Mechanistic studies of PLP-independent racemases

by

Yeong-Chan Ahn

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

Abstract

Pyridoxal phosphate (PLP)-independent racemases, enzymes that are cofactor and metal free invert, the configuration at the α-carbon of amino acids. In nature, diaminopimelate epimerase (DapF), glutamate racemase (GluR), aspartate racemase (AspR), proline racemase (ProR), O-ureidoserine racemase (DcsC), and a putative amino acid racemase from *Bacillus* (RacX) are known as PLP-independent racemases. Comparison of these enzymes, may reveal similarities and differences in their unusual mechanism of action.

The first part of this dissertation describes mechanistic studies of *O*-ureidoserine racemase (DcsC). DcsC is a PLP-independent enzyme that is involved in the biosynthetic route of the antibiotic D-cycloserine. To obtain insight into the mechanism of DcsC, a variant featuring an N-terminal SUMO-tag was expressed possessing significantly enhanced stability and activity. Synthesis of enantiomerically pure inhibitors in combination with site-specific mutation of active site cysteines to serines of this enzyme offers information about the mechanism of this transformation. Homology modeling with a close relative (diaminopimelate epimerase, DapF) inspired C- and N-terminal truncation of DcsC to produce a more compact yet still active enzyme variant. An ongoing goal of this project is obtaining a single X-ray structure of the enzyme to understand the interior active site dimensions and recognition of the distal site.

The second part of this thesis describes efforts on the characterization of another amino acid racemase, namely RacX. RacX is also a PLP-independent enzyme that is found almost exclusively in *Bacillus* species. Even though the enzyme is highly homologous to AspR and GluR, aspartate and glutamate are not good substrate for RacX. Since the

catalytic efficiency for other amino acids with RacX was also low, it is possible that the natural substrate of RacX is not a single amino acid but a derivative thereof. The RacX gene is clustered with penicillin binding protein 4 (PbpE) which catalyzes the process of peptidoglycan hydrolysis. Therefore, it may be that RacX has a substrate which is a part of the peptidoglycan structure. To investigate this, various oligopeptides were synthesized corresponding to peptides in peptidoglycan. Further synthesis of proposed substrates and testing of their catalytic transformation are ongoing to determine if they can be efficiently transformed.

Preface

The content of Chapter 2 was published as Y-C. Ahn et al. *Org. Biomol. Chem.,* 2018, 16, 1126–1133. I performed all the experiments with assistance from the other authors, specifically in the molecular biology for the DcsC enzyme. I wrote the manuscript.

Chapter 3 is an unpublished work. I performed all the experiments described in this chapter.

Acknowledgements

"No one who achieves success does so without acknowledging the help of others. The wise and confident acknowledge this help with gratitude."

– Alfred North Whitehead

This thesis would not have been possible without the help, guidance, and support of several individuals. First and foremost, I am beyond grateful to my supervisor, Prof. John C. Vederas, for giving me the opportunity to do researches in his lab. He advised me for the research, supported and encouraged my work, and also gave us lots of teachings in life.

Thanks also to Vederas group members both past and present, for making the best memories and having unforgettable experiences here. Most of all, special thanks to Dr. Marco van Belkum and Dr. Conrad Fisher, who are my research dad and mom, for being excellent mentors and colleagues. Without their help, I could not succeed. I will never forget their warm advice, and keep them in my heart. Randy Sanichar, Fabricio Mosquera and other coworkers, I am grateful to know you guys and thanks for being wonderful friends both in and out of the lab. It was great time to work and hang out with you guys, and I wish we can share more great memories in the future. Finally, great thanks to Wayne Voung for proofreading this thesis.

I would like to say thanks to the entire scientific services staff in the Department of Chemistry for their contributions. Their help allowed me to perform the experiments of this thesis. A special thanks to Bernie Hippel (General Services), Gareth Lambkin (Biological

Services), Dr. Randy Whittal, Bela Reiz, Jing Zheng (Mass Spectrometry), Mark Miskolzie (NMR Laboratory), and Wayne Moffat (Analytical and Instrumentation Laboratory).

I am also extremely grateful to my family and friends for the encouragement and support. To my parents, thank you for all your sacrifices and love. Specially, thanks to support my plan to study abroad in Canada and help to survive here. Without their sacrifices, I could not think to study in Canada and follow my dream. To my sister, thanks also for your support and sacrifices. I could come to Canada to study because you are staying with my parents and having time together. Also, thanks to give me chances to meet my gorgeous nieces. To all my friends, thank you for encouraging me to not give up the program. I could finish my Ph.D program because you guys are always cheering and helping me. Thanks everyone again, and I will never forget your guys help.

Table of Contents

List of Figures

Chapter 1. Introduction of PLP-independent racemases

Chapter 2. Mechanistic studies of *O*-ureidoserine (DcsC)

Chapter 3. Amino acid racemase in *Bacillus Subtilis* (RacX)

List of Tables

List of Abbreviations

NAG	N-acetylglucosamine
NAM	N-acetylmuramic acid
PLP	Pyridoxal phosphate
AlaR	Alanine racemase
SerR	Serine racemase
NMDARs	N-methyl-D-aspartate receptors
EDTA	Ethylenediaminetetraacetic acid
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
GRIP	Glutamate receptor interacting protein
PICK1	Protein interaction with C kinase 1
Golga3	Golgin subfamily A member 3
PKC	Protein kinase C
DapF	Diaminopimelate epimerase
GluR	Glutamate racemase
AspR	Aspartate racemase
ProR	Proline racemase
DAP	Diaminopimelate
UDP	Uridine diphosphate
***Ec*GluR**	Glutamate racemase from *Escherichia coli*

***Ho*GluR**	Glutamate racemase from *Homo sapiens*
DeGluR	Glutamate racemase from *Deinococcus maricopensis*
***He*GluR**	Glutamate racemase from *Helicobacter pylori*
***Aq*GluR**	Glutamate racemase from *Aquifex pyrophilus*
***La*GluR**	Glutamate racemase from *Lactobacillus fermenti*
***Ba*GluR**	Glutamate racemase from *Bacillus anthracis*
***En*GluR**	Glutamate racemase from *Enterococcus faecalis*
***Py*AspR**	Aspartate racemase from *Pyrococcus horikoshii*
***Ac*AspR**	Aspartate racemase from *Aspergillus niger*
***St*AspR**	Aspartate racemase from *Streptococcus thermophilus*
***Tp*AspR**	Aspartate racemase from *Thermococcus paralvinellae*
EcL-DER	Aspartate/glutamate racemase from *Escherichia coli*
***Bb*AspR**	Aspartate racemase from *Bifidobacterium bifidum*
***Cs*ProR**	Proline racemase from *Clostridium sticklandii*
***Tc*ProR**	Proline racemase from *Typanosoma cruzi*
HypE	Hydroxyproline-2-epimerase (HypE)
PYC	Pyrrole-2-carboxylic acid
GABA	γ-aminobutyric acid
INH	Isoniazid
RIF	Rifampin
EMB	Ethambutol
PZA	Pyrazinamide
Ddl	D-alanyl-D-alanine ligase

DcsA	N-hydroxyl-L-arginine synthase
DcsB	N^{ω}-hydroxy-L-arginine amidinohydrolase
DcsC	*O*-ureidoserine racemase
DcsD	*O*-acetyl-L-serine synthase
DcsE	L-serine-O-acetyltransferase
DcsG	O-ureido-D-serine cyclo-ligase
KOtBu	Potassium *tert*-butoxide
NMR	Nuclear magnetic resonance
CD	Circular dichroism
ESI	Electrospray ionization
TOF	Time of flight
MS	Mass spectrometry
PCR	Polymerase chain reaction
SUMO	Small ubiquitin-like modifier
TIPS	Triisopropylsilyl chloride
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
TBAF	Tetra-*n*-butylammonium fluoride
PBr_3	Phosphorous tribromide
PPh_3	Triphenylphosphine
DIAD	Diisopropyl azodicarboxylate
Et_2O	Diethyl ether
DMF	Dimethylformamide
DCM	Dichloromethane

DTT	Dithiotreitol
RacX	Amino acid racemase in *Bacillus subtilis*
PBP4a	Penicillin-binding protein
Orn	Ornithine
Hse	Homoserine
Aba	Aminobutyric acid
Nva	Norvaline
PbpE	Penicillin-binding protein 4
Ni-NTA	Nickel-Nitrilotriacetic acid
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TFA	Trifluoroacetic acid
Boc	tert-butyloxycarbonyl
HATU	(1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]-pyridinium 3-oxid hexafluorophosphate
DIPEA	Diisopropylethylamine
HOBt	Hydroxylbenzotriazole

Chapter 1

Introduction of PLP-independent racemases

1.1. D-amino acids

1.1.1. General introduction

Except for glycine, all proteinogenic amino acids have a chiral center. The two optical isomers are referred to as D- and L-amino acids. It has been suggested there were evolutionary advantages in favouring L-amino acids as building blocks for proteins.[1,2] However, many organisms also synthesize D-amino acids for various reasons. Peptides containing D-amino acids are more resistant against proteases.[3] Most proteases recognize protein residues with L-amino acids as substrates. Some experiments have shown that peptides formed by D-amino acids can have a left-handed helix structure and are not recognized by proteases in serum.[4,5] As the chirality of amino acids affects the stability of proteins, some organisms such as bacteria use D-amino acids in cell walls to survive in nature.[6]

1.1.2. D-amino acids in the bacterial cell wall

A common place to find D-amino acids is the peptidoglycan layer in the bacterial cell wall.[7] D-amino acids are key constituents of the short peptides in the peptidoglycan

polymer. Peptidoglycan monomers are composed of β-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc), bearing a short peptide chain in the 3-position (Figure 1.1). The sequence of these peptides is different depending on the species, but includes D-configurated entities such as D-glutamate, D-alanine and *meso*-diaminopimelate. The peptidoglycan monomers are polymerized by transglycosylases, and the peptides in peptidoglycan can be crosslinked with other peptidoglycan polymers to construct the bacterial cell wall.[8,9,10]

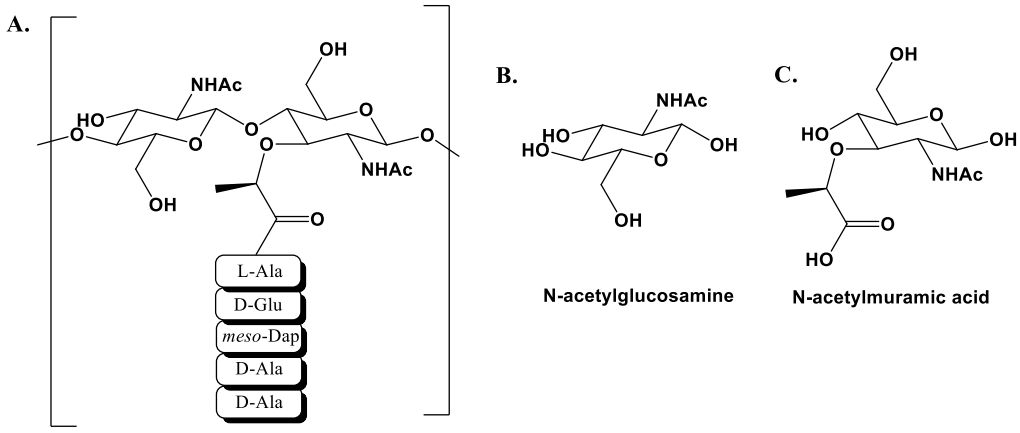


Figure 1.1. A. Basic structure of peptidoglycan monomer in bacteria cell wall in Gram negative bacteria. B. The structure of N-acetylglucosamine. C. The structure of N-acetylmuramic acid.

1.1.3. D-amino acids in other biological processes

D-amino acids are also involved in other important biological processes in bacteria, including biofilm decomposition, spore germination, bacterial capsule formation, and cell signaling.[11,12,13,14] Biofilm, polymeric structures assembled by many bacterial cells, is

formed for survival advantages such as protection, cellular communication, transfer of genetic material, and increasing availability of nutrients.[15,16] D-amino acids are used in widespread signaling of biofilm disassembly, therefore suggesting D-amino acids can break biofilms.[11] D-amino acids also function as inhibitors of spore germination in bacteria.[12] Thirdly, certain bacteria, such as *Bacillus anthracis*, synthesize capsules after spore germination to protect themselves. These capsules are composed of polymers containing β-linked D-glutamic acid.[13] In animals and humans, high concentrations of D-amino acids, especially D-serine and D-aspartate, are found in the central nervous system. D-serine plays a role in excitatory amino acid metabolism, which is associated with neurological disorders such as schizophrenia, ischemia, epilepsy, and neurodegenerative diseases. The role of D-aspartate in unknown, but it is expected to have a role in age-related ocular diseases.[17,18]

1.2. PLP-dependent racemases

1.2.1. General overview

In nature, most D-amino acids are synthesized by racemization of L-amino acids employing racemases or epimerases. There are two types of racemases found in nature: one is pyridoxal phosphate (PLP) dependent racemases, and the other is PLP-independent racemases. Because the *p*Ka of the hydrogen at the α-carbon (α-hydrogen) in amino acids is around 29 and it is geometrically disfavored for removal, nature needs enzymes to decrease the *p*Ka of the α-hydrogen.[19] The conjugated π-system of the PLP cofactor

delocalizes the negative charge of the α-carbanion, thus decreasing the *p*Ka of the α-hydrogen (Figure 1.2).[20] When the substrate amino acid enters the enzyme, it is covalently bound as an imine to the nitrogen of PLP. Next, the base from the racemase enzyme deprotonates the α-hydrogen to make an α-carbanion. Electrons are delocalized in the conjugated π-system of PLP, which provides an "electron sink". Subsequently, the α-carbon is protonated again by a moiety in the active site of the racemase from the opposite face, which results in inversion of the chirality of the α-carbon. Many PLP-dependent racemases are known, including alanine racemase (AlaR) and serine racemase (SerR).

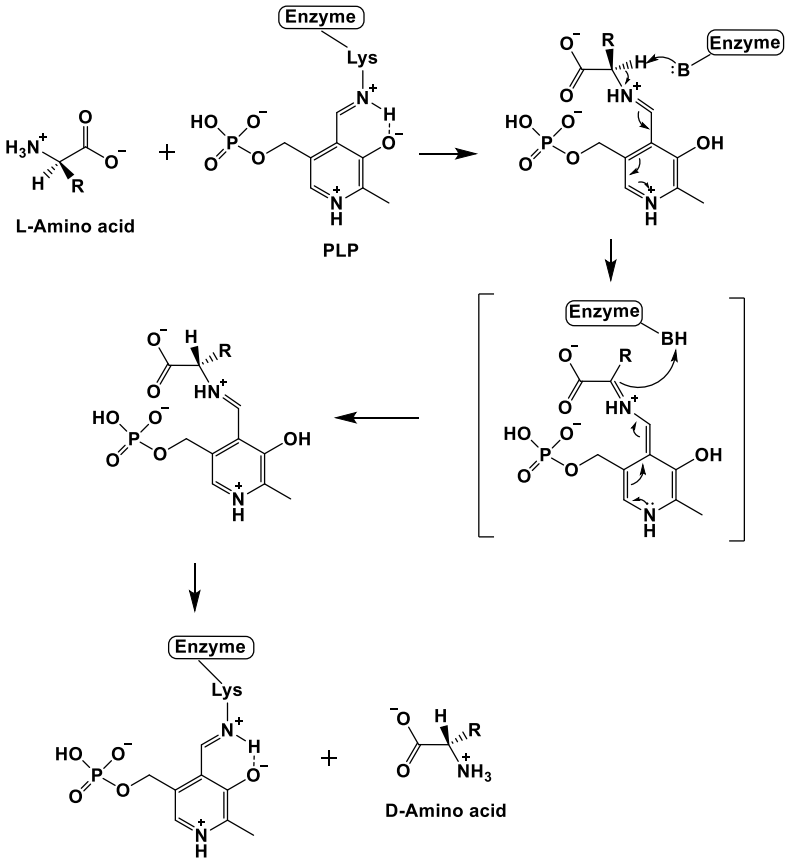


Figure 1.2. Mechanism of PLP-dependent racemases.

1.2.2. Alanine racemase

D-alanine is an important amino acid in the peptidoglycan found in bacteria cell walls (Figure 1.1). Alanine racemase (AlaR), which is found in bacteria and some eukaryotes such as aquatic invertebrates,[21] yeast,[22] fungi,[23] and plants,[24] catalyzes the racemization of L-alanine and D-alanine. AlaR requires PLP as a cofactor to decrease the *p*Ka of the α-hydrogen. The active site residues of AlaR from *Geobacillus stearothermophilus* include Lys39 and Tyr265. The sidechain *p*Ka of the lysine is 10.5, and that of tyrosine is 10. The crystal structure of AlaR from *G. stearothermophilus* helps to explain a detailed mechanism.[19] Initially, Lys39 is covalently attached to PLP as an aldimine. Transimination with L-alanine frees Lys39 and leaves the PLP-alanine adduct bound non-covalently in the active site. Tyr265 deprotonates the α-carbon to form a quinonoid resonance-stabilized intermediate. Finally, Lys39 protonates the opposite site of the α-carbon to produce the corresponding enantiomer of the substrate. The D-alanine product is released, presumably by transamination with Lys39. Other amino acids in the active site help to recognize the substrate and the cofactor by hydrogen bonding (Figure 1.3). The phosphate group of PLP, including negatively charged oxygens, is stabilized by hydrogen bonding with the hydroxyl groups of Ser204 (2.8 Å), Tyr43 (2.9 Å), and Tyr354 (3.0 Å). The nitrogen in the pyridine ring of PLP is stabilized by hydrogen bonding with the guanidinium group of Arg219 (3.6 Å), and this guanidium group is stabilized by His200 (3.0 Å). The hydroxyl group of PLP has a hydrogen bond with the amino group of guanidinium group of Arg136 (3.4 Å). The other active site, Tyr265 has a hydrogen bond with the guanidinium group of Arg136 (3.6 Å), and is also connected with His166 (2.7 Å).

Also, the alanine (substrate) is specifically recognized by the active site of AlaR by hydrogen bonding and other non-covalent interactions.

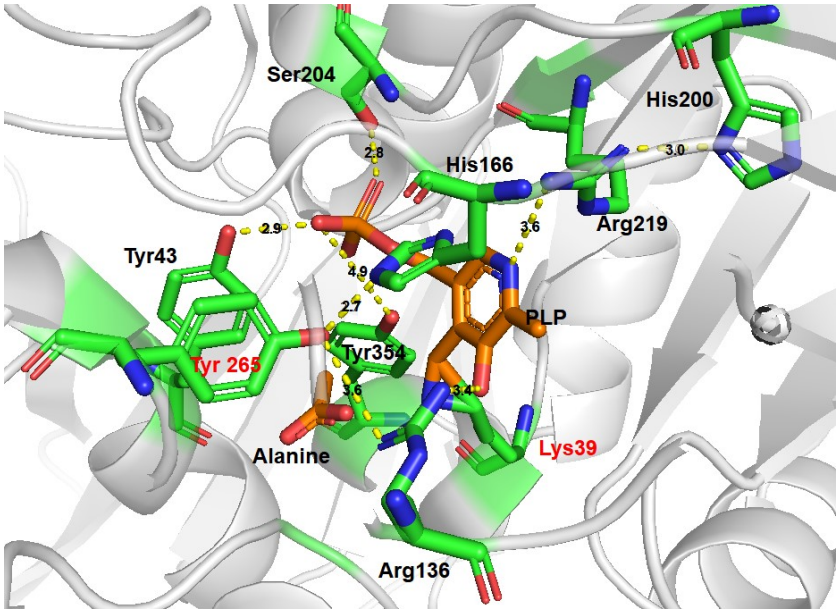


Figure 1.3. Crystal structure of active site of alanine racemase in *Geobacillus stearothermophilus* with PLP and D-alanine. [19]

1.2.3. Serine racemase

1.2.3.1. Function and mechanism

Serine racemase (SerR) is an enzyme that catalyzes reversible conversion of L- to D-serine. It was first discovered in pupae of the silkworm *Bombyx mori* in 1998 and independently in rat brain (mammalian cell) in 1999.[25,26] D-serine is used in neuro-modulation by coactivating N-methyl-D-aspartate receptors (NMDARs) in the mammalian central nervous system.[27] Also, D-serine was also found in peripheral tissues and organs,

where it functions predominantly in corticolimbic brain structures in signaling roles.[28,29] Because of the importance of D-serine, SerR and its mechanism is widely researched. SerR is known as PLP-dependent racemase. The crystal structure of SerR from humans supports the mechanism of serine racemase.[30] SerR in humans has Lys56 and Ser82 in the active site, which work as an acid-base pair to protonate and deprotonate the α-carbon. Initially, Lys56 at the active site is bound to PLP by an aldimine linkage, and the enzyme catalyzes transamination of the substrate L-serine. Since the *p*Ka of the α-hydrogen of the PLP binding serine substrate is decreased by the cofactor portion, Lys56 can act as a base to deprotonate the α-hydrogen and generate the quinonoid intermediate. Finally, Ser82 protonates the PLP-serine complex substrate from the opposite side to change its configuration to D-serine. The mechanism for racemization of D-serine is similar to that of L-serine. Ser82 in SerR deprotonates the α-carbon and Lys56 protonates the α-carbon to racemize D-serine to L-serine. Interestingly, pyruvate is also synthesized by SerR with a similar mechanism. From the quinonoid transition state, the water is lost, and lysine protonates the aminoacrylate which is then hydrolyzed to pyruvate (Figure 1.4). In other words, serine racemases not only generate D-serine, but also convert substrate to pyruvate.[30]

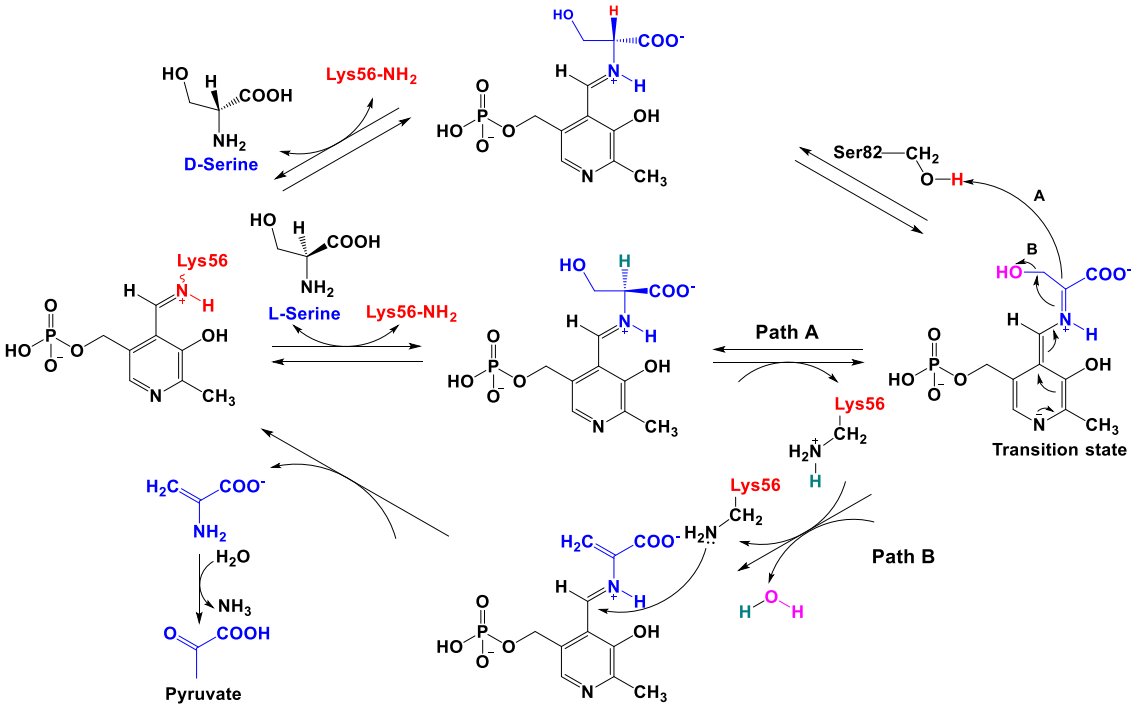


Figure 1.4. Mechanisms of serine racemase.

1.2.3.2. Catalytic regulatory mechanisms of serine racemase

SerR is regulated by interactions with divalent cations, nucleotides, and specific proteins. Firstly, SerR has a cation binding site, where divalent cations such as calcium, magnesium, and manganese bind to SerR via hexavalent coordination.[31] In other words, two carboxylate anions of amino acid side chains, a carbonyl oxygen of the main-chain, and three water molecules have hexavalent coordination with divalent cation. In human SerR, the carbonyl group of Glu210 (2.1 Å), Ala214 (2.2 Å), Asp216 (2.1 Å) and three molecules (2.2, 2.2, 2.3 Å) are coordinated with a magnesium ion (Figure 1.5)[31]. The Mg^{2+} binding increases the activity of SerR more than 10-fold, as shown by inhibition experiments with ethylenediaminetetraacetic acid (EDTA) and SerR.[31] Additional research

has shown that calcium ions (Ca^{2+}) are also regulators of SerR, but more than two times higher concentrations of Ca^{2+} are needed to show the same activity as Mg^{2+}.

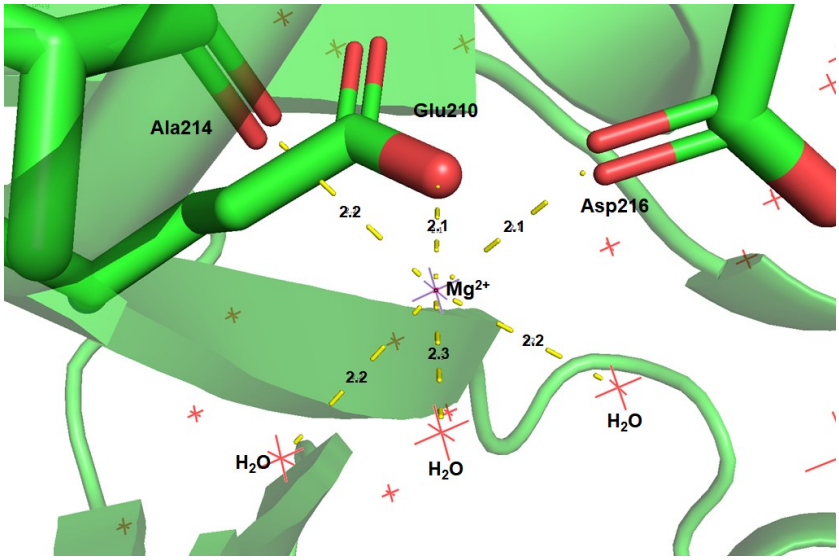


Figure 1.5. Cation binding site of serine racemase from human.[31]

In addition to cations, nucleotides also bind to SerR to regulate the enzyme. Nucleotides including adenosine triphosphate (ATP) and adenosine diphosphate (ADP) bind to the subunit binding site. If there are no nucleotides bound to SerR, the enzymatic activity is lower by more than 10-fold. Also, nucleotides form complexes with Mg^{2+} ion to further to regulate the enzyme in human SerR.[31] According to experiments done by the Wolosker group, binding of the ATP-Mg^{2+} complex showed 2 to 10-fold increased activity compared to binding of only ATP or Mg^{2+} ions. ATP showed the highest binding affinity with SerR (K_m= 3 μM), about 30-fold higher than for ADP (K_m=100 μM). Thus ATP is considered the main nucleotide regulator for SerR, which forms a complex with a Mg^{2+} ion to regulate SerR.

Specific proteins can also bind to SerR to regulate the enzyme. There are three proteins that are known to be SerR binding proteins; glutamate receptor interacting protein (GRIP), [32, 33] protein interacting with C kinase 1 (PICK1)[34] and the Golgin subfamily A member 3 (Golga3).[35] GRIP, which is common subunit in the 2-amino-3-(5-methyl-3-oxo-1,2-oxazol-4-yl) propanoic acid (AMPA) receptors found in the central nervous system, binds to the C-terminus of SerR to physiologically activate SerR.[36] PICK1, which is known to interact with protein kinase C (PKC), also binds to the C-terminus of SerR to regulate the enzyme. However, PKC decreases the activity SerR by phosphorylation which is mediated by PICK1/SerR interactions.[37] Golga3 binds to the N-terminal 66 amino acids of brain SerR. Golga3 significantly increases SerR activity by inducing a conformational change in the enzyme, giving rise to an increased synthesis of D-Serine. [35]

1.3. PLP-independent racemases

1.3.1. General overview

One key difference in the mechanism of racemases is how the *p*Ka of the α-hydrogen (~29) is decreased. As described above, PLP-dependent racemases use a cofactor to decrease the *p*Ka of the α-hydrogen. In contrast, PLP-independent racemases do not require any cofactors or metal ions to achieve this goal. PLP-independent enzymes possess a thiol-thiolate pair, which acts as a base (thiolate) to deprotonate the α-carbon and subsequently as an acid (thiol) to reprotonate the α-carbon from the opposite site (Figure 1.6). However, since the *p*Ka of cysteine is comparatively low (~8), the corresponding anion does not appear basic enough to deprotonate the α-carbon.[38]

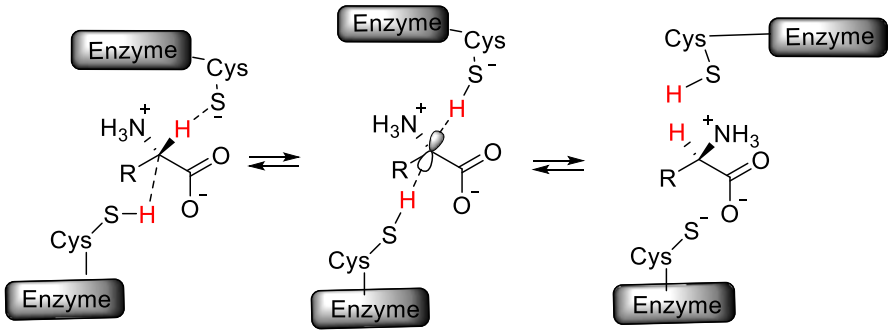


Figure 1.6. Racemization of amino acids by PLP-independent racemases.

There have been numerous studies to reveal how PLP-independent racemases are able to deprotonate the substrate. Several PLP-independent racemases have been characterized including diaminopimelate epimerase (DapF), [39] glutamate racemase (GluR),[40] aspartate racemase (AspR),[41] and proline racemase (ProR).[42] It is known that there are mechanistic and structural differences between them (Table 1.1). Understanding the mechanism of these enzymes will provide insight into how catalysis and control of *p*Ka can be achieved by protein.

Table 1.1. Statistic and structural features of PLP-independent racemases [a]

	DapF	GluR	AspR	ProR	DcsC	RacX
Year of discovery	1957	1952	1972	1957	2010	1993
First PDB structure	1998	1999	2002	2006	-	-
PDB	2gkj	2jfx	2dx7	1w61	-	-
Approx. hydrophobic dimerization interface, Å^2	1000	1015	2540	1100	-	-
Approx. pocket volume, Å^3	160	310	150	100	(200)*	(>200)*
Distance of active site residues, Å	6.3	7.5	9.6 (7.4)	6.8	(6.1)*	(9.3)*
Water molecules included in active site	0	1	1	0	-	-

[a] See discussion below for references

1.3.2. Diaminopimelate epimerase (DapF)

1.3.2.1. Function and localization

Diaminopimelate epimerase (DapF), which catalyzes the interconversion of LL-diaminopimelate (DAP) and DL-(*meso*)-diaminopimelate, is a key enzyme in the lysine biosynthesis pathway of plants and bacteria.[43] It also generates a key precursor, *meso*-DAP, for peptidoglycan biosynthesis in bacteria.[39] According to the NCBI database, there are currently over 140,000 entities of DapF including 137,000 from bacterial species (>99%). Because both *meso*-DAP and lysine are important constituents of the bacterial cell wall, targeting DapF could be a powerful approach to develop next generation antibiotics. DapF is known as PLP-independent racemase, does not require any cofactors or metal ions and possesses two cysteine residues that act as base (thiolate) and acid (thiol) pairs at the active (Figure 1.7). The structure and mechanism of DapF is characterized in detail, presenting the classic example of the bilateral thiol/thiolate moderated mechanism.[39]

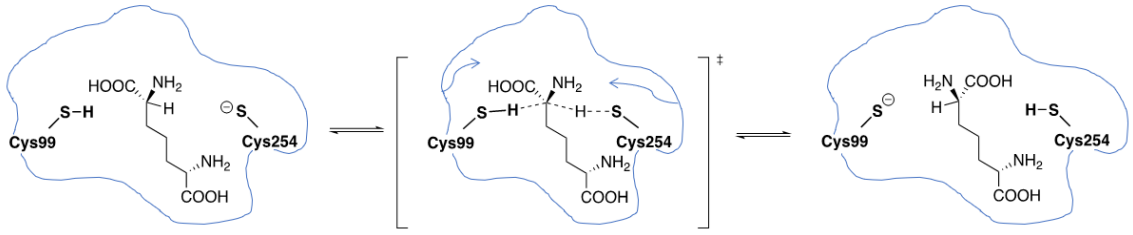


Figure 1.7. Active site bilateral mechanism of epimerization by DapF in *Arabidopsis thaliana*. [39]

1.3.2.2. Structure and mechanism

X-ray crystal structures of DapF and active site mutants have been reported from various species[44,45] with the substrate DAP[39,46] or with aziridine based inhibitors.[47,48] DapF has an overall dimeric structure with two structurally homologous domains including eight antiparallel β-strands and two α-strands in each domain.[47] Each domain contributes one cysteine residue to the active site (C99, C254 for *Arabidopsis thaliana*, Figure 1.7). These domains are flipped open and expose the active site to water in the substrate-free form of the enzyme. When the substrate *meso*-DAP enters the enzyme, the two domains start to close in an 'oyster shell-like' clamping motion, moving both domains closer together by about 26° (18° with azi-DAP inhibitor). This 'oyster shell-like' clamping motion expels water from the active site to create a non-polar active site cage (Figure 1.8). This hydrophobic environment in the active site contributes to conformational restriction, making the thiolate anion more basic. Hydrogen bonding of the substrate's α-carboxyl stabilizes developing negative charge at the α-carbon.[49]

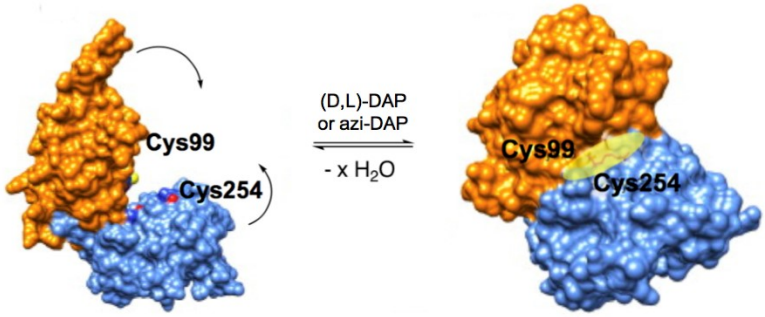


Figure 1.8. Movement of DapF upon ligand binding induces expulsion of water molecules and formation of a non-polar active site cage (highlighted yellow). [48]

In addition, the overall stereoelectronic arrangement of the active site also decreases the *p*Ka of the α-hydrogen. According to mutational studies, both active site cysteine residues are necessary for activity, supporting a bilateral reaction mechanism.[39] The crystal structure of DapF in *Arabidopsis* reveals details about these arrangements (Figure 1.9). There is a tight network of hydrogen bonds to selectively recognize the L-configuration for a perfect fit: Asn37, Gly71, Asp90, Asn101, Glu245, Arg246, Ala253, and Thr256 all contribute hydrogen bonds selective for the L-substrate.

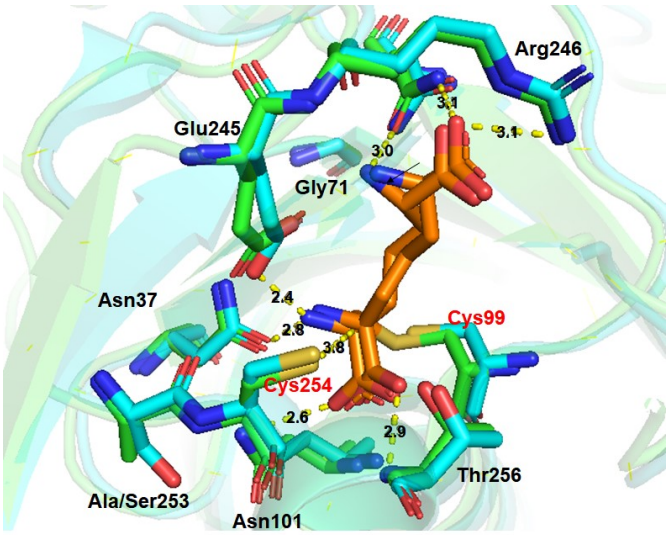


Figure 1.9. Active site hydrogen-bond network of DapF from *A. thaliana* (green), superimposed with the DapF from *Corynebacterium glutamicum* (blue). [47, 48]

Regarding the epimerization site, six structural features have been identified that contribute beneficially to catalysis:

(I) The dipole of an α-helix initially lowers the *p*Ka of the thiol to enhance thiolate formation at neutral pH (Figure 1.10).

(II) Water exclusion upon domain closure induced by substrate binding enhances basicity of thiolate anion to remove the α-hydrogen (Figure 1.8).

(III) The substrate carboxylate group maximizes orbital overlap between its π system and the breaking hydrogen-α-carbon σ bond (stereoelectronic alignment). When the cysteine protonates the α-carbon, the transition state of carbanion becomes a planar (σ→ π system, Figure 1.11).

(IV) The α-NH_2 group of the substrate is fully protonated and switches the hydrogen bonding partner during epimerization (Glu208 ↔ Gln44 in *H. influenzae*). The temporary release during switching of the hydrogen-bonding partners enhances the electron-withdrawing capability of the ammonium moiety (Figure 1.10).

(V) The enzyme stabilizes the negative charge on the α-carboxylate by the dipole of an extended α-helix opposite of the zwitterionic substrate (Figure 1.10).

(VI) Substrates have minimal movement during catalysis. The substrates are tightly held at the active site by hydrogen bond networks to prevent interference in the epimerization reaction (Figure 1.9).

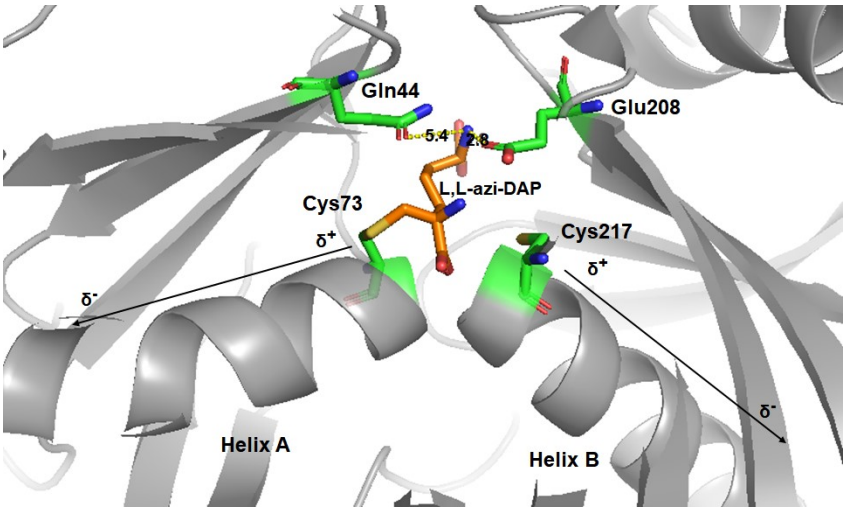


Figure 1.10. The crystal structure of DapF from *Haemophilus influenzae* with L,L-**azi-DAP.** [39]

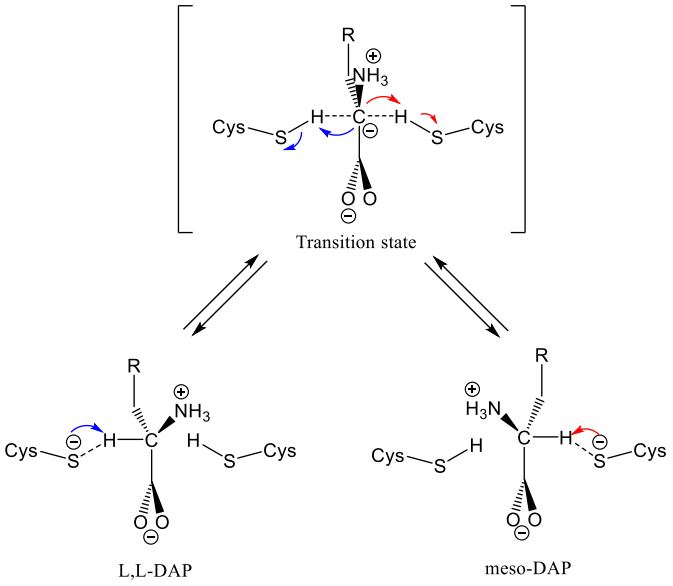


Figure 1.11. Proposed mechanism for the interconversion of LL-DAP and *meso*-DAP by DapF. The transition state bears a planar carbanion. [47]

These catalytic features are found in DapF homologues of various species, with most interactions of the two enzyme domains remaining highly conserved. With exception of variations in the N-terminal region, the overall conformations of the two domains are similar.[41,43]

1.3.2.3. Inhibition strategies

Since γ-methylene-DAP, the first DapF inhibitor, was synthesized in 1986 by Girodeau and coworkers, characterization of DapF inhibitors underlines an ongoing interest in finding potent pharmaceutical agents against tuberculosis due to the importance of DapF in nature (Figure 1.12A).[50]A detailed review outlines many studies on DapF inhibitors.[51] The general strategy relies on mechanism-based inhibition with substrate analogue inhibitors. The first example of a mechanism-based inhibitor is 3-halo DAP. Mechanistically, an imine is formed *in situ* through elimination, and tautomerization to give K_i values between 4-25 μM (Figure 1.12B).[52,53] The second example involves a reactive aziridine functional group which becomes covalently bound to a cysteine in the active site.[54,55] Irreversible inhibition by a 2-(4-amino-4-carboxybutyl)aziridine-2-carboxylic acid (azi-DAP) analogue was first introduced in 1990 and chiral azi-DAP was synthesized in 2005.[56] The (L,L)-azi-DAP is directly bound to the active site Cys73, whereas the (D,L)-azi-DAP form exclusively binds to Cys217 (Figure 1.12C). This bilateral inhibition mechanism was confirmed by X-ray crystal structures. [39, 47]

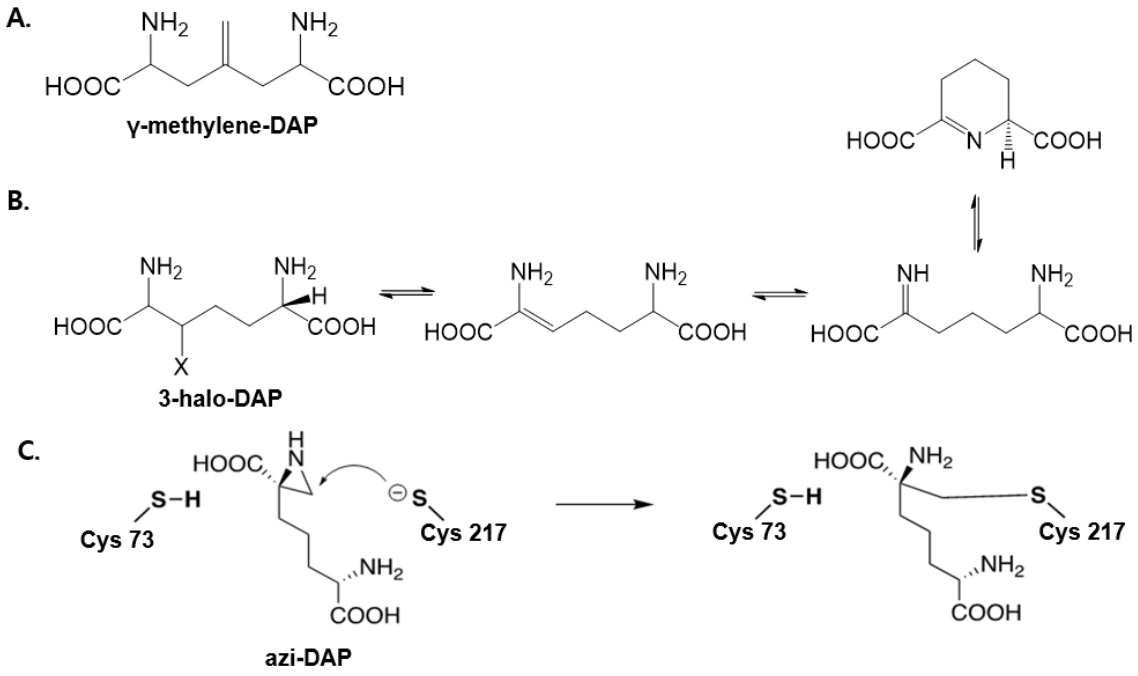


Figure 1.12. Various DapF inhibitors. A. First reported DAP analogue inhibitor. B. The elimination of 3-halo-DAP. C. Inactivation of active site Cys217 by the substrate mimetic DL-azi-DAP.

1.3.3. Glutamate racemase (GluR)

1.3.3.1. Function and localization

Glutamate racemase (GluR) is an essential enzyme in peptidoglycan synthesis for bacterial cell walls, which catalyzes the conversion between L-glutamic acid and D-glutamic acid.[40,57] According to the NCBI database, there are around 20,500 glutamate racemase entries, including 20,100 from bacterial species (>99%). Since GluR is essential for bacterial growth, elucidating the mechanistic details of GluR can aid in developing

selective inhibitors as new antibacterial therapeutics.[58] Similar to DapF, GluR is also co-factor free and possesses two active site cysteines that work to deprotonate the glutamate α-carbon by a thiolate and subsequently reprotonate the intermediate carbanion from the other side via a thiol.[59,60] In comparison with DapF, the macromolecular assembly of the GluR enzyme shows a broader variety of structures.

1.3.3.2. Structure of GluR

Depending on species, GluR can adopt various oligomerization states including monomeric, head-to-head dimeric, tail-to-tail dimeric, and a mixed form dimer.[61] For example, the crystal structure of *E. coli* GluR presents the molecule as an activator-modulated monomer (Figure 1.13). In the activator-free state, *E. coli* GluR shows very low activity for either of the substrate enantiomers. However, when UDP-MurNAc-Ala (Figure 1.13), a product of the peptidoglycan biosynthetic pathway, is added as activator, the activity of GluR is 1,000-fold increased.[62] These results suggest that GluR in *E. coli* plays an important role as feedback regulator.[63] When the concentrations of L-glutamic acid and UDP-MurNAc-Ala are high, the activator binds to GluR, thus catalyzing the conversion of L-glutamic acid to D-glutamic acid. On the other hand, when the concentration of D-glutamic acid is high and the concentration of UDP-MurNAc-Ala is low, there is not enough activator to bind to GluR, and the enzyme remains deactivated.[60,64]

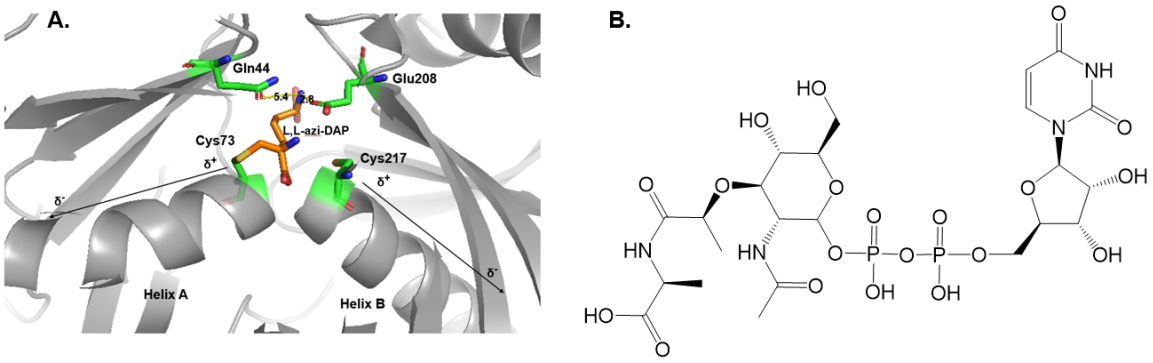


Figure 1.13. The crystal structure of GluR from *E.coli* (A) and the chemical structure of UDP-MurNAc-Ala (B). [61]

In contrast, molecules of GluR from *Helicobacter pylori* arrange in a classical dimeric structure without activator modulation (Figure 1.14). Instead of modulation by UDP-MurNAc-Ala, GluR in *H. pylori* controls epimerization activity by substrate inhibition. GluR has an asymmetric Michaelis constant (Km) for two substrates; Km for D-glutamic acid (63 μM) is around 12-fold lower than L-glutamic acid (740 μM). This asymmetry suggests D-glutamic acid acts as an inhibitor when the concentration of D-glutamic acid is too high. GluR in *H. pylori* has a homo-dimeric structure, with each monomer contributing one active site cysteine. The two active site cysteines are aligned in a face-to-face orientation, creating a pocket that is shielded from the solvent. [61]

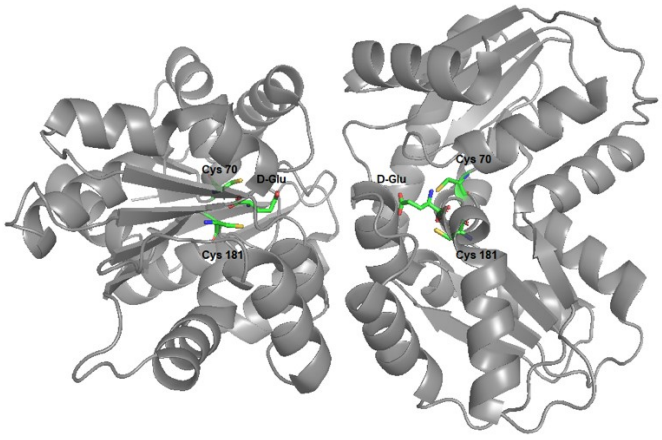


Figure 1.14. The crystal structure of GluR from *H. pylori*. [61]

The third variation on structural assembly of GluR can be found in *Staphylococcus aureus, Enterococcus faecalis* and *Enterococcus faecium*. These Gram-positive species possess similar dimeric structures as *H. pylori* GluR and are unaffected by UDP-MurNAc-Ala activation. However, in contrast to *H. pylori* GluR, each monomer from these species dimerizes in the tail-to-tail orientation, and the active site cysteines are placed on opposite sides. Helices of each monomer create an oligomerization interface (Figure 1.15). This results in an active site pocket that is exposed to the solvent and the enzyme interacts with various salts to initiate a hinge movement of the two subunits. Most GluR structures from Gram-positive species have been shown to oligomerize tail-to-tail. [60, 65] The GluR homologues in Gram-positive species show less substrate inhibition, possibly because Gram-positive species require more D-Glu for a thicker peptidoglycan layer than Gram-negative species. [61]

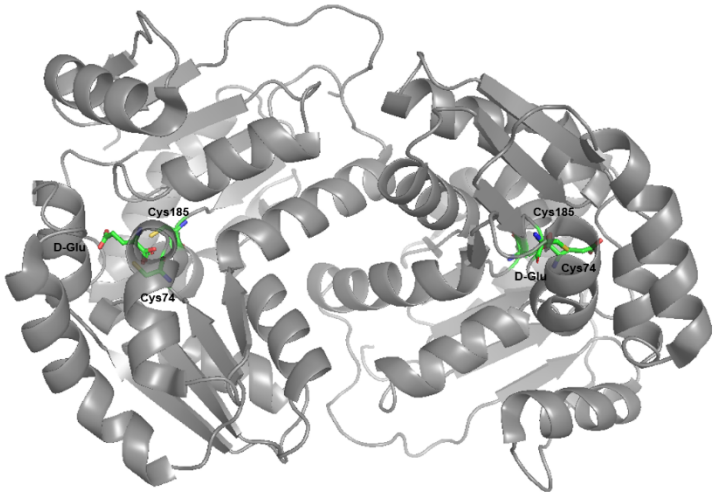


Figure 1.15. The crystal structure of GluR from *E. faecalis*.[61]

1.3.3.3. Mechanism of GluR

The first reported crystal structure of a PLP-independent racemase was glutamate racemase from *Aquifex pyrophilus* by Hwang *et al.*[66] They showed that GluR exists as tightly associated dimer and each monomer contributes one cysteine residue to the active site (Cys70, Cys178). In the structure, the active site is surrounded by various amino acids: Cys 70 is enclosed by Asp7, Ser8, Tyr39, Thr72, and Ala73. When the substrate L-Glu binds to the active site, the carboxyl group of L-Glu is stabilized by a water-mediated interaction with O_δ of Asp7, and hydrogen bonding with O_γ of Ser8. The main-chain amino group of L-Glu is close to O_δ of Asn7. On the other hand, Cys178 is enclosed by Asn71, Thr114, Thr117, Cys139 and Glu147. When D-Glu approaches the active site, the carboxyl group of D-Glu maintains a water-mediated interaction with O_ε of Glu147 and hydrogen bonding with O_γ of Thr114. Identical to L-Glu, the main-chain amino group of D-Glu is in close proximity to O_δ of Asn71 (Figure 1.16).

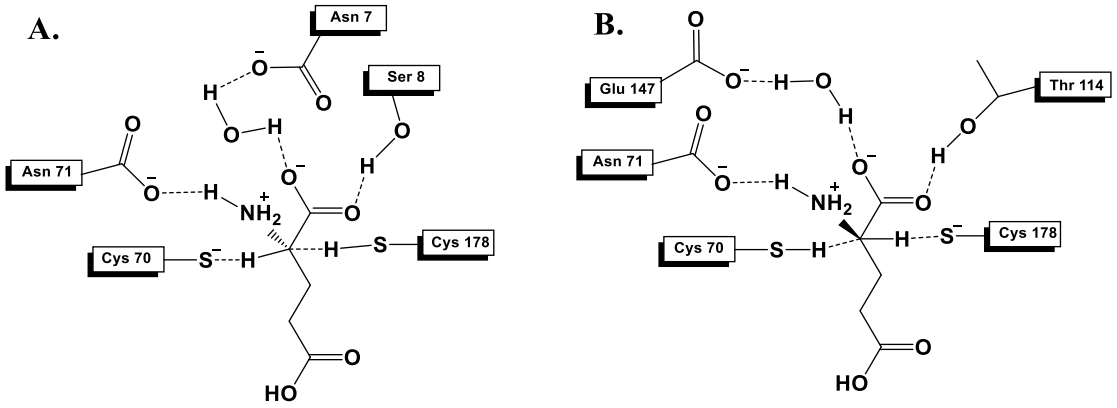


Figure 1.16. The active site of GluR from *A. pyrophilus*.[66]

These interactions between substrate and amino acid residues are considered crucial to recognize the substrate. However, it remained a mystery how the *p*Ka of the α-hydrogen of glutamine can be so fundamentally reduced as to enable deprotonation by an active site cysteine residue.[48] To solve this, Glavas and Tanner performed sequence alignment of 13 distinct isozymes of GluR, and they found that four residues in GluR (Asp10, Asp36, Glu152, and His186 in *La*GluR) are strictly conserved (Figure 1.17).[67] These amino acids seemingly have significant roles to assist the active site cysteines (Cys73 and Cys184 in *La*GluR) in the deprotonation and protonation of glutamate.

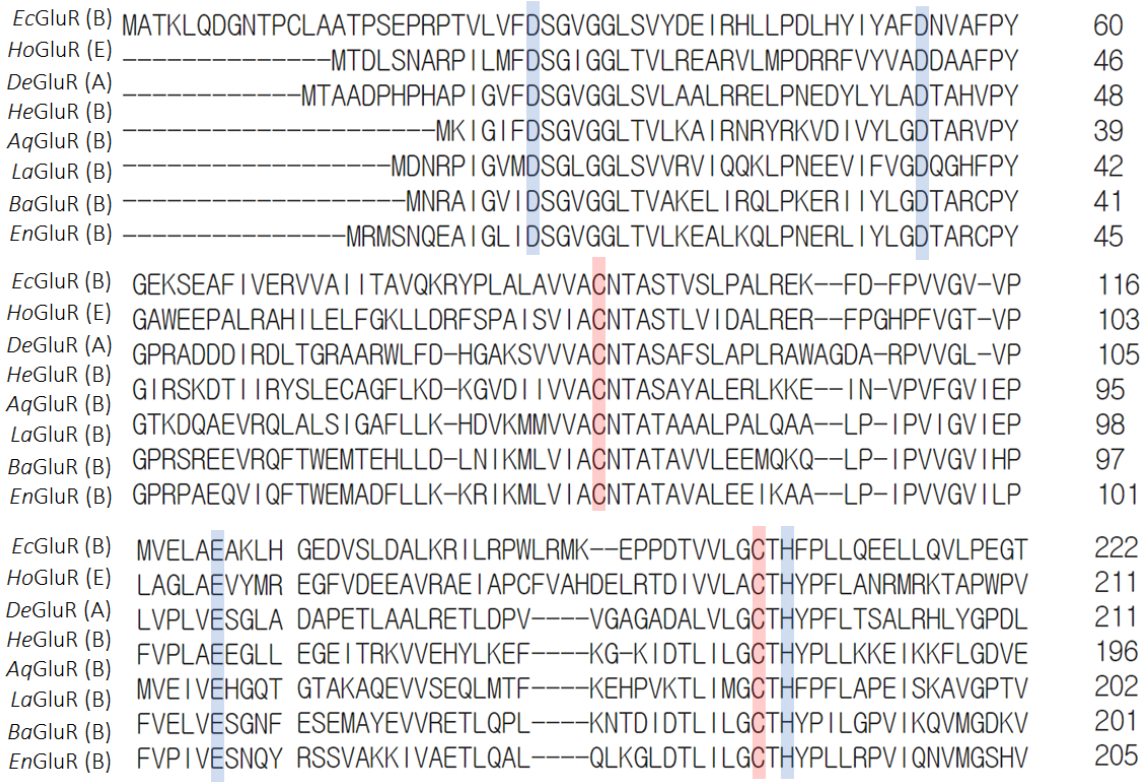


Figure 1.17. Amino acid sequence alignment of GluRs from *Escherichia coli* (*Ec*GluR), *Homo sapiens* (*Ho*GluR), *Deinococcus maricopensis* (*De*GluR), *Helicobacter pylori* (*He*GluR), *Aquifex pyrophilus* (*Aq*GluR), *Lactobacillus fermenti* (*La*GluR), *Bacillus anthracis* (*Ba*GluR), and *Enterococcus faecalis* (*En*GluR). The active site cysteine residues (red) and other conserved catalytic residues (blue) are highlighted.

Kinetic isotope effects of *La*GluR mutants show that Asp10 and His186 support the active site cysteines in deprotonating the α-hydrogen of glutamate. The ratio of the isotopic effect (V_{max} D→L/L→D) of a D10N mutant appears high, which is similar to the C73S mutant. In other words, Asp10 may strongly associate with Cys73, and it assists Cys73 in the deprotonation. On the other hand, in the case of the H186N mutant, the isotopic effect (V_{max} D→L/L→D ratio) is small, much like the C184S mutant. This indicates that His186 assists Cys184 in the deprotonation of the α-hydrogen in D-glutamate (Figure 1.18). The authors explain that the $O_{\delta}1$ of Asp10 is associated with sulfur atom of Cys74, and Asp10 acts as a general base to deprotonate Cys74 in order to increase the *p*Ka of cysteine. This interaction helps to deprotonate the α-hydrogen, which normally has a much higher *p*Ka (~29). Similarly, His 186 acts as a general base to deprotonate Cys184 by interaction between $N_{\delta}1$ atom in His 186 and sulfur atom in Cys184.[68]

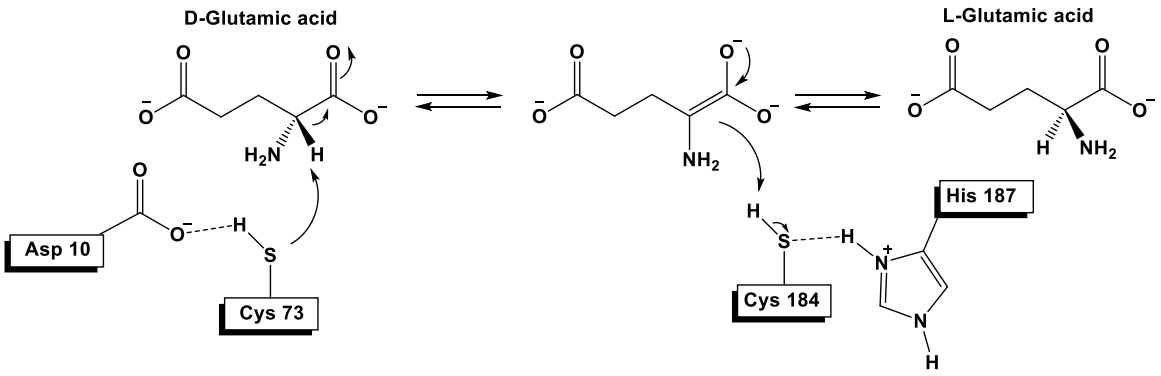


Figure 1.18. Potential roles for residues assisting as catalytic acid/base in GluR.

The crystal structure of *Bacillus anthracis* GluR (*Ba*GluR) supports the concept of residue assisted catalysis.[65] The distance between sulfur atom of Cys74 (Cys73 in *La*GluR) and Oδ1 atom of Asp11 (Asp10 in *La*GluR) is 4.0Å – close enough to allow interaction between them. In addition, the distance between the sulfur atom in Cys184 and the His186-

Nδ1 atom is 4.3 Å, also allowing for a stabilizing interaction between the two residues (Figure 1.19).

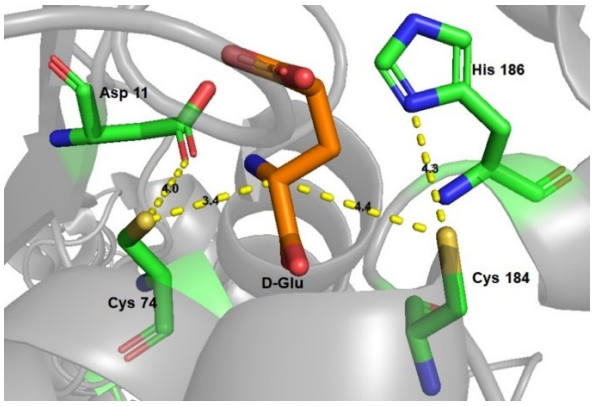


Figure 1.19. The active site of GluR from *B. anthracis*.[65]

1.3.3.4. Inhibition strategies of GluR

Because of the importance of GluR in making D-Glu for bacterial cell wall synthesis, design of efficient GluR inhibitors is a promising target for antibiotic development. The first generation of GluR inhibitors were glutamic acid analogues. The first GluR inhibitor, L-serine-O-sulfate, was irreversible, unstable, and less efficient for the inhibition process.[69] In 1994, aziridino glutamate was reported, which overall has a similar structure to common DapF inhibitors. However, this inhibitor was also unstable, and inhibition efficiency was low.[70] In 2002, Lilly Research Laboratories developed a new class of GluR inhibitors, 4-substituted D-glutamic acid analogues.[71] In this series, 2-naphthylmethyl-D-glutamate is one of the most potent compounds, showing high inhibition efficiency (K_i=16 nM, Figure 1.20).[53] In 2007, the crystal structure of D-2-naphthylmethyl-glutamate-inhibited GluR of *Streptococcus pyogenes* was reported.[72] According to the crystal structure, several amino acids in GluR form hydrogen bonds with the inhibitor, and the naphthyl moiety of the

inhibitor is located in a hydrophobic pocket, resulting in a good inhibitor fit. Another type of GluR inhibitors are non-amino acid analogues. A variety of these inhibitors have been conveniently reviewed by Conti et al.[46] The major difference between amino acid analogue inhibitors and non-amino acid analogue inhibitors is the binding site. Normally, amino acid analogue inhibitors bind to the active site of glutamate racemase, and act as competitive inhibitors or irreversible inactivators. However, non-amino acid analogue inhibitors often target allosteric binding pockets of the enzyme and change the shape of the active site, which then causes a change in its activity.[46]

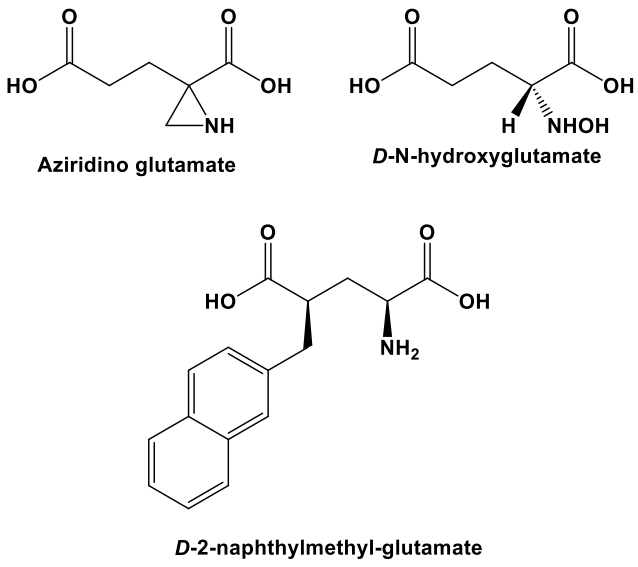


Figure 1.20. Glutamate analogue inhibitors of GluR.

1.3.4. Aspartate Racemase (AspR)

1.3.4.1. Function and localization

In nature, D-aspartate can be found in mammalian neural and endocrine tissues. Specifically, peptidic D-aspartyl residues are interpreted as markers of aging [73] and Alzheimer's disease, which is related to the accumulation of D-aspartate in the brain *β*-amyloid protein.[74] D-aspartate residues originate predominantly from the action of PLP-dependent aspartate racemase (AspR), however PLP-independent AspR enzymes have been found in lactic acid bacteria, including *Streptococcus* and *Lactobacillus* species. In 1991, Yohda and his coworkers first purified AspR from cell-free extracts of *Streptococcus thermophiles*.[75] Bacterial D-Asp is an integral component of the peptidoglycan crossbridge, and increases the resistance of lactic acid bacteria towards proteolytic degradation.

1.3.4.2. Structure and mechanism

In 2002, the first X-ray structure of an AspR was reported by L. Liu and coworkers.[41] The enzyme from *Pyrococcus horikoshii* (*Ph*AspR) has a stable dimeric structure with homologous α/β domains that each consist of four-stranded parallel β-sheets flanked by six α-helices. These two subunits face each other with a disulfide bond between Cys72 from each domain creating twofold symmetry (Figure 1.21).

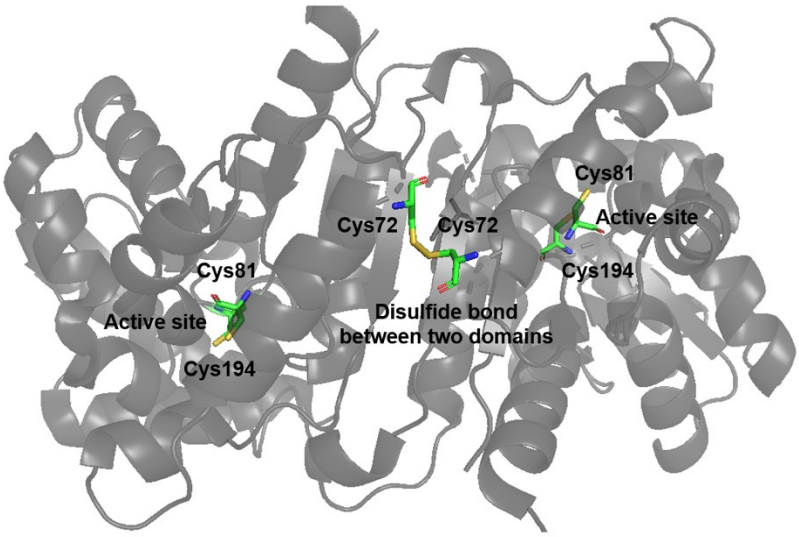


Figure 1.21. The crystal structure of AspR from *Pyrococcus horikoshii.*[41]

Similar to other PLP-independent racemases, *Ph*AspR has two active site cysteine residues (Cys82 and Cys194) in each subunit, supporting the "two base" mechanism (Figure 1.21). The active site of the enzyme is composed of Arg48, Pro81, Cys82, Asn83, Thr84, Lys164, Gly193, Cys194, Thr195 and 5 water molecules, which are highly conserved among AspR homologues (Figure 1.22). These residues have important roles to recognize the substrate and create a positively charged pocket to accept acidic substrate (Figure 1.23). The catalytic Cys residues have a 9.6 Å distance between the two γ-sulfur atoms, which creates a large dimerization interface of 2540 $Å^2$ that could be referred to as the "open form" of the enzyme (Figure 1.23).

Figure 1.22. Amino acid sequence alignment of AspRs from *Aspergillus niger* (*Ac*AspR), *Streptococcus thermophilus* (*St*AspR), *Thermococcus paralvinellae* (*Tp*AspR), *Pyrococcus horikoshii* (*Ph*AspR) and the bifunctional L-**DER racemase from *Escherichia coli* (*Ec*L-DER). The active site cysteine residues (red) and other conserved catalytic residues (blue) are highlighted.**

The inactive C82A mutant of *Ph*AspR with citric acid (aspartate analogue) was crystallized, and it showed the whole enzyme exists in a much more contracted form in this case.[76] With the binding of substrate, the enzyme undergoes hinge movement,[77] causing a shortening of the thiolate distance to an estimated 7.4 Å and reducing the size of the active site cavity. However, since the active site Cys82 is mutated to Ala, only Cys194 in the active site is near the citric acid, which has a binding mode that resembles both L- and D-Asp (Figure 1.23). Therefore, the authors suggest an alternative "one base" mechanism for the Asp racemization involving only Cys194. However, a recent X-ray structure of *Picrophilus torridus* AspR (*Pt*AspR) in complex with tartrate[78] led to revision of this proposal. The enzyme structures with both L- and D-Asp show that only the Cα-position of

the substrate is substantially different, supporting the traditional "two base" mechanism. Interestingly, only one of the two active sites of the dimer is ligand-occupied in this structure.

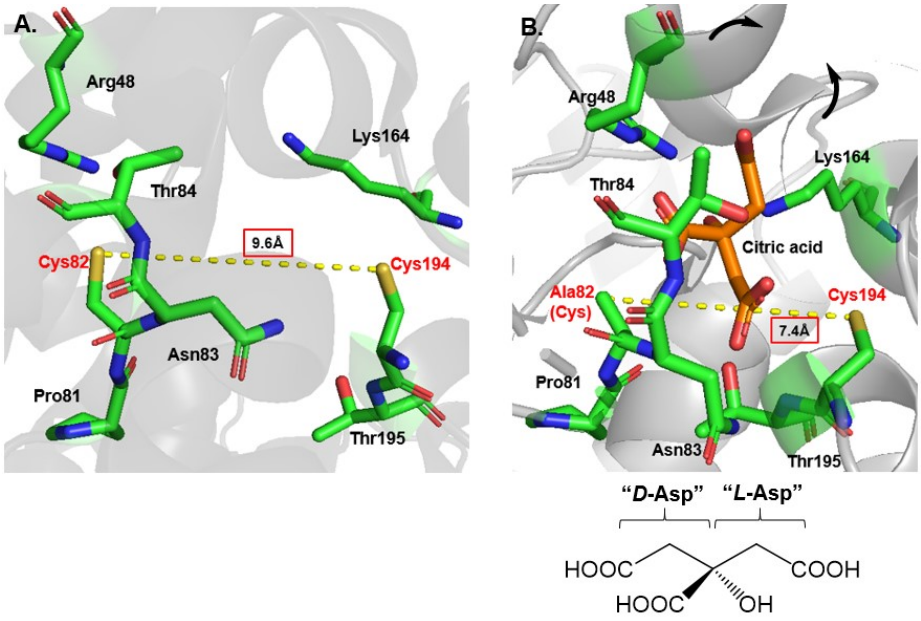


Figure 1.23. A. Active site geometry of *Ph*AspR. B. Active site geometry of citrate bound to inactive C82A mutant *Ph*AspR. The hinge movement is indicated by two arrows.[78]

AspR shows about 26% sequence identity with *A. pyrophilus* GluR, despite the relatively similar overall fold of the two protein families. Because the precise arrangement of the active site residues results in strict specificity recognition of the substrate, GluR exclusively isomerizes D- and L-Glu[79], and AspR acts solely on D- and L-Asp. However, a series of new X-ray structures of racemases from *E. coli* have been reported indicating higher substrate promiscuity. Interestingly, these racemases do not have cysteine pairs at the active site, but rather one of the active site cysteines is exchanged for a threonine

residue.[80,81] Originating from a pathogenic *E. coli* strain, this version of a mixed aspartate/glutamate racemase (EcL-DER) has an unbalanced catalytic pair (Thr83 and Cys193). This enzyme is able to convert both L-Asp and L-Glu, with a 3-fold higher activity for L-Glu.[20] The slightly different orientation of L-Asp and D-Asp in these X-ray structures supports the observed mono-directionality (L→D direction) of this transformation (Figure 1.24). Although the distance between the α-carbon of D-Asp and Thr83 (3.3 Å) is closer than that of L-Asp and Cys197, D-Asp is not racemized due to presence of threonine instead of cysteine. In addition, this exchange may cause less specificity in recognition of the substrate, as EcL-DER can recognize both L-Asp and D-Asp.[77]

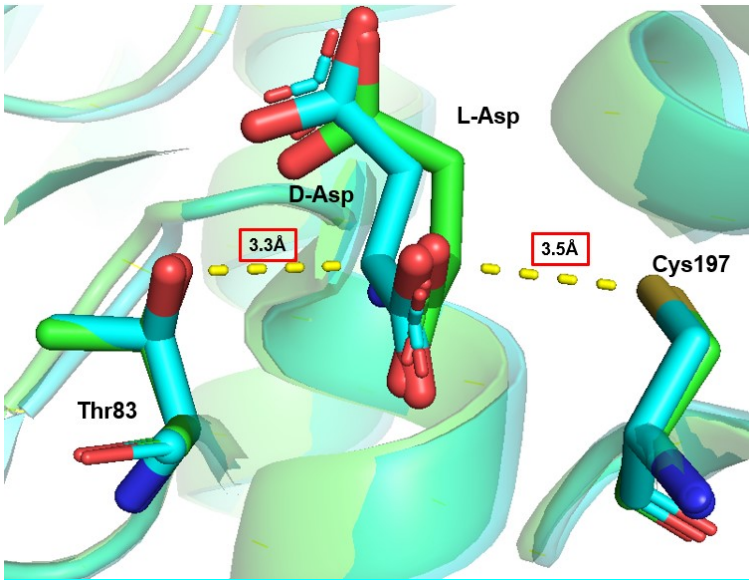


Figure 1.24. Superimposition of substrate molecules in active site of *Ec*L-DER.[79]

1.3.4.3. Inhibition strategies

Surprisingly, there is currently a lack of systematic inhibition studies on PLP-independent AspR. General inhibitors for cysteine residues have been shown to be potent

agents against monomeric AspR from *Bifidobacterium bifidum*. For example, both N-ethylmaleimide and 5,5'-dithiobis(2-nitrobenzoate) inhibit *Bb*AspR completely at 1 mM.[46]

1.3.5. Proline racemase (ProR)

1.3.5.1. Function and localization

D-proline, which is recently reported as an essential amino acid for isomeric adipokinetic neuropeptide hormones in various species of cicadas,[82] is produced by proline racemase (ProR). This enzyme has been demonstrated to have various roles in nature. Firstly, ProR is an effective mitogen for B cell activation, directly involved in Chagas disease caused by the parasite *Trypanosoma cruzi*[83] and the causative agent in livestock trypanosmosis by *Trypanosoma vivax*.[84] Secondly, ProR contributes to evasion of the immune system by the parasite and persistence in the human host through non-specific lymphocyte activation. The conversion of L→D-proline is also a key step in certain *Clostridia* species (Stickland) reaction.[85] D-proline is also found in certain marine antimicrobial cyclopeptides.[86] ProR has been identified in both bacteria and eukaryotes, and recently there is evidence for it in archaea.[87]

1.3.5.2. Structure and mechanism

In 1968, ProR from *Clostridium sticklandii* (*Cs*ProR) was the first isolated and structurally characterized proline racemase.[88,89] The first eukaryotic ProR was identified in the human parasite *Trypanosoma cruzi* (*Tc*ProR), and the X-ray structure of *Tc*ProR was

reported in 2006.[42] Bacterial and eukaryotic structures of ProR are very similar, with 52% homology between *Tc*ProR and *Cs*ProR.[89] According to the sequence alignment, many residues in the catalytic site are conserved throughout different species (Figure 1.25), whereas significant differences are only found in the C- and N-terminal domains.

```
CgProR (E)          GEPRGHYEMFGAIIVSPDIESCDFACIFIHNEGYATMCGDAVISLMR 57
TlProR/HypE(A)      HEPRGHSDQFGAVLVPS--DIADFGVIYMDTSGYLDMCGHATMGVAT 63
FaProR (A)          LEPRGHKDQFGAILLPPVNPDSDYTVIYPTTESYLDMCGHATIGVST 70
TcProR (E)          LEPRGHDDMFGAFLFDPIEEGADLGIVFMDTGGYLNMCGHNSIAAVT 107
CdProR (B)          LEPRGHNDMFGSVMTQPCCPDADFGIIFMDGGGYLNMCGHGTIGAMT 68
SmHypE (B)          FEPRGHDMMSGSILYPPTRPDCDVAVLFIETSGCLPMCGHGTIGTIT 67
hsT3LHypD (E)       FEPRGHRDMYGAVLVPSELPDAHLGVLFLHNEGYSSMCGHAVLALGR 81

CgProR (E)          VYGTILTDG-NDDYSEEITDTMCVYAEKQVDRSPCGSGTSSRIA    224
TlProR/HypE(A)      INLAMLTD--EPEREDSDGKNVVIWGEGSVDRSPCGTGSASRVA    229
FaProR (A)          EPLAMITD--N----NHDYRAIVTFAGNSFDRSPCGTGTAARAA    228
TcProR (E)          VDCVEIYG--PPTNPEANYKNVVIFGNRQADRSPCGTGTSAKMA    278
CdProR (B)          VDLVEIYD--EPTHPEATYKNVVIFGQGQVDRSPCGTGTSAKLA    233
SmHypE (B)          LSHIQWTG--KPTQPEAHARNAVFYGEKAIDRSPCGTGTSARIA    231
hsT3LHypD (E)       LYGTILTDG-KDAYTKEPTTNICVFADEQVDRSPTGSGVTARIA    252
```

Figure 1.25. Structural comparison of amino acid sequences between ProRs from different genera and hydroxyproline-2-epimerase (HypE). Active site cysteines are highlighted in red, other highly conserved residues are highlighted in blue.

Similar to the structure of DapF, ProR is also a homodimer comprised of two α/β units with a cysteine pair at the end of an α-helix, which occurs in the active site of each monomeric unit. The X-ray structures of ProR from *E. coli* with pyrrole-2-carboxylic acid (proline analogue) as the "substrate" shows a network of 7 hydrogen bonds involving the residues Gly131, His132, Asp296, Gly301 and Thr302 (Figure 1.26) [87].

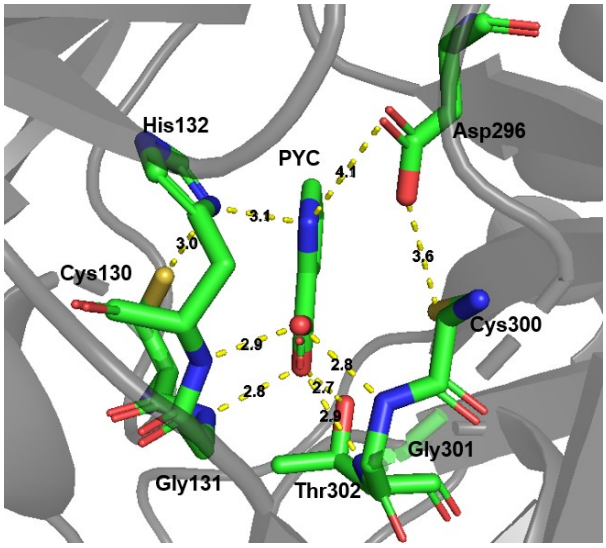


Figure 1.26. Active site of ProR from *E. coli*.[42]

According to QM/MM studies,[90] there are two possible explanations for the racemization mechanism of ProR using thiol-thiolate groups (*p*Ka~8.4): One is that hydrophobic environment modulates the deprotonation of the thiol group, whereas the other is that the amine group of the substrate is involved in the removal of the proton from thiol (Figure 1.27). A computed energy difference for the catalytic effect of 14 kcal/mol causes a reduction of the internal *p*Ka of the α-hydrogen of proline by more than 7 units (*p*Ka~23). This reduction changes the electrostatic field in the closed form of the enzyme. The substrate binding causes a hinge-movement of the two domains, thus the water from the active site is expelled. These conformational changes are supported by recent MD calculations which suggested a reduction of exposed surface area from about 6000 $Å^2$ (open form) to 1500 $Å^2$ (closed form).[91] In addition, nearby His132 and Asp296 residues assist Cys130 and Cys300 in performing alternate acid and base functions on the α-proton of the substrate, similar to GluR (Figure 1.18 and Figure 1.26). These tight electrostatic interactions seem to prevent proline from reorienting in a position that enables ammonium

deprotonation, which should thermodynamically be favored over the α-carbon deprotonation (*p*Ka~9.6 vs. *p*Ka~23). Alternatively, water could act as an external base to help with the proton transfer between Cys130 and Cys300, considering a *p*Ka value of about 8 for both cysteines and the fact that ProR operates at pH 8. The intricate and Taylor-made geometric and electrostatic arrangement of the active site pocket guarantees the sole recognition of proline and allows no known substrate alternatives.

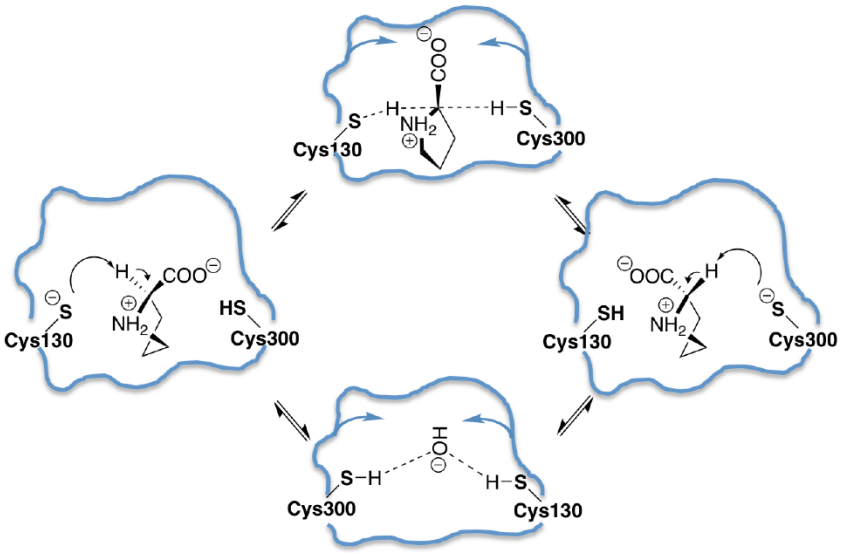


Figure 1.27. Putative catalytic pathways of *Tc*ProR.

1.3.5.3. Inhibition strategies

Recently, new therapeutic strategies and novel drugs addressing the inhibition of ProR have been suggested to treat Chagas disease caused by *Trypanosoma cruzi*. This is one of the more neglected diseases in several regions of Latin America, with at least 13 million people currently infected.[92] Until recently, only two nitroimidazole based drugs, nifurtimox and benznidazole, were used to treat infected patients (Figure 1.28). Both are

connected to major complications like agranulocytosis, thrombogenic purpura and convulsions, and are most effective in the early stages of infection. However, the majority of patients are diagnosed once they reach the chronic phase and have serious side effects, such as liver and kidney failure for benznidazole and of psychiatric and neurological disorders for nifurtimox.[93] The inhibitor pyrrole-2-carboxylic acid (PYC, Ki~0.01mM) is known to block the catalytic crevice, ultimately precluding ProR interaction with B-cells and reducing the number of infected host cells *in vivo*. However, PYC has poor water solubility, thus limiting its therapeutic use.[94] Recently, Minoprio *et al*. suggested several pyrazole analogues as potential new inhibitors (K_i ~0.3-2 mM). Even though some analogues showed better water solubility, none showed greater affinity for ProR than PYC.[95] Derksen and Bearne synthesized tetrahydro-1H-pyrrolizine-7a(5H)-carbonxylate (K_i = 111 mM) and 7-aza-bi-cyclo[2.2.1]heptan-7-ium-1-carboxylate (K_i ~350 mM) which have extended ring systems in the bicyclic compounds (Figure 1.28), but these compounds also exhibit weak inhibition. These results demonstrate the limited plasticity of the ProR active site and often observed failure of the classical docking approach.

By means of the "path optimization and exploration" approach (POE), which encompasses molecular dynamics and transition path calculations, the Minoporo group was able to suggest the path between open and closed forms of the ProR. Based on the available crystal structure data, the movements upon ligand binding were calculated, thereby assisting new inhibitor design.[87] The authors virtually screened the docking of a 661,000-compound library, resulting in two hits that were 2- and 4-fold more potent than PYC, The key compounds are (E)-4-oxopent-2-enoic acid and its derivative (E)-5-bromo-4-oxopent-2-enoic acid (Figure 1.28). A dose-dependent reduction in host cell infection and a lower

mean number of parasites per *T. cruzi* infected host cells (14% and 34% lower compared to PYC at 10 μM) demonstrates the *in vivo* potency of these compounds.

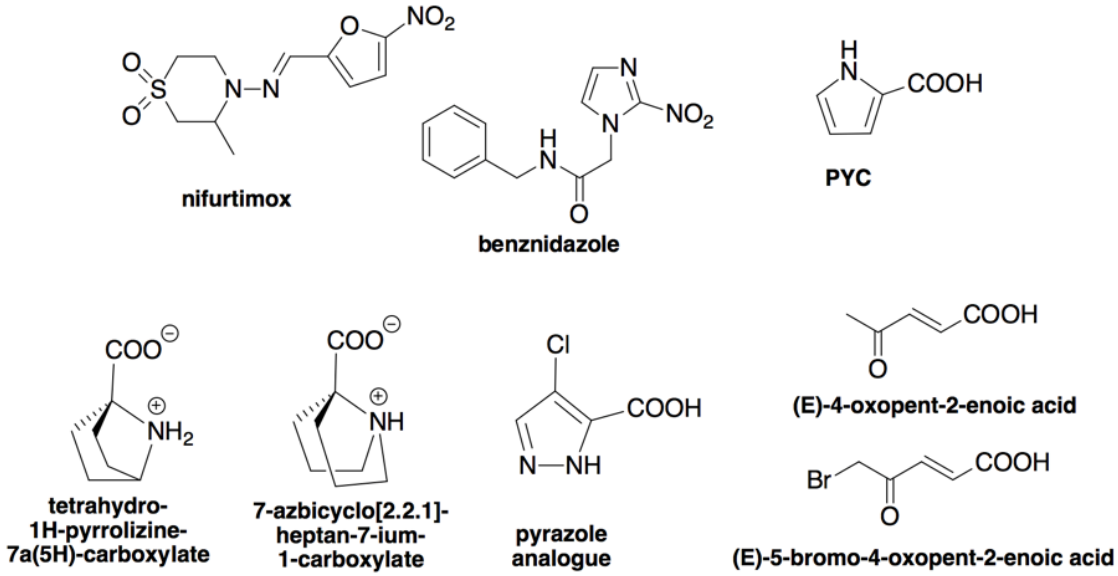


Figure 1.28. Structures of selected ProR inhibitors.

Chapter 2

Mechanistic studies of O-ureidoserine racemase (DcsC)

2.1. D-Cycloserine

2.1.1. General introduction

D-cycloserine, which was first discovered in 1954, is a cyclic analog of the amino acid serine.[96] It is produced by various *Streptomyces* species, including *Streptomyces lavendulae* and *Streptomyces garyphalus*.[97] D-cycloserine has the brand name Seromycin, and is used as an antibiotic for tuberculosis caused by *Mycobacterium tuberculosis*. D-cycloserine is classified as a second-line drug because of numerous neurological side effects.[98] D-cycloserine can penetrate the central nervous system and cause paresthesia, headaches, drowsiness, depression, dizziness, vertigo, confusion, dysarthria, hyperirritability, psychosis, convulsions, and shaking.[99] Also, D-cycloserine can inhibit PLP-dependent enzymes, such as γ-aminobutyric acid aminotransferase in biosynthesis pathway of the neurotransmitter γ-aminobutyric acid (GABA).[100] For these reasons, D-cycloserine is used when the first-line drugs including isoniazid (INH), rifampin (RIF), ethambutol (EMB), and pyrazinamide (PZA) are no longer effective.[101] Even though D-cycloserine has disadvantages as an antibiotic for tuberculosis, the market size of D-

cycloserine is growing and its price has increased from $30 to $1,050 per month of treatment in 2015.[102]

2.1.2. The mechanism of D-cycloserine inhibition

D-cycloserine inhibits alanine racemase (AlaR) and D-alanyl–D-alanine ligase, which are essential enzymes for peptidoglycan biosynthesis in bacterial cell walls (Figure 1.1).

2.1.2.1. D-Cycloserine as an alanine racemase inhibitor

Alanine racemase, which catalyzes racemization of L-alanine and D-alanine, is irreversibly inhibited by D-cycloserine. When the D-cycloserine is added to AlaR, PLP is covalently bound to D-cycloserine and tautomerization to an isoxazole occurs (Figure 2.1).[46, 103] The crystal structure of AlaR in *Peptoclostridium difficile* helps explain this inhibition mechanism (Figure 2.2).[104] Similar to the crystal structure of AlaR with the alanine substrate and PLP, D-cycloserine bound to AlaR shows that the phosphate group of PLP is stabilized by hydrogen bonding with hydroxyl groups on Ser208 (2.4 Å), Tyr43 (2.5 Å), and Tyr356 (2.5 Å). Interestingly, in the active site of AlaR, Lys39 (3.1 Å) and Tyr268 (3.7 Å) have a short distance to the α-carbon of D-cycloserine. This indicates that AlaR recognizes D-cycloserine as substrate and makes the D-cycloserine-PLP complex at the active site. Enzyme-involved tautomerization gives an aromatic isoxazole that no longer has a hydrolyzable imine bond (Figure 2.1).[46, 103] Tyr287 also helps to recognize D-cycloserine by establishing a hydrogen bond with the oxygen in the cycloserine ring.[105] Therefore, it is suggested that D-cycloserine acts as an affinity label inhibitor for AlaR (Figure 2.2).[103, 106]

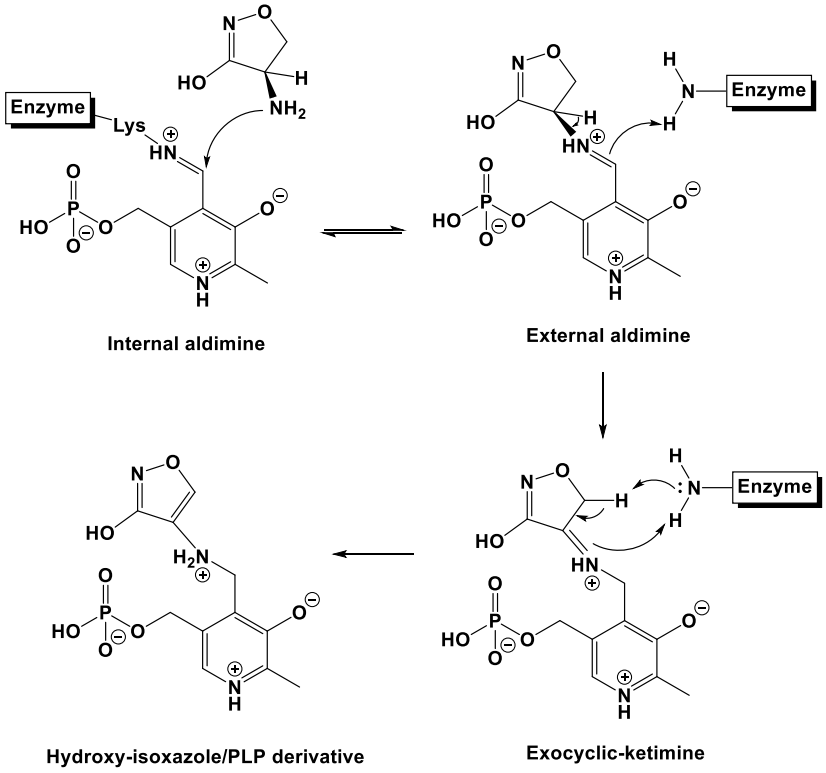


Figure 2.1. Mechanism of inactivation of AlaR by aromatization of D-cycloserine.[46]

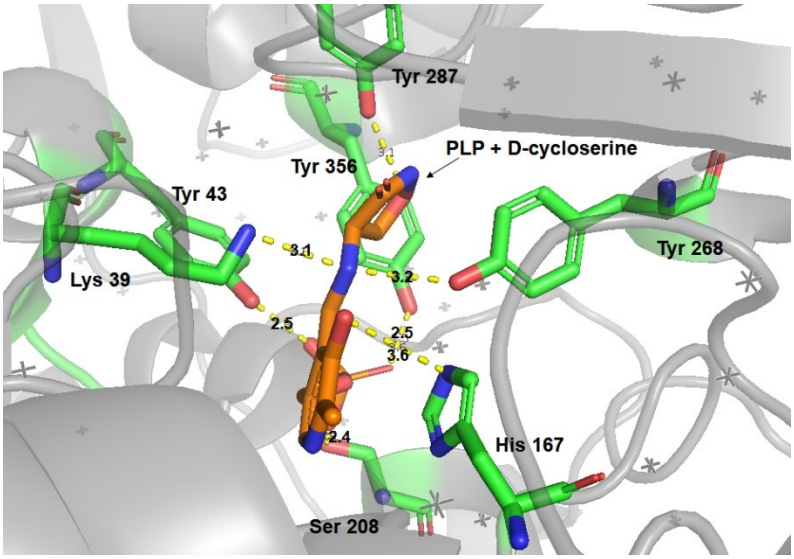


Figure 2.2. Crystal structure of AlaR from *Peptoclostridium difficile* with D-cycloserine.[104]

2.1.2.2. D-Cycloserine as D-alanyl–D-alanine ligase inhibitor

D-alanyl-D-alanine ligase (Ddl) catalyzes the ligation of two D-alanines in the peptide tail of peptidoglycan (Figure 1.1). Without an inhibitor, Ddl recognizes D-alanine and adenosine triphosphate (ATP) to initiate the reaction. The oxygen of the carboxylic acid group in D-alanine attacks the terminal phosphate group of ATP,[105] and the phosphoester bond is synthesized by nucleophilic substitution. Next, the nitrogen from another D-alanine amino group attacks the carbonyl, and an amide bond is established between the two D-alanines (Figure 2.3).[107] According to this mechanism, ATP plays significant roles in enzymatic function, and D-cycloserine binds to ATP to inhibit the Ddl. The amide group of D-cycloserine can be tautomerized to make an iminol. The alkoxide ion of D-cycloserine attacks phosphorous in ATP to synthesize a phosphoester bond.[107] Because of this phosphoester-D-cycloserine complex, the D-alanine cannot react with ATP and Ddl is inhibited (Figure 2.3). The crystal structure of D-cycloserine inhibited Ddl from *E. coli* supports this inhibition mechanism (Figure 2.4).[108] After inhibition by D-cycloserine, only phosphorylated D-cycloserine and ADP are detected. Interestingly, D-cycloserine cannot be placed at the N-terminal D-alanine position, but can be placed C-terminally. In other words, between two substrate binding sites, D-cycloserine is bound to the C-terminus D-alanine binding site to inhibit the Ddl.[109] Therefore, D-cycloserine inhibits Ddl by specifically competing with the C-terminus D-alanine.

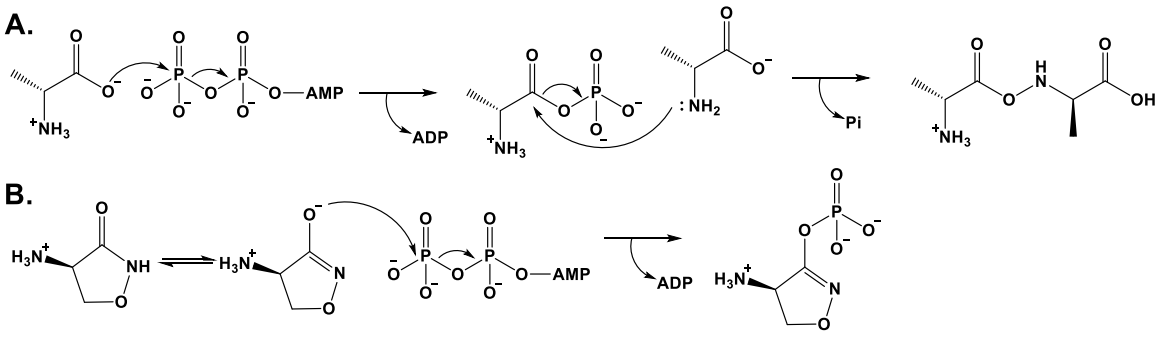


Figure 2.3. D-Alanyl-D-alanine ligase mechanism. A. Mechanism of D-**alanyl-**D-**alanine ligase. B. Inhibition mechanism by** D-**cycloserine.**

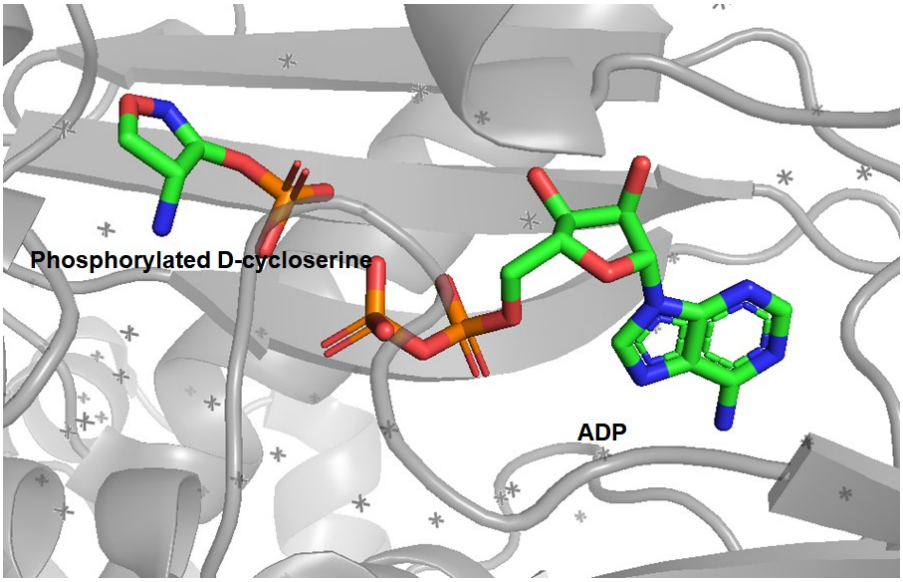


Figure 2.4. The crystal structure of D-cycloserine inhibited D-alanyl-D-alanine ligase from *E. coli*.[108]

2.1.3. Biosynthesis of D-cycloserine

Although D-cycloserine has been used as an antibiotic for a long time and possesses a relatively simple structure, the biosynthesis of D-cycloserine has only recently been explored.[110] Various *Streptomyces* species produce D-cycloserine from L-serine and L-arginine. Six enzymes are proposed to be involved in D-cycloserine biosynthesis in *Streptomyces*; DcsA, DcsB, DcsC, DcsD, DcsE, DcsG (Figure 2.5).[111] First, L-serine is acetylated by acetyl-CoA and L-serine-*O*-acetyltransferase (DcsE), producing *O*-acetyl-L-serine. L-arginine is hydroxylated by *N*-hydroxy-L-arginine synthase (DcsA) to make *N*-hydroxy-L-arginine, followed by N^{ω}-hydroxy-L-arginine amidinohydrolase (DcsB) to synthesize hydroxyurea and L-ornithine. Next, hydroxyurea reacts with *O*-acetyl-L-serine, catalyzed by *O*-ureido-L-serine synthase (DcsD), and *O*-ureido-L-serine is produced. DcsD has a high structural similarity with *O*-acetylserine sulfhydrylase, which is an essential PLP-dependent enzyme for L-cysteine biosynthesis.[112,113] The next reaction, racemization of *O*-ureido-L-serine is catalyzed by *O*-ureidoserine racemase (DcsC), a PLP-independent enzyme. Lastly, *O*-ureido-D-serine is cyclized by *O*-ureido-D-serine cyclo-ligase (DcsG) to produce D-cycloserine. DcsG is suggested to be an ATP-Mg(II)-dependent enzyme with an ATP-grasp fold motif. It can catalyze the formation of other cyclic D-amino acid analogs, such as D-homocysteine thiolactone. Most of the enzyme mechanisms in D-cycloserine biosynthesis are not yet studied, but it is expected that they are similar to known enzymes.

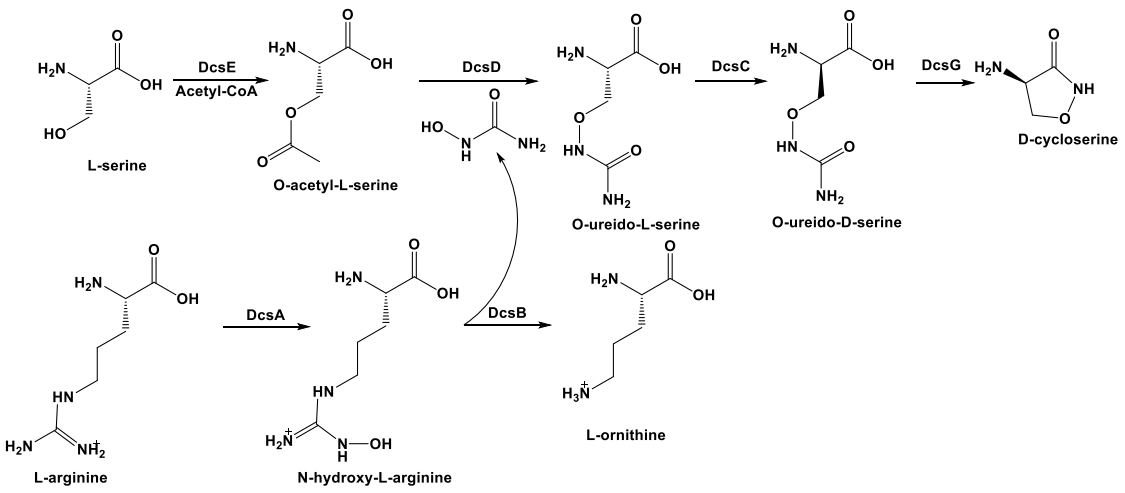


Figure 2.5. Biosynthesis pathway of D-cycloserine.

2.2. Project background

2.2.1 Introduction of DcsC

O-Ureidoserine racemase (DcsC) catalyzes the racemization of *O*-ureido-L-serine and *O*-ureido-D-serine (Figure 2.6). DcsC was reported as a PLP-independent racemase.[114] Previous research from our group showed that DcsC is a divalent metal-independent enzyme since incubation with 1mM EDTA, which is a known metal ion chelating agent, did not cause changes in catalytic performance.[114] Enzymatic activity also remained unaffected by adding PLP or substances that react with it.

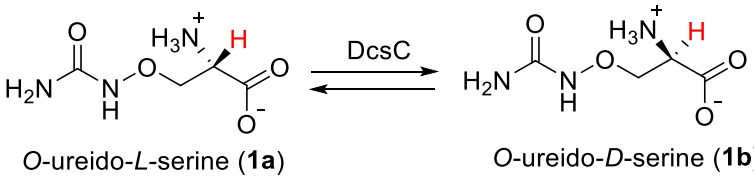


Figure 2.6. Racemization of *O*-ureido-L-serine and *O*-ureido-D-serine by DcsC.

DcsC shows a high degree of homology to diaminopimelate epimerase (DapF), a PLP-independent racemase. The sequence alignment between DcsC and DapF showed that they are 40% identical (Figure 2.7). Similar to DapF, the active site of DcsC has a thiolate-thiol pair, Cys81-Cys227, for the deprotonation-reprotonation at the α-hydrogen of the substrate.

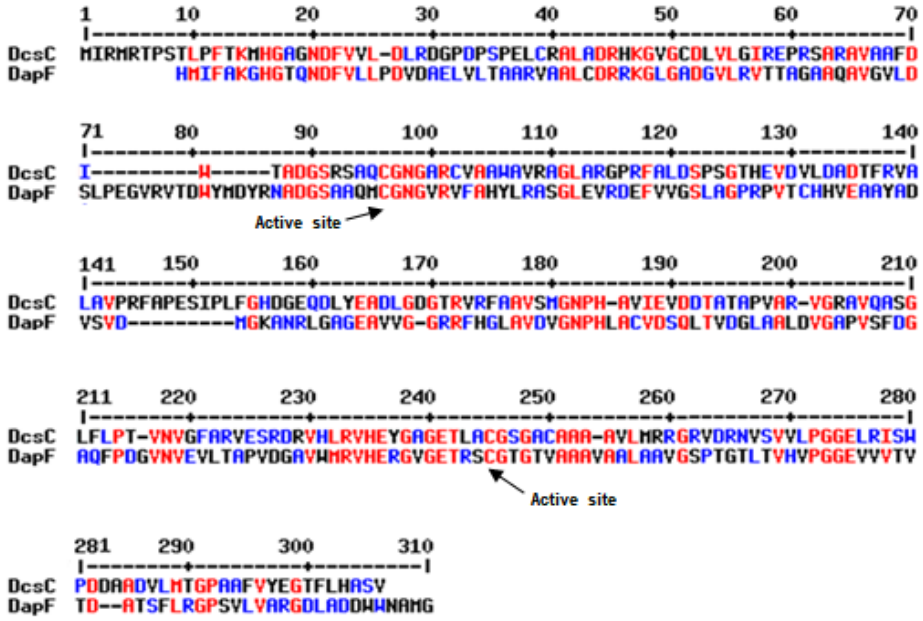


Figure 2.7. Sequence alignment of DcsC and DapF. Red colour shows high sequence similarity, and blue colour shows partial sequence similarity.

Interestingly, many amino acids near the active site are identical between DapF and DcsC, although the structures of substrates for those enzymes are different. This indicates that those amino acids may not be important in the recognition of substrate but rather play important roles in the mechanism of DapF and DcsC. It is expected that DcsC deprotonates-reprotonates the substrate by a thiolate-thiol pair, and the intermediate in this racemization process is stabilized by nearby amino acids in the active site (Figure 2.8).

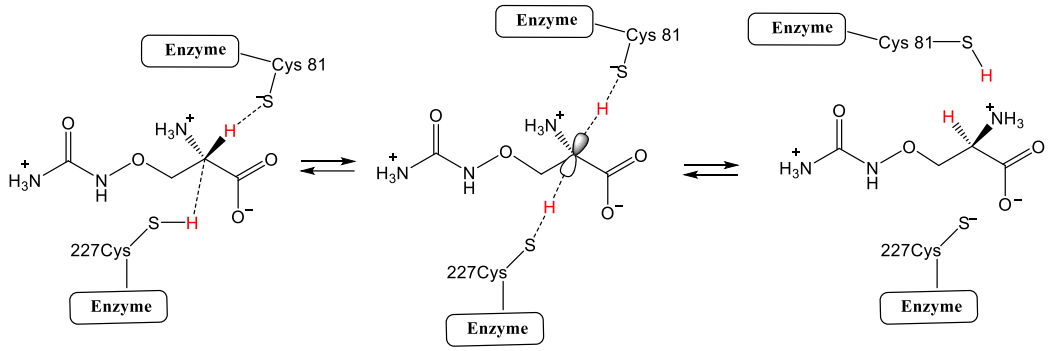


Figure 2.8. Proposed mechanism of DcsC employing a thiolate–thiol pair in the active site.

2.2.2 Inhibition of DcsC

In analogy to aziridine inhibitors for DapF or GluR,[51,54] Dr. David Dietrich of our group examined an irreversible oxirane inhibitor **6** for DcsC.[114] Based on the proposed mechanism of DcsC (Figure 2.8), a suitable substrate involving an electrophile can form a covalent bond with the thiolate in the DcsC active site (Figure 2.9). Since oxirane is a good electrophile to react with a thiolate, oxirane inhibitor **6** could be covalently bound to the DcsC active site.

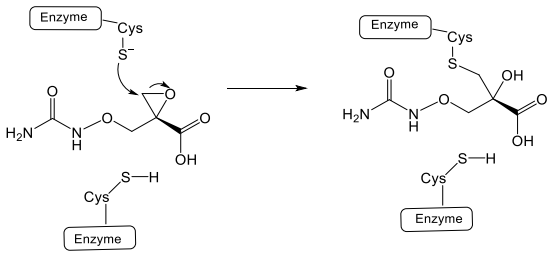


Figure 2.9. Proposed inhibition mechanism of DcsC with oxirane inhibitor.

Dr. Dietrich and coworkers synthesized oxirane inhibitor **6** from ethyl 2-bromomethylacrylate **3**. They first added *tert*-butyl peroxide, using sodium hydride as a

base, to substitute the bromide group. Next, ethyl 2-*tert*-butylperoxomethylacrylate **4** was reacted with hydroxyurea to synthesize the oxirane ring, using potassium *tert*-butoxide (KOtBu) as a base. The ethyl ester of compound **5** was then saponified using lithium hydroxide to produce oxirane **6** (Figure 2.10). Oxirane **6** is quite unstable at room temperature because of the reactivity of epoxide ring, thus it was added immediately to the enzyme solution.

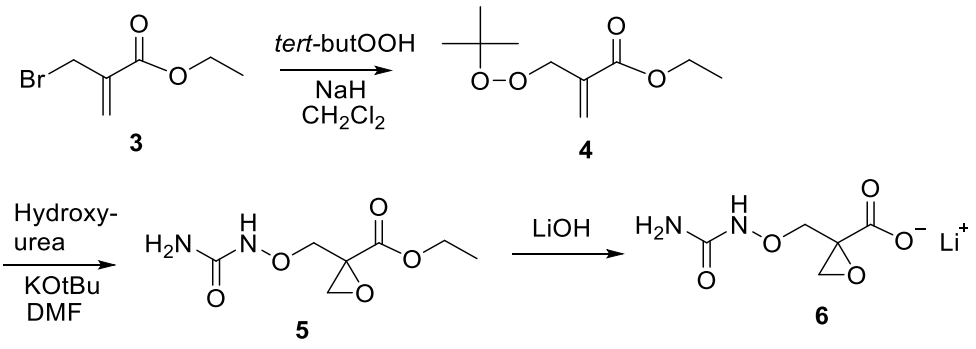


Figure 2.10. Synthesis of the oxirane inhibitor

The binding of inhibitor **6** to DcsC was detected by mass spectrometry, ^{1}H NMR spectroscopy, and circular dichroism (CD). The buffered DcsC-inhibitor solution was incubated for an hour at room temperature, and then analyzed by ESI-TOF-MS. The mass data showed that the enzyme was successfully inhibited by the oxirane inhibitor **6** (Figure 2.11). The m/z ratio of the enzyme with oxirane inhibitor **6** was shown to contain a mixture of two proteins. The difference between the peaks of the two entities peaks is 176 Da, which is the mass of oxirane inhibitor **6**. It indicates that the first peak is the native DcsC, while the second peak is inhibited DcsC.

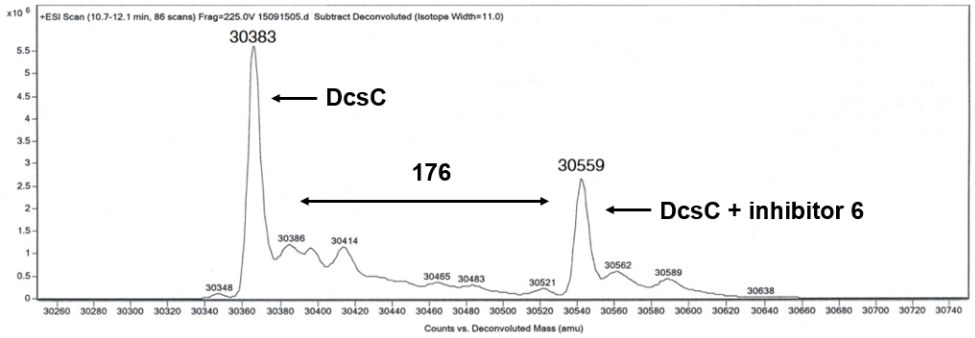


Figure 2.11. Electrospray mass spectra of covalent attachment of epoxide inhibitor 6.

The ^{1}H NMR data also suggested the inhibition of DcsC. When the substrate *O*-ureidoserine was added to a DcsC solution in D_2O, the α-hydrogen is exchanged with deuterium by DcsC. The disappearance of the α-hydrogen signal in the ^{1}H NMR spectrum indicates that DcsC is active. However, when the oxirane inhibitor **6** is added to the DcsC solution in D_2O, no introduction of deuterium to the substrate was detected (Figure 2.12). In agreement with ^{1}H NMR results, when *O*-ureidoserine is incubated with a DcsC solution, an ellipticity change was detected at 206 nm using CD. However, this ellipticity change was not detected when oxirane inhibitor **6** was added. These experiments support the covalent attachment of the inhibitor at cysteine at the active site. [114]

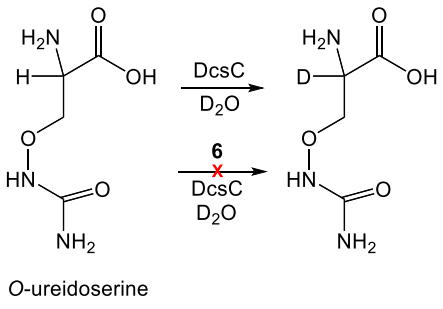


Figure 2.12. Deuterium exchange of *O*-ureidoserine by DcsC without and with oxirane inhibitor 6.

2.2.3. Objectives of the research

Previous research on DcsC suggested that DcsC is a PLP-independent racemase that is expected to have a similar mechanism to DapF. However, more research is required to understand the detailed mechanism, including site-specific inhibitor binding for each active site cysteine residue. In the scope of this thesis, studies are targeted elucidation of the mechanism of DcsC by site-directed mutagenesis, chiral inhibitor synthesis, and calculation of the kinetic parameters from DcsC variants. In addition, one goal was to determine structure and mechanism of DcsC by crystallization and X-ray analysis.

2.3. Site-directed mutagenesis

2.3.1. Small ubiquitin-related modifier (SUMO)

In previous research by our group, the DNA sequence encoding DcsC with a C-terminal His_6-tag was cloned into *E. coli*.[114] After overexpressing the enzyme, His_6-DcsC was extracted and purified by Ni^{2+} affinity chromatography.[114] However, the mutated His_6-DcsC was unstable and degraded after purification. To increase protein stability, a SUMO-fusion protein was targeted for further research. The reversible attachment of a SUMO protein to other proteins is a modification that has been demonstrated in various cellular processes.[115] There is evidence that the use of an N-terminal SUMO-fusion protein aids in the stabilization of overexpressed proteins.[116,117] In addition, it is reported that a SUMO-fusion helps to crystallize proteins.[118] If necessary, the SUMO tag can be cleaved by a SUMO protease.[118] Purification of the pure protein after SUMO-tag cleavage is possible

using a His_6-tag fused N-terminally to SUMO and Ni^+-NTA column chromatography. [119] For these reasons, a SUMO-tag was fused to the N-terminus of DcsC.

2.3.2. Site-directed mutagenesis of DcsC

When racemic inhibitor **6** is added to wild type DcsC, each of the active site cysteines (Cys81 and Cys227) of DcsC can react. However, single-site mutation of these cysteines to serines could allow confirmation of the suggested bilateral epimerization mechanism. Therefore, site-directed mutagenesis was used to independently change Cys81 to Ser81 and Cys227 to Ser227. To create the mutations, site-directed mutagenesis was performed by polymerase chain reaction (PCR) on the gene encoding SUMO-DcsC. Using primers containing a mutation for the active site, PCR produced plasmids were overexpressed in *E. coli* by Dr. Marco van Belkum. The proteins were purified by Ni-NTA column chromatography and size-exclusion chromatography (Figure 2.13). [120]

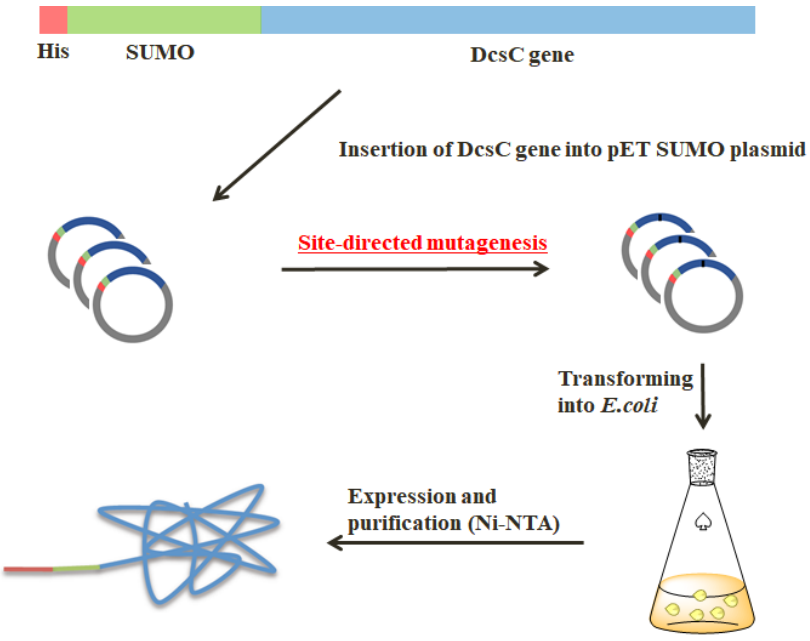


Figure 2.13. Site-directed mutagenesis of DcsC.

2.3.3. Inhibition of SUMO-DcsC and mutants

To confirm the mutation of DcsC, ESI-TOF-MS was used to detect the mass of non-mutated DcsC and mutated DcsC. The loss of 16 m/z units in the respective ESI-TOF-MS (m/z 43.649 kDa for non-mutated DcsC, and m/z 43.633 kDa for both DcsC mutants) confirms the desired Cys to Ser mutation of the DcsC proteins (Figure 2.14). These studies also included both non-mutated DcsC and mutated DcsC with the racemic inhibitor **6**. The difference in the two peaks in each data set is 176 m/z units, which is in accordance with the mass of the racemic inhibitor **6**. Hence, the DcsC mutants were also successfully inhibited by the racemic inhibitor **6**. [120]

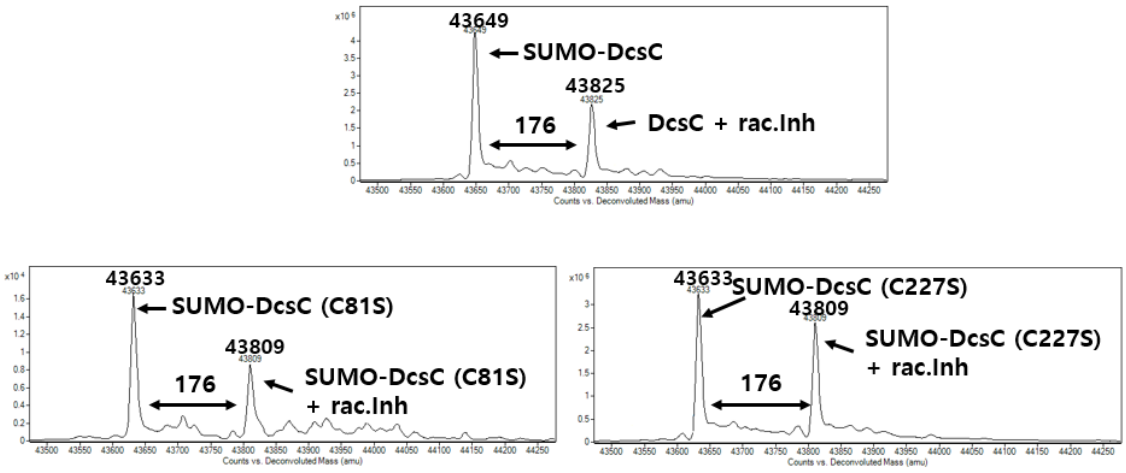


Figure 2.14. ESI-TOF-MS data of non-mutated DcsC and DcsC mutants with racemic inhibitor 6. A. Non-mutated DcsC with racemic inhibitor 6. B. SUMO-DcsC (C81S) with racemic inhibitor 6. C. SUMO-DcsC (C227S) with racemic inhibitor 6.

A trypsin-based MS/MS sequencing approach also confirms desired mutation of SUMO-SUMO-DcsC (C81S) and SUMO-DcsC (C227S). To detect the sequence of DcsC proteins by MS/MS sequencing, DcsC proteins were treated with iodoacetamide for carbamidomethylation of cysteine, and with trypsin to digest the proteins to smaller peptides. These peptides were then analyzed by MS/MS, and sequenced by comparison of mass. The peptide sequence of DcsC mutants with racemic inhibitor **6** also confirmed the mutation of DcsC (Figure 2.15). The SUMO-DcsC (C81S) data demonstrated that Cys81 was successfully mutated to serine (red line in Figure 2.15), and that the serine was not modified by inhibitor **6**. However, Cys227 did react with inhibitor **6** (Figure 2.15A). When Cys227 was mutated to serine in the SUMO-DcsC (C227S), Cys81 reacted with racemic inhibitor **6** (Figure 2.15B). These data sets demonstrate that DcsC can be inhibited with one active site mutation. [120]

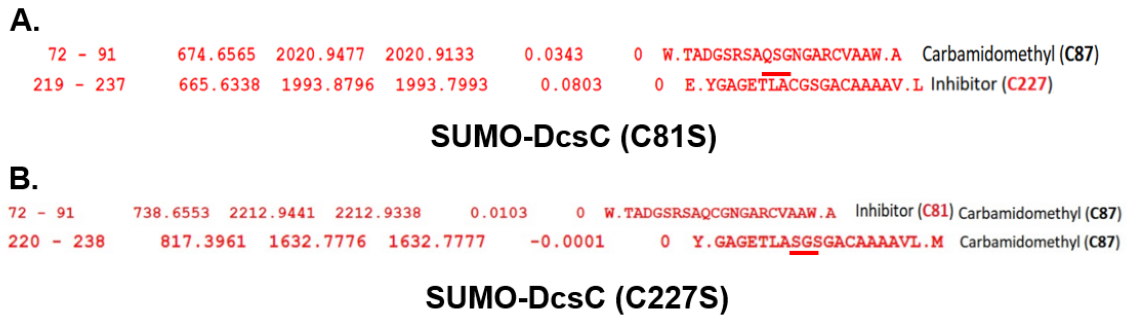


Figure 2.15. MS-MS sequencing data of inhibited SUMO-DcsC (C81S) (A) and SUMO-DcsC (C227S) mutants (B) by racemic inhibitor 6, indicating the site of inhibition.

2.3.4. The activity of non-mutated SUMO-DcsC and mutants

To confirm the activity of the enzymes, the natural substrate isomers, L-*O*-ureidoserine **(1a**) and D-*O*-ureidoserine (**1b**) were synthesized from cycloserine (Figure 2.16). [114]

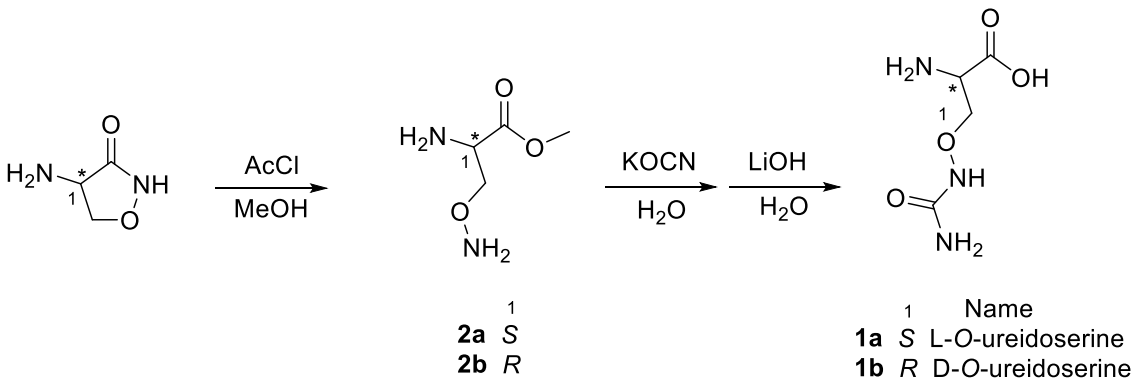


Figure 2.16. Synthesis of L-*O*-ureidoserine (1a) and D-*O*-ureidoserine (1b).

The activity of non-mutated SUMO-DcsC, SUMO-DcsC (C81S) and SUMO-DcsC (C227S) were determined by NMR spectroscopy (20 mM Tris, D_2O, pD 7.6).[120] Non-mutated SUMO-DcsC successfully exchanges the α-hydrogen of **1a** or **1b** to deuterium, as observed by a disappearance of the α-H peak at 4.05 ppm after about 20 min (Figure 2.17). However, neither of DcsC mutants (SUMO-DcsC (C81S) or SUMO-DcsC (C227S)) showed any deuterium exchange with either **1a** or **1b** after 24 hours. According to the proposed mechanism of DcsC, in the native enzyme both active site thiols/thiolates have important roles in deprotonation and re-protonation of the α-position of substrate **1**. Hence in both enzyme mutants, deprotonation of the α-carbon of the substrate by a single active

site thiolate, if it occurs at all, does not result in deuterium exchange with the solvent. The racemization process does not occur, presumably because of the significantly higher *p*Ka of an active site serine as a proton donor on the opposite side.

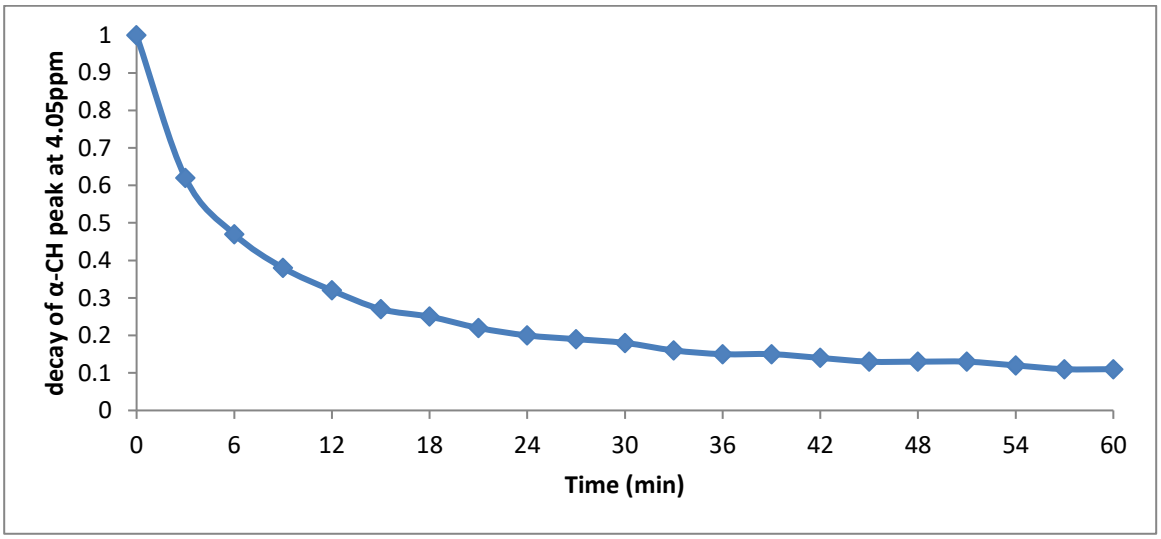


Figure 2.17. ^{1}H NMR assay of non-mutated SUMO-DcsC shown for substrate 1a. The decay of the α-CH peak at 4.05ppm is recorded over time.

2.4. Synthesis of chiral inhibitors

2.4.1. Objectives of the project

Similarity to the suggested mechanism of DapF, each cysteine in the active site of DcsC is expected to only be inhibited by a specific chiral inhibitor. To investigate the mechanism of DcsC in more detail, we aimed to synthesize optically pure inhibitors. Since racemic oxirane inhibitor **6** was unstable under aqueous conditions at room temperature due to epoxide ring opening, we chose to synthesize stable diastereomeric ester precursors to

enable separation of isomers. After saponification of the pure diastereomers, enantiopure inhibitors are in principle obtainable.

2.4.2. Synthetic approaches and procedures

2.4.2.1. Synthesis of neomenthyl derivative

The first attempt to synthesize a stable diastereomeric ester precursor involved a menthyl ester. Since neomenthol is widely used as a chiral auxiliary in asymmetric synthesis, we chose it to synthesize an optically pure oxirane compound (Figure 2.18).[121] To synthesize neomenthyl bromomethylacrylate (**12**) as a precursor of an oxirane, ethyl hydroxymethylacrylate (**7**) was chosen as the starting material. The hydroxyl group of ethyl hydroxymethylacrylate (**7**) was protected by triisopropylsilyl chloride (TIPS) using imidazole as a base. Next, acrylate **8** was saponified using lithium hydroxide. The free carboxylic acid group of TIPS protected acrylic acid **9** was coupled with (+)-neomenthol using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). The TIPS protecting group of neomenthyl ester **10** was deprotected by tetra-*n*-butylammonium fluoride (TBAF) in the following reaction. However, the hydroxyl group of neomenthyl ester **11** was not readily transformed to a bromo group by phosphorous tribromide (PBr_3) to form acrylate **12**. It is maybe that the presence of the neomenthyl ester lowers the reactivity of the hydroxyl group. Therefore, it was decided to switch to another diastereomeric ester precursor.

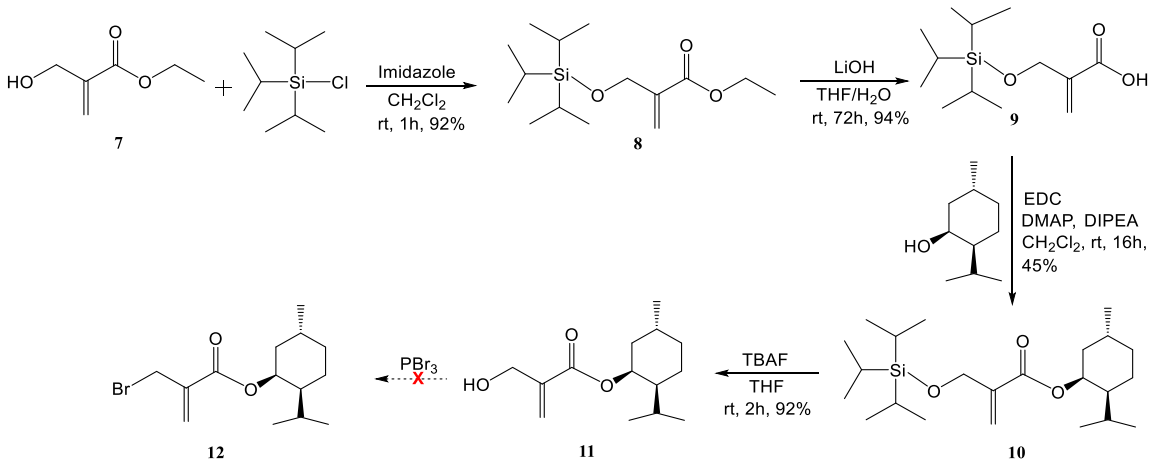


Figure 2.18. Synthesis of neomenthyl bromomethylacrylate.

2.4.2.2. Mitsunobu reaction

In 1967, Professor Oyo Mitsunobu uncovered a new organic reaction that converts hydroxyl groups to other functional groups using triphenylphosphine and an azodicarboxylate, now commonly known as the Mitsunobu reaction.[122] The Mitsunobu reaction is widely used for the esterification of a carboxylic acid group with a primary or secondary alcohol.[123] The mechanism of the Mitsunobu reaction is complex, but well studied in the literature.[120] Triphenylphosphine reacts with the azodicarboxylate *via* a nucleophilic attack to produce a betaine intermediate. This reaction occurs very quickly. The betaine intermediate deprotonates the carboxylic acid and reacts with the substrate alcohol to produce an oxophosphonium ion. Finally, the resulting adduct is attacked by the carboxylate anion to produce the ester, producing triphenylphosphine oxide as a side product. The carboxylate anion attack changes the configuration of the alcohol (inversion, Figure 2.19).[124]

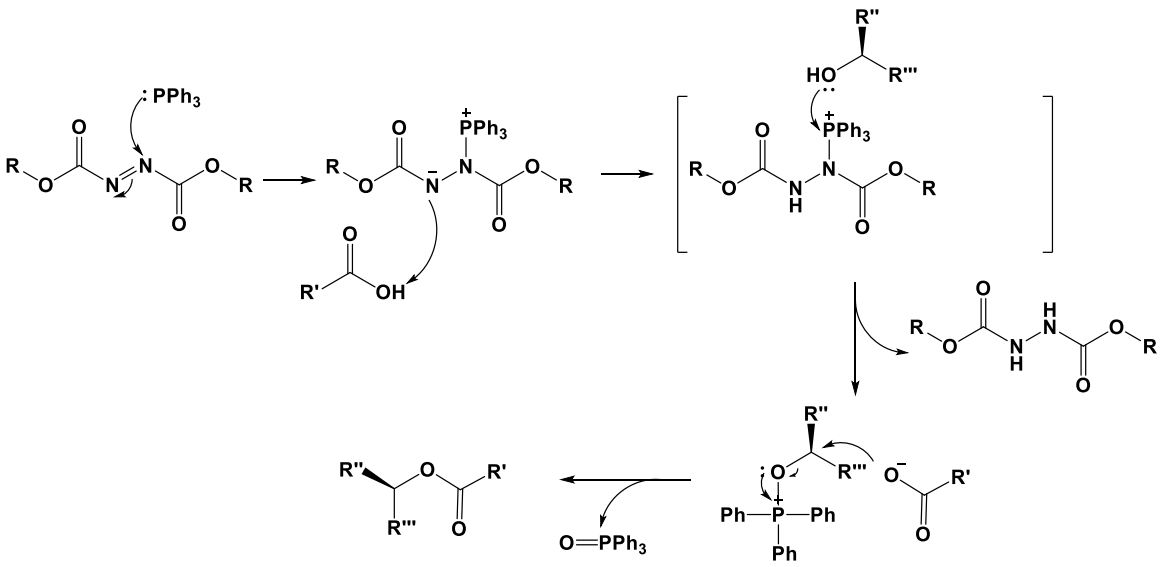


Figure 2.19. General mechanism of the Mitsunobu reaction.

2.4.2.3. Synthesis of 2-methylbutyl inhibitor

To synthesize a stable diastereomeric ester, 2-methylbutanol was used as a chiral auxiliary (Figure 2.20).[125] First, bromomethylacrylic acid (**13**) was esterified with *(S)*-2-methylbutanol *via* the Mitsunobu reaction to give **14**. Next, the bromo group is substituted by a *tert*-butylperoxo group using *tert*-butyl hydroperoxide and sodium hydride as a base. Finally, *tert*-butylperoxo acrylate **15** was treated with hydroxyurea and potassium *tert*-butoxide to synthesize oxirane ester **16** as a diastereomeric mixture (1:1). However, the stereogenic center of the compound has substituents that have similar characteristics. Hence the diastereomers possess very similar properties, and it was therefore difficult to separate the diastereomers of **16** by column chromatography or recrystallization. Therefore, we decided to change the chiral auxiliary to one which allows easier separation of the diastereoisomers.

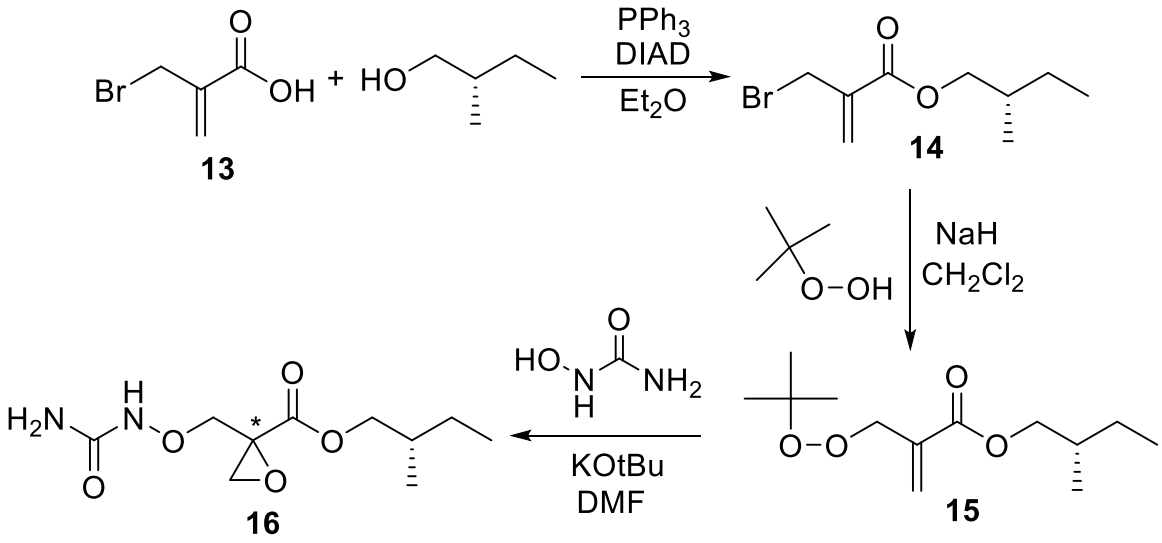


Figure 2.20. Synthesis of 2-methylbutyl oxirane 16.

2.4.2.4. Synthesis of 1-phenylethyl ester

1-Phenylethanol is a well-known chiral auxiliary which contains a chiral center and an aromatic ring. Various 1-phenylethyl esters showed good diastereomeric separation because they contain significantly different substituents: ester, phenyl, and methyl groups are present on the stereogenic center.[126,127] To synthesize the *(S)*-inhibitor **6a** and *(R)*-inhibitor **6b**, *(R)*-1-phenylethanol and *(S)*-1-phenylethanol were used as chiral starting materials for the synthesis (Figure 2.21). Briefly, bromoacrylic acid (**13**) was esterified separately with each alcohol *via* a Mitsunobu reaction. The bromoesters **17a** and **17b** were treated individually with *tert*-butyl hydroperoxide, using NaH as base in diethyl ether. Peroxides **18a** and **18b** were then reacted with hydroxyurea to form the diastereomeric precursor inhibitor mixtures **19a/20a** and **19b/20b**, respectively. The diastereomeric compounds in each mixture were successfully separated by recrystallization. Repeated

recrystallization from a benzene–EtOAc mixture (2:1) gave enantiomerically pure precursors **19a** and **19b**, whose structures were confirmed by X-ray diffraction analysis (Figure 2.22). Saponification of each by LiOH furnishes the pure chiral acids **6a** and **6b**, respectively. [117]

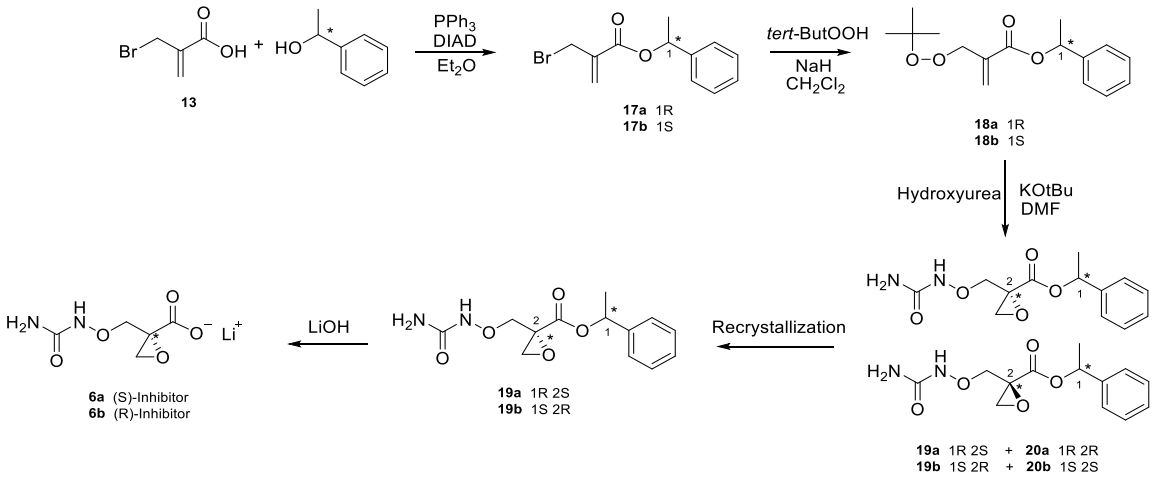


Figure 2.21. Synthesis of chiral oxiranes with 1-phenylethanol.

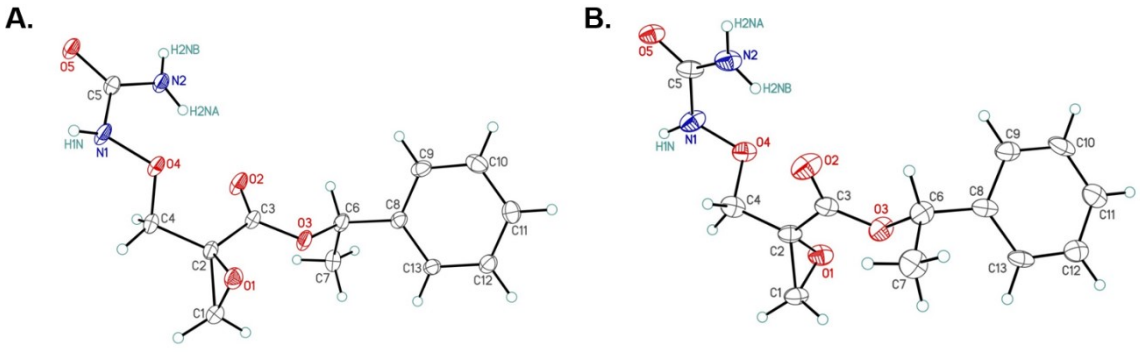


Figure 2.22. X-ray diffraction structure of chiral inhibitors. ORTEP diagram (30 % probability level) (A: 19a, B: 19b).

2.4.3. Inhibition of SUMO-DcsC and mutants with chiral inhibitors

To confirm the inhibition of DcsC and its mutants by chiral compounds **6a** and **6b**, ESI-TOF-MS was used to detect the mass of the enzymes after incubation (Figure 2.23). [117] The gain of 176 m/z units in the respective ESI-TOF-MS confirms the desired inhibition of the DcsC proteins. Interestingly, the SUMO-DcsC (C81S) mutant was not inhibited by the *(S)*-inhibitor, but was inhibited by *(R)*-inhibitor. On the other hand, the SUMO-DcsC (C227S) inhibitor was not inhibited by *(R)*-inhibitor, but was successfully inhibited by *(S)*-inhibitor. This data indicates that the free active site cysteine in both mutants (C227 in SUMO-DcsC (C81S) and C81 in SUMO-DcsC (C227S)) can each only react with one specific enantiomer of the inhibitor, thus supporting the proposed bilateral reaction mechanism.

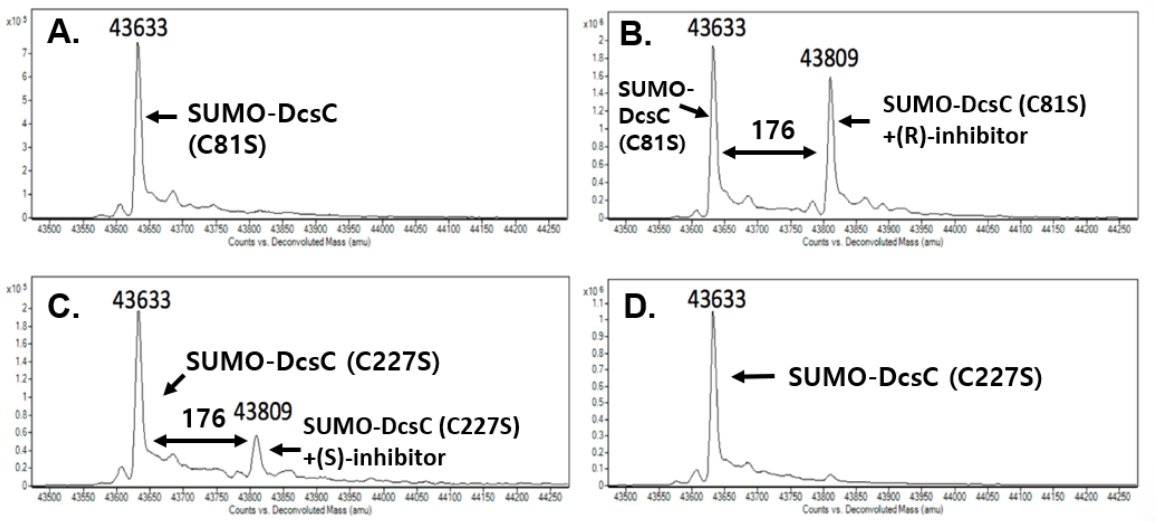


Figure 2.23. Electrospray mass spectra of inhibited DcsC mutants. A. SUMO-DcsC (C81S) with (*S*)-inhibitor. B. SUMO-DcsC (C81S) with (*R*)-inhibitor. C. SUMO-DcsC (C227S) with (*S*)-inhibitor. D. SUMO-DcsC (C227S) with (*R*)-inhibitor

Further support of the bilateral reaction mechanism of DcsC came from the trypsin-based MS–MS sequencing that was performed with the DcsC mutants and associated chiral inhibitors (Figure 2.24). The SUMO-DcsC (C81S) sequence data demonstrated that Cys81 was successfully mutated to serine, and that the mutated residue did not react with either chiral inhibitor. Its Cys227 did not react with *(S)*-inhibitor but was alkylated by *(R)*-inhibitor. In contrast, the SUMO-DcsC (C227S) sequence data showed that Cys227 was successfully mutated to serine, and the mutated residue did not react with either chiral inhibitors. Its Cys81 was not modified by the *(R)*-inhibitor but was alkylated by the *(S)*-inhibitor. These results are in agreement with ESI-TOF-MS data and support a bilateral mechanism of DcsC. [117]

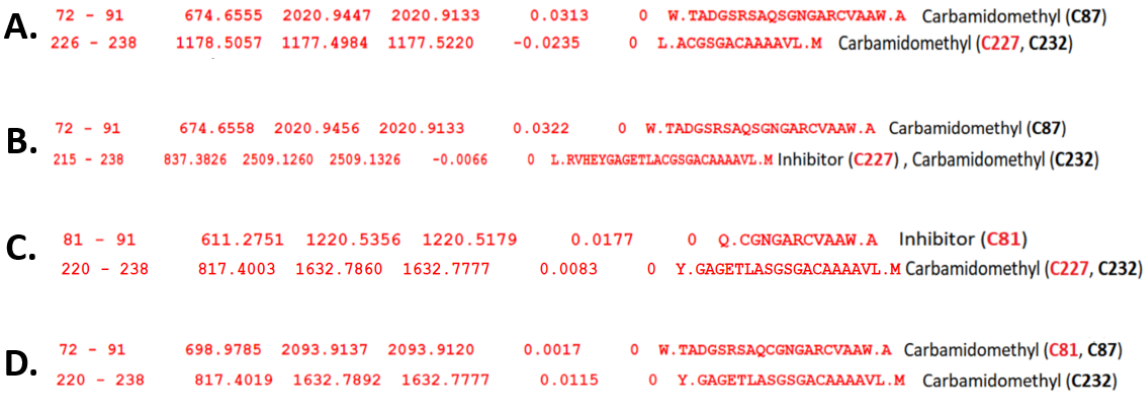


Figure 2.24. MS-MS sequencing data of mutant DcsC with chiral inhibitor 6a and 6b. A. SUMO-DcsC (C81S) with (*S*)-inhibitor. B. SUMO-DcsC (C81S) with (*R*)-inhibitor. C. SUMO-DcsC (C227S) with (*S*)-inhibitor. D. SUMO-DcsC (C227S) with (*R*)-inhibitor.

2.5. Truncation of DcsC

2.5.1. Objectives of the project

To obtain additional structural information about DcsC, the modeled structure of DcsC and the reported structures of DapF from *H. influenzae* were compared (Figure 2.25) with the assistance of Dr. Conrad Fischer and Dr. Marco van Belkum. The structures of the two enzymes exhibited high homology, as expected from sequence alignment (Figure 2.6). Closer examination reveals a flexible N- and C-terminus of DcsC consisting of 9 and 3 amino acids respectively. Truncation of these less-structured segments could result in an enzyme (tDcsC) that maintains catalytic activity with improved stability. Since initial crystallization attempts of the SUMO-DcsC protein and mutants were unsuccessful, we were curious whether truncation of the enzyme could bring positive results.

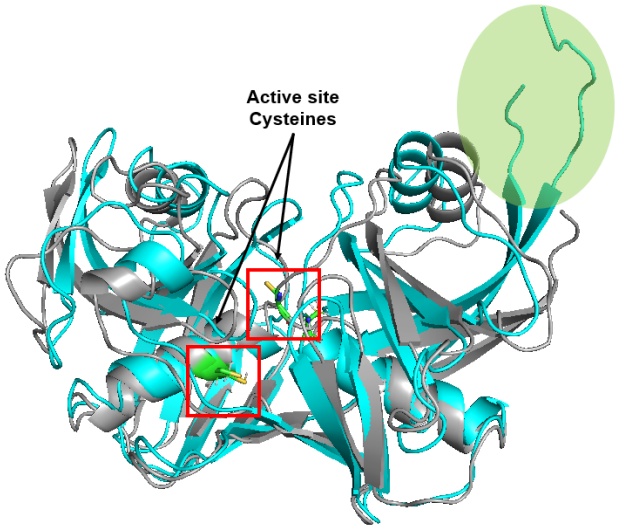


Figure 2.25. Ribbon-model structure of DcsC (lightblue) based on the modeling of DapF from *H. influenzae* (gray, PDB ID: 1bwz, Seq. identity: 42%).[39] The truncated areas (C- and N-terminal) are highlighted in green.

2.5.2. Preparation of truncated enzyme and activities

The truncated DcsC protein with a SUMO-tag at the N-terminus was expressed and purified by Ni-NTA column chromatography and size-exclusion chromatography. Site-directed mutagenesis of the SUMO-truncated DcsC (SUMO-tDcsC) was also performed, using the same procedure as described before, to confirm the bilateral mechanism of DcsC. Hence for activity testing, SUMO-tDcsC was incubated with the natural substrates **1a** and **1b** (20 mM Tris, D_2O, pD 7.6). The activity profiles of SUMO-tDcsC were somewhat lower compared to that of the native enzyme (see the section 2.6.) [117]

2.5.3. Inhibition of SUMO-tDcsC and mutants with chiral inhibitors

To confirm truncation of the enzyme and its inhibition by the chiral inhibitors, we analyzed the mass of SUMO-tDcsC and its mutants with bound inhibitors using ESI-TOF-MS. The loss of 1,331 m/z units in the respective ESI-TOF-MS data (m/z 43.633 kDa for SUMO-DcsC (C81S), and m/z 42.302 kDa for SUMO-tDcsC (C72S)) confirms the desired 12 amino acid truncation of the DcsC proteins (Figure 2.25). As with non-truncated DcsC mutants, SUMO-tDcsC (C72S) was only inhibited by the *(R)*-inhibitor and SUMO-tDcsC (C218S) was only inhibited by the *(S)*-inhibitor. This data again confirmed the bilateral mechanism of truncated DcsC. Interestingly, the inhibition efficiency of the tDcsC enzyme was much higher than that of the non-truncated enzyme. In addition, a second site of alkylation becomes available in the truncated versions (Cys41), as indicated by a small additional mass peak (Figure 2.26). This suggests a somewhat looser structure with greater access for inhibitors to the enzyme interior. [117]

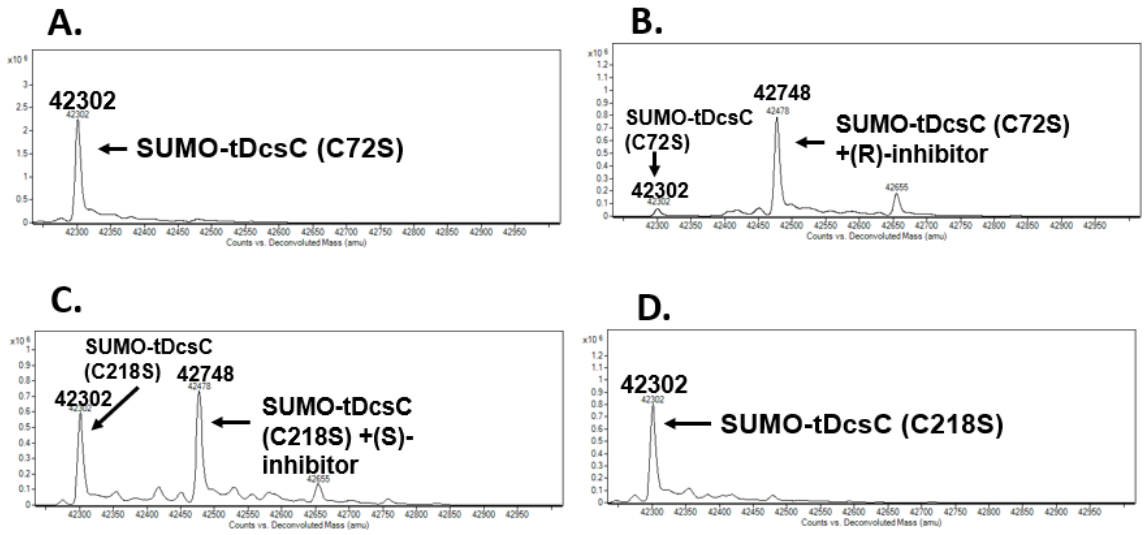


Figure 2.26. Electrospray mass spectra of inhibited SUMO-tDcsC mutants. A. SUMO-tDcsC (C72S) with (*S*)-inhibitor. B. SUMO-tDcsC (C72S) with (*R*)-inhibitor. C. SUMO-tDcsC (C218S) with (*S*)-inhibitor. D. SUMO-DcsC (C218S) with (*R*)-inhibitor.

To gather more evidence about the bilateral reaction mechanism of SUMO-tDcsC, trypsin-based MS/MS sequencing was performed on the tDcsC mutants and associated chiral inhibitors (Figure 2.27). Because of the 9 amino acid truncation at the N-terminus, Cys72 (C81 for non-truncated DcsC) and Cys218 (C227 for non-truncated DcsC) represent the active site of tDcsC. The sequence data demonstrated that Cys72 was mutated to a serine for the SUMO-tDcsC (C72S) mutant and that Cys218 was mutated to a serine for the SUMO-tDcsC (C218S) mutant successfully. The sequence data also indicated that the free active site cysteine in both mutants (C72 in SUMO-tDcsC (C72S) and C218 in SUMO-tDcsC (C218S)) can each only react with one enantiomer of the inhibitor, again suggesting the bilateral mechanism of truncated DcsC. [117]

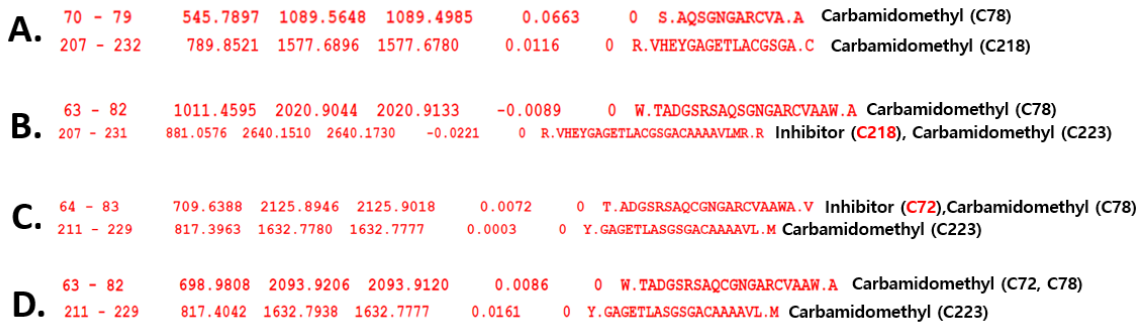


Figure 2.27. MS-MS sequencing data of mutant SUMO-tDcsC. A. SUMO-tDcsC (C72S) with (*S*)-inhibitor. B. SUMO-tDcsC (C72S) with (*R*)-inhibitor. C. SUMO-tDcsC (C218S) with (*S*)-inhibitor. D. SUMO-DcsC (C218S) with (*R*)-inhibitor.

2.6. Kinetics

2.6.1. Kinetic parameters of DcsC variants with ^{1}H-NMR

The catalytic parameters for various DcsC variants were determined by time-dependent ^{1}H-NMR spectroscopy and circular dichroism (CD) spectroscopy, and are summarized in Table 2.1. The catalytic parameters for the C-terminal His_6-tagged version of DcsC published previously[114] are also included in this table for reference.

Table 2.1. Kinetic parameters of DcsC variants [a]

Species	Direction	v_{max} (mM s^{-1})	K_m (mM)	K_{cat} (s^{-1})	K_{cat}/K_m (1 M^{-1} s^{-1})	Ref.
DcsC-His_6	**1a → 1b**	0.084	110	158	1,436	[114]
DcsC-His_6	**1b → 1a**	0.015	17	29	1,706	[114]
SUMO-DcsC	**1a → 1b**	0.022±0.0011	12±1.12	475	41,130	[120]
SUMO-DcsC	**1b → 1a**	0.047±0.0023	32±0.95	1450	45,302	[120]
DcsC (tag-free)	**1a → 1b**	<0.001	n.d.	n.d.	n.d.	[120]
DcsC (tag-free)	**1b → 1a**	<0.001	n.d.	n.d.	n.d.	[120]
SUMO-tDcsC	**1a → 1b**	0.042±0.0075	15±1.34	315	21,137	[120]
SUMO-tDcsC	**1b → 1a**	0.046±0.0031	22±1.23	194	8,642	[120]

[a] The kinetic parameters were measured by ^{1}H NMR and CD.

The data was extracted from the signal ratio between the vanishing α-hydrogen (at δ ~ 4.05 ppm) and the β-hydrogen (at δ ~ 4.20 ppm) (Figure 2.28). The initial velocities (3 min) were used to calculate the Michaelis–Menten parameters. In general, these calculations revealed two trends: the catalytic efficiency (K_{cat}/K_m) in both directions is almost 30-times higher for the SUMO-tagged version of the DcsC in comparison to the His_6-tagged version of the enzyme. One possible explanation for this enhanced catalytic performance could be the predominantly dimeric state of SUMO-DcsC compared to a higher oligomeric state of the His_6-tagged version, which has been described as a factor of catalytic acceleration in other racemases.[128] Truncation of the enzyme (SUMO-tDcsC, Table 2.1) renders these variants 50–80% less active than the native version but still 5-14 times more active than the C-terminally His-tagged variant. Removal of the SUMO-tag results in an unstable protein that quickly precipitates even at 0 °C, preventing

determination of kinetic parameters. Unlike the His-tagged DcsC version, the substrate affinity towards the *O*-ureidoserine enantiomers (direction of transformation, **1a** → **1b** or **1b** → **1a**) is comparable to that of the SUMO-tagged variants.

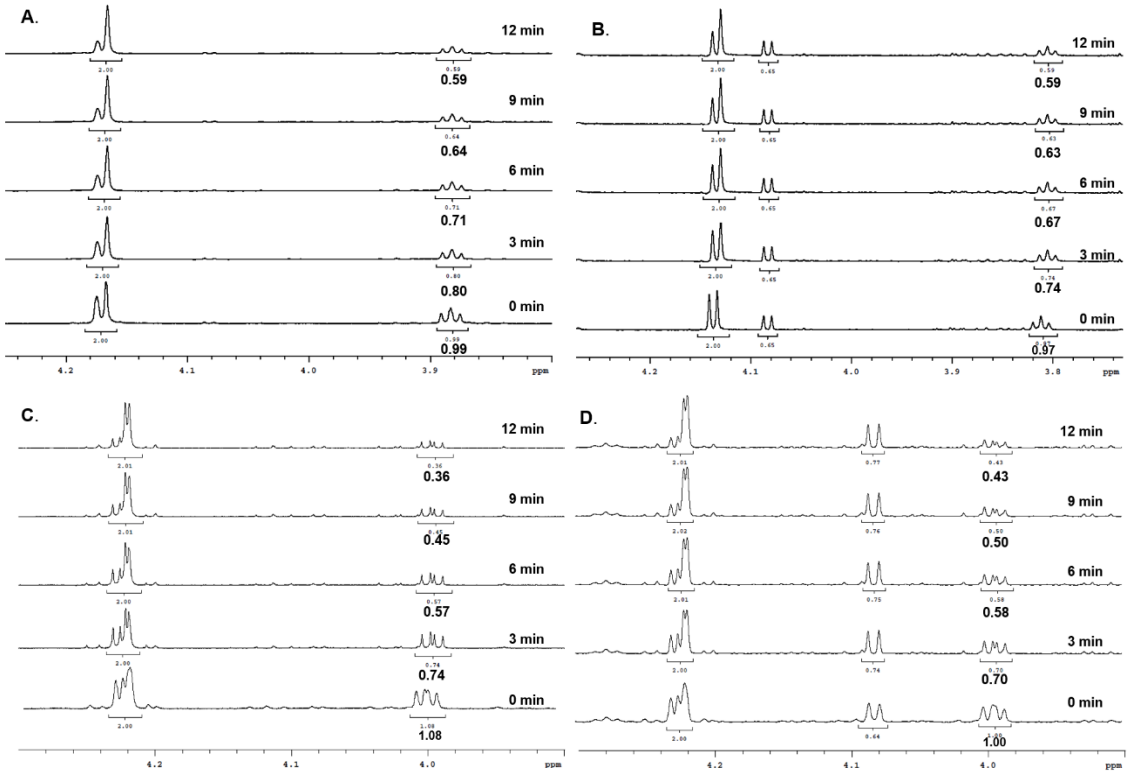


Figure 2.28. Enzyme activities by ^{1}H-NMR. A. SUMO-DcsC with 1a. B. SUMO-DcsC with 1b. C. truncated DcsC with 1a. D. truncated DcsC with 1b.

2.6.2. Kinetic parameters of DcsC variants by Circular Dichroism

To confirm the kinetic parameters of various DcsC, circular dichroism (CD) spectroscopy was employed (Figure 2.29). The ellipticity of **1a** and **1b** showed linear dependence on concentration at 206 nm. The assays were performed at 30 °C in phosphate

buffer (pH= 7.8) with 1mM dithiothreitol (DTT) to prevent disulfide formation at the active site. When DcsC (2-5 μg) was added to the buffer solution with various substrate concentrations (2-15 mM), the loss of CD signal over time was monitored for both substrate **1a** and **1b**. The initial velocities (2 min) of the racemization, calculated from the slope constant of the trend line, were used to determine the Michaelis–Menten parameters of the enzyme. The kinetic parameters of the enzyme as measured by CD also showed similar values as the ^{1}H-NMR data, with the average of values shown in Table 2.1.

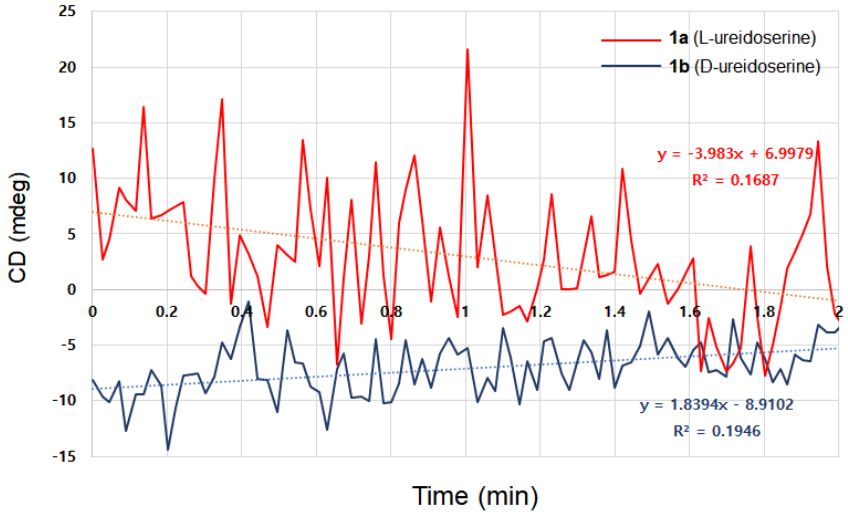


Figure 2.29. Typical changes in CD at 206 nm that were used to determine the kinetics of DcsC on substrate 1a and 1b (12mM).

2.7. Conclusion

2.7.1. Putative mechanism of DcsC

Individual mutation of both active site cysteines and incubation with both enantiomers (**6a** and **6b**) of the inhibitor reveals selective inactivation dependent upon the chirality of the inhibitor, thus supporting the bilateral reaction mechanism of the enzyme. Summarizing these mutation studies in the active site, Cys81 is covalently bound with only the *(S)*-inhibitor due to chiral specificity. In contrast, Cys227 can only recognize the *(R)*-inhibitor as a substrate (Figure 2.30).

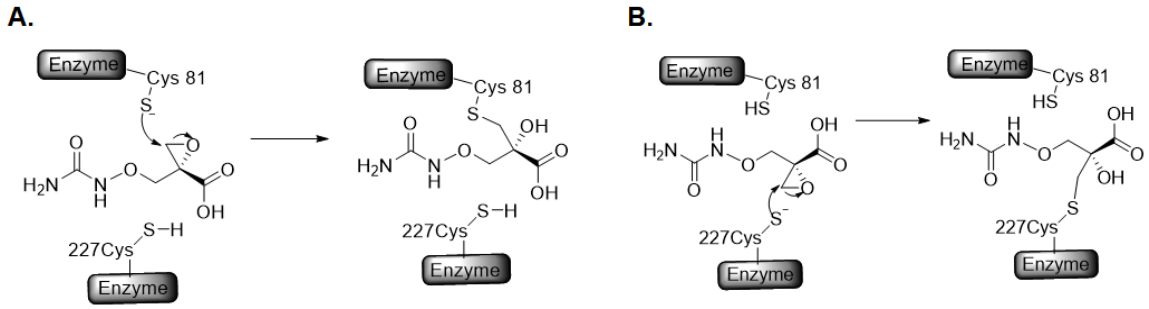


Figure 2.30. Putative inhibition mechanism of DcsC with 6a and 6b. A. Inhibition of Cys 81 by *(S)*-Inhibitor (6a). B. Inhibition of Cys 227 by *(R)*-Inhibitor (6b).

The bilateral inhibition mechanism of the enzyme suggests that the mechanism of substrate recognition by DcsC is also a bilateral mechanism, which is similar to the racemization in DapF. In the active site, it is expected that Cys81 recognizes only L-O-ureidoserine (**1a**, *S* configuration), and Cys227 recognizes only D-O-ureidoserine (**1b**, *R*

configuration, Figure 2.31). This putative bilateral mechanism demonstrates how the enzyme recognizes specific substrates for each active site.

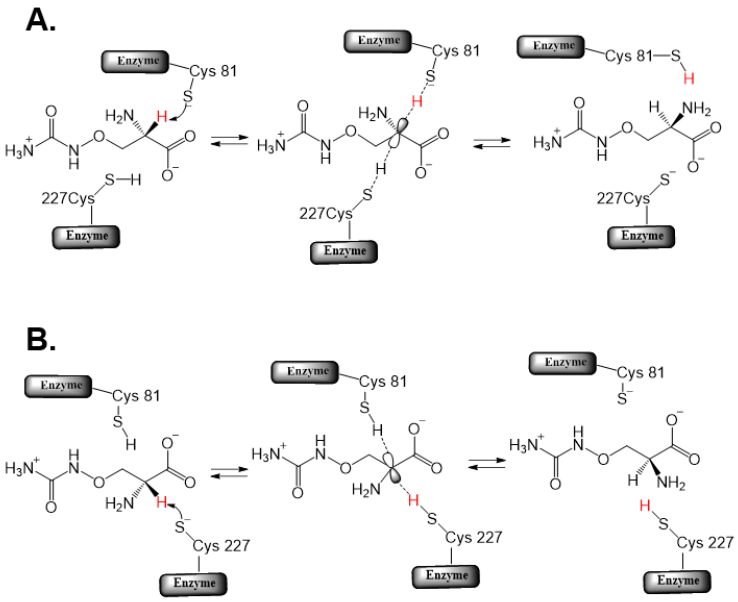


Figure 2.31. Putative mechanism of DcsC. A. Substrate recognition of Cys 81 with 1a. B. Substrate recognition of Cys 227 with 1b.

2.7.2. Future work

In this project, we reported the recombinant expression of O-ureidoserine racemase (DcsC) featuring an N-terminal SUMO-tag. Purified SUMO-DcsC racemizes the natural substrates L- and D-O-ureidoserine **1a** and **1b** with a 30-fold increase in catalytic activity when compared to a C-terminally His_6-tagged version of the enzyme. Truncation of the C- and N-terminus of the protein as well as removal of the SUMO-tag results in a destabilization of the protein structure and subsequent reduction of enzymatic activity. The

mutated enzyme, when incubated with chiral inhibitors **6a** and **6b,** demonstrated a bilateral reaction mechanism. Crystallization studies are underway, ultimately aiming to obtain X-ray structures of the enzyme, inhibited and not, to aid in understanding the interior dimensions and distal site substrate recognition by the active site.

Chapter 3

Amino acid racemase in *Bacillus subtilis* (RacX)

3.1. Introduction of RacX

3.1.1. Previous research

RacX, found exclusively in *Bacillus* species, is a PLP-independent racemase first mentioned in 1993 as a downstream open reading frame attached to a penicillin-binding protein (PBP4a) in *Bacillus subtilis* (*Bs*RacX)[129]. This enzyme has high homology with aspartate racemase, and thus it was expected that the natural substrates of the enzyme would be either aspartate or glutamate. However, both amino acids show no activity with an overexpressed purified enzyme preparation. In 2017, Miyamoto *et al.* reported the various amino acid substrates for RacX (Table 3.1).[130] The kinetic parameters were determined using HPLC with chiral derivatives (Boc-*L*-cysteine and *N*-acetyl-*L*-cysteine) and fluorescent reagent (*o*-phthalaldehyde). In contrast with other PLP-independent racemases, RacX did not have strict specificity in the recognition of particular substrate, with more than 16 amino acids being racemized, including: L,L-diaminopimelate, lysine (D, L), arginine (D, L), ornithine (D, L), L-histidine, L-alanine, L-tyrosine, L-phenylalanine, L-serine, L-glutamine, L-methionine, L-asparagine, L-homoserine, L-norleucine, L-

aminobutyric acid, and L- norvaline. Interestingly, some of these amino acids were known to be racemized by PLP-dependent racemases.

Substrate	Substrate specificity (nmol/min/mg)	Relative activity (%)	k_{cat} (min^{-1})	k_{cat} (sec^{-1})
L,L-**Dap**	1.301	100	0.03461	0.0005768
D-**Lys**	1.217	93.54	0.03238	0.0005396
L-**Arg**	1.178	90.55	0.03134	0.0005223
L-**Orn**	1.122	86.24	0.02985	0.0004975
L-**Lys**	1.056	81.19	0.02809	0.0004682
D-**Arg**	1.016	78.09	0.02702	0.0004505
D-**Orn**	0.607	46.65	0.01615	0.0002691
L-**His**	0.417	32.05	0.01109	0.0001849
L-**Ala**	0.337	25.90	0.008965	0.0001494
L-**Tyr**	0.296	22.75	0.007874	0.0001312
L-**Phe**	0.285	21.90	0.007582	0.0001264
L-**Ser**	0.249	19.14	0.006624	0.0001104
L-**Gln**	0.187	14.37	0.004975	0.00008291
L-**Met**	0.17	13.07	0.004522	0.00007537
L-**Asn**	0.134	10.30	0.003565	0.00005941
L-**Hse**	0.121	9.311	0.003219	0.00005364
L-**Nle**	0.073	5.611	0.001942	0.00003236
L-**Aba**	0.052	3.997	0.001383	0.00002306
L-**Nva**	0.047	3.613	0.001250	0.00002084

Table 3.1. Substrate specificity and kinetic parameters of RacX racemase activity.

However, the reaction rates of RacX with these single amino acids were very low (Table 3.1). For example, the k_{cat} of L,L-diaminopimelate was only 0.03461 min^{-1} even though it showed the highest reaction rate with RacX. Compared with the turnover number of SUMO-DcsC (28,500 min^{-1} with L-ureidoserine), RacX with L,L-diaminopimelate is almost a million times slower. Therefore, we postulated that these single amino acids may not be the natural substrate for RacX.

3.1.2. Gene clusters of RacX and PbpE

According to the reported gene sequences of *B. subtilis*, RacX gene is considered to be clustered with penicillin binding protein 4 (PbpE, Figure 3.1).[131] PbpE is known to catalyze a process of peptidoglycan hydrolysis: it has D,D-endopeptidase activity capable of cleaving peptidoglycan-peptide cross-links for recycling purposes.[132] If the role of RacX is related to PbpE, it is expected that RacX might also be involved in the recycling process of peptidoglycans. Therefore, we chose to test the hypothesis that the true substrates for RacX are peptides that are part of the peptidoglycan structure.

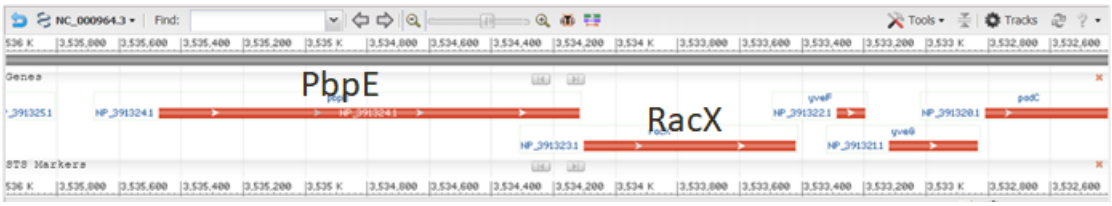


Figure 3.1. Database results showing gene sequences of *B. subtilis* encoding RacX and PbpE.

3.2. Results and discussion

3.2.1. Purification of RacX

To obtain the RacX protein, the DNA sequence encoding RacX with N-terminal SUMO tag was purchased and cloned into the pET SUMO plasmid by Dr. Marco van Belkum.[133] Similar to the case for DcsC, the SUMO-protein fusion was chosen to increase protein overexpression and stability. The plasmid was transformed into *E. coli* to

overexpress the protein. The protein was purified by Ni-NTA column and detected by SDS-PAGE (Figure 3.2). The side products in the SUMO-RacX solution with 300 mM imidazole was removed by dialysis and size-exclusion chromatography.

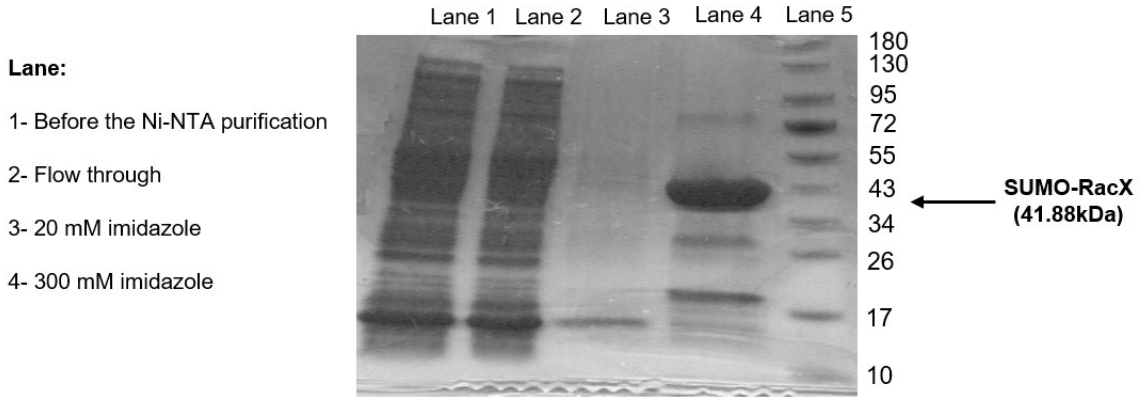


Figure 3.2. Proteins separated by SDS-PAGE after Ni-NTA purification. Flow through includes all proteins except of SUMO-RacX. The eluent with 300 mM imidazole solution includes RacX protein (41.88 kDa).

3.2.2. Synthesis of various peptides

The short peptides from the peptidoglycan in *Bacilli* bacteria cell walls consist of L-alanine, D-glutamate, *meso*-diaminopimelate (*meso*-DAP), and D-alanine. If RacX joins in the recycling process of peptidoglycan with PbpE, one of these short peptides containing of these amino acids could be the substrate for RacX (Figure 3.3). To test this idea, we set out to synthesize peptides with analogues structures in peptidoglycan.

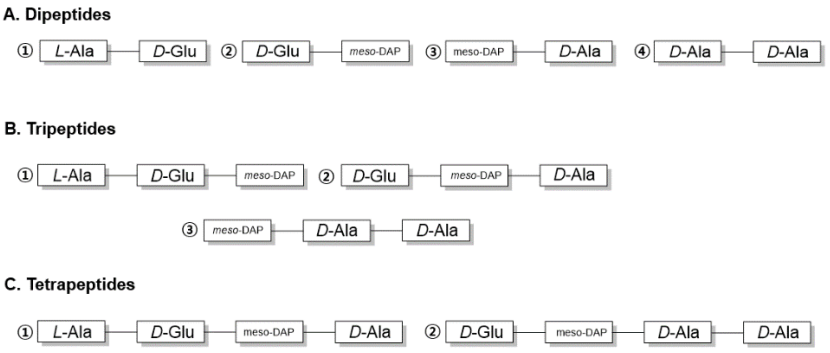


Figure 3.3. Various candidates as natural substrates for RacX.

3.2.2.1. Synthesis of orthogonal protected *meso*-diaminopimelate

To synthesize various peptides, we needed amino acids that are components of peptidoglycan. Protected L-alanine, D-glutamate, and D-alanine were purchased from commercial sources, but orthogonally protected *meso*-Dap was prepared separately using a precedented scheme from Dr. Eva Rodriguez of our group (Figure 3.4).[134] First, D-allylglycine **24** was protected using ethyl trifluoroacetate (amine protection) and methyl iodide (carboxylic acid protection) to make TFA-methyl protected D-allylglycine **25**. Boc-vinylglycine-OBn **27** was prepared from *N*-Boc-L-glutamic acid-α-benzyl ester **26**, lead (IV) acetate, and copper (II) acetate.[135] To synthesize alkene **28**, Compounds **25** and **26** were connected by an olefin metathesis reaction using Hoveyda-Grubbs second generation catalyst. In this reaction, there was a possibility to produce dimers of **25** and **26**, and the yield of this reaction was low (< 40%). The protected alkene **28** was treated with palladium on carbon and hydrogen gas to reduce the alkene, yielding orthogonally protected *meso*-Dap **29**.[134]

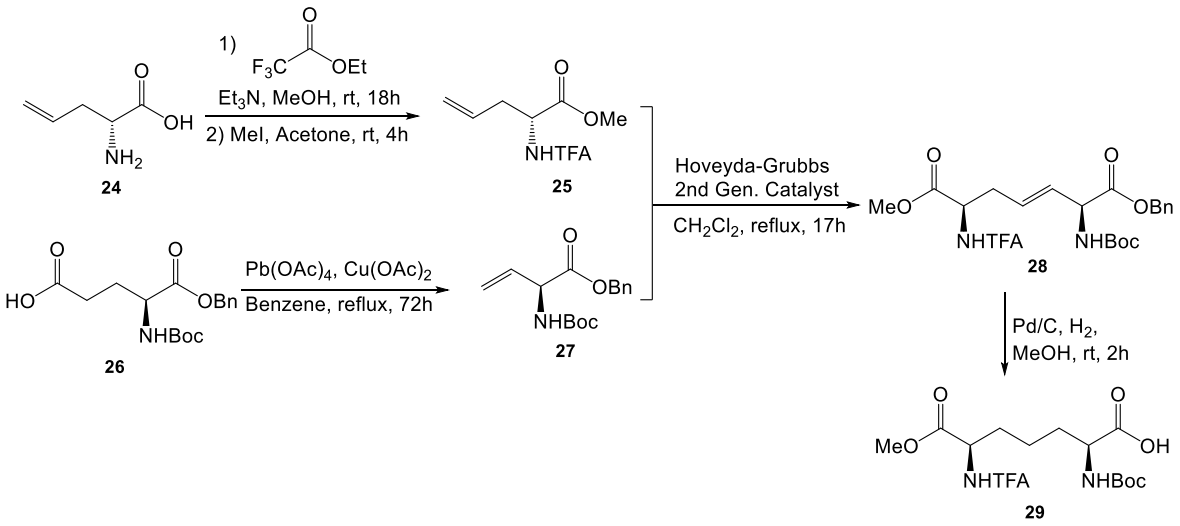


Figure 3.4. Synthesis of orthogonally protected *meso*-Dap.

3.2.2.2. Synthesis of protected D-Glu-*meso*-Dap-D-Ala-D-Ala

To synthesize the various peptides (dipeptide A4, tripeptide B3, and tetrapeptide C2 from Figure 3.3), most of the assembly was performed from the C-terminus to N-terminus of the peptides. First, D-alanine methyl ester (D-Ala-OMe) and Boc-protected D-alanine (Boc-D-Ala-OH) were coupled to produce dipeptide **30**. Next, the amine group of dipeptide **30** was deprotected using TFA, which was then coupled with *meso*-Dap derivative **29**. Finally, tripeptide **31** was coupled with Boc-protected D-glutamate (Boc-D-Glu-OMe) to produce tetrapeptide **32** (Figure 3.5). All coupling reactions were performed using 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluoro-phosphate (HATU) as the coupling reagent and diisopropylethylamine (DIPEA) as the base. Dipeptide **30** (Boc-D-Ala-D-Ala-OMe), tripeptide **31** *(*Boc-*meso*-Dap-D-Ala-D-Ala-OMe), and tetra peptide **32** (Boc-D-Glu-*meso*-Dap-D-Ala-D-Ala-OMe) were deprotected using

trifluoroacetic acid (TFA) and sodium hydroxide (NaOH). However, cyclization and dimerization had occurred after deprotection of the peptides (detected by LC/MS ESI), thus alternative deprotection processes are needed to utilize these peptides as substrates for RacX.

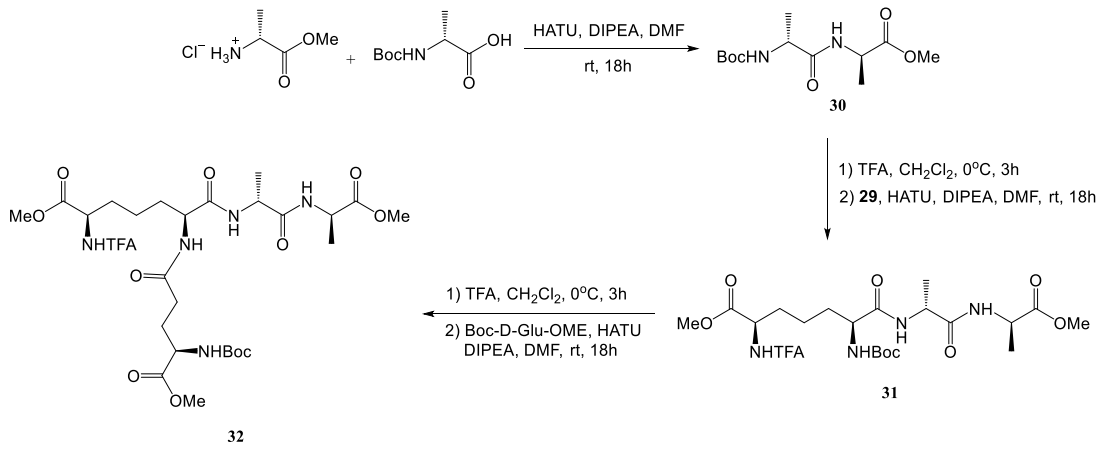


Figure 3.5. Synthesis of D-Glu-*meso*-Dap-D-Ala-D-Ala tetrapeptide.

3.2.2.3. Synthesis of protected L-Ala-D-Glu-*meso*-Dap-D-Ala

To produce different versions of peptides (dipeptide A3, tripeptide B2, and tetrapeptide C1 from Figure 3.3), D-Ala-OMe and orthogonally protected *meso*-Dap **29** coupled to provide dipeptide **33** using HATU and DIPEA (Figure 3.6). Dipeptide **33** was treated with TFA to remove the Boc-protecting group. The crude material was then coupled directly with Boc-D-Glu-OH to give protected tripeptide **34**. Following this, tripeptide **34** was deprotected using TFA and the resulting crude amine coupled directly with Boc-protected L-alanine (Boc-L-Ala-OH), using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexa-fluorophosphate (PyBOP, coupling reagent),

hydroxybenzotriazole (HOBt, racemization suppressing additive) and DIPEA, to give tetrapeptide **35**. Although the desired peptides **33**, **34** and **35** were prepared, novel deprotection strategies are still required to prevent cyclization or dimerization.

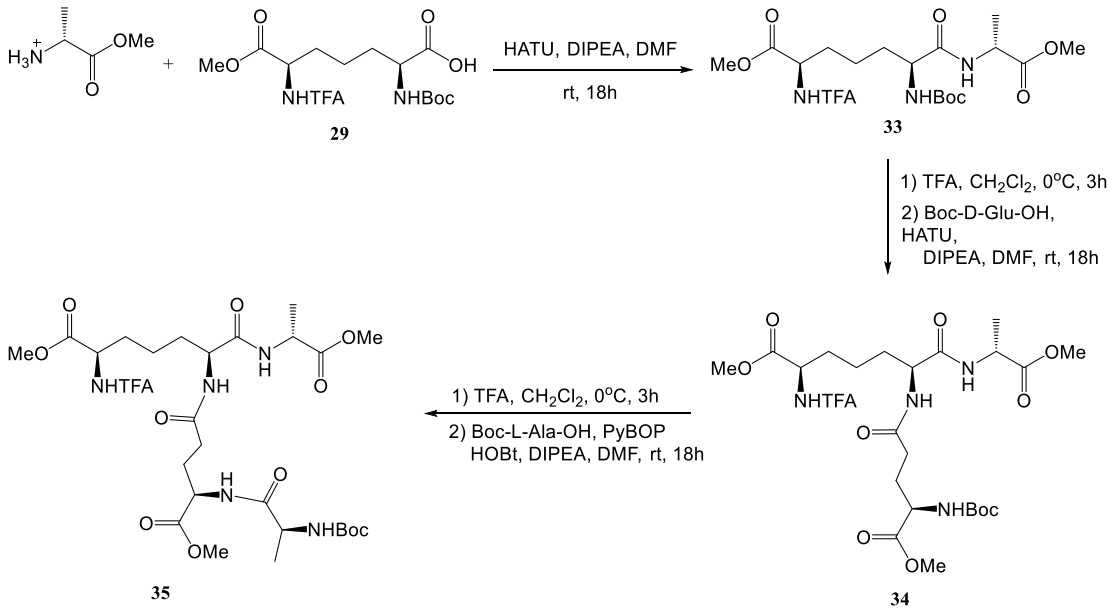


Figure 3.6. Synthesis of L-Ala-D-Glu-*meso*-Dap-D-Ala tetrapeptide.

3.2.2.4. Synthesis of L-Ala-D-Glu

To synthesize an additional substrate candidate for RacX (dipeptide A1 from Figure 3.3), an attempt to synthesize the peptides from the N-terminus to C-terminus was performed (Figure 3.7). First, carboxybenzyl (Cbz, amine protection) protected L-alanine (Cbz-L-Ala) and D-glutamic acid dibenzyl ester (D-Glu(OBn)-OBn) were coupled using HATU and DIPEA. The protected dipeptide **36** (Cbz-L-Ala-D-Glu(OBn)-OBn) was then successfully deprotected using palladium on carbon and hydrogen gas to produce dipeptide

37. However, further peptide synthesis was not performed as different analogues of orthogonally protected *meso*-Dap (one free amine containing differentially protected carboxylic acids) is currently pending due to time restrictions. Additionally, to avoid side reactions, D-glutamic acid should have two different protecting groups on the carboxylic acid. Alternative synthetic strategies are required for the further synthesis.

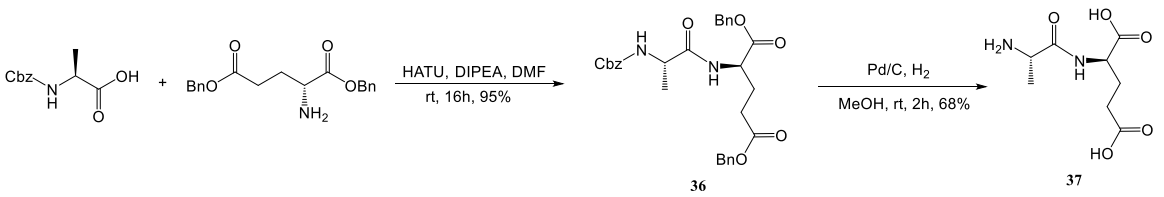


Figure 3.7. Synthesis of L-Ala-D-Glu dipeptide.

3.3. Conclusion and future work

3.3.1. Conclusion

In this project, we perused the recombinant expression of RacX featuring an N-terminal SUMO-tag and the synthesis of various substrates for its enzyme. Although the SUMO-RacX enzyme was successfully prepared with high stability, the activity of the enzyme was not detected because the reaction rate with known substrates (single amino acids) was too slow to monitor by NMR or CD. The prepared protected peptides including dipeptide **30** (Boc-D-Ala-D-Ala-OMe), tripeptide **31** *(*Boc-*meso*-Dap-D-Ala-D-Ala-OMe), tetra peptide **32** (Boc-D-Glu-*meso*-Dap-D-Ala-D-Ala-OMe), dipeptide **33** (Boc-*meso*-Dap-D-Ala-OMe), tripeptide **34** (Boc-D-Glu-*meso*-Dap-D-Ala-OMe), and tetrapeptide **35** (Boc-

L-Ala-D-Glu-*meso*-Dap-D-Ala-OMe) were not tested with RacX due to the presence of protecting groups on the peptides. The unprotected dipeptide **37** was tested as a substrate for RacX. However, even though the reaction mixture was incubated for 3 days at 37 ℃, the racemization was not detected by NMR spectroscopy or CD. Therefore, it appears that dipeptide **37** is not a natural substrate of RacX.

3.3.2. Future work

To find the natural substrate of RacX, the prepared protected peptides **30, 31, 32, 33, 34,** and **35** should be deprotected and tested with RacX. Since TFA and NaOH deprotection resulted in dimerization and cyclization, different deprotection strategies are required. In the event that none of the peptides found in as sub-structures peptidoglycan show activity with RacX, the whole peptidoglycan pentapeptide or peptides with sugar moieties may be substrates for RacX. If the natural substrate for RacX is discovered, further research on the synthesis of relevant inhibitors and crystallization of the enzyme will aid in rationalizing the substrate recognition mechanism of RacX.

Chapter 4

Summary and Conclusions

This thesis contains mechanistic studies of PLP-independent racemases including *O*-ureidoserine racemase (DcsC), and a putative amino acid racemase in *Bacillus subtilis* (RacX). Previous work by others detailed the structural characterization and mechanisms of diaminopimelate epimerase (DapF), glutamate racemase (GluR), aspartate racemase (AspR), and proline racemase (ProR). To further develop an understanding of PLP-independent racemases, an investigation of the two PLP-independent racemases was undertaken. Contained within are the mechanistic and mutational studies of DcsC, an enzyme involved in the biosynthesis of D-cycloserine. Lastly, this thesis describes the research conducted on identification and characterization of potential substrates of RacX, which is an amino acid racemase found in *B. subtilis*.

In contrast with PLP-dependent racemases which require PLP or metal cofactors for enzymatic activity, most PLP-independent racemases do not require any cofactors. An exception is UDP-MurNAc-Ala acting as an activator of GluR in *E. coli*.[61] Most active sites in PLP-independent racemases include a thiolate-thiol pair, with the exception of aspartate/glutamate racemase (*Ec*L-DER), which uses a Thr-Cys pair in the active site to protonate and deprotonate the α-carbon of the coressponding substrates. Because the cysteine active sites are paired, these enzymes are able to recognize both L- and D-

substrates for racemization. On the other hand, *Ec*L-DER was observed to exhibit one-directionality (L→D direction) because of the threonine in the active site (Figure 1.23).

PLP-independent racemases have hydrogen bonding networks to recognize substrates in the active site. According to the sequence alignments and X-ray crystal structures (modeling structure for DcsC) of each enzyme, various amino acids near the cysteine active site are highly conserved between different species, and serve similar roles to recognize specific substrates by hydrogen bonding. Additionally, DapF, AspR, ProR, and DcsC have a narrow pocket size for the active site, thus only natural substrates or analogues are able to fit and be recognized. On the other hand, GluR and *Ec*L-DER have relatively spacious active sites, even though they also possess a hydrogen bonding network. Therefore, these two enzymes can recognize ligands of different sizes as substrate, for example, 2-naphethylmethyl glutamate, GluR inhibitor, which is much larger than glutamate.

The mechanistic details of PLP-independent racemases which are explained in this thesis show numerous similarities. Firstly, these enzymes follow a two-base bilateral mechanism, meaning that each cysteine of the active site only recognizes a specific stereo-configuration of the substrate. For example, Cys99 of DapF in *Arabidopsis thaliana* can only bind L,L-DAP, whereas Cys254 binds to D,L-DAP. Secondly, the enzyme stabilizes the negative charge on the α-carboxylate of the substrate by using opposing α-helical dipole moments and the hydrogen bonding environment. This charge distribution increases the acidity of the α-hydrogen, thus allowing cysteine to deprotonate it. Thirdly, stereoelectronic alignment occurs when the cysteine of these enzymes protonate the α-carbon of the

substrate. Since the carboxylate maximizes orbital overlap between its π system and the breaking σ bond of the hydrogen to the α-carbon, the transition state of the carbanion becomes a planar (σ→ π system), increasing the acidity of the α-hydrogen. Last, the negative charge of thiolate base in the active site is stabilized by the effective positive charge at the N-terminus of the α-helix dipole. Therefore, the thiolate can be maintained as a negative charge to act as base to deprotonate the α-hydrogen of the substrate.

The mechanism of PLP-independent racemases also shows differences depending on the enzyme. First of all, GluR and ProR are reported to have two additional amino acids that act as a catalytic acid/base near the active site cysteines, whereas other enzymes are lacking these. These residues deprotonate/protonate cysteines to realize a thiolate-thiol pair. This pair acts as on acid/base to deprotonate/protonate the α-hydrogen of the substrate. In addition, DapF, ProR, and DcsC exclude water by hinge movements when they bind to substrates, whereas GluR and AspR require one water molecule to maintain a hydrogen-bonding network with the substrate at the active site. These movements create a hydrophobic environment at the active site which helps to decrease the *p*Ka of the α-hydrogen. Further structural studies and substrate specificity experiments on DcsC and RacX are planned.

Chapter 5

Experimental procedures

5.1. General Experimental Detail

5.1.1. Media and bacterial strains

5.1.1.1. Media preparation

All media were purchased from Becton Dickinson and Company (BD, Franklin Lakes, NJ, USA), and prepared following the manufacturer's instructions. Agar plates contained 1.5% (w/v) agar dissolved in the desired broth (1% of tryptone, 0.5% of yeast extract, and 1% of sodium chloride). Antibiotic (kanamycin) was added to the agar solution (50μg/mL) and 20 mL of the solution was poured in each sterile Petri dish in a laminar flow cabinet. The plates were stored at 4-8 ℃.

5.1.1.2. Glycerol stocks

Bacterial strains were stored at -80 ℃ as 20% glycerol stocks. For inoculating fresh cultures, a small amount of the frozen glycerol stock was added into the broth using a sterile pipet tips under flame condition.

5.1.2. General molecular biology techniques

5.1.2.1. Cloning target DNA into expression vector

Purchased target DNA the racemases of interest (BioBasic Inc., Ontario, Canada) were cleaved by restriction enzymes (EcoRI and HindIII) and cleaned using a QIAquick PCR purification kit (Qiagen, Mississauga, ON, Canada). The DNA fragments were ligated with the pET SUMO (digested by the same enzymes) by T4 ligase (Invitrogen) and transformed into *E. coli*. *E. coli* transformants were analyzed by colony PCR and by sequencing the plasmid.

5.1.2.2. Polymerase chain reaction

DNA was amplified by polymerase chain reaction (PCR) using a Techgene Thermal Cycler (Techne Inc., Burlington, NJ, USA) using SUMO F2 primer (5'-TCCGATGGATCTTCAGAGATC-3'), T7 reverse primer (5'-TAGTTATTG-CTCAGCGGTGG-3'). 30 cycles of the PCR reaction were performed, and each cycle was set up at 94 ℃ for 20 seconds (denaturation), 50 ℃ for 30 seconds (annealing), and 68 ℃ for 1 minute (elongation). The PCR product was purified by a QIAquick PCR purification kit (Qiagen).

5.1.2.3. Agarose gel electrophoresis

Agarose gels (0.1% w/v) were prepared using Ultra-Pure Agarose (Invitrogen) in Tris-borate-EDTA (TBE) buffer and were stained with SYBR Safe DNA gel stain (Invitrogen). Sample buffer (6x, Thermo Scientific, Waltham, MA, USA) was added to DNA samples prior to loading into the gel. Gels were run at 84 V and visualized using a Dark Reader Transilluminator (Clare Chemical Research, Dolores, CO, USA).

5.1.2.4. DNA quantification

DNA concentrations were measured by an Implen NanoPhotometer P360 (Implen Inc., Westlake Village, CA, USA), specifically measuring the absorbance at 260 nm.

5.1.2.5. Plasmid isolation

Plasmids from *E. coli* were isolated using QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. The plasmids were eluted in Tris-ethylenediaminetetraacetic acid (EDTA) buffer and stored at -20 °C.

5.1.2.6. DNA sequencing

DNA sequencing was performed at the University of Alberta Molecular Biology Service Unit using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and an ABI 3730 DNA Analyzer (Applied Biosystems).

5.1.2.7. Transformation of pET SUMO plasmid into *E. coli*

A 50µL aliquot of *E. coli* (Invitrogen or New England Biolabs; Ipswich MA, USA) was thawed on an ice bath, and 1-5 µL of plasmid solution was added. After 10 to 15 minutes incubation on ice, the cells were heat-shocked at 42℃ for 30 seconds. The solution was cooled to 0 ℃ using an ice bath, and 250 µL of Super Optimal Broth (SOB medium) was added. The mixture was incubated at 37 ℃, 225 rpm for 45 minutes, spread on antibiotic-coated LB agar plates, and incubated overnight at 37 ℃.

5.1.3. Protein purification

5.1.3.1. Protein expression in *E. coli*

E. coli BL21 (DE3) containing plasmid of interest (glycerol stock) was inoculated in sterile LB broth (1% of Tryptone, 0.5 of Yeast, 1% of sodium chloride and 50 µg/mL Kanamycin) under aseptic conditions. The *E. coli* transformant was grown at 37 °C with shaking (225 rpm) to an optical density (OD600) of 0.6–0.8 using kanamycin as selective pressure. Next, the cells were chilled using an ice bath for 10 minutes before isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.2 mM) was added to the culture to induce protein expression. The culture was then incubated at 23 °C for 16 hours with shaking (225 rpm), and the cells were harvested by centrifugation (6300 rpm, 4 °C, 30 min). The cells were temporarily stored at -80 ℃.

5.1.3.2. Cell lysis

Bacterial cells were resuspended in 20 mL of lysis buffer (100 mM sodium-phosphate (pH 7.8), 300 mM NaCl, 20 mM imidazole, 1 mM EDTA) with 2 mg of lysozyme and 10 units of DNase I (Thermo Scientific). The cells were lysed using a TS Series Bench Top Cell Disruptor (Constant Systems Ltd., Low March, UK) or Branson Sonifier 450 (Branson Ultrasonics Corp., Danbury, CT, USA). The bacterial cell solution was passed through at 20 kpsi, followed with 10–20 mL lysis buffer for washing that was collected as well. The lysate was centrifuged (15,000 rpm, 30 minutes), and the supernatant containing the fusion protein was isolated. For sonication procedure, the bacterial suspension was transferred to 50 mL of falcon tube, and the cells were lysed by 6 cycles of 30 second sonication followed by 30 second rest on ice.

5.1.3.3. Tris-glycine SDS-PAGE

Protein samples were analyzed using 4-20% Mini-PROTEIN TGX precast gel (Bio-Rad, Hercules, CA, USA). Protein samples were mixed with 2x Laemmli Sample Buffer (Bio-Rad) including 2x Laemmli Sample Buffer (Bio-Rad) and heated at 100 ℃ for 5 minutes. The protein samples and protein ladder (Bio-Rad) were loaded into the gels and were run at 90 V for 10 minutes and 180 V for 40 minutes. Gels were stained to be visualized with coomassie stain (0.1% w/v Coomassie R-250, 40% ethanol, 10% acetic acid) for 2 minutes in a microwave (700 W), and destained using destaining solution (10% ethanol, 7.5% acetic acid).

5.1.3.4. Dialysis and concentration of proteins

The protein solution was dialyzed using dialysis tubing (Fischer Scientific, Hampton, NH, United States, 30 kDa). The protein solution was transferred to dialysis tubing and clipped by dialysis tubing closures. The tubing was stirred in 3 L of Tris buffer (pH 7.8) for overnight to dialyze the proteins. The protein solution was transferred to Amicon Ultra centrifugal tubes (Millipore Sigma, Burlington, MA, United States), and centrifuged to concentrate the protein (4,000 G, 20 minutes).

5.1.3.5. Protein quantification

Protein concentrations were measured by recording absorbance at 280 nm using an Implen Nanophotometer P360 (Implen Inc.). Molar extinction coefficients were calculated using the Proparam program (ExPASy proteomics server). Protein concentration was calculated by Beer-Lambert Law: $c=A/b\varepsilon$, c is the concentration (mol/L), A is the absorbacne at 280 nm, b is the path length (cm), and ε is the molar extinction coefficient ($L \cdot mol^{-1} \cdot cm^{-1}$)

5.1.3.6. Site-directed mutagenesis

QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) was used to construct plasmids encoding DcsC and tDcsC with cysteine to serine residues mutations in the active site according to the manufacturer's instructions. Primer pair MVB206 (5'-

CCGTTCCGCGCAGTCTGGCAACGGTGCACG-3') and MVB207 (5'-CGTGCACCGTTGCCAGACTGCGCGGAACGG-3') was used to create the C81S mutation, whereas primer pair MVB208 (5'-GTGAAACCCTGGCGTCCGGTTCCGGTGCATG-3') and MVB209 (5'-CATGCACCGGAACCGGACGCCAGGGTTTCAC-3') was used to create the C227S mutation. With those primers, PCR was performed with pET SUMO-DcsC plasmid as template. The presence of these mutations was verified by nucleotide sequencing of the whole coding region. The resulting plasmids were transformed into *E. coli* BL21(DE3) to overexpress and isolate the mutant enzymes.

5.1.4. Protein purification

5.1.4.1. Ni-NTA affinity chromatography

Ni-NTA (Nickel-nitrilotriacetic acid) Superflow resin (Qiagen) was used to purify His-tagged proteins. Above 1 to 2 mL of resin slurry (50% suspension in 30% ethanol) was added to 40 mL of cell free lysate and stirred for 1 hour. The mixtures of proteins and Ni-NTA resins were loaded on a column (Biorad, USA) by gravity. The resins were washed with lysis buffer with 5 mL of 20 mM imidazole and the proteins were eluted by lysis buffer with 10 mL of 300 mM imidazole.

5.1.4.2. Size exclusion chromatography

Size exclusion chromatography was performed on a BioLogic LP system (Biorad, USA) using a gravity-packed glass column (300 x 18 mm, 180 mL bed volume, Sephadex G-100). The gel bed was prepared according to the manufacturers recommendation. Briefly, 15 g of dry Sephadex G-100 medium was swollen for 72 hours in 200 mL of buffer (25 mM NaCl, 10 mM TrisHCl, pH 7.6, 1 mM DTT) and air-bubble-free transferred into the glass column. The column was equilibrated with 5 column-volumes (CV) of buffer at a flow rate of 1 mL/min. A protein cocktail of BSA (bovine serum albumin, Mw =66 kDa), ovalbumin (Mw = 44 kDa) and lysozyme (Mw =15 kDa) was used as standard for the calibration of retention times. Preparative gel filtration was performed on a 1.5 mL sample of the dialyzed DcsC protein (c = 8 mg/ml) at a flowrate of 1 mL/min and 2 mL fractions were collected with an autosampler. UV-absorbance at 280 nm was monitored throughout the run and used to detect protein fractions. Pure DcsC-containing fractions were pooled and concentrated using amicon filter units (15 mL, 30 kDa cutoff).

5.1.5. SUMO protease digestion

To cleave the SUMO tag, approximately 20 mg of fusion protein was incubated with 25 units of His-tagged SUMO protease (McLab, South San Francisco, CA) in 50 mM Tris-HCl (pH 8.0), 1 mM DTT, 0.2% Igepal CA-360 (Sigma) and 150 mM NaCl in a total volume of 5 mL. Digestion of the fusion protein was performed for 20 h at 4 °C. Cleavage of the fusion protein was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

5.1.6. Mass spectrometry

5.1.6.1. ESI spectrometry

Electrospray ionization (ESI) was performed using a Agilent Technologies 1100 HPLC with G1946A MSD (Agilent Technologies, Santa Clara, CA, USA). Single Quadrupole MS (Agilent Technologies) was used to collect MS data, which were processed using ChemStation LC/MSD software (Agilent Technologies).

5.1.6.2. ESI-TOF MS

Electrospray ionization-Time of flight mass spectrometry (ESI-TOF MS) was performed using Agilent Technologies 6220 oaTOF (Agilent Technologies, Santa Clara, CA, USA). Orthogonal acceleration TOF (oaTOF) was used to collect MS data, and Full-scan MS (high-resolution analysis) with two-point lock mass correction mode was used.

5.1.6.3. LC-MS/MS

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed using a nanoAcquity column (100 pore Å size, 75 µm x 15 cm, 3 µm Atlantis dC18; Waters, MA). Quadrupole-time of flight Premier MS (Micromass, UK) was used to collect MS/MS

data, which were processed using PEAKS 5.1 software (Bioinformatics Solutions, Waterloo, ON, Canada).

5.1.7. NMR spectroscopy

Nuclear magnetic resonance (NMR) spectra were recorded on a Varian Inova 600 MHz spectrometer (equipped with a triple-resonance HCN probe) or a Varian VNMRS 700 MHz spectrometer (equipped with a triple-resonance HCN cold probe and z-axis pulsed-field gradients) at the Department of Chemistry of University of Alberta at ambient temperature. NMR data were analyzed using Agilent VnmrJ 4.2 software (Agilent). Experimental details are described in the subsequent relevant section. Chemical shifts are reported in parts per million (ppm).

5.1.8. Circular dichroism spectroscopy

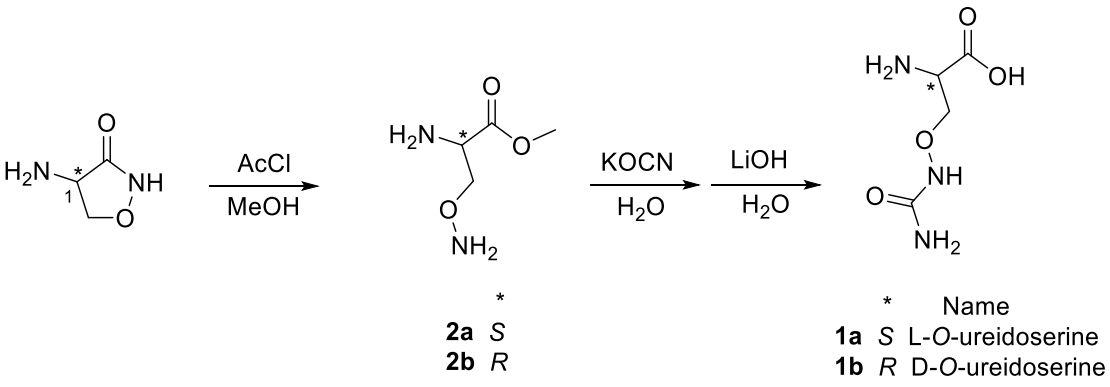
Circular dichroism (CD) data were obtained using a OLIS DSM 17 CD spectrophotometer with a 0.2 mm quartz cuvette. For broad range scan, five scans were run for each sample at 1 nm increment from 195 nm to 400 nm. For time-dependent scan to obtain kinetic data, fixed wavelength (206 nm) scans were performed every 0.5 seconds. Baseline spectra of solvent reference samples were subtracted from the respective protein samples. The temperature and solutions used are described in the relevant sections.

5.1.9. Bioinformatics, chemical and protein structural programs

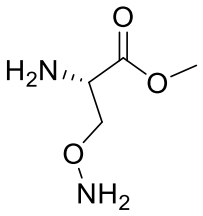
The NCBI-BLAST (National Center for Biotechnology Information-Basic Local Alignment Search Tool) was used for genome mining. Cluster W and Cluster Omega (The European Bioinformatics Institute, EMBL-EBI) were used for sequence alignment. PyMOL was used to generate the 3D structures of protein, and structure homology modeling was done using the SWISS-MODEL server. The crystal structure information was obtained from PDB (Protein Data Bank). Chemdraw Prime 16.0 (CambridgeSoft, Waltham, Massachusetts, USA) was used to draw all chemical structures and for the calculation of the mass of compounds.

5.2. Experimental Procedures for mechanistic studies of DcsC

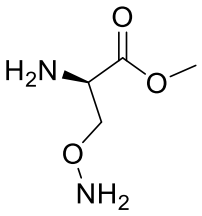
5.2.1. O-ureidoserine synthesis



5.2.1.1. D/L -Aminoserine methyl ester dihydrochloride (2a, 2b)



2a L-aminoserine methyl ester



2b D-aminoserine methyl ester

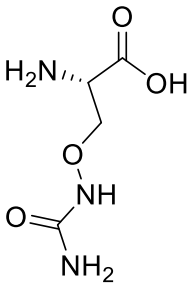
To a solution of dry MeOH (100 mL) in an ice-bath, was added acetyl chloride (7.14 mL, 100 mmol), and stirred under argon for 15 minutes. Cycloserine (1.00 g, 9.80 mmol) was added to the solution, and the cloudy solution was warmed to room temperature for 15 minutes. A reflux condenser was added to the flask, and the reaction mixture was warmed to 70 °C and stirred for 15 hours. Insoluble material was removed by filtration. To a stirring solution of the filtrate was added, ethyl acetate (700 mL) was added dropwise with stirring over 1 hour, at which point the product precipitated as a white solid.

L-aminoserine methyl ester dihydrochloride (**2a**), 74% yield; IR (MeOH, cm^{-1}), 2899, 2656, 1751, 1507, 1249, 1043; $[\alpha]_D^{23}$+12.31 (c = 0.45, MeOH); ^{1}H NMR (400 MHz, CD_3OD), δ (ppm) 4.59-4.55 (m, 3H, α-CH + β-CH_2), 3.90 (s, 3H, -OCH_3); ^{13}C NMR (125 MHz, CD_3OD), δ (ppm) 167.5, 72.7, 54.3, 52.8; ESI-HRMS $[M+H]^+$, 135.0765 (observed), 135.0764 (expected).

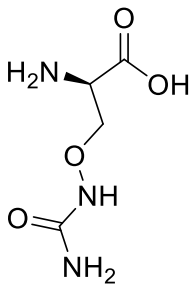
D-aminoserine methyl ester dihydrochloride (**2b**), 98% yield; IR (MeOH, cm^{-1}) 2919, 2665, 1751, 1507, 1248, 1041; $[\alpha]_D^{23}$-12.26 (c = 0.90, MeOH); ^{1}H NMR (400 MHz, CD_3OD), δ (ppm) 4.59-4.55 (m, 3H, α-CH + β-CH_2), 3.90 (s, 3H, -OCH_3); ^{13}C NMR (125 MHz,

CD_3OD), δ (ppm) 167.5, 72.7, 54.3, 52.8; ESI-HRMS $[M+H]^+$, 135.0765 (observed), 135.0764 (expected).

5.2.1.2. *D*/*L*-O-ureidoserine (1a, 1b)



1a L-*O*-ureidoserine



1b D-*O*-ureidoserine

To a stirred solution of aminoserine methyl ester **2a** or **2b** (300 mg, 2.20 mmol) in water (10 mL) was slowly added a solution of potassium isocyanate (180 mg, 2.22 mmol) in water (10 mL). Upon complete addition, the solution was determined to be pH 4-5. The pH was maintained in this range pH 4-5, to obtain maximal yield. The solution was stirred for 2h at 20 °C, and Checked by TLC (7:3 = iPrOH: conc. NH_4OH). Upon completion, the pH was adjusted to ~pH 1 with 1M HCl. Solvent was removed under reduced pressure, and the oily product used in the next step without purification. To the crude product (2.2 mmol) in water (20 mL) was added a solution of 2M KOH (2.2 mL, 4.4 mmol). The reaction was stirred at room temperature until consumption of starting material by TLC (iPrOH:conc. NH_4OH = 7:3). The mixture was adjusted to pH ~ 4 using 1 M HCl, and excess EtOH (10 vol.) was added to precipitate the product. The product was further purified through a Varian BondElut C18 SPE cartridge. The cartridge was washed with 0.1 % aqueous TFA (20 mL) and then the product was applied in 0.1% TFA (1 mL). The product was eluted with 0.1 % aqueous TFA (20 mL) followed by MeCN-H_2O/TFA (20 mL, 1:1:0.001).

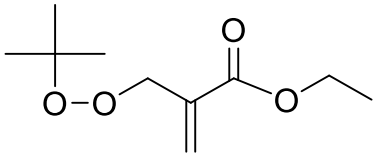
Fractions were collected (1 mL) and those containing product were combined and lyophilized to yield the title compound as a white solid.

L-ureidoserine trifluoroacetate salt (**1a**): 41% yield; IR (MeOH, cm^{-1}) 3182, 1665, 1643, 1504, 1408, 1358, 1100; $[\alpha]_D^{23}$ -15.9 (c = 0.71, H_2O); ^1H NMR (400 MHz, D_2O), δ (ppm) 4.27-4.29 (m, 2H, β-CH_2), 4.18 (dd, 1H, α–CH, J = 5.2, 4.0 Hz); ^{13}C NMR (125 MHz, D_2O), δ (ppm) 172.3, 164.2, 75.1, 54.6; ESI-HRMS $[M-H]^-$, 162.0518 (observed), 162.0520 (expected).

D-ureidoserine trifluoroacetate salt (**1b**): 37% yield; IR (MeOH, cm^{-1}) 3182, 1665, 1643, 1504, 1408, 1358, 1100; $[\alpha]_D^{23}$ +15.1 (c = 0.73, H_2O); ^1H NMR (400 MHz, D_2O), δ (ppm) 4.27-4.29 (m, 2H, β-CH_2), 4.18 (dd, 1H, α-CH, J = 5.2, 4.0 Hz); ^{13}C NMR (125 MHz, D_2O), δ (ppm) 172.3, 164.2, 75.1, 54.6; ESI-HRMS $[M-H]^-$, 162.0518 (observed), 162.0520 (expected).

5.2.2. Synthesis of DcsC inhibitors

5.2.2.1. Ethyl 2-*tert*-butylperoxomethylacrylate (4)



4

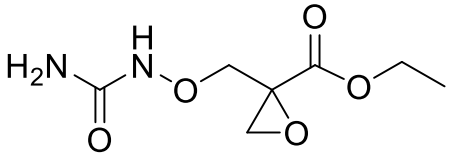
To a solution of *tert*-butylhydroperoxide (3.30 mL, ~5.5 M solution in decanes, 15.5 mmol) in dry CH_2Cl_2 (45 mL), was added sodium hydride (621 mg, 15.5 mmol) at -5 °C and the

reaction stirred for 15 minutes. Ethyl 2-bromomethylacrylate (1.00 g, 7.24 mmol) was added, and the reaction mixture was stirred at -5℃ for 4 hours. The solvent was

evaporated under reduced pressure and white impurities were filtered. The product was purified by column chromatography (hexane:EtOAc = 97:3, R_f = 0.27) to yield a colourless liquid (78%). Product was stored in a -20℃ fridge.

4: IR (CH_2Cl_2, cm^{-1}), 2980, 2935, 1725, 1364, 1303, 1246, 1156, 1027; ^1H NMR (600 MHz, $CDCl_3$), δ (ppm) 6.34 (dt, 1H, J = 1.5, 0.73 Hz, $=CH_2$), 5.89 (d, 1H, J = 1.5 Hz, $=CH_2$), 4.65 (dd, 2H, J = 1.3, 0.7 Hz, tBuOO-CH_2), 4.21 (q, 2H, J = 7.1 Hz, -OCH_2CH_3), 1.29 (t, 3H, J = 7.1 Hz, -OCH_2CH_3), 1.23 (s, 9H, -$OOC(CH_3)_3$); ^{13}C NMR (125 MHz, $CDCl_3$), δ (ppm) 165.9, 136.1, 128.2, 80.6, 60.9, 26.4, 14.3; ESI-HRMS $[M+Na]^+$, 225.1094 (observed), 225.1097 (expected).

5.2.2.2. (+/-) Ethyl 2-(ureidooxymethyl)oxirane-2-carboxylate (5)

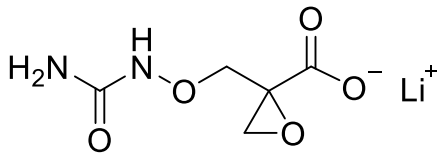


5

Peroxide **4** (576 mg, 2.85 mmol) and hydroxyurea (233 mg, 3.07 mmol) were dissolved in dry DMF (20 mL) and the mixture cooled to 0 °C using an ice-water bath. To this mixture was added potassium *tert*-butoxide (1M in THF, 1.40 mL, 1.40 mmol) and the solution turned light yellow. The reaction mixture was stirred at 0 °C for 1 hour and then at room temperature for 4 hours. The solvent was removed under reduced pressure. The crude product was purified by column chromatography (CH_2Cl_2:MeOH = 97:3, R_f = 0.25), to yield the product as a waxy white solid (33%).

5: IR ($CHCl_3$ cast, cm^{-1}) 3466, 3188, 2986, 1728, 1690, 1588, 1185, 1143; ^1H NMR (400 MHz, $CDCl_3$), δ (ppm) 7.57 (br s, 1H, NH), 5.95-5.70 (br s, 2H, NH_2), 4.37 (d, 1H, J = 11.0 Hz, -CH_2ONH), 4.26 (qd, 1H, J= 7.1, 2.0 Hz, -OCH_2CH_3), 4.08 (d, 1H, J = 11 Hz, -CH_2ONH), 3.17 (d, 1H, J = 5.9 Hz, oxirane-CH_2), 2.98 (d, 1H, J = 5.9 Hz, oxirane-CH_2), 1.30 (t, 3H, J = 7.1 Hz, -OCH_2CH_3); ^{13}C NMR (150 MHz, CDCl3), δ (ppm) 169.1, 161.4, 75.8, 62.6, 54.2, 50.3, 14.1; ESI-HRMS $[M+Na]^+$, 227.0634 (observed), 227.0638 (expected).

5.2.2.3. (+/-) 2-(ureidooxymethyl)oxirane-2-carboxylic acid (6)

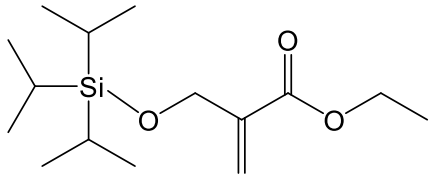


6

To ester **5** (5.50 mg, 0.0230 mmol) in water (0.7 mL) was added LiOH (2 mg, 0.046 mmol) at 20 ℃. The solvent was removed by lyophilization once the reaction was deemed to be complete by TLC (95:5, CH_2Cl_2-MeOH) (96%).

6: IR (MeOH, cm^{-1}) 3325, 1670, 1618, 1429, 1024; ^1H NMR (500 MHz, D_2O), δ (ppm) 4.65 (d, 1H, J = 11.4 Hz, -CH_2ONH), 3.82 (d, 1H, J = 11.4 Hz, -CH_2ONH), 2.98 (s, 2H, oxirane-CH_2); ^{13}C NMR (125 MHz, D_2O), δ (ppm) 174.9, 163.4, 77.5, 57.1, 49.1; ESI-HRMS $[M-H]^-$, 175.0353 (observed), 175.0360 (expected).

5.2.2.4. Ethyl 2-triisopropylsilylhydroxymethylacrylate (8)

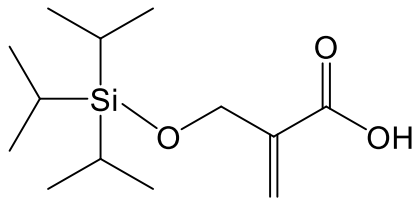


8

Ethyl hydroxymethylacrylate (2.21 g, 17.0 mmol) and imidazole (1.78 g, 26.2 mmol) were dissolved in CH_2Cl_2 (20 mL) and triisopropylsilyl chloride (TIPSCl, 4 mL, 18.7 mmol) was added to the solution at 0°C. The solution was stirred at 0°C for 30 minutes and warmed to room temperature. After 30 minutes stirring, the reaction was quenched by addition of saturated NH_4Cl solution (10 mL). The organic phase was extracted with CH_2Cl_2 (20 mL). The combined organic phases were washed with brine and dried with Na_2SO_4. The solution was concentrated by rotary evaporator, and the residue purified by silica gel column chromatography (hexane:EtOAc = 7:1, R_f = 0.24) to isolate product (92%).

8: IR (CH_2Cl_2, cm^{-1}) 3537, 2942, 2865, 1721, 1463, 1308, 1104, 882; ^1H NMR (400 MHz, CD_3OD) δ (ppm) 6.26 (dd, 1H, $=CH_2$), 5.98 (dd, 1H, $=CH_2$), 4.44 (t, 2H, -OC**H₂**CH3), 4.18 (q, 2H, -OC**H₂**C-), 1.19 (m, 3H, -SiC**H**CH_3), 1.16-1.02 (m, 21H, -SiCHC**H₃** and -CH2C**H₃**); ^{13}C NMR (150 MHz, CD_3OD) δ(ppm) 166.6, 139.8, 124.0, 61.8, 51.8, 18.2, 12.2; ESI-HRMS $[M+Na]^+$, 309.1867 (observed), 309.1862 (expected).

5.2.2.5. 2-Triisopropylmethylacrylic acid (9)

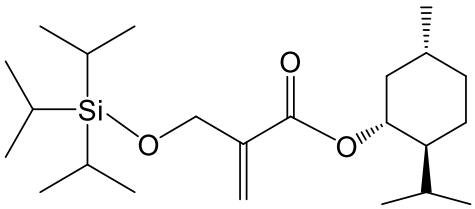


9

To a mixture of THF (20 mL) and water (20 mL) was added ethyl triiopropylhydroxymethylacrylate (**8**) (5.04 g, 17.6 mmol) and lithium hydroxide monohydrate (2.21 g, 52.6 mmol). The mixture was stirred for 72 hours at room temperature and quenched by addition of 5% aq. HCl (20 mL). The organic solution was extracted with diethyl ether, washed with brine solution, and dried with Na_2SO_4. The solution was concentrated under reduced pressure and the residue was purified by column chromatography (hexane:EtOAc = 2:1, R_f = 0.50) to isolate product (94%).

9: IR (CH_3OH, cm^{-1}) 2943, 2866, 1698, 1634, 1463, 1280, 1107, 882; ^1H NMR (500 MHz, $CDCl_3$) δ(ppm) 6.44 (dd, 1H, $=CH_2$), 6.12 (dd, 1H, $=CH_2$), 4.50 (t, 2H, -OC**H₂**CH_3), 1.19 (m, 3H, -SiC**H**CH_3), 1.17 (d, 18H, -SiCHC**H₃**); 13C NMR (150 MHz, CD_3OD) δ (ppm) 170.3, 138.9, 126.2, 61.6, 60.4, 21.1; ESI-HRMS $[M-H]^-$, 257.1569 (observed), 257.1573 (expected).

5.2.2.6. (1*R*, 2*S*, 5*R*)-(-)-Menthyl 2-triisopropylsilylhydroxymethylacrylate (10)



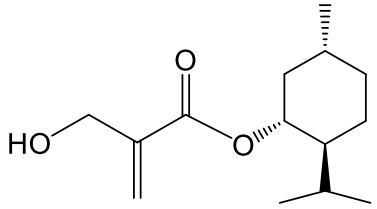
10

To a solution of acrylic acid **9** (0.203 g, 0.786 mmol) in CH_2Cl_2 (20 mL) was added triethylamine (0.119 mL, 0.850 mmol) and (1*R*, 2*S*, 5*R*)-(-)-menthyl chloroformate (0.169 mL, 0.786 mmol) at 0℃. The solution was stirred for 5 minutes at 0 ℃, and the DMAP (12.6 mg, 0.103 mmol) and menthol (126 mg, 0.801 mmol) were added 0 ℃. The mixture was stirred react overnight at room temperature. The reaction mixture was diluted with

CH_2Cl_2 (15 mL), washed with saturated with $NaHCO_3$ (10 mL), 0.1M HCl (10 mL) and brine (10 mL). The aqueous layers were extracted with CH_2Cl_2 (10 mL) and the combined extracts were dried over anhydrous $MgSO_4$, filtered and concentrated under reduced pressure. The desired product **9** was purified by column chromatography (hexane:diethyl ether = 9:1, Rf = 0.32) to yield the product (45%)

10: IR (CH_2Cl_2, cm^{-1}) 3512, 2981, 2945, 2910, 2835, 1712, 1491, 1433, 1312, 1165; ^1H NMR (500 MHz, $CDCl_3$) δ(ppm) 6.23 (dd, 1H, $=CH_2$), 5.98 (dd, 1H, $=CH_2$), 4.80 (td, 1H, -OC**H** in menthyl), 4.45 (t, 2H, -OC**H₂**CH_3), 1.35 - 2.05 (m, 9H, C**H** and C**H₂** in menthol), 1.13 (m, 3H, -SiC**H**CH_3), 1.09 (d, 18H, -SiCHC**H₃**), 0.92 (d, 9H, -$CHCH_3$); ^{13}C NMR (150 MHz, CD_3OD) δ (ppm) 165.6, 140.3, 123.2, 61.8, 47.5, 40.8, 34.3, 31.3, 26.5, 23.7, 22.0, 20.7, 16.5; ESI-HRMS $[M+Na]^+$, 419.2961 (observed), 419.2957 (expected).

5.2.2.7. (1R, 2S, 5R)-(-)-Menthyl 2-hydroxymethylmethylacrylate (11)



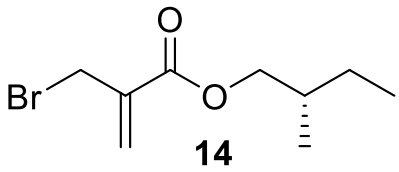
11

To the solution of acrylate **10** (0.560 g, 1.41 mmol) in 15 mL of THF was added tetra-N-butylammonium fluoride solution (1.60 mL, 1.60 mmol, 1M in THF). The solution was stirred for 2 hours at room temperature. The reaction mixture was washed with saturated $NaHCO_3$ (10 mL) and extracted with Et_2O (3 x 5 mL). The organic layers were dried with $MgSO_4$, filtered, and concentrated under reduced pressure to obtain the product (92%).

11: IR (CH_2Cl_2, cm^{-1}) 3415, 2983, 2939, 2907, 2872, 1706, 1636, 1463, 1448, 1387, 1152,

1051; ^{1}H NMR (500 MHz, $CDCl_3$) δ(ppm) 6.27 (dd, 1H, =CH2), 5.84 (dd, 1H, =CH_2), 4.83 (td, 1H, -OC**H** in menthyl), 4.36 (t, 2H, -OC**H₂**CH_3), 1.41 - 2.05 (m, 9H, C**H** and C**H₂** in menthol), 0.92 (d, 9H, -$CHCH_3$); ^{13}C NMR (150 MHz, CD_3OD) δ (ppm) 125.9, 62.2, 47.3, 40.3, 34.1, 31.2, 26.3, 23.1, 22.0, 20.2, 16.0; ESI-HRMS $[M+Na]^+$, 263.1627 (observed), 263.1623 (expected).

5.2.2.8. *(S)*-2-Methylbutyl bromomethylacrylate (14)



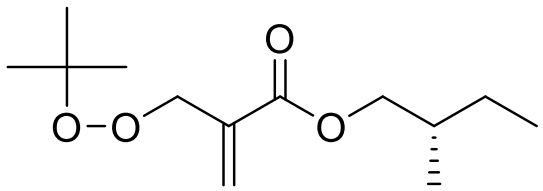
14

To a solution of bromomethylacrylic acid (0.500 g, 3.03 mmol) and diisopropyl azodicarboxylate (DIAD, 0.600 mL, 3.03 mmol) in diethyl ether (10 mL) was added a solution of *(S)*-2-methyl-1-butanol (0.486 mL, 4.50 mmol) and triphenylphosphine (0.787 g, 3.03 mmol) in diethyl ether (10 mL) dropwise at 0 °C. The mixture was stirred at 0 °C for 30 minutes, and then at room temperature for 72 hours. The mixture was filtered and washed with Et_2O (20 mL). After concentration of the filtrate, the crude product was purified by flash column chromatography (Eluent Hexane:EtOAc = 96:4) and confirmed by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.625) to furnish the acrylate as a yellow liquid (1.47 g, 27%).

14: IR (CH_2Cl_2, cm^{-1}) 2981, 2933, 2901, 2869, 1722, 1632, 1453, 1444, 1398, 1186; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 6.35 (d, 1H, J=1.0 Hz, =CH), 5.94 (d, 1H, J=1.0 Hz, =CH), 4.20 (d, 2H, J=1.2 Hz, O-CH_2), 4.13 – 4.01 (dd, 2H, Br-CH_2), 1.60 (d, 1H, CH-CH_3), 1.29 (m, 2H, -C**H₂**CH_3), 0.98 (t, 6H, -CH_2C**H₃** + -CHC**H₃**). ^{13}C-NMR (150 MHz, $CDCl_3$) δ 11.2,

16.5, 26.1, 31.6, 69.9, 128.9, 137.6, 165.0; ESI-HRMS $[M+Na]^+$, 257.0155 (observed), 257.0152 (expected).

5.2.2.9. *(S)*-2-Methylbutyl *tert*-butylperoxomethylacrylate (15)

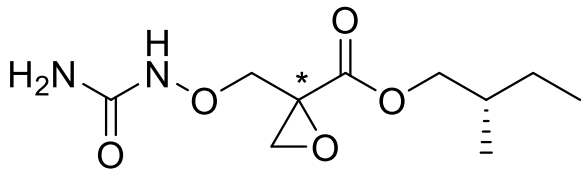


15

A solution of *tert*-butylhydroperoxide (5.5 M solution in decane, 0.350 mL, 2.01 mmol) in dry CH_2Cl_2 (10 mL) was stirred at −5 °C under a stream of argon. To this was added sodium hydride (0.601 g, 1.50 mmol, 60% dispersion in mineral oil). After 15 minutes, acrylate **14** (0.194 g, 0.820 mmol) was added dropwise at −5 °C. After 4 hours, the mixture was filtered and washed with dry CH_2Cl_2. After concentration of the filtrate, the crude product was purified by flash column chromatography (Hexane:EtOAc=96:4) and confirmed to be pure by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.47) to obtain **15** as a yellow liquid (0.72 g, 46%).

15: IR (CH_2Cl_2, cm^{-1}) 2978, 2931, 2912, 2876, 1722, 1631, 1441, 1392, 1186; ^{1}H-NMR (400 MHz, $CDCl_3$) δ 6.38 (d, 1H, J=1.0 Hz, =CH), 5.93 (d, 1H, J=1.0 Hz, =CH), 4.68 (d, 2H, J=1.2 MHz, O-CH_2), 4.10 – 3.98 (dd, 2H, Br-CH_2), 1.56 (d, 1H, CH-CH_3), 1.29 (m, 2H, -C**H₂**CH_3), 1.24 – 1.22 (s, 9H, *tert*-butyl-CH_3), 0.97 – 0.91 (t, 6H, -CH_2C**H₃** + -CHC**H₃**). ^{13}C-NMR (150 MHz, $CDCl_3$) δ 11.2, 16.5, 26.3, 34.2, 69.9, 73.2, 128.0, 137.1, 166.2; ESI-HRMS $[M+NH_4]^+$, 262.2020 (observed), 262.2018 (expected).

5.2.2.10. *(S)*-2-methylbutyl 2-(ureidooxymethyl)oxirane-2-carboxylate (16)

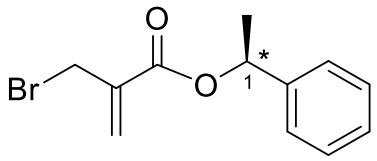


16

Compound **15** (0.106 g, 0.434 mmol) and hydroxyurea (0.651 g, 0.855 mmol) were dissolved in dry DMF (10 mL) and the mixture cooled to 0 °C using an ice-water bath. To this mixture was added a solution of potassium *tert*-butoxide (0.40 mL, 0.40 mmol, 1 M in THF) dropwise and the solution turned light yellow. The mixture was stirred at 0 °C for 30 minutesvand and then at room temperature for 4 hours. The solvent was then removed under reduced pressure. The crude yellow mixture was purified by silica gel chromatography (DCM:MeOH = 98:2, R_f = 0.21) to yield the product **16** as a liquid (0.034 g, 34%).

16: IR (CH_2Cl_2, cm^{-1}) 3411, 3397, 2989, 2941, 2911, 2878, 1711, 1681, 1452, 1374, 1174; ^{1}H-NMR (500 MHz, $CDCl_3$) δ 4.42 (d, 1H, J=11Hz, CH_2ONH), 4.14 – 3.98 (m, 3H, Br-CH_2 + CH_2ONH), 3.20 (d, 1H, J = 5.5 Hz, oxirane-CH_2), 3.02 (d, 1H, J = 6.0 Hz, oxirane-CH_2), 1.78 (m, 3H, CH-C**H$_3$**), 1.57 (m, 1H, CH-CH_3), 1.25 (m, 2H, -C**H$_2$**CH_3), 0.97 – 0.91 (t, 6H, -CH_2C**H$_3$** + -CHC**H$_3$**); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 11.3, 16.5, 26.2, 34.6, 51.7, 65.6, 69.4, 153.7, 170.1; ESI-HRMS $[M+Na]^+$, 269.1116 (observed), 269.1113 (expected).

5.2.2.11. 1-Phenylethyl 2-bromomethylacrylate (17a, 17b)



17a 1R
17b 1S

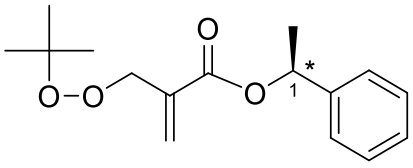
To a solution of bromomethylacrylic acid **13** (1.50 g, 9.09 mmol) and diisopropyl azodicarboxylate (1.80 mL, 9.14 mmol) in diethyl ether (15 mL) was added a solution of *(S)*-1-phenylethanol (to synthesize **17a**) or *(R)*-1-phenylethanol (to synthesize **17b**) (1.20 mL, 9.90 mmol) and triphenylphosphine (2.36 g, 9.09 mmol) in diethyl ether (15 mL), dropwise at 0 °C. The mixture was stirred at 0 °C for 30 minutes and then at room temperature for 48 hours. The mixture was filtered and washed with Et_2O (20 mL). After concentration of the filtrate, the crude product was purified by flash column chromatography (Hexane:EtOAc = 95:5) and confirmed to be pure by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.625) to obtain the acrylate as a liquid (1.47 g, 60%).

17a *(R)*: IR (CH_2Cl_2, cm^{-1}) 3087, 3063, 2960, 2930, 2872, 1722, 1632, 1453, 1186; ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.40 (m, 5H, Ph-H), 6.39 (d, 1H, J=1.0Hz, =CH), 6.01 (d, 1H, J=1.0Hz , =CH), 5.95 (q, 1H, J=6.9 Hz, Ph-CH), 4.21 (dd, 2H, J=2.5, 1.0Hz, Br-CH_2), 1.58 (d, 3H, J=8.5 Hz, CH-CH_3);^{13C}-NMR (150 MHz, $CDCl_3$) δ 20.3, 21.2, 28.5, 73.4, 125.6, 127.5, 128.1, 128.3, 138.3, 141.5, 164.1; $[\alpha]_D^{25}$ = +23.24 (c = 0.500, CH_2Cl_2); ESI-HRMS $[M+Na]^+$, 290.9991 (observed), 290.9991 (expected).

17b *(S)*: Starting with *(R)*-1-phenylethanol (**4b**). IR (CH_2Cl_2, cm^{-1}) 3087, 3064, 2980, 2931, 2872, 1722, 1632, 1453, 1181; ^{1}H-NMR (600 MHz, CDCl3) δ 7.40 (m, 5H, Ph-H), 6.39 (d, 1H, J=1.0 Hz , =CH), 6.01 (d, 1H, J=1.0 Hz , =CH), 5.95 (q, 1H, J=6.9 Hz, Ph-CH), 4.25 (dd, 2H, J=2.5, 1.0 Hz, Br-CH_2), 1.58 (d, 3H, J=6.4 Hz, CH-CH_3). ^{13}C-NMR (150 MHz,

$CDCl_3$) δ 20.1, 21.3, 28.4, 73.5, 125.7, 127.5, 128.0, 128.3, 138.1, 141.5, 164.1; $[\alpha]_D^{25}$ = -26.71 (c = 0.500, CH_2Cl_2). ESI-HRMS $[M+Na]^+$, 290.9991 (observed), 290.9991 (expected).

5.2.2.12. 1-Phenylethyl 2-*tert*-butylperoxomethylacrylate (18a, 18b)



18a 1R
18b 1S

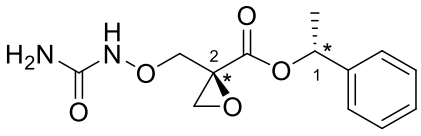
A solution of *tert*-butylhydroperoxide (5.5M solution in decane, 1.50 mL, 8.26 mmol) in dry CH_2Cl_2 (20 mL) was stirred at -5 °C under a stream of argon. To this was added sodium hydride (0.280 g, 7.00 mmol, 60% dispersion in mineral oil). After 15 minutes, **17a** or **17b** (1.47 g, 5.48 mmol) was added dropwise at -5° C. After 4 hours, the reaction mixture was filtered and washed with dry CH_2Cl_2. After concentration of the filtrate, the crude product was purified by flash column chromatography (Hexane:EtOAc = 95:5) and confirmed to be pure by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.375) to obtain **18a** or **18b** as a yellow liquid (0.720 g, 46%).

18a *(R)*: IR (CH_2Cl_2, cm^{-1}) 3089, 3065, 3034, 2980, 2932, 2874, 1723, 1637, 1454, 1363, 1162; ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.38 (m, 5H, Ph-H), 6.35 (d, 1H, J=1.0 Hz , =CH), 5.91 (d, 1H, J=1.0 Hz, =CH), 5.94 (q, 1H, J=6.6 Hz, Ph-CH), 4.63 (dd, 2H, J=3.0, 1.0 Hz, O-CH_2), 1.58 (d, 3H, J=7.0 Hz, CH-CH_3), 1.18 (s, 9H, O-C-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 21.1, 25.1, 72.7, 79.8, 125.8, 127.5, 128.2, 136.7, 141.8, 165.0; $[\alpha]_D^{25}$= +20.73 (c

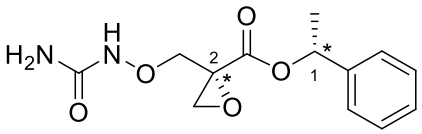
= 0.500, CH_2Cl_2); ESI-HRMS $[M+Na]^+$, 301.1409 (observed), 301.1410 (expected).

18b *(S)*: IR (CH_2Cl_2, cm^{-1}) 3088, 3065, 3034, 2980, 2932, 2873, 1724, 1638, 1454, 1363, 1162; ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.38 (m, 5H, Ph-H), 6.35 (d, 1H, J=1.0 Hz , =CH), 5.91 (d, 1H, J=1.0 Hz, =CH), 5.94 (q, 1H, J=6.6 Hz, Ph-CH), 4.63 (dd, 2H, J=3.0, 1.0 Hz, O-CH_2), 1.58 (d, 3H, J=7.0 Hz, CH-CH_3), 1.18 (s, 9H, O-C-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 21.4, 25.2, 72.9, 79.9, 125.6, 127.5, 128.1, 136.6, 141.7, 165.1; $[\alpha]_D^{25}$ = -19.23 (c = 0.500, CH_2Cl_2). ESI-HRMS $[M+Na]^+$, 301.1410 (observed), found 301.1410 (expected).

5.2.2.13. *(R)*-1-phenylethyl 2-(ureidooxymethyl)oxirane-2-carboxylates 19a (1*R*,2*S*) and 20a (1*R*,2*R*)



19a 1*R* 2*S*



20a 1*R* 2*R*

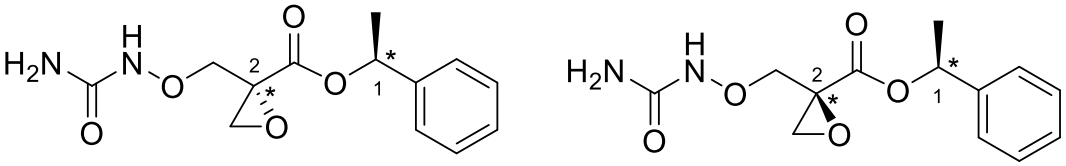
Compound **18a *(R)*** (0.720 g, 2.58 mmol) and hydroxyurea (0.250 g, 3.28 mmol) were dissolved in dry DMF (15 mL), and the mixture cooled to 0 °C using an ice-water bath. To this mixture was added a solution of potassium *tert*-butoxide (0.800 mL, 0.800 mmol, 1M in THF) dropwise, and the solution turned light yellow. The mixture was stirred at 0 °C for 30 minutes, and then at room temperature for 4 hours. The solvent was then removed under reduced pressure. The crude yellow mixture was purified by column chromatography (EtOAc:hexane = 9:1, then pure EtOAc, R_f = 0.21), to yield the diastereomers **19a (1*R*,2*S*)** and **20a (1*R*,2*R*)** as a yellowish solid-liquid mixture (0.170 g, 23%). Repeated re-crystallization from benzene/EtOAc yielded **7a** (1R,2S) as crystalline

material.

19a (1*R*,2*S*): IR (CH_3OH, cm^{-1}) 3415, 3392, 3174, 3040, 2989, 2940, 2878, 1718, 1682, 1457, 1373, 1176; 1H-NMR (600 MHz, $CDCl_3$) δ 7.38 (m, 5H, Ph-H), 5.94 (q, 1H, J=6.5 Hz, Ph-CH), 4.41(d, 1H, J=10.1 Hz, CH_2ONH), 4.13 (d, 1H, J=10.1 Hz, CH_2ONH), 3.16 (d, 1H, J = 6.0 Hz, oxirane-CH_2), 3.02 (d, 1H, J = 6.0 Hz), 1.61 (d, 3H, J=6.5 Hz, CH-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 21.8, 50.1, 74.2, 75.1, 124.8, 128.2, 128.8, 165.4; ESI-HRMS $[M+Na]^+$, 303.0954 (observed), 303.0951 (calculated).

Evaporation of the filtrate after re-crystallization gave pure **20a (1*R*,2*R*)**: ^{1}H-NMR (600 MHz, $CDCl_3$) δ 7.38 (m, 5H, Ph-H), 5.94 (q, 1H, J=6.5Hz, Ph-CH), 4.37 (d, 1H, J=10.1Hz, CH_2ONH), 4.05 (d, 1H, J=10.1Hz, CH_2ONH), 3.23 (d, 1H, J = 6.0 Hz, oxirane-CH_2), 2.96 (d, 0.6H, J = 6.0 Hz), 1.61 (d, 3H, J=6.5 Hz, CH-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 21.8, 50.1, 74.2, 75.1, 124.8, 128.2, 128.8, 165.2; ESI-HRMS $[M+Na]^+$, 303.0952(observed), found 303.0951(calculated).

5.2.2.14. *(S)*-1-phenylethyl 2-(ureidooxymethyl)oxirane-2-carboxylates 19b (1*S*, 2*R*) and 20b (1*S*, 2*S*)



19b 1*S* 2*R*

20b 1*S* 2*S*

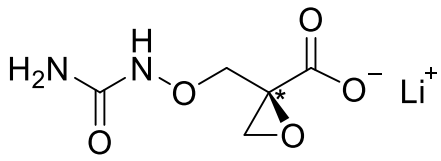
Compound **18b *(S)*** (0.720 g, 2.58 mmol) and hydroxyurea (0.250 g, 3.28 mmol) were dissolved in dry DMF (15 mL), and the mixture cooled to 0 °C using an ice-water bath. To this mixture was added a solution of potassium *tert*-butoxide (0.800 mL, 0.800 mmol, 1M in THF) dropwise, and the solution turned light yellow. The mixture was stirred

at 0 °C for 30 minutes, and at room temperature for 4 hours. The solvent was then removed under reduced pressure. The crude yellow mixture was purified by column chromatography (EtOAc:hexane = 9:1, then pure EtOAc, R_f = 0.21), to yield the diastereomers **19b (1*S*, 2*R*)** and **20b (1*S*, 2*S*)** as a yellowish solid-liquid mixture (0.170 g, 23%). Repeated re-crystallization from benzene/EtOAc yielded **19b (1*S*, 2*R*)** as crystalline material.

19b (1*S*, 2*R*): IR (CH_3OH, cm^{-1}) 3415, 3392, 3171, 3041, 2989, 2939, 2877, 1719, 1683, 1457, 1373, 1177; ^{1}H-NMR(600 MHz, $CDCl_3$) δ 7.38 (m, 5H, Ph-H), 5.94 (q, 1H, J=6.5 Hz, Ph-CH), 4.41(d, 1H, J=10.1 Hz, CH_2ONH), 4.13 (d, 1H, J=10.1 Hz, CH_2ONH), 3.16 (d, 1H, J = 6.0 Hz, oxirane-CH_2), 3.02 (d, 1H, J = 6.0 Hz), 1.61 (d, 3H, J=6.5 Hz, CH-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 21.6, 50.3, 74.7, 124.8, 128.2, 128.8, 165.3. ESI-HRMS $[M+Na]^+$, 303.0950 (observed), 303.0950 (expected).

Evaporation of the filtrate and re-crystallization gave pure **20b (1*S*,2*S*)** showing identical analytical data as the 1R,2R-compound (**19a**) obtained from the *(R)*-precursor.

5.2.2.15. (*S*)-2-(ureidooxymethyl)oxirane-2-carboxylic acid (6a)



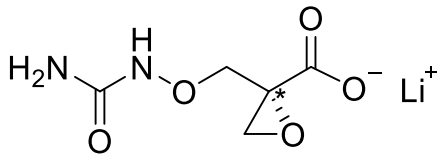
6a (S)-Inhibitor

To ester **19a (1R,2S)** (1.000 mg, 0.00357 mmol) in D_2O (0.7 mL) was added a LiOH solution (0.0100 mL, 0.0100 mmol, 1M solution in D_2O). After 10 minutes, the reaction was deemed to be complete by ^{1}H-NMR and the solvent was removed by lyophilization.

6a *(S)*-inhibitor: IR (CH_3OH, cm^{-1}) 3301 (broad), 2938, 1671, 1619, 1425, 1107; ^{1}H-NMR(600 MHz, D_2O) δ 4.43 (d, 1H, J=10.2Hz, CH_2ONH), 3.72 (d, 1H, J=10.1Hz,

CH_2ONH), 2.95 (s, 2H, oxirane-CH_2); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 49.2, 57.1, 76.7, 163.1, 174.5; melting point 116-118 ℃; ESI-HRMS $[M-H]^-$, 175.0357 (observed), 175.0360 (expected).

5.2.2.16. (*R*)-2-(ureidooxymethyl)oxirane-2-carboxylic acid (6b)

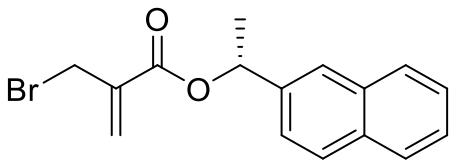


6b (R)-Inhibitor

To ester **19b (1S, 2R)** (1.000 mg, 0.00357 mmol) in D_2O (0.7 mL) was added a LiOH solution (0.0100 mL, 0.0100 mmol, 1M solution in D_2O). After 10 minutes, the reaction was deemed to be complete by ^{1}H-NMR and the solvent was removed by lyophilization.

19b (1S, 2R): IR (CH_3OH, cm^{-1}) 3301 (broad), 2938, 1671, 1619, 1425, 1107; ^{1}H-NMR (600 MHz, D_2O) δ 4.43 (d, 1H, J=10.2Hz, CH_2ONH), 3.72 (d, 1H, J=10.1Hz, CH_2ONH), 2.95 (s, 2H, oxirane-CH_2); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 49.2, 57.2, 76.8, 163.3, 174.7; melting point 116-118 ℃; ESI-HRMS $[M-H]^-$, 175.0358(observed), 175.0360(expected).

5.2.2.17. *(S)*-1-(Naphthalen-2-yl)lethyl 2-bromomethylacrylate (21)



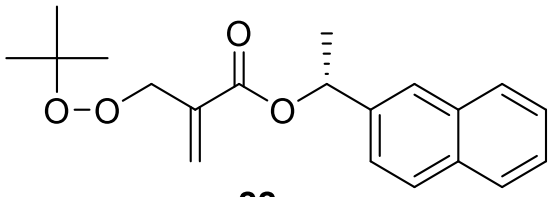
21

To a solution of bromomethylacrylic acid **13** (0.465 g, 2.82 mmol) and diisopropyl

azodicarboxylate (0.670 mL, 3.40 mmol) in diethyl ether (10 mL), was added a solution of *(R)*-1-(2-naphthyl)-ethanol (0.500 g, 2.82 mmol) and triphenylphosphine (0.892 g, 2.82 mmol) in diethyl ether (10 mL), dropwise at 0 °C. The mixture was stirred at 0 °C for 30 minutes, and then at room temperature for 48h. The mixture was filtered and washed with Et_2O (20 mL). After concentration of the filtrate, the crude product was purified by flash column chromatography (Hexane:EtOAc=96:4) and confirmed to be pure by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.65) to obtain the acrylate as a yellow liquid (0.518 g, 58%).

21: IR (CH_2Cl_2, cm^{-1}) 3067, 3054, 3023, 2941, 2923, 2871, 1733, 1612, 1443, 1174; ^{1}H-NMR (400 MHz, CDCl3) δ 7.86 (m, 4H, Ph-H), 7.28 (m, 3H, Ph-H), 6.42 (d, 1H, J=1.0 Hz, =CH), 6.20 (q, 1H, J=6.1 Hz, Ph-CH), 5.99 (d, 1H, J=1.0 Hz , =CH), 4.21 (dd, 2H, J=2.5, 1.0 Hz, Br-CH_2), 1.72 (d, 3H, J=10.2 Hz, CH-CH_3);^{13}C-NMR (150 MHz, $CDCl_3$) δ 20.3, 29.4, 73.9, 124.0, 125.2, 126.1, 127.4, 128.0, 128.3, 129.1, 138.9, 139.7, 164.1; $[\alpha]_D^{25}$ = +19.21 (c = 0.500, CH_2Cl_2). ESI-HRMS [M+Na]+, 341.0160 (observed), 341.0153 (expected).

5.2.2.18. *(S)*-1-(naphthalen-2-yl)ethyl 2-*tert*-butylperoxomethylacrylate (22)



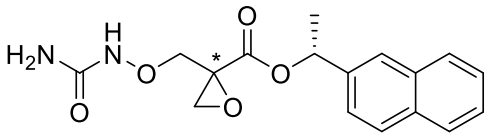
22

A solution of *tert*-butylhydroperoxide (5.5M solution in decane, 0.545 mL, 3.00 mmol) in dry CH_2Cl_2 (10 mL) was stirred at -5 °C under a stream of argon. To this was added sodium hydride (0.118 g, 2.95 mmol, 60% dispersion in mineral oil). After 15 minutes, acrylate **21** (0.312 g, 0.977 mmol) was added by dropwise at -5°C. After 4h, the

mixture was filtered and washed with dry CH_2Cl_2. After concentration of the filtrate, the crude product was purified by flash column chromatography (Hexane:EtOAc=97:3) and confirmed to be pure by TLC (Eluent Hexane:EtOAc = 9:1, R_f = 0.40) to obtain **22** as a yellow liquid (0.255 g, 80%).

22: IR (CH_2Cl_2, cm^{-1}) 3067, 3054, 3023, 2945, 2931, 1722, 1637, 1455, 1364, 1161; ^{1}H-NMR(600 MHz, $CDCl_3$) δ 7.84 (m, 4H, Ph-H), 7.41 (m, 3H, Ph-H), 6.29 (d, 1H, J=1.0 Hz , =CH), 5.87 (d, 1H, J=1.0 Hz, =CH), 6.01 (q, 1H, J=6.6 Hz, Ph-CH), 4.61 (dd, 2H, J=2.7, 1.0 Hz, O-CH_2), 1.55 (d, 3H, J=7.0 Hz, CH-CH_3), 1.18 (s, 9H, O-C-CH_3); ^{13}C-NMR (150 MHz, $CDCl_3$) δ 20.1, 25.6, 71.7, 79.3, 124.5, 125.6, 126.3, 127.1, 128.4, 128.6, 129.2, 139.4, 139.5, 164.3; $[\alpha]_D^{25}$ = +16.43 (c = 0.500, CH_2Cl_2). ESI-HRMS $[M+Na]^+$, 351.1570 (observed), 351.1572 (expected).

5.2.2.19. *(R)*-1-(naphthalen-2-yl)ethyl 2-(ureidooxymethyl)oxirane-2-carboxylate (23, diastereomer)



23

Compound **22** (0.207 g, 0.630 mmol) and hydroxyurea (71.2 mg, 0.936 mmol) were dissolved in dry DMF (15 mL) and the mixture cooled to 0 °C using an ice-water bath. To this mixture was added a solution of potassium *tert*-butoxide (0.800 mL, 0.800 mmol, 1M in THF) dropwise, and the solution turned light yellow. The mixture was stirred at 0 °C for 30 minutes, and then at room temperature for 4 hours. The solvent was then removed under reduced pressure. The crude yellow mixture was purified by column chromatography

(Hexane:EtOAc = 75:25, R_f = 0.33), to yield the diastereomers **23** as a yellowish liquid (0.114 g, 55%).

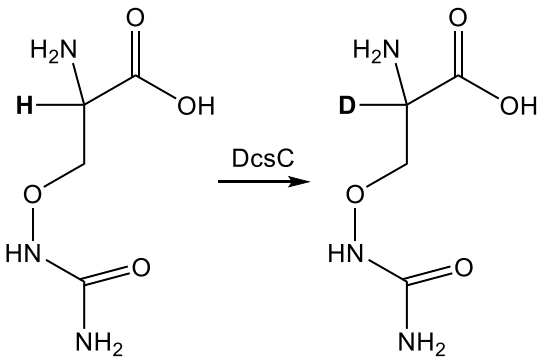
23: IR (CH_3OH, cm^{-1}) 3423, 3387, 3169, 3081, 3033, 2981, 2930, 2872, 1721, 1678, 1465, 1369, 1180; ^{1}H-NMR(400 MHz, CD_3OD) δ 7.87 (m, 4H, Ph-H), 7.43 (m, 3H, Ph-H), 6.10 (q, 1H, J=6.5 Hz, Ph-CH), 4.41(d, 1H, CH_2ONH), 4.03 (d, 1H, CH_2ONH), 3.13 (d, 1H, oxirane-CH_2), 3.03 (d, 1H, oxirane-CH_2), 1.61 (d, 3H, CH-CH_3); ESI-HRMS [M+Na]+, 353.1116 (observed), 353.1113 (calculated).

5.2.3. X-ray diffraction analysis of chiral inhibitor 19a and 19b

Crystals of **19a** (1*R*,2*S*) and **19b** (1*S*,2*R*) suitable for analysis were obtained by slow cooling of a hot solution of each pure isomer in benzene/EtOAc (2 : 1). The intensity data were collected on a Bruker APEX II diffractometer with CuKα radiation (λ = 1.5418 Å) using ω- and φ-scans. Reflections were corrected for background and Lorentz polarization effects. Preliminary structure models were derived by application of direct methods and were refined by fullmatrix least squares calculation based on F2 for all reflections. All hydrogen atoms were included in the models in calculated positions and were refined as constrained to bonding atoms. The crystal data and experimental parameters are summarized in the Appendix and the structures are deposited under CCDC 1588579 (**7a**) and CCDC 1588578 (**7b**).

5.2.4. Enzyme Kinetics of DcsC

5.2.4.1. Enzyme kinetics monitoring by ^{1}H-NMR



O-ureidoserine

To analyze the kinetics of truncated DcsC by NMR, enantiomerically pure *O*-ureidoserine (2-15 mM) was dissolved to Tris buffer (D_2O, 20 mM, pD=7.8). To initiate the reaction, truncated DcsC (3-6 μg) was added and ^{1}H NMR was recorded every 3 minutes. To determine enzyme activity, the peak of β-CH_2 peak (mutiplet, 4.2-4.3 ppm) was compared with the α-CH peak (triplet, 3.9-4.0 ppm). Since truncated DcsC exchanges the α-CH to α-CD, the progress of the reaction was monitored by the disappearance of the α-CH peak.

5.2.4.2. Enzyme kinetics monitored by CD spectroscopy

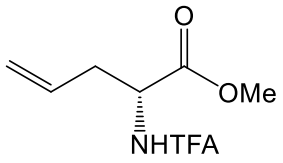
The *O*-ureidoserine samples (300 μL, 2-15 mM, phosphate buffer (pH=7.8), 1 mM DTT) were placed in 1 mm quartz cell, and reactions were monitored at 30°C at 206 nm. The CD signal (in mdeg) was recorded every two seconds, for five minutes. The reaction

was initiated by adding truncated DcsC (2-5 μg), and the quartz cell was directly placed in CD spectrophotometer. The data were processed using Excel and GraphPad prism. The enzyme intial velocity (Vo) was calculated within the linear range of the CD-signal decay (initial 2 minutes), and kinetics were extracted using Michaelis-Menten analysis.

5.3. Experimental procedures for mechanistic studies of RacX

5.3.1. Synthesis of orthogonally protected *meso*-diaminopimelate

5.3.1.1. *N*-Trifluoroacetyl-D-Allylglycine methyl ester (25)



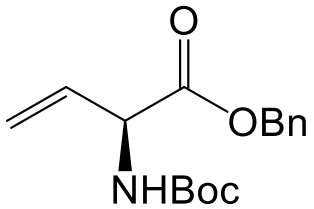
25

To the solution of D-allylglycine (**24**) (3.00 g, 26.0 mmol) in dry methanol (30 mL) was added ethyl trifluoroacetate (6.00 mL, 50.2 mmol) and triethylamine (4.60 mL, 42.7 mmol) at room temperature. The reaction mixture was stirred for 18 hours under argon. When the reaction was complete, the oily residue was dissolved in EtOAc (180 mL) and washed with 10% citric acid (120 mL) and brine (50 mL). The solution was dried with sodium sulfate and concentrate in vacuum, and used for further without purification. The crude product (TFA-D-allylglycine-OH, 4.29 g, 20.3 mmol) was dissolved in dry acetone

(30 mL). To this solution was added potassium carbonate (5.40 g, 39.1 mmol) and methyl iodide (2.70 mL, 43.4 mmol), and the reaction mixture was refluxed for 4 hours. The reaction was monitored by TLC (Hexane:EtOAc=85:15, Rf:0.45). Ethyl acetate (50 mL) was added, and the mixture was washed with saturated $NaHCO_3$ (30 mL) and dried over anhydrous Na_2SO_4. The solvent was evaporated in vacuum, and the residue was purified by column chromatography (Hexane:EtOAc = 85:15, R_f = 0.35) to obtain desired product **25** (4.23 g, 93%).

25: IR ($CHCl_3$ cast, cm^{-1}) 3326, 3087, 3009, 2987, 2960, 1721, 1553, 1441, 1210, 1179; ^1H NMR (500 MHz, $CDCl_3$), δ (ppm) 7.28 (br s, 1H, NH), 5.69 (m, 1H, -CH_2C**H**CH_2-), 5.22 (dd, 1H, -C**H**$_2$CH), 5.20 (dd, 1H, -C**H**$_2$CH), 4.72 (t, 1H, -C**H**COO), 3.83 (t, 3H, -OC**H**$_3$), 2.61-2.74 (m, 2H, -CHC**H**$_2$CH); ^{13}C NMR (150 MHz, $CDCl_3$), δ (ppm) 170.6, 156.6, 130.8, 120.4, 115.6, 60.1, 52.9, 52.0, 35.9; $[\alpha]_D^{25}$=-55.69 (c = 0.8000, CH_2Cl_2); ESI-HRMS [M-H]$^-$, 224.0541 (observed), 224.0534 (expected).

5.3.1.2. Synthesis of *(R)-N*-(*tert*-butoxycarbonyl)-vinylglycine benzyl ester (27)



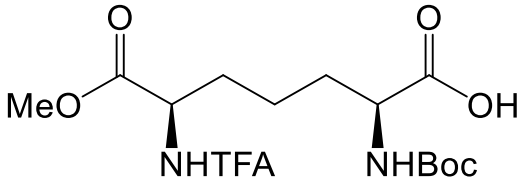
27

A solution of *N*-Boc-L-glutamic acid-α-benzyl ester **26** (9.00 g, 26.7 mmol) and cupric acetate monohydrate (1.20 g, 6.00 mmol) in dry benzene (250 mL) was stirred for 1 hour at room temperature under argon. Lead tetraacetate (23.6 g, 53.4 mmol) was added

and the resulting suspension heated to reflux for 72 hours. The cooled mixture was filtered through Celite, and washed with EtOAc (3 x 300 mL). The filtrate was then washed with water (3 x 180 mL) and brine (180 mL), dried over anhydrous Na_2SO_4 and concentrated in vacuum to afford a brown oil. The residue was purified by column chromatography (Hexane:EtOAc = 9:1, R_f = 0.25) to yield product **27** as a colorless oil (3.23 g, 42%).

27: IR ($CHCl_3$ cast, cm^{-1}) 3373, 3066, 3035, 2978, 2932, 1717, 1643, 1499, 1164; ^1H NMR ($CDCl_3$, 500 MHz) δ 7.39 (m, 5H, -PhH), 5.95 (m, 1H, NH), 5.38 (dd, J = 11.5, 1.0 Hz, 1H, Hβ), 5.27 (dd, J = 10.0, 1.0 Hz, 1H, Hδ), 5.21 (m, 3H, Hδ, CH_2Ph), 4.94 (s, 1H, Hα), 1.47 (s, 9H, Boc); ^{13}C NMR (CDCl3, 125 MHz) δ 170.4, 155.2, 135.1, 132.3, 128.4, 128.6, 128.0, 127.7, 117.8, 80.1, 67.4, 55.7, 28.1; $[\alpha]_D^{25}$ = -6.72 (c = 1.20 g/100 mL, $CHCl_3$); ESI-HRMS $[M+Na]^+$, 314.1363 (observed), 134.1363 (expected).

5.3.1.3. (2*S*,6*R*)-2-(*tert*-butoxycarbonylamino)-7-methoxy-7-oxo-6-(2,2,2-trifluoro-acetamido)heptanoic acid (29)



29

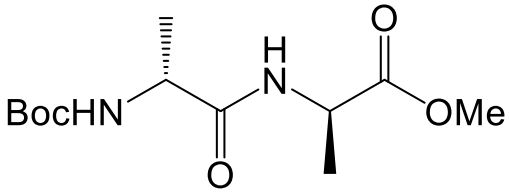
To a solution of **27** (2.00 g, 6.87 mmol) in dry CH_2Cl_2 (40 mL) was added **25** (1.55 g, 6.89 mmol). The mixture was deoxygenated by argon bubbling for 5 minutes. Hoveyda-Grubbs 2nd generation catalyst (0.430 g, 0.686 mmol) was added to the solution, and the mixture was heated to reflux under argon for 16 hours. The mixture was concentrated in vacuum to yield a black oil, which was purified by column chromatography

(Hexane:EtOAc = 95:5, R_f = 0.20) to yield alkene **28** (668 mg, 20%). Palladium (10% on carbon, 300 mg) in dry MeOH (10 mL) was activated under a hydrogen atmosphere for 15 minutes, and alkene **23** (668 mg, 1.36 mmol) was added to the solution. The mixture was stirred for 90 minutes under a hydrogen atmosphere, and then filtered through Celite, eluting with methanol (30 mL). The filtrate was concentrated in vacuum to afford orthogonally protected *meso*-DAP **29** (405 mg, 74%).

29: IR ($CHCl_3$ cast, cm^{-1}) 3306, 3080, 2979, 2958, 2936, 1717, 1555, 1457, 1167; ^1H NMR ($CDCl_3$, 700 MHz) δ 4.45 (m, 1H, Hε), 4.00 (m, 1H, Hα), 3.75 (s, 3H, OCH3), 1.97 (m, 1H, Hδ), 1.79 (m, 2H, Hβ), 1.64 (m, 1H, Hδ), 1.47 (m, 2H, Hγ), 1.45 (s, 9H, Boc); ^{13}C NMR ($CDCl_3$, 150 MHz) δ 177.1, 172.6, 159.1, 158.0, 117.4, 80.5, 54.3, 53.9, 53.0, 32.6, 31.3, 31.2, 28.7, 23.4.; $[\alpha]_D^{25}$ = 33.39 (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS $[M-H]^-$, 399.1387 (observed), 399.1385 (expected).

5.3.2. Synthesis of peptides

5.3.2.1. Boc-D-Ala-D-Ala-OMe (30)



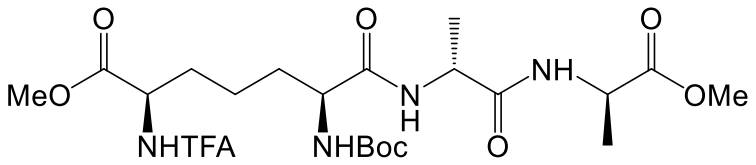
30

H-D-Ala-OMe HCl (0.500 g, 3.58 mmol), Boc-D-Ala-OH (0.678 g, 3.58 mmol) and HATU (1.36 g, 3.58 mmol) were dissolved in dry DMF (20 mL) and cooled to 0 °C. DIPEA (0.625 mL, 10.74 mmol) was added and the reaction mixture was stirred at room

temperature for 18 hours. The solution was concentrated under reduced pressure, and the residue dissolved in EtOAc (20 mL). The organic solution was washed with 0.5 M HCl (10 mL), saturated sodium bicarbonate (10 mL) and brine (10 mL). The organic phase was dried over anhydrous sodium sulfate and concentrated under reduced pressure to yield dipeptide **30** as a white solid (502 mg, 51%)

30: IR ($CHCl_3$ cast, cm^{-1}) 3314, 3080, 2980, 2932, 1747, 1665, 1527, 1456, 1167; ^1H NMR ($CDCl_3$, 700 MHz) δ 6.56 (m, 1H, D-Ala-NH), 4.95 (m, 1H, D-Ala-NH), 4.56 (m, 1H, D-Ala-Hα), 4.15 (m, 1H, D-Ala-Hα), 3.74 (s, 3H, D-Ala-OMe), 1.43 (m, 9H, Boc), 1.39 (d, J = 7 Hz, 3H, D-Ala-Hβ), 1.35 (d, J = 7.7 Hz, 3H, D-Ala-Hβ).; ^{13}C NMR ($CDCl_3$, 150 MHz) δ 173.2, 172.1, 52.6, 48.1, 28.3, 18.3, 18.1; $[\alpha]_D^{25}$= 24.72 (c = 0.550 g/100 mL, CH_2Cl_2); ESI-HRMS $[M+Na]^+$, 297.1420 (observed), 297.1421 (expected).

5.3.2.2. Boc-*meso*-Dap(TFA, OMe)-D-Ala-D-Ala-OMe (31)



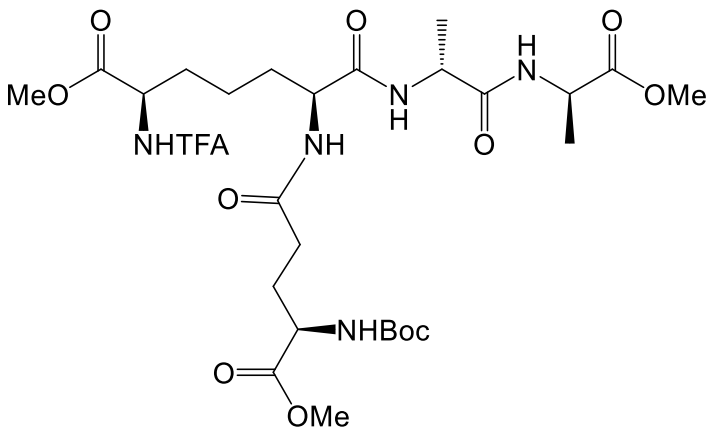
31

To the solution of dipeptide **30** (85 mg, 0.310 mmol) in CH_2Cl_2 (1.5 mL) was added TFA (1.5 mL) at 0 ℃ and the mixture stirred for 3 hours. The solution was concentrated in vacuum, and the resulting oil was co-evaporated with toluene (3 x 1mL). The compound was concentrated under reduced pressure. Protected DAP **29** (120 mg, 0.300 mmol) and HATU (120 mg, 0.316 mmol) were dissolved in dry DMF (5 mL) at 0℃. DIPEA (0.0700

mL, 0.402 mmol) was added to the resulting solution, and this was stirred at 0℃ for 15 minutes. The deprotected dipeptide (0.310 mmol) and DIPEA (0.070 mL, 0.402 mmol) were dissolved in dry DMF (5 mL) in separate flasks, stirred for 5 minutes, and then combined. The combined mixture was stirred for 18 hours at room temperature for the coupling reaction. The mixture was concentrated in vacuum and re-dissolved in EtOAc (5 mL). The solution was washed the with 0.5 M HCl (5 mL), saturated sodium bicarbonate (5 mL) and brine (5 mL). The organic solution was dried with Na_2SO_4 and concentrated in vacuum to obtain tripeptide **31** (110 mg, 64%).

31: IR ($CHCl_3$ cast, cm^{-1}) 3300, 3073, 2979, 2955, 2938, 1745, 1722, 1665, 1529, 1453, 1165; ^1H NMR ($CDCl_3$, 700 MHz) δ 6.70 (d, J = 7.4 Hz, 1H, D-Ala-NH), 6.67 – 6.57 (m, 1H, D-Ala-NH), 5.07 (m, 1H, DAP-γNH), 4.66 – 4.47 (m, 3H, DAP-Hε + D-Ala-NHα + D-Ala-NHα), 4.21 (m, 1H, DAP-Hα), 3.79 (s, 3H, -OMe), 3.76 (s, 3H, -OMe), 2.00-1.91 (m, 1H, DAP-Hδ), 1.93 – 1.71 (m, 2H, DAP-Hδ + DAP-Hβ), 1.73 – 1.53 (m, 2H, DAP-Hβ + DAP-Hγ), 1.42 (m, 16H, DAP-Hβ + Boc + D-Ala-Hβ + D-Ala-Hβ); $[\alpha]_D^{25}$= 33.6 (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS [M+Na]+, 579.2250 (observed), 579.2248 (expected).

5.3.2.3. Boc-D-γ-Glu(OMe)-*meso*-Dap(TFA, OMe)-D-Ala-D-Ala-OMe (32)

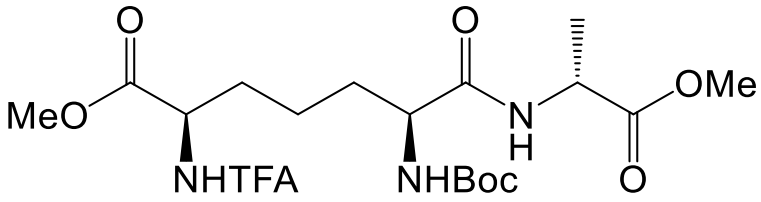


32

To the solution of tripeptide **31** (100 mg, 0.180 mmol) in CH_2Cl_2 (2 mL), TFA (2 mL) was added at 0 ℃ and the mixture was stirred for 3 hours. The solution was concentrated in vacuum, and the resulting oil was co-evaporated with toluene (3 x 1mL). The compound was concentrated under reduced pressure. Boc-D-Glu-α-OMe (57.5 mg, 0.220 mmol) and HATU (84.5 mg, 0.222 mmol) were dissolved in dry DMF (5 mL) at 0 ℃. DIPEA (0.0500 mL, 0.287 mmol) was added to the solution, and the resulting yellow solution at 0 ℃ for 15 minutes. The deprotected tripeptide (0.180 mmol) and DIPEA (0.050 mL, 0.287 mmol) were dissolved in dry DMF (5 mL) in separate a flask, the reaction mixture was stirred for 5 minutes, and then added to the other mixture. The combined mixture was stirred for 18 hours at room temperature for the coupling reaction. The mixture was concentrated in vacuum and re-dissolved in EtOAc (5 mL). The solution was washed with 0.5 M HCl (5 mL), saturated sodium bicarbonate (5 mL) and brine (5 mL). The organic solution was dried with Na_2SO_4 and concentrated in vacuum to get tetrapeptide **32** (114 mg, 91%).

32: IR ($CHCl_3$ cast, cm^{-1}) 3303, 3080, 2980, 2957, 2938, 1747, 1724, 1693, 1657, 1529, 1455, 1165; ^1H NMR ($CDCl_3$, 700 MHz) δ 9.79 (d, J = 7.5 Hz, 1H, DAP-εNH), 8.20 (m, 2H, D-Ala-NH + D-Ala-NH), 7.99 (d, J = 7.6 Hz, 1H, DAP-αNH), 7.22 (d, J = 7.7 Hz, 1H, DGlu-NH), 5.07 (m, 1H, DAP-γNH), 4.66 – 4.47 (m, 3H, DAP-Hε + D-Ala-NHα + D-Ala-NHα), 4.36 – 4.17 (m, 4H, DAP-Hα + DAP-Hε + D-Ala-Hα + D-Ala-Hα), 3.94 (m, 1H, D-Glu-Hα), 3.70 – 3.56 (m, 9H, 3 x OMe), 2.26 – 2.11 (m, 2H, D-Glu-Hγ), 1.94 – 1.67 (m, 4H, D-Glu-Hβ + DAP-Hδ), 1.59 (m, 1H, DAP-Hβ), 1.48 (m, 1H, DAP-Hβ), 1.43 – 1.13 (m, 17H, DAP-Hγ + Boc + D-Ala-Hβ + D-Ala-Hβ); $[\alpha]_D^{25} = 35.0$ (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS $[M+Na]^+$, 722.2834 (observed), 722.2831 (expected).

5.3.2.4. Boc-*meso*-Dap(TFA, OMe)-D-Ala-OMe (33)



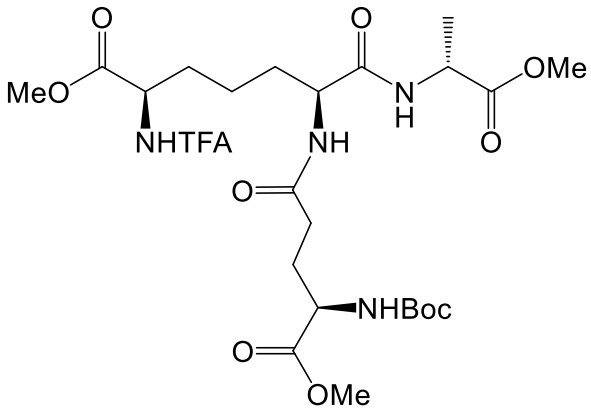
33

H-D-Ala-OMe HCl (62.1 mg, 0.445 mmol), protected DAP **29** (119 mg, 0.298 mmol) and HATU (0.123 g, 0.323 mmol) were dissolved in dry DMF (5 mL) and cooled to 0 °C. DIPEA (0.0850 mL, 0.488 mmol) was added and the reaction mixture was stirred at room temperature for 18 hours. The solution was concentrated under reduced pressure, and the residue was dissolved in EtOAc (20 mL). The organic solution was washed with 0.5 M HCl (10 mL), saturated sodium bicarbonate (10 mL) and brine (10 mL). The organic phase was then dried over anhydrous sodium sulfate and concentrated in vacuum to yield

dipeptide **33** as white solid (93.2mg, 59 %)

33: IR ($CHCl_3$ cast, cm^{-1}) 3308, 3086, 2980, 2957, 1748, 1723, 1665, 1530, 1456, 1368, 1164; ^1H NMR (CD_3OD, 700 MHz) 4.45 – 4.37 (m, 2H, DAP-Hε + D-Ala-NHα), 4.10 (m, 1H, DAP-Hα), 3.74 (s, 3H, -OMe), 3.71 (s, 3H, -OMe), 2.01 - 1.92 (m, 1H, DAP-Hδ), 1.80 – 1.74 (m, 2H, DAP-Hδ + DAP-Hβ), 1.63 – 1.60 (m, 2H, DAP-Hβ + DAP-Hγ), 1.42 (m, 14H, DAP-Hβ + Boc + DAla-Hβ); $[\alpha]_D^{25}$= = 7.40 (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS [M+Na]+, 508.1874 (observed), 508.1877 (expected).

5.3.2.5. Boc-D-γ-Glu(OMe)-*meso*-Dap(TFA, OMe)-D-Ala-OMe (34)



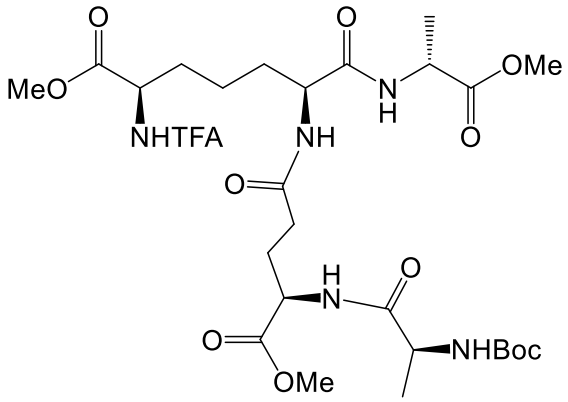
34

To the solution of dipeptide **33** (50.1 mg, 0.103 mmol) in CH_2Cl_2 (1.0 mL), TFA (1.0 mL) was added at 0 ℃ and the mixture was stirred for 3 hours. The solution was concentrated in vacuum, and the resulting oil was co-evaporated with toluene (3 x 1 mL). The compound was dried under high vacuum for 30 minutes. Boc-D-Glu-α-OMe (35.2 mg, 0.135 mmol) and HATU (52.2 mg, 0.137 mmol) were dissolved in dry DMF (4 mL) at 0℃. DIPEA (0.50 mL, 0.287 mmol) was added to the resulting solution, and the mixture was

stirred at 0 ℃ for 15 minutes. The deprotected dipeptide (0.103 mmol) and DIPEA (0.030 mL, 0.172 mmol) were dissolved in dry DMF (5 mL) in separate flasks, stirred for 5 minutes, and the two solutions were combined. The mixture was stirred for 18 hours at room temperature for the coupling reaction. The combined mixture was concentrated in vacuum and re-dissolved in EtOAc (5 mL). The solution was washed the with 0.5 M HCl (5 mL), saturated sodium bicarbonate (5 mL) and brine (5 mL). The organic solution was dried with Na_2SO_4 and concentrated in vacuum to get tripeptide **34** (56.0 mg, 89%).

34: IR ($CHCl_3$ cast, cm^{-1}) 3324, 3084, 2958, 2939, 1747, 1724, 1664, 1528, 1456, 1368, 1163; ^1H NMR (CD_3OD, 700 MHz) δ 4.45 – 4.40 (m, 2H, DAP-Hε + D-Ala-NHα), 4.33 (m, 1H, DAP-Hα), 4.11 – 4.09 (m, 1H, D-Glu-Hα), 3.76 – 3.70 (m, 9H, 3 x OMe), 2.37 – 2.16 (m, 2H, D-Glu-Hγ), 1.94 – 1.67 (m, 4H, D-Glu-Hβ + DAP-Hδ), 1.59 (m, 1H, DAP-Hβ), 1.48 (m, 1H, DAP-Hβ), 1.43 – 1.13 (m, 14H, DAP-Hγ + Boc + D-Ala-Hβ).; $[\alpha]_D^{25}$= = 2.20 (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS $[M+Na]^+$, 651.2457 (observed), 651.2460 (expected).

5.3.2.6. Boc-L-Ala-D-γ-Glu-*meso*-Dap(TFA, OMe)-D-Ala-OMe (35)

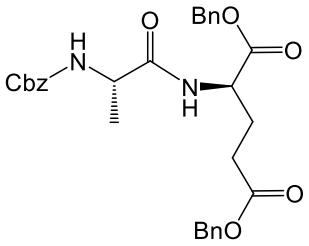


35

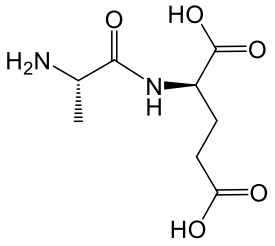
To a solution of tripeptide **34** (40.0 mg, 0.064 mmol) in CH_2Cl_2 (0.6 mL) was added TFA (0.6 mL) at 0 ℃ and the mixture was stirred for 3 hours. The solution was concentrated in vacuum, and the resulting oil was co-evaporated with toluene (3 x 1 mL). The compound was dried under high vacuum for 30 minutes. During this time, Boc-L-Ala-OH (15.0 mg, 0.0793 mmol), PyBOP (41.0 mg, 0.787mmol) and HOBt (12.0 mg, 0.0784 mmol) were dissolved in dry DMF (2 mL) at 0 ℃. DIPEA (0.030 mL, 0.172 mmol) was added, and the resulting yellow solution was stirred at 0 ℃ for 15 minutes. The deprotected tripeptide (0.064 mmol) and DIPEA (0.030 mL, 0.172 mmol) were dissolved in dry DMF (3 mL), stirred for 5 minutes, and the two solutions were combined. The combined mixture was stirred for 18 hours at room temperature. The mixture was concentrated in vacuum and re-dissolved in EtOAc (5 mL). The solution was washed the with 0.5 M HCl (5 mL), saturated sodium bicarbonate (5 mL) and brine (5 mL). The organic solution was dried with Na_2SO_4 and concentrated in vacuum to obtain tetrapeptide **35** (33.3 mg, 74%).

35: IR ($CHCl_3$ cast, cm^{-1}) 3301, 3070, 2979, 2956, 1745, 1722, 1663, 1531, 1454, 1368, 1164; ^1H NMR ($CDCl_3$, 700 MHz) δ 8.04 (d, J = 7.5 Hz, 1H, DAP-εNH), 7.56 – 7.54 (m, 2H, D-Ala-NH + L-Ala-NH), 7.42 – 7.40 (m, 1H, DAP-αNH), 6.90 – 6.83 (m, 1H, D-Glu-NH), 5.01 – 4.98 (m, 1H, DAP-γNH), 4.68 – 4.52 (m, 3H, DAP-Hε + L-Ala-NHα + D-Ala-NHα), 4.45 – 4.11 (m, 4H, DAP-Hα + DAP-Hε + L-Ala-Hα + D-Ala-Hα), 4.10 (m, 1H, D-Glu-Hα), 3.80 – 3.71 (m, 9H, 3 x OMe), 2.26 – 2.03 (m, 2H, D-Glu-Hγ), 1.93 – 1.78 (m, 4H, D-Glu-2Hβ + DAP-3Hδ), 1.48 – 1.20 (m, 19H, DAP3Hβ + DAP3Hγ + Boc + D-Ala-4Hβ + D-Ala5Hβ).; $[\alpha]_D^{25}$= -0.20 (c = 0.100 g/100 mL, CH_3OH); ESI-HRMS $[M+Na]^+$, 722.2830 (observed), 722.2831 (expected).

5.3.2.7. L-Ala-D-Glu-OH (37)



36



37

Cbz-L-Alanine (89.3 mg, 0.400 mmol), D-glutamic acid dibenzyl ester 4-toluenesulfonate salt (200 mg, 0.400 mmol), and HATU (160 mg, 0.420 mmol) were dissolved in dry DMF (5 mL) and the mixture was cooled to 0 °C. DIPEA (0.150 mL, 0.861 mmol) was added and the reaction stirred at room temperature for 18 hours. The solution was concentrated under reduced pressure, and the residue was dissolved in EtOAc (20 mL). The organic solution washed with 0.5 M HCl (10 mL), saturated sodium bicarbonate (10 mL) and brine (10 mL). The organic phase was dried over anhydrous sodium sulfate and concentrated by rotary evaporator to yield dipeptide **36** as a white solid (204 mg, 96%). This was used without further purification. Palladium (10% on carbon, 10 mg) was dissolved in dry MeOH (10 mL) and activated under hydrogen atmosphere for 10 minutes. Dipeptide **36** (204 mg, 0.389 mmol) was added to the solution, and the mixture was stirred under hydrogen atmosphere for 90 minutes. The suspension was filtered through Celite, and washed with methanol. The filtrate was concentrated in vacuum to obtain the deprotected dipeptide **37** (57.8 mg, 68%)

37: IR ($CHCl_3$ cast, cm^{-1}) 3450, 3330, 3283, 2902, 1687, 1628, 1538, 1453; ^1H NMR ($CDCl_3$, 700 MHz) δ 4.33-4.51 (m, 1H, CHCOOH), 4.09 (q, 1H, CH_3CH), 2.33 (t, 2H,

CH_2CH_2); 2.05 (q, 2H, CH-CH_2CH_2); 1.28 (d, 3H, $CHCH_3$); $[\alpha]_D^{25}$= +32.4 (c = 0.100 g/100 mL, CH_3OH), ESI-HRMS $[M+Na]^+$, 241.1970(observed), 241.1973 (expected)/

References

[1] L. Morozov, Mirror symmetry breaking in biochemical evolution, *Orig. Life Evol. Biospheres*, 1979, 9, 187-217

[2] S. Toxvaerd, Origin of homochirality in biological systems, *Int. J. Astrobiol*., 2005, 4, 43–48

[3] S.Y. Hong, J. E. Oh, and K. H. Lee. Effect of D-amino acid substitution on the stability, the secondary structure, and the activity of membrane-active peptide. *Biochem. Pharmacol*., 1999, 58, 1775-1780.

[4] K. Hamamoto, Y. Kida, Y. Zhang, T. Shimizu, and K. Kuwano. Antimicrobial activity and stability to proteolysis of small linear cationic peptides with D-amino acid substitutions. *Microbiol. Immunol*., 2002, 46, 741-749.

[5] G. Carmona, A. Rodriguez, D. Juarez, G. Corzo, and E. Villegas. Improved protease stability of the antimicrobial peptide Pin2 substituted with D-amino acids. *Protein J.,* 2013, 32, 456–466

[6] D. Wade, A. Boman, B. Wåhlin, C. M. Drain, D. Andreu, H. G. Boman, and R. B. Merrifield. All-D amino acid-containing channel-forming antibiotic peptides. *Proc. Natl. Acad. Sci. USA.,* 1990, 87, 4761-4765

[7] J. van Heijenoort. Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat. Prod. Rep.,* 2001, 18, 503-519.

[8] K. Soda, T. Yoshimura, and N. Esaki. Stereospecificity for the hydrogen transfer of pyridoxal enzyme reactions. *Chem. Rec.,* 2001, 1, 373-384.

[9] M. Caparrós, A. G. Pisabarro, and M. A. de Pedro. Effect of D-amino acids on structure and synthesis of peptidoglycan in *Escherichia coli*. *J. Bacteriol.,* 1992, 174, 5549–5559.

[10] F. Cava, H. Lam, M. A. de Pedro, and M. K. Waldor. Emerging knowledge of regulatory roles of D-amino acids in bacteria. *Cell Mol. Life Sci.,* 2011, 68, 817–831.

[11] I. Kolodkin-Gal, D. Romero, S. Cao, J. Clardy, R. Kolter, and R. Losick, D-amino acids trigger biofilm disassembly. *Science*, 2010, 328, 627-629.

[12] G. Fey, G. W. Gould and A. D. Hitchins. Identification of D-alanine as the auto-inhibitor of germination of *Bacillus globigii* spores. *J. Gen. Microbiol*., 1964, 35, 229-236.

[13] M. A. Hubbard, P. Thorkildson, W. H. Welch, and T. R. Kozel. Stereo-selective binding of monoclonal antibodies to the poly-gamma-D-glutamic acid capsular antigen of *Bacillus anthracis*. *Mol. Immunol.,* 2013, 55, 337-344.

[14] G. H. Fisher, A. D'Aniello, A. Vetere, L. Padula, G. P. Cusano, and E. H. Man, E.H. Free D-aspartate and D-alanine in normal and Alzheimer brain. *Brain Res. Bull.,* 1991, 26, 983-985.

[15] D. Dufour, V. Leung, and C. M. Lévesque. Bacterial biofilm: structure, function, and antimicrobial resistance. *Endodontic Topics*, 2010, 22, 2-16.

[16] C. Reichhardt, J. Y. Lim, D. Rice, J. N. Fong, and L. Cegelski. Structure and function of bacterial biofilms by solid-state NMR. *Biophys. J.,* 2014, 106, 192a.

[17] S.A. Fuchs, R. Berger, L. W. Klomp, and T. J. De Koning. D-amino acids in the central nervous system in health and disease. *Mol. Genetics Metab.,* 2005, 85, 168-180.

[18] Y. Kaji, T. Oshika, and Y. Takazawa. Accumulation of D-beta-aspartic acid containing proteins in age-related ocular diseases. *Chem. Biodivers.,* 2010, 7, 1364-1370.

[19] J. P. Richard, T. L. Amyes, B. Goryanova, and X. Zhai. Enzyme architecture: on the importance of being in a protein cage. *Curr. Opin. Chem. Biol.,* 2014, 21, 1–10.

[20] J. C. Vederas, and H. G. Floss. Stereochemistry of pyridoxal phosphate catalyzed enzyme reactions *Acc. Chem. Res.,* 1980, 13, 455–463.

[21] H. Abe, N. Yoshikawa, M. G. Sarower, and S. Okada. Physiological function and metabolism of free D-alanine in aquatic animals. *Biol. Pharm. Bull.,* 2005, 28, 1571-1577.

[22] K. Hoffmann, E. Schneider-Scherzer, H. Kleinkauf, and R. Zocher. Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis. *J. Biol. Chem.,* 1994, 269, 12710-12714.

[23] T. Uo, T. Yoshimura, N. Tanaka, K. Takegawa, and N. Esaki. Functional characterization of aanine racemase from *Schizosaccharomyces pombe*: a eucaryotic counterpart to Bacterial alanine racemase. *J. Bacteriol.,* 2001, 183, 2226-2233.

[24] K. Ono, K. Yanagida, T. Oikawa, T. Ogawa, and K. Soda. Alanine racemase of alfalfa seedlings (*Medicago sativa L.*): first evidence for the presence of an amino acid racemase in plants. *Phytochemistry,* 2006, 67, 856-860.

[25] T. Uo, T. Yoshimura, S. Shimizu, and N. Esaki. Occurrence of pyridoxal 5'- phosphate-dependent serine racemase in silkworm, *Bombyx mori*. *Biochem. Biophys. Res. Commun.,* 1998, 246, 31-34.

[26] H. Wolosker, K. N. Sheth, M. Takahashi, J. P. Mothet, R. O. Brady Jr., C. D. Ferris, and S. H. Snyder. Purification of serine racemase: Biosynthesis of the neuromodulator D-serine. *Proc. Natl. Acad. Sci. U.S.A.,* 1999, 96, 721–725.

[27] Y. Liu, R. H. Hill, P. Arhem, and G. von Euler. NMDA and glycine regulate the affinity of the Mg2+-block site in NR1-1a/NR2A NMDA receptor channels expressed in *Xenopus* oocytes. *Life Sciences,* 2001, 68, 1817–1826.

[28] T. Takarada, E. Hinoi, Y. Takahata, and Y. Yoneda. Serine racemase suppresses chondrogenic differentiation in cartilage in a Sox9-dependent manner. *J. Cell. Physiol.,* 2008, 215, 320–328.

[29] M. C. Ma, H. S. Huang, Y. S. Chen, and S. H. Lee. Mechanosensitive N-methyl-D-aspartate receptors contribute to sensory activation in the rat renal pelvis. *Hypertension.*, 2008, 52, 938–944.

[30] M. A. Smith, V. Mack, A. Ebneth, I. Moraes, B. Felicetti, M. Wood, D. Schonfeld, O. Mather, A. Cesura, and J. Barker. The structure of mammalian serine racemase-evidence

for conformational changes upon inhibitor binding. *J. Biol. Chem.,* 2010, 285(17), 12873–12881.

[31] J. De Miranda, R. Panizzutti, V. N. Foltyn, and H. Wolosker. Cofactors of serine racemase that physiologically stimulate the synthesis of the N-methyl-D-aspartate (NMDA) receptor coagonist D-serine. *Proc. Natl. Acad. Sci. U.S.A.*, 2002, 99, 14542-14527.

[32] C.M. Huang, C. A. Elmets, D. C. Tang, F. Li, and N. Yusuf. Proteomics reveals that proteins expressed during the early stage of Bacillus anthracis infection are potential targets for the development of vaccines and drugs. *Genomics Proteomics Bioinf*., 2004, 2, 143-151.

[33] P. M. Kim, H. Aizawa, P. S. Kim, A. S. Huang, S. R. Wickramasinghe, A. H. Kashani, R. K. Barrow, R. L. Huganir, A. Ghosh, and S. H. Snyder. Serine racemase: Activation by glutamate neurotransmission via glutamate receptor interacting protein and mediation of neuronal migration. *Proc. Natl. Acad. Sci. U.S.A.,* 2005, 102, 2105-2110.

[34] K. Fujii, K. Maeda, T. Hikida, A. K. Mustafa, R. Balkissoon, J. Xia, T. Yamada, Y. Ozeki, R. Kawahara, M. Okawa, R. L. Huganir, H. Ujike, S. H. Snyder, and A. Sawa. Serine racemase binds to PICK1: potential relevance to schizophrenia. *Mol. Psychiatry*, 2006, 11, 150-157.

[35] E. Dumin, I. Bendikov, V. N. Foltyn, Y. Misumi, Y. Ikehara, E. Kartvelishvily, and H. J. Wolosker. Modulation of D-serine levels via ubiquitin-dependent proteasomal degradation of serine racemase. *J. Biol. Chem.,* 2006, 281, 20291-20302.

[36] A. K. Mustafa, M. Kumar, B. Selvakumar, G. P. Ho, J. T. Ehmsen, R. K. Barrow, L. M. Amzel, and S. H. Snyder. Nitric oxide S-nitrosylates serine racemase, mediating feedback inhibition of D-serine formation. *Proc. Natl. Acad. Sci. U.S.A.,* 2007, 104, 2950-2955.

[37] C. Vargas-Lopes, C. Madeira, S. A. Kahn, I. Albino do Couto, P. Bado, J. C. Houzel, J. De Miranda, M. S. de Freitas, S. T. Ferreira, and R. J. Panizzutti. Protein kinase C activity regulates D-serine availability in the brain. *Neurochem*., 2011, 116, 281-290.

[38] E. Awoonor-Williams and C. N. Rowley. Evaluation of methods for the calculation of the pKa of cysteine residues in proteins. *J. Chem. Theory Comput*., 2016, 12, 4662–4673

[39] B. Pillai, M. M. Cherney, C. M. Diaper, A. Sutherland, J. S. Blanchard, J. C. Vederas, M. N. James. Structural insights into stereochemical inversion by diaminopimelate epimerase: An antibacterial drug target. *Proc. Natl. Acad. Sci U.S.A.*, 2006, 103, 8668-8673.

[40] S. Y. Choi, N. Esaki, T. Yoshimura, K. Soda. Bacterial glutamate racemase has high sequence similarity with myoglobins and forms an equimolar inactive complex with hemin. *Protein Expr. Purif.* 1991, 2, 90–93.

[41] L. Liu, K. Iwata, A. Kita, Y. Kawarabayasi, M. Yohda, K. Miki. Crystal structure of aspartate racemase from *Pyrococcus horikoshii* OT3 and its implications for molecular mechanism of PLP-independent racemization. *J. Mol. Biol*. 2002, 319, 479-489.

[42] A. Buschiazzo, M. Goytia, F. Schaeffer, W. Degrave, W. Shepard, C. Grégoire, N. Chamond, A. Cosson, A. Berneman, N. Coatnoan, P. M. Alzari, P. Minoprio. Crystal

structure, catalytic mechanism, and mitogenic properties of *Trypanosoma cruzi* proline racemase. *Proc. Nat. Acad. Sci. U.S.A*. 2006, 103, 1705-1710.

[43] Y. Liu, R. H. White, and W. B. Whitman. *Methanococci* use the diaminopimelate aminotransferase (DapL) pathway for lysine biosynthesis. *J. Bacteriol*., 2010, 192, 3304–3310.

[44] M. Cirilli, R. Zheng, G. Scapin, J. S. Blanchard. Structural Symmetry: The three-dimensional structure of *Haemophilus influenzae* diaminopimelate epimerase. *Biochemistry*, 1998, 37, 16452-16458.

[45] A. J. Lloyd, T. Huyton, J. P. Turkenburg, D.I. Roper. Refinement of *Haemophilus influenzae* diaminopimelic acid epimerase (DapF) at 1.75 A resolution suggests a mechanism for stereocontrol during catalysis. *Acta Cryst. D Biol. Crystallogr*., 2004, 60, 397-400.

[46] H.-Y. Sagong, K. J. Kim. Structural basis for redox sensitivity in *Corynebacterium glutamicum* diaminopimelate epimerase: an enzyme involved in L-lysine biosynthesis. *Sci. Rep*., 2017, 7, 42318-42318.

[47] B. Pillai, M. Cherney, C. M. Diaper, A. Sutherland, J. S. Blanchard, J. C. Vederas, M. N. G. James. Dynamics of catalysis revealed from the crystal structures of mutants of diaminopimelate epimerase. *Biochem. Biophys. Res. Commun*., 2007, 363, 547-553.

[48] B. Pillai, V. A. Moorthie, M. J. van Belkum, S. L. Marcus, M. M. Cherney, C. M. Diaper, J. C. Vederas, M. N. G. James. Crystal structure of diaminopimelate epimerase from

Arabidopsis thaliana, an amino acid racemase critical for L-lysine biosynthesis. *J. Mol. Biol*., 2009, 385, 580-594.

[49] D. G. Isom, C. A. Castañeda, B. R. Cannon, and B. García-Moreno. Large shifts in *p*Ka values of lysine residues buried inside a protein. *Proc. Natl. Acad. Sci. U.S.A.,* 2011, 108(13), 5260-5265.

[50] J. -M. Girodeau, C. Agouridas, M. Masson, R. Pineau, and F. Le Goffict, F. The lysine pathway as a target for a new genera of synthetic antibacterial antibiotics. *J. Med. Chem.* 1986, 29, 1023-1030.

[51] P. Conti, L. Tamborini, A. Pinto, A.Blondel, P. Minoprio, A. Mozzarelli, and C. De Michel. Drug discovery targeting amino acid racemases. *Chem. Rev.* 2011, 111, 6919–6946.

[52] L.K. P. Lam, L. D. Arnold, T.H. Kalantar, J. G. Kelland, P.M. Lane-Bell, M.M. Palcic, M. A. Pickard, J. C. Vederas. Analogs of diaminopimelic acid as inhibitors of *meso*-diaminopimelate dehydrogenase and LL-diaminopimelate epimerase. *J. Biol. Chem*., 1988, 263, 11814-11819.

[53] M.H. Gelb, Y. Lin, M. A. Pickard, Y. Song, J. C. Vederas. Synthesis of 3-fluorodiaminopimelic acid isomers as inhibitors of diaminopimelate epimerase: stereocontrolled enzymatic elimination of hydrogen fluoride. *J. Am. Chem. Soc.,* 1990, 112, 4932-4942.

[54] F. Gerhart, W. Higgins, C. Tardif, J. B. Ducep. 2-(4-Amino-4-carboxybutyl)aziridine-2-carboxylic acid. A potent irreversible inhibitor of diaminopimelic acid epimerase.

Spontaneous formation from alpha-(halomethyl)diaminopimelic acids. *J. Med. Chem*., 1990, 33, 2157-2162.

[55] W. Higgins, C. Tardif, C. Richaud, M. Krivanek, A. Cardin. Expression of recombinant diaminopimelate epimerase in *Escherichia coli*. *Eur. J. Biochem*., 1989, 186, 137-143.

[56] C. M. Diaper, A. Sutherland, B. Pillai, M. N. G. James, P. Semchuk, J. S. Blanchard, J. C. Vederas. The stereoselective synthesis of aziridine analogues of diaminopimelic acid (DAP) and their interaction with DAP epimerase. *Org. Biomol. Chem.,* 2005, 3, 4402-4411.

[57] W. F. Diven. Studies on amino acid racemases II. Purification and properties of the glutamate racemase from *Lactobacillus fermenti*. *Biochim*. *Biophys. Acta*, 1969, 191, 702–706

[58] P. Doublet, J. vanHeijenoort, J. P. Bohin, D. Mengin-Lecreulx. The murI gene of *Escherichia coli* is an essential gene that encodes a glutamate racemase activity. *J. Bacteriol*. 1993, 175, 2970–2979.

[59] K. A. Gallo, M. E. Tanner, J. R. Knowles. Mechanism of the reaction catalyzed by glutamate racemase. *Biochemistry*, 1993, 32, 3991–3997.

[60] S. Glavas, and M. E. Tanner. Catalytic acid/base residues of glutamate racemase. *Biochemistry,* 1999, 38, 4106–4113.

[61] T. Lundqvist, S. L. Fisher, G. Kern, R. H. A. Folmer, Y. Xue, D. T. Newton, T. A. Keating, R. A. Alm, B. L. M. de Jonge. Exploitation of structural and regulatory diversity in glutamate racemases. *Nature,* 2007, 447, 817-822.

[62] P. Doublet. J. van Heijenoort, D. Mengin-Lecreulx. The glutamate racemase activity from *Escherichia coli* is regulated by peptidoglycan precursor UDP-N-acetylmuramoyl-L-alanine. *Biochemistry,* 1994, 33, 5285–5290.

[63] H-T. Ho, P. J. Falk, K. M. Ervin, B. S. Krishnan, L. F. Discotto, T. J. Dougherty, M. J. Pucci. UDP-N-acetylmuramyl-L-alanine functions as an activator in the regulation of the *Escherichia coli* glutamate racemase activity. *Biochemistry,* 1995, 34, 2464–2470.

[64] P. Doublet, J. van Heijenoort, D. Mengin-Lecreulx. the glutamate racemase of *Escherichia coli* investigated by site-directed mutagenesis. *Microb. Drug Resist. Mechanisms Epidemiol. Dis.*, 1996, 2, 43–49.

[65] M. May, S. Mehboob, D. C. Mulhearn, Z. Wang, H. Yu, G. R. J. Thatcher, B. D. Santarsiero, M. E. Johnson, and A. D. Mesecar. Structural and functional analysis of two glutamate racemase isozymes from *Bacillus anthracis* and implications for inhibitor design. *J. Mol. Biol*., 2007, 371, 1219-1237

[66] K. Y. Hwang, C-S. Cho, S. S. Kim, H-C. Sung, Y. G. Yu, and Y. Cho. Structure and mechanism of glutamate racemase from *Aquifex pyrophilus*. *Nature*, 1999, 6, 422-426.

[67] S. Glavas, and M. E. Tanner. Active Site residues of glutamate racemase. *Biochemistry,* 2001, 40*,* 6199-6204.

[68] M. Cirilli, R. Zheng, G. Scapin, and J. S. Blanchard. Structural symmetry: the three-dimensional structure of *Haemophilus influenzae* diaminopimelate epimerase. *Biochemistry,* 1998, 37, 16452-16458.

[69] M. Ashiuchi, T. Yoshimura, N. Esaki, H. Ueno, and K. Soda. Inactivation of glutamate racemase of *Pediococcus pentosaceus* with L-serine O-sulfate. *Biosci. Biotech. Biochem*. 1993, 57, 1978-1979.

[70] M. E. Tanner, and S. Miao. The synthesis and stability of aziridino-glutamate, an irreversible inhibitor of glutamate racemase. *Tetrahedron Lett*. 1994, 35, 4073.

[71] A. de Dios, L. Prieto, J. A. Martín, A. Rubio, J. Ezquerra, M. Tebbe, B. Lopez de Uralde, J. Martín, A. Sanchez, D. L. LeTourneau, J. E. McGee, C. Boylan, T. R. Parr, and M. C. Smith. 4-Substituted D-glutamic acid analogues: the first potent inhibitors of glutamate racemase (MurI) enzyme with antibacterial activity. *J. Med. Chem*. 2002, 45, 4559-4570.

[72] K-H. Kim, Y. J. Bong, J. K. Park, K. J. Shin, K. Y. Hwang, and E. E. Kim. Structural basis for glutamate racemase inhibition. *J. Mol. Biol.* 2007, 372, 434–443

[73] E.H. Man, M.E. Sandhouse, J. Burg, G.H. Fisher. Accumulation of D-aspartic acid with age in the human brain. *Science*, 1983, 220, 1407–1408.

[74] N. Fujii, T. Saito. Homochirality and life. *Chem. Rec.* 2004, 4, 267-278.

[75] H. Okada, M. Yohda, Y. Giga-Haam, Y. Ueno, S. Ohdo, H. Kumagai. Distribution and purification of aspartate racemase in lactic acid bacteria. *Biochem. Biophys*. Acta 1991, 1078, 377.

[76] A. Ohtaki, Y. Nakano, E. Iizuka, T. Arakawa, K. Yamada, M. Yodha. Structure of aspartate racemase complexed with a dual substrate analogue, citric acid, and implications for the reaction mechanism. *Proteins*, 2008, 70, 1167-1174.

[77] T. Yoshida, T. Seko, O. Okada, L. Liu, K. Miki, M. Yohda. Roles of conserved basic amino acid residues and activation mechanism of the hyperthermophilic aspartate racemase at high temperature. *Proteins*, 2006, 64, 502-512.

[78] T. Aihara, T. Ito, Y. Yamanaka, K. Noguchi, M. Odaka, M. Sekine, H. Homma, M. Yohda. Structural and functional characterization of aspartate racemase from the acidothermophilic archaeon *Picrophilus torridus*. *Extremophiles*, 2016, 20, 385-393.

[79] M. Ashiuchi, K. Tani, K. Soda, H. Risono. Properties of glutamate racemase from *Bacillus subtilis* IFO 3336 producing poly- 7-glutamate. *J. Biochem*. 1998, 123, 1156.

[80] J.-W. Ahn, J. H. Chang, K.-J Kim. Structural basis for an atypical active site of an L-aspartate/glutamate-specific racemase from *Escherichia coli*. *FEBS Lett.* 2015, 589, 3842-3847.

[81] X. Liu, F. Gao, Y. Ma, S. Liu, Y. Cui, Z. Yuan, X. Kang. Crystal structure and molecular mechanism of an aspartate/glutamate racemase from *Escherichia coli* O157. *FEBS Lett*. 2016, 590, 1262-1269.

[82] S. König, H. Marco, and G. Gäde. D-Proline: Comment to "An overview on D-amino acids". *Amino Acids*, 2018, 50, 359–361

[83] B. Reina-San-Martín, W. Degrave, C. Rougeot, A. Cosson, N. Chamond, A. Cordeiro-Da-Silva, M. Arala-Chaves, A. Coutinho, P. Minoprio. A B-cell mitogen from a pathogenic trypanosome is a eukaryotic proline racemase. *Nat. Med*. 2000, 6, 890-897.

[84] N. Chamond, A. Cosson, N. Coatnoan, P. Minoprio. *Trypanosoma vivax* Infections: pushing ahead with mouse models for the study of Nagana. I. Parasitological, hematological and pathological parameters. *Mol. Biochem. Parasitol*. 2009,165,170–179.

[85] G. J. Cardinale, R. H. Abeles. Purification and mechanism of action of proline racemase. *Biochemistry*, 1968, 7, 3970-3978.

[86] W.-Y. Fang, E. Dahiya, H.-L. Qin, R. Mourya, S. Maharaj, Natural proline-rich cyclopolypeptides from marine organisms: chemistry, synthetic methodologies and biological status. *Mar. Drugs*, 2016, 14, 194.

[87] E. A. Becker, P. M. Seitzer, A. Tritt, D. Larsen, M. Krusor, A. I. Yao, D. Wu, D. Madern, J. A. Eisen, A. E. Darling, M. T. Facciotti. Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response. *PLoS Genet.* 2014, 10, e1004784.

[88] G. J. Cardinale, and R. H. Abeles. Purification and mechanism of action of proline racemase. *Biochemistry*, 1968, 7, 3970–3978.

[89] G. Rudnick and R. H. Abeles. Reaction mechanism and structure of the active site of proline racemase. *Biochemistry,*1975, 14, 4515–4522.

[90] A. Rubinstein, D. T. Major. Catalyzing racemizations in the absence of a cofactor: The reaction mechanism in proline racemase. *J. Am. Chem. Soc.* 2009, 131, 8513-8521.

[91] C. A. F. de Oliveira, B. J. Grant, M. Zhou, J. A. McCammon. Large-scale conformational changes of *Trypanosoma cruzi* proline racemase predicted by accelerated molecular dynamics simulation. *PLoS Comput. Biol*. 2011, 7, e1002178.

[92] L. V. Kirchhoff, L. M. Weiss, M. Wittner, and H. B. Tanowitz. Parasitic diseases of the heart. Parasitic diseases of the heart. *Front. Biosci.,* 2004, 9, 706-723

[93] J. R. Coura, and J. C. Pinto Dias. Epidemiology, control and surveillance of Chagas disease: 100 years after its discovery. *Mem. Inst. Oswaldo Cruz*, 2009, 104, 31-40

[94] L. Coutinho, M. Alves Ferreira, A. Cosson, M. Meuser Batista, D. da Gama Jaen Batista, P. Minoprio, W. M. Degrave, A. Berneman, And M. de Nazare Correia Soeiro. Mem. Inst. Oswaldo Cruz, Rio de Janeiro. Inhibition of *Trypanosoma cruzi* proline racemase affects host-parasite interactions and the outcome of in vitro infection. *Mem. Inst. Oswaldo Cruz,* 2009, 104, 1055-1062

[95] A. Berneman, M. Alves-Ferreira, N. Coatnoan, N. Chamond, and P. J. Minoprio. Medium/high throughput D-amino acid oxidase colorimetric method for determination of D-amino acids. Application for amino acid racemases. *Microbial. Biochem. Technol*., 2010, 2, 139-146

[96] A. G. Marr, and P. W. Wilson. The alanine racemase of *Brucella abortus*. *Arch. Biochem. Biophys.,* 1954, 49, 424-433

[97] D. A. Harris, M. Ruger, M. A. Reagan, F. J. Wolf, R. L. Peck, H. Wallick, and H. B. Woodruff. Discovery, development, and antimicrobial properties of D-4-amino-3 isoxazolidone (oxamycin), a new antibiotic produced by *Streptomyces garyphalus*. *Antibiot. Chemother.,* 1955, 5, 83–190.

[98] L. M. Humma. Prevention and treatment of drug-resistant tuberculosis. *Am. J. Health Syst. Pharm.,* 1996, 53, 2291–2298.

[99] M. Nitsche, W. Jaussi, D. Liebetanz, N. Lang, F. Tergau, and W. Paulus. Consolidation of human motor cortical neuroplasticity by D-cycloserine. *Neuropsychopharmacology*, 2004, 29, 1573–1578.

[100] P. Storici, D. De Biase, F. Bossa, S. Bruno, A. Mozzarelli, C. Peneff, R. B. Silverman, and T. Schirmer. Structures of gamma-aminobutyric acid (GABA) aminotransferase, a pyridoxal 5'-phosphate, and [2Fe-2S] cluster-containing enzyme, complexed with gamma-ethynyl-GABA and with the antiepilepsy drug vigabatrin. J Biol Chem., 2004, 279, 363-373.

[101] A. R. Frydenberg and S. M. Graham. Toxicity of first-line drugs for treatment of tuberculosis in children: review. *Med. Int. Health*, 2009, 14, 1329-1337.

[102] A. Pollack. Drug goes from $13.50 a tablet to $750, overnight. *The New York Times.* 2015, September 21st, Page B1

[103] T. D. Fenn, G. F. Stamper, A. A. Morollo, and D. Ringe. A side reaction of alanine racemase: transamination of cycloserine. *Biochemistry,* 2003, 2, 5775-5783

[104] O. A. Asojo, S. K. Nelson, S. Mootien, Y. Lee, W. C. Rezende, D. A. Hyman, M. M. Matsumoto, S. Reiling, A. Kelleher, M. Ledizet, R. A. Koski, and K. G. Anthony. Structural and biochemical analyses of alanine racemase from the multidrug-resistant *Clostridium difficile* strain 630. *Acta Crystallogr. Sect.D.,* 2014, 70, 1922-1933

[105] M. P. Lambert and F. C. Neuhaus. Mechanism of D-cycloserine action: alanine racemase from *Escherichia coli. W. J. Bacteriol.,* 1972, 110, 978–987

[106] M. A. Azam and U. Jayaram. Inhibitors of alanine racemase enzyme: a review. *J Enzyme Inhib. Med. Chem.*, 2016, 31, 517–526.

[107] G. D. Wright and C. T. Walsh. D-Alanyl-D-alanine ligases and the molecular mechanism of mancomycin resistance. *Acc. Chem. Res.,* 1992, 25, 468-473.

[108] S. Batson, C. de Chiara, V. Majce, A. J. Lloyd, S. Gobec, D. Rea, V. Fülöp, C. W. Thoroughgood, K. J. Simmons, C. G. Dowson, C. W. G. Fishwick, L. P. S. de Carvalho, and D. I. Roper. Inhibition of D-Ala:D-Ala ligase through a phosphorylated form of the antibiotic D-cycloserine. *Nat. Commun.,* 2017, 8, 1939-1946.

[109] G. A. Prosser and L. P. S. de Carvalho. Reinterpreting the mechanism of inhibition of *Mycobacterium tuberculosis* D-alanine:D-alanine ligase by D-cycloserine. *Biochemistry*, 2013, 52, 7145−7149.

[110] T. Kumagai, Y. Koyama, K. Oda, M. Noda, Y. Matoba and M. Sugiyama. Molecular cloning and heterologous expression of a biosynthetic gene cluster for the antitubercular agent D-cycloserine produced by *Streptomyces lavendulae*. *Antimicrob. Agents Chemother*., 2010, 54, 1132-1139.

[111] N. Uda, Y. Matoba, T. Kumagai, K. Oda, M. Noda, and M. Sugiyama. Establishment of an in vitro D-cycloserine-synthesizing system by using O-ureido-L-serine synthase and D-cycloserine synthetase found in the biosynthetic pathway. *Antimicrob. Agents Chemother.,* 2013, 57, 2603–2612.

[112] T. Kumagai, T. Ozawa, M. Tanimoto, M. Noda, Y. Matoba, and M. Sugiyama. High-Level heterologous production of D-cycloserine by *Escherichia coli*. *Appl. Environ. Microbiol.,* 2015, 81(22), 7881-7887.

[113] W. M. Rabeh and P. F. Cook. Structure and mechanism of O-acetylserine sulfhydrylase. *J. Biol. Chem.,* 2004, 279, 26803–26806.

[114] D. Dietrich, M. J. van Belkum and J. C. Vederas. Characterization of DcsC, a PLP-independent racemase involved in the biosynthesis of D-cycloserine. *Org.Biomol. Chem.,* 2012, 10, 2248–2254.

[115] T. Panavas, C. Sanders, and T. R. Butt. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. *Methods Mol. Biol.,* 2009, 497, 303-317.

[116] C-D. Lee, H-C. Sun, S-M. Hu, C-F. Chiu, A. Homhuan, S-M. Liang, C-H. Leng, and T-F. Wang. An improved SUMO fusion protein system for effective production of native proteins. *Protein Sci.,* 2008, 17, 1241–1248.

[117] M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks, and T. R. Butt. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics*, 2004, 5, 75–86.

[118] D. R. Smyth, M. K. Mrozkiewicz, W. J. Mcgrath, P. Listwan, and B. Kobe. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.,* 2003, 12, 1313–1322.

[119] J. Ren, X. Gao, C. Jin, M. Zhu, X. Wang, A. Shaw, L. Wen, X. Yao, and Y. Xue. Systematic study of protein SUMOylation: Development of a site-specific predictor of SUMOsp 2.0. *Proteomics.,* 2009, 9, 3409–3412.

[120] Y-C. Ahn, C. Fischer, M. J. van Belkum, and J. C. Vederas. PLP-independent racemization: mechanistic and mutational studies of O-ureidoserine racemase (DcsC). *Org. Biomol. Chem*., 2018, 16, 1126–1133.

[121] J. Hartner and U. M. Reinscheid. Conformational analysis of menthol diastereomers by NMR and DFT computation. *J. Mol. Struct*., 2008, 872, 145–149.

[122] O. Mitsunobu, and M. Yamada. Preparation of esters of carboxylic and phosphoric acid via quaternary phosphonium salts. *Bull. Chem. Soc. Jpn.*, 1967, 40, 2380-2382.

[123] O. Mitsunobu. The use of diethyl azodicarboxylate and triphenylphosphine in synthesis and transformation of natural products. *Synthesis,* 1981, 1, 1–28.

[124] K. C. K. Swamy, N. N. B. Kumar, E. Balaraman, and K. V. P. P. Kumar. Mitsunobu and related reactions: advances and applications. *Chem. Rev*., 2009, 109, 2551–2651

[125] D. Potin, F.Dumas, and J. d'Angelo. New chiral auxiliaries: their use in the asymmetric hydrogenation of ß-acetamidocrotonates. *J. Am. Chem. Soc*., 1990, 112, 3483-3486.

[126] S. W. Garry and D. G. Neilson. The preparation and absolute configuration of the optically active forms of the diastereoisomers of 2-(1-phenylethyl)amino-1-phenylethanol and their use as chiral auxiliaries in the asymmetric reduction of acetophenone with modified lithium aluminium hydrides. *J. Chem. Soc., Perkin Trans. 1*, 1987, 601-605.

[127] F. Yongxian, X. Zhangming, Z. Huawei, and Q. Junqing. Kinetic resolution of both 1-phenylethanol enantiomers produced by hydrolysis of 1-phenylethyl acetate with *Candida antarctica* lipase B in different solvent systems. *Kinet. Catal*., 2011, 52, 686–690

[128] L. Hor, R. C. J. Dobson, M. T. Downton, J. Wagner, C. A. Hutton and M. A. Perguini. Dimerization of bacterial diaminopimelate epimerase is essential for catalysis. *J. Biol. Chem.,* 2013, 288, 9238–9248.

[129] D. Popham, and P. Setlow. Cloning, nucleotide sequence, and regulation of *Bacillus subtili*s pbpE operon, which codes for penicillin binding protein 4 and an apparent amino acid racemase. *J. Bacteriol*., 1993, 175, 2917-2925.

[130] T. Miyamoto, M, Katane, Y. Saitoh, M. Sekine, and H. Homma. Identification and characterization of novel broad-spectrum amino acid racemases from *Escherichia coli* and *Bacillus subtilis*. *Amino Acid*s, 2017, 49, 1885–1894.

[131] E. Sauvage, F. Kerff, M. Terrak, J. A. Ayala, and P. Charlier. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.,* 2008, 32, 234–258.

[132] M. M. Palomino, C. Sanchez-Rivas, and S. M. Ruzal. High salt stress in *Bacillus subtilis*: involvement of PBP4 as a peptidoglycan hydrolase. *Res Microbiol.*, 2009, 160, 117-124

[133] V. Barbe, S. Cruveiller, F. Kunst, P. Lenoble, G. Meurice, A. Sekowska, D. Vallenet, T. Wang, I. Moszer, C. Médigue, and A. Danchin. From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. *Microbiology,* 2009, 155, 1758-1775.

[134] S. A. Cochrane, B. Findlay, A. Bakhtiary, J. Z. Acedo, E. M. Rodriguez-Lopez, P. Mercier, and J. C. Vederas. Antimicrobial lipopeptide tridecaptin A1 selectively binds to Gram-negative lipid II. *Proc. Natl. Acad. Sci. U.S.A.,* 2016, 113, 11561–11566

[135] J. D. Bacha, and J. K. Kochi. Alkenes from acids by oxidative decarboxylation. *Tetrahedron*, 1968, 24, 2215-2226.

Appendix 1: X-ray Crystal Structure of (1*R*)-1-phenylethyl (2*S*)-2-(ureidooxymethyl)oxirane-2-carboxylate
STRUCTURE REPORT

XCL Code: JCV1705 **Date:** 21 September 2017

Compound: (1*R*)-1-phenylethyl (2*S*)-2-(ureidooxymethyl)oxirane-2-carboxylate

Formula: $C_{13}H_{16}N_2O_5$

Supervisor: J. C. Vederas **Crystallographer:** M. J. Ferguson

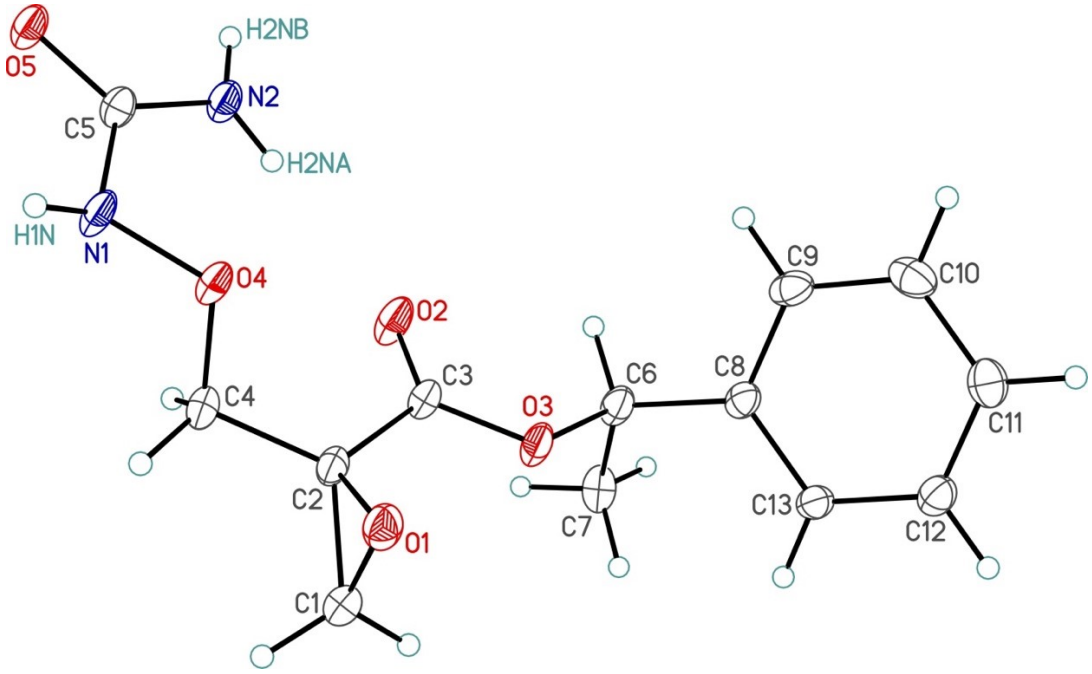
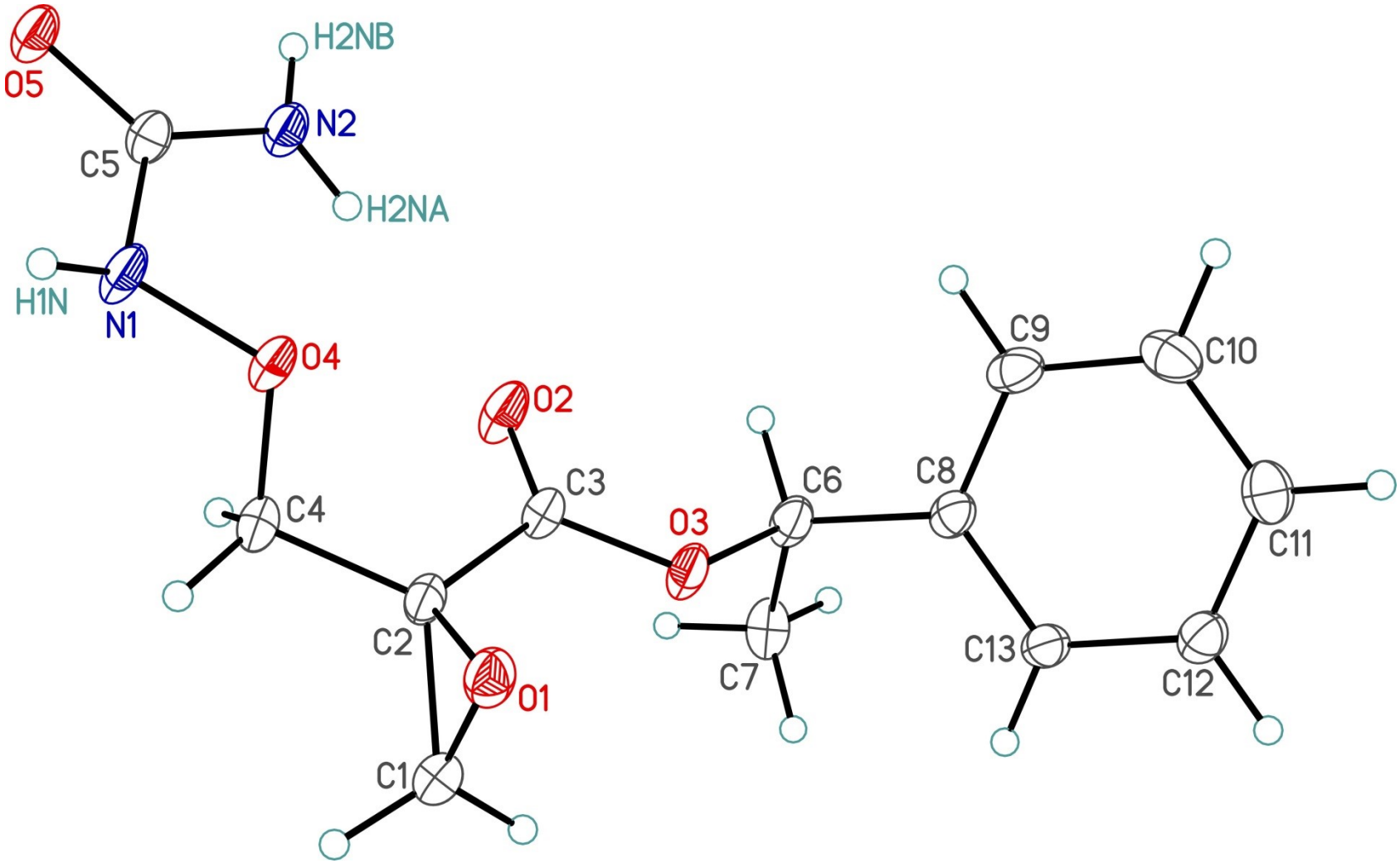


Figure Legends

Figure 1. Perspective view of the (1*R*)-1-phenylethyl (2*S*)-2-(ureidooxymethyl)oxirane-2-carboxylate molecule showing the atom labelling scheme. Non-hydrogen atoms are represented by Gaussian ellipsoids at the 30% probability level. Hydrogen atoms are shown with arbitrarily small thermal parameters.


O5
H2NB
N2
C5
H2NA
H1N
N1
O4
O2
C3
C4
O3
C6
C8
C9
C10
C11
C2
O1
C7
C13
C12
C1

List of Tables

Table 1. Crystallographic Experimental Details

A. Crystal Data	
formula	$C_{13}H_{16}N_2O_5$
formula weight	280.28
crystal dimensions (mm)	0.25 × 0.13 × 0.03
crystal system	monoclinic
space group	$P2_1$ (No. 4)
unit cell parameters[a]	
a (Å)	5.92380 (10)
b (Å)	7.17870 (10)
c (Å)	16.1585 (3)
β (deg)	90.8223 (8)
V (Å^3)	687.07 (2)
Z	2
ρ_{calcd} (g cm^{-3})	1.355
μ (mm^{-1})	0.886
B. Data Collection and Refinement Conditions	
diffractometer	Bruker D8/APEX II CCD[b]
radiation (λ [Å])	Cu Kα (1.54178) (microfocus source)
temperature (°C)	–100
scan type	ω and ϕ scans (1.0°) (5 s exposures)
data collection 2θ limit (deg)	147.90
total data collected	4927 ($-7 \le h \le 7$, $-8 \le k \le 8$, $-20 \le l \le 20$)
independent reflections	2754 (R_{int} = 0.0174)

number of observed reflections (*NO*)	2674 [$F_o^2 \geq 2\sigma(F_o^2)$]
structure solution method	intrinsic phasing (*SHELXT-2014*[c])
refinement method	full-matrix least-squares on F^2 (*SHELXL–2014*[d])
absorption correction method	Gaussian integration (face-indexed)
range of transmission factors	1.0000–0.7597
data/restraints/parameters	2754 / 0 / 194
extinction coefficient (x)[e]	0.0037(11)
Flack absolute structure parameter[f]	0.08(7)
goodness-of-fit (S)[g] [all data]	1.025
final R indices[h]	
R_1 [$F_o^2 \geq 2\sigma(F_o^2)$]	0.0268
wR_2 [all data]	0.0729
largest difference peak and hole	0.171 and –0.137 e Å^{-3}

[a]Obtained from least-squares refinement of 9857 reflections with $5.46° < 2\theta < 147.08°$.

[b]Programs for diffractometer operation, data collection, data reduction and absorption correction were those supplied by Bruker.

[c]Sheldrick, G. M. *Acta Crystallogr.* **2015**, *A71*, 3–8. (*SHELXT-2014*)

[d]Sheldrick, G. M. *Acta Crystallogr.* **2015**, *C71*, 3–8. (*SHELXL-2014*)

[e]$F_c^* = kF_c[1 + x\{0.001F_c^2\lambda^3/\sin(2\theta)\}]^{-1/4}$ where k is the overall scale factor.

[f]Flack, H. D. *Acta Crystallogr.* **1983**, *A39*, 876–881; Flack, H. D.; Bernardinelli, G. *Acta Crystallogr.* **1999**, *A55*, 908–915; Flack, H. D.; Bernardinelli, G. *J. Appl. Cryst.* **2000**, *33*, 1143–1148. The Flack parameter will refine to a value near zero if the structure is in the correct configuration and will refine to a value near one for the inverted configuration. The low anomalous scattering power of the atoms in this structure (none heavier than oxygen) implies that the data cannot be used for absolute structure assignment, thus the Flack parameter is provided for informational purposes only. The stereochemistry was assigned on the basis of the established stereochemistry of the precursor compound.

[g]$S = [\Sigma w(F_o^2 - F_c^2)^2/(n-p)]^{1/2}$ (n = number of data; p = number of parameters varied; $w = [\sigma^2(F_o^2)$

$+ (0.0429P)^2 + 0.0614P]^{-1}$ where $P = [\text{Max}(F_o^2, 0) + 2F_c^2]/3)$.

$^hR_1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$; $wR_2 = [\Sigma w(F_o^2 - F_c^2)^2/\Sigma w(F_o^4)]^{1/2}$.

Table 2. Atomic Coordinates and Equivalent Isotropic Displacement Parameters

Atom	*x*	*y*	*z*	U_{eq}, Å^2
O1	0.3141(2)	0.8250(2)	0.26192(8)	0.0351(3)*
O2	0.3513(3)	0.3626(2)	0.17921(8)	0.0424(4)*
O3	0.2068(2)	0.46650(18)	0.29919(8)	0.0349(3)*
O4	0.6520(2)	0.68546(19)	0.13333(7)	0.0313(3)*
O5	1.0047(2)	0.51275(19)	-0.01896(8)	0.0348(3)*
N1	0.7617(3)	0.6803(2)	0.05578(10)	0.0362(4)*
H1N	0.846(4)	0.787(4)	0.0449(16)	0.041(6)
N2	0.7902(3)	0.3625(2)	0.07595(10)	0.0335(3)*
H2NA	0.681(5)	0.372(4)	0.1073(16)	0.046(7)
H2NB	0.856(5)	0.257(5)	0.0607(18)	0.053(8)
C1	0.1033(3)	0.8089(3)	0.21824(11)	0.0364(4)*
C2	0.2958(3)	0.6887(3)	0.19721(9)	0.0284(4)*
C3	0.2906(3)	0.4870(3)	0.22380(10)	0.0290(4)*
C4	0.4188(3)	0.7291(3)	0.11879(11)	0.0328(4)*
C5	0.8607(3)	0.5141(3)	0.03668(10)	0.0286(3)*
C6	0.1499(3)	0.2736(3)	0.32456(11)	0.0330(4)*
C7	-0.0908(4)	0.2391(3)	0.29683(12)	0.0424(5)*
C8	0.1946(3)	0.2618(2)	0.41664(11)	0.0285(4)*
C9	0.3976(3)	0.1862(3)	0.44551(13)	0.0362(4)*
C10	0.4466(3)	0.1777(3)	0.52942(14)	0.0404(4)*
C11	0.2906(4)	0.2411(3)	0.58607(12)	0.0376(4)*
C12	0.0861(3)	0.3121(3)	0.55814(11)	0.0338(4)*
C13	0.0386(3)	0.3239(2)	0.47390(11)	0.0293(4)*

Anisotropically-refined atoms are marked with an asterisk (*). The form of the anisotropic displacement parameter is: $\exp[-2\pi^2(h^2a^{*2}U_{11} + k^2b^{*2}U_{22} + l^2c^{*2}U_{33} + 2klb^*c^*U_{23} + 2hla^*c^*U_{13} + 2hka^*b^*U_{12})]$.

Table 3. Selected Interatomic Distances (Å)

Atom1	Atom2	Distance	Atom1	Atom2	Distance
O1	C1	1.430(2)	C2	C3	1.511(3)
O1	C2	1.435(2)	C2	C4	1.499(2)
O2	C3	1.206(2)	C6	C7	1.509(3)
O3	C3	1.330(2)	C6	C8	1.510(2)
O3	C6	1.484(2)	C8	C9	1.393(3)
O4	N1	1.4205(17)	C8	C13	1.391(2)
O4	C4	1.432(2)	C9	C10	1.384(3)
O5	C5	1.248(2)	C10	C11	1.387(3)
N1	C5	1.367(3)	C11	C12	1.384(3)
N2	C5	1.330(2)	C12	C13	1.389(2)
C1	C2	1.474(2)			

Table 4. Selected Interatomic Angles (deg)

Atom1	Atom2	Atom3	Angle	Atom1	Atom2	Atom3	Angle
C1	O1	C2	61.92(11)	O5	C5	N1	117.98(17)
C3	O3	C6	116.45(13)	O5	C5	N2	124.06(17)
N1	O4	C4	108.27(13)	N1	C5	N2	117.89(15)
O4	N1	C5	115.18(15)	O3	C6	C7	106.78(16)
O1	C1	C2	59.19(11)	O3	C6	C8	106.69(14)
O1	C2	C1	58.89(12)	C7	C6	C8	115.79(15)
O1	C2	C3	116.61(13)	C6	C8	C9	119.35(16)
O1	C2	C4	116.83(16)	C6	C8	C13	121.92(16)
C1	C2	C3	118.45(15)	C9	C8	C13	118.73(17)
C1	C2	C4	117.95(16)	C8	C9	C10	120.86(17)
C3	C2	C4	115.99(15)	C9	C10	C11	120.00(18)
O2	C3	O3	125.72(17)	C10	C11	C12	119.60(18)
O2	C3	C2	122.17(15)	C11	C12	C13	120.43(17)
O3	C3	C2	112.09(14)	C8	C13	C12	120.34(17)
O4	C4	C2	107.35(14)				

Table 5. Hydrogen-Bonded Interactions

D–H···A	D–H (Å)	H···A (Å)	D···A (Å)	∠D–H···A (deg)
N1–H1N···O5[a]	0.94(3)	1.89(3)	2.826(2)	175(2)
N2–H2NA···O2	0.83(3)	2.29(3)	3.109(2)	170(3)
N2–H2NB···O5[b]	0.89(3)	2.06(3)	2.943(2)	176(3)

[a]At +2, *y*+0.5, .

[b]At +2, *y*–0.5, .

Table 6. Torsional Angles (deg)

Atom1	Atom2	Atom3	Atom4	Angle	Atom1	Atom2	Atom3	Atom4	Angle
C1	O1	C2	C3	-108.66(17)	C4	C2	C3	O3	171.15(15)
C1	O1	C2	C4	107.95(18)	O1	C2	C4	O4	78.67(19)
C6	O3	C3	O2	-10.4(3)	C1	C2	C4	O4	145.90(16)
C6	O3	C3	C2	167.79(15)	C3	C2	C4	O4	-64.9(2)
C3	O3	C6	C7	-88.72(18)	O3	C6	C8	C9	-98.04(19)
C3	O3	C6	C8	146.90(15)	O3	C6	C8	C13	82.1(2)
C4	O4	N1	C5	-120.52(17)	C7	C6	C8	C9	143.31(19)
N1	O4	C4	C2	166.48(15)	C7	C6	C8	C13	-36.5(3)
O4	N1	C5	O5	-162.48(15)	C6	C8	C9	C10	178.29(18)
O4	N1	C5	N2	20.5(2)	C13	C8	C9	C10	-1.9(3)
O1	C1	C2	C3	105.55(17)	C6	C8	C13	C12	-179.54(17)
O1	C1	C2	C4	-106.06(18)	C9	C8	C13	C12	0.6(3)
O1	C2	C3	O2	-154.31(18)	C8	C9	C10	C11	1.5(3)
O1	C2	C3	O3	27.4(2)	C9	C10	C11	C12	0.2(3)
C1	C2	C3	O2	138.38(19)	C10	C11	C12	C13	-1.4(3)
C1	C2	C3	O3	-39.9(2)	C11	C12	C13	C8	1.0(3)
C4	C2	C3	O2	-10.6(3)					

Table 7. Anisotropic Displacement Parameters (U_{ij}, Å^2)

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	0.0453(7)	0.0339(7)	0.0262(6)	-0.0023(5)	0.0047(5)	0.0006(6)
O2	0.0583(9)	0.0352(7)	0.0344(6)	-0.0050(6)	0.0226(6)	-0.0054(6)
O3	0.0507(8)	0.0278(7)	0.0265(6)	0.0018(5)	0.0174(5)	-0.0005(6)
O4	0.0336(6)	0.0364(7)	0.0243(6)	-0.0021(5)	0.0119(5)	-0.0023(5)
O5	0.0417(7)	0.0310(6)	0.0321(6)	-0.0008(5)	0.0176(5)	-0.0011(6)
N1	0.0459(9)	0.0302(8)	0.0331(8)	-0.0009(7)	0.0239(7)	-0.0030(7)
N2	0.0421(9)	0.0296(8)	0.0293(7)	0.0008(6)	0.0146(6)	0.0010(7)
C1	0.0386(10)	0.0418(10)	0.0290(8)	0.0013(8)	0.0083(7)	0.0063(8)
C2	0.0318(8)	0.0325(9)	0.0211(7)	0.0012(7)	0.0063(6)	0.0000(7)
C3	0.0305(8)	0.0333(10)	0.0235(7)	-0.0001(7)	0.0078(6)	-0.0029(7)
C4	0.0366(9)	0.0384(10)	0.0237(7)	0.0045(7)	0.0086(7)	-0.0005(7)
C5	0.0325(8)	0.0308(8)	0.0228(7)	-0.0032(7)	0.0050(6)	-0.0026(7)
C6	0.0437(10)	0.0268(9)	0.0290(8)	0.0015(6)	0.0129(7)	-0.0004(7)
C7	0.0514(12)	0.0466(12)	0.0293(8)	0.0046(8)	0.0025(8)	-0.0063(9)
C8	0.0306(8)	0.0242(8)	0.0309(8)	0.0027(6)	0.0084(7)	-0.0019(6)
C9	0.0301(8)	0.0311(9)	0.0478(10)	0.0061(8)	0.0147(7)	0.0021(7)
C10	0.0298(8)	0.0360(10)	0.0554(11)	0.0147(10)	-0.0008(8)	0.0003(8)
C11	0.0442(10)	0.0337(10)	0.0347(9)	0.0047(7)	-0.0021(8)	-0.0070(8)
C12	0.0388(9)	0.0317(9)	0.0312(8)	-0.0023(7)	0.0060(7)	0.0011(7)
C13	0.0292(8)	0.0274(8)	0.0314(8)	-0.0001(7)	0.0055(7)	0.0033(7)

The form of the anisotropic displacement parameter is:

$$\exp[-2\pi^2(h^2a^{*2}U_{11} + k^2b^{*2}U_{22} + l^2c^{*2}U_{33} + 2klb^*c^*U_{23} + 2hla^*c^*U_{13} + 2hka^*b^*U_{12})]$$

Table 8. Derived Atomic Coordinates and Displacement Parameters for Hydrogen Atoms

Atom	x	y	z	U_{eq}, Å^2
H1A	0.0612	0.9094	0.1790	0.044
H1B	-0.0248	0.7514	0.2475	0.044
H4A	0.3571	0.6522	0.0729	0.039
H4B	0.4018	0.8621	0.1037	0.039

H6	0.2514	0.1843	0.2956	0.040
H7A	-0.1355	0.1126	0.3124	0.051
H7B	-0.1904	0.3293	0.3234	0.051
H7C	-0.1029	0.2529	0.2366	0.051
H9	0.5037	0.1398	0.4071	0.043
H10	0.5871	0.1285	0.5482	0.048
H11	0.3238	0.2357	0.6437	0.045
H12	-0.0224	0.3530	0.5968	0.041
H13	-0.1012	0.3747	0.4553	0.035

Appendix 2: X-ray Crystal Structure of (1*S*)-1-phenylethyl (2*R*)-2-(ureidooxymethyl)oxirane-2-carboxylate
STRUCTURE REPORT

XCL Code: JCV1701 **Date:** 15 March 2017

Compound: (1*S*)-1-phenylethyl (2*R*)-2-(ureidooxymethyl)oxirane-2-carboxylate

Formula: $C_{13}H_{16}N_2O_5$

Supervisor: J. C. Vederas **Crystallographer:** M. J. Ferguson

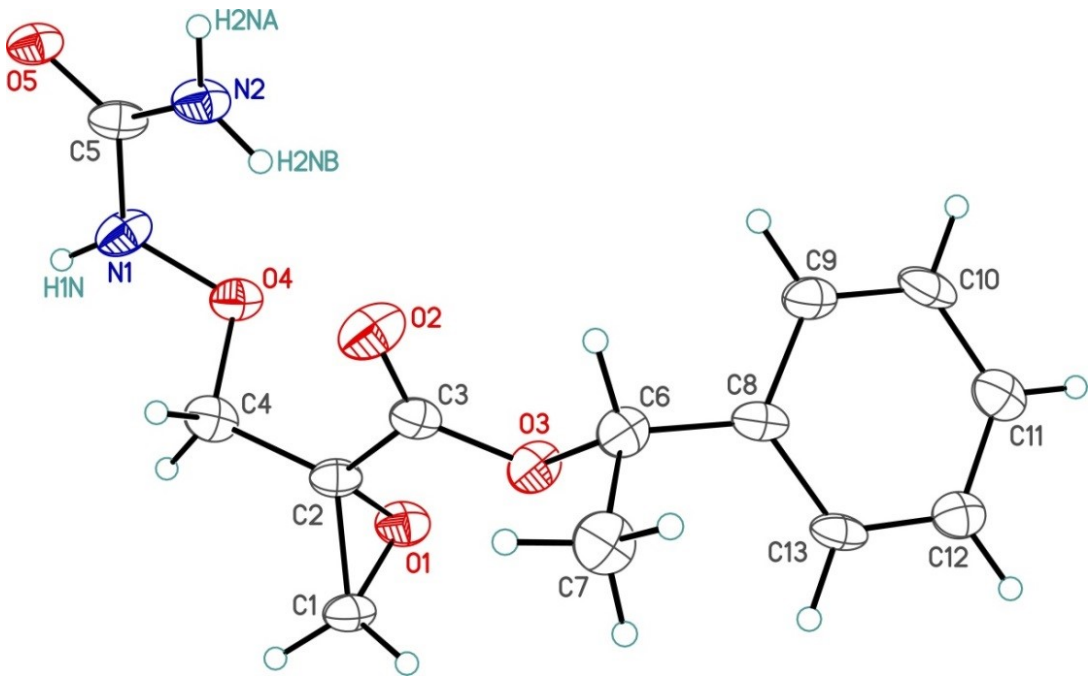
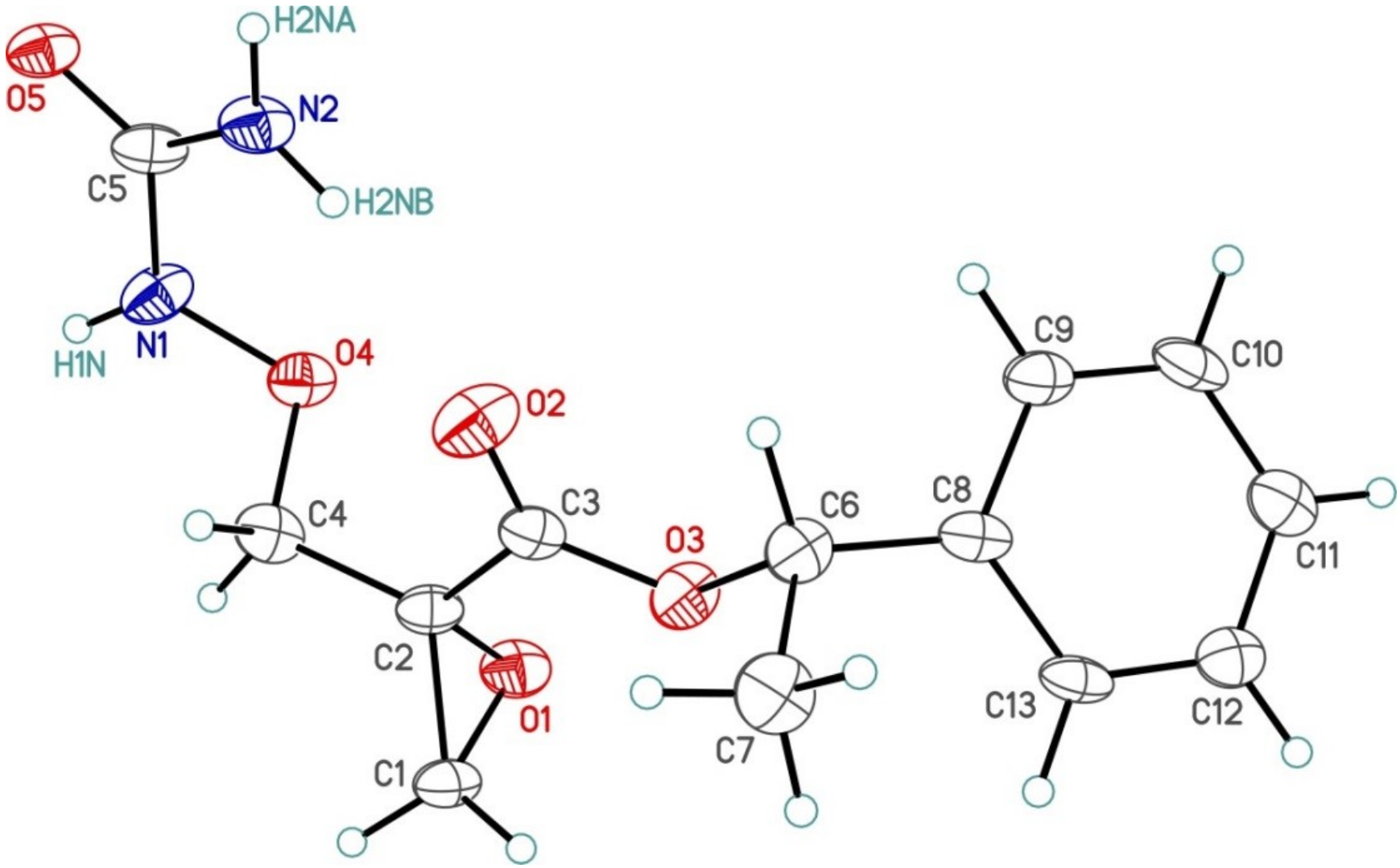


Figure Legends

Figure 1. Perspective view of the (1*S*)-1-phenylethyl (2*R*)-2-(ureidooxymethyl)oxirane-2-carboxylate molecule showing the atom labelling scheme. Non-hydrogen atoms are represented by Gaussian ellipsoids at the 30% probability level. Hydrogen atoms are shown with arbitrarily small thermal parameters.


H2NA
O5
N2
C5
H2NB
H1N
N1
O4
O2
C4
C3
O3
C6
C8
C9
C10
C11
C2
O1
C7
C13
C12
C1

List of Tables

Table 1. Crystallographic Experimental Details

A. Crystal Data

formula	$C_{13}H_{16}N_2O_5$
formula weight	280.28
crystal dimensions (mm)	0.25 × 0.25 × 0.05
crystal system	monoclinic
space group	$P2_1$ (No. 4)
unit cell parameters[a]	
a (Å)	5.9509 (10)
b (Å)	7.1789(10)
c (Å)	16.121 (3)
β (deg)	90.821 (8)
V (Å^3)	688.6 (2)
Z	2
ρ_{calcd} (g cm^{-3})	1.352
μ (mm^{-1})	0.884

B. Data Collection and Refinement Conditions

diffractometer	Bruker D8/APEX II CCD[b]
radiation (λ [Å])	Cu Kα (1.54178) (microfocus source)
temperature (°C)	–100
scan type	ω and ϕ scans (1.0°) (15 s exposures)
data collection 2θ limit (deg)	135.31
total data collected	3111 (-7 ≤ h ≤ 7, -8 ≤ k ≤ 8, -19 ≤ l ≤ 18)
independent reflections	2076 (R_{int} = 0.0492)

number of observed reflections (*NO*)	1582 [$F_o^2 \geq 2\sigma(F_o^2)$]
structure solution method	intrinsic phasing (*SHELXT-2014*[c])
refinement method	full-matrix least-squares on F^2 (*SHELXL–2014*[d])
absorption correction method	Gaussian integration (face-indexed)
range of transmission factors	1.0000–0.5229
data/restraints/parameters	2076 / 0 / 182
Flack absolute structure parameter[f]	1.0(5)
goodness-of-fit (*S*)[g] [all data]	1.096
final *R* indices[h]	
R_1 [$F_o^2 \geq 2\sigma(F_o^2)$]	0.0830
wR_2 [all data]	0.2560
largest difference peak and hole	0.364 and –0.425 e Å^{-3}

[a]Obtained from least-squares refinement of 3325 reflections with $5.48° < 2\theta < 138.66°$.

[b]Programs for diffractometer operation, data collection, data reduction and absorption correction were those supplied by Bruker.

[c]Sheldrick, G. M. *Acta Crystallogr.* **2015**, *A71*, 3–8. (*SHELXT-2014*)

[d]Sheldrick, G. M. *Acta Crystallogr.* **2015**, *C71*, 3–8. (*SHELXL-2014*)

[e]$F_c^* = kF_c[1 + x\{0.001F_c^2\lambda^3/\sin(2\theta)\}]^{-1/4}$ where *k* is the overall scale factor.

[f]Flack, H. D. *Acta Crystallogr.* **1983**, *A39*, 876–881; Flack, H. D.; Bernardinelli, G. *Acta Crystallogr.* **1999**, *A55*, 908–915; Flack, H. D.; Bernardinelli, G. *J. Appl. Cryst.* **2000**, *33*, 1143–1148. The Flack parameter will refine to a value near zero if the structure is in the correct configuration and will refine to a value near one for the inverted configuration. The low anomalous scattering power of the atoms in this structure (none heavier than oxygen) implies that the data cannot be used for absolute structure assignment, thus the Flack parameter is provided for informational purposes only. The stereochemistry was assigned on the basis of the established stereochemistry of the precursor compound.

[g]$S = [\Sigma w(F_o^2 - F_c^2)^2/(n-p)]^{1/2}$ (*n* = number of data; *p* = number of parameters varied; $w = [\sigma^2(F_o^2) + (0.0429P)^2 + 0.0614P]^{-1}$ where $P = [\mathrm{Max}(F_o^2, 0) + 2F_c^2]/3$).

[h]$R_1 = \Sigma||F_o| - |F_c||/\Sigma|F_o|$; $wR_2 = [\Sigma w(F_o^2 - F_c^2)^2/\Sigma w(F_o^4)]^{1/2}$.

Table 2. Atomic Coordinates and Equivalent Isotropic Displacement Parameters

Atom	*x*	*y*	*z*	U_{eq}, Å^2
O1	0.1885(9)	-0.3259(8)	0.2379(4)	0.0484(15)*
O2	0.1429(11)	0.1344(9)	0.3203(4)	0.0575(18)*
O3	0.2967(10)	0.0322(7)	0.2020(4)	0.0483(15)*
O4	-0.1522(9)	-0.1901(8)	0.3665(4)	0.0438(14)*
O5	-0.5045(9)	-0.0176(8)	0.5189(4)	0.0433(14)*
N1	-0.2619(12)	-0.1849(10)	0.4439(5)	0.0490(19)*
N2	-0.2906(12)	0.1322(9)	0.4247(5)	0.0470(18)*
C1	0.3950(12)	-0.3082(11)	0.2821(5)	0.0419(19)*
C2	0.2019(11)	-0.1909(11)	0.3029(5)	0.0377(17)*
C3	0.2103(11)	0.0113(10)	0.2763(5)	0.0372(17)*
C4	0.0799(13)	-0.2293(11)	0.3817(6)	0.0448(19)*
C5	-0.3628(11)	-0.0186(12)	0.4636(6)	0.0416(18)*
C6	0.3517(14)	0.2248(11)	0.1765(6)	0.0420(19)*
C7	0.5908(15)	0.2615(14)	0.2044(7)	0.055(2)*
C8	0.3062(11)	0.2361(10)	0.0846(6)	0.0364(18)*
C9	0.1039(12)	0.3119(11)	0.0546(6)	0.0442(19)*
C10	0.0568(13)	0.3237(12)	-0.0291(6)	0.050(2)*
C11	0.2104(14)	0.2606(12)	-0.0849(6)	0.049(2)*
C12	0.4121(14)	0.1890(11)	-0.0580(6)	0.0438(19)*
C13	0.4622(11)	0.1757(10)	0.0269(5)	0.0319(18)*

Anisotropically-refined atoms are marked with an asterisk (*). The form of the anisotropic displacement parameter is: $\exp[-2\pi^2(h^2a^{*2}U_{11} + k^2b^{*2}U_{22} + l^2c^{*2}U_{33} + 2klb^*c^*U_{23} + 2hla^*c^*U_{13} + 2hka^*b^*U_{12})]$.

Table 3. Selected Interatomic Distances (Å)

Atom1	Atom2	Distance	Atom1	Atom2	Distance
O1	C1	1.417(10)	C2	C3	1.515(11)
O1	C2	1.429(10)	C2	C4	1.497(11)
O2	C3	1.207(10)	C6	C7	1.510(12)
O3	C3	1.318(10)	C6	C8	1.504(12)
O3	C6	1.481(9)	C8	C9	1.401(11)
O4	N1	1.416(9)	C8	C13	1.393(10)
O4	C4	1.427(10)	C9	C10	1.376(13)
O5	C5	1.236(9)	C10	C11	1.370(13)
N1	C5	1.375(11)	C11	C12	1.370(12)
N2	C5	1.326(11)	C12	C13	1.400(12)
C1	C2	1.467(10)			

Table 4. Selected Interatomic Angles (deg)

Atom1	Atom2	Atom3	Angle	Atom1	Atom2	Atom3	Angle
C1	O1	C2	62.1(5)	O5	C5	N1	118.4(8)
C3	O3	C6	116.7(6)	O5	C5	N2	124.4(8)
N1	O4	C4	108.2(6)	N1	C5	N2	117.0(7)
O4	N1	C5	115.8(7)	O3	C6	C7	106.9(7)
O1	C1	C2	59.4(5)	O3	C6	C8	106.7(6)
O1	C2	C1	58.6(5)	C7	C6	C8	116.1(7)
O1	C2	C3	116.4(7)	C6	C8	C9	120.1(7)
O1	C2	C4	118.3(7)	C6	C8	C13	121.9(7)
C1	C2	C3	117.1(6)	C9	C8	C13	118.0(8)
C1	C2	C4	118.6(7)	C8	C9	C10	121.6(8)
C3	C2	C4	115.8(7)	C9	C10	C11	119.7(8)
O2	C3	O3	126.1(7)	C10	C11	C12	120.4(9)
O2	C3	C2	121.5(7)	C11	C12	C13	120.6(8)
O3	C3	C2	112.4(7)	C8	C13	C12	119.8(7)
O4	C4	C2	107.4(7)				

Table 5. Hydrogen-Bonded Interactions

D–H···A	D–H (Å)	H···A (Å)	D···A (Å)	∠D–H···A (deg)
N1–H1N···O5[a]	0.88	2.18	2.832(9)	130.7
N2–H2NA···O2	0.88	2.07	2.944(9)	172.0
N2–H2NB···O5[b]	0.88	2.25	3.096(9)	161.0

[a]At –1, *y*–0.5, +1

[b]At –1, *y*+0.5, +1

Table 6. Torsional Angles (deg)

Atom1	Atom2	Atom3	Atom4	Angle	Atom1	Atom2	Atom3	Atom4	Angle
C1	O1	C2	C3	107.1(7)	C4	C2	C3	O3	-172.6(7)
C1	O1	C2	C4	-108.0(8)	O1	C2	C4	O4	-77.8(8)
C6	O3	C3	O2	11.4(12)	C1	C2	C4	O4	-145.5(7)
C6	O3	C3	C2	-169.0(7)	C3	C2	C4	O4	67.3(8)
C3	O3	C6	C7	89.8(8)	O3	C6	C8	C9	98.5(8)
C3	O3	C6	C8	-145.4(6)	O3	C6	C8	C13	-82.6(8)
C4	O4	N1	C5	119.5(7)	C7	C6	C8	C9	-142.5(8)
N1	O4	C4	C2	-167.1(6)	C7	C6	C8	C13	36.4(11)
O4	N1	C5	O5	162.7(7)	C6	C8	C9	C10	-179.8(7)
O4	N1	C5	N2	-21.1(11)	C13	C8	C9	C10	1.3(11)
O1	C1	C2	C3	-105.8(8)	C6	C8	C13	C12	180.0(7)
O1	C1	C2	C4	107.4(8)	C9	C8	C13	C12	-1.2(10)
O1	C2	C3	O2	152.8(7)	C8	C9	C10	C11	-0.1(12)
O1	C2	C3	O3	-26.8(9)	C9	C10	C11	C12	-1.2(12)
C1	C2	C3	O2	-140.8(8)	C10	C11	C12	C13	1.3(12)
C1	C2	C3	O3	39.7(0)	C11	C12	C13	C8	-0.1(11)
C4	C2	C3	O2	7.0(11)					

Table 7. Anisotropic Displacement Parameters (U_{ij}, Å^2)

Atom	U_{11}	U_{22}	U_{33}	U_{23}	U_{13}	U_{12}
O1	0.049(3)	0.030(3)	0.066(4)	-0.003(3)	0.004(3)	0.003(2)
O2	0.066(4)	0.028(3)	0.080(5)	-0.010(3)	0.026(4)	-0.008(3)
O3	0.060(3)	0.023(3)	0.062(4)	0.003(3)	0.020(3)	-0.001(2)
O4	0.036(2)	0.035(3)	0.061(4)	0.003(3)	0.007(2)	-0.001(2)
O5	0.042(3)	0.024(2)	0.064(4)	-0.001(3)	0.012(3)	0.000(2)
N1	0.056(4)	0.027(3)	0.065(5)	-0.006(3)	0.022(4)	-0.003(3)
N2	0.047(4)	0.024(3)	0.071(5)	0.007(3)	0.013(4)	0.002(3)
C1	0.033(3)	0.033(4)	0.060(5)	0.001(4)	0.008(3)	0.006(3)
C2	0.026(3)	0.033(4)	0.054(4)	-0.004(4)	0.000(3)	-0.001(3)
C3	0.025(3)	0.030(4)	0.056(4)	0.007(4)	0.002(3)	-0.004(3)
C4	0.039(4)	0.031(4)	0.065(5)	0.009(4)	0.002(4)	-0.005(3)
C5	0.024(3)	0.031(4)	0.070(5)	-0.003(4)	0.001(3)	-0.002(3)
C6	0.049(4)	0.021(4)	0.056(5)	0.002(3)	0.012(4)	-0.004(3)
C7	0.052(5)	0.045(5)	0.068(6)	0.004(5)	0.000(4)	-0.013(4)
C8	0.023(3)	0.020(3)	0.066(5)	0.001(3)	0.000(3)	-0.004(2)
C9	0.032(4)	0.032(4)	0.069(5)	0.006(4)	0.010(4)	0.000(3)
C10	0.034(4)	0.029(4)	0.088(7)	0.017(5)	-0.005(4)	0.005(3)
C11	0.043(4)	0.036(5)	0.067(6)	0.007(4)	-0.007(4)	-0.006(3)
C12	0.043(4)	0.031(4)	0.057(5)	-0.004(4)	0.003(4)	-0.001(3)
C13	0.028(3)	0.022(3)	0.067(5)	0.001(4)	-0.002(3)	0.004(3)

The form of the anisotropic displacement parameter is:

$$\exp[-2\pi^2(h^2a^{*2}U_{11} + k^2b^{*2}U_{22} + l^2c^{*2}U_{33} + 2klb^*c^*U_{23} + 2hla^*c^*U_{13} + 2hka^*b^*U_{12})]$$

Table 8. Derived Atomic Coordinates and Displacement Parameters for Hydrogen Atoms

Atom	x	y	z	U_{eq}, Å^2
H1N	-0.2653	-0.2818	0.4772	0.059
H2NA	-0.3441	0.2424	0.4378	0.056
H2NB	-0.1891	0.1221	0.3858	0.056
H1A	0.5222	-0.2488	0.2533	0.050
H1B	0.4375	-0.4088	0.3212	0.050
H4A	0.0997	-0.3612	0.3981	0.054

H4B	0.1392	-0.1493	0.4269	0.054
H6	0.2497	0.3134	0.2055	0.050
H7A	0.6358	0.3866	0.1868	0.066
H7B	0.6905	0.1690	0.1795	0.066
H7C	0.6015	0.2528	0.2605	0.066
H9	-0.0034	0.3563	0.0903	0.053
H10	-0.0817	0.3754	-0.0480	0.060
H11	0.1770	0.2665	-0.1427	0.059
H12	0.5187	0.1479	-0.0973	0.053
H13	0.6023	0.1257	0.0451	0.047