratio of protein to reservoir solution were used, including 2:1 (300 nL total volume), 1:1 (200 nL total volume) and 1:2 (300 nL total volume) and the drop was equilibrated against 80 μL of reservoir solution. The ThnE V153V Δ2–46 protein solution contained 20 mg ml^{-1} protein in 100 mM Tris-HCl (pH 7.5). The screens tested included JCSG-plus, PEG-Ion and INDEX.

Apo-ThnE crystallisation optimisation

Crystals of apo-ThnE V153V Δ2–46 were grown using a sitting drop vapor diffusion methodology at 20 °C as above with an optimised screen. The best crystals grew under the conditions described in Table 6.1.

ThnE V153V co-crystallisation with CoA analogues

A set of crystallisation 'coarse' screens was tested as above for ThnE V153V Δ2–46, this time in the presence of CoA derivatives and analogues, using the JCSG-plus (Molecular Dimensions), Crystal and SaltRX (Hampton Research) screens. Diffraction quality crystals of ThnE in complex with malonyl-CoA **130**, CoASH **133** and L-proline, and acetly-CoA **127** and L-proline were grown as above using an optimised screen with reservoir solution containing 0.6–1.4 M trisodium citrate (pH 8.0 and 9.0) and 0.06–0.2 M Tris HCl (pH 8.5).

Data collection and processing

All X-ray diffraction data collection was performed with the help of Dr Michael A. McDonough on an in-house X-ray source at the Chemistry Research Laboratory in Oxford, U.K., or at the Diamond Light Source, Harwell, U.K. (beamlines I03-1 and I04-1). All X-ray diffraction data processing was performed by Dr Michael A. McDonough at the Chemistry Research Laboratory in Oxford, U.K.

When using the in-house X-ray source, the crystals were cryoprotected using well solution diluted to 25 or 30% v/v glycerol and flash cooled in an Oxford Cryosystems nitrogen gas stream. Data were collected from a single crystal at 100K using a Rigaku FR-E+ Superbright diffractometer equipped with a copper rotating anode, Osmic HF optics, and a Saturn 944+ CCD detector. At the SRS X-ray source, the crystals were processed as above and data collected from a single crystal at 100K.

The data were indexed, integrated, and scaled using HKL2000 [7] and the structure was determined by molecular replacement using the AutoMR (PHASER) [8] subroutine in PHENIX [9] using ThnE coordinates from Prof. Karen Valgard (Uppsala, Sweden) as a search model. Restraints for the ligands were generated by eLBOW [10] using the appropriate SMILES [11] strings. Iterative rounds of model building and refinement using Coot [12] and PHENIX [9] were performed until the decreasing R_{work} and R_{free} no longer converged. The final R-factors for the models were: apo-ThnE, R_{work} = 30% and R_{free} = 27%; ThnE in complex with malonyl-CoA **130**, R_{work} = 15% and R_{free} = 16%; ThnE in complex with acetly-CoA **127** and L-proline **134**, R_{work} = 15% and R_{free} = 16%.

Chapter 7

Characterisation of a Putative Carboxylate Reductase Enzyme-Pair

7.1 Introduction

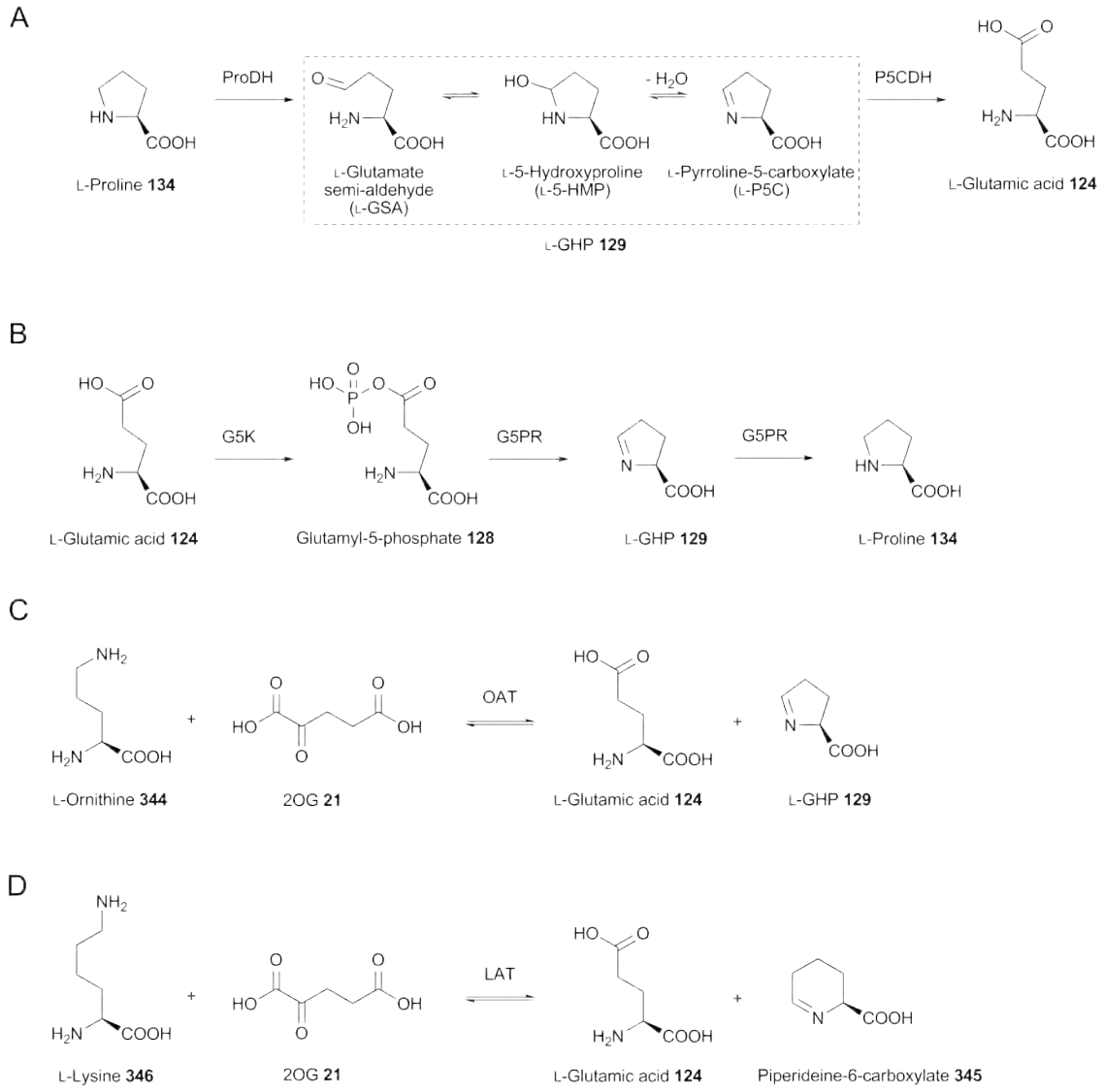
The results for CMPS catalysis using analogues of L-GHP presented in Chapter 4 were encouraging but also pointed to the limitations of using CMPS enzymes as biocatalysts. In particular, the synthesis of L-GHP analogues described in Chapter 3 can be challenging and costly. The possibility of a biocatalytic approach to amino acid aldehyde CMPS substrates was therefore investigated.

7.1.1 L-GHP Occurrence in Nature

L-GHP **129** is a metabolic intermediate in all life forms. A brief description of various enzymes known to generate L-GHP **129** is given below.

L-Proline Metabolism

Two catalytic steps are involved in the conversion of L-proline **134** into L-glutamic acid **124** (Scheme 7.1 A) [1]. These reactions are performed by the bifunctional flavoprotein proline utilisation A (PutA) enzyme. The first step is catalysed by a flavin containing domain that has proline dehydrogenase (ProDH) activity. The second step is catalysed by the NAD^+ dependent Δ^1-L-pyrroline-5-carboxylate dehydrogenase (P5CDH) domain. Based on crystallographic studies, the two



Scheme 7.1: Occurrence of L-GHP in nature. Various metabolic pathways involve L-GHP **129** as an intermediate. (A) The conversion of L-proline **134** into L-glutamic acid **124** proceeds in two steps. The first step is catalysed by proline dehydrogenase (ProDH) and the second step by pyrroline-5-carboxylate dehydrogenase (P5CDH); (B) The reverse process converting L-glutamic acid **124** into L-proline **134** requires three steps. First L-glutamic acid **124** is phosphorylated by γ-glutamate kinase (or glutamate-5-kinase, G5K). This γ-glutamylphosphate intermediate is reduced by γ-glutamylphosphate reductase (or glutamyl-5-phosphate reductase, G5PR) to form L-GHP **129**. L-GHP **129** is finally reduced by pyrroline-5-carboxylate reductase (P5CR); (C) L-GHP **129** is also an intermediate in the metabolism of L-ornithine **344**. Ornithine aminotransferase (OAT) catalyses the reversible P5P-dependent reaction, using L-ornithine and 2OG **21** as substrates, and forming L-GHP **129** and L-glutamic acid **124**; (D) In a similar manner, lysine aminotransferase (LAT) catalyses the reversible P5P-dependent production of piperideine-6-carboxylate **345** from L-lysine **346**.

catalytic centers are proposed to be connected by a substrate channeling domain which prevents the escape of the semi aldehyde intermediate into the cytoplasm [2]. In some Gram-negative bacteria such as *E. coli*, PutA is proposed to be a trifunctional protein because, in addition to the two enzymatic functions, PutA also links metabolism and gene regulation by repressing transcription of the proline utilisation regulon. The PutA regulatory function has been reviewed by Zhou *et al.* [3].

L-Glutamic Acid Metabolism

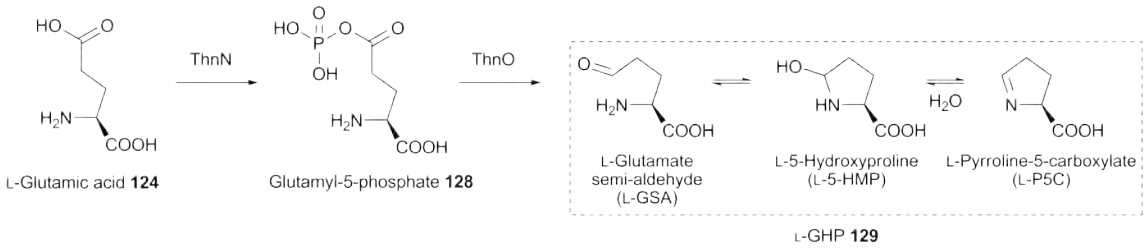
The conversion of L-glutamic acid **124** into L-proline **134** requires three steps (Scheme 7.1 B) [4]. In the first step, L-glutamic acid **124** is phosphorylated by γ-glutamate kinase (or glutamate-5-kinase, G5K) in an ATP-dependent manner. The glutamyl-5-phosphate **128** intermediate is then reduced by the NAD(P)H-dependent γ-glutamylphosphate reductase (or glutamyl-5-phosphate reductase, G5PR) to form L-GHP **129**. In higher organisms, these two functions are encoded by a single pyrroline-5-carboxylate synthase (P5CS) operon which may have evolved from the two separate genes present in prokaryotes. The final step is catalysed by a pyrroline-5-carboxylate reductase (P5CR) (EC:1.5.1.2). This enzyme catalyses the NAD(P)-dependent oxidation of L-GHP **129** into L-proline **134**.

L-Aspartic Acid Metabolism

L-Aspartic acid **344** can be converted to the amino acids L-lysine **346**, L-methionine, L-isoleucine and L-threonine, via L-aspartate semialdehyde **221**. L-aspartic acid **344** is phosphorylated by three aspartate kinases (AK I-III) and β-aspartyl phosphate is converted to aspartate semialdehyde **221** by the enzyme aspartate semialdehyde dehydrogenase (ASA-DH) [5]. But since aspartate semialdehyde **221** was found not to be a substrate for either CarB or ThnE [6], this pathway is not relevant in the context of CMPS biocatalysis.

L-Ornithine and L-Lysine Metabolism

L-GHP is also an intermediate in the metabolism of L-ornithine. Ornithine aminotransferase (OAT) is a pyridoxyl-5′-phosphate (P5P)-dependent enzyme which uses 2-oxoglutarate **21** (2OG)



Scheme 7.2: Proposed roles for ThnN and ThnO in carbapenem biosynthesis. ThnN and ThnO were postulated to be involved in the production of L-GHP from L-glutamate **345** in *S. cattleya*. Rather than L-glutamate-5-adenosylmonophosphate, L-glutamyl-5-phosphate is proposed as an intermediate, based on the observations described in Section 7.2.2.

as a cosubstrate to convert L-ornithine into L-GHP **129** and L-glutamic acid **124** (Scheme 7.1 C). In a similar way, L-lysine is metabolised into 2-aminoadipic acid in two catalytic steps, for example in cephamycin producing bacteria [7, 8]. The first step is catalysed by lysine aminotransferase (LAT) and produces L-1-piperideine-6-carboxylic acid which is then reduced to 2-aminoadipic acid in a NAD^+-dependent manner (Scheme 7.1 D). This pathway is of interest in the context of CMPS biocatalysts since L-1-piperideine-6-carboxylic acid was demonstrated to be a substrate for CarB variants [9]. OAT could be used as a L-1-piperideine-6-carboxylic acid source in the context of a CMPS coupled assay.

7.1.2 Proposed L-GHP Sources in Carbapenem Producers

Analyses of the genes present in the biosynthetic clusters of (5*R*)-1-carbapen-2-em-3-carboxylic acid **115** and thienamycin **116** suggest different origins for L-GHP **129** in *P. carotovorum* (Scheme 2.6) and *S. cattleya* (Scheme 7.2). It is interesting to note that, despite L-GHP **129** being an intermediate of several metabolic pathways, these species seem to have specific enzymes dedicated to the production of L-GHP **129** in the context of carbapenem biosynthesis.

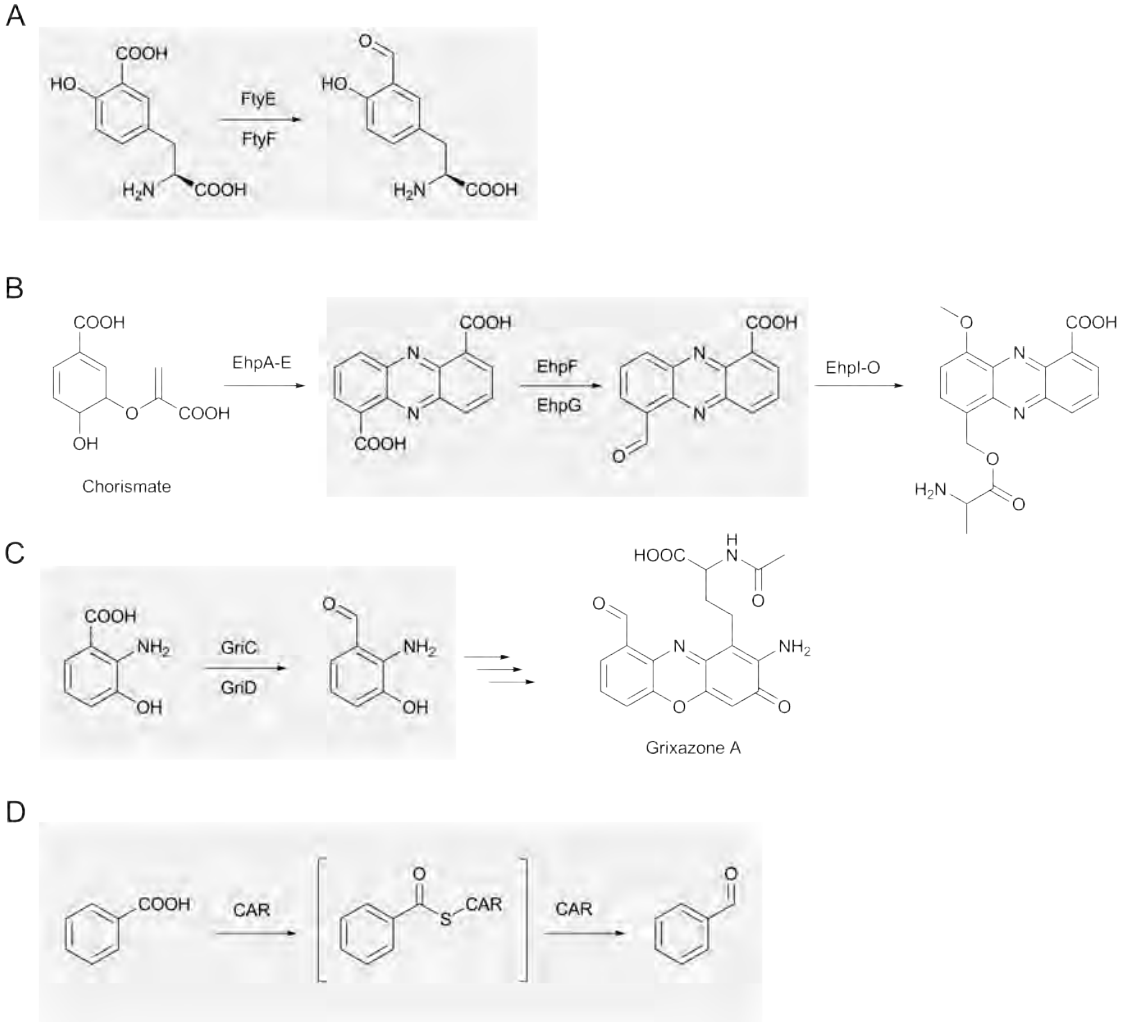
In *P. carotovorum*, CarD and CarE are proposed to be responsible for the conversion of L-proline **134** into L-GHP **129** (Section 2.2.2, Scheme 2.6), whilst in *S. cattleya*, ThnN and ThnO are proposed to use L-glutamic acid **124** as a substrate. The *thnN* and the *thnO* genes show similarities with other pairs of prokaryotic enzymes proposed to catalyse the reduction of carboxylic acids into the corresponding aldehydes [10]. Examples include *ftyE* and *ftyF* from the biosynthetic gene cluster of

3-formyl-tyrosine metabolites in *Pseudoalteomonoas tunicata* (Scheme 7.3 A) [11], *ehpF* and *ehpG* from the D-alanylgriseoluteic acid biosynthetic gene cluster in *Pantoea agglomerans* (also known as *Erwinia herbicola*, Scheme 7.3 B) [12], as well as *griC* and *griD* from the grixazone biosynthetic cluster in *Streptomyces griseus* (Scheme 7.3 C) [13]. The sequence alignments for these enzymes with ThnN and ThnO are given in Appendix E.3 and Appendix E.6, respectively.

The first enzyme of the pair, including ThnN, is proposed to belong to the ANL superfamily of adenylating enzymes which include three main subfamilies, i.e. the acyl-CoA synthetases, the NRPS adenylation domains, and the luciferase enzymes [14]. Members of the ANL superfamily are enzymes that typically catalyse two partial reactions; Firstly, they use ATP as a cosubstrate for the magnesium-dependent adenylation of a carboxylate to form an acyl-AMP intermediate. The adenylate intermediate is then used as a substrate for a second partial reaction, *e.g.* the formation of a thioester via nucleophilic substitution. The nucleophile can be the thiol or hydroxyl group of an internal cysteine or serine, CoASH **133** or a protein-bound phosphopantetheine, *e.g.* in the ACP domains of modular non-ribosomal peptide synthetases (NRPSs). ThnN was therefore proposed to catalyse the reaction between L-glutamic acid **124** and ATP to yield L-glutamate-5-adenosylmonophosphate (G5A). ThnO was proposed to use G5A as a substrate and catalyse the reduction to L-glutamate-5-semialdehyde (GSA) (Scheme 7.2) although there is no reported example of a reductase acting on a carboxylate activated as the adenylate. However, the pair of enzymes involved in the reduction of L-glutamic acid **124** to L-GHP **129** are known to catalyse this reaction via a phosphorylated carboxylate.

7.1.3 The Carboxylate Reductases

Except for the enzymes participating to the L-glutamic acid metabolism (Section 7.1.1), there is little knowledge about the biochemistry of other carboxylate reductase enzyme-pairs. In the case of grixazone biosynthesis, disruption of either the *griC* or *griD* gene resulted in the accumulation of the precursor 3-acetylamino-4-hydroxybenzoic acid (Scheme 7.3 C) [13]. The initial proposal that GriC and GriD could compose a carboxylic acid reductase enzyme-pair was based on sequence similarities of GriC to AMP-binding proteins and of GriD to NAD(P)H-dependent aldehyde dehydrogenases. Although there is no apparent sequence similarity between GriC or GriD and the adenylating *N*-terminal domain or the reducing *C*-terminal domain of the carboxylic acid reductase (CAR) enzyme from *Nocardia sp.* NRRL 5646 (Scheme 7.3 D), CAR was cited as an example of such an aromatic



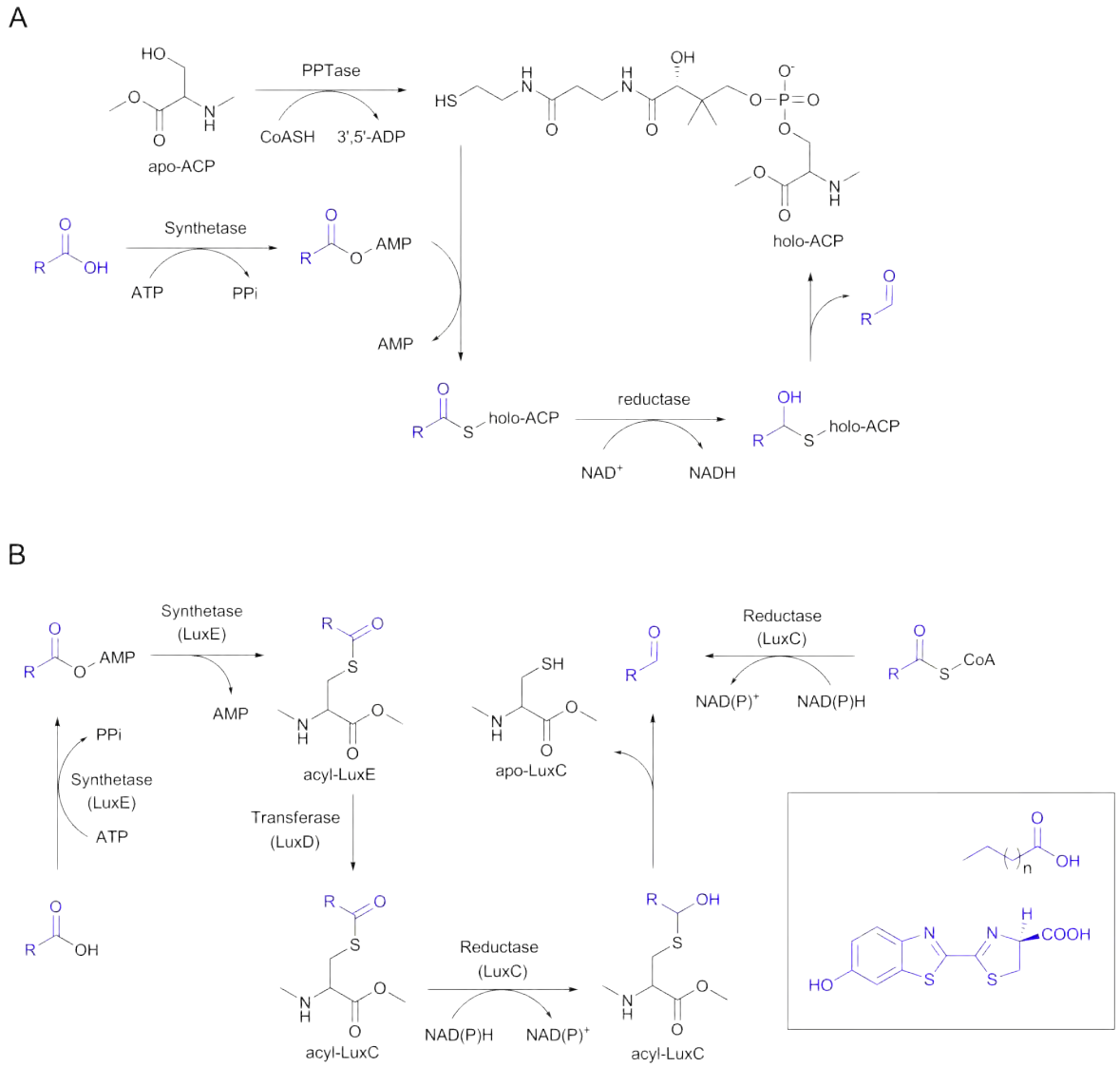
Scheme 7.3: Carboxlic acid reductase systems proposed to be involved in the biosynthesis of various natural products. The carboxylate reductase step is highlighted in a gray box for each pathway. (A) *Pseudoalteomonoas tunicata* FtyE and FtyF; (B) *Pantoea agglomerans* EhpF and EhpG; (C) *Streptomyces griseus* GriC and GriD; (D) *Nocardia* sp. carboxylic acid reductase (CAR).

carboxylate reductase because of the analogies in the reactions catalysed [15, 16, 17]. Rosazza and coworkers have shown that CAR carries both the adenylase and reductase activities within one single polypeptide chain. The two domains are linked by a region carrying a phosphopantetheine attachment site and the activity of CAR was found to be strictly dependent on phosphopantetheinylation at this site [18]. The proposed mechanism for the carboxylic acid reductase activity of CAR is a three step process (Scheme 7.4 A). Both ATP and the carboxylic acid initially bind to the enzyme and react to form a mixed 5′-adenylic acid-carbonylanhydride intermediate [19]. The intermediate is then transfered to the pantetheine arm to form a thioester-linked substrate with the release of AMP. This thioester bond is finally reduced by the NADPH-dependent domain.

Other carboxylate reductase systems have been described, one prominent example being the light producing luciferase enzymes isolated from *Vibrio fischeri* which have had significant impact on molecular biology as reporter genes [20]. Light producing bacteria, such as *Photobacterium phosphoreum* and *Vibrio* species, use luciferin and fatty aldehydes as substrates for their luminescent reaction [21]. The reduction of fatty acids into the corresponding aldehydes is catalysed by the LuxCDE system (Scheme 7.4 B). The free fatty acids are activated into the corresponding fatty acyl adenylate by a synthetase LuxE. The intermediate reacts with a cysteine of LuxE synthetase to form a thioester bond [22]. The acyl group is then transfered onto a cysteine of the reductase protein LuxC by a transferase LuxD [23, 22]. The reduction domain uses NADPH as a hydride donor to reduce the thioester bond and the aldehyde product is released by hydrolysis, regenerating the cysteine thiol on LuxC.

7.1.4 Objectives

This chapter describes the preliminary characterisation of a pair of enzymes, ThnN and ThnO, proposed to be responsible for the formation of L-GHP **129** in *S. cattleya*. Bioinformatic analyses indicate that ThnN and ThnO share significant sequence similarity to other pairs of enzymes proposed to convert carboxylic acids into aldehydes [10]. No molecular characterisation of any of these pairs has been reported so far. It was hoped that active recombinant ThnN and ThnO would be useful for the generation of L-GHP analogues from their corresponding L-glutamate analogues. In addition, the characterisation of ThnN and ThnO might aid in developing an understanding of other predicted



Scheme 7.4: Proposed mechanisms for the reduction of carboxylic acids. The formation of an acyl-adenylate species has been demonstrated in the CAR and LuxCDE-catalysed reduction of carboxylic acids into the corresponding aldehydes. (A) Proposed mechanism for the reduction of aryl-carboxylic acid by CAR. An initial phophopantetheination of the apo enzyme by a transferase (PPTase) is necessary for activity. The adenylation domain catalyses the formation of an aryl-adenylate intermediate which is transfered onto the phosphopantetheine arm. This intermediate can then reach the NADPH-dependent reductase domain which catalyse the hydride transfer and the aldehyde product is released to regenerate the free thiol; (B) Proposed mechanism for the LuxCDE carboxylate reductase system.

carboyxlate reductase enzyme-pairs.

7.2 Results

7.2.1 Gene cloning, protein expression and purification

Initial attempts to clone the *thnN* and *thnO* genes individually from *S. cattleya* genomic DNA were unsuccessful. A successful strategy was to amplify both adjacent genes together from the genomic DNA and clone them as a single *thnON* fragment into a pET-24a(+) vector (personal communication from Dr R. Hamed). Both the *thnN* and *thnO* genes were then amplified from this construct and subcloned in a pET-28a(+) vector expression vector. The *thnN* gene was cloned between the NdeI and BamHI restriction sites and the *thnO* gene between the NdeI and XhoI sites. The transcriptional products are carrying a *N*-terminal hexahistidine tag in this configuration. The plasmids were then transformed into an *E. coli* Rosetta^{TM} 2 strain which contains a chloramphenicol-resistant plasmid carrying the genes for 7 rare codons tRNAs (AGA, AGG, AUA, CUA, GGA, CCC, and CGG). This allows the enterobacterial host to increase the translational efficiency for genes originally from actinomycetes species containing high GC-content genomes such as *S. cattleya*.

Expression trials were performed to optimise expression conditions for ThnN and ThnO proteins. The optimal expression conditions were determined by small scale purification of the hexahistidine tagged proteins using immobilised metal affinity chromatography (IMAC). The ThnN and ThnO expression systems were found to be 'leaky' and no more than 0.1 mM IPTG was needed for expression when the cells were incubated at 16 °C for 16 h.

After large scale expression and cell lysis using these conditions, recombinant ThnN and ThnO were purified using a two-step procedure. Affinity chromatography was performed using a HisTrap^{TM} FF 5 mL column using an imidazole gradient elution (Figure 7.1 B and D). Alternatively, IMAC purification was performed by loading the cell lysate onto a column packed with a resin carrying iminodiacetic acid chelating nickel divalent cations. The resin was washed with lysis buffer, followed by a buffer containing a low imidazole concentration, and then eluted with two sequential increases in imidazole concentration. The purity for ThnN was estimated to be $\geq 90\%$ after the affinity step, as judged by SDS-PAGE analysis. Both samples were further purified using size exclusion chromatography. In order to prevent protein precipitation, 10% v/v glycerol and 1 mM DTT were added to the buffer. Because ThnO was found to precipitate when concentrated to 2 mg ml^{-1}, only

limited amounts could be loaded onto the size exclusion chromatography columns, and this extra purification step was found to have low recovery yield of ThnO (Appendix E.2, Figure E.1 lane O4). The final protein concentration for ThnN was 18.5 mg ml^{-1} (400 μM) and the yield 15.4 g L^{-1} of culture. For ThnO, addition of 10% glycerol and 1 mM DTT allowed concentration to 5 mg ml^{-1} (100 μM) but most aliquots were kept at 2.5 mg ml^{-1} (50 μM) for assay purposes. ThnO overall yield was 3.2 g L^{-1} of culture.

Samples of recombinant ThnN and ThnO proteins were analysed by LC-MS in order to confirm their identity (Figure 7.2). The elution peak for ThnN was found to contain two species (Figure 7.2 C). One has the expected mass for the ThnN polypeptide after removal of the *N*-terminal residue by methionine aminopeptidase in *E. coli* (observed: 43447 Da; calculated: 43437 Da), the other one shows a mass greater than that predicted for ThnN (observed: 43626 Da, + 179 Da), in accordance with the previously described introduction of a gluconoyl group (+ 178 Da) on the *N*-terminus of some recombinant proteins carrying a hexahistidine tag when expressed in *E. coli* [24]. The ThnO sample contained two species, one having a mass similar to that predicted for ThnO (observed: 51465 Da; calculated: 51561 Da). A contaminating band displaying a higher molecular weight was observed by SDS-PAGE analysis (indicated with a asterisk in Figure 7.1 C and D). The contaminating protein was also observed by LC-MS (observed: 51465 Da; calculated: 51561 Da, indicated with a asterisk in Figure 7.2 D).

In order to further investigate the identity of the purified protein, a sample of ThnN was loaded onto SDS-PAGE and the protein content of the gel revealed with Coomassie stain. The band corresponding to the pure protein was cut out and subjected to tryptic digest, followed by MS-MS analysis of the resulting fragments. The protein was identified as ThnN from *S. cattleya* with a sequence coverage of 75% and the score was 4099 as determined using Mascot version 2.3 (Matrix Science). This value represents the negative log of the probability that this identification is a random match (significance threshold filter, $p < 0.01$). The peptide matches are given in Appendix E.3.

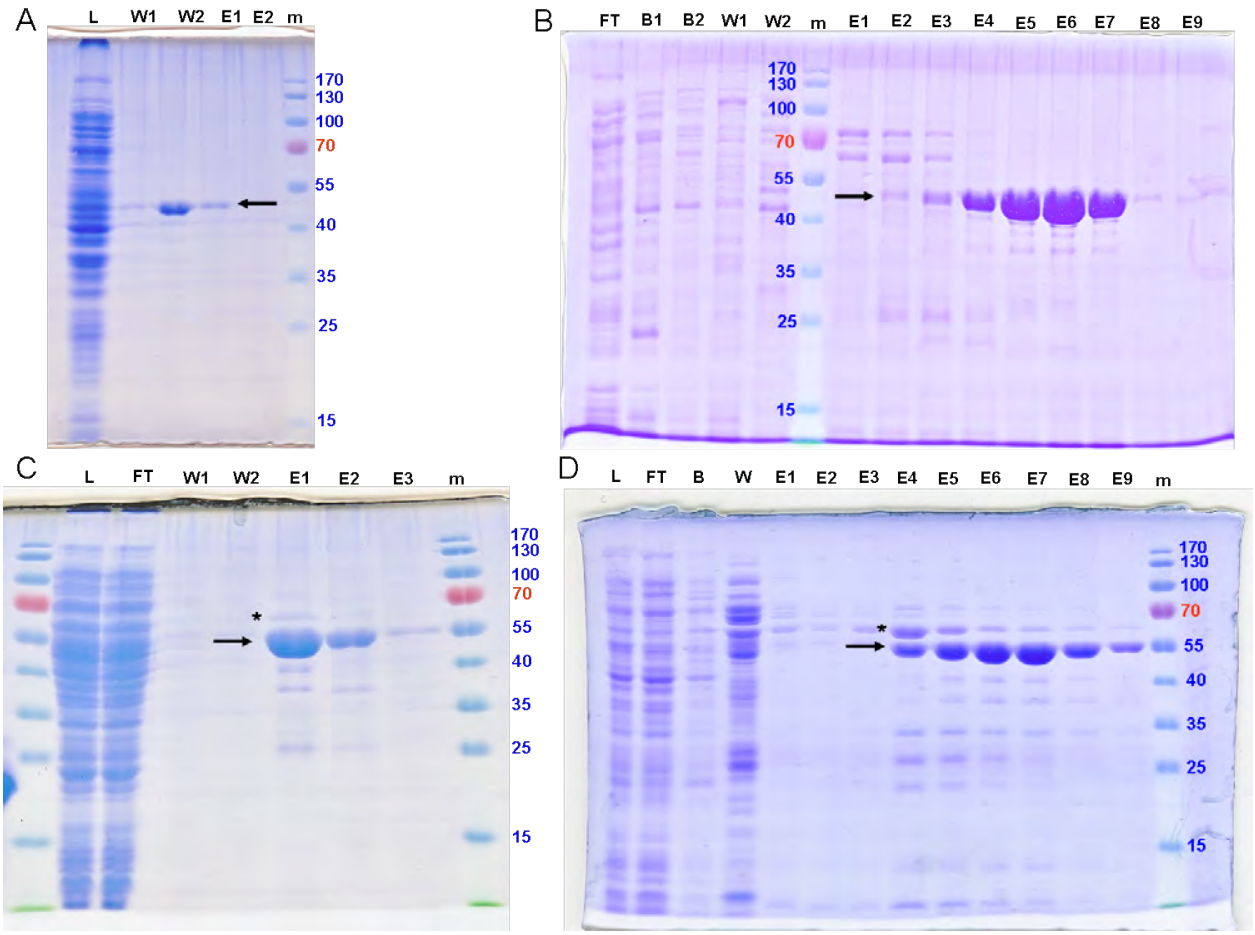


Figure 7.1: SDS-PAGE analysis for representative ThnN and ThnO purification steps. Purification after recombinant expression in *E. coli* of ThnN and ThnO proteins carrying a *N*-terminal hexahistidine tag. All samples were run on 12% acrylamide gels. L, total cell lysate; W, wash fractions; E, elution fractions; m, prestained protein ladder. The expected mass for ThnN was 43568 Da and for ThnO 51693 Da. The molecular weights of the markers are given in kDa. (A) IMAC purification of ThnN; (B) FPLC purification of ThnN using a HisTrap™ FF 5 mL Column; (C) IMAC purification of ThnO. The band marked with an asterix correspond to the impurity also observed in the LC-MS analysis of the ThnO samples (Figure 7.2 D); (D) FPLC purification of ThnN using a HisTrap™ FF 5 mL Column.

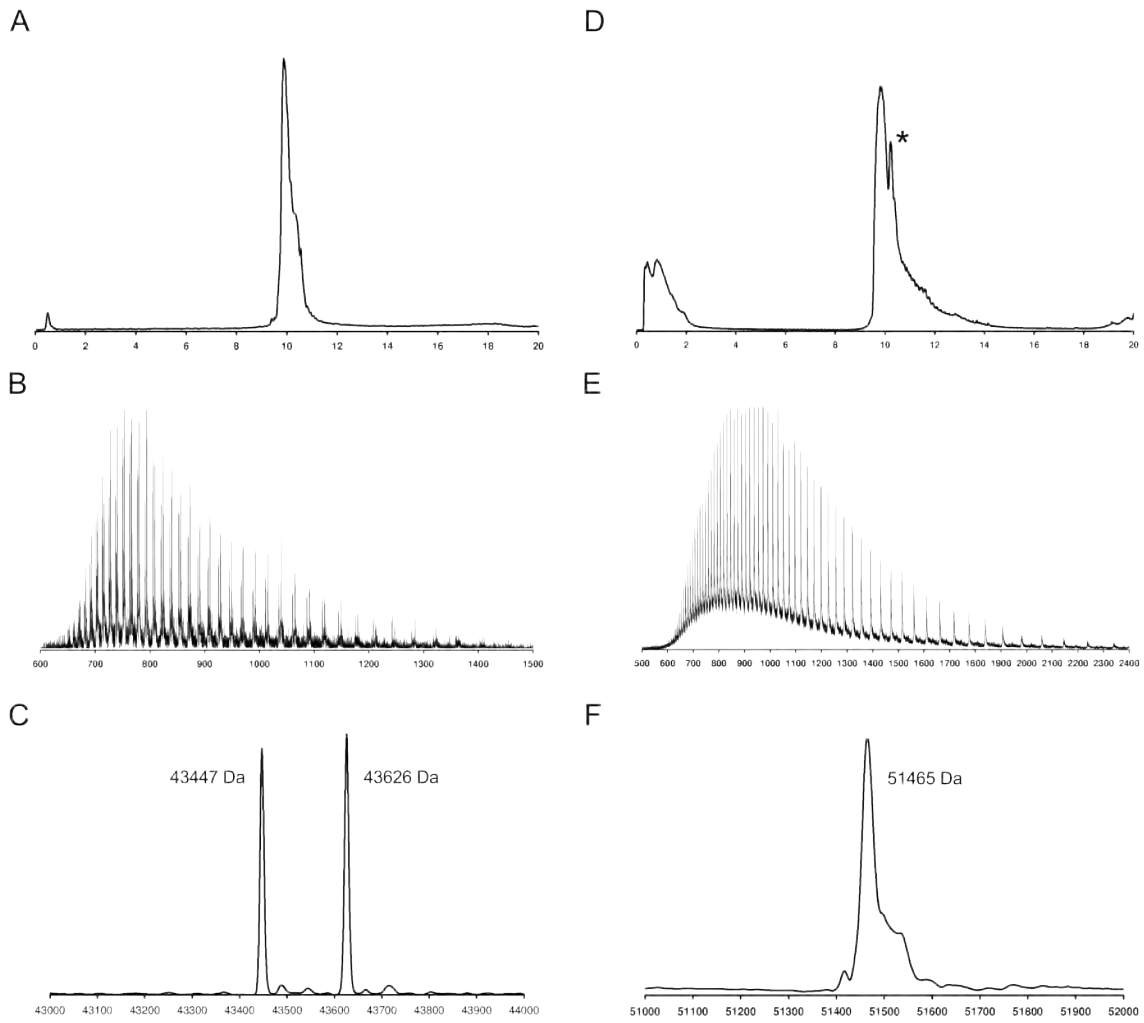


Figure 7.2: LC-MS analysis of recombinant ThnN and ThnO proteins. (A) Chromatogram for ThnN; (B) Mass spectrum for ThnN. Two species, giving each a m/z peak for a given charge, are clearly visible; (C) Deconvoluted mass spectrum for ThnN. The first peak shows a mass of 43447 Da, the second one of 43636 Da corresponding to the addition of a gluconoyl group (+ 178 Da); (D) Chromatogram for ThnO. The main peak is ThnO (51465 Da), the peak marked with an asterix contains an unidentified species with a mass of 57098 Da; (E) Mass spectrum for ThnO; (F) Deconvoluted mass spectrum for ThnO showing a mass of 51465 Da.

7.2.2 Preliminary Characterisation of ThnN and ThnO

Coupled Assays with CMPS do not Produce *t*-CMP

In order to test the hypothesis that ThnN and ThnO catalyse the production of L-GHP, a CMPS coupled assay was investigated. It was hoped that coupling ThnN and ThnO formation of L-GHP with CMPS activity may produce *t*-CMP. *t*-CMP could then be detected by LC-MS as described in Section 4.3. A typical coupled assay contained the CMPS assay components except for L-GHP. Instead, ThnN and ThnO were added, as well as their predicted substrate and cofactors, i.e. L-glutamate, ATP, Mg^{2+}, and NADH or NADPH. Addition of L-GHP to the assay restored t-CMP production, excluding that the ThnN and ThnO components inhibit CMPS activity (data not shown). Different conditions were tested for the CMPS coupled assay. In one case, CMPS and cosubstrate malonyl-CoA were immediately co-incubated with the ThnN and ThnO assay components. In the other case, preincubation of L-glutamic acid with ThnN and ThnO, ATP, Mg^{2+} and NADH/NADPH for a given time was followed by addition of the CMPS reaction mixture. No *t*-CMP formation could be observed by LC-MS analysis in any of the above cases, neither in the presence of NADH nor NADPH. ThnN was later investigated by itself, as well as in the presence of ThnO, as described below.

ThnN Assay Development

Various assays were investigated in order to monitor the activity and substrate requirement of ThnN. ATP-utilising enzymes play an important role in biology and various strategies have been developed to monitor their activity. For safety reasons as well as convenience, assays involving the use of radiolabelled substrates were not pursued. Based on bioinformatic analyses, ThnN was anticipated to catalyse the conversion of ATP into AMP and pyrophosphate. The ATP hydrolysis reaction has previously been extensively monitored in the case of ANL superfamily members, using ^{31}P NMR [25] or HPLC monitoring [26].

^{31}P NMR spectrometry and HPLC techniques have different advantages and disadvantages and were therefore used in a complementary way. The broadness of ^{31}P NMR signals for phosphate groups is dependent on the pH as well as the magnesium concentration present in the assay [27],

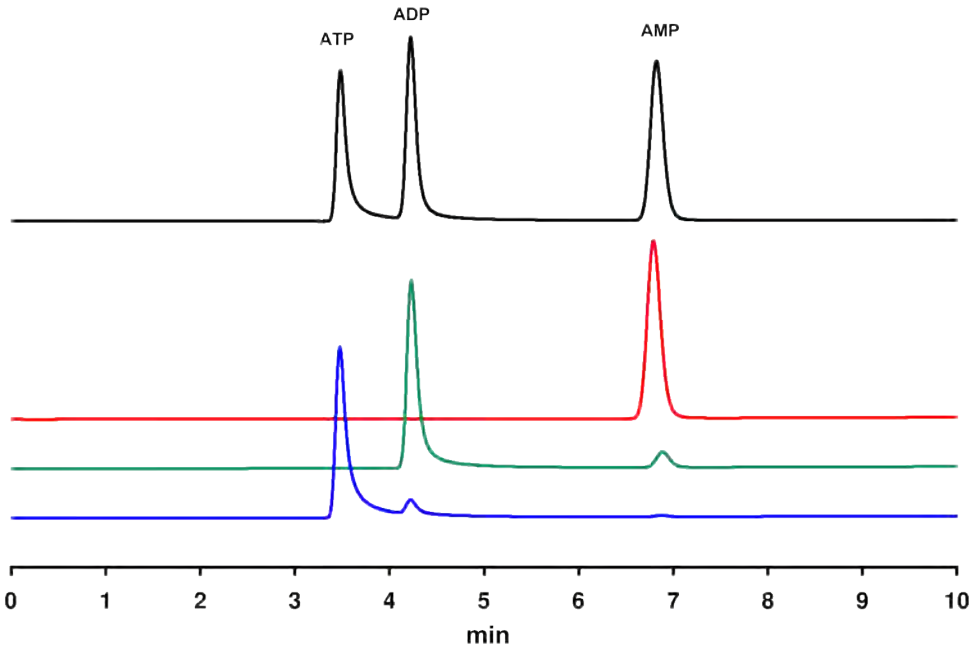


Figure 7.3: HPLC traces for adenosine nucleotides standards. Elution profiles for ATP (blue trace), ADP (green trace) and AMP (red trace) together with a stoichiometric mixture (black trace). The samples were injected on a Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μ particle size) and eluted at 1 mL min^{-1} using a 2–20% gradient of MeOH in 20 mM NH_4OOCCH_3 (pH 6.3) over 30 min.

and can vary substantially for different phosphorus atoms in the same molecule. Therefore, the integration values for ^{31}P NMR signals are not always reliable due to differences in relaxation times and inconsistent nuclear Overhauser effects [27]. This would make quantitative analysis of enzyme kinetics based on ^{31}P NMR rather difficult, but the advantages of ^{31}P NMR include a relatively good sensitivity and a relative simplicity of the spectrum. The 1/2 spin ^{31}P nucleus has a 100% natural abundance. HPLC analysis on the other hand allows the detection of very low levels of compounds in an assay. The observation of the UV absorbance of the adenosine nucleoside, coupled to HPLC separation is highly sensitive, but might not allow the observation of labile intermediates.

An analytical HPLC method for the detection of ThnN activity was developed by adapting a previously reported methodology used for the study of two adenylating enzymes, AuaE and AuaEII [26]. The HPLC separation of ATP, ADP, and AMP was optimised using commercially available standards until their retention times were different enough for identification (Figure 7.3).

In a similar manner, the ^{31}P NMR chemical shift for ATP (Figure 7.4 B), ADP (Figure 7.4 C), AMP (Figure 7.4 D), pyrophosphate (Figure 7.4 E) and phosphate (Figure 7.4 F) were recorded for commercially available standards under conditions identical to the assay. The level of spontaneous

hydrolysis of ATP and ADP was also measured under the same conditions in order to anticipate any non-enzymatic formation of products. ATP and ADP were found to be stable and no hydrolysis could be observed under these conditions for at least 2 days, and less than 10% hydrolysis was observed for ATP after 2 weeks at room temperature (data not shown). The non-enzymatic formation of the products was therefore assumed to be negligible.

Measuring Pantothenate Kinase and Phosphopantetheine Adenylyltransferase Activity

It is unclear whether ThnN catalyses the adenylation or the phosphorylation of its substrate. In the former case, the uncoupled reaction would produce AMP and pyrophosphate (PP_i) while the latter would produce ADP and phosphate (P_i). The pantothenate kinase (CoaA) and phosphopantetheine adenylyltransferase (CoaD) proteins were purified from *E. coli* in Chapter 5 and they were anticipated to catalyse the production of P_i and PP_i, respectively, in the presence of their respective substrates. CoaA and CoaD were therefore investigated as positive controls for ATP hydrolysis.

Pantothenate kinase (CoaA) was found to catalyse the rapid hydrolysis of ATP to ADP as judged by ^{31}P NMR and HPLC analyses; the reaction was coupled to the phosphorylation of D-pantothenate **246**. 65% turnover was observed by ^{31}P NMR after 15 minutes at room temperature (Figure 7.4 C). Complete turnover was obtained under the same conditions after 4 hours (Figure 7.4 D). In the presence of a slight excess of D-pantothenate **246**, no free phosphate could be observed suggesting a highly coupled mechanism. Very low levels of ATP hydrolysis activity could be observed by NMR in the absence of D-pantothenate **246** (Figure 7.4 B). The ATP conversion in the presence of CoaA alone was found to be lower than 10% after 24 hours at room temperature (data not shown). This is in accordance with the lower activity observed by HPLC analysis in the absence (blue trace, Figure 7.4 B) than in the presence of D-pantothenate **246** (red trace, Figure 7.4 B). No ATP hydrolysis was observed in the absence of enzyme (black trace, Figure 7.4 B).

Phosphopantetheine adenylyltransferase (CoaD) was expected to hydrolyse ATP into AMP and pyrophosphate in the presence of Mg^{2+} ions. Because phosphopantothenate **345** is not readily available, CoaD-catalysed ATP hydrolysis was monitored in the absence of the substrate. When

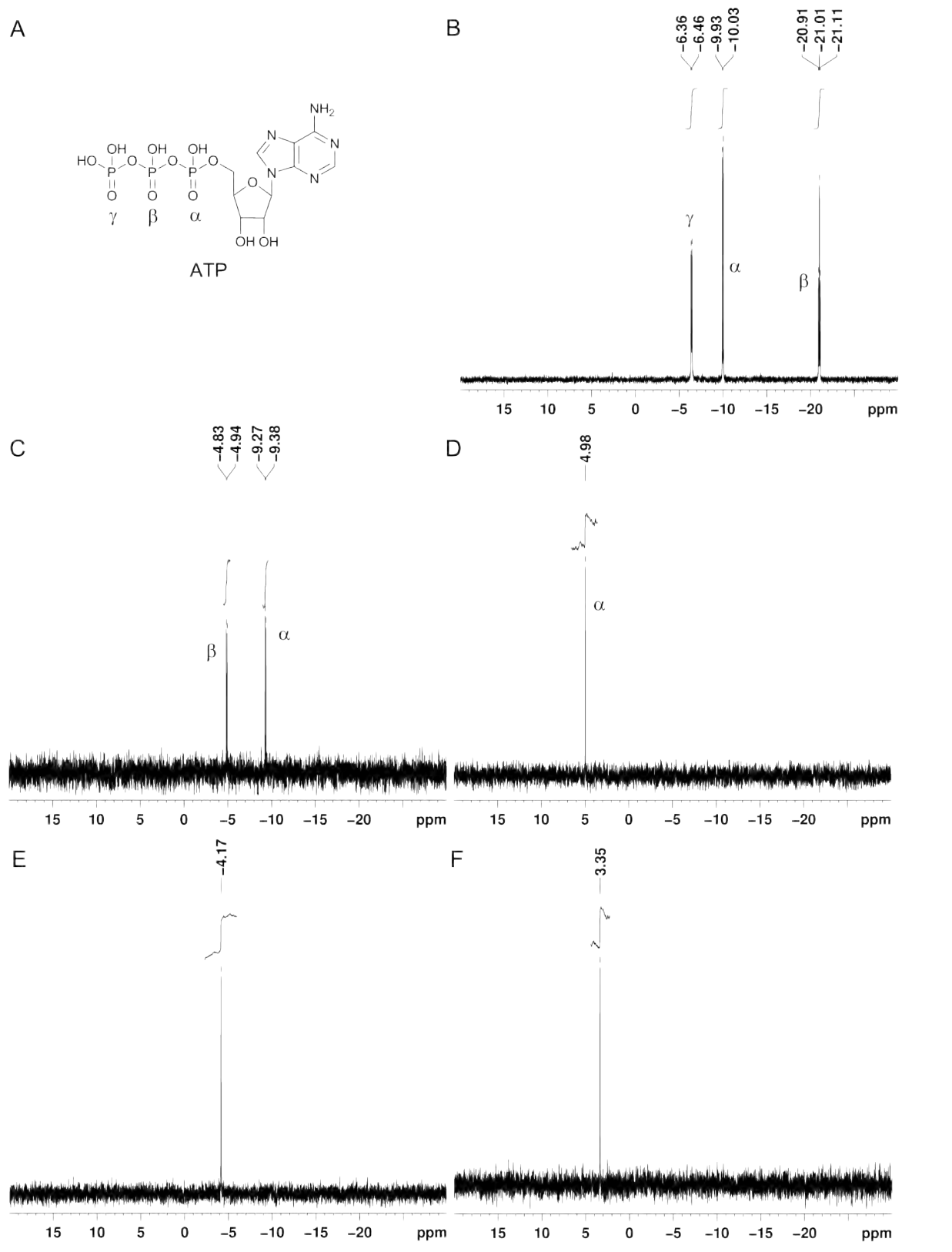


Figure 7.4: ^{31}P chemical shifts for standard compounds. Samples were prepared in 10% D_2O, 10 mM HEPES pH 8.3 and contained 5 mM of the corresponding phosphate, 10mM $MgCl_2$ and 25 mM NaCl. (A) Chemical structure of adenosine triphosphate (ATP) in its fully protonated state. (B) ^{31}P NMR spectrum for ATP (202 MHz, 10% D_2O): δ = - 6.41 (*d*, *J* = 20, P_γ), - 9.98 (*d*, *J* = 20, P_α), - 21.0 (*dd*, *J* = 20 x 2, P_β); (C) ^{31}P NMR spectrum for ADP (202 MHz, 10% D_2O): δ = - 4.89 (*d*, *J* = 22, P_α), - 9.33 (*d*, *J* = 22, P_β); (D) ^{31}P NMR spectrum for AMP (202 MHz, 10% D_2O): δ = 4.98 (*s*, P_α); (E) ^{31}P NMR spectrum for PP_i (202 MHz, 10% D_2O): δ = - 4.17 (*s*, 2P); (F) ^{31}P NMR spectrum for P_i (202 MHz, 10% D_2O): δ = 3.35 (*s*, 1P).

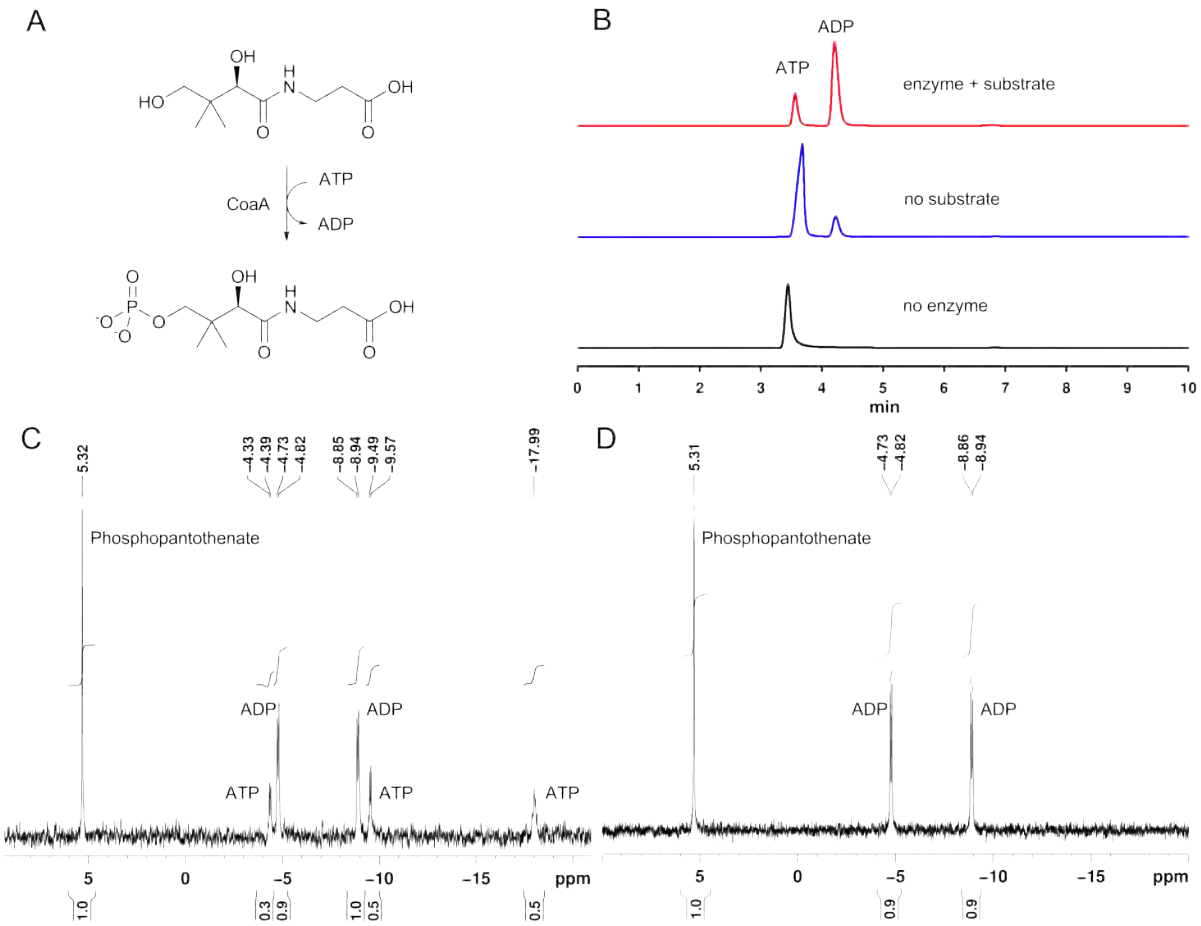
the assay was carried out at 37 °C and monitored by ^{31}P NMR or HPLC, no ATP consumption was observed, even after 24 hours, suggesting that pyrophosphate production only occurs when coupled to adenyl-5′-phosphate transfer activity.

Three out of Four Nucleoside Triphosphates are Hydrolysed by ThnN

Nine common nucleoside triphosphates are present in nature. They differ by the nature of the nucleobase moiety, i.e. the five ribonucleotides (ATP, GTP, m5UTP, UTP and CTP) and the four deoxyribonucleotides (dATP, dGTP, dTTP, dCTP). While most carboxylate activating enzymes use ATP as a cosubstrate, some have been demonstrated to have a preference for other ribonucleosides triphosphates, *e.g.* the *E. coli* phosphopanthothenoylcysteine synthetase domain of the protein encoded by the coaBC gene uses CTP rather than ATP as the carboxylate activating species [28]. In contrast, the human phosphopanthothenoylcysteine synthetase enzyme has no preference for CTP over ATP [29]. Four nucleoside triphosphates, i.e. ATP, CTP, GTP, and UTP were tested for ThnN-catalysed hydrolysis using ^{31}P NMR spectroscopy and HPLC analyses (Figure 7.5 A). In the absence of any amino acid substrate, ThnN was found to hydrolyse CTP (Figure 7.5 D) more efficiently than ATP (Figure 7.5 C). GTP was found to be only hydrolysed to a low extent and UTP seemed to be untouched by the enzyme (Figure 7.5 E and F). Although addition of the L-glutamate substrate did not change the extent to which ThnN hydrolyses ATP, it had an inhibitory effect on CTP hydrolysis (Figure 7.5 D).

ThnN Catalyses the ATP-Dependent Formation of L-Pyroglutamic Acid

Since the ATP hydrolysis reaction catalysed by CoaA was only found to proceed in the presence of the D-pantothenate substrate, several amino acid substrates were tested in the ThnN assay, i.e. L- and D-glutamic acid, *N*-acetyl-L-glutamic acid, L-pyroglutamic acid, L- and D-aspartic acid as well as L- and D-glutamine. The rate of ATP hydrolysis was found to be identical in the presence of any of these amino acids, as judged by ^{31}P NMR spectroscopy (data now shown) although there was a difference in the ^{1}H NMR spectrum. A signal corresponding to the α-proton of L-pyroglutamic acid (δ = 4.09 ppm, *dd*, *J* = 9.0, 6.0) appeared when L-glutamic acid was incubated with ThnN and ThnO



Scheme 7.5: CoaA activity monitoring using ^{31}P NMR spectroscopy and HPLC. CoaA catalyses the magnesium-dependent transfer of a phosphate group from ATP onto the terminal hydroxyl group of D-pantothenate **246**. ^{31}P NMR samples were prepared in 10% D_2O, 10 mM HEPES pH 8.3 and contained 5 mM ATP, 5 mM D-pantothenate, 10 mM $MgCl_2$ and 25 mM NaCl and were recorded at 202 MHz (ns = 512). HPLC samples were prepared in 100 mM HEPES pH 8.3 and contained 1 mM ATP, 5 mM D-pantothenate, 5 mM $MgCl_2$ and 25 mM NaCl. The samples were injected on a Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μ particle size) after removal of the proteins and the absorbance was recorded at λ = 254 nm. (A) Scheme for the reaction catalysed by CoaA; (B) HPLC traces for the CoaA catalysed hydrolysis of ATP in the presence (red trace) and in the absence (blue trace) of the D-pantothenate substrate. The control with no enzyme is shown as a reference (black trace). The reaction was left at 25 °C for 30 min. The samples were injected on a Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μ particle size) after removal of the protein.; (C) ^{31}P NMR spectrum for the CoaA catalysed hydrolysis of ATP in the presence of D-pantothenate and after 15 min at 25 °C; (D) ^{31}P NMR spectrum for the CoaA catalysed hydrolysis of ATP in the presence of D-pantothenate and after 4h at 25 °C.

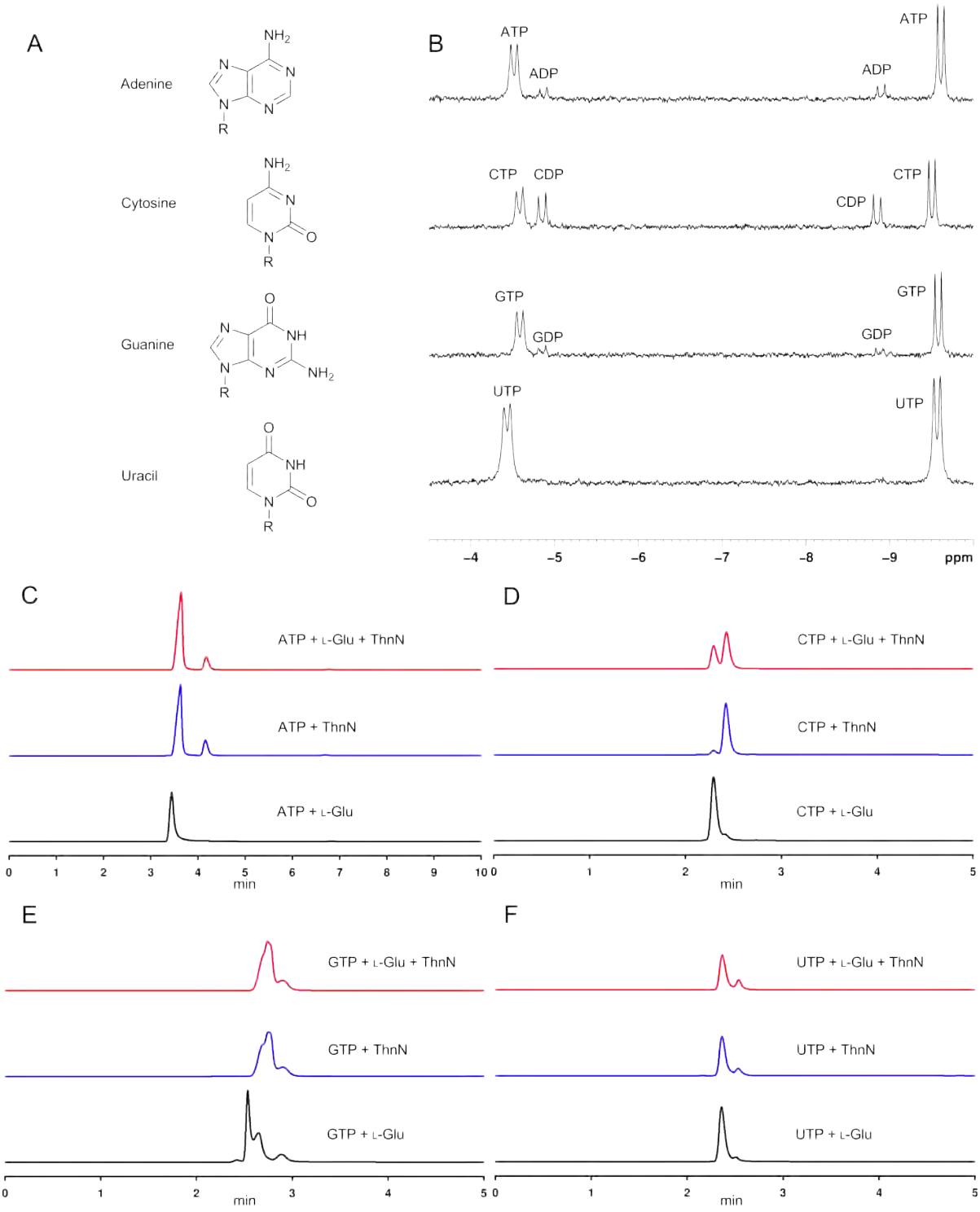
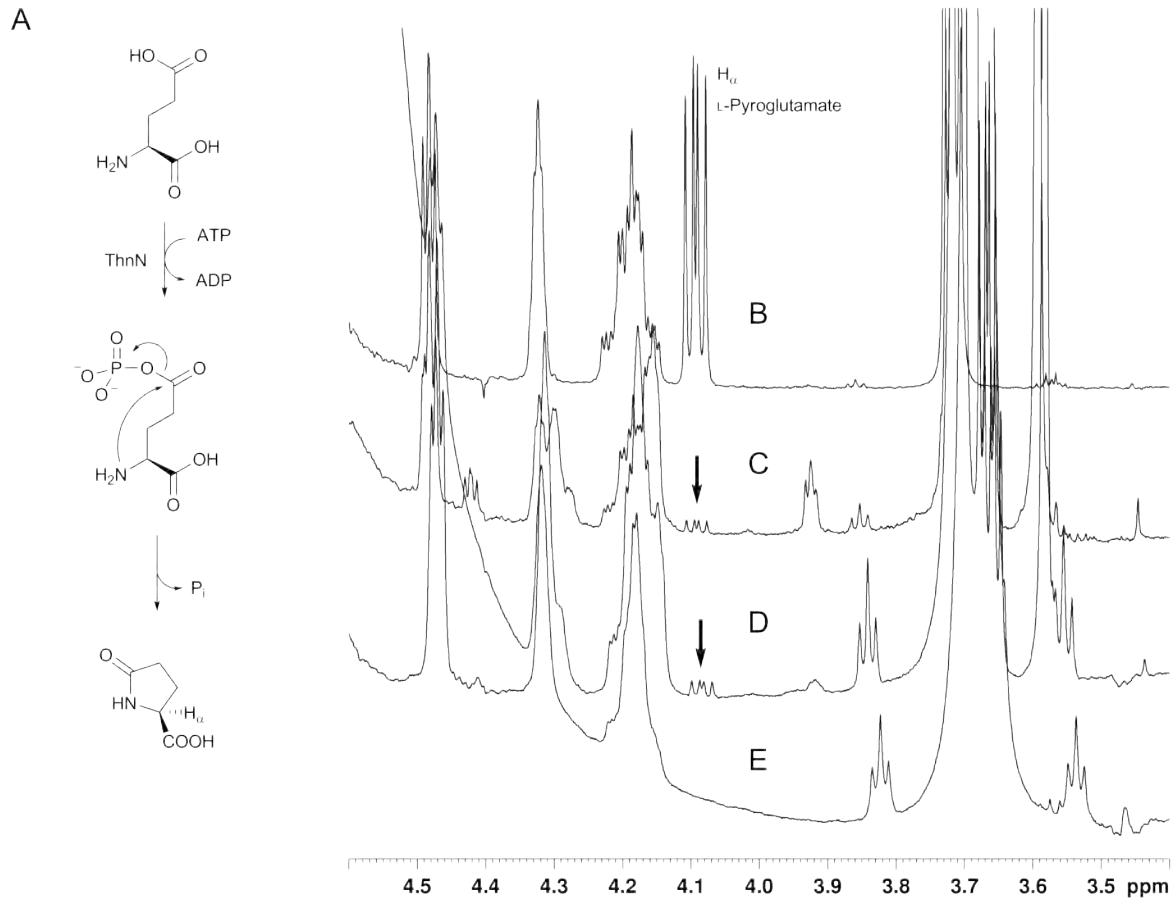


Figure 7.5: The four nucleoside triphosphates shown were tested as ThnN substrates using ^{31}P NMR spectroscopy and HPLC-based assays. Reactions were incubated at 25 °C for 4 h. ^{31}P NMR samples were recorded at 202 MHz (ns = 512). HPLC analysis was performed using a Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μ particle size) after removal of the proteins and the absorbance was recorded at λ = 254 nm. (A) Chemical structures for the nucleobases of the four nucleosides triphosphates tested as ThnN substrates. R = ribose-5′-triphosphate; (B) Expansions of the ^{31}P NMR spectra for the 4 h reaction between ThnN and the four different nucleoside triphosphates; HPLC traces for the assay in the presence of ATP (C), CTP (D), GTP (E) or UTP (F).



Scheme 7.6: ^{1}H NMR spectra expansions showing the formation of L-pyroglutamate. Arrows indicate the characteristic signal corresponding to the H_α proton of L-pyroglutamate (δ = 4.09 ppm, *dd*, *J* = 9.0, 6.0). The reaction was incubated at 25 °C for 16 h. All samples were prepared in 10% D_2O, 10 mM HEPES (pH 8.3) and contained 5 mM ATP, 10mM $MgCl_2$ and 25 mM NaCl. ^{1}H NMR spectra were recorded at 500 MHz (ns = 1024). (A) Proposed mechanism for the ThnN catalysed formation of L-pyroglutamate in the presence of L-glutamate and ATP; (B) No enzymes, 5 mM L-pyroglutamate; (C) 10 μM ThnN, 3 μM ThnO, 5 mM L-glutamate; (D) 10 μM ThnN, no ThnO, 5 mM L-glutamate; (E) 10 μM ThnN, 3 μM ThnO, no L-glutamate.

in the presence of ATP (Scheme 7.6 C and D). This signal was absent when D-glutamic acid or *N*-acetyl-L-glutamic acid were used as substrates (data not shown) suggesting that only L-glutamic acid is a ThnN substrate.

Evidence for Cooperative Activity between ThnN and ThnO

In order to investigate a possible effect of ThnO on ThnN-catalysed ribonucleoside triphosphate hydrolysis, all the experiments described above were repeated in the presence of ThnO. It is possible that the two proteins form a complex in solution, allowing substrate channeling between the ThnN

and ThnO active sites. This would allow a labile phospho-L-glutamate, or adenosyl-L-glutamate intermediate to be reduced by ThnO.

As expected, and whether ThnO was added or not to the assay, there was no observable nucleoside triphosphate hydrolysis activity for ThnN in the absence of magnesium ions. This was also true for the CoaA catalysed ATP hydrolysis. In the absence of L-glutamate substrate, the addition of ThnO to the ThnN assay was found to have a positive effect on the extent of ATP hydrolysis (Figure 7.6 D), suggesting that the presence of ThnO increases ThnN activity. However, this hypothesis is questioned by the observation that ThnO alone has ATP hydrolysis activity (Figure 7.6 B). However, it is possible that this activity is due to the contaminating protein. The protein impurity present in the ThnO preparations was not identified (Figure 7.2 D) and ATP hydrolysis activity associated with this impurity cannot be excluded.

Another observation was made when L-glutamate was added to the assay containing both ThnN and ThnO. An additional ^{31}P NMR signal (δ = 4.4 ppm, *s*) appeared (Figure 7.6 A), and even though AMP (δ = 5.0 ppm, *s*) and pyrophosphate (δ = - 4.2 ppm, *s*) could be excluded by comparison with standard chemical shifts (Figure 7.6 D and E, respectively), the new species could not be identified. No additional species could be observed by HPLC suggesting that it is either labile or that it does not contain any chromophore, supporting the hypothesis that it could be L-glutamyl-5-phosphate.

Coenzyme A Does not Seem to Play a Role in ThnN Catalysis

In the light of some other families of carboxylate reductase enzymes, *e.g.* luciferase enzymes [30], the possible role of coenzyme A **133** in the formation of an acyl-coenzyme A intermediate was investigated using the HPLC Assay. When CoASH **133** was added in stoichiometric amount to the assay containing ThnN, ThnO, and the L-glutamate substrate, no new species and no difference in the ATP hydrolysis rate could be observed (data not shown).

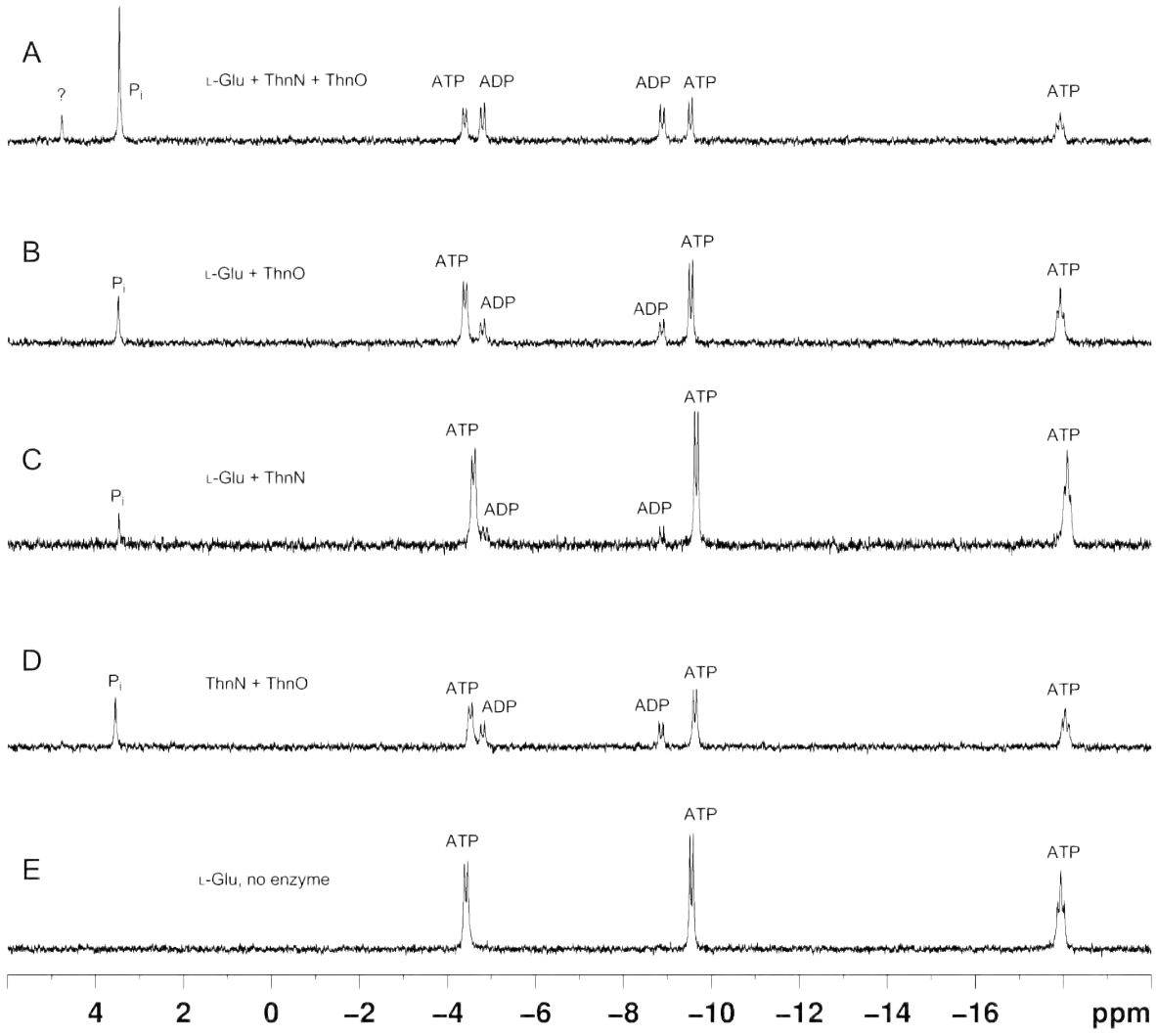


Figure 7.6: ^{31}P NMR spectra showing ThnN activity in conjugation with ThnO. The reaction was performed at 25 °C for 16 h. All samples were prepared in 10% D_2O, 10 mM HEPES (pH 8.3) and contained 5 mM ATP, 10mM $MgCl_2$ and 25 mM NaCl. ^{31}P NMR spectra were recorded at 202 MHz (ns = 512). The question mark indicates an unindentified species. (A) 10 μM ThnN, 3 μM ThnO, 5 mM L-glutamate; (B) No ThnN, 3 μM ThnO, 5 mM L-glutamate; (C) 10 μM ThnN, no ThnO, 5 mM L-glutamate; (D) 10 μM ThnN, 3 μM ThnO, no L-glutamate; (E) No enzymes, 5 mM L-glutamate.

7.3 Summary for Chapter 7

The work described in Chapter 7 is an attempt to characterise ThnN and ThnO, two enzymes proposed to be responsible for the formation of L-GHP **129** in *S. cattleya* in the context of thienamycin biosynthesis. Based on sequence similarities, ThnN and ThnO were proposed to act as an enzyme-pair and together carry the carboxylate reductase activity required to produce L-GHP **129** from L-glutamic acid **124**. Although these preliminary results do not allow the drawing of any conclusion regarding the activity of ThnN and ThnO and their role in the thienamycin biosynthesis pathway, some observations are worth noting; the recombinant expression of the *thnN* and *thnO* genes was carried out in *E. coli* but the corresponding protein expression levels found to be rather low, possibly due to their high GC content. *Streptomyces* species are known to carry a genome containing an unusually high proportion of guanine and cytosine bases compared to other bacteria who typically have a GC content of about 50%. ThnN displays 71% GC and ThnO 76% GC, giving an extremely high codon bias. However, using *E. coli* strain containing a plasmid carrying additional genes coding for rare codons tRNAs did not improve the protein expression. ThnO was found to precipitate when concentrated, probably because of a reduced stability under the purification conditions.

In an initial attempt to detect the enzymatic production of L-GHP **129**, ThnN and ThnO were incubated in the presence of their anticipated cofactors and cosubstrates, and their activity coupled to a CMPS assay. None of the conditions tested lead to detectable levels of *t*-CMP. After developing ^{31}P NMR and HPLC methods to identify the different possible products of the ThnN-catalysed utilisation of ATP, ThnN alone or in combination with ThnO was incubated in the presence of magnesium ions, ATP and various amino acids including L-glutamate. Although slowly, ATP was found to be consumed and its hydrolysis coupled to ADP and phosphate production. This observation pointed towards a phosphorylated intermediate rather than an adenylated one, potentially in contradiction with the sequence-based prediction [13]. When ^{1}H NMR spectra were recorded for the assay in the presence of ThnN, ThnO, magnesium ions and ATP, a new signal corresponding to the L-pyroglutamate α-proton was observed. The observation that L-pyroglutamate is formed is of interest because it indicates that L-glutamic acid is activated by ThnN, probably through the phosphorylation of the 5-carboxylate, followed by cyclisation (Figure 7.6). It also suggests that the production of L-GHP involves two closely coupled reactions to avoid the formation of the L-pyroglutamate shunt product. There is a parallel in human pathology. The inherited disease cystinosis is characterised

by a state of ATP depletion and cystine accumulation. This condition was proposed to be due to the production of L-pyroglutamate from L-glutamate by an ATP-dependent enzyme of the glutathione biosynthesis pathway, γ-glutamyl cysteine synthase (γ-GCS), in the absence of the cosubstrate L-cysteine [31]. A second equivalent of ATP is then consumed by the enzyme 5-oxoprolinase which converts L-pyroglutamate (also called 5-oxoproline) back to L-glutamate.

The presence of an unidentified species in ^{31}P NMR-based assay in the presence of ThnN and ThnO should be further investigated. LC-MS analysis of the assay mixture could for example lead to the identification of this species and confirm the identity of the L-pyroglutamate species observed by ^{1}H NMR. In addition, oxygen transfer studies using γ-[^{18}O]-L-glutamic acid could potentially rule out the involvement of an acyl-adenylate intermediate and confirm the formation of γ-glutamylphosphate.

References

[1] Tanner, J. J. (2008) Structural biology of proline catabolism. *Amino Acids*, **35**, 719–730.

[2] Singh, R. K., Larson, J. D., Zhu, W., Rambo, R. P., Hura, G. L., Becker, D. F., and Tanner, J. J. (2011) Small-angle X-ray scattering studies of the oligomeric state and quaternary structure of the trifunctional proline utilization A (putA) flavoprotein from *Escherichia coli*. *J Biol Chem*.

[3] Zhou, Y., Zhu, W., Bellur, P. S., Rewinkel, D., and Becker, D. F. (2008) Direct linking of metabolism and gene expression in the proline utilization A protein from Escherichia coli. *Amino Acids*, **35**, 711–718.

[4] Pérez-Arellano, I., Carmona-Alvarez, F., Martínez, A. I., Rodríguez-Díaz, J., and Cervera, J. (2010) Pyrroline-5-carboxylate synthase and proline biosynthesis: From osmotolerance to rare metabolic disease. *Protein Sci*, **19**, 372–382.

[5] Viola, R. E. (2001) The central enzymes of the aspartate family of amino acid biosynthesis. *Acc Chem Res*, **34**, 339–349.

[6] Sorensen, J. L., Sleeman, M. C., and Schofield, C. J. (2005) Synthesis of deuterium labelled L- and D-glutamate semialdehydes and their evaluation as substrates for carboxymethylproline synthase (CarB)–implications for carbapenem biosynthesis. *Chem Commun*, **9**, 1155–1157.

[7] Coque, J. J., Liras, P., Laiz, L., and Martín, J. F. (1991) A gene encoding lysine 6-aminotransferase, which forms the β-lactam precursor α-aminoadipic acid, is located in the cluster of cephamycin biosynthetic genes in *Nocardia lactamdurans*. *J Bacteriol*, **173**, 6258–6264.

[8] Tobin, M. B., Kovacevic, S., Madduri, K., Hoskins, J. A., Skatrud, P. L., Vining, L. C., Stuttard, C., and Miller, J. R. (1991) Localization of the lysine ϵ-aminotransferase (*lat*) and δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase (*pcbAB*) genes from *Streptomyces clavuligerus* and production of lysine ϵ-aminotransferase activity in *Escherichia coli*. *J Bacteriol*, **173**, 6223–6229.

[9] Hamed, R. B., Mecinović, J., Ducho, C., Claridge, T. D. W., and Schofield, C. J. (2010) Carboxymethylproline synthase catalysed syntheses of functionalised *N*-heterocycles. *Chem Commun*, **46**, 1413–1415.

[10] Núnez, L. E., Méndez, C., Brana, A. F., Blanco, G., and Salas, J. A. (2003) The biosynthetic gene cluster for the β-lactam carbapenem thienamycin in *Streptomyces cattleya*. *Chem Biol*, **10**, 301–311.

[11] Blasiak, L. C. and Clardy, J. (2010) Discovery of 3-formyl-tyrosine metabolites from *Pseudoalteromonas tunicata* through heterologous expression. *J Am Chem Soc*, **132**, 926–927.

[12] Giddens, S. R., Feng, Y., and Mahanty, H. K. (2002) Characterization of a novel phenazine antibiotic gene cluster in *Erwinia herbicola* Eh1087. *Mol Microbiol*, **45**, 769–783.

[13] Suzuki, H., Ohnishi, Y., and Horinouchi, S. (2007) GriC and GriD constitute a carboxylic acid reductase involved in grixazone biosynthesis in *Streptomyces griseus*. *J Antibiot (Tokyo)*, **60**, 380–387.

[14] Gulick, A. M. (2009) Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem Biol*, **4**, 811–827.

[15] Gross, G. G. (1972) Formation and reduction of intermediate acyladenylate by aryl-aldehyde. *Eur J Biochem*, **31**, 585–592.

[16] Li, T. and Rosazza, J. P. (1997) Purification, characterization, and properties of an aryl aldehyde oxidoreductase from *Nocardia* sp. strain NRRL 5646. *J Bacteriol*, **179**, 3482–3487.

[17] He, A., Li, T., Daniels, L., Fotheringham, I., and Rosazza, J. P. N. (2004) *Nocardia* sp. carboxylic acid reductase: Cloning, expression, and characterization of a new aldehyde oxidoreductase family. *Appl Environ Microbiol*, **70**, 1874–1881.

[18] Venkitasubramanian, P., Daniels, L., and Rosazza, J. P. N. (2007) Reduction of carboxylic acids by *Nocardia* aldehyde oxidoreductase requires a phosphopantetheinylated enzyme. *J Biol Chem*, **282**, 478–485.

[19] Li, T. and Rosazza, J. P. (1998) NMR identification of an acyl-adenylate intermediate in the aryl-aldehyde oxidoreductase catalyzed reaction. *J Biol Chem*, **273**, 34230–34233.

[20] Engebrecht, J. and Silverman, M. (1984) Identification of genes and gene products necessary for bacterial bioluminescence. *Proc Natl Acad Sci U S A*, **81**, 4154–4158.

[21] Marques, S. M. and da Silva, J. C. G. E. (2009) Firefly bioluminescence: A mechanistic approach of luciferase catalyzed reactions. *IUBMB Life*, **61**, 6–17.

[22] Soly, R. R. and Meighen, E. A. (1991) Identification of the acyl transfer site of fatty acyl-protein synthetase from bioluminescent bacteria. *J Mol Biol*, **219**, 69–77.

[23] Lee, C. Y. and Meighen, E. A. (1997) Cysteine-286 as the site of acylation of the Lux-specific fatty acyl-CoA reductase. *Biochim Biophys Acta*, **1338**, 215–222.

[24] Geoghegan, K. F., et al. (1999) Spontaneous α-*N*-6-phosphogluconoylation of a His tag in *Escherichia coli*: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal Biochem*, **267**, 169–184.

[25] Song, Z., Parker, K. J., Enoh, I., Zhao, H., and Olubajo, O. (2009) Myosin-catalyzed ATP hydrolysis elucidated by ^{31}P NMR kinetic studies and ^{1}H PFG-diffusion measurements. *Anal Bioanal Chem*, **395**, 1453–1459.

[26] Pistorius, D., Li, Y., Mann, S., and Mller, R. (2011) Unprecedented anthranilate priming involving two enzymes of the acyl adenylating superfamily in aurachin biosynthesis. *J Am Chem Soc*, **133**, 12362–12365.

[27] Burt, C. T., Cheng, H. M., Gabel, S., and London, R. E. (1990) ATP, ADP, and magnesium mixed solutions: In vitro ^{31}P NMR characterization and *in vivo* application to lens. *J Biochem*, **108**, 441–448.

[28] Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P. (2001) Phosphopantothenoylcysteine synthetase from *Escherichia coli*. Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J Biol Chem*, **276**, 13513–13516.

[29] Yao, J. and Dotson, G. D. (2009) Kinetic characterization of human phosphopantothenoylcysteine synthetase. *Biochim Biophys Acta*, **1794**, 1743–1750.

[30] Fraga, H., da Silva, J. C. G. E., and Fontes, R. (2004) Identification of luciferyl adenylate and luciferyl coenzyme a synthesized by firefly luciferase. *ChemBioChem*, **5**, 110–115.

[31] Kumar, A. and Bachhawat, A. K. (2010) A futile cycle, formed between two ATP-dependant γ-glutamyl cycle enzymes, γ-glutamyl cysteine synthetase and 5-oxoprolinase: The cause of cellular ATP depletion in nephrotic cystinosis? *J Biosci*, **35**, 21–25.

[32] Gottlieb, H. E., Kotlyar, V., and Nudelman, A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *J Org Chem*, **62**, 7512–7515.

7.4 Materials and Methods

7.4.1 General Material and Methods for Assays

Unless otherwise stated, chemicals were from Aldrich and used without further purification. NMR tubes (5 mm) were obtained from Aldrich. Deuterated solvents were obtained from Aldrich and Apollo Scientific Ltd. All ^{1}H, ^{31}P and ^{13}C NMR Spectra were recorded using a Bruker AV 500 MHz spectrometer with ^{13}C cryoprobe. All chemical shifts are given in ppm relative to the solvent peak [32] or relative to an internal standard. Coupling constants *J* are given in Hz to the nearest 0.5 Hz value. Analytic HPLC was performed on a 1290 Infinity™ system (Agilent Technologies) using various columns and mobile phases as described in each case.

7.4.2 General Material and Methods for Molecular Biology

All the molecular biology protocols were carried out according to standard procedures in place in the group of Prof Christopher J. Schofield at the University of Oxford. The general material and methods for molecular biology were essentially identical to those described for Chapter 4.

7.4.3 ThnN and ThnO Cloning

The *thnN* and *thnO* genes were amplified from a pET-24a(+) plasmid containing the *thnON* fragment between NdeI and BamHI restriction site (generously provided by Dr Refaat B. Hamed) using the PCR protocol described in Chapter 4 Materials and Methods (Section 4.5.1). The following oligonucleotides were used for *thnN* amplification with the introduction of a 5' XhoI restriction site and a 3' BamHI restriction site: forward primer sequence (33 mer): `AAAAcatatgCCAAGTCCGTCCAAGCTCCGTCC`; reverse primer sequence (30 mer): `TTTTggatccTCAGTAGACCCCCTCGATCC`. The following oligonucleotides were used for *thnO* amplification with the introduction of a 5' XhoI restriction site and a 3' NdeI restriction site: forward primer sequence (33 mer): `AAAActcgagTCACCCGGCCGCGCCCGGG`; reverse primer sequence (33 mer): `TTTTcatatgACGACCGTCACCACCGGACAC`.

The PCR amplicons were restricted with the corresponding enzymes for 90 min at 37 °C. After vector amplification in *E. coli* BL21 DE3 cells and purification using a miniprep kit (QIAgen), a pET-28a(+) plasmid (1.3 μg) was also restricted with the corresponding enzymes for 90 min at 37 °C. This restricted vector was separated from unrestricted sample and unwanted fragments on agarose gel electrophoresis and purified using a gel extraction kit (QIAgen). The digested PCR fragments were purified using a PCR clean kit (QIAgen) and ligated (16 ng) into the digested pET-28a(+) plasmid (10 ng) using T4 ligase for 2 h at 37 °C. The ligation mix was transformed into *E. coli* XL10 gold cells following the transformation protocol described in Section 4.5.1, and the positive clones were grown in 6 mL 2TY medium for plasmid purification. The correct insertion of the gene in the positive clones was verified by sequencing.

7.4.4 ThnN and ThnO Expression and Purification

Protein Expression

Large scale protein expression cultures were carried out in 2 L conical flasks containing 600 mL of 2TY growth medium. The pET-28a(+) plasmids containing either the *thnN* or *thnO* gene were transformed into *E. coli* Rosetta 2 (DE3) cells. The antibiotic corresponding to the resistance gene present in the expression plasmid was added to a final concentration of 3.0–5.0 mg L^{-1} from a 3.0–5.0 mg mL^{-1} sterile stock in water (ampicillin) or ethanol (kanamycin). The flasks were inoculated with

6 mL of overnight culture and incubated at 37°C in a shaker (180 rpm) until the end of the linear growth curve was reached, as judged by optical density ($\cong$ 0.6). The protein expression was induced by addition of IPTG to the optimal final concentration from a sterile 1M stock in water. Expression trials showed that the best expression conditions for both constructs were when cells were grown at 37 °C and then left for 18 hours at 18 °C after induction using 0.5 mM IPTG final concentration. The cultures were pelleted by centrifugation on an Avanti instrument using a JA-10 rotor at 8 krpm for 5 min and the pellets stored at -80 °C until lysis and protein purification. Typically, each flask would produce 4–6 g of bacterial pellet (6–10 g L^{-1}).

Protein Purification

Proteins were purified using chromatographic methods, including affinity and size exclusion (gel filtration) chromatography as described below. All protein purification protocols were carried out at 4°C using Äkta FPLC™ systems (GE Healthcare). Buffers were freshly prepared using Milli-Q water and filtered through a 0.2 μm polyamide filter (Millipore U.K. Ltd.) before use. Purification procedures were followed by monitoring absorbance at λ = 280 nm with the UPC-900 monitor on the FPLC and by SDS-PAGE. All columns were cleaned after each use according to the manufacturers' instructions. All columns were stored in 20% ethanol.

Nickel affinity chromatography

When the protein of interest contained a hexahistidine tag, a nickel affinity chromatography was performed. The following buffer were used:

Ni Lysis buffer	
Tris	50 mM
Imidazole	10 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Wash buffer	
Tris	50 mM
Imidazole	40 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Elution buffer	
Tris	50 mM
Imidazole	500 mM
Sodium chloride	500 mM
β-mercaptoethanol	1 mM

Ni Strip buffer	
Tris	20 mM
Sodium chloride	100 mM
EDTA	100 mM

The one-step purification of the recombinant ThnN and ThnO proteins was carried out as follow. The filtered cell lysate was loaded at a 1 mL min^{-1} flow rate using the FPLC pump on a 5 mL HisTrap™ FF column (GE healthcare) previously charged with with nickel ions and, after washing the unbound proteins with buffer, the tagged protein was eluted with a 60–600 mM imidazole gradient

over 10 column volumes. Analysis of the flow through and wash fractions indicated a fully efficient binding of the target protein to the resin. The fractions containing eluted proteins, as judged by monitoring the UV absorbance, were screened by SDS-PAGE electrophoresis. Fractions containing the proteins of interest were pooled together, mixed with 20 mM Tris-HCL pH 7.5 containing 300 mM NaCl and 100 mM EDTA to reach a final concentration of 1 mM EDTA and incubated at 4°C for 1 h to chelate nickel ions leached off the column. The EDTA and the imidazole were removed from the protein samples by overnight dialysis at 4 °C against 50 mM HEPES buffer pH 8.0 containing 250 mM NaCl, 1 mM DTT and 10% glycerol.

IMAC purification was performed as follows. The lysate was incubated at 4 °C in 1 mL nickel affinity beads (Biorad) for 1 h. The beads were then washed with wash buffer and a stepwise elution was performed with 10 mL of 50 mM Tris-HCl (pH 7.5) buffer containing 100 mM NaCl, 1 mM DTT and 10% glycerol and a medium (250 mM), followed by a high (500 mM) concentration of imidazole.

Size Exclusion Chromatography

Size exclusion chromatography, also commonly called gel filtration chromatography, was used to further purify the proteins after previous chromatography steps. To prevent precipitation, 1 mM DTT and 10% glycerol were added to all buffers. All protein samples ($\leq$ 2 mL) were applied on a 300 mL Superdex^{TM} S200 column, equilibrated with 100 mM Tris-HCl (pH 7.5) containing 100 mM NaCl. The proteins were eluted using 1 column volume of 100 mM Tris-HCl (pH 7.5) containing 100 mM NaCl. The fractions containing the proteins of interest and with purity $\geq$ 95%, as judged by SDS-PAGE analysis, were pooled and concentrated as described below.

7.4.5 ThnN and ThnO MS Analyses

Protein Electrospray Ionisation Mass Spectrometry

Intact protein masses were recorded by liquid chromatography/mass spectrometry (LC/MS) using a ZMD (Waters®) single quadrupole mass spectrometer, interfaced with a Hewlett Packard® Series 1050 liquid chromatography and sample handling system. The protein sample (30 μM in 50 mM Tris pH 7.5 M, 5 l) was injected onto a Dionex Proswift RP-4H 1 x 50 mm Monolithic reverse phase HPLC column and eluted at 0.4 ml min^{-1} using a gradient system from Solvent A (95% water, 4.9% acetonitrile, 0.1% (v/v) formic acid) to Solvent B (95% acetonitrile, 4.9% water, 0.1% (v/v) formic acid) according to the following conditions:

Time (min)	Solvent A (%)	Solvent A (%)	Flow Rate (ml min^{-1})
0:00	100	0	0.4
1:00	100	0	0.4
5:00	0	100	0.4
7:00	0	100	0.4
7:50	100	0	0.4
10:00	100	0	0.4

Table 7.1: HPLC method for LC/MS protein analysis.

The eluant was injected directly into the mass spectrometer. The following MS parameters were used: Capillary voltage, 3,000 V; Sample cone voltage, 35 V; Extraction cone voltage, 1 V; Desolvation temperature, 250 °CC; Source temperature, 100 °CC; Cone gas flow, 100 L h^{-1}; Desolvation gas flow (N_2), 830 L h^{-1}. Sodium formate was used to calibrate the instrument. Spectra

were processed using the MassLynx^{TM} v4.0 software (Waters Corporation) using the Maximum Entropy method (MaxEnt1).

Protein in-gel digestion

Protein samples were separated by 12.5% SDS-PAGE and isolated gel bands were prepared for MS analysis as follow. The gel bands stained with Coomassie Blue were cut into small pieces (ca. 1 mm^3). Destaining solution (50% methanol, 5% acetic acid) was added, and the solution was shaken (350 rpm) for 3 h, the destaining solution was then removed, 200 μl of the same solution added and left overnight under the same conditions. The de-staining solution was then removed, 200 μl of acetonitrile was added and the solution left for 5-15 min (until the gel pieces had shrunk and whitened). The acetonitrile was then pipetted off and the gel pieces dried by centrifugation in vacuo for 5 min. Subsequently, reduction was carried out with DTT (10 mM, 30 μl), followed by alkylation with iodoacetamide (100 mM, 30 μl). Trypsin digestion for MS analysis was then carried out whereby the excised gel pieces were subjected twice to hydration and dehydration using de-staining solution and acetonitrile, respectively, followed by overnight incubation at 37 °C with trypsin (20 ng/μl) in 100 mM ammonium hydrogen carbonate (pH 7.8) digestion buffer. The final samples were redissolved in 0.1% formic acid, 2% acetonitrile and stored at -20 °C until analysis.

MS/MS Protein analysis

Protein analysis by MS/MS was performed by Dr Mukram M. Mackeen at the Henry Wellcome Building for Molecular Physiology, Department of Clinical Medicine, University of Oxford. LC-MS/MS analysis of the digested material was carried out by nano-ultra performance liquid chromatography tandem MS (nano-UPLC-MS/MS) using a 75 μm-inner diameter x 25 cm C18 nanoAcquity ^{TM} UPLC ^{TM} column (1.7 μm particle size, Waters) with a 45 min gradient of 2–40% solvent B (solvent A: 99.9% H_2O, 0.1% HCOOH; solvent B: 99.9% MeCN, 0.1% HCOOH). The Waters nanoAcquity UPLC system (final flow rate, 250 nl min^{-1}; ca. 7000 p.s.i.) was coupled to a Q-TOF Premier tandem mass spectrometer (Waters) run in positive ion mode. MS analysis was performed in data-directed analysis (DDA) mode with MS to MS/MS switching at precursor ion counts greater than 10 with a return from MS/MS to MS survey after 1s (MS/MS collision energy is dependent on precursor ion mass and charge state). All raw MS data were processed to generate pkl files using the PLGS version 2.3 (Waters) software with deisotoping and deconvolution (converting masses with multiple charge states to m/z = 1). The mass accuracy of the raw data was corrected using Glu-fibrinopeptide for the Waters. MS/MS spectra of the digested biological samples were searched against the NCBInr database (nr/nr. v2011.03.27; 13,366,630 bacteria (eubacteria) sequences) using Mascot version 2.3.01 (Matrix Science) with the following parameters: peptide tolerance, 50 ppm; 13C = 1; fragment tolerance, 0.1 Da; missed cleavages, 3; instrument type, ESI-Q-TOF-IMM; fixed modification, carbamidomethylation (C); and variable modifications: deamidation (Asp, Glu) and methionine oxidation. A local *in-house* Mascot server used for this study is supported and maintained by the Computational Biology Research Group at the University of Oxford.

7.4.6 ThnN, ThnO and CMPS Coupled Assays

Assay performed with ThnN and ThnO preincubation

All assays were performed in 100 mM HEPES buffer at a final pH of 7.9 to a final volume of 125 μL. A 100 mM solution of L-glutamic acid (5 μL) was mixed with a 200 mM solution of $MgCl_2$ (5 μL), 100 mM ATP (5 μL), 10 mM DTT (1 μL) and 100 mM NAD(P)H reduced (5 μL) in 200 mM HEPES pH 8.3 (62 μL). ThnN (5 μL) and ThnO (35 μL) were added to a final concentration of 8 μM.

The assay mixture was kept at 37 °C for 1 h and malonyl-CoA (80 nM, 8 μL of 10 mM solution) and CMPS (0.3 nmol, 3 μL of 1.5 mM solution) were added. After 1 h at 37 °C the assay was quenched with a 250 μM *p*-aminosalicylic acid solution in acetonitrile (120 μL). The samples were kept on ice for 15 min, and then centrifuged at 13000 rpm for 15 min at 4 °C to remove the precipitated protein. The supernatant was transfered in a vial for analysis.

Assay performed without preincubation

The assays performed without preincubation were essentially carried out as above except malonyl-CoA (80 nM, 8 μL of 10 mM solution) and CMPS (0.3 nmol, 3 μL of 1.5 mM solution) were added immediately. After 1 h at 37 °C the assay was quenched with a 250 μM *p*-aminosalicylic acid solution in acetonitrile (120 μL). The samples were kept on ice for 15 min, and then centrifuged at 13000 rpm for 15 min at 4 °C to remove the precipitated protein. The supernatant was transfered in a vial for analysis.

ThnN/ThnO/CMPS coupled Assay LC-MS Analysis

The formation of *t*-CMP **131** was monitored as in Section 4.5.3. The resulting chromatograms were analysed using the MassLynx™ v4.0 software.

7.4.7 ThnN and ThnO HPLC Activity Assays

All HPLC assays were performed in 100 mM HEPES buffer at a final pH of 7.9 in a final volume of 100 μL. A 50 mM solution of amino acid substrate (10 μL) was mixed with a 100 mM solution of $MgCl_2$ (5 μL), 100 mM nucleoside triphosphate (1 μL), 100 mM CoASH (1 μL) and 10 mM DTT (1 μL) in 200 mM HEPES pH 8.3 containing 50 mM NaCl (50 μL). ThnN (5 μL) and ThnO (10 μL) were added to a final concentration of 3 μM. The assay mixture was kept at 37 °C for 4 h and the proteins removed using an Amicon® Ultracel® 3K centrifugal filter. The filtrate was transfered in a vial for HPLC analysis.

Analytical HPLC analysis of the nulceoside triphosphates was monitored using a Phenomenex Synergi Hydro-RP column (250 mm x 4.6 mm, 5 μm pore size). Eluent A was 20 mM ammonium acetate pH 6.3, eluent B was MeOH. The samples (10 μL) were eluted using a flow rate of 1 mL min^{-1} with 5% eluent over 1 minutes, followed by a gradient of 1–20% B over 20 minutes, and then 20-95% B over 5 minutes. The column was washed with 95% B for 5 minutes and the column was re-equilibrated at 5% B for 10 minutes before the next sample was injected.

7.4.8 ThnN and ThnO NMR Activity Assays

All NMR assays were performed in 100 mM HEPES buffer in 10% D_2O at a final pH of 7.9 in a final volume of 750 μL. A 50 mM solution of amino acid substrate (50 μL) was mixed with a 100 mM nucleoside triphosphate solution (25 μL), to a mixture (600 μL) containing 10 mM HEPES pH 8.3, 25 mM NaCl, 10 mM $MgCl_2$ and 10 mM DTT in 10% D_2O. ThnN (25 μL) and ThnO (50 μL) were added to a final concentration of 10 and 3 μM, respectively. The assay mixture was kept at 37 °C for the given time, and the proteins removed using an Amicon® Ultracel® 3K centrifugal filter. The filtrate was transfered in an 5 mm NMR tube for analysis. ^{1}H NMR spectra were recorded at 500 MHz using standard ^{1}H NMR Brucker pulse sequence with water suppression by presaturation (ns = 1024). ^{31}P NMR spectra were recorded at 202 MHz using the ^{31}P Brucker pulse sequence with proton decoupling (ns = 512).

List of Schemes

List of Figures

List of Tables

Appendices

Appendix A

Supplementary Information for Chapter 1

A.1 Human BBOX Gene and Protein Sequences

>gi|48145741|emb|CAG33093.1| BBOX1 [Homo sapiens]

```
  1 MACTIQKAEA LDGAHLMQIL WYDEEESLYP AVWLRDNCPC SDCYLDSAKA
 51 RKLLVEALDV NIGIKGLIFD RKKVYITWPD EHYSEFQADW LKKRCFSKQA
101 RAKLQRELFF PECQYWGSEL QLPTLDFEDV LRYDEHAYKW LSTLKKVGIV
151 RLTGASDKPG EVSKLGKRMG FLYLTFYGHT WQVQDKIDAN NVAYTTGKLS
201 FHTDYPALHH PPGVQLLHCI KQTVTGGDSE IVDGFNVCQK LKKNNPQAFQ
251 ILSSTFVDFT DIGVDYCDFS VQSKHKIIEL DDKGQVVRIN FNNATRDTIF
301 DVPVERVQPF YAALKEFVDL MNSKESKFTF KMNPGDVITF DNWRLLHGRR
351 SYEAGTEISR HLEGAYADWD VVMSRLRILR QRVENGN
```

Table A.1: hBBOX1 protein sequence.

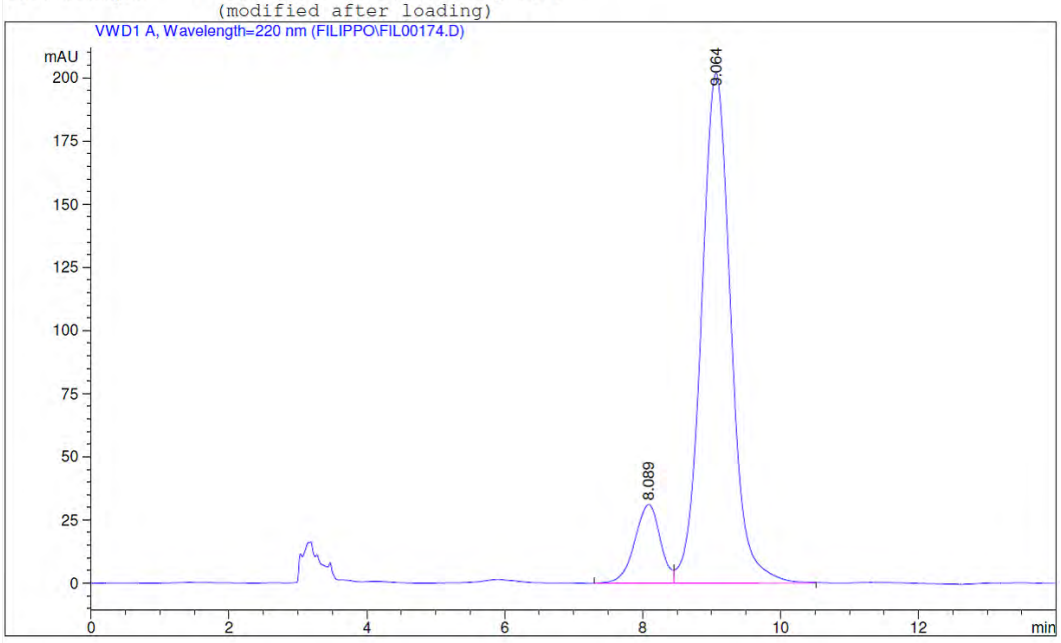
>gi|48145740|emb|CR456812.1| Homo sapiens full open reading frame
cDNA clone RZPDo834H0118D for gene BBOX1 incl. stop codon

```
   1 atggcttgta ccatccaaaa ggcagaagca cttgacgggg ctcatttgat gcagatcctc
  61 tggtatgatg aggaagagtc tctctaccca gctgtatggt tgagagacaa ctgtccgtgc
 121 tctgattgct acctggattc tgcaaaagca cggaaacttc tagtggaagc tcttgatgtg
 181 aacattggaa ttaaaggctt gatatttgac agaaaaaagg tgtacatcac atggcccgat
 241 gagcattaca gtgaattcca ggctgattgg ctgaagaaaa gatgcttttc caagcaggcc
 301 agagcaaagc tccaaagaga attgtttttt ccagaatgcc aatactgggg ctcagagctc
 361 cagctaccca ctttggattt tgaagatgtt ttaagatatg atgaacacgc atacaagtgg
 421 ctctccaccc tcaagaaagt aggcatagta agactcaccg gagcatctga caaaccagga
 481 gaagtttcaa aacttgggaa aaggatgggt ttcctctatc tcacatttta tggacatact
 541 tggcaagtgc aagacaaaat cgatgcaaac aatgtggctt acacaactgg gaagctaagc
 601 tttcacactg attatccagc cctccatcat ccacccgggg ttcagcttct tcactgcata
 661 aagcaaacag tcacaggggg tgattcagaa attgtagatg ggtttaatgt gtgccaaaaa
 721 ctaaagaaaa ataatcctca ggcattccag attttgtcct ctacctttgt ggactttaca
 781 gacattggag tggattactg tgatttttct gtacaatcaa aacataaaat tatagagtta
 841 gatgataaag gccaagtggt tcgcatcaac ttcaataacg caactaggga cacaatattt
 901 gatgtgcctg ttgaaagagt tcagcctttt tatgctgctc tgaaggagtt tgttgacctc
 961 atgaacagca aagaatccaa gtttaccttc aagatgaatc caggtgatgt gattactttt
1021 gataactggc gcttacttca tggccgacgt agctatgaag caggaactga gatatcccgc
1081 catctagaag gagcttatgc tgactgggat gtggtcatgt caaggcttcg tatcttaagg
1141 cagagggtgg agaatggaaa ttaa
```

Table A.2: hBBOX1 gene sequence.

A.2 Chiral HPLC Analyses

```
================================================================
Injection Date  : 18/11/2011 17:25:31 PM        Seq. Line :   2
Sample Name     : luc-05-Sample                 Location : Vial 8
Acq. Operator   : filippo                         Inj :   1
Acq. Instrument : Instrument 1                Inj Volume : 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\DJD26\95HEX220.M
Last changed    : 18/11/2011 17:24:45 PM by filippo
                  (modified after loading)
Analysis Method : C:\HPCHEM\1\METHODS\DJD26\90HEX220.M
Last changed    : 22/02/2012 09:48:56 PM by Dave
                  (modified after loading)
```

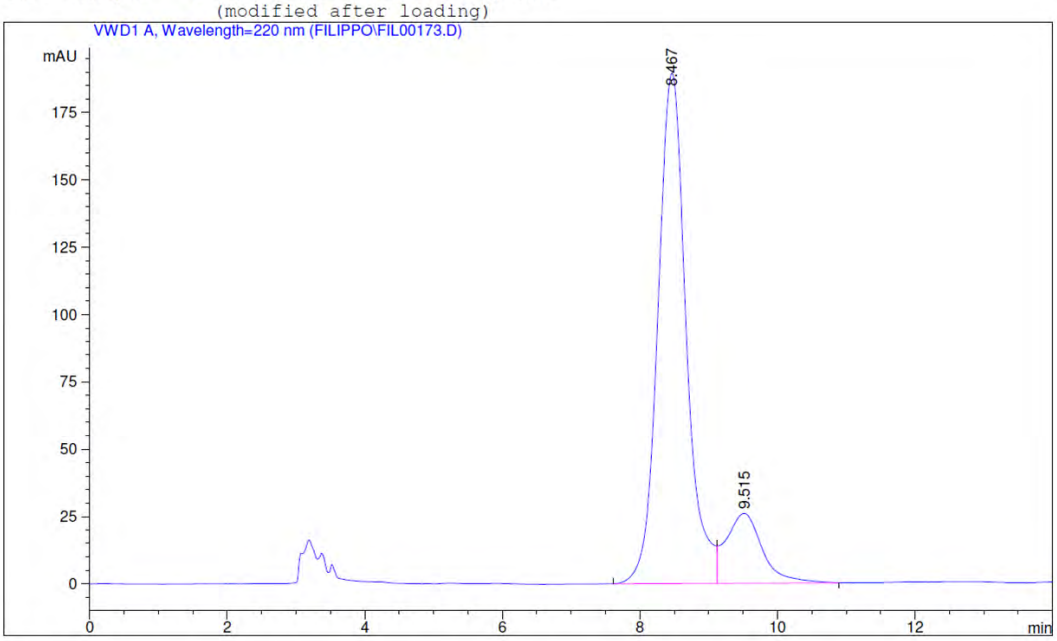



```
================================================================
                      Area Percent Report
================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=220 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   8.089 PV    0.4012  803.58893   31.12475  11.8721
   2   9.064 VB    0.4569 5965.13330  201.77008  88.1279

Totals :                  6768.72223  232.89483
```

Figure A.1: Chiral HPLC trace for (3*S*)-*tert*-butyl 4-(benzyl(methyl)amino)-3-(*tert*-butoxycarbonylamino)butanoate **105**. Integration shows ee = 77%. The sample was injected on a Chiracel AD column (250 mm x 4.6 mm x 5 μ) and eluted using 1 mL min^{-1} of 5% isopropanol in *n*-hexane. The trace represent the absorbance at λ = 220 nm.

luc, AD, 95-5, 1mL-min, 220 nm

```
======================================================================
Injection Date  : 30/11/2011 15:51:42 PM         Seq. Line :   2
Sample Name     : Luc enantipure R                Location : Vial 8
Acq. Operator   : filippo                            Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\DJD26\95HEX220.M
Last changed    : 30/11/2011 15:20:12 PM by filippo
Analysis Method : C:\HPCHEM\1\METHODS\DJD26\90HEX220.M
Last changed    : 22/02/2012 09:48:43 PM by Dave
                  (modified after loading)
```

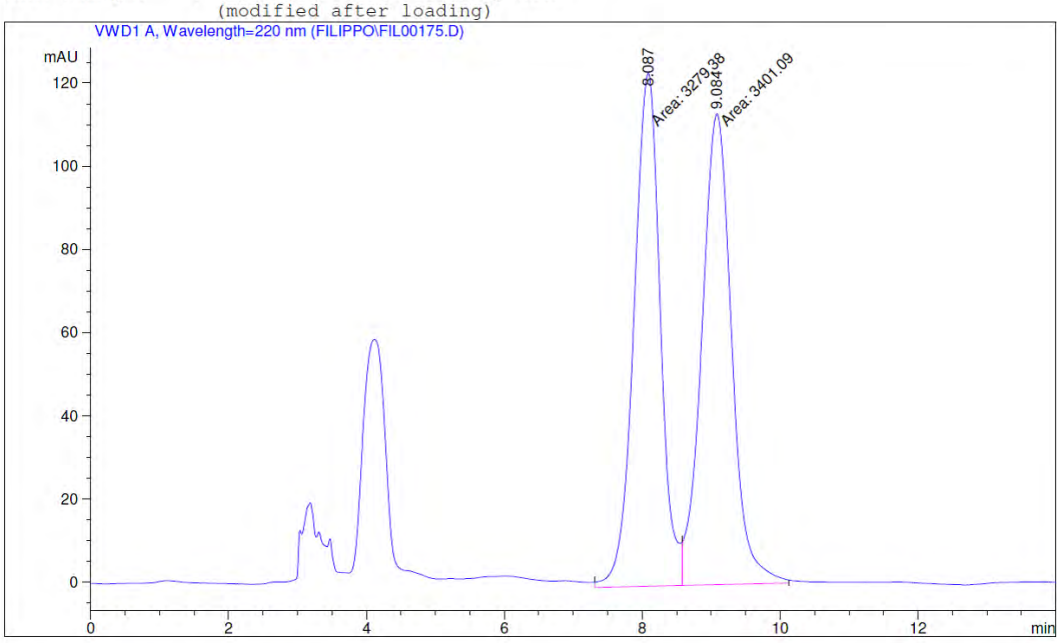



```
======================================================================
                        Area Percent Report
======================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=220 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   8.467 BV    0.4412 5409.93799  189.42065  85.0646
   2   9.515 VB    0.5396  949.86353   25.91426  14.9354

Totals :                  6359.80151  215.33491
```

Figure A.2: Chiral HPLC trace for (3*S*)-*tert*-butyl 4-(benzyl(methyl)amino)-3-(*tert*-butoxycarbonylamino)butanoate **112**. Integration shows ee = 71%. The sample was injected on a Chiracel AD column (250 mm x 4.6 mm x 5 μ) and eluted using 1 mL min^{-1} of 5% isopropanol in *n*-hexane. The trace represent the absorbance at λ = 220 nm.

A.3 NMR Spectra for 3-Amino-4-(methylamino)butanoic Acid Enantiomers

```
================================================================
Injection Date  : 18/11/2011 17:41:17 PM        Seq. Line :   3
Sample Name     : luc-67 Sample                  Location : Vial 7
Acq. Operator   : filippo                          Inj :   1
Acq. Instrument : Instrument 1                 Inj Volume : 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\DJD26\95HEX220.M
Last changed    : 18/11/2011 17:24:45 PM by filippo
                  (modified after loading)
Analysis Method : C:\HPCHEM\1\METHODS\DJD26\90HEX220.M
Last changed    : 22/02/2012 09:49:39 PM by Dave
                  (modified after loading)
```




```
================================================================
                      Area Percent Report
================================================================

Sorted By             :      Signal
Multiplier            :      1.0000
Dilution              :      1.0000
Use Multiplier & Dilution Factor with ISTDs


Signal 1: VWD1 A, Wavelength=220 nm

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]   [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1   8.087 MF    0.4430 3279.38232  123.38100  49.0890
   2   9.084 FM    0.5005 3401.09448  113.26304  50.9110

Totals :                  6680.47681  236.64404
```

Figure A.3: Chiral HPLC trace for racemic *tert*-butyl 4-(benzyl(methyl)amino)-3-(*tert*-butoxycarbonylamino)butanoate **106**. Integration shows no ee. The sample was injected on a Chiracel AD column (250 mm x 4.6 mm x 5 μ) and eluted using 1 mL min^{-1} of 5% isopropanol in *n*-hexane. The trace represent the absorbance at λ = 220 nm.

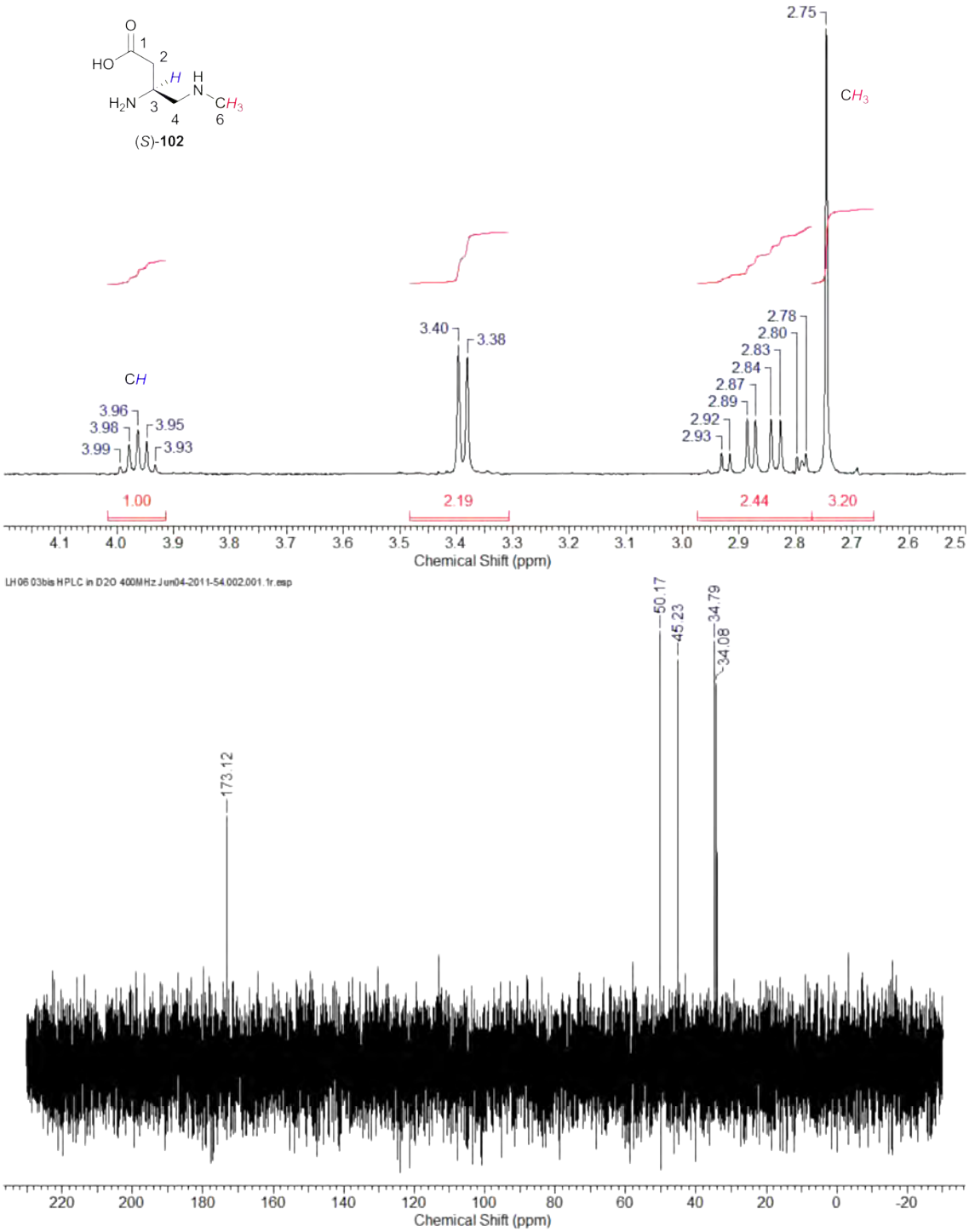


Figure A.4: (3*S*)-3-Amino-4-(methylamino)butanoic acid **102**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400MHz and 100MHz respectively.

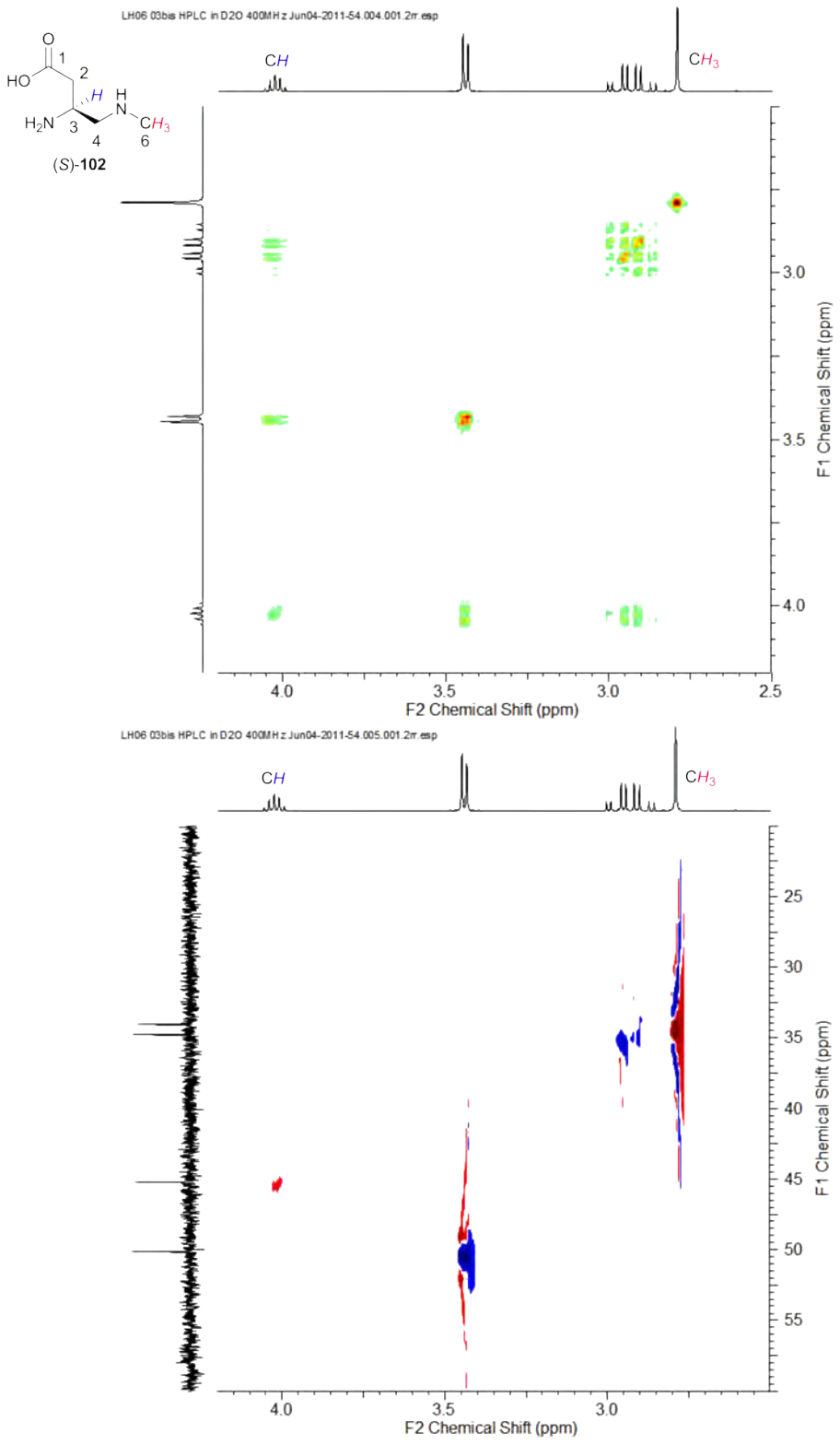


Figure A.5: (3*S*)-3-Amino-4-(methylamino)butanoic acid **102**. 2D COSY and ^{1}H-^{13}C HSQC NMR spectra were recorded in D_2O at 400MHz.

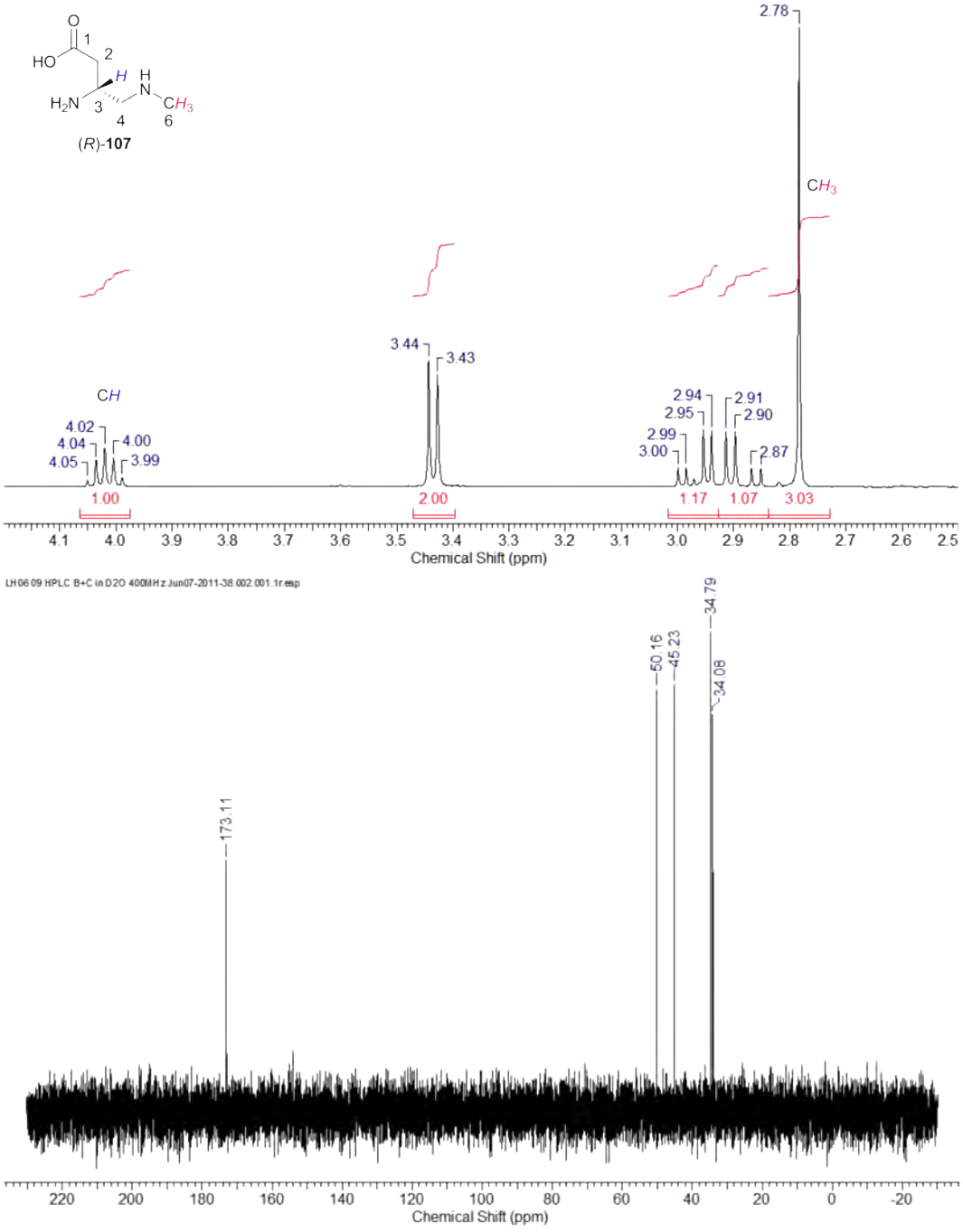


Figure A.6: (3*R*)-3-Amino-4-(methylamino)butanoic acid **107**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400MHz and 100MHz respectively.

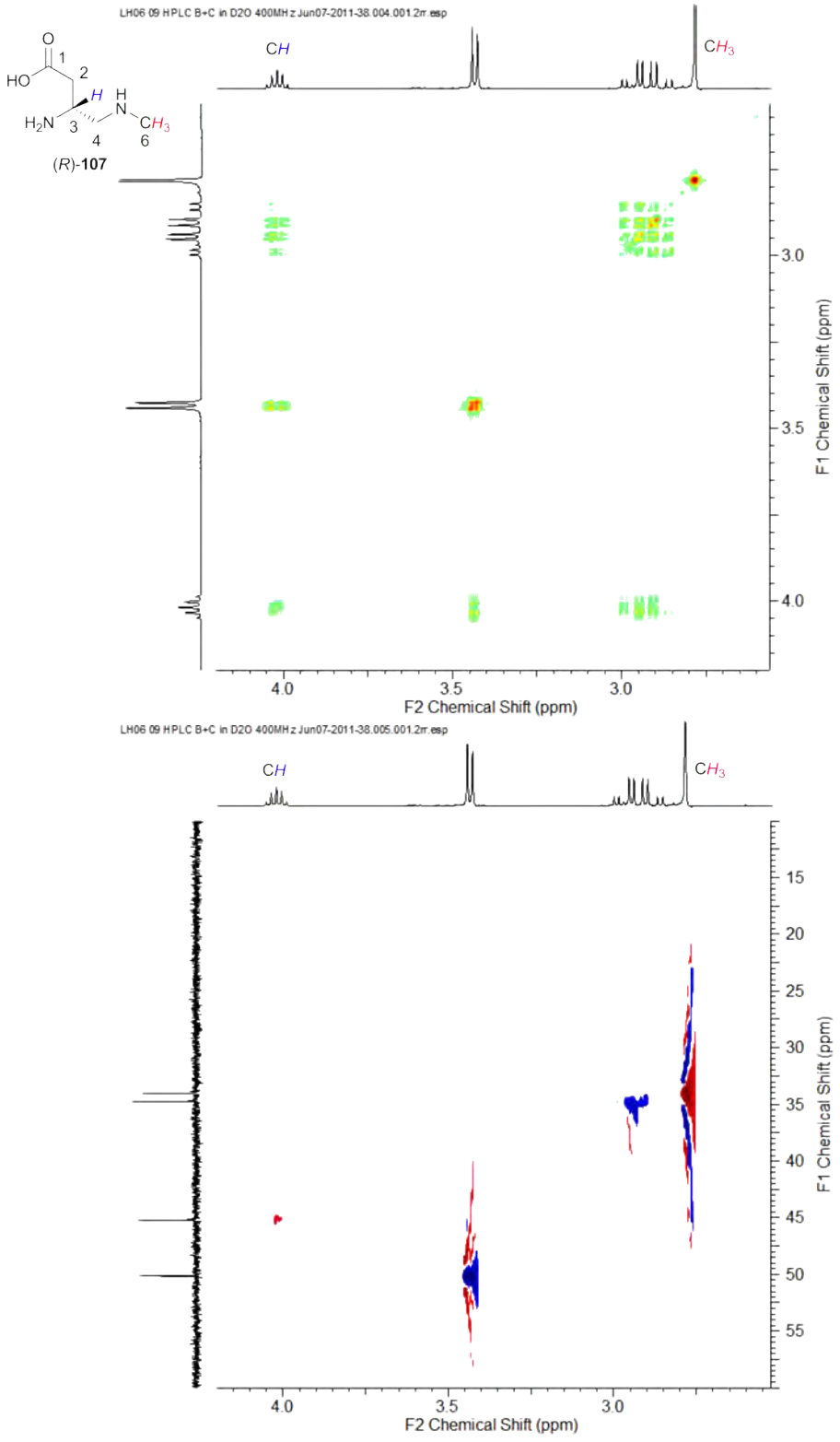


Figure A.7: (3*R*)-3-Amino-4-(methylamino)butanoic acid **107**. 2D COSY and ^{1}H-^{13}C HSQC NMR spectra were recorded in D_2O at 400MHz.

Appendix B

Supplementary Information for Chapter 2

B.1 CMPS Multiple Sequence Alignment

```
ThnE   MGAAAGERRTLSRKRGTTVDVPVKQESRQDRFEDQDEDRFAAQESIECSRLGDGIALAEFSGAHEQNPFSRARMRELTAL  80
ThnE*  ------------------MSVPIQPSQKN------------QKSVGCSHVGDGVVLVEFSGAHEQNPFSRARMQELTAL  49
CarB   ---------------------------------------------MVFEENSDEVRVITLDHPNKHNPFSRTLETSVKDA  35
CpmB   ---------------------------------------------MVIERDFSSIRVIILNHENKHNPFSESLENSIKKS  35
                                                       :  ..  . : :  :.  ::::***.:   .:.

ThnE   MRELDADEKVRCVVLYGGAGRSFGVGGDFHEVSEFTGGDEVNAWIDDITDLYTTVAAISKPVIAAIDGYAIGVGLQISLC  160
ThnE*  MHDLDQDASVRSVVLHGGSGRSFGVGGDFNEVSAFTGGDEVSDWIDDITDLYTAIAGISKPVVAAIDGYAIGIGLQIALC  129
CarB   LARANADDSVRAVVVYGGAERSFSAGGDFNEVKQLSRSEDIEEWIDRVIDLYQAVLNVNKPTIAAVDGYAIGMGFQFALM  115
CpmB   LIKADQDDNVKAIVVYGGENRSFSSGGDFSEVKNLS-GESVERWIDRVIDLYCAVLNVNKPTVAAVDGYAIGMGFQFSLM  114
       :   : * .*:.:*::**  ***. **** **. :: .:.:. *** : *** ::  :.**.:**********:*:::*

ThnE   CDYRLGSEQARLVMPEFRVGIACNFGGFMLEAAAGRTVMQRMLLTCDEWPAERALADGLLHETVASPRLLDRALELARTI  240
ThnE*  CDYRVAADTARLVMPELRVGIACNFGGFMLRYTVGTSVMQRMLFTCEEWPADRALEDGLLHEVRPSEDVLDRALELARTI  209
CarB   FDQRLMASTANFVMPELKHGIGCSVGAAILGFTHGFSTMQEIIYQCQSLDAPRCVDYRLVNQVVESSALLDAAITQAHVM  195
CpmB   FDQRIVSSEAKFIMPELKHGIGCSVGAAILSFTHGHNIMKKIVFECEEISSEMCIEYNIANQVIDKEILLETAIERANLL  194
        * *: :. *.:****:: **.*..*. :*  : * . *:.::  *:.  :  .:   : ::.  .  :*: *:  *. :

ThnE   SGYTAEAVQSTRPRVNAPFVAGLERIRREAKESHRRAFAAGEAQVRMRRVIGRS-                          294
ThnE*  AGYTPAAVQGTRPRVNAPFVEGLERVRTEAKESHRRAFSAGEAQARMRRIIGKD-                          263
CarB   ASYPASAFINTKRAVNKPFIHLLEQTRDASKAVHKAAFQARDAQGHFKNVLGKKY                          250
CpmB   ATYPKTAYLNTKKFMNKKFVDILEDSRKESKLVHKNSFGSRDSQKHFKKVLGEKY                          249
       : *.  *  .*:  :*  *:  **  *  :*  *: :* : ::* :::.::*..
```

Sequence alignment for *Streptomyces cattleya* ThnE (gi|337763142), *Streptomyces flavogriseus* ATCC 33331 ThnE* (gi|357409413), *Pectobacterium carotovorum* CarB (gi|5042366) and *Photorhabdus luminescens* CpmB (gi|37524205) generated using Clustal W [1]. The proposed oxyanion hole forming residues (Gly-108 and Val-153 in ThnE) are highlighted in blue. The carboxylate proposed to be responsible for catalysis (Glu-176) is highlighted in red.

Table B.1: CMPS multiple sequence alignment.

Appendix C

Supplementary Information for Chapter 4

C.1 Sequences of Genes and Proteins used in Chapter 4

>gi|5042366|gb|AAD38230.1| CarB [Pectobacterium carotovorum]

```
  1  MVFEENSDEV RVITLDHPNK HNPFSRTLET SVKDALARAN ADDSVRAVVV
 51  YGGAERSFSA GGDFNEVKQL SRSEDIEEWI DRVIDLYQAV LNVNKPTIAA
101  VDGYAIGMGF QFALMFDQRL MASTANFVMP ELKHGIGCSV GAAILGFTHG
151  FSTMQEIIYQ CQSLDAPRCV DYRLVNQVVE SSALLDAAIT QAHVMASYPA
201  SAFINTKRAV NKPFIHLLEQ TRDASKAVHK AAFQARDAQG HFKNVLGKKY
```

Table C.1: *Pectobacterium carotovorum* CarB protein sequence.

>gb|U17224.2|ECU17224:3120-3872 CarB [Pectobacterium carotovorum]

```
  1  atggtttttg aagagaattc cgatgaggtc cgagtgatca cgctcgatca tccgaacaag
 61  cataacccct ttagtcgaac gctggaaacc agcgtgaagg acgcgctggc gcgagccaac
121  gctgacgaca gcgtacgagc tgtcgtggtg tatggcgggg ccgagcgttc cttttccgct
181  ggcggagatt tcaacgaagt taagcagcta tcgcgcagcg aggacatcga agagtggatt
241  gaccgcgtga ttgatttgta tcaggccgtc ctgaatgtga ataaaccgac gatcgcagcg
301  gtggatggct atgcgattgg tatgggtttc cagttcgccc tgatgtttga ccagagactg
361  atggcgtcga cagcaaattt tgtgatgccg gaacttaagc atggtatcgg gtgctcggtt
421  ggggcggcca ttctgggatt cacccacgga tttagcacaa tgcaggaaat catctaccag
481  tgccagtcgc tcgatgctcc acgctgtgtg gactaccggc tggttaatca ggtggtggag
541  agcagtgcgc tgctggatgc cgctatcaca caggcgcacg tgatggctag ctacccggct
601  tctgcattca tcaatacgaa acgagcggtt aacaaaccgt tcatccatct actggaacaa
661  acccgtgacg cttccaaagc tgtccataag gcagcgttcc aggctcgtga cgctcaggga
721  catttcaaaa atgtgcttgg caaaaaatac tga
```

Table C.2: *Pectobacterium carotovorum* CarB gene sequence.

>gi|357408193|ref|YP_004920116.1| ThnE [Streptomyces cattleya NRRL 8057]

```
  1 MKQESRQDRF EDQDEDRFAA QESIECSRLG DGIALAEFSG AHEQNPFSRAR
 51 MRELTALMRE LDADEKVRCV VLYGGAGRSF GVGGDFHEVS EFTGGDEVNAW
101 IDDITDLYTT VAAISKPVIA AIDGYAIGVG LQISLCCDYR LGSEQARLVMP
151 EFRVGIACNF GGFMLEAAAG RTVMQRMLLT CDEWPAERAL ADGLLHETVAS
201 PRLLDRALEL ARTISGYTAE AVQSTRPRVN APFVAGLERI RREAKESHRRA
251 FAAGEAQVRM RRVIGRS
```

Table C.3: *Streptomyces cattleya* ThnE protein sequence.

>gi|357407371:906899-907717 ThnE [Streptomyces cattleya NRRL 8057]

```
  1 gtgaagcagg agtcgcggca ggaccggttc gaggatcagg acgaggaccg gttcgcggcg
 61 caggagtcca tcgagtgctc gcggctcggc gacgggatcg ccctggcgga gttctcgggg
121 gcgcacgagc agaacccgtt cagccgcgcg cggatgcgcg aactcaccgc gttgatgcgg
181 gagttggacg cggacgagaa ggtcaggtgc gtggtgctgt acggcggcgc cgggcggtcg
241 ttcggggtgg ggggcgactt ccacgaggtg tcggagttca ccggcgggga cgaggtcaac
301 gcgtggatcg acgacatcac cgacctgtac accacggtgg cggcaatctc caagccggtg
361 atcgccgcca tcgacggtta tgccatcggt gtcgggttgc agatctcgtt gtgttgtgac
421 taccggctcg gcagtgagca ggcacggctg gtgatgccgg agttccgggt gggcatcgcc
481 tgcaacttcg gcggtttcat gctggaggcc gcggccggcc gcaccgtgat gcagcgcatg
541 ctgctgacct gcgacgagtg gccggccgaa cgggcgctgg cggacgggct gctgcacgag
601 acggtggcct cgccccggct gctggaccgg gccctggagc tggcccggac catctccggc
661 tacaccgccg aggccgtgca gtccacccgt ccgcgggtca acgcgccttt cgtggccggg
721 ctggagcgga tccgccgcga ggcgaaggag tcgcaccgca gagccttcgc ggccggcgag
781 gcccaggtgc ggatgcgccg cgtcatcggg cggagctga
```

Table C.4: *Streptomyces cattleya* ThnE gene sequence.

C.2 Sequences of ThnE Expression Construct

```
  1 MIECSRLGDG IALAEFSGAH EQNPFSRARM RELTALMREL DADEKVRCVV
 51 LYGGAGRSFG VGGDFHEVSE FTGGDEVNAW IDDITDLYTT VAAISKPVIA
101 AIDGYAIGVG LQISLCCDYR LGSEQARLVM PEFRVGIACN FGGFMLEAAA
151 GRTVMQRMLL TCDEWPAERA LADGLLHETV ASPRLLDRAL ELARTISGYT
201 AEAVQSTRPR VNAPFVAGLE RIRREAKESH RRAFAAGEAQ VRMRRVIGRS
251 *
```

Table C.5: ThnE V153V Δ2-46 expression construct protein sequence (250 a.a. calc. MW = 27314 Da).

Appendix D

Supplementary Information for Chapter 5

D.1 NMR Spectra for Pantetheine Analogues

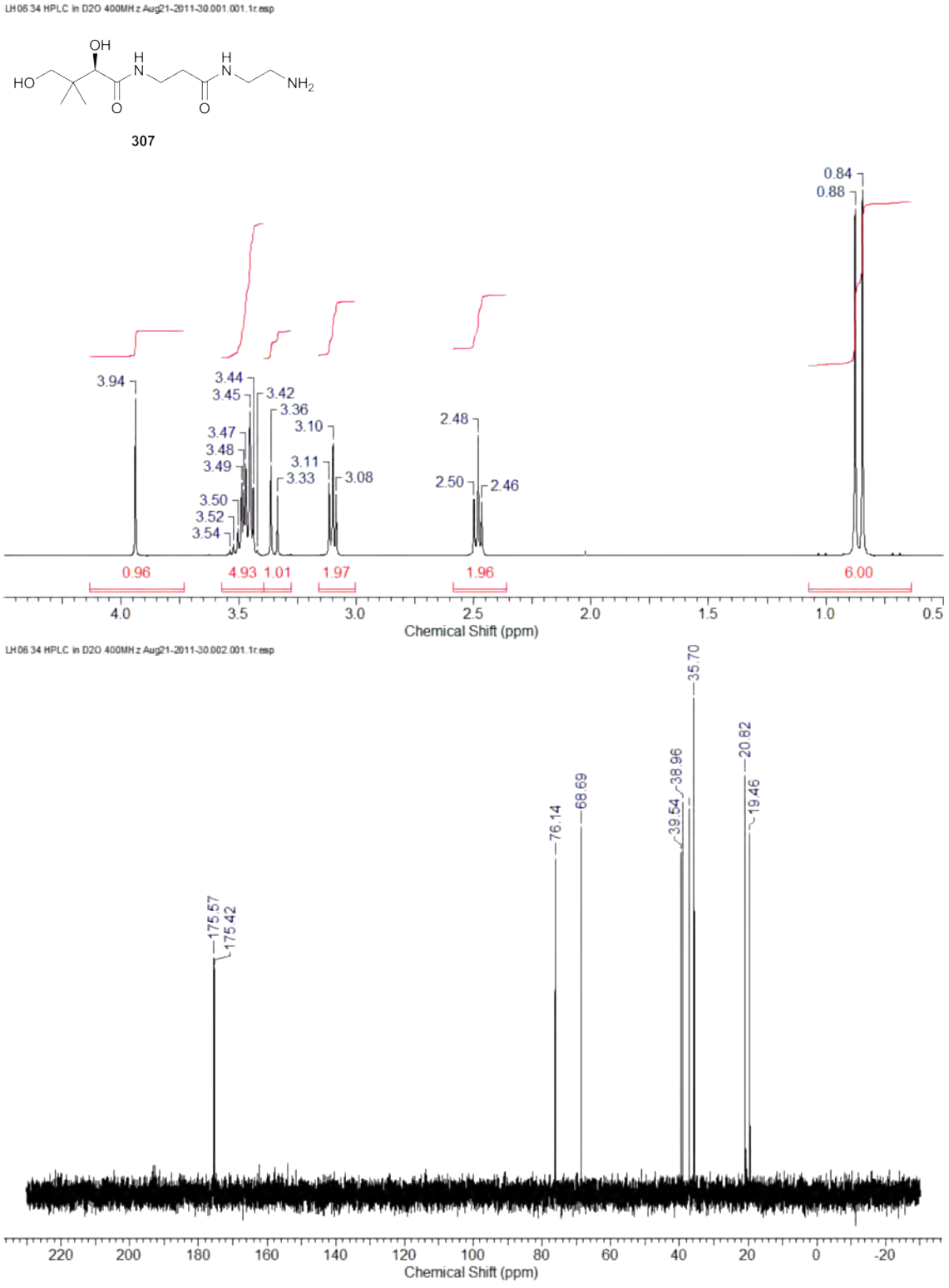


Figure D.1: (*R*)-*N*-(3-(2-Aminoethylamino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide **307**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400 MHz and 100 MHz, respectively.

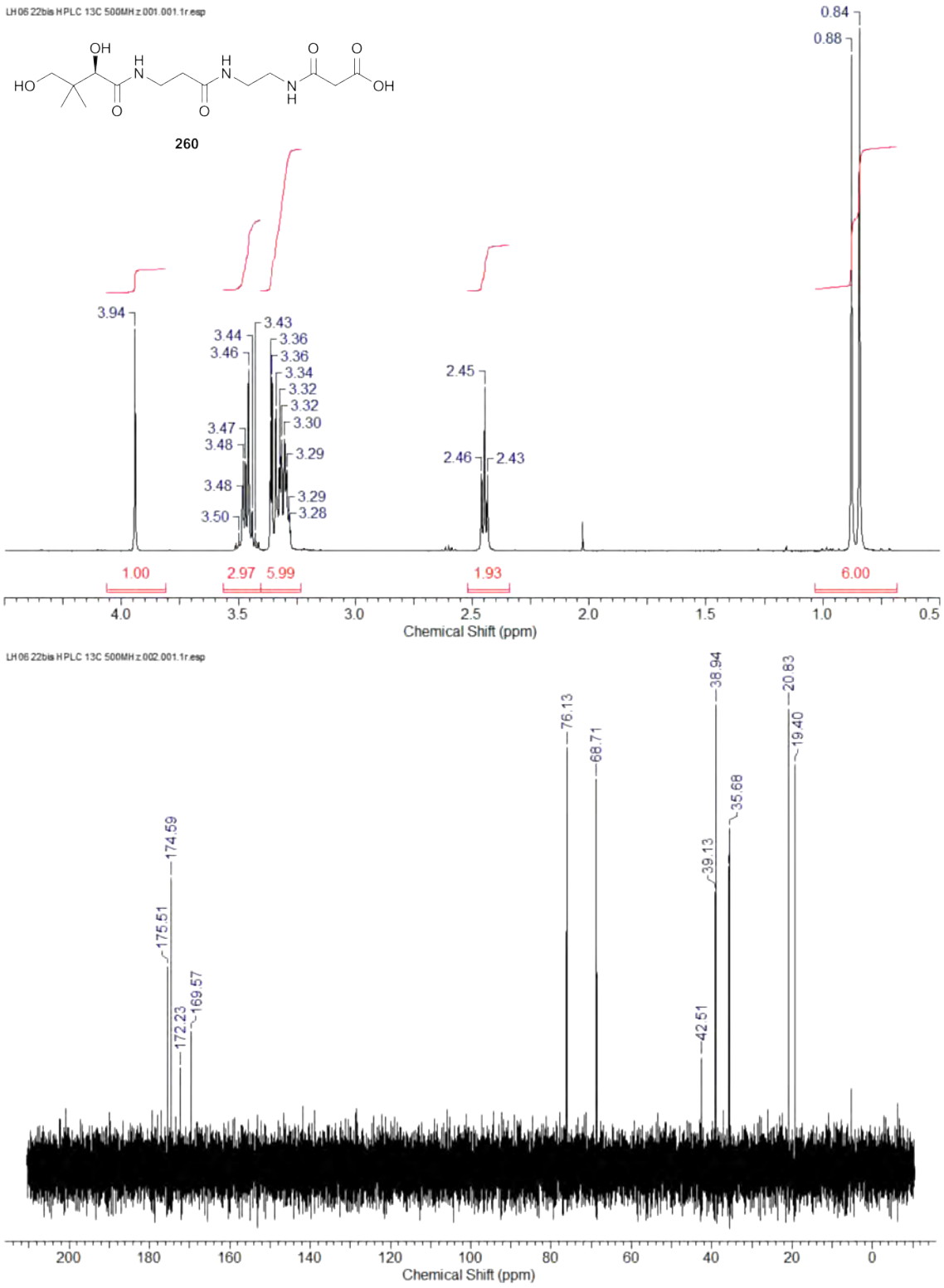


Figure D.2: (*R*)-3-(2-(3-(2,4-Dihydroxy-3,3-dimethylbutanamido)propanamido)ethylamino)-3-oxopropanoic acid **260**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 500 MHz and 125 MHz, respectively.

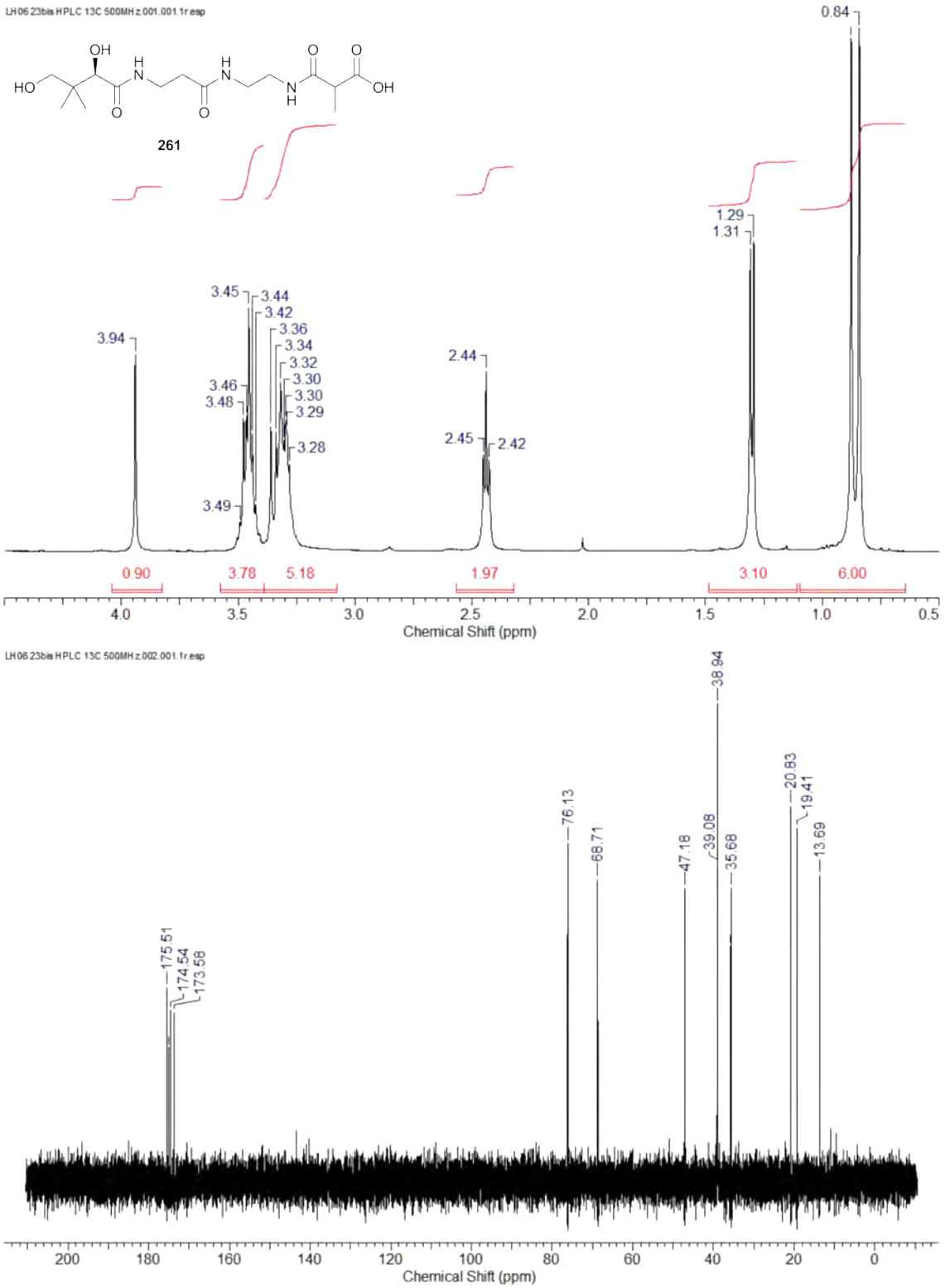


Figure D.3: 3-(2-(3-((*R*)-2,4-Dihydroxy-3,3-dimethylbutanamido)propanamido)ethylamino)-2-methyl-3-oxopropanoic acid **261**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 500 MHz and 125 MHz, respectively.

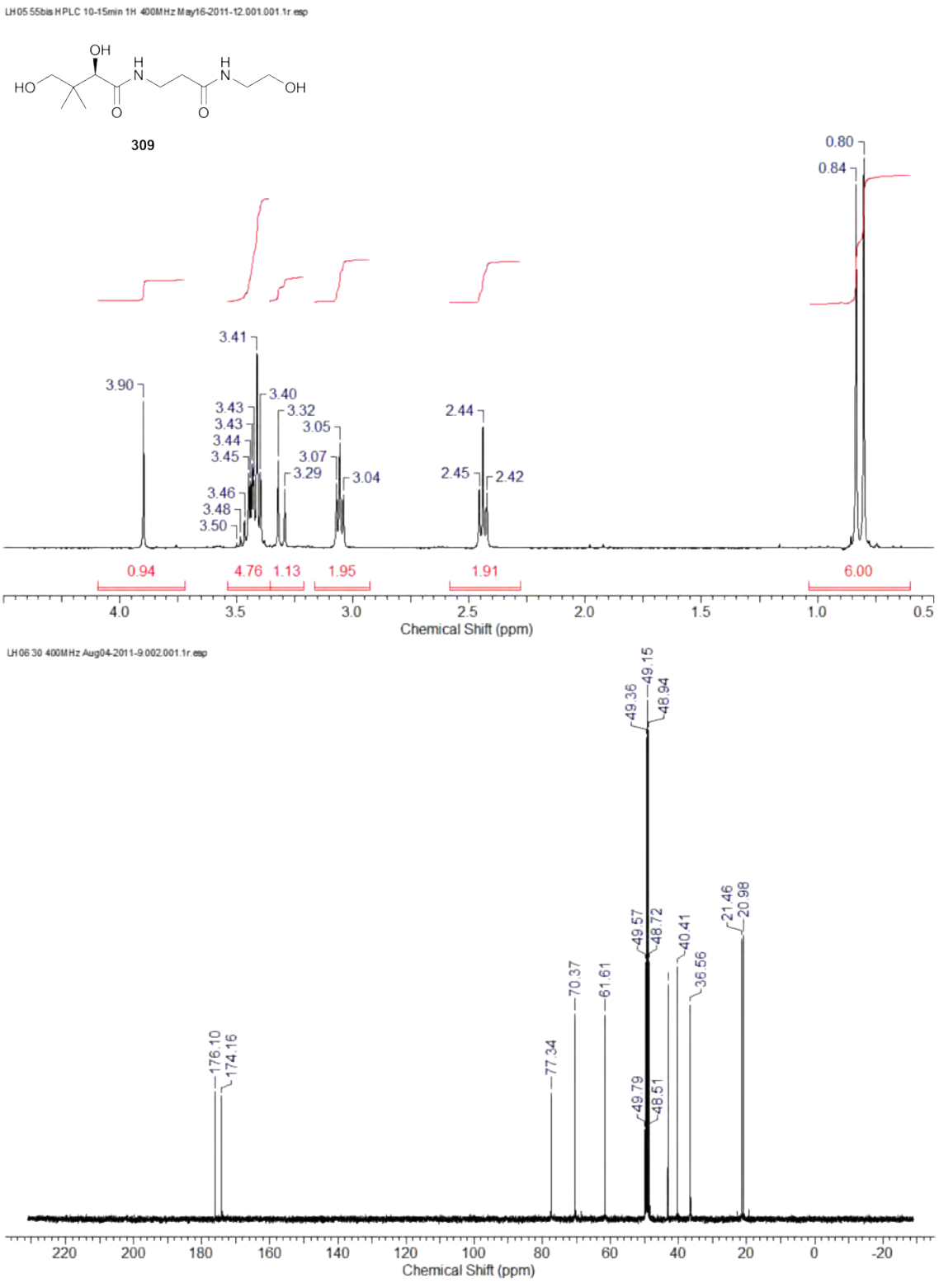


Figure D.4: (*R*)-2,4-Dihydroxy-*N*-(3-(2-hydroxyethylamino)-3-oxopropyl)-3,3-dimethylbutanamide **309**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400 MHz and 100 MHz, respectively.

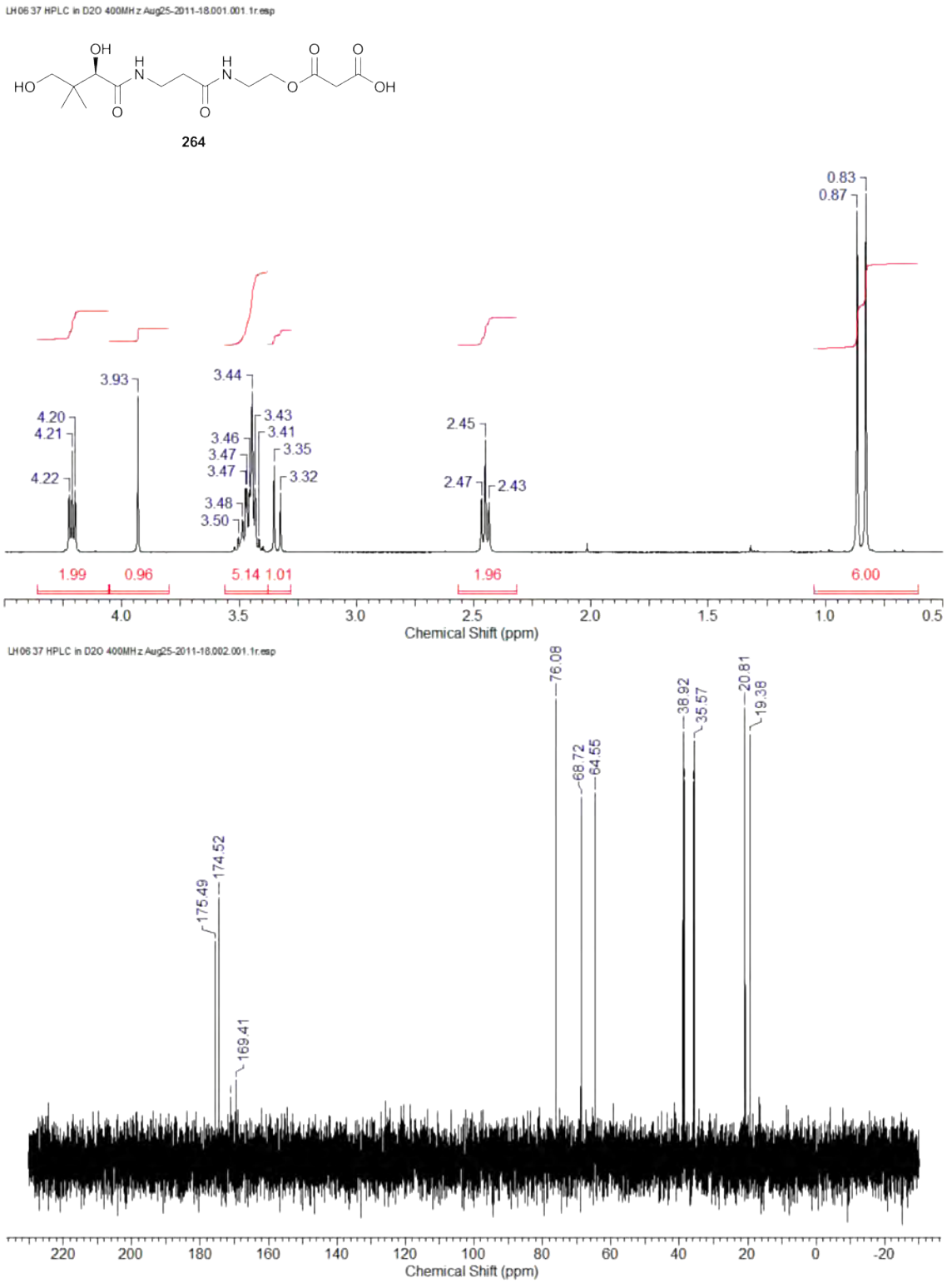


Figure D.5: (*R*)-3-(2-(3-(2,4-Dihydroxy-3,3-dimethylbutanamido)propanamido)ethoxy)-3-oxopropanoic acid **264**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400 MHz and 100 MHz, respectively.

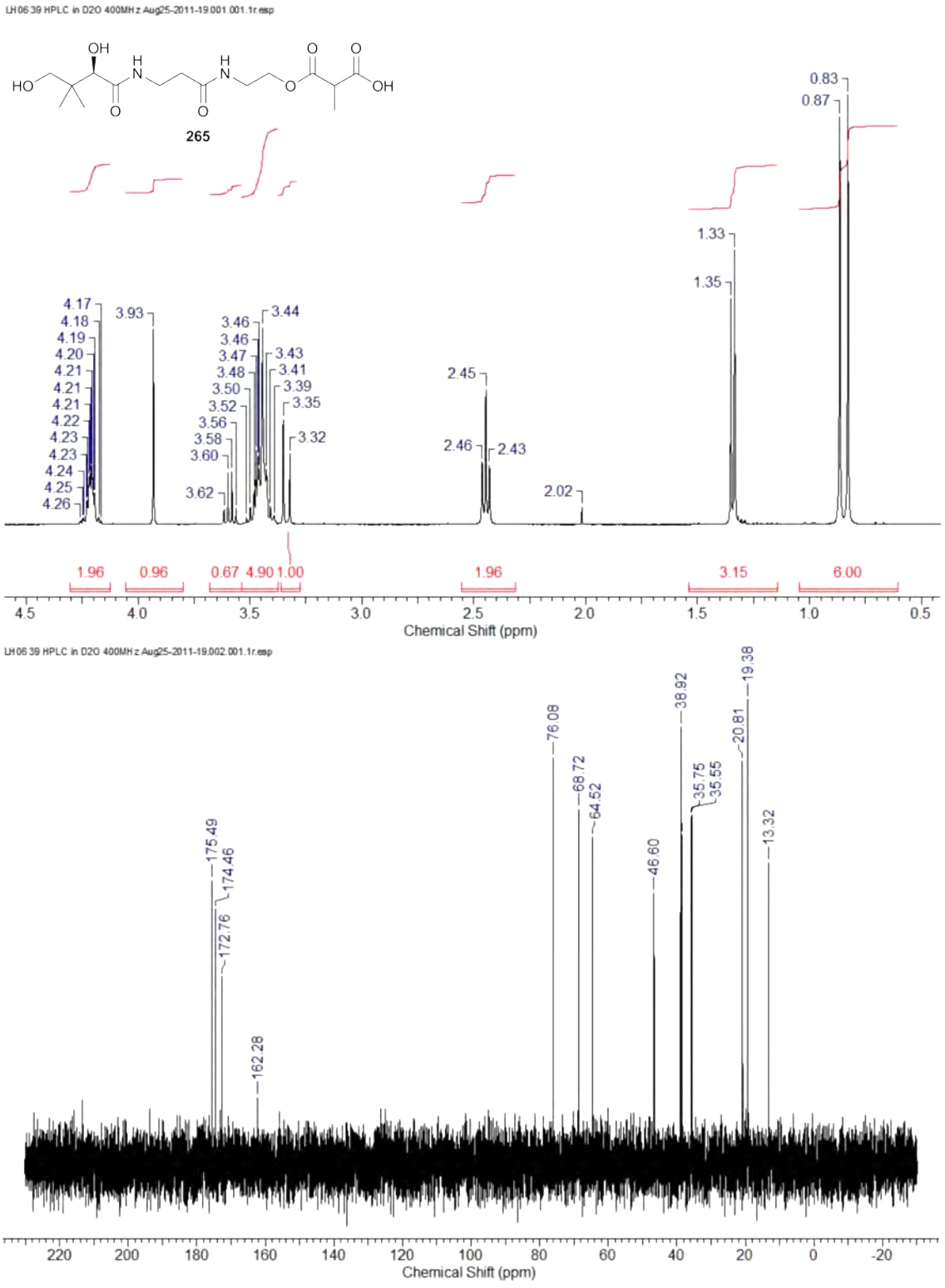


Figure D.6: 3-(2-(3-((*R*)-2,4-Dihydroxy-3,3-dimethylbutanamido)propanamido)ethoxy)-2-methyl-3-oxopropanoic acid **265**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400 MHz and 100 MHz, respectively.

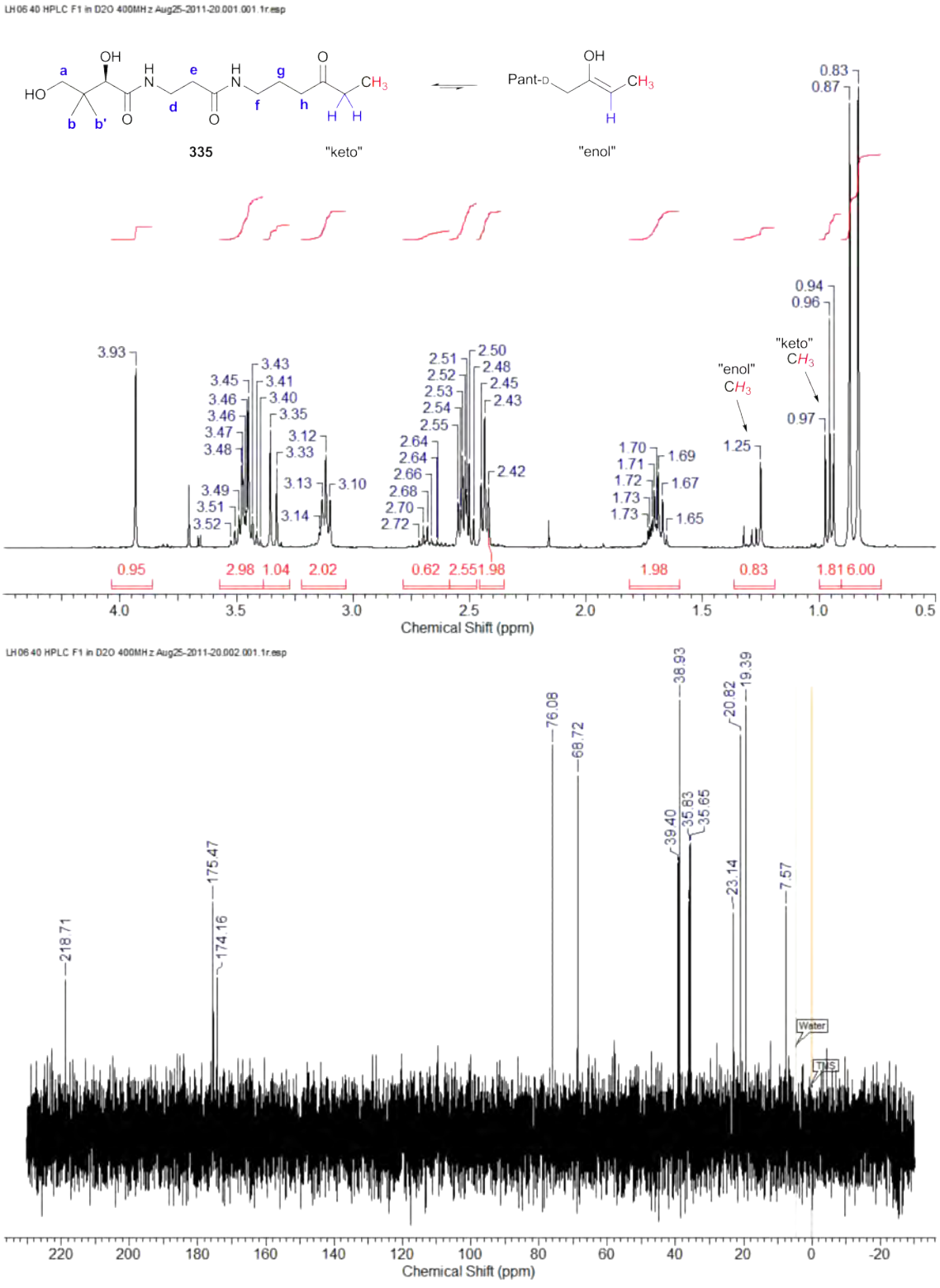


Figure D.7: 2,4-Dihydroxy-3,3-dimethyl-*N*-(3-oxo-3-(4-oxohexylamino)propyl)butanamide **335**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 400 MHz and 100 MHz, respectively.

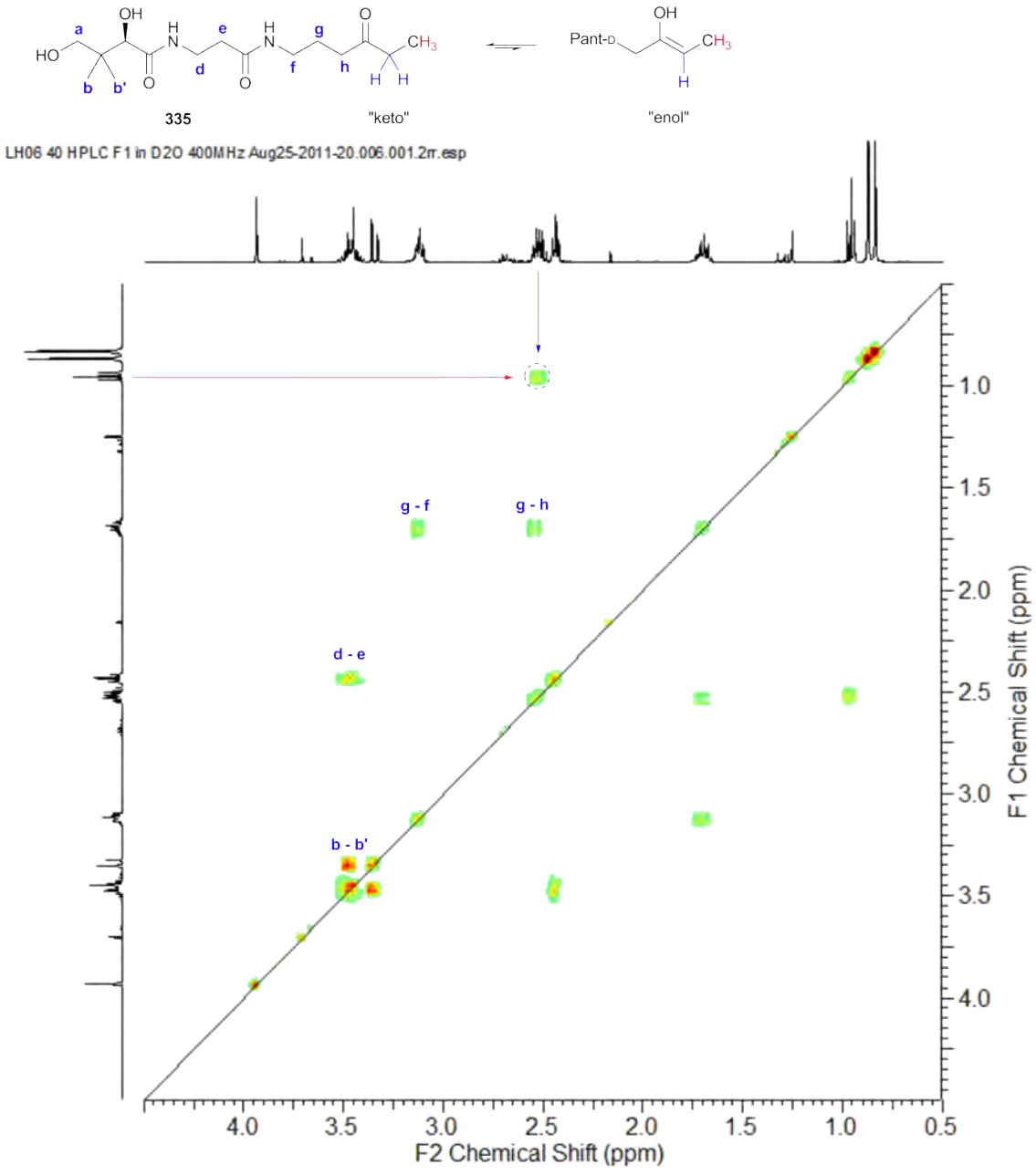


Figure D.8: 2,4-Dihydroxy-3,3-dimethyl-*N*-(3-oxo-3-(4-oxohexylamino)propyl)butanamide **335** recorded in D_2O at 400 MHz.

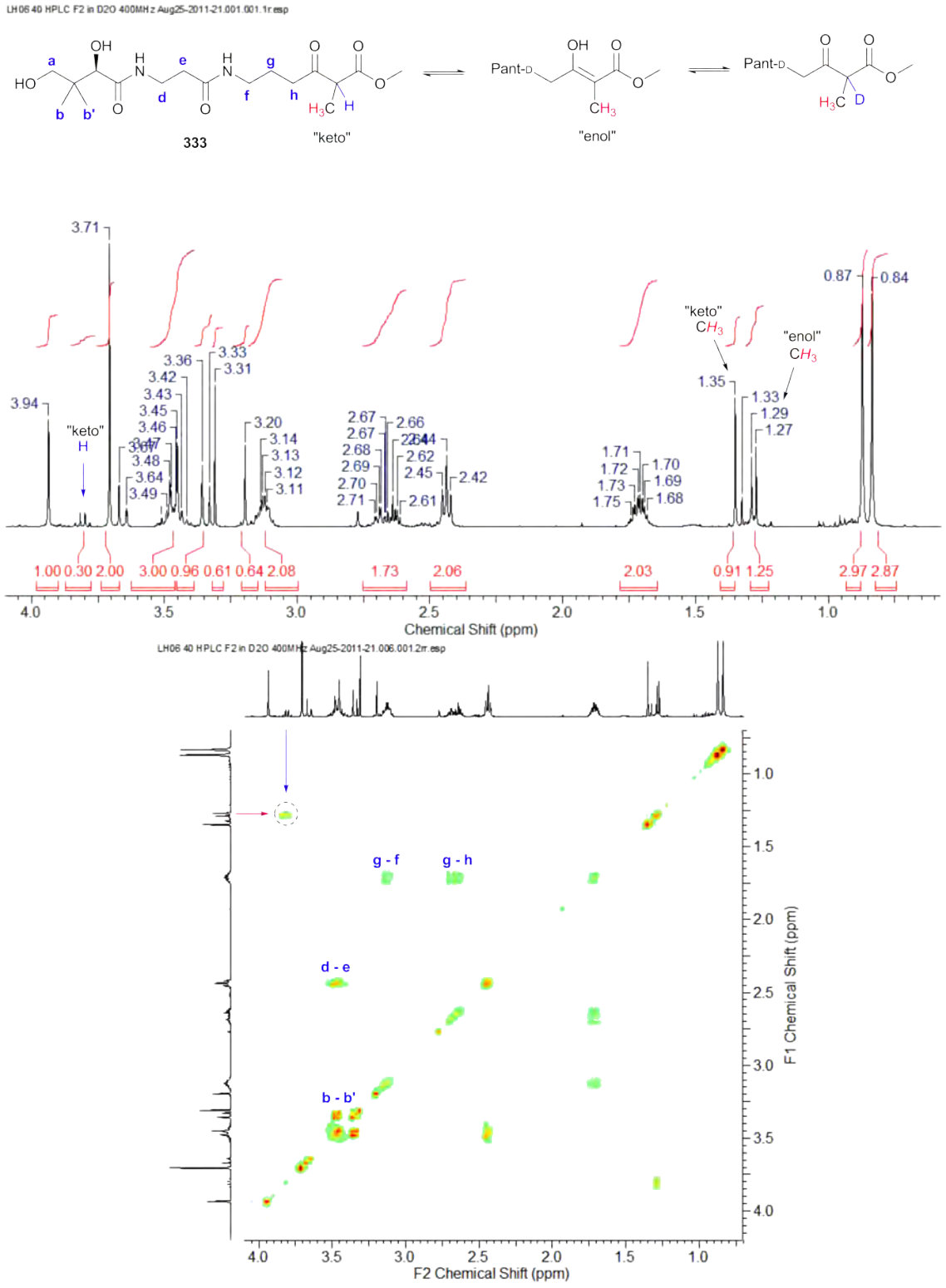


Figure D.9: Methyl 6-(3-((*R*)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)-2-methyl-3-oxohexanoate **333**. ^{1}H NMR and COSY NMR spectra were recorded in D_2O at 400 MHz.

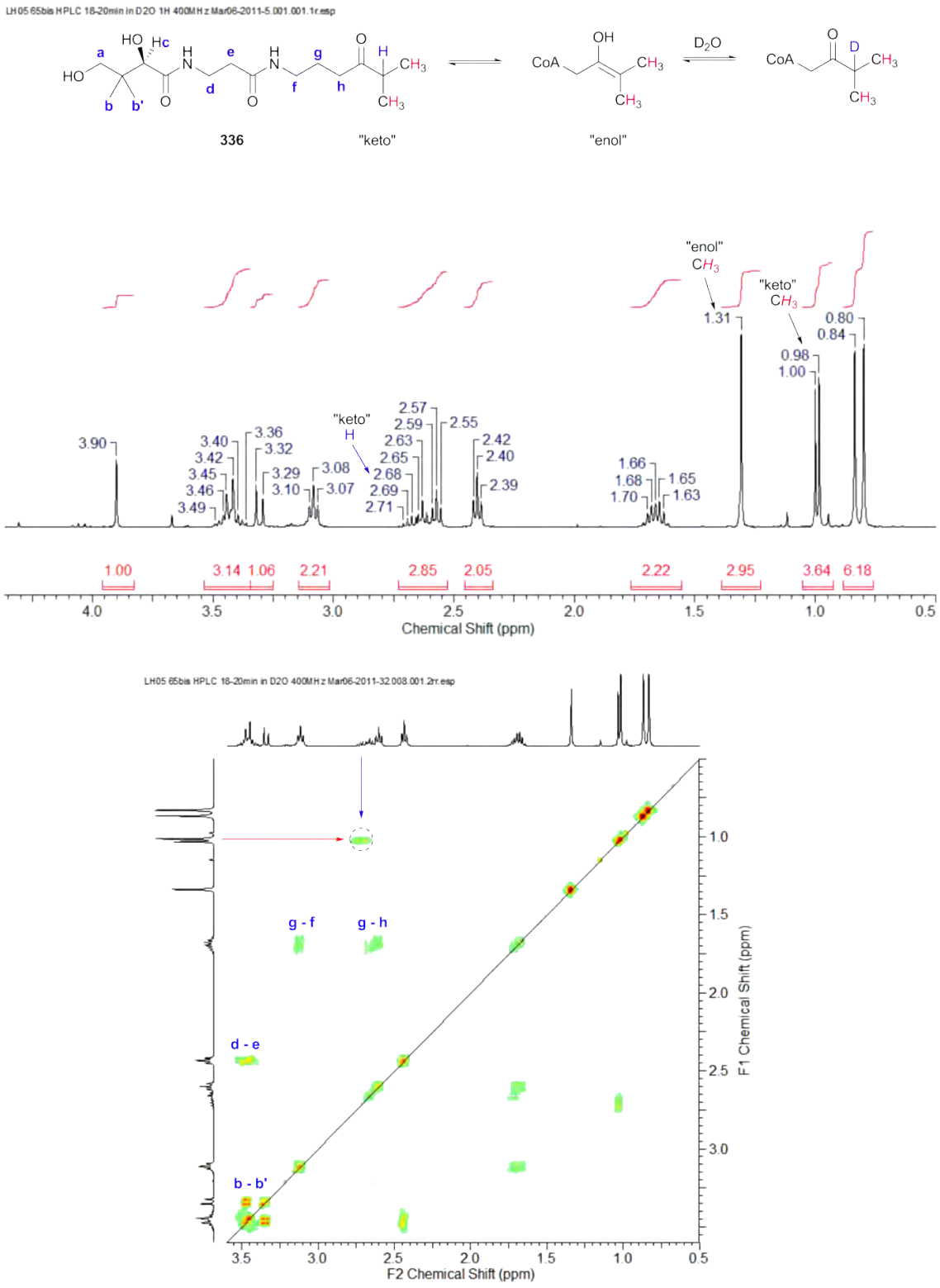
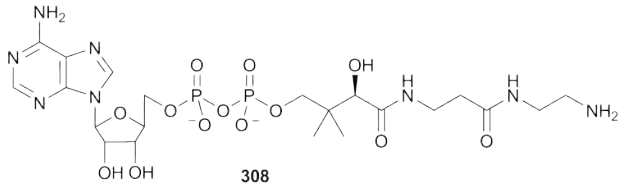


Figure D.10: 2,4-Dihydroxy-*N*-(3-(4-hydroxy-5-methylhex-4-enylamino)-3-oxopropyl)-3,3-dimethylbutanamide **336**. ^{1}H NMR and 2D COSY NMR spectra were recorded in D_2O at 400 MHz.

D.2 HPLC Traces and NMR Spectra for CoA Analogues



308

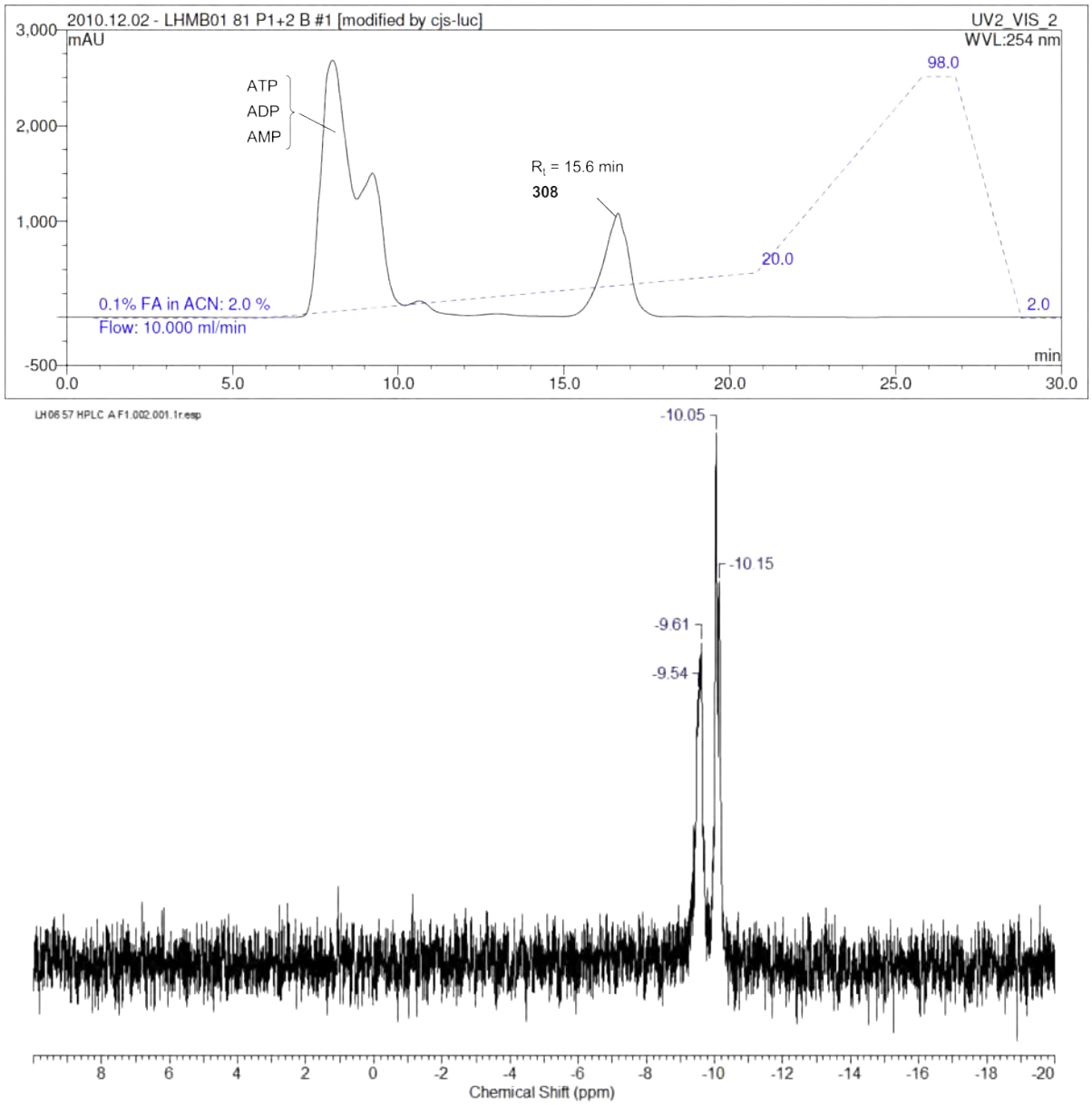


Figure D.11: Preparative HPLC trace and ^{31}P NMR spectrum for dephospho-amino(dethia)CoA **308**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). The trace represents the UV absorbance at λ = 254 nm. ^{31}P NMR spectrum was recorded in D_2O at 202 MHz.

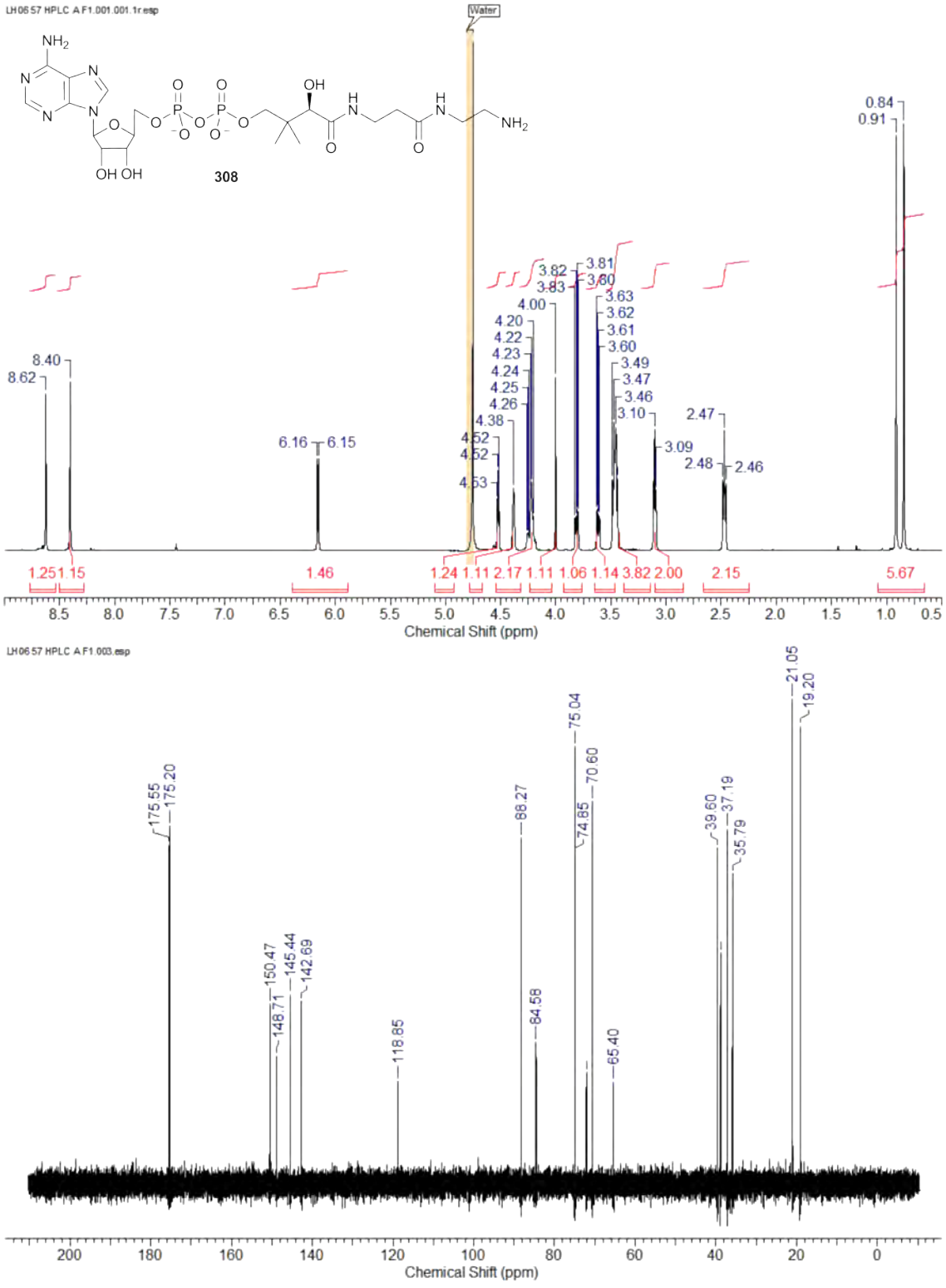


Figure D.12: ^{1}H and ^{13}C NMR spectra for dephospho-amino(dethia)CoA **308** recorded in D_2O at 500 MHz and 125 MHz respectively.

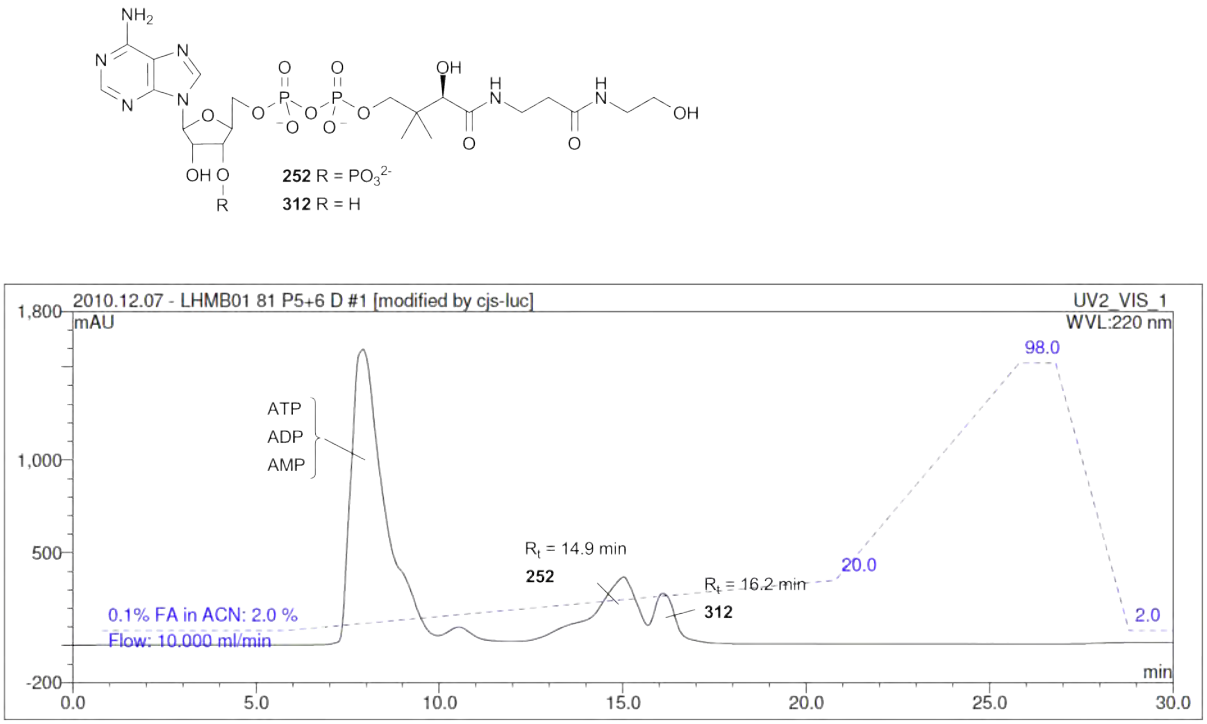


Figure D.13: Preparative HPLC trace for compound **252**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). The trace represents the UV absorbance at λ = 220 nm.

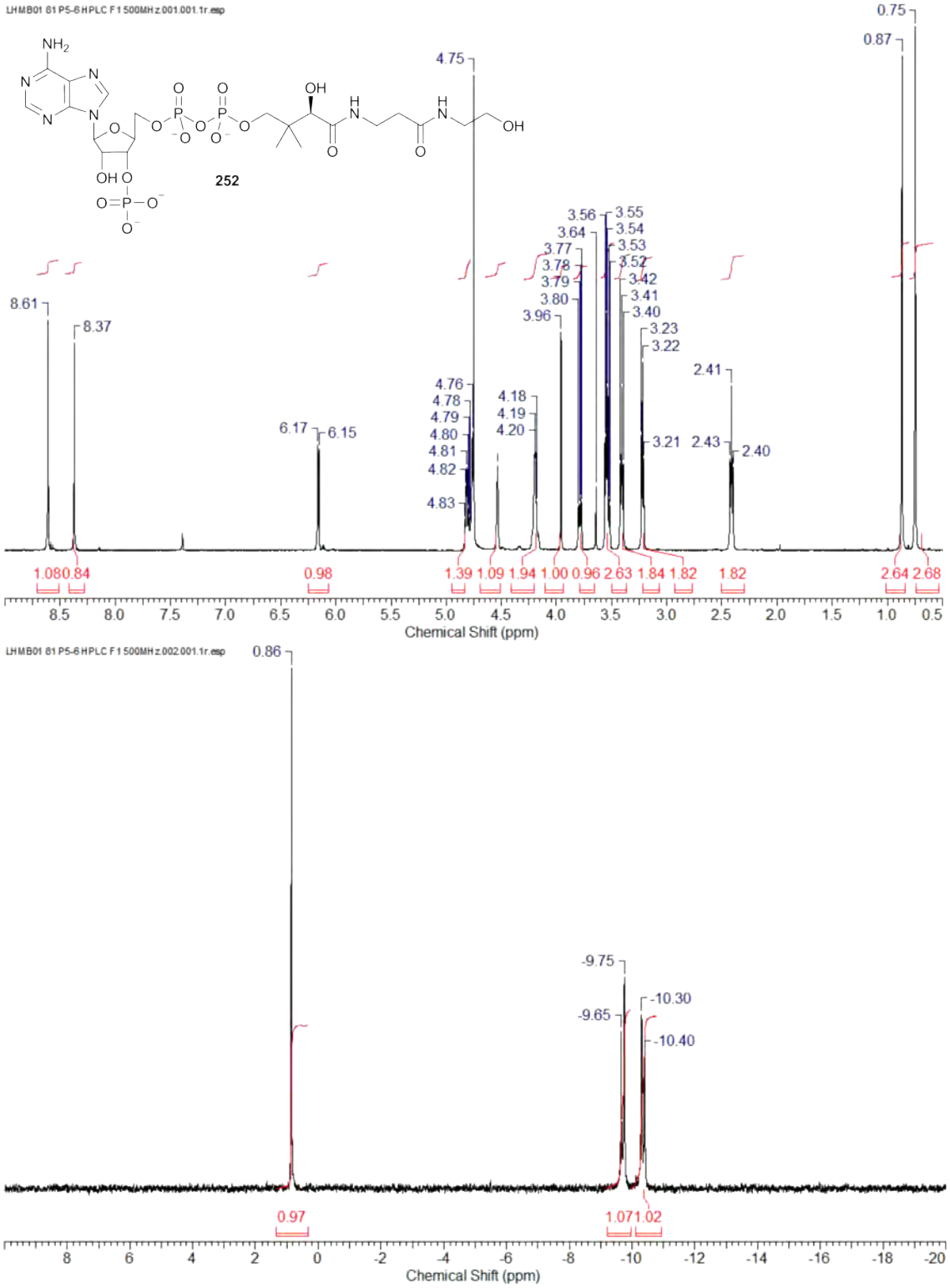


Figure D.14: ^{1}H and ^{31}P NMR spectra for hydroxy(dethia)CoA **252** recorded in D_2O at 500 MHz and 202 MHz, respectively.

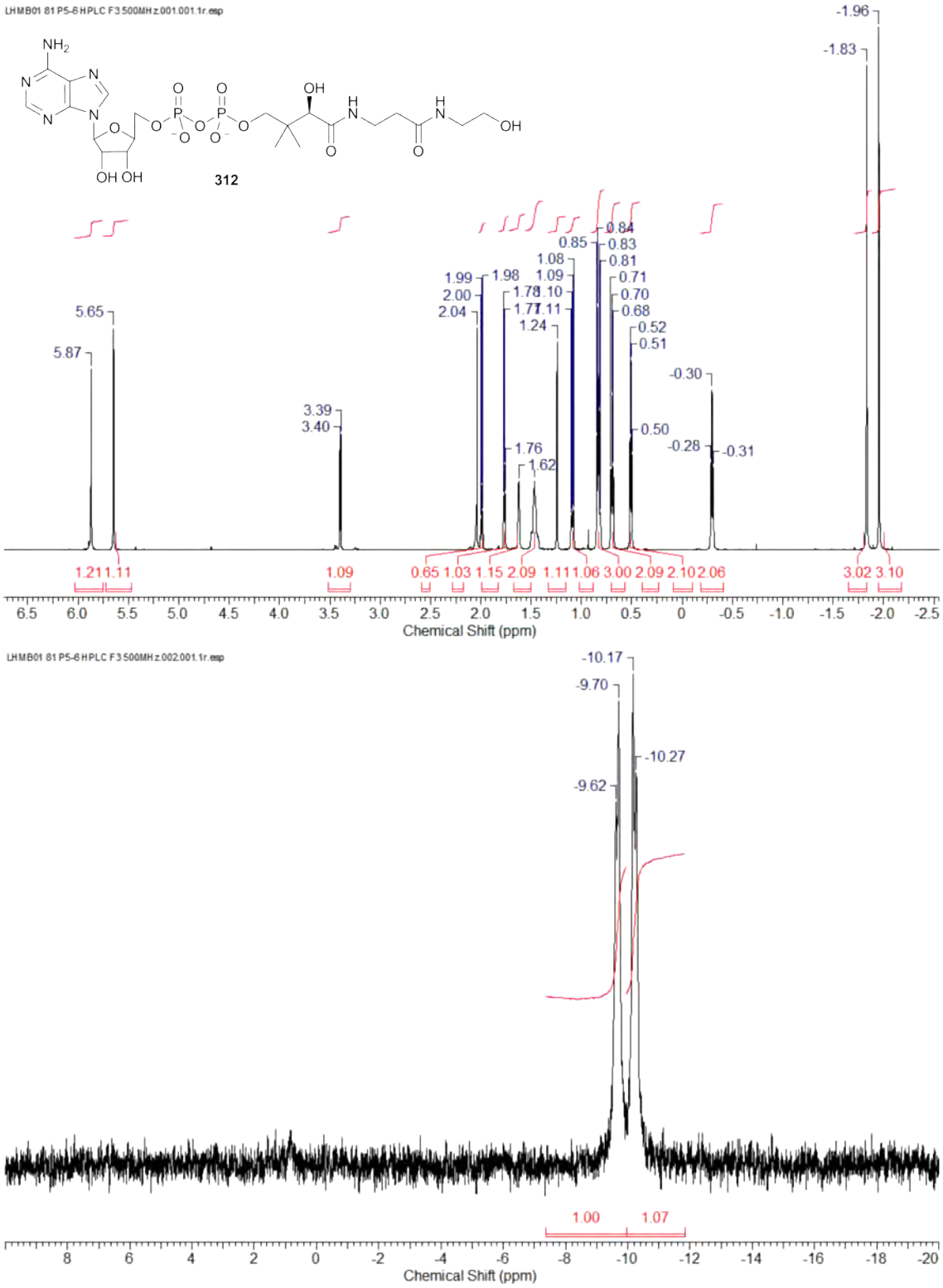


Figure D.15: ^{1}H and ^{31}P NMR spectra for dephospho-hydroxy(dethia)CoA **312**. recorded in D_2O at 500 MHz and 202 MHz, respectively.

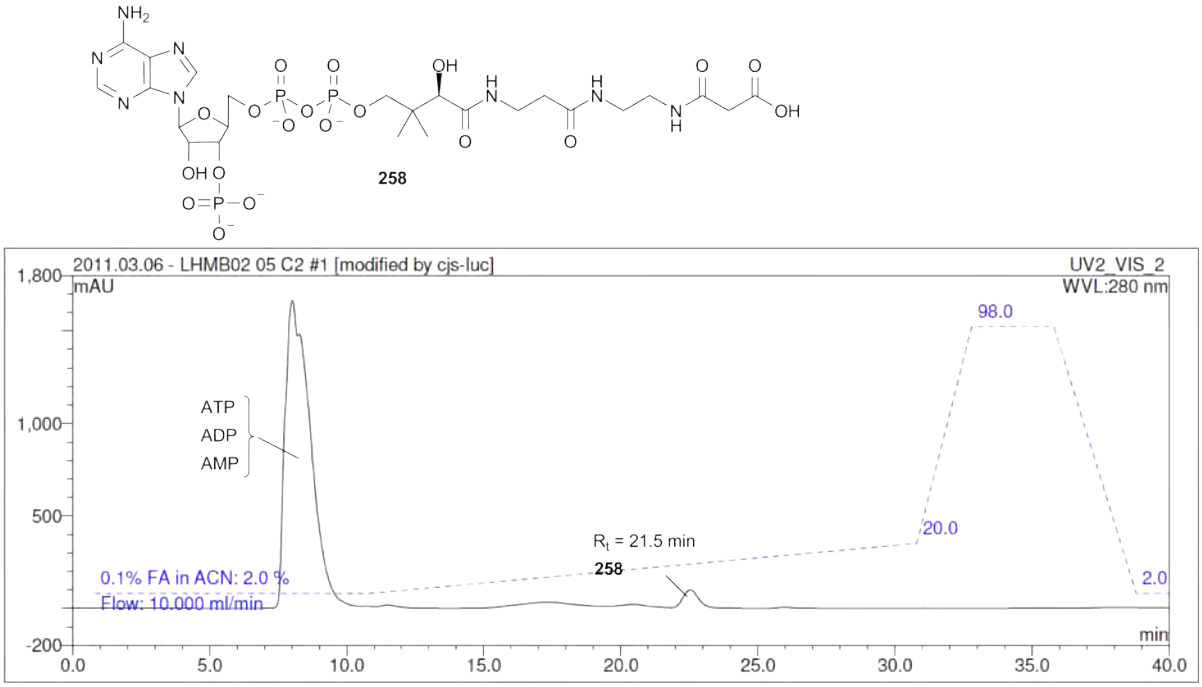


Figure D.16: Preparative HPLC trace for compound **258**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

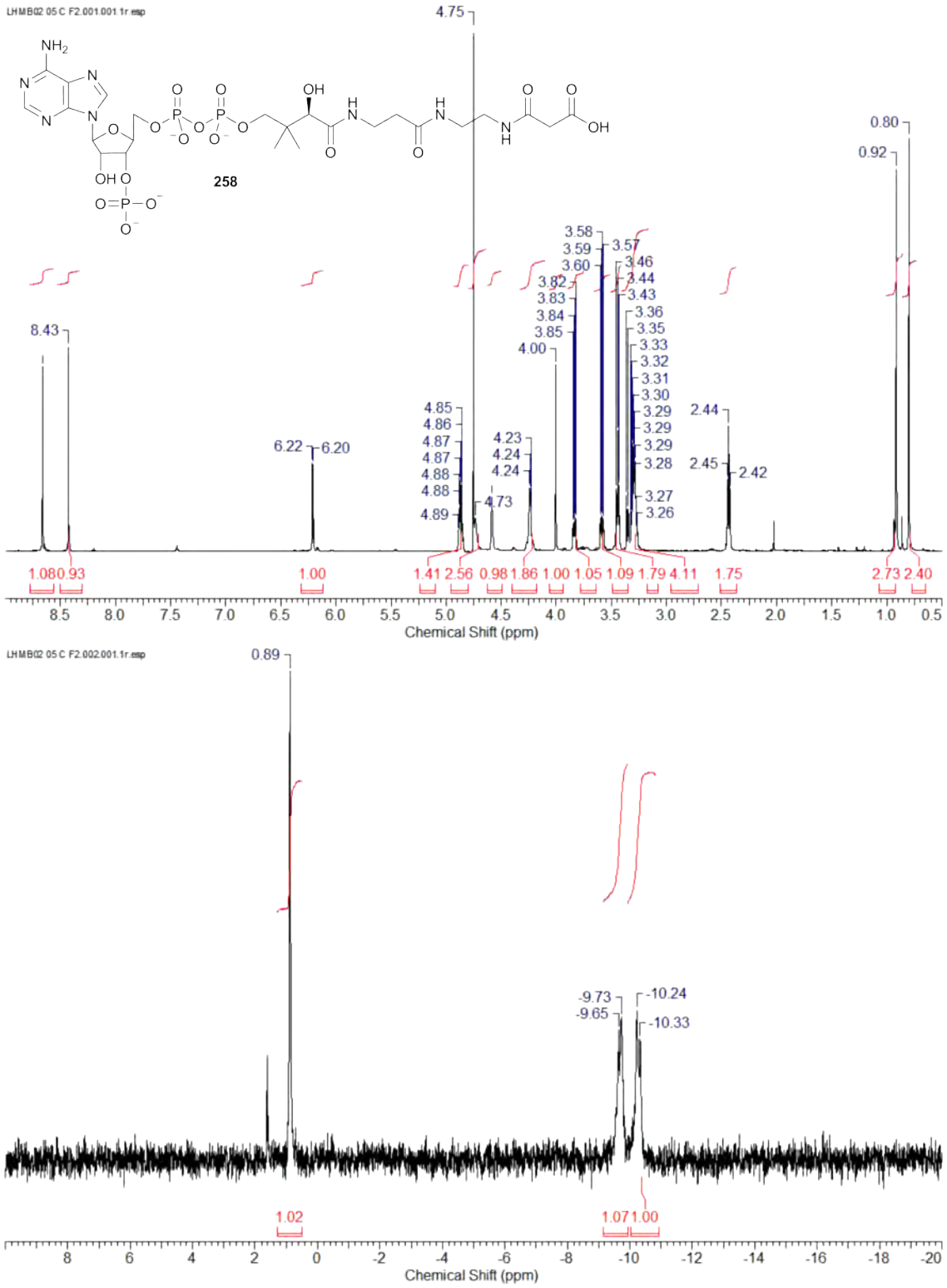


Figure D.17: ^{1}H and ^{31}P NMR spectra for compound **258**. Slow exchange of the malonate protons for deuterium was observed abolishing the singlet signal for malonyl-aza(dethia)CoA over 12 h. ^{1}H NMR and ^{31}P NMR spectra were recorded in D_2O at 500 MHz and 202 MHz, respectively.

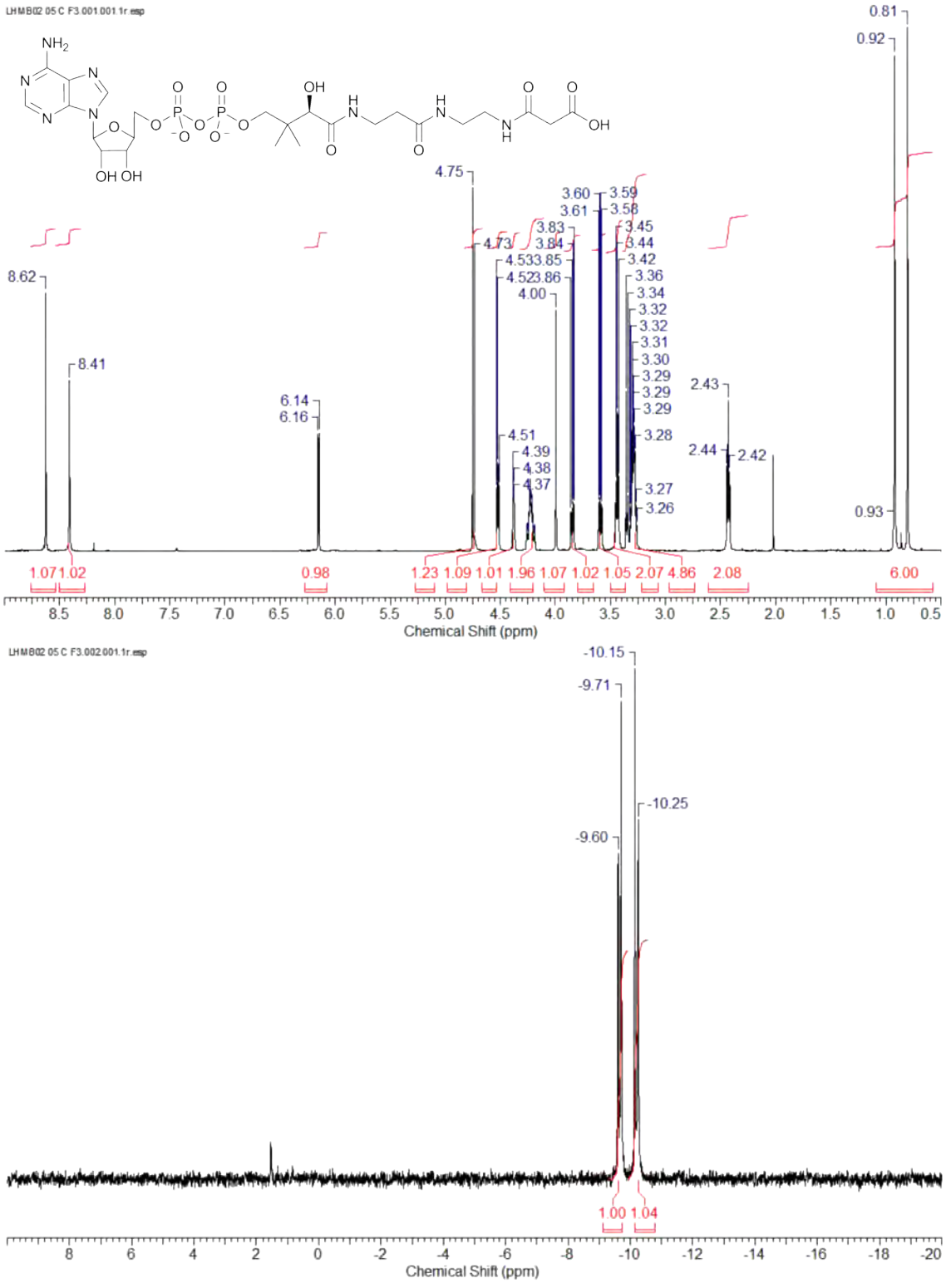


Figure D.18: ^{1}H and ^{31}P NMR spectra for compound **342** recorded in D_2O at 400 MHz and 202 MHz, respectively.

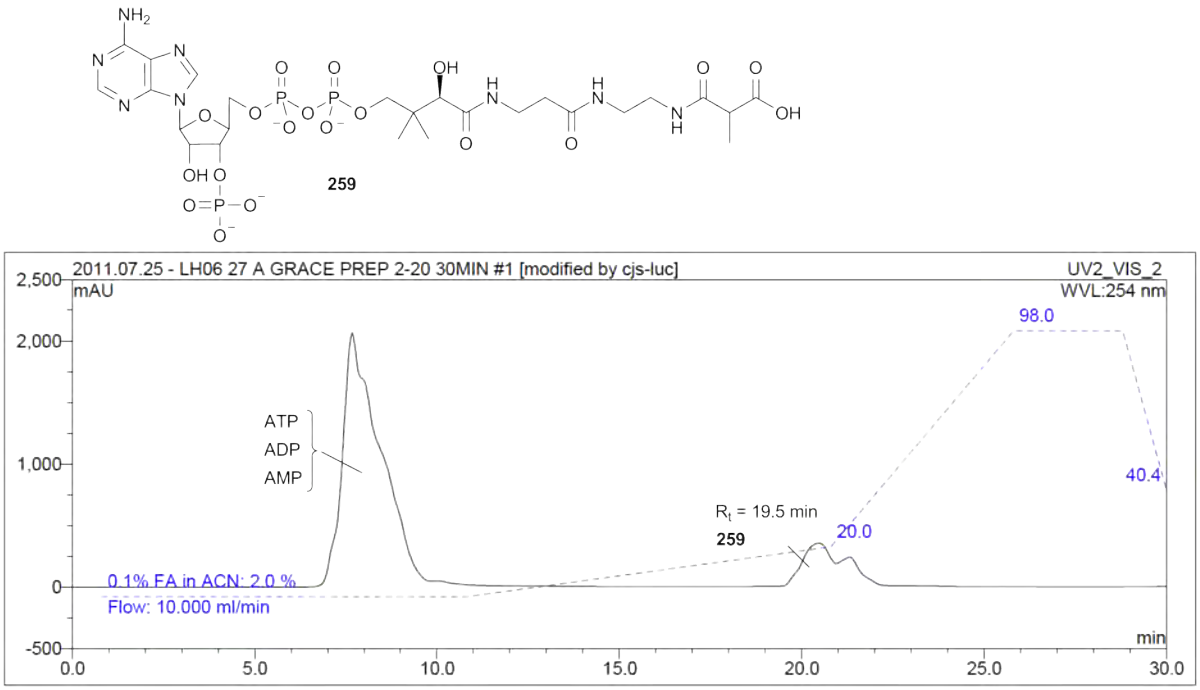


Figure D.19: Preparative HPLC trace for compound **259**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 254 nm.

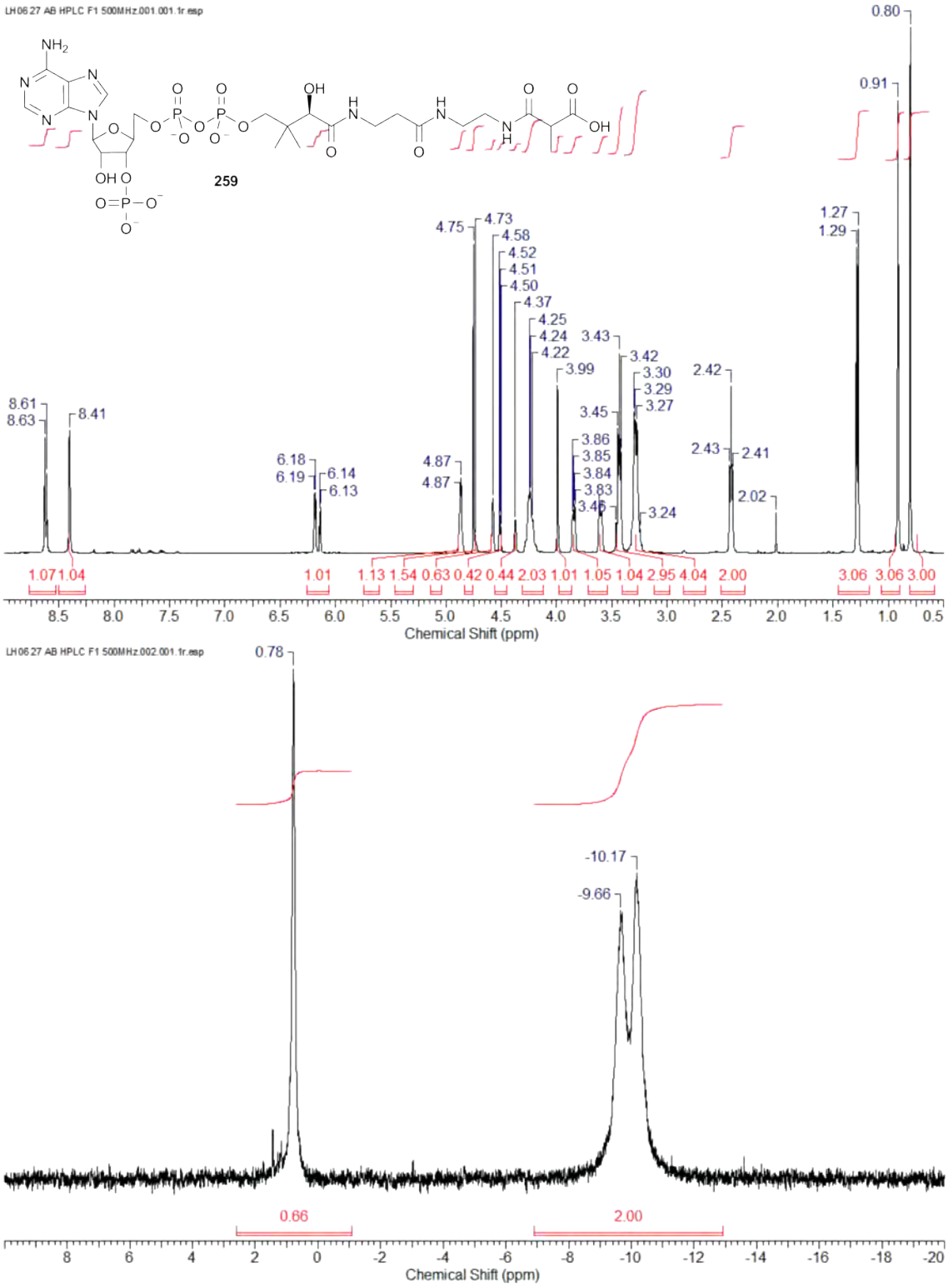


Figure D.20: ^{1}H and ^{31}P NMR spectra for compound **259**. When the ^{1}H NMR spectrum was recorded immediately after sample preparation, the malonate proton was still visible, i.e. the anticipated quartet comes at 3.43 ppm together with a triplet from a methylene CH_2. The corresponding methyl group signal at 1.28 ppm has a J = 7.5 Hz coupling constant. ^{1}H NMR and ^{31}P NMR spectra were recorded in D_2O at 500 MHz and 202 MHz, respectively.

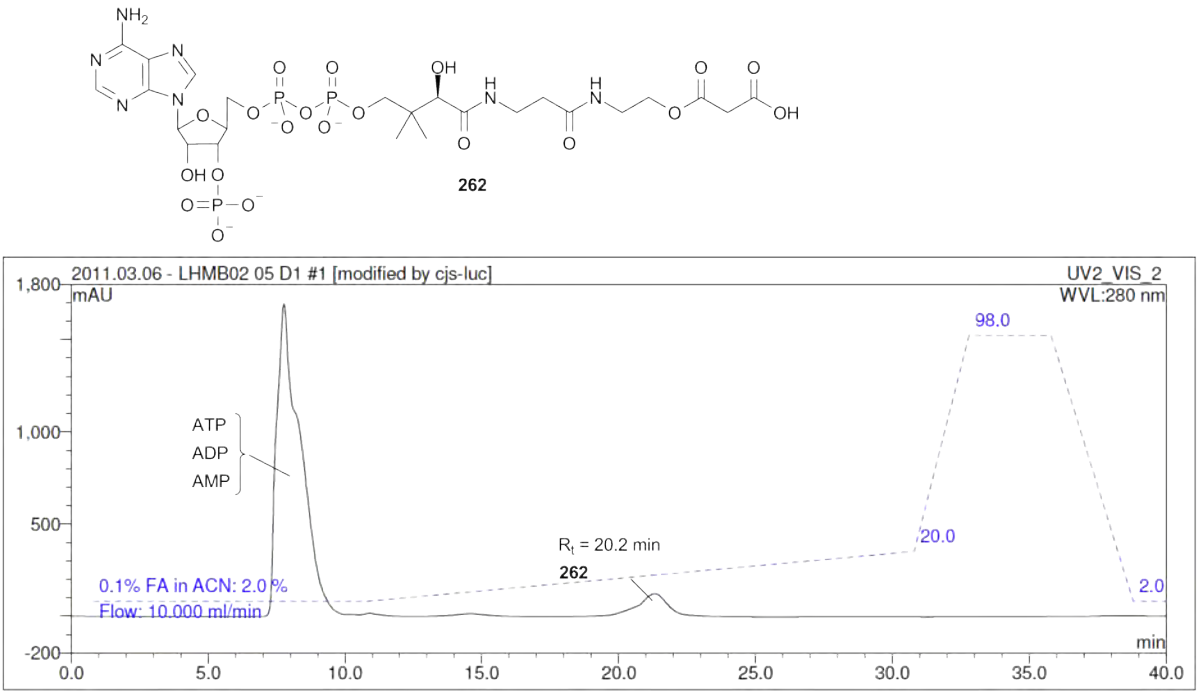


Figure D.21: Preparative HPLC trace for compound **262**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

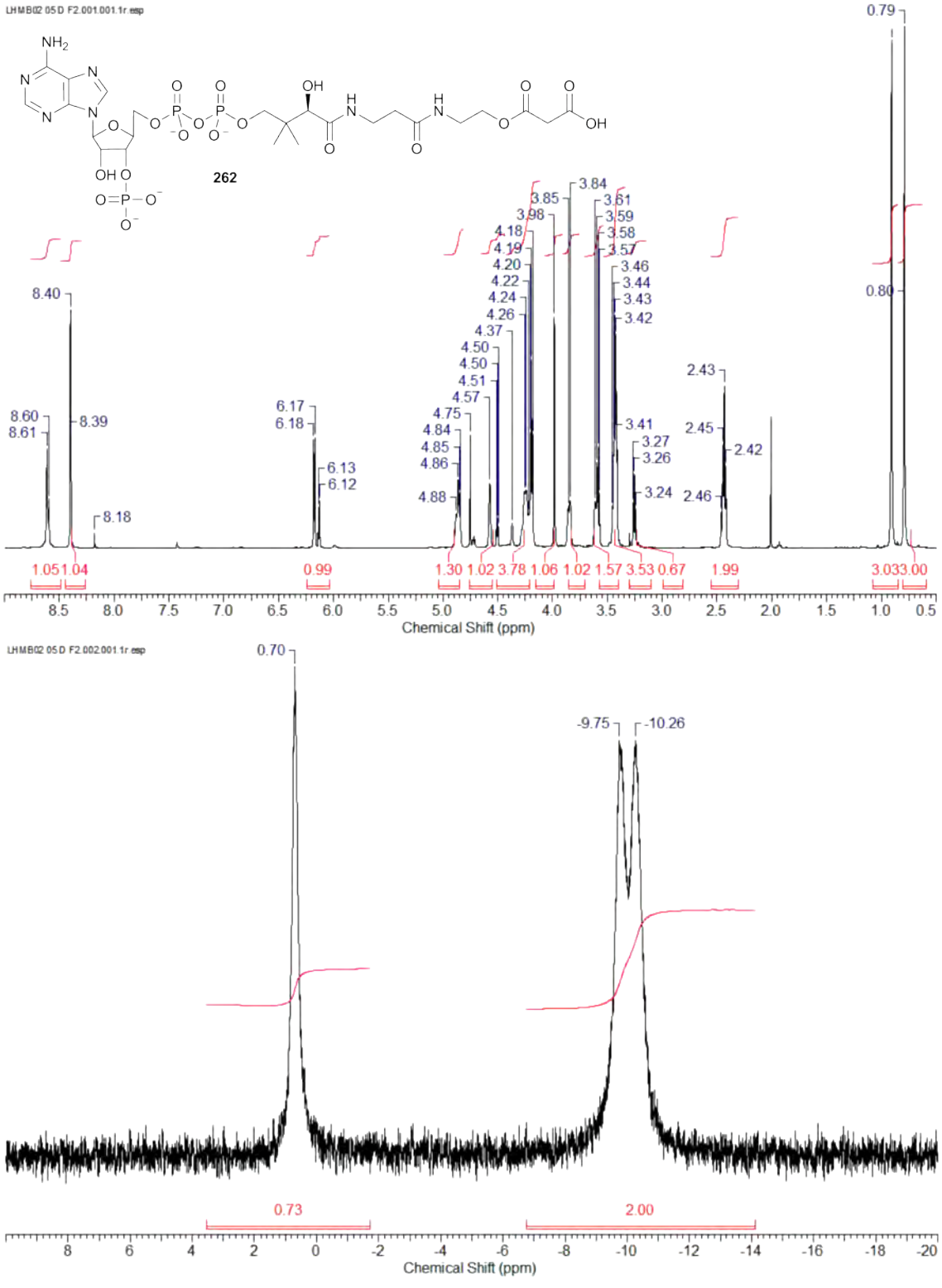


Figure D.22: ^{1}H and ^{31}P NMR spectra for compound **262** recorded in D_2O at 500 MHz and 202 MHz, respectively.

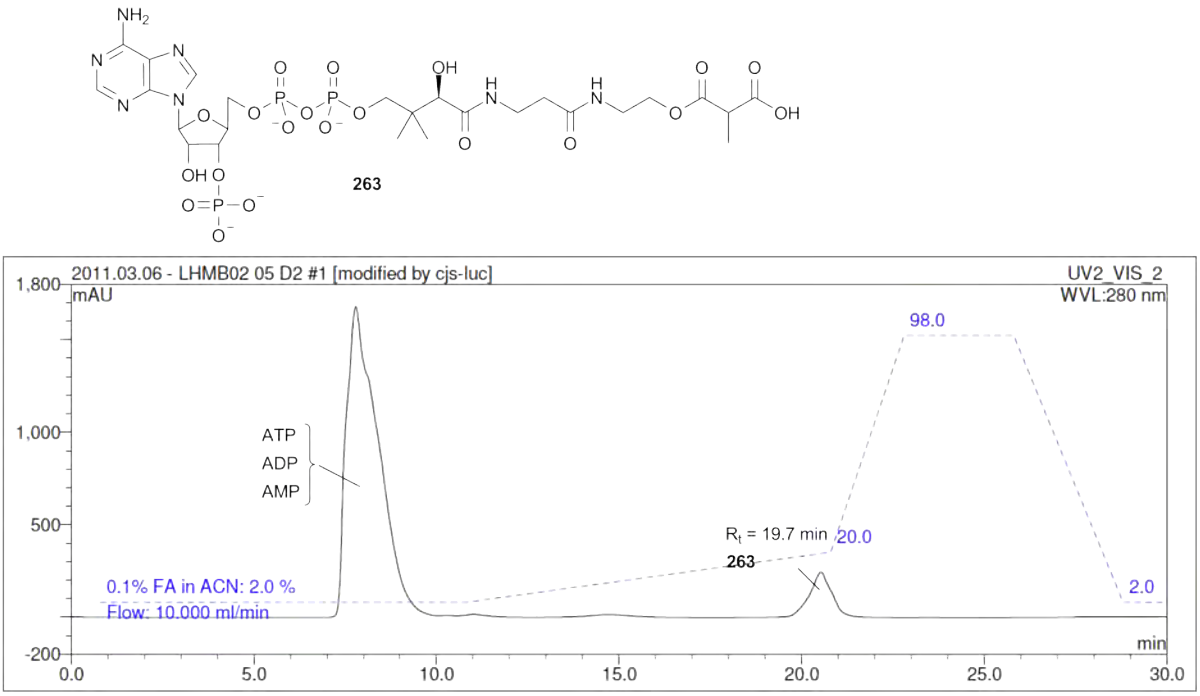


Figure D.23: Preparative HPLC trace for compound **263**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

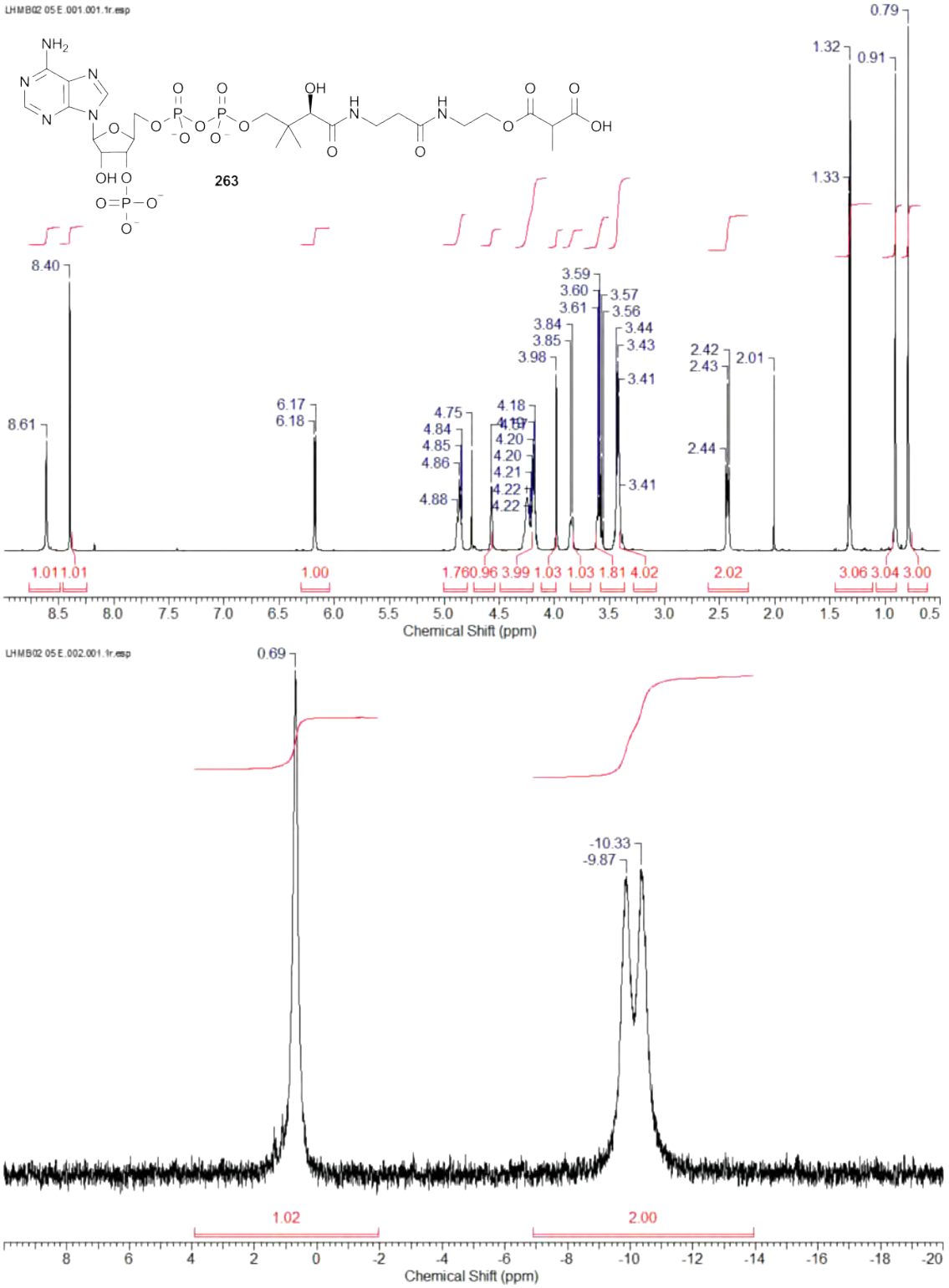


Figure D.24: ^{1}H and ^{31}P NMR spectra for compound **263** recorded in D_2O at 500 MHz and 202 MHz, respectively.

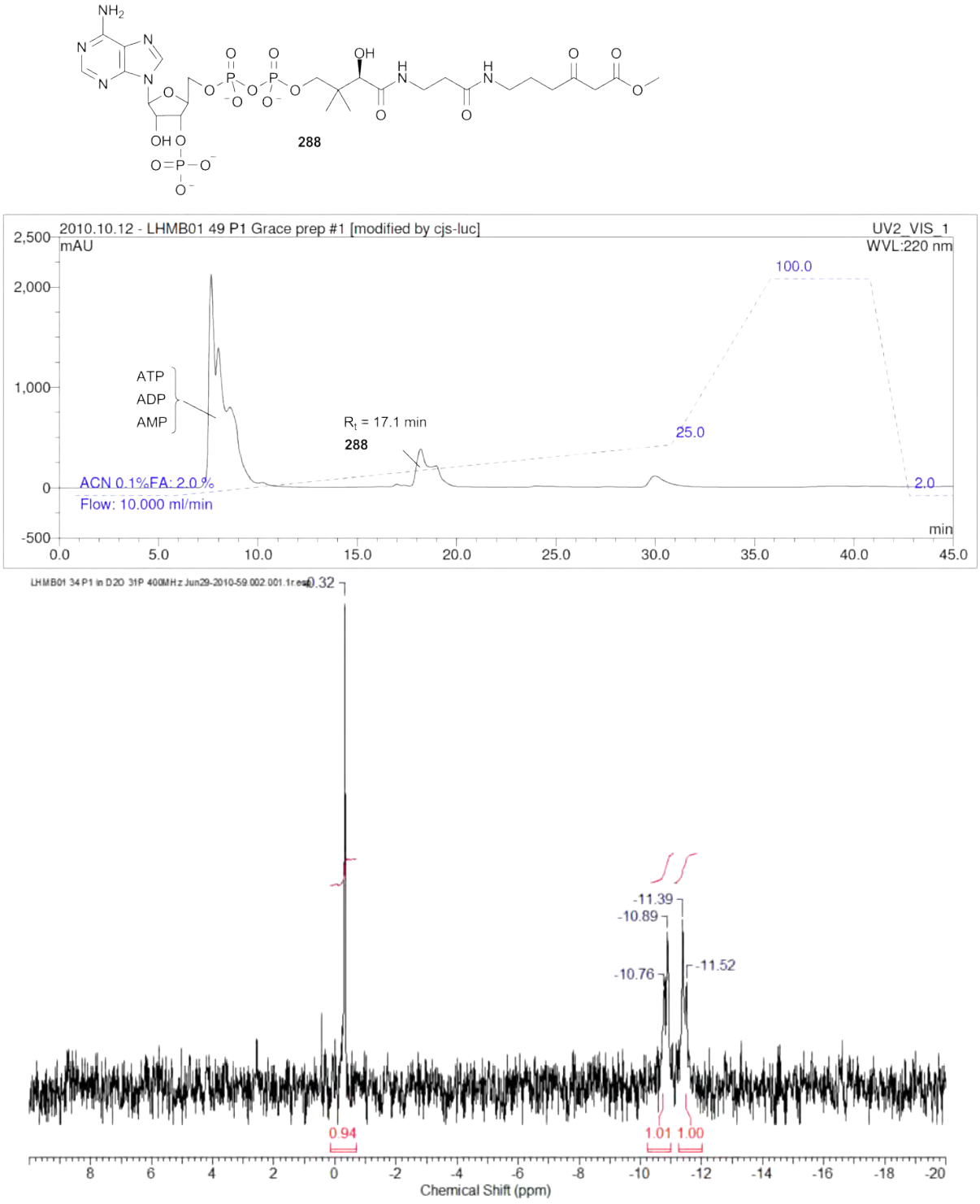


Figure D.25: Preparative HPLC trace and ^{31}P NMR spectrum for compound **288**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 220 nm. ^{31}P NMR spectrum was recorded in D_2O at 202 MHz.

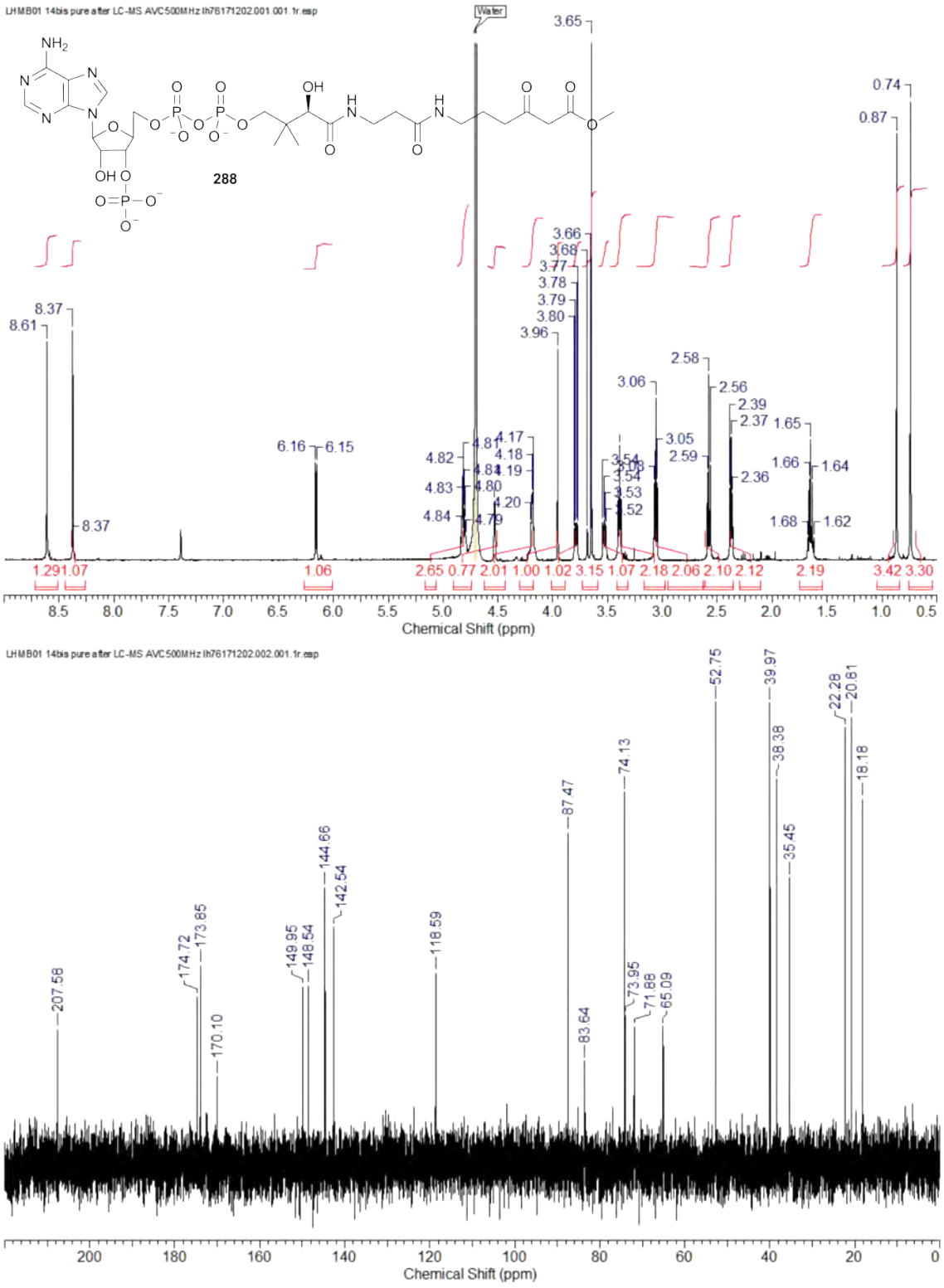


Figure D.26: ^{1}H and ^{13}C NMR spectra for compound **288**. ^{1}H NMR and ^{13}C NMR spectra were recorded in D_2O at 500 MHz and 125 MHz, respectively.

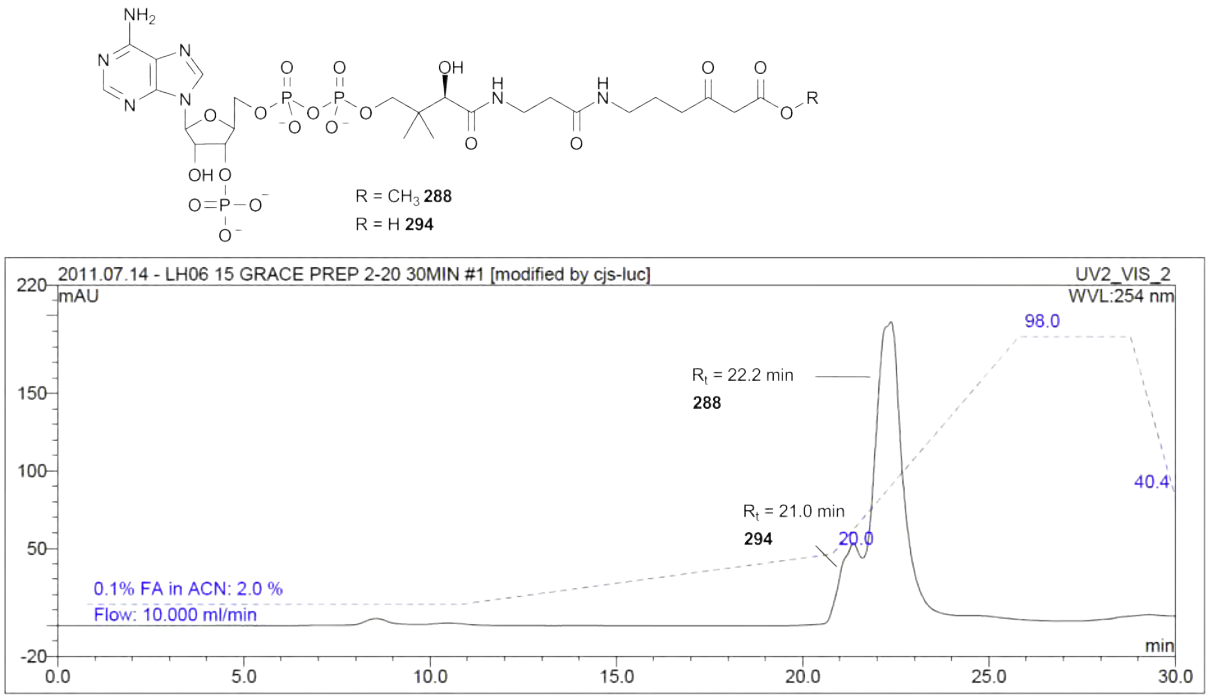


Figure D.27: HPLC trace for compound **294**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 254 nm.

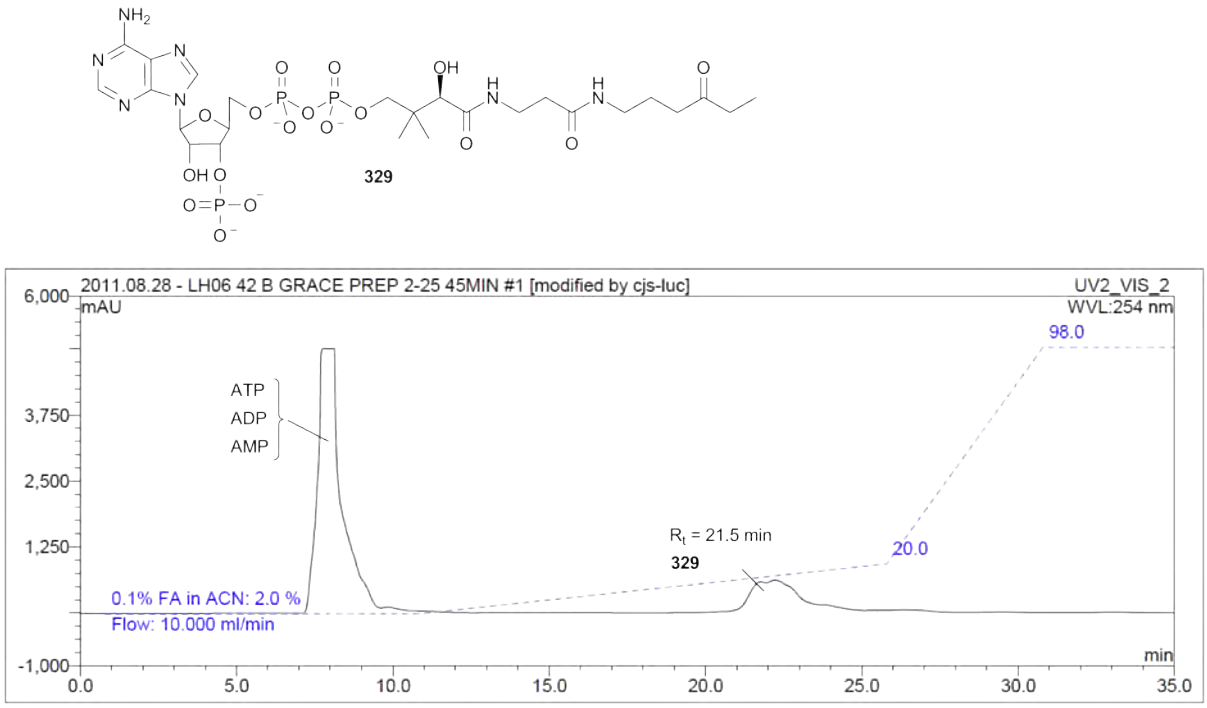


Figure D.28: HPLC trace for compound **329**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 254 nm.

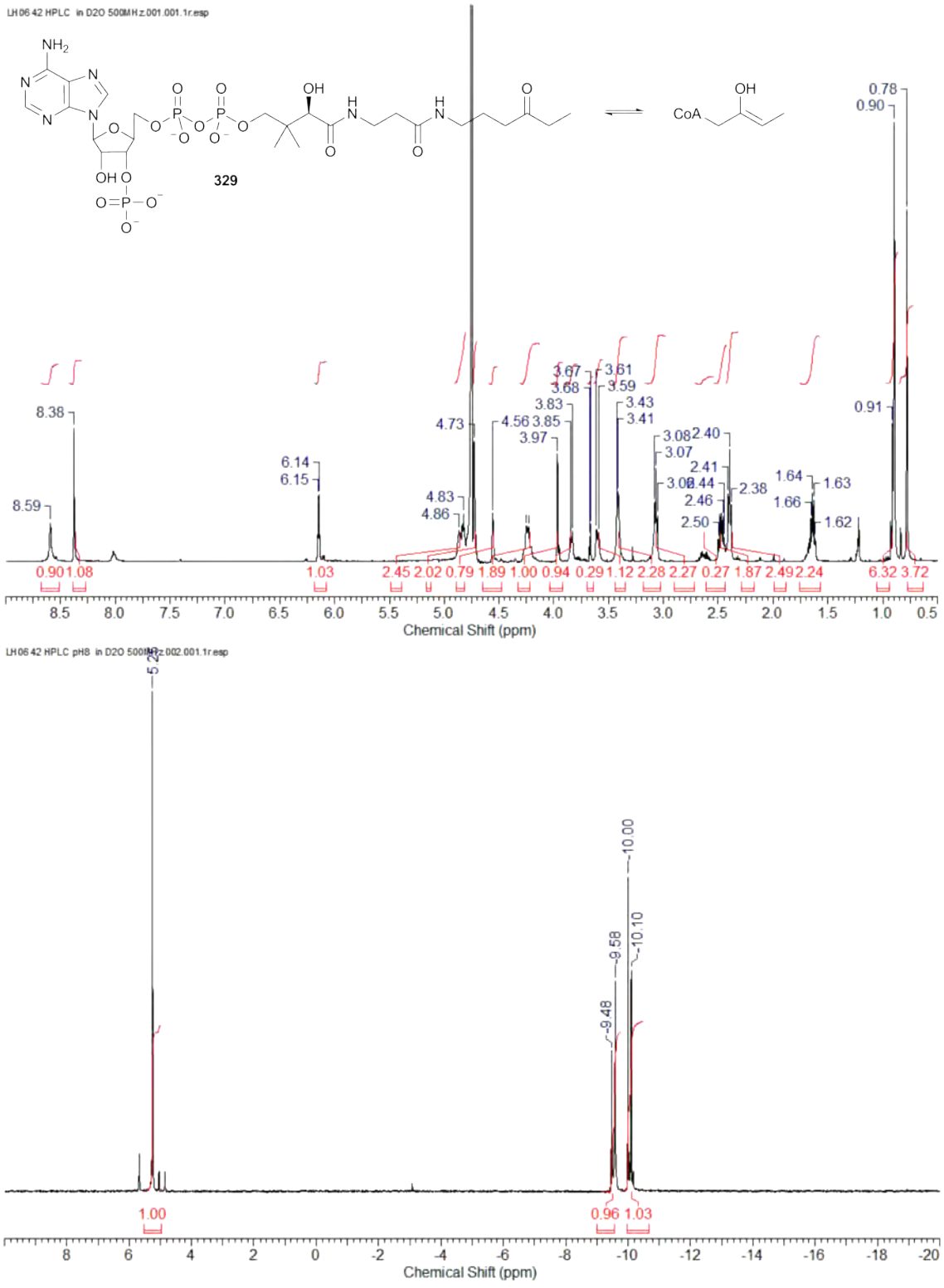


Figure D.29: ^{1}H and ^{31}P NMR spectra for compound **329** recorded in D_2O at 500 MHz and 202 MHz, respectively.

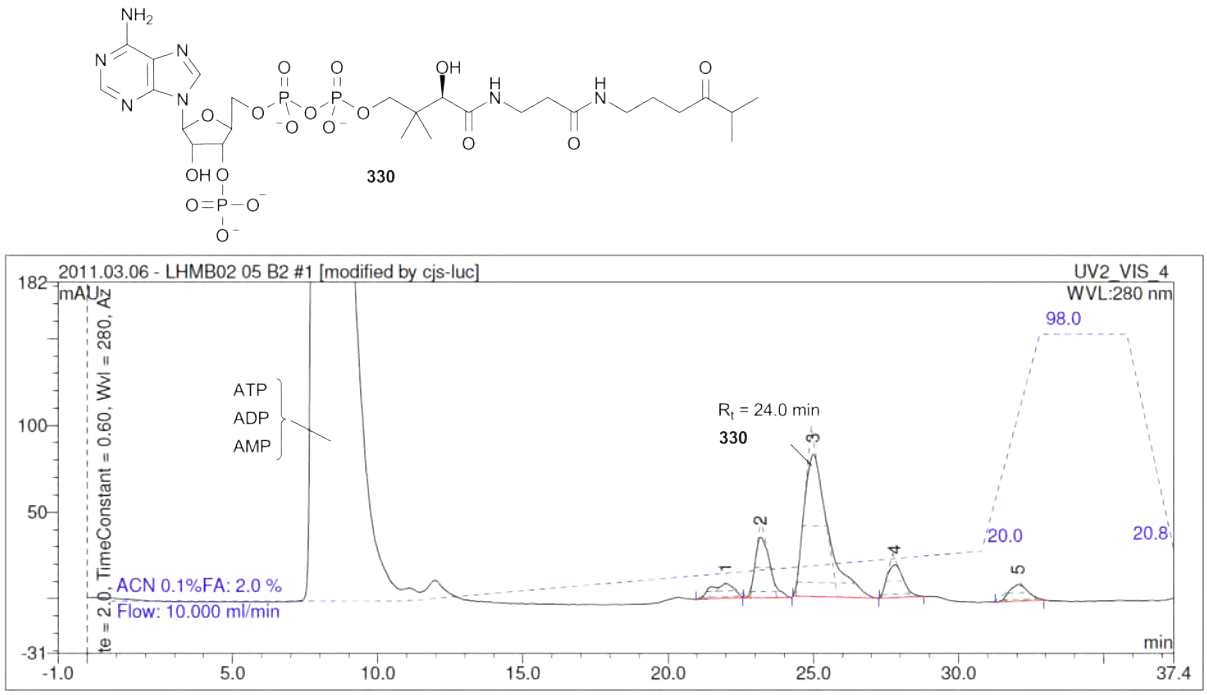


Figure D.30: HPLC trace for compound **330**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

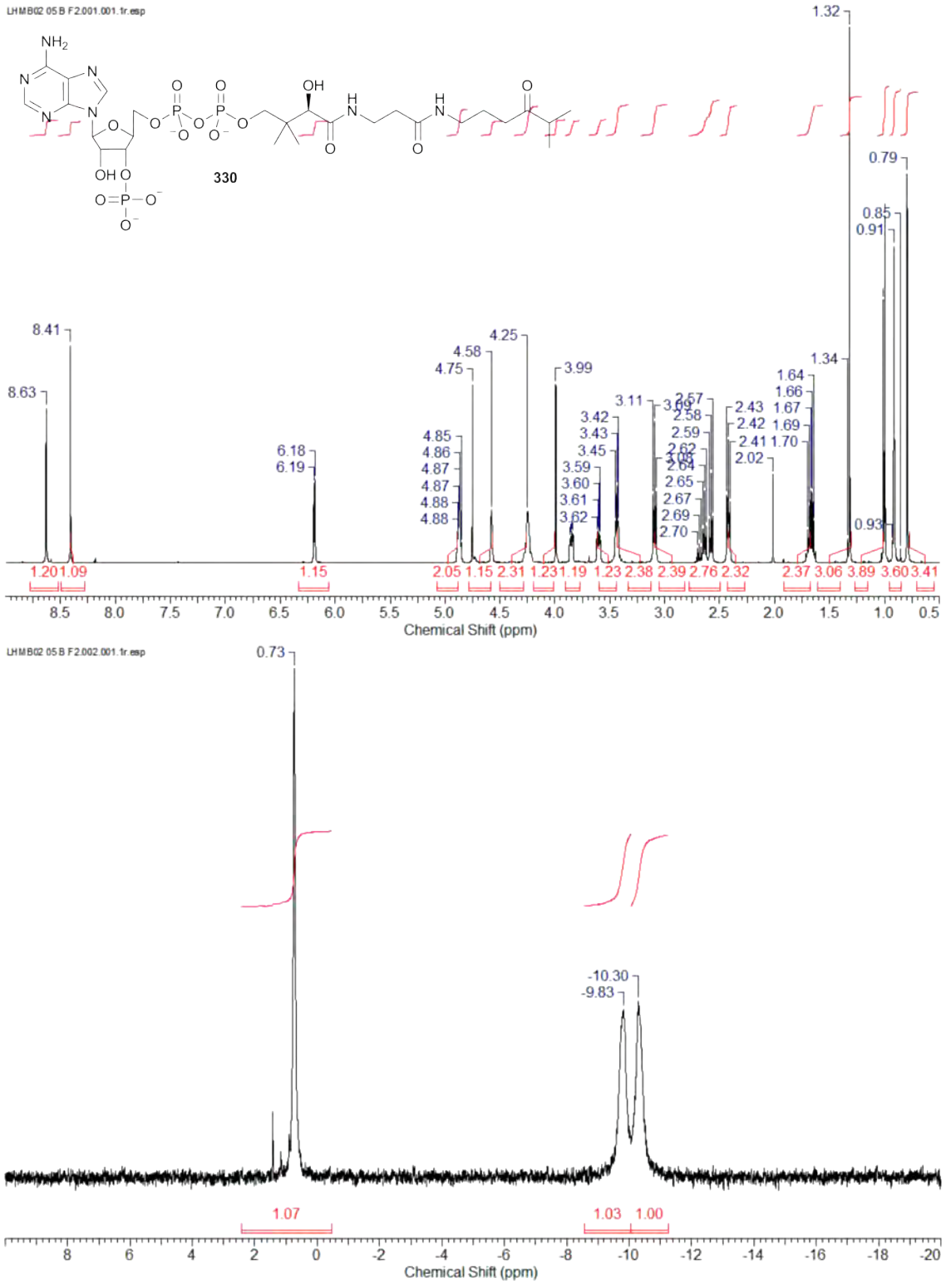


Figure D.31: ^{1}H and ^{31}P NMR spectra for compound **330** recorded in D_2O at 500 MHz and 202 MHz, respectively.

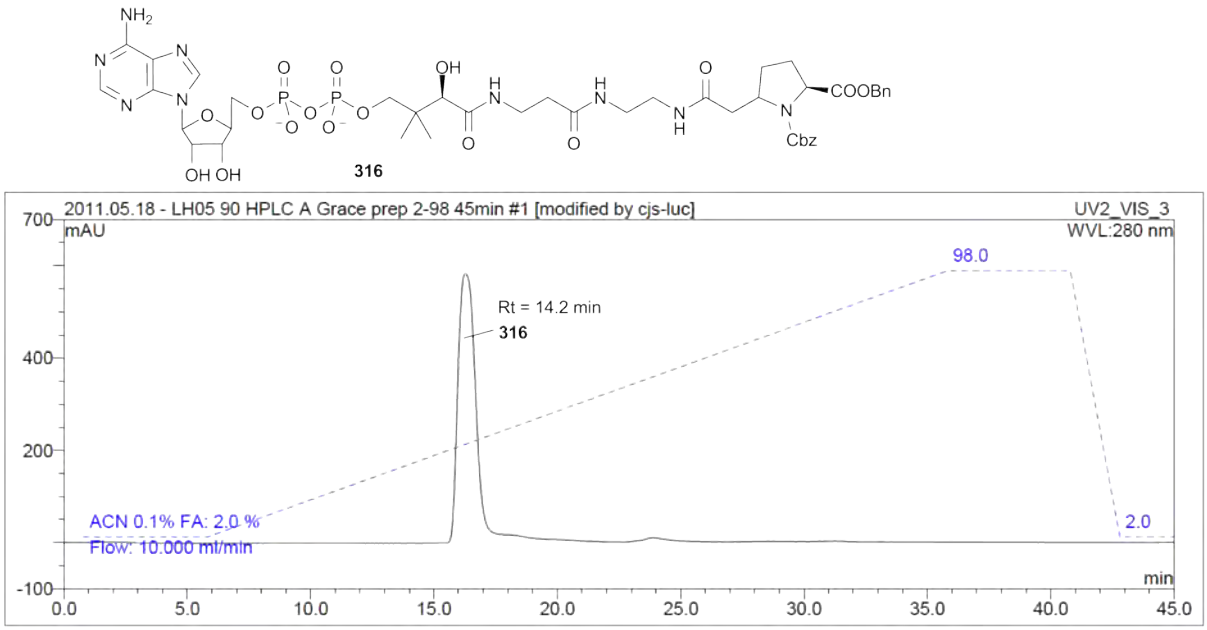


Figure D.32: HPLC trace for compound **316**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

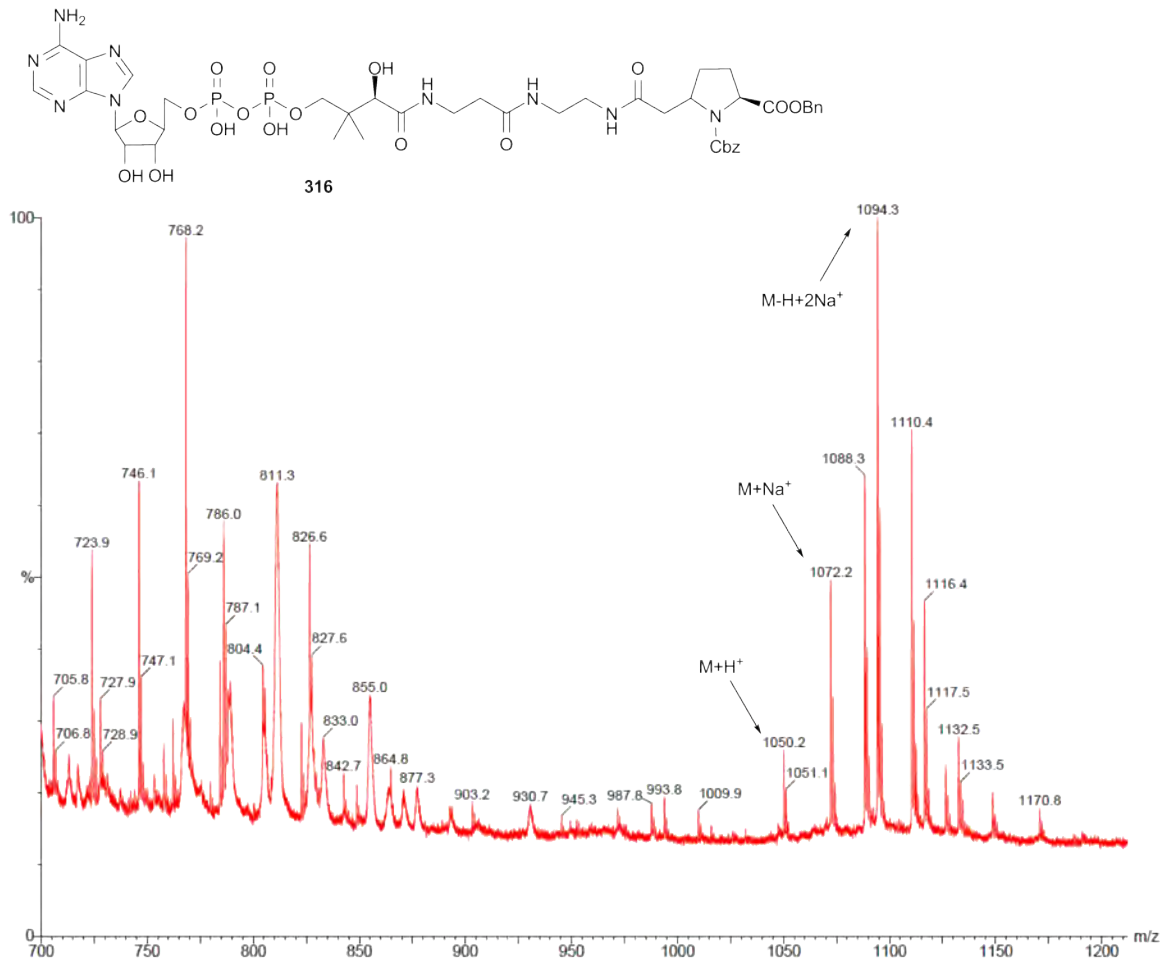


Figure D.33: MALDI-TOF MS spectrum for compound **316**.

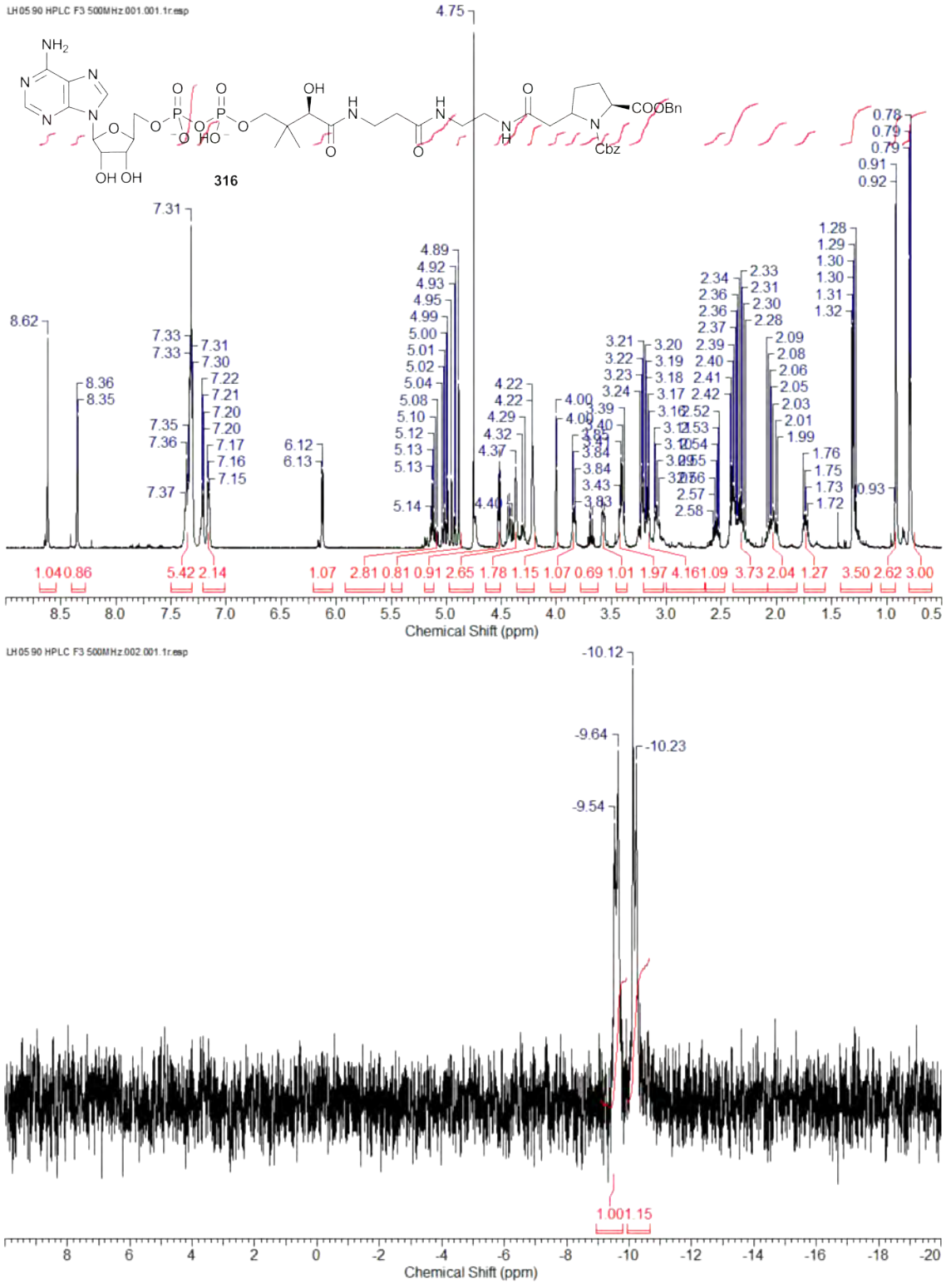


Figure D.34: ^{1}H and ^{31}P NMR spectra for compound **316** recorded in D_2O at 500 MHz and 202 MHz, respectively.

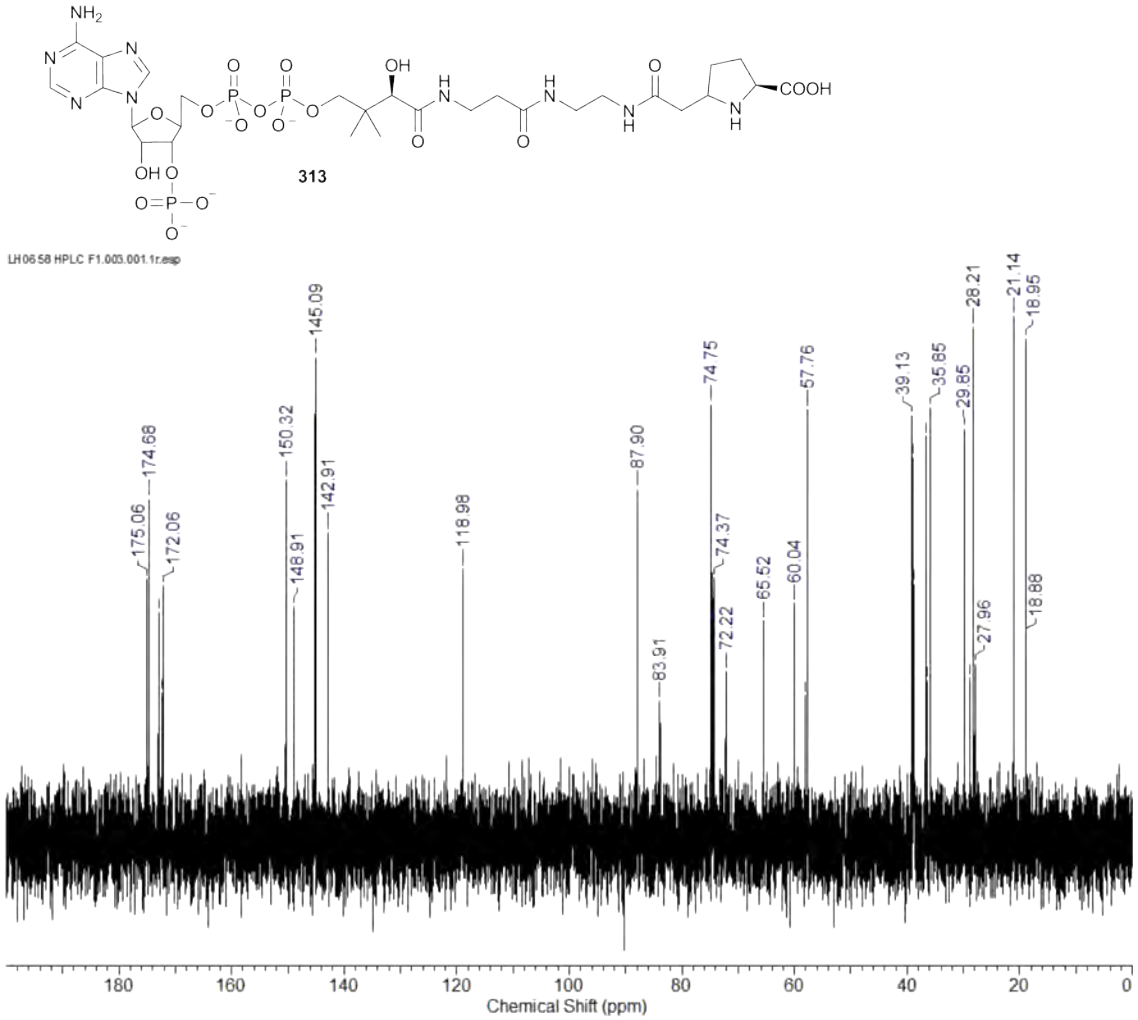


Figure D.35: Preparative HPLC trace and ^{13}C NMR spectrum for compound **313**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 220 nm. ^{13}C NMR spectrum recorded in D_2O at 125 MHz.

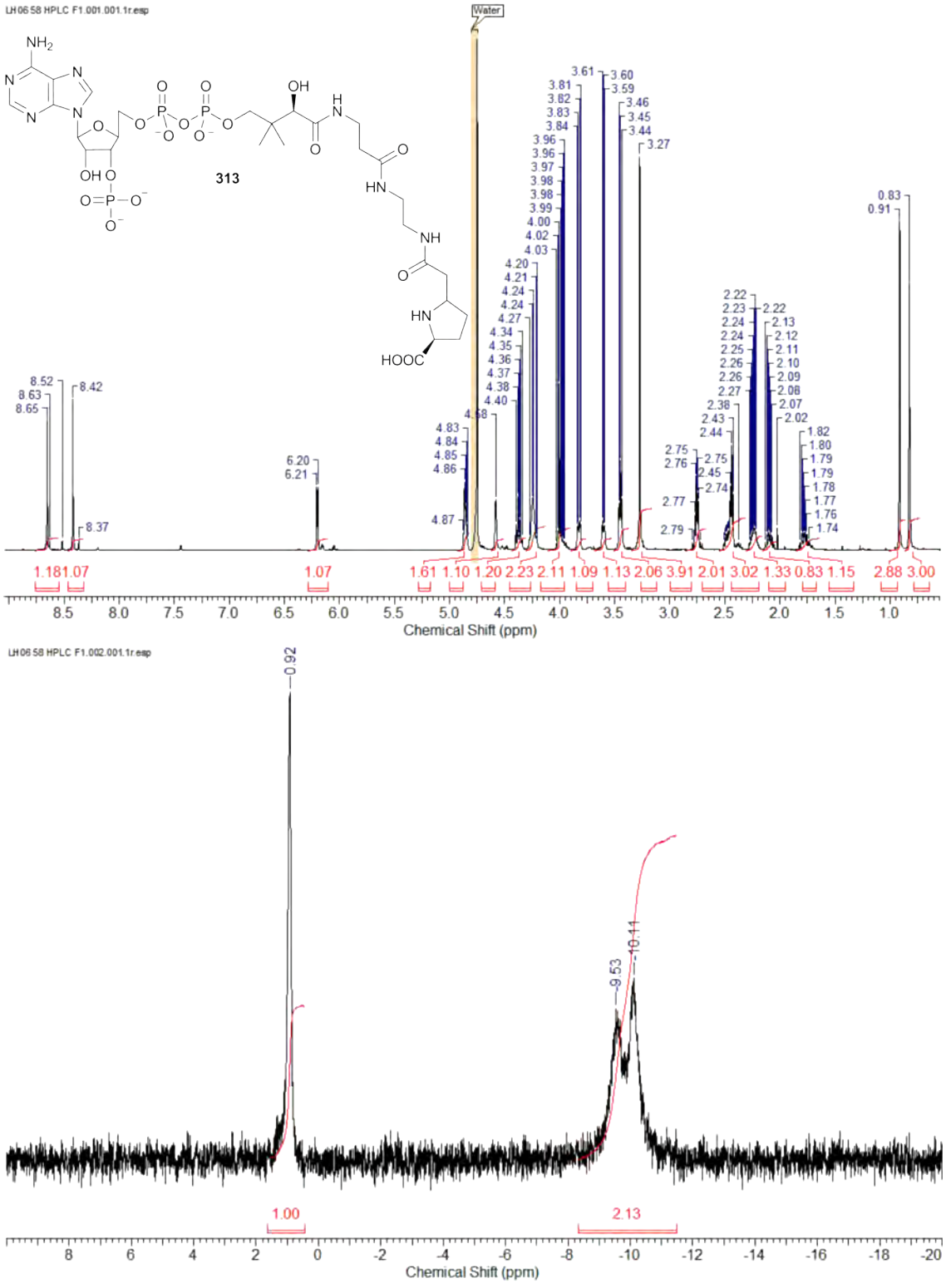


Figure D.36: ^{1}H and ^{31}P NMR spectra for compound **313** recorded in D_2O at 500 MHz and 202 MHz, respectively.

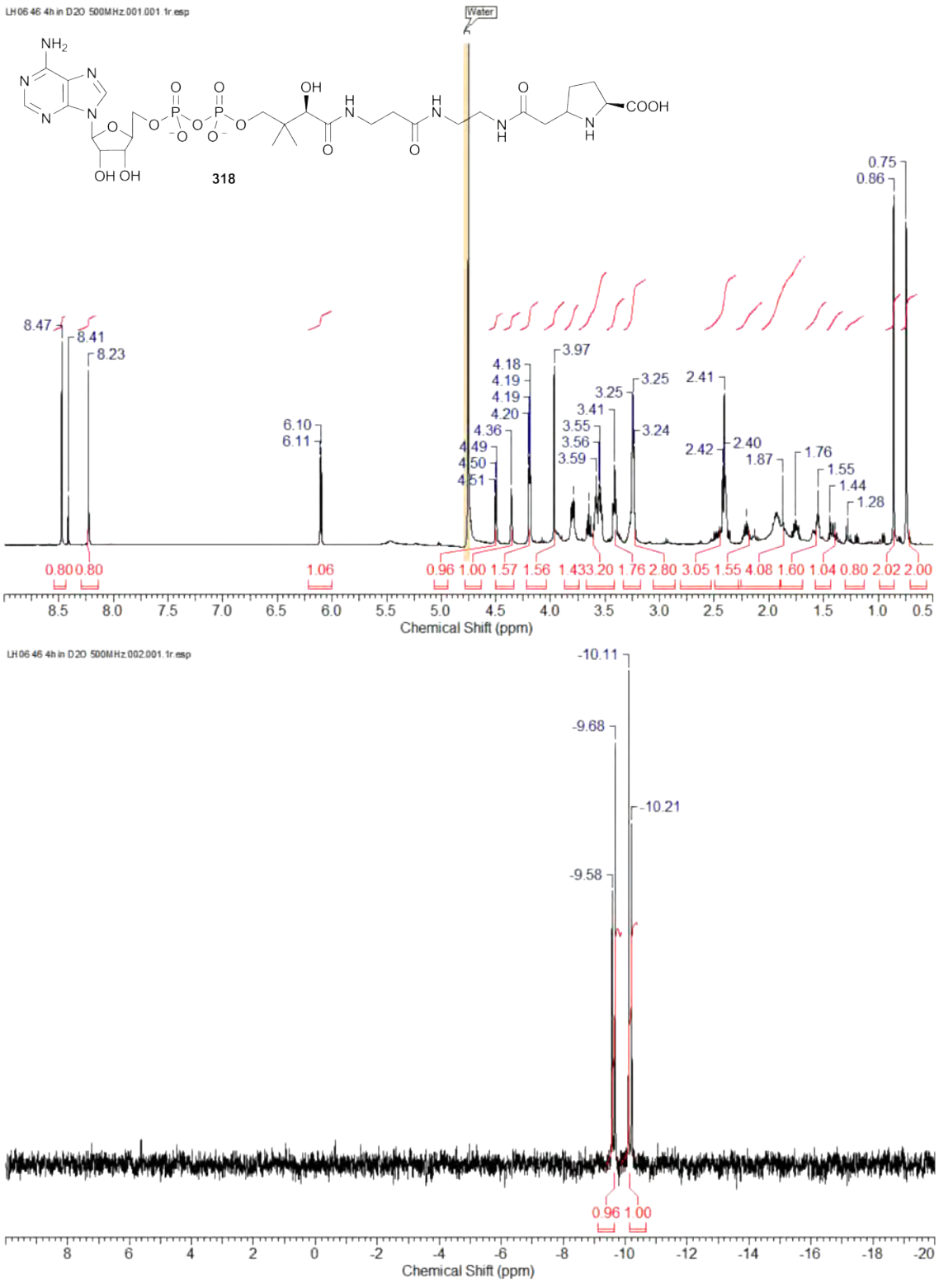


Figure D.37: ^{1}H and ^{31}P NMR spectra for compound **318** recorded in D_2O at 500 MHz and 202 MHz, respectively.

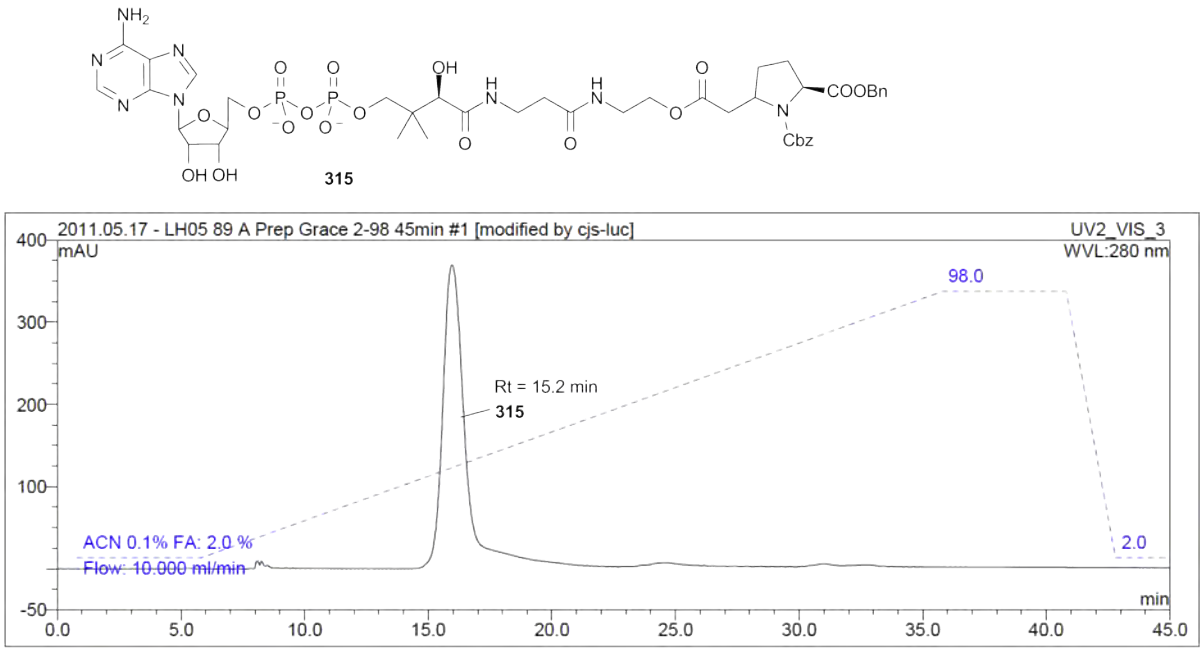


Figure D.38: HPLC trace for compound **315**. Preparative HPLC was performed using a Vydac 218TP C18 column (250 mm x 25 mm 10-15 μ). HPLC Trace of UV absorbance at λ = 280 nm.

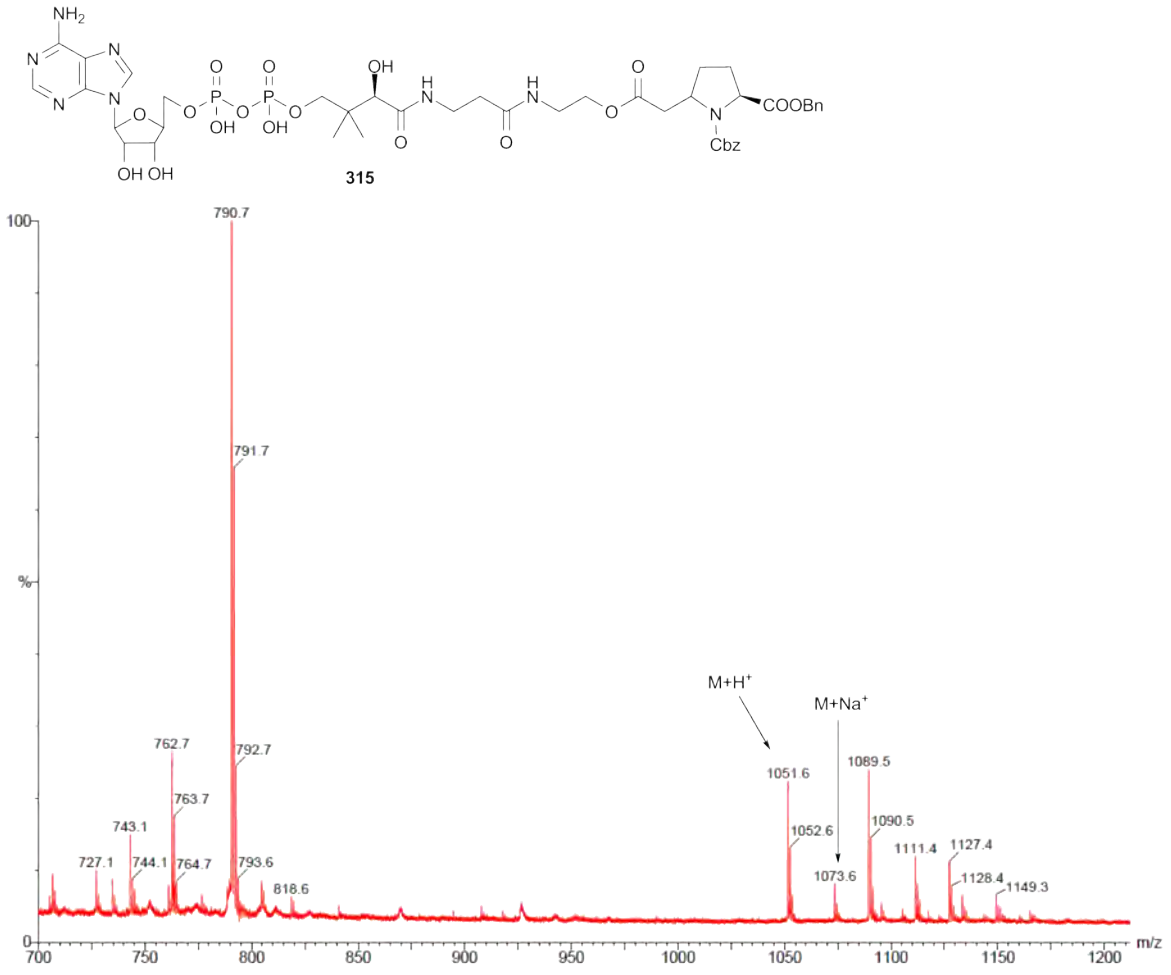


Figure D.39: MALDI-TOF MS spectrum for compound **315**.

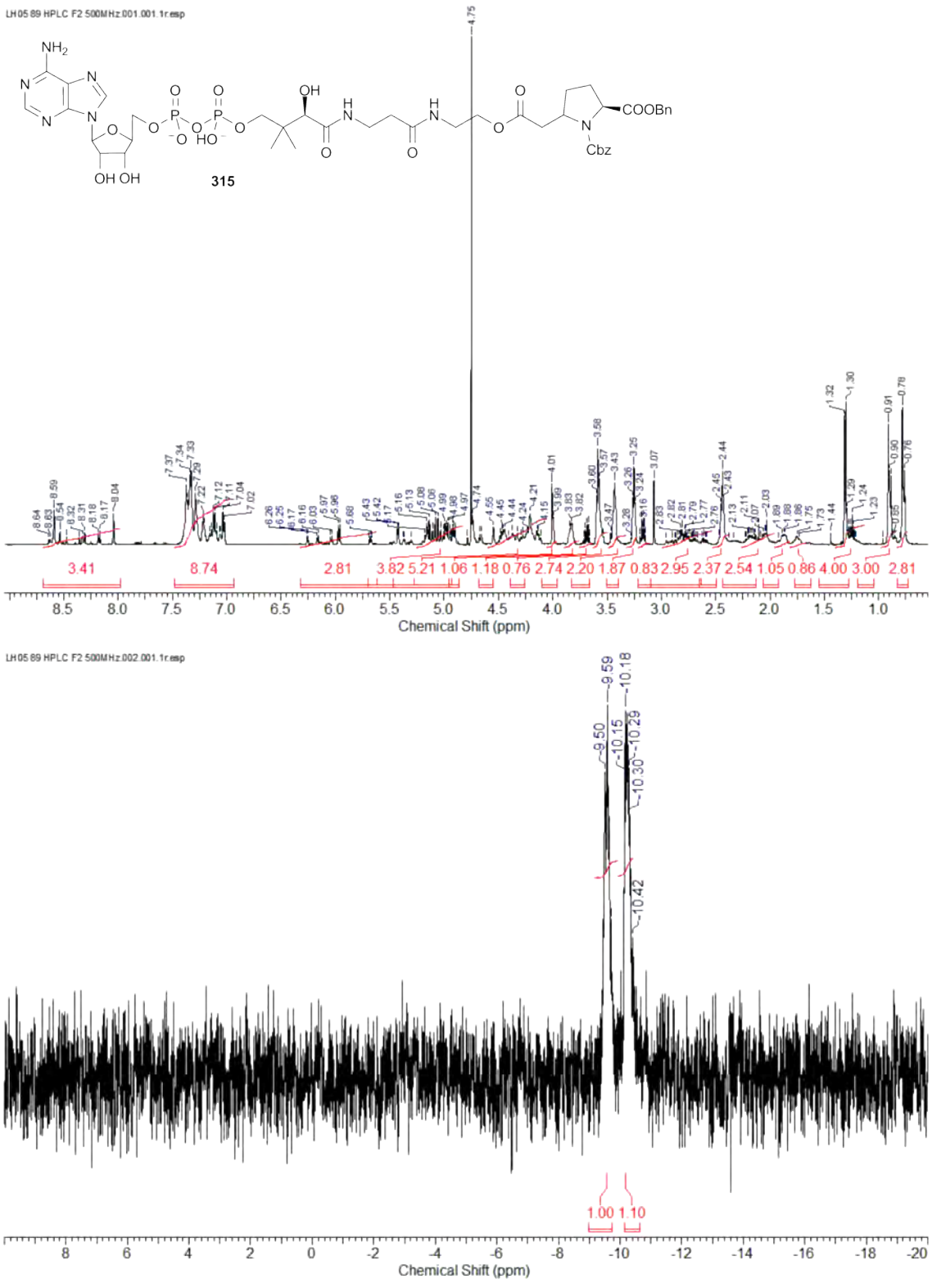


Figure D.40: ^{1}H and ^{31}P NMR spectra for compound **315** recorded in D_2O at 500 MHz and 202 MHz, respectively.

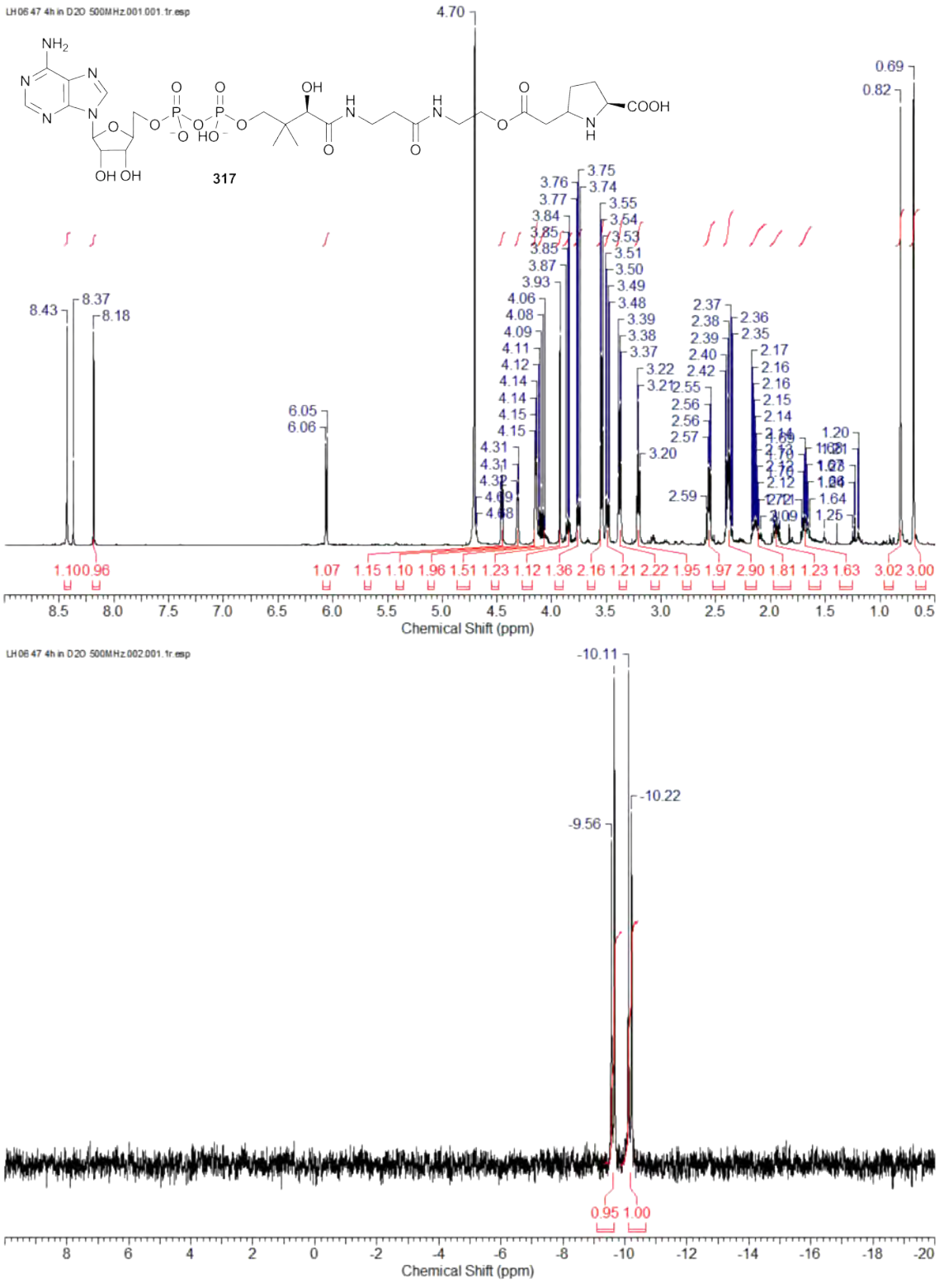


Figure D.41: ^{1}H and ^{31}P NMR spectra for compound **317** recorded in D_2O at 500 MHz and 202 MHz, respectively.

D.3 Gene and Protein Sequences of CoA Biosynthetic Enzymes

>gi|253979977|gb|ACT45647.1| CoaA [Escherichia coli BL21(DE3)]

```
  1 MSIKEQTLMT PYLQFDRNQW AALRDSVPMT LSEDEIARLK GINEDLSLEE
 51 VAEIYLPLSR LLNFYISSNL RRQAVLEQFL GTNGQRIPYI ISIAGSVAVG
101 KSTTARVLQA LLSRWPEHRR VELITTDGFL HPNQVLKERG LMKKKGFPES
151 YDMHRLVKFV SDLKSGVPNV TAPVYSHLIY DVIPDGDKTV VQPDILILEG
201 LNVLQSGMDY PHDPHHVFVS DFVDFSIYVD APEDLLQTWY INRFLKFREG
151 AFTDPDSYFH NYAKLTKEEA IKTAMTLWKE INWLNLKQNI LPTRERASLI
201 LTKSANHAVE EVRLRK
```

Table D.1: CoaA protein sequence.

>gb|CP001509.3|:4082127-4083077 CoaA [Escherichia coli BL21(DE3)]

```
   1 ttatttgcgt agtctgacct cttctaccgc atgattagca cttttcgtca ggattaaact
  61 ggcgcgctca cgagtaggta gaatattttg ctttaagttc agccagttga tctctttcca
 121 caatgtcatg gcagtcttaa tcgcttcttc tttagttaat ttcgcgtagt tatgaaaata
 181 ggaatccggg tcggtaaaag ccccttcgcg gaatttcaga aaacggttga tataccacgt
 241 ctgaagtaag tcttccggtg catcaacata tatcgaaaaa tcgacaaaat cagaaacaaa
 301 tacatgatgt ggatcgtgtg gataatccat cccgctctgt aagacattta acccttcaag
 361 aattaaaata tcaggctgaa caaccgtttt atctccatcc gggatcacat cataaataag
 421 atgcgagtaa acaggggctg taacgtttgg cacgccggat ttgagatcgg aaacaaactt
 481 caccaggcga tgcatatcat acgattccgg gaagcctttc ttcttcatca gaccacgttc
 541 tttcagaacc tgattagggt gaaggaagcc atctgtagtg atcagttcaa cacgacgatg
 601 ttccggccaa cggcttaata gcgcctgcaa tacacgggcg gttgtacttt tccccaccgc
 661 gacactgcca gcaatactga taatgtaagg aatgcgttgc ccgttagtac caagaaactg
 721 ttccagaact gcctgacggc gcagattcga gcttatatag aagttcagca aacgtgacaa
 781 aggtaaatag atctcggcaa cttcttctaa cgagagatct tcattaatac ctttgagacg
 841 ggcgatctca tcttccgata acgtcatagg tacggaatca cgcagagctg cccactggtt
 901 gcggtcaaac tgtaggtaag gcgtcattaa cgtttgctct tttatactca t
```

Table D.2: CoaA gene sequence.

>gi|253979619|gb|ACT45289.1| CoaD [Escherichia coli BL21(DE3)]

```
  1 MQKRAIYPGT FDPITNGHID IVTRATQMFD HVILAIAASP SKKPMFTLEE
 51 RVALAQQATA HLGNVEVVGF SDLMANFARN QHATVLIRGL RAVADFEYEM
101 QLAHMNRHLM PELESVFLMP SKEWSFISSS LVKEVARHQG DVTHFLPENV
151 HQALMAKLA
```

Table D.3: CoaD protein sequence.

>gb|CP001509.3|:3676615-3677094 CoaD [Escherichia coli BL21(DE3)]

```
  1 atgcaaaaac gggcgattta tccgggtact ttcgatccca ttaccaatgg tcatatcgat
 61 atcgtgacgc gcgccacgca gatgttcgat cacgttattc tggcgattgc cgccagcccc
121 agtaaaaaac cgatgtttac cctggaagag cgtgtggcac tggcacagca ggcaaccgcg
181 catctgggga acgtggaagt ggtcgggttt agtgatttga tggcgaactt cgcccgtaat
241 caacacgcta cggtgctgat tcgtggcctg cgtgcggtgg cagattttga atatgaaatg
301 cagctggcgc atatgaatcg ccacttaatg ccggaactgg aaagtgtgtt tctgatgccg
361 tcgaaagagt ggtcgtttat ctcttcatcg ttggtgaaag aggtggcgcg ccatcagggc
421 gatgtcaccc atttcctgcc ggagaatgtc catcaggcgc tgatggcgaa gttagcgtag
```

Table D.4: CoaD gene sequence.

>gi|253976334|gb|ACT42004.1| CoaE [Escherichia coli BL21(DE3)]

```
  1 MRYIVALTGG IGSGKSTVAN AFADLGINVI DADIIARQVV EPGAPALHAI
 51 ADHFGANMIA ADGTLQRRAL RERIFANPEE KNWLNALLHP LIQQETQHQI
101 QQATSPYVLW VVPLLVENSL YKKANRVLVV DVSPETQLKR TMQRDDVTRE
151 HVEQILAAQA TREARLAVAD DVIDNNGAPD AIASDVARLH AHYLQLASQF
201 VSQEKP
```

Table D.5: CoaE protein sequence.

>gb|CP001509.3|:115404-116024 CoaE [Escherichia coli BL21(DE3)]

```
  1 ttacggtttt tcctgtgaga caaactgcga cgcaagctgc aaatagtgtg cgtgcaggcg
 61 ggcaacatcc gatgcgatag catccggtgc gccgttatta tcaatgacgt catctgccac
121 ggcaaggcgg gcttcgcgcg ttgcctgagc agcaaggatt tgttcgacat gctcgcgagt
181 tacatcatcg cgctgcatgg tgcgcttaag ttgcgtttct gggctgacat ccaccaccag
241 cactcgattc gctttttttat acagtgagtt ttctaccagc aatggcacaa cccacagtac
301 atagggggaa gtagcttgct ggatctggtg ttgcgtctct tgctgaatca gcggatgcag
361 cagggcgtta agccagtttt tctcttccgg gttggcgaag atccgctcgc gcaaggcccg
421 gcgctgcaat gttccatcag cagcaatcat gttagcgcca aagtgatcag caatggcatg
481 tagcgcaggt gcacctggtt caaccacctg acgcgcaata atatcggcat caatgacgtt
541 aattccgaga tcagcaaacg cattggcaac ggtactcttg ccactgccaa tgcctcccgt
601 taaggcaact atatacctca t
```

Table D.6: CoaE gene sequence.

D.4 Sequences of CoaA, CoaD and CoaE Expression Constructs

```
  1 MGSSHHHHHH SSGLVPRGSH MTARNMLMSI KEQTLMTPY LQFDRNQWAAL
 51 RDSVPMTLSE DEIARLKGIN EDLSLEEVAE IYLPLSRLL NFYISSNLRRQ
101 AVLEQFLGTN GQRIPYIISI AGSVAVGKST TARVLQALL SRWPEHRRVEL
151 IATDGFLHPN QVLKERGLMK KKGFPESYDM HRLVKFVSD LKSGVPNVTAP
201 VYSHLIYDVI PDGDKTVVQP DILILEGLNV LQSGMDYPH DPHHVFVSDFV
251 DFSIYVDAPE DLLQTWYINR FLKFREGAFT DPDSYFHNY AKLTKEEAIKT
301 AMTLWKEINW LNLKQNILPT RERASLILTK SANHAVEEV RLRK*
```

Table D.7: CoaA expression construct protein sequence (343 a.a. MW = 39312 Da).

```
  1 MKYLLPTAAA GLLLLAAQPA MAMARYIVAL TGGIGSGKST VANAFADLGI
 51 NVIDADIIAR QVVEPGAPAL HAIADHFGAN MIAADGTLQR WALRERIFAN
101 PEEKNWLNAL LHPLIQQETQ HQIQQATSPY VLWVVPLLVE NSLYKKANRV
151 LVVDVSPETQ LKRTMQRDDV TREHVEQILA AQATREARLA VADDVIDNNG
201 APDAIASDVA RLHAHYLQLA SQFVSQEKPA LEHHHHHH*
```

Table D.8: CoaD expression construct protein sequence (238 a.a. MW = 26070 Da).

```
  1 MGSSHHHHHH SSGLVPRGSH MASMQKRAIY PGTFDPITNG HIDIVTRATQ
 51 MFDHVILAIA ASPSKKPMFT LEERVALAQQ ATAHLGNVEV VGFSDLMANF
101 ARNQHATVLI RGLRAVADFE YEMQLAHMNR HLMPELESVF LMPSKEWSFI
151 SSSLVKEVAR HQGDVTHFLP ENVHQALMAK LA*
```

Table D.9: CoaE expression construct protein sequence (182 a.a. MW = 20290 Da).

Appendix E

Supplementary Information for Chapter 7

E.1 Sequences of genes and proteins used in Chapter 7

>gi|337763151|emb|CCB71859.1| ThnN [Streptomyces cattleya NRRL 8057]

```
  1  MPSPSKLRPI TDEDVRRAVR LHFDPQDGTP YWLERDRRDG TDALRKVHGM
 51  LDAQQLIGLR DGPDQAHFEE ASRRTPLANF VPRRVLTEGG PLWAAQTGGT
101  TGPAKHGTWG ARYWADILEF SDEFLDLHGV PRGVDWLFVG PMGPHTTGRL
151  VVAFAERRGG MCFSVDLDPR IVKIFGEEGM TAAYDRYVQH IWDQVAEVAR
201  AQRIGVLFCT SRLLEMLPER LGTEALPGLR AIVHAGTTME PESHRQLRED
251  FFPGVPVVGM YGTSTTGISW QKPFEPEDDH HVVYVPCQPH IALDAVDDDG
301  NEVPFGEEGR VRVWRLTDDA LLPGFLERDR ARRVQPYGAA AERYPWPWLG
351  DPYSPEFTEGRRIEGVY
```

Table E.1: ThnN protein sequence.

>gi|337762320:918157-919260 ThnN [Streptomyces cattleya NRRL 8057]

```
   1  tcagtagacc ccctcgatcc ggcgtccctc ggtgaattcg gggctgtacg ggtcgcccag
  61  ccagggccac gggtagcgtt ccgcggccgc tccgtagggc tggacgcggc gggcccggtc
 121  gcgttcgagg aagccgggca gcagggcgtc gtcggtgagc cgccagacgc ggacccggcc
 181  ctcctcgccg aacggcacct cgttgccgtc gtcgtcgacc gcgtccaggg cgatgtgcgg
 241  ctggcagggc acgtagacca cgtggtggtc gtcctccggc tcgaacggct tctgccagct
 301  gatgccggtg gtggaggtgc cgtacatgcc gaccacgggg accccgggga agaagtcctc
 361  gcgcagttgg cggtgggatt ccggctccat ggtggtgccg gcgtggacga tggcgcgaag
 421  ccccggcagc gcttcggtgc cgaggcgttc cggcagcatc tccagcagcc gtgaggtgca
 481  gaagagcacc ccgatccgct gggcgcgggc gacttcggcg acctggtccc agatgtgctg
 541  gacgtagcgg tcataggcgg ccgtcatgcc ttcctcgccg aagatcttca cgatccgcgg
 601  gtcgaggtcg acggagaagc acatcccgcc ccggcgttcc gcgaaggcga ccaccagccg
 661  tccggtggtg tgcggtccca tcgggccgac gaagagccag tcgacgccgc gcggcacgcc
 721  gtgcaggtcg aggaactcgt cggagaactc caggatgtcg gcccagtaac gggcgcccca
 781  ggtgccgtgt ttggcgggtc cggtggtgcc gccggtctgc gcggcccaca gcggtccgcc
 841  ctcggtgagg acgcggcgcg gcacgaagtt cgccagtggc gtgcgccggg acgcctcctc
 901  gaagtgtgcc tggtccggcc cgtcgcgcaa gccgatcagt tgctgggcgt ccagcatgcc
 961  gtgcaccttg cgcagtgcgt cggtgccgtc gcgccggtcg cgctccagcc agtacggggt
1021  gccgtcctgc gggtcgaagt ggagtcgtac ggcccgtcgc acgtcctcgt cggtgatcgg
1081  acggagcttg gacggacttg gcat
```

Table E.2: ThnN gene sequence.

```
ThnN   ---------MPSPSKLRPITDEDVRRAVRLHFDPQDGTPYWLERDRRDGTDALRKVHGMLDAQQLIGLRDGPDQAHFEEA  71
ThnN*  -------MRCPSESRLRPITDEDVRRLVRLHFDPQHGTPYWLERDRRQGTDALGKVHGLADAVALIGLLDGSDQMHFEEA  73
GriC   -----------MSLIDDPEFDSHVVKLMAWHFGAETGSPFWLGKLAGLGFDPVTDVRGLADLTRFP---------DVSAE  69
EhpF   MKDYSLEIDAVMKAAQINDTNNFVQALMRWHFSKETGSPFWLGMREQLNFDPIKDVKTINDLRQFS---------DISHC  80
                          :. *   :  **. : *:*:**      . *.: .*: : *   :          ...

ThnN   SRRTPLANFVPRRVLTEGGPLWAAQTGGTTGPAKHGTWGARYWADILEFSDEFLDLHGVPRGVDWLFVGPMGPHTTGRLV  151
ThnN*  ARRTPLENFVPRSVLAEGGTLWAAQTGGTTGPAKHGTWGSRYWDDILAFSDEFLDLHGVPRGKNWLFVGPMGPHTTGRLV  153
GriC   LRSVPAQDLIP-RGLS-DRPFRVYDSGGTTGSPKRIVDS-GYRSLVTEWARERLIANGVPADGNWLHLGPGGPHVIGFDT  137
EhpF   LRQEPVANLVP-QGLPADSHPQVYESGGTTGAPKYVVAYDAWIEALISWRMSGYQHRPGRPSGNTLAAIPTGPHIVGAIN  150
        *  *  :::*   *. .    . ::******..*  .    :   :  :  .    .    . : *   * ***  *

ThnN   VAFAERRGGMCFSVDLDPRIVKIFGEEGMTAAYDRYVQHIWDQVAEVARAQRIGVLFCTSRLLEMLPER--LGTEALPGL  229
ThnN*  VAFAERRGGMCFSIDLDPRIVKIFGEEGMTAAYDRYVDHIWAQIAEVARSQDIGVLFCTSRLLEMLPER--LGTEALAGV  231
GriC   ARYAALGGGVFYTVDLDPRWVKRLLSQGRGEESEAYVQHLLDQAETILESQEIAVLNTTPPLLEAICARPRLYELVRSGV  217
EhpF   KERALRLGGMFFSIDIDPRWVKRSLSEGDTATVRKYTHHLVDQVQNTLMNQDIRFLVTTPPVLRELLKRPEVVLQMKQSL  230
          *   **: :::*:*** **   .:*       *..*:  *       * * .*  *. :*. :  *  :

ThnN   RAIVHAGTTMEPESHRQLREDFFPGVPVVGMYGTSTTGISWQKPFEPEDDHHVVYVPCQPHIALDAVD-DDGNEVPFGEE  309
ThnN*  EAIVHAGTTLDPETHRQLREDFFPGVPVVGMYGTSTTGISWQKPFEAEDEHRVVYVPCQPHVSLDVVD-GQGDDVPFGEE  311
GriC   RAIIWAGTSISRETLTQLDDVFFPAASVVGVYGNSLMGVAPQRSGRPGDAHRCVFEPYPETTRLHLVD-EDGTPVAYGER  297
EhpF   AQITLGGTELNLDEIKFIASEILPDCEFSASYGSTSALGVSRSLLITSESQQVIYDSFSPFITYDVVDSITAQTVEYGER  310
         *  .** :. :    : . ::*   . . **.:     :    . : :: :: .      . **   .  * :**.

ThnN   GRVRVWRLTDDALLPGFLERDRARRVQPYGAAAERYPWPWLGDPYSPEFTEGR-RIEGVY                    367
ThnN*  GRVRVWRLTDDALLPGFWERDRARRVRPYGVAAERYPWPWVGDPYSPEFTEGR-RVEGVY                    369
GriC   GRVRLHLVSEEMFLPNVLERDTAIRVEPG----EGSTVDGLADVQTYRSLDDVAIIEGVY                    352
EhpF   GNVIVTHLSPWAFYPRVAERDTAIRLPGV----SGFAGDRLADIEPLKISEGRKVIEGVY                    366
       *.* :  ::   : * . *** * *:       .    . :.*  . .  :.   :****
```

Sequence alignment for *Streptomyces cattleya* ThnN (gi|357408202), *Streptomyces flavogriseus* ATCC 33331 ThnN* (gi|357409404), *Streptomyces griseus* GriC (gi|182438036) and *Pantoea agglomerans* EhpF (gi|24021017) generated using Clustal W [1].

Table E.3: Sequence alignment between ThnN and other predicted ANL superfamily members.

>gi|337763152|emb|CCB71860.1| ThnO [Streptomyces cattleya NRRL 8057]

```
  1 MTTVTTGHAG PQRAGAVLDV PAYAGRRPVM SSRPAALTGT TGETVARVGAA
 51 GRLVQFQMAQ EAAAVARRLR ATDDDAWFAL LRRAADTLTA RVADGTAQPWL
101 TALSAASGLP PRRAARGIET VAADLARMDE ILAAQSPDGT VAAYRTGGHHP
151 RWSWRPAGRS VLVKVADNFP TINIEWLQAL AARRPVLLST SRHDPFTPVLL
201 TEALYEAGLP EHAVSVVHGD APALRRLADQ VLWPGEDVPA DLPPGKAKTYH
251 FGRSRAVVGA GAADDAWPRL ARAAFAGCGR LCTNVSSVIA LGDARQAADRL
301 AEEFATRPVL PLDDPAATVP AFPDRARRDA LATRIEREIA AGAVDVTEAVT
351 GVPLRVEVDG AAFLRPTVLL VDPGSPLFGT ELPFPFTAVA HVPRSKAVAAC
401 AGSLIVSVLG DAAGLLGALA EEPGVDKVFG AEEYDRGYHP CDPHEGYLADF
451 LFRKKAVIPGAAG
```

Table E.4: ThnO protein sequence.

>gi|337762320:919299-920717 ThnO [Streptomyces cattleya NRRL 8057]

```
   1  tcacccggcc gcgcccggga tgaccgcctt cttccggaac aggaagtcgg cgaggtagcc
  61  ctcgtgcggg tcacacgggt ggtagccgcg gtcgtactcc tcggcgccga agaccttgtc
 121  caccccgggc tcctcggcca gcgcgccgag cagtccggcc gcgtcgccga ggaccgagac
 181  gatcagcgaa ccggcgcagg cggccaccgc cttggagcgc ggcacgtggg caacggccgt
 241  gaacgggaac ggcagttcgg tgccgaagag cggtgacccc gggtcgacaa ggagcacggt
 301  ggggcgcagg aaggcggcgc cgtccacctc gacccgcagc ggcaccccgg tgacggcctc
 361  ggtgacgtcg acggcaccgg ccgcgatctc ccgttcgatg cgggtggcga gggcgtcgcg
 421  ccgcgcccgg tccgggaacg ccggtacggt ggccgccggg tcgtccagcg gcagcaccgg
 481  ccgggtggcg aactcctcgg cgaggcggtc ggcggcctgg cgggcgtccc cgagggcgat
 541  caccgagctg acgttggtgc acagccgtcc gcagccggcg aaggcggccc gggccagccg
 601  gggccacgcg tcgtcggcgg cgccggcgcc caccaccgcc cggctgcggc cgaagtggta
 661  ggtcttcgcc ttgcccggcg gcaggtcggc ggggacgtcc tcgccgggcc agagcacctg
 721  gtcggcgagg cggcgcagcg ccggggcgtc gccgtgcacc accgagacgg cgtgctcggg
 781  cagcccggcc tcgtacagcg cctcggtgag cagcaccggg gtgaacgggt cgtgccggga
 841  ggtggacagc agcaccgggc ggcgggcggc gagcgcctgg agccattcga tgttgatggt
 901  ggggaagttg tcggcgacct tgaccagcac cgagcggccg gccgggcgcc agctccagcg
 961  ggggtggtgg ccgccggtgc ggtaggccgc gacggtgccg tccggtgact gggcggcgag
1021  gatctcgtcc atccgggcca ggtcggcggc gacggtctcg atgccgcggg cggcgcgccg
1081  gggaggcagc ccgctcgccg cggagagggc ggtcagccac ggctgcgcgg tgccgtcggc
1141  gacgcgtgcg gtgagcgtgt ccgcggcgcg gcgcagcagc gcgaaccagg cgtcgtcgtc
1201  ggtggcgcgc agccgccggg cgaccgcggc ggcctcctgg gccatctgga actggacgag
1261  ccggccggcc gccccgaccc gggccaccgt ctcgccggtg gtgccggtga gcgcggccgg
1321  cctgctggac atcaccggcc ggcgcccggc gtaggccggg acgtcgagga cggcgccggc
1381  ccgctgcggg ccggcgtgtc cggtggtgac ggtcgtcac
```

Table E.5: ThnO gene sequence.

```
ThnO   MTTVTTGHAGPQRAGAVLDVPAYAGRRPVMSSRPAALTGTTGETVARVGAAGRLVQFQMAQEAAAVARRLRATDDDAWFA  80
ThnO*  MSAPTAALG-----TEVLDVPAYDGARAVMSHRPDVVRDGEGRPLARLGSTGRLTQFRMAEEAGRAARRLRGIDDADWFG  75
GriD   --MSTVAPP-----AVVLDPRVRG--RLVASSDRTRLPSVLGGDLADVGTAPRLLAVAALNEIRAHADGR----PPASEV  71
           *..          ***   .    * * *     : .  *  :* :*:: **  .   :*    *

ThnO   LLRRAADTLTARVADGTAQPWLTALSAASGLPPRRAARGIETVAADLARMDEILAAQSPDGTVAAYRTGGHHPRWSWRPA  160
ThnO*  LLRRAADRVTTKLDNGDLDVWLRALSAVSGLPRRRAARGLRTAAQDLARMEEILAAQAPDGTVSAYRTGGSGLGWSWRPA  155
GriD   FTEAARLFATATLDGETPREYARRVCHATGLTATAVEQAVTDLTGELRALPATTAAELP------ATGFGDGFDTRWVPR  147
       : . *    *: : .     :   :. .:**.   . :.:   : :*  :    **: *          *      * *

ThnO   GRSVLVKVADNFPTINIEWLQALAARRPVLLSTSRHDPFTPVLLTEALYEAGLPEHAVSVVHG----DAPALRRLADQVL  236
ThnO*  GRSVLVRQPGNFPTINIEWLQALATRRPVLLSTAAADPFTPVVLTEELYAAGLPEHGVSVVHG----DAPALRRLADQVL  231
GriD   GRTFAAVMASNHPVPNISWAQALFHGYSVLVKPGSRDPFTPARLLAALTAAGLPPDRAAFLPCSREVGAYLLREADRGIV  221
       **:. .  ..*.*. **.* ***    .**:...  *****. *   *  **** . .:.:      .*  **.    ::

ThnO   WPGEDVPADLPPGK-AKTYHFGRSRAVVGAGAADDAWPRLARAAFAGCGRLCTNVSSVIALGDARQAADRLAEEFATRPV  315
ThnO*  WPGEDVPADLPAGK-VKTYHFGRSRAVIGDRLPDDAWPRLAREAFAGCGRLCVNMSSLVVLGDPREAAEALAHEFAARPV  310
GriD   YGGDSAVATWHQRESVAVRGPGRTKAFLDRAPDDAVIDHLALSASFDGGTRCTNLSAVLTTGPVEEVADRLAERLARLPS  301
       : *:.. *     : . .   **::*.:.    * .  :**  *  . *  *.*:*:::. *  .:.*: **..:*  *

ThnO   LPLDDPAATVPAFPDRARRDALATRIEREIAAGAVDVTEAVTGVPLRVEVDGAAFLRPTVLLVD-PGSPLFGTELPFPFT  395
ThnO*  LPLDDPRAAVPAFPDRAARDSLALRIEREVAAGAVDVTAAATGVPLRVELDGKAFLRPTVLLVD-PGSPLFGTELPFPFT  390
GriD   LPATDEAATLLVVN-RARADGLREQVAALRASLTDHSARAGDPETVVELPDGSFLPRPVVLSADRADHPAVGTELPFPFV  340
       **  *  *:: ..  **  *.*  ::    *: : . : *    .:    **  : **.** .* .. * .********.

ThnO   AVAHVPRSKAVAACAGSLIVSVLGDAAGLLGALAEEPGVDKVFGAEEYDRGYHPCDPHEGYLADFLFRKKAVIPGAAG    472
ThnO*  AVAHVPRSKAVAACAGSLIVSVVGEADGLVRALAEEPGVDKVFAGDQFDRGYHPCDPHEGHLADFLFKKKAVIPGALA    467
GriD   VVAPWAEADGVEPLRDSLVLNLLTDREELVEAAVREPSVRKVTRGPVLPWTATPGIPHDDNYTQFLLEPKGVVARR--    456
       .**  ..:..* .  .**::.:: :   *: * ..**.* **  .        *  **:.  ::**:. *.*:.
```

Sequence alignment for *Streptomyces cattleya* NRRL 8057 ThnO (gi|337763152), *Streptomyces flavogriseus* ATCC 33331 ThnO* (gi|357409403), and *Streptomyces griseus* GriD (gi|182438037).

Table E.6: Sequence alignment between ThnO and other predicted reductases.

E.2 ThnN and ThnO Sequences of Expression Constructs and Purification

```
  1 MGSSHHHHHH SSGLVPRGSH MPSPSKLRPI TDEDVRRAVR LHFDPQDGTP
 51 YWLERDRRDG TDALRKVHGM LDAQQLIGLR DGPDQAHFEE ASRRTPLANF
101 VPRRVLTEGG PLWAAQTGGT TGPAKHGTWG ARYWADILEF SDEFLDLHGV
151 PRGVDWLFVG PMGPHTTGRL VVAFAERRGG MCFSVDLDPR IVKIFGEEGM
201 TAAYDRYVQH IWDQVAEVAR AQRIGVLFCT SRLLEMLPER LGTEALPGLR
251 AIVHAGTTME PESHRQLRED FFPGVPVVGM YGTSTTGISW QKPFEPEDDH
301 HVVYVPCQPH IALDAVDDDG NEVPFGEEGR VRVWRLTDDA LLPGFLERDR
351 ARRVQPYGAA AERYPWPWLG DPYSPEFTEG RRIEGVY*
```

Table E.7: ThnN expression construct Protein Sequence (387 a.a. MW = 43568 Da).

```
  1 MGSSHHHHHH SSGLVPRGSH MTTVTTGHAG PQRAGAVLDV PAYAGRRPVM
 51 SSRPAALTGT TGETVARVGA AGRLVQFQMA QEAAAVARRL RATDDDAWFA
101 LLRRAADTLT ARVADGTAQP WLTALSAASG LPPRRAARGI ETVAADLARM
151 DEILAAQSPD GTVAAYRTGG HHPRWSWRPA GRSVLVKVAD NFPTINIEWL
201 QALAARRPVL LSTSRHDPFT PVLLTEALYE AGLPEHAVSV VHGDAPALRR
251 LADQVLWPGE DVPADLPPGK AKTYHFGRSR AVVGAGAADD AWPRLARAAF
301 AGCGRLCTNV SSVIALGDAR QAADRLAEEF ATRPVLPLDD PAATVPAFPD
351 RARRDALATR IEREIAAGAV DVTEAVTGVP LRVEVDGAAF LRPTVLLVDP
401 GSPLFGTELP FPFTAVAHVP RSKAVAACAG SLIVSVLGDA AGLLGALAEE
451 PGVDKVFGAE EYDRGYHPCD PHEGYLADFL FRKKAVIPGA AG*
```

Table E.8: ThnO expression construct protein sequence (492 a.a. MW = 51693 Da).

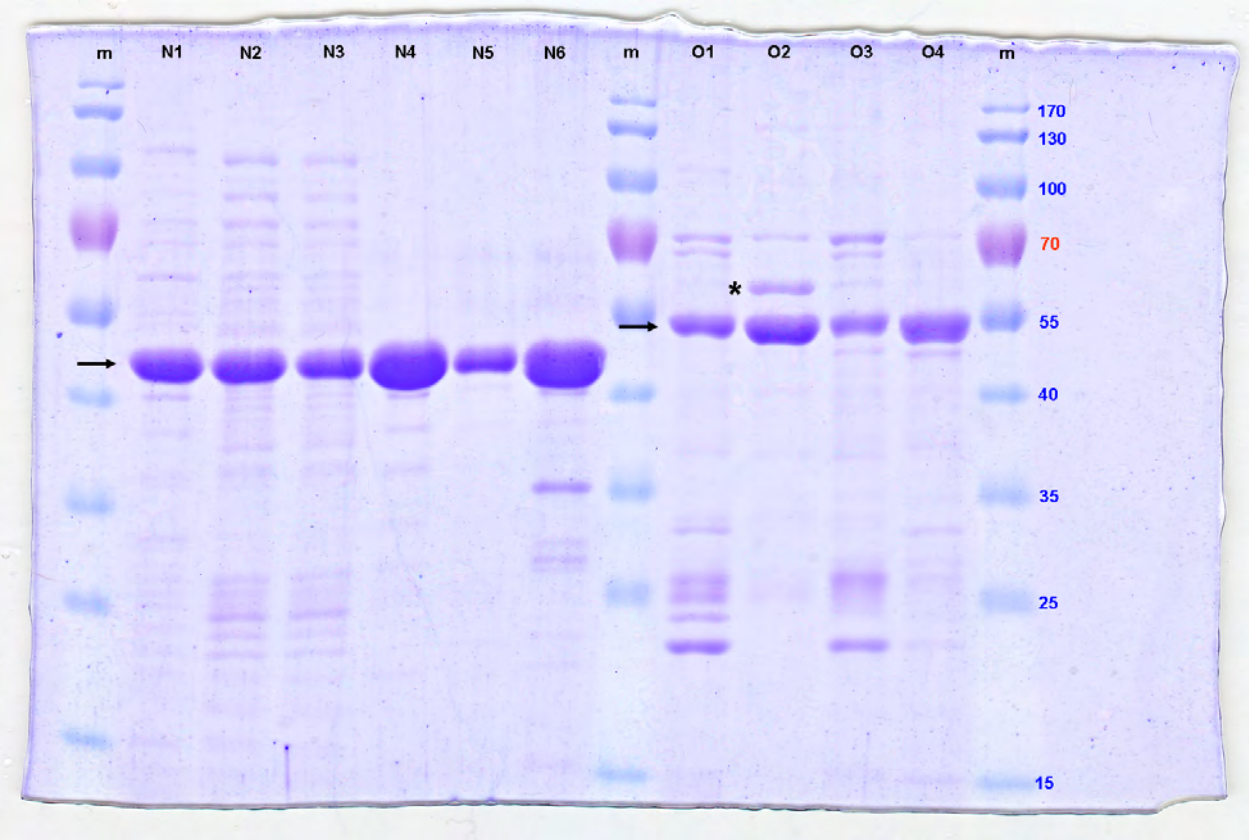


Figure E.1: SDS-PAGE analysis for different batches of ThnN and ThnO purification. Six different purification batch were performed for ThnN (lanes N1-6) and four for ThnO (lanes O1-4). The highest purity could be obtained after a size exclusion chromatography step, but a large portion of the protein was lost (lanes N4-6 and O4). The band marked with an asterix corresponds to an impurity also observed in the LC-MS analysis of the ThnO samples (Figure 7.2 D). All samples were run on 12.5% acrylamide gels. The molecular weights of the markers are given in kDa.

E.3 Confirmation of ThnN Identity

Protein name	ThnN	Peptide matches
NCBI Accession No.	>gi|30577691	K.LRPITDEDVR.R
Protein Score†	4099	K.LRPITDEDVRR.A
Sequence coverage (%)	75	R.LHFDPQDGTPYWLER.D
		R.DRRDGTDALR.K
		R.RDGTDALR.K
		R.KVHGMLDAQQLIGLR.D
		R.KVHGMLDAQQLIGLRDGPDQAHFEEASR.R
		K.VHGMLDAQQLIGLR.D
		K.VHGMLDAQQLIGLRDGPDQAHFEEASR.R
		R.DGPDQAHFEEASR.R
		R.DGPDQAHFEEASRR.T
		R.RTPLANFVPR.R
		R.TPLANFVPR.R
		R.RVLTEGGPLWAAQTGGTTGPAK.H
		R.VLTEGGPLWAAQTGGTTGPAK.H
		R.VLTEGGPLWAAQTGGTTGPAKHGTWGAR.Y
		K.HGTWGAR.Y
		R.YWADILEFSDEFLDLHGVPR.G
		R.GVDWLFVGPMGPHTTGR.L
		R.LVVAFAER.R
		R.RGGMCFSVDLDPR.I
		R.GGMCFSVDLDPR.I
		K.IFGEEGMTAAYDR.Y
		K.IFGEEGMTAAYDRYVQHIWDQVAEVAR.A
		R.YVQHIWDQVAEVAR.A
		R.IGVLFCTSR.L
		R.LLEMLPER.L
		R.LGTEALPGLR.A
		R.AIVHAGTTMEPESHR.Q
		R.LTDDALLPGFLER.D
		R.RVQPYGAAAER.Y
		R.VQPYGAAAER.Y
		R.YPWPWLGDPYSPEFTEGR.R
		R.YPWPWLGDPYSPEFTEGRR.I
		R.RIEGVY.-
		R.IEGVY.-

† Protein scores were determined using Mascot version 2.3 (Matrix Science). The values represent the negative log of the probability that this identification is a random match (significance threshold filter, $p < 0.01$).

Table E.9: Confirmation of ThnN identity by tryptic digest followed by MS/MS analysis.

References

[1] Larkin, M. A., et al. (2007) Clustal W and clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.