2010

INSIGHTS INTO THE CATALYTIC MECHANISM OF RETRO-ALDOL CLEAVAGE OF β-HYDROXY AMINO ACIDS BY ESCHERICHIA COLI L-THREONINE ALDOLASE

Remsh Soumya Govinda
Virginia Commonwealth University

INSIGHTS INTO THE CATALYTIC MECHANISM OF RETRO-ALDOL CLEAVAGE OF β-HYDROXY AMINO ACIDS BY *ESCHERICHIA COLI* L-THREONINE ALDOLASE

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

SOUMYA GOVINDA REMESH
Bachelor of Pharmacy, University of Mumbai, Mumbai, India, 2008

Director: Martin K. Safo, PhD
Associate Professor, Department of Medicinal Chemistry

Virginia Commonwealth University
Richmond, Virginia
July, 2010

Table of Contents

Page

List of Tables

Page

List of Figures

Page

Acknowledgement

I would like thank my adviser Dr. Martin Safo for being a patient teacher. Dr. Mohini Ghatge has provided tremendous support and is like family to me. Thank you for being such a wonderful supervisor!!. I am also thankful to Dr. Faik Musayev in helping me understand crystallography in depth. I would also like to thank my committee members, Dr. Umesh Desai and Dr. Walter Holmes for trying to make me a better scientist!! Last but far from least I would like to thank my friends and my family for their unconditional support throughout this emotionally draining journey that is graduate studies.

Abstract

INSIGHTS INTO THE CATALYTIC MECHANISM OF RETRO-ALDOL CLEAVAGE OF β-HYDROXY AMINO ACIDS BY *ESCHERICHIA COLI* L-THREONINE ALDOLASE

By Soumya Govinda Remesh, M.S

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2010

Major Director: Martin K. Safo. PhD
Associate Professor, Department of Medicinal Chemistry

With over 140 vitamin B_6 (Pyridoxal 5'-phosphate, PLP) dependent enzymes, serving vital roles in various transamination, decarboxylation, retro-aldol cleavage and synthesis pathways these enzymes constitute the most versatile catalytic systems in nature. Enzymes of this group have an inherent reaction as well as substrate specificity. A single co-factor namely, PLP is used by specific enzymes of this group to serve distinct roles during the catalytic reaction. An ordered evolutionary adaptation in these enzymes has led to specialization achieved by each enzyme for catalyzing specific reactions.

L-Threonine aldolase (L-TA) is one such PLP- dependent enzyme that catalyzes the retro-aldol cleavage of several β-hydroxy amino acids, although its natural substrates are L-threonine and L-*allo*-threonine with the enzyme having significant preference for L-*allo*-threonine. It also catalyzes racemization and transamination of D-alanine but not of the β-hydroxy amino acids. Thus, the enzyme exhibits both substrate and reaction

specificity. Although, L-TA is frequently employed for stereoselective synthesis of pharmaceutically useful compounds, its reaction mechanism and associated specificity is still not clearly understood.

L-TA from *Escherichia coli* (eTA) is being studied in our laboratory. Our objective is to elucidate the catalytic mechanism of eTA and its mode of substrate and reaction specificity using X-ray crystallography. Another objective is to establish evolutionary relationship of L-TA with other B_6-dependent enzymes, such as serine hydroxymethyltransferase (SHMT) and *Thermatoga maritima* L-TA (TTA) that have the same fold and catalyze similar reactions.

Our structural studies show that while the crystal structures of the two L-TAs are similar, they are significantly different from that of SHMT, especially at the active site. In the L-TA structures, a loop with proposed important active site residue, His126 is replaced by tetrahydrofolate (THF) in SHMT.

The crystal structures of eTA in its native form and in complex with substrate or product have highlighted the importance of His126 in ensuring substrate specificity during retro-aldol cleavage of various β-hydroxy amino acids and His83 or a conserved water molecule to be active site base. Our study emphasizes the molecular level implications of the catalytic mechanism of eTA.

CHAPTER 1 General Introduction

1.1 L-Threonine and L-Threonine aldolase:

L-Threonine is one of the amino acids required for protein synthesis. Since there is no *de novo* pathway for the synthesis of L-threonine in human beings, the amino acid needs to be supplemented for in a protein rich diet. Also, its catabolism (in general, amino acid catabolism) supports gluconeogenesis during starvation. Besides the amino acid is ketorgenic entering important metabolic cycles at various locations.

L-threonine degradation pathway (Figure 1) includes the enzymes L-threonine dehydratase (L-TD) as well as L-threonine 3-dehydrogenase (TDH) and L-threonine aldolase (L-TA). The dehydratase pathway generates α-ketobutyrate from L-threonine which enters the tri-carboxylic acid cycle (TCA) as succinyl-CoA. TDH metabolizes L-threonine to 2-keto 3-ketobutyrate and subsequently yields acetyl-CoA which is also an intermediate in TCA. The two enzymes collectively account for more than 80% of L-threonine metabolism.[1]

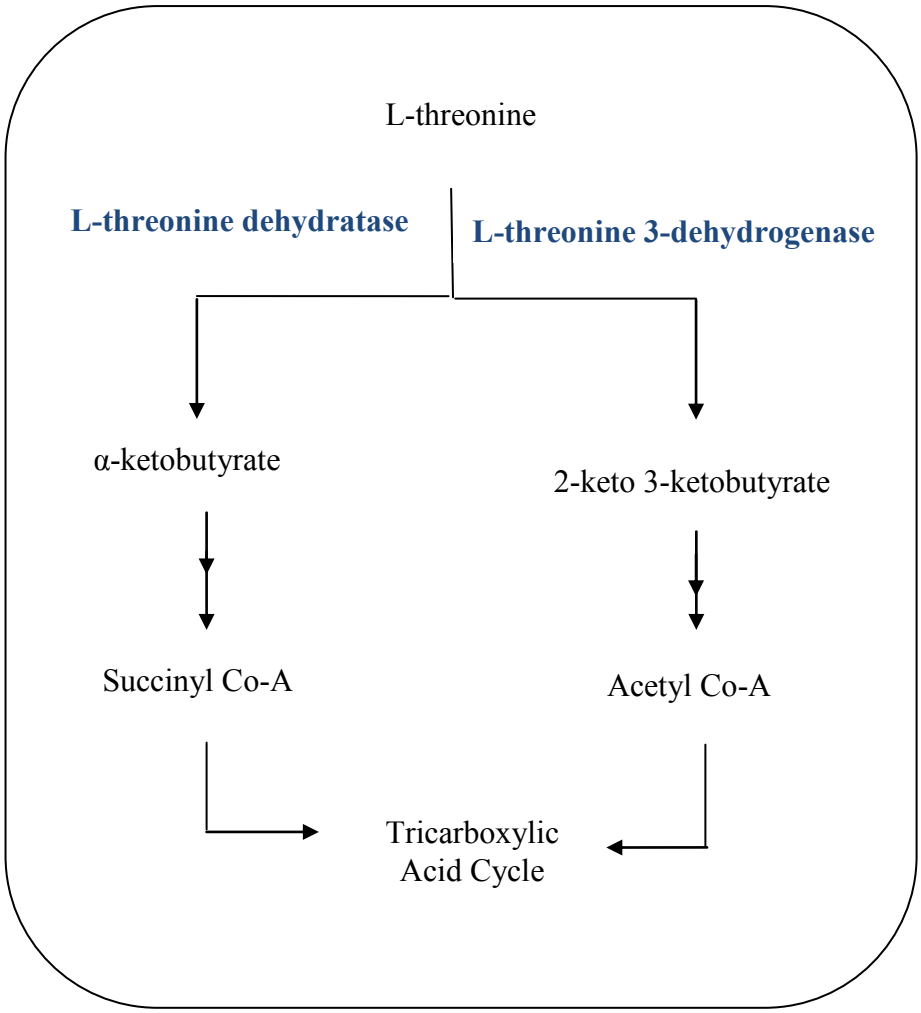


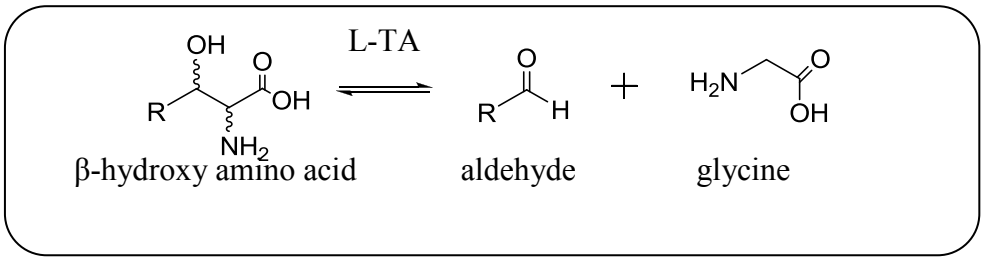
Figure 1: L-Threonine degradation pathway. The major enzymes that are involved in the pathway are shown here.

The third enzyme in threonine degradation, L-TA, has minimal role in threonine metabolism.[1] L-TA belongs to the large group of enzymes that require the cofactor, pyridoxal 5'-phophate (PLP), to function. It catalyzes the retro-aldol cleavage of β-

hydroxy amino acids which may be L-threonine, L-*allo*-threonine or L-serine, although the natural substrates of the enzyme are L-threonine and L-*allo*- threonine. Our interest in the enzyme arises from the fact that, L-TA catalyzes reactions which are considered as 'forced errors' for other enzymes that use PLP as a co-factor.[2]

Orthologous genes encoding L-threonine aldolase from varied microorganisms such as *Escherichia coli*, yeast and *Candida albicans*, have been cloned and expressed.[3] Enzyme activity was detected in tissue preparations and in liver cells in vertebrates. In humans the enzyme is a non-transcribed pseudogene.[4]

The general reaction catalyzed by the enzyme may be represented as;



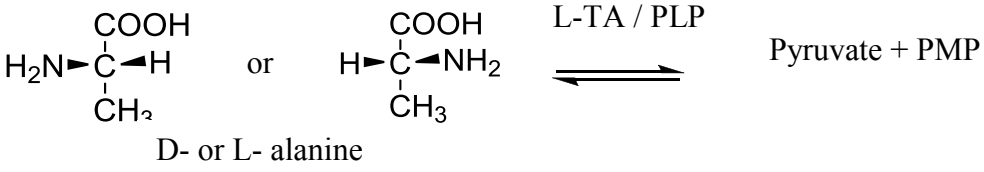
Depending on the source of the enzyme, L-TAs show significantly different substrate specificities. The enzyme isolated from *Lactobacillus bulgaricus* is highly specific for L-threonine while the enzyme isolated from *Aeromonas jandaei* is very selective towards L-*allo*-threonine.[3] *Escherichia coli* L-TA (eTA) being studied in our laboratory, is a low specificity L-TA capable of cleaving both L-threonine and L-*allo*-threonine. Nonetheless, kinetic studies (Table 1) have revealed that L-*allo*-threonine is the preferred substrate.[2]

Table 1: Kinetic data depicting substrate specificity for L-TAs from different sources.[2, 5] *Aeromonas jandaei* selectively catalyzes the retro-aldol cleavage of only L-*allo*-Thr as indicated by the absence of Km value for L-Thr cleavage by *A. jandaei* L-TA

Source Reaction catalyzed	*Escherichia coli* K_m (mM)	*Aeromonas jandaei* K_m (mM)
L-threonine cleavage	10	—
L-allo-threonine cleavage	0.19	0.37

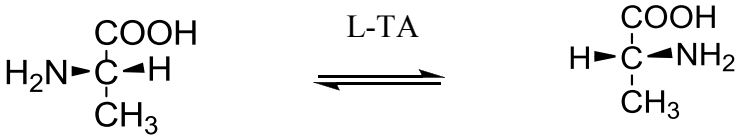
In addition to the retro-aldol cleavage reaction, eTA also catalyzes transamination and racemization of D-alanine.[2] Representative reactions are shown below;

1.Transamination:



D- or L- alanine

2. Racemization:



L- alanine D-alanine

Kinetic studies (Table 2) of the transamination and racemization have revealed that these reactions proceed at much lower rates as compared to the retro-aldol cleavage reaction.

Table 2. Kinetic data depicting rate of transamination and racemization reaction.[2]

Reaction	$k_{cat}(min^{-1})$	K_m (mM)	k_{cat}/K_m (mM^{-1} min^{-1})
L-alanine transamination	0.114	225	0.0005
D-alanine transamination	0.109	77	0.0014
Alanine racemization	0.27	77	0.0035

In summary, eTA preferentially catalyzes the retro-aldol cleavage of L-*allo*-threonine over other β-hydroxy amino acids. It also selectively catalyzes the retro-aldol cleavage of β-hydroxy amino acids over transamination or racemization of D-alanine. It is quite intriguing how the enzyme achieves both reaction and substrate specificity.

1.2 L-Threonine aldolase as a biocatalyst:

Biocatalysis or controlled enzymatic catalysis provides a unique avenue for the stereo- and regio-selective synthesis of chemical moieties. With the advancements in protein and enzyme engineering, biocatalysis has developed into a very promising tool for synthetic chemists. Enzymes purified to homogeneity as well as whole host cells have been employed in synthesis of chemical and pharmaceutical intermediates.[6]

L-Threonine aldolase (L-TA) is widely used as a biocatalyst. L-TA has been employed in the resolution of stereo-isomers as well as in the single step catalytic synthesis of specific stereo-isomers in excess of the other. The latter method generally is more productive in terms of yield and the enantiomeric excess of the required product.[3]

The enzyme finds use in the synthesis of polyoxins[7] and of chiral building blocks for drugs like L-threo-β-(3, 4-dihydroxy) serine for treatment of Parkinson's disease.[8]

The enzyme has found use in the synthesis of ω-carboxy- β-hydroxy-(L)-α-amino acids.[9] L-threonine aldolase has also been utilized in the synthesis of glycoconjugate precursors.[10]

L-TA has been reported as a biocatalyst for the stereoselective synthesis of β-hydroxy amino acids from glycine and aliphatic or aromatic aldehydes. If glycine and aliphatic aldehydes are used as substrates, the *erythro*-β-hydroxy amino acids are the major products, while the use of aromatic aldehydes leads to *threo*-β-hydroxy amino acids as the major products.[11] L-TA purified from *Pseudomonas putida* has provided new routes for the synthesis of industrially important amino acids like L-β-(3, 4-dihydroxyphenyl) serine and L- β-phenylserine.[12] L-TA purified from *Pseudomonas putida* has also been applied to achieve diastreoselectivity and better yields of the versatile γ –halogenated- β-hydroxy-α- amino acids. These compounds have found use as enzyme inhibitors, for control of allergies and for tumor suppression. With glycine, the fluoro-aldehydes gave better diastereoselectivity (in favor of syn conformation) and yield compared to other halogenated aldehydes. Use of higher carbon analogues (C9 and above) of the aldehyde resulted in reversal of diastereoselectivity with the anti conformation being favored. The reversal in diastereoselectivity is attributed to the size of the hydrophobic pocket of the enzyme that lacks favorable interactions with long chain aldehydes.[13]

L-TA catalyzed retro-aldol cleavage reaction is an equilibrium reaction and controlling the reaction in favor of the formation of an aldol product is a challenge. Also, the aldehydes have reduced solubility in the reaction conditions generally employed for carrying out catalysis. The utility of L-TA in organic synthesis is immense at the

laboratory scale but to make it industrially viable further optimization of the enzyme is required. Future work to improve L-TA as a biocatalyst may employ protein engineering to improve the enzyme to achieve better diastereoselectivity and synthetic activity. Also, to solve the problem of solubility of aldehydes, L-TA may be immobilized or a two-phase reaction system may be employed.[3]

1.3 L-TA, a B_6 requiring enzymes:

There are over 140 vitamin B_6 (PLP-dependent) enzymes, serving vital roles in various transamination, decarboxylation, retro-aldol cleavage and synthesis pathways involving glucose and lipid metabolism, amino acid and homocysteine metabolism, heme and DNA/RNA synthesis and neurotransmitter production.[14] Understanding of the functioning of PLP- dependent enzymes has increased remarkably over the last twenty years owing to the large number of sequences and crystal structures of the enzymes becoming available.[15, 16] Enzymes of this group have an inherent reaction as well as substrate specificity. Each enzyme possesses the necessary apparatus to modulate the use of a single co-factor, PLP, in the catalytic reaction. Selectivity in substrates and reactions for each enzyme is a consequence of an ordered evolutionary adaptation in these enzymes. There are extensive reviews available in the literature on PLP-dependent enzyme, including catalysis, reaction specificity and evolutionary relationship between the enzymes.[14-19] These enzymes are described in greater details hereunder;

1.3.1 Structural Diversity within PLP-dependent enzymes with emphasis on fold type I:

PLP-dependent enzymes are categorized into five or six structurally distinct fold types. It has been recognized that there is no correlation between fold types and reaction

types, i.e., enzymes belonging to same fold may represent catalysis of versatile reactions. The most well understood of the six different fold types of this evolutionarily conserved family of enzymes is Fold Type I. Aspartate aminotransferase is the prototype for the enzymes in this fold type.[15, 16] Serine hydroxymethyltransferase (SHMT) and L-TA are other examples of B6 enzymes belonging to Fold Type I.

Figure 2 is a schematic representation of structure of Fold type I enzymes[18]. Fold Type I enzymes catalytically active as homodimers with each subunit made of two domains, namely the large and the small domain. The large domain features seven β-stranded sheets with interlinking α-helices. The small domain comprises of the C-terminus folded into three to four β-stranded sheets with α-helices on one side. The N-terminal also usually contributes to the small domain. The aldehyde group of PLP binds to a conserved lysine residue of the large domain forming an internal aldimine linkage with the amine group of lysine.[18] In general, active site is located at the interface of the two closely associated monomeric subunits forming a tight dimer.[14, 16] Residues from both the subunits contribute to the catalytic mechanism.[18]

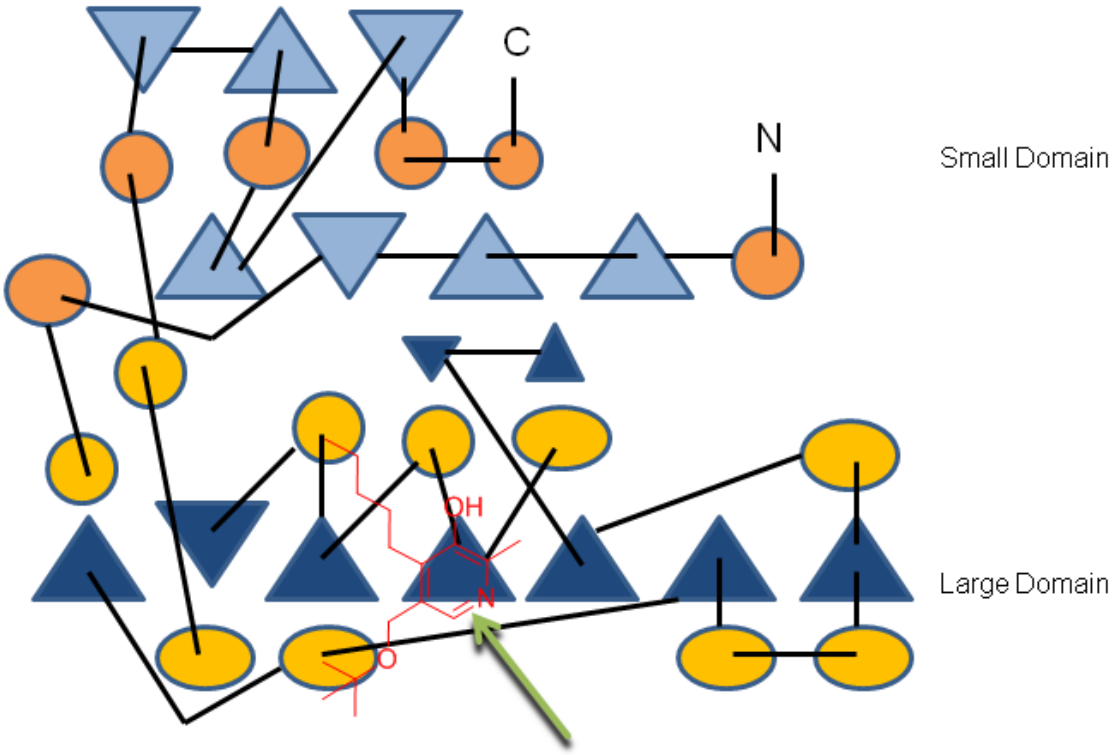


Figure 2: Schematic representation[17] of overall structure of Fold Type I enzymes. Yellow circles are the α-helices and blue triangles are the β-sheets of the large domain. Orange circles are the α-helices and the cyan triangles are the β-sheets of the small domain

1.3.2 Achieving Reaction and Substrate Specificity in PLP-dependent enzymes:

PLP is very reactive and readily forms a Schiff base (an aldimine linkage) with the N-terminus of amino acids through its aldehyde group. This is termed external aldimine. When it forms an aldimine linkage with the active site Lys in the B_6 enzymes, it is called an internal aldimine (Figure 3). PLP is also an 'electron sink' which can withdraw electrons from a wide variety of substrates accounting for its success as a co-enzyme for such a wide array of reactions. But the protein (enzyme) to which PLP is bound as a Schiff base provides the necessary apparatus to ensure reaction and substrate specificity.[14]

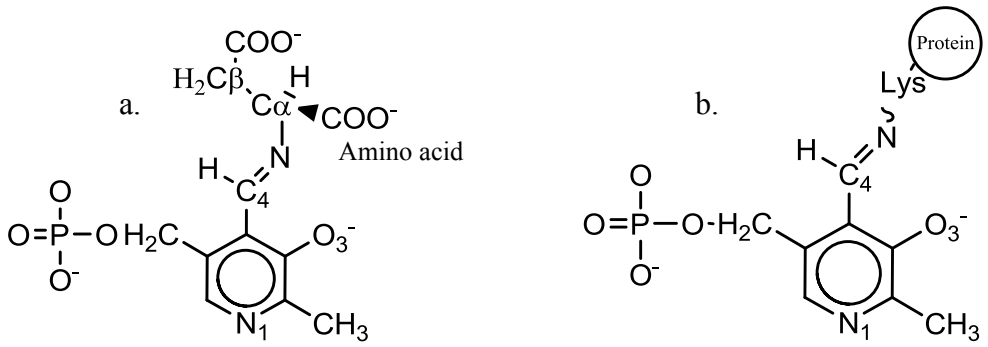


Figure 3: Schematic representation[14] of Schiff base between a. with N-terminus of amino acid (forming an external aldimine) and b. with N-terminus of Lys residue of the protein (forming an internal aldimine)

Several general hypotheses have been proposed to explain the reaction and substrate specificity achieved by the PLP-dependent enzymes. Two of these are briefly explained below.

Dunathan hypothesis:

Dunathan proposed that the geometry of aldimine linkage directs the bond cleavage in reactions catalyzed by enzymes using PLP as a co-factor. He suggested that the bond to be broken aligns perpendicular to the plane of the aldimine linkage and pyridine ring of PLP. He reasoned that this ensures maximum σ-π interaction during transition state.[14] Thus, one way to distinguish between deprotonation and decarboxylation, for example is by orienting the substrate differently in these different reactions (Figure 4).

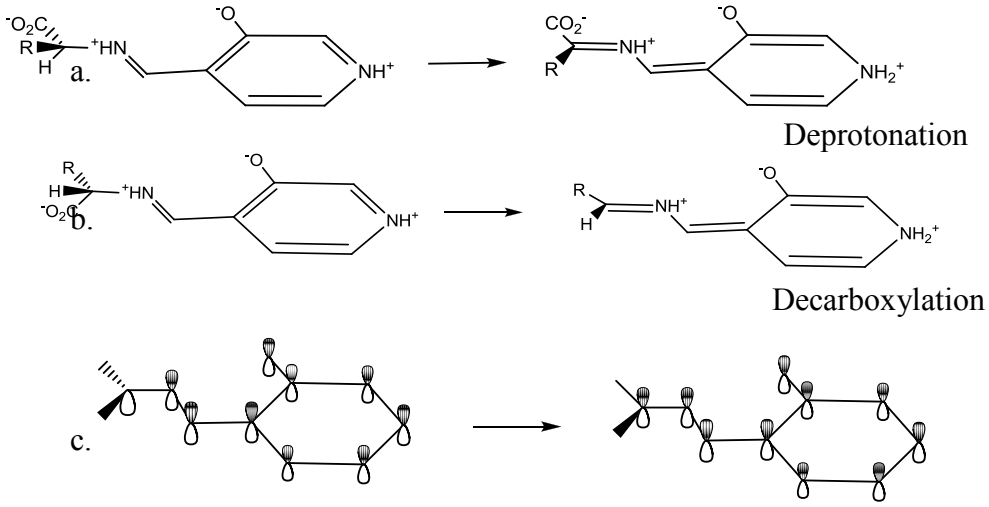


Figure 4: a. Substrate orientation for deprotonation reaction; b. Substrate orientation for decarboxylation reaction; c. Schematic representation of σ-π orbital overlap for maximum interaction[17]

Closed conformation:

Studies with aspartate aminotransferase[20] and SHMT[18, 19, 21] indicate that during the enzyme catalyzed reaction there is an equilibrium between an open and closed state of the enzymes. It is generally accepted for aspartate aminotransferase that with C4 substrates like L-aspartate the enzyme undergoes a conformational change from open to closed form and the entire reaction proceeds in the closed state. In case of SHMT, reaction specificity is achieved when the enzyme is in the closed conformation.[18, 19, 21]

The schematic in Figure 5 describes the open and closed structures specific for aspartate aminotransferase.[20]

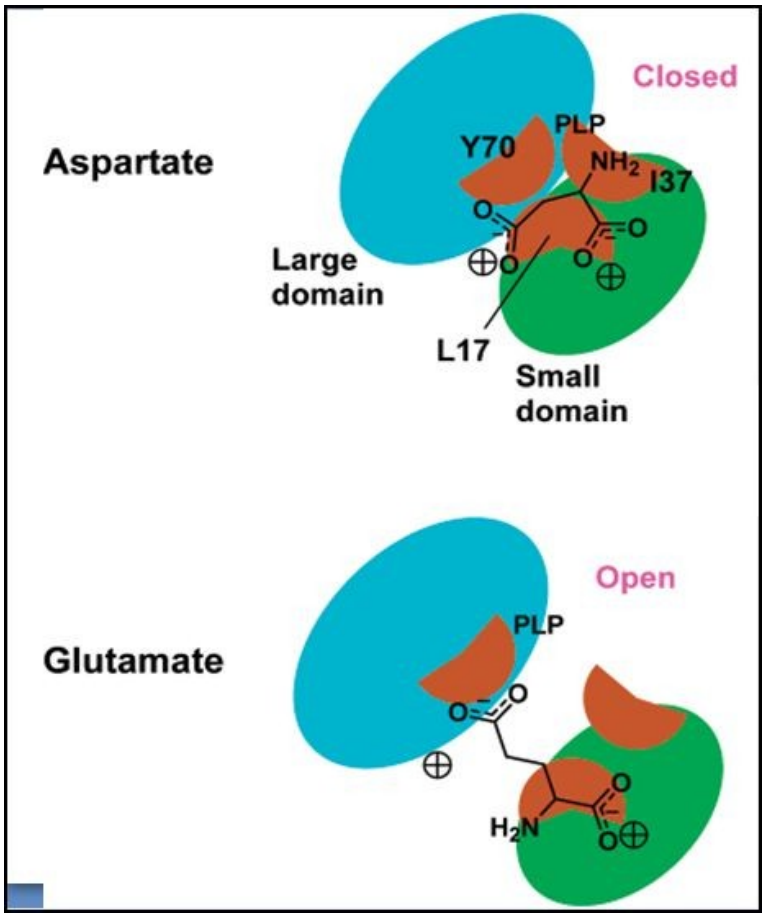


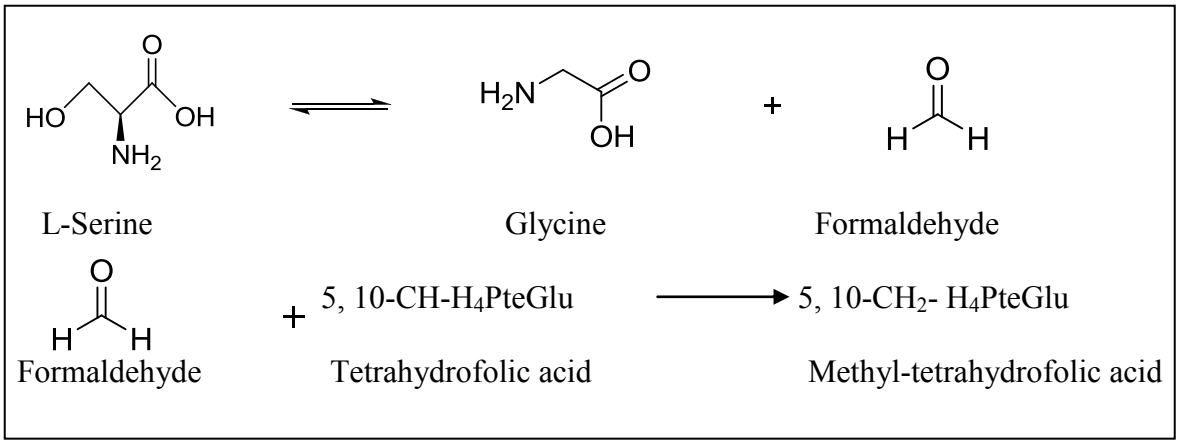
Figure 5: Schematic representation of **open and closed s**tructures in aspartate aminotransferase.[20]

1.3.3 Serine hydroxymethyltransferase and L-TA:

There are excellent and extensive reviews available on the studies of SHMT.[22-24] Most of the SHMTs characterized to date are either homotetramers (rabbit SHMT) or homodimers in solution (*E. coli* SHMT).

Historically, SHMT was thought to possess residual L-TA activity[25]. Recently, L-TA has been established as a 'genuine' enzyme with L-TA from several microbial sources being cloned and expressed.[3, 26-28]

SHMT is a well established enzyme in serine metabolism pathway. The physiological reaction catalyzed by SHMT is the retro-aldol cleavage of L-Serine to form glycine and formaldehyde in the presence of a second co-factor, tetrahydrofolate (THF), a one-carbon donor that participates in nucleic acid biosynthesis. The formaldehyde then reacts non-enzymatically with the bound THF at the active site.[21-24, 29] The enzyme fails to catalyze this reaction in the absence of THF. The retro-aldol cleavage reaction of L-serine by SHMT is represented below.



In the absence of THF, SHMT can function like L-TA, catalyzing the retro-aldol cleavage of other amino acids like L-threonine and L-*allo*-threonine.[30] Racemization and transamination reactions of D-alanine by SHMT are also known.[31] The specificity of SHMT towards retro-aldol cleavage of L-Serine has perplexed scientists and has led to detailed study of SHMT to understand its mechanistic implications.

Figure 6 depicts the crystal structure of eSHMT which shows external aldimine linkage to PLP with product glycine after a retro-aldol cleavage of the substrate L-Serine.[28]

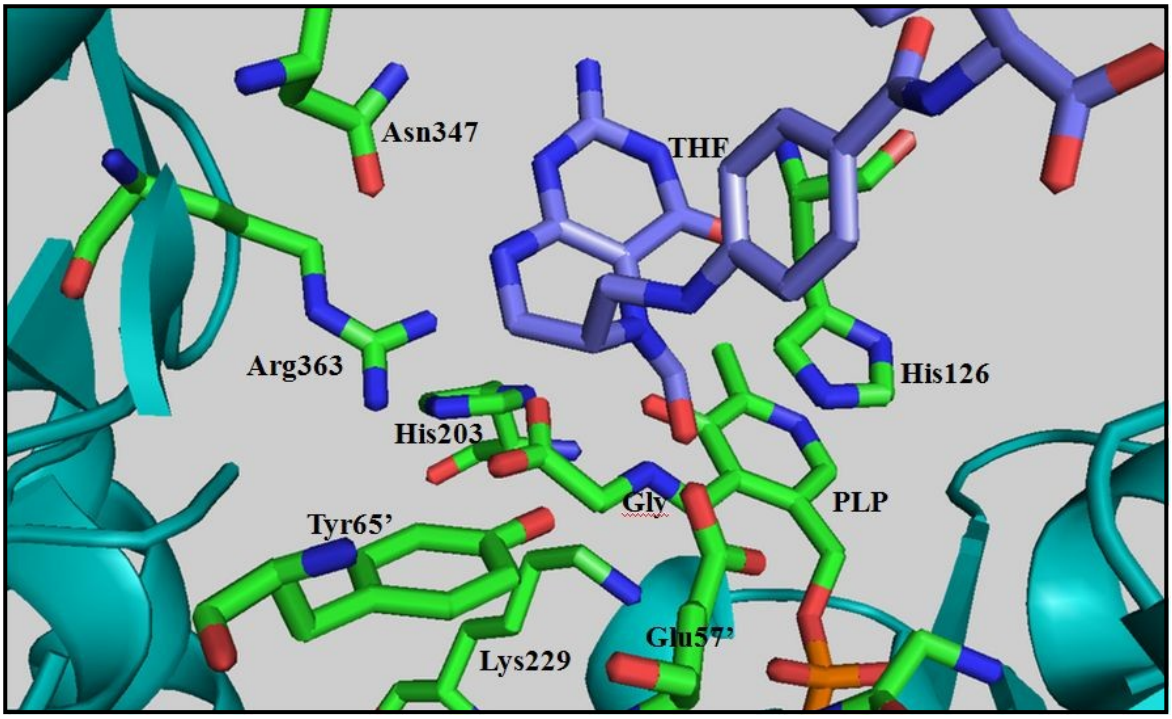


Figure 6: THF (purple) is seen in the figure along with other active site residues (green)

Specific residues at the active site of SHMT have been established as controllers of the reaction of the enzyme and research has been focused on the stereochemistry at Cβ carbon of β-hydroxy amino acids (Figure 6).[32-41] The catalytically important unit of eSHMT is a tight dimer formed by two monomeric subunits. A conserved lysine residue (Lysine 229 of eSHMT) forms an internal aldimine linkage to PLP (not shown in Figure 6). The pyridoxal ring of PLP and the side chain of His 126 of eSHMT are coplanar and stack on top of each other. THF binds close to PLP at the active site and its pteridine ring

lies parallel to the pyridoxal ring of PLP although not stacking on it directly. Carboxylate group of the product glycine bound as an external aldimine to PLP makes hydrogen bond interactions with Arg363. One of the oxygens of the carboxylate group of glycine also makes hydrogen bond interaction with Ser35 from the same subunit and with Tyr65 from the other subunit while the other oxygen interacts with His203 from the same subunit. Glu57 at the active site has been implicated to be important in the transfer of the formyl moiety to THF.[22, 37]

It is generally accepted that SHMT achieves reaction specificity by maintaining a closed conformation of the active site during catalysis of L-Serine in the presence of THF, the physiological reaction catalyzed by SHMT. Nevertheless, the open structure accepts other β-hydroxy-amino acids for retro-aldol cleavage in the absence of THF. Racemization and transamination reaction also proceeds with the enzyme in the open conformation.[21, 22]

A generally accepted mechanism for the retro-aldol reaction catalyzed by SHMT is illustrated below (Figure 7).[32-41] Reaction is initiated after the substrate forms an external aldimine with the aldehyde group of PLP with the Cβ hydroxyl group of L-Serine perpendicular and synperiplanar with the π-system of PLP.[32] The N_5 of THF seems to make direct nucleophilic attack on the Cβ carbon of bound L-Serine. A quinonoid intermediate is formed after the proton abstraction is complete which is detected at a wavelength of 490nm.[33] This intermediate rearranges to the product bound PLP-external aldimine. This step involves a proton donor, B3H (in Figure 7) which saturates the Cα of the quinonoid intermediate. The identity of B3H has is not clear yet. The formaldehyde

released subsequently reacts with THF and Glu57 has been identified as B_1H, the residue involved in the transfer.[22, 37]

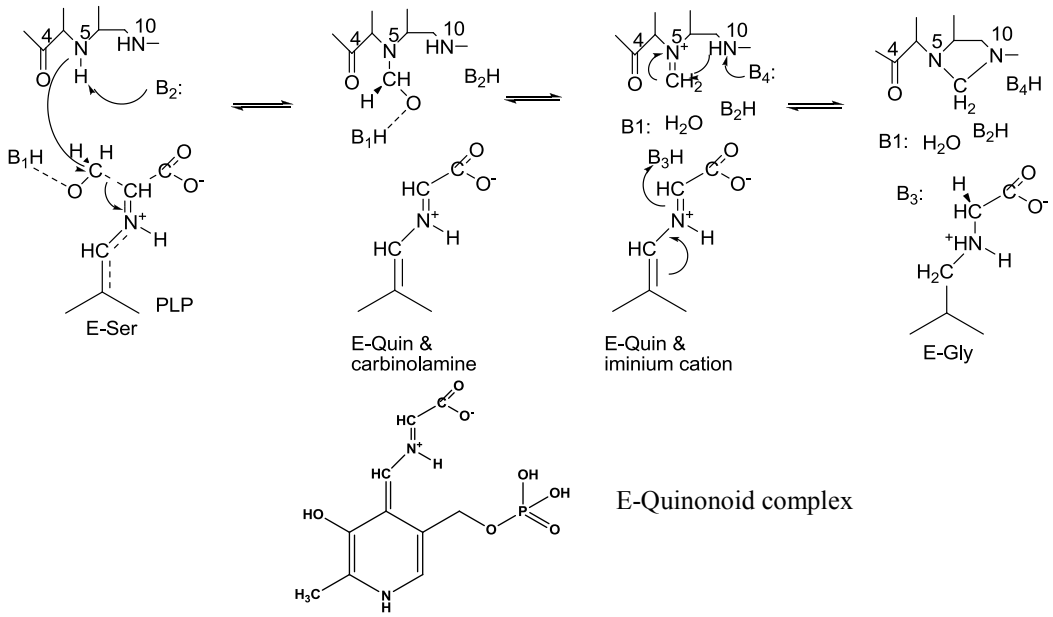


Figure 7: Schematic representation of proposed mechanism of conversion of L-Serine to glycine in the presence of THF.[22]

Threonine aldolase does not participate in the one carbon transfer reactions as SHMT since L-TA does not bind THF. Kinetic studies by our collaborators at the University of Rome have provided us with a wealth of information about the substrate preferences for the two enzymes.[2] Table 3 summarizes the results of comparative study between the two enzymes.

Table 3: Kinetic studies of eTA and eSHMT summarizing the substrate preference for either in a retroaldol cleavage reaction. L-Thr or L-*allo*-Thr cleavage is catalyzed by SHMT in the absence of THF.[2]

	eTA			eSHMT		
Reaction catalyzed	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1}/min^{-1})	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1}/min^{-1})
L-threonine cleavage	62	10	6.2	4.3	43	0.1
L-allo-threonine cleavage	376	0.19	1980	30	1.5	20
L-serine cleavage						
1.(-) THF	1.9	16	0.11	ND	0.8	ND
2.(+) THF	-	-	-	640	0.3	2130

As seen from the kinetic data in Table 3, SHMT is specific for L-serine cleavage whereas L-TA has preference for L-*allo*-threonine. Even though SHMT has different substrate specificity as compared to L-TA, it provides an excellent model for critical evaluation of L-TA and to evaluate the origin of reaction and substrate specificity in PLP-dependent enzymes.

1.3.4 *E.coli* L-TA (eTA) and *Thermatoga maritima* (TTA):

The two most well characterized L-TAs are those from *E. coli* and *T. maritima.* The sequence identity between TTA and eTA is 24.1% and the alignment of the sequences is shown below in Figure 8.[42] The TTA structure has already been solved.[43] TTA enzyme like eTA is a low specificity L-TA. The TTA structures show both internal as well as external aldimine formation (Figure 9, only internal aldimine is shown in the

Figure) and also asymmetry in the substrate binding to different subunits of this tetrameric protein. Approximately 15^{o} rotation of the PLP ring is seen between the internal and external aldimine structures. Carboxylate group of the substrate bound to PLP as an external aldimine interacts with Arg171, Arg316 and Ser6. These residues are thought to hold the substrate in place. Also, His83 stacks on top of the PLP pyridoxal ring and hydrogen bonds to the substrate. Another residue at the active site is His125, but has not been implicated in any mechanistic function.

Another active site residue, Tyr87 has been suggested to be involved in ensuring preference for L-*allo*-threonine through hydrophobic interaction with the methyl group of the amino acid.[43] Nevertheless, our analysis of the TTA structure did not show any close interaction between the L-*allo*-threonine and this hydrophobic residue. Even with low sequence similarity the TTA structure serves as a good model for determination and evaluation of eTA structure.

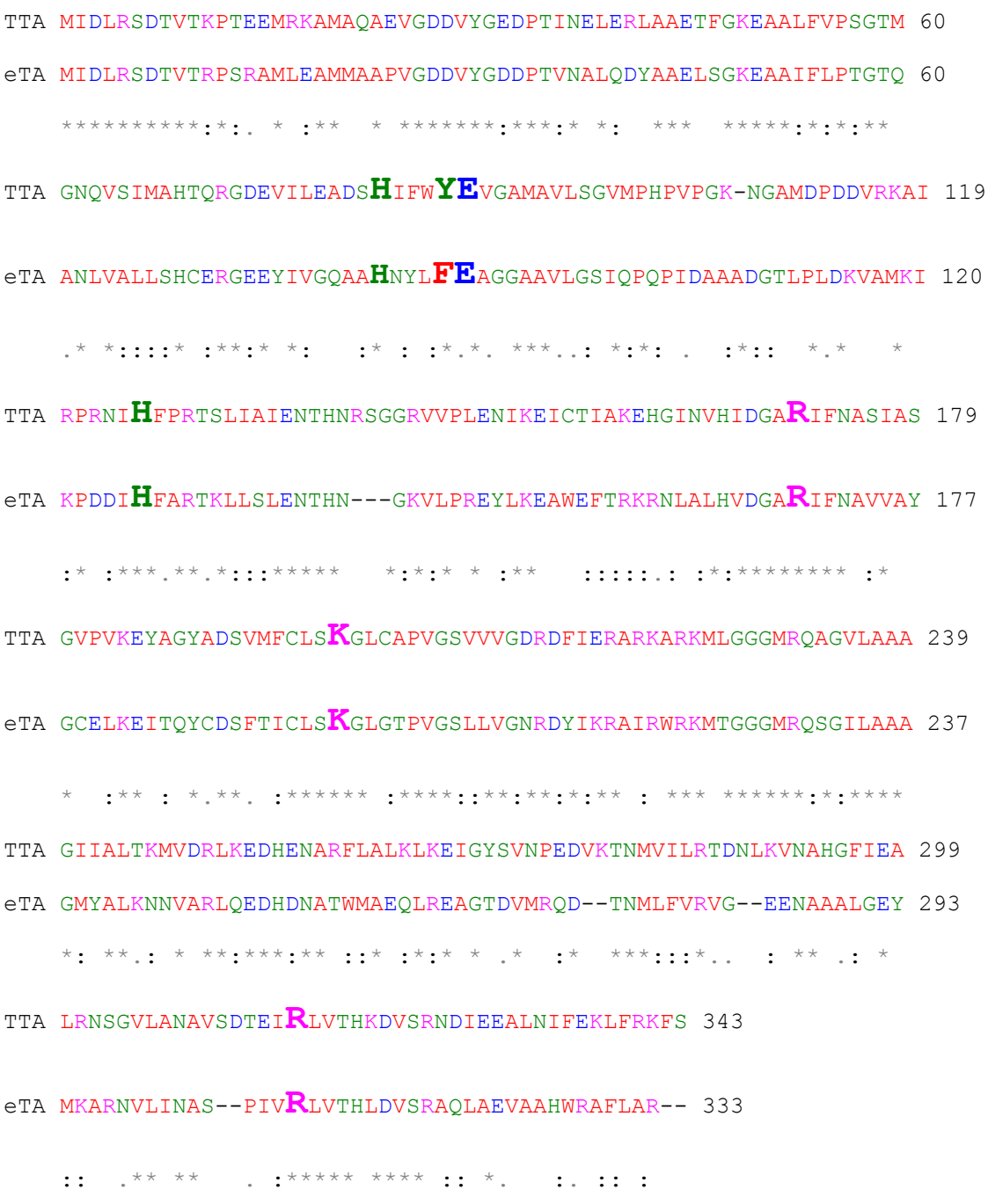


Figure 8: Sequence alignment diagram of eTA with TTA. The residues in bold are conserved in the active both the enzymes.[42]

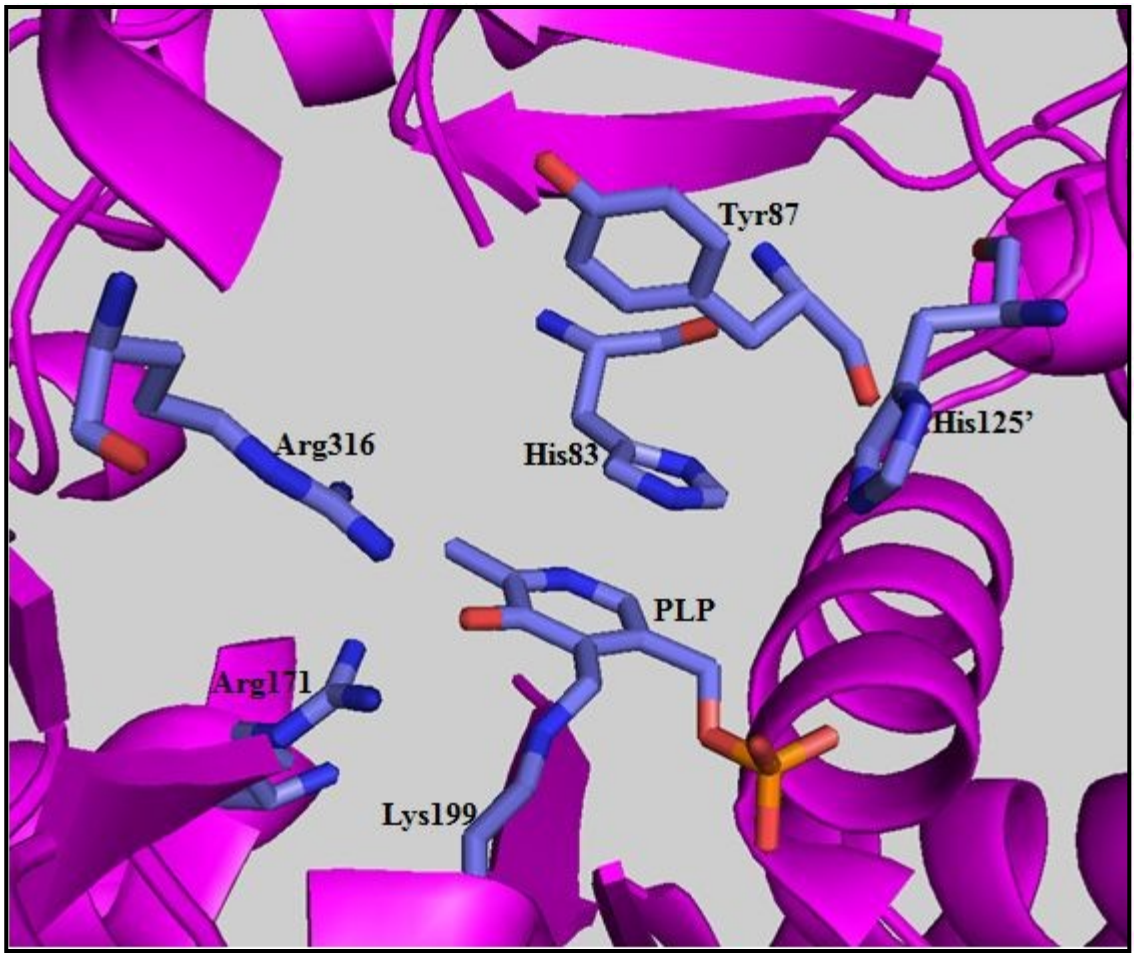


Figure 9: Shown above is the active site of native structure of *Thermatoga maritima* L-TA. Only internal aldimine formation is shown here.

This enzyme is thermophilic and the mechanism of catalysis may be different for the enzyme at elevated temperatures than that proposed at room temperature. Nevertheless, the enzyme served as a good candidate to further our understanding of the catalytic machinery which is similar to that for eTA.

1.4 Conclusion:

Although, L-threonine aldolase does not contribute significantly to threonine metabolism, it has the potential to be an industrially important biocatalyst since good stereoselectivity is achieved during the single step L-TA catalyzed retro-aldol cleavage of glycine and assorted aldehydes at the laboratory scale. Immobilized and engineered enzymes would improve diasteroselectivity as well as synthetic activity of L-TA and make it industrially viable.

L-TA being studied in our laboratory is obtained from *Escherichia coli (*eTA) and belongs to Fold Type I of the large family of PLP-dependent enzymes. It has selectivity towards retro-aldol cleavage of L-*allo*-threonine although other β-hydroxy amino acids may also act as substrates. eTA also catalyzes transamination and racemization reactions. The enzyme shows both substrate and reaction specificity. Elucidating the mechanism of the enzyme catalyzed reaction would aid in understanding how eTA achieves such stereoselectivity. Other closely related PLP-dependent enzymes like SHMT and L-TA from *Thermatoga maritima* are used for comparative analysis with eTA to develop a better appreciation of its reaction mechanism. Also, comprehension of the enzymatic reaction in greater details would help in developing the enzyme as a biocatalyst as well as establish an evolutionary relationship between the PLP-dependent enzymes of Fold Type I.

Specific Aim:

Elucidate the reaction mechanism of *Escherichia coli* L-threonine aldolase using X-ray crystallography:

Our objective is to use site-directed mutagenesis, enzyme kinetics and X-ray crystallography, as well as comparative analysis with other PLP-dependent enzymes that catalyze similar reactions as eTA to understand its catalytic mechanism at the molecular level. The specific aim of the current study is to **elucidate the reaction mechanism of *Escherichia coli* L-threonine aldolase using X-ray crystallography.**

CHAPTER 2 Structural Study of *Escherichia coli* L-threonine aldolase (eTA)

2.1 Crystallography and its utility in establishing the reaction mechanism of eTA:

Crystallography:

X-ray crystallography is a well established tool in structural biology. Of around 65,000 macromolecular structures deposited in the Protein Data Bank, more than 56,000 have been determined using X-ray crystallography as the experimental method (http://www.pdb.org/pdb/search/advSearch.do). Development of high speed computers and crystallography related programs and the necessary software have made the job of a crystallographer much easier as compared to the early days of Max Perutz who won a Nobel prize for solving the structure of haemoglobin in 1962.[44]

A brief description of macromolecular X-ray crystallography is provided here. In general, an object diffracts radiation whose wavelength is comparable to the size of the object. Thus, X-rays whose wavelength is comparable to ultra small protein molecules serve as the radiation source to obtain images of these molecules. Also, since a single molecule does not diffract strongly enough to be detected, there is a need for a crystal composed of a wide array of ordered molecules oriented identically that diffract together to produce strong detectable reflections. A lens system to converge X-rays (as seen in a microscope where the radiation source is white light) is unavailable and readily interpretable images are not obtained directly. To circumvent this problem, computer simulated mathematical model of the protein is developed using a wide array of programs and related software.

In crystallography, information about the intensity as well as the phases of the recorded reflections is needed for structure determination. The intensity of individual reflections is obtained directly from the diffraction patterns recorded on the detectors. The phase information is a little more difficult to obtain. Isomorphous Replacement (Heavy atom replacement), Anomalous Scattering and Molecular Replacement are employed to solve the phase problem. In Isomorphous Replacement, a heavy atom (eg: selenium) is introduced into the protein and the positions of the atoms of the protein are calculated based on the diffraction by the heavy atom. In Anomalous Scattering, the property of heavy atoms to absorb X-rays of specific wavelengths is taken advantage of and the positions of atoms of the protein are then determined based on the positions of the heavy atoms. Molecular replacement is used when the unknown structure of a protein is solved using the known structure of a similar protein.[45]

Crystallography in determining the mechanism of eTA:

Since crystallography provides a molecular picture of the protein with a wealth of information about the active site we explored structural studies with X-ray crystallography to understand the catalytic mechanism of eTA. We have used the structures of eTA in its native form as well as bound to its substrate(s) to draw inferences about the differences at the active site, if any in the presence and absence of ligands and understand the mechanism of the enzyme at the molecular level.

Obtaining ligand bound crystals:

Protein crystals contain more than 50% solvent and readily allow the diffusion of ligands (substrate or product) to reach the active site. Co-crystallization of protein and ligand is extensively used to obtain crystals with substrate or product bound. Another

technique frequently employed to achieve crystals with bound ligands is soaking the native crystals in a buffer with a predetermined concentration of the ligand. This technique is simple but not very elegant in that the crystal may dissolve, become disorganized and/or lose its diffraction ability.[46, 47] Another approach is to use a cryo-protectant with a predetermined concentration of the ligand. Cryo-protectants are used to protect the crystal from degrading when impinged by high intensity X-rays. Cryo-protectants are employed just prior to setting the crystal in X-ray beam and can readily aid in penetration of the ligand to the active site prior to data collection.[48] Yet another technique used to obtain ligand bound crystals is to synthesize the substrate bound to the active site of an enzyme from a crystal containing its product if the enzyme catalyzes a reversible reaction.[43] We have used co-crystallization to obtain eTA bound to different substrates and have solved structures of eTA with product as well as substrate bound.

There is only one other L-TA structure from *Thermatoga maritima* deposited in the Protein Data Bank in the native form as well as bound to the substrate.

2.2 Purification, Crystallization, Data Collection and Structure Determination of eTA:

Clone of *Escherichia coli* L-threonine aldolase (eTA, molecular weight per monomer is 36,500 Dalton) was provided by our collaborators. Expression and purification of eTA was carried out using a procedure developed by our collaborators in University of Rome[4] but slightly modified to suit our laboratory set up.

2.2.1 Purification of E.coli L-TA (eTA):

HMS174 cells transformed with L-threonine aldolase (from *Escherichia coli)* over expressing pET22 plasmid was used to grow an overnight culture (300mL). The culture was inoculated (1:20) into 6L of Luria-Bertani broth (non-buffered, containing 100mg/mL ampicillin and 30mg/mL pyridoxal 5'-phosphate). Bacteria were grown at 37^0 C under aerobic conditions until the OD_{600} reached 0.5. The cells were induced with 0.05mM isopropyl-thio-β-D-galactosidase (IPTG) to express eTA. Bacteria were harvested after 20hrs. Cells were suspended in a buffer containing 10mM Tris/HCl (pH 7.6), 1mM NaCl and 1mM EDTA and cell lysis was carried out in Avestin French Press followed by centrifugation at 12,000 rpm for 20 mins. Streptomycin sulfate (10 g/L) was added to the supernatant in order to precipitate DNA. The cell extract was centrifuged at 12,000 rpm for 20 mins. , the pellet discarded while ammonium sulfate was added to the supernatant to a final concentration of 313 g/L (50% saturation). The solution was centrifuged at 12,000 rpm for 20 mins. and ammonium sulfate was added to the supernatant to 75% saturation (176 g/L). After centrifugation, the precipitant was dissolved in 20mM Potassium Phosphate buffer pH7.5 which was dialyzed against the same buffer with an overnight dialysis change of 1L. The diasylate was centrifuged and

the supernatant was loaded onto DEAE-Sepahdex column (5 x 15 cm.) which was equilibrated with 20 mm potassium phosphate. 100 mL of the equilibration buffer was used to wash the column. Elution was carried out with a 1L linear gradient of 20mM potassium phosphate pH 7.5 with increasing salt concentration up to 0.4M NaCl. Fractions with highest A_{420}/A_{280} were pooled and precipitated by ammonium sulfate addition to 75% saturation followed by centrifugation. The pellet was dissolved in 50mM potassium phosphate pH 7.0 and was loaded onto phenyl-Sepharose (3 x 12 cm.) The column was pre-equilibrated with 50 mM potassium phosphate containing 176 g/L ammonium sulfate. The elution buffer was a 600mL linear gradient of equilibration buffer and 20mM potassium phosphate pH 7.0 with no salt. Fractions of highest purity as judged by single band on SDS-page gel (Figure 10) were pooled and concentrated with 75% ammonium sulfate. Following centrifugation, the pellets were stored at -20^0C and dialyzed just prior to crystallization.

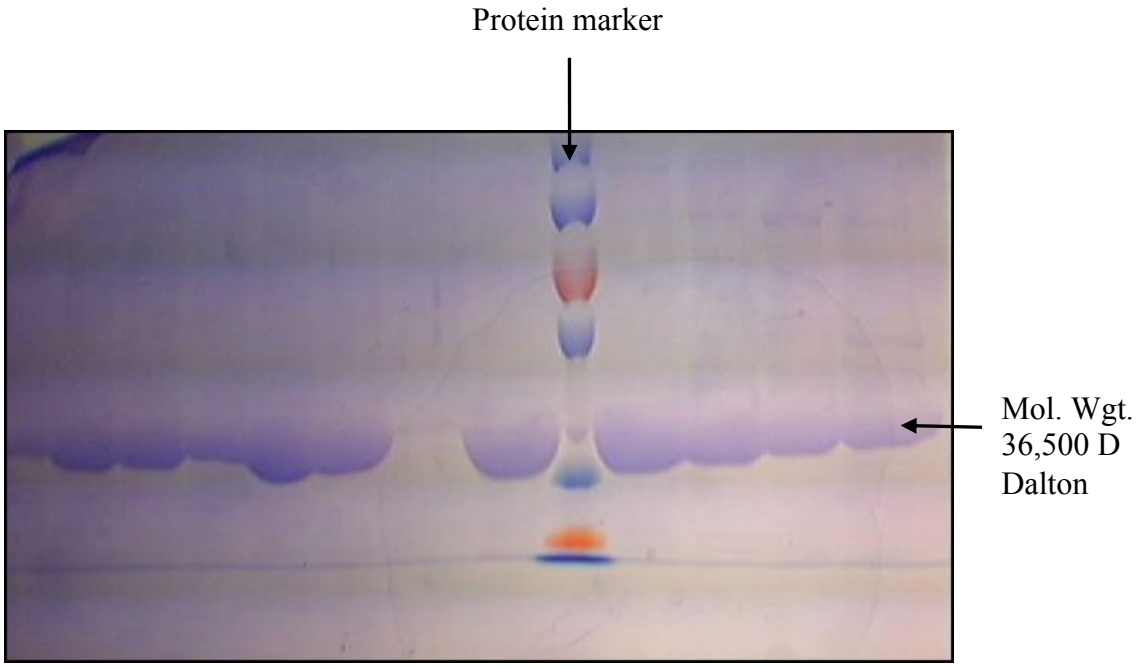


Figure 10: SDS polyacrylamide gel with each lane showing the different fractions of eTA collected after purification with Phenyl-Sepharose column. These fractions were pooled together and the protein was precipitated with ammonium sulfate. The final dialysis was carried out in 20mM Potassium Phosphate pH 7.0.

2.2.2. Crystallization of native as well as ligand bound eTA:

Protein Quantification, Purity and Activity:

Pure and homogenous protein solution is a pre-requisite for successful crystallization. Thus, eTA was dialyzed in 20mM potassium phosphate and its purity was checked using SDS-Page gel electrophoresis. The enzyme was found to be 90% pure. Shown below are the pictures of the SDS-page gel before as well as after purification (Figure 11).

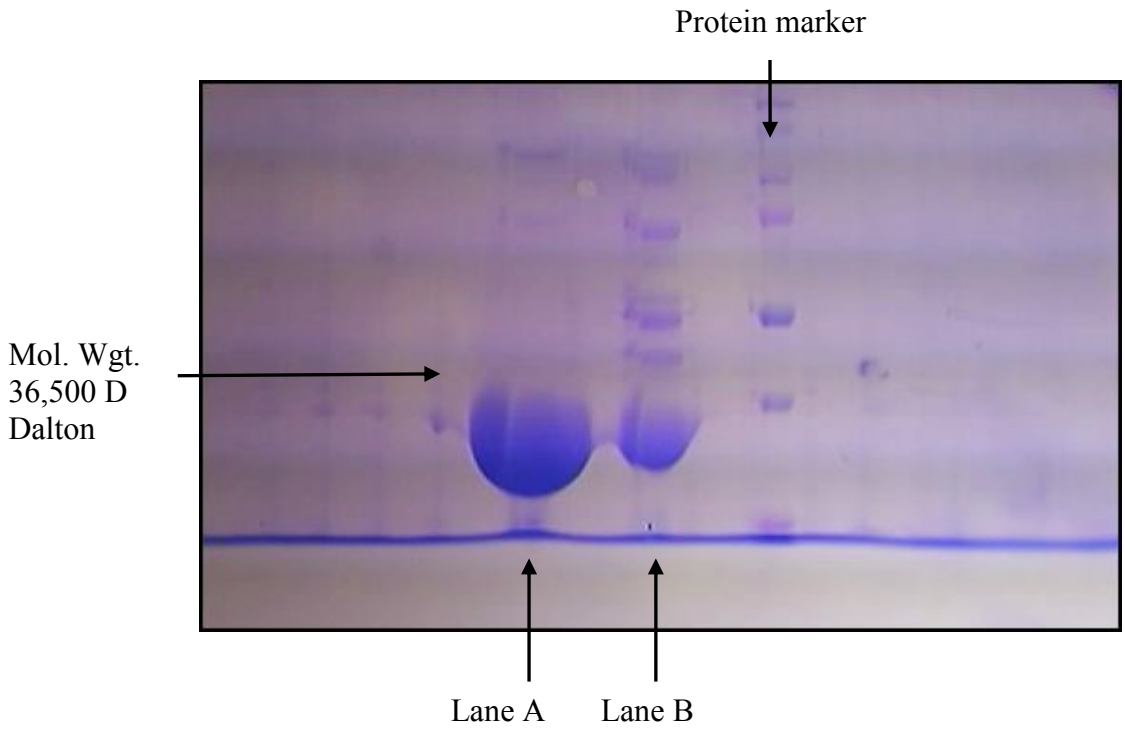


Figure 11: SDS polyacrylamide gel showing a fraction before purification with Phenyl-Sepharose column (Lane B) and after purification with Phenyl-Sepharose column followed by final dialysis in 20mM Potassium Phosphate pH 7.0 (Lane A)

To ensure that the enzyme is stable and functional, the enzyme activity was determined prior to crystallization. The activity of the enzyme dialyzed against 20mM potassium phosphate was determined as per the previously reported procedure for SHMT.[49]

eTA activity: L-*allo*-threonine which is a substrate for eTA undergoes retro-aldol cleavage in the presence of enzyme. The enzyme alcohol dehydrogenase also present in the reaction mixture converts the product of the retro-aldol cleavage reaction, acetaldehyde to ethanol in turn converting NADH into NAD^+. The decrease in absorbance at 340nm corresponding to the conversion of NADH to NAD^+ is followed and

thus activity of eTA is determined. Table 4 shows the observed activity for the purified eTA.

Table 4: Table shows the rate measurement for the activity assay performed for eTA

Enzyme (10mg/mL)	ADH (3mg/mL) + NADH(6 mg/mL)	L-*allo*-threonine (50mM)	Buffer (Na-BES pH 7.4)	Rate (min^{-1})
20 µL	30 µL	50 µL	900 µL	-2.104

Crystallization Experiment:

The enzyme was dialyzed against 20mM potassium phosphate pH 7.0. Following dialysis the enzyme was quantified using the relation;

1AU = 0.983 mg/mL

Molar concentration was calculated using;

13mg/mL = 90 µM (90% pure)

Crystallization Buffer*:* The enzyme (eTA) was screened against the crystallization buffer conditions in Screen I (40 conditions) and Screen II (40 conditions) provided in the Crystal Screen Macromolecular Crystallization kit by Hampton Research. Each of the conditions is composed of a precipitant in a buffer with or without the presence of a salt.

Crystallization*:* Both 'hanging drop' and 'sitting drop' vapor diffusion methods were attempted. Hanging drop method provided better results. A schematic representation of the 'hanging drop' method is provided below (Figure 12).

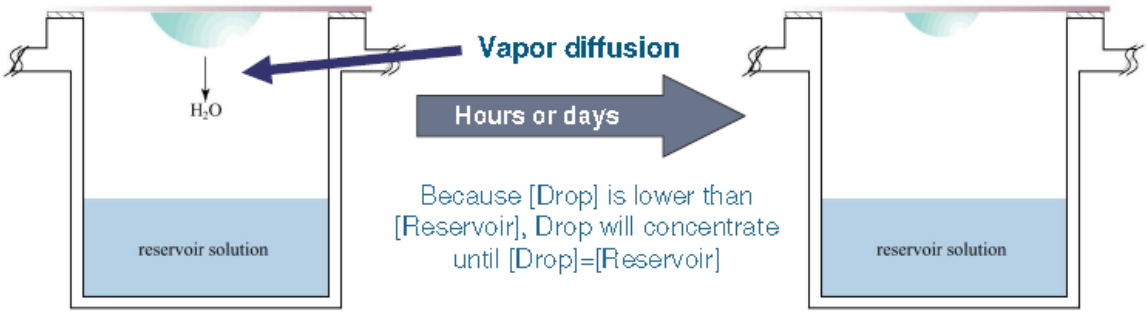


Figure 12: Schematic representation of Hanging Drop vapor diffusion method for protein crystallization. The schematic is reproduced from Qiagen handbook for Protein crystallization. The drop on the cover slip contains the protein and precipitant, while the reservoir solution contains the precipitant.

The reservoir contains the precipitant (550µL) while each drop is comprised of a mixture of protein and precipitant. Several combinations of protein (22mg/mL, 154µM or 10mg/mL, 70µM) to crystallization buffer ratios were used in the drop of which 1:1 ratios provided better crystals. Schematic representation of a crystallization set up is shown in Figure 12. In one drop 2µL of protein of concentration 22mg/mL (154µM) was mixed with 2µL of the crystallization buffer while in the other drop 2µL of protein of concentration 10 mg/mL (70µM) was mixed with 2µL of the crystallization buffer (Figure 13). To obtain product or substrate bound crystals, protein was equilibrated with the amino acid substrates (L-Thr and L-Ser) just prior to crystallization. The ligand concentration was between 6-30 mM.

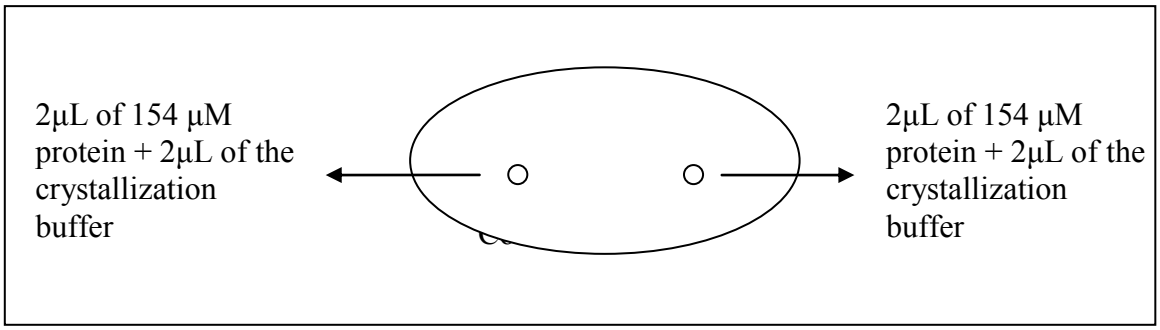


Figure 13: Schematic representation of a crystallization set up.

The crystallization conditions that provided single diffractable crystals are summarized in the results section (Section 2.3.1). Crystals were obtained for native eTA as well as eTA with L-Ser at pH 5.6 (eTA/ Ser5.6), eTA with L-Ser at pH 7.5 (eTA/ Ser7.5) and eTA with L-Thr at pH 7.5 (eTA/ Thr7.5).

2.2.3 Data Collection:

The general procedure followed for data collection of each of the four crystals mentioned above (native eTA; eTA/ Ser5.6; eTA/ Ser7.5; eTA/ Thr7.5) was similar with minor modifications. The crystals were cryo-protected, and flash cooled in a nitrogen stream. For three of the four crystals the crystallization buffer contained PEG 400 which served as a cryo-protectant whereas 5uL of 100% glycerol was needed for protecting the fifth one (see table Table 5 for details about the data collection protocol).

Diffraction data was collected for the cryo-protected crystals at 100K using a Molecular Structure Corporation (MSC) X-Stream Cryogenic Crystal Cooler, a Rigaku IV ++ image plate detector, and a Rigaku MicroMax-007 X-Ray source (copper) fitted with MSC Varimax Confocal optics operating at 40 kV and 20 mA. Collection of

redundant data was imperative since all the crystals obtained were thin and flat and data was lost when the plane of the crystals were aligned parallel to the X-ray beam.

Initial Indexing:

Initial still frames were indexed for each of the four crystals to determine their respective space groups and unit cell dimensions. The space groups and the unit cell dimensions were then used to decide the number of oscillation frames to be collected as well as the crystal-detector distance. Three of the crystals (native, eTA/ Ser7.5 and eTA/ Thr7.5) indicated an orthorhombic space group requiring a minimum data of 90^0. On an average 240 oscillation frames were collected from φ 0^0 to 120 ^{0} at 0.5^0 intervals for each of the three crystals. The crystal of eTA/ Ser5.6 complex indicated a monoclinic cell requiring a minimum data of 180^0. 380 oscillations frames ($0\text{-}190^0$) were collected for this crystal to ensure adequate data for scaling and averaging. The data collection protocol is summarized in Table 5.

Table 5: Data Collection protocol for each of the four crystals is summarized in this table.

Parameter	Native (pH 7.5)	eTA/ Ser5.6	eTA/ Ser7.5	eTA/ Thr7.5
Resolution (Å)	2.3	2.3	2.2	2.3
Crystal to detector distance	150mm	110mm	150mm	150mm
Exposure time	6 mins.	6 mins.	6 mins.	6 mins.
Start to end angle	0^0 to 120^0	0^0 to 190^0	0^0 to 120^0	0^0 to 120^0
Cryo-protectant	Mother liquor	40 µL Mother liquor + 5µL 100% glycerol	Mother liquor	Mother liquor

Data reduction:

The diffraction data were processed using the using the d*trek program in the CrystalClear software package provided by Rigaku Americas Co. and the CCP4 suite.[50] Statistics are listed in Table 6.

Table 6: Data collection and reduction statistics.

Parameter	Native crystal (pH 7.5)	eTA/ Ser 5.6	eTA/ Ser7.5	eTA/ Thr7.5
Space group	$C222_1$	$P2_1$	$C222_1$	$C222_1$
Unit cell parameters (a, b, c) Å	76.6,100.8, 175.6	77.2,104.1, 84.9	77.1,101, 176.6	74.4, 101, 176.1
Resolution range (Å)	38.32-2.3 (2.38-2.3)	34.46-2.3 (2.38-2.3)	29.16-2.2 (2.28-2.2)	29.34-2.3 (2.38-2.3)
Total number of observed reflections	320884	204371	99927	132270
Unique reflections	28335	59351	30725	28720
Completeness overall/outer shell (%)	92.3% (88%)	99% (97.7%)	86.8% (90.7%)	93.6% (94.2%)
R_{merge} (%)	10.4 (28.8)	13.7 (35)	8.5 (24.7)	9 (23.3)
Cell Volume (cubic Å)	1358075.612	685323.7	1376069.7	1357834.8
Average I/σ(I)	16.9 (8.7)	10.2 (4.6)	7 (3.4)	11.8 (5.8)
Mean redundancy	11.32 (11.59)	3.44 (3.34)	3.25 (3.5)	4.61 (4.58)

$R_{merge} = \sum| I - \{I\}|/\sum (I)$, where I is the observed intensity and {I} is the weighted mean of the reflection intensity. The values in parentheses are for the highest resolution shell
Matthew's Co-efficient [51] (M.C):

Matthew's Co-efficient is calculated to determine the solvent content in a protein structure. A solvent content of 30-70% with the corresponding number of molecules indicated in the asymmetric unit is acceptable.

M.C was determined for the native eTA structure as well as eTA/ Ser5.6 complex and the solvent content was found to be 47% for both the structures and indicated two and four molecules in the asymmetric unit, respectively. Since the co-crystals with L-Ser and L-Thr at pH 7.5 had the same cell parameters as the native eTA their respective asymmetric unit was also concluded to contain two molecules.

2.2.4 Structure Determination:

Native eTA ($C222_1$):

Structure solution with Molecular Replacement:

The structure determination of the native eTA was carried out using the WEB-based molecular replacement program, **CaspR.**[52] The structure of L-TA from *Thermatoga maritima* (PDB code 1LW5, 24.1% sequence identity to eTA) was used as a search model to carry out the molecular replacement. Multiple structure sequence alignment was carried out with T-coffee in CaspR[52] followed by homology model building with MODELLER. Out of 15 models generated one model gave the best solution (correlation coefficient of 26%, *R*-factor of 53.4%) which was subsequently refined using the CNS (Crystallographic and NMR Systems)**.**

The final model obtained from CNS in CaspR[52] did not have the sequence of eTA. The program **Phenix (*P*ython-based *H*ierarchical *EN*vironment for *I*ntegrated**

Xtallography)[53] was used to autobuild the eTA sequence into the electron density map of the solved model. The output model from Phenix[53] was used for further refinement. Phenix is a software suite for automated determination of protein structures.

Structure Refinement with **CNS 1.2**[54]: CNS is a software suite used for structure determination, in particular, structure refinement. The starting model obtained from Phenix[53] was subjected to alternate rigid body refinement, gradient minimization followed by simulated annealing using CNS 1.2.[54]Simulated annealing was performed to remove any phase bias that might have been introduced from molecular replacement. The model was manually corrected using **COOT 0.5.2**[55] and **TOM.**[56] Almost all the residues were well defined in the electron density maps after the first refinement itself except for a loop of five residues from 284 to 288 and a loop of three residues from 136 to 138. These recurrently appeared in the Ramachandran plot (generated using **PROCHECK**[57]) as outliers. The model was corrected with subsequent refinements. After several cycles of refinement with CNS 1.2[54], manual model correction and water additions were done using COOT 0.5.2[55] and TOM[56], the current R_{work}/ R_{free} for the native eTA structure is 24.5(%)/ 31(%). Further refinement with model correction and addition of water and other ligands are ongoing.

eTA/ Ser5.6 complex ($P2_1$):

Structure solution using molecular replacement and refinement: The refined native eTA structure was used to solve for the eTA/ L-serine complex at pH 5.6 (space group, $P2_1$).

Using **AMoRe** (an automated package for molecular replacement)[58], a single solution was obtained with correlation co-efficient of 26% and *R*-factor of 53.4%.

The model was subjected to rigid body refinement followed by alternate minimization and simulated annealing using CNS 1.2.[54] Manual model correction was done using COOT 0.5.2[55] and TOM.[56] After several cycles of refinement with CNS 1.2[54] and model correction (especially for the loop of residues described in native eTA structure) above and water addition using COOT 0.5.2[55] and TOM[56], the current R_{work}/ R_{free} is 24.8(%)/ 30.6(%). Further refinement with model correction and addition of water and other ligands are ongoing.

eTA/ Ser7.5 complex ($C222_1$):

Structure determination and refinement: Since eTA/ Ser7.5 complex is isomorphous with native eTA structure with same space group, $C222_1$ and similar cell parameters, the refined model of native eTA structure was used as an input model for refinement of the complex using CNS 1.2.[54] After several cycles of alternate refinements, including rigid body, minimization and simulated annealing, addition of water and other ligands, and intermittently correcting the model manually with COOT 0.5.2[55] and/or TOM[56], the current R_{work}/ R_{free} is 26(%)/ 30.6(%). Further refinement with model correction and addition of water and other ligands are ongoing.

eTA/ Thr7.5 complex ($C222_1$):

Structure determination and refinement: Structure determination was performed as described for eTA/ Ser7.5 complex since this complex is also isomorphous with the native structure. The model was subjected to several cycles of refinement with CNS 1.2[54]

and model correction (especially for the loop of residues described in native eTA structure above) and water addition using COOT 0.5.2[55] and TOM[56] with the final R_{work}/ R_{free} 20.33(%)/ 25(%)

Refined model statistics for eTA/ Thr7.5 complex, as well as each of the other crystals is summarized in Table 7. The root mean square deviations for the bond angles and bond lengths as well as the average B-factors for protein, ligands and water are also provided in Table 7.

Table 7: Refined model statistics for each of the four crystals is summarized in this table.

	Native eTA	eTA/ Ser5.6	eTA/ Ser7.5	eTA/ Thr7.5
R_{work} (95% data)	24.5	24.8	26	20.33
R_{free} (5% data)	31	30.6	30.6	25.5
Rmsd standard Geometry:				
Bond Length (Å)	0.007	0.006	0.007	0.006
Bond Angle (^{0})	1.3	1.3	1.3	1.4
Average B factor				
(i) Non -hydrogen atoms	25.5	25.1	28.5	21.9
(ii) Protein	25.5	24.8	28.4	21.4
(iii) Water	26.4	26.4	29.5	26
(iv) Cation (Mg^{2+}, Ca^{2+}, Na^{+})	22.7	19.5	34.8	14.9
(v)PLP	23.5	25.5	-	-
PLG	-	-	36.2	-
PLT	-	-	-	24.7
(vi) Ligands:				
L-Ser	Nil	37.9	Nil	Nil

$R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. R_{free} is the crystallographic R_{work} calculated with 5% of the data that were excluded from the structure refinement. PLT is L-Thr bound to PLP as an external aldimine. PLG is glycine bound to PLP as an external aldimine.

Validation:

The Ramachandran plot was also checked at the later stages of the refinement to inspect for unrealistic conformations in the models. The residues that occur in the generously allowed as well as disallowed regions of the plot are shown for each of the structures in red font (Figures 14-17) and Ramachandran plot statistics are summarized in Table 8.

Table 8: Ramachandran plot statistics for all the four structures are presented here.

	Native eTA	eTA/ Ser5.6	eTA/ Ser7.5	eTA/ Thr7.5
Residues in the most favored region	89.1%	91.1%	88.6%	90.6%
Residues in additionally allowed region	9.7%	8%	10.2%	8.3%
Residues in generously allowed region	1%	0.6%	1%	1.1%
Residues in disallowed region	0.2%	0.3%	0.2%	0.0%

For each of the structures the residues that occur in the disallowed regions of the plot were checked in their respective electron-density maps. These residues seem to recur in the Ramachandran plot since they appear disordered in the maps.

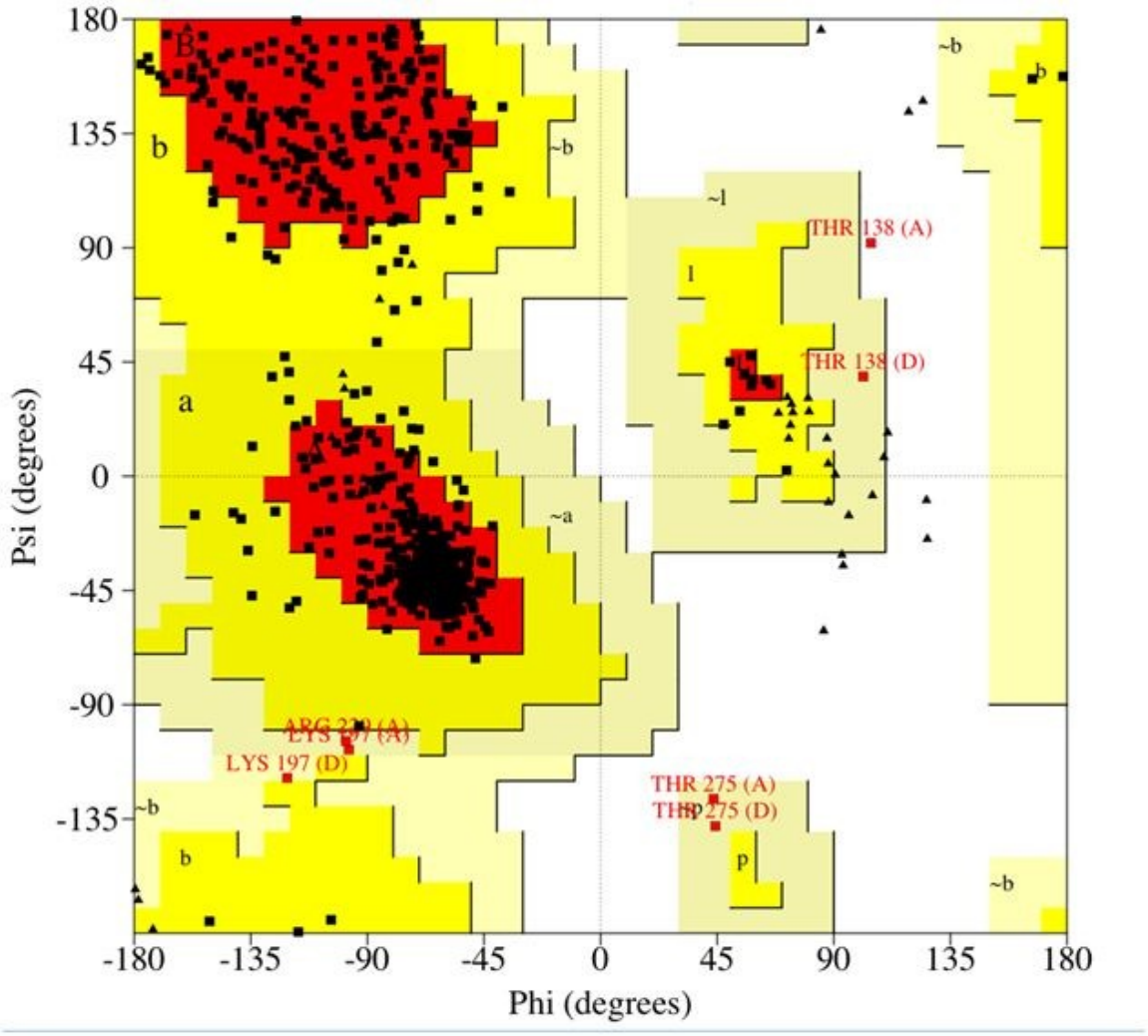


Figure 14: Ramachandran plot of refined structure of native eTA. The plot was generated using PROCHECK.[57]

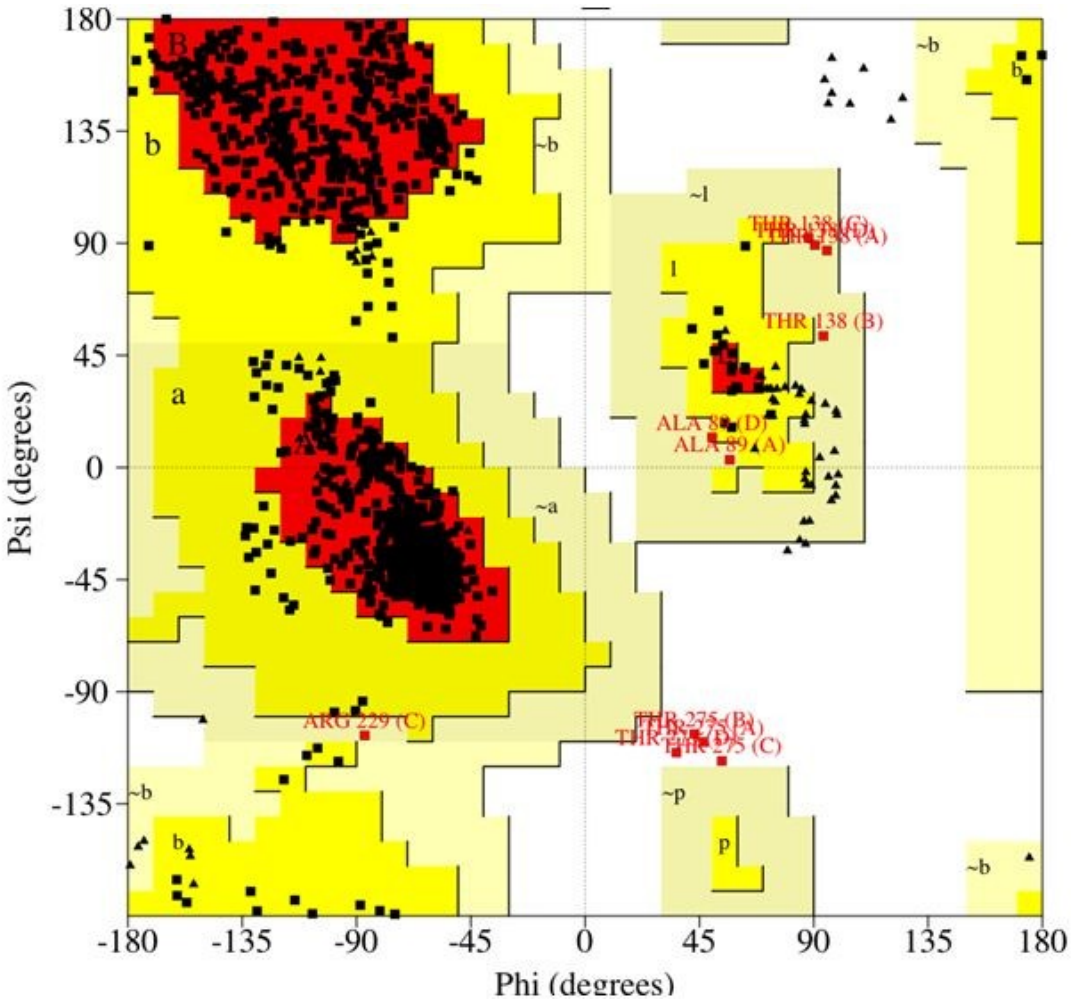


Figure 15: Ramachandran plot of refined structure of eTA/ Ser5.6 complex. The plot was generated using PROCHECK.[57]

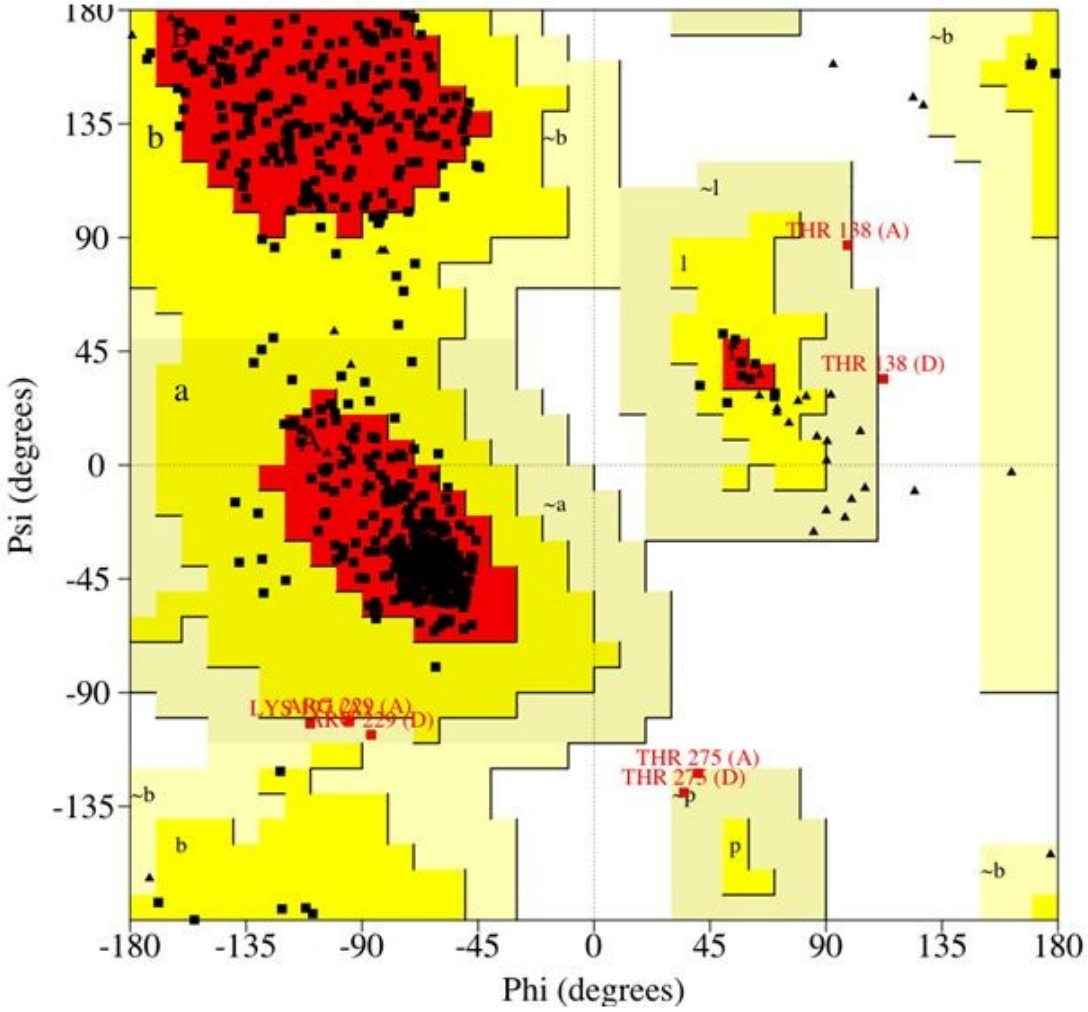


Figure 16: Ramachandran plot of refined structure of eTA/ Ser7.5 complex. The plot was generated using PROCHECK.[57]

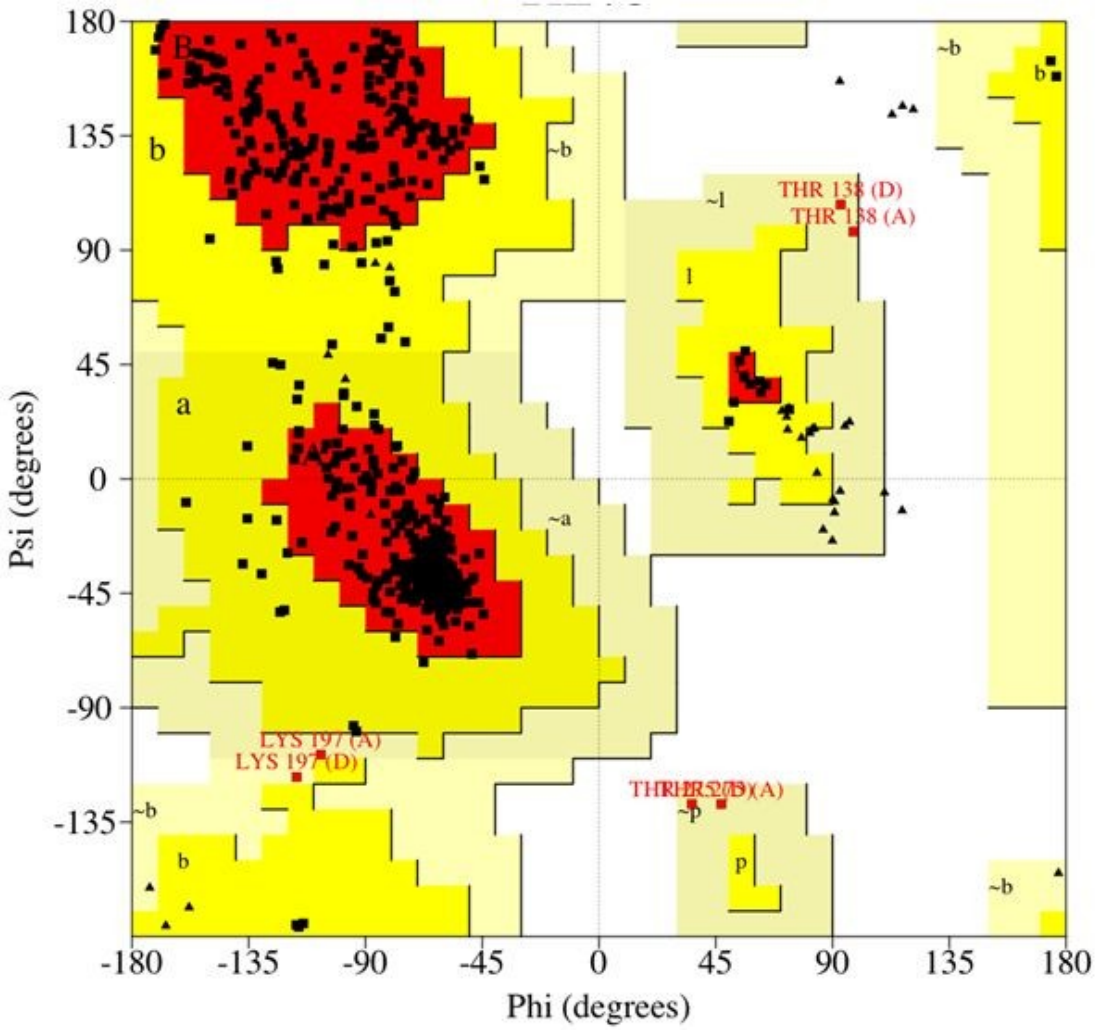


Figure 17: Ramachandran plot of refined structure of eTA/ Thr7.5 complex. The plot was generated using PROCHECK.[57]

2.3 Results and Discussion:

2.3.1 Crystallization:

Initial Crystallization: Initial crystals were obtained in some of the conditions provided in the Crystal Screen Macromolecular Crystallization kit (Screen I and II) by Hampton Research for native eTA as well as eTA bound with ligands. The conditions are summarized in the tables below (Table 9 and Table 10).

The protein concentration in each of the conditions was either 10 mg/mL or 22mg/mL. The ligand concentration is between 6-30mM. Crystallization was carried out at room temperature. All the crystals obtained were flat and very thin.

Table 9: Conditions from Screen I that provided crystals for native eTA and eTA bound to ligands is summarized here.

Crystallization Condition				Observation (after 7 days)
Salt	Buffer	Precipitant	Ligand	
A) 0.2M ammonium sulfate	0.1M sodium cacodylate pH 6.5	30% w/v PEG 8000	Nil	Clear drop
B) 0.2M magnesium chloride hexahydrate	0.1M HEPES pH 7.5	30% PEG 400	Nil Or L-Thr	Flat plate crystals
C) 0.2M Calcium chloride hexahydrate	0.1M HEPES pH 7.5	28% PEG 400	L-Ser	Flat plate crystals Spherical agglomerate
D) None	0.1M HEPES pH 7.5	0.8M sodium phosphate monobasic monohydrate	Nil	Salt crystal
E) None	0.1M HEPES pH 7.5	1.4M sodium citrate tribasic dihydrate	Nil	Flat, plate crystals
F) None	0.1M sodium citrate tribasic dehydrate pH 5.6	20% v/v 2-propanol 20% w/v PEG 4000	L-Ser (6.25mM)	Flat, plate crystals

Table 10: Conditions from Screen II that provided crystals for native eTA and eTA bound to ligands is summarized here

Crystallization Condition			Observation (after 7 days)
Salt	Buffer	Precipitant	
A) 2M ammonium sulfate	None	5% v/v 2-propanol	Spherical agglomerate
B) 0.01M zinc sulfate heptahydrate	0.1M MES pH 6.5	25% v/v PEG monoethylene ether 550	Needle Cluster
C) None	0.1M HEPES pH 7.5	20% w/v PEG 10,000	Flat plate crystals

The initial crystals obtained with eTA (10mg/mL) in condition C) [0.2M calcium chloride hexahydrate in 0.1M HEPES pH 7.5 with 28% PEG 400 and L-Serine (2mM)] and condition F) [0.1M sodium citrate tribasic dehydrate pH 5.6 with 20% v/v 2-propanol 20% w/v PEG 4000 and L-Serine (6.25mM)] from Screen I (Table 9) were single and of diffraction quality. Although, the ligand in both the crystals was L-Serine, the pH at which the crystals were obtained was different. In condition C) the pH was 7.5 while in condition F) the pH was 5.6. This difference in pH of crystallization buffer has consequences in our interpretation of the structures as well as the mechanism that is ultimately proposed.

Single crystals were also obtained with eTA (17.5mg/mL) in condition B) [0.2M magnesium chloride hexahydrate in 0.1M HEPES pH 7.5 with 30% PEG 400] from Screen I (Table 9) in the presence of L-Threonine (28mM) as the substrate and were used for data collection.

The initial crystal obtained with eTA (22mg/mL) in the absence of any ligand in condition B) [0.2M magnesium chloride hexahydrate in 0.1M HEPES pH 7.5 with 30% PEG 400] from Screen I (Table 9) was not single and hence, optimization of the conditions was required.

Optimization of conditions: Since condition B) from Screen I provided twinned crystals in the absence of a ligand, optimization of the condition was necessary. Addition of 5% glycerol to the reservoir solution of condition B) to slow down nucleation only resulted in a thin film formation and did not aid the formation of single crystals. Small amount of a detergent (n-octyl-β-D-glucoside) was added to prevent agglomeration and hence aid single crystal formation. Although a few single crystals were obtained on addition of the detergent, the crystals were not diffractable. PEG 400 concentration was also varied between 25-30% for condition B). Use of 25% PEG 400 in condition B) instead of 30% with fresh protein (10mg/mL) resulted in single crystals of diffraction quality.

Final crystallization conditions that provided single native as well as ligand bound crystals are summarized in Table 11.

Table 11: Final conditions that provided single crystals are summarized here. The protein concentration used in conditions I, II and III is 10mg/mL while in condition IV the protein concentration is 17.5mg/mL. The protein to crystallization buffer ration in all the cases is 1:1

Crystallization Condition				Observation (after 7days)
Salt	Buffer	Precipitant	Substrate	
I) 0.2M Calcium chloride hexahydrate	0.1M HEPES pH 7.5	28% PEG 400	L-Ser	Flat, single plate crystals
II) None	0.1M sodium citrate tribasic dehydrate pH 5.6	20% v/v 2-propanol 20% w/v PEG 4000	L-Ser	Flat, single plate crystals
III) 0.2M magnesium chloride hexahydrate	0.1M HEPES pH 7.5	25% PEG 400	Nil	Flat, single plate crystals
IV) 0.2M magnesium chloride hexahydrate	0.1M HEPES pH 7.5	30% PEG 400	L-Thr	Flat, single plate crystals

2.3.2. Overall Structural Description:

eTA has been established as a tetrameric protein in solution[59]. It binds one PLP molecule per monomer. In the structures reported here, two of the monomers form an intimate tight catalytic dimer. As with other Fold-type I enzymes , residues from both the monomers form the active site and each monomer consists of a large domain with seven

β-strand structures flanked by α-helices and a small domain composed of three β-strand sheet structures with interlinking α-helices.

The eTA structures show PLP at the active site either as an internal aldimine bound to the conserved Lysine residue of the active site (covalent bond between C4' of PLP with amide N_1 of Lys197) or as an external aldimine linked to the substrate or the product (covalent bond between C4' of PLP with amide N of a β-hydroxy amino acid). The pyridoxal ring of PLP rotates by ~20^0 between the internal and external aldimine form.

Monovalent or divalent cation, whichever is present in the crystallization buffer (Na^{+1}, Ca^{+2} or Mg^{+2}) is seen close to the active site. But these are not suggested to be involved in catalysis[43], but possibly in maintaining the structure of the active site. Each of these ions readily form an octahedral coordination sphere flanked by the residues Thr8, Thr10, Ser196, Thr201 and Gln230 with a water molecule in close proximity in all the structures. Well coordinated ions are also present at the dimer interface to stabilize the quaternary structure.[43] Detailed description of native as well as ligand bound structures is provided hereunder;

Structural details of native eTA:

Although eTA is a tetrameric protein, the asymmetric unit of native eTA shows the presence of a dimer (Figure 18). The tetramer can be generated using symmetry operators. A divalent cation (Mg^{+2}) is located between the large and the small domains. Twenty nine residues from monomer A interact with 31 residues of monomer D.[60] The dimer buries a total area of 3990 $Å^2$ at the A/D interface.[60]

Each of the two molecules of PLP (related by non-crystallographic symmetry) are located at the interface of the large and small domains of individual monomeric unit to make a Schiff-base with Lys197 ***forming an internal aldimine*** (Figure 18)**.** Nevertheless, a loop of about ten residues from the opposite monomer (from Lys121 to Lys131) wraps around the active site to position His126 in such a way that it has access to the active site (Figure 18). The position of His126 is well defined in the native structure.

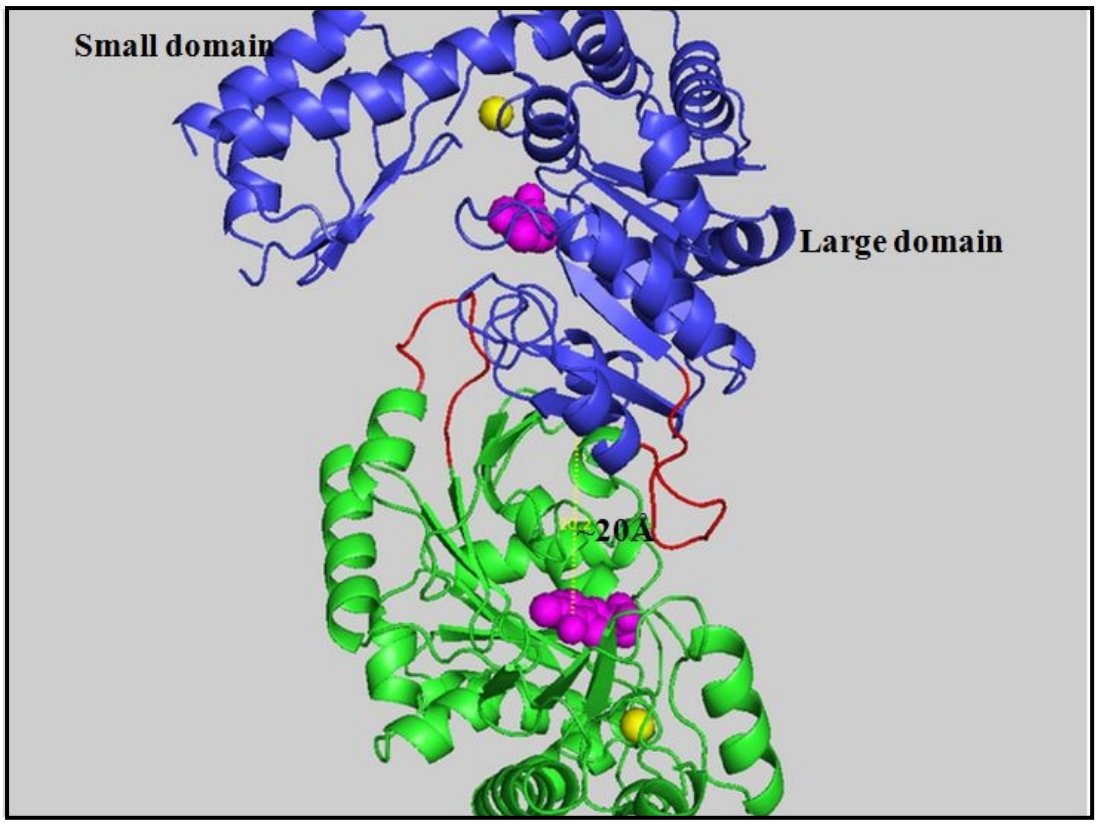


Figure 18: The catalytic dimer is represented here. The two monomers have been depicted in green and blue ribbons. Mg^{+2} ions (yellow) at the interface of the small and the large domain; PLP (magenta) bound to Lys197 with a loop (red) of about ten residues from the other monomer wrapping around the active site.

Figure 19 shows the active site of native eTA structure with the residues that interact with PLP. His83 stacks parallel on top of the PLP ring from the *re* face while Ala168 is present at the *si* face (His83 and Ala168 are not shown in Figure 19). The PLP is recognized via hydrogen bond interactions between its phenolic group (O3) and the guanidinium group of Arg169, and its pyridine nitrogen (N1) with the side chain of Asp166. The position of Asp166 is maintained by hydrogen bond interaction with Asn62. Other residues that recognize the PLP phosphates via hydrogen bond interactions are the

amide nitrogen of Gly58 and hydroxyl group of Thr59, as well as water-mediated hydrogen bond interactions with the protein. The residues Arg169, Asp166, Asn62, Glu58 and Thr59 are conserved in L-TAs.

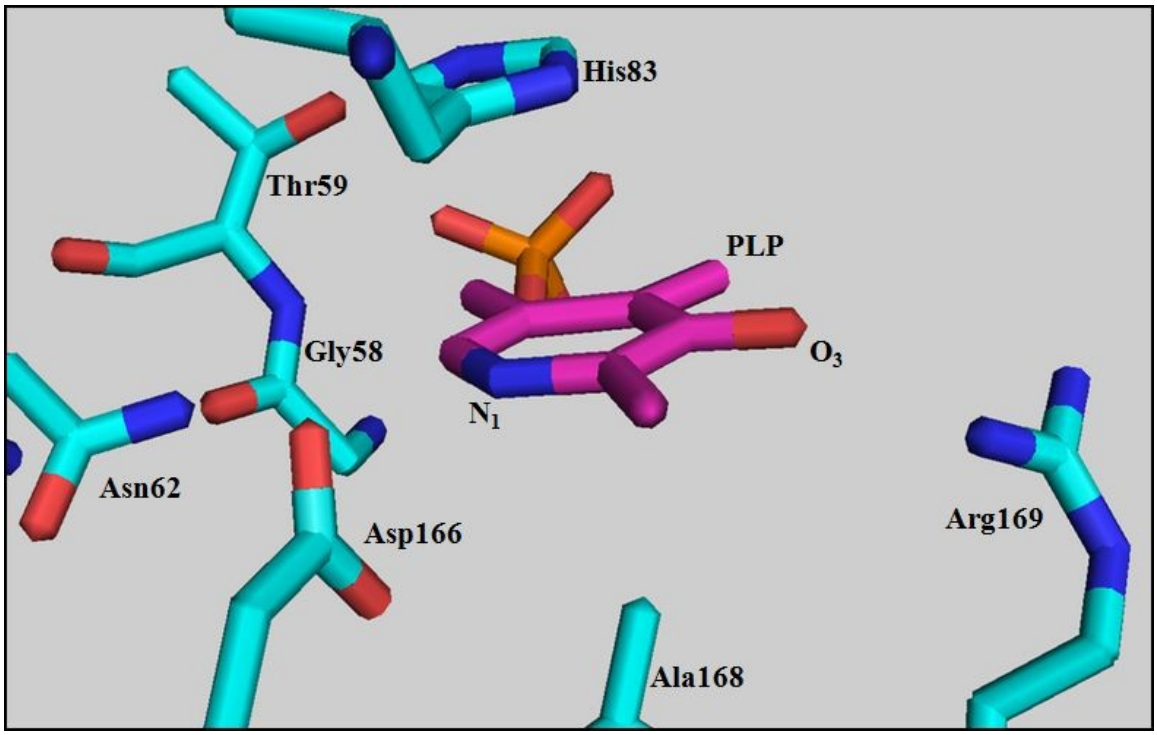


Figure 19: Active site of native eTA showing the residues interact with PLP. These residues are conserved amongst various L-TAs as well as other PLP-dependent enzymes.

Figure 20 shows the active site of native eTA with the residues other than those involved in PLP recognition. Arg308 makes close hydrogen-bond interactions with two water molecules via its guanidinium groups. Other residues including His83, His126, Glu88 and Phe87 have been implicated in the catalytic function of eTA and are shown in Figure 20.

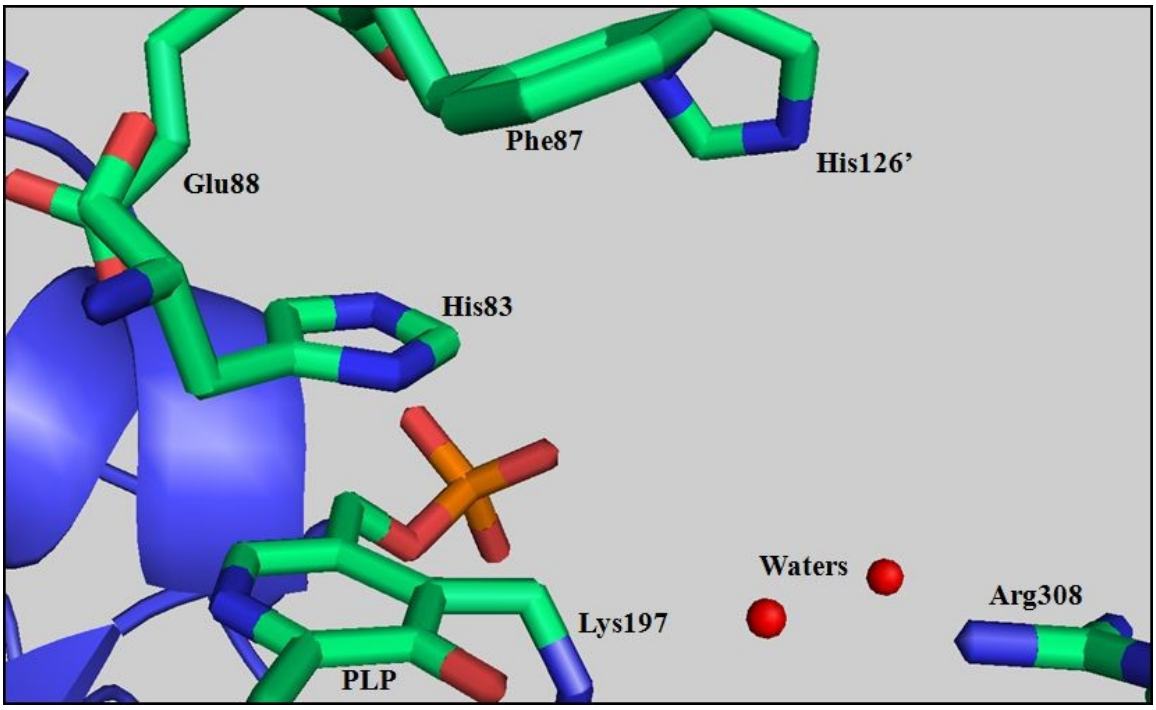


Figure 20: Active site of native eTA structure. The active site residues as well as the two well co-ordinated water molecules are shown.

Structural details of eTA in complex with L-Serine at pH 5.6 (eTA/ Ser5.6):

The complex between eTA and L-Ser was crystallized at a pH of 5.6 which is much lower than pH 7.4 at which physiologically relevant reactions are carried out in the cell. At pH 5.6, the catalytic activity of the enzyme is reduced almost by 10-fold[4]. To understand the molecular basis for the reduced enzymatic activity at pH 5.6 we co-crystallized the protein with L-Ser.

Unlike the native structure, this complex crystallized with a tetramer in the asymmetric unit (Figure 21). Each monomer binds a molecule of PLP through internal aldimine formation, similar to the native/ eTA structure. As with the native structure, we observe ions (Na^{+}instead of Mg^{+2} in native eTA structure since the former was used in the

crystallization) close to the active site, and also at the inter-dimer interface (Figure 21).The four subunits are denoted A-D. The A/D (or B/C) is the catalytic dimer (distance between N1-N1 of PLP in A/D is 34Å) while A/C (or B/D) is less intimate (N1-N1 distance of 45Å).

The tertiary structure of the complex is very similar to the native structure with root mean square deviation (rmsd) of 0.3Å. Nevertheless, subtle changes in the active site are evident on examination of the structure.

The residues Asp166, Arg169, Gly58 and Thr59 (these residues are not shown in Figure 22) make hydrogen bond interactions with various atoms of PLP similar to that described above for the native eTA structure. PLP is bound to Lys197 as an internal aldimine (Figure 22). Interestingly, co-crystallization of the ligand L-Ser at pH 5.6 has resulted in the substrate being trapped at the active site. The amide nitrogen of L-Ser is about 3.3Å away from the C4' of PLP suggesting absence of any external aldimine formation. Consistently, we do not observe presence of product glycine. Although the ligand density is very strong in all four subunits, it is disordered into several orientations especially in two of the subunits (A and C). Thus, the structure shows asymmetry in binding.

The carboxylate of L-Ser makes hydrogen-bond interactions with the guanidinium groups of Arg308 and Arg169, as well as the hydroxyl of Ser6. The two water molecules that were found to hydrogen-bond to the guanidinium group of Arg308 (Figure 22) in the native structure are now absent in this structure because the carboxylate group of L-Ser replaces the two water molecules. The carboxylate group also makes close hydrogen bond interactions with the guanidinium group of Arg169. The hydroxyl group

of L-Ser makes hydrogen bond interactions with the side chains of both His83 and His126, the interaction with the latter being stronger. His126 is well defined in the electron density map.

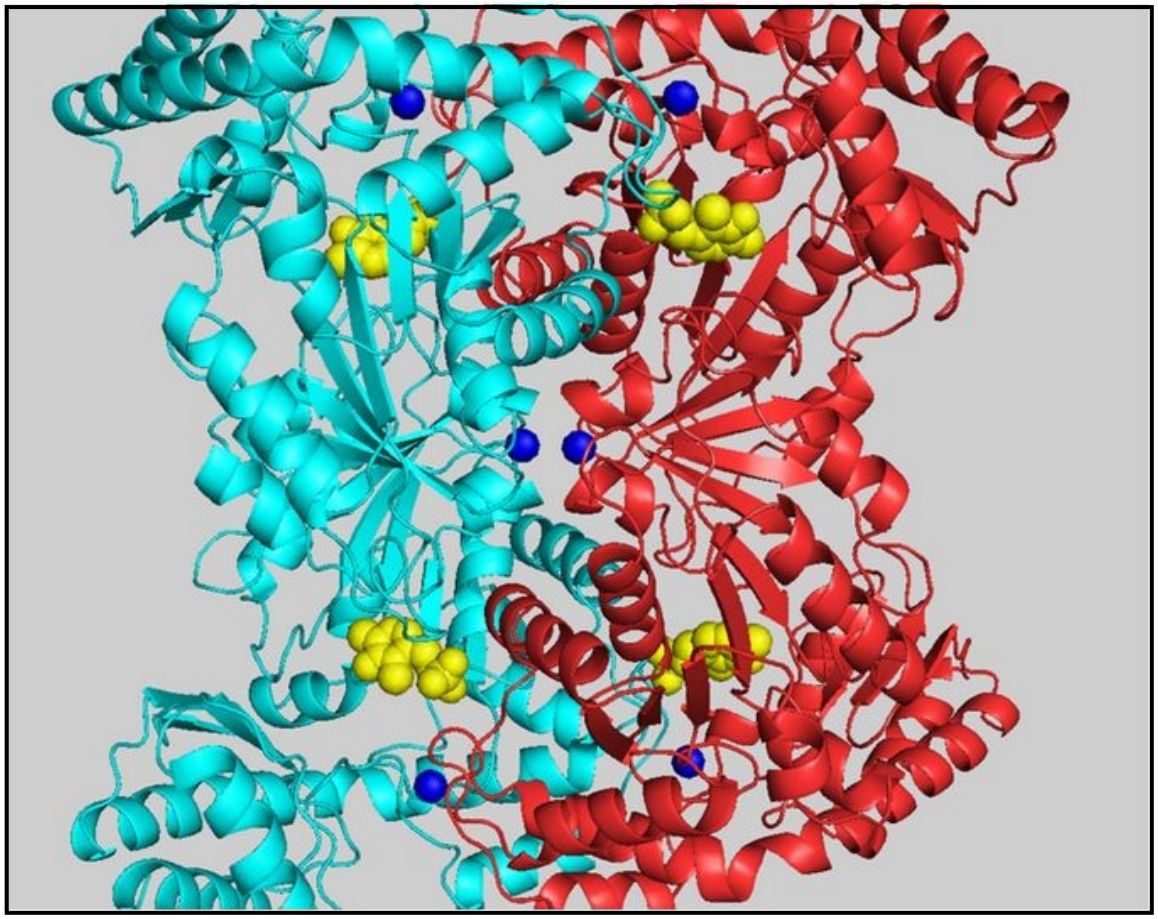


Figure 21: Tetrameric eTA/ L-serine complex at pH 5.6 is shown here. Sodium ions (blue) and PLP molecules (yellow) and the catalytic dimers A/D (cyan) and B/C (red) are also depicted here.

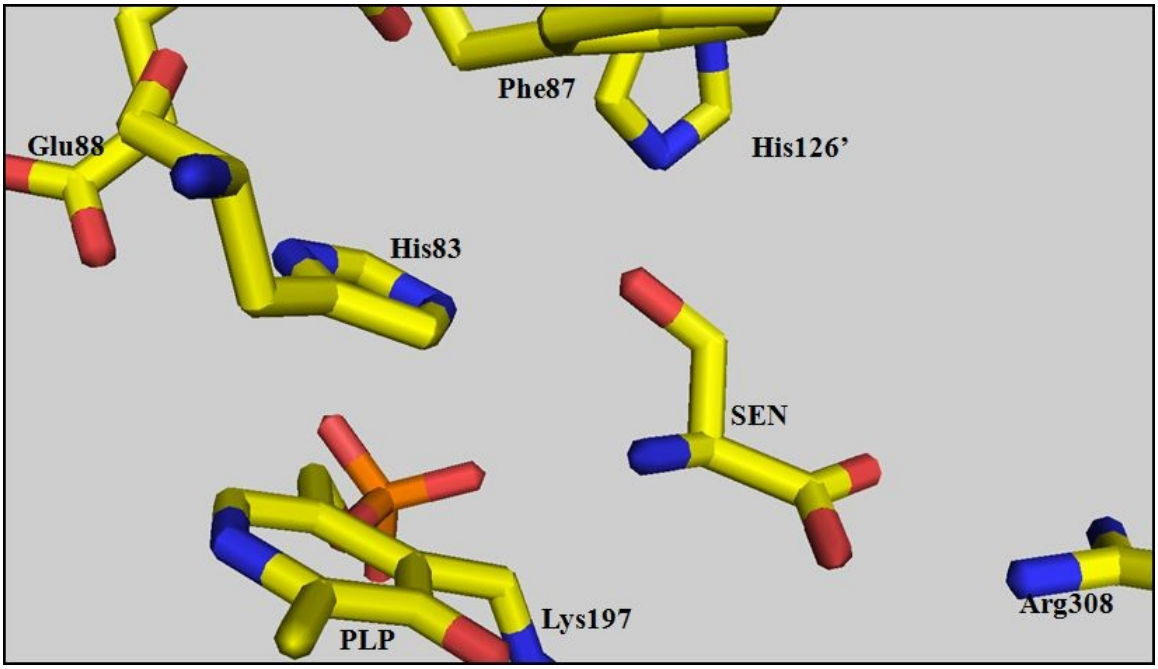


Figure 22: Shown above is the active site of eTA/ L-Serine complex at pH 5.6. L-serine (SEN) is shown trapped at the active site. The carboxylate group of SEN displaces the two well co-ordinated water molecules seen in the native structure. Not all residues, including Ser6, Asp166, Arg169, Glu58, Thr59 and Arg229 described in the text are shown here.

Structural details of eTA in complex with L-Serine at pH 7.5 (eTA/ Ser7.5):

We have also studied the complex structure between eTA and L-Ser at physiological pH of 7.5, where the enzyme shows optimum activity. The eTA/ Ser7.5 complex structure is isomorphous to the native structure, crystallizing with a dimer in the asymmetric unit. The rmsd between the two structures is 0.2Å, suggesting similar conformation. As with the native eTA structure, cations (Ca^{+2} instead of Mg^{+2} in native eTA structure) are present at the interdomain interface of each monomer.

As shown in Figure 23, eTA/ Ser7.5 also shows bound PLP at the active site, making similar protein interactions as described for the structures above. Also, the PLP

essentially makes an aldimine linkage with product Gly (i.e. an external aldimine linkage); however there is evidence of a low occupancy presence of a PLP-internal aldimine with Lys197. In addition to the product, we also observe a weak presence of the substrate, L-Ser that superimposes closely with the Gly. The Cα-Cβ bond of L-Ser bound as an external aldimine to PLP is nearly orthogonal with the plane of the pyridoxal ring (81^0) to make maximum electronic interactions with the π system of PLP. The carboxylate group of glycine (and L-Ser) displaces the water molecules seen co-ordinated to Arg308 in the native structure and makes hydrogen bond interactions with Arg308 and Arg169. The bound L-Ser makes additional interactions through its hydroxyl moiety with the side chain of His126 and His83, as well as with a water molecule which mediates the interaction with the PLP phosphate. Unlike the native and the eTA/ Ser5.6 structures, His126 of this structure is disordered. The OH of L-Ser is also located between the side-chains of His83 and His126, but compared to the eTA/ Ser5.6 structure, the interaction between the OH of L-Ser and His126 seems to be weaker and His83 stronger

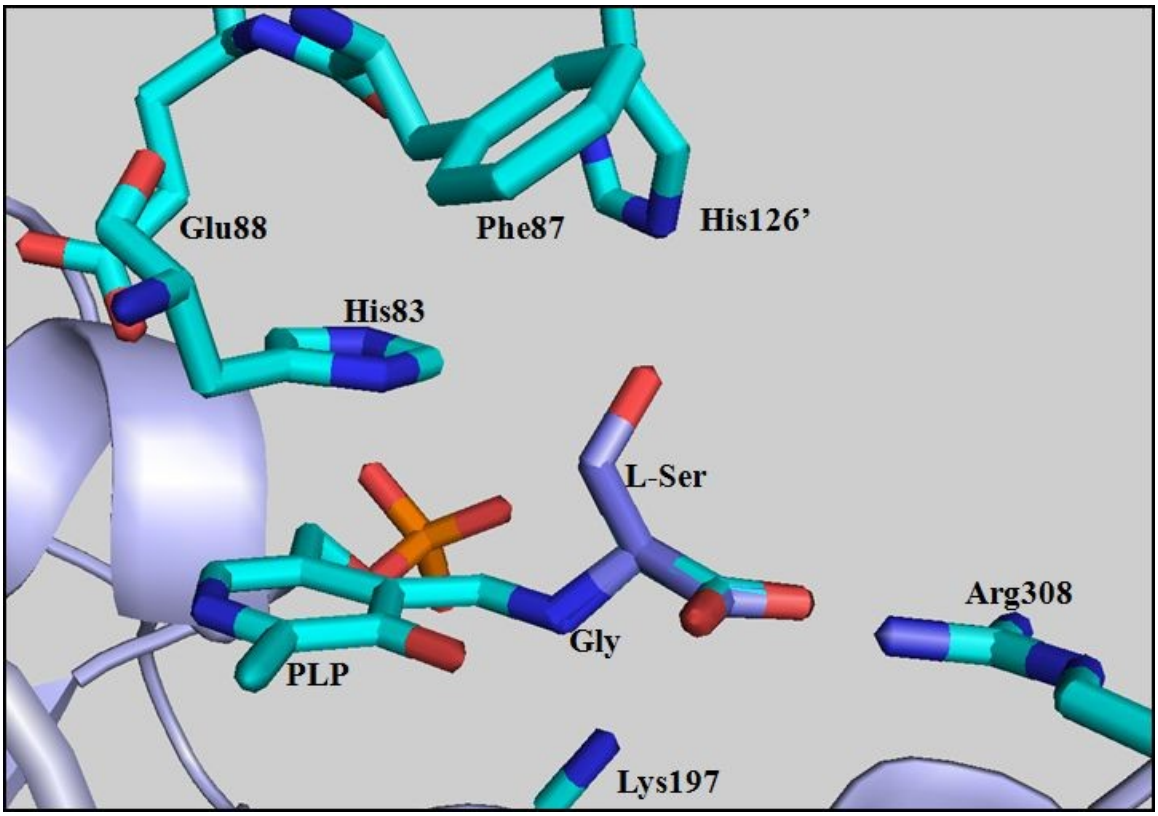


Figure 23: The active site residue of eTA/ L-Serine complex at pH 7.5. The product Gly is seen bound to PLP as an external aldimine. Gly and L-Ser are superimposed with each other at the active site and bound as external aldimine to PLP.

Structural details of eTA in complex with L-Threonine at pH 7.5 (eTA/ Thr7.5):

We have also studied the complex structure between eTA and L-Thr at physiological pH of 7.5, to understand the difference in mode of binding and subsequent catalysis (if any) of varied substrates by eTA. The eTA/ Thr7.5 complex structure is isomorphous to the native structure and crystallized with a dimer in the asymmetric unit as well. The rmsd between the two structures is 0.2Å, suggesting similar conformation.

Shown below in Figure 24 is the active site of eTA/ Thr7.5. Similar to the structures described above, eTA/ Thr7.5 shows bound PLP at the active site, recognized

by residues described for the native eTA structure. As observed in eTA/ Ser7.5 structure, PLP forms an external aldimine with the product glycine, with evidence of low occupancy of substrate L-Thr bound external aldimine. There is no evidence of internal aldimine formation in either of the two subunits in eTA/ Thr7.5 structure. The angle between the Cα-Cβ bond of L-Thr and the plane of the pyridoxal ring is 88.5^0, consistent with the Dunathan hypothesis that for optimum reactivity the Cα-Cβ should be nearly orthogonal to the plane of the pyridoxal ring of PLP. Again, the carboxylate group of glycine (and L-Thr) displaces the water molecules seen co-ordinated to Arg308 in the native structure and makes hydrogen bond interactions with Arg308 and Arg169. The bound L-Thr makes hydrogen bond interactions with the side chains His83, weak hydrogen bond interactions with His126 and a water molecule as described above for the L-Ser in eTA/ Ser7.5 structure. His126 is also disordered in this structure. In comparison with the eTA/ Ser5.6 or eTA/ Ser7.5 structures, the hydroxyl group of L-Thr seems to be located closest to His83 but furthest from His126.

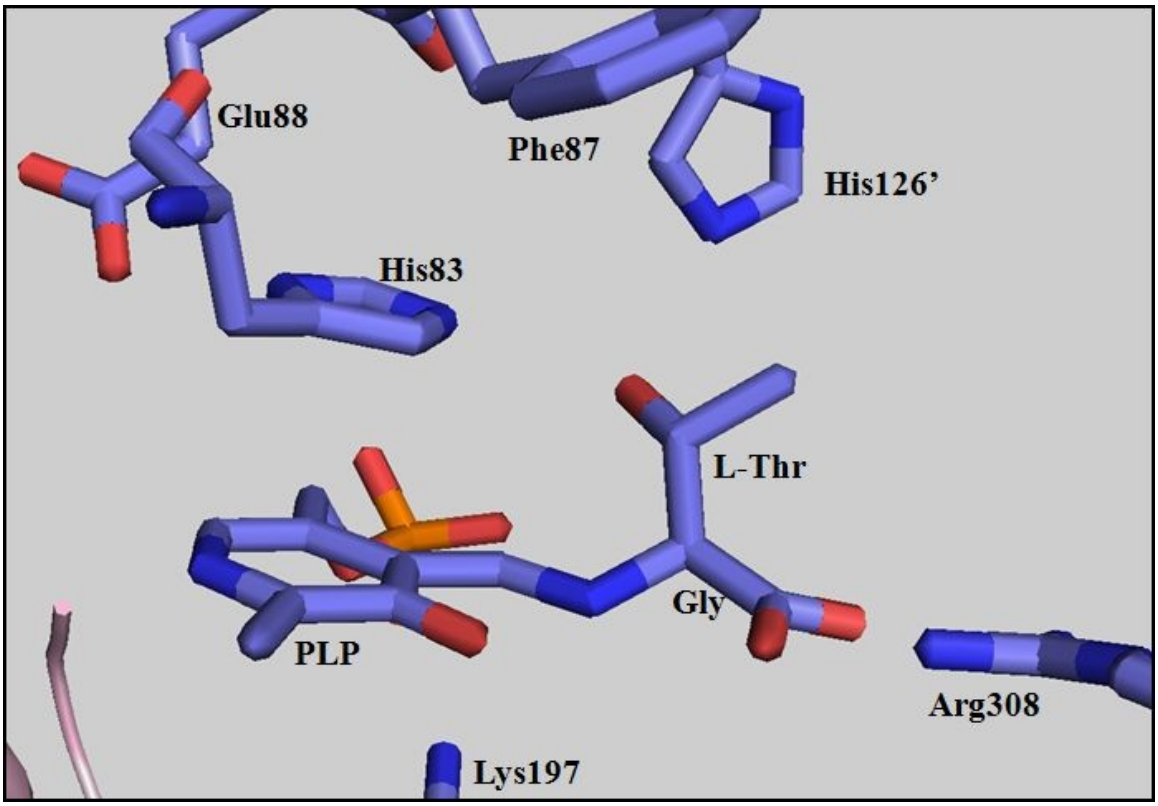


Figure 24: The active site residue of eTA/ L-Threonine complex at pH 7.5. The product Gly and substrate L-Thr bound to PLP as external aldimine.

Modeling L-allo-Thr into the active site of eTA/ Thr7.5:

Although the natural substrates of eTA are L-threonine and L-*allo*-threonine, the enzyme has significant preference for L-*allo*-threonine. To understand the basis of this substrate specificity, and in the absence of a co-crystal structure of L-*allo*-threonine with eTA, we modeled this ligand into the active site of eTA/ Thr7.5 complex based on the binding mode of L-Thr (Figure 25). If the two hydroxyl groups are superimposed, the methyl group of L-allo-Thr which is disposed in a different direction from that of L-Thr is found to sterically clash with the side chain carbon of His83 (Figure 25). Due to the freely rotatable bond at Cβ carbon of L-*allo*-Thr, the methyl group probably moves away

from His83 to relieve the steric clash. A consequence of this action would be the hydroxyl group of L-*allo*-Thr moving closer to His83 but further away from His126 (Figure 26). Note that the methyl group of L-Thr shows no such steric clash and its hydroxyl group is almost equidistant from both His83 and His126 (Figure 25 and Figure 26). Thus, it seems that the enzyme selects between L-Thr and L-*allo*-Thr based on the methyl group position.

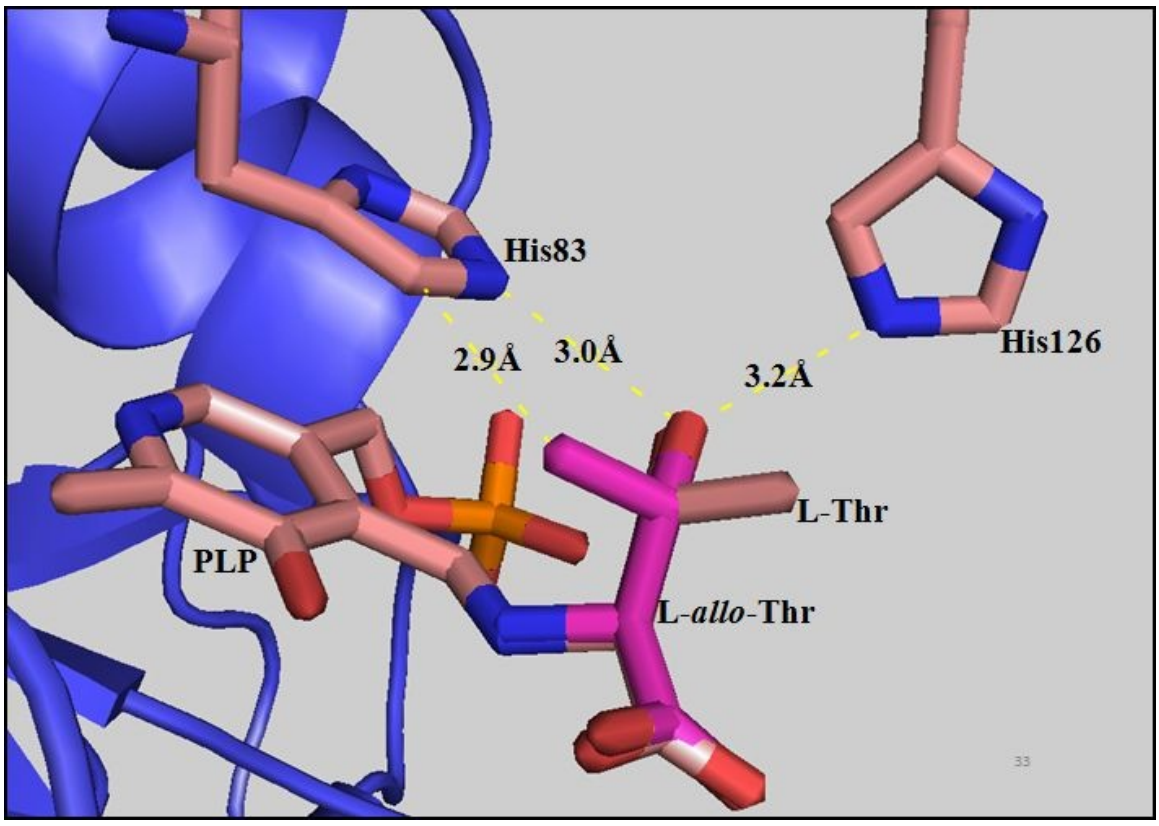


Figure 25: Hydroxyl group of L-threonine bound as external aldimine (Pink) is 3Å and 3.1Å from His83 and His126, respectively. L-*allo*-threonine (magenta) has been modeled into the density observed for L-threonine. Hydroxyl group is at hydrogen bonding distance from His83 but the methyl group makes steric interaction with the side chain carbon of His83.

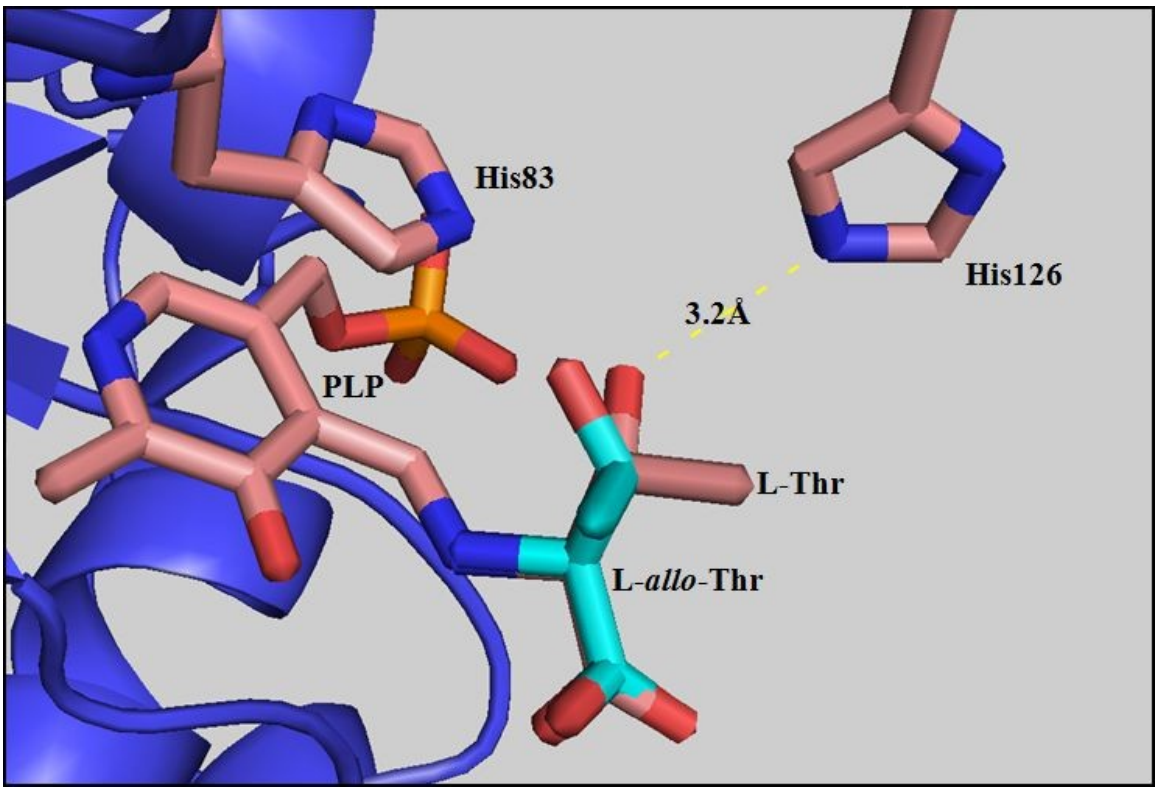


Figure 26: Hydroxyl group of L-threonine bound as external aldimine (Pink) and 3.2Å from His83 and His126, respectively. L-*allo*-threonine (cyan) has been modeled to be at an acceptable distance from the side chain carbon of His83. Hydroxyl group is now closer to His83 but further away from His126, while the steric clash between the methyl group and the side chain carbon of His83 is relieved

2.3.3 Catalytic Mechanism of eTA:

To draw a clear picture of the catalytic mechanism and substrate specificity of eTA, we have determined the crystal structures of native eTA and its complex with substrate and/or product at different pH's for comparative studies. Comparative analysis of eTA structures with SHMT as well as with TTA structures also provided important information about probable roles of active site residues in the catalytic mechanism of eTA.

Kinetic data of mutants of the active site residues of eTA was also very useful in understanding the mechanism of the enzyme in greater details. The mutation studies were performed by our collaborators at the University of Rome and their findings (unpublished data) is used to make meaningful interpretations of our structures.

Catalytically Significant Residues:

The role of the active site residues His83, His126, Phe87 and Glu88 is examined in greater details to propose a mechanism for the retro-aldol cleavage of β-hydroxy amino acids. An overview for the mechanism of the reaction involves abstraction of the hydroxyl proton from the substrate bound external aldimine, followed by formation of product glycine bound external aldimine and subsequent regeneration of the internal aldimine. The enzyme also selectively catalyzes the retro-aldol cleavage of L-*allo*-Thr > L-Thr > L-Ser. The residues involved in the mechanism of retro-aldol cleavage as well as the selection process have been identified here.

His83 as the catalytic base:

His83 at the active site has been proposed to be the catalytic base which is involved in the abstraction of the hydroxyl proton of the β-hydroxy amino acid during the retro-aldol cleavage reaction.

In the crystal structure of eTA/ Ser5.6, the hydroxyl group of L-Serine makes very weak interaction with His83, preventing optimal abstraction of the hydroxyl proton, and as a result the enzyme does not catalyze L-Ser. In comparison, the crystal structure of eTA/ Thr7.5 and eTA/ Ser7.5 have the hydroxyl groups of L-Thr and L-Ser, respectively, closer to His83, which resulted in the turnover of these substrates to their corresponding

products. The hydroxyl group of L-Thr in eTA/Thr7.5 complex is closer to His83 when compared to the hydroxyl group of L-Ser in eTA/ Ser7.5, and this is reflected in higher k_{cat} for the L-Thr (Table 12). The hydroxyl group of the modeled L-*allo*-Thr is even closer to His 83, and expectedly L-*allo*-Thr has the highest kcat (Table 12).

Since, His83 is closer to the hydroxyl group of L-*allo*-Thr , this residue seems to influence the affinity of the enzyme for L-*allo*-Thr significantly. Mutation of His83 to Phe loses the hydrogen bonding capability to this ligand resulting in a 30-fold increase in the K_m of L-*allo*-Thr. Asn is capable of making weak hydrogen bond interactions with the hydroxyl group of L-*allo*-Thr, hence only an eight fold decrease in affinity is observed with the H83N mutant (Table 12).

With L-Thr catalysis, hydrogen bond interactions with both His83 and His126 seem to influence the Km values. H83F mutant shows a three-fold decrease in Km values when L-Thr is the substrate (Table 12, Figure 27). Although Phe lacks the ability of forming hydrogen bond interactions with the hydroxyl group of L-Thr, we speculate that the hydroxyl group of L-Thr is pushed closer to His126 in H83F mutants due to steric interaction with Phe. This results in a stronger hydrogen bond interaction between the hydroxyl group and His126 that increased the affinity of the substrate for the mutant, nevertheless the kcat is reduced significantly because of the loss of the catalytic base. In the H83N mutants, Asn obviously can still make hydrogen-bond interaction with the hydroxyl group of L-Thr but probably weaker compared to the His, resulting in the observed two-fold increase in Km.

Table 12: The table shows the kinetic data of His83 mutants. Also, shown is the kinetic data of L-Ser cleavage in wild type enzyme

Substrate	WT			H83F			H83N		
	k_{cat}	K_m	k_{cat} / K_m	k_{cat}	K_m	k_{cat} / K_m	k_{cat}	K_m	k_{cat} / K_m
L-threonine	112	19.4	5.8	1.2	7	0.17	17	38	0.4
L-*allo*-threonine	213	0.24	887	3.7	7	0.53	77	1.7	45
L-Serine	1.9	16	0.11	-	-	-	-	-	-

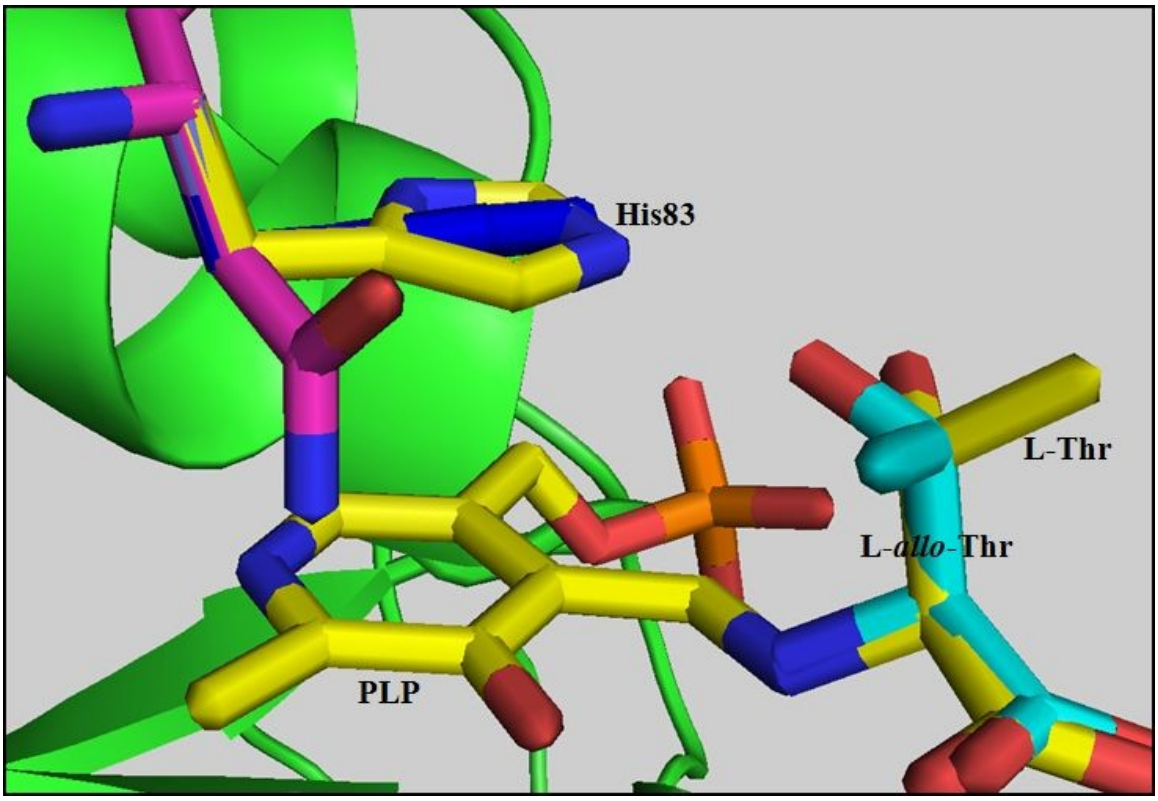
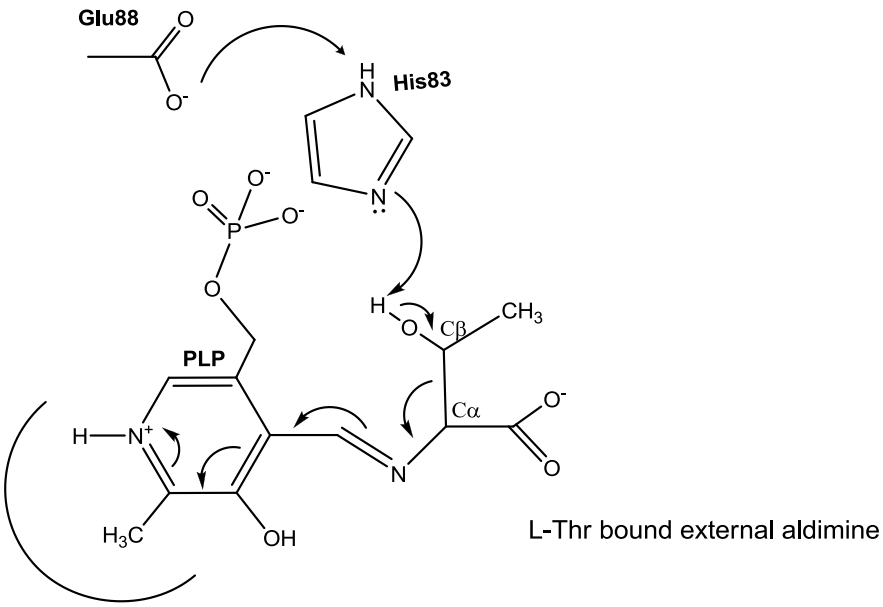


Figure 27: Hydroxyl group of L-threonine (yellow) bound to PLP is close to His83 to make hydrogen bond interactions with it. The hydroxyl group of the modeled L-*allo*-Threonine (cyan) is closer to His83. This figure also depicts the mutation of His83 to Asn (magenta) and Phe (blue) in the structure of eTA/ Thr7.5 complex.

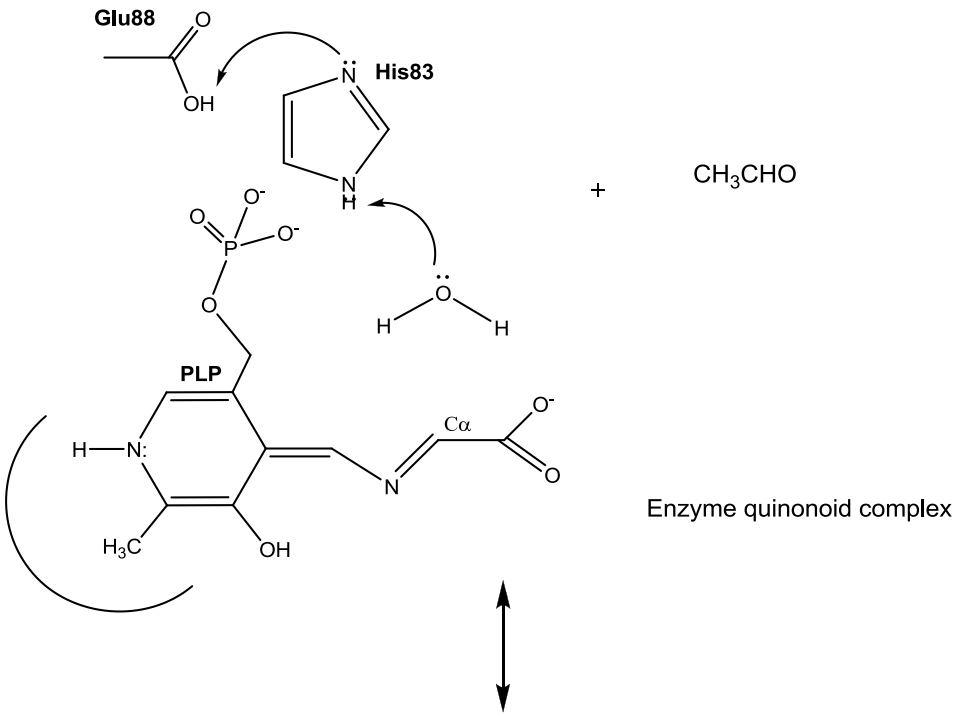
A mechanism has been proposed here based on the structural and mutagenesis studies involving His83 as the catalytic base (Figure 28). The external aldimine formed between the substrate L-Thr and the aldehyde group of PLP positions the hydroxyl group of L-Thr close to His83. The mechanism based on the earlier proposal of His83 to be the catalytic base involves abstraction of the L-Thr hydroxyl proton and generation of the quinonoid complex (Step I and II). The abstraction of the hydroxyl proton by His83 will be enhanced by a residue with an acidic side chain, Glu88 in Step I. The bond angle and bond distances between the hydroxyl proton and His83 are optimum for proton abstraction. On examination of the L-Thr and Gly bound external aldimines, we observe

a water molecule to occupy the position of hydroxyl group of L-Thr in Step II. This water molecule could be involved in generation of the Gly bound external aldimine. The basicity of this water molecule to act as a nucleophile in Step II is enhanced by a water mediated interaction with the phosphate group of PLP.

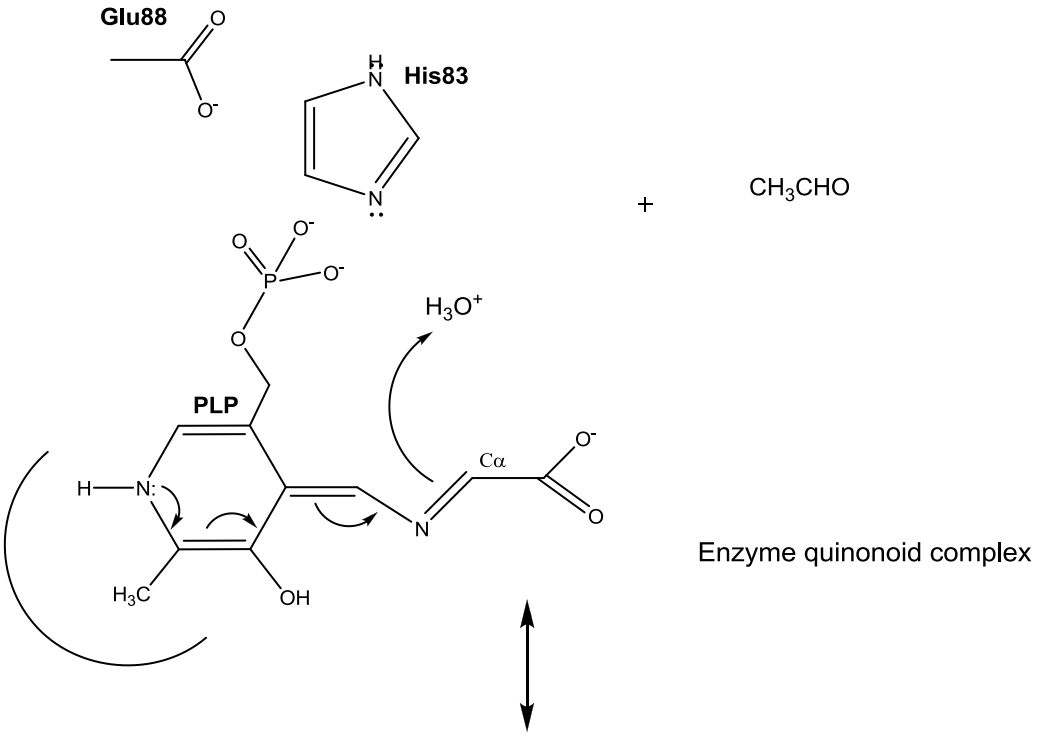
Step I



Step II



Step III



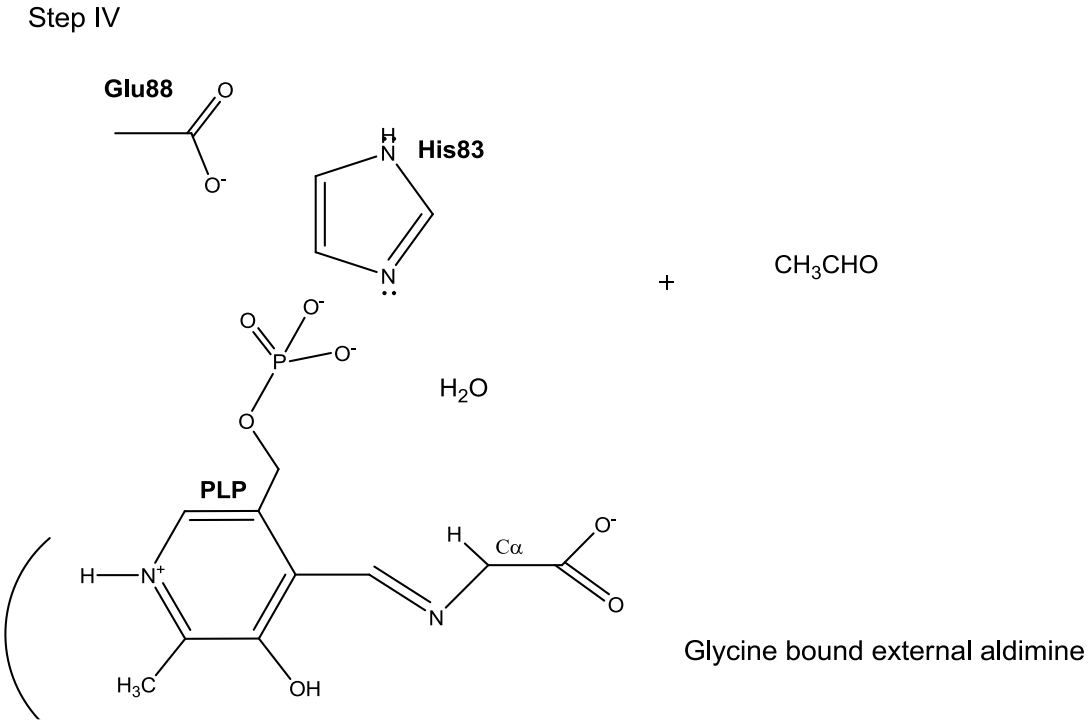
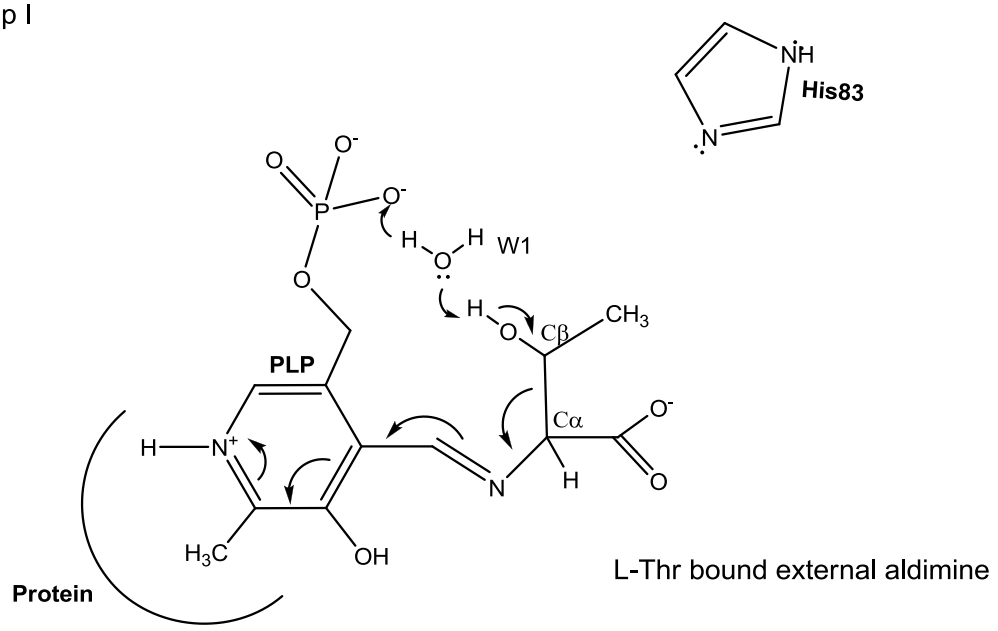


Figure 28: Retro-aldol cleavage mechanism with His83 as the active site base

The site-directed mutational studies by our collaborators shows some decrease in the catalytic function of eTA with L-Thr (L-*allo*-Thr or L-Ser) when His83 is replaced with either Asn or Phe (Table 12). However, this decrease in catalytic function is not significant enough to render His83 as the sole base in the catalysis of β-hydroxy amino acids. Thus, His83 does not seem to be the active site base but it seems to be in the involved in catalysis and may have some different, nevertheless important function. We believe that His83 might be involved in substrate position at the active site owing to its favorable position in the catalytic apparatus

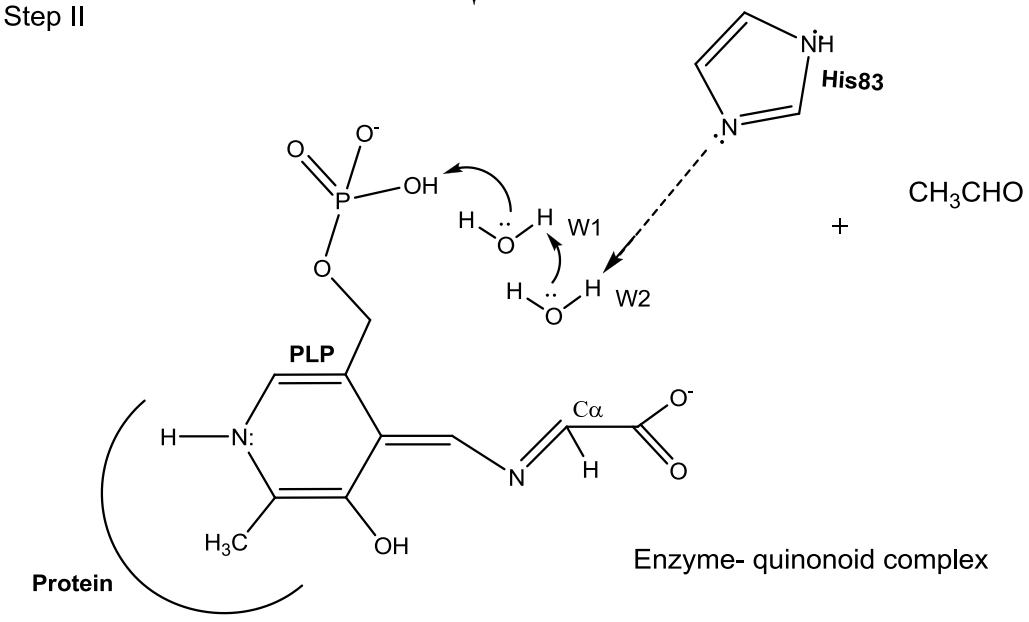
It seems unlikely that His83 could act as the catalytic base in the retro-aldol cleavage reaction. Hence, an alternate mechanism with a more plausible basic candidate was proposed (Figure 29). The reaction commences as pointed earlier, with the substrate (L-Thr) forming an external aldimine with the co-factor PLP. We propose that a spacially close water molecule acts as the catalytic base that abstracts the hydroxyl proton from the bound L-Thr (Step I). This water molecule, W1 is a good enough nucleophile owing to its proximity to the negatively charged phosphate group of PLP. The concomitant release of the aldol reaction acceptor aldehyde generates a quinonoid complex (Step II). A second water molecule, W2 acts as the acidic group involved in generation of the product glycine bound external aldimine from the quinonoid complex (Step II and IV). The acidity of W2 is enhanced by His83 which is at a distance of around 2.9Å. The possibility of a single water molecule, W1 that interposes between phosphate group of PLP and the Cα of the quinonoid complex performing simultaneously the functions of the catalytic base cannot be ruled out. However, W1 has to move closer to the Cα of the quinonoid complex to also perform the function of W2. Our proposed mechanism involving two water molecules instead of one or His83 is more reasonable and probable. Additional kinetic experiments need to be performed to substantiate this mechanism.

Step I



L-Thr bound external aldimine

Step II



CH_3CHO +

Enzyme- quinonoid complex

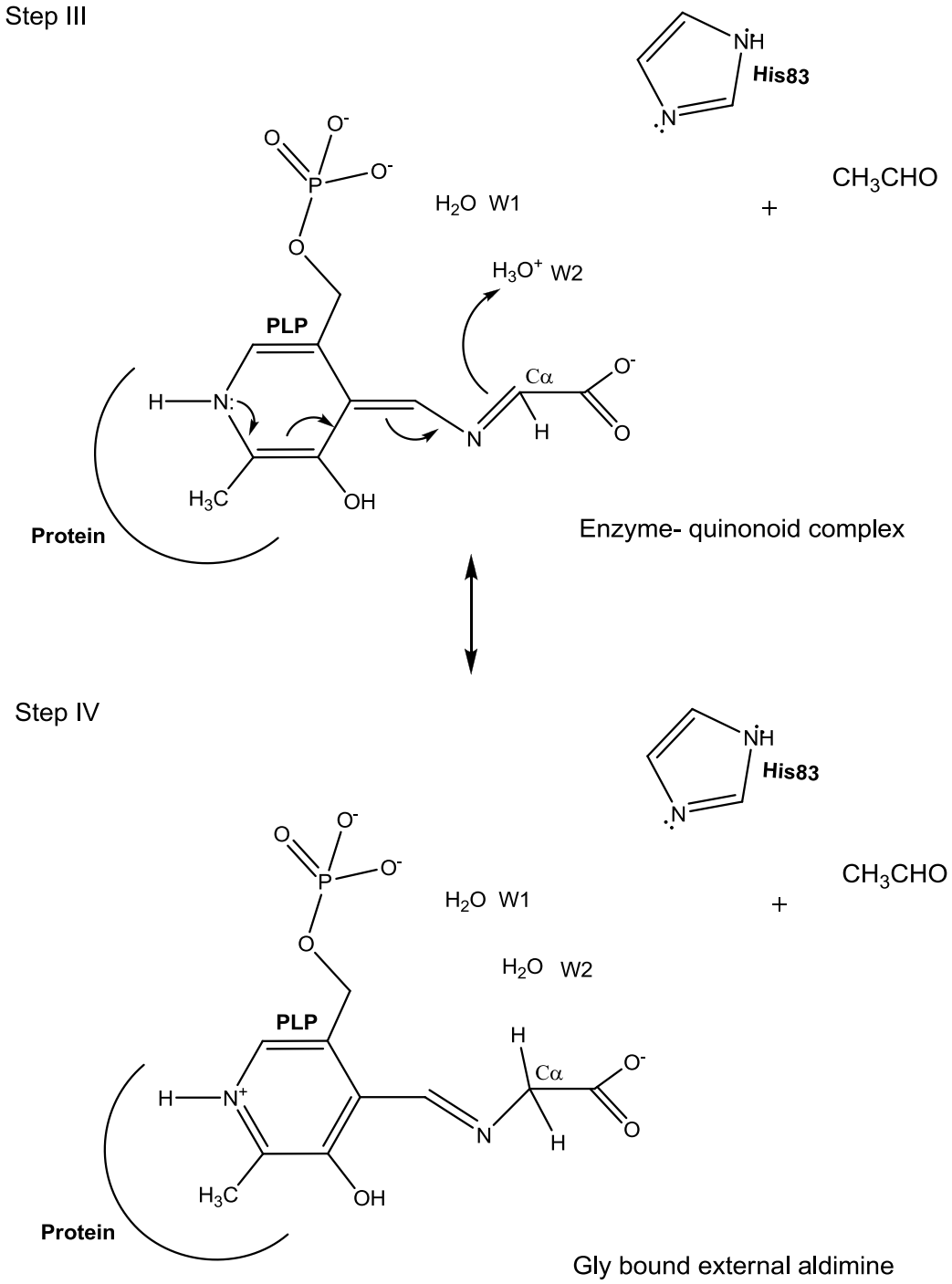


Figure 29: Alternate mechanism proposed for the retro-aldol cleavage reaction with water molecule acting as the active site base

His126 as Substrate Specificity Determinant:

Like His83, His126 has also been proposed to be the catalytic base. The mutational studies clearly suggest otherwise since mutating the residue results in increase in kcat (Table 13). Rather we propose that His126 may be acting as the substrate specificity determinant. Comparative analysis of native eTA structure with eTA/ Thr7.5, eTA/ Ser7.5 and eTA/ Ser5.6 (Figure 30) suggest that His126 may be acting as a 'gating residue', affecting substrate entry and catalysis at the active site.

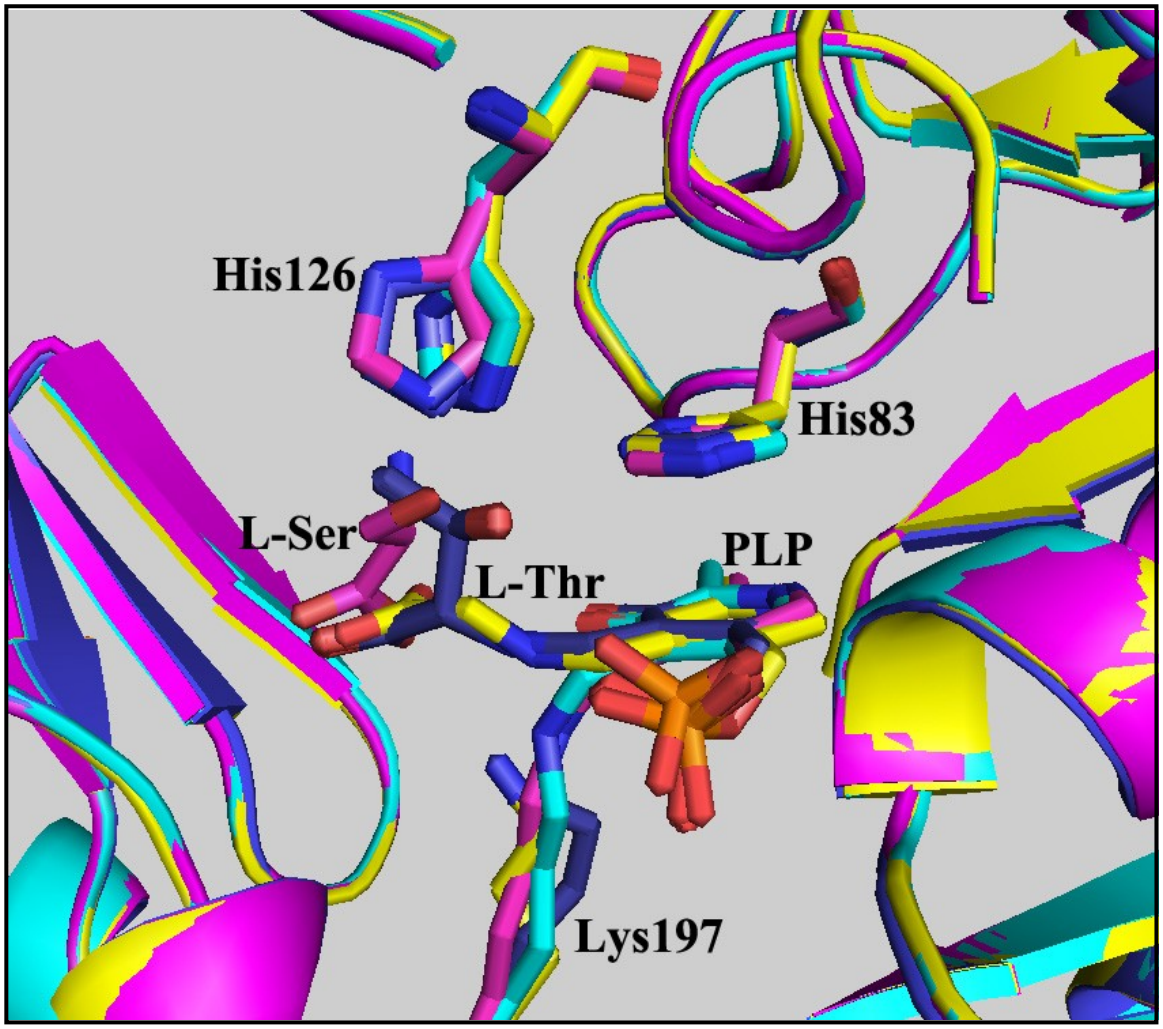


Figure 30: eTA native (yellow) structure at pH7.5 superimposed on eTA/ L-Serine complex at pH 5.6 (magenta), eTA/ L-threonine complex at pH 7.5 (blue) and eTA/ L-serine complex at pH 7.5 (cyan)

Our hypothesis is that the hydrogen bond interactions between the substrate and His126 dictate the entry of substrates into the active site, as well as optimal positioning of the substrate for catalysis. Consistently, in the eTA/ Ser5.6 structure where the hydroxyl group of L-Ser makes closer interaction with His126 the substrate is further removed from the proposed base, His83 and shows almost no catalytic activity with the substrate

being trapped at the active site. eTA/ Ser7.5 structure shows the hydroxyl group of L-Ser to be making weaker hydrogen bond interactions to His126 and in eTA/ Thr7.5 structure interaction of the hydroxyl group of L-Thr with His126 is even weaker, and consistently the catalysis of L-Ser is slower than L-Thr at pH 7.5. When L-*allo*-Thr is modeled at the active site, there is negligible interaction of the hydroxyl group of L-*allo*-Thr with His126 consistent with the maximum catalytic turnover of this substrate. Thus, the strength of the interaction between the hydroxyl group of the substrates and the side chain of His126 decreases from L-Ser at pH 5.6 > L-Ser at pH 7.5 > L-Thr at Ph 7.5. This trend is reversed with respect to interaction with His83 as described above and concomitantly with kcat.

The theory of His126 being a substrate specificity determinant, is substantiated by the mutation study since mutation of His126 results in decrease in substrate specificity from 150 [$(k_{cat} / K_m)_{L\text{-}allo\text{-}Thr}$ to $(k_{cat} / K_m)_{L\text{-}Thr}$] to 13 and 120 for H126F and H126N mutants respectively (Table 13).

As expected, the k_{cat} for L-*allo*-Thr and L-Thr increase with the H126F mutant which abrogates hydrogen bond interaction if any with residue at position 126. Replacement with Asn126 also leads to increase in kcat but expectedly smaller than observed with the Phe mutant as Asn126 can still make weak attraction to the substrate (Table 13.).

We observe no significant change or slight increase in the Km of L-*allo*-Thr in the H126Fand His126N mutants. This is also expected since in case of L-*allo*-Thr, interaction with His126 is very weak and therefore any change at this position should not significantly affect the affinity of L-*allo*-Thr. Nevertheless, with L-Thr we observe a

decrease in Km for the H126F mutant but an increase in the Km in the H126N mutant. Phe seems to sterically push the hydroxyl group to His83 which may explain the decrease in Km. Asn, on the other hand still make interaction with the hydroxyl group but weaker, explaining the increase in Km.

Table 13: The table shows the kinetic data of His126 mutants. Also, shown is the kinetic data of L-Ser cleavage in wild type enzyme

Substrate	WT			H126F			H126N		
	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m
L-threonine	112	19.4	5.8	360	1.7	212	262	61	4
L-*allo*-threonine	213	0.24	887	541	0.2	2705	469	0.96	488
L-Serine	1.9	16	0.11	-	-	-	-	-	-

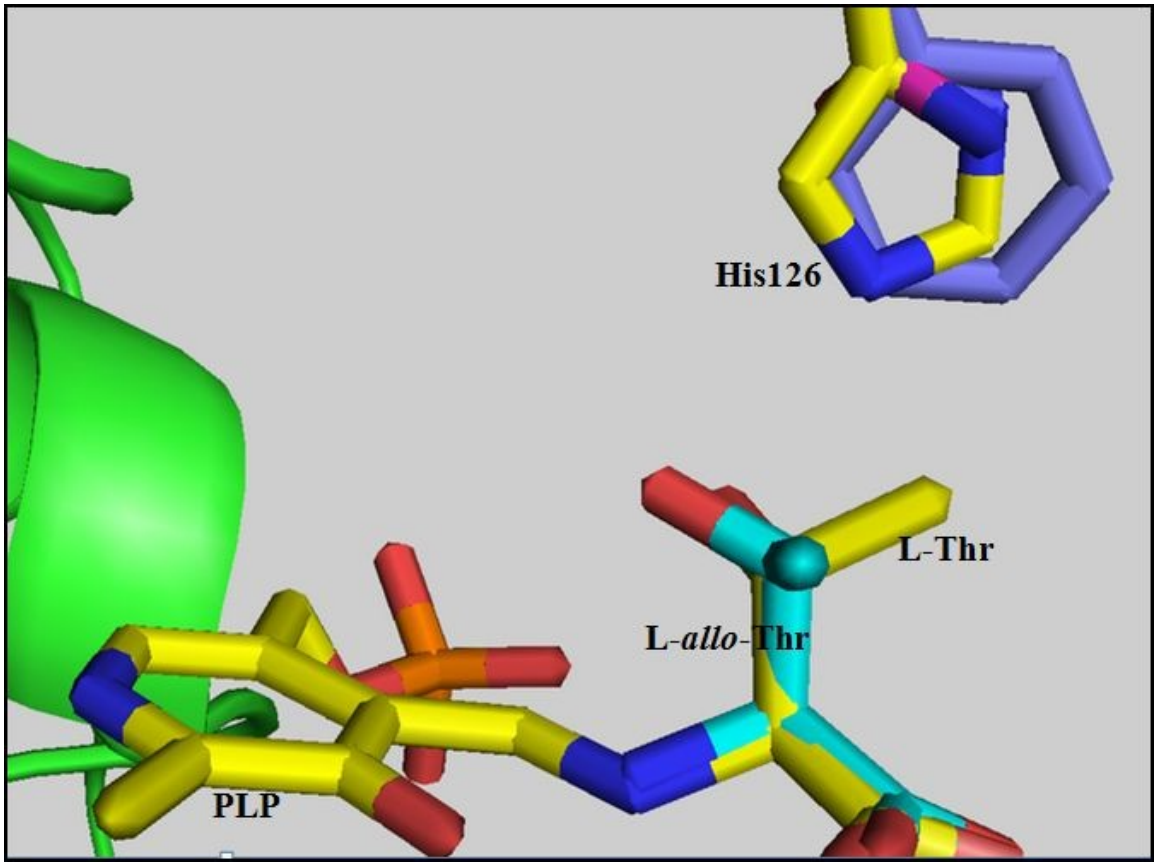


Figure 31: Hydroxyl group of L-threonine (yellow) bound to PLP is close to His126 to make hydrogen bond interactions. The hydroxyl group of the modeled L-*allo*-Threonine (cyan) is closer to His83. This figure also depicts, His126 mutated to Phe (purple) and Asn (magenta).

Thus, to summarize, there is an interplay between the role of His83 and His126 to influence the affinity of the enzyme for L-Thr and its subsequent catalysis (Figure 31). In the case of L-*allo*-Thr, His83 seems to be the determinant for both affinity of eTA for the substrate and rate of catalysis since the hydroxyl group of L-*allo*-Thr is closer to His83 (Figure 32).

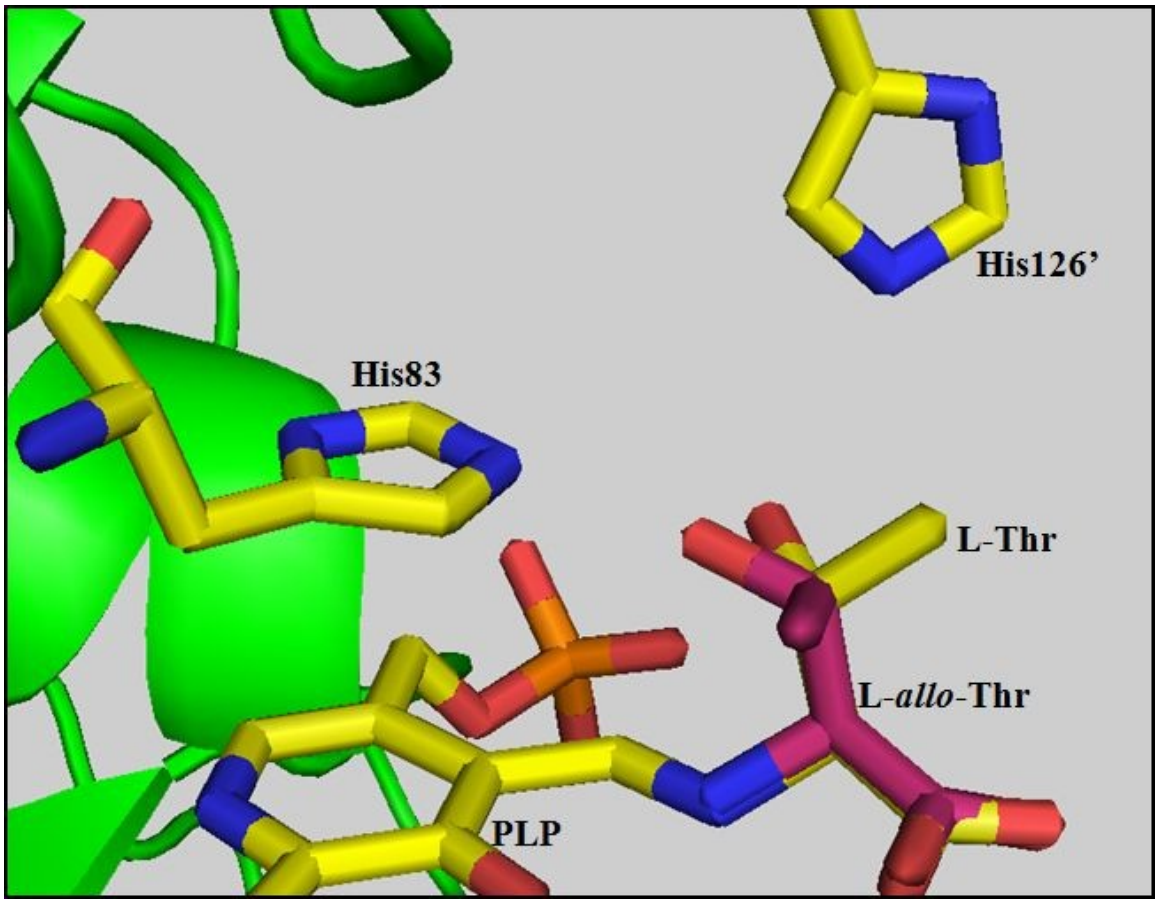


Figure 32: L-*allo*-threonine (magenta) is closer to His83 while L-Threonine (yellow) is equidistant from His83 and His126.

Other residues at the active site*:*

Another important residue at the active site that indirectly influences the catalysis is Glu88. Glu88 (Figure 33) makes close hydrogen bond interactions (between 2.6-2.8 Å) with the side chain of His83. This acidic residue seems to be involved in increasing the basicity of His83 and enhance its ability to abstract protons. Mutation of Glu88 with Gln (Table 14) results in reduced rate of catalysis with both, L-Thr and L-*allo*-Thr. This is most likely due to the fact that the neutral Gln is not as capable in enhancing the basicity of His83 compared to Glu. The associated increase in Km with L-Thr in E88Q mutants is not explained by our structural findings.

Phe87 (Figure 33) is another residue present at the active site that seems to be important in catalysis. Mutation of this residue has deleterious effect on the rate of catalysis of either of L-Thr and L-*allo*-Thr (Table 14). Interestingly, Phe87 is too far away from the active site (more than 5 Å away) from the bound ligands to make any direct and meaningful interactions with the substrate. Nevertheless, its close association with other hydrophobic residues adjacent to the active site highly suggests this residue to be important in maintaining the structural integrity of the active site. Most likely, mutation of this residue results in significant change in the active site structure with a concomitant effect on the catalysis.

Table 14: The table shows the kinetic data of Phe87 and Glu88 mutants.

Substrate	WT			F87A			E88Q		
	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m
L-threonine	112	19.4	5.8	43	8.7	4.9	6.4	49	0.13
L-*allo*-threonine	213	0.24	887	160	0.31	516	67.2	0.21	319

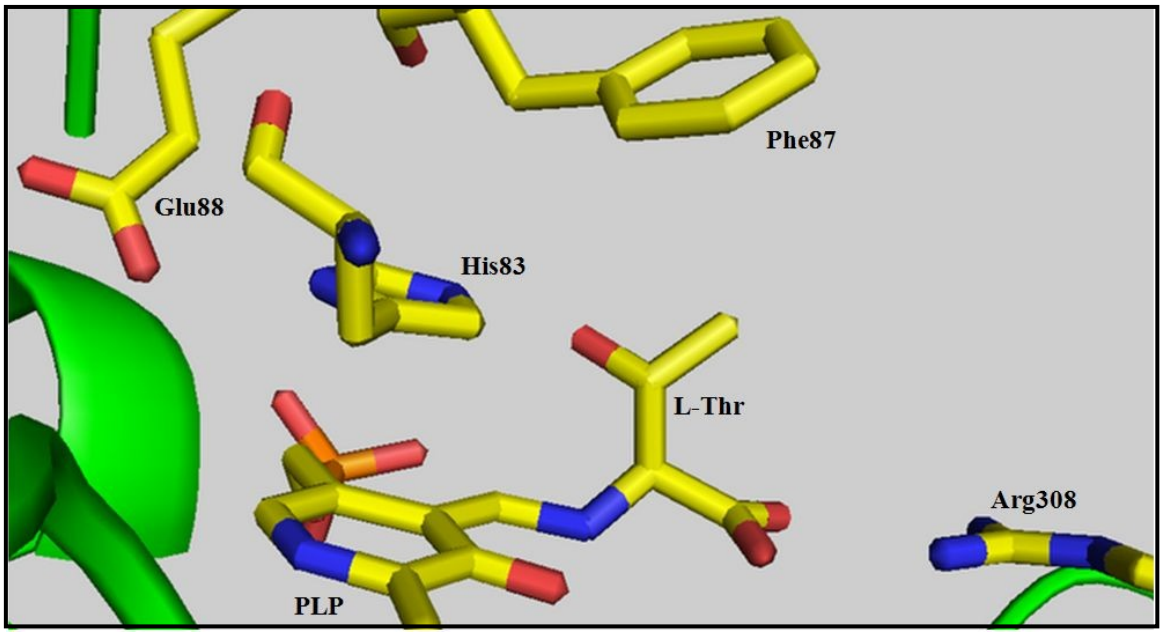


Figure 33: Active site residues depicted in eTA/ L-threonine complex. Other active site residues of eTA are depicted here, namely Glu88 and Phe87.

During external aldimine formation with PLP, Arg308 (Figure 33) forms salt bridges with the carboxylate group of the substrates and positions the hydroxyl group for ready abstraction of the proton by the base for the retro-aldol reaction, His83. Although, mutation studies to substantiate this observation is lacking, the residue is conserved in SHMT structures and has the same role as indicated here.

We have also observed, Cys194 (not shown in Figure 31) that seems to hydrogen bond (3.17Å) to the free Lys197 (when PLP is in the external aldimine form). We think that Cys194 positions Lys197 into the active site for regeneration of internal aldimine. Cys194 makes close hydrogen bond interactions with the residue Ser205 in the native structure while in case of an eTA/ substrate complex (E.g.: eTA/ Thr7.5) it makes hydrogen bond interactions with the free Lys197 as well as Ser205.

Structural comparison between eSHMT and eTA:

eSHMT structure has been studied in great details and several residues have been implicated to have varied functions with an ultimate result of achieving substrate and reaction specificity. eTA active site is more open as compared to the active site of eSHMT. In eSHMT structure, the initial amino acids at amino terminus extend away from their respective monomer and wrap around the other monomer to form a tight catalytic dimer. This is not seen with eTA where the amino terminus is held within the monomer (Figure 34). This is probably the reason that the active site of eSHMT is extensively made up of residues from both the sub-units of the catalytic dimer while in eTA, only a His126 from the other sub-unit is accessible to the active site.

Figure 34: Monomer of native eTA (cyan) superimposed on monomer of eSHMT (magenta) showing an extra loop from the amino terminal of eSHMT that wraps around the other monomer to form the tight dimer.

The rmsd between the monomers of eSHMT and native/ eTA is 1.2 Å. The catalytic dimers are also very different with an rmsd of 4.3 Å. Even with large structural differences the two enzymes belong to the same fold, Fold Type I of PLP-dependent enzymes and can catalyze retro-aldol cleavage of several β-hydroxy amino acids and also racemization and transamination reactions. Consistent with the difference in substrate specificity of the two enzymes, their active sites are designed to accept the respective physiological substrate.

Comparative analysis of the two structures has revealed that some of the residues at the active site of eSHMT are conserved while a few are seem to be unique to eSHMT. A conserved Lys229 (Lys197 in eTA) forms an internal aldimine with PLP. His126 of eSHMT (different from His126 in eTA but corresponding to His83 in eTA) forms π-

stacking interactions with the pyridoxal ring of PLP from the *re* face and Ala202 (Ala 168 in eTA) interacts with PLP from the *si* face. A threonine residue (Thr128) is in close proximity to His126 in eSHMT while in eTA a glutamate residue (Glu88) is close to the conserved His83. Glu88 of eTA seems to enhance the proton abstraction by His83 of eTA and thus, His83 has been proposed to be the base in the retro-aldol cleavage reaction. His126 has not been implicated as a base in eSHMT catalysis probably because a threonine residue (Thr128 in close proximity to His126) lacks the ability to increase the proton abstraction capability of the basic residue, His126 as compared to Glutamine residue in eTA. Other residues like Asp200 (Asp166 in eTA) and Ser175 (Arg169 in eTA) that interacts with the PLP are also conserved.

Tyr65' in eSHMT interposes between L-Ser (substrate for SHMT) and THF (another co-factor required for SHMT) and this interposition is associated with a conformational change in the catalytic dimer of eSHMT leading to the conversion of the enzyme from an 'open' to a 'closed' structure to ensure **reaction specificity**[61]. A residue that corresponds to Ty65' of eSHMT has not been identified for eTA and the occurrence of an open and a closed structure in eTA has not been ascertained (although His126 has been identified as a **substrate specificity** determinant). But some degree of displacement of the small domains of the monomers of native and the substrate bound structures of eTA is evident during comparative analysis of the structures and may be indicative of conformational changes associated with achieving reaction specificity in eTA as well.

The most significant difference between the two enzymes is the conspicuous absence of the binding site for the co-factor, THF in eTA (Figure 33). eTA can catalyze the retro-aldol of L-Ser and does not require this second co-factor. eSHMT catalyzes the

retro-aldol cleavage of L-Serine in the presence of THF. In the absence of this co-factor, eSHMT (and all other SHMTs) catalyzes the retro-aldol cleavage of other β-hydroxy amino acids at increasingly high rates. It may be inferred that THF is involved in the substrate selectivity. His126 of eTA occupies the spacial position of THF found bound to eSHMT (Figure 35). This above observation substantiates the hypothesis from our current structural studies that His126 is a substrate specificity determinant.

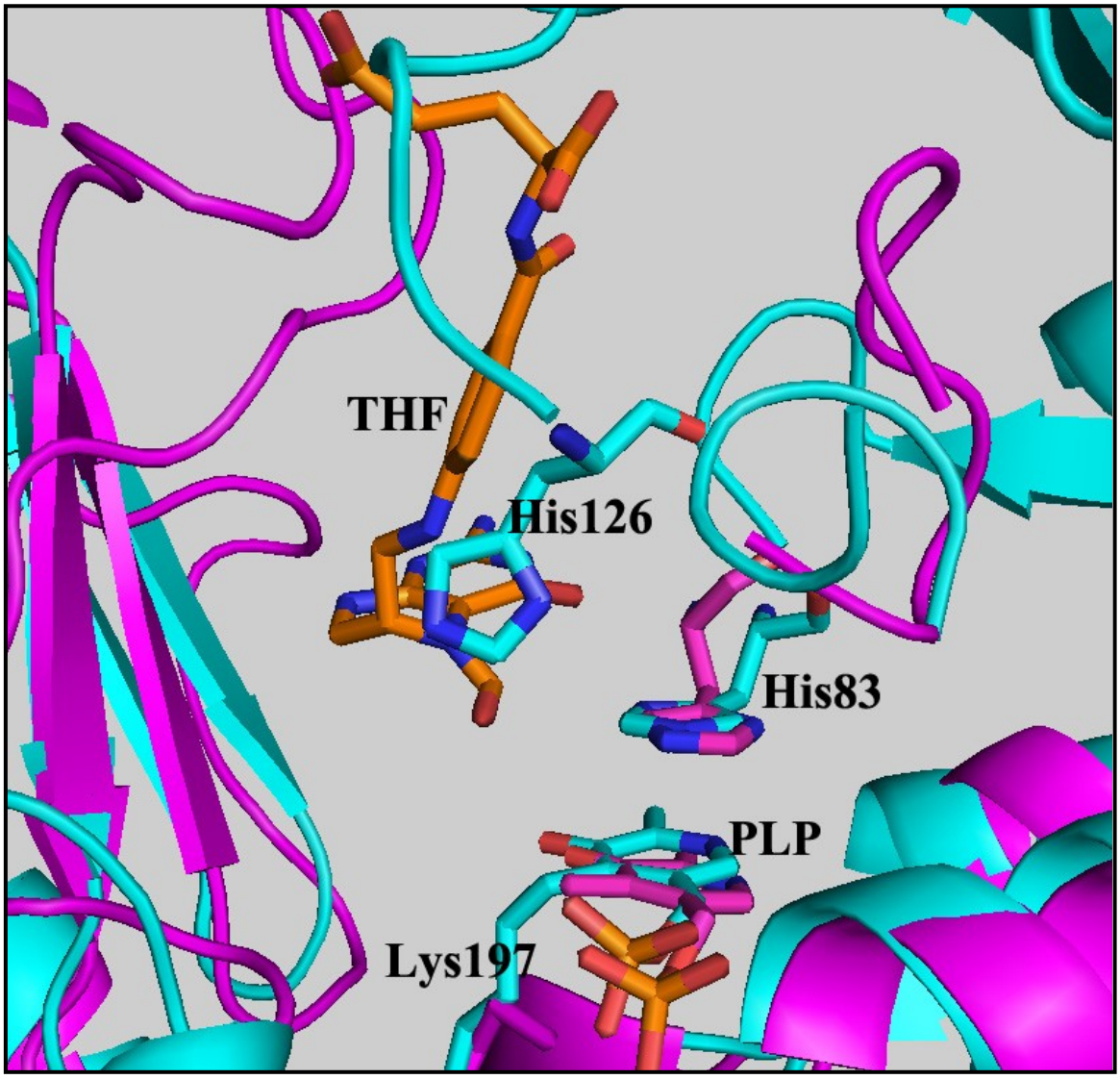


Figure 35: THF (orange) of eSHMT (magenta) is an important co-factor in the retro-aldol cleavage of L-serine. Important active site residues of native eTA (cyan) are also shown in the figure.

Structural comparison of Thermatoga maritima L-TA (TTA) with eTA:

Comparative studies were also done between TTA and eTA to understand the mechanism and the function of the active site residues in greater details. The monomers of TTA and eTA are similar with an rmsd of 0.9 Å but the catalytic dimers of the two enzymes are quite different with an rmsd of 1.1 Å.

The residues of the active site are also conserved between the two enzymes (Figure 36). PLP forms a Schiff base with Lys199 (Lys197 of eTA) of each monomer. PLP is recognized by Asp168 and Arg171 (Asp166 and Arg169 in eTA). The Cα-Cβ bond is almost orthogonal (100^0) to the pyridoxal ring of PLP which compares to 88.5^0 in eTA/ Thr7.5 structure.

As seen in Figure 36, His125 in TTA is localized into the active site from the other monomer (His126 in eTA). This residue has not been implicated to be of significance in substrate selection in TTA as we propose here for eTA.

Tyr 87 (Phe87 in eTA) as shown in Figure 36 has been implicated in as the substrate selectivity determinant in TTA, although it is about 5Å away from the substrate. Phe87 of eTA is also more than 5Å away from the active site for it to make any meaningful interactions with either the substrate or with any residue directly important for catalysis. As discussed, it may be that Phe87 maintains the structure of the active site which indirectly effects substrate catalysis but its possible proximity to the substrate during inter-domain movement cannot be ruled out. His83 in TTA (His83 in eTA) structure stacks on top of pyridoxal ring of PLP and has been implicated as a base of the retro-aldol cleavage reaction. His83 in eTA has also been suggested to be the base for the

retro-aldol cleavage reaction in our structural studies and this hypothesis is supported by the results of mutation studies performed by our collaborators.

In the native TTA structure, five water molecules are seen at the active site. Two of the water molecules form hydrogen bond interactions with Arg316. These waters are also seen in the eTA structure co-ordinated to Arg 308. Another water molecule makes close hydrogen bond interactions with the phosphate group of PLP and slightly weaker hydrogen bond interactions with His125 in TTA. This water molecule is also conserved in eTA structures and has been implicated to be important in the catalysis in the current study (See Alternate Mechanism, Figure 29).

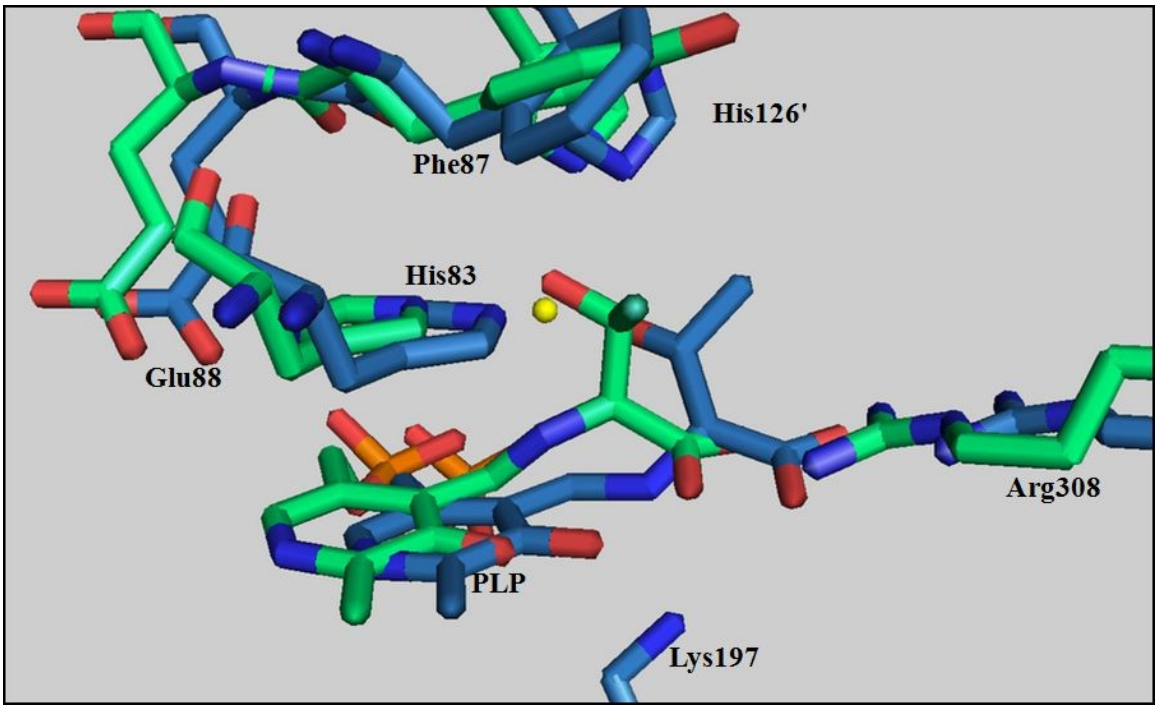


Figure 36: Active site of native eTA (blue) structure is superimposed on active site of native TTA (green) structure (eTA residues have been named in this figure). Corresponding TTA residues are mentioned in the text above. The conserved water molecule is shown in yellow.

2.4 Conclusion:

Crystal structures of eTA were solved in the presence and in the absence of ligands (β-hydroxy amino acids) at pH 7.5 as well as pH 5.6. Pyridoxal ring PLP seems to undergo some degree of rotation between its position in the native and substrate bound structures. eTA/ Ser7.5 and eTA/ Thr7.5 show the presence of low occupancy presence of substrate and product bound external aldimine at the active site. eTA/ Ser7.5 structure shows the presence of very low percentage of internal aldimine with PLP as well. Crystal structure of eTA/ Ser5.6 showed the presence of an abortive complex of L-Serine with eTA at the active site.

Orientation of His126 seems to be altered between native and substrate bound structures and His126 is proposed to be the substrate specificity determinant in the retro-aldol cleavage of β-hydroxyl amino acids. His83 stacks on top of the pyridoxal ring of PLP and is deduced to be the base that abstracts the proton from the hydroxyl group of β-hydroxyl amino acids. Glu88 is positioned to make favorable interactions with His83 and seems to enhance proton abstraction by His83. Phe87 seems to provide structural integrity to the active site. Salt bridges between Arg308 and the carboxylate groups of the substrates are observed in substrate bound structures and may be involved in positioning the substrate at the active site. The proposed roles of the residues are substantiated by the wealth of information provided by mutagenesis studies performed by our collaborators at the University of Rome.

An appreciable degree of movement in evident in the small domains of individual monomers of native eTA and eTA/ substrate complexes but the rmsd between the structures is very low to make conclusions about the occurrence of open and closed

structures in eTA. But based on the movements seen in the small domains it may be inferred that the enzyme achieves reaction specificity by maintaining a closed conformation of the active site as was seen in eSHMT.

eSHMT structure served as a very good model to draw comparison between eTA and other PLP-dependent enzymes. eSHMT catalyzes retro-aldol cleavage of L-Serine in the presence of THF. eTA catalyzes the same reaction in the absence of THF. Through our current study and future work we are trying to understand the evolution of THF binding site in eSHMT.

A mechanism for retro-aldol cleavage reaction of β-hydroxy amino acids inspired by the mechanism of L-Serine cleavage by SHMT is proposed in our current study. His83 is suggested as the base that abstracts a proton from the hydroxyl group of the substrate bound to PLP as an external aldimine. This is followed by the generation of a quinonoid intermediate which is converted to product bound external aldimine. An acidic moiety is involved in this conversion. A water molecule adjacent to the phosphate group of PLP, redundant in the native/ eTA and eTA/ substrate bound structures is suggested to be the candidate that performs this function.

2.5 Future Directions:

1. Molecular dynamics studies may be performed to confirm our hypothesis of hydrogen bonding of the substrates to His126 during their entry into the active site for their subsequent selection for catalysis as well as the occurrence of a closed conformation for the retro-aldol cleavage reaction to ensure reaction specificity.
2. Since eTA is a PLP accepting enzyme and future work may be directed towards understanding the mechanism of PLP transfer from PLP-donor enzymes, the pyridoxal kinases to eTA.
3. The origin of THF binding site may be studied to draw correlation between PLP-dependent enzymes.
4. Engineering of eTA to serve as a better biocatalyst is also a significant future work.

Literature cited

1. Bird, M. and Nunn, P. B. Metabolic homeostasis of L-threonine in the normally fed rat. Importance of liver threonine dehydrogenase activity. *Biochem. J.* **1983**, *214,* 687-694
2. Contestabile, R.; Paiardini, A.; Pascarella, S.; di Salvo, M. L.; D'Aguanno, S.; Bossa, F. L-Threonine aldolase, Serine hydroxymethyltransferase and fungal alanine racemase A subgroup of strictly related enzymes specialized for different functions. *Eur. J. Biochem*. **2001**, *268*, 6508-6525.
3. Liu, J-Q.; Dairi, T.; Itoh, N.; Katoaka M.; Shimizu, S.;Yamada, H. Diversity of microbial threonine aldolases and their application. *J. Mol. Catal.: B Enzy*. **2000**, *10*, 107-115
4. Edgar, A. J. Mice have transcribed L-threonine aldolase/ GLY1 gene, but the human GLY1 gene is a non-processed pseudogene. *BMC Genomics* **2005**, *06*, 32-43
5. Liu, J-Q.; Dairi, T.; Katoaka, M.; Shimizu, S.; Yamada, H. L-allo-Threonine aldolase from *Aeromonas jandaei* DK-39: Gene cloning, nucleotide sequencing, and identification of Pyridoxal-5'-phosphate binding lysine residue by site-directed mutagenesis. *J. Bacteriol*. **1997**, *179*, 3555-3560
6. Schimid, A.; Dordick, J. S.; Hauer, B.; Kieners, A.; Wubbolts, M.; Witholt, B. Industrial biocatalyst today and tomorrow. *Nature* **2001**, *409*, 258-268

7. Vassilev, V. P.; Uchiyama, T.; Kajomoto, T.; Wong, C-H. L-Threonine aldolase in organic synthesis: Preparation of novel β-hydroxy-α-amino acids. *Tetrahedron Lett*. **1995**, *36*, 4081-4084
8. Maruyama, W.; Naoi, M.; Narabayashi H. The metabolism of L-DOPA and L-threo-3,4-dihydroxyphenylserine and their effects on monoamines in the human brain: analysis of intraventricular fluid from parkinsonian patients. *J. Neurol. Sci*. **1996**, *139*, 141-148
9. Sagui, F.; Conti, P.; Roda, G.; Contestabile, R.;Riva, S. Enzymatic synthesis of ω-carboxy-β-hydroxy-(L)-α-amino acids. *Tetrahedron* **2008**, *64*, 5079-5084
10. Vassilev, V. P.; Kajimoto, T.; Wong C-H. Applications of L-allo-threonine aldolase in synthesis of glycoconjugate precursors. *Riken Rev*. **1995**, *08*, 25-26
11. Kimura, T.; Vassilev, V. P.; Shen, G-J.; Wong, C-H. Enzymatic synthesis of β-hydroxy-α- amino acids based on recombinant D- and L-threonine aldolase. *J. Am. Chem. Soc.* **1997**, *119*, 11734-11742
12. Liu, J-Q.; Ito, S.; Dairi, T.; Itoh, N.; Shimizu, S.; Yamada, H. Low specificity L-Threonine aldolase of *Pseudomonas* sp. NCIMB 10558: purofocation, characterization and its application to β-hydroxy-α- amino acid synthesis. *Appl. Microbiol. Biotechnol.* **1998**, *49*, 702-708
13. Steinreiber J.; Fesko, M. C.; Reisinger, C.; Schümann, M.; Griengl, H. Synthesis of γ-halogenated and long chain β-hydroxy-α- amino acids and 2-amino-1,3-diols using threonine aldolase. *Tetrahedron* **2007**, *63*, 8088-8093
14. Elliot. A. C. and Kirch J. F. Pyridoxal phosphate enzymes: Mechanistic, Structural and evolutionary considerations. *Annu. Rev. Biochem*. **2004**, *73*, 383-415

15. Mehta, P. and Christen, P. The molecular evolution of pyridoxal-5'-phosphate dependent enzymes. *Adv. Enzymol. Relat. Areas Mol. Biol*. **2000**, *74*, 129-184

16. Hayashi, H. Pyridoxal enzymes: mechanistic diversity and uniformity. *J. Biochem*. **1995**, *118*, 463-473

17. Schneider, G.; Käck, H.; Lindqvist, Y. The manifold of Vitamin B_6 dependent enzymes. *Structure* **2000**, *08*, R1-R6

18. Jansonius, J. N. Structure, evolution and action of vitamin B_6-dependent enzymes. *Curr. Opinion in Chem. Biol.* **1998**, *08*, 759-769

19. John, R. A. Pyridoxal phosphate-dependent enzymes. *Biochimica et Biophysica Acta*. **1995**, *1248*, 81-96

20. Islam, M. M.; Goto, M.; Miyahara, I.; Ikushiro, H.; Hirotsu K.; Hayashi, H. Binding of C5-dicarboxylic substrate to Aspartate Aminotransferase: Implications for conformational change at the transaldimination step. *Biochemistry* **2005**, *44*, 8218-8229

21. Schirch, V.; Shostak, K.; Zamora, M.; Gautam-Basak, M. The origin of reaction specificity in Serine hydroxymethyltransferase. *J. Biol. Chem*. **1991**, *266*, 759-764

22. Schirch V. and Szebenyi D. M. E. Serine hydroxymethyltransferase revisited, *Curr. Opinion in Chem. Biol.* **2005**, *09*, 482-487

23. Rao N. A., Ambili M., Jala V. R., Subramanya H. S. and Savithri, H.S. Structure-function relationship in serine hydroxymethyltransferase, *Biochimica. Et. Biophysica. Acta* **2003**, *1647*, 24-29

24. Schirch. J. Serine hydroxymethyltransferase, *Adv. Enzymol. Relat. Areas Mol. Biol*. **1982**, *53*, 83-112

25. Schirch, L and Gross, T. Serine transhydroxymethylase. Identification as the threonine and allothreonine aldolase. *J. Biol. Chem*. **1968**, *242*, 5651-5655

26. Ogawa, H.; Gomi, T.; Fujioka, M. Ogawa, H., Gomi, T. and Fujioka, M. Serine hydroxymethyltransferase and threonine aldolase: are they identical? *Int. J. Biochem. Cell Biol.* **2000**, 32, 289-301

27. Yamada, H.; Kumagai, H.; Nagate, T.;Yoshida, H. Formation of threonine aldolase by bacteria and yeast., *Agric. Biol. Chem.* **1971**, 35, 1340-1345.

28. Bell, S. C.;Turner, J. M. Bacterial catabolism of threonine. Threonine degradation initiated by L-threonine acetaldehyde-lyase (aldolase) in species of *Pseudomonas. Biochem. J*. **1977**, 166, 209-216

29. Kallen, R. G.; Jencks, W. P. The mechanism of condensation of formaldehyde with tetrahydrofolate, *J. Biol. Chem*. **1966**, 241, 5851-5863

30. Matthews, R. B.; Drummond, J. T. Providing one-carbon units for biological methylations: mechanistic studies on serine hydroxymethyltransferase, methytetrahydrofolate reductase, and methytetrahydrofolate-homocysteine methyltransferase, *Chem. Rev*. **1990**, 90, 1275-1290

31. Shostak, K.; Schirch, V. Serine hydroxymethyltransferase: Mechanism of racemization and transamination of D- and L-alanine. *Biochemistry* **1988,** 27, 8007-8014

32. Trivedi V., Gupta A., Jala V. R., Sravanan P., Rao G. S. J., Rao N. A., Savithri H. S., Subramanya H. S., Crystal structure of binary and ternary complexes of serine hydroxymethyltransferase from *Bacillus stearothermophilu*s: insights into the catalytic mechanism. *J. Biol. Chem*. **2002**, 277, 17161-17169

33. Strover, P.; Schirch, V. 5-Formyltetrahydrofolate polygltamates are slow tight binding inhibitors of Serine hydroxymethyltransferase. *J. Biol. Chem*. **1991**, *266*, 1543-1550

34. Ulevithch, R. J.; Kallen, R. G. Studies of the reactions of substituted D, L-erythro-β-phenylserine with lamb liver Serine hydroxymethyltransferase. Effects of substitutions upon the dealdolization step. *Biochemistry* **1977**, *16*, 5342-5349

35. Scarsdale, N. J.; Radaev, S.; Kazanina, G.; Schirch, V.; Wright, H. T. Crystal structure at 2.4 Å resolution of *E. coli*. Serine hydroxymethyltransferase in complex with glycine substrate and 5-formyl tetrahydrofolate, *J. Mol. Biol*. **2000**, *296*, 155-168

36. Strover, P.; Shostak, K.; Zamora, M.; Gautam-Basak, M.; Schirch, V. *E. coli*. Serine hydroxymethyltransferase: The role of Histidine 228 in determining reaction specificity. *J. Biol. Chem.* **1992,** *267*, 17679-17687

37. Szebenyi, D. M. E.; Musayev, F. N.; di Salvo, M. L.; Safo, M. K.; Schirch, V. Serine hydroxymethyltransferase: role of Glu75 and evidence that serine is cleaved by retro-aldol cleavage mechanism. *Biochemistry* **2004**, *43*, 6865-6867

38. Scarsdale, J. N.; Radaev, S.; Kazarina, G.; Schirch, V.; Wright, H. T. Crystal structure of rabbit cytosolic Serine hydroxymethyltransferase at 2.8 Å resolution: mechanistic implications. *Biochemistry* **1999**, *38*, 8347-8358

39. Szebenyi, D. M. E.; Liu, X; Kriksunov, I.A.; Strover, P. J.; Theil, D. J. Structure of a murine cytoplasmic Serine hydroxymethyltransferase quinonoid ternary complex: evidence for asymmetric obligate dimers. *Biochemistry* **2000**, *39*, 13313-13323

40. Rao, J. V.; Prakash, V.; Rao, N. A.; Savithri, H. S. The role of Glu74 and Tyr82 in the reaction catalyzed by sheep liver cytosolic Serine hydroxymethyltransferase. *Eur. J. Biochem*. **2000**, *267*, 5967-5976

41. Renwick, S. B.; Snell, K.; Baumann, U. The crystal structure of human cytosolic Serine hydroxymethyltransferase; a target for cancer chemotherapy. *Structure* **1998**, *06*, 1105-1116

42. ClustalW WWW Service at the European Bioinformatics Institute http://www.ebi.ac.uk/Tools/clustalw2/ (accessed on Jul 07, 2010)

43. Keilkopf, C. L.; Burley, S. K. X-ray structure of Threonine aldolase complexes: Structural Basis of substrate recognition. *Biochemistry*, **2002**, *41*, 11711-11720

44. Campbell I. D., The march of crystallography. *Nat. Rev. Mol. Cell. Biol*. **2002,** *03*, 377-381

45. Rhodes G., *Crystallography made crystal clear;* 2nd ed. Academic Press: California, 1993

46. Drenth J., Mesters J. *Principles of protein X-ray crystallography*; 3rd ed.; Springer: New York, 2007; p 332

47. McPherson, A. Introduction to macromolecular crystallography; Wiley-Liss; Hoboken, N. J., 2003; p 237

48. Pai, V. R.; Rajaram, V.; Bisht, S. Structural and functional studies of *Bacillus stearothermophilus* serine hydroxymethyltransferas: the role of Asn^{341}, Tyr^{60} and Phe^{351} in tetrahydrofolate binding. *Biochem. J*. **2009**, *418*, 635-642

49. Schirch, V.; Peterson, D. Purification and properties of mitochondrial Serine hydroxymethyltransferas, *J. Biol. Chem*, **1980**, *255*, 7801-7806

50. Collaborative Computational Project, Number 4 The CCP4 suite: programs for protein crystallography. *Acta Cryst. D*. **1994**, *50*, 760-763.

51. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol*., **1968**, 33, 491-497

52. Claude, J. B.; Suhre, K.; Notredame, C.; Claverie, J-M.; and Abegrel, C. CaspR: A Web server for automated molecular replacement using homology modeling. *Nucleic Acids Res*. **2004**, *32*, W606-W609

53. Adams, P. D.; Grosse-Kunsteleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Cryst. D*. **2002**, *58*, 1948-1954

54. Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; .Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.–S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren G. L. Crystallography and NMR system, *Acta Cryst. D*. **1998**, *54*, 905-921

55. Emsley, P.; Cowtan, K.; COOT: model-building tools for molecular graphics. *Acta Cryst. D*. **2004,** *60*, 2126-2132

56. Cambillau, C; Horjales, E. TOM: a frodo subpackage for protein-ligand fitting with interactive energy minimization, *J. Mol. Graph*. **1987**, *05*, 74-177

57. Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **1996**, *08*, 477-486.

58. Navaza, J.; AMoRe: An automated package for molecular replacement. *Acta Cryst. A*. **1994**, *50*, 157-163.

59. Liu, J.Q.; Dairi, T.; Itoh, N.; Kataoka, M.; Shimizu, S.; Yamada, H. Gene cloning, biochemical characterization and physiological role of a thermostable low-specificity L-threonine aldolase from *Escherichia coli*., *Eur. J. Biochem.* **1998**, *255*, 220-226.

60. Protein interfaces, surfaces and assemblies service PISA at European Bioinformatics Institute http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html (accessed on Jul 07, 2010)

61. Contestabile, R.; Angelaccio S., Bossa, F.,Wright. H. T., Scarsdale, N., Kazanina, G. and Schirch, V. Role of Tyrosine 65 in the mechanism of Serine hydroxymethyltransferase. *Biochemistry* **2000**, *39*, 7492-7500