University of Alberta

LL-Diaminopimelate Aminotransferase: The Mechanism of Substrate Recognition and Specificity

by

Nobuhiko Watanabe

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

Edmonton, Alberta, Canada

Spring, 2011

Examining Committee

Dr. Michael N. G. James, Biochemistry

Dr. Mark J. N. Glover, Biochemistry

Dr. Charles F. B. Holmes, Biochemistry

Dr. John C. Vederas, Chemistry

Dr. Wayne F. Anderson, Molecular Pharmacology and Biological Chemistry, Northwestern University, Chicago

Abstract

Amino acid biosynthesis is an essential process in living organisms. Certain amino acids can be synthesized by some organisms but not by others. L-Lysine is one of the essential amino acids that bacteria can synthesize but humans cannot. This is somewhat inconvenient for humans as much of their L-lysine must come from their diet. However, the lack of the lysine biosynthetic pathway in humans makes the bacterial enzymes within the pathway attractive drug targets. Recently, a novel lysine biosynthetic pathway was discovered in plants, *Chlamydia*e and some archaea. It is called the "diaminopimelate aminotransferase (DAP-AT) pathway". In this pathway, LL-DAP-AT plays a key role by directly converting L-tetrahydrodipicolinate to LL-DAP in a single step. This is a quite interesting characteristic of LL-DAP-AT as the above conversion takes three sequential enzymatic steps in the previously known lysine biosynthetic pathways. Due to its absence in humans, LL-DAP-AT would be an attractive target for the development of novel antibiotics. In order to understand the catalytic mechanism and substrate recognition of LL-DAP-AT, the structural characterization of LL-DAP-AT is of paramount importance. In this thesis, the overall architecture of LL-DAP-AT and its substrate recognition mechanism revealed by the crystal structures of LL-DAP-AT from *Arabidopsis thaliana* and *Chlamydia trachomatis* will be discussed.

The crystal structure of the native LL-DAP-AT from *A. thaliana* (AtDAP-AT) presented in this thesis is the first structure of LL-DAP-AT to be determined. This structure revealed that LL-DAP-AT forms a functional

homodimer and belongs to the type I fold family of PLP dependent aminotransferases. The subsequent determination of the substrate-bound AtDAP-AT structure showed how the two substrates, (LL-DAP and L-Glu) significantly different in size, are recognized by the same set of residues without significant conformational changes in the backbone structure. In addition, the LL-DAP-bound AtDAP-AT structure shows that the C^{ε}-amino group of LL-DAP is recognized stereospecifically by the active site residues that are unique to the family of LL-DAP-AT enzymes.

Lastly, the chlamydial LL-DAP-AT presented in this thesis shows a new "open" conformation for LL-DAP-AT. The implications of the conformational flexibility of CtDAP-AT on the differences in substrate specificities among LL-DAP-AT are discussed.

Acknowledgement

My graduate studies could not have been completed with the help of many people during the last five and a half years. Due to space limitation, I can not list everyone here, but I would like to take this opportunity to express my appreciation.

I am very thankful and think of myself as extremely fortunate for being able to complete my Ph.D. study under Dr. Michael James. The last five and a half years working in the Michael's lab was very enjoyable. He has given me an excellent Ph.D. project and the freedom to explore whatever I wanted to do. I really appreciate all the constructive comments and advise that he has given me for my research papers and research presentations. I have really learned and improved a lot from these.

All the members of the James lab have provided me with a very supportive environment throughout my graduate degree. In the last five and a half years, many people came and went. But, everyone was so helpful, from data processing to protein purifications. I would like to thank Craig, Grace, Jiang, Joanne and Sarah for all the helpful advice in cloning, protein expression, and purifications in the wet lab. Especially, I thank Craig for taking me through many different methods/protocols in protein expression and purifications. I would also like to thank Maia for her help with crystallizations. She made some beautiful crystals.

As I did not know much about data processing before I joined the James lab, many people have helped me analyze diffraction data. Biswal, Sankar and Bindu were always there for me to ask questions regarding data collection and processing. I am certain that I could not have completed my Ph.D. without their everyday help. I would also like to thank Kathy and Sheraz for all their help with the X-ray machine and synchrotron trips. Thank you to Vera who was always helpful for tackling difficult computer problems.

The collaborations with Dr. John Vederas and his laboratory members could not have been any better. Much of the data presented in this thesis is the result of good collaborations with the Vederas laboratory. Especially, Matt, Marco, Sandra and Chenguang have contributed significantly to the preparation of AtDAP-AT. Matt, Marco and Sandra have purified and analyzed AtDAP-AT by the biochemical assay and the mass-spectroscopy. Matt and Chenguang have provided generous amount of substrate analogues for obtaining the substrate-analogue bound AtDAP-AT structures. I really enjoyed working with them. Also, I thank my committee members, Dr. Mark Glover and Dr. Charles Holmes for the constructive advice throughout my graduate carrier. I also thank Dr. Wayne Anderson for kindly accepting to review my thesis and provide valuable comments.

On the non-academic side, I would like to thank Wei and my parents for their continuous support. Wei, my fiancée, has always been supportive of what I do and cheered me up when experiments were not working so well. I also really appreciate my parents for letting me stay and study in Canada all of these years since Grade 11. The biochemistry slo-pitch team, “Kiss my fatty acids” or “Biochemical disasters”, made my graduate life in Edmonton very enjoyable. I am glad that I won at least one T-shirt before finishing my Ph.D.

Lastly, I thank all of the synchrotron beamlines: the Canadian Light Source (CLS), the Stanford Synchrotron Radiation Laboratory (SSRL) and the American Light Source (ALS)) and the funding agencies (Canadian Institute for Health Research (CIHR) and Alberta Innovates Health Solutions) for their technical and financial support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

*	Indicates the residues from neighbour subunit
Å	Angstrom (10^{-10} meters)
AAA	α-aminoadipic acid
AIDS	Acquired immune deficiency syndrome
AspAT	Aspartate aminotransferase
AtDAP-AT	*Arabidopsis thaliana* LL-DAP-AT
AroAT	Aromatic aminotransferase
B-factor	Temperature factor
CtDAP-AT	*Chlamydia trachomatis* LL-DAP-AT
DHDP	Dihydrodipicolinate
DapD	THDP *N*-succinyltransferase
LL-DAP	LL-diaminopimelate
m-DAP	*meso*-diaminopimelate
DTT	Dithiothreitol
E_L	Unprotonated PLP-bound enzyme
E_LH^+	Protonated PLP-bound enzyme
E_M	Unprotonated PMP-bound enzyme
GABA	γ-aminobutyric acid
GSA	Glutamate-1-semialdehyde aminomutase
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HisP-AT	Histidinol-phosphate aminotransferase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K270N-DAP	Asn270 variant of AtDAP-AT in complex with LL-DAP forming true external aldimine

K270N-Glu	Asn270 variant of AtDAP-AT in complex with L-Glu forming true external aldimine
LD	Large domain of LL-DAP-AT
MAD	Multiple-wavelength anomalous dispersion
mAspAT	Mitochondrial aspartate aminotransferase
MtbDAP-AT	*Mycobacteria tuberculosis N*-succinyl DAP aminotransferase
NADPH	Nicotinamide adenine dinucleotide phosphate
PcDAP-AT	*Protochlamydia amoebophila* LL-DAP-AT
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PLP	Pyridoxal 5'-phosphate
PLP-Glu	*N*-(5'-phosphopyridoxyl)-L-glutamate
PLP-DAP	*N*-(5'- phosphopyridoxyl)-LL-diaminopimelate
PLP-THDP	*N*-(5'-phosphopyridoxyl)-tetrahydrodipicolinate
PMP	Pyridoxamine 5'-phosphate
r.m.s.d.	Root-mean-square deviation
S (or S^+)	Unprotonated substrate (or protonated substrate)
SD	Small domain of LL-DAP-AT
Se-Met	Seleno-methionine
STD	Sexual transmitted disease
tAspAT	*Thermus thermophilus* HB8 aspartate aminotransferase
TM1255	*Thermotoga martima* aspartate aminotransferase
THDP	L-Tetrahydrodipicolinate

Chapter 1: Introduction

1.1 Historical perspectives and biological diversity of PLP dependent enzymes

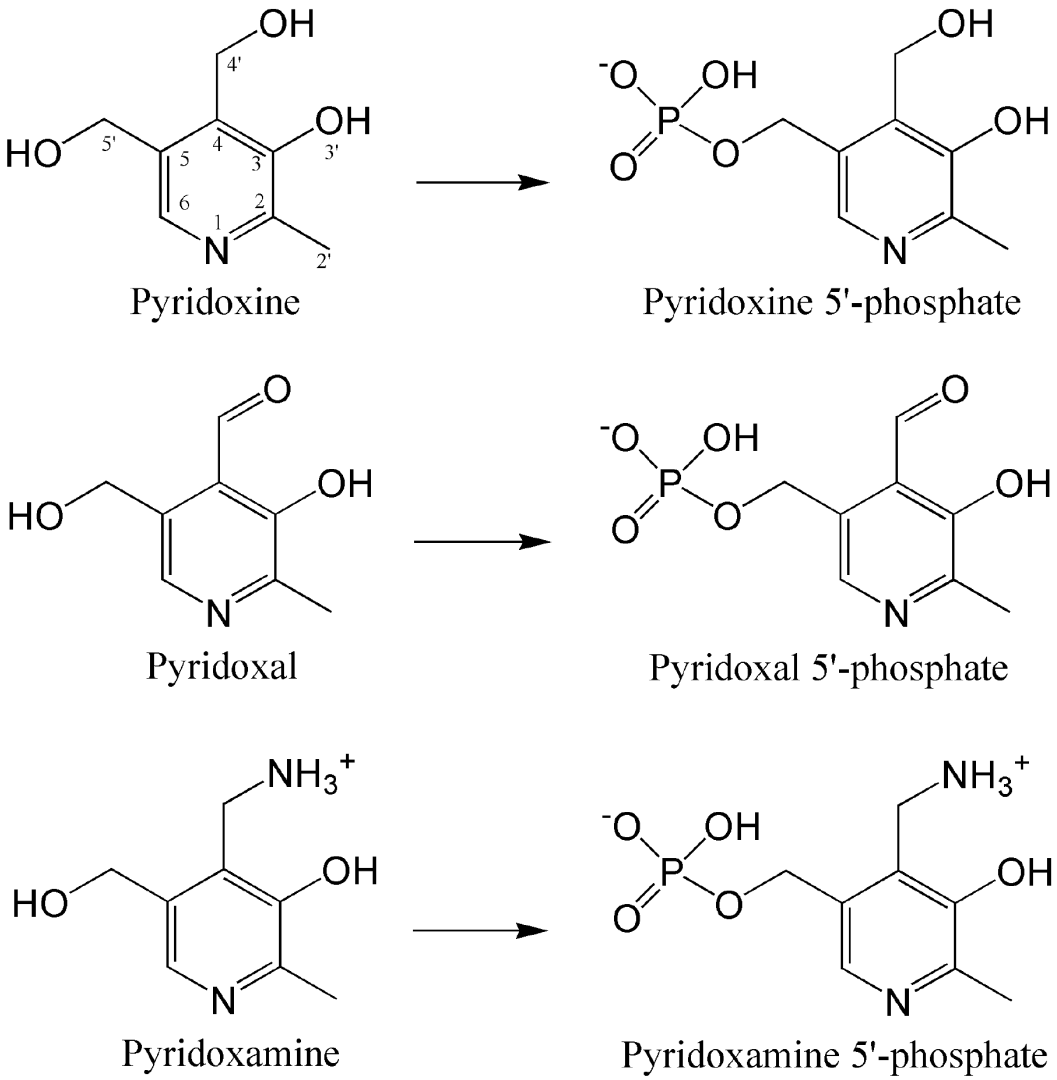
Discovery of pyridoxal-5'-phosphate (PLP)

The discoveries of vitamins are surely one of the greatest findings in science. Currently, thirteen distinct vitamins have been identified, each necessary for the growth and development of living organisms. Deficiencies or overdoses of vitamins can cause a wide variety of diseases in humans such as nyctalopia (Vitamin A deficiency) or rickets (Vitamin D deficiency). With important roles in many metabolic processes, vitamins are of great interest to researchers studying the biochemistry of biological pathways.

Vitamin B_6 is one of the thirteen vitamins currently identified. It was discovered by Paul György in 1935.[1-3] At that time, he was working on a new substance called "pellagra-preventive factor" to cure rats affected by acrodynia, a dermatological problem.[1] Based on his experiments, he believed that the pellagra-preventive factor would represent a new type of vitamin, and tentatively named that substance vitamin B_6.[1] After acquiring the capacity to synthesize vitamin B_6, the chemical structure of the synthetic vitamin B_6 was determined.[4] The structure of synthetic vitamin B_6 was shown to be closely related to that of the pyridine molecule.[4] Due to the similarity of their structures, the synthetic vitamin B_6 molecule was named "pyridoxine" (Figure 1-1), a name that was synonymous to vitamin B_6 for several years after its initial discovery.[5]

While pyridoxine cured the skin disease in György's rats, the effect of pyridoxine on prokaryotic organisms was not well known at that time. Early research had suggested that pyridoxine was required for the growth of some bacteria[6]. However, a few years later, Snell *et al*. showed that pyridoxine itself can

Figure 1-1 The chemical structures of various forms of vitamin B_6. The precursors are shown on the left, and the phosphorylated active forms are shown on the right.



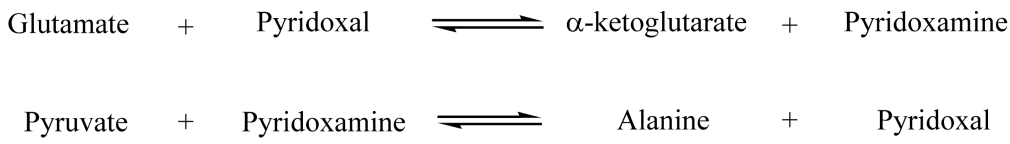
not support growth in those bacteria.[7, 8] In other words, he implied that another active compound, related to pyridoxine, was supporting growth in those bacteria.[7]

Shortly after, the other forms of vitamin B_6[9, 10] including the physiologically active forms of vitamin B_6[11-14] were discovered. Currently, six forms of vitamin B_6 exit in nature (Figure 1-1). In order for vitamin B_6 to participate in biological reactions, the three precursors (pyridoxal, pyridoxine, and pyridoxamine) of vitamin B_6 shown in Figure 1-1 must be converted into the phosphorylated active forms called "pyridoxine 5'-phosphate", "pyridoxal 5'-phosphate" (PLP) and pyridoxamine 5'-phosphate (PMP), respectively.

PLP as a cofactor of biological enzymes

The elucidation of PLP's role in biological systems did not take too long after its discovery. Well before the discovery of PLP, the existence of enzymatic transamination was discovered through the biochemical investigations of pigeon muscle extracts by Braunstein and Kritzmann in the late 1930s.[15-18]

A year later PLP was determined as the active component of Vitamin B_6, Snell reported the mechanism of the reversible PLP dependent non-enzymatic transamination using amino acids and α-keto acids as shown below.[5, 19]



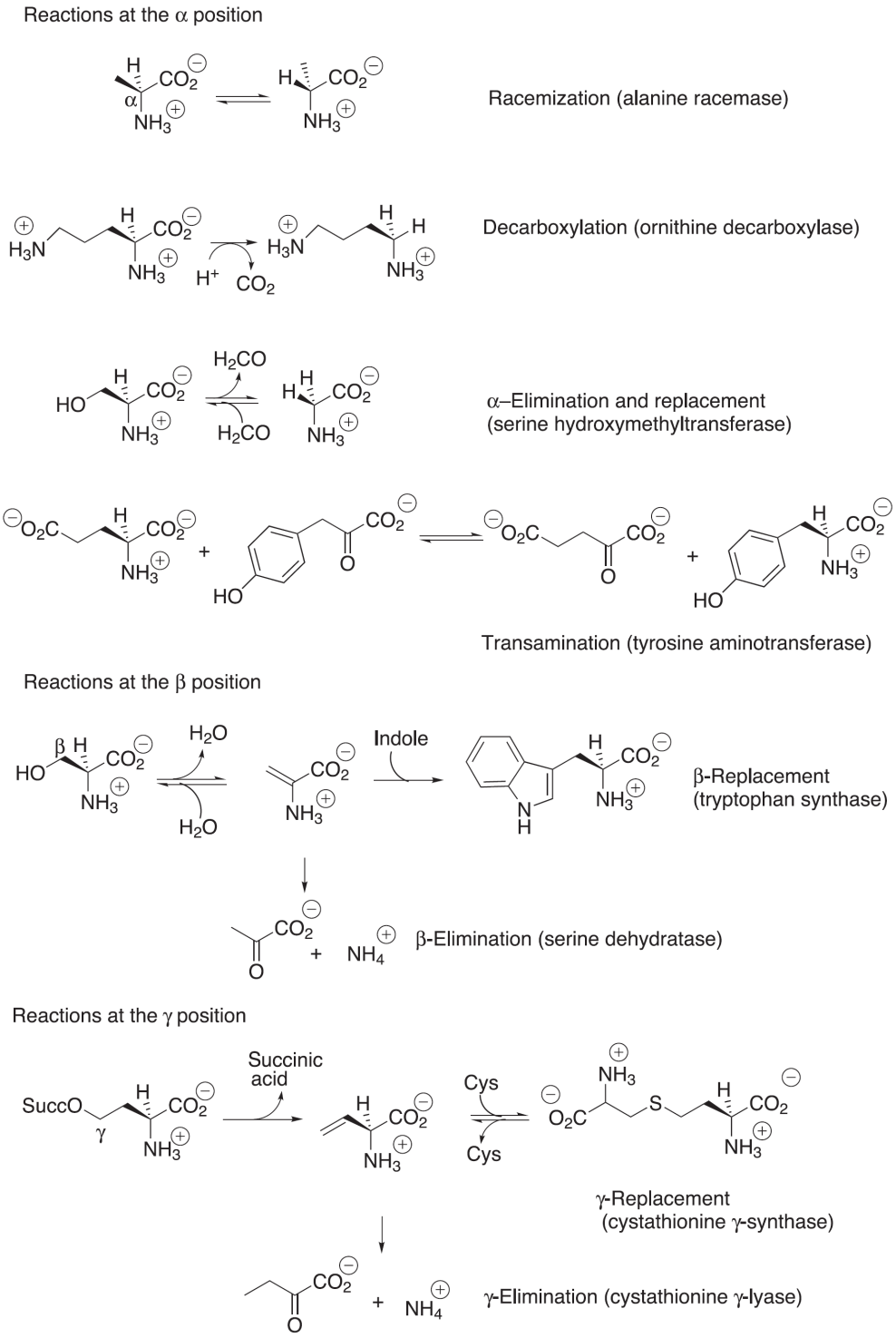
In the same literature, Snell had somewhat intuitively mentioned that the above reaction scheme may exist in biological systems.[19] Subsequently, this hypothesis was proven correct, as the activity of aminotransferases was shown to be well correlated with the presence of PLP.[20-22] Several other experiments were performed to support the above hypothesis.[23-25] Through accumulation of many years of research, PLP became known as the cofactor of several biological enzymes.

Distributions of PLP dependent aminotransferases in biological systems

Over the course of half a century, the number of known PLP dependent enzymes has increased significantly. Currently, more than 140 distinctive PLP dependent enzymes with nine different reaction types (Figure 1-2) have been identified in living organisms.[26, 27] The function of PLP in biological processes is broad. One example is the involvement of transaminases in amino acid biosynthesis. Transaminases are ubiquitously found in amino acid metabolism due to its frequent use for the addition or elimination of amino groups.[28] Another important example of PLP activity in humans is in the production of neurotransmitters. Neurotransmitters such as histamine, serotonin, and γ-aminobutyric acid (GABA) are produced from amino acids through PLP dependent catalysis.[28] The above neurotransmitters are the result of decarboxylation of histidine, tryptophan, and glutamate respectively by distinctive PLP dependent enzymes.[28] In addition to transamination and decarboxylation, PLP dependent enzymes are known to catalyze the racemization of amino acids, α, β, γ-replacement, and α, β, γ-elimination reactions (Figure 1-2).[27] It should be also noted that because PLP's activity is so widely spread over many aspects of biological systems, an imbalance of vitamin B_6 level can cause a variety of illnesses such as dermatitis or neuropathy.

As described above, the involvement of PLP dependent enzymes in biological systems is extremely versatile. At their discovery 70 years ago, no one would have expected PLP to play a role in so many enzymatic reactions. Yet, despite all that is known, there are still uncharacterized PLP dependent enzymes in nature. As the imbalance of vitamin B_6 levels can cause a variety of illnesses, much research has been carried out to identify and understand the mechanism of PLP dependent enzymes.

Figure 1-2 Various types of chemical reaction catalyzed by PLP dependent enzymes. This figure is adopted from Elliot *et al.* (2004). *Annu. Rev. Biochem.* **73,** 383 - 415.[27]



1.2 Classification of PLP dependent enzymes

Early structural studies of PLP dependent enzymes

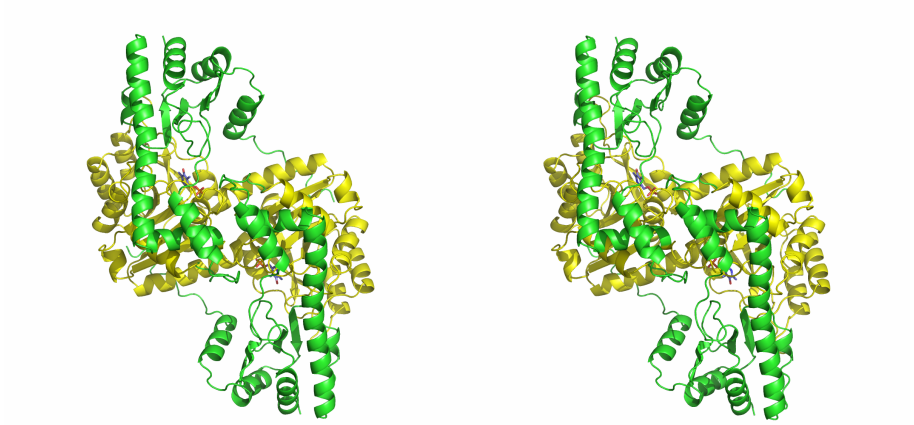
In the last few years, the increase of structural information available on protein molecules has been remarkable. Currently, over 500 PLP associated structures have been deposited into the Protein Data Bank (PDB). X-ray crystallography has been one of the most powerful techniques to understand protein functions. In fact, it has been critical to the understanding of PLP dependent enzymes.

The first crystal structure of the PLP dependent aminotransferases determined was the mitochondrial aspartate aminotransferase (mAspAT) in 1980.[29] Prior to that, there were many mechanisms proposed for PLP dependent aminotransferases based on over 40 years of biochemical studies.[30-37] However, it was not until the first crystal structure of mAspAT that the uncertainties at that time around the catalytic mechanisms of the aminotransferases were answered.[38]

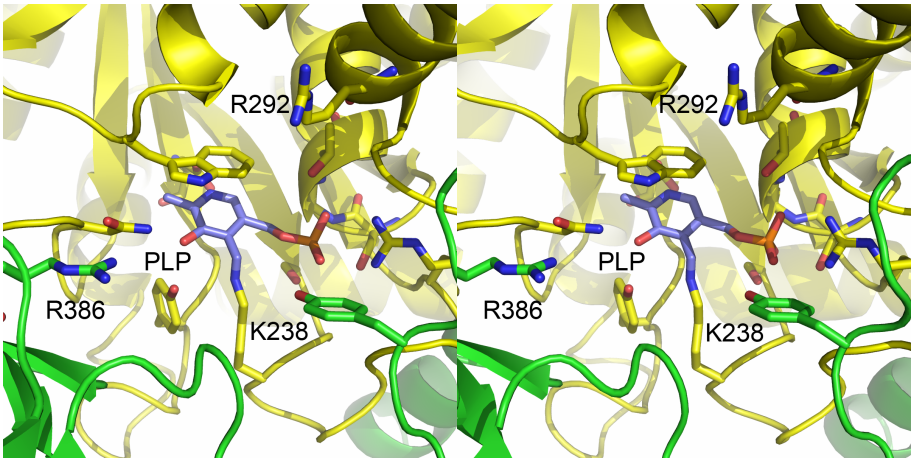
The crystal structure of mAspAT revealed several unique features of aspartate aminotransferases (AspAT) (Figure 1-3a). AspAT consists of a homodimer where PLP is found on each subunit.[29] Each subunit contains large ("PLP binding domain") and small domains. PLP is held in the active site through a number of non-covalent interactions with the surrounding residues and a covalent linkage with the active site lysine situated on the *si* face of PLP (Figure 1-3b). On the *re* face, two arginine residues are located for the binding of dicarboxylic substrates. Including those arginine residues, the active site is composed of a mixture of residues from both subunits. This structure showed the importance of the enzyme being a homodimer. Early structural information helped to answer many important aspects of aminotransferases' catalytic mechanism such as how PLP and substrates are accommodated in the active site.

Figure 1-3 Stereo diagrams of the AspAT structure (7AAT)[39]. a). The homodimeric structure of AspAT. Each subunit contains two domains. The large domain is coloured in yellow, and the small domain is coloured in green. b). The active site of AspAT. PLP binding residues are shown in stick format. Two arginine residues (Arg292 and Arg386) that are important for the substrate binding are also shown in stick format.

a).



b).



Classification of PLP dependent enzymes

As the structural information for PLP dependent enzymes increased, a variation of the tertiary structures within PLP dependent enzymes became apparent. Currently, PLP dependent enzymes are classified into five different superfamilies according to their tertiary structures (Table 1-1).[27, 40, 41] This classification scheme does not correlate with the type of reactions the enzymes catalyze as each group contains enzymes from different reaction types (i.e. both transamination and decarboxylation are found in type I fold type).

The type I fold family is called the aspartate aminotransferase family (Figure 1-4a). Although this group is mainly represented by AspAT, many other enzymes are also members of this family as shown in Table 1-1. As described previously, the characteristics of AspAT are that it is a homodimer with each subunit containing two distinctive domains.[29, 42] The large domain (PLP binding domain) adopts the α-β-α sandwich, and the small domain folds into the α-β complex. In some cases, a significant domain movement is observed upon binding of substrates.[42] This family is the most well studied and the structurally best characterized family.

The type II fold family is the tryptophan synthase β family (Figure 1-4b). This family includes three different enzymes, tryptophan synthase, threonine deaminase, and *O*-acetylserine sulfhydrylase. The main difference from the type I fold family is that in the type II fold family, the active site is composed entirely from one subunit. Nevertheless, the enzymes in this family still generally function as a homodimers. In some cases, higher oligomeric states have been observed.[43] In addition to the main catalytic scaffold, some enzymes in this family have a regulatory domain that is regulated allosterically.[44, 45]

The type III fold family is the alanine racemase family (Figure 1-4c). In this family, the amino acid residues stabilizing the pyridine ring of PLP are

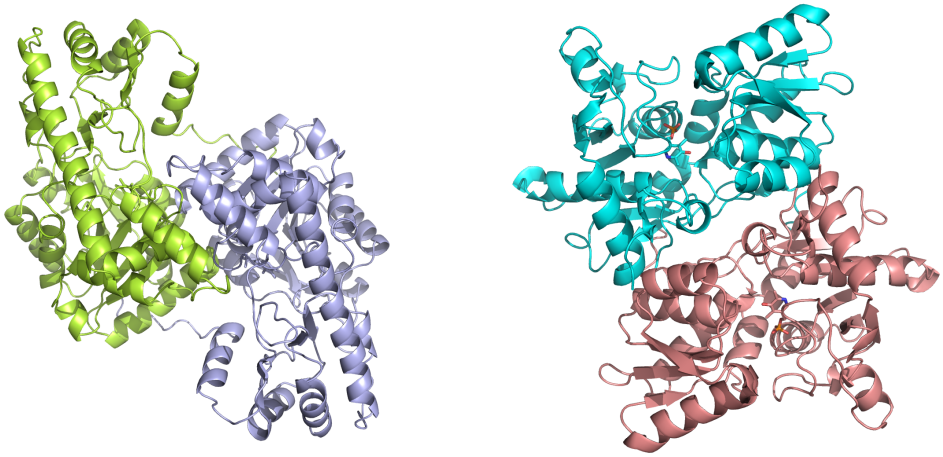
Table 1-1 Fold type classification of PLP dependent enzymes

	Subclass	Enzyme
Fold type I	Aminotransferase subclass I	Aspartate aminotransferase Aromatic acid aminotransferase Tyrosine aminotransferase
	Aminotransferase subclass II	Dialkylglycine decarboxylase Glutamate-1-semialdehyde aminomutase Ornithine aminotransferase ω-Amino acid:pyruvate aminotransferase 8-Amino-7-oxononanoate Diamino-pelargonic acid synthase GABA aminotransferase
	Phosphoserine aminotransferase subclass	Phosphoserine aminotransferase
	Tyrosine-phenol lyase subclass	Tyrosine-phenol lyase Tryptophan-indole lyase
	Cystathionine β-lyase subclass	Cystathionine-β-lyase Cystathionine-γ-synthase
	Ornithine decarboxylase subclass	Prokaryotic ornithine decarboxylase
	Serine hydroxymethyl transferase subclass	Serine hydroxymethyl transferase
	3-Amino5-hydroxybenzoic acid synthase subclass	3-Amino5-hydroxybenzoic acid synthase
Fold type II		Tryptophan synthase O-acetylserine sulfhydrylase Threonine deaminae
Fold type III		Alanine racemase Eukaryotic ornithine decarboxylase Yeast hypothetical protein
Fold type IV		D-amino acid aminotransferase Branched-chain amino acid transferase
Fold type V		Glycogen phosphorylase Maltodextrin phosphorylase

- This table is mostly adopted from Schneider G. *et al.* (2000). *Structure.* **8,** R1-R6.[40]

Figure 1-4 Representative structures of each fold family of PLP dependent enzymes. Type I to IV enzymes form homodimer. Each monomer of the homodimer is representated by the different colours. a). Fold type I: aspartate aminotransferase (1ASN)[42] b). Fold type II: *O*-acetylserine sulfhydrylase (1OAS)[46] c). Fold type III: alanine recemase (1SFT)[47] d). Fold type IV: D-amino acid aminotransferase (1DAA)[48] e). Fold type V: glycogen phosphorylase (1A8I)[49]

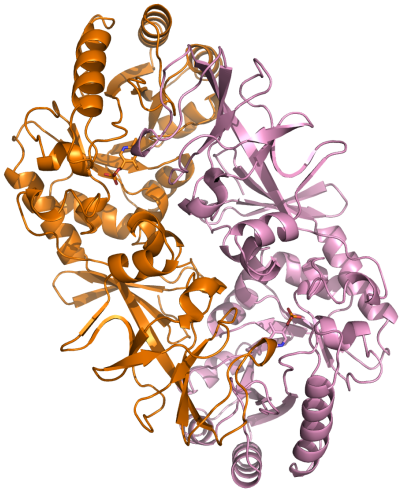
a). b).



Type I fold Type II fold

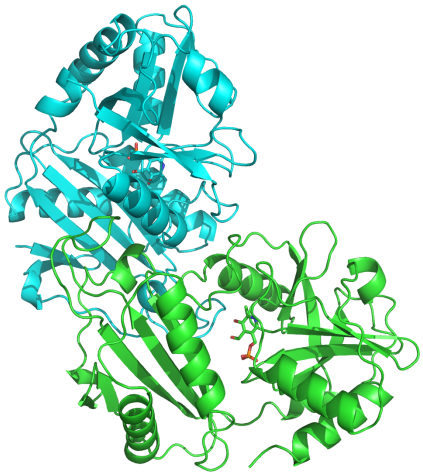
Figure 1-4 (continued)

c).



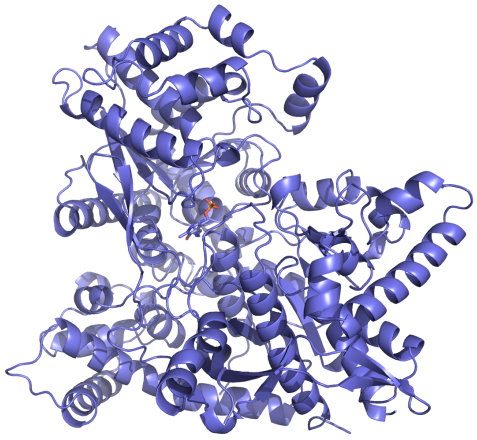
Type III fold

d).



Type IV fold

e).



Type V fold

different from the other families. Normally, the nitrogen atom of the pyridine ring is positively charged and maintained by the negative charge of the well-conserved aspartate side chain. This feature is important to PLP's catalytic competency. In the alanine racemase family, a slightly different mechanism is used. The pyridine nitrogen is kept unprotonated by the proximity of the positive charge of the arginine side chain.[47] The enzymes in this family have a characteristic α/β barrel for the PLP binding scaffold.

The type IV fold family is the D-amino acid aminotransferase family (Figure 1-4d). The interesting feature of this family is that the enzymes catalyze the amino-transfer reaction of D-amino acid instead of L-amino acid. The stereochemical relationship of the two amino acid isomers is directly reflected in the aminotransferase protein structure. Its active site is essentially the mirror image of that in AspAT of the type I fold family.[48] The active site lysine is situated on the *re* face of PLP in D-amino acid aminotransferases instead of on the *si* face of PLP as in AspAT. This arrangement at the active site allows the D-amino acid aminotransferases to accommodate stereochemically opposite substrates.

The last group is the type V fold family, which is the glycogen phosphorylase family (Figure 1-4e). The enzymes in this family are quite different from other families. Typically, the pyridine ring of PLP plays an important role in catalysis in the other fold families. However, enzymes in the type V fold family use the phosphate moiety of PLP for catalysis.[49] Therefore, the fundamental mechanism of PLP chemistry in this family is different from the other fold families. Additionally, the type V fold enzymes tend to be much larger in size than the enzymes from the other families. They contain approximately 800 amino acid residues per subunit as opposed to the approximately 400 amino acid residues in other families.

Although the folding patterns of the above five fold types are quite different from each other, the cofactor based structural alignment has shown some common features of the PLP binding site. In all folding types, the PLP is covalently linked to the active site lysine. This is a landmark feature of the PLP dependent enzymes. Also, the negatively charged phosphate group of PLP is accomodated in the positively charged phosphate binding pocket. The positive charge provided by the helical dipole moment of the α-helix is a conserved feature among all fold types. With the few exceptions of the common features described above, most of the PLP binding residues are in fact quite variable among different fold types.

Recent advances in X-ray crystallography and structural genomics research have increased our knowledge about the appearance of protein molecules and facilitated understanding of the interesting structural relationships among different reaction types. Continued structural and biochemical research on the PLP dependent enzymes would not only increase our understanding of their catalytic mechanism, but will also show how substrates are specifically recognized by these structurally similar but catalytically different enzymes.

1.3 Catalytic mechanism of PLP dependent aminotransferases (Asp AT)

The role of PLP as a cofactor

PLP is one of the most widely used cofactors in biological systems. Despite the fact that several different types of PLP dependent reactions exist in living organisms, the role of PLP in most reactions is virtually the same. The primary role of PLP as a cofactor is to act as an “electron sink” for the stabilization of the carbanion formed during catalysis. Many of the PLP dependent reactions require the deprotonation of the hydrogen atom at the C^{α} position of the substrates. This deprotonation event results in the formation of a carbanion at the

C^{α} atom. Normally, this proton cannot be removed at physiological conditions as its pKa is somewhere around 30.[27, 50] However, PLP is able to facilitate the reaction through the transient delocalization of the non-bonding electrons onto the positively charged nitrogen of its pyridine ring.

Among all the known PLP dependent enzymes, AspAT is mechanistically and structurally the best characterized. Since its discovery, it has been the model system for studying PLP dependent amino-transfer reactions. In this section, the currently accepted catalytic mechanism of the PLP dependent aminotransferases will be discussed using AspAT as a model system.

AspAT catalyzes the reversible amino-transfer reaction shown in Figure 1-5. The kinetic mechanism is known to follow a ping-pong bi-bi mechanism. In this case, the aspartate first binds and donates the amino group to PLP. It then leaves the active site. Subsequently, α-ketoglutarate binds to the active site and accepts the amino group from PMP to become glutamate. Overall, the amino-transfer reaction can be divided into two main stages: 1). Donation of the amino group to PLP and 2).Transfer of the amino group from PMP (Figure 1-6).

Donation of the amino group to PLP

As described previously, PLP is held in the active site through a covalent linkage (imine bond) to the side chain of the active site lysine. This lysine is almost completely conserved within the superfamily of PLP dependent enzymes and plays an essential role during catalysis. In the unliganded state, the covalent complex between PLP and the ε-amino group of lysine is often referred to as an "internal aldimine". At the physiological pH, the imine nitrogen can be protonated (E_LH^+) or unprotonated (E_L) as its pKa is approximately 6.8 (Step 1).[51] Therefore, two kinds of Michaelis complex ($E_LH^+ \bullet SH^+$ & $E_L \bullet SH^+$) can be formed in solution (Step 2 and 1'). However, only the latter form can undergo subsequent

Figure 1-5 The chemical reaction catalyzed by aspartate aminotransferase.

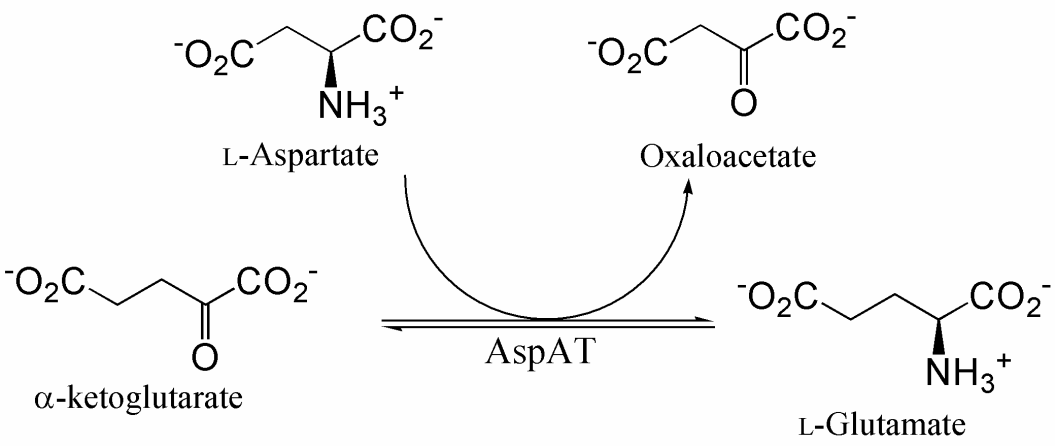
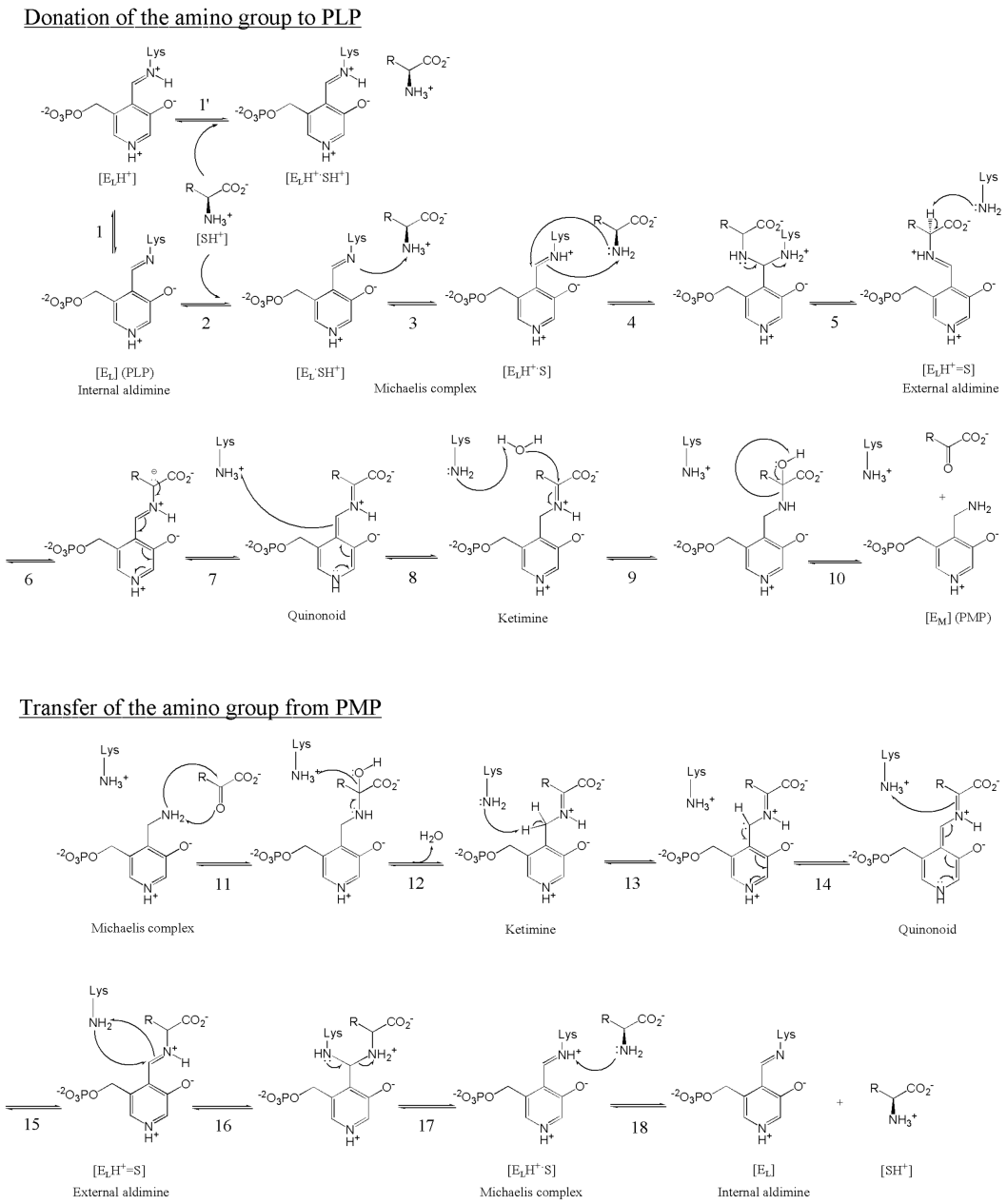


Figure 1-6 The catalytic mechanism of PLP dependent aminotransferase. This scheme is produced based on the currently accepted catalytic mechanism of aspartate aminotransferase. "R" represents "CH_3COO^-" in the "donation of the amino group to PLP" and "$(CH_3)_2COO^-$" for the "transfer of the amino group from PMP".



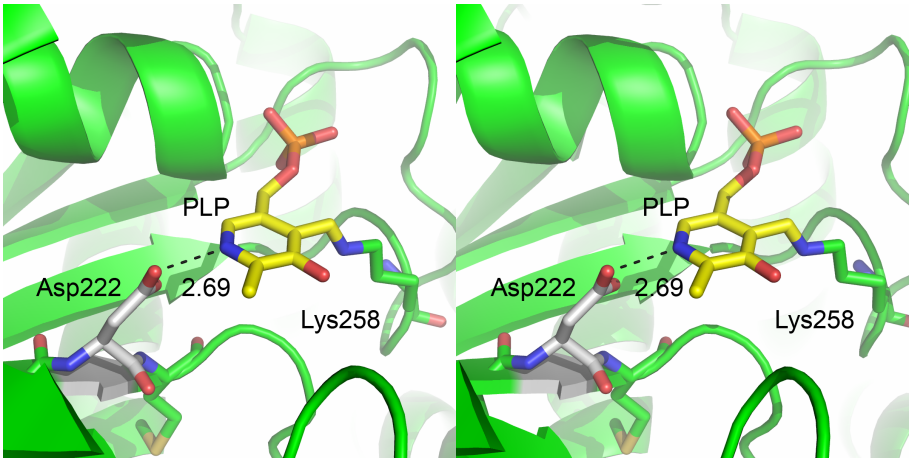
nucleophilic attacks as the α-amino group of aspartate has to be deprotonated to become a nucleophile (Step 3).

This abstraction of the proton from the aspartate α-amino group is accomplished by the unprotonated imine nitrogen of the internal aldimine. Upon binding of aspartate, the chemical environment of the active site is altered so that the pKa of the imine nitrogen increases approximately 2.0 units (pKa ~ 8.8).[30, 52, 53] The change of the imine nitrogen pKa increases the population of the protonated imine nitrogen. (E_L•SH^+→ E_LH^+•S: Step 3).

After the formation of E_LH^+•S, the nucleophilic α-amino group attacks the C4' of PLP and forms a tetrahedral intermediate (Step 4). Subsequently, the electron from the α-amino nitrogen will be pushed around to release the side chain of lysine from PLP; now PLP makes a covalent linkage with aspartate (Step 5). This complex between PLP and the substrate is called an "external aldimine" (E_LH^+=S) as opposed to the previously described internal aldimine (E_LH^+ or E_L).

The next step is the formation of the carbanion at the α-carbon of aspartate (Step 6). In order to break the C^α-N bond through hydrolysis, a double bond must be created at the C^α-N bond shown in Figure 1-6 (Quinonoid intermediate between step 7 and step 8). The deprotonation of the C^α-proton is essential for the creation of the double bond between the C^α carbon and the imine nitrogen atom. As shown in Figure 1-6, the unprotonated ε-amino group of lysine acts as a general base to abstract a proton from C^α (Step 6). Then, the non-bonding electrons of C^α are delocalized onto the pyridine ring to become a quinonoid intermediate (Step 7). It is very important for the pyridine nitrogen to be positively charged because it enhances the pyridine ring's electron withdrawing property. The positive charge of the pyridine nitrogen is maintained through the ion pair interactions with the side-chain of the well-conserved aspartate residue, in this case Asp222 (Figure 1-7).

Figure 1-7 Electrostatic interactions between the nitrogen atom of pyridine ring and the side chain of Asp222 residue (7AAT)[39]. The distance is shown in Å and illustrated with a black dotted line.



The delocalized electrons are quickly pushed back out from the pyridine ring; the C4-C4’ bond becomes a single bond after taking up a hydrogen atom from the ε-amino group of lysine (Step 8). The resultant molecule is called a ketimine. The next step, the last in the first half of the amino-transfer reaction, is the hydrolysis of C^{α}-N double bond (Step 9). The unprotonated ε-amino group of lysine acts as a general base once again to increase the nucleophilicity of a water molecule for the nucleophilic attack on C^{α} (Figure 1-6). Via the formation of a tetrahedral intermediate, oxaloacetate and PMP will be produced at the end of the first half of the aminotransferase reaction (Step 10).

Transfer of the amino group from PMP

The transfer of the amino group from PMP to α-ketoglutarate is shown in the bottom portion of Figure 1-6. This half of the amino-transfer reaction is essentially the reverse of the first half. The amino group of PMP attacks α-ketoglutarate to form the ketimine intermediate (Step 11 and 12). Through stabilization of the carbanion via the quinonoid intermediate (Step 13 and 14), an external aldimine with α-ketoglutarate is formed (Step 15). The release of the end product (glutamate) is accomplished by the nucleophilic attack of the ε-amino group of lysine (Step 16 to 17). This last step restores the initial state of the unliganded enzyme with the lysine forming an internal aldimine with PLP (Step 18). The enzyme then becomes ready for another cycle of the amino-transfer reaction.

Through many years of extensive research, the overall picture of PLP dependent catalysis has become quite clear. However, more research must be done to understand the details of each catalytic step, especially the protonation status of PLP at the different reaction stages. Although it is a simple transfer of amino groups from one substrate to another, quite a few catalytic steps are needed. The

above catalytic mechanism shows how PLP is involved and why it is important for the entire catalytic pathway of PLP dependent enzymes.

1.4 LL-Diaminopimelate aminotransferase: a novel enzyme in the lysine biosynthetic pathway

Introduction to the lysine biosynthetic pathway

The involvement of PLP dependent enzymes in amino acid metabolism has been mentioned briefly in the section 1.1 of this thesis. In bacteria, 17 out of the 20 amino acid biosynthetic pathways are known to contain at least one PLP dependent enzyme.[28] The lysine biosynthetic pathway is one of the 17 PLP dependent amino acid biosynthetic pathways. It has been of great interest to many scientists for many years.

The lysine biosynthetic pathway is an important area of research as it offers many attractive opportunities in antibiotics design[54-57] and plant bioengineering[58]. It serves as an attractive drug target for the following reasons. Firstly, mammals lack the lysine biosynthetic pathway. Drugs specifically targeted to this pathway will therefore only work on the pathogens and should not cause significant side effects in the human host. Secondly, L-lysine and its precursor *meso*-DAP (*m*-DAP) are important building blocks of peptidoglycan layers. Depending on the bacterium, either L-lysine or *m*-DAP will be incorporated into the peptidoglycan structure for cross-linking of the polyglycan chains.[59] Cross-linking produces a strong, mesh-like structure that is very important for bacterial survival. Therefore, the absence of L-lysine or *m*-DAP can significantly inhibit the growth of pathogenic bacteria.

In addition to the development of antibiotics in the medical field, research on the lysine biosynthetic pathway can be applied to the agricultural field.

Drugs targeting the lysine biosynthetic pathway not only act as anti-microbial agents but also can be used as herbicides as plants too have the lysine biosynthetic pathway.

Another important area of research involving the lysine biosynthetic pathway is in plant bioengineering.[58, 60, 61] As implied previously, mammals are not capable of producing their own lysine. They rely on food sources instead. Therefore, many plant researchers are examining ways to increase lysine production in crops or plants so as to assist in increasing the available lysine for the human diet.[58, 60, 61]

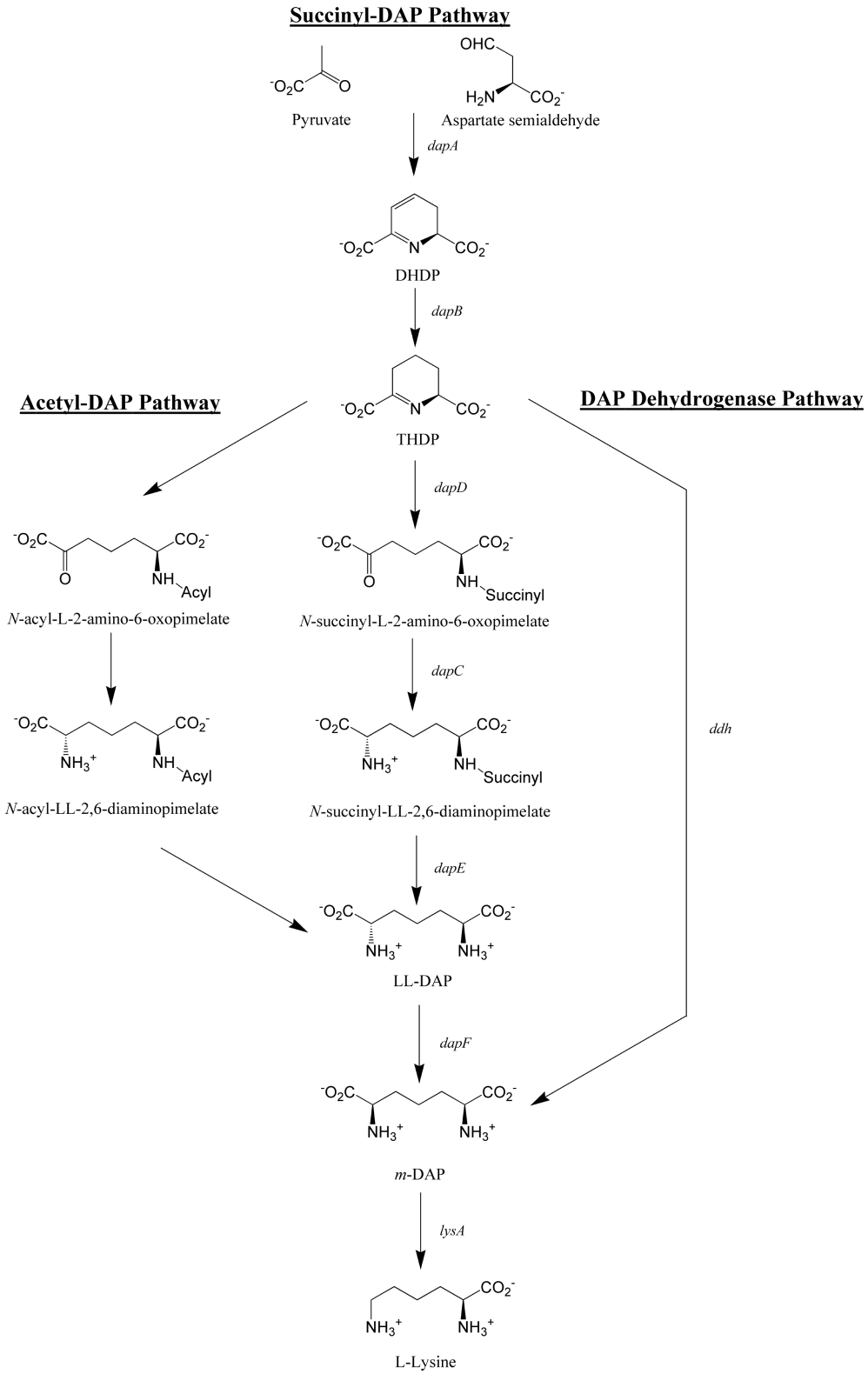
Prokaryotic and eukaryotic lysine biosynthetic pathways

There are five different lysine biosynthetic pathways that have been identified by the early 2000s. Three of them belong to the "diaminopimelic acid (DAP) pathway" and the other two belong to the "α-aminoadipic acid (AAA) pathway". The nomenclature of these pathways is based on the common intermediate used in each pathway.

The three variants of the DAP pathway are shown in Figure 1-8. This pathway is mainly found in prokaryotes. It uses diaminopimelic acid as an intermediate molecule. The first one of the three DAP pathways to be discovered was the *N*-succinyl DAP pathway (Figure 1-8).[62, 63] This pathway is the most widely used lysine biosynthetic pathway in bacteria.

The *N*-succinyl DAP pathway starts with the condensation of aspartate semialdehyde with pyruvate to become dihydrodipicolinate (DHDP) plus water. DHDP is then reduced to tetrahydrodipicolinate (THDP) by DHDP reductase (*dapB*). Next, the THDP succinyl-coA dependent transferase (*dapD*) opens up the six-membered ring of THDP through *N*-succinylation. The resulting product, *N*-succinyl-L-2-amino-6-oxopimelate, becomes an amino-acceptor molecule for

Figure 1-8 Three variants of lysine biosynthetic pathway that use DAP as an intermediate.



the amino-transferring reaction by *N*-succinyl DAP aminotransferase (*dapC*). The following step, catalyzed by *N*-succinyl DAP desuccinylase (*dapE*), removes the succinyl group from *N*-succinyl-2,6-diaminopimelate to produce LL-DAP. LL-DAP is then isomerized to *m*-DAP by DAP epimerase (*dapF*). Finally, the decarboxylation of *m*-DAP by DAP decarboxylase (*lysA*) results in L-lysine.

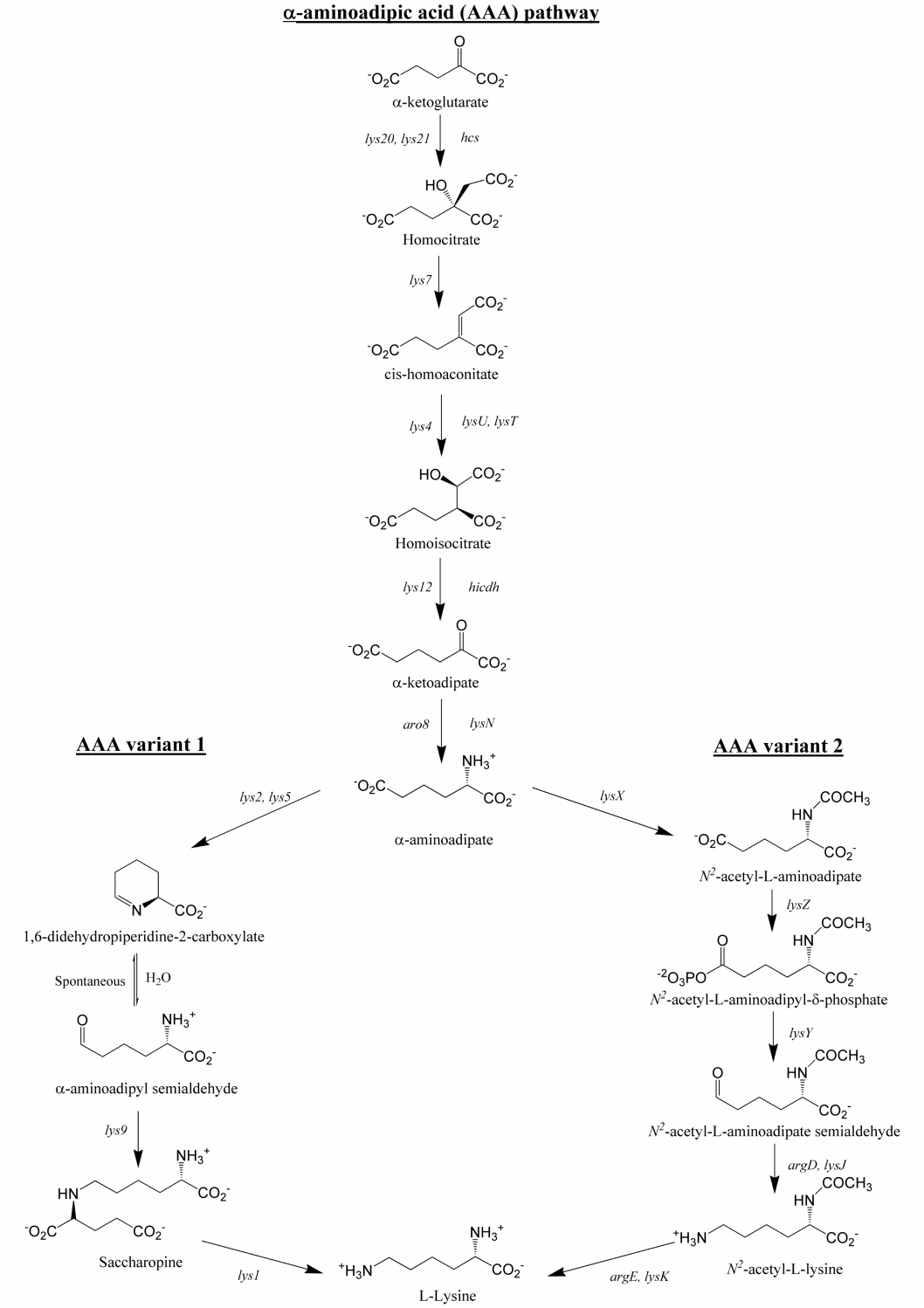
The second DAP pathway, which is used by some *Bacillus* sp., is called the *N*-acetyl DAP pathway (Figure 1-8).[64, 65] This pathway is identical to the above described *N*-succinyl DAP pathway except that *N*-acetylation is used during the conversion of THDP to LL-DAP instead of *N*-succinylation.

Although the *N*-succinyl or *N*-acetyl DAP pathways are used by the majority of bacteria, a few Gram-positive bacteria are known to use the third DAP pathway called the DAP dehydrogenase pathway (Figure 1-8).[66, 67] In this pathway, THDP is directly converted into *m*-DAP in a single step by *m*-DAP dehydrogenase (*ddh*). This pathway often coexists with the *N*-succinyl DAP pathway in some Gram-positive bacteria.[68-70]

The second lysine biosynthetic pathway is called the AAA pathway. This pathway evolved independently from the DAP pathway and uses α-aminoadipic acid as the intermediate molecule. The AAA pathway was first identified in the fungus *Neurospora crassa*.[71, 72] For a long time, it was believed that fungi and euglenoids were the only organisms that used this pathway.[73-78] However, recent findings suggested that particular species of bacteria (i.e. *Thermus thermophilus*)[79-81] and some archaea[82] also use the AAA pathway for lysine biosynthesis.

Two variants of the AAA pathway are shown in Figure 1-9. The first five enzymatic steps converting α-ketoglutarate to α-aminoadipate are exactly the same in the two variants. However, depending on the species, the pathways leading to lysine from the α-aminoadipate intermediate differ significantly. In

Figure 1-9 Two variants of lysine biosynthetic pathway that use AAA as an intermediate.



fungi and euglenoids, α-aminoadipate is converted to lysine via saccharopine. In bacteria, a slightly longer pathway analogous to the arginine biosynthetic pathway is used.[81] So far, no organism is known to possess both the DAP and the AAA pathways.[83]

For over half a century, an incredible amount of research has been dedicated to understanding lysine biosynthetic pathways. As a result, five different lysine biosynthetic pathways have been identified and characterized in different organisms. The variations of the lysine biosynthetic pathways in biological systems show that organisms have evolved to use different pathways to synthesize lysine. Throughout this thesis, the DAP pathway will be the main focus of discussion.

Discovery of LL-diaminopimelate aminotransferase

Plants have been known to synthesize their own lysine for a long time. However, plants do not seem to utilize any of the five lysine biosynthetic pathways described above. Earlier studies with C^{14}-labeling experiments by Vogel H. J. indicated that plants use the DAP pathway instead of the AAA pathway.[84] Since then, one of the bacterial DAP pathway was believed to be utilized in plants for a long period of time.[85, 83] For over half a century, the exact DAP pathway used by plants was not well established as there was no concrete evidence to support a specific pathway.

Nearly 50 years after Vogel's initial experiment, analysis of the *Arabidopsis* genome and biochemical studies of plant extracts have shown that *dapC*, *dapD*, *dapE* and *ddh* activities were missing in plants.[86] These results indicated that the DAP dehydrogenase pathway is not present in plants, but that an incomplete *N*-succinyl or *N*-acetyl DAP pathway may exist.[86] These findings prompted researchers to look for a novel lysine biosynthetic pathway in plants.

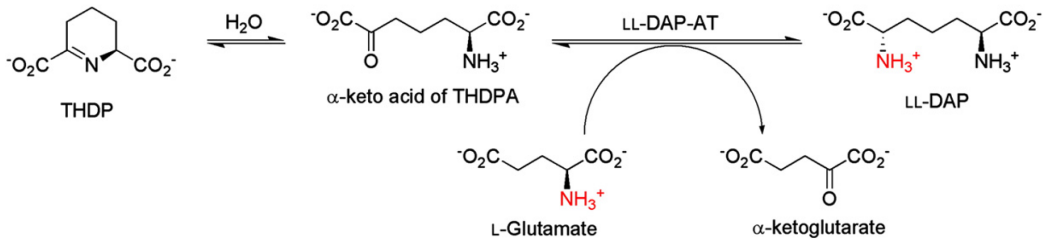
Shortly after the genomic analysis of *Arabidopsis* was reported, Leustek and colleagues discovered the missing piece in the plant lysine biosynthetic pathway- a novel enzyme called LL-diaminopimelate aminotransferase (LL-DAP-AT).[87]

LL-DAP-AT from *Arabidopsis thaliana* (AtDAP-AT) is a PLP dependent aminotransferase that consists of 461 amino acid residues (the first 36 residues is a signal peptide for AtDAP-AT to localize in plastids).[87] The discovery of AtDAP-AT was significant and unique in that this enzyme can bypass three enzymatic steps of the previously described *N*-acetyl or *N*-succinyl DAP pathways.[87] This new lysine biosynthetic pathway that converts THDP to LL-DAP in a single enzymatic step is called the "DAP aminotransferase pathway".[87]

The chemical reaction catalyzed by LL-DAP-AT is shown in Figure 1-10. LL-DAP-AT uses PLP as a cofactor for the reversible conversion of THDP to LL-DAP. The mechanism of PLP dependent amino-transfer by LL-DAP-AT is thought to follow the previously described ping-pong bi-bi mechanism of AspAT. Glutamate first binds to the active site of LL-DAP-AT and donates the amino group to PLP. After glutamate is converted to α-ketoglutarate through the donation of the amino group, PLP becomes PMP. Once α-ketoglutarate leaves the active site, THDP enters and accepts the amino group from PMP to become LL-DAP (Figure 1-10).

The major uncertainty regarding the DAP aminotransferase pathway is whether THDP binds to LL-DAP-AT in a ring-opened or a ring-closed form (Figure 1-10). THDP is known to exist in equilibrium between the ring-opened (hydrolyzed) or ring-closed form in solution.[89] In the *N*-acetyl or *N*-succinyl DAP pathway, THDP is in the open-conformation due to its acetylation or succinylation by DapD. The DAP aminotransferase pathway however, does not contain any DapD homologues for acetylation or succinylation. Therefore, either forms of THDP could be candidates for the substrate of LL-DAP-AT. This issue will be examined more closely in Chapter 2 and 3.

Figure 1-10 The chemical reaction catalyzed by LL-DAP-AT. This figure is adopted from Watanabe *et al.* (2007). *J. Mol. Biol.* **384,** 1314-1329.[88]



Identification of LL-DAP-AT in other organisms

Perhaps, plants were not the only organisms that use the DAP aminotransferase pathway for lysine biosynthesis. Soon after the discovery of LL-DAP-AT in *A. thaliana*, the same DAP aminotransferase pathway was found in *Chlamydia*.[90] The discovery of LL-DAP-AT in *Chlamydia* was not entirely surprising since previous research had suggested that many of *Chlamydia*'s gene products are very closely related to those of plants.[91] This finding not only showed the existence of the DAP aminotransferase pathway outside of plants, but also provided strong support for the evolutionary relationship between plants and *Chlamydia*.[90]

Following the discovery of LL-DAP-AT in *Chlamydia*, LL-DAP-AT gained increased attention from the scientific community. A large microbial genomic search was performed in order to identify the presence of this enzyme in other organisms. Based on the phylogenetic study, LL-DAP-AT was found in specific lineages of bacteria (*Cyanobacteria*, *Desulfuromonadales*, *Firmicutes*, *Bacteroidetes*, *Chlamydiae*, *Spirochaeta*, and *Chloroflexi)* and archaea (*Methanobacteriaceae* and *Archaeoglobaceae)*.[92] These organisms represent only 14% of the currently sequenced microbial genome;[92] thus the presence of the DAP aminotransferase pathway in the bacterial kingdom seems to be quite limited, but definitely present.

Interestingly, the phylogenetic studies revealed that there are two variants of LL-DAP-AT present in above organisms, namely DapL1 and DapL2.[92] The DapL1 and DapL2 variations are not species-specific. They are differentiated based on their amino acid sequence identity.[92] The amino acid sequence identity within each variant is approximately 50-60% compared to less than 30% between the two variants.[92] Whereas the sequence identity suggested that the two variants differ approximately 70% in their amino acid sequences, those differences likely

do not affect the substrate recognition mechanism and the overall protein folding. Based on the currently known crystal structures of PLP dependent enzymes, enzymes with sequence identity of less than 30% can have the same or very similar protein folding, particularly if they catalyze the same chemical reaction.[42, 93]

LL-DAP-AT as a target for novel anti-Chlamydia drugs

The identification of LL-DAP-AT in *Chlamydia* not only supported the plant-*Chlamydia* evolutionary relationship, but also excited researchers interested in the development of antibiotics against *Chlamydia*. *Chlamydia* infections are devastating diseases around the world. Every year, more than 90 million new cases of *Chlamydia* infections are reported worldwide.[94] *C. trachomatis* causes the most frequently reported sexually transmitted disease (STD) in the world.[95] In women, untreated *Chlamydia* infections can cause pelvic inflammatory disease which can result in scarring of the fallopian tubes and subsequent infertility.[95] Moreover, women with *Chlamydia* infections have greater chance of being infected by HIV and developing AIDS.[94, 96] *C. trachomatis* can also cause trachoma, which is one of the commonest causes of infectious blindness in the world.[94, 97] Despite the effort to control *Chlamydia* infections, a large portion of the developing world is still at high risk of developing trachoma and millions of people are still losing their vision.[94, 97] In addition to the serovar causing trachoma, the other serovar (L2) of *C. trachomat*is can cause lymphogranuloma venereum (LGV), which can trigger a fatal infection of the lining of the brain.[95]

Although several antibiotics are available for *Chlamydia* infections, there have been reports of multi-drug resistance *Chlamydia*.[98] Such resistance is not only an emerging problem for *Chlamydia*, but also for many other infectious bacteria. The development of new therapeutic agents is desperately needed to help

prevent morbidity related to *Chlamydia* infections.

LL-DAP-AT appears to be an attractive target for anti-*Chlamydia* drugs. This enzyme is found in very narrow taxonomic groups and importantly, is absent in humans. Understanding the substrate specificity and the inhibition mechanism of LL-DAP-AT can significantly assist in the development of anti-*Chlamydia* drugs.

1.5 Research goals and objectives

My thesis will focus on the structural characterization and elucidation of the substrate recognition mechanism of LL-diaminopimelate aminotransferase from *Arabidopsis thaliana* and *Chlamydia trachomatis*.

Chapter 2: The crystal structures of the native and malate-bound AtDAP-AT will be discussed. As LL-DAP-AT is an important drug target and a relatively new enzyme, the structural information of LL-DAP-AT will be greatly appreciated. Based on the determined native structure of AtDAP-AT, the overall architecture of LL-DAP-AT and the coordination of PLP in the active site will be discussed. The malate-bound AtDAP-AT structure will be used to explain how substrates may bind to the active site of AtDAP-AT as the structure of malate mimics the binding of the key dicarboxylic substrates.

Chapter 3: The mechanism of substrate recognition used by LL-DAP-AT will be discussed. Seven new crystal structures of AtDAP-AT and its active site variants in complex with different substrates or substrate analogues are reported here. Based on analysis of these structures, the substrate binding mode of AtDAP-AT will be discussed. In addition, the structure of AtDAP-AT with modeled *m*-DAP in the active site will be described to explain how AtDAP-AT can differentiate between LL-DAP and *m*-DAP stereospecifically.

Chapter 4: The crystal structures of apo- and PLP-bound LL-DAP-AT from *Chlamydia trachomatis* will be reported in this chapter. The differences in overall architecture between AtDAP-AT and CtDAP-AT will be discussed. Through comparing those two structures and the other PLP dependent enzymes, the potential mechanism of the broad substrate specificity observed in the CtDAP-AT will be discussed.

1.6 References

1. György, P. (1935). Investigations on the vitamin B(2) complex: The differentiation of lactoflavin and the "rat antipellagra" factor. *Biochem. J.* **29**, 741-759.
2. György, P. (1935). Investigations on the vitamin B(2) complex: The distribution of lactoflavin and of the "pellagra-preventing factor" (vitamin B(6)) in natural products of animal origin. *Biochem. J.* **29**, 760-766.
3. György, P. (1935). Investigations on the vitamin B(2) complex: The inactivation of lactoflavin and vitamin B(6) by visible light. *Biochem. J.* **29**, 767-775.
4. Harris, S. A., and Folkers, K. (1939). Synthetic vitamin B6. *Science*. **89**, 347.
5. Dolphin, D., Poulson, R., and Avramovic, O. (1986). Vitamin B6 pyridoxal phosphate: chemical, biochemical and medical aspects. Wiley Interscience, New York, New York.
6. Moeller, E. F. (1938). Vitamin B6 (Adermin) als Wuchsstoff fur Milchsaurebakterien. *Z. Physiol. Chem.* **254**, 285-286.
7. Snell, E. E., Guirard, B. M., and Williams, R. J. (1942). Occurrence in natural products of a physiologically active metabolite of pyridoxine. *J. Biol. Chem*. **143**, 519.
8. Snell, E. E. (1942). Effect of heat sterilization on growth-promoting activity of pyridoxine for *Streptococcus lactis* R. *Proc. Soc. Exp. Biol. Med.* **51**, 356-358.
9. Snell, E. E. (1944). The vitamin activities of "pyridoxal" and "pyridoxamine". *J. Biol. Chem*. **154**, 313.
10. Harris, S. A., Heyl, D., and Folkers, K. (1944). The structure and synthesis of pyridoxamine and pyridoxal. *J. Biol. Chem*. **154**, 315.

11. Gunsalus, I. C., and Bellamy, W. D. (1944). A function of pyridoxal. *J. Biol. Chem*. **155**, 357.
12. Gunsalus, I. C., and Bellamy, W. D. (1944). The function of pyridoxine and pyridoxine derivatives in the decarboxylation of tyrosine. *J. Biol. Chem*. **155**, 557.
13. Gunsalus, I. C., Bellamy, W. D., and Umbreit, W. W. (1944). A phosphorylated derivative of pyridoxal as the coenzyme of tyrosine decarboxylase. *J. Biol. Chem*. **155**, 685.
14. Gunsalus, I. C., Umbreit, W. W., Bellamy, W. D., and Foust, C. E. (1945). Some properties of synthetic codecarboxylase. *J. Biol. Chem*. **161**, 743.
15. Braunstein, A., and Kritzmann, M. (1937). Formation and breakdown of amino-acids by inter-molecular transfer of the amino group. *Nature*. **140**, 503-504.
16. Braunstein, A., and Bychkov, S. (1939). A cell-free enzymatic model of L-amino-acid dehydrogenase ('l deaminase). *Nature*. **144**, 751-752.
17. Braunstein, A. E., and Kritzmann, M. G. (1937). *Biokhimiya 2*. 242-262.
18. Braunstein, A. E., and Kritzmann, M. G. (1937). *Biokhimiya 2*. 859-874.
19. Snell, E. E. (1945). The vitamin-B6 group .5. The reversible interconversion of pyridoxal and pyridoxamine by transamination reactions. *J. Am. Chem. Soc.* **67**, 194-197.
20. Schlenk, F., and Fisher, A. (1945). Note on the purificaiton and properties of glutamic-aspartic transaminase. *Arch. Biochem*. **8**, 337-338.
21. Schlenk, F., and Snell, E. E. (1945). Vitamin-B6 and transamination. *J. Biol. Chem*. **157**, 425-426.
22. Lichstein, H. C., Gunsalus, I. C., and Umbreit, W. W. (1945). Function of the vitamin B6 group: pyridoxal phosphate (codecarboxylase) in transamination. *J. Biol. Chem*. **161**, 311.

23. Metzler, D. E., and Snell, E. E. (1952). Deamination of serine: II. D-Serine dehydrase, a vitamin B6 enzyme from *Escherichia coli*. *J. Biol. Chem*. **198**, 363.
24. Wood, W. A., Gunsalus, I. C., and Umbreit, W. W. (1947). Function of pyridoxal phosphate: resolution and purification of the tryptophanase enzyme of *Escherichia coli*. *J. Biol. Chem*. **170**, 313.
25. Yanofsky, C. (1952). D-Serine dehydrase of Neurospora. *J.Biol. Chem*. **198**, 343.
26. Percudani, R., and Peracchi, A. (2003). A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO reports*. **4**, 850−854.
27. Eliot, A. C., and Kirsch, J. F. (2004). PYRIDOXAL PHOSPHATE ENZYMES: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73**, 383-415.
28. Nelson, D. L., and Cox, M. M. (2000). Lehninger Principles of Biochemistry. 3rd ed. Worth Publisher, New York, New York.
29. Ford, G. C., Eichele, G., and Jansonius, J. N. (1980). Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2559-2563.
30. Karpeisky, M. Y., and Ivanov, V. I. (1966). A molecular mechanism for enzymatic transamination. *Nature*. **210**, 493-496.
31. Metzler, D. E., Ikawa, M., and Snell, E. E. (1954). A General Mechanism for Vitamin B6-catalyzed Reactions1. *J. Am. Chem. Soc*. **76**, 648-652.
32. Jenkins, W. T., and Sizer, I. W. (1957). Glutamic Aspartic Transaminase. *J. Am. Chem. Soc.* **79**, 2655-2656.
33. Fischer, E. H., Kent, A. B., Snyder, E. R., and Krebs, E. G. (1958). The Reaction of sodium borohydride with muscle phosphorylase. *J. Am. Chem. Soc*. **80**, 2906-2907.

34. Cooper, A. J., and Meister, A. (1989). An appreciation of Professor Alexander E. Braunstein. The discovery and scope of enzymatic transamination. *Biochimie*. **71**, 387-404.

35. Hughes, R. C., Jenkins, W. T., and Fischer, E. H. (1962). The site of binding of pyridoxal-5'-phosphate to heart glutamic-aspartic transaminase. *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1615-1618.

36. Dunathan, H. C. (1966). Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **55**, 712.

37. Ivanov, V. I., and Karpeisky, M. Y. (1969). Dynamic three-dimensional model for enzymic transamination. *Adv. Enzymol. Relat. Areas Mol. Biol*. **32**, 21-53.

38. Kirsch, J. F., Eichele, G., Ford, G. C., Vincent, M. G., Jansonius, J. N., Gehring, H., and Christen, P. (1984). Mechanism of action of aspartate aminotransferase proposed on the basis of its spatial structure. *J. Mol. Biol*. **174**, 497-525.

39. McPhalen, C. A., Vincent, M. G., and Jansonius, J. N. (1992). X-ray structure refinement and comparison of three forms of mitochondrial aspartate aminotransferase. *J. Mol. Biol*. **225**, 495-517.

40. Schneider, G., Käck, H., and Lindqvist, Y. (2000). The manifold of vitamin B6 dependent enzymes. *Structure*. **8**, R1−R6.

41. Jansonius, J. N. (1998). Structure, evolution and action of vitamin B6-dependent enzymes. *Curr. Opin Struct Biol*. **8**, 759−769.

42. Jäger, J., Moser, M., Sauder, U., and Jansonius, J. N. (1994). Crystal structures of *Escherichia coli* aspartate aminotransferase in two conformations. Comparison of an unliganded open and two liganded closed forms. *J. Mol. Biol*. **239**, 285-305.

43. Lee, S. J., Ogasahara, K., Ma, J., Nishio, K., Ishida, M., Yamagata, Y., Tsukihara, T., and Yutani, K. (2005). Conformational Changes in the tryptophan synthase from a hyperthermophile upon $\alpha_2\beta_2$ complex formation: crystal structure of the complex. *Biochemistry*. **44**, 11417-11427.

44. Madison, J. T., and Thompson, J. F. (1976). Threonine synthetase from higher plants: stimulation by S-adenosylmethionine and inhibition by cysteine. *Biochem. Biophys. Res. Commun*. **71**, 684-691.

45. Finkelstein, J. D., Kyle, W. E., Martin, J. L., and Pick, A. M. (1975). Activation of cystathionine synthase by adenosylmethionine and adenosylethionine. *Biochem. Biophys. Res. Commun*. **66**, 81-87.

46. Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F., and Jansonius, J. N. (1998). Three-dimensional structure of *O*-acetylserine sulfhydrylase from *Salmonella typhimurium*. *J. Mol. Biol*. **283**, 121-133.

47. Shaw, J. P., Petsko, G. A., and Ringe, D. (1997). Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-A resolution. *Biochemistry*. **36**, 1329-1342.

48. Sugio, S., Petsko, G. A., Manning, J. M., Soda, K., and Ringe, D. (1995). Crystal structure of a D-amino acid aminotransferase: how the protein controls stereoselectivity. *Biochemistry*. **34**, 9661−9669.

49. Gregoriou, M., Noble, M. E., Watson, K. A., Garman, E. F., Krulle, T. M., de la Fuente, C., Fleet, G. W., Oikonomakos, N. G., and Johnson, L. N. (1998). The structure of a glycogen phosphorylase glucopyranose spirohydantoin complex at 1.8 Å resolution and 100 K: the role of the water structure and its contribution to binding. *Protein Sci*. **7**, 915-927.

50. Rios, A., Amyes, T. L., and Richard, J. P. (2000). Formation and Stability of Organic Zwitterions in Aqueous Solution: Enolates of the Amino Acid Glycine and Its Derivatives. *J. Am. Chem. Soc.* **122**, 9373-9385.

51. Kallen, R. G., Korpela, T., Martell, A. E., Matsushima, Y., Metzler, C. M., Metzler, D. E., Morozov, Y. V., Ralston, I. M., Savin, F. A., Torchinsky, Y. M., and Ueno, H. (1985). Chemical and spectroscopic properties of pyridoxal and pyridoxamine phosphates. in *Transaminases*, pp 37-108. John Wiley & Sons, New York, New York.
52. Hayashi, H., Mizuguchi, H., Miyahara, I., Islam, M. M., Ikushiro, H., Nakajima, Y., Hirotsu, K., and Kagamiyama, H. (2003). Strain and catalysis in aspartate aminotransferase. *Biochim. Biophys. Acta*. **1647**, 103-109.
53. Mizuguchi, H., Hayashi, H., Okada, K., Miyahara, I., Hirotsu, K., and Kagamiyama, H. (2001). Strain is more important than electrostatic interaction in controlling the pKa of the catalytic group in aspartate aminotransferase. *Biochemistry*. **40**, 353-360.
54. Vederas, J. C. (2006). Diaminopimelate and lysine biosynthesis - An antimicrobial target in bacteria. *Canad. J. Chem.* **84**, 1197-1207.
55. Cox, R. J., Sutherland, A., and Vederas, J. C. (2000). Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorg. Med. Chem*. **8**, 843-871.
56. Scapin, G., and Blanchard, J. S. (1998). Enzymology of bacterial lysine biosynthesis. *Adv. Enzymol. Relat. Areas Mol. Biol*. **72**, 279-324.
57. Hutton, C. A., Perugini, M. A., and Gerrard, J. A. (2007). Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Mol. BioSyst.* **3**, 458.
58. Azevedo, R. A., and Arruda, P. (2010). High-lysine maize: the key discoveries that have made it possible. *Amino Acids*. 979-989.
59. van Heijenoort, J. (2001). Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod. Rep*. **18**, 503-519.
60. Mazur, B., Krebbers, E., and Tingey, S. (1999). Gene discovery and product development for grain quality traits. *Science*. **285**, 372-375.

61. Mertz, E. T., Bates, L. S., and Nelson, O. E. (1964). Mutant Gene That Changes Protein Composition and Increases Lysine Content of Maize Endosperm. *Science*. **145**, 279-280.

62. Gilvarg, C. (1961). *N*-Succinyl-α-amino-6-ketopimelic acid. *J. Biol. Chem*. **236**, 1429-1431.

63. Gilvarg, C. (1959). *N*-Succinyl-L-diaminopimelic acid. *J. Biol. Chem*. **234**, 2955-2959.

64. Sundharadas, G., and Gilvarg, C. (1967). Biosynthesis of α, ε-diaminopimelic acid in *Bacillus megaterium*. *J. Biol. Chem*. **242**, 3983-3984.

65. Weinberger, S., and Gilvarg, C. (1970). Bacterial distribution of the use of succinyl and acetyl blocking groups in diaminopimelic acid biosynthesis. *J. Bacteriol*. **101**, 323-324.

66. Misono, H., Togawa, H., Yamamoto, T., and Soda, K. (1976). Occurrence of meso-α, ε-diaminopimelate dehydrogenase in *Bacillus sphaericus*. *Biochem. Biophys. Res. Commun*. **72**, 89-93.

67. White, P. J. (1983). The essential role of diaminopimelate dehydrogenase in the biosynthesis of lysine by *Bacillus sphaericus*. *J. Gen. Microbiol*. **129**, 739.

68. Shaw-Reid, C. A., McCormick, M. M., Sinskey, A. J., and Stephanopoulos, G. (1999). Flux through the tetrahydrodipicolinate succinylase pathway is dispensable for L-lysine production in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol*. **51**, 325-333.

69. Wehrmann, A., Phillipp, B., Sahm, H., and Eggeling, L. (1998). Different modes of diaminopimelate synthesis and their role in cell wall integrity: a study with *Corynebacterium glutamicum*. *J. Bacteriol*. **180**, 3159-3165.

70. Schrumpf, B., Schwarzer, A., Kalinowski, J., Pühler, A., Eggeling, L., and Sahm, H. (1991). A functionally split pathway for lysine synthesis in *Corynebacterium glutamicium*. *J. Bacteriol*. **173**, 4510-4516.

71. Abelson, P. H., and Vogel, H. J. (1955). Amino acid biosynthesis in *Torulopsis utilis* and *Neurospora crassa*. *J. Biol. Chem*. **213**, 355-364.

72. Mitchell, H. K., and Houlahan, M. B. (1948). An intermediate in the biosynthesis of lysine in Neurospora. *J. Biol. Chem*. **174**, 883-887.

73. LéJohn, H. B. (1971). Enzyme regulation, lysine pathways and cell wall structures as indicators of major lines of evolution in fungi. *Nature*. **231**, 164-168.

74. Matsuda, M., and Ogur, M. (1969). Separation and specificity of the yeast glutamate-α-ketoadipate transaminase. *J. Biol. Chem*. **244**, 3352-3358.

75. Rowley, B., and Tucci, A. F. (1970). Homoisocitric dehydrogenase from yeast. *Arch. Biochem. Biophys*. **141**, 499-510.

76. Weidner, G., Steffan, B., and Brakhage, A. A. (1997). The *Aspergillus nidulans* lysF gene encodes homoaconitase, an enzyme involved in the fungus-specific lysine biosynthesis pathway. *Mol. Gen. Genet*. **255**, 237-247.

77. Bhattacharjee, J. K. (1985). α-Aminoadipate pathway for the biosynthesis of lysine in lower eukaryotes. *Crit. Rev. Microbiol*. **12**, 131-151.

78. Vogel, H. J. (1965). Lysine biosynthesis and evolution. in *Evolving Genes and Proteins*, pp 25-40. Academic Press, New York, New York.

79. Nishida, H. (2001). Distribution of genes for lysine biosynthesis through the aminoadipate pathway among prokaryotic genomes. *Bioinformatics*. **17**, 189-191.

80. Nishida, H., and Nishiyama, M. (2000). What is characteristic of fungal lysine synthesis through the α-aminoadipate pathway? *J. Mol. Evol*. **51**, 299-302.

81. Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T., and Yamane, H. (1999). A prokaryotic gene cluster involved in synthesis of lysine through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. *Genome Res*. **9**, 1175-1183.

82. Schäfer, S., Paalme, T., Vilu, R., and Fuchs, G. (1989). ^{13}C-NMR study of acetate assimilation in *Thermoproteus neutrophilus*. *Eur. J. Biochem*. **186**, 695-700.

83. Velasco, A. M., Leguina, J. I., and Lazcano, A. (2002). Molecular evolution of the lysine biosynthetic pathways. *J. Mol. Evol*. **55**, 445-459.

84. Vogel, H. J. (1959). On biochemical evolution: lysine formation in higher plants. *Proc. Natl. Acad. Sci. U.S.A.* **45**, 1717-1721.

85. Azevedo, R. A., Lancien, M., and Lea, P. J. (2006). The aspartic acid metabolic pathway, an exciting and essential pathway in plants. *Amino Acids*. **30**, 143-162.

86. Hudson, A. O., Bless, C., Macedo, P., Chatterjee, S. P., Singh, B. K., Gilvarg, C., and Leustek, T. (2005). Biosynthesis of lysine in plants: evidence for a variant of the known bacterial pathways. *Biochim. Biophys. Acta*. **1721**, 27-36.

87. Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol*. **140**, 292-301.

88. Watanabe, N., Cherney, M. M., van Belkum, M. J., Marcus, S. L., Flegel, M. D., Clay, M. D., Deyholos, M. K., Vederas, J. C., and James, M. N. G. (2007). Crystal structure of LL-diaminopimelate aminotransferase from *Arabidopsis thaliana*: a recently discovered enzyme in the biosynthesis of L-lysine by plants and *Chlamydia*. *J. Mol. Biol*. **371**, 685-702.

89. Caplan, J. F., Sutherland, A., and Vederas, J. C. (2001). The first stereospecific synthesis of L-tetrahydrodipicolinic acid; a key intermediate of diaminopimelate metabolism. *J. Chem. Soc., Perkin Trans. 1*. **18**, 2217-2220.

90. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T., and Maurelli, A. T. (2006). L,L-diaminopimelate aminotransferase, a

trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17909-17914.

91. Brinkman, F. S. L., Blanchard, J. L., Cherkasov, A., Av-Gay, Y., Brunham, R. C., Fernandez, R. C., Finlay, B. B., Otto, S. P., Ouellette, B. F. F., Keeling, P. J., Rose, A. M., Hancock, R. E. W., Jones, S. J. M., and Greberg, H. (2002). Evidence that plant-like genes in *Chlamydia* species reflect an ancestral relationship between *Chlamydiaceae*, cyanobacteria, and the chloroplast. *Genome Res*. **12**, 1159-1167.
92. Hudson, A. O., Gilvarg, C., and Leustek, T. (2008). Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase. *J. Bacteriol*. **190**, 3256-3263.
93. Nakai, T., Okada, K., Akutsu, S., Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S., and Hirotsu, K. (1999). Structure of *Thermus thermophilus* HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry*. **38**, 2413-2424.
94. Belland, R., Ojcius, D. M., and Byrne, G. I. (2004). Focus: *Chlamydia*. *Nat Rev Micro*. **2**, 530-531.
95. Manavi, K. (2006). A review on infection with *Chlamydia trachomatis*. *Best Practice & Research Clinical Obstetrics & Gynaecology*. **20**, 941−951.
96. Galvin, S. R., and Cohen, M. S. (2004). The role of sexually transmitted diseases in HIV transmission. *Nat. Rev. Microbiol*. **2**, 33-42.
97. Thylefors, B., Négrel, A. D., Pararajasegaram, R., and Dadzie, K. Y. (1995). Global data on blindness. *Bull. World Health Organ*. **73**, 115-121.
98. Somani, J., Bhullar, V. B., Workowski, K. A., Farshy, C. E., and Black, C. M. (2000). Multiple-drug-resistant *Chlamydia trachomatis* associated with clinical treatment failure. *J. Infect. Dis*. **181**, 1421−1427.

Chapter 2: Crystal Structure of LL-Diaminopimelate Aminotransferase from *Arabidopsis thaliana*: a Recently-Discovered Enzyme in the Biosynthesis of L-Lysine by Plants and *Chlamydia*

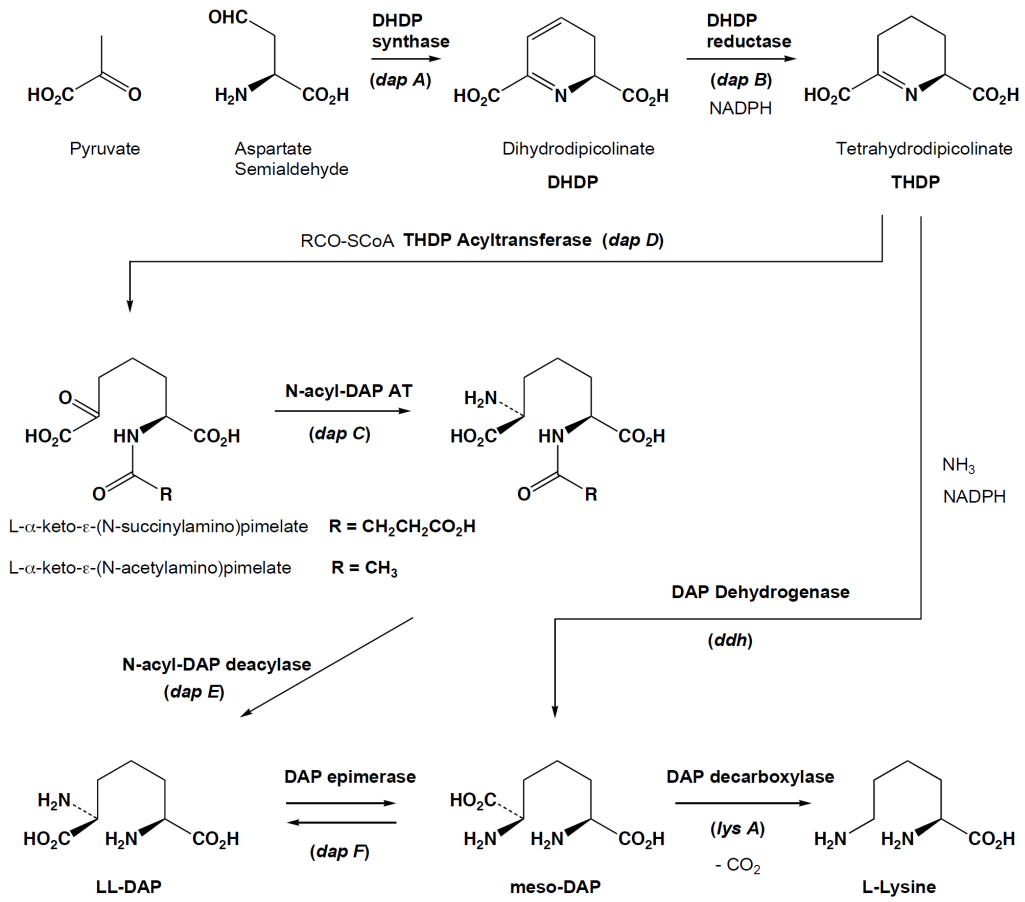
2.1 Introduction

The diaminopimelic acid (DAP) pathway to L-lysine in bacteria and plants has been extensively investigated by many groups because inhibition of the enzymes within it presents an attractive opportunity for development of antimicrobial agents or herbicides.[1-3] This is especially true as L-lysine is an essential amino acid for mammals and must be acquired in their diet. It can also be a limiting nutrient in food, and considerable effort has been devoted to the generation of plants with high lysine content.

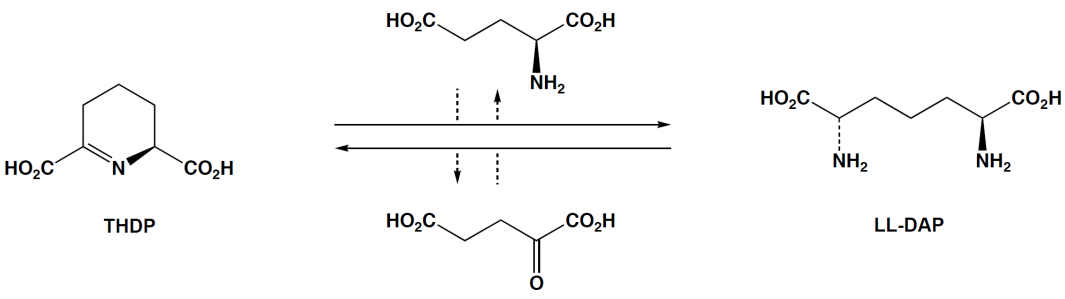
Bacteria have long been known to use one (or more) of three biosynthetic routes to L-lysine (Figure 2-1a)[3]. All routes proceed from the condensation of L-aspartate semialdehyde and pyruvate to give L-dihydrodipicolinate, which is then reduced to L-tetrahydrodipicolinate (THDP). At this stage, most bacteria use *N*-succinylation followed by transamination and desuccinylation to make LL-diaminopimelate (LL-DAP), which is then epimerized by DAP epimerase[4], to form *meso*-DAP (*m*-DAP). This latter compound is utilized directly in the formation of peptidoglycan of virtually all Gram-negative bacteria[5] and is also enzymatically decarboxylated with inversion of configuration to form L-lysine.[6] L-Lysine is incorporated into the peptidoglycan of many Gram-positive organisms[5] as well as being used to make proteins. Some *Bacillus* sp. use *N*-acetylation rather than *N*-succinylation, whereas a few Gram-positive organisms rely on DAP dehydrogenase to directly transform THDP to *m*-DAP. In a few cases the latter pathway coexists with an acetylation route to LL-DAP.[7]

Figure 2-1 a). Biosynthetic pathways of L-lysine in most bacteria, b). Reaction catalyzed by LL-diaminopimelate aminotransferase.

a).



b).



For a considerable time, it was believed that plants and bacteria use the same approach to biosynthesize L-lysine. Indeed, the early enzymes up to THDP formation have been purified and well-studied from plants with the goal of making high lysine producers.[8] In 2006 Leustek, Gilvarg and coworkers reported that higher plants and cyanobacteria have yet another and shorter pathway, namely direct conversion of THDP to LL-DAP catalyzed by LL-DAP aminotransferase (LL-DAP-AT) (Figure 2-1b).[9] LL-DAP-AT has <20% sequence identity when compared to bacterial DAP-ATs (eg, *N*-succinyl-DAP aminotransferase, DapC). It has been suggested that LL-DAP-AT shares a distinct linage which is different from DapC.[9] The plant LL-DAP-AT is a pyridoxal-5'-phosphate (PLP) dependent enzyme that uses L-glutamate (L-Glu) as the amino donor, presumably via the typical PLP mechanism for transamination. An uncertain feature is whether the true substrate is THDP or its hydrolysis product, L-α-keto-ε-aminopimelate, which exists in equilibrium with it in aqueous solution.[10]

Very recently, Leustek, Maurelli and co-workers reported that *Chlamydia trachomatis* and *Protochlamydia amoebophila* also use the LL-DAP-AT pathway, thereby supporting an evolutionary relationship between cyanobacteria and *Chlamydiae*.[11] As *Chlamydiae* causes a variety of diseases in animals and humans, including sexually-transmitted infections (*C. trachomatis*) and are associated with atherosclerosis and coronary heart disease (*C. pneumoniae*),[11] LL-DAP-AT represents an interesting enzyme to target for possible drug development. Here, we describe the three-dimensional structure of this protein from *Arabidopsis thaliana* as a prelude to inhibitor design and testing.

2.2 Material and methods

Cloning, expression and purification of AtDAP-AT

DNA encoding AtDAP-AT fused to a C-terminal histidine-tag (His_6) was obtained from BioBasic Inc. (Ontario, Canada). The design of the gene for AtDAP-AT was based on the sequence of At4g33680 whereby the codons of the 36 amino acid transit peptide were replaced by a methionine start codon preceded by a ribosome binding site. Furthermore, the codon usage of the DNA encoding AtDAP-AT was optimized for *E. coli*. The DNA fragment was designed with EcoRI and HindIII sites adjacent to the gene enabling the fragment to be cloned into the EcoRI and HindIII sites of the expression vector pQE60 (Qiagen Inc.). The resulting plasmid was transformed into *E. coli* strain M15 (pREP4) (Qiagen Inc.) and grown in 2×YT medium at 37°C to an A_{600} of 0.7 with shaking. IPTG was added (1 mM) and culture growth was continued for 4 h after which the cells were harvested. For seleno-methionine (Se-Met) labelling, cells were grown under conditions that inhibit methionine production according to established procedures[12, 13] except that the growth temperature was reduced to 25°C before the addition of amino acids and IPTG and incubation was continued for an additional 20 h before cell harvest.

To purify AtDAP-AT, cells were suspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, 10 mM 2-mercaptoethanol (pH 7.7)) and lysed by French press. Soluble protein was purified by affinity chromatography on Talon Superflow Metal Affinity Resin (Clontech). The column was washed in lysis buffer and eluted in 50 mM sodium phosphate, 300 mM NaCl, 200 mM imidazole, 10 mM 2-mercaptoethanol (pH 8.0). Protein was dialyzed and concentrated to 10 mg/ml in 200 mM NaCl, 20 mM Hepes-KOH (pH 7.6), 3 mM DTT.

Enzyme assay

Enzymatic activity was quantitatively measured in the physiologically reverse direction in 1.00 mL cells containing 100 μmol Hepes-KOH (pH 7.6), 0.4 μmol NADPH, 4 μmol α-ketoglutarate, 40 μmol NH_4Cl, 11.2 μg DAP dehydrogenase, and 0.32 μg of AtDAP-AT. The concentration of LL-DAP was varied from 10 μM - 200 μM. The reactions were incubated at 30°C and the decrease in absorbance at 340 nm was monitored. Kinetic constants were calculated by non-linear regression analysis using GraphPad Prism version 4.03 for Windows (GraphPad Software, San Diego, California, USA).

Crystallization of AtDAP-AT

The purified native AtDAP-AT was concentrated to 10 mg/mL and dialyzed against 100 mM NaCl, 20 mM Hepes (pH 7.6), and 1 mM DTT. The initial screening of the possible AtDAP-AT crystallization conditions was conducted with the Index Crystal Screening Kit (96-well, Hampton Research) at room temperature. AtDAP-AT (0.4 μL) was mixed with an equal volume of each reservoir solution. A small crystal of AtDAP-AT appeared within a week in a presence of 2.1 M D, L-malic acid (malate at pH 7.0) as the precipitant. After optimizing the initial crystallization conditions, X-ray-diffraction quality crystals were obtained in a drop containing 1 μL of protein solution and 1 μL of reservoir solution (2.1 M D, L-malic acid, pH 7.0) with the hanging drop vapour diffusion method at room temperature. AtDAP-AT crystals grew to the dimensions of 0.10 mm x 0.04 mm x 0.02 mm.

Purified Se-Met substituted AtDAP-AT was also concentrated to 10 mg/mL and dialyzed against the same screening solutions as native AtDAP-AT. Se-Met AtDAP-AT was crystallized in 45% (w/v) $(NH_4)_2SO_4$, 0.1 M Hepes (pH 7.5), 3% (w/v) PEG400 with the hanging drop vapour diffusion method. Within a

week, X-ray diffraction quality crystals of Se-Met AtDAP-AT appeared in a drop containing 1 μL of the protein solution and 1 μL of the reservoir solution. The crystals grew to the dimensions of 0.10 mm x 0.10 mm x 0.03 mm.

Data collection

Prior to data collection, the native crystal was flash-cooled to 100 K with liquid nitrogen in a cryo-protectant containing 30% (v/v) glycerol and 2.1 M D, L-malic acid (pH 7.0). In a same manner, the Se-Met derivative crystal was flash-cooled in a cryo-protectant containing 30% (v/v) glycerol, 45% (w/v) $(NH_4)_2SO_4$, 0.1 M Hepes (pH 7.5), 3% (w/v) PEG400.

Data collection of the native crystal grown in D, L-malic acid was carried out at the Advanced Light Source (ALS) beamline 8.3.1, Berkeley, CA, USA. Multiple-wavelength anomalous dispersion (MAD) data sets of Se-Met substituted AtDAP-AT crystals were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2, CA, USA. Se-Met MAD data were collected at three different wavelengths (peak, inflection, and remote) with inverse beam geometry in order to obtain sufficient phase information. In addition to the MAD data set, a high resolution data set using the same Se-Met substituted crystal was also collected at SSRL. The crystallographic data for the native and the Se-Met substituted crystals are given in Table 2-1. All data sets were scaled and processed with the HKL2000 package[14]. Detailed data collection statistics are presented in Table 2-1.

Structure determination and refinement

Scaled MAD datasets of AtDAP-AT were analyzed with SHARP[15]. In SHARP, the coordinates of the selenium sites were determined and refined. A total of 18 out of the 20 expected selenium sites in the AtDAP-AT dimer were found.

Table 2-1 Data collection and crystallographic statistics

	Se-Met-DAP AT		
Crystal	Remote	Peak	Inflection
Space group	$P3_221$	$P3_221$	$P3_221$
a (Å)	102.91	102.87	102.88
b (Å)	102.91	102.87	102.88
c (Å)	171.54	171.44	171.47
α = β (º)	90	90	90
γ(º)	120	120	120
Data collection (SSRL beamline 9-2, ALS beamline8.3.1)	SSRL	SSRL	SSRL
Temperature (K)	100	100	100
Detector	ADSCQ315	ADSCQ315	ADSCQ315
No. of crystals	1	1	1
Total rotation range (º)	240	240	240
Wavelength (Å)	0.918370	0.979128	0.979298
Resolution (Å)	40.0 - 2.30	40.0 - 2.30	40.0 - 2.30
High resolution shell (Å)	2.38 – 2.30	2.38 – 2.30	2.38 – 2.30
Total number of observations	661713	661575	663424
Unique reflections	90534 (9063)[a]	90425 (9066)	90475 (9052)
Multiplicity	7.3 (6.9)	7.3 (6.8)[a]	7.3 (6.9)
<I/σ(I)>	16.5 (3.8)	16.8 (3.8)	16.3 (3.8)
Completeness (%)	100 (100)	100 (100)	100 (100)
R_{sym} (%)[b]	11.9 (54.6)	12.0 (54.9)	11.9 (54.0)
Phasing			
Resolution (Å)		40.0 -2.30	
Number of Selenium		18	
Figure of Merit			
acentric		0.530	
centric		0.435	

[a] Values within parentheses refer to the highest resolution shell.

[b] $R_{sym} = \sum_h \sum_i(|I_i(\boldsymbol{h})-\langle I(\boldsymbol{h})\rangle|) / \sum_h \sum_i I_i(\boldsymbol{h})$, where $I_i(\boldsymbol{h})$ is the i^{th} intensity measurement and $\langle I(\boldsymbol{h})\rangle$ is the weighted mean of all measurements of $I_i(\boldsymbol{h})$.

The selenium coordinates were refined and initial 2.3 Å resolution MAD phases for AtDAP-AT were calculated with SHARP. Automated model building with ARP/wARP[16] built ~90% of protein. However, the N- and C-termini and some loop residues were placed incorrectly into the electron density. Therefore, ~25% of residues had to be placed into the electron density manually. All the manual fittings were accomplished using the XFIT program from the XtalView package[17]. The initial model was first refined with REFMAC5[18, 19]. This model was then refined against the 1.95 Å high resolution dataset. After several cycles of manual adjustment and refinement with REFMAC5, water molecules were picked with ARP/wARP solvent building. After the confirmation of the positions of water molecules, PLP, glycerol, and sulphate ion were manually placed into the positive density near to the active site. The final refinement with REFMAC5 resulted in R_{work} of 0.182 and R_{free} of 0.227. A total of five C-terminal histidine residues were observed in the asymmetric unit, four on subunit A and one on subunit B. The first 18 residues of the N-terminus were not observed in the final electron density map.

The L-malate-bound structure was solved using MOLREP[20] and the refined 1.95 Å native structure as the search model. The resultant structure was refined with REFMAC5, and water molecules were built with ARP/wARP solvent building. After the subsequent refinement, L-malate was manually fitted into the positive density in the active site. Final refinement with REFMAC5 resulted in R_{work} of 0.180 and R_{free} of 0.251. This model also lacks the first 18 residues of the N-terminus, and no His-tags were observed at the C-termini of either monomer.

Stereochemical analysis of both the native and the L-malate-bound structure was done by PROCHECK[21]. All the refinement statistics are presented in Table 2-2. PyMOL (http://www.pymol.Sourceforge.net) or ChemDraw were used to prepare all figures unless specifically noted.

Table 2-2 Refinement statistics

	Native DAP AT	
Crystal	PLP / SO_4 / Glycerol	PLP / Malate
Space group	$P3_221$	$P3_221$
a (Å)	102.91	102.59
b (Å)	102.91	102.59
c (Å)	171.45	172.94
α = β (º)	90	90
γ(º)	120	120
Data collection (SSRL beamline 9-2, ALS beamline8.3.1)	SSRL	ALS
Temperature (K)	100	100
Detector	ADSCQ350	ADSCQ210
No. of crystals	1	1
Total rotation range (º)	120	100
Wavelength (Å)	0.979462	1.115879
Resolution (Å)	20.0 - 1.95	30.0 – 2.40
High resolution shell (Å)	2.02 – 1.95	2.49 – 2.40
Total number of observations	449392	151459
Unique reflections	72271	37831
	(7204)[a]	(3838)
Multiplicity	6.2 (5.3)	4.0 (3.7)
<I/σ(I)>	15.4 (2.3)	13.5 (1.9)
Completeness (%)	93.6 (94.4)	90.4 (93.6)
R_{sym} (%)[b]	9.4 (66.4)	8.3 (55.3)
Refinement resolution (Å)	40.0 - 1.95	40.0 – 2.40
$R_{working}$ [c]	0.182	0.180
R_{free} [c]	0.227	0.251
Number of atoms		
Protein	6343	6288
Water	632	223
PLP	30	30
Sulfate	10	-
Glycerol	24	-
Selenium	18	-
Malic acid	-	9
Average *B*-factor (Å^2)	21.7	36.6

R.m.s. deviations from the ideal geometry		
Bond lengths (Å)	0.017	0.028
Bond angle (°)	1.679	2.374
Percentage of residues in Ramachandran plot		
In most favoured regions	92.7	90.5
In additional allowed regions	7.0	8.9
In generously allowed regions	0.3	0.4
In disallowed regions	0	0.1

[a] Values within parentheses refer to the highest resolution shell.

[b] $R_{sym} = \sum_h \sum_i(|I_i(\boldsymbol{h})-<I(\boldsymbol{h})>|) / \sum_h \sum_i I_i(\boldsymbol{h})$, where $I_i(\boldsymbol{h})$ is the i^{th} intensity measurement and $<I(\boldsymbol{h})>$ is the weighted mean of all measurements of $I_i(\boldsymbol{h})$.

[c] $R_{working}$ and $R_{free} = \sum_h ||F(\boldsymbol{h})_o| - |F(\boldsymbol{h})_c|| / \sum_h |F(\boldsymbol{h})_o|$, $|F(\boldsymbol{h})_o|$ and $|F(\boldsymbol{h})_c|$ are the observed and calculated structure factor amplitudes respectively. R_{free} was calculated using 5% of data.

Protein Data Bank accession number

The atomic coordinates and structure factors have been deposited with RCSB Protein Data Bank as entry PDB ID 2Z20 for native AtDAP-AT and 2Z1Z for L-malate-bound AtDAP-AT.

Genbank accession number

The optimized nucleotide sequence of AtDAP-AT used for protein expression in our study has been deposited with Genbank data library. Genbank accession number is EF543644.

2.3 Results and discussion

Sequence alignment

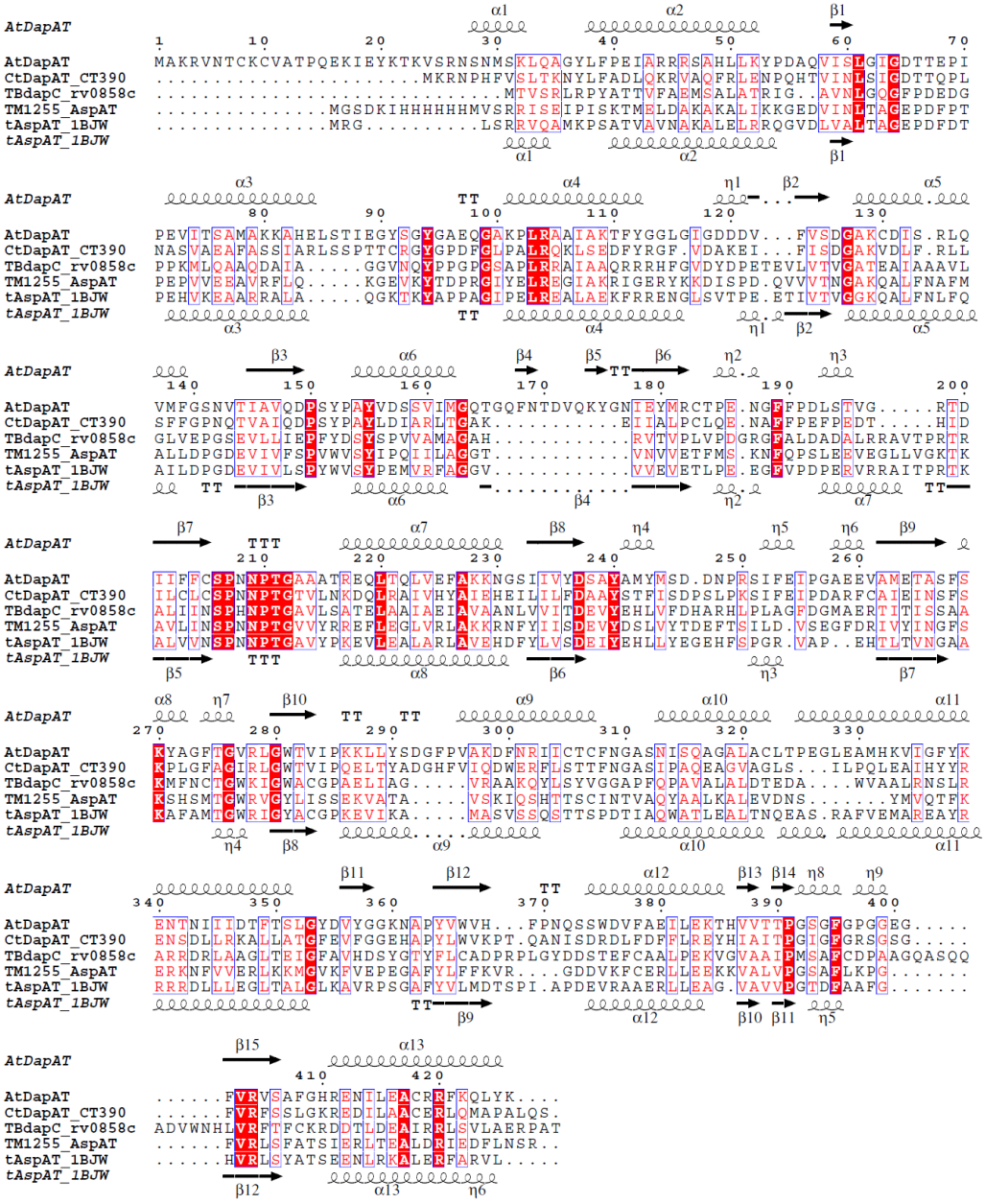
AtDAP-AT belongs to a family of PLP dependent aminotransferases. It is a novel enzyme that bypasses three of the enzymatic steps in the bacterial lysine biosynthesis pathway.[9] Any differences in its primary sequence and its tertiary structure from its relatives will provide some insights into AtDAP-AT activity and its relationship to the other aminotransferases.

In order to obtain the sequence identity with other aminotransferases, multiple sequence alignments with ClustalW[22] were conducted (Figure 2-2). Although the overall sequence identities between AtDAP-AT and the aspartate aminotransferases (AspATs) are quite low (<25%), the residues participating in the PLP dependent transamination reaction are well conserved. This result clearly indicates that AtDAP-AT, like other aminotransferases, depends on PLP for transamination. However, the low sequence similarities also suggest that AtDAP-AT might have evolutionally distinct origins from other AspATs.[9]

Recently, LL-DAP-AT has been characterized in three different species: *A. thaliana*, *C. trachomatis*, and *P. amoebophila*.[9, 11] Among these species, the pair-wise sequence identity approaches 40 - 50%.[11] However, the *N*-succinyl-DAP-AT (MtbDAP-AT, DapC)[23] present in the succinyl-DAP lysine biosynthesis pathway of *Mycobacterium tuberculosis* possesses only 17% sequence identity with AtDAP-AT. Since MtbDAP-AT is likely to catalyze a similar reaction that AtDAP-AT does (the difference being the presence of the distal *N*-acetyl group), it was surprising to see such a low sequence conservation between the two DAP-ATs. However, it should be noted that the biochemical function of this *M. tuberculosis* enzyme has not been rigorously demonstrated. As pointed out by Hudson, A. O. *et al.*[9], this sequence alignment clearly shows that LL-DAP-AT is distinct from DapC (MtbDAP-AT).

Figure 2-2 Sequence alignment of DAP-ATs with structurally similar aspartate aminotransferases. Sequence alignment of AtDAP-AT[9], CtDAP-AT[11], PcDAP-AT[11], MtbDAP-AT (2O0R)[23], TM1255 AspAT (1O4S)[24] and tAspAT (1BJW)[25] was carried out with ClustalW (http://www.ebi.ac.uk/clustalw/)[22]. TM1255 AspAT was chosen as the AspAT being a structurally known aminotransferase and having the highest sequence identity. tAspAT was chosen as the closest known structural homolog. Secondary structure elements of AtDAP-AT and tAspAT are shown on the top and at the bottom of the sequence alignment, respectively. All the secondary structure elements were determined with the program DSSP[26]. Identical residues are shown in white on red background, and similar residues are shown in red on white background. This figure was prepared with ESPript v2.2[27].

Figure 2-2 (continued)



Production and assay of Arabidopsis thaliana DAP-AT

AtDAP-AT with a C-terminal hexa-histidine tag was obtained by expression in *E. coli* using the vector pQE60. The gene encoding AtDAP-AT was optimized for bacterial codon usage, and the codons for the 36 amino acid plant transit peptide were replaced by a methionine start codon preceded by a ribosome binding site (Genbank EF543644). Standard protein purification typically gave 5 - 10 mg of active AtDAP-AT per litre of culture. Enzymatic activity of AtDAP-AT was confirmed by measuring the conversion of LL-DAP to THDP using a coupled assay with DAP dehydrogenase.[10] In the presence of α-ketoglutarate, AtDAP-AT converts LL-DAP to THDP, which is then transformed to *m*-DAP by DAP dehydrogenase with concomitant oxidation of NADPH. The decrease in absorbance at 340 nm, corresponding to the conversion of NADPH to $NADP^+$, was used to monitor the rate of reaction. Using various concentrations of LL-DAP, AtDAP-AT was found to have an activity of 3.3 μmol min^{-1} mg^{-1} protein and a *K*m value of 47 μM for LL-DAP, in reasonable agreement with recently reported values for both the plant and chlamydial enzymes.[9, 11]

Subunit structure

AtDAP-AT (M_r 46,650) crystallized in space group $P3_221$. The three dimensional structure of AtDAP-AT was determined using the MAD phasing method with a Se-Met derivative. An initial electron density map was computed at 2.3 Å resolution by SHARP[15], and the amino acid residues were fitted into the electron density by ARP/wARP[16] and XFIT[17]. The initial model was then refined against the high resolution data at 1.95 Å resolution with REFMAC5[18, 19] (Tables 2-1 & 2-2).

Two protomers of AtDAP-AT are present in the crystallographic asymmetric unit. They are related by a non-crystallographic 2-fold axis, and

comprise a homodimer (Figure 2-3). Many aminotransferases are known to exist as dimers because the aminotransferase active site is composed of residues from both subunits[28].

The native AtDAP-AT is composed of 426 amino acid residues, but the electron density belonging to the first 18 residues (Met1 - Ile18) of each subunit was not observed due to the disordered N-terminus. The electron density of the C-terminal His-tag was present on both subunit A (His_4) and subunit B (His_1). The root-mean-square deviation (r.m.s.d.) between the 409 C^{α} atom pairs (whole subunit) of the two subunits was calculated as 0.386 Å. The only significant (3σ) difference between the two subunits was a displacement of the α2 helices by 1.68 Å maximum displacement, and the r.m.s.d. of α2 helices (Tyr37 - Gln57) between two subunits was calculated as 1.09 Å. The r.m.s.d. between the 388 C^{α} atom pairs of the two subunits without the α2 helices was calculated as 0.304 Å. The C-terminus of helix α2 in subunit A is far from the molecule related by a crystallographic 3_2-screw axis (Figure 2-4a, the shortest distance between backbone carbon atoms is 8.95 Å). On the other hand, the C-terminus of helix α2 and the proceeding loop in subunit B is much closer (~5 Å) to the 3_{10} helix (η3) of the molecule related by a crystallographic 3_2-screw axis (Figure 2-4b). This interaction displaces the C-terminus of helix α2 in subunit B towards the active site cavity compared to the residues of helix α2 in subunit A. Therefore, the slight displacement of helix α2 caused by the crystal packing has contributed to the relatively large r.m.s.d. between the two subunits in the AtDAP-AT.

The domains and the overall fold of AtDAP-AT were analyzed by the CATHEDRAL server[29]. The analysis showed that the monomer of AtDAP-AT consists of two domains; a large domain (LD) and a small domain (SD) (Figure 2-5). The LD (Ile70 - Lys333) belongs to an alpha-beta class, which folds into an α-β-α sandwich. Seven of the nine β-strands (β2-β10) in the LD form helically

Figure 2-3 Overall structure of AtDAP-AT homodimer in cartoon representation. Subunit A and B are coloured blue and yellow, respectively. PLP, sulphate ions and glycerol molecules bound to each subunit are represented as sticks. The two monomers are related by a non-crystallographic two-fold axis.

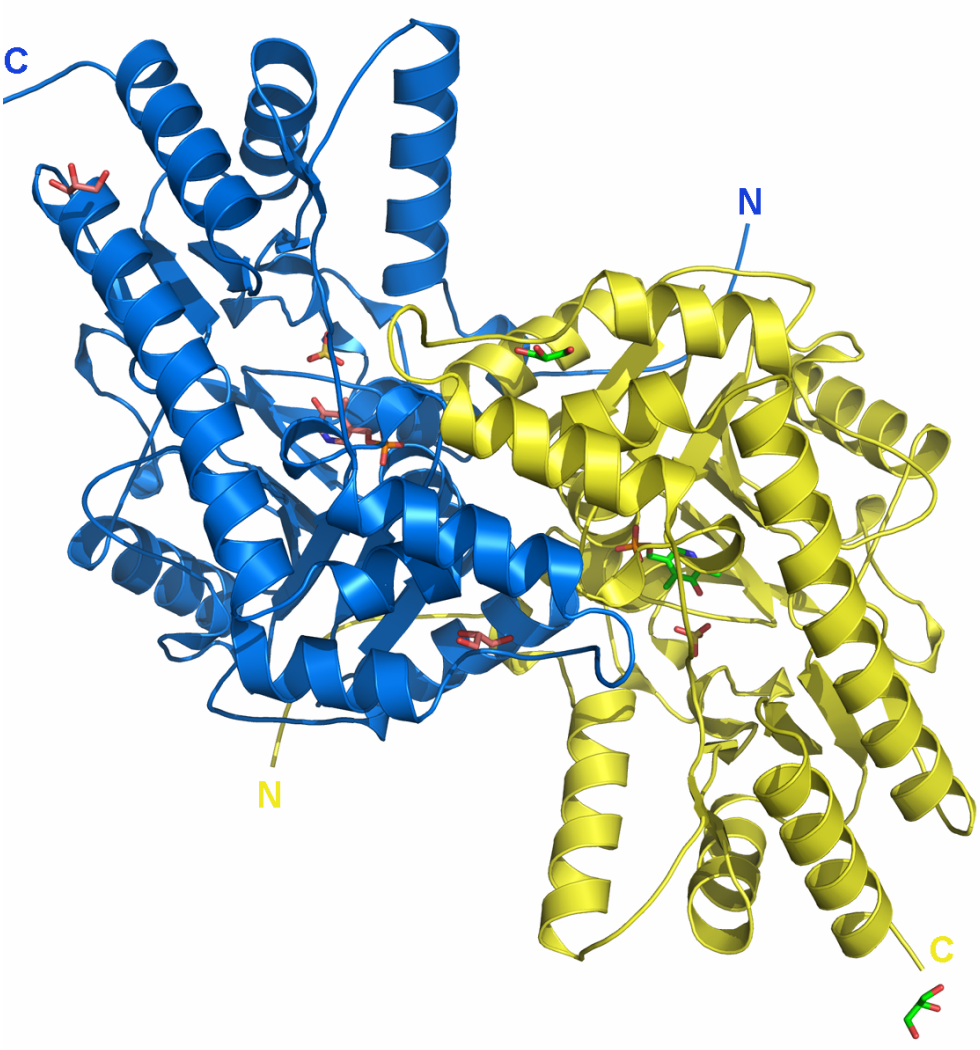
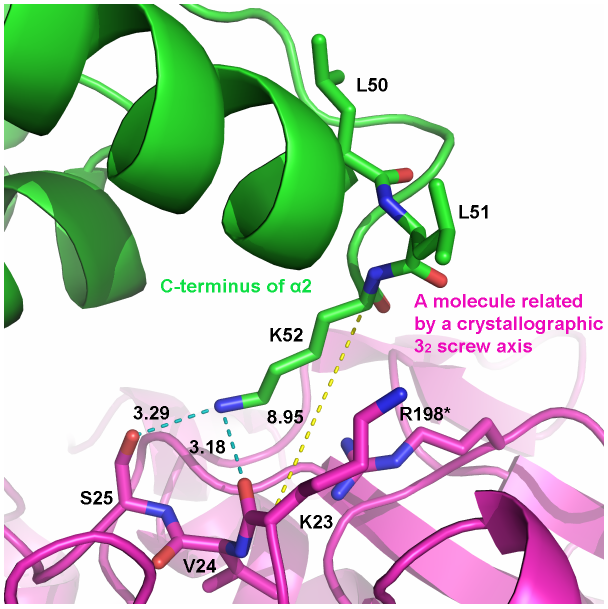
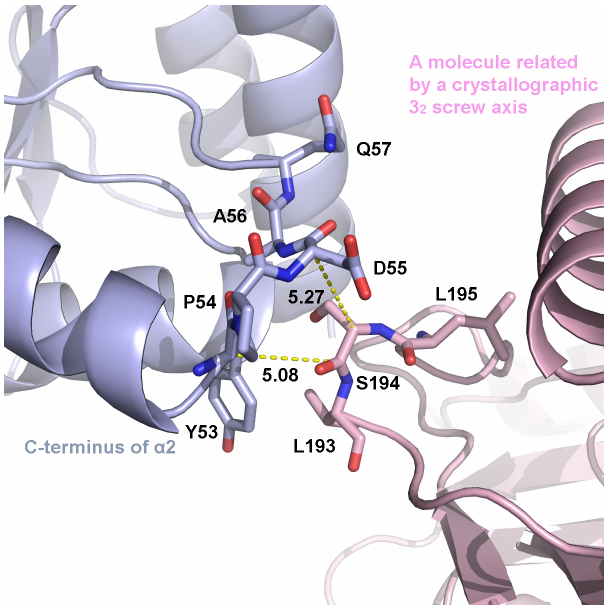


Figure 2-4 Interaction of helix α2 with the molecule in the neighbouring asymmetric unit. Blue dotted lines indicate hydrogen bonds, and the yellow dotted line shows the intermolecular distance between main chain carbon atoms. Distance is in Å. a). Subunit A (green), two-fold related molecule (magenta), b). Subunit B (light grey), two-fold related molecule (pink).

a).



b).



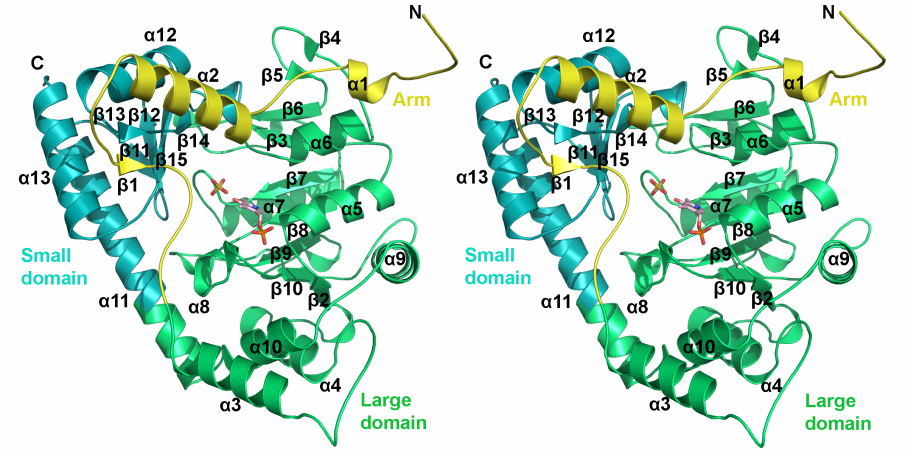
twisted β-sheets that comprise the core of LD. Other than the β10 strand, these strands are all parallel. Helix α6 and a loop with a short antiparallel β-sheet (β4, β5) connect β3 to β6. The core seven-stranded β-sheet is surrounded by a total of eight α-helices (α3 - α10) and seven short 3_{10}-helices (η1 - η7). Three of the α-helices (α5, α6 and α9) reside on the PLP-binding side of the seven-stranded β-sheet; the other helices including the 3_{10}-helices are on the opposite side of the seven-stranded β-sheet.

The SD (Glu18 - Pro69 plus Val334 - Lys426) also belongs to the alpha-beta class of protein fold and forms an α-β complex. The core of the SD is composed of a four-stranded antiparallel β-sheet (β11, β12, β14 and β15), which is surrounded by three α-helices (α11 - α13) and two 3_{10}-helices (η8 and η9). The SD also contains an "arm" region (α1, α2 and β1) at the N-terminus. The arm interacts with the rest of the SD mainly by forming a parallel β-sheet interaction between strands β1 and β13.

Comparison with a homologous structure (tAspAT)

The overall fold of AtDAP-AT showed that it is closely related to that of the AspAT family. Among the five different fold types present in the PLP dependent enzymes, the AtDAP-AT belongs to the type I fold.[30, 31] The DALI server[32] determined that the AspAT from *Thermus thermophilus* HB8 (tAspAT, PDB ID: 1BJW)[25] was structurally the most similar to AtDAP-AT. tAspAT also exists as a biologically active homodimer. The overall fold and the secondary structural elements between AtDAP-AT and tAspAT are almost identical. However, the r.m.s.d. between AtDAP-AT and tAspAT was quite large at 1.87 Å for 358 C^{α} atom pairs; this is certainly why the AtDAP-AT structure determination required the MAD phasing method instead of molecular replacement using tAspAT as a search model.

Figure 2-5 Stereodiagram of PLP-bound subunit structure of AtDAP-AT in cartoon representation. PLP is represented in stick. The large domain, the small domain, and the arm regions are coloured blue, green, and yellow, respectively. All α-helices and β-strands are labelled accordingly.



There are major structural differences between the two enzymes; these are shown in Figure 2-6a. The first one is the 11 (or 12)-residue insertion comprising the two-stranded antiparallel β-sheet (β4 and β5). Based on the sequence alignment among the several DAP-ATs (Figure 2-2), this extra loop with the two-stranded antiparallel β-sheet is only present in *A. thaliana;* it is not in *Chlamydia sp.* Therefore, it may not be required for LL-DAP-AT activity.

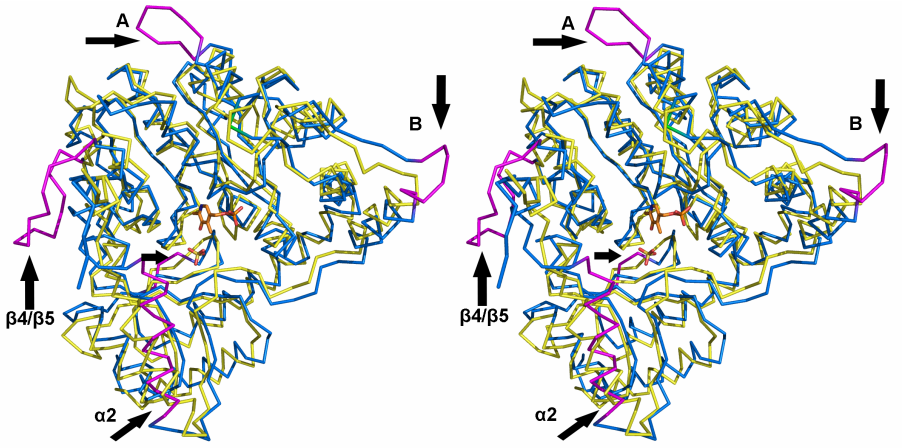
The second difference is the β11 strand located in the core of the SD. The core of tAspAT SD contains a three-stranded β-sheet, whereas that of AtDAP-AT SD contains a four-stranded β-sheet with β11 being the additional strand. Although the amino acid sequence of β11 is reasonably conserved among different LL-DAP-ATs, the effect of the extra β-strand in the SD of AtDAP-AT on the catalytic activity is uncertain at this point.

Loop A and B indicated in Figure 2-6a are considerably longer in AtDAP-AT than those of tAspAT. Both loops are involved in dimer formation in AtDAP-AT, but only the residues corresponding to AtDAP-AT LoopB in tAspAT are involved in dimer formation. LoopA in AtDAP-AT is involved in dimer formation by interacting with the N-terminal arm region of the crystallographic symmetry related subunit.

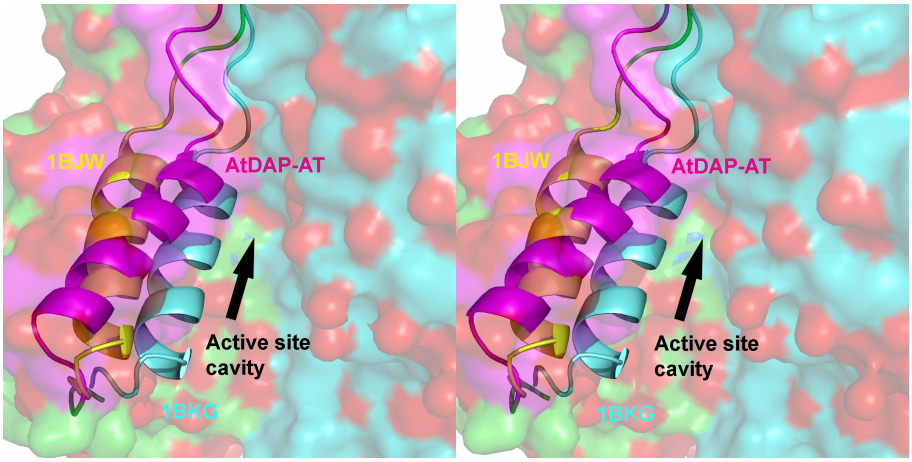
Lastly, the position of helix α2 is different in AtDAP-AT and tAspAT. Figure 2-6b shows the conformations of the α2-helices from AtDAP-AT, tAspAT[25] and the maleate-bound tAspAT (PDB ID: 1BKG)[25]. In tAspAT, a movement of helix α2 was observed upon binding the maleate anion. It is obvious that helix α2 in AtDAP-AT does not fit into either one of the conformations of the tAspAT α2-helices. Also, there were no conformational differences observed upon binding of malate to AtDAP-AT, which will be discussed in a following section. Therefore, the difference in helix α2 conformation can not be correlated with a substrate-induced conformational change at this point, and further study will be required.

Figure 2-6 a). Superposition of AtDAP-AT and tAspAT main chain α-carbon atom structures in stereoview. AtDAP-AT and tAspAT are coloured in blue and yellow, respectively. The five major differences observed between AtDAP-AT and tAspAT are coloured in magenta, and indicated by an arrow. The two structures are aligned with ALIGN[33]. β11 strands are located in the middle and of the molecule and represented by the short arrow. b). Superimposed α2-helices from AtDAP-AT (magenta, 2Z20), tAspAT (yellow, 1BJW) and maleate-bound tAspAT (sky blue, 1BKG).

a).



b).



Despite the low sequence identity between AtDAP-AT and tAspAT (21%), a remarkable similarity in three-dimensional structure was observed; this observation emphasizes the structural conservation among aspartate aminotransferases having the type I fold.

Dimer interface

Type I fold aminotransferases are known to be active as dimers because the active site of each subunit is composed of residues from both subunits.[30] AtDAP-AT is no exception. The two subunits of AtDAP-AT interact tightly to form two active sites. Details of the solvent accessible surface area and buried surface area of the AtDAP-AT dimer are summarized in Table 2-3. The buried surface area of AtDAP-AT was calculated using NACCESS (Hubbard, S. J. & Thirnton, J. M. (1993). NACCESS. 2.1.1. University College of London, Department of Biochemistry and Molecular Biology.) and the PISA server[34]. Comparison of the buried surface area with the total solvent accessible area of the AtDAP-AT dimer reveals that about 21% of the total solvent accessible surface area is involved in the dimer interaction. A total of 65% of the residues in the buried surface area are non-polar, whereas 35% of the residues in the buried surface area are polar. This indicates that the dimer is mainly held together by hydrophobic interactions. Although hydrophobic interactions are the main associative force at the dimer interface, extensive electrostatic interactions account for more than one third of the interfacial interactions. A total of 54 hydrogen bonds and 15 salt bridges are formed between the two subunits, and these interactions are mostly two-fold symmetrically identical between the two subunits.

Tight interactions of the AtDAP-AT dimer can also be appreciated from the protrusion of the N-terminal arm region (Glu18 - Pro40) into the surface of the

Table 2-3 Summary of solvent accessible area and buried surface area of AtDAP-AT

Solvent accessible area	
Homodimer (Å^2)	28518.3
Subunit A (Å^2)	18036.8
Subunit B (Å^2)	18059.0
Dimer interface	
Buried surface area (Å^2)	7577.5
Average buried surface area / protomer (Å^2)	3788.7
Average polar buried surface area (Å^2)	2637.0 (34.8%)
Average non-polar buried surface area (Å^2)	4940.5 (65.2%)
Residues present at the dimer interface	25% of the total number of residues (300 residues)

other subunit of the dimer. The N-terminal arm makes a total of 14 hydrogen-bond interactions and two salt bridges with the neighbour. Of the total number of hydrogen bonds that exist at the interface, more than a quarter of the interfacial hydrogen bonds are contributed by this N-terminal arm. Similar observations were made with histidinol-phosphate aminotransferase (HisP-AT) dimeric structure[35], however not to as large an extent as in AtDAP-AT. Therefore, the N-terminal arm also contributes significantly towards dimer stability. Another significant interfacial interaction can be seen at the α3 helices (Glu72 - Glu84) related by the non-crystallographic 2-fold axis of symmetry. The α3 helices from both subunits interact with each other in an antiparallel way. The association between these two helices is provided mainly by hydrophobic forces.

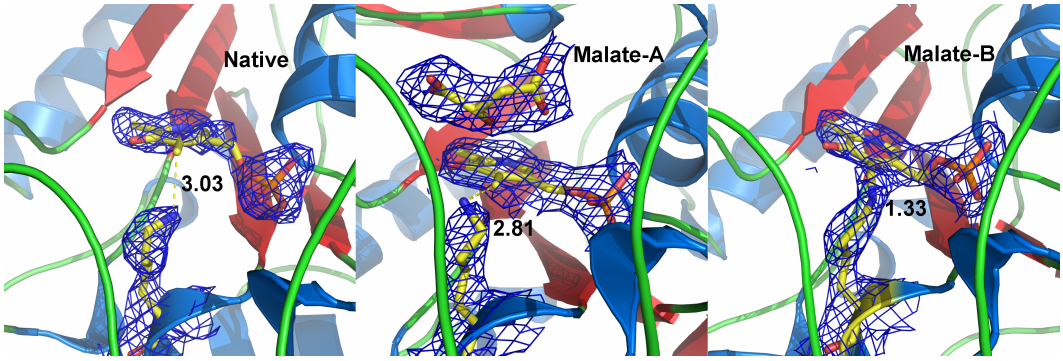
The large buried surface area of dimer is conserved in many type I fold aminotransferases[28, 25]; AtDAP-AT is not an exception. Our three-dimensional structure of AtDAP-AT emphasizes that the type I fold aminotransferases form a biologically stable dimer through the large dimer interface and the protrusion of N-terminal arm region.

PLP-binding site

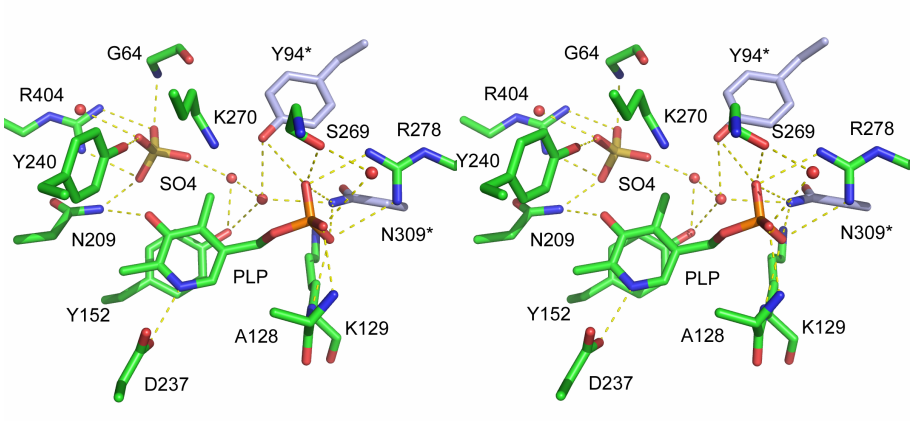
A sequence alignment of several aminotransferases (Figure 2-2) shows that AtDAP-AT contains the residues that are conserved among the PLP dependent enzymes. Also, AtDAP-AT in aqueous solution exhibited the pale yellow color characteristic of PLP-bound enzymes. Therefore, an internal aldimine was expected to be present at the active site Lys270. However, there was no electron density for the covalent bond between Lys270N^{ζ} and C4’ of PLP present in the electron density map after several cycles of refinement, even though an excellent electron density for PLP was observed at the active site (Figure 2-7a).

Figure 2-7 a). 2|F_o|-|F_c| electron density map (contour level = 1σ) for PLP, Lys270 and malate from the native (Native) and malate-bound (Malate-A) AtDAP-AT. Internal aldimine observed in malate-bound structure is also shown (Malate-B). The numbers beside Lys270 are the distance between Lys270N$^{\zeta}$ and C4' of PLP. b).PLP- and sulphate ion-bound active site of AtDAP-AT. Active site residues, PLP and sulphate ion are represented in stick. Residues from subunit A are coloured in green, and B are in grey. PLP is coloured in same colour as the bound subunit. Water molecules are represented as small red spheres. Hydrogen bonds are shown as dotted lines and coloured yellow. The cut off for hydrogen bonds is 3.6 Å.

a).



b).



Radiation damage could explain the loss of the covalent bond in the Schiff base. It has been shown that a high dosage of radiation could damage a Schiff base bond between LysN^{ζ} and PLP.[36] During the data collection for AtDAP-AT at the synchrotron beamline, a single crystal was used to obtain a complete three-wavelength MAD data set (with the inverse beam strategy) and a high resolution native data set (estimated total exposure time = 8640 s). Therefore, the high dose of radiation likely has disrupted the covalent bond between Lys270N^{ζ} and PLP-C4'.

A data set from a crystal of malate-bound AtDAP-AT was collected on a separate occasion using far less exposure time (estimated total exposure time = 1245 s). In this structure, the subunit that is not bound by malate exhibited electron density corresponding to the covalent aldimine linkage; however the other subunit showed the disrupted covalent bond (Figure 2-7a). A similar result was also observed in the crystal structure of glutamate-1-semialdehyde aminomutase (GSA).[37] In the GSA structure, only one of the two subunits showed the electron density for the covalently bound lysine and PLP.[37] They suggested that the subunit without the covalent bond would contain pyridoxamine 5'-phosphate (PMP) instead of PLP. In their PMP-bound structure of GSA, the active site lysine is pointing away from PMP and about 2.7 Å apart from PMP.[37] This observation in GSA closely resembles our malate-bound subunit of AtDAP-AT structure (Figure 2-7a, 2.81 Å apart). Therefore, the disruption of the covalent linkage in the malate-bound structure might not be caused by radiation damage, but it could indicate the presence of PMP instead of radiation damage.

In addition, the electron density for PLP and Lys270 in the malate-bound structure is quite well ordered compared to that of the native structure. Despite the fact that both structures contain non-covalently bound PLP, one possible reason for the better ordered density of the malate-bound structure may be the binding of

malate. It has been shown that non-covalently bound PLP in the active site tends to be disordered and does not show clear electron density even at high resolution (our unpublished data)[38]. Therefore, the malate binding in the active site would likely order the residues and non-covalently bound PLP in the active site.

Although a covalent interaction between Lys270 and PLP in the high resolution structure was not observed, PLP was clearly bound to the active site of AtDAP-AT, and the ε-amino group of Lys270 was pointing towards PLP as expected from other aminotransferase structures[23, 25, 24] (Figure 2-7a).

Figure 2-7b portrays a structural diagram of PLP binding in the AtDAP-AT active site. The pyridine ring of PLP is stabilized by an aromatic ring stacking interaction with Tyr152. The O3' of PLP is stabilized by hydrogen bonds to Tyr240, Asn209 and a bound sulphate ion. The side chain of Asp237 is situated so that it can accept a hydrogen bond from the positively charged nitrogen (N1) in the pyridine ring. The phosphate moiety of PLP is mainly stabilized by the positively charged residues and a helix dipole from helix α5. The phosphate oxygen forms a salt bridge and two hydrogen bonds with the Arg278 guanidinium group. Also, the N-terminus of helix α5 is positioned so that the positively charged helix dipole is accepting the negatively charged phosphate group. In addition, the phosphate group also interacts with a water molecule, Ala128N, Lys129N, Ser269O^{γ}, Tyr94*O^{η} and Asn309*$N^{\delta 2}$ through hydrogen bonds. The sulphate ion and the Asn309*$O^{\delta 1}$ are bridged by two water molecules. According to the sequence alignment and the active site structures, all of the PLP-binding residues are well conserved between AtDAP-AT and tAspAT but with an additional interaction with Asn309* on AtDAP-AT. As observed in the other type I fold aminotransferases, the PLP-binding site requires residues from both subunits and extensive interactions to keep PLP in the active site.[23, 25, 24]

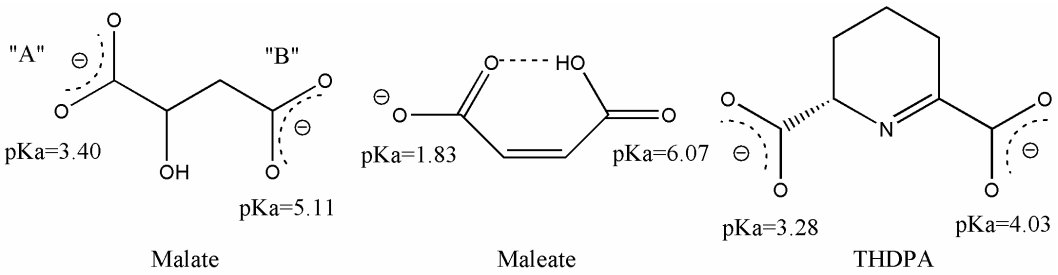
Malate binding

Our AtDAP-AT crystals were crystallized under two different conditions. The high-resolution data were from a crystal crystallized in ammonium sulphate, whereas the moderately diffracting crystal was obtained with racemic malic acid (malate at pH 7.0, Figure 2-8a). In the crystal obtained from racemic malic acid, only the L-malate anion was bound to one of the AtDAP-AT subunits. It would be expected to observe only the L-isomer of malate in the active site because the natural substrates (LL-DAP or L-Glu) also have the L-configuration. Substrate-bound structures have been studied with several of the AspATs. Some of these complex structures with bound maleate (Figure 2-8a), a similar compound to malate, have been reported.[25, 39-41] As maleate was shown to mimic the substrate binding in AspAT, malate-bound structures may provide useful insights into substrate binding in the active site of AtDAP-AT.

A stereo representation of the malate-bound active site of AtDAP-AT is shown in Figure 2-8b. The carboxylate ("carboxylate-A") corresponding to the α-carboxylate of the natural substrate (i.e. L-Glu) is stabilized by a salt bridge with the guanidinium group of Arg404 and two hydrogen bonds with Asn209$N^{\delta 2}$ and main chain Gly64N. In most of the well-studied AspATs, recognition of carboxylate-A seems to be conserved. The combination of a salt bridge from a guanidinium group of Arg404 and hydrogen bonds from the neighbouring polar amino acids is employed by most AspATs in order to stabilize carboxylate-A. The second carboxylate ("carboxylate-B") corresponding to the side chain carboxylate of the natural substrates (i.e. aspartate or glutamate) is stabilized by a number of direct and water-mediated hydrogen bonds with Lys129N^{ζ}, Glu97$O^{\varepsilon 2}$ and Tyr152O^{η}. Interaction with the substrate is also mediated by the side chain of Ile63 which provides a hydrophobic environment for the malate methylene carbon atoms. The OH group of malate, which may mimic the amino donor (NH_3^+) of the

Figure 2-8 Stereodiagram of malate-bound active site of AtDAP-AT. Active site residues, PLP and malate are represented in stick. Residues from subunit A, B and malate is coloured in green, grey and sky blue, respectively. Water molecules are represented in small red sphere. Hydrogen bonds are shown as dotted lines and coloured yellow. The cut-off for hydrogen bonds is 3.6 Å. a). Chemical structures of malate, maleate and THDP at pH 7.0. The structures were drawn with ChemDraw. b). Stereodiagram of malate-bound active site of AtDAP-AT.

a).



natural substrates, is pointing towards the C4' of PLP. It is stabilized by a hydrogen bond with Tyr152O$^{\eta}$.

The closest structural homolog of AtDAP-AT (tAspAT) belongs to subclass Ib of the AspATs.[25] Subclass Ia and Ib differ in their degree of substrate-induced conformational change and in their substrate recognition residues.[25] In subclass Ib (tAspAT), a small movement of the small domain towards the active site was observed.[25] In contrast, subclass Ia enzymes show a large movement of the small domain towards the active site.[42, 43] In terms of the conformational changes, no significant change was observed upon binding of malate to AtDAP-AT (r.m.s.d. between AtDAP-AT-malate complex and tAspAt (1BKG) = 1.83 Å for 355 C$^{\alpha}$ pairs). However, some unique aspects of AtDAP-AT were observed in its recognition of the carboxylate-B of malate. In the subclass Ib, carboxylate-B is directly recognized by the ε-amino group of a Lys residue (Lys101 in tAspAT)[25] whereas in subclass Ia, it is recognized by the guanidinium group of an Arg residue[42, 43]. In the AtDAP-AT, Lys129 seems to correspond to Lys101 of tAspAT; there are no Arg residues found in the vicinity of the carboxylate-B in AtDAP-AT. Therefore, in terms of residue similarity and sequence conservation, AtDAP-AT seems more closely related to subclass Ib of the AspAT family than to subclass Ia.

Although the position of the Lys residue is conserved, Lys129 of AtDAP-AT does not make a direct interaction with malate carboxylate-B. Instead, it interacts with carboxylate-B through a water-mediated hydrogen bond. Moreover, further comparison of the active sites of AtDAP-AT and tAspAT revealed that the side chain of Asn309* and Glu97* of AtDAP-AT are positioned for possible substrate recognition. Although Lys129, Glu97* and Asn309* are too far among to interact directly with carboxylate-B, they form an array of hydrogen bonds in the active site and are positioned likely for the longer methylene carbon

substrates like LL-DAP. The hydrogen bonding network formed by Lys129, Asn309* and Glu97* also stabilizes each other for substrate recognition. Interestingly, Asn and the negatively charged amino acid (Glu or Asp) at those positions are conserved and only present in the LL-DAP-AT family. (In Figure 2-2, Asn of TM1255 and Asp of tAspAT are aligned with Asn309* of AtDAP-AT but Asn of TM1255 is located about five residues away from the position of Asn309* in three-dimensional space.) The above observations suggest that Lys129, Asn309* and Glu97* are likely to be important for substrate recognition in LL-DAP-AT.

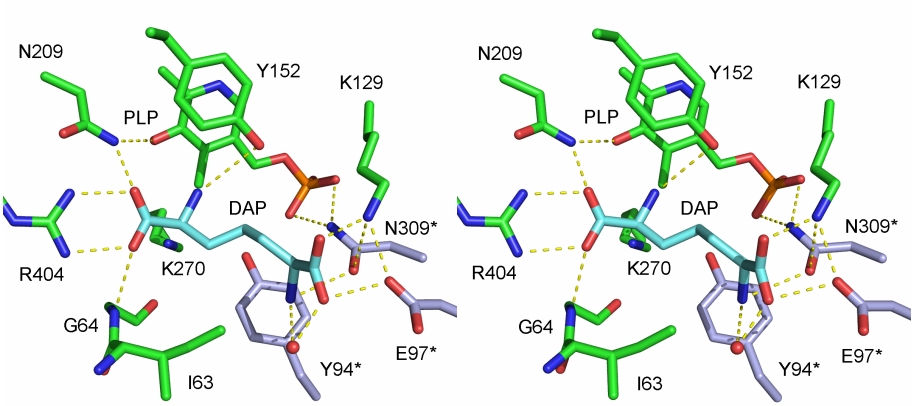
Prediction of LL-DAP and L-Glu binding

Based on the malate-bound structure of AtDAP-AT, LL-DAP and L-Glu were modelled into the active site to explore whether the positions of residues Asn309*, Glu97* and Lys129 can accommodate LL-DAP and L-Glu. The orientations of the predicted LL-DAP binding in the active site of AtDAP-AT are shown in Figure 2-9a. Upon LL-DAP binding, a water molecule that was bridging between Lys129N^{ζ} and the malate's carboxylate-B is now displaced by the LL-DAP's carboxylate. Glu97*$O^{\varepsilon 2}$**,** and Lys129N^{ζ} now can make direct hydrogen bonds with the LL-DAP carboxylate oxygen atoms. Furthermore, the chiral, primary amine nitrogen can be positioned in between Asn309*$O^{\delta 1}$ and a water molecule. It is stabilized well by hydrogen bonds. Similar stereospecific recognition of two inhibitors by DAP epimerase has been observed.[4] Based on the predicted position of LL-DAP bound in the active site, Asn309*, Glu97*, and Lys129 can provide extensive hydrogen bonds for stereospecific recognition of LL-DAP in the active site.

In the same manner as LL-DAP, L-Glu was modelled into the active site of AtDAP-AT (Figure 2-9b). L-Glu is one methylene carbon atom longer than

Figure 2-9 Stereodiagram of AtDAP-AT active site with manually docked a) LL-DAP, b) L-Glu and c) THDP. LL-DAP / L-Glu were modelled into the active site based on the structure of malate binding in the active site of AtDAP-AT. Residues from subunit A, B and LL-DAP / L-Glu are coloured in green, grey and sky blue, respectively. Two different orientations of THDP are coloured yellow and sky blue. Water molecules are represented as small red spheres. Hydrogen bonds are shown as dotted lines and coloured yellow. The cut-off for hydrogen bonds is 3.6Å

a).



b).

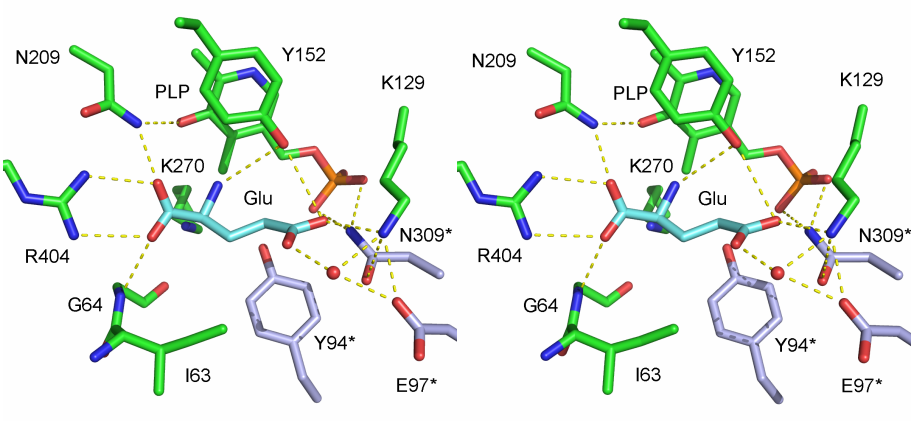
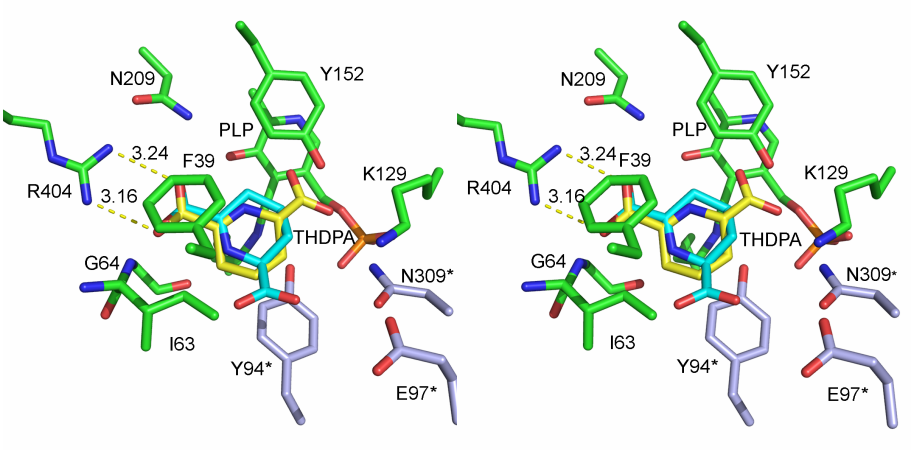


Figure 2-9 (continued)

c).



malate. As it was observed in LL-DAP binding, one of the side-chain oxygen atoms of L-Glu is in the appropriate position for a direct hydrogen bond to Lys129, and additionally with Tyr152. The additional hydrogen bond with Tyr152 can be formed because L-Glu is two methylene carbon atoms shorter than LL-DAP. Having this side chain length in L-Glu, the other side chain oxygen atom of L-Glu is not properly positioned to displace the water molecule for a direct hydrogen bond with Glu97*$O^{\varepsilon 2}$. Therefore, the L-Glu side-chain oxygen atom will likely make a water-mediated hydrogen bond to that water molecule which is slightly shifted toward Glu97*$O^{\varepsilon 2}$. We have also attempted to model THDP into the active site of AtDAP-AT shown in Figure 2-9c. However, the steric clashes at the chiral carboxyl group could not be overcome in either conformation of THDP shown in Figure 2-9c. The active site cavity of the present AtDAP-AT structure seems to be best for accommodating acyclic, elongated molecules, and not for cyclic molecules like THDP. Therefore, we would expect a conformational change in order to accommodate THDP as substrate.

In addition, it has been shown that CtDAP-AT (*Chlamydia trachomatis* LL-DAP-AT) and PcDAP-AT (*Protochlamydia amoebophila* LL-DAP-AT) have broader substrate specificities than AtDAP-AT does despite the fact that the active site residues are highly conserved.[11] Therefore, Asn309*, Glu97* and Lys129 residues are important and may also participate in the substrate binding. However, they are not likely to be responsible for the entire substrate recognition mechanism.

Normally, AspATs possess dual specificity in order to acquire and transfer the amino group from two different substrates. The dual specificity would be sufficient for the specific transamination reaction if the shape of the substrates is similar[30]; this is accomplished mainly by the arginine-switch or by the rearrangement of hydrogen bonds.[30] As THDP in aqueous solution is in rapid

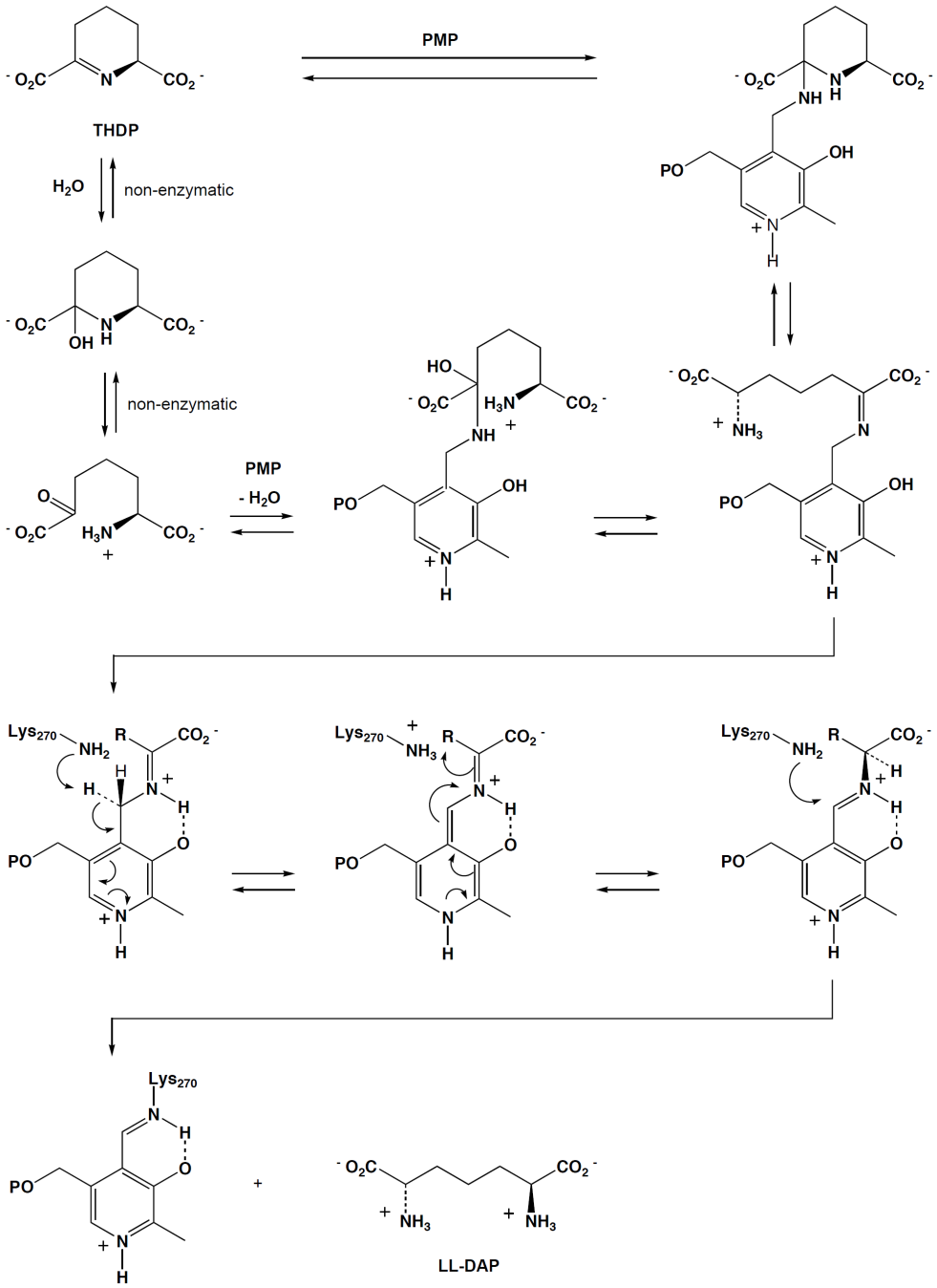
equilibrium with its open (hydrolyzed) form, namely L-α-keto-ε-aminopimelate,[10] it may be that the enzyme accepts only the non-cyclic substrate. However, if AtDAP-AT binds the cyclic THDP directly, it would be required to accommodate at least two substrates that are considerably different in their shape (i.e. THDP and LL-DAP). Moreover, the opening of THDP ring would also have to be catalyzed by AtDAP-AT. Such enzyme-catalyzed cleavage of the THDP ring could involve either hydrolysis in the active site or direct attack by the amino group of PMP.

Our modeling studies show that elongated substrates like LL-DAP and L-Glu can be accommodated in the active site of AtDAP-AT. It is accomplished by the rearrangement of hydrogen-bonds among the substrates, active site residues (Asn309*, Glu97*, Lys129) and water molecules. However, our present structure of AtDAP-AT indicates that the rearrangement of hydrogen bonds is not sufficient for the accommodation of THDP or for the catalysis of THDP ring opening. If unhydrolyzed THDP is the true substrate, significant conformational rearrangement of protein domains would be necessary to fit the molecule into the active site. In this case, LL-DAP-AT would possess a more complex catalytic mechanism than other members of its aminotransferase family such as tAspAT.

Proposed catalytic mechanism of LL-DAP-AT

An unusual feature of the LL-DAP-AT compared to other enzymes in the main bacterial lysine biosynthetic pathways (Figure 2-1) is its ability to produce LL-DAP directly from THDP in a PLP dependent process without prior *N*-acetylation to keep the substrate in a non-cyclic α-keto form. Consideration of the possible catalytic mechanism of LL-DAP-AT (Figure 2-10) suggests that the first half of the reaction produces PMP from L-Glu and the PLP-Lys270 internal aldimine in the same fashion as AspAT (e.g. tAspAT). Once α-ketoglutarate leaves

Figure 2-10 Proposed catalytic mechanism of AtDAP-AT



the active site, LL-DAP-AT is ready for the formation of LL-DAP from THDP. As mentioned above, rapid and reversible non-enzymatic hydrolysis[10] of THDP will form L-α-keto-ε-aminopimelate, which would be a "normal" substrate for the PMP form of an aminotransferase. Although THDP is also the substrate for THDP *N*-succinyltransferase (DapD), attempts to produce THDP-DapD complex by soaking have not given the cyclic compound in the active site.[44] Instead, the linear L-α-keto-ε-aminopimelate was seen bound to the enzyme.[44] It is presently not clear whether the direct binding and ring opening of THDP can be catalyzed by AtDAP-AT. If so, this would probably proceed via nucleophilic attack on the imine bond of THDP by the 4'-amino group of PMP in a transimination process, and would need to be preceded by conformational reorganization of the active site to accommodate the cyclic substrate. Alternatively, if L-α-keto-ε-aminopimelate is the substrate, Asn309*O$^{\delta 1}$ and the water molecule would recognize the primary amine at the C2 position of L-α-keto-ε-aminopimelate stereospecifically as shown in Figure 2-9a (LL-DAP model). PMP in the active site would then attack C6 of this substrate to form the ketimine intermediate. Deprotonated Lys270N$^{\zeta}$ would then abstract a proton from the ketimine and form the quinonoid intermediate. Tautomerization via abstraction of the proton from protonated Lys270N$^{\zeta}$ would form the aldimine. Lastly, transimination by Lys270N$^{\zeta}$ will displace the amino group of LL-DAP thus forming the enzyme Schiff base with PLP again to regenerate the initial PLP-Lys270 internal aldimine. On-going studies with enzyme-inhibitor complexes may clarify whether THDP can be used directly by AtDAP-AT without prior hydrolysis to the linear L-α-keto-ε-aminopimelate.

Concluding remarks

In this chapter, the crystal structure of AtDAP-AT from *A. thaliana* to 1.95 Å resolution has been reported. The structure of AtDAP-AT shares many

similarities with other type I fold aminotransferases, confirming the structural conservation among the type I fold family of PLP dependent enzymes. Finding structural conservation with other aminotransferases was surprising because plant and *Chlamydia* LL-DAP-AT possess low sequence identity with other aminotransferases. At the same time, several unique features of LL-DAP-AT were also identified by examining the malate-bound AtDAP-AT structure in comparison with several known aminotransferase structures. The malate-bound AtDAP-AT structure revealed the possible residues required for substrate recognition. These structural studies form the initial step in understanding the catalytic mechanism of LL-DAP-AT and its application in the development of new herbicides for plants or novel antibiotics for *Chlamydia* species.

2.4 References

1. Vederas, J. C. (2006). Diaminopimelate and lysine biosynthesis - An antimicrobial target in bacteria. *Canad. J. Chem.* **84**, 1197-1207.
2. Cox, R. J., Sutherland, A., and Vederas, J. C. (2000). Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorg. Med. Chem*. **8**, 843-871.
3. Scapin, G., and Blanchard, J. S. (1998). Enzymology of bacterial lysine biosynthesis. *Adv. Enzymol. Relat. Areas Mol. Biol*. **72**, 279-324.
4. Pillai, B., Cherney, M. M., Diaper, C. M., Sutherland, A., Blanchard, J. S., Vederas, J. C., and James, M. N. (2006). Structural insights into stereochemical inversion by diaminopimelate epimerase: An antibacterial drug target. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8668.
5. van Heijenoort, J. (2001). Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod. Rep*. **18**, 503-519.
6. Kelland, J. G., Palcic, M. M., Pickard, M. A., and Vederas, J. C. (1985). Stereochemistry of lysine formation by meso-diaminopimelate decarboxylase from wheat germ: use of ^{1}H-^{13}C NMR shift correlation to detect stereospecific deuterium labeling. *Biochemistry*. **24**, 3263-3267.
7. Shaw-Reid, C. A., McCormick, M. M., Sinskey, A. J., and Stephanopoulos, G. (1999). Flux through the tetrahydrodipicolinate succinylase pathway is dispensable for L-lysine production in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol*. **51**, 325-333.
8. Mazur, B., Krebbers, E., and Tingey, S. (1999). Gene discovery and product development for grain quality traits. *Science*. **285**, 372-375.
9. Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol*. **140**, 292-301.

10. Caplan, J. F., Sutherland, A., and Vederas, J. C. (2001). The first stereospecific synthesis of l-tetrahydrodipicolinic acid; a key intermediate of diaminopimelate metabolism. *J. Chem. Soc., Perkin Trans. 1*. **18**, 2217-2220.
11. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T., and Maurelli, A. T. (2006). L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17909-17914.
12. Doublié, S. (1997). Preparation of selenomethionyl proteins for phase determination. *Meth. Enzymol*. **276**, 523-530.
13. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol*. **229**, 105-124.
14. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307-326.
15. de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol*. **276**, 472-494.
16. Perrakis, A., Morris, R., and Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol*. **6**, 458-463.
17. McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol*. **125**, 156-165.
18. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr*. **50**, 760-763.
19. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr*. **55**, 247-255.

20. Vagin, A., and Teplyakov, A. (1997). MOLREP : an Automated Program for Molecular Replacement. *J. Appl. Crystallogr*. **30**, 1022-1025.

21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*. **26**, 283-291.

22. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**, 4673-4680.

23. Weyand, S., Kefala, G., and Weiss, M. S. (2007). The three-dimensional structure of *N*-succinyldiaminopimelate aminotransferase from *Mycobacterium tuberculosis*. *J. Mol. Biol*. **367**, 825-838.

24. Schwarzenbacher, R., Jaroszewski, L., von Delft, F., Abdubek, P., Ambing, E., Biorac, T., Brinen, L. S., Canaves, J. M., Cambell, J., Chiu, H., Dai, X., Deacon, A. M., DiDonato, M., Elsliger, M., Eshagi, S., Floyd, R., Godzik, A., Grittini, C., Grzechnik, S. K., Hampton, E., Karlak, C., Klock, H. E., Koesema, E., Kovarik, J. S., Kreusch, A., Kuhn, P., Lesley, S. A., Levin, I., McMullan, D., McPhillips, T. M., Miller, M. D., Morse, A., Moy, K., Ouyang, J., Page, R., Quijano, K., Robb, A., Spraggon, G., Stevens, R. C., van den Bedem, H., Velasquez, J., Vincent, J., Wang, X., West, B., Wolf, G., Xu, Q., Hodgson, K. O., Wooley, J., and Wilson, I. A. (2004). Crystal structure of an aspartate aminotransferase (TM1255) from *Thermotoga maritima* at 1.90 A resolution. *Proteins*. **55**, 759-763.

25. Nakai, T., Okada, K., Akutsu, S., Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S., and Hirotsu, K. (1999). Structure of *Thermus thermophilus* HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry*. **38**, 2413-2424.

26. Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. **22**, 2577-2637.
27. Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. **31**, 3320-3323.
28. Schneider, G., Käck, H., and Lindqvist, Y. (2000). The manifold of vitamin B6 dependent enzymes. *Structure*. **8**, R1−R6.
29. Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997). CATH--a hierarchic classification of protein domain structures. *Structure*. **5**, 1093-1108.
30. Eliot, A. C., and Kirsch, J. F. (2004). PYRIDOXAL PHOSPHATE ENZYMES: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73**, 383-415.
31. Mehta, P. K., Hale, T. I., and Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur. J. Biochem*. **214**, 549-561.
32. Holm, L., and Sander, C. (1997). Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res*. **25**, 231-234.
33. Cohen, G. E. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallogr*. **30**, 1160-1161.
34. Krissinel, E., and Henrick, K. (2005). Detection of protein assemblies in crystals. *In Computational Life Sciences, Proceedings*. **vol. 3695**, 163-174, Springer-Verlag, Berlin, Heidelberg.
35. Sivaraman, J., Li, Y., Larocque, R., Schrag, J. D., Cygler, M., and Matte, A. (2001). Crystal structure of histidinol phosphate aminotransferase (HisC) from *Escherichia coli*, and its covalent complex with pyridoxal-5’-phosphate

and L-histidinol phosphate1. *J. Mol. Biol.* **311**, 761-776.

36. Dubnovitsky, A. P., Ravelli, R. B., Popov, A. N., and Papageorgiou, A. C. (2005). Strain relief at the active site of phosphoserine aminotransferase induced by radiation damage. *Protein Sci.* **14**, 1498-1507.
37. Hennig, M., Grimm, B., Contestabile, R., John, R. A., and Jansonius, J. N. (1997). Crystal structure of glutamate-1-semialdehyde aminomutase: An α_2-dimeric vitamin B6-dependent enzyme with asymmetry in structure and active site reactivity. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4866.
38. Biswal, B. K., Au, K., Cherney, M. M., Garen, C., and James, M. N. G. (2006). The molecular structure of Rv2074, a probable pyridoxine 5'-phosphate oxidase from *Mycobacterium tuberculosis*, at 1.6 angstroms resolution. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62**, 735-742.
39. Jeffery, C. J., Gloss, L. M., Petsko, G. A., and Ringe, D. (2000). The role of residues outside the active site: structural basis for function of C191 mutants of *Escherichia coli* aspartate aminotransferase. *Protein Eng.* **13**, 105-112.
40. Vacca, R. A., Christen, P., Malashkevich, V. N., Jansonius, J. N., and Sandmeier, E. (1995). Substitution of apolar residues in the active site of aspartate aminotransferase by histidine. Effects on reaction and substrate specificity. *Eur. J. Biochem.* **227**, 481-487.
41. Hayashi, H. (2003). Conformational Change in Aspartate Aminotransferase on Substrate Binding Induces Strain in the Catalytic Group and Enhances Catalysis. *J. Biol. Chem.* **278**, 9481-9488.
42. Jäger, J., Moser, M., Sauder, U., and Jansonius, J. N. (1994). Crystal structures of *Escherichia coli* aspartate aminotransferase in two conformations. Comparison of an unliganded open and two liganded closed forms. *J. Mol. Biol.* **239**, 285-305.

43. Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H. (1994). X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate aminotransferases from *Escherichia coli* in open and closed form. *J. Biochem. (Tokyo)*. **116**, 95-107.

44. Beaman, T. W., Blanchard, J. S., and Roderick, S. L. (1998). The conformational change and active site structure of tetrahydrodipicolinate *N*-succinyltransferase. *Biochemistry*. **37**, 10363−10369.

Chapter 3: Mechanism of Substrate Recognition and PLP-induced Conformational Changes in LL-Diaminopimelate Aminotransferase from *Arabidopsis thaliana*

3.1 Introduction

Lysine biosynthesis in plants and bacteria occurs via the diaminopimelate (DAP) pathway, and has been extensively investigated because its inhibition represents an attractive target for antibiotic or herbicide development.[1-3] Mammals lack the enzymes of this biosynthetic route and require L-lysine in their diet. Considerable research has also focussed on the development of lysine-rich food crops because this substance is often the limiting amino acid in the human diet.[4]

The DAP pathway (Figure 3-1) to lysine begins with the condensation of aspartate semialdehyde with pyruvate to give dihydrodipicolinic acid (DHDP), that is then reduced to L-tetrahydrodipicolinic acid (THDP). THDP is converted to LL-diaminopimelate (LL-DAP), which after epimerization to *meso*-DAP (*m*-DAP) undergoes a decarboxylation at the D-stereocenter to produce L-lysine. In most bacteria, the conversion of THDP to LL-DAP occurs via a three step reaction sequence involving *N*-succinylation (or *N*-acetylation) of THDP, followed by transamination and desuccinylation to provide LL-DAP. For many years it was assumed that plants used a similar approach; however, Leustek and coworkers discovered recently that in higher plants and in *Chlamydia* the conversion of THDP to LL-DAP proceeds directly without *N*-acetylated intermediates through transamination by LL-diaminopimelate aminotransferase (LL-DAP-AT). This enzyme uses L-glutamate (L-Glu) as the source of the amino group (Figure 3-2).[3, 5-8]

Figure 3-1 Lysine biosynthetic pathway

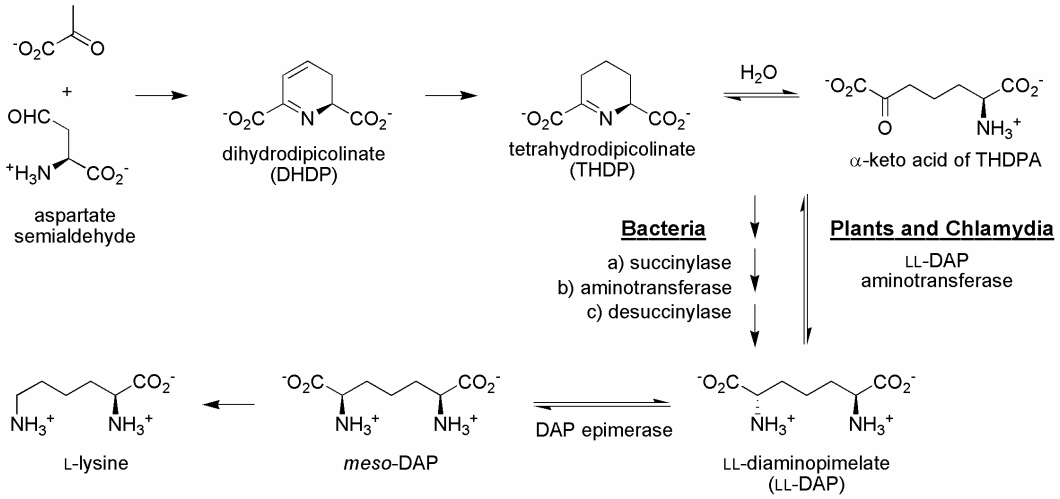
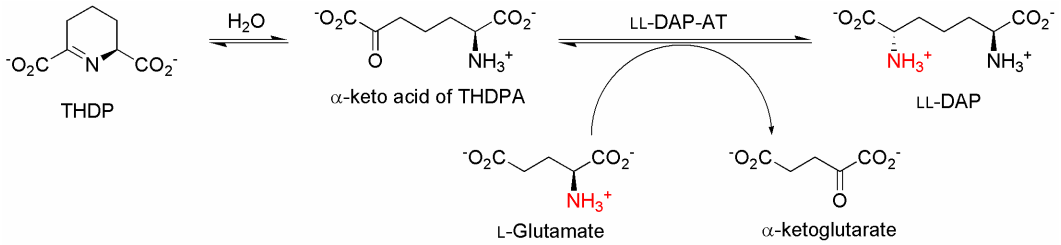


Figure 3-2 The chemical reaction catalyzed by LL-DAP-AT

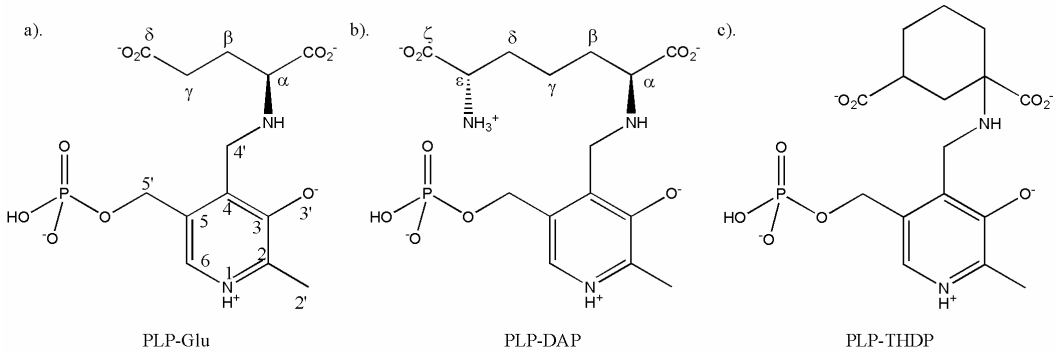


Recently, two variants of LL-DAP-AT have been discovered (DapL1 and DapL2) that differ significantly in sequence.[7] LL-DAP-AT enzymes from plants and *Chlamydia* belong to the DapL1 variant of LL-DAP-AT and share approximately 50% amino acid sequence identity.[7, 8] The DapL2 variant is primarily found in archaea and shares approximately 30% amino acid sequence identity with the DapL1 variant.[7, 8]

Our previous analysis of the crystal structure of LL-DAP-AT from *Arabidopsis thaliana* (AtDAP-AT) revealed that the enzyme is a homodimer[9] and belongs to the type I fold family of PLP-dependent aminotransferases (the aspartate aminotransferase (AspAT) family).[9-11] In particular, it closely resembles subclass Ib aminotransferases, such as *Thermus thermophilus* HB8 AspAT (1BJW).[12] Each subunit consists of a large domain (LD) and a small domain (SD). Both domains belong to the α-β class of protein fold; the LD and the SD fold into an α-β-α sandwich and an α-β complex, respectively. Because of the functional and structural similarities with those of AspAT, the kinetic mechanism of LL-DAP-AT is thought to resemble that of AspAT (ping-pong bi-bi mechanism). Despite the similarity in folding, the actual modes of binding of LL-DAP and L-Glu remained unknown. Modelling of these substrates into the active site of AtDAP-AT suggested that Glu97*, Asn309* and Lys129 may be positioned for the specific recognition of the distal carboxylate groups of L-Glu and LL-DAP and for the stereospecific recognition of the C^{ε}-amino group of LL-DAP. However, AspATs undergo a conformational change upon substrate binding,[12] and it seemed possible that AtDAP-AT could also undergo active site reorganization upon exposure to substrates. Here, in order to assist in the understanding of substrate recognition and catalysis by AtDAP-AT, we have determined the crystal structures of LL-DAP-AT from *A. thaliana* in complex with two analogues of the external aldimines, PLP-Glu (*N*-(5'-phosphopyridoxyl)-L-glutamate) and PLP-DAP

(*N*-(5'- phosphopyridoxyl)-LL-diaminopimelate) (Figure 3-3), in which the imine bond between the substrate and the cofactor has been reduced. A reduced PLP-Glu analogue has been used in the study of AspAT structures with a mimic of the cofactor-substrate complex in the active site.[13] We also report the crystal structures of the asparagine and glutamine variants of the active site lysine, K270N and K270Q, with the bound substrate-cofactor complexes. In contrast to the native enzyme complexes having the reduced analogues, these variant enzymes contain the unreduced external aldimine of PLP with L-Glu and LL-DAP in the active site. Together with an apo-AtDAP-AT structure, the results provide new insights into the mechanism of substrate/cofactor binding and the associated conformational changes in the enzyme.

Figure 3-3 Chemical structures of a) PLP-Glu external aldimine analogue, b) PLP-DAP external aldimine analogue, c) PLP-THDP external aldimine analogue



3.2 Material and methods

Crystallization and data collection of AtDAP-AT

Solutions of the purified AtDAP-AT with the bound PLP-DAP or PLP-Glu aldimine analogues were concentrated to 10 mg/mL and dialyzed against 100 mM NaCl, 20 mM Hepes (pH 7.6), and 1 mM DTT. These AtDAP-AT complexes were then crystallized in 45% (w/v) $(NH_4)_2SO_4$, 0.1 M Hepes (pH 7.5), 3% (w/v) PEG400 by the hanging-drop vapour diffusion method. Within two weeks, X-ray diffraction quality crystals of AtDAP-AT appeared in a drop containing 1 μL of the protein solution and 1 μL of the crystallizing agent. The crystals for both complexes grew to the approximate dimensions of 0.10 mm x 0.10 mm x 0.03 mm within two weeks.

The crystals of the K270N and K270Q variants were prepared using the same procedure as described for the crystals of AtDAP-AT with bound reduced aldimine analogues. All of the K270N variant crystals grew to the dimensions of 0.10 mm x 0.10 mm x 0.03 mm within two weeks. The K270Q variant crystals grew to the dimensions of 0.1 mm x 0.1 mm x 0.1 mm within two weeks.

Before data collection, all of the AtDAP-AT crystals were flash-cooled to 100 K with liquid nitrogen in a cryo-protectant containing 30% (v/v) glycerol, 45% (w/v) $(NH_4)_2SO_4$, 0.1 M Hepes (pH 7.5), 3% (w/v) PEG400. Diffraction data for all of the AtDAP-AT crystals were collected either at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2 or 7-1, CA, USA, or at the Advanced Light Source (ALS) beamline 8.3.1, Berkeley, CA, USA. All data sets were scaled and processed with the HKL2000 package[14]. Detailed data collection statistics are presented in Table 3-1.

Table 3-1 Data collection and refinement statistics

	Native- reduced complex		K270N			K270Q	Native
Crystal	PLP-Glu	PLP-DAP	PLP only	Glu	DAP	Glu	Apo - no PLP
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$
a (Å)	102.56	102.67	103.05	102.58	102.48	102.68	103.10
b (Å)	102.56	102.67	103.05	102.58	102.48	102.68	103.10
c (Å)	171.95	172.15	171.98	171.18	171.68	172.01	171.83
α = β (º)	90	90	90	90	90	90	90
γ(º)	120	120	120	120	120	120	120
Beamline	SSRL 9-2	ALS	SSRL 7-1	ALS 8.3.1	ALS 8.3.1	ALS 8.3.1	ALS 8.3.1
Temperature (K)	100	100	100	100	100	100	100
Detector	ADSCQ350	ADSCQ210	ADSCQ350	ADSCQ350	ADSCQ350	ADSCQ350	ADSCQ210
No. of crystals	1	1	1	1	1	1	1
Total rotation range (º)	132	108	120	120	119.5	120	131
Wavelength (Å)	0.978483	1.115872	0.976067	1.115872	1.115872	1.115872	1.115872

Resolution (Å)	40.0 - 2.05	40.0 - 1.90	40.0 - 1.85	40.0 - 1.55	40.0 - 1.60	40.0 - 1.85	40.0 - 1.99
High resolution shell (Å)	2.12 - 2.05	1.97 - 1.90	1.92 - 1.85	1.61 - 1.55	1.66 - 1.60	1.92 - 1.85	2.06 - 1.99
Total number of observations	424991	458416	369765	1523841	1200738	640890	572704
Unique reflections	62726	78472	83053	151099	137740	86778	73013
	(4464)[a]	(5443)	(8069)	(14973)	(13538)	(7686)	(7219)
Multiplicity	6.8 (3.6)	5.8 (1.7)	4.5 (4.2)	10.1 (6.5)	8.7 (5.7)	7.4 (4.3)	7.8 (7.4)
<I/σ(I)>	15.5 (1.9)	27.3 (2.9)	20.0 (2.3)	24.1 (4.3)	27.6 (2.9)	24.5 (3.2)	25.9 (2.4)
Completeness (%)	94.5 (68.2)	94.1 (66.1)	91.4 (89.9)	99.9 (100.0)	99.8 (99.2)	95.9 (86.1)	100 (100)
R_{sym} (%)[b]	11.6 (58.6)	5.3 (25.6)	6.2 (68.7)	8.2 (44.3)	7.3 (64.4)	7.5 (65.3)	7.8 (79.7)
Refinement resolution (Å)	88.7 – 2.05	89.1 - 1.90	89.1 - 1.85	88.7 - 1.55	88.7 - 1.6	89.1 - 1.85	89.4 - 1.99
$R_{working}$ [c]	0.170	0.157	0.162	0.164	0.167	0.186	0.180
R_{free} [c]	0.221	0.189	0.197	0.191	0.188	0.230	0.224
Number of atoms							
Protein	6308	6338	6356	6346	6346	6338	6338
Water	482	567	1016	1101	1016	754	628

PLP-Glu	50	-	-	-	-	-	-
PLP-DAP	-	56	-	-	-	-	-
PLP	-	-	32	-	16	-	-
K270N-Glu	-	-	-	50	-	-	-
K270N-DAP	-	-	-	-	28	-	-
K270Q-Glu	-	-	-	-	-	50	-
Sulfate	10	15	25	15	10	10	10
Glycerol	24	18	6	36	36	-	-
Average *B*-factor (Å^2)	29.9	25.0	25.0	19.5	22.4	32.1	36.1
Average *B*-factor for protein (Å^2)	29.2	23.3	22.7	16.8	20.0	30.9	35.3
Average *B*-factor for ligands (Å^2)	PLP-Glu	PLP-DAP	PLP-K270N	Glu-K270N	DAP-K270N	Glu-K270Q	-
	30.4	26.1	35.6	21.9	23.3	38.9	-
Average *B*-factor for water (Å^2)	36.9	37.3	37.9	33.4	37.2	40.4	43.0

R.m.s. deviations from the ideal geometry							
Bond lengths (Å)	0.019	0.014	0.015	0.009	0.009	0.022	0.024
Bond angle (°)	1.640	1.328	1.344	1.198	1.222	1.853	1.858
Percentage of residues in Ramachandran plot							
In most favoured regions	92.7	93.1	92.9	93.0	93.0	92.7	92.1
In additional allowed regions	7.0	6.6	6.8	7.0	6.7	7.2	7.7
In generously allowed regions	0.3	0.3	0.3	0	0.3	0.1	0.1

[a] Value within parentheses refer to the highest resolution shell.

[b] $R_{sym} = \sum_h \sum_i(|I_i(\boldsymbol{h})-\langle I(\boldsymbol{h})\rangle|) / \sum_h \sum_i I_i(\boldsymbol{h})$, where $I_i(\boldsymbol{h})$ is the i^{th} intensity measurement and $\langle I(\boldsymbol{h})\rangle$ is the weighted mean of all measurements of $I_i(\boldsymbol{h})$.

[c] $R_{working}$ and $R_{free} = \sum_h ||F(\boldsymbol{h})_o| - |F(\boldsymbol{h})_c|| / \sum_h |F(\boldsymbol{h})_o|$, $|F(\boldsymbol{h})_o|$ and $|F(\boldsymbol{h})_c|$ are the observed and calculated structure factor amplitudes respectively. R_{free} was calculated using 5% of data.

Structure determination and refinement

All of the AtDAP-AT structures presented in Table 3-1 were determined by molecular replacement using the program Phaser[15] or by MOLREP[16] with the native AtDAP-AT as the search model (PDB ID: 2Z20)[9]. All of the AtDAP-AT crystals belong to space group $P3_221$ and have two molecules in the asymmetric unit. The resultant structures were refined with REFMAC5[17, 18] and the water molecules were located using the ARP/wARP solvent building program[19]. After the initial refinement, the molecular models of PLP-DAP and PLP-Glu or the external aldimine with LL-DAP and L-Glu were manually fitted using $|F_o|$-$|F_c|$, α_c electron density maps as found in the active site with the XFIT program from the XtalView package[20]. In a similar manner, glycerol and the sulphate ions were placed in the all of AtDAP-AT structures. The models of AtDAP-AT including the K270N and K270Q variants lack the first 18 residues of the N-terminus as was observed for the native structure. Stereochemical analysis of all complex structures was done by PROCHECK[21]. All of the refinement statistics are presented in Table 3-1. PyMOL (http://pymol.sourceforge.net) or ChemDraw were used to prepare all figures unless specifically noted.

Protein Data Bank accession numbers

The atomic coordinates and structure factors have been deposited with the RCSB Protein Data Bank as entry PDB ID 3EI5 for the PLP–Glu complex, 3EI6 for the PLP–DAP complex, 3EIB for the K270N variant, 3EI9 for the K270N-Glu complex, 3EI8 for the K270N–DAP complex, 3EIA for the K270Q–Glu complex and 3EI7 for apo-AtDAP-AT.

3.3 Results and discussion

Structure of AtDAP-AT in complex with reduced PLP-Glu

The reduced PLP-Glu analogue was synthesized by treatment of a mixture of PLP and L-Glu with sodium borohydride as described.[22] Introduction of this analogue into the AtDAP-AT active site was accomplished by first removing PLP from the native enzyme with phenylhydrazine followed by dialysis against a buffer containing PLP-Glu. Crystallization of the enzyme containing the reduced complex was achieved by the hanging-drop vapour diffusion method.

The structure of AtDAP-AT with reduced PLP-Glu bound in the active site was determined at 2.0 Å resolution. Figure 3-4a shows an excellent $2|F_o|-|F_c|$, α_c electron density map with the reduced PLP-Glu analogue in the active site of AtDAP-AT. In order to determine any conformational changes induced by the binding of reduced PLP-Glu, the structure was aligned with that of our previously-determined native AtDAP-AT[9] having only PLP in the active site using the program ALIGN[23]. The root-mean-square deviation (r.m.s.d.) for backbone atoms between the two was 0.203 Å for 805 C^{α} pairs. This relatively small value holds for residues in the AtDAP-AT active site and indicates that the binding of reduced PLP-Glu by the active site of AtDAP-AT does not result in any significant changes in the protein backbone structure.

The PLP portion of PLP-Glu is bound in the same manner as it is in native AtDAP-AT (Figure 3-4b and 3-4c).[9] Interactions of the glutamate moiety with protein residues illustrate the mode of recognition and binding of this amino acid. The C^{α}-carboxylate group of L-Glu replaces a sulphate anion bound to the native structure and is bound principally by a salt bridge with the guanidinium group of Arg404 and by hydrogen bonds from Arg404$N^{\eta 1\&\eta 2}$, Asn209$N^{\delta 2}$ and Gly64N, as was predicted by our modelling based on the previously observed binding of malate[9]. However, the binding of the C^{γ}-carboxylate group is quite

Figure 3-4 View of the structure of PLP-Glu bound AtDAP-AT a). 2|F_o|-|F_c|, α_c electron density map (contour level = 1σ) for the PLP-Glu aldimine analogue found in the active site of AtDAP-AT. The distance (2.94 Å) between O3' and the nitrogen atom of the Glu is represented by the black dashed line. b). A stereo diagram of the active site of AtDAP-AT in the complex with the PLP-Glu aldimine analogue. Active site residues and PLP-Glu are represented in stick format. Residues from subunit A, B and PLP-Glu are colored green, grey and yellow, respectively. Water molecules are represented by small red spheres. Hydrogen bonds are shown as dashed lines and colored black. The cut-off distance for hydrogen bonds is 3.6 Å. c). A stereo diagram of the superimposed residues in the active sites of native and PLP-Glu-bound AtDAP-AT. Residues and water molecules from native and PLP-Glu complex are coloured yellow and light blue, respectively.

a).

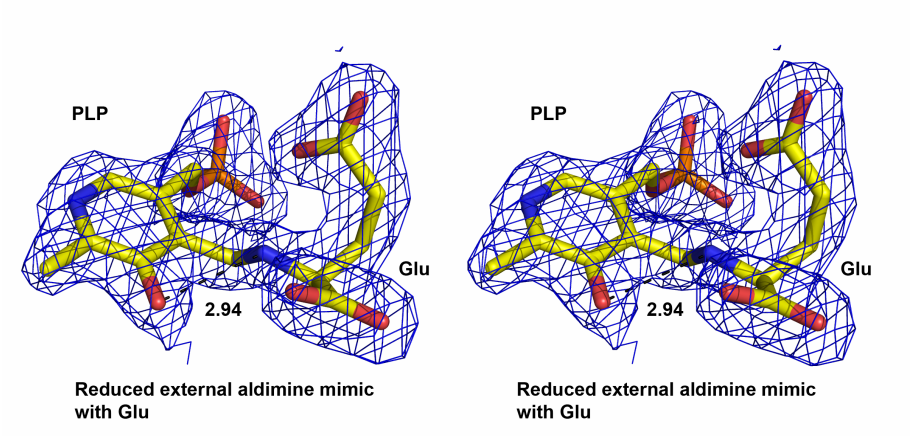
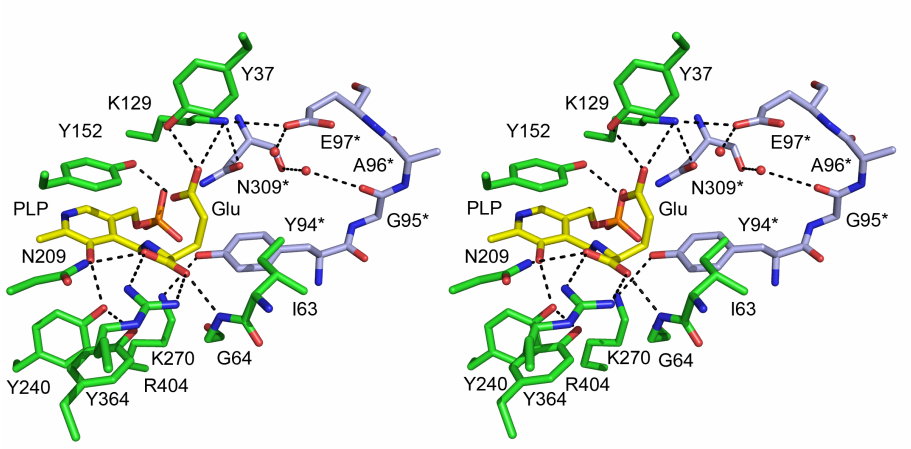
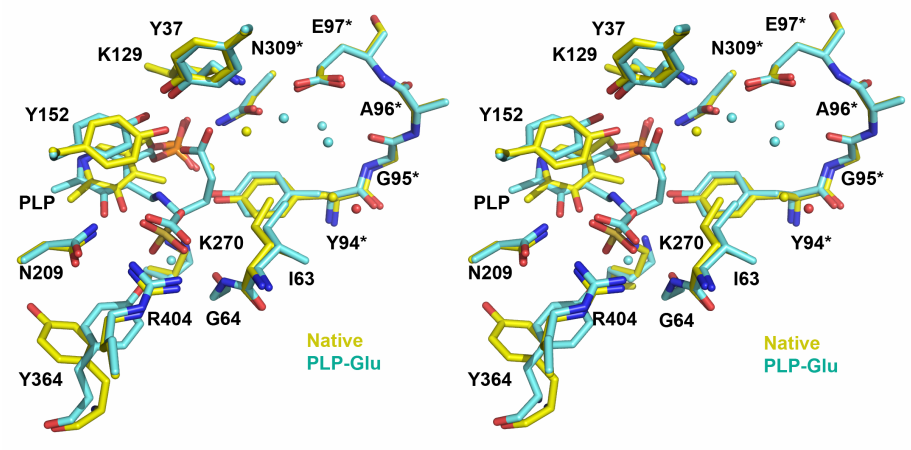


Figure 3-4 (continued)

b).



c).



different from that previously predicted on the basis of the structure containing malate[9]. Rather than the glutamate backbone adopting an extended conformation to form direct and water-mediated hydrogen bonds with Lys129N^{ζ} and Glu97*$O^{\varepsilon 2}$ as had been expected, the distal carboxylate group of L-Glu is curved towards two tyrosine residues, forming a twisted "V-shaped" conformation (χ_1 = 72.1°, χ_2 = -87.5°, χ_3 = -169.2°) (Figure 3-4b). The C^{γ}-carboxylate group of L-Glu is coordinated to three well-conserved residues, Tyr37O^{η}, Tyr152O^{η} and Lys129N^{ζ}, through hydrogen bonds (Figure 3-4b). Glu97* is not involved in the recognition of the γ-carboxylate group. A similar, although not identical, conformation of L-Glu has been observed for an AspAT.[13] Recognition of the distal carboxylate group by Tyr37O^{η}, Tyr152O^{η} and Lys129N^{ζ} has not been seen before in the type I fold family of PLP-dependent aminotransferases. These three residues are highly conserved in the family of LL-DAP-ATs and the mechanism of distal carboxylate binding may be unique to these enzymes.

Although the binding of the reduced PLP-Glu analogue did not result in any significant changes in the backbone structure, there are some conformational changes in the side chains of active site residues. A stereo diagram of the superimposed active site residues between the native (only PLP-bound) and the reduced PLP-Glu-bound AtDAP-AT is shown in Figure 3-4c. A significant change in the conformation of Tyr152 was observed upon binding the reduced PLP-Glu. In the native structure, Tyr152O^{η} makes hydrogen-bonding interactions with Tyr37O^{η} and water molecules bound in the active site. However, those water molecules are displaced by the glutamate moiety of the PLP-Glu analogue and Tyr152O^{η} moves away from the active site cavity by approximately 2.2 Å and forms a hydrogen bond to $O^{\varepsilon 1}$ of the distal carboxylate group of L-Glu. This change also makes the aromatic ring of Tyr152 and the pyridine ring of PLP nearly parallel with one another. A significant difference is also observed for

Tyr364 (Figure 3-4c). In the native unbound enzyme, the side chain of Tyr364 points away from the active site. However, in the reduced PLP-Glu-bound structure it has rotated by ~90^{o} towards the active site cavity. These conformational changes were observed consistently in all refined structures. Similar changes are also observed in the K270N and K270Q variants (see below).

Structure of AtDAP-AT in complex with reduced PLP-DAP

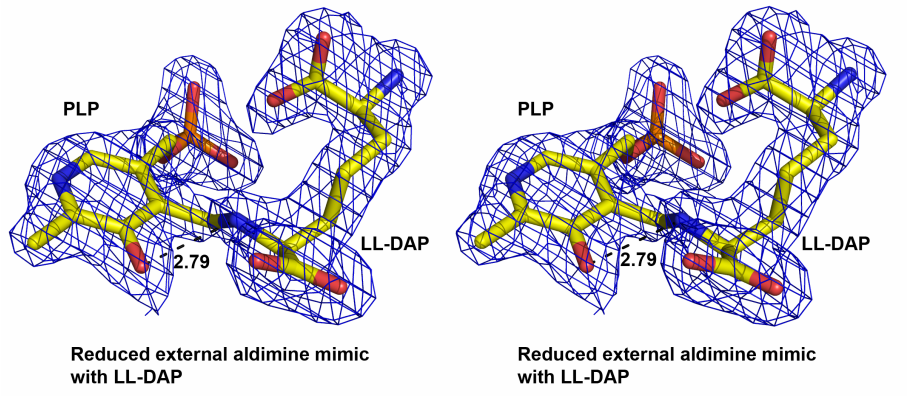
The reduced PLP-DAP analogue was synthesized and incorporated into the active site of AtDAP-AT in a fashion analogous to that for the reduced PLP-Glu analogue. The crystal structure of the resulting complex was determined at 1.9 Å resolution. Figure 3-5a shows an excellent $2|F_o|-|F_c|$, α_c electron density map for the reduced PLP-DAP analogue in the active site of AtDAP-AT. Conformational changes of AtDAP-AT upon binding of PLP-DAP in the active site were again examined by using the structural alignment program ALIGN[23]. The r.m.s.d. between native and reduced PLP-DAP-bound AtDAP-AT was calculated as 0.238 Å for 789 C^{α} atom pairs, indicating that there is no large change in the backbone structure upon the binding of this cofactor-substrate analogue in the active site, as observed with the binding of the reduced PLP-Glu. The difference in C^{α} backbone structures between the reduced PLP-Glu-bound and the reduced PLP-DAP-bound AtDAP-AT is also small with a r.m.s.d. of 0.188 Å for 795 C^{α} atom pairs. Hence the overall structure of AtDAP-AT is relatively rigid and no significant backbone conformational change occurs upon the binding of these analogues of substrate-cofactor complexes.

Figure 3-5b shows PLP-DAP binding in the active site of AtDAP-AT. Surprisingly, both carboxylate groups of LL-DAP are recognized by AtDAP-AT in the same way as those of L-Glu (LL-DAP: χ_1 = 67.8^{o}, χ_2 = -169.9^{o}, χ_3 = 79.1^{o}, χ_4 = 47.9^{o}, χ_5 = 86.6^{o}). Specifically, the C^{α}-carboxylate group of LL-DAP is bound by

Figure 3-5 Views of the structure of PLP-DAP bound AtDAP-AT. a). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) for the PLP-DAP aldimine analogue found in the active site of AtDAP-AT. The distance (2.79Å) between O3' and the nitrogen atom of LL-DAP is represented by the black dashed line. b). A stereo diagram of the active site of AtDAP-AT in complex with the PLP-DAP aldimine analogue. Active site residues and PLP-DAP are represented in stick format. Residues from subunit A, B and PLP-DAP are coloured green, grey and yellow, respectively. Water molecules are represented by small red spheres. Hydrogen bonds are shown as dashed lines and coloured black. The cut-off distance for hydrogen bonds is 3.6 Å. c). A stereo diagram showing the active site residues of superimposed native AtDAP-AT, AtDAP-AT with PLP-Glu bound and PLP-DAP bound. Residues from native AtDAP-AT, PLP-Glu bound and PLP-DAP bound AtDAP-AT are coloured yellow, light blue and emerald green, respectively. d). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) that shows the hydrogen-bonding interactions for the binding of the chiral C^{ε}-amino group of the PLP-DAP substrate analogue. e). A stereo diagram of the active site of AtDAP-AT showing why *m*-DAP is unable to bind due to the too short contact between the ε-amino group and Tyr94* (2.34 Å). Residues from subunit A, B and *m*-DAP are coloured green, grey and light blue, respectively. Water molecules are represented by small red spheres.

Figure 3-5 (continued)

a).



b).

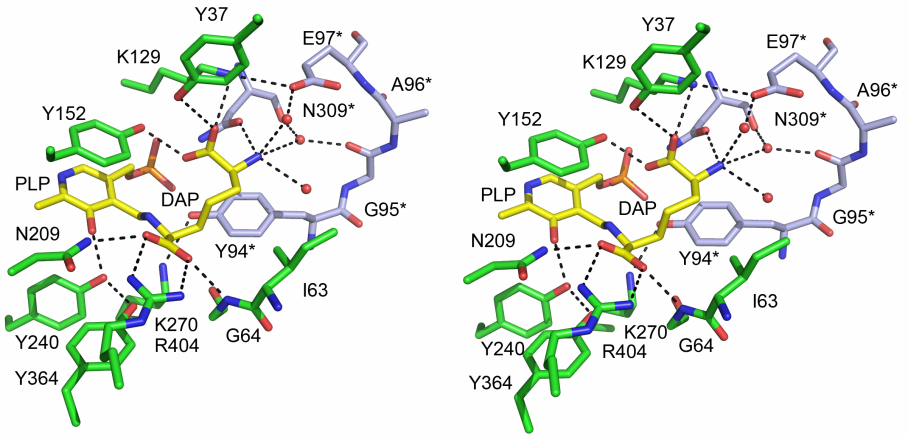
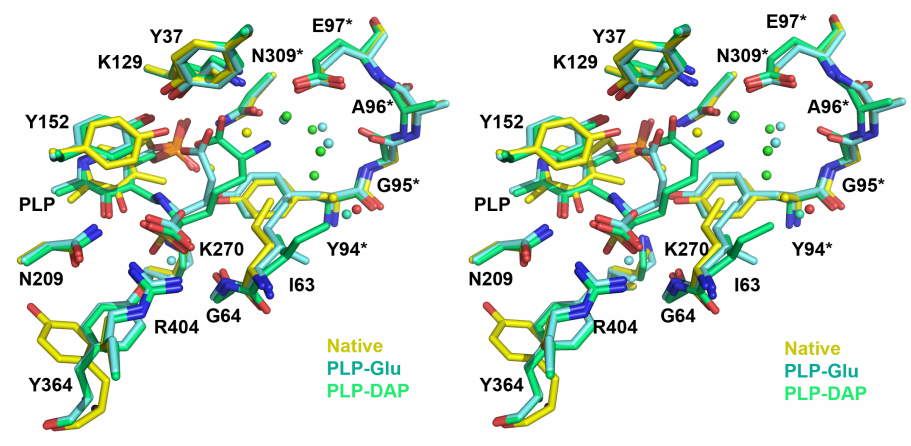
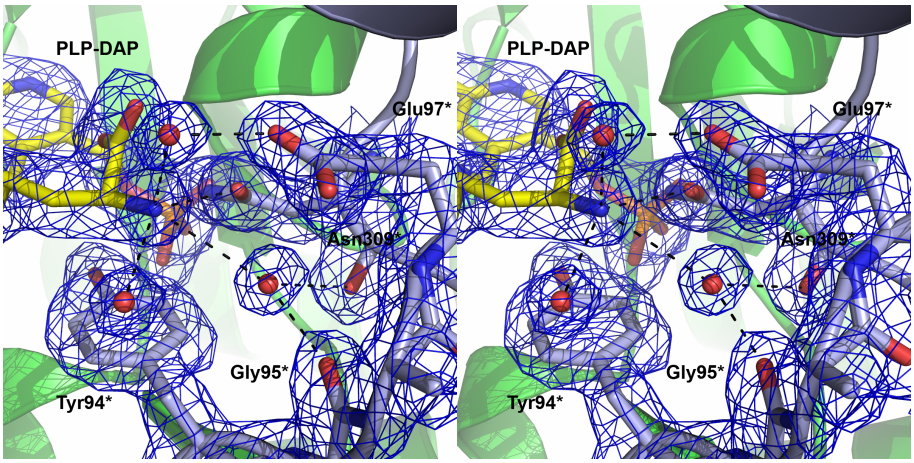


Figure 3-5 (continued)

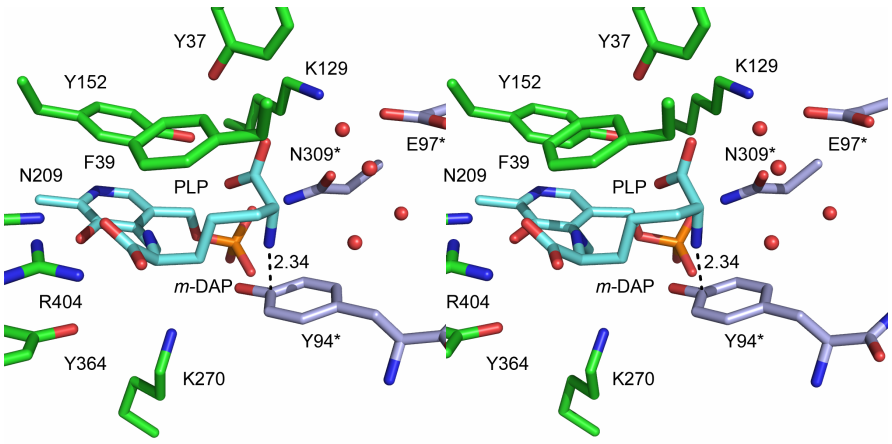
c).



d).



e).



the salt bridge with the guanidinium group of Arg404, and the distal carboxylate group is bound by hydrogen bonds from Tyr37O^{η}, Tyr152O^{η} and Lys129N^{ζ}. In spite of the fact that the distal carboxylate group of LL-DAP is located two methylene groups farther than that of the distal carboxylate in L-Glu, LL-DAP adopts a twisted "V-shaped" conformation similar to that of L-Glu. Thus the longer backbone of LL-DAP requires considerably more space than L-Glu in the active site. This is provided by a conformational change of the side chain of Ile63 (Figure 3-5c), which moves away by approximately 1.2 Å and rotates about the χ_1 bond by ~110^o (Native-Ile63: χ_1 = -169.2^o, DAP-Ile63: χ_1 = -58.9^o). The $C^{\gamma 1}$ and C^{δ} atoms of Ile63 are thus rotated away to provide space for the LL-DAP backbone and to allow both carboxylate groups of LL-DAP to be recognized in the same way as those of L-Glu. Although *Escherichia coli* AspAT contains isoleucine in a similar position, the conformational change upon substrate binding in that enzyme is completely different.[24] This recognition mechanism among the type I fold PLP-dependent aminotransferases appears to be unique to the LL-DAP-ATs as the difference in backbone length of for AspATs is normally only one methylene carbon atom (i.e., Asp and Glu).

An interesting feature of AtDAP-AT is its ability to transform the LL-isomer of DAP exclusively. The *meso*-isomer is not a substrate for the *Arabidopsis* enzyme, although the corresponding *Chlamydia* enzyme apparently can accept both substrates.[6] Figure 3-5b and 3-5d show that the C^{ε}-amino group is bound by Asn309*$O^{\delta 1}$ and by three additional water molecules through hydrogen-bonding. One of these water molecules is hydrogen-bonded to Glu97* and the C^{ε}-amino group, whereas another is hydrogen-bonded to Gly95*O and to Asn309*O. A total of four hydrogen-bonding pairs are present for the recognition of the C^{ε}-amino group of LL-DAP. In addition, the side chain of Tyr94* is located directly below the C^{ε}-amino group and provides additional stabilization through

electrostatic interactions between the aromatic ring and the positive charge on the C^{ε}-amino group (Figure 3-5d). Such interactions with *m*-DAP appear unlikely. A manually modelled structure of the *meso*-isomer in the active site is shown in Figure 3-5e. In this case the C^{ε}-amino group sterically clashes with the Tyr94* side chain and would occupy an electrostatically unfavourable position as Asn309*$N^{\delta 2}$ is in close proximity. The side chain of Tyr94* is held rather rigidly by hydrogen-bonding interactions to the phosphate group on the PLP and to the C^{ε}-amino group of Lys270. The amino acid residues involving the C^{ε}-amino group of LL-DAP in *A. thaliana* are also well-conserved in *Chlamydia trachomatis*. Thus, the mechanism by which chlamydial LL-DAP-AT accepts *m*-DAP is currently unknown. Structural flexibility in the active site may be the factor; the future structural study of chlamydial enzyme should help to solve this problem. Unlike the AspATs, for which differentiation between stereogenic groups at the distal position is unnecessary,[10, 13] AtDAP-AT has specific residues for interaction with the C^{ε}-amino group that allow the stereospecific recognition of the LL-DAP.

Structures of K270N and K270Q variants of AtDAP-AT with L-Glu and LL-DAP

The structures of native enzyme with reduced PLP-DAP and PLP-Glu are not fully representative of the species on the catalytic pathway. They are stable to further transformation and resemble the intermediates, so analogous derivatives have previously been used to characterize substrate-cofactor complexes in other PLP dependent enzymes via crystallography.[13] However, the greater mobility resulting from reduction of the imine bond may change the active site geometry. Fortunately, substitution of the active site lysine can create a catalytically inactive variant in which the true external aldimines – those having a Schiff base linkage between PLP and the substrate – are stable and observable in the crystalline state.[25] Hence, the crystal structures of an active site K270N variant were

determined with: only PLP bound (holoenzyme), in complex with L-Glu and in complex with LL-DAP. A structure of the K270Q variant in complex with L-Glu was also obtained.

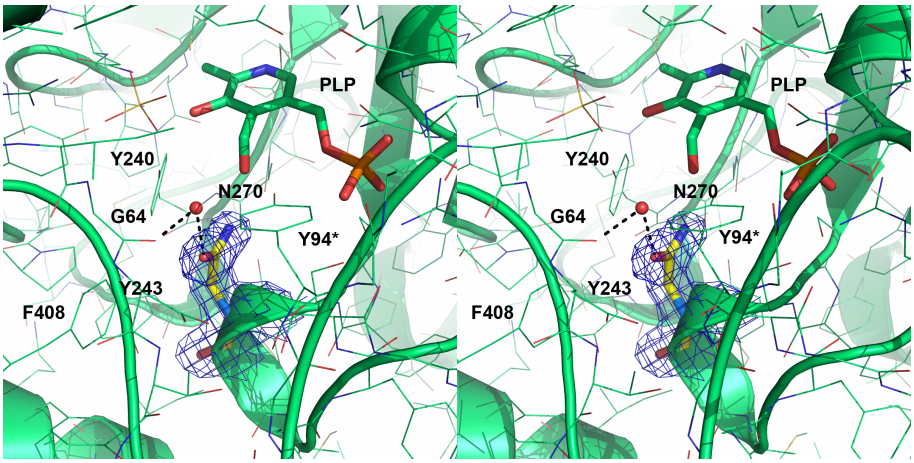
High resolution electrospray mass spectral analysis of the K270N and K270Q variants gave a molecular ion with a mass that was ~117 Da less than expected due to excision of the N-terminal methionine followed by methylation of alanine as the new N-terminus. These post-translational modifications occurred also in the over-expression of native AtDAP-AT. This has been observed previously for other recombinant proteins over-expressed in *E. coli*.[26] However, these modifications were not observed in the present crystal structures due to the disordered N-terminus.

The overall backbone conformations of the K270N and K270Q variants are very similar to those of the native AtDAP-AT containing the reduced substrate-PLP analogues (Figure 3-6a and 3-6b). The r.m.s.d. among native, K270N and K270Q variant structures are all calculated as less than 0.22 (native-K270N: 0.158 for 793 C^{α} and native-K270Q: 0.216 for 798 C^{α}, K270N-K270Q: 0.179 for 814 C^{α}). These values indicate that the substitution did not significantly affect the overall conformation of AtDAP-AT.

Figure 3-7a and 3-7b show $2|F_o|-|F_c|$, α_c electron density maps for the external aldimines between PLP and the two substrates, L-Glu and LL-DAP, respectively. External aldimines with L-Glu and LL-DAP could be trapped successfully in the K270N variant, whereas only the external aldimine with L-Glu was observable for the K270Q variant. Another noticeable point from the structure of the K270N variant is that only partial density for LL-DAP is observed in subunit A of the K270N-DAP complex, whereas its subunit B shows an excellent electron density map of LL-DAP, as presented in Figure 3-7a. The modes of binding of L-Glu and LL-DAP in the active sites of the K270N and K270Q

Figure 3-6 The active site structures of the K270N and K270Q variants. a). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) showing the Asn270 variant in the active site of AtDAP-AT. Hydrogen bonds are shown by black dashed lines, b). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) for the Gln270 variant in the active site of AtDAP-AT. Hydrogen bonds are shown by black dashed lines.

a).



b).

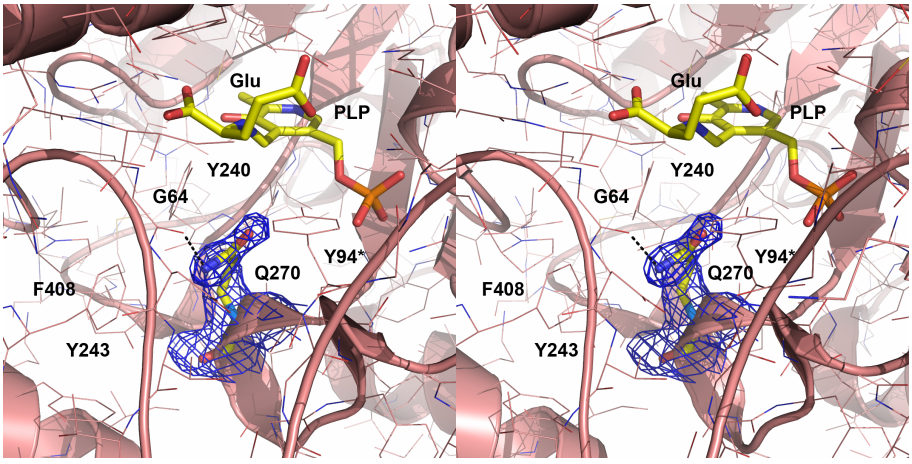
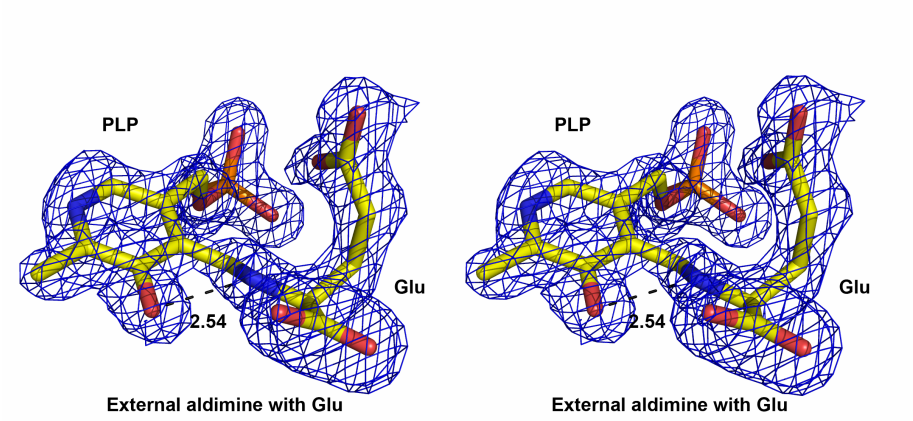


Figure 3-7 a). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) for the external aldimine between PLP and L-Glu as found in the K270N variant structure of AtDAP-AT. The distance between O3' and the nitrogen atom (2.55 Å) is represented by a black dashed line. b). $2|F_o|-|F_c|$, α_c electron density map (contour level = 1σ) for the external aldimine between PLP and LL-DAP as found in the structure of the K270N variant of AtDAP-AT. The distance between O3' and the nitrogen atom (2.54 Å) is represented as a black dashed line. c). Newman projections of the C3-C4-C4'-N bond. C4 is represented as a grey filled circle.

a).



b).

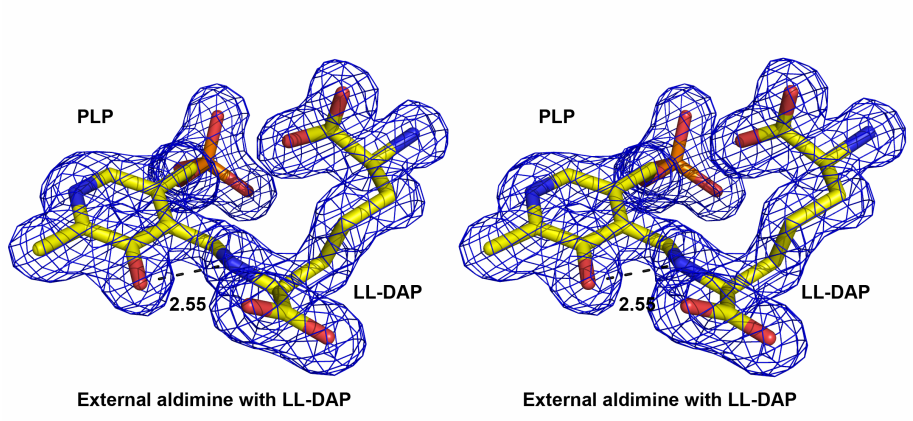
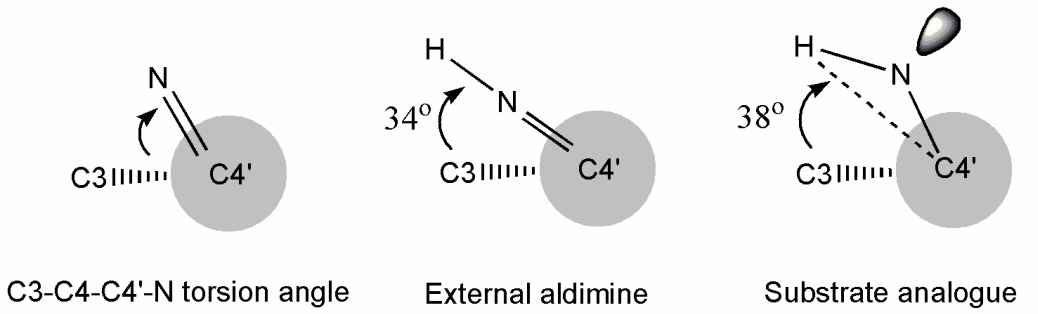


Figure 3-7 (continued)

c).



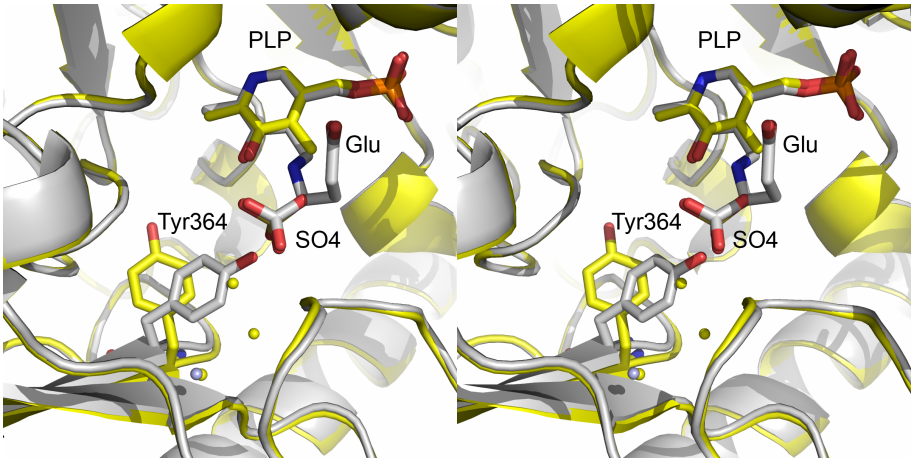
variants generally match those found for the reduced PLP-Glu and PLP-DAP analogues. However, several differences are observed. The torsion angles of C3-C4-C4'-N between the substrate and PLP are as follows: reduced PLP-Glu (-61.8º), reduced PLP-DAP (-55.4º), aldimine PLP-Glu (-40.3º), aldimine PLP-DAP (-33.7º). It would be expected that the Schiff base linkage of the external aldimines would be close to coplanar with the pyridine ring due to a hydrogen bond between the positively charged imine nitrogen and O3' of PLP.[25, 27] This conformation allows optimal stereoelectronic stabilization of the cleaving bond,[28] and is in good agreement with previous studies.[25, 27] However, in the structures of substrate analogues, the torsion angles of C3-C4-C4'-N are approximately 20º larger than that of the external aldimines. These differences in torsion angle can be explained by the presence of hydrogen bond formed between the secondary amine and O3' of PLP in the substrate analogues. Figure 3-7c shows the Newman projection of C3-C4-C4'-N bond. Unlike the imine hydrogen of the external aldimines, the amine hydrogen of the substrate analogues is not located in the same plane as that of C4-C4'-N. This is because of the tetrahedral geometry of the secondary amine of the substrate analogues. In this conformation, the amine hydrogen of the substrate analogues is located at approximately the same position as that of the external aldimines to form a hydrogen bond with O3' of PLP (Figure 3-7c). Although the position of the hydrogen atoms can not be determined by the X-ray crystallography at this resolution, well defined electron density of C3-C4-C4'-N bonds is observed. This implies that the C3-C4-C4'-N bond is quite ordered, and the non-covalent interaction such as hydrogen bond is present to keep the C3-C4-C4'-N bond in well-ordered state. Therefore, we expect that hydrogen bond is formed between amine hydrogen and O3' of PLP of substrate analogues as it is seen in the external aldimines.

Comparison of all the structures indicates two different conformations of the Tyr364 side chain (Figure 3-8a). In AtDAP-AT with the reduced PLP-Glu or PLP-DAP analogues and in the K270Q PLP-Glu external aldimine structure, the side chain of Tyr364 points into the active site cavity. However, in the native AtDAP-AT lacking the substrate and in all of the K270N variant structures, the Tyr364 side chain points away from the active site cavity. In the latter, the volume occupied previously by the Tyr364 side chain is occupied by a bound water molecule (Figure 3-8a). In other type I fold PLP dependent aminotransferases, this position is typically occupied by aromatic amino acid side chains (AspAT (Ia): Phe,[29] AspAT (Ib): Tyr,[12] HisP-AT: Tyr [30]) (Figure 3-8b). In HisP-AT, the tyrosine residue points away from the active site regardless of whether a substrate is bound.[30] In AspATs Ia and Ib, the tyrosine or phenylalanine residue points towards the active site regardless of substrate binding.[12, 29] This suggests that the different conformations of Tyr364 are probably not important for the catalytic function of LL-DAP-AT.

Helix α8 is moved slightly away from the active site cavity in the K270Q variant compared to its position in the native and the K270N variant structures (the r.m.s.d. for residues 270-275 between K270Q and native = 0.46 Å for 12 C^{α}, Max. displacement = 0.72 Å at Tyr271C^{α}) (Figure 3-9). The r.m.s.d. at helix α8 and subsequent loop region (residue 270 to 275) is less than 0.25 Å among the native and all of K270N variant structures. The K270Q mutant did not yield crystals with the bound PLP-DAP external aldimine. It is possible that this variant is not as stable as the K270N variant in solution. This is supported by the relatively large r.m.s.d. of helix α8, as well as by the overall *B*-factor, which is substantially higher than that of the other AtDAP-AT structures (Table 3-2). It may be that the mutant enzyme's glutamine side chain has unfavourable interactions with the larger substrate, LL-DAP, whereas the smaller glutamate moiety can be accommodated successfully.

Figure 3-8 Two distinct conformations of Tyr364 in AtDAP-AT. a). An overlay of the crystal structures of the native AtDAP-AT and the PLP-Glu aldimine analogue complex. The native structure is coloured yellow. The PLP-Glu aldimine analogue structure is coloured grey. Two different conformations of Tyr364 in the active site are shown. b). An overlay of aromatic residues equivalent to Tyr364 of AtDAP-AT from the other type I fold aminotransferases. AtDAP-AT, AspAT (Ia), AspAT (Ib) and HisP-AT are represented in grey, yellow, blue and emerald green, respectively.

a).



b).

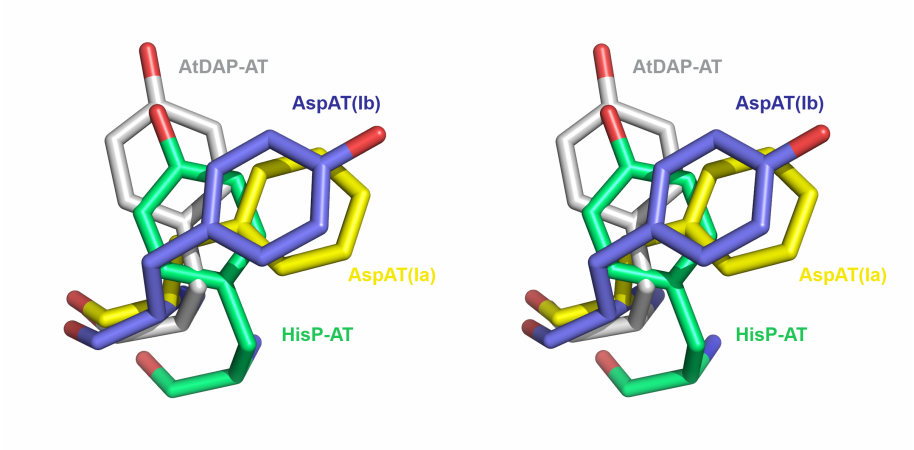
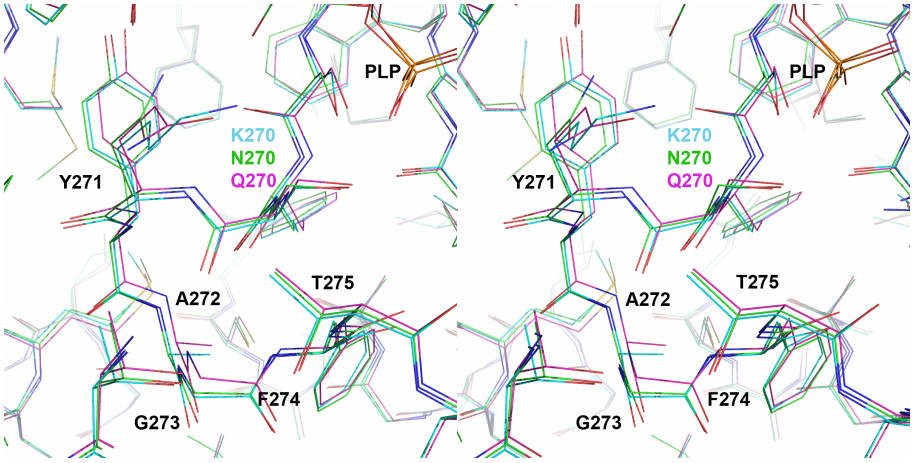


Figure 3-9 An overlay of the structures of native AtDAP-AT, K270N and K270Q variants of AtDAP-AT around active site Lys270 residue shown in line representation. The native, K270N and K270Q variants are coloured cyan, green and magenta, respectively.



LL-DAP-AT catalyzes the reversible conversion of THDP and LL-DAP. However, in solution THDP is in equilibrium with the linear α-keto acid resulting from hydrolysis of the imine bond,[31] and it was unclear which form was accepted by LL-DAP-AT. Previous modeling of THDP with AtDAP-AT having PLP bound indicated that the cyclic form could not be fit into the active site, but it is possible that the enzyme conformation could change to accommodate this molecule.[9] A PLP-THDP analogue (Figure 3-3c) in which the ring nitrogen of THDP was replaced by a methylene group was synthesized to determine if it could be bound in the enzyme active site after removal of PLP. However, in contrast to our results with the reduced PLP-Glu and PLP-DAP derivatives, it was not possible to demonstrate binding nor to obtain a crystal structure of AtDAP-AT with any stereoisomer of this analogue in the active site. This could be due to the replacement of the ring nitrogen of THDP by a methylene, but it seems more likely that the linear α-keto-ε-aminopimelate is the actual substrate for LL-DAP-AT.

Structure of apo-AtDAP-AT and the PLP-induced conformational change at the active site

Attempts to co-crystallize AtDAP-AT with the PLP-THDP analogue did provide a crystal structure of apo-AtDAP-AT that was missing the cofactor (and substrate). This structure revealed that the binding of PLP is the primary factor in the formation of an ordered active site of AtDAP-AT. One of the most noticeable features of apo-AtDAP-AT is that loops “A” and “B” (loop A: residues 91 to 98 and loop B: residues 306 to 312) that partly form the active site, are completely disordered (Figure 3-10a). In addition, loop “C” (residues 62 to 66) in the active site is partially disordered. Furthermore, a large conformational change was observed for residues 267 to 278. In the PLP-bound structure (Figure 3-10b),

Figure 3-10 The structure of apo-AtDAP-AT. a). A cartoon representation of apo-AtDAP-AT. Disordered loops A, B and C are coloured orange. Loops and helices where significant conformational changes were observed are coloured red. The active site Lys270 residue is shown in a stick representation. b). Active site conformations of residues 267 to 279 in the PLP-bound structure represented as sticks. Secondary structural elements are shown in a transparent cartoon representation. PLP, AtDAP-AT residues 267 to 279, loop A and loop C are coloured yellow, grey, cyan and green, respectively. For reason of clarity, the side chain of Y91* is not shown in the figure. Hydrogen-bonds are represented by dashed black lines. c). An overlay of the apo- and the PLP-bound structures of AtDAP-AT. Apo-AtDAP-AT is coloured yellow, and PLP-bound AtDAP-AT is coloured blue. The arrow indicates that helix α8 is shifted towards the active site by approximately 3.4 Å. d). The active site conformation of residues 267 to 279 in the apo structure shown in a stick representation. Secondary structural elements are shown in a transparent cartoon representation. Residues 267 to 279, loop A and loop C are coloured yellow, cyan and green, respectively. Note that the side chain of Arg278, loop A and C are disordered. For reason of clarity, the side chain of Y91* is not shown in the figure.

a).

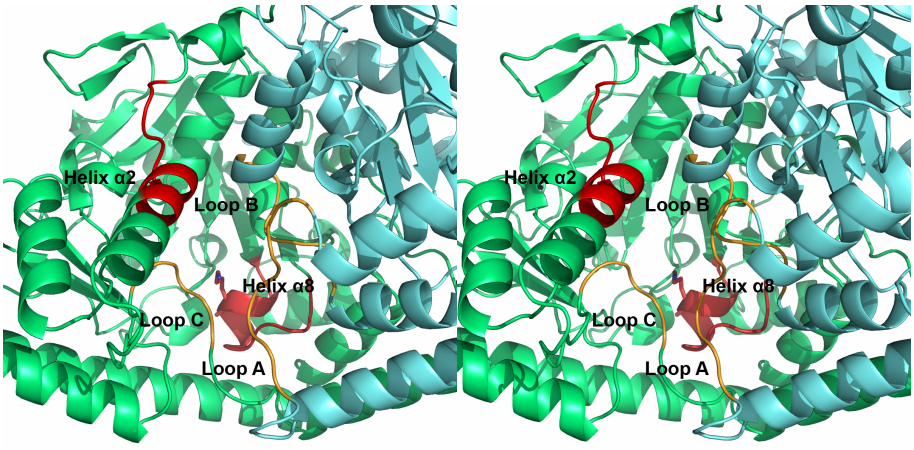
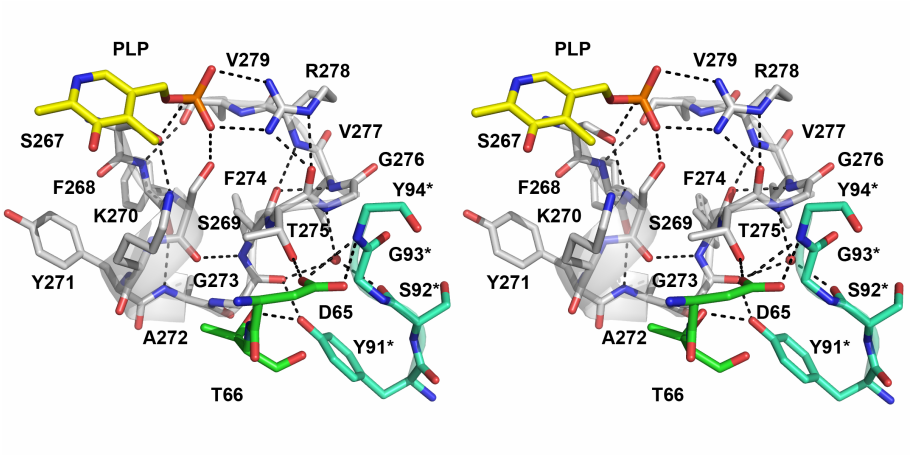
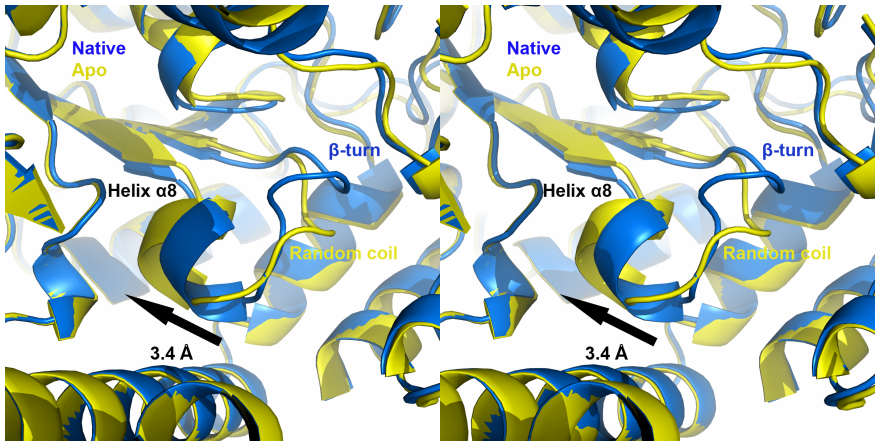


Figure 3-10 (continued)

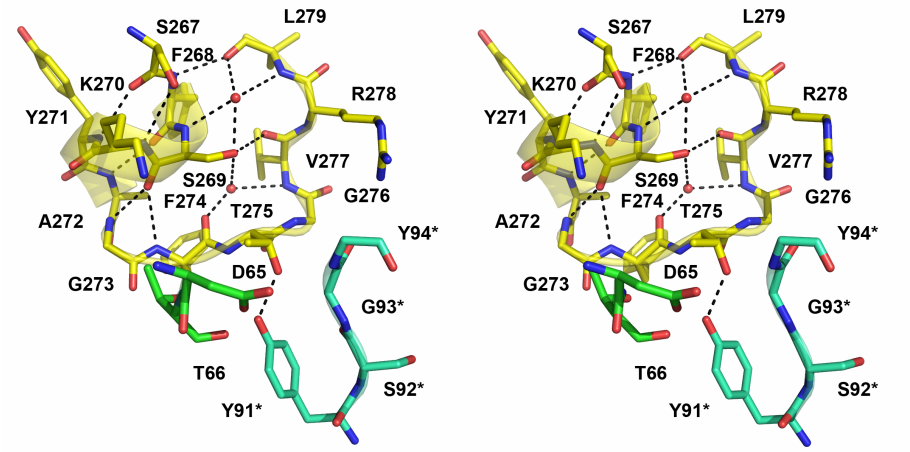
b).



c).



d).



residues 269 to 272 form a short α-helix (α8) with the side chains of Ser267 and Ser269 making two hydrogen bonds with the phosphate moiety of PLP. Residues 274 to 277 form a β-turn that makes several non-covalent interactions with neighbouring residues on loops A and C, including direct (Asp65$O^{\delta1}$ to Thr275O^{γ}, Gly273O to Tyr91*O^{η}, Thr275O to Arg278$N^{\varepsilon\&\eta1}$) and water-mediated (Gly276N/Gly273O to Ser92O) hydrogen bonds. Loops A and C also interact with each other via hydrogen bonds between Thr66O^{γ} and Tyr91*O^{η}, and between Asp65 $O^{\delta1\&\delta2}$ and Tyr94*N. In the absence of PLP, helix α8 (residues 269 - 272) moves towards the active site cavity by approximately 3.4 Å (Figure 3-10c). This allows Ser267O to form a new hydrogen bond with Tyr271N (Figure 3-10d), and that causes Phe268 to become part of helix α8. In the region of residues 274 to 277, the direct hydrogen bond between Phe274O and Val277N is broken in the apo-enzyme and the β-turn structure becomes a random coil structure (Figure 3-10b, 3-10c and 3-10d). Some of the non-covalent interactions with loop A and C are also lost, causing these loops to become disordered. Loop B is also disordered due to the loss of interactions between the phosphate of PLP and Asn309$N^{\delta2}$, and with residues of loop A. The absence of PLP also greatly affects the conformation of the Arg278 side chain as it becomes completely disordered. This results in the loss of two hydrogen bonds between Arg278$N^{\varepsilon\&\eta1}$ and Thr275O that normally provide additional stabilization for the β-turn conformation (Figure 3-10b and 3-10d). Additionally, Ser267O^{γ} and Ser269O form new hydrogen bonds with K270N, and with V277O and water molecules in the absence of PLP, respectively.

The disorder of loops A and B in the apo-enzyme prompted us to examine more closely the conformation of these loops in the other AtDAP-AT structures. The *B*-factors of residues in loops A and B in the native, the K270N, the K270Q and the apo-AtDAP-AT structures are summarized in Table 3-2. As can be seen, the differences in *B*-factor between the overall protein and loop B do

Table 3-2 Differences in *B*-factors between the overall protein and loop A or B.

Protein name	Overall protein	Loop A Loop B	Difference from overall protein (Loop - protein)
Native	20.5	32.0	**+12.5**
		20.9	+0.4
PLP-Glu	29.2	38.2	**+9.0**
		24.9	-4.3
PLP-DAP	23.3	23.3	**0**
		18.9	-4.4
K270N	22.7	24.1	**+1.4**
		18.5	-4.2
K270N-Glu	16.8	18.4	**+1.6**
		13.3	-3.5
K270N-DAP	20.0	19.6	**-0.4**
		15.5	-4.5
K270Q-Glu	30.9	31.6	+0.7
		26.5	-4.4
Apo-enzyme	35.3	88.6	+53.3
		79.9	+44.6

not change significantly as substrate binds. However, the differences in *B*-factor between the overall protein and loop A are relatively high in the native and L-Glu complexes, but decrease significantly upon LL-DAP binding. The same trend in *B*-factors can also be seen in the K270N variant structures. The phosphate moiety of PLP forms a hydrogen-bond with Asn309N$^{\delta 2}$ that stabilizes loop B. For loop A, binding of LL-DAP results in the formation of water-mediated hydrogen bonds between the chiral C$^{\varepsilon}$-amino group of LL-DAP and Glu97 and Gly95 (Figure 3-5d). Since these residues are not involved in the L-Glu binding, the *B*-factor for loop A decreases mainly upon LL-DAP binding. These results imply that loop A may act as a "door" for substrate entry. Its flexibility permits substrates to enter the active site, but it becomes rigid upon substrate binding, essentially preventing access to the active site.

Concluding remarks

In this chapter, the crystal structures of AtDAP-AT in complex with two reduced substrate analogues, PLP-Glu and PLP-DAP have been reported. The structures of these complexes have revealed a novel mechanism employed by AtDAP-AT to accommodate two different substrates in the active site without the need for major conformational changes in the enzyme. Two well conserved tyrosine residues (Tyr37 and Tyr152) and Lys129 are used for binding the distal carboxylate group of both L-Glu and LL-DAP. A single rotation of χ_1 in the side chain of Ile63 accommodates the different lengths of the methylene carbon atoms of these substrates. Glu97* and Asn309* are specifically located to form arrays of hydrogen bonds with the C$^{\varepsilon}$-amino group of LL-DAP, accounting for the recognition of only this isomer. Three structures of a K270N variant (PLP only, PLP-Glu aldimine and PLP-DAP aldimine) and one of a K270Q variant (PLP-Glu aldimine) allow visualization of the binding of the true external aldimines in the

active site of AtDAP-AT. Together with the apo-AtDAP-AT structure and that of the previously reported native holoenzyme,[9] these results clarify the conformational changes induced by PLP binding and the stabilization of loop A at the active site through binding the substrate. These results provide a basis for understanding the mechanism and in assisting in the design of effective inhibitors that may act as herbicides or antibiotics against chlamydial infections.

3.4 References

1. Vederas, J. C. (2006). Diaminopimelate and lysine biosynthesis - An antimicrobial target in bacteria. *Canad. J. Chem.* **84**, 1197-1207.
2. Cox, R. J., Sutherland, A., and Vederas, J. C. (2000). Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorg. Med. Chem*. **8**, 843-871.
3. Scapin, G., and Blanchard, J. S. (1998). Enzymology of bacterial lysine biosynthesis. *Adv. Enzymol. Relat. Areas Mol. Biol*. **72**, 279-324.
4. Mazur, B., Krebbers, E., and Tingey, S. (1999). Gene discovery and product development for grain quality traits. *Science*. **285**, 372-375.
5. Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol*. **140**, 292-301.
6. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T., and Maurelli, A. T. (2006). L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proc. Natl. Acad. Sci. U.S.A*. **103**, 17909-17914.
7. Hudson, A. O., Gilvarg, C., and Leustek, T. (2008). Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase. *J. Bacteriol*. **190**, 3256-3263.
8. Graham, D. E., and Huse, H. K. (2008). Methanogens with pseudomurein use diaminopimelate aminotransferase in lysine biosynthesis. *FEBS Lett*. **582**, 1369−1374.
9. Watanabe, N., Cherney, M. M., van Belkum, M. J., Marcus, S. L., Flegel, M. D., Clay, M. D., Deyholos, M. K., Vederas, J. C., and James, M. N. G. (2007). Crystal structure of LL-diaminopimelate aminotransferase from *Arabidopsis*

thaliana: a recently discovered enzyme in the biosynthesis of L-lysine by plants and *Chlamydia*. *J. Mol. Biol*. **371**, 685-702.

10. Eliot, A. C., and Kirsch, J. F. (2004). PYRIDOXAL PHOSPHATE ENZYMES: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73**, 383-415.
11. Mehta, P. K., Hale, T. I., and Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur. J. Biochem*. **214**, 549-561.
12. Nakai, T., Okada, K., Akutsu, S., Miyahara, I., Kawaguchi, S., Kato, R., Kuramitsu, S., and Hirotsu, K. (1999). Structure of *Thermus thermophilus* HB8 aspartate aminotransferase and its complex with maleate. *Biochemistry*. **38**, 2413-2424.
13. Malashkevich, V. N., Jäger, J., Ziak, M., Sauder, U., Gehring, H., Christen, P., and Jansonius, J. N. (1995). Structural basis for the catalytic activity of aspartate aminotransferase K258H lacking the pyridoxal 5'-phosphate-binding lysine residue. *Biochemistry*. **34**, 405-414.
14. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307-326.
15. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr*. **40**, 658-674.
16. Vagin, A., and Teplyakov, A. (1997). MOLREP : an Automated Program for Molecular Replacement. *J. Appl. Crystallogr*. **30**, 1022-1025.
17. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr*. **50**, 760-763.
18. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr*. **55**, 247-255.

19. Perrakis, A., Morris, R., and Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol*. **6**, 458-463.
20. McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol*. **125**, 156-165.
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*. **26**, 283-291.
22. Watanabe, A. (2002). Reaction Mechanism of Alanine Racemase from *Bacillus stearothermophilus*. X-ray crystallographic studies of the enzyme bound with *N*-(5'-phosphopyridoxyl)-alanine. *J. Biol. Chem*. **277**, 19166-19172.
23. Cohen, G. E. (1997). ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J. Appl. Crystallogr*. **30**, 1160-1161.
24. Hayashi, H. (2003). Conformational Change in Aspartate Aminotransferase on Substrate Binding Induces Strain in the Catalytic Group and Enhances Catalysis. *J. Biol. Chem*. **278**, 9481-9488.
25. Rhee, S., Parris, K. D., Hyde, C. C., Ahmed, S. A., Miles, E. W., and Davies, D. R. (1997). Crystal structures of a mutant (βK87T) tryptophan synthase $\alpha_2\beta_2$ complex with ligands bound to the active sites of the α- and β-subunits reveal ligand-induced conformational changes. *Biochemistry*. **36**, 7664-7680.
26. Apostol, I., Aitken, J., Levine, J., Lippincott, J., Davidson, J. S., and Abbott-Brown, D. (1995). Recombinant protein sequences can trigger methylation of N-terminal amino acids in *Escherichia coli*. *Protein Sci*. **4**, 2616-2618.
27. Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H. (1994). X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate

aminotransferases from *Escherichia coli* in open and closed form. *J. Biochem. (Tokyo)*. **116**, 95-107.

28. Dunathan, H. C. (1966). Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **55**, 712.
29. Smith, D. L., Almo, S. C., Toney, M. D., and Ringe, D. (1989). 2.8-Å-resolution crystal structure of an active-site mutant of aspartate aminotransferase from *Escherichia coli*. *Biochemistry*. **28**, 8161-8167.
30. Haruyama, K., Nakai, T., Miyahara, I., Hirotsu, K., Mizuguchi, H., Hayashi, H., and Kagamiyama, H. (2001). Structures of *Escherichia coli* Histidinol-Phosphate Aminotransferase and Its Complexes with Histidinol-Phosphate and *N*-(5'-Phosphopyridoxyl)-L-Glutamate: Double Substrate Recognition of the Enzyme. *Biochemistry*. **40**, 4633-4644.
31. Caplan, J. F., Sutherland, A., and Vederas, J. C. (2001). The first stereospecific synthesis of l-tetrahydrodipicolinic acid; a key intermediate of diaminopimelate metabolism. *J. Chem. Soc., Perkin Trans. 1*. **18**, 2217-2220.

Chapter 4: The Open Conformation of LL-DAP-AT from *Chlamydia trachomatis*: Implications on the Broad Substrate Specificities

4.1 Introduction

The recent increase in numbers of drug resistant bacteria has posed significant health threat to our society. Bacterial infections by *Mycobacterium tuberculosis* or *Staphylococcus aureus* were previously treated successfully with a number of existing antibiotics. However, the appearance of the multi-drug resistant strains of these bacteria has made these infections hard to treat.[1] Moreover, the recent identification of infectious species with NDM-1 (New Delhi metallo-beta-lactamase) gene has caused even more health concerns as we battle a growing number of drug resistant bacteria.[2]

Chlamydia is another infectious bacterial species that has started to acquire drug resistances.[3] Several antibiotics are currently available for the treatment of *Chlamydia*. However, there are several incidences of observation of the drug-resistant species of *Chlamydia* in hospitals reported.[3]

Bacterial species that belong to the genus *Chlamydiae* have been causing significant numbers of severe human infections around the world. The most frequently reported chlamydial infection is the sexual transmitted disease.[4] It is estimated that 90 million people are infected with *Chlamydia* world wide.[5] This is the most frequently reported sexual transmitted diseases in the world. If left untreated, certain women can develop pelvic inflammatory disease with potential complications of infertility.[4] Aside from sexually transmitted diseases, *Chlamydia* can also cause trachoma, a common cause of blindness in the world.[5, 6] Knowing that *Chlamydia* causes a number of serious human diseases, increasing drug resistance among this genus can cause a severe worldwide health problem.

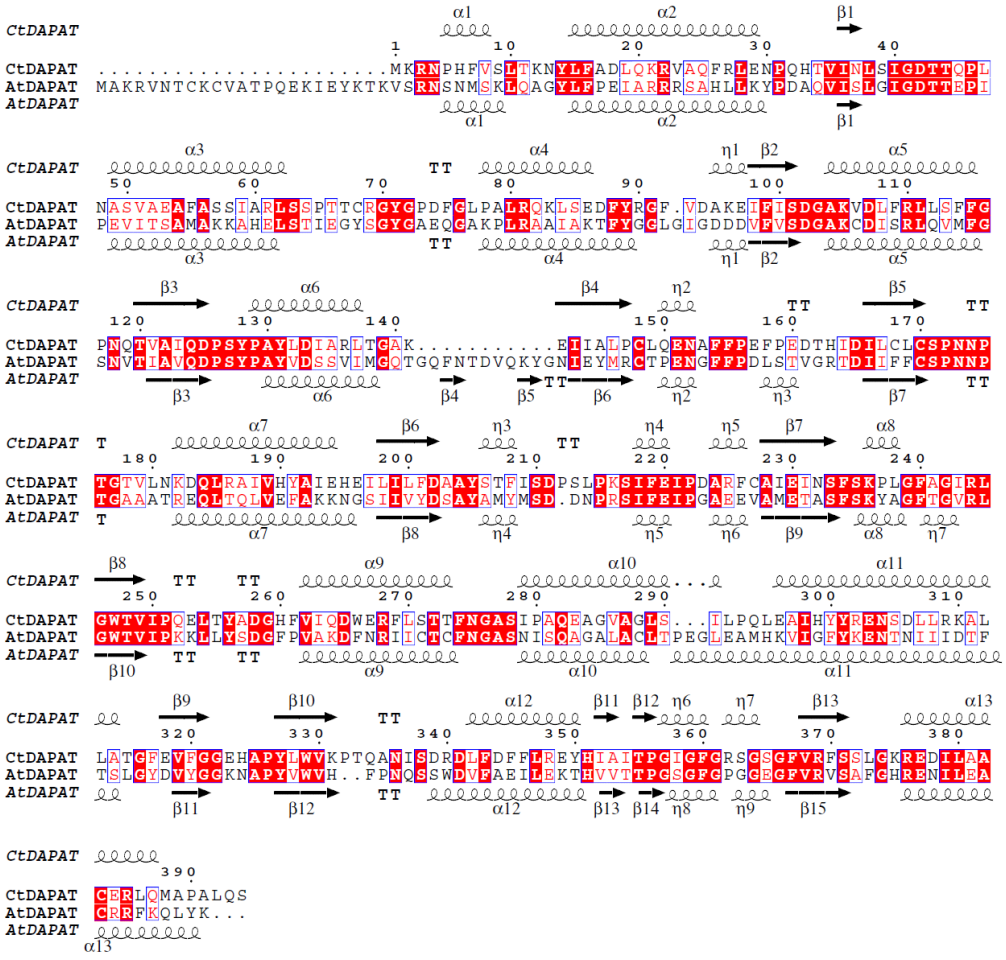
Therefore, the urgent development of novel antibiotics against *Chlamydia* is desperately needed.

Recently, it was found that *Chlamydia* takes an alternative route for lysine biosynthesis from the previously known bacterial lysine biosynthetic pathways.[7] This new pathway is called the "DAP aminotransferase pathway" and involves a novel enzyme, LL-diaminopimelate aminotransferase (LL-DAP-AT). This enzyme is very interesting in the way that it bypasses three enzymatic steps in the previously known bacterial lysine biosynthetic pathways for the conversion of THDP to LL-DAP.

The enzymes involved in the lysine biosynthetic pathways are known to be the attractive drug targets.[8-11] Firstly, these pathways are absent in humans who do not synthesize lysine, but rather obtain it from diet. Secondly, products of the lysine biosynthetic pathway such as *meso*-DAP (*m*-DAP) and L-lysine are needed to maintain microbial structural integrity as they are involved in the important process of cross-linking polyglycan chains in the peptidoglycan cell wall.[12] Therefore, the drugs specifically targeted to the enzymes such as LL-DAP-AT can potentially be effective novel antibiotics.

In order to assist in the development of drugs against LL-DAP-AT, structural investigations of LL-DAP-AT has been undertaken. Recently, several crystal structures of LL-DAP-AT from *Arabidopsis thaliana* (AtDAP-AT) have been reported.[13, 14] AtDAP-AT is a homodimeric enzyme that belongs to the type I fold family of pyridoxal 5'-phosphate (PLP) dependent enzymes.[13] Subsequently, the crystal structures of AtDAP-AT in complex with several substrates (LL-DAP and L-glutamate) have provided valuable information on how LL-DAP-AT might recognize those substrates in the active site.[14] As AtDAP-AT and *Chlamydia* LL-DAP-AT (CtDAP-AT) are more than 40% identical in primary sequences (Figure 4-1), these studies represent the initial important steps toward the development of novel antibiotics against *Chlamydia* infections.

Figure 4-1 Primary sequence alignment between CtDAP-AT and AtDAP-AT. The secondary structures are shown on the top (CtDAP-AT) and the bottom (AtDAP-AT) of the aligned sequences. Sequence alignment was carried out using ClustalW[15]. Identical residues are shown in white on red background, and similar residues are shown in red on white background. This figure was prepared with ESPript v2.2[16].



Although CtDAP-AT and AtDAP-AT likely have similar overall structural folds, they are significantly different in their substrate specificities. As it has been shown in the previous structural and biochemical studies of AtDAP-AT,[7, 14] AtDAP-AT has a very narrow substrate specificity. AtDAP-AT specifically recognizes the L-isomer of DAP (LL-DAP) over its diastereomer, *m*-DAP. With *m*-DAP as an amino donor, AtDAP-AT's catalytic activity is almost completely abolished.[7] On the other hand, CtDAP-AT is capable of using *m*-DAP almost as efficient as LL-DAP as an amino donor. Furthermore, it is also capable of using several other substrates to a lesser extent.[7]

In order to understand the mechanism for the differences in the substrate specificities and to use the structural information for the development of anti-*Chlamydia* drugs, the determination of the three-dimensional structure of CtDAP-AT is crucial. Here, we report the crystal structure of CtDAP-AT in apo- and PLP-bound forms. These structures of CtDAP-AT show several stunning differences in the overall conformation from that of AtDAP-AT and provide some insights into the differences in substrate specificities between AtDAP-AT and CtDAP-AT.

4.2 Material and methods

Cloning, expression and purification of CtDAP-AT

The full-length CtDAP-AT with a C-terminus hexa-histidine tag (His_6) was PCR amplified using pQE60 plasmid (Qiagen Inc.) containing the nucleotide sequence for CtDAP-AT as a template. The PCR product was cloned into the pET3a expression vector (Novagen) using NdeI and BamHI restriction sites. The resultant vector containing the C-terminus His_6 tagged CtDAP-AT gene was transformed into *E. coli* BL21 (DE3) cells (Novagen).

The overexpression of CtDAP-AT-(His_6) was accomplished by using the auto-induction method. The modified auto-induction media (2xYT-5052)[17] supplemented with ampicillin and chloramphenicol was used to support the growth of bacteria and expression of the target proteins. 2xYT-5052 contained the conventional 2xYT media, 5052 (0.5% glycerol, 0.05% Glucose, 0.02% Lactose), NPS (25 mM $(NH_4)_2SO_4$, 50 mM KH_2PO_4, 50 mM NH_2PO_4), trace metal mix (0.05 mM $FeCl_3.6H_2O$, 0.02 mM $CaCl_2.2H_2O$, 0.01 mM $MnC_{l2}.4H_2O$, 0.01 mM $ZnSO_4.7H_2O$, 0.002 mM $CoCl_2$, 0.002 mM $CuSO_4$, 0.002 mM $NiSO_4.6H_2O$, 0.002 mM $Na_2MoO_4.2H_2O$, 0.002 mM H_3BO_3) and 2 mM $MgSO_4$ per 1L of culture media. The *E. coli* BL21 (DE3) cells harbouring the expression vectors were grown approximately for 24 h with vigorous shaking at 25 ^oC.

The overnight grown cells were collected by centrifugation at 7000 rpm for 15 min. The pellet was resuspended into the buffer A (50 mM sodium phosphate (pH 7.5), 10 mM imidazole, 2 mM 2-mercaptoethanol). The resuspended cells were lysed by using Emulsi-flux (15000 psi). After the removal of unbroken cells and insoluble materials by centrifugation, the supernatant was separated for the affinity purification using Talon Superflow Affinity Resin. A 20 column volume of the buffer A was used to wash out unbound or weakly bound proteins. The bound proteins were eluted using a step-wise increase of imidazole

concentration (30 mM, 100 mM, 300 mM and 500 mM imidazole with 50 mM sodium phosphate (pH 7.5) and 2 mM 2-mercaptoethanol). The 30 mM, 100 mM and 300 mM fractions were pooled and concentrated to 20 - 30mg/mL using the 30 kDa cut-off membrane concentrator.

Crystallization of the native and PLP-bound CtDAP-AT

The initial crystallization conditions for CtDAP-AT were found using the Index Crystal Screening Kit from Hampton Research at room temperature. 0.3 μL of CtDAP-AT (20 mg/mL) was mixed with an equal volume of the reservoir solution. The crystals appeared in 1-3 days in the presence of 2.0 M NaCl and 10% (w/v) PEG6000 at room temperature. The X-ray-diffraction quality crystals were grown using the sitting drop vapour diffusion method in the presence of 1.5 - 2.0 M NaCl and 4 - 12 % (w/v) PEG6000 as the precipitants. The final crystals of CtDAP-AT were grown to the size of 0.20 mm x 0.30 mm x 0.20 mm.

The PLP-bound CtDAP-AT crystals were produced using the soaking method. The apo-CtDAP-AT crystals were manually transferred into the soaking solution containing 2.0 M NaCl, 10 % (w/v) PEG6000 and 5 mM PLP for 2 - 3 h.

Data collection and structural determination

The crystals of the apo- and PLP-bound CtDAP-AT were flash-cooled prior to data collection in a cryoprotectant containing 2.0 M NaCl, 10% (w/v) PEG6000 and 30% (v/v) glycerol for the apo crystals. For the PLP soaked crystals, the above cryoprotectant plus 5 mM PLP was used. X-ray diffraction data from both types of crystals were collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-2, CA, USA and at the Canadian Light Source beamline CMCF2, Saskatoon, SK, Canada. All data sets were processed and scaled using the HKL2000 package.[18] Detailed data collection statistics are shown in Table 4-1.

Table 4-1 Data collection and refinement statistics

	CtDAP-AT	
Crystal	Apo	PLP-bound
Space group	$I4_122$	$I4_122$
a (Å)	102.56	102.67
b (Å)	102.56	102.67
c (Å)	171.95	172.15
α = β = γ (º)	90	90
Data collection (SSRL beamline 9-2, CLS beamline CMCF 2)	SSRL 9-2	CLS CMCF 2
Temperature (K)	100	100
Detector	ADSCQ350	MAR300
No. of crystals	1	1
Total rotation range (º)	80	133
Wavelength (Å)	0.978483	0.97949
Resolution (Å)	40.0 - 2.05	40.0 - 2.70
High resolution shell (Å)	2.12 - 2.05	2.80 - 2.70
Total number of observations	281061	143010
Unique reflections	40054 (3945)[a]	17651 (1733)
Multiplicity	7.0 (6.5)	8.1 (8.2)
<I/σ(I)>	41.7 (2.0)	35.8 (3.3)
Completeness (%)	99.9 (100.0)	99.0 (99.8)
R_{sym} (%)[b]	4.2 (87.6)	5.3 (67.5)
Refinement resolution (Å)	97.1 – 2.05	40.0 - 2.70
$R_{working}$ [c]	0.227	0.222
R_{free} [c]	0.277	0.287
Number of atoms		
Protein	3060	2850
Water	216	39
PLP	--	30
Average *B*-factor (Å^2)	51.0	79.9
Average *B*-factor for protein (Å^2)	50.7	79.8
Average *B*-factor for ligands (Å^2)	--	119.4
Average *B*-factor for water (Å^2)	55.3	73.8

R.m.s. deviations from the ideal geometry		
Bond lengths (Å)	0.012	0.020
Bond angle (º)	1.433	2.087
Percentage of residues in Ramachandran plot		
In most favored regions	88.7	83.5
In additional allowed regions	8.0	14.9
In generously allowed regions	2.7	1.3
In disallowed regions [d]	0.6[d]	0.3 [d]

[a] Values within parentheses refer to the highest resolution shell.

[b] $R_{sym} = \sum_h \sum_i(|I_i(\boldsymbol{h})-\langle I(\boldsymbol{h})\rangle|) / \sum_h \sum_i I_i(\boldsymbol{h})$, where $I_i(\boldsymbol{h})$ is the i^{th} intensity measurement and $\langle I(\boldsymbol{h})\rangle$ is the weighted mean of all measurements of $I_i(\boldsymbol{h})$.

[c] $R_{working}$ and $R_{free} = \sum_h ||F(\boldsymbol{h})_o| - |F(\boldsymbol{h})_c|| / \sum_h |F(\boldsymbol{h})_o|$, $|F(\boldsymbol{h})_o|$ and $|F(\boldsymbol{h})_c|$ are the observed and calculated structure factor amplitudes respectively. R_{free} was calculated using 5% of the data.

[d] Residues in disallowed regions are located in the flexible loop areas.

The structure of the apo-CtDAP-AT was determined by molecular replacement using Balbes[19]. The search model (the large domain of AtDAP-AT) was automatically created using structures from the Protein Data Bank (PDB). The crystal of CtDAP-AT belongs to $I4_122$ space group. The asymmetric unit contains one molecule of CtDAP-AT. All structure refinements were carried out using REFMAC5.[20, 21] After the initial refinement, the small domain of CtDAP-AT was manually built into the $|F_o|-|F_c|$, α_c electron density map using XFIT from the XtalView package[22]. Water molecules were mostly built using ARPw/ARP solvent building program[23]. The final model does not include the C-terminal His_6 tag and residues Ala389 - Ser394 due to the disordered in these regions of the electron density map. Final refinement of the apo-CtDAP-AT structure resulted in $R_{work} = 0.226$ and $R_{free} = 0.275$.

The structure of the PLP-bound CtDAP-AT was determined by molecular replacement method using program Phaser[24]. The apo-CtDAP-AT was used as a search model. The PLP was manually built into the active site according to the $|F_o|-|F_c|$, α_c electron density map. The PLP-bound CtDAP-AT structure does not contain residues Thr11 to Leu15, His33 to Thr34, Asp42 to Gln45, Thr67 to Pro73, Thr333 to Ser338, Ala389 to Ser394 and the C-terminal His_6 tag as they are disordered. Final refinement of PLP bound-CtDAP-AT structure resulted in $R_{work} = 0.222$ and $R_{free} = 0.287$. Stereochemical analysis of CtDAP-AT structures was carried out using PROCHECK[25]. All refinement statistics for the apo- and the PLP-bound CtDAP-AT structures are shown in Table 4-1.

Protein Data Bank accession number

The atomic coordinates and structure factors have been deposited with RCSB Protein Data Bank as entry PDB ID XXXX for apo-CtDAP-AT and XXXX for PLP-bound CtDAP-AT.

4.3 Results and discussion

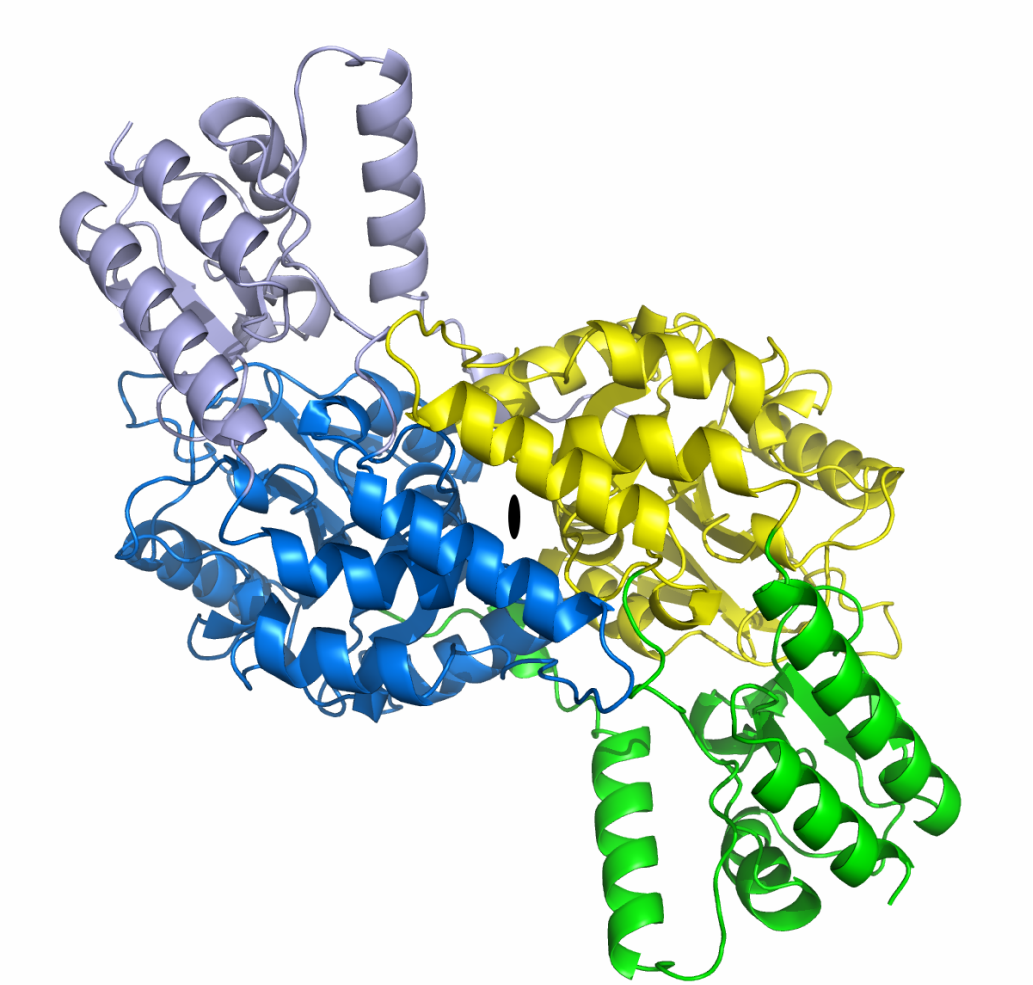
Crystal structures of apo- and PLP-bound CtDAP-AT

The crystal structure of the apo-CtDAP-AT was determined to 2.05 Å resolution using the molecular replacement method. The large domain (LD) of AtDAP-AT was used as the search model due to the failure to obtain correct solution with whole AtDAP-AT molecule. The crystal of apo-CtDAP-AT belongs to the space group $I4_122$ and the asymmetric unit contains one molecule of CtDAP-AT. Although only one molecule of CtDAP-AT was found in the asymmetric unit, the crystallographic two-fold axis produces a homodimeric arrangement of CtDAP-AT (Figure 4-2).

The crystal structure of CtDAP-AT consists of two domains: the LD and the small domain (SD) (Figure 4-2). The LD (Asn48 - Gln294) consists of 246 amino acid residues and folds into the α-β-α sandwich. This domain acts as the PLP binding scaffold and involves in the majority of the dimer forming interactions. The SD (Met1 - Leu47, Leu295-Ser394) consists of the N-terminal arm region and the last 100 residues of the C-terminus. This domain forms an α-β complex. The electron density map for Ala389 to Ser394 and the C-terminal hexa-histidine tag was not observed as they are exposed to the solvent region and are disordered. The overall folding pattern and the homodimeric arrangement of CtDAP-AT are in good agreement with the typical folding pattern of the type I fold family of PLP dependent aminotransferases[26, 27].

In order to see how PLP binds to the active site of CtDAP-AT, we have also determined the crystal structure of the PLP-bound CtDAP-AT to 2.70 Å resolution. The PLP-bound CtDAP-AT crystals were produced by soaking PLP into the apo-CtDAP-AT crystals for two to three hours in the presence of mother liquor. The PLP-bound CtDAP-AT structure was solved by molecular replacement using the apo-CtDAP-AT structure as the search model. The mode of PLP binding

Figure 4-2 Crystal structure of LL-DAP-AT from *Chlamydia trachomatis*. The large domains of each subunit are coloured blue and yellow, and the small domains are coloured grey and green. The 2-fold crystallographic axis is shown by the black oval.



in the active site of CtDAP-AT is shown in Figure 4-3. The $2|F_o|-|F_c|$, α_c electron density map clearly shows that the PLP cofactor is covalently linked to the active site Lys236 through a Schiff base linkage.

The binding mode of PLP in CtDAP-AT is similar to that of AtDAP-AT (Figure 4-3)[13]. The pyridine ring of PLP is stabilized by the aromatic ring stacking interactions with Tyr128, and the hydrogen bonding interaction with Tyr205O^{η}. The positive charge on the pyridine N1 atom is stabilized by the negative charge of the carboxylate group of Asp202. The phosphate moiety of PLP is stabilized by the helical dipole moment from helix α5 and by the several hydrogen bonding interactions with Ser233O^{γ}, Ser235O^{γ}, Asn275*$N^{\delta 2}$, Ala104N and Lys105N. The only major difference in PLP interactions from those of AtDAP-AT is that the guanidinium group of Arg244, which forms a salt bridge with the phosphate moiety of PLP in the AtDAP-AT structure (Arg278), is not facing toward the phosphate group of PLP in the structure of CtDAP-AT (Figure 4-3). Instead, it faces toward the neighbouring subunit, and forms hydrogen-bonding interactions with carbonyl oxygen atoms of Asn275 and Ala277. Although the exact reason for this alternative conformation of Arg244 is not known, it suggests that the interaction with the arginine side chain is not absolutely necessary for the stabilization of PLP in the active site.

The overall structure of the enzyme portion of the apo- and PLP-bound CtDAP-AT is essentially identical (root-mean-square deviation (r.m.s.d.) = 0.335 Å for 342 C^{α} atom pairs) except for the conformation of the active site loop (Ser235 - Leu238) (Figure 4-4a and 4-4b). In the apo-enzyme, Ser235 - Leu238 makes a short helical structure by making three hydrogen-bonds between the main chain oxygen atoms and the nitrogen atoms. In this conformation, the side chain of the active site Lys236, which normally forms a covalent bond with PLP, is pointing out toward the solvent region. However, in the PLP-bound form, the

Figure 4-3 The PLP binding mode of CtDAP-AT. The $2|F_o|-|F_c|$, α_c electron density map (contour level at 0.8σ) for the active site Lys236 and PLP cofactor. Active site residues and PLP are represented in stick format. Residues from subunit A, B and the PLP cofactor are coloured in green, grey and yellow, respectively.

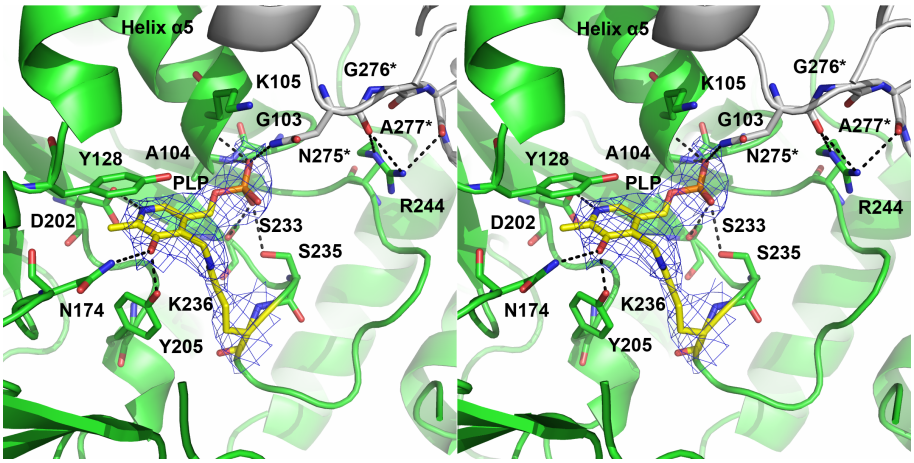
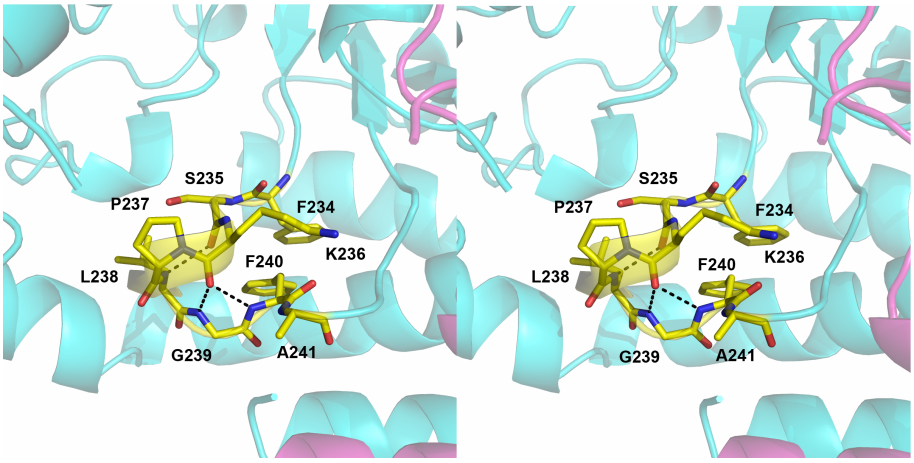
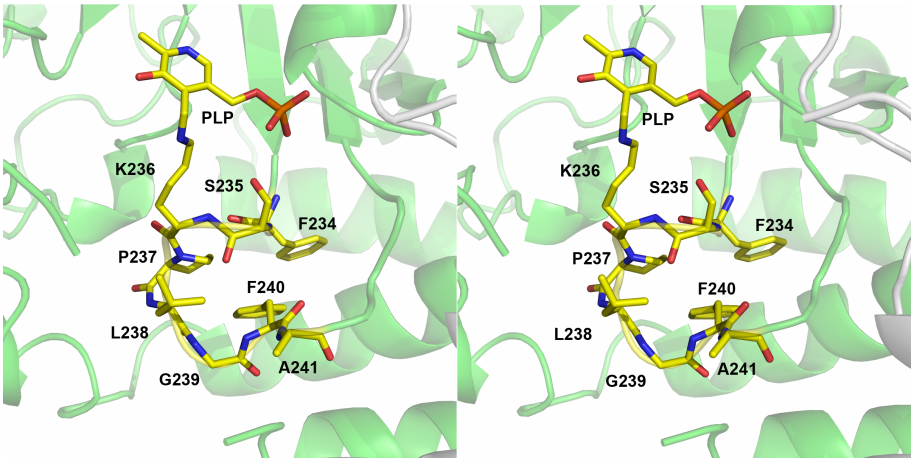


Figure 4-4 The conformational changes observed at the active site helix/loop region of CtDAP-AT. a). Apo-CtDAT-AT (blue and pink) active site helix, b). PLP-bound CtDAP-AT (green and grey) active site loop. Phe234 to A241 are shown as stick model and coloured yellow.

a).



b).



binding of PLP brings Lys236 towards the active site cavity and disrupts the helical structure observed in the apo-form. All of the hydrogen bonding interactions observed in the apo-structure are now lost. This results in Pro237 to move closer to Phe234 and Phe240 and making good hydrophobic interactions with them. This PLP-induced conformational change in the active site region seems to be the common mechanism in LL-DAP-AT. Similar conformational change at the active site helix was also observed between the apo- and the PLP-bound AtDAP-AT structures previously.[14]

The apo-CtDAP-AT structure will be used for the subsequent analysis instead of the PLP-bound CtDAP-AT since those structures are essentially identical and the several disordered loops in PLP-bound form are visible in the apo-form.

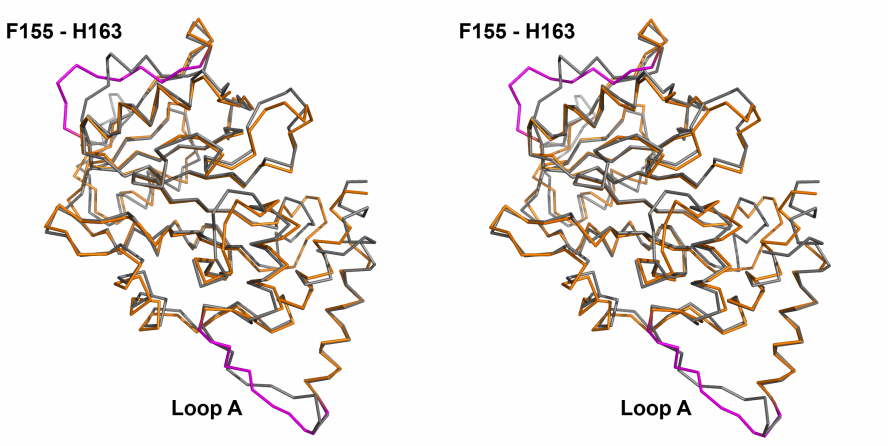
Comparison with the AtDAP-AT structure

The initial attempts to solve the CtDAP-AT structure using the entire AtDAP-AT structure as the search model was unsuccessful. However, subsequent attempts using only the large domain of AtDAP-AT produced the correct phase information for the CtDAP-AT crystal. This indicates that there are significant differences in the overall structures of AtDAP-AT and CtDAP-AT, or that the relative positions of the two domains (LD and SD) in the structures of AtDAP-AT and CtDAP-AT are different. In order to observe conformational differences between AtDAP-AT and CtDAP-AT, the two structures were structurally aligned using the program ALIGN[28].

Each domain was independently aligned to assess any differences in the domain structures (Figure 4-5a and 4-5b). The alignments showed that the secondary structural elements of each domain (r.m.s.d. of SD = 1.50 Å for 124 C^{α} atom pairs, r.m.s.d. of LD = 1.21 Å for 221 C^{α} atom pairs) were well conserved

Figure 4-5 The superposition of the domain structures between CtDAP-AT and AtDAP-AT. a). The superposition of the large domains (CtDAP-AT: orange, AtDAP-AT: grey), b). The superposition of the small domains (CtDAP-AT: orange, AtDAP-AT: grey). The flexible loops (Flexible Loop A, B, and C) of CtDAP-AT are coloured magenta, c). The relative spatial relationship of the flexible loops. The Loop A, B, C are coloured green. Subunit A and B are coloured grey and pink, respectively.

a).



b).

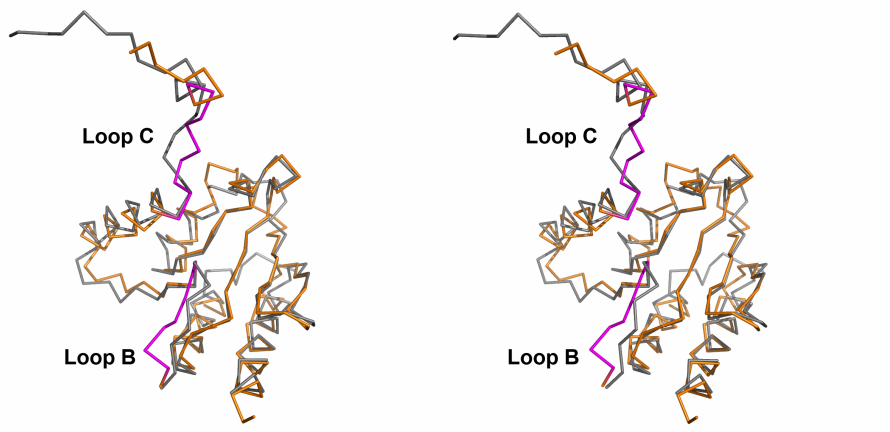
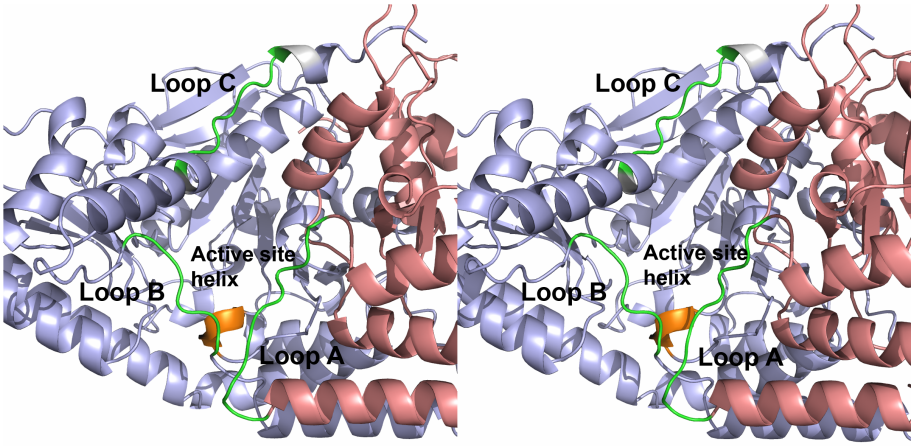


Figure 4-5 (continued)

c).



between AtDAP-AT and CtDAP-AT. The small differences in the r.m.s.d. values were most likely due to the differences in the conformations of the many loop regions.

In CtDAP-AT, there are several loops (Loop A, B, and C) that are quite flexible (Figure 4-5c). The *B*-values for Loop A (Ser64 - Asp74), Loop B (Ser39 - Gln45) and Loop C (Ser9 - Leu15) are determined as 76.4 $Å^2$, 65.2 $Å^2$ and 89.7 $Å^2$, respectively. These values are quite large compared to the rest of the protein molecule (50.9 $Å^2$).

After several cycles of model building and refinement, it was found that the electron density for the residues in Loop A are quite disordered as the main chain atoms of this loop are barely traceable in the apo-structure. Although not as flexible as Loop A, Loop B and C also have disordered electron densities for the side chains. The above loops are much more disordered in the PLP-bound CtDAP-AT structure as most of the above loops show no corresponding electron densities. Because these loops contain the amino acid residues that are important for substrate binding (Loop A: Tyr71 - Asp74, Loop B: Ile40 - Gly41, Loop C: Tyr14) the flexibility of these residues are considered to be very significant and will be discussed in a later section.

In addition to the above differences in the conformation of the loops, the conformation of the residues Phe155 to His163 in CtDAP-AT are different from the corresponding region (Pro191 - Thr199) of AtDAP-AT (Figure 4-5a). In AtDAP-AT, the residues Pro191 to Thr199 adopt a short α-helical structure,[13] whereas in CtDAP-AT, Phe155 to His163 have a more or less unstructured loop. This difference in the conformation is not considered to be catalytically significant because it is likely caused by the crystal packing of molecules in the AtDAP-AT crystal. The above regions of AtDAP-AT are involved in the crystal contacts,[13] whereas the corresponding loop of CtDAP-AT is completely exposed

to the solvent. Therefore, the above region of AtDAP-AT is well structured by packing interactions, whereas that of CtDAP-AT becomes more unstructured due to the exposure to the bulk solvent.

Open/closed conformations of CtDAP-AT

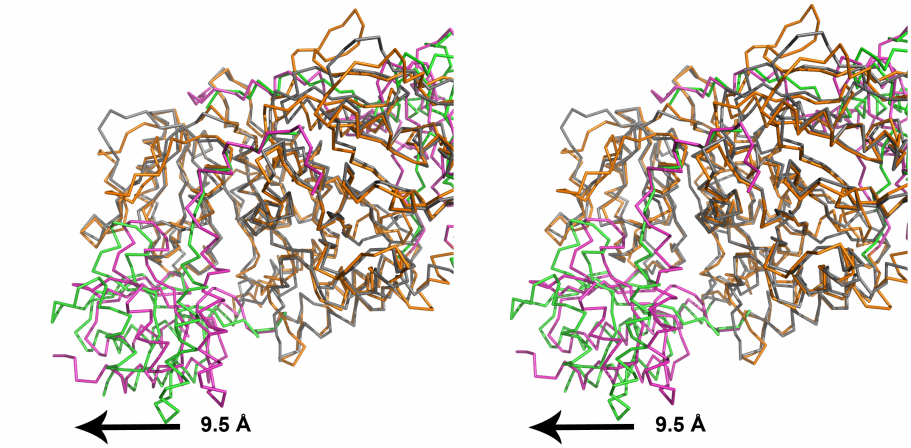
In the next step, the entire structures of apo-CtDAP-AT and PLP-bound AtDAP-AT were structurally aligned. In order to see the differences in the relative positions of each domain, only the LD was used for the alignment of the entire molecule (Figure 4-6a). Based on this alignment, it was found that the entire SD of CtDAP-AT had moved approximately 9.5 Å (maximum displacement) outward compared to that of the AtDAP-AT structure. In this conformation, the active site of CtDAP-AT had opened up and was more exposed to the bulk solvent.

This SD movement is quite dramatic and has shown some effect on the conformation of the active site structure. The superposition of the active site structures of CtDAP-AT and AtDAP-AT is shown in Figure 4-6b. This figure shows that all of the substrate binding residues are well conserved between the two enzymes. However, due to the outward movement of the SD in the CtDAP-AT structure, Asn174 and Arg369, which coordinate the main chain carboxylate groups of the dicarboxylic acid substrates, have moved approximately 2.0 Å away from the active site. Also, two tyrosine residues (Tyr128 and Tyr14 (disordered)) that are important for the distal carboxylate binding show different conformations as well. These results indicate that the open conformation of CtDAP-AT must undergo a certain degree of closing after the binding of substrate. Otherwise, those substrate recognizing residues are too far away to coordinate the substrates in front of PLP for the transamination reaction.

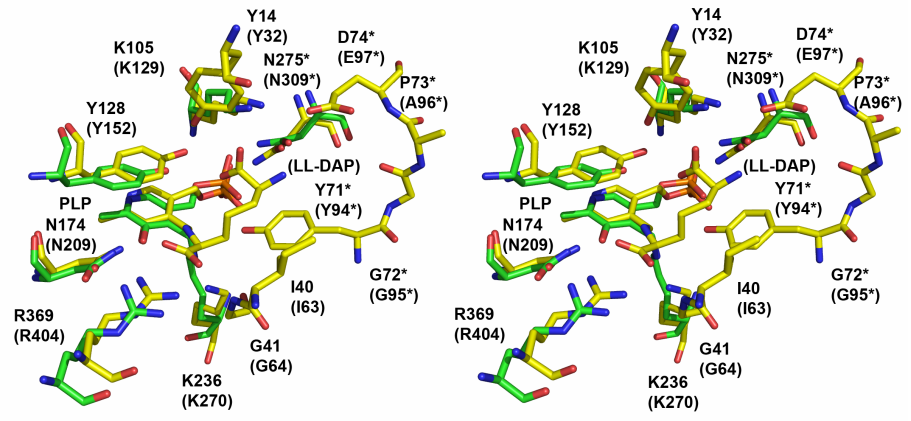
Similar domain opening/closure has been observed in several other aminotransferases in the type I fold family such as AspAT and AroAT.[29-32]

Figure 4-6 The superposition of the entire molecules between CtDAP-AT and AtDAP-AT. a). The displacement of SD is shown by the black arrow. LD of CtDAP-AT and AtDAP-AT are coloured grey and orange, respectively. SD of CtDAP-AT and AtDAP-AT are coloured green and magenta respectively. b). the over-lay of the active site structures between CtDAP-AT (green) and AtDAP-AT (yellow) in stick model. Residue names in bracket are from AtDAP-AT. Residues from Loop A and B of CtDAP-AT are not shown as they are disordered in the PLP-bound CtDAP-AT structure.

a)



b).



However, their SD displacements are much smaller, approximately 3.0 to 4.0 Å. In AspAT, the similar domain movement occurs upon binding of substrates. When the dicarboxylic substrates (Asp or Glu) bind to the active site, coordinated by two Arg residues through electrostatic and hydrogen bonding interactions, the small domain rotates toward the active site and encapsulates the substrates from the bulk solvent.[33, 29, 30] Because each Arg residue comes from different domains (the LD and the SD), the binding of substrates somewhat brings each domain together to make the closed conformation.[34] The interactions between the two Arg residues and the dicarboxylic acids are so important for the domain closure that the lack of one of the carboxylate groups in substrates would hinder formation of the closed conformation.[34] Therefore, the open/closed conformational change in AspAT is strictly governed by substrate binding.

It is quite natural to consider that a similar type of mechanism also exists in LL-DAP-AT since their tertiary structures and their transamination mechanisms are quite similar to each other. A number of different crystal structures of AtDAP-AT has been determined to date including several substrate bound structures.[13, 14] All of the crystal structures of AtDAP-AT are in the closed conformation. However, the substrate binding cavity is quite buried and inaccessible to the bulk solvent. The inaccessibility of the active site is supported by the failed attempts to soak the substrates into the AtDAP-AT crystals (unpublished results). The substrate-bound structures are only obtainable by co-crystallization even though the solvent channel toward the active site entrance can be observed in the crystals of AtDAP-AT.[14] This indicates that a small degree of opening of the active site cavity is necessary for the entry of substrates.

Similarly, the CtDAP-AT is thought to show open/closed conformational changes upon binding of substrate. As described earlier, the open-conformation of CtDAP-AT is not the substrate binding conformation. In other words, the active

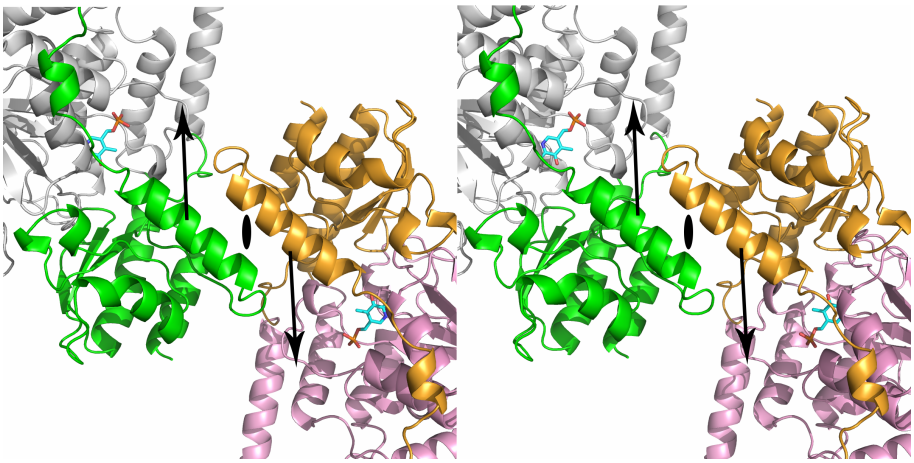
site residues of the open-conformation are not in positions capable of correctly orientating the substrates for the transamination reaction. Therefore, the closing motion of the SD is absolutely necessary for CtDAP-AT to catalyze the transamination reaction. Attempts to soak substrates into the active site of CtDAP-AT in the preformed crystals were also made, but they were not successful either. (The accessibility to the active site has been proven by the soaking of PLP into the active site.) Even though the substrates were able to reach the active site through soaking, the possible closing movement of SD in the CtDAP-AT crystal is prevented by its crystal packing. As shown in Figure 4-7, there is a crystallographic two-fold axis present at helix α2. If the SD were to close upon the soaking of substrates, there would be an opposing force at the crystallographic two-fold axis. This opposing force possibly prevents the closing movement of the SD in the CtDAP-AT crystal.

After examining the active site conformation of CtDAP-AT, it is clear that the substrate will have difficulty in binding to the active site, and that the closure of SD relative to LD is necessary. Therefore, based on the above observations, we anticipate that LL-DAP-AT would undergo a similar domain open/closure mechanism as present in the other type I fold family of PLP dependent aminotransferases.

Implications of the wide-open conformation of CtDAP-AT

Although the domain open/closed mechanism would likely be the common mechanism in LL-DAP-AT, the degree of opening observed in CtDAP-AT is quite large (9.5 Å). Considering the displacement observed in CtDAP-AT is almost three times larger than that observed in AspAT (3.0-4.0 Å)[35], the large degree of opening in CtDAP-AT is probably not necessary for substrate entry. Therefore, the extensive degree of opening observed in CtDAP-AT may

Figure 4-7 The 2-fold crystallographic axis observed in the CtDAP-AT crystals. One monomer is coloured in green (SD) and grey (LD), and the crystallographically related monomer is coloured in orange (SD) and pink (LD). The directions of the force caused by the possible closing movement of SD are shown by the black arrows.



reflect its high internal flexibility. This unique characteristic of CtDAP-AT is supported by: 1). the unusually high overall *B*-values compared to those of AtDAP-AT, and 2). the flexible conformation at the hinge region between the SD and the LD.

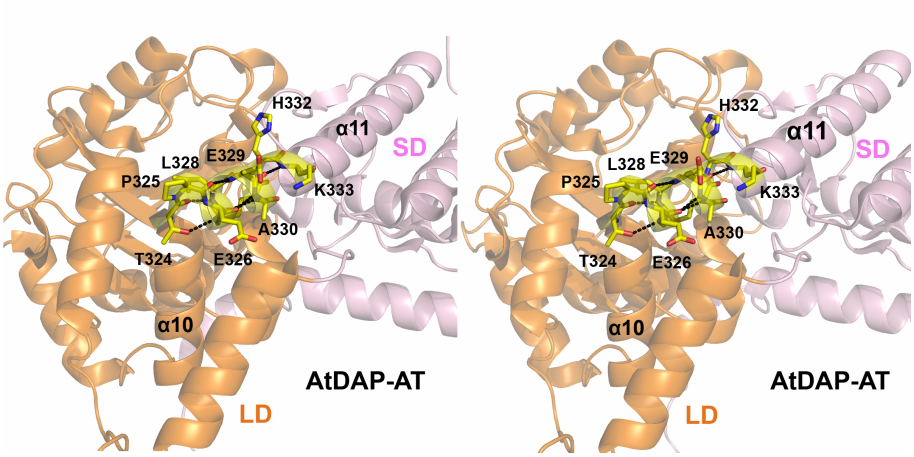
The overall *B*-factors of CtDAP-AT and AtDAP-AT were found to be 50.9 $Å^2$ and 21.7 $Å^2$, respectively. Even though both structures were determined at similar resolution ranges, CtDAP-AT's *B*-value is more than twice that of the *B*-value of AtDAP-AT. More interestingly, the differences in *B*-values between LD and SD in each structure (CtDAP-AT: 19.9 $Å^2$ and AtDAP-AT: 3.3 $Å^2$) show that the SD of CtDAP-AT is much more flexible than its LD, whereas the SD of AtDAP-AT is almost as stable as its LD.

The structures of the hinge region (helix α10 to α11) between the SD and the LD of CtDAP-AT and AtDAP-AT are shown in Figure 4-8a and 4-8b. In AtDAP-AT, the length of helix α11 is much longer than the equivalent helix in CtDAP-AT. This results in a very sharp turn from helix α10 to helix α11 by Thr324 and Pro325 in AtDAP-AT. Proline is a residue well known to cause sharp turns and reduce flexibility. Additionally, the hydrogen bond between Thr324O and Gly326N also reduces flexibility between helix α10 and α11.

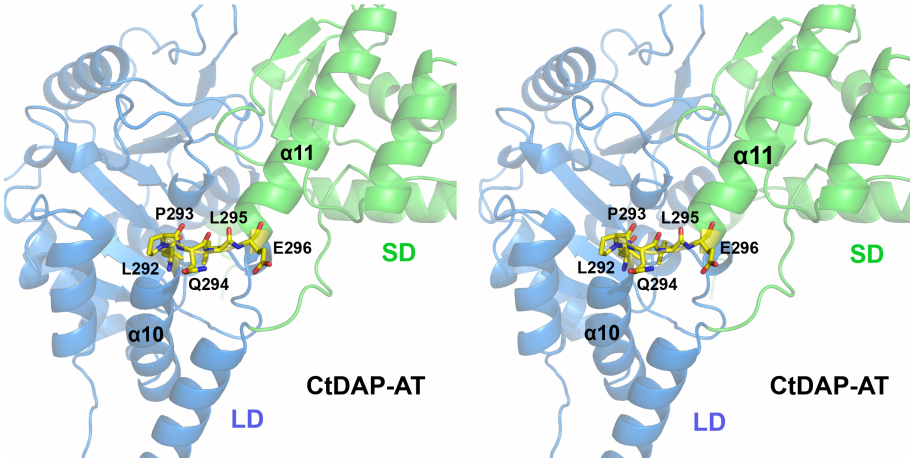
In contrast, in CtDAP-AT, the residues between helix α10 and helix α11 form an unstructured loop (Gln294 to Glu296) (Figure 4-8b). These residues do not form any particular non-covalent interactions, hydrogen bonding or otherwise, with each other, and seem to provide much more flexibility than the corresponding region of AtDAP-AT. Thus, a larger movement of the SD is more likely to be permitted compared to that in AtDAP-AT. Therefore, LL-DAP-AT's domain open/closure mechanism may be universal, but AtDAP-AT or other LL-DAP-ATs may not show as flexible domain movements as CtDAP-AT does.

Figure 4-8 The structures of the hinge region between LD and SD in a). AtDAP-AT and (SD: pink and LD: orange), b). CtDAP-AT (SD: green and LD: blue). Hydrogen bonds are shown as dotted lines. The hinge regions are coloured yellow.

a).



b).



The flexibility and substrate specificity of CtDAP-AT

Since the discovery of LL-DAP-AT in *A. thaliana*, LL-DAP-ATs from several different organisms have been identified and biochemically characterized.[36] From those studies, differences in the substrate specificities among the LL-DAP-ATs became apparent.[36] Depending on the organism, there are two types of LL-DAP-ATs present that exhibit either narrow or broad substrate specificity.

AtDAP-AT and CtDAP-AT are two of the first LL-DAP-ATs to be characterized.[7, 37] They represent the narrow and the broad substrate specificity LL-DAP-AT, respectively. In terms of binding the amino donor molecule, AtDAP-AT is only capable of using LL-DAP; its diastereomer, *m*-DAP, would not be accepted in the active site.[7, 13] On the other hand, CtDAP-AT is capable of using *m*-DAP almost as efficiently as LL-DAP.[7] To a lesser extent, it can also use other structurally similar substrates such as L-lysine or L-ornithine.[7]

Based on the comparison of the primary sequences among different LL-DAP-ATs, the active site residues that are involved in substrate binding are almost completely conserved.[36] This indicates that the variance of the substrate binding residues does not account for the differences in the substrate specificities among LL-DAP-ATs. This is an interesting property of LL-DAP-AT. Understanding this property would be helpful and important to design novel inhibitors. In this last section, the implication of CtDAP-AT's conformational flexibility on the differences in substrate specificity will be discussed.

For several years, there have been attempts to change the substrate specificity of *E. coli* AspAT into that of *E. coli* aromatic aminotransferases (AroAT) using directed evolution and site-directed mutagenesis.[38-40] Both AspAT and AroAT belong to the type I fold family of PLP dependent enzymes.[26] Their active site residues and their spatial dispositions are almost identical. However,

AspAT mainly uses dicarboxylate substrates, whereas AroAT is capable of using both dicarboxylate substrates as well as aromatic amino acids as substrates in the transamination reaction.[41, 42]

Based on the above studies, it was found that in AroAT, the flexibility of the residues lining the active site is important in accommodating both dicarboxlate substrates and aromatic amino acids.[38, 41, 42] For recognition of the distal carboxylates of the dicarboxylic acid substrates, Arg292 makes direct electrostatic and hydrogen bonding interactions. When aromatic substrates bind to the active site, the Arg292 side chain flips away from the active site to make room for the bulky aromatic side chain.[38, 40-42] That conformational change also increases the hydrophobicity of the active site. Although flipping the Arg292 side chain is necessary, it is not sufficient for the aromatic substrates to bind.[43, 42] Instead, slight rearrangement of the hydrogen bonding network formed by the surrounding residues must accompany the above Arg292 conformational change to make recognition of both aromatic and dicarboxylate substrates possible.[41, 42] These studies suggest that the key to AroAT's broad substrate specificity is the flexibility of the residues lining the active site to undergo rearrangement of its hydrogen bonding network.

In LL-DAP-AT, the major structural difference between AtDAP-AT and CtDAP-AT is their conformational flexibility. As described previously, the structure of CtDAP-AT is much more flexible than AtDAP-AT. The loops lining the active site lining loops (Loop A, B, and C) are quite flexible in the CtDAP-AT structure. As in the case of AroAT, the high degree of flexibility in the SD and the active site lining loops may also be important for the broad substrate specificities of CtDAP-AT.

One of the highly disordered loops lining the active site, Loop A, is considered to be the critical loop for the stereospecific recognition of LL-DAP in

AtDAP-AT. The Gly95*, Ala 96* and Glu97* residues of Loop A form an array of hydrogen bonding interactions with the C^{ε}-amino group of LL-DAP.[14] Tyr94* forms an electrostatic interaction with the C^{ε}-amino group and its presence sterically hinders *m*-DAP from binding in the active site.[14] However, if Loop A has increased flexibility as observed in the CtDAP-AT structure, it is very likely that *m*-DAP could be accommodate in the active site since the tyrosine (Tyr71) side chain can be shifted slightly to avoid a steric clash and to provide similar amino-aromatic interactions as it does with LL-DAP. It is also possible that the hydrogen bonding network formed by the above residues can be rearranged for the stabilization of the C^{ε}-amino group of *m*-DAP. Therefore, the flexibility of Loop A has great implications on the substrate specificity of LL-DAP-AT.

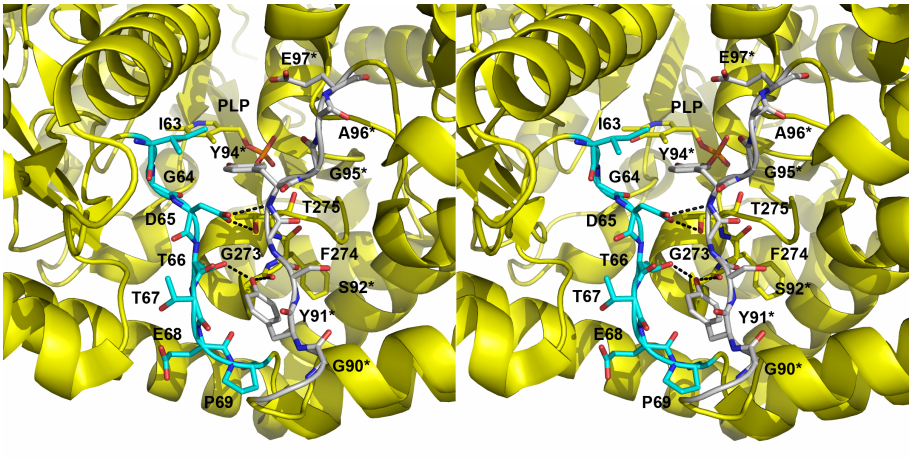
Based on the analysis of the CtDAP-AT structures, the increased flexibility of Loop A is partly due to the large movement of the SD. In AtDAP-AT, Loop A is stabilized by several hydrogen-bonding interactions with Loop B (Figure 4-9a).[14] These interactions are thought to keep those loops in relatively stable conformations. On the other hand, in CtDAP-AT, the large movement of SD causes Loop B to be pulled away from Loop A (Figure 4-9b). This results in the complete loss of hydrogen bonding interactions between Loop A and B. Thus, both loops, with Loop A in particular, are highly disordered in the CtDAP-AT structure. Although LL-DAP-AT's Loop A and Loop B are considered as generally somewhat flexible,[14] CtDAP-AT's overall greater flexibility may allow a higher degree of flexibility in the Loop A conformation allowing it, ultimately, to accommodate a greater number of substrates.

Concluding remarks

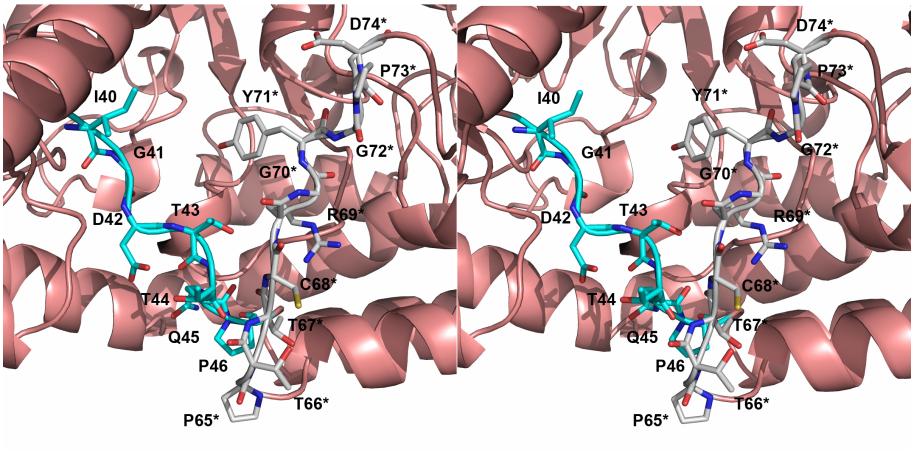
The crystal structure of the open conformation of CtDAP-AT reported here is an important initial step towards understanding the interesting differences

Figure 4-9 The interactions between Loop A and Loop B in a). AtDAP-AT (yellow), and b). CtDAP-AT (pink). Loop A and B are coloured grey and cyan, respectively. The hydrogen bonding interactions are shown as dotted lines. Note that side chains of Loop A in CtDAP-AT are disordered.

a).



b).



in the substrate specificities among the LL-DAP-ATs. Future studies of the closed conformation of CtDAP-AT in complex with various substrates or the crystal structures of LL-DAP-AT from other species would be quite interesting and provide more insights into the above mechanism in order to understand fully the catalytic mechanism of LL-DAP-AT.

4.4. References

1. Mandavilli, A. (2007). Virtually incurable TB warns of impending disaster. *Nat. Med*. **13**, 271.
2. Kumarasamy, K. K., Toleman, M. A., Walsh, T. R., Bagaria, J., Butt, F., Balakrishnan, R., Chaudhary, U., Doumith, M., Giske, C. G., Irfan, S., Krishnan, P., Kumar, A. V., Maharjan, S., Mushtaq, S., Noorie, T., Paterson, D. L., Pearson, A., Perry, C., Pike, R., Rao, B., Ray, U., Sarma, J. B., Sharma, M., Sheridan, E., Thirunarayan, M. A., Turton, J., Upadhyay, S., Warner, M., Welfare, W., Livermore, D. M., and Woodford, N. (2010). Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect Dis*. **10**, 597-602.
3. Somani, J., Bhullar, V. B., Workowski, K. A., Farshy, C. E., and Black, C. M. (2000). Multiple-drug-resistant *Chlamydia trachomatis* associated with clinical treatment failure. *J. Infect. Dis*. **181**, 1421−1427.
4. Manavi, K. (2006). A review on infection with *Chlamydia trachomatis*. *Best Practice & Research Clinical Obstetrics & Gynaecology*. **20**, 941−951.
5. Belland, R., Ojcius, D. M., and Byrne, G. I. (2004). Focus: *Chlamydia*. *Nat Rev Micro*. **2**, 530-531.
6. Thylefors, B., Négrel, A. D., Pararajasegaram, R., and Dadzie, K. Y. (1995). Global data on blindness. *Bull. World Health Organ*. **73**, 115-121.
7. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T., and Maurelli, A. T. (2006). L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17909-17914.
8. Hutton, C. A., Perugini, M. A., and Gerrard, J. A. (2007). Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Mol. BioSyst.* **3**, 458.

9. Scapin, G., and Blanchard, J. S. (1998). Enzymology of bacterial lysine biosynthesis. *Adv. Enzymol. Relat. Areas Mol. Biol*. **72**, 279-324.
10. Vederas, J. C. (2006). Diaminopimelate and lysine biosynthesis - An antimicrobial target in bacteria. *Canad. J. Chem.* **84**, 1197-1207.
11. Cox, R. J., Sutherland, A., and Vederas, J. C. (2000). Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorg. Med. Chem*. **8**, 843-871.
12. van Heijenoort, J. (2001). Recent advances in the formation of the bacterial peptidoglycan monomer unit. *Nat Prod. Rep*. **18**, 503-519.
13. Watanabe, N., Cherney, M. M., van Belkum, M. J., Marcus, S. L., Flegel, M. D., Clay, M. D., Deyholos, M. K., Vederas, J. C., and James, M. N. G. (2007). Crystal structure of LL-diaminopimelate aminotransferase from *Arabidopsis thaliana*: a recently discovered enzyme in the biosynthesis of L-lysine by plants and *Chlamydia*. *J. Mol. Biol*. **371**, 685-702.
14. Watanabe, N., Clay, M. D., van Belkum, M. J., Cherney, M. M., Vederas, J. C., and James, M. N. G. (2008). Mechanism of substrate recognition and PLP-induced conformational changes in LL-diaminopimelate aminotransferase from *Arabidopsis thaliana*. *J. Mol. Biol*. **384**, 1314-1329.
15. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**, 4673-4680.
16. Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. **31**, 3320-3323.
17. Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif*. **41**, 207-234.

18. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307-326.
19. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr*. **64**, 125-132.
20. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr*. **50**, 760-763.
21. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. D Biol. Crystallogr*. **55**, 247-255.
22. McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol*. **125**, 156-165.
23. Perrakis, A., Morris, R., and Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol*. **6**, 458-463.
24. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr*. **40**, 658-674.
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*. **26**, 283-291.
26. Schneider, G., Käck, H., and Lindqvist, Y. (2000). The manifold of vitamin B6 dependent enzymes. *Structure*. **8**, R1−R6.
27. Eliot, A. C., and Kirsch, J. F. (2004). PYRIDOXAL PHOSPHATE ENZYMES: Mechanistic, Structural, and Evolutionary Considerations. *Annu. Rev. Biochem.* **73**, 383-415.
28. Cohen, G. E. (1997). ALIGN: a program to superimpose protein coordinates,

accounting for insertions and deletions. *J. Appl. Crystallogr*. **30**, 1160-1161.

29. Islam, M. M., Goto, M., Miyahara, I., Ikushiro, H., Hirotsu, K., and Hayashi, H. (2005). Binding of C5-dicarboxylic substrate to aspartate aminotransferase: implications for the conformational change at the transaldimination step. *Biochemistry*. **44**, 8218-8229.
30. Jäger, J., Moser, M., Sauder, U., and Jansonius, J. N. (1994). Crystal structures of *Escherichia col*i aspartate aminotransferase in two conformations. Comparison of an unliganded open and two liganded closed forms. *J. Mol. Biol*. **239**, 285-305.
31. Malashkevich, V. N., Strokopytov, B. V., Borisov, V. V., Dauter, Z., Wilson, K. S., and Torchinsky, Y. M. (1995). Crystal structure of the closed form of chicken cytosolic aspartate aminotransferase at 1.9 Å resolution. *J. Mol. Biol*. **247**, 111−124.
32. Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. (1998). Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: a substrate recognition site constructed by rearrangement of hydrogen bond network. *J. Mol. Biol*. **280**, 443-461.
33. Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., and Arnone, A. (1997). Refinement and comparisons of the crystal structures of pig cytosolic aspartate aminotransferase and its complex with 2-methylaspartate. *J. Biol. Chem*. **272**, 17293.
34. Marković-Housley, Z., Schirmer, T., Hohenester, E., Khomutov, A. R., Khomutov, R. M., Karpeisky, M. Y., Sandmeier, E., Christen, P., and Jansonius, J. N. (1996). Crystal structures and solution studies of oxime adducts of mitochondrial aspartate aminotransferase. *Eur. J. Biochem*. **236**, 1025-1032.
35. Okamoto, A., Higuchi, T., Hirotsu, K., Kuramitsu, S., and Kagamiyama, H.

(1994). X-ray crystallographic study of pyridoxal 5'-phosphate-type aspartate aminotransferases from *Escherichia coli* in open and closed form. *J. Biochem. (Tokyo)*. **116**, 95-107.

36. Hudson, A. O., Gilvarg, C., and Leustek, T. (2008). Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase. *J. Bacteriol*. **190**, 3256-3263.
37. Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol*. **140**, 292-301.
38. Chow, M. A., McElroy, K. E., Corbett, K. D., Berger, J. M., and Kirsch, J. F. (2004). Narrowing substrate specificity in a directly evolved enzyme: the A293D mutant of aspartate aminotransferase. *Biochemistry*. **43**, 12780-12787.
39. Onuffer, J. J., and Kirsch, J. F. (1995). Redesign of the substrate specificity of *Escherichia coli* aspartate aminotransferase to that of *Escherichia coli* tyrosine aminotransferase by homology modeling and site-directed mutagenesis. *Protein Sci*. **4**, 1750-1757.
40. Malashkevich, V. N., Onuffer, J. J., Kirsch, J. F., and Jansonius, J. N. (1995). Alternating arginine-modulated substrate specificity in an engineered tyrosine aminotransferase. *Nat. Struct. Biol*. **2**, 548-553.
41. Okamoto, A., Ishii, S., Hirotsu, K., and Kagamiyama, H. (1999). The active site of *Paracoccus denitrificans* aromatic amino acid aminotransferase has contrary properties: flexibility and rigidity. *Biochemistry*. **38**, 1176-1184.
42. Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. (1998). Crystal structures of *Paracoccus denitrificans* aromatic amino acid aminotransferase: a substrate recognition site constructed by rearrangement of hydrogen bond network. *J. Mol. Biol*. **280**, 443-461.

43. Hayashi, H., Kuramitsu, S., Inoue, Y., Morino, Y., and Kagamiyama, H. (1989). [Arg292----Val] or [Arg292----Leu] mutation enhances the reactivity of *Escherichia coli* aspartate aminotransferase with aromatic amino acids. *Biochem. Biophys. Res. Commun*. **159**, 337-342.

Chapter 5: Conclusion

5.1 Structural investigations of LL-DAP-AT from *Arabidopsis thaliana* and *Chlamydia trachomatis*

The development of antibiotics is a challenging but necessary task. As the number of multi-drug resistant strains of bacteria increases, the need for novel antibiotics becomes urgent. There are several criteria that must be considered in order to design good antibiotics. One of which is that the molecule targeted by the antibiotic should be absent in humans. By fulfilling this requirement, the potential toxicity to humans can be decreased significantly. The biosynthetic pathway of essential amino acids such as L-lysine is a very attractive drug target since the entire pathway is absent in humans.

The discovery of LL-diaminopimelate aminotransferase (LL-DAP-AT) and the new lysine biosynthetic pathway in plants and *Chlamydia*[1, 2] have made a significant impact and have provided new hope for scientists working toward the development of new antibiotics. In particular, the presence of LL-DAP-AT in *Chlamydia* was very exciting as *Chlamydia* has started to acquire some multi-drug resistances. In order to use LL-DAP-AT as a target in anti-Chlamydial drugs, the elucidation of its three-dimensional structure and the details of its catalytic mechanism would be a significant step towards achieving the ultimate goal.

The crystal structures of the native LL-DAP-AT from two different species, *Arabidopsis thaliana* (AtDAP-AT)[3] and *Chlamydia trachomatis* (CtDAP-AT) have been determined and reported in this thesis. These structures have revealed that LL-DAP-AT belongs to the family of the pyridoxal 5'-phosphate (PLP) dependent enzymes. PLP is one of the most widely used cofactors in biological systems. In fact, many aminotransferases depend heavily on PLP for their catalysis of the transamination reaction. As described in the

introduction to this thesis, PLP plays a central role in the transfer of an amino group from one substrate to another.

The importance of PLP to the family of PLP dependent aminotransferases is shown by the evolutionarily conserved PLP binding site. The primary sequence alignment and the structural comparison of PLP binding sites among different PLP dependent aminotransferases including LL-DAP-AT have shown that the types of interactions and PLP binding residues are well conserved among different PLP dependent aminotransferases. The active site lysine residue is present in almost every PLP dependent aminotransferase and it forms a covalent linkage with the C4' of PLP. The non-covalent interactions such as the ring stacking interactions between aromatic residues of the enzyme and the pyridine ring of PLP are commonly observed in the PLP dependent aminotransferases. The positively charged environment surrounding the phosphate moiety of PLP is also a well conserved characteristic of the PLP binding cavity. These evolutionarily conserved PLP binding sites reflect how important PLP is for those enzymes to function properly.

PLP dependent enzymes are known to adopt a limited number of structural folds. The crystal structures of the LL-DAP-ATs have shown that its structural fold belongs to the type I fold family of PLP dependent enzymes. Within the type I fold family, the tertiary structure of LL-DAP-AT resembles that of aspartate aminotransferase (AspAT), which represents subclass I of the type I fold family. This subclass of enzymes has been studied for a long period of time as AspAT was the first crystal structure of the PLP dependent aminotransferases to be determined.[4] Many of the unique characteristics of the subclass I of the type I fold family of PLP dependent aminotransferases previously described for the structures of AspAT are also observed in the structures of LL-DAP-AT.

One of the commonly observed features among the subclass I of the type

I fold family enzymes is domain movement. LL-DAP-AT is composed of two domains; the large domain (LD) and the small domain (SD). These domains are known to change their relative orientation upon binding of substrates. This results in the “open” and the “closed” conformations.

In the structures of LL-DAP-AT described in this thesis, a substrate induced conformational change is not observed. However, two different conformations of LL-DAP-AT, the “open” and the “closed” conformations, are observed in the crystal structures from two different species. The analysis of these structures has indicated that LL-DAP-AT undergoes similar open / closed conformational changes observed in other subclass I, type I fold family of PLP dependent aminotransferases.

The structures of the both native and substrate-bound AtDAP-AT show the closed conformation even in the absence of substrates. However, the detailed examination of the several crystal structures of AtDAP-AT has indicated that the opening of the active site is necessary for the substrate to enter into the active site. The active site is not accessible without a certain degree of opening. Similarly, the structure of CtDAP-AT has shown that change to its conformation is necessary for efficient catalysis. The wide-open conformation observed in the crystal structure of CtDAP-AT cannot orient its substrates into the correct position without a closing motion by the SD. Therefore, these structures have indicated that the open/closed conformational changes is a common mechanism in LL-DAP-AT, consistent with the other type I fold family PLP dependent enzymes.

Although the domain movement and the overall architecture of LL-DAP-AT is quite similar to the other type I fold family enzymes, its mode of substrate recognition distinguishes it from the other similarly folded enzymes. While the size of substrates (L-Glu and L-Asp) for AspAT is similar to each other, those of LL-DAP-AT (L-Glu and LL-DAP) are quite different. L-Glu and

LL-DAP differ in size by two methylene carbons and LL-DAP also has an extra C^{ε}-amino group. The key to understanding this unique substrate recognition mechanism of LL-DAP-AT is drawn from the structures of AtDAP-AT in complex with the LL-DAP and L-Glu substrate analogues (PLP-DAP and PLP-Glu).[5] These substrate analogues are covalent adducts of PLP with the two substrates; they have been effectively used to study the mode of substrate binding in the other PLP dependent aminotransferases.[6]

Due to the two methylene carbon length differences between LL-DAP and L-Glu, a large conformational rearrangement of the active site residues in AtDAP-AT was initially expected. However, despite the significant differences in size of the two substrates, the dicarboxylic acids of LL-DAP and L-Glu are recognized by the same set of residues without significant conformational changes in the enzyme's backbone structure. As previously predicted by malate binding, the guanidinium group of Arg404 provides the major stabilization force for the α-carboxylate group through a salt bridge and hydrogen bonding interactions. The distal carboxylate groups of LL-DAP and L-Glu are recognized by the two well-conserved tyrosine residues (Tyr37 and Tyr152) and Lys129 through hydrogen-bonding interactions.

In order for AtDAP-AT to recognize both substrates in the same fashion, a rotation of the side chain of Ile63 is found to be the key mechanism for accommodating the long methylene carbon chains of LL-DAP. When L-Glu is bound in the active site of AtDAP-AT, the bulky side of Ile63 side chain ($C^{\gamma 1}$ and C^{δ} atoms) faces towards the substrate binding cavity. In contrast, when LL-DAP is bound in the active site, the rotation at the χ_1 bond of Ile63 makes the $C^{\gamma 1}$ and C^{δ} atoms move away from the active site to make extra room for the long methylene carbon chain of LL-DAP. This simple rotation of the Ile63 side chain provides the extra room for LL-DAP and allows good van der Waals contacts to be made

between the Ile63 side chain and each substrate. These substrate binding residues are well-conserved among all LL-DAP-ATs. Therefore, the above substrate recognition mechanism is expected to be universal to all LL-DAP-ATs.

Such a recognition mechanism has never been observed and is novel to the subclass I, type I fold family of PLP dependent aminotransferases. One interesting point arose from those structures with the bound substrate analogues is that all of the residues recognizing the dicarboxylate groups came from the same subunit. One of the most important structural characteristics observed in the type I fold family of PLP dependent enzyme is the formation of the homodimer. In AspAT or the aromatic aminotransferase (AroAT), the formation of the homodimer is essential because each carboxylate group of the dicarboxylate substrates is recognized by the residues from different subunits. In fact, LL-DAP-AT is also a functional homodimer, but the recognition of the dicarboxylate groups of its substrates does not require residues from the neighbouring subunit. Perhaps, in LL-DAP-AT, the importance of homodimer formation can be attributed to the recognition of the C^{ε}-amino group of LL-DAP.

As shown by the structure of AtDAP-AT with bound PLP-DAP, all of the residues interacting with the C^{ε}-amino group of LL-DAP essentially come from the neighbouring subunit. Gly95*, Glu97* and Asn309* form direct and water-mediated hydrogen-bonds with the C^{ε}-amino group of LL-DAP. Tyr94*, situated just beneath of the C^{ε}-amino group of LL-DAP, provides some stabilization for LL-DAP through aromatic ring - amino interactions. These residues are found to play a key role in the stereospecific recognition of LL-DAP that effectively discriminates LL-DAP from its diastereomer, *m*-DAP in AtDAP-AT. Therefore, despite the fact that the recognition of the dicarboxylate groups of their substrates are accomplished by the residues from one subunit, the structures of AtDAP-AT with bound PLP-DAP have clearly demonstrated the

absolute requirement of the formation of homodimer in LL-DAP-AT.

Although Tyr94*, Gly95*, Glu97* and Asn309* are definitely important in the recognition of the C^{ε}-amino group of LL-DAP, the presence of those residues themselves can not explain the apparent differences in the substrate specificities observed among different LL-DAP-ATs. There are two types of LL-DAP-AT that exist: one has a broad substrate specificity and the other, a narrow substrate specificity.[7] AtDAP-AT exhibits a narrow substrate specificity and CtDAP-AT shows broad substrate specificity. For example, CtDAP-AT is known to be capable of using *m*-DAP as efficiently as LL-DAP, whereas AtDAP-AT is not.[2] The above residues are almost completely conserved between the two species. Therefore variation of residues is not a factor.

Comparison of the CtDAP-AT and AtDAP-AT structures has provided some of the initial important clues to understanding the differences in substrate specificities. The structure of CtDAP-AT is considerably more flexible than AtDAP-AT. In particular, the loop region containing residues Tyr94*, Gly95* and Glu97* is significantly more flexible in CtDAP-AT. Based on the combined analysis of the structures presented in this thesis and of the other enzymes exhibiting broad substrate specificity, the increased flexibility in that loop region is found to play an important role in accommodating slightly different substrates such as *m*-DAP. In AtDAP-AT, the binding of *m*-DAP is hindered mainly by the steric clash with Tyr94* and the complete loss of hydrogen-bonding interactions. However, in CtDAP-AT, the increased flexibility of the loop region allows those residues to rearrange the hydrogen bonding network and to adopt into the alternative conformations that can comfortably accommodate *m*-DAP.

The structural studies described in this thesis have provided a detailed investigation of LL-DAP-AT in order to understand the overall architecture and the substrate recognition mechanism of LL-DAP-AT. Several structures, including

the substrate analogues-bound structures of AtDAP-AT, have provided a clear picture of the mode of substrate recognition used by LL-DAP-AT. The information obtained from those structures will be very important to the future development of antibiotics targeted against LL-DAP-AT.

At the same time, the investigation of the crystal structure of CtDAP-AT in the flexible, “open” conformation has provided some interesting insights into the differences in the substrate specificities observed among LL-DAP-ATs from different species. The observation of high flexibility in the CtDAP-AT structure was quite stunning and must be considered when designing novel antibiotics targeting this enzyme. Future structural studies including that of CtDAP-AT in complex with different substrates, or the determination of LL-DAP-AT from different species would certainly be needed to increase the understanding of the differences in the substrate specificities in LL-DAP-AT.

5.2 References

1. Hudson, A. O., Singh, B. K., Leustek, T., and Gilvarg, C. (2006). An LL-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant Physiol.* **140**, 292-301.
2. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T., and Maurelli, A. T. (2006). L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17909-17914.
3. Watanabe, N., Cherney, M. M., van Belkum, M. J., Marcus, S. L., Flegel, M. D., Clay, M. D., Deyholos, M. K., Vederas, J. C., and James, M. N. G. (2007). Crystal structure of LL-diaminopimelate aminotransferase from *Arabidopsis thaliana*: a recently discovered enzyme in the biosynthesis of L-lysine by plants and *Chlamydia*. *J. Mol. Biol.* **371**, 685-702.
4. Ford, G. C., Eichele, G., and Jansonius, J. N. (1980). Three-dimensional structure of a pyridoxal-phosphate-dependent enzyme, mitochondrial aspartate aminotransferase. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2559-2563.
5. Watanabe, N., Clay, M. D., van Belkum, M. J., Cherney, M. M., Vederas, J. C., and James, M. N. G. (2008). Mechanism of substrate recognition and PLP-induced conformational changes in LL-diaminopimelate aminotransferase from *Arabidopsis thaliana*. *J. Mol. Biol.* **384**, 1314-1329.
6. Watanabe, A. (2002). Reaction Mechanism of Alanine Racemase from *Bacillus stearothermophilus*. X-ray crystallographic studies of the enzyme bound with *N*-(5'-phosphopyridoxyl)-alanine. *J. Biol. Chem.* **277**, 19166-19172.
7. Hudson, A. O., Gilvarg, C., and Leustek, T. (2008). Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase. *J. Bacteriol.* **190**, 3256-3263.