Structure and Kinetics of Gamma Glutamyl Transferases from *Bacillus species*

A thesis submitted to the University of Pune for the award of a Ph.D. degree in Biotechnology

SHARATH B.

Division of Biochemical Sciences,
National Chemical Laboratory
Pune, India

May 2008

DECLARATION BY THE RSEARCH GUIDE

Certified that the work incorporated in the thesis entitled: " Structure and Kinetics of Gamma Glutamyl Transferases from *Bacillus species*' submitted by Mr. Sharath B, for the degree of Doctor of Philosophy, was carried out by the candidate under my supervision at the Division of Biochemcial Sciences, National Chemcial Laboratory, Pune – 411008. Material acquired from other sources has ben duly acknowledged.

Dr. (Mrs.) Asmita. A. Prabhune
Research Guide

DECLARATION OF THE CANDIDATE

I hereby declare that this thesis entitled "Structure and Kinetics of Gamma Glutamyl Transferases from *Bacillus species*" is entirely original and was carried out by me under the supervision of Dr. Asmita A. Prabhune in the Division of Biochemical Sciences, NCL, Pune. Any material borrowed from other sources has been duly acknowledged. I further declare that this material presented here has not formed the basis for the award of any degree or diploma of any University or Institution.

Sharath B.

(Research Scholar)

May 2008

Acknowledgements

Unlike art, scientific research is an endeavor that cannot be successful without the help and friendship of a number of people. In case of doctoral studies, this influence extends not just to the academics but also to the social life. It is imperative that these people be thanked profusely as they offered help without anticipating anything in return. This altruism not only makes doctoral life conformable but also a memorable experience.

First and foremost, my sincere thanks are due to Dr. Asmita Prabhune, Prof. Tony Wilkinson and Dr. Jim Brannigan for providing advice and laboratory facilities. Sincere thanks to Prof Guy Dodson for help in securing the Commonwealth Scholarship. Dr. Archana Pundle, Dr. C. G. Suresh, Dr. M. I. Khan, Prof. Eleanor Dodson, Dr. Johan Turkenburg, and Prof. Marek Brzozowski are thanked for technical insights at various stages.

The community of fellow scholars (PhD students, post-docs, project assistants and trainees) constitutes the backbone for successful work. This work may not have seen the light of the day without their apparently small but innumerable. My deepest thanks are due to fellow scholars (both past and present) at National Chemical Laboratory, University of Pune and University of York. Among these, my sincerest thanks are due to Satyanarayana, Suresh Kumar, Nagraj, Claudia Schnick, Axel Muller, Asmita Tingare, Nupur Hirve, Atul Kumar, Sridevi, Reetika, Sachin Shah, Shashidhar, Feroz Khan, Atul Thakur, Rohtas Rangera, Anish, Ramesh Raghavan, Nookaraju, Rohini Devi, Sunil Gotwal, Ajit Kanwal, Satish Yadav, Elangovan, Anil Suresh Kumar, Servesh Soni, Sachin Kadam, Ashwini Shete, Vishnu Menon, Jayashree Kantak, Ana Silva, Justyna Korczynska, Claudia Szolkowy, Jitka Vevodova, Andrea Rawlings, Carlos Martinez, Leong, Jean Whittingham.

Sincere thanks are due to Prof. S. Mazumdar, TIFR, Mumbai for permitting the use of CD spectrometer. Sally, Simon, Sam and Caroline are thanked for help in lab. CSIR, New Delhi and Commonwealth Scholarship Commission, London are thanked for financial support. British Council, Manchester and Mumbai units are thanked for arranging the visit to UK.

CONTENTS

ABSTRACT

CHAPTER 1: INTRODUCTION

Gamma Glutamyl Transferases (GGT) are highly conserved enzymes that act on substrates of the form Glu-γCO-NHR. The enzyme removes the terminal glutamate and transfers it to water (hydrolase activity) or an acceptor compound with a free amine group (transferase activity). Under physiological conditions, mammalian GGTs appear to act on glutathione and contribute to its homeostatic level. Human GGT has been implicated in a number of pathological conditions like diabetes, neurodegenerative diseases, induction of apoptosis and cancer. The physiological function of plant and bacterial GGTs is not known unambiguously. The homologue constitutes a virulence factor in *Helicobater pylori* and *Bacillus anthracis*. Structurally, GGT belongs to the superfamily of N-terminal nucleophile (Ntn) hydrolases. The enzyme is produced as an inactive single chain precursor which undergoes self-activation to form an active heterodimer. The side chain hydroxyl group of the Thr residue that occurs at the newly exposed N terminus functions as the catalytic nucleophile. *B. subtilis* GGT is an extracellular enzyme composed of two subunits weighing 40 and 20KDa. The enzyme is speculated to support nitrogen nutrition during adverse conditions.

CHAPTER 2: EXPERIMENTAL PROCEDURES

This chapter encompasses the experimental methods employed during the course of the present study. The chapter is organised into two sections. Section 1 deals with the description of molecular,

biochemical and kinetic methods. *B. subtilis* GGT was cloned into pET 26 vector, expressed in *E. coli* strain BL21 and purified by nickel affinity chromatography. The kinetics was determined using γ-glutamyl-(3-carboxyl)-4-nitroaniline as the substrate. Glycylglycine was used as model acceptor for studying transferase activity. Section 2 deals with the determination of the crystal structure by x-ray diffraction method.

CHAPTER 3: RESULTS

B. subtilis GGT hydrolysed γ-glutamyl-(3-carboxyl)-4-nitroaniline with a K_M of 27mM and k_{cat} of 366 s^{-1} at pH 11.0 The pH optimum for the hydrolase activity was found to be dependent on the concentration of the substrate. It was 9.0 and 11.0 at low and high concentrations respectively. Furthermore, an interesting phenomenon involving the kinetics of hydrolase activity was observed. The kinetics was hyperbolic in the pH range 7.0-9.0 but transformed into sigmoid form at and above pH 10.0.

Inclusion of 50mM glycylglycine to induce transferase activity resulted in reduction of K_M for donor by about 25 fold. However, the impact on V_{max} was marginal. Thermodynamics of transpeptidation and hydrolysis were compared by preparing Eyring plots. The change in enthalpy of activation was found to be the primary difference between the two reactions.

A single crystal of the enzyme belonging to primitive orthorhombic space group ($P2_12_12_1$) was obtained. This crystal was used to collect

diffraction data to a spacing of 1.84Å. The data was solved by molecular replacement method. The final model had R_{cryst}/R_{free} (%) of 21/25. The chemistry of enzyme-substrate interaction was analysed by building glutamate, acylated glutamate and γ-tetraglutamate in the model using molecular dynamics.

The active site was compared with that of *E. coli* and *H. pylori* GGTs. The lid loop, which forms a hydrophobic barrier near the inlet to the active site, was found to be missing in *B. subtilis* GGT due to truncation of the polypeptide chain. The structure, especially the active site, was also compared with class IV cephalosporin acylases and aberrant GGTs.

CHAPTER 4: DISCUSSION

The transferase activity of *B. subtilis* GGT is unique as is it affects only the K_M with negligible impact on V_{max}. This is in contrast to the transferase activity of mammalian homologues wherein both the parameters are affected. Also, the nature of change is contrasting. In case of mammalian homologues the K_M in increase by about 130 fold while it decrease in case of *B. subtilis* GGT by nearly 25 times.

The hydrogen bonds in the microenvironment of Thr^{403} were analyzed to examine the possible involvement of forces with potential contribution towards nucleophilicity. Formation of charge at Oγ and NH_2 of Thr^{403} appear to be promoted by stabilizing interactions. Thr^{421} Oγ balances the negative charge on the nucleophile while the positive charge on α-NH_2 is stabilized by an immobilized water

molecule and the backbone oxygen atom of Met^{88}. Thus displacement of proton from the side chain from $Thr^{403}O\gamma$ involves not only 'pull' from the general base catalyst but also 'push' from a general acid catalyst.

In comparison to its homologues, the K_M of *B. subtilis* GGT for γ-glutamyl-4-nitroaniline is relatively higher. This coincides with the absence of the so called 'lid loop'. The absence appears to be an adaptation to permit binding of larger substrates like γ-tetra-glutamic acid. The affinity of *B. subtilis* for this γ-glutamyl compound is considerably higher than for γ-glutamyl-(3-carboxyl)-4-nitroaniline. In addition to the absence of lid loop, the higher affinity appears to arise due to interactions between the enzyme and the leaving group. These results ascribe a chain of gamma linked glutamic acids as the probable physiological substrate of *B. subtilis* GGT.

Structural comparison of GGT with class IV cephalosporin acylases reveals close homology in the overall structure as well as in the configuration of the catalytic machinery. However, the two families differ in the nature of the binding pocket. These observations indicate that the two families evolved from a common ancestor by differential specialization of the binding pocket.

CHAPTER 5: SUMMARY AND CONCLUSIONS

This chapter discusses the salient features of the present investigation with respect to kinetics and structure-function relationships of the enzyme.

APPENDIX

This chapter records the biochemical and structural studies on CapD enzyme from *B. anthracis*. CapD is a member of the GGT family, but is incapable of hydrolyzing γ-glutamyl-4-nitroaniline, the common GGT substrate. The enzyme was obtained by cloning the cognate gene in pET 28 vector. There are indications that the enzyme uses diaminopimelic acid as acceptor. Experimental analysis indicates that this compound enhances the enzyme's hydrolytic activity on γ-D-poly-glutamic acid. Attempts were made to crystallize the enzyme. Crystals were obtained that could diffract up to 2Å. However, the crystals were not single and therefore could not be used for data collection.

List of publications:

1. Kinetic and Structural Evaluation of catalysis in Gamma Glutamyl Transferase from *Bacillus subtilis*. Manuscript under preparation.
2. Nature of transpeptidation catalyzed by Gamma Glutamyl Transferase from *Bacillus subtilis*. Manuscript under preparation

LIST OF ABBREVIATIONS USED

AGA	Glycosylasparaginase
Bis-Tris	1,3-bis(tris(hydroxymethyl)methylamino)propane
BSH	Bile Salt Hydrolases
CA	Cephalosporin Acylase
CAPS	*N*-cyclohexyl-3-aminopropanesulfonic acid
CD	Circular Dichroism
Da	Dalton
DEAE	Di ethyl amino ethyl
DmpA	L-aminopeptidase D-Ala-esterase/amidase
DON	6-diazo-5-oxo-L-norleucine
EcAIII	Isoaspartyl Aminopeptidase
GGL	Gamma-Glutamyl Leuckotrinase
GGT	Gamma Glutamyl Transpeptidase
GGT-rb	GGT variant from rat brain
GGT-rel	GGT variant from human placenta
Grpp	Glutamine-PPRP-Amidotransferase
Glu-1…Glu-4	Position of γ-glutamyl residues in γ-tetraglutamic acid from N-terminus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High Performance Liquid Chromatography
HSL V	Heat Shock Locus V
k_{cat}	Catalytic Efficiency
K_M	Michaelis constant
MALDI	Matrix assisted laser desorption ionization
MME	Monomethyl ether

MPD	2 -methyl-2, 4-pentanediol
Ntn	N Terminal Nucleophile
OAT	Ornithine Acetyltransferase
ORF	Open Reading Frame
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PGA	Penicillin G Acylases
PRO	β subunit of 20 S Proteasome
psi	Pressure per square inch
PVA	Penicillin V Acylases
R_{factor}	Reliability factor
rmsd	Root mean square deviation
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
TOF	Time of flight mass spectrometry
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultra Violet
v	Initial velocity
V_{max}	Maximum velocity

Objectives

Analysis of mechanism of catalysis and substrate specificity of Gamma Glutamyl Transferase from *Bacillus subtilis* by a combination of crystallographic, biochemical and kinetic methods.

INTRODUCTION

Chapter 1

Definition and catalytic reaction

Gamma glutamyl transferases (GGT; E.C.2.2.3.2) are highly conserved enzymes that occur in bacteria, yeast, plants and in animals from nematodes to humans[1]. The enzyme catalyses the removal of the terminal γ-glutamyl moiety from a donor molecule of the general form Glu-γCO-NH-R by lysing the γ-amide bond and transfers it to a receptive molecule. Some of the common donor substrates are glutathione[2], γ-poly glutamic acid[3] or glutamine[4]. Three types of reactions are possible based on the destination of the γ-glutamyl moiety: (1) transfer to water results in hydrolysis, (2) while a transpeptidation reaction ensues on transfer to an 'acceptor' like amino acids or peptides; (3) transfer to another molecule of the substrate results in auto-transpeptidation. The three reactions are shown in figure 1.1. Glutathione appears to be the likely physiological substrate of mammalian GGTs[5]. However, little is known about the physiological substrate of the plant and bacterial homologues.

Nomenclature

The Nomenclature Committee of International Union of Biochemistry and Molecular Biology recommended that this enzyme, E.C.2.3.2.2, (5-L-glutamyl) peptide: amino-acid 5-glutamyl transferase, be referred to as 'γ-glutamyl transferase'[6]. However, many authors continue to use the older name 'gamma-glutamyl transpeptidase'. There are no conclusive evidences to say which of the three reactions listed above occurs under physiological conditions. Thus the connotation to transferase activity in the title may be misleading.

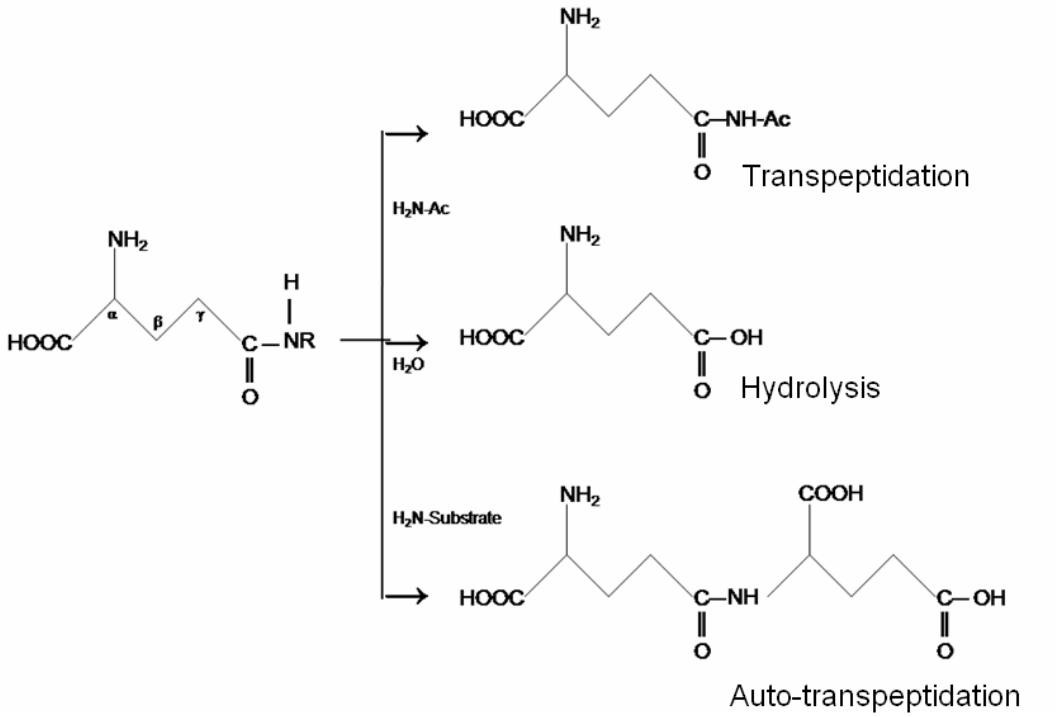


Figure 1.1: Schematic representation of the reactions catalysed by GGT

Prokaryotic and eukaryotic GGTs

Prokaryotic and eukaryotic GGTs differ in both structural and functional aspects. While the prokaryotic GGTs occur as soluble protein either in the periplasmic or extracellular space, the eukaryotic homologues are type II transmembrane proteins. A typical eukaryotic GGT has a large extracellular domain anchored to the membrane by a transmembrane hydrophobic anchor (~8 residues) and a short cytoplasmic tail (containing 4 polar residues). The large extracellular domain, where the catalytic function is located, can be separated from the membrane by papain treatment[7]. The procedure does not affect the enzyme function and is therefore employed in the purification of the enzyme

from animal tissues. The hydrophobic anchor can be removed without affecting the enzyme activity[8]. Furthermore, eukaryotic GGTs have N and O linked glycans whose composition varies in tissue specific manner. The sugars have no catalytic function, as their removal by enzymatic deglycosylation does not affect the activity[9]. The carbohydrates appear to confer protection against proteases. Furthermore, bacterial and eukaryotic GGTs show marked differences in the nature of kinetics. Eukaryotic GGTs are catalytically more powerful than their bacterial homologues. The specific activity of rat kidney GGT is nearly 100 fold higher than that of *E. coli* GGT[10]. Further differences occur in the nature of transpeptidation. The catalytic rate of eukaryotic GGT can be stimulated up to 100 fold by acceptors[10]. In contrast, the acceptors produce marginal change in the rate of prokaryotic GGTs. The fold increase is just ~2 in *E. coli* GGT [11], ~11 in *B. subtilis* GGT [12], and almost none in *H. pylori* GGT [13].

Cellular location

Mammalian GGTs occur on the surface of the epithelial cells mostly in tissues involved in secretion or absorption such as kidneys, bile duct, intestine, pancreas and epididymis[14]. The enzyme is located on the cell membrane even in plant cells[15]. On the contrary, the homologue in yeast is localised on the inner face of the vacuolar membrane[16]. The bacterial homologues are soluble proteins and occur outside the cytoplasm. The enzyme occurs in the periplasm of gram-negative bacteria like *Escherichia coli, Proteus vulgaris* and *Helicobacter pylori* [17,18,19]. In contrast, the homologue in *Neisseria meningitidis* is located in the cytoplasm on the inner face of the inner membrane[20]. The

enzyme in gram-positive bacteria like *B. subtilis* is secreted into the extracellular medium[21].

Expression in mammals

In humans, GGT is represented by atleast 7 different genes, located mostly on chromosome 22. On the contrary, a single *ggt* gene occurs in rat and mouse. In both the cases, the presence of multiple promoters and splicing variations result in a variety of transcripts[22]. The diversity in rodents, which arises due to tissue specific expression, is confined to the 5' untranslated region (5' UTR) with the coding region being essentially same. However, human transcripts extend the variation even to the coding region. Some of the transcripts are truncated and may contain sequence only for heavy or light subunits. It is not known if these incomplete transcripts are eventually translated. Catalytic activity requires the participation of both the subunits. These potentially truncated proteins may therefore serve a non-enzymatic function. Interestingly, human GGT has been shown to induce osteoclasts by a process that is independent of its catalytic activity[23]. Human GGT does not form isozymes and the apparent electrophoretic heterogeneity arises due to tissue specific variation in glycosylation[24].

Catalytic mechanism

Active site: Specificity studies support the classification of the active site into three sub-sites. The γ-glutamyl moiety and the leaving group of the donor bind to sub-site 1 and sub-site 2 respectively. The nature of acceptor binding site (sub-site 3) is ambiguous and will be discussed later.

The γ-glutamyl sub-site exhibits broader optical specificity by accepting both L and D isomers of glutamic acid. Both α-ammonium and α-carboxylate groups of the γ-glutamyl moiety are involved in binding to sub-site 1. However, α-ammonium group appears to be critical as substitution of its nitrogen atom significantly diminishes the affinity. On the contrary, α-carboxylate group tolerates derivatization into uncharged, isosteric or bulkier (e.g., methyl, t-butyl) forms [25,26].

The presence of a discrete site for the binding of acceptor (sub-site 3) is still inconclusive. Some results tend to indicate that the acceptor binds to the site occupied by the leaving group[27, 28]. There are three possible binding mechanisms in reactions that involve two substrates[29] : (1) non-sequential mechanism — both the substrates have discrete binding sites and therefore bind independently; (2) random mechanism — both the substrates occupy the same site by entering randomly (3) sequential mechanism — the native enzyme has affinity only for one of the substrates ; the higher affinity substrate binds to the enzyme and modifies it in a way that improves the affinity for the second substrate; thus the binding of the second substrate is dependent on the first substrate. A transfer reaction can proceed by either sequential or non-sequential mechanism. These two reactions are distinguished by the nature of reciprocal plots produced with various fixed concentrations of one of the substrate. The plots are converging for non-sequential reaction but parallel for the sequential type. Parallel lines have been observed in transpeptidation catalysed by human and rat kidney GGTs [2, 30, 31, 32,33,34,35]. However, the validity of these plots is questionable

because of the imprecision involved in determining the reciprocal plots; thus, plots that appear to be parallel might be converging gradually. Furthermore, interpretation of the secondary plots is complicated by the simultaneous occurrence — in some conditions — of hydrolysis, transpeptidation and autotranspeptidation. However, the catalytic mechanism in GGT is considered to proceed through ping-pong mechanism (a type of sequential reaction) as there are strong evidences for the presence of a modified enzyme in the form of an acyl-enzyme complex.

Acyl-enzyme intermediate: Formation of γ-glutamyl complex has been demonstrated by chemical, kinetic and crystallographic methods. Treatment with glutamine analogs like 6-diazo-5-oxo-L-norleucine (DON)[36] and *O*-diazoacetyl-L-serine (azaserine)[37] inactivates the enzyme by covalent and stoichiometric binding to the γ-glutamyl site. Thus the reaction is assumed to proceed by nucleophilic attack on the amide bond resulting in a tetrahedral transition state whose collapse leads to the formation of γ-glutamyl-enzyme complex concomitant with the expulsion of the leaving group. The free enzyme is regenerated by the reverse reaction involving a non-enzymic group as the nucleophile. Catalysis thus proceeds in two steps with an initial 'acylation' followed by 'deacylation'. The mechanism for hydrolysis is schematically represented below:

$$E + Glu\text{-}R \rightleftharpoons E.Glu\text{-}R \rightarrow E\text{-}Glu + R \rightarrow E + Glu$$

Formation of an intermediate is additionally supported by stop flow studies[38]. Under pre-steady state conditions, the activity follows a

biphasic pattern that has been interpreted to represent a faster acylation step followed by a rate limiting deacylation step. The supposed intermediate has been confirmed by the recent crystallographic studies on *E. coli* GGT, wherein γ-glutamylation of the active site was observed in glutathione soaked crystals[39].

Catalytic nucleophile: Treatment with labelled DON results in localization of radioactivity in the light chain, thus mapping the site of covalent attachment [36]. Furthermore, the attachment was found to be through O-ether bond indicating the involvement of side chain hydroxyl group of either Ser or Thr. Participation of a critical hydroxyl group is in agreement with the inhibitory effect of serine-borate complex[40]. In borate buffer, the affinity of L-serine for the γ-glutamyl binding site is greatly enhanced and results in competitive inhibition. Borates are known to form reversible complexes with vicinal hydroxyl groups. Therefore, the affinity of L-serine is believed to be due to a borate-bridge between its side chain and a hydroxyl group in active site. The nucleophilic residue was finally identified by trapping with a mechanism based inhibitor: 2-amino-4-fluorophosphono-butanoic acid[41]. The treatment results in phosphonylation of a Thr residue that occurs at the N-terminus of the light chain. This residue is conserved in all GGTs. The candidacy of the N-terminal threonine is in agreement with the substrate complexed crystal structure of *E. coli* GGT where a covalent link was observed between the Thr Oγ and the γ-glutamyl moiety[39].

Catalytic mechanism: Insights provided by chemical, kinetic and crystallographic studies have allowed elucidation of the mechanism by which GGT promotes hydrolysis and transpeptidation. The activated site chain of the catalytic Thr ($-O^-$) attacks the carbonyl carbon of the scissile bond, thus forming an unstable anionic tetrahedral transition state. The back bone nitrogen atoms of two tandem Gly residues donate hydrogen bonds to offset the negative charge developed on the carbonyl oxygen. The two Gly residues thus form an oxyanion hole. C-N bond cleavage occurs with the collapse of the tetrahedron simultaneous with the expulsion of the leaving group. The bond breakage occurs concomitantly with the protonation of the amide nitrogen by general acid catalysis[42]. The additional proton satisfies the valency of the nitrogen atom and thus stimulates it to withdraw from the amide bond. At the end of acylation, the γ-glutamyl moiety is left behind by an ester link with the Thr nucleophile. The free enzyme is regenerated from the esterified nucleophile in a reverse process catalysed by a non-enzymic nucleophile. The deacylating function is provided by water or the free amino group of an acceptor in case hydrolysis and transpeptidation respectively. The deacylating agent is motivated to attack the ester intermediate by a general base catalysed deprotonation[43]. The catalytic mechanism is schematically represented in figure 1.2.

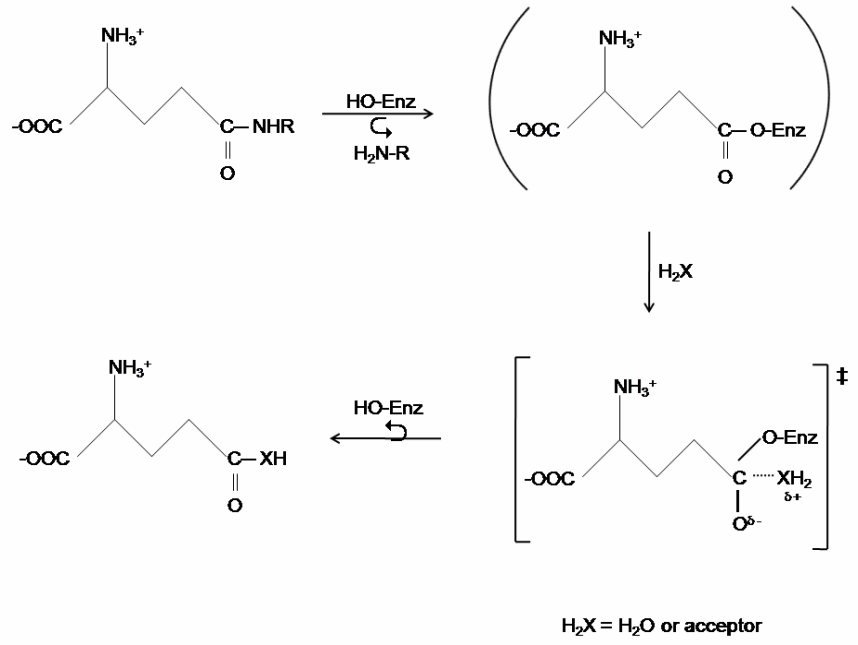


Figure 1.2: Schematic representation of catalytic mechanism of GGT

Measurement of GGT activity

A number of substrates have been reported for assaying GGT activity. Of these, γ-glutamyl-4-nitroaniline is the popular substrate[44]. GGT breaks the molecule into glutamic acid and 4-nitroaniline.The bright yellow colour of 4-nitroaniline (ε_{410} = 8800M^{-1}) allows quantification by spectrophotometry. For conditions requiring higher concentrations, the limited solubility of the anilide is overcome by using γ-glutamyl-(3-carboxyl)-4-nitroaniline. Glycylglycine is used as the standard acceptor for analysing the kinetics of transpeptidation. In clinical diagnostics, the reaction between γ-glutamyl-4-nitroaniline and glycylglycine is used for assaying serum GGT based on the recommendation of International Federation of Clinical Chemistry (IFCC)[45]. The transferase activity is preferred as it produces relatively higher activity than the hydrolytic

counterpart. Accurate measurement, which is crucial in diagnostics, is impeded by autotranspeptidation[46]. The aberrant reaction becomes significant at higher substrate concentrations and results in downward curling of the saturation curve. Thus clinical assays use substrate at a concentration well below the K_M.

Catalytic modulators

Both inhibitors and activators have been reported for GGT. The well known inhibitors are DON[36], serine-borate complex[40], L-azaserine (*O*-diazo-acetyl-L-serine)[38], acivicin (L-(a*S*,5*S*)-a-amino-3-chloro-4,5-dihydro-5-isoxazoleacetic acid)[47], L-methionine sulphoxide (S_cS_s)[48] and γ-phosphono diester analogues of glutamate[49]. These compounds have been useful in the elucidation of the catalytic mechanism. The mode of inhibition by DON and serine-borate complex was described earlier. Ser and borate associate through a reversible link to form the inhibitory complex. It was believed that a permanent link might improve the affinity of the inhibitor and thus a boronate derivative of glutamine was designed and synthesised[50]. As predicted, the affinity of the derivative was improved from 20μM to 17nM. Like DON, γ-phosphono diester derivatives of glutamate inactivate the enzyme by forming a complex with the Thr nucleophile. The sulphoxide moiety of methionine sulphoxide mimics the gamma carbonyl of the substrate and thus interacts with the oxyanion hole to form a reversible complex[51]. These inhibitory mechanisms are shown in figure 1.3.

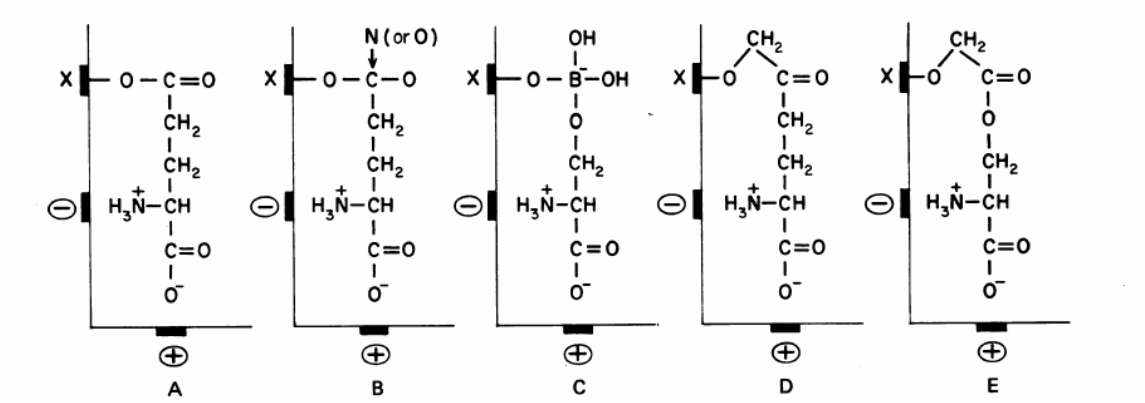


Figure 1.3: Schematic representation of active site of GGT depicting the γ-glutamyl-enzyme intermediate (A), the prospective transition state during transfer to an acceptor or water (B), the prospective serine-borate complex (C), covalent intermediate obtained by treatment with DON (D), and azaserine (E). From ref 40.

Maleate stimulates the hydrolase activity of rat kidney GGT but decreases its transferase function[52]. The differential effect appears to originate from the increased availability of water for deacylation of the γ-glutamyl-intermediate. The consequential improvement of hydrolysis apparently inhibits the transpeptidation reaction. However, the effects are slightly different when glutamine, which is a poor substrate than glutathione, is used in the reaction. In the presence of maleate, the hydrolysis of glutamine increases by >10 fold while its transpeptidation to hydroxylamine rises by ~5 fold. Increase, though differentially, in both the cases might reflect facilitated binding of an otherwise unfavourable substrate. This explains the maleate induced increase in the inactivation rate of γ-glutamyl analogues like DON. Hippurate and its derivatives are also capable of uncoupling hydrolytic and transpeptidation reactions[53]. Unlike maleate, hippurate occurs in

significant amounts under *in vivo* conditions, thus indicating the potential importance of the uncoupling under physiological conditions.

Free bile acids and their glycine and taurine conjugates affect the kinetics of hydrolysis and transpeptidation catalysed by rat GGT[54]. The two reactions are stimulated by free bile acids (cholate, chenodeoxycholate and deoxycholate), supposedly by inducing conformational change upon binding to an allosteric site. In contrast, conjugated bile acids, like maleate and hippurate, stimulate hydrolysis but inhibit transpeptidation. It is believed that the site at which these inhibitors bind partially overlaps with the substrate binding site. Free bile acids and conjugated bile salts predominantly occur in bile duct and intestine. Also, significant levels of GGT are expressed in the epithelium of these tissues. The coincidence in conjunction with other evidences appears to highlight the physiological importance of the modulatory effect.

Genetic deficiency in humans and mice

GGT deficiency is a rare autosomal recessive disease and has been that has been documented in 5 patients[5]. The affected patients show increased concentration of glutathione in blood (glutathionemia) and urine (glutathionuria) and mental retardation. GGT knock out mice have been developed by embryonic stem cell technique[55]. The mutant mice are normal at birth but grow slowly, develop cataract, are half the weight of the wild type, produce grey instead of agouti fur and fail to develop sexually. In the deficient mice, glutathione level in plasma and urine were increased by 6-fold and 2500-fold respectively while

the plasma Cys level was just 20% of the wild type. The debilitating effects were shown to be secondary to cysteine deficiency as these could be reversed by dietary supplementation with N-acetyl cysteine.

Physiological functions

Many lines of evidences point to glutathione as the likely physiological substrate of mammalian GGTs. Glutathione (γ-L-glutamyl-L-cysteinyl-L-glycine) is a thiol peptide occurring ubiquitously in eukaryotic cells at levels as high as 0.5-10Mm[56]. The free sulphydryl moiety enables glutathione to function as the major antioxidant to maintain a reducing intracellular environment. Glutathione neutralises oxidants like peroxides by forming the respective disulphide, while electrophiles are negated by formation of S-conjugates[5]:

$$2GSH + H_2O_2 \rightarrow GS{=}SG + 2H_2O$$

$$GSH + \text{electrophile} \rightarrow \text{GS-conjugate}$$

The physiological availability of Cys, the critical residue in glutathione appears to be limited[46]. Mice fed with protein deficient diet have lower glutathione but higher GGT levels and the digression could be remedied by supplementation with methionine. These findings demonstrate the dietary importance of sulphur containing amino acids in maintaining glutathione homeostasis and the reciprocal role of GGT in the process. However, the mechanism by which GGT contributes to glutathione homeostasis is not clear. GGT was proposed to mediate the transmembrane transfer of amino acids by participating in the 'γ-glutamyl cycle'[57] (figure 1.4). The cycle begins with the transfer of γ-L-glutamyl moiety from glutathione to any amino acid (other than proline) to form γ-L-glutamyl-amino acid. It was assumed that γ-glutamylation

somehow facilitates the transport of the amino acid into the cytoplasm. In this view, glutathione serves as a donor of the vital glutamyl residue. While in the cytoplasm, the glutamylated amino acid forms 5-oxoproline with the concomitant release of the amino acid. Glutamic acid is then reformed from 5-oxoproline (pyroglutamic acid) for reuse in glutathione synthesis. The predominant localization of GGT in epithelia active in absorptive function apparently supports the putative role in amino acid absorption. However, the hypothesis may not be valid as it is in conflict with the observation of unimpaired amino acid transport in GGT deficient humans and animals. Furthermore, the 'γ-glutamyl cycle' was proposed on the basis of circumstantial evidences and was not demonstrated by radioactive tracer method. Thus it appears that the primary function of GGT is in the liberation of Cys from glutathione for the salvage pathway.

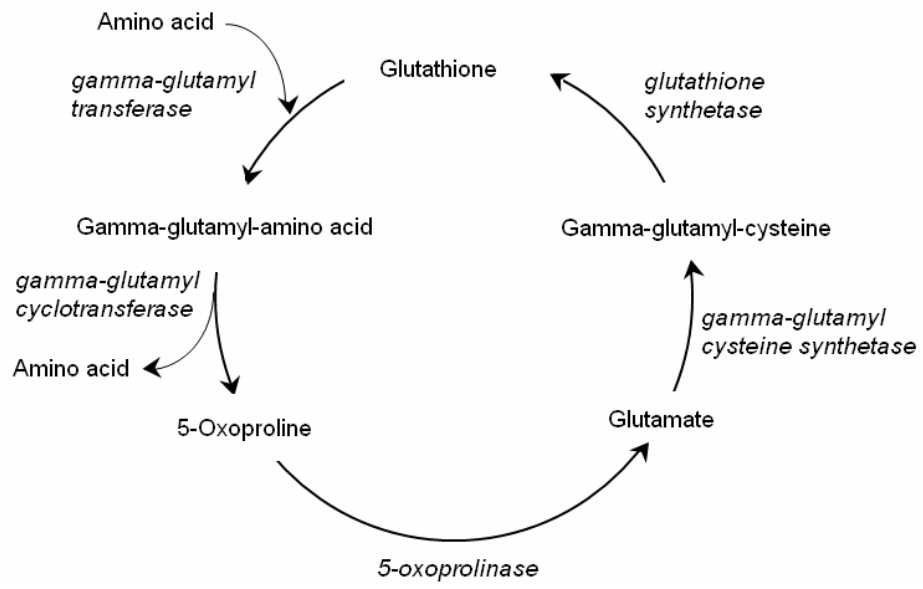


Figure 1.4: Glutamate cycle

GGT also appears to be involved in the formation of mercapturic acids (N-acetyl S-substituted cysteine derivative) from glutathione[58]. Mercapturic acid derivative improves the solubility of the xenobiotics thus enabling their removal by the excretory system. Biosynthesis of mercapturic acids begins in the cytoplasm of livers cells by conjugation with glutathione, a reaction catalysed by glutathione *S*-transferase. The conjugates are then transported into the extracellular space where they are converted into cysteine *S*-conjugates by membrane bound GGT and dipeptidases. These trimmed conjugates are then returned to the cytoplasm for acetylation by N-acetytransferases.

The role of GGT in plant tissues is still unclear as there is no evidence for the presence of glutamyl cycle in plant tissues[59]. Furthermore, some plant GGTs have low affinity for glutathione under *in vitro* conditions. The enzyme is speculated to participate in the biosynthesis of γ–glutamyl dipeptides that are formed during fruit ripening and accumulate in storage tissues such seeds and bulbs in certain plants[60, 61, 62, 63]. In onion, GGT is supposed to catalyse the last step in the formation of precursors of volatile compounds[64]. Studies with suspension cultures of tobacco cells indicate potential participation of GGT in glutathione catabolism[15]. Likewise, little is known about the physiological role of bacterial GGTs. In *E. coli,* inactivation of *ggt* gene impairs the cells from utilizing glutathione as a nitrogen source [65]. However, the apparent association may not be physiologically important as the tripeptide is absent in bacteria like *B. subtilis*[66]. Extensive conservation of an enzyme would be possible only if the

cognate function is likewise conserved. Thus the functions ascribed for the various homologues will have to be considered with caution.

Pathological role

GGT has been implicated in many physiological disorders like neurodegenerative diseases[67], inflammation[68], diabetes[69] and cardiovascular diseases[70]. Overexpression of GGT is often observed in human tumours, and its role in tumour progression[71] and the expression of mailganant phenotypes of cancer cells, such as drug resistance[72] and metastasis [73, 74, 75] have been suggested.

Brugia malayi, a lymphatic filarial parasite, causes tropical pulmonary eosinophilia (TPE), a syndrome involving inflammation of the airway of the epithelium[76]. The filarial GGT functions as an allergen to induce humoral response which subsequently cross-reacts with the homologue present of the pulmonary epithelium of the host. Thus molecular mimicry between the filarial and human enzymes forms the basis for autoimmune response during the infection.

Helicobacter pylori, a gram negative spiral bacterium, causes gastritis and ulcers in humans. GGT is expressed constitutively by *H. pylori* cells and is used as a marker for identification. This enzyme was found to be the virulence factor as its deficiency disables the bacterium from colonising Swiss mice[19]. Subsequent studies with other mice strains and piglets indicate that the GGT may not be primarily necessary for colonization but when present greatly enhances the extent colonisation. *Niesseria meningitidis* is a gram negative bacterium that colonises the

nasopharynx of humans. It sometimes spreads into the bloodstream and reaches cerebrospinal fluid to induce meningitis. GGT produced by this bacterium appears to promote growth by deriving Cys from extracellular γ-glutamyl-cysteinyl compounds[78].

GGT appears to be involved both positively and negatively in the physiology of oxidative stress[46]. The enzyme protects the cells against oxidative effects by its involvement in glutathione synthesis. Induction of oxidative stress results in upregulated expression of GGT. Surprisingly, GGT also appears to generate prooxidant effects. The thiol of Cys-Gly (leaving group) is more reactive than that in the parent molecule. The enhanced reactivity promotes reduction of ferric ion Fe (III) to ferrous ion (II), thus starting a redox–cycling process that could ultimately lead to the production of reactive oxygen species (ROS) and thiyl ($-S^{\bullet}$) radicals. Pathological implications of GGT in inducing oxidative process have been explored in liver carcinogenesis, human atherosclerosis, kidney ischemia and GGT expressing cancer cells.

Diagnostic marker

GGT level in the serum is measured in clinical laboratories as a marker for liver function[79]. The diagnostic assay is based on empirical studies and has been in practise for over four decades. Though precise physiological function of the enzyme is still nebulous, a large number of data is available on factors influencing the activity in serum. GGT assay is a sensitive test but lacks specificity. Many conditions like hepatitis C infection, cholestasis, pancreatitis, diabetes, obesity, excessive-alcohol intake and enzyme-inducing drugs elevate serum

GGT. GGT is particularly sensitive to alcoholic liver diseases and is elevated in large proportion of alcoholics. The levels are lower in 'moderate' and occasional drinkers than in those consuming potentially hazardous volume. Due to this correlation, GGT is useful as a marker for monitoring alcohol deaddiction[80]. Many attempts have been made to correlate abnormal serum GGT with drink-drivers (drivers with non-permissible levels of blood alcohol potentially leading to debilitating driving skills due to neurological impairment) with the objective of restricting the issue of licenses to safe drivers. A significant correlation has been noticed in studies conducted in Germany, Scotland and Norway [81, 82, 83]. The lack of specificity restricts use of GGT test for decision on individuals. The mechanism leading to rise in serum GGT is not well understood. The rise in case of obstructive jaundice is mostly due to solubilization of membranes by bile salts. The rise, on the contrary, is due to enhanced expression in case of alcoholism[47].

Crystal structure of GGT

Recently, the crystal structures of *E. coli* [39] and *H. pylori* [84] GGTs were determined at 1.95 and 1.9Å respectively. The tertiary structure is characterised by a tetralamellar aββa core comprising of two central antiparallel β-sheets embedded between two layers of a-helices. The enzyme is kidney-shaped and encloses a shallow groove in the region below the sheets. The active site is located in a depression located in the groove. The composition of the active site was analysed from crystal complexes with glutamate and γ-glutamyl acyl intermediate in *E. coli* GGT and with glutamate and *S*-(nitrobenzyl) glutathione in *H. pylori* GGT. The substrate is bound to the active site by a number of

hydrogen bonds and salt bridges formed between the enzyme and the α-amino and α-carboxylate groups of the substrate. The interactions are represented in figure 1.5. The glutamyl Cγ receives two hydrogen bonds from the backbone N atoms of two tandem Gly residues. These interactions form an oxyanion hole that stabilises the negative charge formed on the γ-carbonyl oxygen in the tetrahedral transition state. In both *E. coli* and *H. pylori* GGTs, the mouth of the active site is partially closed by a Tyr residue that sits at the apex of the so called lid loop. The conformation of this loop is same in both native and complex structures. There are indications that the loop might be intrinsically mobile and might move transiently to facilitate the entry of the substrate into the active site[85].

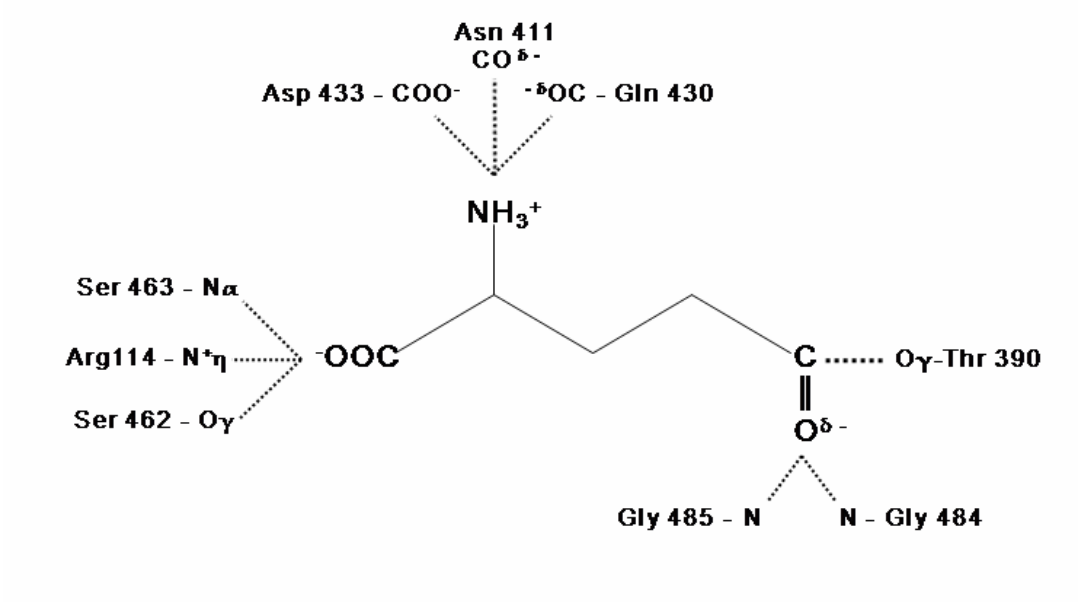


Figure 1.5: Diagrammatic representation of interactions that bind the γ-glutamyl moiety of the substrate to the active site [39].

GGT from *B. subtilis*

B. subtilis GGT is a heterodimer of two chains weighing 45 and 22 kDa[12]. The enzyme is produced mostly in the stationary phase and secreted into the extracellular medium[86]. The enzyme is kinetically unusual as its K_M of 10mM for γ-glutamyl-(3-carboxyl)-4-nitroaniline is relatively higher than that of its homologues[87]. In contrast, the enzyme shows higher affinity for tetra- γ-glutamic acid (K_M= 9μM) and poly- γ-glutamic acid (K_M= 9μM) [3].

Poly- γ-glutamic acid (PGA) is commonly produced by bacteria that belong to the genus *Bacillus*[88]. PGA is a sticky material composed of linear unbranched chain of γ-linked glutamic acids. It is polyanionic and thus highly soluble in water. The polymer may be composed of only D, or only L or both the enantiomers of glutamic acid. *B. subtilis* produces two types of PGA: PLGA (contain only L glutamate) and PLDGA (contains both D and L glutamate). The PGA filaments in *B. subtilis* can vary from 160-1500kDa. The synthesis of the polymer is coded by three *pgs* genes (for polyglutamate synthase): *pgs B, pgs C* and *pgs AA*. There are indications supporting the involvement of a fourth gene called *pgs E*.

During the late stationary phase, PGA is degraded and the resultant monomers are used as a source of nitrogen. GGT is considered to mediate the process as it can hydrolyse PGA under *in vitro* conditions and is expressed predominantly during the stationary phase[3]. Furthermore, *ggt* gene in *B. subtilis* is under the control of quorum sensing system[3]. The hypothetical involvement of GGT in nitrogen

metabolism is additionally supported by the increase in relative frequency of sporulation of mutants lacking GGT[3]. In the wild type cells, sporulation is induced by the nutritional deficiency of the medium[89].

CapD

CapD is a member of a multi-enzyme series called the Cap (capsule) system[90]. This system is responsible for the synthesis of capsule in *B. anthracis,* the aetiologic agent of anthrax. The capsule, which is composed of γ-poly-glutamic acid, forms a mucilaginous lining on the external face of the cell wall. The capsule confers protection to the bacilli against the host immune system by reducing the interaction with phagocytes and antibodies. Most of the PGAs are racemic mixtures of D and L enantiomers. However, the polymer from *B. anthracis* is exclusively made of D-form. The capsule is non-immunogenic due to the D-form of the glutamyl residues and the absence of chemical complexity. Thus, capsule constitutes the principal virulence factor along with the anthrax toxin. Its removal either by genetic intervention or heat treatment reduces the virulence. The genes for the capsule and the toxin are located on pXO-2 and pXO-1 plasmids respectively[92].

The cap system comprises of five enzymes: capB, capC, capA, capE and capD. Cap B and Cap C catalyses the synthesis of PGA in the cytoplasm[91]. The nascent polymer is transported to the cell surface by CapA and CapE. The topical PGA is then covalently linked to peptidoglycan by CapD. Covalent anchorage is vital for the anti-immunoproperty as mutants lacking CapD show reduced virulence[90].

The immobilized nature of PGA is unique to *B. anthracis* as most other bacillary PGAs are released into the medium. Because of the crucial association with virulence, enzymes of the Cap series are potential targets for development of drugs.

CapD is considered to be a member of the GGT family because of sequence homology[90]. The enzyme is synthesized as an inactive precursor that matures by an autocatalytic process. The active enzyme is a heterodimer with Thr at the N-terminus of the light chain. Despite these similarities, CapD is unable to unable to hydrolyze γ-glutamyl-p-nitroaniline, the common chromogenic substrate for assaying GGT activity.

Other γ-glutamyl hydrolases

Earlier, GGT was thought to be the only enzyme capable of hydrolysing the γ-glutamyl bond. This had been the basis for the indiscriminate association of the enzyme with every process that involves break and/or transfer of γ-glutamyl moiety. GGT was considered as a participant in the biosynthesis of leuckotrienes (LTs). LTs are arachidonate derivatives that form the active mediators of inflammatory and allergic reactions. Biosynthesis of LTs occurs mostly in basophils, eosinophils, mast cells and macropahges. The synthesis occurs in many steps, one of which involves deglutamylation of LTC4 to produce LTD4, supposedly by GGT [92]. The putative importance of GGT was further supported by the occurrence of higher levels of LTC4 in the urine of individuals with genetically deficient GGT [93]. However, contrasting results were observed in case of *ggt* knock out mutants of

mouse. The mutant mice are stunted, grow slowly, fail to mature sexually, develop cataract and die prematurely by ~12 weeks of age [94]. But no serious impairment was observed in the cellular conversion of LTC4 to LTD4, thus indicating the presence of an alternate enzyme for the process[95]. Recently, an enzyme capable of supporting LTC4→LTD4 conversion was cloned and characterised[96]. The primary structure of this enzyme, named as γ–glutamyl leuckotrinase (GGL), is homologous to mouse GGT by nearly 41%. However, it fails to act on γ–glutamyl-p-nitroaniline, the common substrate for assaying GGT activity.

An enzyme (called GGT-rel) related to but distinct from human GGT has been cloned from human placental cDNA library[97]. GGt-rel shows 39.5% sequence homology with human GGT, localises on cell membrane, acts on glutathione and converts LTC4 to LTD4. However, it is inactive on the foresaid chromogenic substrate of GGT.

Variant of GGT that expresses exclusively in rat brain (GGT-rb) has been isolated and characterised. It is homologius to the predomiant form of GGT (which is expressed in the liver) 33% identity and 73% similarity. It does not act on γ-glutamyl-4-nitroaniline but can accept glutathione as substarte[98].

In addition to GGT, *B subtilis* produces another γ-poly-glutamate hydrolase (Ywr D) that cleaves the chain between two D-glutamates[99]. The amino acid sequences of Ywr and GGT from *B. subtilis* are

homologous by 27%. The former enzyme is considered to be cytoplasmic as there is no evidence for a signal peptide in its gene.

φNIT1 — a bacteriophage that infects *B. subtilis* — produces a γ-poly-glutamate hydrolase that lacks stereospecificity[100]. It breaks both a and γ linked polyglutamates into tri, tetra and pentamers. The enzyme serves to perforate the bacilli's capsule during infection.

Other enzymes include PGA hydrolases from *Flavobacterium*[101] and *Myrothecium*[102], carboxypeptidase G (3.4.17.11)[103], γ-glutamyl hydrolase (GGH) from animal tissues (EC 3.4.19.9)[104] and glutamate caboxypeptidase II (EC 3.4.17.21)[105]. GGH is a thiol enzyme that hydrolyses γ-glutamyl tail of folyl- γ-PGA. Glutamate caboxypeptidase II is a metalloenzyme that uses Zn^{2+} for endopeptidic lysis of folyl-γ-PGA. Carboxypeptidase G is also a zinc metalloenzyme that releases C-terminal glutamate residues from γ-glutamyl peptides and folyl- γ -PGA.

Ntn hydrolases

The αββa fold of GGT is shared by several enzymes that are listed in table 1.1. These enzymes form a unique structural superfamily called 'N terminal nucleophile hydrolases' (Ntn hydrolases) [106]. The family is associated with five diagnostic features:

1. The core of the enzyme is characterised by a 'αββa' tetralamellar fold.

2. They are produced as inactive precursor whose maturational processing involves an autocatalytic chain cleavage that results in the formation of a new N terminus.

3. The newly formed N terminal residue, which can be Thr, Ser or Cys, functions as the catalytic nucleophile.

4. Their catalytic machinery comprises of a 'single residue nucleophile' system. The general base catalyst required for deprotonation of the side chain -OH or -SH function is provided by the a-amino group of the nucleophilic residue itself.

5. The enzymes catalyse amidolytic reactions.

Table 1.1: List of Ntn hydrolases[39, 85,109-120]

Hydrolase	Nucleophile State	oligomer
Bile Salt Hydrolases (BSH)	Cys	$(\alpha)_4$
Cephalosporin Acylases (CA)*	Ser/Thr	$(\alpha\beta)$, $(\alpha\beta)_2$
Gamma-Glutamyl Transferases (GGT)	Thr	$(\alpha\beta)$, $(\alpha\beta)_2$
Glucosamine-6-Phosphate Synthase	Cys	$(\alpha)_4$
Glutamine-PPRP-Amidotransferase (Grpp)	Cys	$(\alpha)_4$
Glycosylasparaginase (AGA)	Thr	$(\alpha\beta)$
Heat Shock Locus V (HSl V)	Thr	$2(\alpha\beta)_6$
L-aminopeptidase-D-Ala-esterase/amidase(Dmp)	Cys	$(\alpha\beta)_4$
Isoaspartyl Aminopeptidase (EcAIII)	Thr	$(\alpha\beta)_2$
MTH1020	N/D	$(\alpha)_4$
Ornithine Acetyltransferase (OAT)	Thr	$(\alpha\beta)_2$
Penicillin G Acylases (PGA)	Ser	$(\alpha\beta)$
Penicillin V Acylases (PVA)	Cys/Ser	$(\alpha)_4$
β subunit of 20 S Proteasome (PRO)	Thr	$4(\alpha\beta)_7$
U34 Peptidase	Cys	N/D

* Class V Cephalosporin acylase have single polypeptide chain and do not belong to Ntn hydrolase family. homodimer $(\alpha)_2$, homotetramer (α), heterodimer $(\alpha\beta)$, dimer of heterodimers $(\alpha\beta)$, $2(\alpha\beta)_6$ dimer of two hexamers stacked head to head, $4(\alpha\beta)_7$ tetramer of hepatmeric rings. N/D not determined.

Functions of Ntn Hydrolases

The antibacterial property of β-lactams has enabled their use in medicine as an antibiotic. However, the lactam produced by various molds have very low antibacterial property and therefore have little therapeutic use. The antibiotic strength of the lactam nucleus can be improved by increasing the complexity of the acyl side chain. The clinically used β-lactam antibiotics are semisynthetic as they involve chemical modification of the naturally produced β-lactam nucleus. Unsubstituted β-lactam nucleus is derived from the natural β-lactam compounds by removal of the acyl side chain. Both chemical and enzymatic methods are available for the deacylation process. Of these, the enzymatic route is employed extensively because of economic and environmental reasons[121]. Penicillin G, penicillin V and cephalosporin C are the commonly available natural β-lactams. They are deacylated by penicillin G acylases, penicillin V acylases and cephalosporin acylases respectively. Deacylation of penicillin G and penicillin V liberates 6-amino-cephalosporanic acid while 7-amino-cephalsopranic acid is generated from cephalosporin C. However, the physiological substrate and function of these deacylating enzymes is not known.

L-aminopeptidase D-Ala-esterase/amidase (DmpA) from *Ochrobactrum anthropi* is an aminopeptidase. It releases N-terminal Ala residues from peptide substrates[113]. The exopeptidase activity has remarkably relaxed stereospecificity as it can recognise both D and L isomers. Among various amniopeptidases, only DmpA is the known lack

stereospecificity. The proposed functions of aminopeptidases include protein maturation, degradation of N-terminus, regulation of hormone level and cell-cycle control.

Bile salt hydrolases (BSH) occur predominantly in enteric bacteria and catalyse deconjugation of bile salts[107]. However, the physiological significance of the enzyme either to the host or the symbiont is not known. The enzyme is believed to modulate the intestinal levels of bile salts. Deconjugation of intestinal bile salts appears to lower the plasma levels of cholesterol[122].

20S proteasome is a multi-subunit, barrel like complex, whose cavity forms the seat for most of the non-lysosomal protein degradation[124]. The degradative function is responsible for cytosolic clearance of misfolded proteins, short-lived regulatory proteins, and alien proteins coded by viruses and intracellular parasites. Similar function is served by heat shock locus V protein in *E. coli* [112].

Aged proteins may spontaneously form isoaspartyl derivatives by transferring the peptide backbone to the side chain through a β-peptide. The structural rearrangement can lead to protein dysfunction and therefore enzymes like isoaspartyl aminopeptidases have been evolved for their repair [114].

Glutamine PRPP amidotransferase is the first enzyme in the *de novo* biosynthesis of nucleotides [110]. The modulatory effect of adenine and

guanine enables the enzyme to serve as a regulator for the synthetic pathway.

Glycosylasparaginases, also known as aspartylglucosaminidase, serve to delink asparagine linked glycoproteins[111]. This reaction forms a part of lysosomal degration of proteins. Deficiency of GA results in lysosomal accumulation of glycoasparagines. The clinical situation, described as 'aspartylglucosaminuria', involves abnormalities in central nervous system, skeleton and connective tissues.

Post-translational maturation

Ntn hydrolases are produced as inactive 'proenzyme' (also called as precursor or zymogen) which is processed into the active form by an autocatalytic mechanism. The processing involves lysis of the peptide bond preceding the future catalytic nucleophile. The processing event can be classified into three patterns:

1. In enzymes like PVA [118], BSH [107] and PRO[123], the cleavage occur just 1-9 residues away from the primary N terminus and the resultant pro-peptide is lost. The mature enzyme is thus a single chain protein shorter than the proenzyme by a few residues.

2. The cleavage in hydrolases like GGT [124] and ASN [125] occurs towards the centre of the chain and leads to the formation of a

heterodimer. The two chains of the mature enzyme are thus products of single gene.

3. In hydrolases like penicillin G acylases[127] and cephalosporin acylases [128], cleavage occurs towards the centre of the chain but involves two closely set lyses, resulting in the excision of about 8-10 residues long pro-peptide fragment called the 'spacer'. In PGA, the pro-peptide functions as a folding catalyst for the precursor. In absence of the pro-peptide, PGA forms a molten globule like intermediate that has native-like secondary but little tertiary structure. The pro-peptide, which is structure-less when isolated in solution, forms an a-helix in the presence of the mature PGA. The three forms represented in figure 1.6.

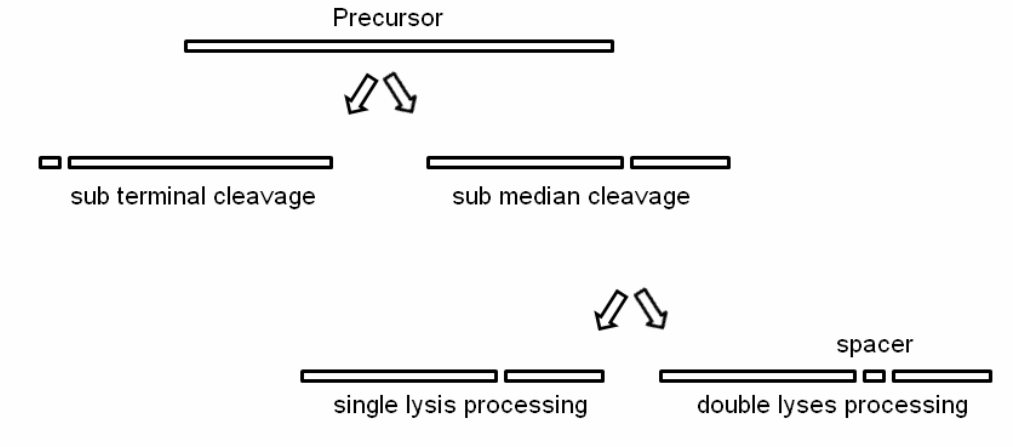


Figure 1.6: Diagrammatic representation of maturational types in Ntn hydrolases

In addition to Ntn hydrolases, autoproteolytic activation has been noticed in subtilin[128], pepsin[129], papain[130], calpain[131], thrombin proenzymes[132], mouse prohormone convertase[133] and hedgehog proteins[134].

The mechanism of processing has been analysed in *E. coli* GGT by developing maturationally blocked mutants[124] and from the crystal structure of the precursor[85]. The precursor differs from the mature enzyme mostly in the region corresponding to the active site. The C-terminus of the heavy chain, called as P-segment, occupies the future active site. During processing, the peptide bond preceding the catalytic Thr (Thr-391) is cleaved. Thereafter, the P segment rotates about Ile-378 —which occurs at the base of the segment— (ψ changes from -45° to127°) thus displacing the segment from the catalytic Thr by a distance of ~34Å. Finally, the space emptied by the retrieval of the P-segment is occupied by the lid loop. Also, the nascent active site contracts due to inward translocation of residues between 411 and 416, thus defining a complementary γ-glutamyl binding site. The maturational proteolysis is mediated by the side chain of the catalytic Thr as Ala substitution completely blocks the processing. Also, the processing of catalytic Thr-Cys mutant can be affected by Cys modifiers like DTNB, PCMB, iodoacetamide. These reagents do not affect the processing of wild type or Thr-Ser mutant enzymes. Furthermore, the processing is intramolecular as incubation of maturationally blocked Thr-391 to Ala mutant with the active enzyme does not induce

processing. Thus the integrity of the catalytic Thr appears to be invariant for processing. In contrast, the C-terminal residue, Gln in *E. coli*, appears to be unimportant as its substitution with Ala does not produce any maturational defects. Perhaps because of this the C-terminal residue does not show conservation (table 1.2).

Table 1.2. Amino acid sequence at the precursor cleavage site

GGT homologue	Scissile bond
Rat	DDGG * TAHL
Pig	DDAG * TAHL
Mouse	DDGG * TAHL
Yeast	N PHG * TAHF
E. coli	E SNQ * TTHY
Pseudomonas	EGS N * TTHY
B. subtilis	VEG Q * TTHF

* site of cleavage

In the mature enzyme, the catalytic function is imparted to the side chain –OH of the N-terminal Thr by a general base catalysed deprotonation. This function is provided by the free a-amino group of the catalytic residue itself. This moiety is not available in the precursor as it is engaged in the peptide bond. In the crystal structure of precursor, a water molecule appears to be in a competent position to mediate the process. Involvement of water has also been noticed in the processing of cephalosporin acylases[137] and β subunit of yeast 20S proteasome[138]. In contrast, the basic potential is provided by protein residues in AGA and β subunit of mammalian 20S proteasome. The peptide bond cleavage is believed to proceed by N-O acyl shift mechanism as there are evidences for the formation of a critical ester

bond during the process. These results have allowed the elucidation of the processing mechanism: a water molecule activates Thr-391 Oγ thereby enabling it to attack on Gln-390 C; the resultant tetrahedral transition state collapses thus transferring the amide link between Gln-390 and Thr-391 into an ester link between Gln-390 C and Thr-391 Oγ (N-O acyl shift); hydrolysis of the ester intermediate separates Gln-390 and Thr-391. The mechanism is illustrated in figure 1.7.

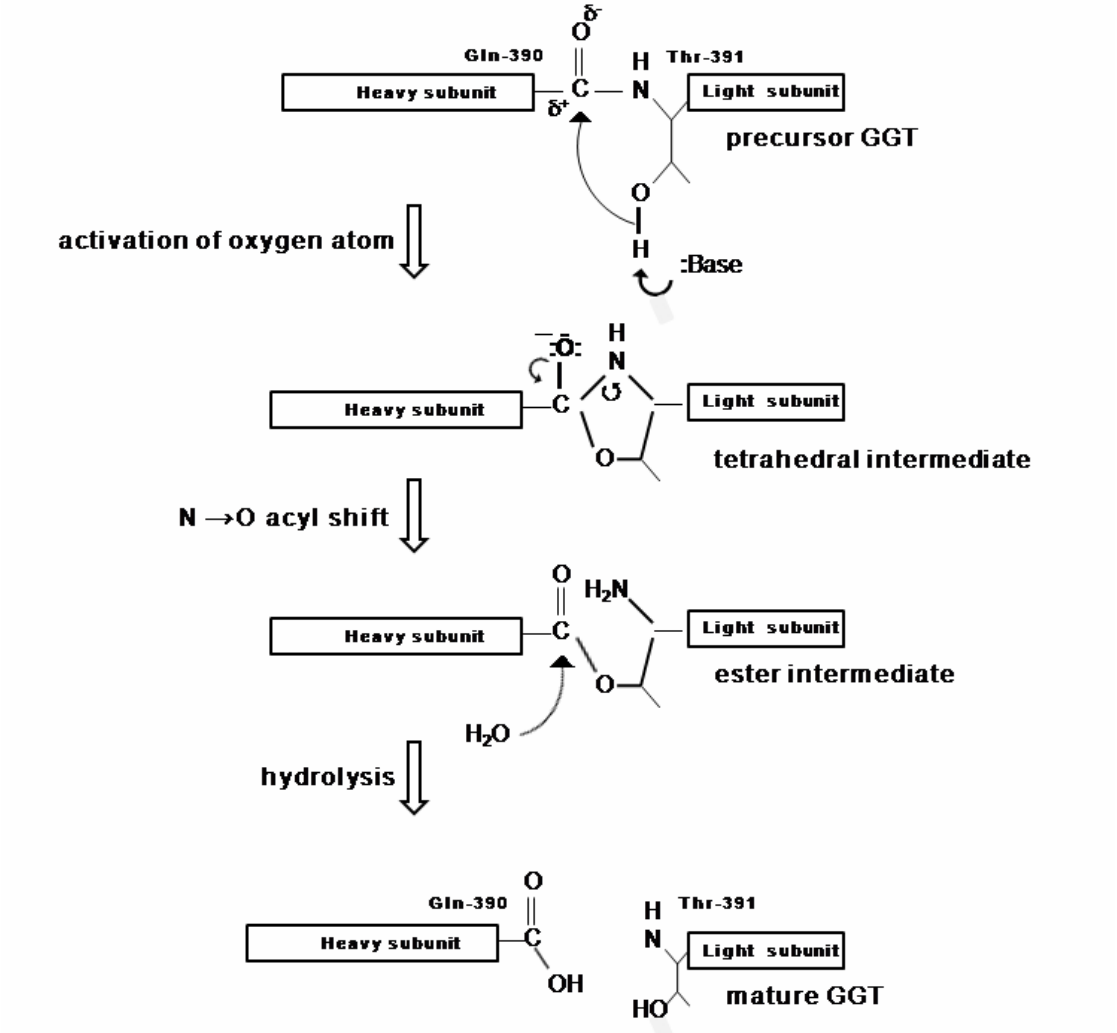


Figure 1.7: Maturation of *E. coli* GGT precursor by N-O acyl mechanism (from ref. 124).

N-terminal nucleophile

Amino acid residues in the functional surfaces of enzymes have unusual p*K*a values. The perturbation transforms hitherto unreactive residues into highly activated forms. Thus perturbations are at the root

of catalytic mechanism. Nucleophilic substitution is employed by many hydrolases that break amide or ester bonds. In serine proteases, where the catalytic apparatus is well documented, the nucelophilicity of the critical Ser is induced by the concerted action of a His and Asp residues. The seryl O_γ which occurs as –OH under ordinary conditions due to elevated p*K*a of ~14 is activated into powerful nucleophile ($-O^-$) by deprotonation. The critical His functions as a general base catalyst to abstract the proton provided the critical Asp is in a competent state to neutralise the resultant positive charge. Ntn hydrolases use the side chains of Ser, Thr or Cys as the catalytic nucleophile. The nucleophilic residue occurs at the N terminus that is produced as a consequence of maturational processing. The only base in the vicinity of the nucleophile is the free αNH_2 group of the nucleophilic residue itself. Therefore, the moiety is considered to be the general base catalyst[106]. Because of the catalytic importance, chemical modification of free αNH_2 can inactivate the enzyme[137]. The catalytic apparatus in Ntn hydrolases are said to form a 'single residue nucleophile system', because of the union of both nucleophile and general base catalyst in the same residue. The precursor is believed to protect the catalytic αNH_2 group from detrimental modifications till maturation[138].

In general, the type of nucleophile is invariant in each Ntn family. However, cephalosporin and penicillin V acylases appear to digress from the usual pattern. Cephalosporin acylases are classified into five groups of which the class V enzyme does not belong to Ntn family[127].

The nucleophilic residue is Ser in Class I, II and III acylases and Thr in class IV acylases. There are strong evidences to show that class IV enzymes are actually GGTs (to be discussed later). This explains the apparent digression as Thr is the characteristic nucleophile in GGTs. Sequence comparison indicates that Ser employing PVA appears to be distinct from cysteinyl PVAs as it show little sequence homology[139,140].

The basis for the choice of a particular nucleophile is not known. Nucleophile substitutions like Ser→Cys in PGA[141], Thr→Ser in Grpp[142], human lysosomal AGA[143], Ser→Thr/Cys in CA[127] and Cys→Ser in PVA[144] inactivates the enzyme. However, Thr→Ser substitution in PRO [145] does not inactive the enzyme but reduces the activity. These results appear to indicate a close connection between nucleophile type and the particular catalytic reaction.

Quaternary and domain structure

As described earlier, precursor processing can form either a monomer or a heterodimer. In heterodimers, both the chains contribute to the formation of the active site. Interestingly, Cys nucleophile occurs only in the monomeric type. The monomeric units associate either into dimer or tetramer and never occur in a free state. Heterodimers may or may not form an oligomer. GGT heterodimer is mostly free but the homologue from *H. pylori* appears to form a dimer in solution[146]. However, the crystal structure of *H. pylori* GGT does not show extensive interactions that would be necessary for dimerization[84]. The importance of

oligomeric association is not known; the effect of oligomer disruption on the catalytic function needs to be examined. Structurally, the active site of the individual units of an oligomer is independent.

Ntn enzymes may fold either intimately with or without organization into domains. CAs employing Ser nucleophile have a cup-shaped structure with two small 'knob' like domains on the 'rim'[108]. aββa fold occurs in the main domain. OAT forms two unequal sized domains, the larger of which contains the aββa fold[116]. Glutamine amidotransferases like Grpp and glucosamine-6-phosphate synthase have two domains[109, 110]. The domain that bears the typical Ntn fold is responsible for the hydrolytic function; transferase activity is located in the non-Ntn domain.

Topology

The aββa core consists of two antiparallel beat-sheets packed against each other and sandwiched between layers of a-helices on either side. Topological comparison indicates conservation of eight secondary structural elements in three layers of the four-layered core structure[147]. Of the two central sheets, one is essentially flat while the other one may be twisted in the distal end. The second sheet is more or less flat in AGA and PRO, slightly twisted in Grpp and highly twisted in PGA. Further variation can be seen in the packing angle between the two central beta sheets. The sheets are mostly parallel in AGA with a β-β packing angle of 5° while it is 35° in case of proteasome and Grpp.

The variation in the packing angle is associated with difference in the ratio of small and large hydrophobic residues in the topologically conserved β-strands in the interface. It is 50:20 in AGA and 50:50 in PRO and Grpp. Furthermore, the number of contacts between the sheets in PRO and Grpp is nearly double that in AGA. Also, the contact residues are smaller in AGA thereby shortening the β-β distance by about 1-1.6Å.

Comparative analysis reveals topological conservation of the catalytic machinery (catalytic nucleophile and residues interacting with it. However, the topological location of residues involved in the formation of the binding pocket and the oxyanion hole differ to some extent.

Evolution of Ntn Hydrolases

Ntn enzymes show little sequence homology despite extensive conservation of the structure. The paucity has hampered determination of evolutionary affinity between the members. The first glimpse into phylogenetic connection within the superfamily was recently shown between penicillin V acylases and bile salt hydrolases[107]. The two enzymes are homologous in their primary, tertiary and quaternary structures. Both the enzymes are homotetramers and use Cys as catalytic nucleophile. The variation occurs chiefly in the nature of the substrate binding pocket.

The topology of Dmp and OAT differs from the consensus pattern in the connectivity and direction of the secondary structure elements. The digressions were thought to constitute a separate subclass of Ntn-enzymes. However recent analysis indicates that the two enzymes along with molybdenum cofactor-binding domain may form a separate structure called the DOM-fold[148]. The apparent similarities with Ntn hydrolases in the core of the fold and the catalytic apparatus is proposed to be a consequence of convergent evolution.

Another interesting case is that of DCase[149]. This homotetrameric enzyme forms the typical tetralamellar fold. But the two central β plates are composed of parallel instead of antiparallel strands. Furthermore, the enzyme does not employ the N-terminal residue as the catalytic nucleophile. A Cys residue occurring in the interior of the chain serves as the nucleophile. As a-NH_2 is involved in the backbone chain, it is incapable of functioning as a general base. The function is instead provided by the side chain of Glu residue.

Relationship between GGT and CA

Cephalosporin acylases (CA) are Ntn hydrolases that are employed in pharmaceutical industry for the deacylation of naturally produced cephalosporin-C into 7-amino-cephalsoporanic acid (7-ACA). These enzymes are classified into five categories. Acylases of class I, II and III employ Ser as the catalytic nucleophile while Class IV acylases employ Thr as in GGTs. Class I-IV acylases are heterodimers while class V

acylase is a single chain protein. Recent homology modelling studies indicates that the class V enzyme forms α/β fold and not the αββα fold, diagnsotic of Ntn hydrolases[150]. Additionally, the primary structures of class IV CAs and GGTs are homologous by about 30%. Furthermore, some of the class IV CAs are active on both CA and GGT substrates. In such dually active enzymes, the affinity for GGT substrate tends be higher than for the CA substrate. Thus the speculation that class IV CAs are primarily GGTs with adventitious secondary CA activity[127]. In *E. coli* GGT, Asp-433 serves to bind the α-amino group of the substrate. Substitution of this residue with Asn abolishes the transferase activity and diminishes the hydrolytic function by 13% but imparts cephalosporin acylase activity which was hitherto absent[151].

Importance of Ntn family

The grouping of enzymes related in structure and catalytic apparatus has provided a framework for answering various biochemical events by analogy. Some Ntn-enzymes are involved in serious pathological conditions and thus the comprehensive knowledge of their catalytic mechanism facilitates are vital for drug development. For instance, identity of the general base catalyst in PRO was initially ambiguous despite the availability of the crystal structure. The critical group was identified after the description of the common family[152]. The general pattern of the family has also been helpful in the identification of the catalytic nucleophile in GGTs[152]. The speculative importance of N-terminal Thr of the light chain led to the development of a suitable

mechanism based inhibitor that finally trapped the catalytic nucleophile of *E. coli* GGT. The inhibitor was recently employed to demonstrate the conservation of the critical Thr in human GGT. Further improvement of this compound has resulted in a series of highly powerful inhibitors which assume significance in light the pathological role of GGTs[41, 49].

References

1. Rawlings, N. D., Morton, F. R. and Barrett, A. J. (2006) *Nucleic Acid Res* **34**, D270-D272
2. Elce, J.,S., and Broxmeyer, B.(1976) *Biochem J* **153**, 223-232
3. Kimura, K., Phan Tra, L-S., Uchida, I and Itoh ,Y. (2004) *Microbiology* **150**, 4115-4123
4. Minami, H., Suzuki, H., and Kumagai, H. (2003) *FEMS Microbiol Lett* **224**, 169-173
5. Zhang, H., Forman, H. J. and Choi, J. (2005) *Methods Enzymol* **401**, 468-479
6. Webb, E., C., Enzyme nomenclature 1992: recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes. San Diego: Academic
7. Tate, S. S. and Meister, A. (1985) *Methods Enzymol* 113, 400-419
8. Ikeda, Y., Fujii, J, Taniguchi, N., and Mesiter, A. (1995) *Proc Natl Acad Sci USA* **92**, 126-130
9. Smith,T.,K., and Mesiter,A. (1994) *FASEB J* **8**, 661-664

10. Ikeda, Y., Anderson, F. J., Anderson, M. E., *et al.* (1995) *J Biol Chem* **270**, 22223-22228
11. Suzuki,H., Kumagai, H., and Tochikura, T. (1986) *J Bacteriol* **168**, 1325-1331
12. Ogawa, Y., Hosoyama, H., Hamano, M., and Motai, H. (1991) *Agric Biol Chem* **55**, 2971- 2977
13. Boanca, G., Sand, A and Barycki, J. J. (2006) *J Biol Chem* **281**, 19029-19037
14. Ikeda, Y. and Taniguchi, N. (2005) *Methods Enzymol* **401**, 408-425
15. Storozhenko, S., Belles-Boix, E., Babiychuk, E., *et al.* (2002) *Plant Physiol.* **128**, 1109-1119
16. Mehdi, K., Thierie, J., and Penninckx, M., J., (2001) *Biochem J* **359**, 631-637
17. Suzuki,H., Kumagai, H., and Tochikura, T. (1986) *J Bacteriol* **168**, 1332-1335
18. Nakayama, R., Kumagai, H and Tochikura, T. (1984) *J Bacteriol* **160**, 1031-1036
19. Chevalier, C., Thiberge, J.,M., Ferrero, R., L., and Labigne, A. (1999) *Mol Microbiol* **31**, 1359-1372
20. Takahashi, H., and Watanabe, H. (2004) *FEMS Microbiol Lett* **234**, 27-35
21. Xu, K., and Strauch, M. A. (1996) *J Bacteriol* **178**, 4319-4322

22. Chikhi, N., Holic, N., Guellaen, G., et al. (1999) *Comp Biochem Physiol B Biochem Mol Biol* **122**, 367-389
23. Niida, S., Kawahara, M., Ishizuka, Y., et al. (2003) *J Biol Chem* **279**, 5752-5756
24. Nenesansky, E., Lott, J. A. (1985) *Clin Chem* **31**, 797-803
25. Keillor, J. W., Castonguay, R., and Lherbet, C. (2005) *Methods Enzymol* **401**, 449-467
26. Cook, N. D., Upperton, K. P., Challis, B. C., and Peters, T. J., (1987) *Biochimica et Biophysica Acta* **914**, 240-245
27. Taniguchi, N. and Ikeda, Y. (1998) *Adv Enzymol Relat Areas Mol Biol* **72**, 239-278
28. Thompson, G. A. and Mesiter, A. (1977) *J Biol Chem* **252**, 6792-6798
29. Palmer, T. Understanding Enzymes (1991) Ellis Horwood, England
30. Tate, S. S. and Meister, A. (1974) *J Biol Chem* **249**, 7593-7602
31. Karkowski, A. M., Bergamini, M. V. W and Orloski, M. (1976) *J Biol Chem* **251**, 4736-4743
32. Stromme, J. H. and Theodorsen, L. (1976) *Clin Chem* **22**, 417-421
33. Huseby, N. E. (1977) *Biochim Biophysics Acta* **483**, 46-56
34. Shaw, L. M., London, J. W., Fetterolf, D. and Garfinkel, D. (1997) *Clin Chem* **23**, 79-85

35. Shaw, L. M., London, J. W. and Tetersen, L. E (1978) *Clin Chem* **24**, 905-915
36. Tate,S.,S., and Meister, A. (1977) *Proc Natl Acad Sci USA* **74**, 931-935
37. Smith, T. K., Ikeda, Y., Fujii, J *et al.* (1995) *Proc Natl Acad Sci USA* **92**, 2360-2364
38. Keillor, J., W., Menard, A., Castonguay, R., *et al.* (2004) *J Phys Org Chem* **17**, 529-536
39. Okada, T., Suzuki, H., Wada, K., Kumagai, H., and Fukuyama, K. (2006) *Proc Natl Acad Sci USA* **103**, 6471-6476
40. Tate, S.,S., and Mesiter, A. (1978) *Proc. Natl.Acad.Sci. USA.* **75**, 4806-4809
41. Inoue,M., Hiratake, J., Suzuki,H., *et al.* (2000) *Biochemistry* **39**,7764-7771
42. Ménard, A., Castonguay, R., Lherbet, C. *et al* (2001) *Biochemistry* **40**, 12678-12685
43. Castonguay, R., Lherbet, C., and Keillor, J.W. (2003) *Biochemistry* **42**, 11505-11513
44. Szasz, G. (1969) *Clin Chem* **15**, 124-136
45. Shaw, L.M., Stromme, J. H., London, J. L., *et al.* (1983) *Clin Chim Acta* **135**, 315F-338F
46. Whitfield, J.B. (2001) *Critical Reviews in Clinical Laboratory Sciences* **38**, 263-355
47. Gardell, S. J. and Tate, S. S. (1980) *FEBS Lett* **122**, 171

48. Lherbet, C., and Keillor. (2004) *Org Biomol Chem* **2**, 238-245
49. Han, L., Hiratake, J., Tachi, N., *et al.* (2006) *Bioorg. Med. Chem.* **14**, 6043-6054
50. London, R. E. and Gabel, S.A. (2001), *Arch. Biochem. Biophys*. 385, 250-258
51. Lherbet, C. and Keillor. (2004) *Org Biomol Chem* **2**, 238-245
52. Tate, S. S. and Mesiter, A. (1974) *Proc Natl Acad Sci USA* **71**, 3329-3333
53. Thompson, G. A. and Meister, A. (1980) *J Biol Chem* **255**, 2109-2113
54. Gardel, S. J. and Tate, S. S. (1983) *J Biol Chem* **258**, 6198-6201
55. Lieberman, M.,W., Wiseman, A.,L., Shi, Z.,Z., *et al.* (1996) *Proc Natl Acad Sci USA* **93**, 7923-7926
56. Mesiter, A., and Anderson, M., E. (1983) *Ann Rev Biochem* **52**, 711-760
57. Meister, A. (1973) *Science* **6**, 33-39
58. Hinchman, C., A., Matsumoto, H., Simmons, T., W., and Ballatori, N. (1991) *J Biol Chem* **266**, 22179-22185
59. Leustek, T., Martin, M. N., Bick, J. A., Davies, J. P. (2000) *Ann Rev Plant Physiol Plant Mol Biol* **51**, 141-165
60. Kean, E. A. and Hare, E. R. (1980) *Phytochemistry* **19**, 194-203
61. Ishikawa, T. (1967) *Agric Biol Chem* **31** 490-493

62. Kasai, T, Ohmiya, A. and Sakamura, S. (1982) *Phytochemistry* **21** 1233-1239
63. Kawasaki, Y., Ogawa, T., Sasaoka, K. (1982) *Biochim Biophys Acta* 716, 194-200
64. Martin, M. N., and Slovin, J. P. (2000) *Plant Physiol* **122**, 1417-1426
65. Suzuki, H., Hashimoto, W. and Kumagai. H (1993) *J Bacteriol* **175**, 6038-6040
66. Fahey, R. C. Brown, W. C., Adams, W. B. and Worsham, M. B. (1978) *J Bacteriol* **133** 1126-1129
67. Sian, J., Dexter, D. T., Lees, A. J., Daniel, S., Jenner, P., Marsden, C. D., (1994) *Ann Neurol* **36**, 356-361
68. Singh, J. C., Singh, S., Singh, G., and Atal, C. K. (1986) *Biochem Pharmacol* **35,** 3753–3760
69. Lee, D.-H., Ha, M.-H., Kim, J,-H., Christiani, D. C., Gross, M. D., Steffes, M., Blomkoff, R., and Jacobs, D. R. (2003) *Diabetologia* **46**, 359-364
70. Hashimoto, Y., Futamura, A., Nakarai, H., and Nakahara, K. (2001) *Atherosclerosis* **158**, 465-470
71. Hanigan, M. H. (1995) *Carcinogenesis* **16**, 181-185
72. Pompella, A., De Tata, V., Paolicchi, A., and Zunino, F. (2006*) Biochem Pharmacol* **71**, 231-238
73. Benlloch, M., Ortega, A., Ferrer, P., *et al.* (2005) *J Biol Chem* **280**, 6950-6959

74. Ortega, A. L., Carretero, J., Obrador, E., *et al.* (2003) *J Biol Chem* **278**, 13888-13897
75. O'Brador, E., Carretero, J., Ortega, A., *et al.* (2002) *Hepatology* **35**, 74-81
76. Gounni, A., S., Borowski, K., S., Palacios, M., *et al.* (2001) *Mol Med* **7**, 344-354
77. Chevalier, C., Thiberge, J.,M., Ferrero, R., L., and Labigne, A. (1999) *Mol Microbiol* **31**, 1359-1372
78. Takahashi, H., Hirose, K., and Watanabe, H.(2004) *J Bacteriol* **186**,244-247
79. Whitfield, J. B., Pounder, R. E., Neale, G., *et al.* (1972) *Gut* **13**, 702-708
80. Scouller, K., Conigrave, K. M., Macaskill, P., *et al.* (2000) *Clin Chem* **46**, 1894-1902
81. Dunbar, J. A., Ogston, S. A., Ritchie, A., *et al.* (1985) *Br Med J* **290**, 827-30
82. Gjerde, H., Sakashaug, J., Morland, J. (1986) *Alcohol Clin Exp Res* **10**, 209-212
83. Niederau, C., Niederau, M., Stromeyer, G., *et al.* (1990) *Digestion* **45**, 115-20
84. Boanca, G., Sand, A., Okada, T., Suzuki, H., Kumagai, H., Fukuyama, K., and Barycki, J. J. (2007) *J Biol Chem* **282**, 534 – 541
85. Okada, T., Suzuki, H., Wada, K., Kumagai, H., and Fukuyama, K. (2007) *J Biol Chem* **282**, 2433-2439

86. Xu, K., and Strauch, M. A. (1996) *J Bacteriol* **178**, 4319-4322

87. Minami, H., Suzuki, H., and Kumagai, H. (2003) *FEMS Microbiol Lett* **224**, 169-173

88. Thorne, C. B. (1993) *Bacillus subtilis and Other Gram-Positive Bacteria*, 113-124 (ed: Sonenshein, A. L., Hock, J. A. and Losick, R.) American Society for Microbiology, Washington, DC

89. Phillips, Z. E. and Strauch, M. A. (2002) *Cell Mol Life Sci* **59**, 392-402

90. Candela, T. and Fouet, A. (2005) *Mol Microbiol* **57**,717-726

91. Candela T. and Fouet, A. (2006) *Mol Microbiol* **60**, 1091-1098

92. Anderson, M. E., Allison, D. R., and Meister, A. *Proc Natl Acad Sci USA* (1982) **79**, 1088-1091

93. Mayatepek, E., Okun, J. G., Meissner, T., *et al.* (2000) *J. Lipid Res.* (2004) 45, 900-904

94. Lieberman, M. W., Wiseman, A. L., Shi, Z. Z., *et al.* (1996) *Proc Natl Acad Sci USA* **93**, 7923-7926

95. Carter,B.,Z., Wiseman, A.,L., Orkiszewski, R. *et al.* (1997) *J Biol Chem* **272**, 12305-12310

96. Carter, B. Z., Shi, Z. Z., Barrios, R., Lieberman, M. W. (1998) *J Biol Chem* **273**, 28277- 28285

97. Heisterkamp, N., Meyts, E., R., Uribe, L., *et al.* (1991) *Proc Natl Acad Sci USA* **88**, 6303-6307

98. Yamaguchi, T., Takei, N., Araki, K. *et al.,* (2000) *J Biochem* **128**, 101-106

99. Minami, H., Suzuki, H., and Kumagai, H. (2004) *J Biol Chem* (2004) **186**, 1213-1214

100. Kimura, K., and Itoh, Y. (2003) *App Env Microbiol* **69**, 2491-2497

101. Volcani, B. E. and Margalith, P. (1957) *J Bacteriol* **74**, 646-655

102. Tanaka, T.O., Hirata, T., Futamura, K. et al. (1993) *Biosci Biotech Biochem* **57**, 2148-2153

103. Sherwood, R., Melton, R. G. Alwan, S. M. and Hughes, P. (1985) *Eur J Biochem* **148**, 447-453

104. Yao, R., Schneider, E., Ryan, T. J. and Galvin, J. (1996) *Proc Natl Acad Sci USA* **93**, 10134-1-138

105. Galivan, J., Ryan, T. J., Chave, K., Rhee, M., Yao, R., and Yin, D. (2000) *Pharmacol Ther* **85**, 207-215

106. Brannigan, J. A., Dodson, G., Duggleby, H. J., *et al.* (1995) *Nature* **378**, 416-419

107. Suresh Kumar, R., Brannigan,J.,A., Prabhune,A.,A., *et al.* (2006) *J Biol Chem* **281**, 32516-32525

108. Kim, Y., Yoon, K-H., Khang, Y., *et al.* (2000) *Structure* **8**, 1059-1068

109. Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denistot, M-A., Badet, B., Plikarpov, I., Littlechild, J. A., and Teplyakov, A. (1996). *Structure* 4, 801-810

110. Kim, J., H., Krahn, J., M., Tomchick, D., R., *et al.*(1996) *J Biol Chem* **271**, 15549-15557

111. Guo, H.,C., Xu, Q., Buckley, D., and Guan, C. (1998) *J. Biol. Chem.* **27**, 20205-20212

112. Bochtler, M., Ditzel, L., Groll,M., and Huber, R. (1997) *Proc Natl Acad Sci USA* **94**, 6070-6074

113. Bompard-Gilles, C., Villeret, V., Davies, G. J., *et al.* (2000) *Structure* **8**, 153-162

114. Michalska, K., Brzesinski, K., and Jaskolski, M. (2005) *J. Biol. Chem.* **280**, 28484-28491

115. Saridakis,V., Christendat, D., Thygesen, A., *et al.* (2002) *Proteins* **48**,141-143

116. Elkinds, J., Kershaw, N., J., and Schofield C., J. (2005) *Biochem J* **385**, 656-573

117. Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, G. and Moody, P. C. E. (1995) *Nature* **373**, 264-268

118. Suresh, C.,G., Pundle, A.,V., SivaRaman,H., et al. (1999) *Nat Struct Mol Biol* **6**, 414-416

119. Löwe, J., Stock, D., Jap, B., Zwicki, P., Baumeister, W. and Huber, R. (1995) *Science* **268**, 533-539

120. Pei, J. and Grishin, N. V. (2003) *Protein Sci* **12** 1131-1135

121. Shewale, J. G. and SivaRaman, H. (1989) *Process Biochem* 24, 146-154

122. De Smet, T., van Hoorde, L., De Saeyer, N., vande Woewtyne, M and Verstraete, W (1994) *Microb Ecol Health Dis* **7**, 315-329

123. Chen, P. and Hochstrasser, M. (1996) Cell **86**, 961-972

124. Suzuki, H. and Kumagai, H. (2002) *J Biol Chem* **277**, 43536-43543

125. Fisher, K. J., Klein, M., Park, H., Vettese, M. B. and Aronson, N. N. (1993) *FEBS Lett* **323**, 271-275

126. Choi, K. S., Kim, J. A. and Kang, H. S. (1992) *J Bacteriol* **174** 6270-6276

127. Li, Y., Chen, J., Jiang, W. *et al* (1999) *Eur J Biochem* **262**, 713-719

128. Gallagher, T., Gilliland, G., Wang, L. And Bryan, P. (1995) *Structure* **3**, 907-914

129. Bustin, M. And Conway-Jacobs, A. (1971) *J Biol Chem* **246**, 615-620

130. Vernet, T., Khouri, H. E., Laflamme, P., *et al*. (1991) *J Biol Chem* **266**, 21451- 21457

131. Zimmerman, U. J. And Schlaepfer, W. W. (1991) *Biochim Biophys Acta* **1078**, 192-198

132. Petrovan R. J., Grovers-Riemslag, J. W. P., Nowak, G., et al. (1998) *Biochemistry* **37**, 1185-91

133. Goodman, L. J. And Gorman, C. M. (1994) *Biochem Biophys Res Commun* **201**, 795-804

134. Hall, T. M., Porter, J. A., Young, K. E., et al. (1997) *Cell* **91**, 85-97

135. Kim, Y., Kim, S., Earnest, T. N. and Hol, W. G. (2002) *J Biol Chem* **277**, 2823-2829

136. Ditzel, L., Huber, R., Mann, K., et a. (!998) *J Mol Biol* **279**, 1187-1191

137. Lee, Y.S., Kim, H. W. and Park, S. S. (2000) *J Biol Chem.* **275**, 39200-39206

138. Arendt, C. S. and Hochstrasser, M. (1999) *EMBO J* **18**, 3575-3583

139. Zhang, D., Mayuko, K., Hiroyuki, I., *et al.* (2007) *J Biotechnol* **128**, 788-800

140. Kumar, A., Prabhune, A., Suresh, C. G. and Pundle, A. (2008) *Process Biochem* (in Press)

141. Choi, K. S., Kim, J. A. and Kang, H. S. (1992) *J Bacteriol* **174**, 6270-6276

142. Li, Songtao, L., Smith, J. L. and Zalkin. H (1999) *J Bacteriol* **181**, 1403-1408

143. Fisher, K. J., Klein, M., Park, H., *et al.* (1993) *FEBS Lett* **323**, 271-275

144. Chandra, P. M., Brannigan, J, A. Prabhune, A., *et al.* (2005) *Acta Crystallogr Sect F* **62** 124-127

145. Seemuller, E., Lupas, A. and Baumeister, W. (1996) *Nature* **382**, 468-470

146. Boanca, G., Sand, A., and Barycki, J.J. (2006) *J. Biol. Chem.* **258**, 6193-6197

147. Oinonen, C., and Rouvinen, J. (2000) *Protein Sci.* **9**, 2329-2337

148. Cheng, H., and Grishin, N., V. (2005) *Protein Sci.* **14**, 1902-1910

149. Nakai, T., Hasegawa, T., Yamashita, E., *et al.* (2000) *Structure* **8**, 729-739

150. Yau, M., Wang, J., Tsang, P.,W.,K., and Fong, W-P. (2006) *FEBS Lett* **580**, 1465-1471

151. Suzuki,H., Miwa,C.,Ishihara,S., and Kumagai,H. (2004) *Appl Enviorn Microbiol* **70**, 6324-6328

152. Bochtler, M., Ditzel, L., Groll, M. et al. (1999) *Annu Rev Biophys Biomol Struct* **28**, 295-317

EXPERIMENTAL PROCEDURES

Chapter 2

Materials:

B. subtilis var. *spinizeni* strain used for purification of wild type was procured from National Collection of Industrial Microorganisms (NCIM), National Chemical Laboratory, Pune. The genomic DNA for cloning experiment was from *B. subtilis* var. *subtilis.*

γ-Glutamyl-(3-carboxyl)-4-nitroaniline ammonium salt and glycylglycine were from Fluka (Switzerland) Sigma Chemicals (USA) respectively

Crystallization screens were from either Hampton Research (USA) or Molecular Dimensions (UK). Crystallographic data was processed with HKL 2000. All other programs employed were from CCP4 program suite 6.1 (Widows version).

pET vectors and *E. coli* host strains XL-Blue and BL21 were from Novagen (USA) and Stratagene (USA) respectively. All reagents used in genetic engineering were from Promega (USA) and the manufacturer's instructions were followed for their use.

Chromatography matrices were either from Amersham Biosciences (UK) or Sigma Chemicals (USA). Gel filtration markers were from Sigma Chemicals (USA).

All other reagents were of analytical grade and were prepared in milli-Q water.

Structures used for comparison were downloaded from Protein Data Bank (www.rcbs.org). The accession codes are: *E. coli* GGT (2dbu), *H. pylori* GGT (2nqo), *Bacillus halodurans* cephalosporin acylase (2nlz), GGT-like protein from *Thermoplasma acidophilum* (2i3o). Amino acid sequences for comparison were obtained from Genbank (www.ncbi.nlm.nih.gov.). Accession numbers are: GGL (AF077765), GGT-rel (M64099), GGT-rb (AF244973), CapD (D14037).

Section 2.1: Molecular, biochemical and kinetic methods

Purification of GGT by fermentation of *B. subtilis*

10ml of overnight culture of *B. subtilis* was inoculated into 200ml of LB broth contained in 1L Erlenmeyer flask and incubated at 28°C with orbital shaking for a period of 3 days. Thereafter the fermented broth was centrifuged at 6000 rpm to remove cell mass and the clarified broth was concentrated to 1/10th original volume by ultra-filtration (Amicon) over 10kDa cut-off membrane. The concentrate was dialysed against 20mM phosphate buffer pH 7.0 (loading buffer-A) and loaded onto DEAE sepharose column pre-equilibrated with loading buffer-A. The unbound material was washed by passing 5 bed volumes of the loading buffer-A and the bound protein eluted by passing a gradient of NaCl in the range 0-0.5M. The fractions showing GGT activity were pooled and treated with solid ammonium sulphate to a final concentration of 15% and loaded onto phenyl sepharose column pre-equilibriated with 50mM phosphate buffer pH 7.0 and 15% ammonium sulphate (loading buffer-B). The unbound

material was removed by washing with 5 bed volumes of the loading buffer-B. The bound proteins were then eluted by passing 10% ammonium sulphate in 50mM phosphate buffer pH 7.0. The active fractions were pooled and concentrated before loading onto sephadex G-200 column pre-equlibriated with 50mM phosphate buffer pH 7.0 and eluted with the same buffer. The purity of the preparation was analysed by SDS-PAGE.

Cloning and expression of *B. subtilis* GGT

The GGT ORF was amplified from the genomic DNA of *B. subtilis* strain with the primers 5'ACG CGG CCA TGG TAA AGC CGC CCA AAA GCT3' (forward) and 3'GTT TTT CTC GAG TTT ACG TTT TAA ATT GCC GAT 5'(reverse). The primers append 5' and 3' termini of the amplicon with flanks bearing recognition sites for the restriction enzymes *Nco*I and *Xho*I, respectively. The amplicon was produced by 25 cycles of polymerase chain reaction with each thermal cycle comprising 30s denaturation at 94°C, 30s annealing at 50°C and 90s extension at 72°C. Taq polymerase was used to catalyse polymerization. Both amplicon and pET26b (Novagen) plasmid DNA were digested with *Nco*I and *Xho*I and then ligated to generate pET26-BsGGT construct. The construct substitutes the N-terminal signal peptide with *E. coli* specific equivalent and appends a C-terminal hexa-histidine tag.

The construct was transformed into *E. coli* expression strain BL21. For protein production, the cells were grown in LB medium supplemented

with 30μg/ml kanamycin. When the OD_{600} reached 0.5, the culture was induced by adding isopropyl-β-D-thio galactopyranoside (IPTG) to a final concentration of 1mM. Thereafter, the incubation was continued at 16°C overnight. The cells were pelleted by centrifugation at 6000 rpm and stored at -20°C till further use.

Purification of recombinant GGT

IPTG induced cell pellet was suspended in 20mM sodium phosphate buffer pH 7.0 and 500mM NaCl (buffer A) containing 20mM imidazole. The cell suspension was homogenised by two cycles of extrusion through French press at 1500 psi. The homogenate was clarified by centrifugation at 15000 rpm for 1 hr and filtered through 0.45μ membrane. The filtrate was loaded onto 1 mL HisTrap column containing precharged high performance nickel sepharose, pre-equilibrated with buffer A and washed with the same buffer till OD_{280} of the effluent reached the baseline. The bound protein was eluted with a linear gradient of 20 - 500mM imidazole contained in the binding buffer. The fractions corresponding to the single major peak were pooled and loaded onto Hi-Load Superdex S200 column pre-equilibrated with 100mM tris HCl pH 7.5 and 100mM NaCl. Pure GGT eluted as a single prominent peak and was confirmed by activity test and SDS-PAGE analysis. The protein was concentrated to 10 mg/ml over a 10 kDa cutoff membrane (Pall) and preserved at -20°C.

Enzyme assay

The enzyme was assayed by the method of Suzuki[1] with some modifications. Briefly, ~0.4μg of enzyme was incubated at 30°C for

5 mins with ammonium salt of γ-glutamyl-(3-carboxyl)-4-nitroaniline in 100mM buffer in a final volume of 100μl. Buffers used were tricine.HCl for pH 7.0 and 9.0 and CAPS for pH 11.0 The reaction was arrested with 900μl of 2M acetic acid and the absorbance measured at 410nm. Suitable blanks were used to correct for absorbance due to substrate. One unit of enzyme was defined as μmoles of 3-carboxyl-4-nitroaniline (ε_{410} = 8800M^{-1}) formed in one minute. The assay was linear with respect to time of incubation.

Steady state kinetics

Michaelis-Menten plots were prepared by varying γ-glutamyl-(3-carboxyl)-4-nitroaniline concentration between 1-128mM. 50mM glycylglycine was used for measuring transferase activity after adjusting the pH to assay condition. The kinetics constants (K_M and V_{max}) were determined by fitting the raw data to Michaelis-Menten equation by non-linear regression method using Origin 6.1. Standard deviation of all the values was less that 10%. For measurements made at pH 11.0, the velocities were fitted to Hill's equation of the form:

$$v = V_{max} \frac{[S]^h}{k_{0.5} + [S]^h}$$

where v is the initial velocity, V_{max} is maximal velocity, $k_{0.5}$ is the substrate concentration for half maximal activity and h is the measure of cooperatively between n interacting sites. The data was fitted to the equation by Levenberg-Marquardt's non-linear regression method using Origin 6.1.

Thermodynamic anlaysis

The value of k_{cat} for the respective reaction was determined at 20, 30, 40 and 50°C. Arrhenius plot was used to determe of activation energy. The data were fitted by linear regression to the equation:

$$\ln k_{cat} = -Ea/RT + \ln A$$

where Ea is the activation energy, R is the gas constant, T is the absolute temperature and *A* is the frequency of collision. The slope of this plot is equal to – Ea/R.

Eyring's plot was used for the determination of enthalpy and entropy of activation. The data were fitted by linear regression to the equation:

$$\ln \frac{kcat}{T} = \ln \frac{k_B}{h} + \frac{\Delta S^{\ddagger}}{R} - \frac{\Delta H^{\ddagger}}{RT}$$

where *h* is Planck's constant, k_B is Boltzmann's constant, R is gas constant and T is absolute temperature. From the plot, $\Delta H^{\ddagger}$ was calculated from the slope (=$-\Delta H^{\ddagger}/R$) and $\Delta S^{\ddagger}$ from the y-axis intercept.

Inhibition studies

Effect of cholate and glycocholate was analysed by incubating their sodium salts with 25mM substrate in either 100mM CAPS buffer pH 11.0 or 100mM tricine buffer 9.0 under standard assay conditions. 25mM glycylglycine was used for measurement of transferase activity. The reaction velocities were measured as before.

CD measurement

Circular dichroism spectra were collected on a Jasco J-715 spectropolarimeter. Near and far UV CD spectra (200-250nm) were collected with 0.1 and 1 mg/ml of the enzyme placed in a path length of 0.1 and 1cm respectively. The final spectrum was an average of 5 accumulations measured at a rate of 100nm/min at 1nm resolution.

Mass Analysis

The mass spectrum was recorded by using a time-of-flight delayed extraction MALDI mass spectrometer (Applied Biosciences). The 5 μL sample was mixed with 35 μl matrix solution (15 mg/ml sinapic acid in 30 % acetonitrile). About 10 μL of the mixture was applied to a stainless steel sample holder and introduced into the mass spectrometer after drying. The spectrum was obtained in the reflectron mode by scanning 5 laser shots with an ion source voltage 1of 19 kV, ion source 2 of 16.27 kv, 100 ns delay and the low mass gate at m/z 600.

Gel Filtration Chromatography

Native molecular weight was estimated by gel filtration chromatography. A column of sephacryl S 200 (1cm x 100cm) was equilibrated with 150mM NaCl in 50mM Na-K phosphate buffer pH 7.0. The elution volume of blue dextran was considered as the void volume of the packed column. The column was calibrated by passing

cytochrome C (12000 Da), carbonic anhydrase (24000 Da), bovine serum albumin (66000 Da), alcohol dehydrogenase (15000 Da) and beta-amylase (200000 Da). A calibration curve was constructed by plotting logarithm of the molecular weight of the individual marker against the ratio of the void (v_o) and respective elution volumes (v_e).

For determining the molecular weight of *B. subtilis* GGT, a sample was passed through the calibrated column and its elution volume was noted. This was used to calculate the molecular weight from the calibration curve. The molecular weight was also determined at pH 11.0. For this, the above column was thoroughly equilibrated with 100mM CAPS buffer pH 11.0 and 150mM NaCl.

Reversibility studies

A sample incubated in 50mM CAPS buffer pH 11.0 and 150mM NaCl for atleast 24 hrs was used to reversibility studies either by 10 fold dilution or buffer change.

Section 2.2: Determination of crystal structure of *B. subtilis* GGT by x-ray crystallographic method

Crystallization

E. coli GGT was first crystallised in 1986 by a very simple procedure[2]. It involved addition of solid ammonium sulphate to the enzyme solution (present in 20mM tris.HCl pH 8.0) till the development of turbidity and thereafter incubation in a refrigerator. Needle like crystals appeared after one day and reached their

maximum size within a week. Subsequently (in 1995) crystals belonging to the primitive orthorhombic space group were obtained from citrate buffer (pH 5.3)[3]. The pH corresponds to the pI of the enzyme. However, the structure could not be solved because of the failure to obtain isomophous heavy atom derivative for MIR experiment. Recently, *E. coli* and *H. pylori* GGT crystals belonging to $P2_12_12_1$ space groups were used to determine the structure.

Clear Strategy Screen I and II were used for the initial screening[4]. The conditions available in the screen are shown in figure 2.1. The enzyme solution (10 mg/ml in 100mM tris HCl pH 7.5 and 100mM NaCl) was centrifuged at 10,000 rpm for 10 mins at 4°C and the resulting supernatant was used for crystallization trials. The screening trials were done by sitting drop vapour diffusion method using 96 well Greiner plates. In these plates, a typical well comprises of a reservoir and a platform for the protein drop (Fig 2.1). In a typical experiment, 80 μl of the mother liquor was placed in the reservoir from which 0.2 μl was withdrawn and mixed with 0.2μl of the enzyme solution placed on the platform. Mosquito robotic liquid handling system (Molecular Dimensions, UK) was used for setting up drops. The plates were sealed with plastic films and incubated at 293K.

Plate-like crystals were obtained in conditions containing 0.2M KSCN. Best crystals were obtained with 0.2M KSCN, 20% polyethylene glycol 20 KDa and 0.1M citrate buffer pH 5.3

(isoelectric point of the ezyme). They diffracted up to 7 Å. The crystals grew as radiating clusters or as irregular stacks. Further improvements were done by varying the pH and precipitant size. Good quality crystals were obtained from 0.1M sodium HEPES pH 7.5, 25% PEG 2000K MME and 0.2M KSCN. For data collection, single crystals were treated with 10% hexanetriol for cryoprotection, mounted in nylon loops and flash cooled in liquid nitrogen.

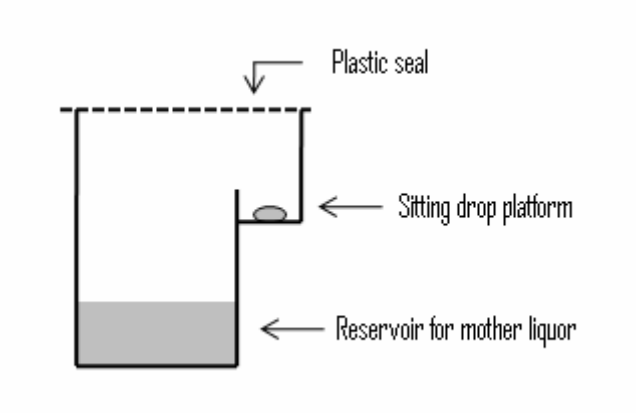


Figure 2.1: Diagrammatic representation of the well format employed for crystallization trails by sitting drop vapour method.

1 0.3 *M* Na acetate 25% PEG 2K MME	2 0.2 *M* Li_2SO_4 25% PEG 2K MME	3 0.2 *M* $MgCl_2$ 25% PEG 2K MME	4 0.2 *M* KBr 25% PEG 2K MME	5 0.2 *M* KSCN 25% PEG 2K MME	6 0.8 *M* Na formate 25% PEG 2K MME
7 0.3 *M* Na acetate 15% PEG 4K	8 0.2 *M* Li_2SO_4 15% PEG 4K	9 0.2 *M* $MgCl_2$ 15% PEG 4K	10 0.2 *M* KBr 15% PEG 4K	11 0.2 *M* KSCN 15% PEG 4K	12 0.8 *M* Na formate 15% PEG 4K
13 0.3 *M* Na acetate 10% PEG 8K+ 10% PEG 1K	14 0.2 *M* Li_2SO_4 10% PEG 8K+ 10% PEG 1K	15 0.2 *M* $MgCl_2$ 10% PEG 8K+ 10% PEG 1K	16 0.2 *M* KBr 10% PEG 8K+ 10% PEG 1K	17 0.2 *M* KSCN 10% PEG 8K+ 10% PEG 1K	18 0.8 *M* Na formate 10% PEG 8K+ 10% PEG 1K
19 0.3 *M* Na acetate 8% PEG 20K+ 8% PEG550 MME	20 0.2 *M* Li_2SO_4 8% PEG 20K+ 8% PEG550 MME	21 0.2 *M* $MgCl_2$ 8% PEG 20K+ 8% PEG550 MME	22 0.2 *M* KBr 8% PEG 20K+ 8% PEG550 MME	23 0.2 *M* KSCN 8% PEG 20K+ 8% PEG550 MME	24 0.8 *M* Na formate 8% PEG 20K+ 8% PEG550 MME

(A)

A

1 1.5 *M* $(NH_4)_2SO_4$	2 0.8 *M* Li_2SO_4	3 2 *M* Na formate	4 0.5 *M* KH_2PO_4
7 2.7 *M* $(NH_4)_2SO_4$	8 1.8 *M* Li_2SO_4	9 4 *M* Na formate	10 1.0 *M* KH_2PO_4

E

5 25% PEG 2KMME 0.2 *M* Ca acetate	6 15% PEG 4K 0.2 *M* Ca acetate
11 10% PEG 8K +10% PEG 1K 0.2 *M* Ca acetate	12 8% PEG 20K +8% PEG 550MME 0.2 *M* Ca acetate

13 40% *v/v* MPD	14 40% *v/v* Butanediol	15 20% PEG 4K 5 m*M* $CdCl_2$	16 20% PEG 550 MME 0.15 *M* KSCN	17 20% PEG 600 0.15 *M* KSCN	18 20% PEG 1.5K 0.15 *M* KSCN
19 35% *v/v* Isopropanol	20 30% *v/v* Jeffamine 600M	21 20% PEG 4K 5 m*M* $NiCl_2$	22 18% PEG 3550 0.15 *M* KSCN	23 18% PEG 5KMME 0.15 *M* KSCN	24 15% PEG 6K 0.15 *M* KSCN

B **C** **D**

(B)

Figure 2.1: (A) Clear strategy screen I, (B) Clear Strategy Screen II. PEG concentrations are in w/v. 1.5K, 2K, 4K, 5K, 6K, 8K and 20K correspond to molecular weight of PEG in thousands of Daltons.

Data collection

X-ray diffraction experiment requires a source of x-rays and a means for detecting the diffracted radiation. Rotating anode generator and synchrotrons are the two common sources of x-rays. Rotating anode generator is commonly used for in house studies. A high voltage electric field is applied to a filament to eject a stream of electrons. The electron beam hits a metal anode (copper or molybdenum) causing excitation of electrons to higher orbital. The radiation emitted depends on the emitting element. The excess heat generated is removed by means of circulating water. Dissipation of heat is augmented by the rotation of the anode. Unwanted radiation is filtered with a nickel-foil filter or a single crystal graphite monochromator. A set of aligned mirrors then focus the radiation. The most commonly used radiation from this source is the CuKa (λ = 1.54 Å) radiation. Synchrotrons produce powerful x-rays that are useful when working with crystals that are very small in size, have extremely large unit cell volume and are sensitive to radiation. Synchrotron light starts with an electron gun. A heated element or cathode produces free electrons which are pulled through a hole in the end of the gun by a powerful electric filed. This produces an electron stream which is then fed into a linear accelerator or linac. Here the electrons assume almost the speed of light. The linac feeds electrons into the booster ring which uses magnetic fields to force the electrons to travel in a circle. The booster ring ramps up the energy in the electron stream to between 1.5 and 2.9 giga electron volts (GeV). The booster ring feeds electron into the storage ring which is a polygonal tube

maintained under vacuum. Synchrotron radiation is emitted when a fast electron interacts with a magnetic field. A magnetic field in an area an electron is traveling in will cause the electron to change direction by exerting a force on it perpendicular to the direction of electron's motion. As a result, the electron will be accelerated, causing it to radiate electromagnetic energy. If the electrons and the magnetic field are energetic enough, the emitted radiation can be in the form of X-rays.

Presently, the commonly used detectors are image plate and charge-coupled device (CCD) detectors. An image plate detector[5] comprises of four components: a phosphor screen, an image reader, an image processor and an image writer. The screen is a plastic plate coated with photosensitive phosphor crystals (often europium doped barium halide) which absorbs the incident x-ray photons and stores them in their F centres. The stored image is read by a scanning with helium-neon laser beam of wavelength λ = 632.8nm. The digitised image of the diffraction pattern produced by the image processor can be directly analysed by a computer. Image plate can be regenerated for further use by erasing the data with intense white light.

CCDs[6] are commonly used in synchrotrons. A light emitting phosphor is bonded to the large end of a tapered fibre optic device and a CCD chip is bonded to the small end. A typical CCD is a 2.5 cm square array of 25μm pixels. Photons incident on CCD generate electron hole (e-h) pairs by photoelectric effect. If a small voltage is

maintained on the pixel, the charge separates to yield an electric charge on the pixel that is proportional to the energy of the incident photon. The CCD array consists of two registers: a vertical (parallel) register and a horizontal (serial) register. The vertical register is a two dimensional array of photosensitive pixels, each of which contains a measurable charge. After exposure of the vertical registers to light, the charge distribution is shifted out by repeatedly shifting rows of vertical register onto the horizontal register, which can be read out by shifting its charges serially into a pre-amplifier, an amplifier an analog to digital converter.

In the present study, a Rigaku RU2000 rotating anode x-ray generator mounted with MAR 300 scanner and Oxford Cryostream was used for in house testing of the crystal. A single crystal from the foresaid condition diffracted up to 2.5 Å. This crystal was used to collect x-ray diffraction data at ID 14-2 beamline at European Synchrotron Radiation Facility, France. The beamline was equipped with a mini-diffractometer comprising a single phi spindle goniometer with motorised z-translation and x/y sample alignment translations, ADSC Q4 CCD detector, an Oxford Cryostream for 100K data collection and an automated sample changer. Data were collected using a crystal cooled to 100K employing a solution of 10% hexanetriol as cryoprotectant. The crystal diffracted up to a resolution of 1.84 Å.

Processing of x-ray diffraction data

Data processing involves four major steps: (1) indexing, (2) integration, (3) scaling and (4) merging. In the present study, HKL 2000 was used for data processing[7]. This package comprises of DENZO, XDISPLAYF and SCALEPACK. Steps 1 and 2 are done with DENZO and XDISPLAYF and steps 3 and 4 with SCALEPACK.

The first step in the processing involves the determination of the crystal system and the unit cell dimensions. Then, each spot on the image is assigned an *hkl* index. In addition, at this stage, the orientation of the crystal in the beam is determined. The autoindexing routine deduces the unit cell parameters and the crystal orientation parameters from a single oscillation image. This is usually carried out with the first image as in most cases is the best image. The parameters that specify the orientation of the crystal in the x-ray beam are the vertical axis, the spindle axis and the crystal rot x, rot y and rot z values. Distortion index and unit cell parameters consistent with all possible Bravais lattices are listed by the program. The lattice with the highest symmetry that fits the data with minimal distortion is chosen. The autoindexing program also gives the crystal orientation parameters.

After a successful autoindexing step, the following parameters are refined in the order : (1) crystal rot x, rot y , rot z, (2) Y beam and X beam, (3)unit cell, (4)crossfire x, y, xy, (5) cassette rot x rot y, (6) radial offset and angular offset, (7) distance. After refinement,

DENZO gives χ^2 values for the X and Y positions of the predicted spots. The χ^2 value represents the average ratio of the squared error in the fitting and the expected error. A good refinement will have χ^2 values near 1.0. Finally, the data is integrated by measuring the difference intensity between the Bragg spot and the local background. In the present case, autoindexing routine identified the space group as $\mathbf{P2_12_12_1}$ and the unit cell dimensions after refinement were found to be: **a = 72.26, b = 108.77, c = 161.35.**

The intensities on all the images are not in same scale due to fluctuations in beam intensity, diffracted crystal volume, crystal decay, absorption, polarization etc. Therefore, a scale factor must be allocated so that the intensities of all the images in the data set can be related. Also, multiple measurements of same reflection (*hkl*) and symmetry related reflections are averaged. The goodness of merging is evaluated by symmetry or merging *R* factor which is defined as:

$$R_{sym} = \frac{\Sigma|I - < I >|}{\Sigma I}$$

χ^2 is another index to assess the quality of the scaled and merged output. This uses σ which is standard deviation of the error in integration of the reflections. At the end of scaling and merging, the goodness of fit or χ^2 is evaluated is verify if σ have been estimated correctly. The error in I value should approximately be equal to I - < I

> if the < I > are nearly correct. So, if σ^2 is close to this error, then χ^2 will be near 1.0. χ^2 is defined as:

$$\chi^2 = \frac{\Sigma(I-<I>)^2}{\sigma^2}$$

Determination of structure factor amplitudes

The program TRUNCATE was used to obtain structure factor amplitudes from the reflection intensities[8]. Also, the program converts the SCALEPACK output file into .mtz file which CCP4 recognizes.

Analysis of asymmetric unit

Usually, protein crystals contain about 50% water. So we calculate the number of protein molecules that will allow the crystal to have 50% water. Cell Content Analysis module from the CCP4 suite was used for the purpose[9,10]. The molecular weight of the heterodimer is about 60 KDa. The Mattew's coefficient V_m was determined to be 2.6 ÅDa^{-1}. The solvent content was estimated to be 53%. This was consistent with the presence of two heterodimers in the asymmetric unit.

Structure determination by molecular replacement method

Determination of the three dimensional image of the molecule contained in the unit cell of a crystal by x-ray diffraction method involves computation of electron density ρ(x,y,z) throughout the volume (V) of the cell:

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

where F_{hkl} is the structure factor defined as : $F_{hkl} = |F_{hkl}|\, e^{ia_{hkl}}$. The amplitude $|F_{hkl}|$ is proportional to the square root of intensity I_{hkl} and therefore can be obtained from the diffraction data. However, the data does not contain the corresponding phases (a_{hkl}) thus impeding the calculation of structure factors. This problem originating because of the incoherence of x-rays constitutes the 'phase problem' in crystallography. The three standard methods for the determination of phase are multiple isomorphous replacement (MIR)[11], multiwavelength anomalous diffraction (MAD)[12] and molecular replacement (MR)[13]. In MIR, atleast two heavy atoms (mercury, platinum, gold) are incorporated without altering the unit cell dimension; thus the native and derivative crystals are isomorphous. Atoms contribute to the scattered intensity in proportion to the square of the number of electrons they contain. So the contribution of the heavy atoms to the intensity is disproportionately large. The differences between the scattered intensities of the native and derivate crystals will largely reflect the scattering contribution of the heavy atoms, and these differences can be used to compute a Patterson map. As there are only a few heavy atoms, the Patterson map will be simple and easy to deconvolute (i.e., work out the original positions that produced the observed Patterson peaks). Once we know where the heavy atoms are located in the crystal, we can compute their contribution to the structure factors. This allows us to make some deductions about possible values for the protein phase angles.

MAD exploits anomalous scattering produced due to incorporation of heavy atoms like selenium. Commonly, this is achieved by introducing selenomethionine instead of methionine by recombinant DNA methods. The incident photon is scattered when its energy is insufficient to cause any electronic transitions in an atom. If however, the incident photon has energy close to the transition energy which can bring an atom to an excited state, only a part of the photon is scattered. The remaining energy is absorbed and re-emitted at lower energy (fluorescence). The scattered photon is no longer in phase with the incident photon and also the intensity is reduced to some extent. This is called anomalous scattering and leads to an anomalous scattering increment, f^{Δ}, to the normal scattering factor, f°. The f° component is purely real, is independent of wavelength and fall off with scattering angle because of diffuseness in outer shells. In contrast f^{Δ} is complex, has a marked wavelength dependence due to the resonance, and is virtually independent of scattering angle because of its origin in core electrons The total scattering factor is $f = f^{\circ} + \left| f\Delta \right| e^{i\delta} = f^{\circ} + f' + if''$, where f' and if'' are the real and imaginary components of the anomalous scattering respectively. According to Friedel's law a pair of reflections h,k,l and $-h,-k,-l$ have the same amplitude but opposite phase. But the law fails at a wavelength close to the absorption edge of an atom, where anomalous scattering comes into effect. The violation of Friedel's law can be observed directly from the diffraction measurements and used to obtain phase information.

MR is the simplest method for phase determination when the structure of a homologous protein is known. The objective of MR is to place the search molecule in the unit cell of the target crystal so as to account for the diffraction pattern. This requires proper orientation and positioning of the molecule in the target unit cell. Orientation requires specification of the three rotational angles [αβγ (Eulerain angles) or φωκ (spherical polar angles)] while positioning requires specification of three translational parameters. So, placement of one molecule in the unit cell is a 6-dimesnional problem.

The intensities calculated from the experimental data can be used to perform a Fourier transform to give a Patterson function or Patterson map. The Patterson function gives a map of the vectors between atoms. Intramolecular vectors (from one atom in the molecule to another atom in the same molecule) depend only on the orientation of the molecule and not on its position in the cell, so are exploited in the rotation function. The Patterson map of the target is the stationary map over which the map of the search model is rotated till maximum superposition is obtained. Rotation function is a product function[14] and the agreement between the search and target Patterson functions is evaluated by an index (R) defined as:

$$\mathbf{R}(\kappa,\phi,\psi) = \int_{r_{\min}}^{r_{\max}} \mathbf{P}_{nat}(\mathbf{u})\,\mathbf{P}_{\mathbf{mod}}(\kappa,\phi,\psi,\mathbf{u})\,d\mathbf{u}$$

where Patterson function P_{nat} is rotated w.r.t P_{mod}. The integral is over a spherical volume U centered at the Patterson origin. In the integration over a spherical shell, a region near the origin is typically omitted to avoid the large origin peak of the Patterson, which would add a large constant term. The sphere is limited in radius because the intramolecular vectors are more concentrated near the origin. Therefore the intramolecular vectors maximized while simultaneously minimizing the intermolecular vectors. A maximum value for R is obtained when the two Patterson maps are well superimposed.

In contrast, intermolecular vectors depend both on the orientation of the molecule and on its position, so are exploited in the translation function. The full Patterson map is used without the origin. As in the previous case, translation function is quantified by comparing the observed and calculated Pattersons. Translation function (T) is again a product function[15] which is evaluated by Fourier transform using the correlation theorem:

$$T(t) = \int_{cell} P_{2\to1}\,(u\text{-}t)P_{nat}(u)\,du$$

where t is the intermolecular vector, u is vector in Patterson space and the integration is over the whole unit cell. When the model is correctly positioned, the maximum value of T is obtained at t.

The structure of *B. subtilis* GGT was solved my molecular replacement method using Molrep program from the CCP4 suite[16]. The homologue from *E. coli* (PDB code: 2dbu) was used as the search model. The asymmetric unit of PDB entry 2dbu contain 4 chains corresponding to

two heterodimers. One copy of the heterodimer (chain ID A and B) was used. The pdb file containing the search model was processed with Chainsaw program. Chainsaw is a molecular replacement utility which takes an alignment between target and model sequences and modifies the model pdb file by pruning non-conserved residues.

Structure refinement and model building

Crystallographic refinement aims at optimising the agreement of an atomic model with both observed diffraction data and chemical restraints. The structure presented in this thesis was refined using REFMAC 5 which is based on maximum likelihood method[17]. Initially, a round of rigid body refinement was done. COOT was used to visualise and correct the protein model. *2Fo-Fc* and *Fo-Fc* maps were used as guide for model building. The model was taken trough iterative rounds of refinement and manual correction. During these cycles, refinement was done in the restrained mode. In this mode, the refinement program sets up *restraints* between related atoms so that they conform to expected geometry. Geometric restraints include bond lengths, bond angles, planar groups, chirality *etc* for the common chemical monomers which are found in macromolecules (*e.g.* amino acids and nucleic acids). In addition NCS restraint was applied over equivalent residues in the two heterodimers present in the asymmetric unit. Restraints are necessary as macromolecular structure determination suffers from poor observation: parameter ratio at the resolutions to which crystals normally diffract. If we simply refine the positions and B-factors of all

the atoms, the refinement will be poorly behaved, the data will be terribly overfit, and the resulting atomic model will probably be very poor. The way we deal with this is either to add "observations" in the form of restraints, or reduce the number of parameters by constraining the model in some way. The restraints are entered as terms in the refinement target, and are weighted so that the deviations from ideal values match the deviations found in databases of high resolution structures.

Cross-validation

The quality of the fit of a model to the diffraction data is given by the R factor, which measures the discrepancy between the observed (Fo) and calculated (Fc) structure factor amplitudes.

$$R = \frac{\Sigma \, | \, |Fo| - |Fc| \, |}{\Sigma \, |Fo|}$$

This value can be made arbitrarily low by increasing the number of adjustable parameters used to describe the model (overfitting of the data). The method of statistical cross validation using free *R* factor for crystallographic refinement is a more faithful indicator of model quality[18,19]. In this method, a small fraction (5-10%) of the diffraction data is randomly selected and kept aside. The flagged off set is called the test set while the rest of the data constitutes the working set. Only the working data is used for model building. Changes to the model that does not conform to the diffraction data does not improve the fit of the model to test set. Thus the R factor calculated with the test

set (R_{free}) is free from overfitting bias as it is not used for model building. In the present study, 5% of the data (5358 out of 101881 reflections) were used to calculate R_{free}. The R factor calculated using the working set of the data during refinement is denoted as R_{cryst}.

Modeling of water

Water molecules were added when the R_{cryst} dropped to 25%. FIND WATER function available in COOT was used for the purpose. Further additions or deletions were made manually with COOT[20]. Water was added upon satisfaction of the following criteria:

1. A clear peak was visible in both 2*Fo-Fc* (contoured at 1σ level) and *Fo-Fc* (contoured at 3σ level) electron density maps.

2. The picked water was with hydrogen bonding distance (2.7-3.5 Å) of another water or protein O or N atom

After water picking, iterative cycles of refinement and model building from the newly improved phases were carried out to give the final model.

Validation of the final model

The stereochemical quality of the model was analysed with PROCHECK[21] available in the CCP4 suite.

Modelling of ligands in the active site

The structures were modelled using Quanta (Accelrys Inc, San Diego). The Charmm22 polar hydrogen potential energy was used to mimic the inter-atomic interactions with the electrostatic and van der Waals interactions cut off at 12Å and a distance-dependent continuum electrostatic model that was attenuated by a factor of 4. Glutamate and γ-glutamyl ester were modelled with the guidance of crystal complexes from *E. coli* GGT (PDB codes: 2dbw, 2dbx); minimizations were carried out with the protein atoms restrained by harmonic forces. The model of γ-tetra-glutamic acid was constructed so that Glu-1 remained in the position specified by the glutamate and γ-glutamyl ester, while the conformation was subject to molecular dynamics simulation at 600K for 100ps under the influence of a protein held fixed. The sampled conformations were then energy minimized with the protein atoms constrained by harmonic forces. The lowest energy structures were chosen as representatives.

Structure deposition:

The coordinates and the structures factors of the structure presented in this thesis have been deposited in Protein Data Bank with accession codes 2v36

Note on pictures:

All molecular graphics appearing in this thesis were prepared either with CCP4mg[22,23] or Pymol[24].

Sequence analysis:

Multiple sequence alignments were done with CLUSTAL W [25] available through European Bioinformatics Institute Server (www.ebi.ac.uk) and displayed with Espript[26]. Secondary structure was defined with DSSP[27].

References

1. Ogawa, Y., Hosoyama, H., Hamano, M., and Motai, H. (1991) *Agric Biol Chem* **55**, 2971-2977
2. Suzuki, H., Kumagai, H., and Tochikura, T. (1986) *J Bacteriol* **168**, 1325-1331
3. Hashimoto, W., Suzuki, H., Yamamoto, K. and Kumagai, H. (1995) *J Biochem (Tokyo)* **118**, 75-80
4. Brzozowski, A.W. and Walton, J. (2001) *J Appl Cryst* **34**, 97-101
5. Miyahara, J., Takahashi, K., Amemiya, Y., *et al.* (1986) *Nucl Instr Meth In Phy res* **A46**, 572-578
6. Westbrook, E.M. and Naday, I.(1997) *Methods Enzymol* **276**, 244-268
7. Otwinowski, Z. and Minor, W. *Methods Enzymol* (1997) **276,** 307-326
8. French, G.S., and Wilson, K.S. (1978) *Acta Crystallogr.* **A34**, 517
9. Mathews, B. W. *J Mol Biol* (1968) **33**, 491-497

10. "The CCP4 Suite: Programs for Protein Crystallography". *Acta Crystallogr Section* ***D*** **50**, 760-763
11. Ke, H. (1997). *Methods Enzymol* **267**, 448-60
12. Hendrickson, W.A. and Ogate, C.M. (1997). *Methods Enzymol* **276**, 494-523
13. Rossmann, M.G. and Arnold, E. (2001) International tables for crystallography, Vol. B (ed. U. Shumueli), 235-263, Iinternational Union of crystallography/ Kluwer, Dordrecht
14. Rossmann, M.G., and Blow, D.M. *Acta Crystallogr* (1962) **15**, 23-31
15. Crowther, R.A., and Blow, D.M. *Acta Crystallogr* (1967) **23**, 544-548
16. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr Sect D* **56**, 1622-1624
17. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) *Acta Crystallogr Sect D* **53**, 240-255
18. Brünger, A.T. *Nature* (1992) **355**, 458-460
19. Kleywegt, G.J. and Brünger, A.T. (1996) *Structure* **4**, 897-904
20. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr Sect D* **60**, 2126-2132
21. Laskowski, R.A., McArthur, M.W., Moss, D.S., and Thornton, J. (1993) *J Appl Crystallogr* **26**, 282-291
22. Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K., and Noble, M. (2002) *Acta Crystallogr Sect D* **58**, 1955-1957
23. Potterton, L., McNicholas, S., Krissinel, E., *et al*. (2004) *Acta Crystallogr Sect D* **60**, 2288-2294

24. DeLano, W.L. (2002) *The PyMol Molecular Graphics System* (DeLano Scientific, San Carlos, CA)
25. Thompson, J.D., Higgins, D. G. and Gibson, T. J. (1994) *Nucleic Acids Res* **22**, 4673-4680
26. Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999) *Bioinformatics* **15**, 305-308
27. Kabsch, W. and Sander, C. (1983) *Biopolymers* **22**, 2577-2637

RESULTS

Chapter 3

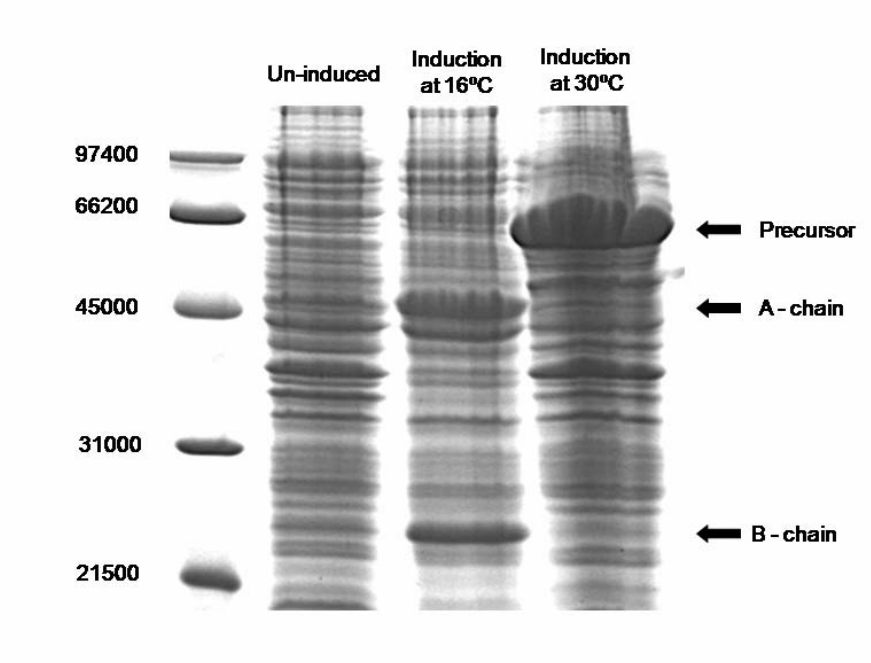


Figure 3.1: SDS-PAGE gel showing the effect of temperature on the expression of pET26-BsGGT construct in *E. coli* BL21 (DE3) host. Position of precursor, A and B chains are marked. Note the absence of processed protein (A and B chains) in sample incubated at 30 C.

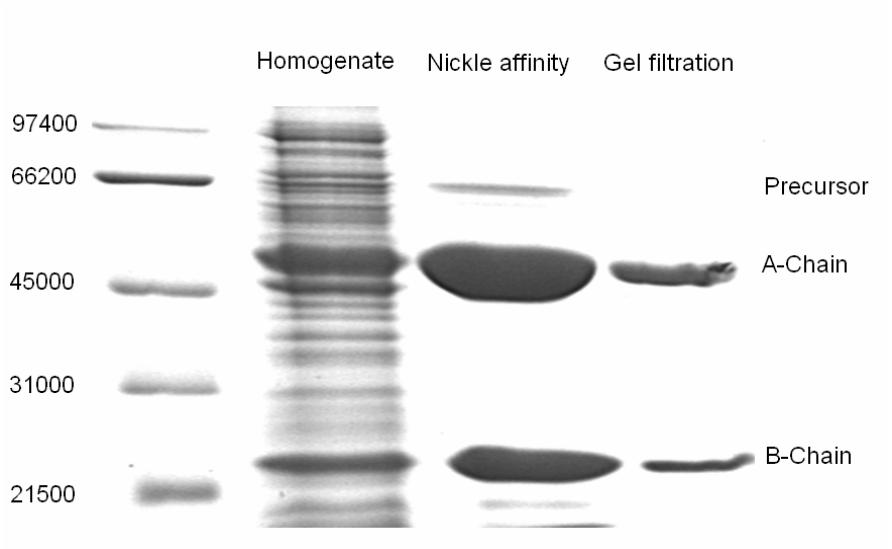


Figure 3.2: SDS-PAGE gel showing stages of enzyme purification. Molecular weight markers are in Daltons.

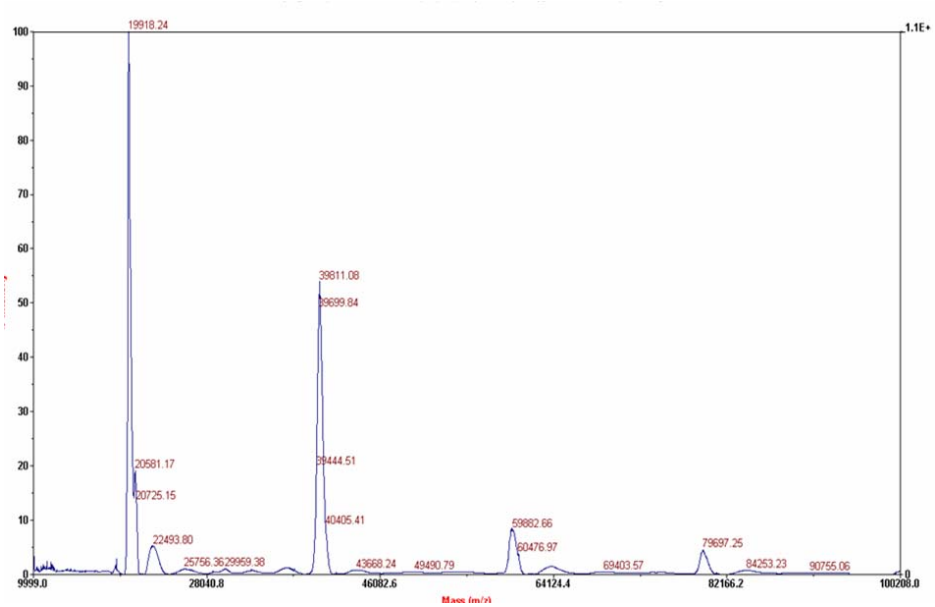


Figure 3.3: MALDI-TOF mass spectrum of *B. subtilis* GGT showing the molecular weight of the two chains. Chain A ≈40,000 Da and Chain B ≈20,000 Da.

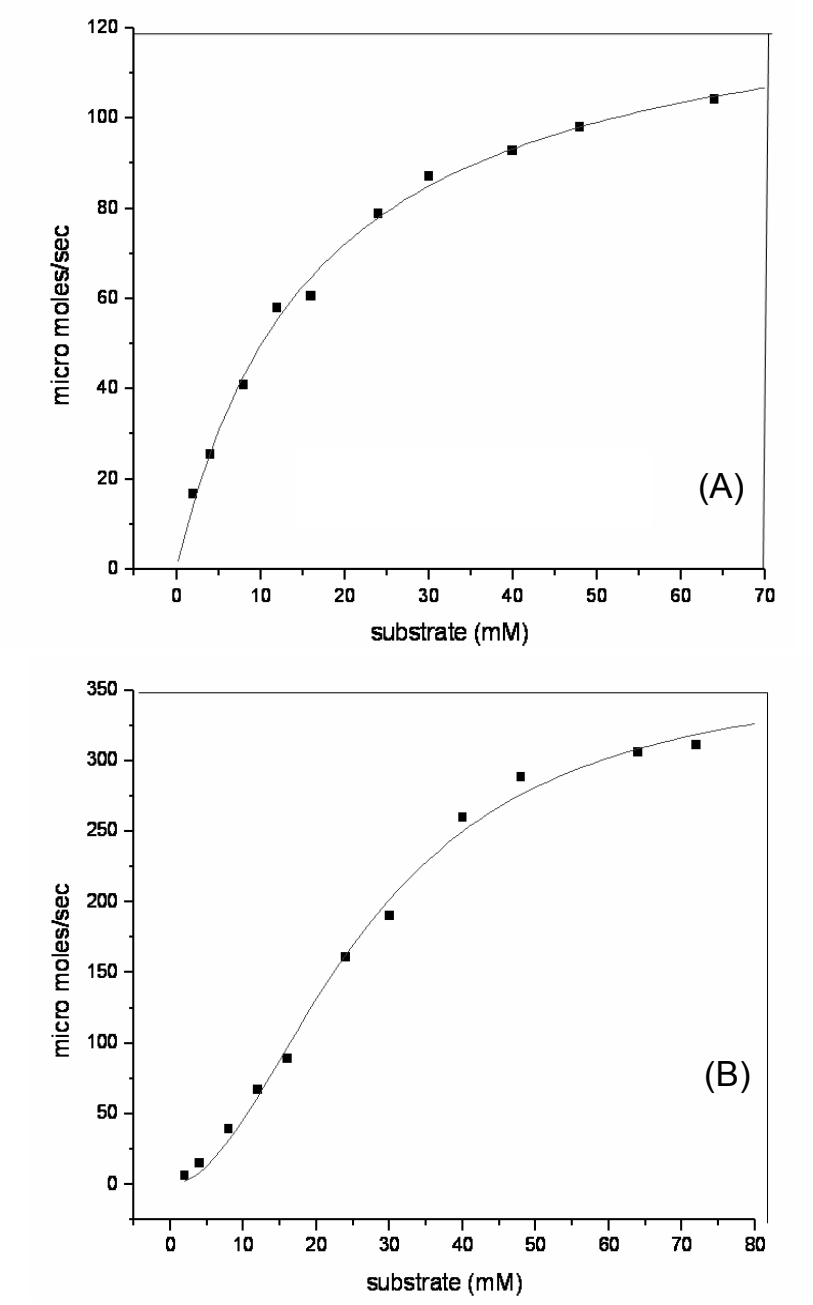


Figure 3.4: Effect of pH on the nature of substrate saturation curve. (A) pH 9.0 and (B) pH 11.0

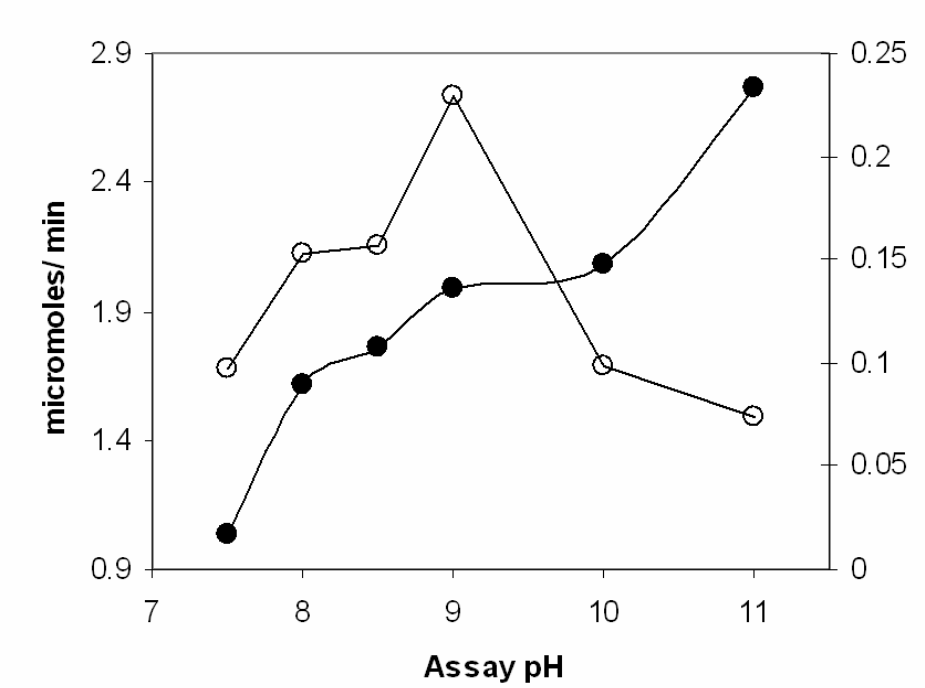


Figure 3.5: Effect of substrate concentration on pH optima of *B. subtilis* GGT. The enzyme was assayed with 64mM (●) and 2mM (○) concentration of substrate.

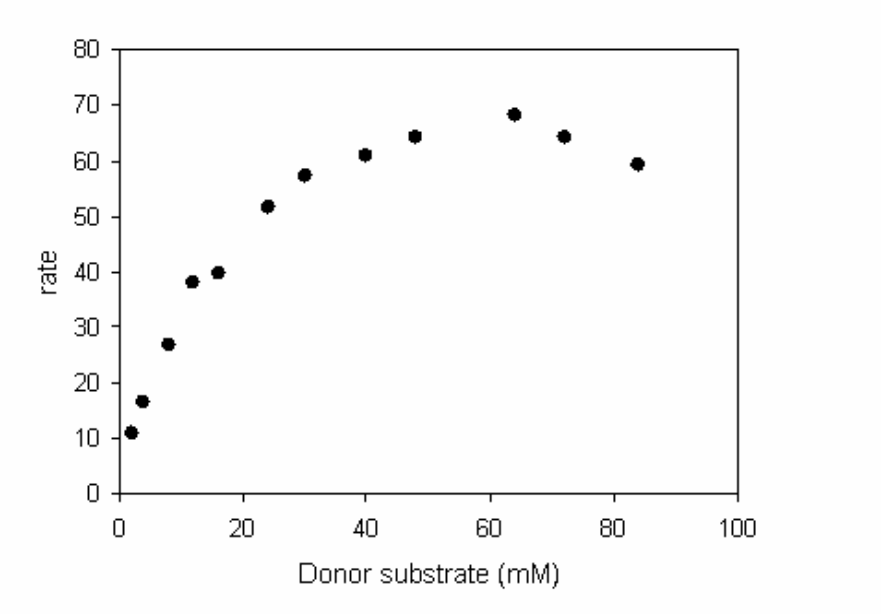


Figure 3.6: Inhibition of hydrolase activity at higher concentrations of donor substrate. The saturation curve was determined at pH 9.0.

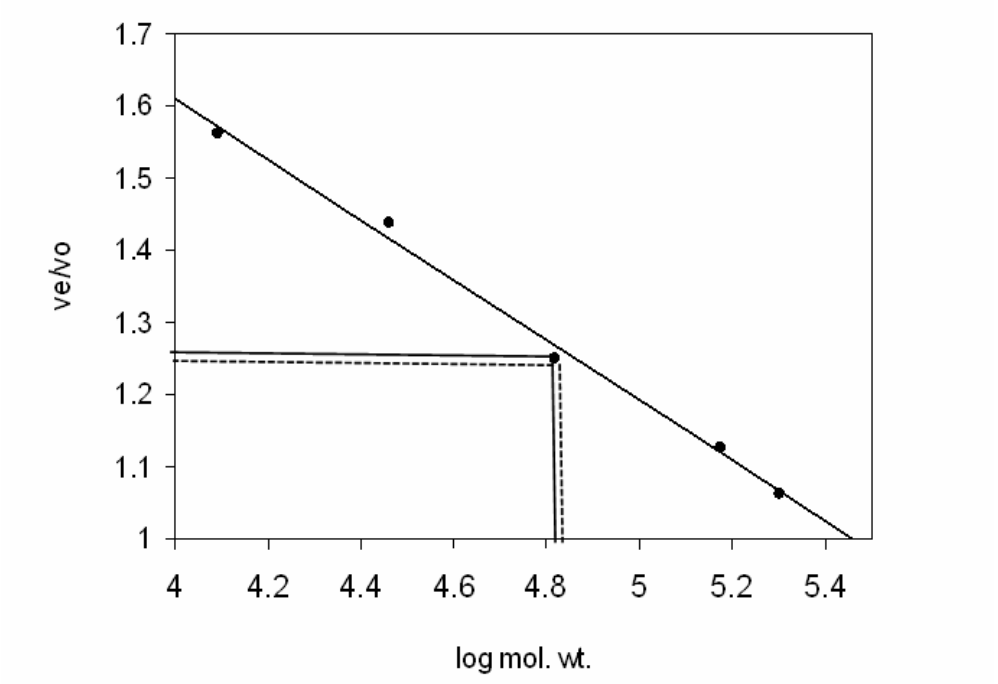


Figure 3.7: Calibration curve for estimation of molecular weight by gel filtration chromatography. Elution of test sample in pH 7.0 (continuous line) and pH 11.0 (discontinuous line) are marked.

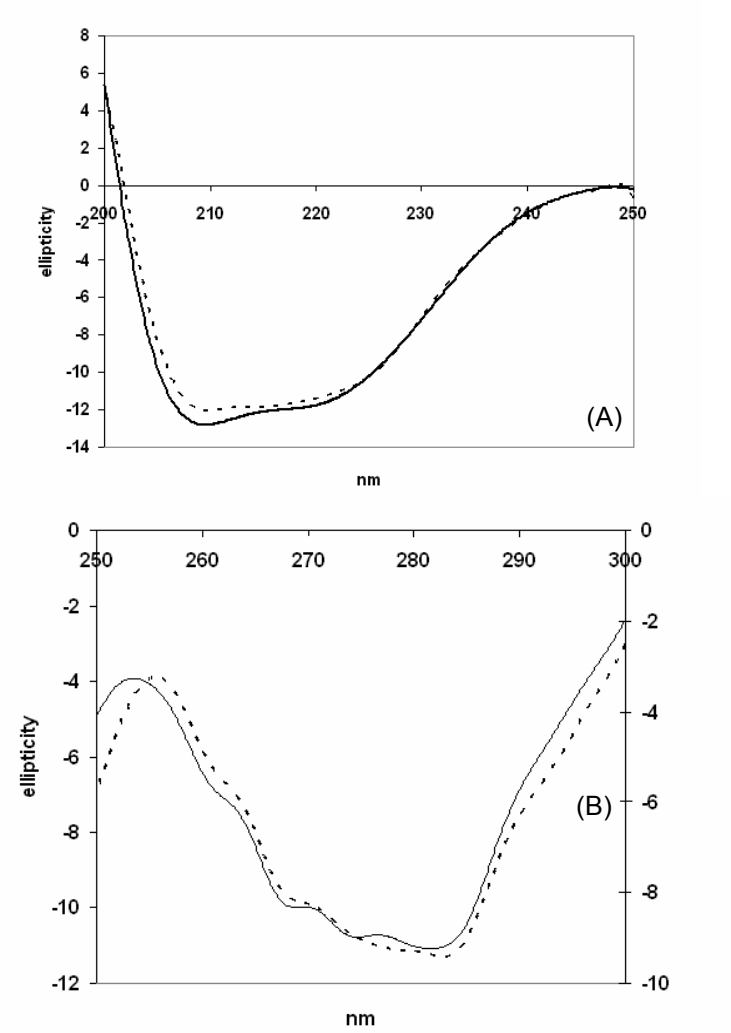


Figure 3.8: Near (A) and far (B) UV-CD in pH 7.0 (continuous curve) and pH 11.0 (discontinuous curve).

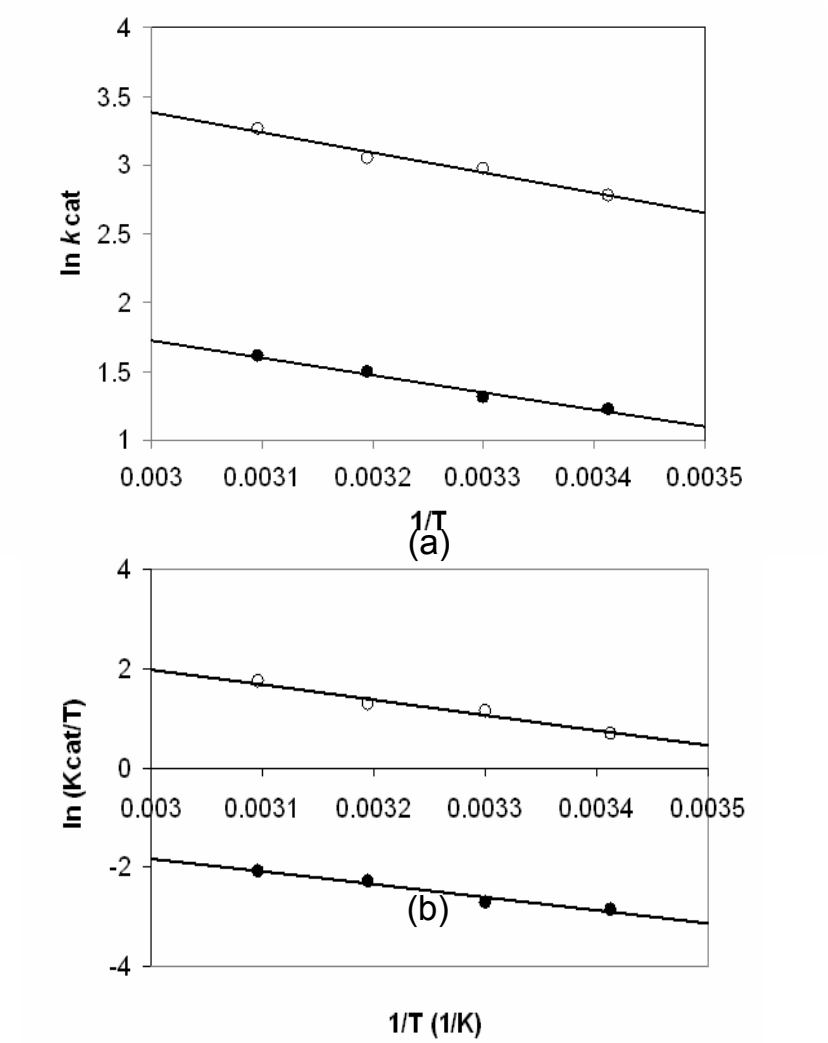


Figure 3.9: Arrhenius (a) and Eyring's (b) plots for hydrolysis at pH 7.5(●) and 11.0 (○).

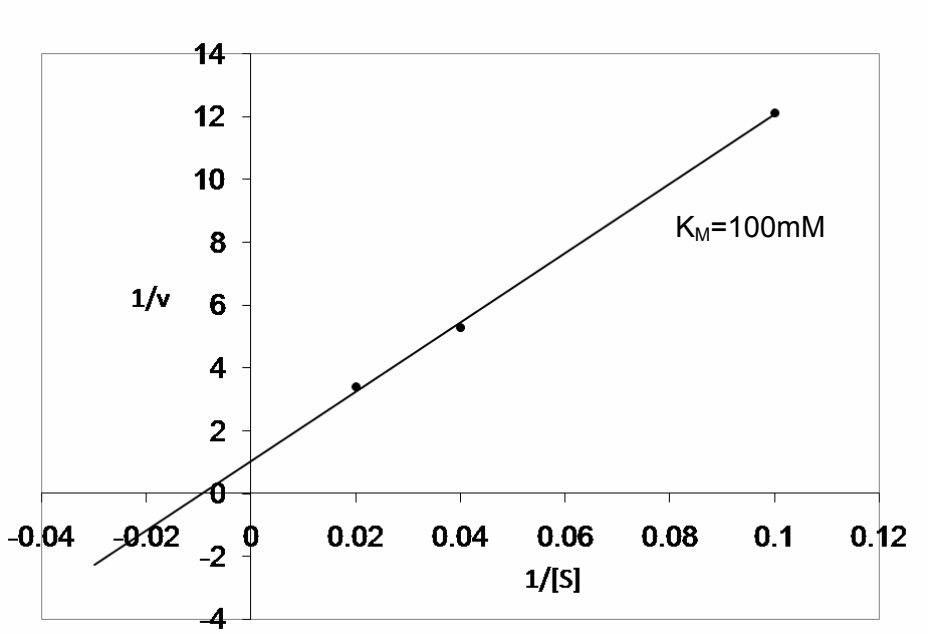


Figure 3.10: Secondary plot for determination of K_M for glycylglycine. Saturation curves for donor consumption was determined at pH 7.5 in the presence of 10, 25 and 50mM glycylglycine and the apparent K_M was determined by non-linear method. By extrapolation, the K_M was found to 100mM.

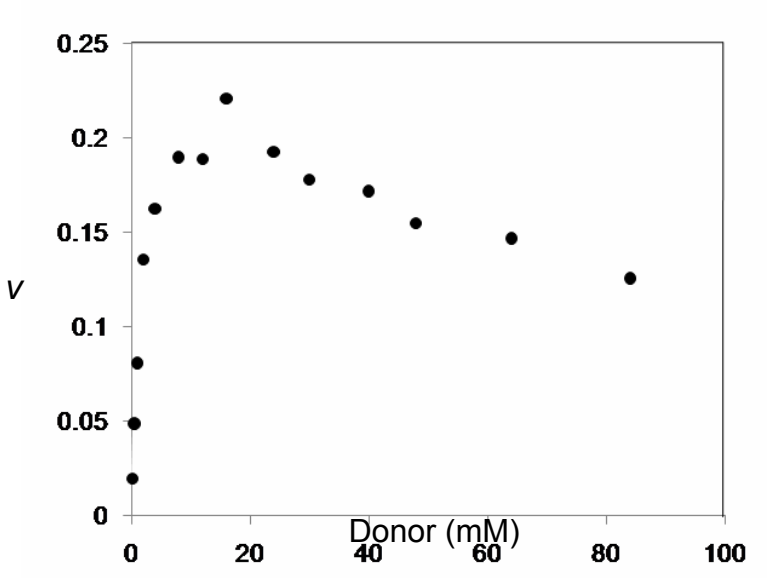


Figure 3.11: Saturation curve for donor consumption at pH 9.0 in the presence of 50mM glycylglycine. Note the decrease in velocity at higher concentrations of the donor substrate.

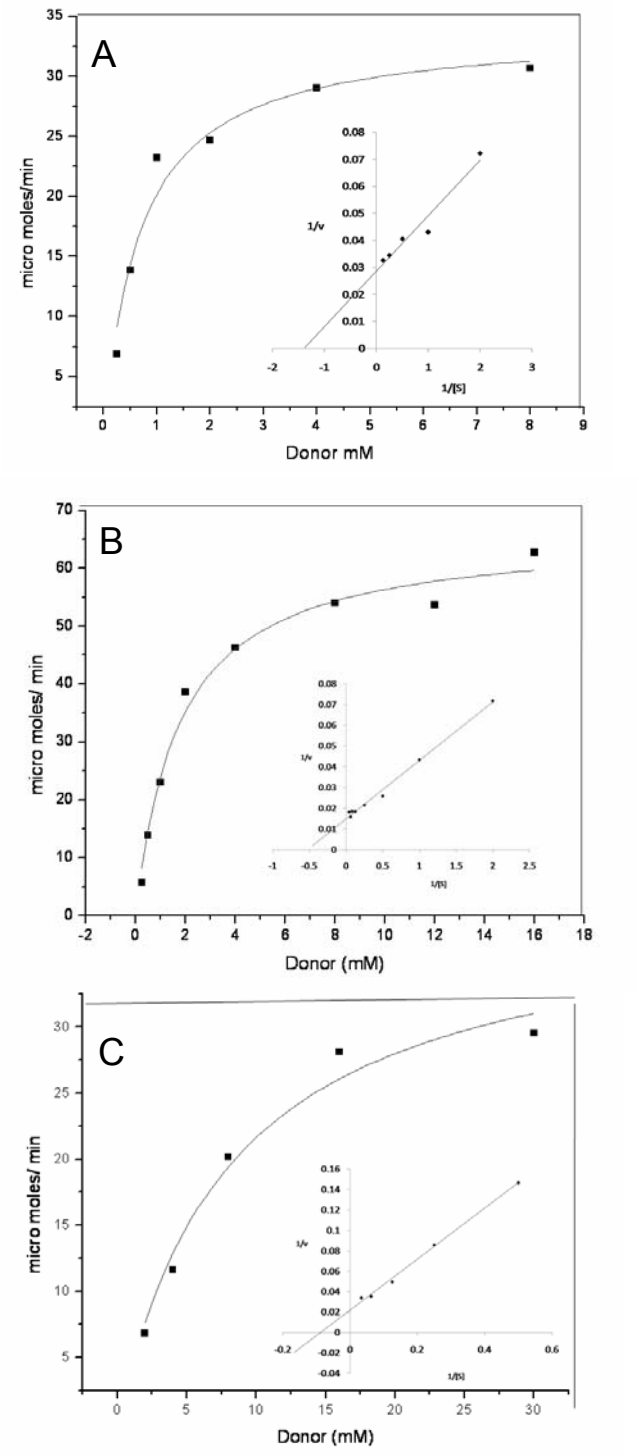


Figure 3.12: Kinetics of transpeptidation at pH 7.5 (A), 9.0 (B) and 11.0 (C). Inset shows Lineweaver Burk plot for the corresponding data.

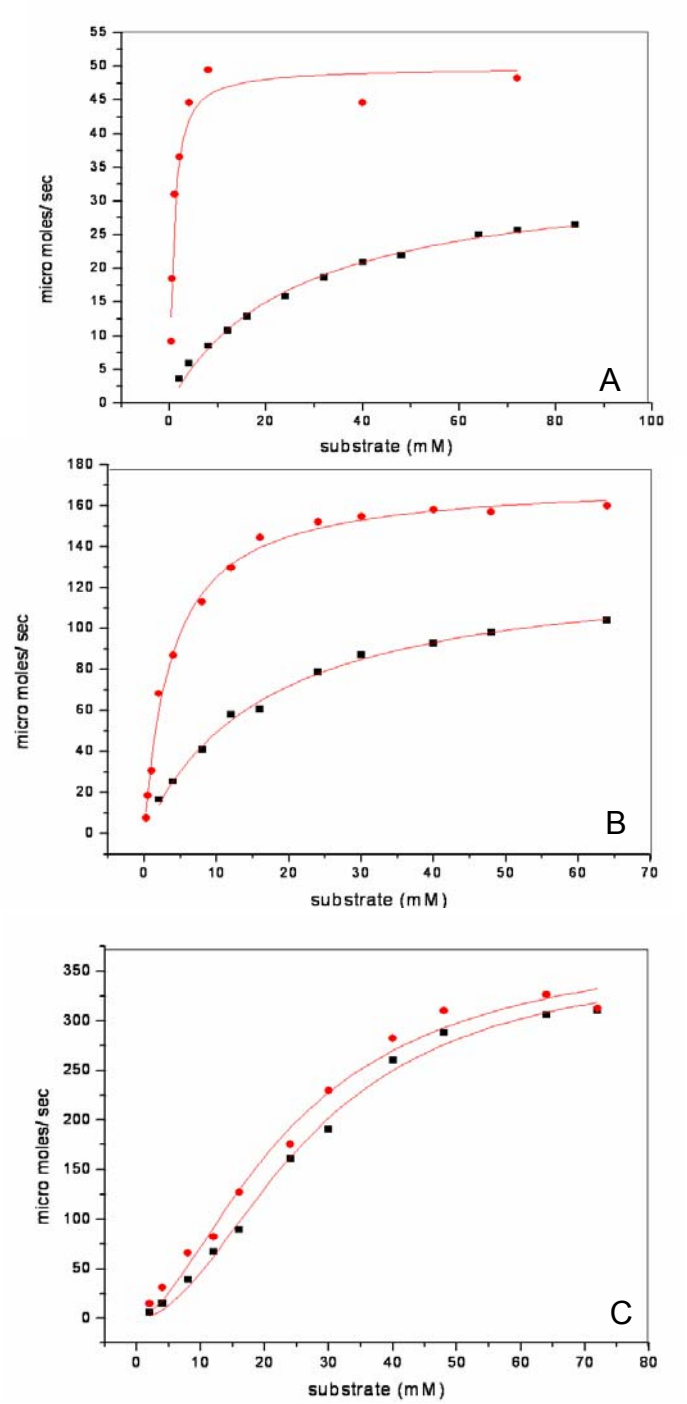
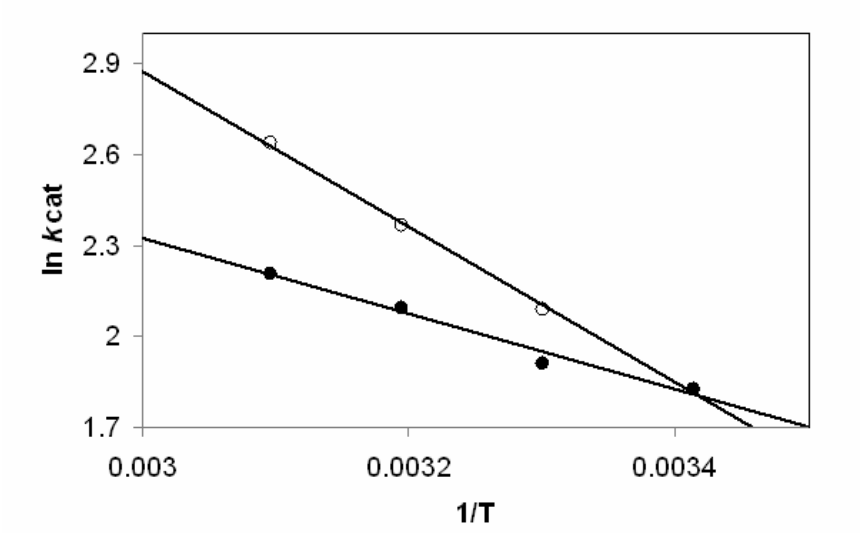
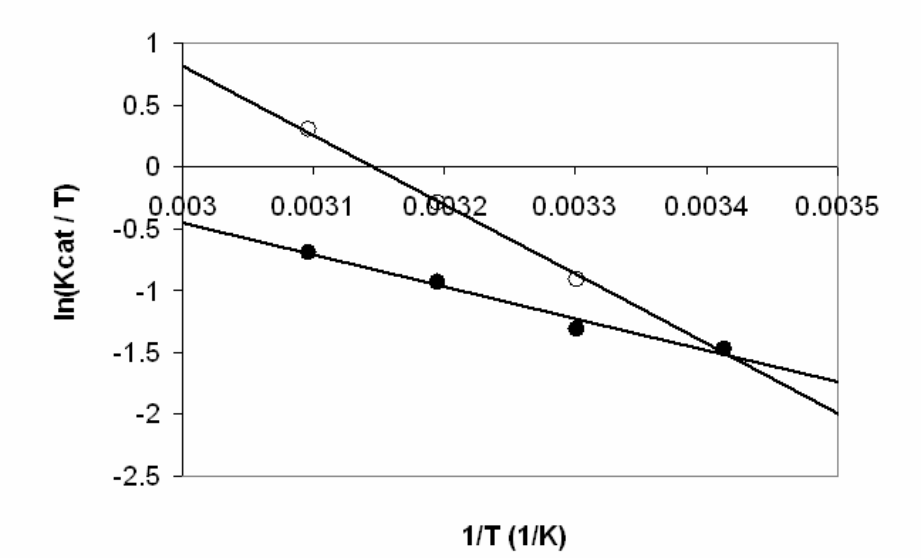


Figure 3.13: Michaelis-Menten plots comparing donor hydrolysis in the presence (●) and absence (■) of glycylglycine at pH 7.5 (A), 9.0 (B) and 11.0 (C).

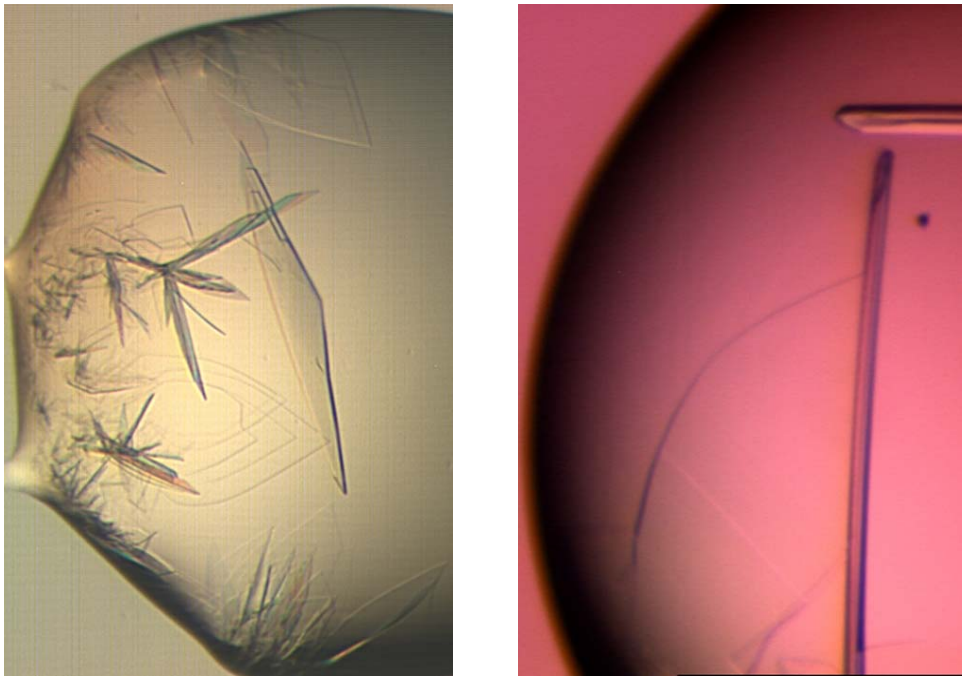


(a)



(b)

Figure 3.14: (a) Arrhenius and Eyring's plots showing thermodynamic differences between hydrolysis (●) and transpeptidation (○).



(A) (B)

Figure 3.15: *B. subtilis* GGT crystals before (A) and after improvement (B). A plate like crystal from the drop shown in picture (B) was used for collecting the diffraction data.

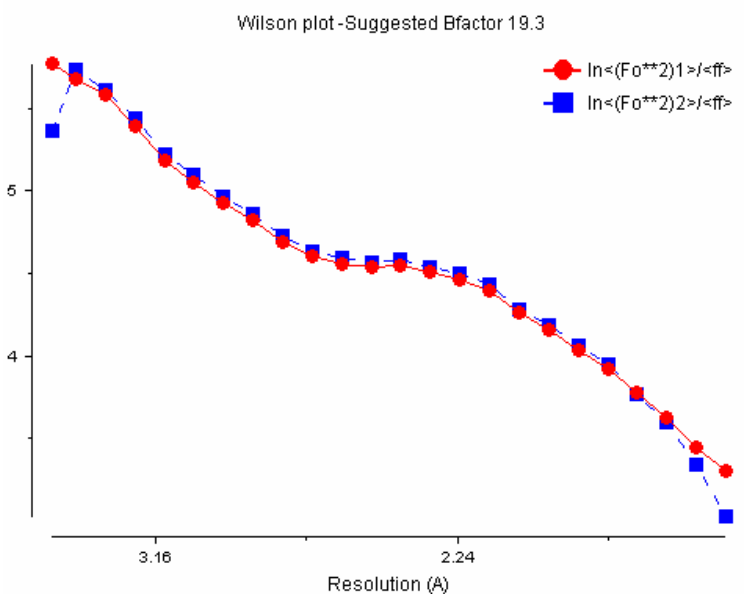


Figure 3.16: Wilson plot of data used for determination of the crystal structure of *B. subtilis* GGT

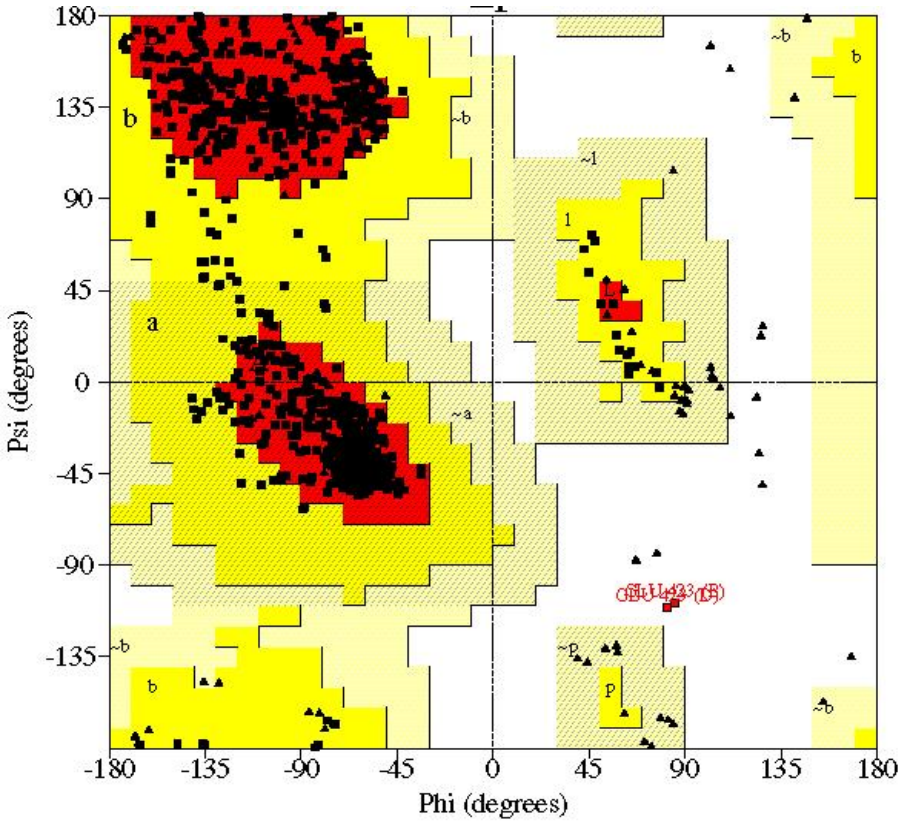


Figure 3.17: Ramachandran plot of the final model of *B. subtilis* GGT. Of the total residues, 91.1% are in the favoured region while 8.7% are in the allowed region. 0.2% of non-glycine residues are in the disallowed region. The digression is due to Glu^{423} from the two heterodimers in the asymmetric unit.

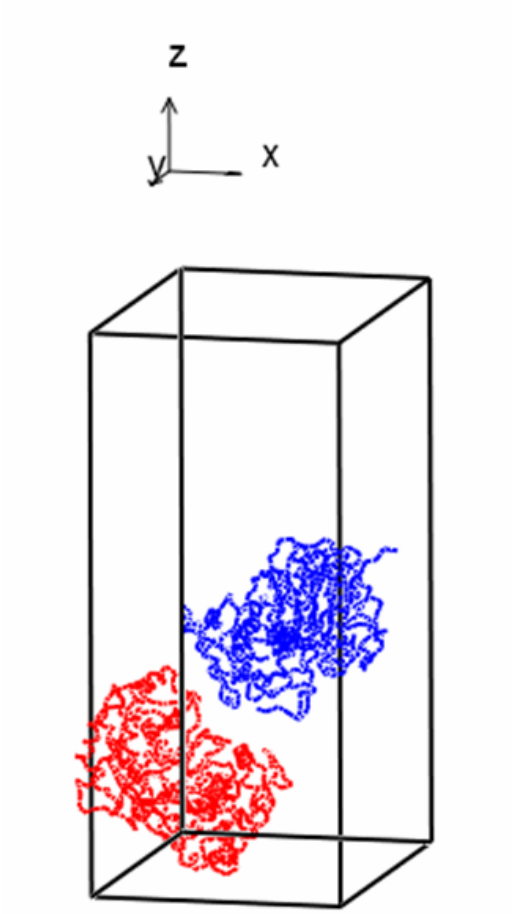


Figure 3.18: Diagram showing the packing of two heterodimers in the asymmetric unit. The space group is $P2_12_12_1$.

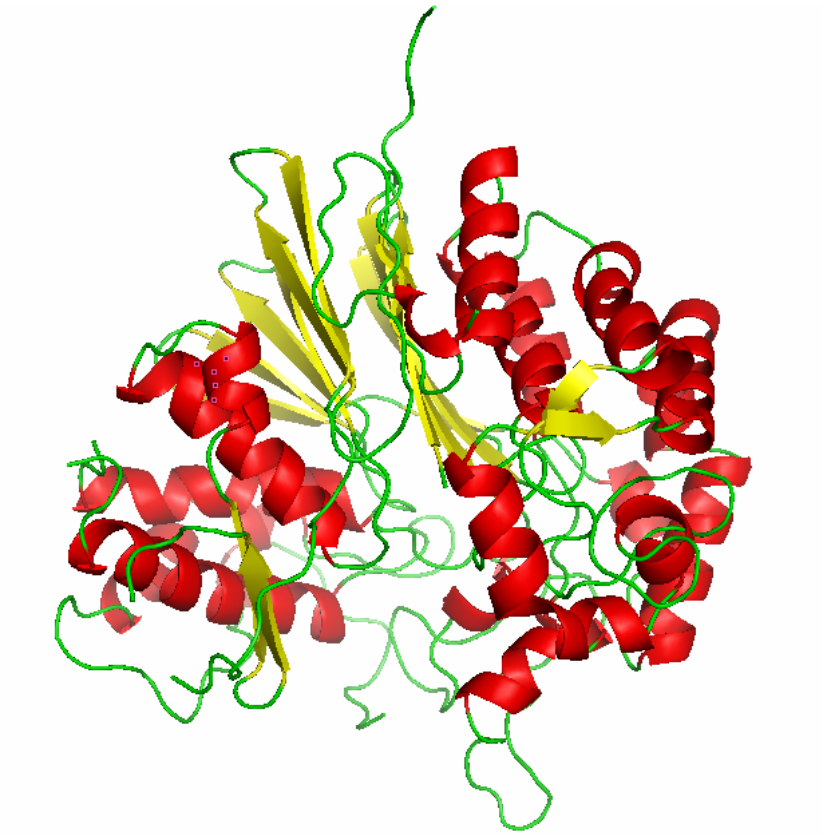


Figure 3.19. Ribbon representation of *B. subtilis* GGT model. α helices and β strands are shown in yellow and green respectively. The view is along the central β-β plate.

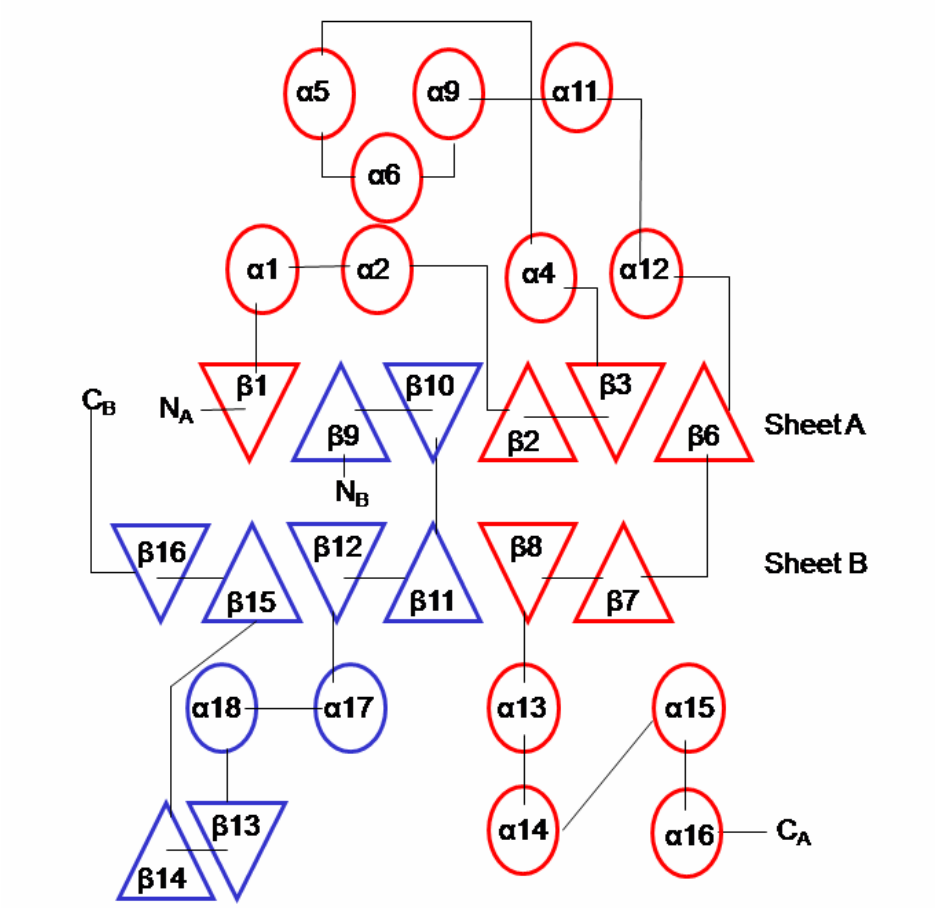


Figure 3.20: Topology diagram of secondary structural elements of *B. subtilis* GGT. Triangles and circles represent β-strands and α-helices respectively. Elements from the chain A and B are in red and blue respectively. The termini ofA and B chains are indicated by subscripts

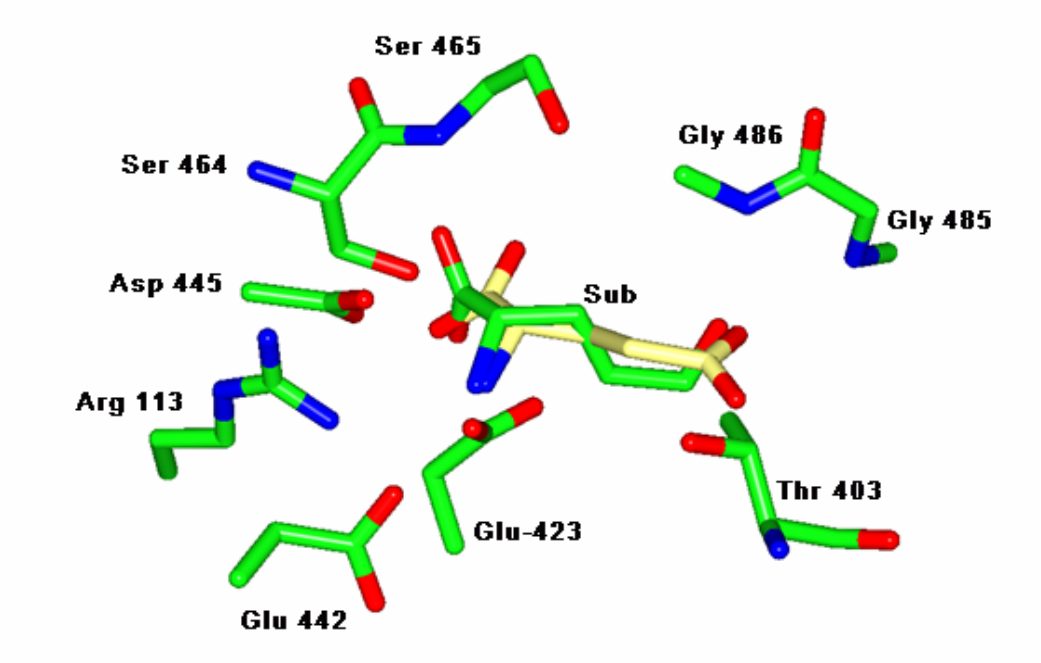


Figure 3.21: Diagram showing interaction between bound substrate and the active site residues. The model was generated by dynamic simulation. Glutamate is shown with yellow carbon backbone and its acyl complex with the catalytic nucleophile with green carbon backbone.

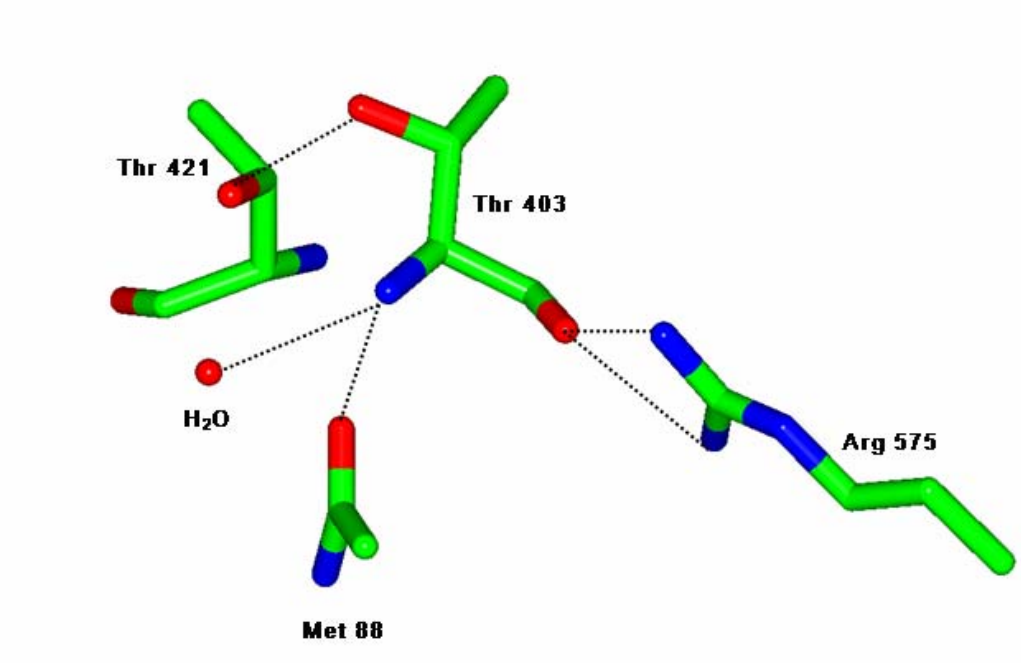


Figure 3.22: Diagram of *B. subtilis* GGT model showing the interactions formed by the catalytic nucleophile (Thr 403). Potential hydrogen bonds are shown as discontinuous line.

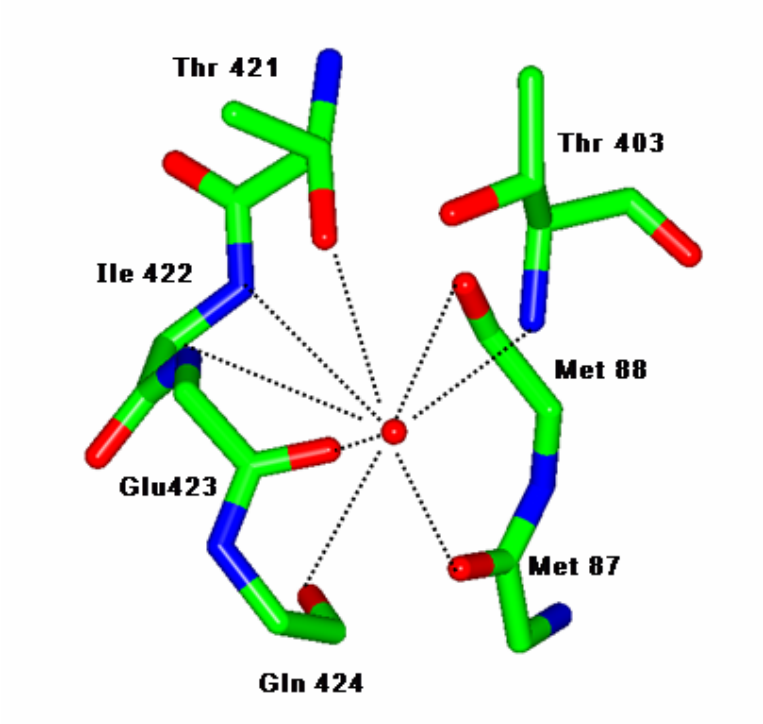


Figure 3.23: Diagram of *B. subtilis* GGT model showing the interactions formed by the α-amino group of the catalytic nucleophile (Thr 403). The water molecule is represented as a red sphere. Potential hydrogen bonds are shown as discontinuous line.

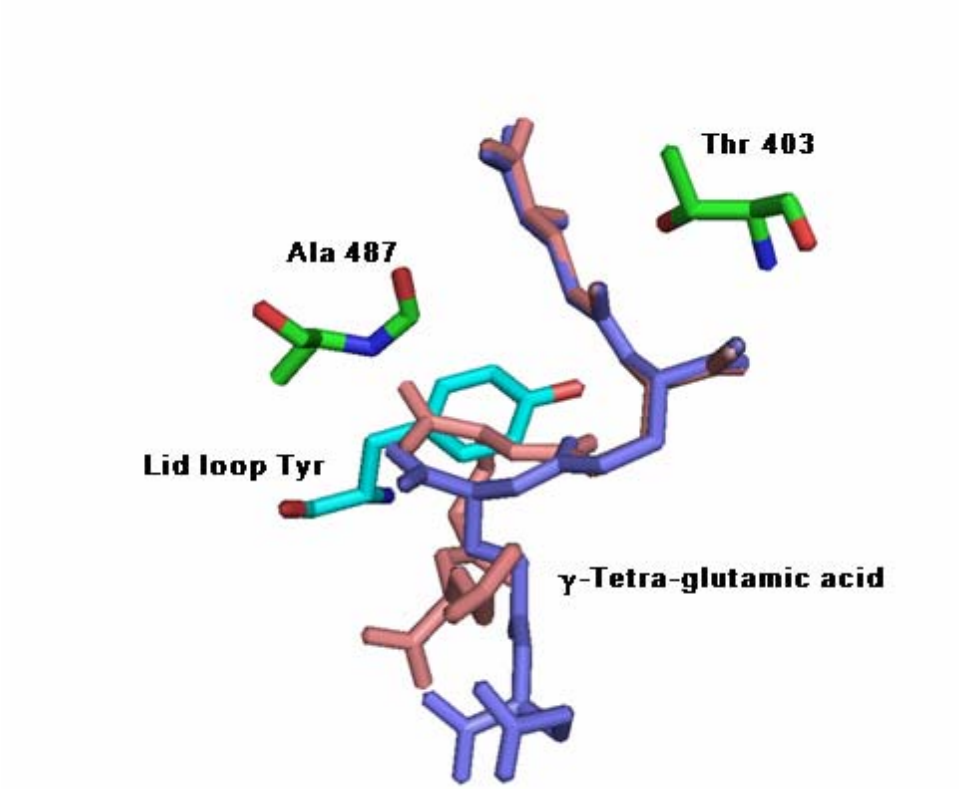


Figure 3.24: Diagrammatic representation of *B. subtilis* GGT model with γ-tetra glutamate bound in the active site. The model was prepared by molecular dynamics simulation. Two extreme cases of the five posssible conformations are shown. Thr 403 represents the catalytic nucleophile. Lid loop Tyr obtained by superposition with *E. coli* GGT (PDB code 2dbu) is shown. Note the steric and charge conflict between the lid loop Tyr and α-carboxylate group of Glu-3 of the substrate.

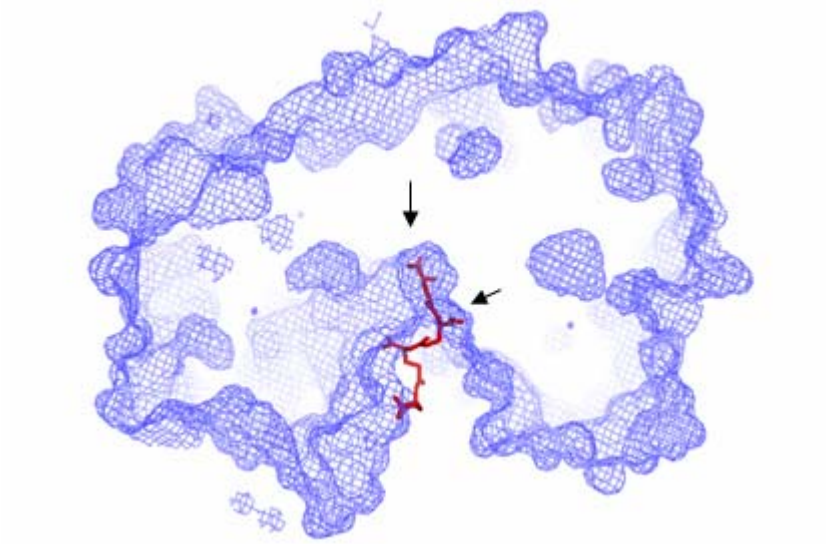


Figure 3.25: Diagrammatic representation of *B. subtilis* GGT model with γ-tetra glutamate bound in the active site. The enzyme model (cross section) is shown in mesh covered surface diagram; substrate is shown in red coloured cylinder model. Note that the structure is kidney-shaped with the active site located in the basal groove. Top arrow represents the cup-shaped cavity for binding Glu-1 residue and the oblique arrow represented a shallow saucer-shaped cavity for binding α-carboxylate of Glu-2.

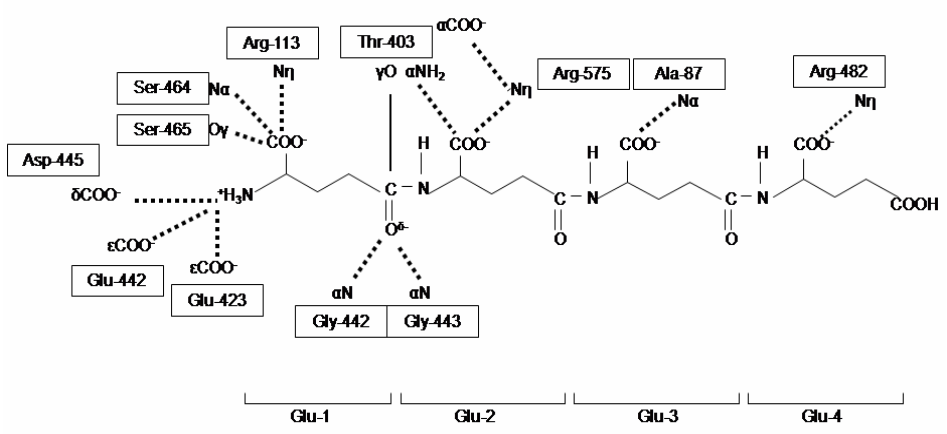


Figure 3.26: Diagrammatic representation of interactions between γ-tetra-glutamic acid and the active site of *B. subtilis* GGT determined from the model prepared by molecular dynamics simulation. Potential bonds are shown as discontinuous lines. The four glutamyl residues of the substrate are labeled numerically from the N-terminus.

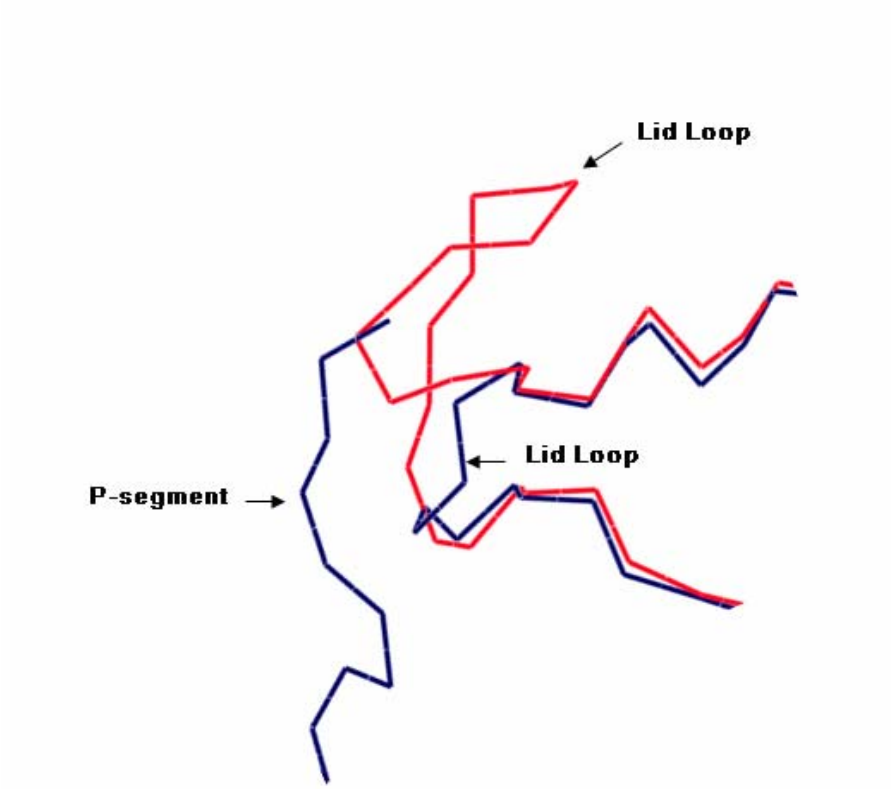
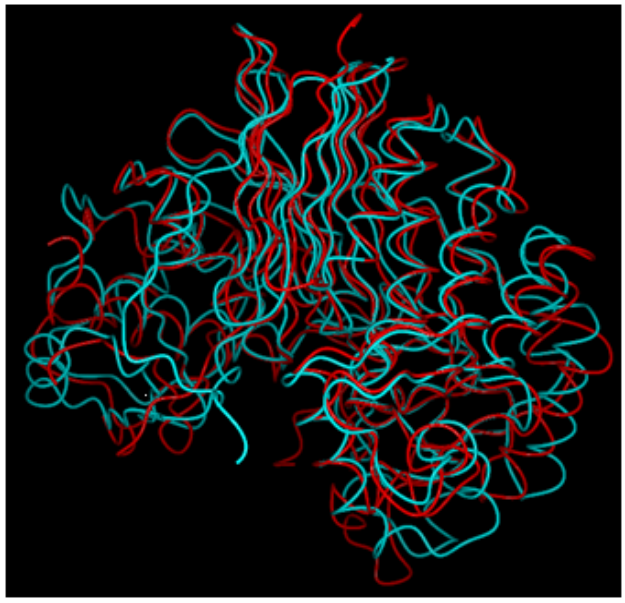
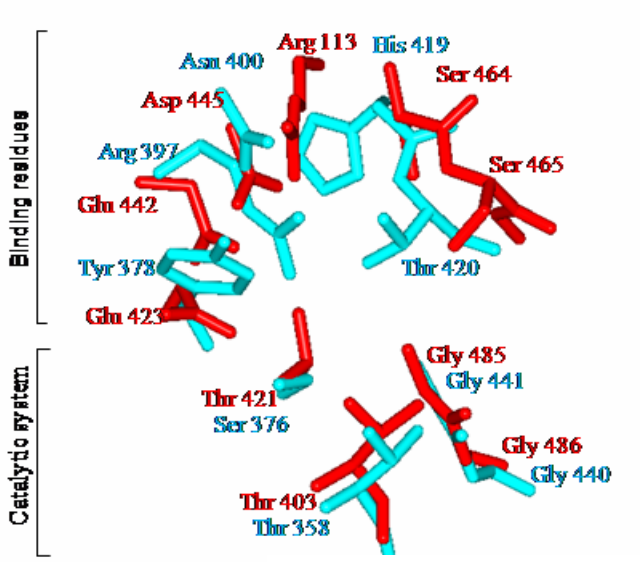


Figure 3.27: Diagrammatic representation of superposed Cα-Cα back bones of *B. subtilis* GGT (blue) and *E. coli* GGT (red) at the lid loop region. Note that the lid loop equivalent in *B. subtilis* GGT is truncated. P-segment refers to the C-terminus of chain A.

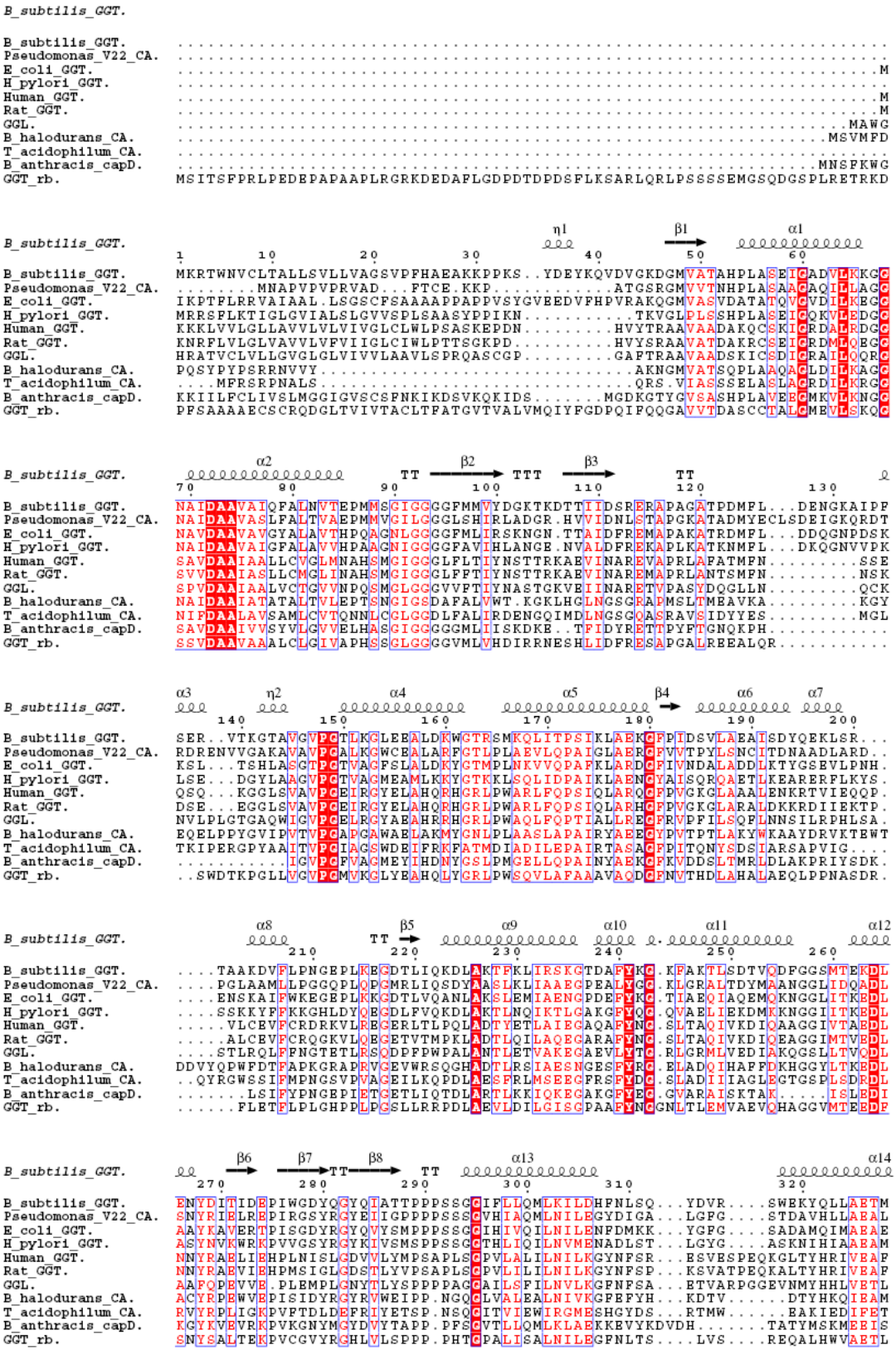


(A)



(B)

Figure 3.28: Diagrams showing superposition of Cα-Cα backbones (A) of *B. subtilis* GGT and *B. halodurans* CA and their active sites (B)


B_subtilis_GGT.
Pseudomonas_V22_CA.
E_coli_GGT.
H_pylori_GGT.
Human_GGT.
Rat_GGT.
GGL.
B_halodurans_CA.
T_acidophilum_CA.
B_anthracis_capD.
GGT_rb.

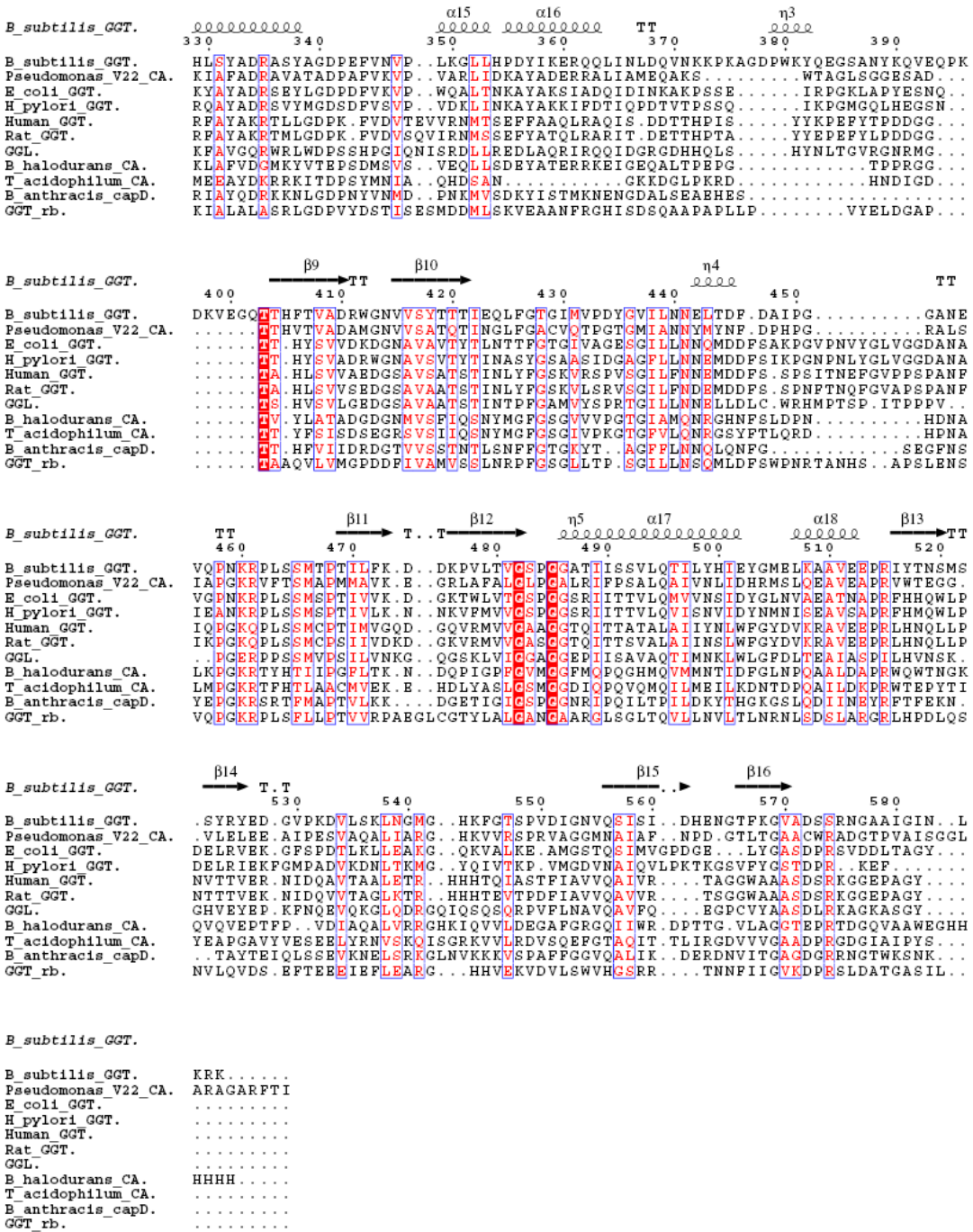


Figure 3.29: Multiple sequence alignment of *B. subtilis* GGT with various GGTs and CAs. Secondary structure of *B. subtilis* GGT as defined by DSSP is shown.

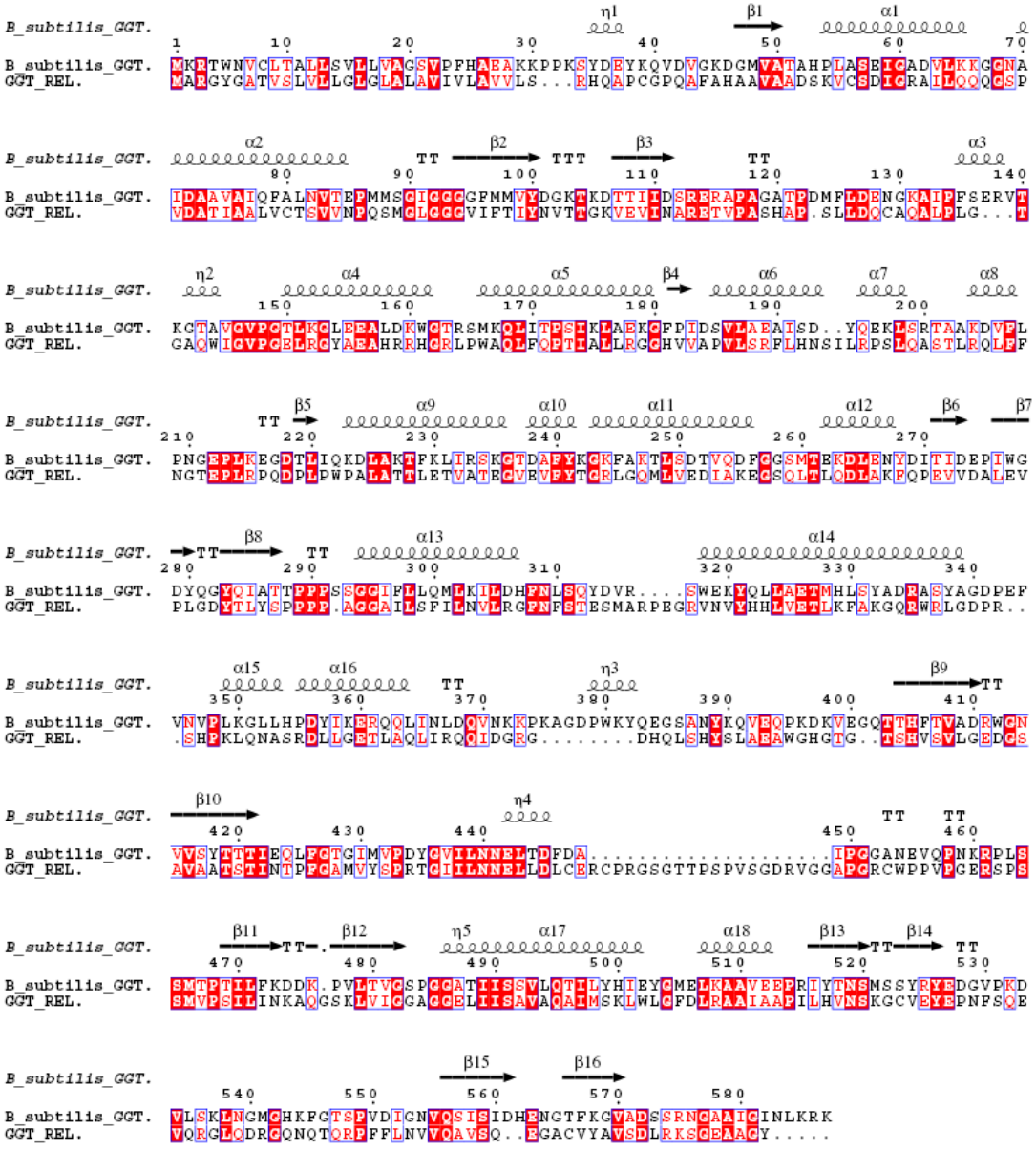


Figure 3.30: Multiple sequence alignment of *B. subtilis* GGT with GGT-rel. Secondary structure of *B. subtilis* GGT as defined by DSSP is shown.

Cloning, expression and protein purification

In the beginning, attempts were made to purify GGT by fermentation of *B. subtilis*. The procedure was tedious, requiring up to two weeks, and in the end yielded just about 100μg of pure protein from 1L of fermented broth. Therefore, *ggt* gene was cloned from the genomic DNA of *B. subtilis* and over-expressed. About 10mg of purified protein was obtained starting from 1L of fermented broth.

The *ggt* ORF was amplified with specific primers and cloned into pET 26 vector. The amplicon, devoid of the region coding for the signal peptide, was inserted between *Nco I* and *Xho I* sites. *Nco I* site occurs downstream of the region coding for *pelB* leader peptide. These arrangements substitute the native signal peptide, which is specific for gram positive bacterium, with a peptide viable in the gram negative *E. coli* cells. Due to the signal peptide, the expression product accumulates in the periplasm. Periplasmic localization was chosen for two reasons: (1) in gram negative bacterium like *E. coli,* the periplasm represents the nearest equivalent of the extracellular medium; this consideration was relevant in context of the extracellular nature of GGT in *B. subtilis*; (2) as periplasm contains little protein, osmotic shock induced permiabilization of the cell wall would be expected to liberate nearly pure enzyme. Choice of *Xho I* resulted in insertion of the ORF upstream of the region coding for hexahistidine-tag.

The cloned ORF was analysed for sequence integrity by sequencing. L302P mutation was inadvertently introduced as an artefact of PCR amplification with the non-proof-reading *Taq* polymerase. From the structure, determined later, it was evident that the locus does not participate in the formation of the active site and is well removed from the catalytic nucleophile by a distance of 32Å (Ca-Ca). The substitution occurred in a13 helix with no destabilising effect.

For expression, the construct was transformed into BL21, grown in LB broth and induced with IPTG. The solubility of the expression product was found to be temperature dependent. Soluble enzyme was formed on induction at 16°C and not at 30°C (figure 3.1). At 30°C, the expression product was composed of insoluble precursor. The pure enzyme was analysed by SDS-PAGE (figure 3.2) and MALDI-TOF (figure 3.3) and found to be a heterodimer of subunits weighing 40KDa and 20 KDa.

Kinetics of hydrolysis

The kinetics of GGT catalysed hydrolysis was studied with γ-glutamyl-(3-carboxyl)-4-nitroaniline as the substrate. The enzyme was active between pH 7.0-11.0. The kinetics was Michaelian between pH 7.0 to 9.0 and became sigmoid at pH 10.0 and 11.0 (figure 3.4). As the change occurs between pH 9.0 and 10.0, the p*K*a of transformation was assumed to be ~9.5. The sigmoid saturation curve was fitted to Hill's equation by non-linear regression and from this the Hill's coefficient (h) was determined to be 1.6-2.0. The kinetic

transformation was reversible as enzyme sample incubated at pH 11.0 for ~48 hours could produce hyperbolic saturation curve when assayed at pH 9.0.

The pH optimum for activity was dependent on the concentration of the substrate. Maximum activity was at pH 9.0 and 11.0 at lower (~2mM) and higher (~64mM) concentrations respectively (figure 3.5). K_M was found to be 17 and 26mM at pH 9.0 and 11.0 respectively. It was earlier reported to be 10.0mM at pH 9.0[1]. Also, the velocity of hydrolysis is subject to inhibition at higher concentrations of the substrate (figure 3.6). This was observed at pH 7.5, 9.0 and 11.0. The kinetics parameters mentioned above were calculated with data points that occur in the linear portion of the respective Lineweaver-Burk plot.

Effect of pH on structure

The influence of pH on the structure was determined to see if any changes accompanied the kinetic transformation. The native molecular weight was analysed by gel filtration and found to be a heterodimer in both pH 7.0 and 11.0 (figure 3.7). Changes in secondary and tertiary structure were monitored by CD spectroscopy. Both near and far UV CD spectra were same in pH 7.0 and 11.0 (figure 3.8). These results indicate that secondary, tertiary and quaternary structures are conserved between pH 7.0 and 11.0.

Thermodynamics of hydrolysis across the pH range

Thermal dependence of k_{cat} was determined in the range 20-50°C and used to prepare Arrhenius (figure 3.9 a) and Eyring (figure 3.9 b) plots. The experiments were performed at pH 7.5 and 11.0. The plots were linear in the experimental range. Activation energy (Ea) was determined from the slope of the Arrhenius plot. Ea for hydrolysis at pH 7.5 and 11.0 were 23.9 and 27.9 KJ/mol respectively. Enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) of activation were determined from the slope and *y*-axis intercepts of Erying plot respectively. $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ were 49.7 KJ/mol and -213.8 J/molK at pH 7.5 and 58.6 KJ/mol and -182.3 J/molK respectively. The activation enthalpies are positive as reactions involving bond breakage are endothermic. The negative value of $\Delta S^{\ddagger}$ indicates that the reactants must assume precise conformation and approach each other at precise angle to form to form the transiton state and that the molecules are more constrained in the activated state. Free energy of activation ($\Delta G^{\ddagger}$) was calculated from $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ at 25°C. The values were 113.4 and 112.9 KJ/mol. The positive value for $\Delta G^{\ddagger}$ are expected as the formation of transition state is non-spontaneous. Thus, hydrolysis at pH 7.5 and 11.0 are more or less thermodynamically similar except for a small improvement in $\Delta S^{\ddagger}$ at pH 11.0.

Kinetics of transpeptidation

Not much is known about the transferase activity in bacterial GGTs. Therefore, glycylglycine which serves as a good acceptor for

mammalian GGTs was used to determine the kinetics. In the transpeptidation reaction, γ-glutamyl moiety from the donor substrate is transferred to the a-amino group of the acceptor to form γ-glutamyl-glycylglycine. The direct method for the detection of this product by HPLC analysis is cumbersome for use in extensive kinetic analysis. Quantitative formation of the transfer product in *B. subtilis* GGT catalysed transpeptidation has been demonstrated previously by this method[2]. The difference between donor hydrolysis in the presence and absence of glycylglycine was taken as the measure of transpeptidation. Saturation curve for donor hydrolysis was determined in the presence of 10, 25 and 50mM glycylglycine. From this, apparent V_{max} was determined by non-linear regression and used to prepare a linear secondary plot (figure 3.10). K_M for glycylglycine was calculated from the *x*-axis intercept to be 100mM. All further kinetic analyses were performed in presence of 50mM glycylglycine. As the concentration is 0.5 x K_M, the hydrolysis is unlikely to be inhibited and the change in activity is due to additional contribution from the transpeptidation reaction. Thus the overall velocity is a sum of velocities due to hydrolysis and transpeptidation i.e., $V = V_H + V_T$; in other words, the reaction in the presence of 50mM glycylglycine is 'mixed'. Transpeptidation was therefore calculated as the difference between the velocities observed in the absence and presence of glycylglycine.

Effect of 50mM glycylglycine was determined at pH 7.5, 9.0 and 11.0 (table 3.1). Maximum transpeptidation was found to occur at

pH 9.0 as reported earlier[2]. It must be recalled that pH 9.0 is also the optimum for hydrolysis under sub-saturating conditions of the substrate. However, the behavior of K_M was different. It increased gradually from pH 7.5 to 11.0. The saturation curve showed a downward bend at higher concentrations indicating inhibition by the donor (figure 3.11). The kinetic parameters were determined by non-linear regression using data point from the linear portion (figure 3. 12). Thus, transpeptidation is occurs prominently in conditions that are unfavourable for hydrolysis. The apparent correlation might be because of inhibitory effect of hydrolysis on transferase activity.

Table 3.1: Effect of pH on the kinetic parameters of transpeptidation

pH	K_M (mM)	V_{max} (µmoles/min)
7.5	0.7±0.14	33.9±1.9
9.0	1.8 ± 0.25	66.2 ± 2.6
11	8.3 ± 1.9	39.6 ± 3.6

Comparison of donor hydrolysis in the absence and presence of glycylglycine reveals that the acceptor primarily affects the K'_M without significant impact on the V'_{max} (figure 3.13). At pH 7.5, 50mM glycylglycine causes the decrease of K'_M by about 25 fold. The effect on K'_M was pH dependent; maximum change was observed at pH 7.5 and minimum at pH 11.0. In contrast, the hydrolysis (in absence of glycylglycine) increases from pH 7.5 to

11.0 (table 3.2). The decrease in the effect of glycylglycine with increasing alkalinity could be due to inhibition by the hydrolytic reaction or due to titration of a critical group. At pH 7.5, presence of glycylglycine increased the specificity constant (k_{cat}/K_M) by nearly 40 fold.

Table 3.2: Donor hydrolysis in the presence and absence of glycylglycine (GG) as a function of pH

pH	Donor hydrolysis in absence of GG			Donor hydrolysis in presence of GG			A/H*
	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})	k'_{cat} (s^{-1})	K'_M (mM)	k'_{cat}/K'_M (s^{-1} mM^{-1})	
7.5	34.5 ±1.03	25.9±1.98	1.33	55.5± 2.18	0.97± 0.12	57.2	43
9.0	132 ± 5.4	16.7 ±1.8	7.9	171 ± 2.1	3.7 ±0.2	46.2	5.8
11.0	365.8± 26.3	26.9 ±2.7	13.6	386.4±28	24.1± 2.8	16.1	1.2

* Fold increase in K_{cat}/K_m of transpeptidation over hydrolysis

Thermodynamics of transpeptidation

Thermal dependence of K_{cat} was determined in the range 20-50°C and used to prepare Arrhenius (figure 3. 14 a) and Eyring (figure 3. 14 b) plots. The plots were linear in the experimental range. Activation energy (Ea) was determined from the slope of the Arrhenius plot. Ea for hydrolysis and transpeptidation were 23.9 and 49.1 KJ/mol respectively. Enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) of activation were determined from the slope and Y-axis intercepts of Erying plot respectively. $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ were 49.2 KJ/mol and -201.3 J/molK for hydrolysis and 107.3 KJ/mol and -190.9 J/molK respectively. Both the activation enthalpies are positive as the reactions involving bond breakage are endothermic. Free energy of activation ($\Delta G^{\ddagger}$) was calculated from $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ at 25°C. The values for hydrolysis and transpeptidation were 85 and 107 KJ/mol. The positive value for $\Delta G^{\ddagger}$ are expected as the formation of transition state is non-spontaneous. Thus, hydrolysis and transpeptidation differ primarily in the enthalpy of activation. Change in enthalpy represents bond formation and breakage and thus the mechanism of reaction. In contrast, activational entropy represents the extent of order of the activated complex in the transition state. A large negative values of $\Delta S^{\ddagger}$ (unfavorable) indicates that the activated complex has a more rigid structure than the reactants in the ground state. This is usually due to restrictions in the degrees of freedom (of translation, rotation, vibration). A positive values (or less negative values) of $\Delta S^{\ddagger}$ (favorable) indicates that the transition state is highly disordered

compared to the ground state. Translational, rotational and vibrational degrees of freedom are liberated on going from the ground state to the transition state. Thus the reaction proceeds fast. Therefore, the presence of glycylglycine appears to change the rate limiting step.

Effect of modulators

Maleate, deoxycholate and glycocholate are reported to modulate the activity of rat kidney GGT (see introduction for details). The influence of these compounds on the kinetics of *B. subtilis* GGT was examined. Inclusion of 10mM maleate, 10mM deoxycholate or 10mM glycocholate did not affect the activity. Thus the activity of *B. subtilis* GGT is not subject to modulation.

Crystallization

A suitable condition for crystallization of the purified protein was screened using with CSS I and II. The trials were setup with 100mM citrate-phosphate buffer pH 5.3 and 100mM phosphate buffer pH 7.0. A buffer with pH = 5.3 was chosen as it corresponds to the isolectric point (pI) of the *B. subtilis* GGT; proteins usually have higher predisposition to crystallise at their pI.

Presence of KSCN, a good chaeotropic salt, was invariant in all the positive conditions. The cation of the salt appears to be important as use of NaSCN did not produce any crystals. The crystals grew as very thin pentagonal plates in intricately packed clusters (figure 15 A). The size of the final crystal was highly variable and depended

more or less on the number of nuclei in the drop. Most often, a large number of nuclei were formed resulting in innumerable tiny crystals. Thus, the next strategy was to improve the condition to allow the growth of single and thicker crystals. Glycerol (10%, 20%), which is known to act as a nucleation inhibitor, did not reduce the extent of nucleation in the present case. Also, solvents (ethanol, methanol, isopropanol; 10% and 20% final) which serve as crystal contact breakers did not produce any improvement. Variation in the concentration of PEG and KSCN were tried but produced no improvement. All the crystallization trials were performed at 19°C. Changing incubation temperature to 5°C or 30°C did not affect the nature of crystallization. Also, crystal seeding was employed to obtained better crystals. Crystals were disrupted with cat-whisker and were transferred to freshly set-up drops. By this method, atleast some drops with fewer and large crystals were obtained. Either single crystals or a piece of a large crystal was used for testing. The exercise resulted in crystals that diffracted up to 3Å.

Finally thicker and larger crystals were obtained by growing in 100mM HEPES buffer pH 7.5. Reduction in the number of nuclei appears to have cause for thicker crystals. Apart from pH, the nature of the buffering compound also appears to be important for crystallization as tris.HCl and phosphate buffers failed to produce the improvements observed with HEPES. However, the crystals still grew as cluster of plates. A single plate like crystal from this condition was used to collect data (figure 15 B).

Quality of the diffraction data

A data set was collected from a single crystal up to a resolution of 1.84Å. The crystal belonged to the space group $P2_12_12_1$. The unit cell dimension was:

a = 72.26, b = 108.77, c = 161.35

The overall value of χ^2 is 1.03 which is close to the ideal value of 1.0. Also, R_{sym} is 9.8%. Usually, R_{sym} ≈5% for an excellent data set and a value up to 10% is acceptable. Furthermore, the data set should be complete by 85-90% to solve the structure. The overall completeness of the data used in this work is 97.8%. Ideally, I/σI value (signal to noise ratio) should typically be ≈20 and a value up to 10 is acceptable. In the present case, the overall I/σI has the acceptable value of 13.8. Statistics of the data are summarised in table 3.3. Furthermore, the data set has a multiplicity of 4.4. A typical data set will have an average multiplicity of ≈4. Also, R_{square} is closer to $R_{linear.}$ If R_{square} is larger than R, then the distribution of the deviations (I - <I>) is skewed implying the presence of too many outliers in the data. The Wilson plot for the data set is shown in figure 3. 16 which follows the characteristic pattern for proteins.

Table 3.3. Summary of scaling statistics

Number of reflections	
total	468030
unique	107363
Multiplicity of data set	4.4
Completeness of the data (%)	
overall(30-1.85 Å)	97.8
outer shell(1.92 -1.85 Å)	82.7
R_{sym} (%)	
overall	9.8
outer shell	43.8
$I/\sigma I$	
overall	13.7
outer shell	1.9

Quality of the final model of *B. subtilis* GGT

The final refined protein model consisted of 8374 non hydrogen protein atoms in the asymmetric unit and 349 water molecules. The crystallographic R_{cryst} and R_{free} for the refined model is 22.0% and 24.9% respectively. The overall G factor, which is considered as a measure of stereochemical quality of the model was estimated to be -0.14 by PROCHECK[3]. Ideally, scores are expected to be above -0.5. The Ramachandran plot (φ,ψ) of the final model and the statistics of the final model are given in figure 3.17 and table 3.4. The plot shows that 99.8% of the residues are in the permissible regions while

0.2% are in the disallowed regions. Glu^{423} is the exceptional residue which occupies the disallowed region. The likely origin of digression will be discussed later. The atomic coordinates and the structure factors have been deposited in the Protein Data bank with the accession code 2v36.

Table 3.4. Statistics of refinement and model validation

R_{cryst}	22.0
R_{free}	24.9
No. of atoms	
Proteins	8023
Water	349
Ramachandran Map	
Favoured	91.1%
Allowed	8.7%
Disallowed	0.2%
Average B factors	
main chain atoms	16.64
side chain and water atoms	18.37
all atoms	17.48
RMS deviations	
bonds(Å)	0.017
angles(°)	2.050

Crystal packing

The asymmetric unit contained two heterodimers (AB and CD) (figure 3.18). The heavy and light subunits were designated A or C and B or D respectively. The two molecules could be superimposed onto each other with a rmsd of 0.24Å for the main-chain atoms. Each heterodimer has a surface area of ~18990 Å. Dimerization of the

heterodimers results in burial of a total surface area of about ~540 Å. This area is too low to allow significant interactions under dilute conditions. Thus the dimer of heterodimers observed in the asymmetric unit is a crystallographic phenomenon.

Crystal structure

The two chains of the heterodimer are intimately intertwined and bury a surface area of ~12750Å. The folded molecule is reniform in cross-sectional view with approximate dimensions of 65 x 22 x 34Å. There were two heterodimers (A chain= 45KDa and B chain= 22KDa) in the asymmetric unit. Six N terminal and seven C terminal residues of chain A and three C terminal residues and $(His)_6$ tag of the B chain were invisible in both the molecules in the asymmetric unit. The main-chain could be superimposed on to that of *E. coli* [4] and *H. pylori*[5] GGTs with a rmsd of 1.36 and 1.62Å respectively. The secondary structural elements form a distinct 'aββa' tetralamellar fold, diagnostic of Ntn hydrolases (figure 3.19). The β-strands form two juxtaposed sheets in the centre, while the a-helices occupy the flanks. A wedge-shaped groove runs below the central β-β layer. Each sheet of the central β-β plate is composed of 6 antiparallel strands in the topological order: $(NH_2)_{\text{heavy chain}}$-β1, $(NH_2)_{\text{light chain}}$-β9, β10, β2, β3 and β6 in sheet A and $(COOH)_{\text{light chain}}$-β16, β15, β12, β11, β8 and β7 in sheet B. Both the N termini are located in sheet A. The C terminus of light chain arises from sheet B ends near the N terminus of the heavy chain. The C terminus of the heavy chain is located in the adjacent layer of a helices (figure 3.20). In contrast, the central

β-β plate of *E. coli* GGT is composed of 7 strands. This is due to additional strands at the N and C termini. The two sheets are dissimilar: sheet A is more or less rectangular with a conspicuous twist at the posterior end while sheet B is relatively flat and rectangular in outline with β10 as the longest strand. Of the 16 β strands, 12 are located in the central plate. The remaining four that occur away from the plate are relatively smaller. The burial of the plate in the interior of the protein appears to be consequence of its amino acid composition. The strands of the β-β plate are predominantly composed of hydrophobic (39%) and electrostatically neutral (46%) residues. Charged residues account for only 15% of the total residues.

Active site

Active site is represented by the region about Thr-403, the N-terminal residue of the light chain that functions as the catalytic nucleophile[6,7]. The residue occurs towards the mouth of depression in the ventral groove. The composition of the active site was analysed by modelling glutamic acid and glutamyl-enzyme complex with the guidance of crystal complexes obtained with *E. coli* GGT. The total binding energy involved in the docking was -44 Kcal/mol. The glutamate (end product) bound model gives a near equivalent of the Michaelis complex. In the other model, Cγ is linked to the side chain of Thr-403 by an ester link. This model represents the acyl-enzyme intermediate formed during the ping pong mechanism. The torsion angles of the two bound ligands are listed in table 3.5. In the glutamate complex,

the distance between Cγ and the side chain OH of Thr-403 is 2.7 Å which reduces to 2.4Å in the acyl complex.

Table 3.5: Torsion angles of bound substrates

Angle	Ligand	
	Glutamate	Acylated Glutamate
NH_2-Cα-Cβ-Cγ	-63	47
Cα-Cβ-Cγ-Cδ	179	-167
Cβ-Cγ-Cδ-Oε′	-149	-68
Cβ-Cγ-Cδ-Oε′′	30	-

Binding residues

In the models, the ligands were bound to the active site through their α-COOH and α-NH_2 groups (figure 3.21). The α-NH_2 interacts with Glu-423 Oε, Glu-442 Oε and Asp-445 Oδ while the α-COOH group bonds with Arg-113 Nη, Ser-464 Oγ and Ser-465 N. In human GGT, substitution of Arg-113 equivalent (Arg-107) with Gln or His resulted in complete loss of activity while substitution with Lys causes increase in K_M by about 8.4 fold[8]. Also, substitution of Asp-445 equivalent (Asp-423) with Ala increases the K_M by ~1000 fold[9]. However, replacement of Ser-464 and Ser-465 equivalents with Ala increases K_M by only 2 fold[10]. These Ser residues may therefore play a minor role in substrate biding. It is interesting to note that all the residues involved in substrate binding, except Arg-113, are from the B chain and are more or less concentrated in the N-terminal half.

Catalytic residues

Thr-403, the catalytic nucleophile emerges from the base of β9 strand and projects into the basal groove. This residue interacts with other residues through its Oγ, NH_2a and COOHa moieties (figure 3.22). The side chain Oγ residue interacts with Thr-421 Oγ by a H-bond. This interaction also occurs in *E. coli* and *H. pylori* GGTs. From the multiple sequence alignment it is evident that Thr-421 is conserved in the GGT family (figure 3.29).

The free α-COOH interacts with Arg-575 Nη (figure 3.22). Arg-B575 is in turn oriented by H-bonding between η2 and Glu-A85 Oε. α-COOH/Arg-575 interaction is absent in *E. coli* GGT due to substitution of Arg-B575 and Glu-85 by Ser-572 and His-87 respectively. In *H. pylori* GGT, Arg-575 is represent by Lys-565. However, Lys-565 does not interact with α-COOH of the catalytic Thr. α-COOH in *E. coli* and *H. pylori* GGTs are instead bonded with a water molecule. Arg-575 residue occurs near the C terminus of the light chain where the sequence is poorly conserved. Thus we are unable to ascertain its fate in the other homologues.

The free α-NH_2 function forms H-bonds with Met-88 O and a conserved water molecule (figure 3.23). In turn, the water forms H-bonds with Thr-421 Oγ and the backbone atoms of Ile-422, Glu-423, Gln-424, Met-87, and Met-88. These interactions appear to orient the electronegative oxygen centre of the water towards the α-NH_2 group. In *E. coli* and *H. pylori* GGTs, the equivalent of Met-88 O is slightly

displaced and thus forms no bond with either the coordinated water or α-NH_2.

Thr^{403} is freely suspended in the basal groove from the bottom of β9 strand. Therefore, the interactions of its Oγ, Na and Oa atom may provide mechanical stability.

The nucleophilic attack on Cγ is possible only if a stabilizing force is available to neutralise the development of partial negative charge on the carbonyl oxygen atom. The backbone N atoms of Gly-485 and Gly-486 are in suitable geometry to donate stabilising H-bonds to the carbonyl oxygen atom. The nitrogen atoms of the tandem glycines thus constitute an oxyanion hole.

Complex with γ-tetra-glutamic acid

B. subtilis GGT is known to be active on γ-poly-glutamic acid, a mucilaginous material secreted by the bacilli[11]. The nature of interaction was analysed by modelling γ-tetra-glutamic acid in the active site. The total binding energy for the interaction was -88 Kcal/mol. The molecule could be docked in five possible conformations of similar energies (figure 3.24). Overall, the ligand assumes a sigmoid form with Glu-1 being full stretched. All the conformers were identical at Glu-1, almost similar at Glu-2, slightly similar at Glu-3 and variable at Glu-4. The binding propensities of Glu-1 were similar to that of glutamic acid and γ-glutamyl ester. Glu-1 binding site is a narrow depression in the ventral groove while the

rest of the molecule, more or less occupies the groove itself (figure 3.25). The mode of binding is in agreement with the exopeptidic nature of hydrolase activity[11].

During catalysis, the isopeptide bond between Glu-1 and Glu-2 will be cleaved expelling Glu-2-Glu-3-Glu-4 tripeptide. Thus Glu-1 and Glu-2-3-4 represent acyl moiety and leaving group respectively. In all the five conformers, α-COOH of Glu-2, Glu-3 and Glu-4 are bonded to enzymic residues. Glu-2 α-COOH interacts with Thr-403 Nα and Arg-575 Nη. The group was at an average of 3.1Å from the α-COOH group of Thr-403. The potential electrostatic conflict appears to be ameliorated by guanidium of Arg-575, which interacts with both the groups. Arg-575 may therefore serve not only to stabilise Thr-403, but also to bind peptide substrates and neutralise the local electrostatic environment. The guanidium of Arg-575 is oriented by interaction with Glu-85 Oε which is additionally bonded with Thr-404 Oγ. Thus, Arg-575 is well poised and suitably stabilised to discharge its function. Glu-3 α-COOH interacts with Ala-487 N atom. This residue occurs next to the glycines that stabilise the oxyanion (Glu-1 Oγ) and yet interacts with a moiety that is well separated from the latter's partner. This is made possible by the sigmoid conformation of the substrate. Glu-4 α-COOH is bonded with Arg-525 Nη. The interactions of γ-tetra-glutamic acid are shown diagrammatically in figure 3.26. The oligomer biding propensity of *H. pylori* and *E. coli* GGTs was analysed by superposing them onto *B. subtilis* GGT complex with γ-tetra-glutamic acid. The binding appears to be

impeded by the critical Tyr of the lid loop. The critical Tyr occurs below Ala-487 equivalent thus forming a steric hindrance against the interaction with Glu3 a–COOH (figure 3.18). In addition to forming a steric barrier, hydrophobicity of the critical Tyr can increase the hindrance by its incompatibility with the charged carboxylate.

B. subtilis GGT lacks a lid loop

In mature *E. coli* GGT, the loop defined by residues 438-449 runs along the floor of the active site in such a way that Tyr-444 — which occurs at the tip of the loop — forms a hydrophobic wall near its mouth[4]. The loop thus reduces the solvent exposure of the active site and remains unchanged upon substrate binding. A similar loop occurs even in *H. pylori* GGT[5]. The lid loop is absent in *B. subtilis* GGT due to a gap corresponding to ten residues in the sequence alignment. However, the conformation of the region near the lid loop is conserved in both *E. coli* and *B. subtilis* enzymes. The missing residues shorten the loop and prevent its ingress into the active site (figure 3.27).

Incompletely withdrawn P-segment

In *E. coli* GGT, the C-terminal region (375-390) of A-chain is described as P-segment. During processing, the bond between the equivalents of Asn-402 and Thr-403 is autocatalytically cleaved by N-O acyl shift mechanism, transforming Thr-403 and Asn-402 into secondary N and C termini respectively[12]. Upon processing, the P-segment is moved away from Thr-403 by a distance of 34Å.

Subsequently, the lid loop occupies the space emptied by the withdrawn P-segment indicating a probable mutuality of the two events. Such mutuality would require retention of P-segment in *B. subtilis* GGT as it lacks a distinct lid loop.

The P-segment in *B. subtilis* GGT, corresponding to residues 375-402, forms the secondary C-terminus with the last residue of the segment (Asn-402) being linked to Thr-403 in the precursor form. The segment involves an insertion of 12 residues and consequently a large gap can be seen in the corresponding region in other GGTs. Exact conformation of the P-segment in *B. subtilis* GGT could not be ascertained due to disorder of the six terminal residues. In order to examine the possibility of P-segment serving as lid loop, we tried to build the missing residues onto the main-chain of the lid loop from *E. coli* GGT superposed with *B. subtilis* GGT. No stable conformation could be built due to steric hindrance by the bulkier side chains of Lys-396, Lys-398, Glu-400, and Gln-402. The map disorder, side chain bulkiness and random coil nature may be reflective of the segment's flexibility. It is therefore likely that the segment is drawn away from the active centre. On the contrary, the lid loop in *E. coli* GGT is pinned to the neighbourhood by a number of interactions.

Unusual geometry

The torsion angles of Glu-423 deviates from ideal geometry ($\varphi = 85°$, $\psi = -109°$) despite the good quality of the electron density and low B factor. Thus it appears that the torsion angles are natural and

unusual. The equivalent angles are likewise unusual in *E. coli* (φ = 92°, ψ = –103°) and *H. pylori* GGTs (φ = 84°, ψ = –106°). This residue appears to be important as its substitution with Ala causes insolubility of the enzyme[13]. Inspection of potential bonding pattern in the local environment reveals a H-bond between main-chain N atom of Glu-423 and Thr-421 Oγ and a water molecule. In Ntn enzymes like isoaspartyl aminopeptidases (EcAIII) and aspartylglucosaminidases (AGA), the equivalent of Thr-421 forms a H-bond with a strained residue[14]. Furthermore, Thr-421 and its equivalent in EcAIII and AGA is an important residue as its side chain interacts with the side of the catalytic Thr. Thus it was believed that the strained geometry is involved in orienting Thr-421. However, the bond appears to be missing in *E. coli* and *H. pylori* GGTs as the corresponding interatomic distances are ~2.6 Å. Thus the strained geometry may not be involved in the orientation of Thr-421 in GGT. Furthermore, Glu-423 is itself an important residue as it is involved in binding α-NH_2 of the substrate. So the special geometry might be an adaptation for optimal substrate binding.

Structural comparison with cephalosporin acylases (CA)

B. subtilis GGT was compared with class IV cephalosporin acylases (CA) like *B. halodurans* CA and *Pseudomonas sp.* strains V22 CA[15]. The crystal structure of *B. halodurans* CA has been determined at 2.7 Å (PDB code 2nlz). Interestingly, close homology could be seen between *B. halodurans* CA and GGT-like protein from *Thermoplasma acidophilum* (PDB code 2I3O). The two structures could be

superposed with rmsd of 1.47Å for 459 main chain atoms. This protein will be considered as CA in further discussion due to extensive similarity — as will be shown later — with the active site of class IV CA. This protein is a zymogen as its chain is uncleaved into heterodimer. The amino acid sequences of *B. halodurans*, *T. acidophilum* and *Pseudomonas sp.* strains V22 CAs are homologous to that of *B. subtilis* GGT by 26%, 22% and 37% respectively.

B. halodurans CA could be superposed onto 482 main-chain atoms of *B. subtilis* GGT with a rmsd of 1.6Å (figure 3.28 A). Digressions were mostly outside the Ntn fold. It could be superposed onto seryl cephalosporin acylase (PDB code: 1gk1) with a rmsd of 4.8Å for 204 main chain atoms. Similarities between *B. subtilis* GGT, and *B. halodurans* and *T. acidophilum* CAs start in their gross structures. Like GGT, *B. halodurans* and *T. acidophilum* CAs are renoid in shape and fold intimately while seryl CAs are pyramidal with two small knob-like domains[16]. The basal groove in seryl CAs is deeper while that in *B. halodurans* and *T. acidophilum* CAs is shallower like in GGT. Sequence alignment shows extensive conservation of the catalytic system. The catalytic nucleophile in Ntn hydrolases is represented by an N-terminal residue. By structural superposition (figure 3.28) and sequence alignment (figure 3.29), this residue is identified as Thr-358 in *B. halodurans* as Thr-331 in *T. acidophilum* and as Thr-368 in *Pseudomonas sp.* strain V22 CA. Thus, both the enzyme groups employ Thr as the catalytic nucleophile. Furthermore, the 2 tandem Gly residues that form the oxyanion hole in *B. subtilis*

GGT is also conserved in *B. halodurans* (Gly-440 & Gly-441), *T. acidophilum* (Gly-413 & Gly-414) and *Pseudomonas sp.* strain V22 CA. In contrast, the oxyanion hole in seryl CA is formed by NH of a main chain (Val) and a side chain (Asn)[16]. Thr-421, whose side chain forms a hydrogen bond with the side chain of catalytic nucleophile (Thr-403) is conserved in *Pseudomonas sp.* strain V22 CA but isosterically substituted with Ser in *B. halodurans* (Ser-376) and *T. acidophilum* (Ser-349) CAs. In seryl CA, the side chain of the catalytic Ser forms hydrogen bond with the main chain N atom of a His residue (His23β in 1gk1). Superposition of the structures of *B. halodurans* and *T. acidophilum* CAs on to *B. subtilis* GGT reveal that not only the critical residues are conserved but even their spatial geometries are similar (figure 3.28 B).

However considerable differences were noticeable in the binding pocket. The backbones of *B. halodurans* CA and *T. acidophilum* are similar to that of *B. subtilis* GGT near the catalytic nucleophile. This is expected due to the location of the active site at the bottom of the conserved β-β plate. So the differences are mostly in primary rather than tertiary structures. Equivalents of residues involved in substrate binding in GGT are shown in table 3.6. The variation is conspicuous in the absence of complements for α-NH_2. This is expected as CAs are specific for glutaryl amides and therefore differ from GGT substrates by the absence of α-amino group. Arg-113, which binds to α-carboxylate of the substrate, is substituted by Ser. It does not enter the active site and therefore the binding residues are exclusively from

the light chain. The region corresponding to Ser-464 Oγ and Ser-465 N is slightly translated inwards. This, along with the inward projection of Glu[442] equivalent (Arg-397 in *B. halodurans* CA and Arg-370 in *T. acidophilum* CA) appears to reduce the volume of the active site. Therefore, the angular disposition of Michaelis complex is likely to be different from that in GGT. Also, it was interesting to note the conservation of unusual torsion angles of Glu-423 equivalent: (Tyr378: φ = 61°, ψ = –98°) in *B. halodurans* CA and (Tyr-351: φ = 87°, ψ = –86°) in *T. acidophilum* CA.

Table 3.6: Comparison of amino acid sequences for binding residue equivalents

Enzyme	*B. subtilis* GGT	*Pseudomonas* V22 CA*	*B. halodurans* CA‡	*T. acidophilum* CA‡
α-Amino group binders	Glu 423 #	Asn 388	Tyr 378 #	Tyr 351#
	Asp 445	Asn 410	Asn 400	Tyr 373
	Glu 442	Tyr 407	Arg 397	Arg 370
α-Carboxylate group binders	Arg 113	Leu 90	Ser 88 ¢	Ser 78 ¢
	Ser 464	Thr 429	His 419	His 392
	Ser 465	Ser 430	Thr 420	Thr 393

* Equivalent residues were deduced from sequence alignment. ‡ Equivalent residues were deduced from structural superposition and sequence alignment. ¢ occur at locus corresponding Arg[133] in the sequence but do not enter the active site. # backbone torsion angles of these residues are unusual.

Comparison with aberrant GGTs

The primary structure of *B. subtilis* GGT was compared with that of CapD, GGL, GGT-rel and GGT-rb by multiple sequence alignment. *B. subtilis* GGT was homologous to Cap D (30%), GGL (21%), GGT-rel (25%) and GGT-rb (21%). Sequence alignment (figure 3.29 & 3.30) also reveals that the critical residues that form the catalytic machinery in *B. subtilis* GGT are conserved in the four aberrant enzymes. The conserved elements include the catalytic nucleophile, oxyanion hole and the nucleophile stabiliser (table 3.7).

Furthermore, in all the four enzymes, residues involved in binding a amino and a carboxylate groups of the substrate are conserved (table 3.8). However, in CapD, Glu-423 is substituted by Ser-372. It must be recalled that Glu-423 appears to be important for the binding of a amino group as its backbone torsion angles are unusually strained. The tandem Ser residues that interact with a carboxylate group are substituted by Thr-409 and Phe-410 in CapD and the second Ser by Phe-521 in GGT-rb enzyme. These substitutions might be unimportant as mutagenesis of these Ser residues in human GGT increases the K_M only by a small margin. Arg-113, which is crucial for binding of a amino group is conserved in all the four enzymes. The potential impact of the bulkier side chains of Phe-410 in CapD and Phe-521 in GGT-rb on substrate binding was analysed by substituting Ser-465 in *B. subtilis* GGT with Phe by *in silico* method (COOT). Arg-113, which is crucial for binding of a amino group is conserved in all the four enzymes. The hydrophobic side chain orients away from the γ-

glutamyl moiety binding site, thus imposing no steric conflict. However, the major differences occur in the lid loop. As in *B. subtilis* GGT, the loop is truncated in CapD and GGL. In GGT-rel, the loop appears to be elongated due to insertion with about 10 residues (figures 3.29 and 3.27). The loop in *E. coli* GGT almost approaches the lip of the active site. Therefore, insertion of additional residues would result in complete blockage of the entrance to the active site. It is therefore likely that the lid loop equivalent in GGT-rel assumes a different conformation. The lid loop appears to be present in GGT-rb, but lacks a distinct aromatic residue.

Table 3.7: Catalytic residues of aberrant GGTs deduced from multiple sequence alignment

Enzyme	Nucleophile	Nucleophile	Oxyanion
CapD	Thr^{352}	Thr^{370}	Gly^{429}/Gly^{430}
GGL	Thr^{389}	Thr^{407}	Gly^{476}/Gly^{477}
GGT-rel	Thr^{388}	Thr^{406}	Gly^{491}/Gly^{492}
GGT-rb	Thr^{451}	Ser^{470}	Gly^{544}/Ala^{545}

Table 3.8: Binding residues of aberrant GGTs deduced from multiple sequence alignment

Enzyme	*B. subtilis*	CapD	GGL	GGT-rel	GGT-rb
α-Amino group binders	Glu 423 #	Ser 372	Asn 409	Asn 408	Asn 472
	Asp 445	Asn 392	Asp 431	Asp 430	Asp 493
	Glu 442	Gln 389	Glu 428	Glu 427	Gln 490
α-Carboxylate	Arg 113	Arg 113	Arg 116	Arg 116	Arg 185

References

1. Minami, H., Suzuki, H., and Kumagai, H. (2003) *FEMS Microbiol Lett* 224, 169-173
2. Ogawa, Y., Hosoyama, H., Hamano, M., and Motai, H. (1991) *Agric Biol Chem* **55**, 2971-2977
3. Laskowski, R.A., McArthur, M.W., Moss, D.S., and Thornton, J. (1993) *J. Appl. Crystallogr* 26, 282-291
4. Okada, T., Suzuki, H., Wada, K., Kumagai, H., and Fukuyama, K. (2006) *Proc Natl Acad Sci USA* **103**, 6471-6476
5. Boanca, G., Sand, A., Okada, T., Suzuki, H., Kumagai, H., Fukuyama, K., and Barycki, J. J. (2007) *J Biol Chem* **282**, 534 – 541

6. Inoue, M., Hiratake, J., Suzuki, H., Kumagai, H., Sakata, K. (2000) *Biochemistry* **39**, 7764-7771

7. Castonguay, R., Halim, D., Morin, M., et al. (2007) *Biochemistry* **46**, 12253-12262

8. Ikeda, Y., Fujii, J., and Taniguchi, N. (1993) *J Biol Chem* **268**, 3980-3985

9. Ikeda, Y., Fujii, J., Taniguchi, N., and Meister, A. (1995) *J Biol Chem* **27**0, 12471-12475

10. Ikeda, Y., Fujii, J., Anderson, M.E., Taniguchi, N., and Mesiter, A. (1995), *J Biol Chem* **270**, 22223-22228

11. Kimura, K., Tran, L. S.P., Uchida, I., and Itoh, Y. (2004) *Microbiology*. **150**, 4115-4123

12. Suzuki,H., and Kumagai, H. (2002) *J Biol Chem* **277**, 43536-43543

13. Morrow, A.L., Williams, K., Sand, A., Boanca, G., and Barycki, J. J. (2007) *Biochemistry* **46**, 13407-13414

14. Michalska, K., Brzesinski, K., and Jaskolski, M. (2005) *J Biol Chem* **280**, 28484-28491

15. Ishiye, M. And Niwa, M. (2000) *Biochimica et biophysica acta* **1132**, 233-239

16. Kim, Y., Yoon, K-H., Khang, Y., et al. *Structure* (2000) **8**,1059-1068

DISCUSSION

Chapter 4

Origins of sigmoid kinetics

The pH dependent transformation of kinetics of donor hydrolysis from Michealian to sigmoid form is an interesting feature. Previously, it was noticed in the homologue from human[1] and rabbit liver[2]. The kinetics of most enzymes follow hyperbolic pattern and thus obey Michaelis-Menten equation. When more than one active species with discreet affinity for the substrate occur in the solution, the hyperbolic curve merge and manifest as a sigmoid curve. Sigmoid curve is typically a feature of enzymes exhibiting co-operative binding[3]. Some enzymes have more than one binding site whose affinity to the substrate is not uniform. Sometimes, binding of a substrate at one site may improve the affinity of the other site leading to co-operative binding. It is also possible that the first substrate might inhibit the binding at the second site. The phenomenon where binding of a ligand is affected by another molecule of its own species is called homotropic co-operativity. In the other form, called the heterotropic co-operativity, the modulatory effect is due to a ligand other than the substrate. The structure presented in this thesis was determined from a crystal grown at pH 7.5. Thus, the structure represents the form which exhibits typical Michelian kinetics. As N-terminal Thr of the B chain forms the catalytic nucleophile, there is only one active site in this form. Structural comparison of samples in pH 7.0 and 11.0 indicates conservation at both secondary and tertiary levels. Co-operative behaviour is mostly seen in oligomeric enzymes and there are indications that *H. pylori* GGT forms a dimer of heterodimer in solution[4]. Such an oliogmerization was not evident in *B. subtilis* GGT

either at pH 7.5 or 11.0. Therefore, the kinetic co-operativity is not due to interaction between multiple active sites. Kinetic co-operativity in monomeric enzyme has been noticed in wheat germ hexokinase type L_1. The co-operativity is this case is considered to be due to mnemonical mechanism. According to this mechanism, the enzyme exists in two interconvertible conformations with discreet affinity for the substrate[5].

Uniqueness of transferase activity

Inclusion of 50mM glycylglycine in *B. subtilis* GGT catalysed hydrolysis decreased the K_M significantly with minor effect on V_{max} (table 4.1). From the figure 3.10 it can seen that V_{max} is just over 4 fold even when glycylglycine is at 200mM (=2K_M) glycylglycine. Thus, acceptor primarily affects the apparent affinity and not the velocity. The asymmetric change in K_M and V_{max} increases the specificity constant (K_{cat}/ K_M). This is in contrast to the changes observed in case of mammalian homologues. Glycylglycine induces increase in both K_M and V_{max} of GGTs from human[6], hog[7], and rat[8] tissues. Transpeptidation in bovine GGTs is an unusual case; K_M decreases at lower concentrations of glycylglycine, reaches a minimum and thereafter increases[9]. It is thus evident that transpeptidation in *B. subtilis* and mammalian GGTs are fundamentally different (table 4.1).

Table 4.1: Comparison of changes induced in the kinetics of donor hydrolysis in the presence and absence of acceptor.

GGT	Absence of acceptor			Presence of acceptor		
	V_{max}	K_M	V_{max}/K_M	V_{max}	K_M	V_{max}/K_M
B. subtilis	25.9	25900	1×10^{-3}	41.7	950	40×10^{-3}
Human[6]	4.3	7.2	0.6	800	1000	0.8

V_{max} and K_M are expressed in μmol/min/mg and μM respectively

The kinetics of other bacterial GGTs has not been explored in detail. Reduction of K_M from 65 to 35 (μM) has been noticed in case of *E. coli* GGT in presence of 60mM glycylglycine[10]. These findings warrant exercise of caution while extrapolating structural details from bacterial GGTs to human GGT, where crystallographic studies have been unsuccessful.

Thermodynamically, presence of glycylglycine induces nearly two fold change in activation energy and enthalpy of activation. Thus the acceptor seems to affect the mechanism of donor hydrolysis. In the two step reaction mechanism by which GGT catalyses hydrolysis, deacylation is the rate limiting step. In case of rat kidney GGT, glycylglycine modulates the reaction by promoting deacylation so

that in presence of saturating concentrations acceptor, acylation becomes the rate limiting step[10].

Significance of transpeptidation in bacterial GGTs

Transpeptidation has been extensively studied in mammalian GGTs, as it is considered to represent the physiological reaction of the enzyme. Nothing is known about the *in vivo* role of this reaction in bacterial GGTs. However, there are some indications that transferase activity of CapD might be crucial for anchorage of capsule to the cell wall[12]. It would be difficult to generalize the importance for bacterial GGTs as CapD is not a typical GGT. Some authors hold that high K_M value for acceptors reflect the unimportance of transferase activity in bacterial homologues. However, the compounds that have been tested as acceptors are mostly those that are competent with mammalian GGTs. These compounds are either proteinaceous amino acids or dipeptides likes methionine or glycylglycine. It might be possible that the true acceptor may not necessarily resemble these compounds. For instance, the putative acceptor of CapD (diaminopimelic acid) is not a proteinaceous amino acid[13].

Nucleophile generation

Activation of Thr-403Oγ into a nucleophile (–OH □ $–O^-$) requires the mediation of a deprotonating agent. As is case in most Ntn hydrolases, the only competent moiety within the relevant geometric sphere is the α-NH_2 of the nucleophilic residue itself. Therefore, it is reasonable to consider that the moiety functions as a general base

catalyst[14]. However, competency of a-NH_2 to serve as a general base catalyst has been questioned as the group is a weak abstractor of proton from the nucleophilic Oγ and forces a strained geometry on the intramolecular H-bond[15]. Therefore, the hydrogen bond network in the vicinity of Thr^{403} was examined to see if there were any additional forces that could augment the process of nucleophile generation. Our analysis revealed interactions that could potentially stabilise the nucleophile and the a-amino group and thus contributes to the development of nucleophile at Thr-403Oγ.

An external mediator can potentially ameliorate the geometric strain involved in deprotonating the nucleophile. This has lead to the speculation of Thr-403 Oγ…Thr-421 Oγ contact as a conduit for transfer of proton from Oγ to a-NH_2 of Thr-403 [16]. However, in our model Thr-421 Oγ is at 2.7Å and 3.4 Å distance from the Oγ and Na of Thr-403 respectively. Thus, the interaction between Thr-421 Oγ and Thr-403 Na may not be significant. We believe that the Thr-403 Oγ…Thr-421 Oγ bond serves to counterbalance the unstable negative charge in the nucleophile. The stabilization would be expected to lower the p*K*a of the threnoyl side chain and thus favour its deprotonation. Modulation may be necessary in GGT as the threonyl alcohol is recalcitrant to ionize under physiological conditions (p*K*a ≈14). Thr-421 Oγ might therefore function as a general acid catalyst.

The a-amino group of Thr-403 interacts with the back bone carbonyl of Met-88 and a conserved water molecule which is held in place by a number of hydrogen bonds with the protein. These two interactions are potentially capable of stabilising the protonated a-amino group. Favourability for the formation of the charged form would improve the moiety's capacity to function as a general base. Two possible fates can befall the coordinated water: it can pick up the proton from aNH_2 and diffuse away, thereby regenerating aNH_2 or pass on the proton to the leaving group. The latter possibility appears to be the likely as kinetic studies indicate protonation of leaving group nitrogen atom concomitantly with amide bond cleavage[16]. Furthermore, we cannot see any other competent group in the relevant geometric sphere. The overall mechanism of catalysis is represented in figure 4.1

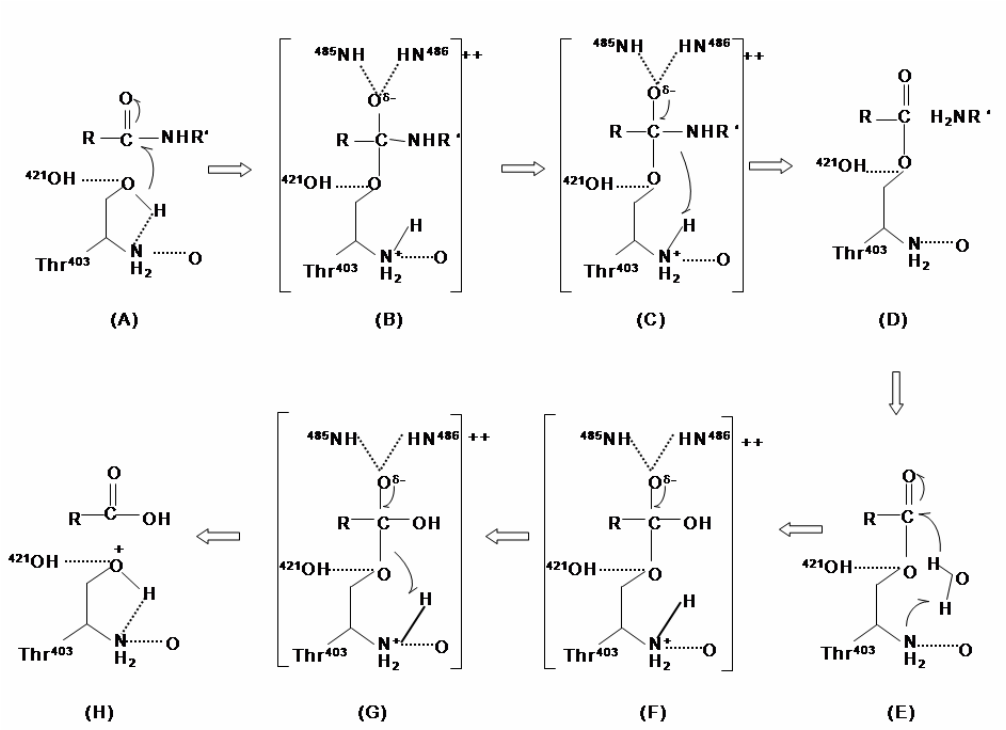


Figure 4.1: Diagrammatic representation of mechanism by which *B. subtilis* GGT hydrolyses a γ-glutamyl compound represented by the general formula RCONHR'; where R and R' represents the γ-glutamyl moiety and amine (leaving group) respectively. Top panel (A-D) represents acylation; bottom panel (E-H) represents deacylation. The side chain of Thr-403 is activated by the free a-NH_2 group. This is promoted by stabilizing interaction from the side chain of Thr-421. The basicity of a-NH_2 group is modulated by stabilizing interaction with the oxygen atoms a conserved water and Met-88 backbone. The negatively charged carbanion formed in tetrahedral transition state is stabilized by oxyanion hole formed by the backbone nitrogen atom of Gly-485 and Gly-486. Deacylation is accomplished by the reverse reaction involving a water molecule as the nucleophile.

Organization of the catalytic system

The involvement of nucleophile stabilization indicates that displacement of proton from Oγ involves not only 'pull' by aNH_2 but also 'push' from Thr-421. In catalytic triads of serine proteases, competency of the general base catalyst is maintained by the conjugate base of an acidic residue. The coordinated water and Met-88 O may be considered as the equivalent in *B. subtilis* GGT. Thus the charge interactions in GGT can be described as [acid □ nucleophile □ base/acid □ base]. The arrows denote the direction of interaction and do not imply proton relay. Even in serine proteases it is now know that the critical Asp does not withdraw proton from His but merely stabilises it[17].

The catalytic apparatus in Ntn hydrolases is described as 'single-residue nucleophile system' due to the union of nucleophile and general base catalyst in the same residue. However, analysis in terms of active components implies a 'tetrad' for GGT. Stabilising interactions of nucleophile and the a-amino group appear to be conserved in all Ntn hydrolases[18] and perhaps therefore the tetrad nature. However, the triad nomenclature arose due the involvement of three residues with discreet function. In terms of numbers, the accessory acidic and basic components imply far more than one residue in Ntn hydrolases.

Implications of truncated lid loop

The substrate affinity of *B. subtilis* (K_M = 26mM), *H. pylori* (K_M = 12.5 μM)[19], and *E. coli* (K_M = 68μM)[20]. GGTs show a large difference despite remarkable similarity in their active sites. It is compelling to associate the reduction with the absence of lid loop, as this is the only significant difference between the three enzymes. In general, the lid loop appears to be a conserved feature in the GGT family and perhaps because of this, only the enzyme from *B. subtilis* shows a deviant K_M value. Then how does the lid loop contribute to affinity?

In *E. coli* and *H. pylori* GGTs, the lid loop interacts both directly and through intermediaries with the portion of the biding pocket that interacts with the α-amino group of the γ-glutamyl moiety. The side chain of the critical Tyr interacts with the equivalent of Glu^{423}. However, this interaction may not be significant as the locus is represented by Phe in eukaryotic GGTs. The enzyme activity is significantly affected by mutations to residues that contribute to structural integrity of the loop or indirectly link it with residues critical for binding α-amino group of the substrate. However, the adverse effect is more pronounced on catalysis than on affinity[21]. Thus, it appears that the lid loop contributes non-specific support to enzyme activity by supplementing the conformational stability of the active site. Studies with rat kidney GGT indicate the participation of both α-amino and α-carboxylate groups in the binding of γ-glutamyl moiety. However, integrity of the α-amino group is invariant as its

derivatization produces significantly negative impact on activity. On the contrary, the a-carboxylate group tolerates derivatization with uncharged, isosteric or bulkier (like methyl ester, buty) groups[22, 23]. These evidences compel us to ascribe the lid loop with an indirect, non-critical and supplementary function in determining the affinity (probably for the a-amino group) of the active site without the involvement of the critical Tyr.

Substrate specificity

In contrast to γ-glutamyl-(3-carboxyl)-4-nitroaniline, *B. subtilis* GGT shows exceptionally high affinity for oligo and polymers of glutamic acid. The Michaelis constant for γ-tetra-glutamic acid and γ-poly-glutamic acid are 8 μM and 9 μM respectively[24]. The difference in kinetic affinity is in agreement with the differential binding enthalpies noticed during substrate-complex simulation (-44 Kcal/mol and -88 Kcal/mol for glutamic acid and γ-tetra-glutamic acid respectively). These results indicate potential adaptation of *B. subtilis* GGT for larger substrates. The additional bonds formed by the leaving group may compensate the reduction in the binding strength of Glu-1 site. However, the affinity for glutathione — a tripeptide which is considered as the putative physiological substrate of mammalian GGTs — is comparable to that of γ-glutamyl-(3-carboxyl)-4-nitroaniline[25]. In glutathione, the position represented by Glu-2 a-carboxylate is occupied by the cysteinyl side chain which is incapable of interacting with Arg-575 Nη and Thr-403 Na. This appears to indicate the invariance of Glu-2 a-carboxylate group in

the binding process. Due to its proximity to the scissile bond, the critical interaction might serve to present the substrate in a manner best suited for nucleophilic attack. These observations appear to indicate an oligo or polymer of gamma linked glutamic acids as the putative physiological substrate of *B. subtilis* GGT.

Functions of lid loop

Nothing is known about the role of lid loop in *E. coli* and *H. pylori* GGT. Superposition studies indicate that the critical Tyr of the lid loop might reduce the affinity for substrates like γ-tetra-glutamic acid due to steric conflict. Also, hydrophobicity of the critical residue might contribute to the conflict as γ-tetra-glutamic acid is acid rich and thus highly hydrophilic. Therefore, the lid loop, particularly the apical hydrophobic residue, appears to affect the chain length specificity of the substrate. The potential involvement of lid loop in the determination of chain length specificity and the absence of Arg-575 may imply a different substrate profile for *E. coli* and *H. pylori* GGTs. The variation might be associated with the inability of these bacteria to produce γ-poly-glutamic acid.

Cap D, gamma-glutamyl leukotrinase (GGL), GGT variant from rat brain (GGT-rb) and GGT variant from human placenta (GGL-rel) show extensive sequence homology to the form that occurs predominantly in the liver and kidneys of the mammals. They act on various γ-glutamyl substrates but not on γ-gluatmyl-4-nitroaniline, the common substrate for spectroscopic assay. Sequence comparison

indicates that the differential activity arises mostly due to variation in the lid loop and to some extent in the residues that are responsible for binding α-amino group in the regular GGTs. This highlights the importance of lid loop in determining the differential substrate specificity in the GGT family.

Evolutionary link between GGTs and CAs

Usually, the K_M of class IV CAs is lower for γ-glutamyl compounds than for glutaryl amides favouring the speculation that they might primarily be GGTs with adventitious CA activity[26]. As shown in chapter 3, the transition of GGT into class IV CA occurs by changes in the substrate binding site. The acidic residues which occur in the region involved in binding α-amino group are distinctly absent in *Pseudomonas* sp. strain V22, *B. halodurans* and *T. acidophilum* CAs. Replacement of Asp-445 equivalent in *E. coli* GGT with Asn, diminishes activity by 18% but imparts glutarylamidase activity that was hitherto absent[28]. This mutation occurs naturally in *Pseudomonas* sp. strain V22 and *B. halodurans* CAs. In *Pseudomonas* V22 CA, Arg-113 is substituted by Leu[90]. Yet, the enzyme shows γ-glutamylamidase activity. This is in agreement with the putative unimportance of α-carboxylate group of the substrate in the binding process. In *B. halodurans* and *T. acidophilum* CAs, the equivalent of amino binding site is completely cationic due to the introduction of an Arg residue. Therefore, these two enzymes appear to be incapable of supporting γ-glutamylamidase activity and might thus constitute true class IV CAs. In this light, it can be argued that class IV CAs are

distinct enzymes but are similar to GGTs in structure and catalytic apparatus because of common descent. The two groups appear to have diverged from the ancestral form by differential specialization of the binding pocket.

References

1. PetitClerc, C., Schiele, F., Bagrel, D., Mahassen, A., and Siest, G. (1980) *Clin Chem* **26**, 1688-1693
2. Bagrel, D., PetitClerc, C., Schiele, F., and Siest, G. (1981) *Biochimica et Biophysica Acta* **658**, 220-231
3. Fersht, A. (1998) *Structure and Mechanism in Protein Science,* W.H. Freeman and Company, New York
4. Boanca, G., Sand, A., and Barycki, J.J. (2006) *J Biol Chem* **258** 6193-6197
5. Rabin, B. R. (1967) *Biochem J* **102**, 22
6. Ikeda, Y., Anderson, M.E., Taniguchi, N., and Meister, A. (1995) *J Biol Chem* **270**, 22223-22228
7. London, J.W., Shaw, L.M., Fetterolf, D., and Garfinkel, D. (1976) *Biochem J* **157**, 609-617
8. Elce, J.S., and Broxmeyer, B. (1976) *Biochem J* **153**, 223-232
9. Gololobov, M., Y., and Bateman, R., C. (1994) *Biochem J* **304**, 869-876
10. Suzuki, H., Kumagai, H., and Tochikura, T. (1986) *J Bacteriol* **168**, 1325-1331

11. Boanca, G., Sand, A., and Barycki, J. J. (2006) *J Biol Chem* **281**, 19029-19037
12. Candela, T., and Fouet, A. (2005) *Mol Microbiol* **57**, 717-726
13. Candela, T., and Fouet, A. (2006) *Mol Microbiol* **60**, 1091-1098
14. Duggleby, H.J., Tolley, S.P., Hill, C.P., et al. (1995) *Nature* **373**, 264-268
15. Michalska, K., Brzesinski, K., and Jaskolski, M. (2005) *J Biol Chem* **280**, 28484-28491
16. Ménard, A., Castonguay, R., Lherbet, C. *et al* (2001) *Biochemistry* **40**, 12678-12685
17. Kossiakoff, A. A. and Spencer, S. A. (1981) *Biochemsitry* **20**, 6462-6474
18. Oinonen, C., and Rouvinen, J. (2000) *Protein Sci* **9**, 2329-2337
19. Boanca, G., Sand, A., Okada, T., Suzuki, H., Kumagai, H., Fukuyama, K., and Barycki, J.J. (2007) *J Biol Chem* **282**, 534-541
20. Suzuki,H., Kumagai, H., and Tochikura, T. (1986) *J Bacteriol* **168**, 1325-1331
21. Morrow, A.L., Williams, K., Sand, A., Boanca, G., and Barycki, J. J. (2007) *Biochemistry* **46**, 13407-13414
22. Cook, N.D., Upperton, K.P., Challis, B.C., and Peters, T. J. (1987) *Biochimica et Biophysica Acta* **914**, 240-245--22,23

23. Keillor, J.W., Castonguay, R., and Lherbet, C. (2005) *Methods in Enzymology* **401**, 449-467

24. Kimura, K., Phan Tra, L-S., Uchida, I and Itoh, Y. (2004) *Microbiology*. **150**: 4115-4123

25. Minami, H., Suzuki, H., and Kumagai, H. (2003) *Enzyme and Microbial Technology* **32**, 431-438

26. Li, Y., Chen, J., Jiang, W., Mao, X., Zhao, G., and Wang, E. (1999) *Eur J Biochem* **262**, 713-719

27. Suzuki,H., Miwa,C.,Ishihara,S., and Kumagai,H. (2004) *Appl Enviorn Microbiol* **70**, 6324-6328

SUMMARY & CONLCUSIONS

Chapter 5

The kinetics of hydrolysis is unusual as the K_M of 25mM for γ-glutamyl-(3-carboxyl)-4-nitroaniline is relatively higher than that of most other GGTs. The hydrolytic kinetics of *B. subtilis* GGT is also unusual as it transforms from typical Michaelian saturation curve into sigmoid form with a p*K*a of 9.5. The change is reversible and involves no structural modifications. The pH induced kinetic transformation greatly influences the pH profile at low and high substrate concentrations. The optimum pH at low and high concentrations of the donor substrate is 9.0 and 11.0 respectively.

Glycylglycine acts as an acceptor in transpeptidation with a K_M of 100mM. Maximum transpeptidation occurs at pH 9.0. Presence of glycylglycine decreases the apparent K_M of donor by ~25 fold without affecting the apparent V_{max}. The change in K_M results in increase of the specificity constant by ~40 fold. The change in specificity constant is pH dependent with maximum being at pH 7.5 and minimum at pH 11.0.

The tetralamellar αββa fold — a diagnostic feature of Ntn hydrolases— occurs in the core of the *B. subtilis* GGT structure. The active enzyme is a heterodimer of heavy (40KDa) and light (20KDa) chains. The two chains fold intimately, without forming any domains, into a kidney shaped molecule.

The active site occurs in floor of the ventral groove. Thr-403, whose side (Oγ) serves as the catalytic nucleophile, occurs at the base of β strand (β9). The a-NH_2 group of Thr-403 is the only base in the vicinity of the nucleophile. This is in accordance with the general property of Ntn-hydrolases wherein the free amino group of the catalytic nucleophile itself serves as the general base. The strength of a-NH_2 group appears to be improved due to hydrogen bonds with the backbone carbonyl group of Met-88 and a conserved water molecule. This water molecule is held in place by hydrogen bonds from Met-87 O, Met-88 O, Thr-421 Oγ, Ile-422 N, Glu-423 O and N, Glu-424 O. The catalytic nucleophile, Thr-403 Oγ, receives a hydrogen bond from the side chain of Thr-421. This interaction appears to promote nucleophile development by stabilizing the negative charge. The formation of negatively charged oxyanion in the transition state is stabilized by the backbone N atoms of Gly-485 and Gly-486.

The substrate binding site is represented by a depression extending below the catalytic Thr. The γ-glutamyl moiety of the substrate is bound to the active site by interactions with its a-amino and a-carboxylate groups. The acidic side chains of Glu423, Glu442 and Asp233 interacts with the a-amino group. Arg^{113} binds to the a-carboxylate group.

The binding of γ-tetra-glutamic acid follows that of glutamate at Glu-1 site. The rest molecule (Glu-2 to Glu-4), representing the leaving group, also appear to be involved in the binding process. Of these, the interaction of Glu-2 α-COOH with Thr-403 Nα and Arg-575 Nη appears to be important. *E. coli* and *H. pylori* GGTs appear to be restricted from accepting γ-tetra-glutamic acid by the steric bulk and hydrophobicity of the lid loop Tyr.

B. subtilis GGT lacks a distinct lid loop. This is the only major difference between the active sites of *B. subtilis* GGT and its homologue from *E. coli* and *H. pylori*. Therefore, this is speculated to be the apparent source for the aberrantly high K_M of *B. subtilis* GGT.

The torsion angles of Glu-423 are unusual and this appears to be an adaption to facilitate substrate binding.

GGTs and class IV CAs show extensive structural homology. Also, nature of catalytic machinery is shared between these enzymes. However, extensive differences are evident in the binding pocket apparently coinciding with their differential substrate profile. Thus both GGTs and class IV CAs appear to be distinct enzymes emanating from a common ancestral stock.

Some GGTs (aberrant GGTs) are inactive on γ-glutamyl-4-nitroaniline, the common chromogenic substrate for assaying GGT activity.

However, they retain the ability to hydrolyse other larger γ-glutamyl substrates. Comparative analysis reveals the difference in the substrate profile arises mostly due to variation in the binding pocket and the lid loop. The catalytic machinery of these enzymes is more or less similar to that of normal GGTs.

NOVEL FINDINGS OF THIS WORK

1. The kinetics of hydrolysis is subject to pH dependent transformation.

2. The optimum for hydrolysis is dependent on the concentration of the substrate.
3. The presence of acceptor reduces the apparent K_M of donor without affecting the apparent V_{max}.

4. The side chain of Thr-421 promotes development of nucleophile at Thr-403 Oγ by donating a hydrogen bond that can stabilize the formation of a negative charge.

5. Thr-403 aNH_2, the general base catalyst, is modulated by hydrogen bonding with Met-88 O and a conserved water molecule.

6. *B. subtilis* GGT appears to prefer an oligomer or polymer of γ-linked glutamic acids as substrate.

7. Lid loop is absent in *B. subtilis* GGT and this appears to be the cause for unusually high K_M.

8. GGTs and CAs share structural and catalytic details but difference in the nature of binding pocket.

9. The substrate profile of aberrant GGTs appears to originate from differences in the binding pocket and the lid loop.

APPENDIX

Materials

CapD–pQE vector and γ-poly-D-glutamate were obtained as *materia gratis* from Dr. Agnès Fouet of Pasteur Institute, Paris.

Cloning and expression of Cap D

The gift construct was used as template for PCR amplification with the primers 5' ATC GGG GCC ATG GCT TCT TTA AAT AAA ATA AAA GAC AGT3' (forward) and 5'CAG GTC CTC GAG TTT ATT TGA TTT CCA AGT TCC ATT 3' (reverse). The amplicon and pET28 plasmid were digested with *Nco*I and *Xho*I and ligated together to give CapD-pET28 construct.

For expression, CapD-pET28 construct was transformed into *E. coli* BL21 (DE3) strain. The cells were grown in autoinduction medium supplemented with 30μg/ml kanamycin at 37°C up to A_{600} = 1.0-1.5 Thereafter, incubation was continued at 16°C for about 12-16 hours.

Enzyme purification

The induced cell pellet was harvested by centrifugation at 6000 rpm and suspended in 20mM phosphate buffer pH 7.0, 300mM NaCl (buffer A) + 10mM imidazole. For cell disruption, the suspension was extruded through French press at 1500 psi. The homogenate was clarified by centrifugation at 16000 rpm for 1hr and passed through HisTrap nickel affinity column. The column was washed with buffer A till A_{280} reached the baseline. The bound protein was eluted with a

linear gradient (20-500mM) of imidazole (pH7.0) in buffer A. Fraction containing CapD were pooled, treated with solid ammonium sulphate to a final of 10% and passed through phenyl sepharose column equilibrated with 10% ammonium sulphate in 50mM sod phosphate buffer pH 7.0 and 150mM NaCl. The flow through was collected as fractions to obtain pure CapD. Purity was established by SDS-PAGE (figure A.1).

Structural Analysis

Analytical Ultracentrifugation was done at the Molecular Interactions lab, Dept. of Biology, University of York, UK. N-terminal sequencing was performed at Astbury Centre, University of Leeds, UK. Mass analysis was done by MALDI-TOF method.

Enzyme assay

5μg of γ-poly-D-glutamate and 1μg of CapD preparation were incubated for 5min at 37°C. After that, the samples were inactivated by heating at 100°C for 5 min. Heat-inactivated CapD was used as control. Samples were run on 1% agarose gel and visualized by staining with 0.5% methylene blue acidified with 3% acetic acid.

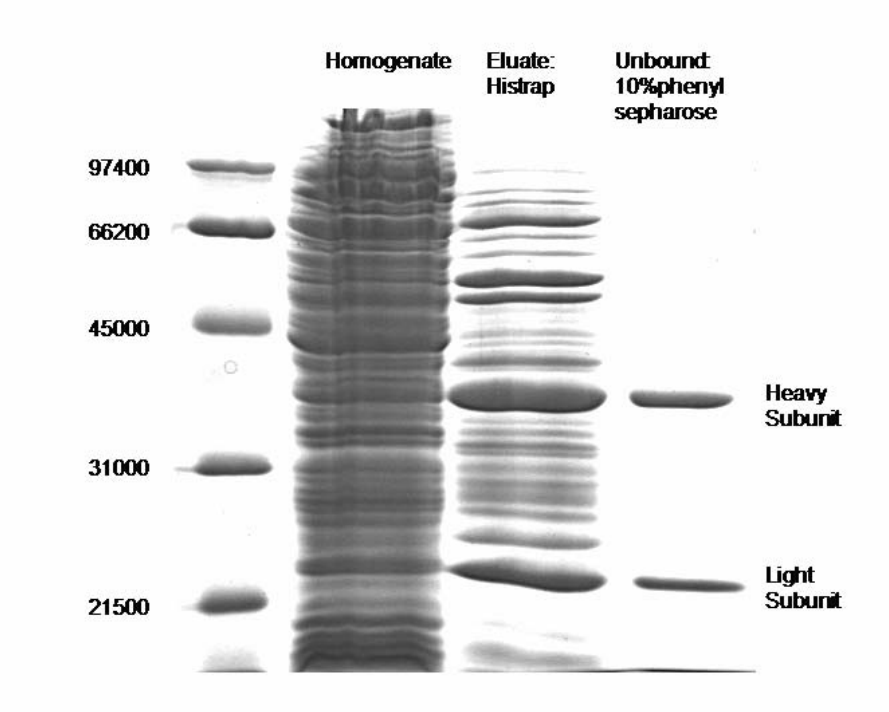


Figure A.1: SDS-PAGE gel showing various stages in the purification of Cap D.

\

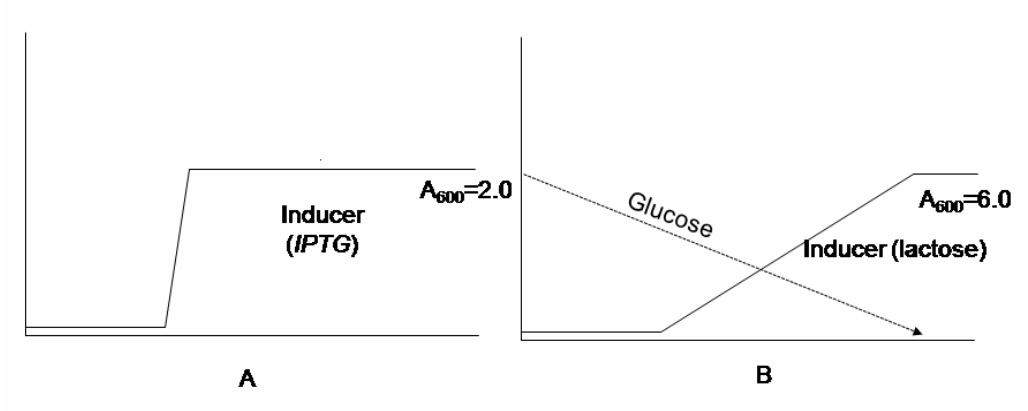


Figure A.2: Diagrammatic representation expression (Y-axis) as a function of time (X-axis) in LB and autoinduction medium. Solid line represents the magnitude of gene expression. Note that the expression occurs as a “burst” in case of LB medium (A) but is gradual in case of autoinduction medium (B). The discontinuous line in (B) represents the gradual depletion of glucose concomitant with gradual induction.

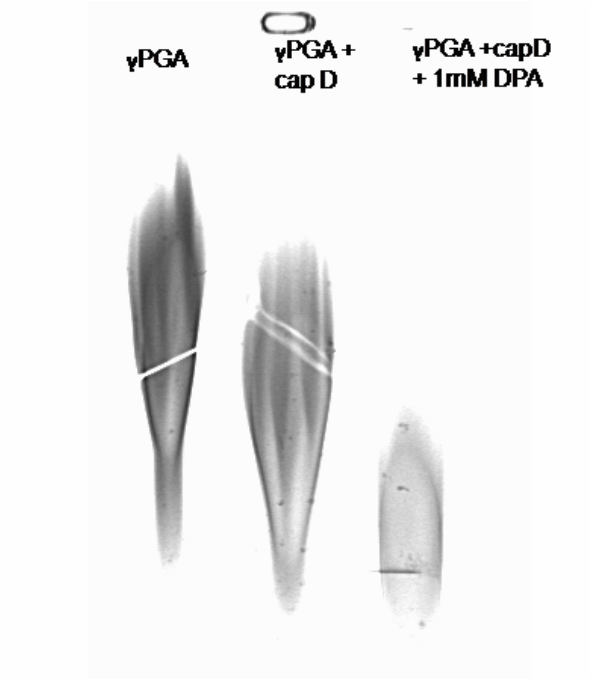


Figure A.3: Activity test for cap D. Enzyme sample was incubated with γ-poly-D-glutamic acid and the extent of digestion was analyzed by running on agarose gel. Notice activity enhancement in the presence of diaminopimelic acid

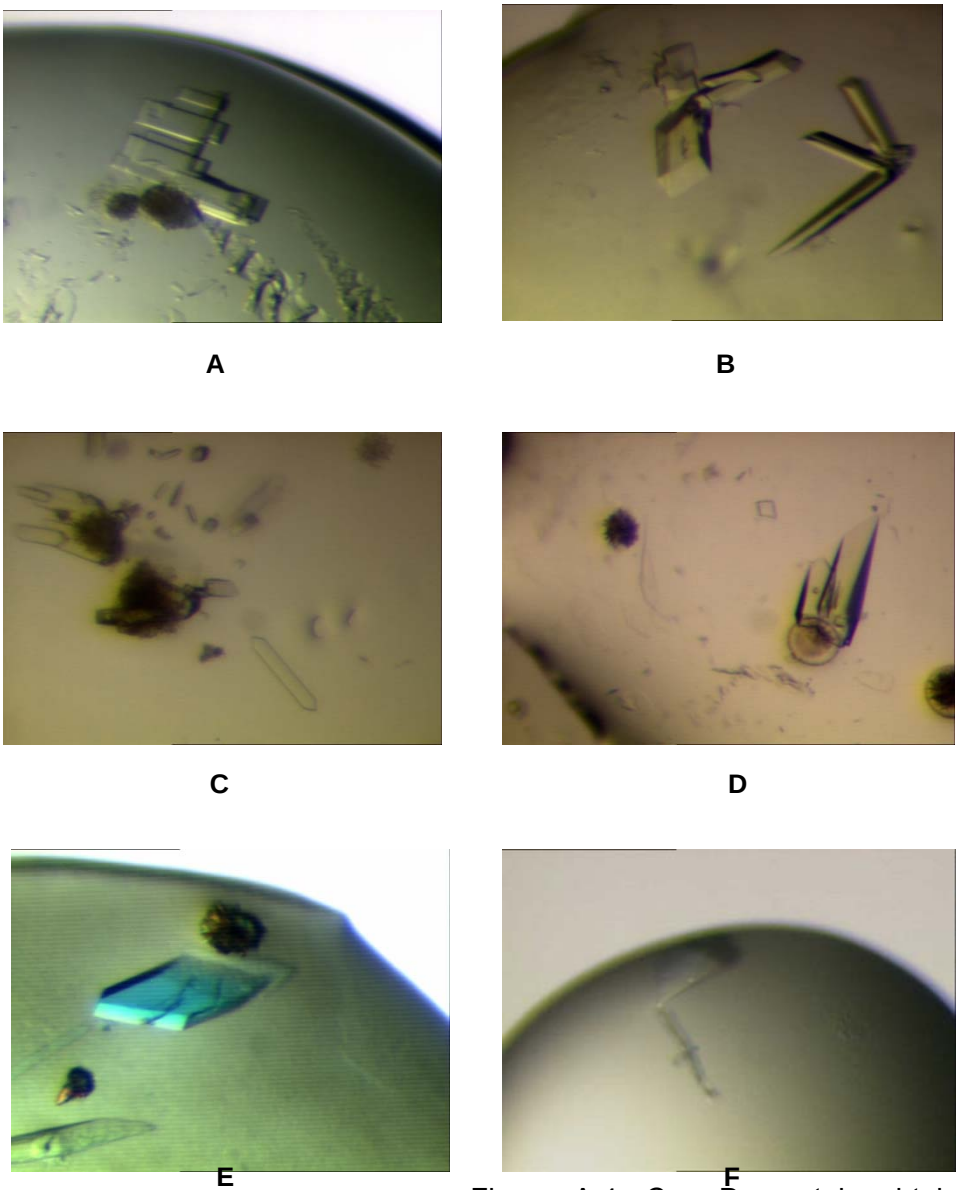


Figure A.4: Cap D crystals obtained from screen condition Index 1(A), Index 2(B), Index 3(C), Index 4(D), Index 5(E), Peg ion 1(F) and PEG Ion 2(G). See text for composition of the conditions.

Results and Discussion

The common method of induction involving addition of IPTG to a mid log culture raised in LB broth was a failure as the expression product was insoluble. The enzyme could bind to nickel column only in the presence of 8M urea. However, the enzyme could be obtained in soluble form when grown in autoinduction medium[1]. This is a special medium containing glucose (1%) and lactose (1%) in addition to other nutritive compounds. During the early phase of growth, the cells utilize glucose as the source of carbon and the utilization of lactose is prevented by catabolite repression. The repression is gradually ameliorated by the receding levels of glucose. This allows consumption of lactose. Once inside the bacterial cell, lactose also serves as an inducer for the T_7 promoter which regulates the expression of CapD-pET28. Thus induction in autoinduction medium is gradual (figure A.2). In contrast, the induction due to IPTG reaches full strength within few minutes of addition. The high rate of expression supercedes the rate of folding and maturation which is particularly extensive in an Ntn hydrolase. Accumulation of unfolded proteins aggregate may form insoluble material. The effect of the bottle neck appears to have been overcome by lowering the rate of expression in autoinduction medium.

The pure enzyme was a heterodimer of two chains weighing 36 KDa and 20 KDa. From analytical ultracentrifugation it was ascertained that the enzyme as occurs as a 56 KDa monomer in solution. The N

terminal sequences of the heavy and light chains were: Ala-Ser-Phe-Asn-Lys-Ile and Thr-Thr-His-Phe-Val-Ile. By comparison with the amino acid sequence deduced from the cloned DNA, the N-terminal Thr of the light chain can be identified as Thr-352. Thus, chain cleavage that accompanies maturation of GGTs occurs between Thr-352 and Ser-351. A conserved Thr occurring at the N-terminus of the light chain functions as the catalytic nucleophile in GGTs. Therefore, Thr-352 appears to be catalytic nucleophile in Cap D.

The activity of the enzyme was determined by incubating with γ-poly-D-glutamic acid. This resulted in increase in the relative mobility of the polymer on 1% agarose indicating shortening of the polymer. Inclusion of 1mM diaminopimelic acid in the reaction mixture resulted in further increase in relative mobility (figure A.3). Thus, diaminopimelic acid appears to enhance the hydrolysis of γ-poly-D-glutamic acid. Transpeptidation reaction catalysed by GGTs usually involves enhancement in the hydrolytic rate. However, the extent of improvement in case of bacterial GGTs is marginal when compared to the mammalian homologues. Previous studies have shown that CapD mediates formation of a covalent link between γ-poly-D-glutamic acid and the peptidoglycan of the cell wall[2]. Diaminopimelic acid occurs at the tip of the peptide moiety of peptidoglycan of *B. anthracis*. Diaminopimelic acid is considered to serve as an acceptor for Cap D mediated transpeptidation as it is the only moiety with a

free amino group in the peptidoglycan[3]. The rate improvement observed in the present study supports this hypothesis.

CRYSTALLIZATION

A suitable condition for obtaining CapD crystals was screened with Index, PEG Ion, Hampton, PACT and CSS (I & II) screens. CSS screen was set up with 100mM citrate phosphate buffer pH 5.3, 100mM Na phosphate buffer pH 7.0and 100mM PCB buffer pH 9.5. Hits were obtained in Index (5 conditions) and PEG ion (2 conditions) (figure A.4). The composition of these conditions were:

Index Screen

1.	-	0.1M Sod acetate pH 4.5	25% PEG 3350
2.	-	0.1M Bis-Tris pH 5.5	25% PEG 3350
3.	0.2MNaCl	0.1M Bis-Tris pH 5.5	25% PEG 3350
4.	0.2M MgCl2	0.1M Bis-Tris pH 5.5	25% PEG 3350
5.	0.2M MgCl2	0.1M HEPES pH 7.5	25% PEG 3350

PegIon Screen

1.	0.2M KBr	20% PEG3350
2.	0.2M NaCl	20% PEG3350

The favourable conditions tended to be either neutral or acidic in pH and preferred a relatively lower molecular weight PEG as precipitant. In general, the crystals grew after about 6-8 weeks of incubation. The crystal grew as intimate clusters. Some of the crystals were apparently single in external appearance but produced split reflections in the diffraction experiment. Usually, cluster formation is due to excessive crystal contact. Also, the quality of diffraction was very poor. Furthermore, crystallization could not be reproduced in all the conditions. Structural studies were abandoned due to failure to obtain

a diffraction quality crystal. The stability of the protein was analysed by dynamic light scattering. Polydispersion of fresh and 2 weeks old samples were 16.5 and 33% respectively. A preparation with polydispersion up to 25% is considered to be stable and suitable for crystallization. Thus instability appears to be the cause for the failure of crystallization experiment. Presence of clusters of amino acid residues with long side chains on surface loops generally increases the entropy of the molecule. Increased entropy is unfavourable as crystal formation depends on systematic or regular stacking of molecules. The amino acid sequence was analysed for excessively high surface entropy spots with Surface Entropy Reduction server available at University of California, Los Angeles, USA[4].

Locus	*Score*
1.ALSEAEHE	5.04
2.QK	4.75
3.QMLKLAEKKEVYK	4.49
4.EKNTAYTEIQ	3.87
5.KDKE	3.75

Previous studies have shown that substitution of residues at high entropy spots these residues with Ala can potentially improve the crystallizability of a protein. Therefore, substitution of one of the critical clusters mentioned above might facilitate crystallization of capD.

References

1. Groman, T.H., Kawasaki, E. S., Punreddy, S. R., and Osburne, M. S. (1998) *Gene* **209**, 95-103

2. Candela, T., and Fouet, A. (2005) *Mol Microbiol* 57, 717-726

3. Candela, T., and Fouet, A. (2006) *Mol Microbiol* **60**, 1091-1098

4. Goldschmidt, L., Cooper, D, R., Derewenda, Z, S., and Eisenberg, D. (2007) *Protein Sci* 16, 1569-1576.