STRUCTURAL INVESTIGATIONS OF SELECTED MICROBIAL PENICILLIN G ACYLASES AND A PROTOZOAL N-MYRISTOYLTRANSFERASE

Thesis Submitted to University of Pune

For the degree of

DOCTOR OF PHILOSOPHY

In

BIOTECHNOLOGY

By

NISHANT KUMAR VARSHNEY

Under the guidance of

Dr. C. G. SURESH

DIVISION OF BIOCHEMICAL SCIENCES

CSIR-NATIONAL CHEMICAL LABORATORY

PUNE – 411008, INDIA

MAY 2013

"Structural investigations of selected microbial penicillin G acylases and a protozoal N-myristoyltransferase"

Thesis Submitted to University of Pune

For the degree of

DOCTOR OF PHILOSOPHY

In

BIOTECHNOLOGY

By

NISHANT KUMAR VARSHNEY

Under the guidance of

Dr. C. G. SURESH

DIVISION OF BIOCHEMICAL SCIENCES

CSIR-NATIONAL CHEMICAL LABORATORY

PUNE – 411008, INDIA

MAY 2013

Dedicated to.........

My Mom, Abha, Pranshi

& the ALMIGHTY

CERTIFICATE

Certified that the work incorporated in this thesis entitled, **"Structural investigations of selected microbial penicillin G acylases and a protozoal N-myristoyltransferase"** submitted by **Mr. Nishant Kumar Varshney** was carried out by the candidate under my supervision. The data obtained from other sources has been duly acknowledged in the thesis.

Date: (Dr. C. G. Suresh)

Pune   Research Guide

DECLARATION OF THE CANDIDATE

I hereby declare that the thesis entitled "**Structural investigations of selected microbial penicillin G acylases and a protozoal N-myristoyltransferase**" is entirely original and was carried out by me under the supervision of Dr. C. G. Suresh at the Division of Biochemical Sciences, National Chemical Laboratory (NCL), Pune, India. The thesis is submitted by me to the University of Pune for the degree of Doctor of Philosophy in Biotechnology.

I further declares that the scientific contents of this thesis has not formed the basis for award of any degree, diploma, associateship, fellowship, titles in this or any other University or other institution of Higher learning. The material obtained from other sources has been duly acknowledged in the thesis.

Nishant Kumar Varshney   Date:

Biochemical Sciences Division,

CSIR-National Chemical Laboratory,

Dr. Homi Bhabha Road,

Pune- 411008. India

CONTENTS

i

LIST OF ABBREVIATIONS

ΔN	:	N-terminal deletion mutant
6-APA	:	6- aminopenicillanic acid
7-ACA	:	7-amino cephalosporanic acid
7-ADCA	:	7-amino 3-deacetoxy cephalosporanic acid
Å	:	Angstrom
*Af*PGA	:	*Alcaligenes faecalis* Penicillin G Acylase
AGA	:	Aspartylglucosaminidase
AIM	:	AutoInduction Medium
AmoRe	:	Automated Molecular Replacement
ANS	:	8-anilino-1-naphthalene sulfonic acid
ATCC	:	American Type Culture Collection
BLAST	:	Basic Local Alignment Search Tool
BSH	:	Bile Salt Hydrolase
CAs	:	Cephalosporin acylases
CCD	:	Charge-Coupled Device
CCP4	:	Collaborative Computational Project No. 4
conc.	:	Concentrated
CPB	:	Citrate Phosphate buffer
CV	:	Column Volume
DLS	:	Dynamic Light Scattering
DLS	:	Dynamic light scattering
DMSO	:	Dimethyl Sulfoxide
DNA	:	Deoxyribonucleic Acid
dNTPs	:	deoxynucleotide triphosphates
DTT	:	Dithiothreitol
E. coli	:	*Escherichia coli*
*Ec*PGA	:	*Escherichia coli* penicillin G acylase
EDTA	:	Ethylenediaminetetraacetic acid
F(hkl)	:	structure factor
g, mg, µg	:	gram, milligram, microgram
GeV	:	Giga Electron Volts
h	:	hour
HCl	:	Hydrochloric acid
HEPES	:	N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid
IC_{50}	**:**	Inhibitor concentration at 50% inhibition
IMAC	:	Immobilized metal ion affinity chromatography
IPTG	:	Isopropyl-β-D-thiogalactopyranoside
IU	:	International Unit
K	:	Kelvin
k_{cat}	:	Turn over number
*Kc*PGA	:	*Kluyvera citrophila* penicillin G acylase
kDa	:	kilo Dalton
K_i	:	Inhibition constant

K_m	:	Michaelis-Menten constant
L, mL, μL, nl	:	liter, milliliter, microliter, nanoliter
LB	:	Luria-Bertani
*Ld*NMT	:	*Leishmania donovani* NMT
LIC	:	Ligation-Independent Cloning
LLG	:	Log Likelihood Gain
M, mM, μM	:	molar, millimolar, micromolar
min	:	Minute
MALDI-TOF-TOF	:	Matrix assisted laser desorption ionization-Time of flight- Time of flight
MOPS	:	3-(N-morpholino) propanesulfonic acid
MPD	:	2-Methyl-2, 4-pentanediol
MR	:	molecular replacement
N_2	:	Nitrogen
NCS	:	non-crystallographic symmetry
NHM	:	Non hydrolysable myristoyl CoA analogue
nm	:	Nano metre
NMR	:	Nuclear magnetic resonance
NMT	:	Myristoyl-CoA:protein N-myristoyltransferase
Ntn	:	N-terminal nucleophile
ºC	:	degree centigrade
OD	:	Optical density/Absorbance
ORF	:	Open Reading Frame
PAA	:	Phenyl Acetic Acid
pac/PAC	:	Penicillin G Acylase gene
PAGE	:	Poly Acrylamide Gel Electrophoresis
PAs	:	Penicillin acylases
PCR	:	Polymerized Chain Reaction
PDAB	:	p-dimethylaminobenzaldehyde
PDB	:	Protein Data Bank
PEG	:	Polyethylene Glycol
PGA	:	Penicillin G acylase
PMSF	:	Phenylmethanesulfonyl fluoride
*Pr*PGA	:	*Providencia rettgeri* penicillin G acylase
PVA	:	Penicillin V acylase
r.m.s.d	:	Root Mean Square Deviation
rpm	:	revolutions per minute
RNA	:	Ribonucleic Acid
SDS	:	Sodium Dodecyl Sulphate
SOC	:	Super Optimal broth with Catabolite repression
Tris	:	tris-hydroxymethyl amino methane
Vm	:	Matthews number
α	:	alpha
β	:	beta
σ	:	sigma
Σ	:	summation

ABSTRACT

Penicillin acylases (penicillin amidohydrolase; **EC 3.5.1.11**) are a group of enzymes responsible for the major commercial production of 6-APA (a β-lactam nucleus), an intermediate used in the preparation of semisynthetic antibiotics. Large-scale production of semisynthetic penicillins and cephalosporins focuses on the condensation of the appropriate D-amino acid derivative with a β-lactam nucleus, which is catalyzed by penicillin acylases. In addition, penicillin acylases can also be employed in other useful biotransformation, such as peptide synthesis, removal of protecting groups and the resolution of racemic mixtures of chiral compounds.

Almost 85% of 6-APA production for the manufacture of semi-synthetic penicillins is by using penicillin G acylase (PGA), the enzyme that hydrolyses penicillin G (benzyl penicillin), with the major share of the enzyme sourced from *E. coli*. However, for pharmacological applications when factors such as tolerance for environmental parameters such as temperature, pH and nature of solvent, along with ease of immobilization are important, attention has turned to PGA enzymes from other sources such as *K. citrophila* (*Kc*PGA). Another PGA produced by *Alcaligenes faecalis* (*Af*PGA) has clear industrial advantage over other well-characterized penicillin acylases in β-lactam conversion because of its higher thermostability and synthetic efficiency in enantioselective synthesis. Such characteristics make *Af*PGA a more attractive biocatalyst for both hydrolysis and synthetic conversions.

PGAs are members of the N-terminal nucleophilic (Ntn) hydrolase superfamily of enzymes, that share a common fold of αββα- core structure possessing N-terminal catalytic nucleophile residue (Thr/Cys/Ser). The Ntn hydrolases invariably produce their active form from the corresponding precursor by an intramolecular autocatalytic cleavage to create a free amino group at the nucleophile residue.

Another enzyme relevant for therapeutics, studied in this thesis as a drug target is N-myristoyltransferase (NMT; **EC 2.3.1.97**). Studies in the past have identified Myristoyl-CoA–protein NMT as a suitable candidate for drug development against protozoan parasitic infections, including those caused by *Leishmania major*, the causative agent of

cutaneous leishmaniasis, as well as *Plasmodium falciparum* and *Trypanosoma brucei*, causative agents of human malaria and African sleeping sickness, respectively. NMT is ubiquitous in eukaryotic cells in which it catalyses the co-translational addition of the C14:0 fatty acid, myristate, via an amide-bond, to the N-terminal glycine residue of a subset of proteins. N-Myristoylation plays a role in targeting proteins to membrane locations, mediating protein–protein interactions and stabilizing protein structures.

With the aim of exploring the application potential of PGA from *Kluyvera citrophila* (*Kc*PGA) and PGA from *Alcaligenes faecalis* (*Af*PGA) in pharmaceutical industry for the production of semi-synthetic penicillins and to exploit myristoyl-CoA–protein N-myristoyltransferase from *Leishmania donovani* (*Ld*NMT) as a potential drug target against leishmaniasis, the research carried out is presented in the thesis is divided into seven chapters, the contents of which are explained below.

Chapter 1: General Introduction and Review of Literature

This chapter presents a survey of research-literature on penicillin G acylases and N-myristoyltransferase enzymes. Discussion focuses mainly on their occurrence, physical and chemical properties, gene organization, comparison of evolutionary related enzymes and the existing/suggested applications.

Chapter 2: Materials and Methods

This chapter provides information on chemicals, resources, methods and procedures used in the preparation of proteins, their purification, crystallization, X-ray data collection, data processing, structure determination, refinement, analysis of the refined structure, biochemical/biophysical experiments and kinetic studies. To facilitate purification, the clones were prepared with His-tag attached either at C-terminal or N-terminal. Standard assays were used to quantify enzyme activity, biochemical and biophysical techniques such as Fluorescence spectroscopy was used to study the stability of the processed *Kc*PGA and the monodispersity of the purified protein was assessed using Dynamic Light Scattering. Screening of crystallization conditions was carried out using

sitting-drop or hanging-drop vapour-diffusion methods. X-ray diffraction data were collected using a laboratory X-ray source: rotating anode at NCL, Pune or using synchrotron radiation either at SSRL, Stanford or at Diamond Light Source, Oxford, UK. The images were processed using programs *HKL*, *XDS* or *Xia*2. The crystal structures were determined using molecular replacement technique implemented in *MOLREP*, *PHASER* and *AutoMR*. The program *BUCCANEER* and *AutoBuild* was used to build the initial model, *REFMAC5* for structure refinement and *QUANTA* and *COOT* for model fitting.

Chapter 3: Cloning and Over-expression of *pac* gene from *Kluyvera citrophila* and Purification and Biochemical characterization of the expressed *Kc*PGA

The chapter describes different strategies used in cloning the penicillin acylase (*pac*) gene from chromosomal DNA of *Kluyvera citrophila* DMSZ 2660 (ATCC 21285), over- expression, purification and biochemical characterization of expressed wild-type enzyme and selected mutants.

K.citrophila pac gene was amplified from chromosomal DNA using primers designed from the gene sequence available in the databank. The purified and amplified gene was then ligated with pET26b (+) and expressed in the BL21 (DE3) pLysS. The enzyme was purified using affinity chromatography on Nickel beads; the resultant protein was more than 90% pure. Subsequent gel exclusion chromatography on Sephacryl S-200 / Sephadex G-200 yielded pure sample for crystallization as well as biochemical and biophysical characterization. Molecular weight determination by SDS-PAGE and mass spectroscopy techniques of Electrospray Ionization (ESI) and MALDI-TOF has confirmed the expected size 23.6 and 62.9 kDa of α and β chain corresponding to processed form of PGA. Various parameters like optimum pH and temperature for maximum activity, enzyme stability and kinetic parameters were determined. The optimum pH and temperature was found to be 7.5 and 65°C, respectively. The enzyme was stable in the pH range of 4.0 to 11.0. The enzyme was more than 75% stable at 50°C for 2hrs. Higher temperature beyond 60°C resulted in aggregation of the enzyme. The specificity towards different substrates was also tested.

The selected site-directed mutation Serβ1Gly yielded unprocessed precursor, whereas Serβ1Cys resulted in processed but catalytically inactive enzyme, useful for preparing enzyme-substrate complexes for structural studies. The effect of conditions such as pH, temperature and time on the *in vitro* processing of precursor enzyme was also studied.

Chapter 4: Structural and folding studies of wild-type enzyme and slow-processing mutant of *Kc*PGA

The structural changes accompanying denaturation by chaotropic agents-guanidium hydrochloride (Gdn-HCl), temperature and pH and resultant protein stability were monitored using fluorescence spectroscopy. Sequence showed the presence of 26 Tryptophan residues in *Kc*PGA. The steady state fluorescence spectrum had a λ_{max} of 346 nm. The λ_{max} showed a red shift (361nm) on treatment with 6M Gdn-HCl indicating exposure of tryptophans to the solvent by opening up the structure. Thermal denaturation of *Kc*PGA led to aggregation at 55 ^{o}C. Hydrophobic dye (ANS) binding at this temperature could be due to exposure of hydrophobic patches upon heating. The denaturation was found to be irreversible.

Incubation of *Kc*PGA in alkaline or acidic buffer decreased the intrinsic fluorescence intensity. ANS binding was observed only at extreme acidic pH; maximum at pH 2.0 with a blue shift in λ_{max} from 520 to 482 nm and ANS fluorescence intensity increased seven times. Surprisingly, at pH 1.0 there was six times increase in the ANS fluorescence intensity with λ_{max} 493 nm indicating possible rearrangement of structure towards native conformation. At pH 3.0 the ANS binding was only twice with λmax of 484 nm.

Solute quenching studies to investigate the microenvironment of tryptophans was carried out. The Trp residues were differentially exposed to the solvent and not fully accessible to the neutral quencher indicating heterogeneity of environment. The fluorescence quenching by acrylamide was collisional-type. The Stern–Volmer plots for quenching of native protein with iodide and cesium ions exhibited negative curvature

showing selective quenching of some of the tryptophans over others. A similar quenching pattern is observed in many multi-tryptophan proteins.

Denaturation of the protein exposes more tryptophan residues to solvent with consequent increase in quenching. The fluorescence quenching of 6M Gdn-HCl treated protein by acrylamide showed an upward curvature indicating collisional and static mechanisms of quenching, as a result of altered conformations, whereas quenching by iodide and cesium ions showed linear Stern–Volmer plots. The estimated K_{sv} values were higher in the case of both the ions (K_{sv}=6.14 M^{-1} and 1.67 M^{-1} for KI and CsCl, respectively) an indication that even buried tryptophans has become exposed. Since the charged quenchers can access only surface exposed tryptophans, the higher quenching by iodide ions is a sign of more positive charge around the surface exposed tryptophans. Thus, the environment is different for different tryptophans in *Kc*PGA; some are partially or fully exposed to solvent while others are buried in a hydrophobic environment.

Only crystals of precursor *Kc*PGA (slow processing mutant Serβ1Gly) have been successfully grown to date. X-ray diffraction data for two crystal forms of this mutant *Kc*PGA belonging to space groups P1 and C2 were collected at resolutions 2.5Å and 3.5Å, respectively. The crystal structure was determined using molecular replacement method using coordinates of *E. coli* precursor PGA (PDB code: 1E3A) as input model. The structure was refined in space group P1 using data at 2.5 Å resolution.

Chapter 5: Structural studies of PGA from *Alcaligenes faecalis*

This chapter describes the structure determination, refinement and analysis of penicillin G acylase from *Alcaligenes faecalis* (*Af*PGA). As already mentioned, the higher thermostability of *Af*PGA is a clear advantage over other well-characterized penicillin acylases in β-lactam conversions. Crystals were grown using enzyme solution of 15 mg/ml concentration in condition 15% (w/v) PEG 8,000, 0.1 M Tris-HCl pH 7.5 and (0.5 %w/v) β-octyl-glucopyranoside. Two crystal forms were obtained depending on the amount of detergent (β-octyl-glucopyranoside) used; addition of 60 µl in the well solution led to orthorhombic crystals, whereas 100 µl resulted in tetragonal crystals. The orthorhombic

data (3.3 Å resolution) could be processed in space group C222$_1$, with unit-cell parameters $a = 72.9$, $b = 86.0$, $c = 260.2$ Å. Data of tetragonal form (3.5 Å resolution) were indexed in space group P4$_1$2$_1$2, with unit-cell parameters $a = b = 85.6$, $c = 298.8$ Å. The structure was determined using molecular replacement method using co-ordinates of *E. coli* PGA (PDB code: 1GK9). Although the resolution was modest features like the presence of a disulfide bridge and calcium binding site could be identified.

Certain features responsible for thermostability of *Af*PGA could also be ascertained. A disulfide bond exists between the residues 492 and 525 of β chain, more number of salt bridges are present compared to other PGAs with known three-dimensional structures, content of uncharged polar amino acids (N+Q+S+T) is less, *Af*PGA has a larger Arg/Lys ratio and presence of more number of prolines decreases the conformational flexibility of the folded state.

Chapter 6: Cloning, Over-expression, Purification and Crystal Structure studies of N-myristoyltransferase from *Leishmania donovani*

This chapter describes Myristoyl-CoA–protein N-myristoyltransferase (NMT), a potential drug target for protozoan parasitic infections. The new construct, an N-terminal mutant with non-cleavable histidine tag, was prepared using full length construct of *LdNMT* gene cloned in modified pET28b (+) with 3C cleavage site. The construct was cloned in the similar vector using the technique of Ligation Independent Cloning (LIC cloning) lacking cleavage site and expressed using pLEMO expression cells in auto induction medium. The enzyme was purified with combination of affinity and anion exchange chromatography followed by size exclusion chromatography using HiLoad Superdex 75 16/60pg column (Amersham Biosciences) yielding 3-4 mg enzyme from one litre culture.

The protein solution (6 mg/ml) was used in the presence of non hydrolysable Myristoyl-CoA analogue (NHM) in two crystallization conditions selected from preliminary trials using commercial screens, which were further optimized to grow diffraction quality crystals. Single crystals were soaked with reported inhibitor of *T. brucei*

NMT and data sets were collected at 100 K at diamond synchrotron facility, Oxford, UK using ADSC Q315r CCD detector and 0.979 Å wavelength. Diffraction data were processed using automated processing softwares and scaled using *SCALA* programme in *CCP4* suite 6.2.0. Crystallization in two different optimized conditions: 2.0 M ammonium sulphate, 0.2 M NaCl, 0.2 M tri-sodium citrate tetrahydrate and 25% (w/v) PEG 3350, 0.1 M sodium acetate pH 4.6 resulted in two crystals forms: P1 with unit cell a = 70.08, b = 79.79, c = 92.29Å, α = 75.18°, β = 73.07°, γ = 72.68° and P2$_1$ with unit cell a = 92.58, b = 90.19, c = 121.71Å, β = 110.45°, respectively.

Structure was determined using molecular replacement method of *MOLREP* and *PHASER* using the coordinates of *L .major* NMT (PDB code: 2WSA) as search model. Initial model was refined in cycles using *REFMAC* 5.6 and manual fitting was done in *COOT*. Electron density for the bound NHM molecule could be found in the structure however soaked crystals failed to show density for the inhibitor. Co-crystallization of NMT with both substrate as well as inhibitor resulted in poorly diffracting crystals.

Chapter 7: Conclusions

The thesis provides an account of the study carried out on three pharmaceutically and medicinally important enzymes, whose action concerns the amide bond. Main objective in carrying out these investigations was to understand the structure-function relation of these enzymes that can facilitate their applications. The enzymes were cloned, expressed and their structures were studied using various biochemical and biophysical techniques. This chapter summarizes how the above studies have provided insights into several chemical, physical and activity characteristics of the enzymes including three-dimensional structures that are important for improving their application potential.

CHAPTER: 1

INTRODUCTION AND REVIEW OF LITERATURE

The account of research presented in this thesis provide information on the detailed study carried out on three pharmaceutically and medicinally important enzymes, two of them classified as penicillin G acylases, from *Alcaligenes faecalis* (*Af*PGA) and *Kluyvera citrophila* (*Kc*PGA), and the third one N-myristoyltransferase from a protozoan *Leishmania donovani*, all of their action concerns an amide bond. Thus the work in this thesis can be broadly classified into two parts. A major part of the work concerns penicillin G acylases, the enzymes used in the industrial production of semi-synthetic penicillins. The other enzyme studied in this research, Myristoyl-CoA:protein N-myristoyltransferase (NMT) from the protozoal parasite *Leishmania donovani*, is proposed to be a potential drug target for protozoan parasitic infections. Accordingly, the background of this work is presented in two parts; the first part is on PGAs and the next on NMT.

1.1. Introduction to Enzymes

Enzymes (pron.: /enzaimz/) are large biological molecules responsible for a number of chemical inter-conversions that sustains life. At any given moment, all chemical reactions in biological systems are essentially catalysed by enzymes. Enzymes are highly selective catalysts, accelerating the specificity and rate of metabolic reactions, from the digestion to the synthesis of DNA, RNA and protein. The enzymes are polypeptide chains, which after translation folds into a unique three-dimensional structure. This three-dimensional structure is essential for carrying out specific chemical reactions with or without organic (e.g., biotin) and inorganic (e.g., magnesium ion) cofactors. Understanding the enzymes indirectly provides insight into cellular function.

In 1833, French chemist Anselme Payen discovered the first enzyme, diastase. In 1877, German physiologist Wilhelm Kühne (1837–1900) first used the term *enzyme*, which comes from Greek *énsymo*, *én* meaning "at" or "in" and *symo* means "leaven" or "yeast". In 1926, James B. Sumner purified and crystallized the enzyme urease; likewise crystallized the enzyme catalase in 1937. The hypothesis that pure proteins can be enzymes was definitively proved by Northrop and Stanley, who proved the digestive

enzymes pepsin, trypsin and chymotrypsin, as proteins. These three scientists were awarded the Nobel Prize in Chemistry in 1946.

Enzymes serve a wide variety of functions inside living organisms. They are indispensable for cell regulation and signal transduction often via kinases and phosphatases; they also generate movement, with myosin utilizing ATP to generate muscle contraction and also to move cargo around the cell as part of the cytoskeleton. Enzymes are used by viruses as well for infecting cells, such as the retroviral integrase and reverse transcriptase, or for viral release from cells, like lysins and neuraminidase. Another important function of enzymes is in the digestive systems of animals. Enzymes such as proteases and amylases break down proteins or starch, respectively into smaller molecules, facilitating their absorption by the intestines. Several enzymes can work together in a specific order, creating metabolic pathways. Enzymes are used in the chemical industry, pharmaceutical industry and other industrial applications when extremely specific catalysts are required. Some enzymes are used commercially, e.g., in the synthesis of antibiotics like penicillin acylases; in fermentation industries; in biological washing powders to remove protein or fat stains on clothes; in meat industries to break down proteins into smaller molecules, making the meat easier to chew. Enzymes are indispensable as therapeutics, e.g., as proteolytic agents to remove blood clots, as clinical diagnoses of certain disease such as diabetes.

Enzyme activity is affected by temperature, pressure, chemical environment (e.g., pH), and the substrate concentration. Activity can be affected by other molecules, for example, inhibitors decreases enzyme activity whereas activators increase it. Many drugs and toxins are inhibitors. Generally enzymes are specific and reaction conditions are mild; these desirable features make the enzymatic reactions a preferred system for commercial application. Immobilization of enzymes has facilitated repeated use of costly enzymes. Enzyme catalyzed reactions are thus cheaper than their chemical counterparts and the process is more environment-friendly also.

The first crystal structure of an enzyme determined was that of lysozyme by David Phillips' group in 1965 (Blake *et al.,* 1965). Elucidation of this high-resolution

structure of an enzyme marked the beginning of the efforts to understand the chemistry and mechanism of enzyme activity at molecular level.

Part I

1.2. β-lactam Antibiotics

Modern healthcare strongly depends on antibiotics. Research in the past several years for discovering new antibiotics has yielded not only many novel compounds, but also provided insight into their mechanisms of action. Different antibiotics affect several vital processes in the prokaryotic cell. They can interfere with DNA, RNA or protein synthesis, or destroy the cell wall. Antibiotics are chemicals produced by microorganisms or fungi that act on other microorganisms. Antibiotics can kill the bacteria (bactericidal) or arrest its growth (bacteriostatic). One of the first identified and still most important classes of antibiotics are the β-lactam antibiotics, which interfere with cell wall synthesis. They account for about 65% of the worldwide antibiotic prescriptions and their total sales were estimated to be US$ 15 billion (Elander *et al.,* 2003). β-lactam antibiotics alone constitute most of the world's antibiotic sales: 3×10^7 kg/year out of a total 5×10^7 kg/year produced worldwide (de Souza *et al*., 2005). As of the date of this writing, FDA has approved over 34 beta-lactam compounds as active ingredients in drugs for human use (FDA's Approved Drug Products with Therapeutic Equivalence Evaluations, generally known as the Orange Book (http://www.accessdata.fda.gov/scripts/cder/ob/default.cfm).

1.3. History of β-lactam Antibiotics

Before the early 20th century, treatments for infections were based primarily on medicinal folklore. Mixtures with antimicrobial properties used in treatments of infections were described over 2000 years ago. Many ancient cultures, such as Egyptians and Greeks, used specially selected mold or plant materials and their extracts to treat infections. The term "antibiosis", meaning "against life", was coined by the French bacteriologist Jean Paul Vuillemin as a descriptive name of the phenomenon exhibited by these early antibacterial drugs (Foster, 1974). Antibiosis was first described in 1877 in bacteria when Louis Pasteur and Robert Koch observed that an airborne bacillus could inhibit the growth of Bacillus *anthracis* (Landsberg,

1949). These drugs were later renamed antibiotics by Selman Waksman, an American microbiologist, in 1942 (Waksman, 1947). Although Alexander Fleming had already discovered penicillin in 1928 (Fleming, 1929), it was not until the second world war that penicillin G (Pen G) was isolated, characterised and produced at a larger scale. The use of penicillin as a medicine had to wait until 1940s when Howard Florey and Ernst Chain prepared a powdery form of the antibiotic. The effectiveness of penicillin against bacterial infection had been first demonstrated in mice. Later sufficient material was prepared with great difficulty to treat a few human patients (Nayler, 1991a). It was indeed a significant finding that antibiotics could attack bacteria specifically and caused no harm to the organism that produced them.

Following the discovery of penicillins many other antibiotics have also been discovered and have made it possible to cure diseases such as pneumonia, tuberculosis, and meningitis caused by bacteria, saving the lives of millions of people around the world. Majority of the clinically used antibiotics have been obtained from actinomycetes, especially *Streptomyces* species. *Bacillus* species and fungi also have yielded few useful antibiotics. They can be classified based on their chemical structure, microbial origin, spectrum of activity or mode of action. One class of antibiotics that work by inhibiting the synthesis of peptidoglycan in bacterial cell walls is β-lactam antibiotics. They have superior inhibitory action on the bacterial cell wall synthesis, broad spectrum of antibacterial activity, low toxicity and outstanding efficacy against various bacterial strains. β-lactam antibiotics can be classified based on their core structure (nucleus) as:

1. **Penicillins:** They are also called **penams**. They can be classified into two types according to the source. Natural penicillins are penicillin G (Pen G) and penicillin V (Pen V) and the semi-synthetic penicillins e.g., amoxicillin, penicillinase-resistant cloxacillin, methicillin.

2. **Cephalosporins:** These are classified in generations I, II, III, IV and V. Along with cephamycin, they form a sub-group called **cephems**.

3. **Carbapenems:** e.g., imipenem, meropenem, entrapenem

4. **Monobactams:** e.g., aztreonam
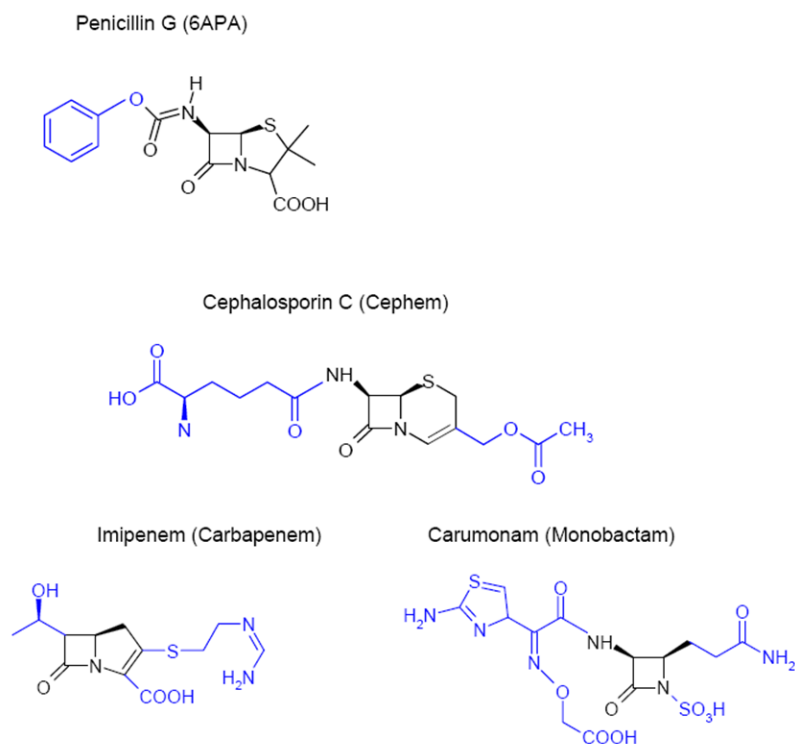
Figure 1.1: Structures of different β-lactams. The nuclei of the different antibiotics are coloured black and indicated in brackets. The side chains are in blue.

During the synthesis of bacterial cell wall, the cross linking between two peptide chains attached to polysaccharide backbone is an important step. In such cross-linking, the terminal alanine from each peptide is firstly hydrolyzed and secondly one alanine is joined to lysine through an amide bond. This cross-linking is catalysed by the enzyme transpeptidase. Structural resemblance of penicillin to D-alanyl-D-alanine portion of cell wall glycopeptide makes it potent antibiotic as it specifically inhibits linking of neighbouring subunits via transpeptidation. Penicillin irreversibly binds at the active site of the transpeptidase enzyme by mimicking the substrate D-alanyl-D-alanine. Transpeptidation reaction gives structural support to cell wall. Thus, the growth of the bacterial cell wall is effectively prevented and at the same time the antibiotic molecule destroys its stability also.

The first crystal structure of naturally occurring penicillin, Pen G, was solved by Dorothy Hodgkin and colleagues (Crowfoot *et al.,* 1949). It has also been established that several penicillins exist with different acyl side chains (RCO-) attached to a common bicyclic (penam ring) nucleus (**Figure 1.2 and 1.3**). The discovery that

the nucleus 6- aminopenicillanic acid (6-APA) is common in different penicillins has lead to the development of semi-synthetic approaches towards preparing novel penicillins starting from 6-APA (Nayler, 1991a, b).
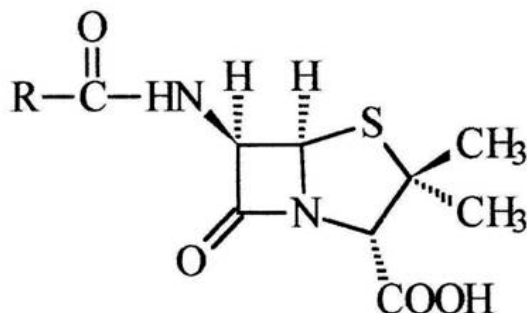
Figure 1.2: The drawing of penam ring nucleus 6-APA, a common feature found in all natural penicillins.
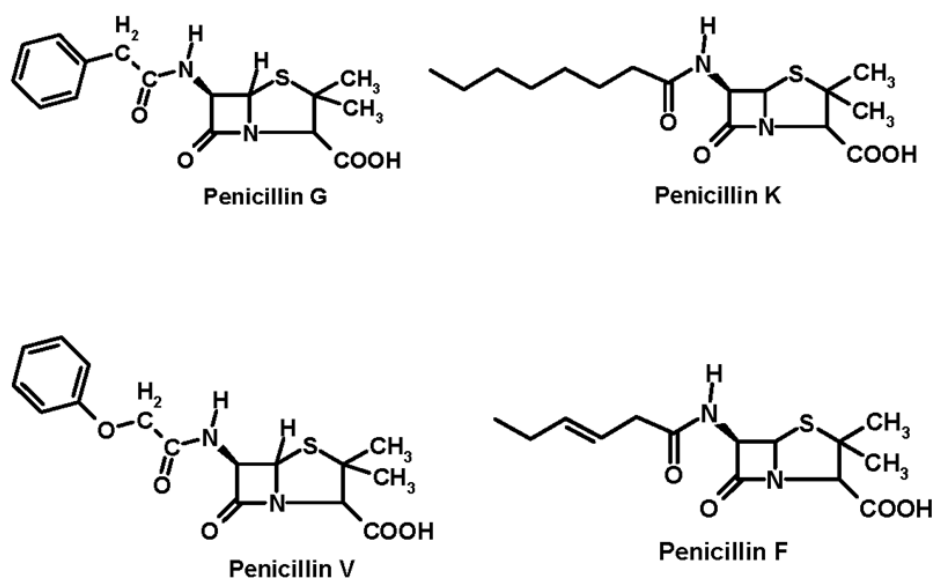
Figure 1.3: Structural drawings of some natural penicillins.

1.4. Bacterial Resistance and Need for Semi-synthetic Antibiotics

Penicillins are the most widely used β-lactam antibiotics, with a share of about 19% of the estimated worldwide antibiotic market (Parmar *et al.*, 2000). Unfortunately, since this wonder drug killed a lot of different bacteria, it was used almost unlimited although effects on the natural equilibrium of micro organisms had not been properly investigated. Research revealed despite the effectiveness of penicillin in curing a wide

range of diseases, infections caused by certain strains of *Staphylococci* cannot be cured because the latter produces an enzyme, penicillinase (β-lactamase), capable of destroying the antibiotic (Jacoby & Munoz-Price, 2005) (**Figure 1.4**). Moreover, the abundant use has resulted in the occurrence of novel variants of pathogenic bacteria resistant towards this compound. *Enterococci* and other bacteria known to cause respiratory and urinary tract infections were found intrinsically resistant to the action of penicillins. Now, diseases like tuberculosis, claimed to be completely eliminated, are making come back with renewed resistance. Bacteria acquire genes conferring resistance by spontaneous DNA mutation, transformation, or through plasmids. Resistance may also arise from change in the structure of penicillin binding proteins such that the antibiotic does not bind efficiently (Essack, 2001; Antignac *et al.,* 2003).

Although bacterial antibiotic resistance is a natural phenomenon, other factors can accelerate it. The most serious preventable cause is inappropriate antibiotic use. The need of the hour is firstly, to prevent the spread of resistance and tackling resistant organisms, using narrow spectrum antibiotics and following proper regimen for the usage of antibiotics. Secondly, newer and effective antibiotics have to be found or developed that are more stable. In the search for new antibiotics, semi-synthetic penicillins with a β-lactam nucleus and a custom-designed side chain looked promising, though resistance to some semi-synthetic penicillins are already reported.
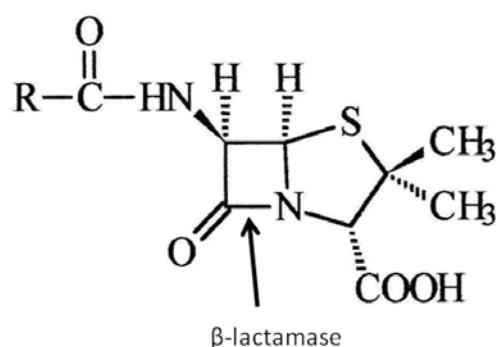
Figure 1.4: Site of action of pencillinase (β-lactamase).

1.5. Semi-synthetic Penicillins

In order to produce novel penicillins effective against resistant bacteria, a combined fermentation and chemical approach is used which leads to the production of semi-synthetic penicillins. Large-scale fermentations also helped to identify a mixture

of penicillin isoforms produced and it was demonstrated that the relative amounts of these naturally occurring penicillins depended on the composition of the fermentation medium (Vandamme & Voets, 1974). Subsequently, the selective fermentation of penicillin G (Pen G) was achieved by the addition of phenylacetic acid as an inducer into the fermentation medium. It was observed that variations in the side chain alter the properties of a β-lactam antibiotic e.g., penicillin V (Pen V) was produced supplying phenoxyacetic acid, N-acetylglycine, triglycine or phenylacetamide (Abraham, 1981; Vandamme & Voets, 1974) providing the first clue for more effective antibiotics. Appropriate chemical treatment of biological precursor penicillin, isolated from bacterial cultures, resulted in a number of semi-synthetic penicillins. After the mid 1950s, a wide range of new penicillins has been developed this way (**Figure 1.5**). The substitution of the side chain of Pen G, phenyl acetic acid, with other side chains like D-phenylglycine and D-p-hydroxyphenylglycine, resulted in the semi-synthetic β-lactam antibiotics Ampicillin and Amoxicillin, respectively, which are more stable and can be orally administered.
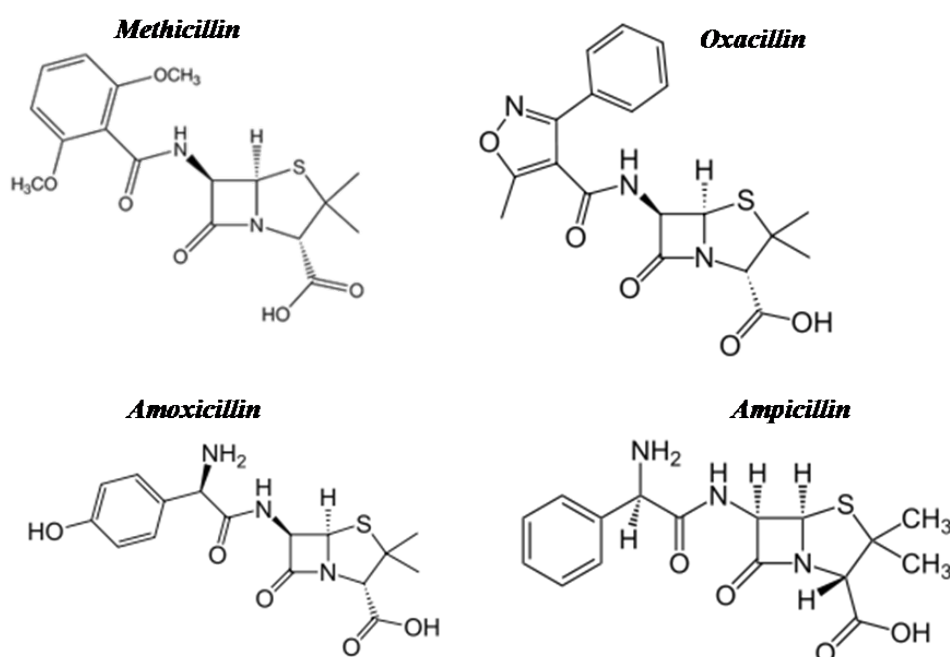
Figure 1.5: Molecules of some of the semi-synthetic penicillins. Each of them has specific spectrum of activity against a wide range of organisms.

Ampicillin and amoxicillin are both active against most aerobic gram-positive cocci. *S. aureus* is usually resistant to ampicillin and amoxicillin. The anaerobic gram

positive cocci and rods are generally susceptible to ampicillin, so are some aerobic gram negative bacilli such as *E. coli*, *Proteus mirabilis*, and *Haemophilus influenzae*. Ampicillin and penicillin are the drugs of choice for the treatment of infections caused by Group B Streptococci. Ampicillin, when combined with an aminoglycoside, is effective in treating intra-amniotic infections. Oxacillin or nafcillin are used in the treatment of complicated skin infections.

Carbenicillin has a similar spectrum of activity as that of ampicillin but is also active against *Pseudomonas aeruginosa*, whereas *Klebsiella pneumoniae* is resistant to carbenicillin. Ticarcillin is two to four times more potent than carbenicillin against *P. aeruginosa*. Piperacillin and mezlocillin have greater activity against gram-negative enteric organisms and also provide coverage against most anaerobes and *Enterococcus*. Dicloxacillin is orally-active semi-synthetic penicillin used to treat infections by bacteria such as *Staphylococcus aureus*.

The most important of semi-synthetic penicillin are methicillin and ampicillin, the former is remarkably effective against penicillinase-producing Staphylococci and the latter is not only active against all organisms normally killed by penicillin, but also inhibits growth of enterococci and many other bacteria. In fact, methicillin was the first successful semi-synthetic penicillin introduced in 1959 to overcome the problems that arose from the increasing prevalence of penicillinase-producing *S. aureus* (Livermore, 2000).

In 1980s, CI-867, a semi-synthetic penicillin was identified that exhibited broad-spectrum activity *in vitro* against gram-positive *cocci*. It was active against *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* (Weaver & Bodey, 1980). All of these newly prepared or identified antibiotics (**Table 1.1**) have varying effect on different types of microbial pathogens. This has provided evidence for the direct influence of the side chain on spectrum of penicillins and researchers started concentrating on the chemical structure of the side chain.

Table 1.1: Some semi-synthetic β-lactam antibiotics in the market

Amoxicillin	Dicloxacillin	Oxacillin
Ampicillin	Flucloxacillin	Piperacillin
Bacampicillin	Methicillin	Pivampicillin
Carbenicillin	Mezlocillin	Pivmecillinam
Cloxacillin	Nafcillin	Ticarcillin

1.6. β-lactam acylases

Classification of β-lactam acylases is based on their substrate specificities since their role *in vivo* is not known. They are broadly grouped into two classes; penicillin acylases and cephalosporin acylases (Deshpande *et al.,* 1994). Penicillin acylases (also known as penicillin amidases, penicillin amidohydrolases (**EC 3.5.1.11**) specifically hydrolyse the amide bond connecting the β-lactam nucleus to the side chain of penicillins.

A. Penicillin acylases

- Penicillin V acylase (type I, EC 3.5.1.11)

- Penicillin G acylase (type II, EC 3.5.1.11)

- Ampicillin acylase (type III, EC 3.5.1.11)

Type I, the phenoxy-methyl-penicillin acylase (PVA) has little activity towards Pen G or other derivatives. PVA occurs mainly in moulds and actinomycetes, although bacterial PVA also has been isolated.

Type II, the benzylpenicillin acylase (PGA) has broader substrate specificity. In addition to Pen G they act on a range of N-phenyl acetyl compounds. PGAs are mainly found in bacteria as periplasmic enzymes.

Type III, the ampicillin acylases specifically act on ampicillins (D-aminobenzyl penicillins). It has no action on Pen V or Pen G or related compounds.

B. Cephalosporin acylase

- Cephalosporin C acylase or glutarylamidase (CPC acylase, E.C. 3.5.1.11)

- Glutaryl –7-aminocephalosporanicacid acylase (Gl-7ACA acylase) (EC 3.5.1.93)

Much of the work reported on cephalosporin acylases is relatively recent and although the other penicillin acylases and cephalosporin acylases will be reviewed, the main discussion will be limited to PGA.

1.7. β-lactam acylases in the Production of Semi-synthetic Penicillins

The most commonly used antibiotics are based on the derivatives of 6-APA and 7- ADCA, which are manufactured from penicillin. In a multi-step reaction, natural penicillins are cleaved to yield the β-lactam nucleus 6-APA and next a desired side chain is chemically attached to 6-APA. Thus, most semi-synthetic penicillins used to be produced from 6-APA (Chisti & Moo-Young, 1991). Chemical ring expansion of penicillin plus enzymatic removal of the phenylacetyl side chain is currently being used in industry to obtain 7-aminodeacetoxycephalosporanic acid (7-ADCA) that is used for the manufacture of semi-synthetic cephalosporins. The chemical methods for producing 6-APA or 7-ADCA are environmentally unfriendly and require the use of hazardous chemicals such as pyridine, phosphorous pentachloride, and nitrosylchloride (Matsumoto, 1993; Vandamme, 1988). A significant discovery was of a particular *Penicillium* that produces free β-lactam nucleus in the absence of any precursors (Bruggink & Roy, 2001; Vandamme & Voets, 1974). This is due to penicillin acylase selectively hydrolyzing the amide bond leaving the β-lactam ring intact. Penicillin acylase from *E. coli* could catalyze the breakdown of Pen G to 6-APA and its organic

acid (Shewale & Shivaraman, 1989) (**Figure 1.6**). Comparative economics of 6-APA manufacture by chemical and enzymatic processes indicate that the enzymatic process is cheaper by at least 9% (Anon, 1992). The savings can be even higher (20%) if 6-APA manufactured by the enzymatic route is integrated with the production of penicillin G by fermentation.

Later it was realized that PGA catalyzes even the reverse reaction, under different conditions (Vandamme & Voets, 1974) (**Figure 1.7**). The enzymatic conversion is regio- and stereo-specific and the reaction conditions are usually mild. Thus, enzymatic process is cheaper and safer than chemical processes. The β-lactam acylases commonly used in pharmaceutical industry include penicillin acylases, cephalosporin acylase and glutaryl 7-aminocephalosporanic acid acylase from *Arthobacter viscosus, Bacillus laterosporus, Bacillus megaterium, Bacillus sphaericus, Kluyvera citrophila, Proteus rettgeri* and *Pseudomonas* sp. (Deshpande *et al.,* 1994).
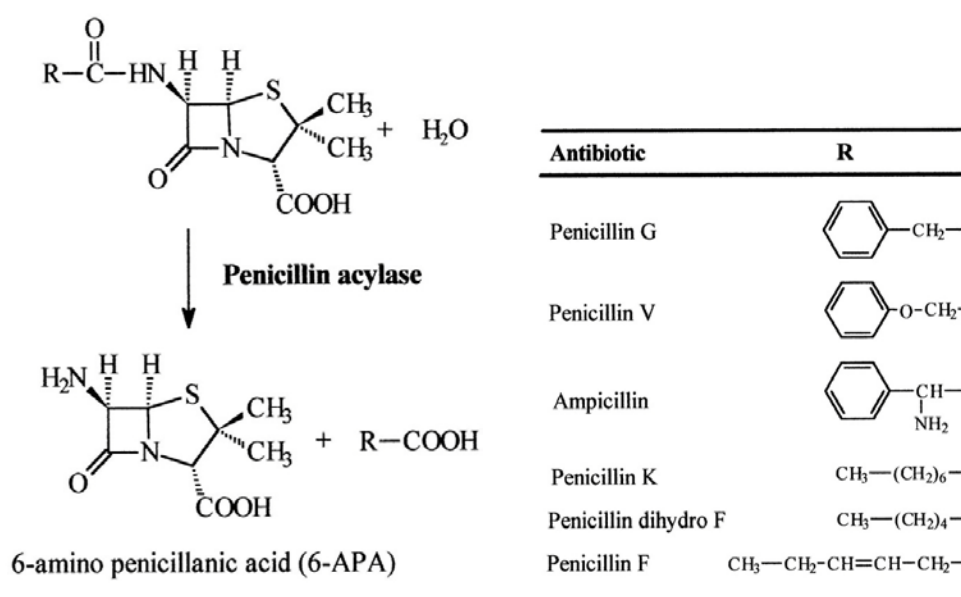
Figure 1.6: Penicillin acylase catalyzed reaction leading to the production of 6-amino penicillanic acid and corresponding side chain acid (Adopted from Arroyo *et al.,* 2003).
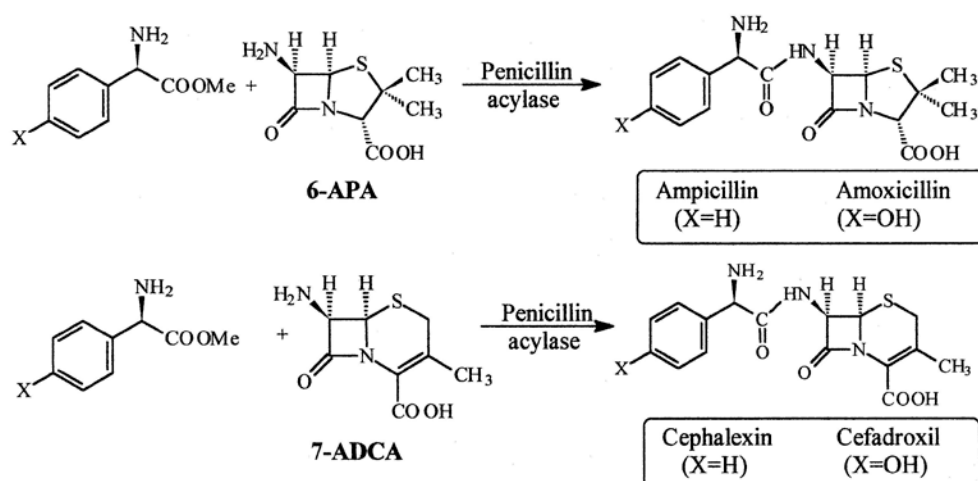
Figure 1.7: The addition of appropriate d-amino acid to 6-APA catalyzed by penicillin acylases at low pH (Adopted from Arroyo *et al*., 2003).

1.8. Cephalosporin acylases

Cephalosporin acylases (CAs) hydrolyze cephalosporin C (CPC) and/or glutaryl 7-aminocephalosporanic acid (GL-7ACA) to produce 7-amino cephalosporanic acid (7-ACA) and β-lactams with a charged side chain (Fritz-Wolf *et al.,* 2002). Two types of cephalosporin acylases have been found: glutaryl 7-ACA acylase (GL-7ACA acylase) and cephalosporin C acylase (CPC acylase). GL-7ACA acylase has high activity on GL-7ACA but much lower activity on CPC. CPC acylase is active on both CPC and GL-7ACA. Irrespective of the low sequence similarity and different substrate specificity between PAs and CAs, the structural similarity at their active sites is impressive. A detailed account of various aspects of cephalosporin acylases has been presented by Kumar *et al.* (1993) and Sonawane, (2006). Several genes encoding cephalosporin acylases from different sources have been isolated and cloned (Matsuda & Komatsu, 1985; Matsuda *et al.,* 1987; Yang *et al.,* 1991; Ishiye & Niwa, 1992; Ishii *et al.,* 1994).

Cephalosporin C acylase from *Pseudomonas* sp. strain N176, a heterodimer of 25 kDa and 58 kDa has been crystallized (Kinoshita *et al.,* 2000). The native and mutant precursor glutaryl 7-aminocephalosporanic acid acylase (GL-7ACA acylase) was crystallized by Kim *et al.* (2003). The nascent enzyme is synthesized as a 74 kDa

polypeptide containing sequences coding for a signal peptide, 16 kDa α-subunit, 9 amino acids spacer peptide, and 54 kDa *β*-subunit. Like PGAs, the enzymes are heterodimer of two non-identical subunits, α and β, derived from a nascent precursor polypeptide that is cleaved proteolytically through a two-step autocatalytic process. The two subunits of the acylase separately are inactive (Kim *et al.,* 2003, Li *et al.,* 1999). The catalytic mechanism of CA is same as PGA (Kim *et al.,* 2003; Mao *et al.,* 2004). The crystal structures of intracellular heterodimeric CA from *Pseudomonas diminuta* **(Figure 1.8)** (Fritz-Wolf *et al.,* 2002; Kim *et al.,* 2000) and *Pseudomonas* sp. similar to the *E. coli* PGA structure and belong to Ntn-hydrolase fold **(Figure 1.9)** (Kim *et al.,* 2000).

The semi-synthetic cephalosporins represent one of the commonly prescribed antibiotic classes. Consequently, there is a huge demand for the intermediates of semi-synthetic cephalosporins, 7-ACA and 7-ADCA (Oh *et al.,* 2003). In order to meet this need, alternative routes for the production of the intermediates are investigated to replace the currently used processes (Bruggink *et al.,* 1998). One of these alternative routes comprises a one-step enzymatic deacylation of either the fermentation product adipyl-7-ADCA (Crawford *et al.,* 1995) to 7-ADCA or the fermentation product Cephalosporin C (CPC) to 7-ACA by cephalosporin acylases. However, in spite of the fact that the available cephalosporin acylases are highly efficient in the deacylation of glutaryl-7-ACA, they are much less efficient in deacylating adipyl-7-ADCA and upfront inefficient for the hydrolysis of CPC (Shibuya *et al.,* 1981; Kim *et al.,* 1999; Li *et al.,* 1999).
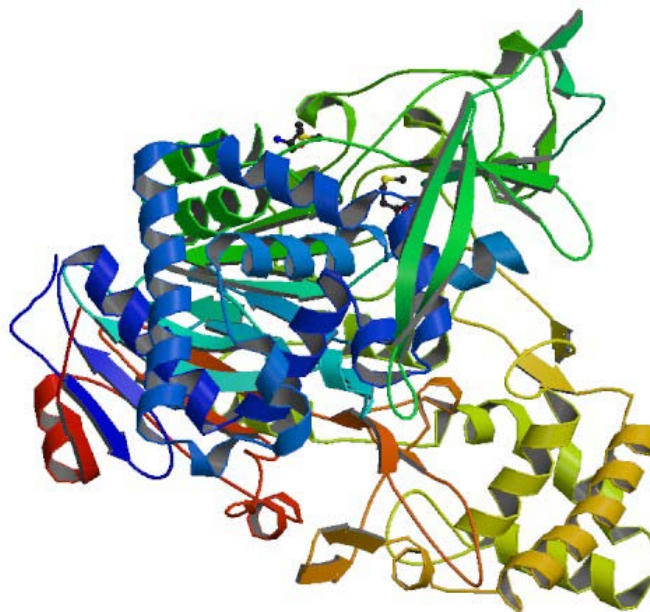
Figure 1.8: Crystal structure of the Cephalosporin C acylase from *Pseudomonas diminuta* (PDB ID: 1FM2, Kim *et al.,* 2000). This figure and the other figures have been prepared using *CCP4MG* (McNicholas *et al.,* 2011).
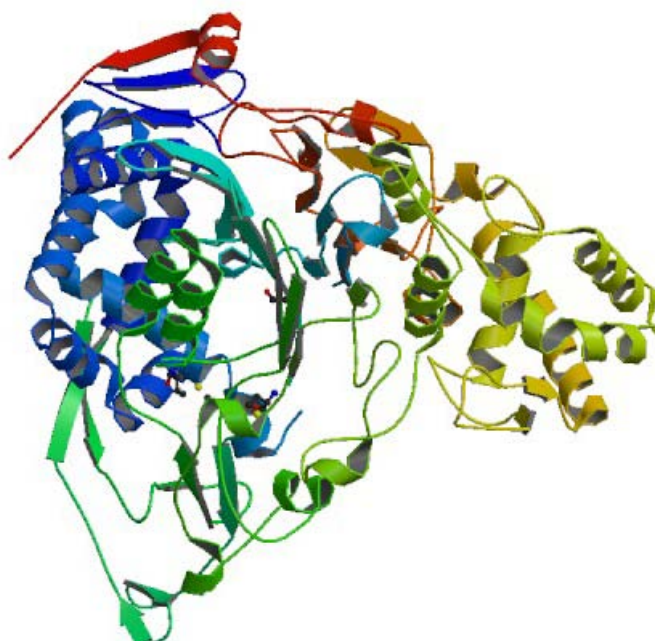
Figure 1.9: Crystal structure of glutaryl 7-aminocephalosporanic acid acylase from *Pseudomonas* sp. strain GK16 (PDB ID: 1OR0, Kim *et al.,* 2003).

1.9. Penicillin acylases

Enzyme responsible for the large amount of production of 6-APA or 7-ADCA has been given several different names including "penicillin amidase" (Sakaguchi & Murao, 1950; Murao, 1955; Claridge *et al.,* 1960, Murao & Kashida, 1961; Claridge *et al.,* 1963), "penamidase", "penicillin deacylase" (Cole, 1964), "penicillin splitting and synthesizing enzyme" (Rolinson *et al.,* 1960; Kaufman & Bauer, 1960), "acyltransferase" (Kaufmann & Bauer, 1964) and, perhaps the least satisfactory name of all, the official (Enzyme Commission) nomenclature "penicillin amidohydrolase" (**EC 3.5.1.11**). Commonly they are called as Penicillin acylases (PAs). This enzyme belongs to the family of hydrolases, those acting on carbon-nitrogen bonds other than peptide bonds. Enzymatic hydrolysis of penicillins to 6-APA was first reported by Sakaguchi & Murao (1950) with the enzymes obtained from *Penicillium chrysogenum* Q176 and *Aspergillus oryzae*. Microorganisms have been extensively screened for penicillin acylase production. Screening of penicillin acylase-producing microorganisms has shown that these enzymes are produced by bacteria, actinomycetes, yeasts and filamentous fungi (Sudhakaran & Borkar 1985a, b; Savidge, 1984; Shewale & Sivaraman, 1989).

As mentioned before, the penicillin acylases are classified according to the type of substrate preferably hydrolyzed such as penicillin V acylase (PVA) hydrolyze pen V, Penicillin G acylase (PGA) pen G and ampicillin acylase act on ampicillins. However, some reports on PAs have mentioned that they have broad substrate specificity, hydrolyzing more than one type of penicillin. PVA from *Streptomyces lavendulae* also acts on aliphatic penicillins like penicillin F, dihydro F and K (Torres *et al.,* 2002). The authors further suggested that the enzyme should be termed as penicillin K acylase since the catalytic constant for penicillin K is larger than that for Pen V (Torres *et al.,* 2002). A newly discovered penicillin acylase from *Actinoplanes utahensis* is described (Torres-Bacete *et al.,* 2007), which can hydrolyze various natural aliphatic penicillins and penicillin K was found to be the best substrate.

1.9.1. Penicillin V acylases (type I)

Penicillin V acylase (PVA) reversibly cleave the amide bond between the side chain phenoxymethyl group and the β-lactam nucleus of Pen V. Immobilized PVA is

mainly involved in this process, which accounts for 15% of the worldwide 6-APA production (Vandamme, 1988). Earlier perception was that PGA being produced mainly by bacterial cultures and PVA by moulds (Claridge *et al.,* 1963), but later investigations established that both PGAs and PVAs are distributed in bacteria, actinomycetes, yeasts, and fungi (Vandamme & Voets, 1974; Sudhakaran & Borkar, 1985, b; Shewale & Sudhakaran, 1997). PVA occur mainly in moulds and actinomycetes, although bacterial PVA has been isolated from *Erwinia aroideae* (Vandamme & Voets, 1975), *Beijerinckia indica var.* (Ambedkar *et al.,* 1991), *Rhodotorula glutinis* var. *glutinis* (Vandamme & Voets, 1973), *Bacillus sphaericus* (Olsson & Uhlen, 1986), *Bacillus cereus* ATUAVP1846 (Sunder *et al*., 2012) and *Pseudomonas acidovorans* (Lowe *et al.,* 1981).

Most of the PVAs are mainly intracellular enzymes, however, PVA from *Fusarium* sp. SKF 235 (Sudhakaran & Shewale, 1995) and actinomycete *Streptomyces lavendulae* ATCC 13664 (Torres-Bacete *et al.*, 2000) are extracellular acylases. The production is either constitutive or enhanced by the presence of phenoxyacetic acid in growth medium. The molecular weight normally ranges from 83.2 kDa in *Fusarium* sp. SKF 235 to 140 kDa in *Bacillus sphaericus*, and their subunit composition varies from monomer to tetramer. However, intracellular PVA from *Rhodotorula aurantiaca* has been reported as a smallest active monomeric penicillin V acylase (Kumar *et al*., 2008). Most of the commercially used cultures are either mutants or genetically engineered strains. Both physical and chemical mutagenic techniques and modern genetic engineering approaches have been exploited for improving the production of PVA (Sudhakaran & Borkar, 1985b; Savidge, 1984; Deshpande *et al.,* 1994). *B. indica var penicillanicum* UREMS-5 producing 168% more PVA was obtained by treating natural isolate successively with UV, γ-irradiation, and ethyl methane sulfonate (EMS) (Ambedkar, 1991).

The optimum pH values for PVA range between pH 5.6 - 8.5, as compared to 6.5-8.5 for PGA (Margolin et al., 1980; Schumacher et al., 1986). Optimum pH of PVA from fungi range between 7.0-8.5; bacteria range between 5.5-7.5 except for PVAs from P. acidovorans and Achromobacter sp. which have pH optima at 8.0 and 8.5, respectively, and for yeasts it ranges between 6.5- 7.5.

The PVA gene from *B. sphaericus* (Olsson *et al.,* 1985) has been cloned and studied for biochemical and structural (Suresh *et al.,* 1999) details. PVA is a homotetramer of 140 kDa. According to sequence similarity, PVA belongs to choloylglycine hydrolase family of the superfamily. The crystal structure information (Suresh *et al.,* 1999) placed *Bsp*PVA in the N-terminal nucleophile (Ntn) hydrolase superfamily (Brannigan *et al.,* 1995). The functional molecule is a well-defined homotetramer of 222 organization formed by four monomers, which generates a flat disc-like assembly (**Figure 1.10A**). The X-ray crystal structure analysis revealed that the *Bsp*PVA monomer contains two layers of central anti-parallel β-sheets, sandwiched above and below by pairs of anti-parallel helices, a hallmark of members of Ntn-hydrolase superfamily. There are two extensions, one from the upper pair of helices and the other from the C-terminal segment that interact with the neighbouring monomers of the tetramer and help to stabilize the quaternary structure (Suresh *et al.,* 1999). The organization and connectivity of β-sheets and α-helices are characteristic of the members of Ntn-hydrolase family (**Figure 1.10B**). PVA crystal structure revealed that the N-terminal nucleophile is a cysteine. DNA sequence had three more amino acids before Cys including the N-terminal methionine, indicating that the three preceding residues including Met were removed during maturation, a post-translational processing shared with other Ntn-hydrolases. The inactive mutants were prepared to study autocatalytic processing of precursor PVA molecule (Chandra *et al.,* 2005).

YxeI gene encoding for an un-annotated protein from *Bacillus subtilis* has been expressed in *Escherichia coli*, purified and confirmed to possess PVA activity (Rathinaswamy *et al.,* 2005). Comparison with the structurally characterized *B. sphaericus* PVA (Suresh *et al.,* 1999) and the bile-salt hydrolase from *Bifidobacterium longum* (Kumar *et al.,* 2006) revealed that like CBH, the catalytic Cys is unmasked by simple removal of the preceding $_f$Met and the protein thus does not possess an N-terminal pro-peptide as present in *B. sphaericus* PVA (Chandra *et al.,* 2005). The sequence comparison also highlights that YxeI shares 40% identity with *B. sphaericus* PVA and 30% identity to the more distantly related *Bifidobacterium longum* bile-salt hydrolase. The enzyme was a homotetramer of 148 kDa with subunit size of 37 kDa which is similar to that of PVA from *B. sphaericus* of size 140 kDa with subunit of 35

kDa and the crystal structure showed the homotetrameric arrangement similar to PVA from *B. sphaericus* and CBH from *Bifidobacterium longum.*
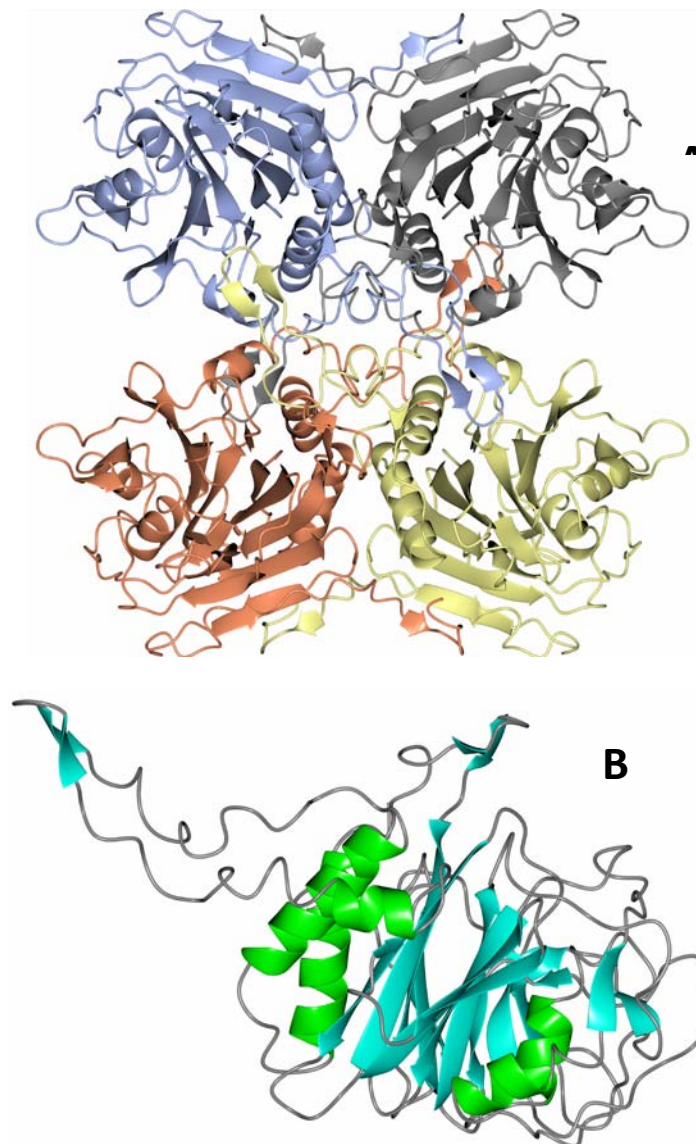
Figure 1.10: (A) Structure of the PVA tetramer from *B. sphaericus*, (B) Structure of the monomer showing the characteristic Ntn-hydrolase fold (PDB ID: 3PVA).

1.9.2. Penicillin G acylases (Type II)

Penicillin G acylase (PGA) catalyses the hydrolysis of Pen G and the condensation of β-lactam nucleus with Cα-substituted phenylacetic acids. In addition to Pen G these enzymes act on a range of N-phenyl acetyl compounds. PGA represents

the only industrial enzyme that has been commercialized within two decades of its discovery. In the semi-synthetic penicillin industry, PGA is being used in 88% of the 6-APA produced, either in the soluble form or in the immobilised form (Rajendhran & Gunasekaran, 2004). It is estimated that 10-30 million tons of immobilized PGA are utilized each year (de Souza *et al.*, 2005).

PGAs are found throughout the whole kingdom of prokaryotes (Arroyo *et al.,* 2003). PGA has been studied from *E. coli* (Cole, 1969; Schumacher *et al.,* 1986), *Kluyvera citrophila* (Barbero *et al.,* 1986; Martin *et al.,* 1991), *Arthrobacter viscosus* (Ohashi *et al.,*1988, 1989), *Alcaligenes faecalis* (Verhaert *et al.,* 1997), *Bacillus megaterium* (Chiang & Bennet, 1967; Martin *et al.,* 1995), *Proteus rettgeri* (Klei *et al.,* 1995) and *Achromobacter xylosoxidans* (Cai *et al.,* 2004). Bacterial PGAs can also catalyze the reverse reaction i.e. the synthesis of penicillins from 6-APA and phenyl acetic acid and its derivatives. The study of PGA remained the basis for all research in penicillin acylases. A list of penicillin G acylase producing microorganisms is summarized in **table 1.2.**

1.9.3. Ampicillin acylase (Type III)

Ampicillin acylases act on the antibiotics with phenylglycine derived side chain, such as ampicillin and cephalexin. Ampicillin acylase from *Kluyvera citrophila* was reported by Nara *et al.* (1971a) and from *Pseudomonas melanogenum* by Okachi *et al.* (1973). This enzyme has a completely different substrate range as it catalyzes both the synthesis and hydrolysis of ampicillin but shows activity neither towards Pen G nor towards Pen V. Ampicillin acylase from *Pseudomonas melanogenum*, a homodimer of 72 kDa size, has been studied by Kim & Byun (1990).

Table 1.2: List of Microorganisms that produce penicillin G acylase

Organism	Reference
Escherichia sp.	Rolinson *et al.,* (1960)
Pseudomonas sp.	Huang *et al.,* (1960)
E. coli ATCC 9637	Kaufman & Bauer (1960)
Aerobacter cloacae	Claridge *et al.,* (1960)
Bacillus subtilis var *niger*	Claridge *et al.,* (1960)
Mycobacterium phlei	Claridge *et al.,* (1960)
Nocardia F D 46973, ATCC 13635	Huang *et al.,* (1960)
Xanthomonas sp.	Huang *et al.,* (1963)
Flavobacterium	Huang *et al.,* (1963)
Proteus rettgeri F D 13424	Huang *et al.,* (1963)
E. coli NCIB 9465	Holt & Stewart (1964)
Alcaligenes faecalis BRL 1237, 1238	Cole & Sutherland (1966)
Proteus rettgeri ATCC 9919, 9250	Cole *et al.,* (1967)
Bacillus megaterium ATCC 14945	Chiang & Bennett (1967)
E. coli NCIB 8134, 8879, 8949	Cole (1967)
Rhodopseudomonas Spheroids	Nara *et al.,* (1971b)
Streptomyces ambofaciens SPSL 15	Nara *et al.,* (1971b)
E. coli ATCC 11105	Bauer *et al.,* (1971)
Kluyvera citrophila KY 3641	Okachi *et al.,* (1972)
Peudomonas aeruginosa KY 3591, KY 8501	Okachi *et al.,* (1972)
E. coli BMN, KY 8219, KY 8268, KY 8275, KY 8289	Okachi *et al.,* (1973)
Proteus morganii KY 4035, KY 405	Okachi *et al.,* (1973)
Arthrobacter viscosus 8895GU	Ohashi *et al.,* (1989)
Achromobacter sp. CCM 4824	Skrob *et al.,* (2003)
Bacillus badius	Rajendhran *et al*., (2002)

1.10. Industrial Applications of Penicillin acylases

Since the discovery of penicillin acylases (PAs) in the 1950's, they have been incorporated throughout the pharmaceutical industry and exploited for industrial purposes. Biotechnological processes for the large-scale production of semi-synthetic penicillins and cephalosporins are focused on the condensation of the appropriate D-amino acid derivative with a β-lactam nucleus catalyzed mainly by penicillin acylases (Bruggink *et al.* 1998) (**Figure 1.7**). Such biotransformations can be performed by either thermodynamically or kinetically controlled synthesis (Kasche, 1986). The combined application of an immobilized biocatalyst (Fernandez-Lafuente *et al.,* 1995), optimized pH value of the reaction medium (Ospina *et al.,* 1996), addition of a suitable solvent (Rosell *et al.,* 1998), and a high concentration of activated side chain donor and β-lactam nucleus (Kaasgaard & Veitland, 1992) have together achieved best results.

Besides penicillins, penicillin acylases catalyze the hydrolysis of amides, acylamino acids, and ester derivatives of the carboxylic acid side chains of penicillins. Substrate susceptibility is determined by the acyl moiety and some of the non-penicillin compounds being better substrates than the corresponding penicillins (Cole, 1964; Kaufmann & Bauer, 1964; Kutzbach & Rauenbusch, 1974).

Apart from penicillin production, PAs are also used in other industries, e.g., in peptide synthesis or acyl group transfer reactions (Van Langen *et al.,* 2000). Amidase from *E. coli* is used in the synthesis of artificial sweetener aspartame (Fuganti *et al.,* 1986) and diphenyl dipeptides, whose derivatives are, used as food additives, fungicidal, antiviral and anti-allergic compounds (Van Langen *et al.,* 2000). PAs can be used to resolve racemic mixtures of chiral compounds such as amino acids (Bossi *et al.,* 1998), β-amino esters, amines and secondary alcohols (Svedas *et al.,* 1996).

1.11. Biological Function of Penicillin acylases

The question of the physiological role of penicillin acylase has been overshadowed by the attention paid to its industrial applications. The question still remains essentially unanswered although evidence suggests that the penicillin acylase gene is related to pathways involved in the assimilation of aromatic compounds as carbon sources.

Expression of the enzyme *in vivo* is regulated by three major mechanisms: cAMP receptor protein (CRP), phenylacetic acid and temperature giving rise to the hypothesis that penicillin acylase is involved in the degradation of phenoxyacetylated compounds for the generation of phenoxyacetic acid, which may be used as a carbon source and could act as an inducer of the degradative pathway (Valle *et al*, 1991). Since its initial suggestion, the location of penicillin acylase gene in *E. coli* near a gene encoding for an aromatic hydroxylase has been discovered (Prieto *et al.,* 1993). Later the nucleic acid sequence of a 14855 base pair region that contains the complete gene cluster encoding 4-hydroxyphenylacetic acid degradative pathway of *E. coli* W ATCC 11105 was determined (Prieto *et al.,* 1996). The gene cluster is located in a region close to the gene encoding for PGA. It has been therefore suggested that PGA is present in *E. coli* W to improve its ability to metabolize a wider range of substrates. Penicillin acylases are able to hydrolyze phenylacetylated compounds, the products that may then be fed into 4- hydroxyphenylacetic acid degradative pathways, thus enhancing the catabolic versatility of *E. coli*.

1.12. Ntn-hydrolase Superfamily

The members of the Ntn-hydrolase superfamily have an N-terminal nucleophile residue, Ser, Thr or Cys whose side chain atom O or S acts as nucleophile and free α-amino group acts as base in catalysis. The members of this family catalyses the hydrolysis of substrate's amide bond and each enzyme has its own substrate specificity (Brannigan *et al.,* 1995). The N-terminal nucleophile residue with free α-amino group to act as proton donor in the catalytic process is often created by a post-translational autocatalytic cleavage at the nucleophile residue. The representative fold of this superfamily is composed of a four-layered catalytically active αββα core structure. This core structure consists of two closely packed four or five-stranded antiparallel β-sheets and they are sandwiched by a layer each of α-helices on either side. One of the sandwiched antiparallel β-sheets is flatter and the other one generally twisted. This family contains enzymes of markedly varying size and complexity, with a wide range of substrates and operating in very different biological contexts; it includes for example penicillin G acylase (Duggleby *et al.,* 1995), cephalosporin acylase (Kim *et al.,* 2000), glutaryl 7-aminocephalosporanic acid acylase (Lee *et al.,* 2000a, b) with a catalytic

serine; the 20S subunit of the proteasome (Lowe *et al.,* 1995; Groll *et al.,* 1997*)* and aspartylglucosaminidase (Oinonen *et al.,* 1995) with threonine; and penicillin V acylase (Suresh *et al.,* 1999), bile salt hydrolase (Rossocha *et al*., 2005), glutamine PRPP amidotransferase (Smith *et al.,* 1994), glucosamine-6-phosphate synthase (Isupov *et al.,* 1996) with cysteine. All are processed autocatalytically from inactive precursors (Brannigan *et al.,* 1995). Recently human acid ceramidase (Shtraizent *et al*., 2008) and acyl-homoserin-lactone acylase pvdQ from *Pseudomonas aeruginosa* (Bokhove *et al*., 2010a) containing cysteine and serine residue, respectively as nucleophile has been added to the superfamily.

Their structures show variability in the number of secondary structural elements and the details of arrangement (Brannigan *et al.,* 1995; Oinonen & Rouvinen, 2000). However, the common connectivity of sequential elements of secondary structure in these enzymes suggests that the Ntn-hydrolases have diverged from a common ancestor (Brannigan *et al.,* 1995). The evolutionary relationship between these hydrolases, although hidden by the absence of sequence similarity, has thus been unveiled on determining their structures.

According to the MEROPS database classification (Rawlings *et al.,* 2004), the Ntn-hydrolases form the peptidase clan PB, which is the only clan with three different catalytic types. Ntn-hydrolases represent a superfamily of amidohydrolases that have developed great divergence in sequence, structure, and substrate specificity (Pei & Grishin, 2003).

Ntn-hydrolases are also emerging as important targets for therapy e.g., the subunit 20 S of the proteasome, is inhibited by the anti-cancer drug bortezomib (Groll *et al*., 2006), N-acylethanolamine-hydrolyzing acid amidase (NAAA) (Tsuboi *et al*., 2007), a potential target for anti-inflammatory and analgesic drugs (Solorzano *et al*., 2009), and acid ceramidase (AC) (Shtraizent et *al*., 2008), a potential target for cancer chemosensitizing agents (Zeidan *et al*., 2008).

1.13. Members of the Ntn-hydrolase Family

The hydrolases and references mainly to their three-dimensional structures are listed below according to their importance and utility in research reported in the thesis:

- Penicillin G acylase from *E. coli* (*Ec*PGA) (Duggleby *et al.,* 1995), *Providencia rettgeri* (McDonough *et al.,* 1999)

- Penicillin V acylase from *Bacillus sphaericus* (*Bsp*PVA) (Suresh *et al.,* 1999)

- Bile salt hyrolase from *Clostridium perfringens* (*Cp*BSH) (Rossocha *et al.,* 2005)

- Aspartylglucosaminidase (AGU) from human (Oinonen *et al.,* 1995; Perakyla & Rouvinen, 1996)

- Glycosylasparaginase (GA) from *Flavobacterium meningosepticum* (Guan *et al.,* 1998; Xu *et al.,* 1999)

- Glutamine phosphoribosyl-pyrophosphate (PRPP) amidotransferase from *Bacillus subtilis* (Smith *et al.,* 1994), *E. coli* (Kim *et al.,* 1996)

- L-Aminopeptidase-D-Ala-amidase (APA) from *Ochrobactrum anthropi* (Bompard-Gilles *et al.,* 2000)

- N-carbamyl-D-aminoacid amidohydrolase (DCase) from *Agrobacterium sp.* (Nakai *et al.,* 2000)

- Glutaminase domain of glucosamine-6-phosphate synthase (GPS) from *E. coli* (Isupov *et al.,* 1996).

- Ornithine acetyltransferase (OA) (Abadjieva *et al.,* 2000; Elkins *et al.,* 2005; Sankaranarayanan *et al*., 2010)

- U34 peptidase (Pei & Grishin, 2003)

- MTH1020 conserved protein from *Methanobacterium thermoautotrophicum* (Saridakis *et al.,* 2002)

- Plant-type L-asparaginase from *E. coli* (Michalska *et al*., 2008a)

- Glutaryl 7-aminocephalosporanic acid acylase from *E. coli* (Lee *et al.,* 2000a; 2000b)

- Heat shock locus V (HslV) from *E. coli* (Bochtler *et al.,* 1997)

- Gamma-glutamyl transferase (GGT) from *E. coli* (Okada *et al.,* 2006)

- 20S subunit of the proteasome *Saccharomyces cerevisiae* (Groll *et al.,* 1997), *Thermoplasma acidophilum* (Lowe *et al.,* 1995)

- lysosomal 66.3 kDa protein from *Mus musculus* (Lakomek *et al*., 2009)

- Acyl-homoserin-lactone acylase pvdQ from *Pseudomonas aeruginosa* (Bokhove *et al*., 2010a)

Table 1.3 lists the chronological order in which the three-dimensional structure of Ntn-hydrolases deposited in protein databank.

Table 1.3: Chronological order in which the three-dimensional structure of Ntn hydrolases deposited in protein databank.

Enzyme	PDB Code	Organism	Resolution(Å)	Catalytic type	Year of deposition and Reference
Glutamine phosphoribosylpyrophosphate (PRPP) amidotransferase (amidophosphoribosyltransferase)	1GPH	*Bacillus subtilis*	3.0	Cysteine	1994-04-20, Smith *et al.*, 1994
20S Proteasome	1PMA	*Thermoplasma acidophilum*	3.4	Threonine	1994-12-19, Lowe *et al.*, 1995
Penicillin G acylase	1PNK	*Escherichia coli*	1.90	Serine	1995-03-16, Duggleby *et al.*, 1995
Human lysosomal aspartylglucosaminidase (AGA)	1APY	*Homo sapiens*	2.00	Threonine	1995-06-14, Oinonen *et al.*, 1995
Glutamine phosphoribosylpyrophosphate (PRPP) amidotransferase (amidophosphoribosyltransferase)	1ECG	*Escherichia coli*	2.30	Cysteine	1996-04-23, Kim *et al.*, 1996
20S Proteasome	1RYP	*Saccharomyces cerevisiae*	2.4	Threonine	1997-02-26, Groll *et al.*, 1997
Heat shock locus V (HslV) (ATP-dependent Protease)	1NED	*Escherichia coli*	3.8	Threonine	1997-04-04, Bochtler *et al.*, 1997
Glycosylasparaginase	1AYY	*Elizabethkingia meningoseptica (Flavobacterium meningosepticum)*	2.32	Threonine	1997-11-12, Xuan *et al.*, 1998
Penicillin V acylase	2PVA	*Lysinibacillus sphaericus (Bacillus sphaericus)*	2.5	Cysteine	1998-11-13, Suresh *et al.*, 1999

L-Aminopeptidase D-ala-esterase/amidase (DmpA)	1B65	*Ochrobactrum anthropi*	1.82	Serine	1999-01-20, Bompard-Gilles *et al.*, 2000
Penicillin G acylase	1CP9	*Providencia rettgeri*	2.50	Serine	1999-06-12, McDonough *et al.*, 1999
Glycosylasparaginase Precursor	9GAF	*Elizabethkingia meningoseptica (Flavobacterium meningosepticum)*	1.9	Threonine	1999-06-15, Xu *et al.*, 1999
Asparagine synthetase B	1CT9	*Escherichia coli*	2.0	Cysteine	1999-08-20, Larsen *et al.*, 1999
N-carbamyl-D-amino acid amidohydrolase	1ERZ	*Agrobacterium sp.*	1.70	Cysteine	2000-04-06, Nakai *et al.*, 2000
Penicillin G acylase precursor	1E3A	*Escherichia coli*	1.80	Serine	2000-06-07, Hewitt *et al.*, 2000
Glutaryl-7aminocephalosporanic acid (Glutaryl-7ACA)	1FM2	*Brevundimonas diminuta (Pseudomonas diminuta)*	2.0	Serine	2000-08-15, Kim *et al.*, 2000
N-carbamyl-D-amino acid amidohydrolase	1FO6	*Agrobacterium radiobacter*	1.95	Cysteine	2000-08-25, Wang *et al.*, 2001
ATP-dependent protease subunit HslV (Heat shock locus V)	1G3K	*Haemophilus influenzae*	1.90	Threonine	2000-10-24, Sousa *et al.*, 2000
Glutamate synthase	1EA0	*Azospirillum brasilense*	3.0	Cysteine	2000-11-02, Binda *et al.*, 2000
Plant-type L-asparaginase	1JN9	*Escherichia coli*	2.30	Threonine	2001-07-23, Michalska *et al.*, 2008a
Glutarylamidase	1GK0	*Pseudomonas sp.*	2.50	Serine	2001-08-07, Fritz-Wolf *et al.*, 2002

Glucosamine 6-phosphate synthase	1JXA	*Escherichia coli*	3.10	Cysteine	2001-09-06, Teplyakov *et al.*, 2001
20S Proteasome complex	1IRU	*Bos taurus*	2.75	Threonine	2001-10-24. Unno *et al.*, 2002
Cephalosporin acylase precursor	1KEH	*Brevundimonas diminuta*	2.5	Serine	2001-11-16, Kim *et al.*, 2002
Glutamate synthase2	1LM1	*Synechocystis sp.*	2.8	Cysteine	2002-04-30, van Den Heuvel *et al.*, 2002
ATP-dependent protease subunit HslV(Heat shock locus V)	1M4Y	*Thermotoga maritima*	2.10	Threonine	2002-07-05, Song *et al.*, 2003
20S Proteasome	1J2Q	*Archaeoglobus fulgidus*	2.83	Threonine	2003-01-08, Groll *et al.*, 2003
L-Asparaginase Precursor	1P4K	*Elizabethkingia meningoseptica(Flavobacterium meningosepticum)*	1.9	Inactive	2003-04-23, Qian *et al.*, 2003
Carbapenam synthetase (CarA)	1Q15	*Pectobacterium carotovorum*	2.3	Serine	2003-07-18, Miller *et al.*, 2003
20S Proteasome	1Q5Q	*Rhodococcus erythropolis*	2.60	Threonine	2003-08-08, Kwon *et al.*, 2004
Ornithine acetyltransferase (OAT2)	1VZ8	*Streptomyces clavuligerus*	2.75	Threonine	2004-05-14, Elkins *et al.*, 2005
Putative Glutamine amidotransferase	1TE5	*Pseudomonas aeruginosa*	2.0	Cysteine#	2004-05-24, Patskovsky and Almo
Conjugated bile acid hydrolase (BSH)	2BJF	*Clostridium perfringens*	1.67	Cysteine	2005-02-02, Rossocha *et al.*, 2005

Ornithine acetyltransferase (OAT)	1VRA	*Bacillus halodurans*	2.0	Threonine	2005-02-17 ,Joint Center for Structural Genomics
ATP-dependent protease subunit ClpQ (CodW)	1YYF	*Bacillus subtilis*	4.16	Serine	2005-02-24, Wang *et al.*, 2005
Threonine aspartase 1(Taspase1)	2A8J	*Homo sapiens*	1.9	Threonine	2005-07-08, Khan *et al.*, 2005
Glutaryl 7-aminocephalosporanic acid acylase	2AE4	*Pseudomonas sp. GK16*	2.30	Serine	2005-07-21, Kim *et al.*, 2006
Gamma-glutamyltranspeptidase (GGT)	2DBU	*Escherichia coli*	1.95	Threonine	2005-12-16, Okada *et al.*, 2006
20S Proteasome	2FHG	*Mycobacterium tuberculosis*	3.23	Threonine	2005-12-23, Hu *et al.*, 2006
L-Asparaginase (AsnB)	2GEZ	*Lupinus luteus*	2.60	Threonine	2006-03-21, Michalska *et al.*, 2006
Bile Salt Hydrolase (BSH)	2HF0	*Bifidobacterium longum*	2.30	Cysteine	2006-06-22, Kumar *et al.*, 2006
Penicillin V acylase Precursor	2IWM	*Lysinibacillus sphaericus (Bacillus spharericus)*	2.50	Inactive	2006-07-01, Chandra *et al.*
Gamma-glutamyltranspeptidase (GGT)	2I3O	*Thermoplasma acidophilum*	2.03	Threonine	2006-08-19, Rao *et al.*
Gamma-glutamyltranspeptidase (GGT) precursor	2E0W	*Escherichia coli*	2.55	Inactive	2006-10-16, Okada *et al.*, 2007
Gamma-glutamyltranspeptidase (GGT)	2NQO	*Helicobacter pylori*	1.9	Threonine	2006-10-31, Boanca *et al.*, 2007
Penicillin V acylase (PVA)	2OQC	*Bacillus subtilis*	2.50	Cysteine	2007-01-31, Suresh *et al.*

Gamma-glutamyltranspeptidase (GGT)	2V36	*Bacillus subtilis*	1.85	Threonine	2007-06-13, Sharath *et al*., Wada *et al*., 2010
Plant-type L-asparaginase precursor	2ZAK	*Escherichia coli*	2.01	Inactive	2007-10-07, Michalska *et al*., 2008b
lysosomal 66.3 kDa protein	3FGR	*Mus musculus*	1.8	Cysteine	2008-12-08, Lakomek *et al*., 2009
Transpeptidase enzyme (Capsule biosynthesis protein, CapD)	3G9K	*Bacillus anthracis*	1.79	Threonine	2009-02-13, Wu *et al*., 2009
Proteasome 20S core particle	3H4P	*Methanocaldococcus jannaschii*	4.10	Threonine	2009-04-20, Zhang *et al*., 2009
Choloylglycine Hydrolase	3HBC	*Bacteroides thetaiotaomicron*	2.27	Cysteine#	2009-05-04, Kim *et al*.
Ornithine acetyltransferase (OAT)	3IT4	*Mycobacterium tuberculosis*	1.7	Threonine	2009-08-27, Sankaranarayanan *et al*., 2010
Penicillin G acylase	3K3W	*Alcaligenes faecalis*	3.3	Serine	2009-10-05, Varshney *et al*., 2012
N-acyl homoserine lactone acylase PvdQ	2WYE	*Pseudomonas aeruginosa*	1.8	Serine	2009-11-16, Bokhove *et al*., 2010
acyl coenzyme A:isopenicillin N acyltransferase (acyl-coenzyme A:6-aminopenicillanic acid acyl-transferase) Precursor	2X1C	*Penicillium chrysogenum*	1.85	Cysteine	2009-12-23, Bokhove *et al*., 2010

acyl coenzyme A:isopenicillin N acyltransferase (acyl-coenzyme A:6-aminopenicillanic acid acyl-transferase) Mature enzyme	2X1D	*Penicillium chrysogenum*	1.64	Cysteine	2009-12-23, Bokhove *et al*., 2010
Glutamine aminotransferase class-II domain protein (SPO2029)	3MDN	*Ruegeria pomeroyi (Silicibacter pomeroyi)*	2.09	Cysteine#	2010-03-30, Ramagopal *et al*.
beta-aminopeptidase (BapA)	2N2W	*Sphingosinicella xenopeptidilytica*	1.45	Serine	2010-05-19, Heck *et al*
N-terminal beta-aminopeptidase (BapA) precursor	3N5I	*Sphingosinicella xenopeptidilytica*	1.80	Serine	2010-05-25, Heck *et al*.
Nylon hydrolase (Endotype 6-aminohexanoat-oligomer hydrolase)	3AXG	*Agromyces sp.*	2.0	Threonine	2011-04-04 Negoro *et al*., 2011
Uncleaved Pantetheine hydrolase (Putative cysteine transferase) precursor	3S3U	*Streptomyces cattleya*	1.60	Threonine	2011-05-18, Buller *et al*., 2012
Pantetheine hydrolase (Putative cysteine transferase)	3TM1	*Streptomyces cattleya*	1.80	Threonine	2011-05-18, Buller *et al*., 2012
Human asparaginase-like protein 1(hASRGL1) (L-asparaginase)	3TKJ	*Homo sapiens*	2.30	Threonine	2011-08-26, Li *et al*., 2012
ATP-dependent protease subunit HslV(Heat shock locus V)	3TY6	*Bacillus anthracis*	2.50	Threonine	2011-09-23, Kim *et al*.
20S Proteasome	3UNB	*Mus musculus*	2.90	Threonine	2011-11-15, Huber *et al*., 2012

1.14. Post-translational Activation of Ntn-hydrolases

The members of Ntn-hydrolase family are expressed as precursors which undergo intramolecular autocatalytic cleavage to remove a spacer peptide to generate the active enzyme. The maturation of precursor protein to functional proteins is long known in viral proteins (Douglass *et al*, 1984) for example, helper component proteinase of tobacco etch virus (Dougherty & Carrington, 1988, Carrington *et al.,* 1989) and peptides such as prothrombin and meizothrombin in higher eukaryotes (Petrovan *et al.,* 1998). Caspases, the important regulators of apoptosis, are activated by the removal of a linker peptide in the pro-enzyme (Stennicke & Salvesen, 1998). This maturation phenomenon was considered a hallmark of eukaryotes (Bussey, 1988) until a the post-translational processing in prokaryotic systems such as in *Bradyrhzobium japonicum* cytochrome *bc1,* (Trumpower, 1990), *Bacillus subtilis* spore coat proteins (Aronson *et al,* 1989), *Bacillus polynyxa* amylase (Uozumi *et al.,* 1989), penicilin G acylases (Thony-Meyer *et al.,* 1992) and gamma-glutamyltranspeptidase (Okada *et al.,* 2007) became available.

Detailed three-dimensional structural information of many precursors is available from mutant constructs that reduce the speed of processing e.g., glycosylasparaginase (Xu *et al.,* 1999; Wang & Guo, 2010), the β-subunit of the proteasome (Ditzel *et al.,* 1998), the PGA precursor from *E.coli* (Hewitt *et al.,* 2000), a plant L-asparaginase (Michalska *et al.,* 2008b), glutaryl 7-aminocephalosporanic acid acylase (Kim *et al.,* 2003) and cephalosporin acylase (Kim *et al.,* 2001). These precursor structures have helped to decipher the mechanism of processing. During post-translational activation of Ntn-hydrolases such as AGA, PGA, cephalosporin acylase, and glutaryl 7-aminocephalosporanic acid acylase, the precursor polypeptide undergo cleavage removing a spacer chain that results in the chain splitting into two in the processed protein, whereas in PVA, proteasome and glutamine PRPP amidotransferase, the activation is achieved by the removal of a small pro-peptide. This autocatalytic processing considered a common feature of Ntn-hydrolases, is absent in the conjugated bile salt hydrolases (BSH), where reported structures are available. Sequence shows that it is produced with the nulcleophile residue present at the second position in the N-

terminus, which gets exposed after enzymatic removal of initiation formyl-methionine without autocatalytic processing.

A range of autocatalytic processing events proceed *via* protein splicing in many organisms from prokaryotes to eukaryotes (Perler *et al*. 1994), where the initial steps are thought to involve an N $\rightarrow$ S or N $\rightarrow$ O acyl rearrangement to generate a branched intermediate (Shao *et al.,* 1996). The hydrolysis of the peptide bond preceding the N-terminal nucleophile residue in Ntn-hydrolases is similar to the initial step in the protein splicing pathway involving internal peptide bond rearrangement through an N$\rightarrow$O or N$\rightarrow$S acyl shift (Lee & Parks, 1998; Lee *et al.,* 2000a, b; Xu *et al.,* 1999) and cleavage step that generates the catalytic α-amino group. Here we would concentrate exclusively on the specific roles of residues surrounding the nucleophile residue SerB1 or Serβ1, in PGAs.

1.15. Autocatalytic Processing of PGAs

Considerable progress has been made in the understanding of post-translational modification process in prokaryotic systems (Thony-Meyer *et al.,* 1992). The most studied maturation pathways of PGA have been that of *E. coli* ATCC 11105. The enzyme from *E.coli* ATCC 11105 was the first to be purified to homogeneity and crystallized. PGA from *E.coli* is produced as 96 kDa cytoplasmic single chain precursor pre-pro-protein consisting of 846 amino acids. The gene structure of the open reading frame of *E.coli* PGA consists of a 26 amino acid signal peptide, which directs the protein and exports into the cytoplasm and a 54 amino acid spacer between the A and B chains, removed post-translationally. The A and B chains are 209 and 557 amino acids in length, respectively, and form a periplasmic 86 kDa heterodimer (Schumacher *et al.,* 1986; Duggleby *et al.,* 1995). Here, the post-translational processing essentially consists of two steps: translocation of the precursor to the periplasmic membrane utilising the twin arginine translocation (tat) machinery and using the 26-amino acid signal peptide which is subsequently cleaved off (Schumacher *et al.,* 1986; Hewitt *et al.,* 2000; Ignatova *et al.,* 2002) followed by autocatalytic peptide-bond cleavage of 54-amino-acid spacer peptide that connects the A and B chains (Oliver *et al.,* 1985). At least four proteolytic cleavages have been proposed for the autocatalytic removal of the spacer peptide to generate mature enzyme (Lee *et al.,* 2000a; Bock *et al.,* 1983; Oh *et*

al., 1987; Sizmann *et al.,* 1990). This autocatalytic processing leads to the formation of a heterodimeric enzyme with two polypeptide chains, A and B, and a free N-terminal serine in the B-chain which is essential for activity of the enzyme (Duggleby *et al.,* 1995). However, the B-chain alone is not catalytic (Daumy *et al.,* 1985), since elements of the substrate specificity site are contributed also by the A chain. Kinetic studies showed that the autoproteolysis in PGA is intramolecular (Kasche *et al.,* 1999).

Thus, the maturation pathway of the precursor protein passes through the following steps:

A. The synthesis of precursor polypeptide (pre-pro-PGA) containing a signal peptide, A-chain, spacer peptide and B-chain from the *pac* gene.

B. The precursor protein is transported across the cytoplasmic membrane into the periplasmic space followed by removal of signal peptide (pro-PGA);

C. Proteolytic cleavage at the N-terminus of the B-chain.

D. Removal of spacer peptide from the A-chain by sequential C-terminal proteolytic processing in two or three steps.

1.16. Structural Studies of PGA from *E.coli*

The crystal structure of the processed form of PGA from *E. coli* at 1.9 Ắ was reported by Duggleby *et al.* (1995) (**Figure 1.11**). The mature *Ec*PGA is kidney shaped in cross-section, contains two anti-parallel β-sheets stacked against one another, one showing a normal twist and the other essentially planar. A β-sandwich comprising the N-terminal residues of A-chain and much of B-chain is flanked by α-helices. The A-chain packs on the surface of the B-chain and the two chains are closely intertwined. The packing of the two chains results in a deep depression on the enzyme's surface, at the bottom of which lies Ser264, the catalytic residue at the N-terminus of the B-chain. The reported structure of the enzyme complexes with PAA and Phenylmethylsulphonoyl fluoride (PMSF) located the binding site for the side chain of the substrate. The phenyl moiety of each of these compounds directed towards a deep depression on the surface of the enzyme. The depression contains a hydrophobic pocket, which is lined with a number of aromatic residues and hydrophobic side chains. The phenyl moiety in each of these compounds is located in this hydrophobic binding

pocket. The complementary fit explains the specificity of PGAs towards the phenyl moiety of a broad range of substrates. The serine residue on the N-terminal of the B-chain is found at the mouth of the binding pocket and is involved in catalysis. Absence of any catalytic base residue at the active site of *E. coli* PGA led to the hypothesis that penicillin acylase has a single-amino-acid catalytic centre (Duggleby *et al.,* 1995). It has been shown that there is no histidine equivalent as found in serine proteases, distinguishing PGAs from them. The base is contributed by the N-terminal residue's own α-amino group. This led to the classification of these types of enzymes as a new type of hydrolases- the N-terminal nucleophile (Ntn) hydrolases (Brannigan *et al.,* 1995).

Structure of a Thr263Gly mutant (Kasche *et al.,* 1999), the slow processing precursor of *E. coli* penicillin acylase, showed that both the mature and precursor structures possessed very similar shape and unit cell dimension with only 2 Å increase in b-axis and 5 Å increase in c -axis (Hewitt *et al.,* 2000). Global superposition of Cα-atoms of both the structures showed that folds of A and B chains in the two structures are extremely similar, confirming that they are arranged in the precursor essentially as in the mature enzyme. Structure has revealed the presence of spacer peptide that fills the depression seen the processed enzyme and blocks the active site cleft (**Figure 1.12**). The presence of spacer peptide accounts for the slight increase in cell volume. Careful inspection of the scissile peptide bond and its environment in the mutated precursor molecule helped in understanding the chemistry of the autocatalytic processing event in PGAs. The catalytic serine side chain showed two conformations in the crystal, one of which corresponds to the Oγ in the active enzyme. The second Oγ conformation puts it in a position appropriate for a direct nucleophilic attack on the carbonyl carbon of the scissile bond and this is obviously relevant to processing. Inspection of the environment around Oγ shows however there are no groups that can act as a base to extract the proton. Thus the reaction is presumably driven by the physical approach of Oγ close to the peptide brought about by the local structure (Hewitt *et al.,* 2000).

A calcium ion was present at the interface between A and B chains in the mature PGA enzyme structure from *E. coli* and *P. rettgeri* as well as in the precursor structure of *E.coli*. Calcium ion is coordinated by the residues, namely the side chains of aspartyl residues 336, 339 and 515, and the main-chain carbonyl groups of Val338

and Pro468 from chain B and the side-chain of Glu156 from chain A in the case of *E.coli*. Presence of calcium ion at same position coordinating by same residues in all the PGA structures reported so far and residues from both chains coordinate this ion, suggesting its importance in PGAs. Ignatova *et al*. (2005) showed that in a refolding experiment of enzyme denatured by 8M urea, without Ca^{2+}, the tertiary structure was established to a large extent, but the autoproteolytic removal of the pro-peptide is severely retarded. Further addition of Ca^{2+} stabilizes the pro-PA structure by protecting some hydrophobic patches. Thus Ca^{2+} is not directly involved in the folding process but has essentially stabilizing role. It shifts the folding equilibrium towards the stable native conformation and facilitates autoprocessing (Ignatova *et al*., 2005).

1.17. Catalytic Mechanism of PGAs

All Ntn-hydrolase enzymes carry out nucleophilic attack on the carboxy carbon of amide bond. The substrates however vary from peptides (the proteasomes), amino acid side chains, (Glutamine amido transferase, GATs), amidated sugars (Aspartyl glucosyl amidase, AGAs) and amides (penicillin G acylases, PGAs, and penicillin V acylases, PVAs).

The reported native structure of PGA from *E. coli* revealed a single amino acid serine as catalytic centre (Duggleby *et al.,* 1995). The enzyme structure solved with complexes including substrate analogues and inhibitors with native enzyme and mutants (Done *et al.,* 1998; Brannigan *et al.,* 2000) has helped to advance our understanding of the enzyme's mode of substrate binding, specificity, mechanism of action and stability. The amide bond of Pen G, the preferred substrate of PGA, between the penam ring and the side chain is the site of cleavage by PGA. The active site of PGA is a deep binding pocket mainly lined with aromatic amino acid residues and hydrophobic side chains involved in substrate binding (Done *et al.,* 1998; Alkema *et al.,* 2004). The hydrophobic residues that form the acyl group binding site in PGA are the main contributors to substrate recognition and binding, resulting in its specificity for Pen G (McVey *et al.,* 2001; Alkema *et al.,* 2000). PGA catalysis proceeds via an acyl-enzyme intermediate. Their mode of action is similar to serine proteases like chymotrypsin (Kato, 1980) but lacks the catalytic triad. The role of intermediary water molecule in the catalytic mechanism has been disputed (McVey *et al.,* 2001). The

hydroxyl group of serine that is located at the N-terminal of B-chain (SerB1) acts as a nucleophile.

There are two mechanisms proposed for the initial reaction. The hydroxyl group of the N-terminal serine (SerB1) is activated by its own α-amino group via a bridging water molecule (Duggleby *et al.,* 1995; Dodson, 2000). However, in 2001, McVey *et al.* (2001) reported the structure of PGA in complex with Pen G sulphoxide, a poor substrate. In that complex, no water molecule was found adjacent to the Oγ atom of SerB1, opening up the possibility of catalysis proceeding after direct activation of nucleophile atom by α-amino group instead of the previously suggested tightly hydrogen bonded water molecule acting as a "virtual" base.

Once activated, the serine oxygen attacks the carbonyl carbon atom of the substrate, forming a tetrahedral intermediate that is stabilized by the main chain amide of AlaB69 and the side chain nitrogen of AsnB241. Rearrangement of electrons leads to the collapse of this intermediate resulting in the release of leaving group and covalent acyl-enzyme intermediate remaining. The acyl-enzyme intermediate is subsequently deacylated by a nucleophile water molecule, which again pass the formation of a tetrahedral intermediate to free the enzyme and the product (**Figure1.13**) (Duggleby *et al.,* 1995). The enzymatic reaction is reversible, although the synthetic reaction is only favoured in acidic conditions, where the protonation of the carboxylic acid will increase the susceptibility of the carbonyl carbon to nucleophilic attack. In this way, β-lactam antibiotics can be cleaved to yield the free β-lactam nucleus and the side-chain carboxylic acid when the nucleophilic attack is performed by water (hydrolysis). However, when the acyl donor is an activated synthetic side chain (either an amide or an ester), a nucleophilic attack by a β-lactam nucleus such as 6-APA will yield a semi-synthetic β-lactam antibiotic through the process called aminolysis. A similar catalytic mechanism is identified in several other Ntn-hydrolases as well.
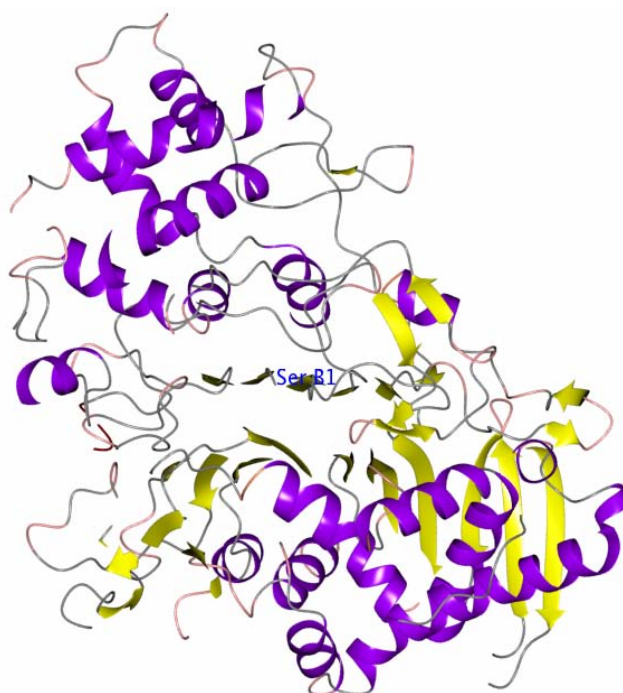
Figure 1.11: Three-dimensional structure of PGA from *E.coli* (PDB code: 1PNK, Duggleby *et al.,* 1995) showing the heteromeric association of subunits and possessing αββα fold.
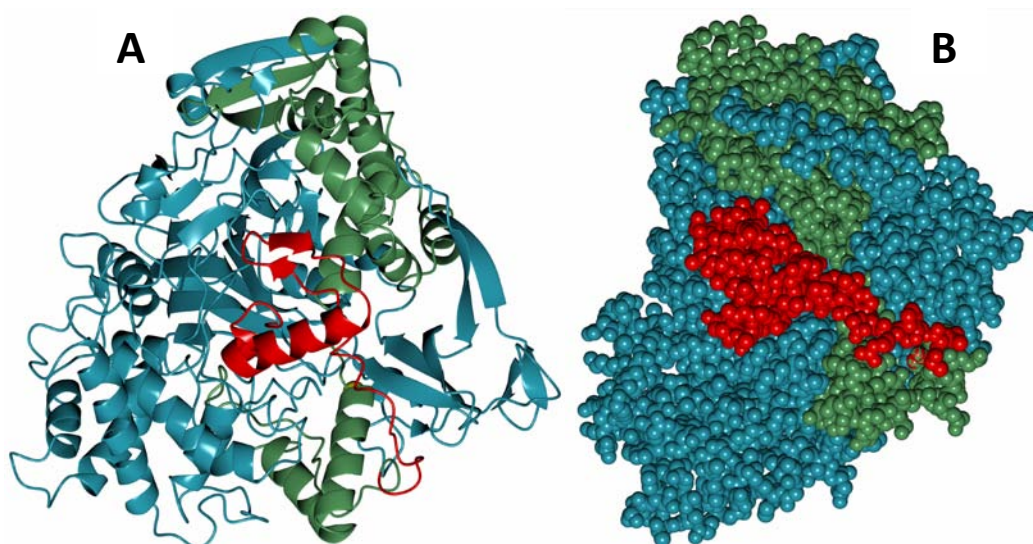
Figure 1.12: Structure of slow processing mutant precursor of PGA from *E .coli* (PDB code: 1E3A; Hewitt *et al.,* 2000) (A) in cartoon model; (B) in space filled model. The A-chain is shown in green, the B-chain in blue and the spacer peptide in red. The nucleophilic serine is shown in yellow although barely visible at the base of the deep active-site cleft. The space filled model clearly shows how the spacer masks the entrance to the active site.
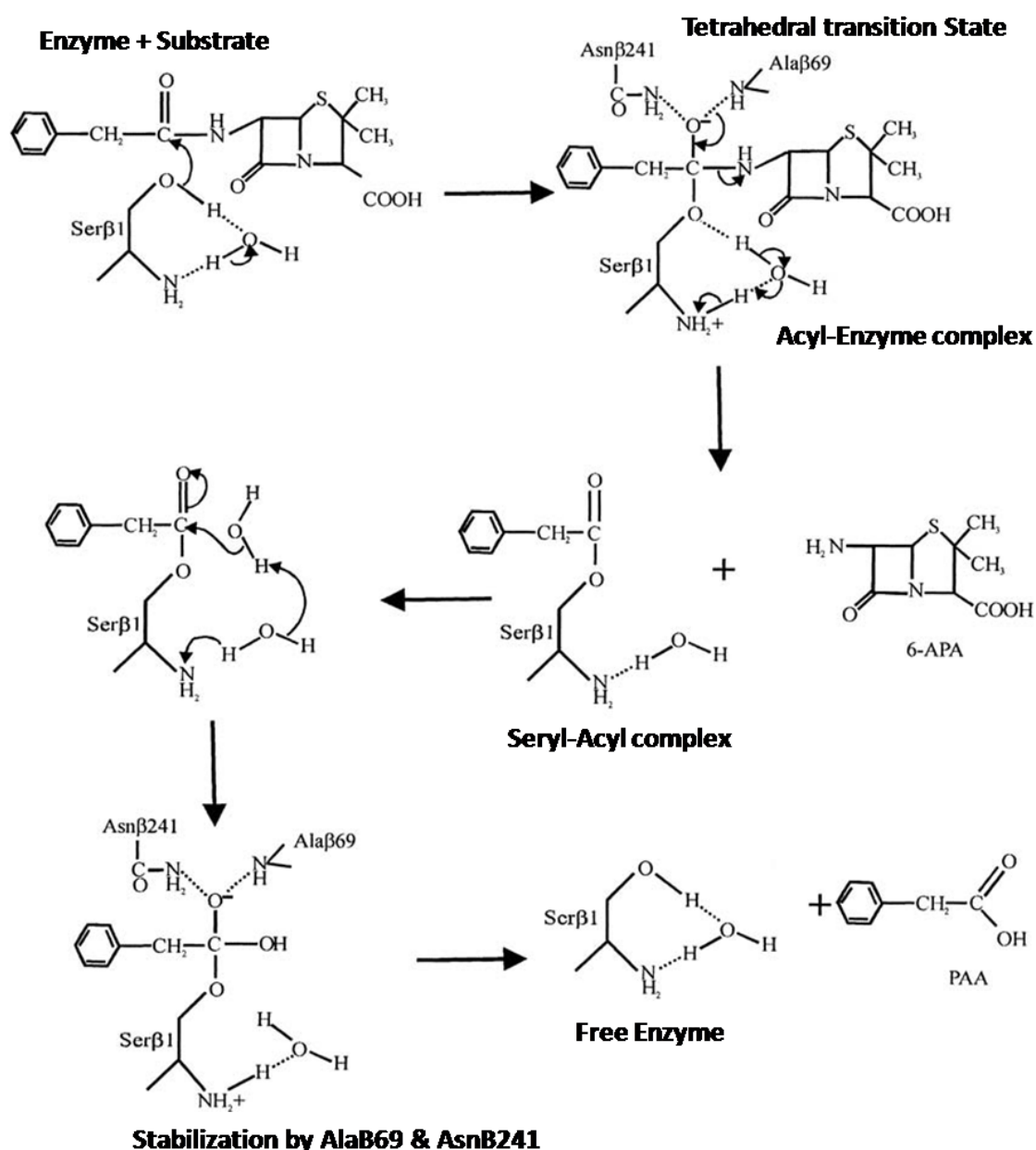
Figure 1.13: Proposed mechanism for action of hydrolysis by PGA. The carbonyl carbon atom of the amide bond of pen G is attacked by the SerB1 and a covalent acyl-enzyme is formed through a tetrahedral transition state, which is further stabilized by H-bonding to AsnB241 and AlaB69. The tetrahedral intermediate then forms a seryl-acyl enzyme and releases the 6-APA. The acyl enzyme is attacked by water to form a second tetrahedral intermediate by the same mechanism with the interaction of AsnB241 and AlaB69. Finally, this intermediate collapses to release the PAA. Dotted lines represent possible hydrogen bonds (adapted and modified from Duggleby *et al.,* 1995).

1.18. Studies on PGAs from other Sources

PGA from *E. coli* is the most studied and commercially commonly used enzyme (Demain, 2000; Gabor *et al.,* 2005). This protein has guided research on PA from other organisms. The crystal structure of PGA from the *Providencia rettgeri* (*Pr*PGA) is similar to that of *Ec*PGA (**Figure 1.14**) (McDonough *et al.,* 1999). The wild-type *Pr*PGA shows 64% sequence identity and 75% sequence similarity with the *Ec*PGA. PGA from *Proteus rettgeri* is an 85.8 kDa enzyme composed of two essential non-identical chains of 23.7 kDa and 62.1 kDa. Together the A- and B-chains were reported to be made up of five distinct structural regions: a β-sandwich, the largest, two predominantly helical regions, a partial β-barrel, and a small β-strand region. Like *Ec*PGA, the B-chain of *Pr*PGA contained a serine residue required for the enzymatic activity and possessed αββα fold characteristic of Ntn-hydrolases. Three-dimensional structural homology between *Pr*PGA and *Ec*PGA also revealed the analogy of auto catalytic processing for this enzyme. The four residues: the catalytic SerB1, plus GlnB23, AlaB69, and AsnB241 present in the active site of the *Ec*PGA are found conserved. The hydrophobic pocket binding substrate is formed by residues from both the A- and B-chains (McDonough *et al.,* 1999).
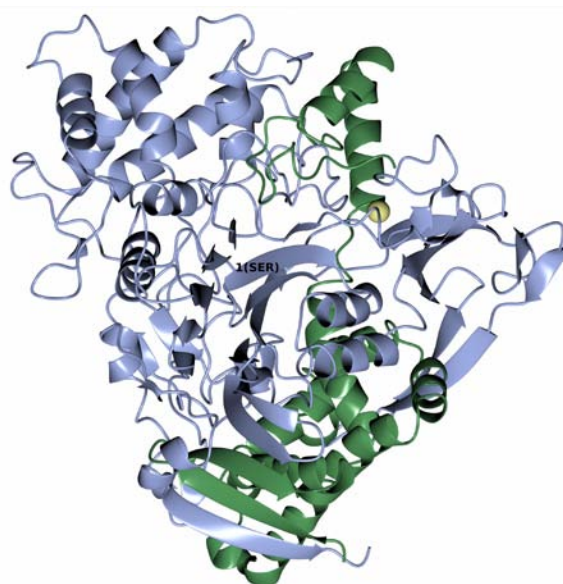
Figure 1.14: Crystal structure of PGA from *P. rettgeri* (PDB code: 1CP9; McDonough *et al.,* 1999) showing αββα fold. The A-chain is shown in green and the B-chain in blue. The calcium ion present is shown as sphere in yellow colour.

There is interest in PGA enzyme from *K. citrophila* (*Kc*PGA), which tolerates harsher conditions such as higher temperature, acid/alkaline pH and changes in solvent composition. The enzyme is easier to immobilise for pharmaceutical industry applications (Alvaro *et al*., 1992; Fernandez-Lafuente *et al.,* 1991 & 1996; Liu *et al.,* 2006). The stabilization of the enzyme by immobilization is also reported (Guisan *et al*., 1993). PGA from *Kluyvera citrophila* ATCC 21285 is reported to be composed of two non-identical subunits of 23 kDa and 62 kDa (Barbero *et al.,* 1986), in contrast to molecular weight of around 63 kDa as reported by Shimizu *et al*. (1975). Its nucleotide sequence is 80% similar to *E. coli* ATCC 11105 PGA (Schumacher *et al.,* 1986), indicating a common ancestor. The cloning of this enzyme and overexpression in *E. coli* had been reported (Garcia *et al*., 1985; Barbero *et al*., 1986; Wen *et al.*, 2004 & 2005).

In addition, MEROPS database (Rawlings *et al.,* 2004) also shows presence of penicillin acylase homolgues in *Sulfolobus solfataricus*, *Aeropyrum pernix*, *Pyrobaculum aerophilum*, *Archaeoglobus fulgidus*, *Thermoplasma acidophilum* and *Thermoplasma volcanium*. PGA has been isolated from two fungal sources, *Aspergillus fumigatus* H/6.17.3 and *Mucor griseocyanum* H/55.1 (Jose *et al.,* 2003). PGA from *Bacillus megaterium* has been cloned (Mc Cullough, 1983) and being used in industry for chemical synthesis. A thermostable PGA from *Achromobacter xylosoxidans* has been cloned and characterized (Cai *et al.,* 2004). The enzyme is a heterodimer with α and β subunit with molecular masses of approximately 27.0 kDa and 62.4 kDa, respectively. The study has suggested that the increased number of buried ion pair networks, lower Asn and Gln contents, excessive Arg residues, and remarkably high content of Pro residues in the structure of PGA can contribute to its high thermostability. The presence of protein-coding genes of penicillin acylase in *Bacillus anthracis* has been predicted by analytical work of Read *et al.* (2002). Unlike other gram negative microorganisms that accumulate the PGA in the periplasmic space, an acylase from gram positive *Arthrobacter viscosus* secretes enzyme into the culture medium (Ohashi *et al*., 1988, 1989). The enzyme was successfully cloned and expressed in *E. coli* and *Bacillus subtilis* (Ohashi *et al*., 1989) and the sequence was reported by Konstantinovic *et al*. (1994). Penicillin acylase from *A. viscosus* is a heterodimer (Ohashi *et al*., 1988), produced as a single polypeptide that is further

processed; it is constituted by the A (24 kDa) and the B (60 kDa) chains having a serine as terminal amino acid in the B-chain. Both *A. viscosus* and *B. megaterium* expresses the enzyme extracellularly. Rajendhran *et al*. (2007) cloned a *pac* gene from *Bacillus badius* and expressed in *Escherichia coli* DH5α. Purified enzyme showed two subunits with apparent molecular masses of 25 kDa and 62 kDa.

During last two decades, a PGA from *A. faecalis* (*Af*PGA), which showed 40% protein sequence identity with *Ec*PGA (Kasche *et al.,* 2003), became a subject of intensive investigations (Svedas *et al.,* 1997; Verhaert *et al.,* 1997). It has a clear industrial advantage over other penicillin acylases in β-lactam conversion because of its higher thermostability and synthetic efficiency in enantioselective synthesis. This makes *Af*PGA a more attractive biocatalyst for both hydrolysis and synthetic conversions.

The enzymes from different sources had similar substrate specificity but different molecular weights, isoelectric point, and electrophoretic mobility in polyacrylamide gels and did not antigenically cross-react. However, most PGAs have a similar subunit configuration, structure and substrate range as *Ec*PGA. This indicates divergent evolution although they have evolved beyond any obvious sequence homology.

1.19. Important Residues for Catalysis and Autoproteolysis of PGAs

In Ntn-hydrolases the nucleophile residue responsible for enzyme catalysis also acts as the nucleophile for the autocatalytic processing (Schmidtke *et al.,* 1996; Lee & Park, 1998; Li *et al.,* 1999; Hewitt *et al.,* 2000; Kim *et al.,* 2002). Other residues were also predicted to be taking part in both the catalytic and autocatalytic reactions. In order to gain insight into the mechanism of activation of the precursor, several mutational studies have been performed using mutations of the residues where bond cleavage occurs during the activation process. Site-directed mutagenesis of the residue at the enzyme's active site (Choi *et al.,* 1992), or at the processing site (Choi *et al.,* 1992; Hewitt *et al.,* 2000) affects the processing of the precursor suggesting that the two processes use the same set of residues (Brannigan *et al.,* 2000). In the case of *Ec*PGA a mutation Thr263Gly in pro-enzyme (numbering as in pro-PGA) generated a slow processing precursor (Kasche *et al.,* 1999). Thus, Thr263 must be playing an important

role in the rate of autoproteolytic reaction. Another mutant Ser264Cys resulted in precursor processing, but the resulting mature enzyme was inactive (Choi *et al.,* 1992). Further importance of the Ser264 residue was demonstrated by sequential mutations of this residue to Argine, Threonine, Glycine (Choi *et al.,* 1992) or Alanine (McVey *et al.,* 1999) that always yielded the unprocessed PGA precursor. In addition, the enzyme gets inactivated by PMSF, an affinity labelling for serine residue in serine proteases (Kutzbach & Rauenbusch, 1974) and the covalent modification of this Ser264 by PMSF has been shown in the electron density of the PGA-PMSF complex crystal structure (Duggleby *et al.*, 1995). Hence Ser264 is important in both autocatalytic processing as well as in the hydrolysis of substrates.

Residue ArgB263 is also implicated in processing events. The mutant ArgB263Ile did not process from the precursor form. ArgB263 is also involved in substrate binding, and helps to position AsnB241. Both ArgA145 and PheA146 are important in substrate positioning in the active site. Mutational studies using AsnB241Ala mutant, showed that binding of the substrate induces a conformational change due to steric interactions between the substrate and PheA146. The conformational change involves displacement of residues PheA146 and ArgA145 from the active site by about 3 Å and 8 Å, respectively, reorienting them toward the solvent to form an open configuration, allowing the enzyme to accommodate penicillin G (Alkema *et al.,* 2000). Substrate binds to PheB24, SerB67 and PheB57 residues via hydrogen bonding (Oinonen & Rouvinen, 2000). AsnB241 contributes to the oxyanion hole as well as in positioning both substrate and the active site SerB1 (McVey *et al.,* 2001; Done *et al.,* 1998). Some of the conserved important amino acid residues are listed out in **Table 1.3.**

Table 1.3: Catalytically important amino acid residues in PGAs. Numbering is according to *Ec*PGA.

Nucleophile	Processing	Oxyanion hole	Substrate binding	Calcium ion coordination
Ser B1	Thr263	Ala B69	Phe B24	Asp B73
	Ser B1 (Ser 264)	Asn B241	Phe B57	Val B75
	Lys B6 (Lys 299)	Gln B23	Ser B67	Asp B76
	B263			Pro B205
				Asp B252
				Glu A152 (A150 in *P.rettgeri* PGA & A153 in *A. faecalis* PGA)

1.20. Industrial Applications of PGAs

In the 1950's it had been shown that an enzyme, later named as penicillin acylase, could hydrolyze penicillin (Sakaguchi & Murao, 1950; Murao, 1955). Later Rolinson *et al*. (1960) demonstrated the practical importance of this enzymatic reaction, in which 6-APA was produced through the hydrolysis of natural penicillin, providing an elementary starting point for the production of semi-synthetic penicillins. Indeed the naming of penicillin acylase has been a reflection of this reaction and its industrial application in the synthesis of antibiotics.

1.20.1. 6-APA Production

The worldwide demand for semi-synthetic penicillins, as the most successful and cost-effective antibiotics, has brought 6-APA into central position as a major pharmaceutical product (Lowe *et al*., 1981). Historically the β-lactam nucleus 6-APA was first isolated on an industrial scale from fermentation by *Penicilliun chrysogenum* (Ballio *et al*., 1961; Batchelor *et al*., 1959). It was then produced by the enzymatic synthesis of precursors Pen V or Pen G by the supplementation of phenoxyacetate or phenylacetate, respectively to commercial *Penicillium chrysogenum* fermentation (Claridge *et al*., 1960, 1963; Kaufmann & Bauer, 1964; Rolinson *et al*., 1960) and then

by the removal of their side chain either by chemical splitting or enzyme catalysed cleavage. The enzymatic cleavage of natural penicillins that was initially performed using cell free extracts was not found commercially competitive with the more efficient chemical deacylation process. High cost, thermal and pH degradation, nonreusability and difficulty to recover the solution form of PGA have prompted immobilization of the enzyme on solid supports. The introduction of immobilized enzymes, which allows multiple use of the catalyst, is now the method of choice in the industrial production of 6-aminopenicillanic acid (6-APA) (Savidge & Cole, 1975; Shewale *et al.*, 1990; Sudhakaran & Shewale, 1993). Immobilized PGA is mainly utilized in this process, which accounts for 88% of the worldwide production of 6-APA.

1.20.2. Ring Expansion

The manufacture of therapeutically important cephalosporins can be achieved also by a chemical ring expansion of the thiozolidine ring to a dihydrothiazine in Pen V or Pen G (De Koning *et al.,* 1975; Verweij & de Vroom, 1993). During the course of this reaction the amino group remains protected as phenoxyacetyl or phenylacetyl amide which is finally removed by PGA. Of particular importance is the choice of a suitable protecting function for the COOH group. It must be stable during the ring expansion but could be removed without damaging the cephalosporin nucleus. This has been achieved using phenylacetyloxymethylene (Baldaro *et al.,* 1988).

1.20.3. Industrial Production of *β*-Lactam Antibiotics

Semi-synthetic penicillins exhibit better properties– such as increased stability, easier absorption and fewer side effects– than penicillin G and V, and represent a practical solution to the problem of adaptive microbial resistance to antibiotics. Use of penicillin acylase as catalyst in the synthetic direction was first demonstrated in 1960 by Kaufman and Bauer, who reported the *E. coli* penicillin acylase catalyzed formation of penicillin G from 6-APA and phenylacetic acid (Kaufmann & Bauer, 1960; Bruggink *et al.,* 1998). Industries that produce β-lactam antibiotics have introduced PGA as biocatalyst by replacing multistep conversion processes with cheap and promising enzymatic conversion, which has an efficiency of ~80–90% (Elander, 2003). 6-APA obtained by the enzymatic deacylation of natural pencillins and 7-ADCA

obtained by chemical or enzymatic ring expansion followed by deacylation, are transformed into semi-synthetic penicillins and semi-synthetic cephalosporins, respectively by condensation with a D-(-)-phenylglycine or a D-(-)-4-hydroxyphenylglycine derivative catalyzed by penicillin acylases, to generate semi-synthetic penicillin or semi-synthetic cephalosporins respectively (Bruggink *et al*., 1998).

1.20.4. Penicillin acylase-catalyzed Peptide Synthesis

Penicillin acylases can be used for the protection and deprotection of the amino groups of amino acids by direct enzymatic synthesis and acyl group transfer reactions. These strategies have been applied to the synthesis of peptides and their derivatives. As a classical example, penicillin G acylase from *E. coli* proved to be an efficient biocatalyst in the synthesis of the sweetener aspartame (Fuganti *et al*., 1986). The same enzyme has recently been used to prepare d-phenyl dipeptides, whose esters readily undergo ring closure to the corresponding diketopiperazines, which find application as food additives, chitinase inhibitors and as synthons for fungicidal, antiviral and anti-allergic compounds (van Langen *et al*., 2000).

1.20.5. Penicillin acylase-catalyzed Resolution of Racemic Mixtures

Penicillin acylases have been shown to efficiently resolve racemic mixtures of chiral compounds such as amino acids (Fadnavis *et al*., 1997; Bossi *et al*., 1998), β-amino esters (Roche *et al*., 1999), amines (Guranda *et al*., 2001) and secondary alcohols (Svedas *et al*., 1996) in aqueous media as well as in water-cosolvent mixtures and anhydrous organic media. The resulting pure enantiomers can be used as intermediates for the synthesis of biologically active compounds. For instance, a racemate mixture of ethyl-3-amino-4-pentynoate can be resolved through an enantioselective acylation catalyzed by penicillin G acylase from *E. coli* to yield the S-isomer, which can be used as a chiral synthon for the synthesis of the anti-platelet agent Xemilofiban (Topgi *et al*., 1999). Another important application of penicillin G acylase is represented by the kinetic enantioselective acylation of the racemic azetidinone intermediate for the synthesis of the carbacephalosporin antibiotic Loracarbef, an analogue of Cefaclor (Zmijewski *et al*., 1991). Enantioselective acylation of the L-

enantiomers of methyl esters of phenylglycine and 4-hydroxyphenylglycine in organic solvents was shown to be catalysed by penicillin G acylase at controlled water activity (Basso *et al*., 2000).

Part II

1.21. Post-translational Modification and Myristoylation

Post-translational modification in eukaryotes result changes in protein sequences and sometimes induces new molecular functions that cannot be directly encoded by the gene sequence. This includes internal protein changes e.g., disulfide bond formation, removal of amino acid residues through proteolysis, N-terminal initiator methionine removal, attachment of other proteins e.g., ubiquitination, N-terminal or side chain modifications e.g., acetylation, arginylation, lipidation by myristoylation or palmitoylation, methylation, phosphorylation, glycosylation and sulfonic acid or nucleotide attachment. Modifications such as these have incredibly diverse biological functions and are known to be involved in signalling cascades, protein maturation, protein trafficking and localisation, extracellular communication, protein regulation etc.

Myristoylation is the attachment of a 14-carbon saturated fatty acid, myristic acid, to the N-terminal glycine of a subset of eukaryotic proteins (Gordon *et al*., 1991; Boutin, 1997). N-Myristoylation plays a role in targeting proteins to membrane locations, metabolism, mediating protein–protein interactions, stabilizing protein structures and effecting turnover number (Resh, 1999, 2006). Although often referred to as a post-translational modification it also occurs co-translationally when fewer than 100 residues have been polymerised by the ribosome and follows removal of the leader methionine residue by a methionine aminopeptidase to expose an N-terminal glycine (Wilcox *et al*., 1987). Myristate is present in only small amounts (less than stearic acid), suggestive of a special role for myristoylation which would not be served by the addition of any other lipid. This high degree of specificity of myristate for amide linkages, suggests the possible existence of distinct classes of protein acyltransferases.

1.22. Myristoyl-CoA:protein *N*-myristoyltransferase (NMT)

The ubiquitous enzyme that catalyses the co- and posttranslational transfer of C14:0 fatty acid, myristate, via an amide bond formation, is known as myristoyl-CoA:protein N-myristoyltransferase (NMT) (**Figure 1.15**). NMT has been first characterised in 1987 in *Saccharomyces cerevisiae* (Towler *et al*., 1987), subsequently has been isolated and characterised from yeast and fungi (*Candida albicans,* (Wiegand *et al*., 1992)*, Cryptococcus neoformans* (Lodge *et al.,* 1994), *Aspergillus nidulans* (Shaw *et al*., 2002), parasitic protozoans (*Plasmodium* species (Gunaratne *et al*., 2000), *Leishmania major* (Panethymitaki *et al*., 2006), *Trypanosoma brucei* (Panethymitaki *et al*., 2006), *Leishmania donovani* (Brannigan *et al.,* 2010)), *Drosophila melanogaster* (Ntwasa *et al*., 2001), plants (wheat germ (Heuckeroth *et al.,* 1988)*, Arabidopsis thaliana* (Boisson *et al*., 2003) and mammals (including mouse, rat, cow and human). NMT has been shown to be essential for the survival of *S. cerevisiae* (Duronio *et al*., 1989), *C. albicans* (Weinberg *et al*., 1995), *C. neoformans* (Lodge *et al*., 1994) and the bloodstream forms of the parasites *L. donovani* (Brannigan *et al*., 2010), *L. major* and *T. brucei* (Price   *et al*., 2003). It is shown to be important in the development of Drosophila (Ntwasa *et al*., 2001) and mice (Yang *et al*., 2005) as well. However, all mammalian species, including humans, appear to have two NMT enzymes encoded by different copies of genes (Giang & Cravatt, 1998). The human isozymes, *Hs*NMT1 and *Hs*NMT2, share 77% identity suggesting two distinct families of NMT and have homologues in mouse and rat (Giang & Cravatt, 1998; Rioux *et al*., 2006). Each isozyme is highly conserved amongst mammals: for example, *Hs*NMT1 and mouse NMT1 share 97% identity and *Hs*NMT2 and mouse NMT2 share 96% of their amino acid sequence (Giang & Cravatt, 1998), bovine NMTs showed 76% identity with human NMT2 (Selvakumar *et al*., 2007), suggesting a specific role for each isozyme in mammalian physiology or development.

The phylogenetic tree analysis of the NMT family reveals that these proteins can be grouped into seven distinct families based on the conservation of amino acids between various NMT species of divergent origin. Bovine NMTs, human NMT1, mouse NMT1 can be grouped as subfamily-1; human NMT2 and mouse NMT2 can be grouped as subfamily-2; Drosophila as subfamily-3; *C. elegans* as subfamily-4; *S. cerevisiae* and *C. albicans* as subfamily-5; *H. capsulatum* as subfamily-6 and *C.*

neoformans as subfamily-7 (Selvakumar *et al*., 2007). It is interesting to note that the essential amino acids, which are involved in the catalytic activity of the NMT, are highly conserved in various species of divergent origin (Raju *et al*., 1996; Ntwasa *et al*., 1997; Zhang *et al*., 1996). Though the species might have diverged many years ago, there is a high degree of conservation of the NMT structure at the nucleotide as well as at the amino acid level. This high conservation might have some evolutionary and functional implications concerning this gene family.

There is no precisely defined "myristoylation motif", although substrate specificity can be rationalised to a certain extent and the enzyme has an absolute requirement for a peptide or protein substrate with glycine at the N-terminus. NMT also appears to be very specific for transfer of myristate, at least *in vivo* (Wright *et al*., 2010; Farazi *et al*., 2001).
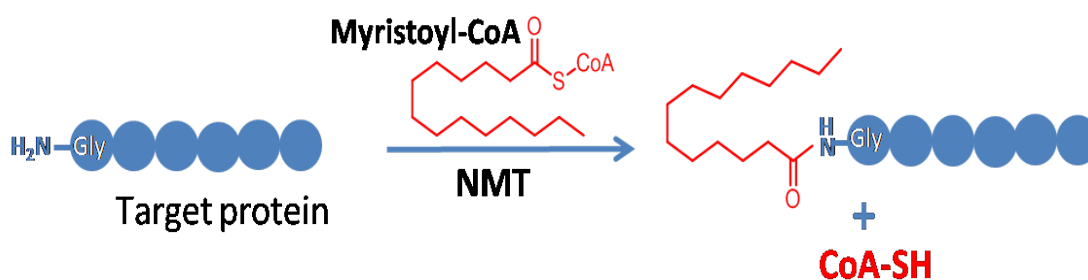
Figure 1.15: Scheme showing Myristoylation of proteins by NMT.

1.23. *In vivo* role of N-terminal Myristoylation of Proteins

N-terminally myristoylated proteins consist of large family of essential eukaryotic and viral proteins with many different functions, and they are located either in the inner membrane of cells or in the cytosol, or both. The prevalence of N-terminally myristoylated proteins in eukaryotes has been estimated at between 0.5% and 3% of the cellular proteome (Maurer-Stroh *et al*., 2002; Martinez *et al*., 2008). These proteins have a broad range of functions and include protein kinases and phosphatases, Gα proteins (Schultz *et al*., 1987), nitric oxide synthase (Braam & Verhaar, 2007), calcineurin B (Aitken *et al*., 1982) and the catalytic subunit of cyclic AMP-dependent protein kinase (Carr *et al*., 1982), several ARF proteins involved in ADP ribosylation (ARFs and ARLs; Johnson *et al*., 1995), various oncoproteins including pp60[src], pp56[lck] (Schultz *et al*., 1985; Semba *et al*., 1986; Sukegawa *et al*.,

1987), a calpain-type protease, the infective stage-specific hydrophilic acylated surface proteins (HASPs), flaggelar calcium binding proteins and membrane or cytoskeleton-associated structural proteins such as myristoylated alanine-Rich C Kinase Substrate (MARCKS) (Stumpo *et al.*, 1989; Resh, 1999; Arbuzova *et al.*, 1998). Many of these proteins are involved in signalling cascades. HASPB, which requires both N-myristoylation and palmitoylation for transport to the parasite plasma membrane, is a vaccine candidate for Viceral Leishmaniases (Stager *et al.*, 2000; Moreno *et al.*, 2007). In addition, several viral proteins e.g., Gag, VP4 coat protein, membrane fusion protein p15 etc and more recently, some bacterial proteins have been reported to be myristoylated, usually by the eukaryotic host cell NMT *in vivo* (Maurer-Stroh & Eisenhaber, 2004). About 150 human proteins are known to be myristoylated including tyrosine kinases.

There is evidence that post-translational myristoylation of proteins may be an important mechanism in the regulation of programmed cell death, or apoptosis (Mishkind, 2001). In humans, protein myristoylation is connected with several diseases, including cancer (Felsted *et al.*, 1995), genetic disorders (Cordeddu *et al.*, 2009), and infection (Provitera *et al.*, 2006). However, the role of myristoylation is still not completely understood, and not all myristoylated protein structures have been experimentally determined (Heal *et al.*, 2008a, b).

The myristoyl moiety can perform one or a combination of several roles in the cell:

❖ regulating protein activity presumably by modifying or stabilizing their conformations,

❖ targeting soluble proteins to the membranes and to appropriate receptors.,

❖ aiding protein-protein interactions.

1.24. Structural Studies on NMTs

First crystal structure of NMT was published for *C. albicans* NMT (*Ca*NMT) and the enzyme is described as a monomer with a compact globular structure and a large, saddle-shaped β-sheet flanked by several α-helices dominating the core (Weston *et al.,* 1998). The protein appears to have internal two fold symmetry, with two distinct

but structurally similar regions corresponding to the N- and C-terminal halves (**Figure 1.16**). Crystal structure of *S. cerevisiae* NMT (*Sc*NMT) with bound substrate analogues of myristoyl-CoA and peptides confirmed the observations that the NMTs have two domains, one implicated in myristoyl-CoA binding and the second for *N*–terminal protein substrate recognition (**Figure 1.17**) (Bhatnagar *et al*., 1998, 1999). These and subsequent co-crystal structures have given insight into the binding of myristoyl-CoA and peptide substrates and has allowed the residues involved in the active site to be identified (Farazi *et al*., 2001; Sogabe *et al*., 2002; Wu *et al*., 2007). Structures of Myristoyl-CoA and analogues bound to *Ca*NMT and *Sc*NMT, and more recently also *L. major* and *L. donovani* NMTs (*Lm*NMT and *Ld*NMT) indicate that myristoyl-CoA binds in a bent "question mark" conformation (**Figure 1.18**). The thioester carbonyl is positioned to interact with the main chain amides of two enzyme residues that create an oxyanion hole. This oxyanion hole polarises the carbonyl and activates it towards nucleophilic substitution, as well as stabilising the tetrahedral intermediate formed (Bhatnagar *et al*., 1998). One of the bends directs the end of the fatty acyl chain into a deep, narrow pocket in the enzyme where the methyl group interacts with the pocket floor. This provides a way for the enzyme to "measure" the length of the acyl-CoA chain and position the CoA and thioester moieties appropriately, and goes to the extent of explaining the high specificity of NMT for myristoyl-CoA (Bhatnagar *et al*., 1998; Wright *et al*., 2010). All NMTs characterised to date have very high affinity for myristoyl-CoA.

Figure 1.16: Ribbon representation of NMT from *C. albicans* (PDB code: 1NMT, Wetson *et al*., 1998) showing pseudo two-fold symmetry in the structure dividing the structure into two symmetrical halves. α-helices are coloured blue and β-stands as red.
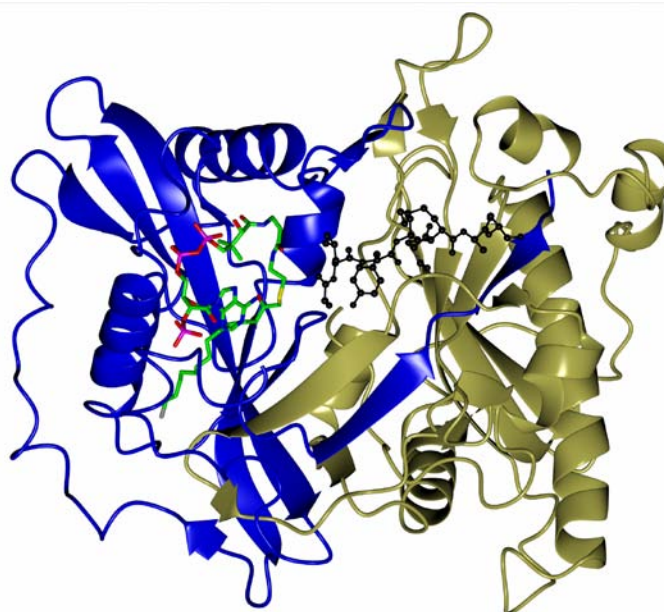
Figure 1.17: Ribbon diagram illustrating the *S. cerevisiae* (PDB code: 1IID, Farazi *et al*., 2001) NMT's pseudo-twofold symmetry. The N-terminal symmetry-related half (Ala34 to Asn225) is coloured blue while the C-terminal symmetry-related half (Trp226 to Leu455) is coloured gold. Non hydrolysable analogue of Myristoyl CoA, known as NHM (S-(2-Oxo)pentadecyl-CoA) is shown as cylinder (carbon as green) while the bound dipeptide inhibitor, SC-58272, is shown in as ball and stick and coloured black.
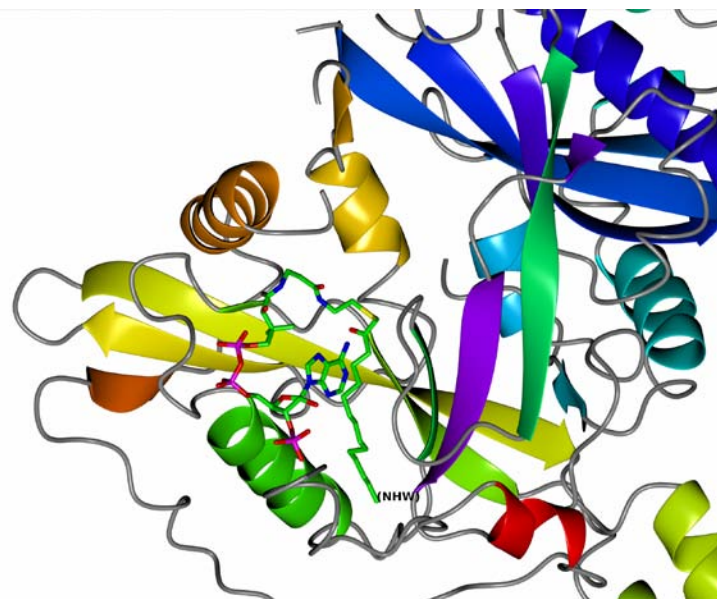
Figure 1.18: Zoom in view of the binding conformation of non hydrolysable analogue of Myristoyl CoA, known as NHW (S-(2-Oxo)pentadecyl-CoA) as in NMT from *L. donovani* (PDB code: 2WUU, Brannigan *et al*., 2010).

1.25. Catalytic Mechanism of N-Myristoylation

Detailed structural studies of NMTs with substrate analogues of myristoyl-CoA and bound peptides have proposed the mechanism of action for *Sc*NMT (Farazi *et al.,* 2001), which may be extrapolated to other NMTs as well. N-Myristoylation by NMT proceeds via an ordered Bi–Bi reaction mechanism; firstly the myristoyl-CoA binds to the protein, binding of myristoyl-CoA resulting in ordering of the "Ab" loop which partially occludes the binding site for the N-terminal peptide of the substrate protein, resulted in opening up of a second pocket for docking of the substrate protein. The apo enzyme has no affinity for the substrate proteins or peptides, implying that the myristyl-CoA to the enzyme is required for the formation of functional N-terminal peptide-binding site. The terminal carboxylate group of Leu421 acts as a base to deprotonate the α-amino group of the terminal glycine residue of the substrate, activating it for nucleophilic attack on the carbonyl carbon of the thioester linkage in myristoyl-CoA. The charge in the transition state of the reaction is dissipated through the carbonyl oxygen, the oxyanion, with stabilizing interactions with the amide groups of Phe170 and Leu171. Myristate group is then transferred through nucleophilic attack

on the carbonyl carbon by the glycine nitrogen. Free CoA has low affinity for NMT apo-enzyme and leaves the complex first. This results in disordering of the "Ab" loop and release of myristoylated peptide (Farazi *et al*., 2001; Towler *et al*., 1987; Rudnick *et al*., 1991; Brannigan *et al*., 2010) (**Figure1.19 and 1.20**).
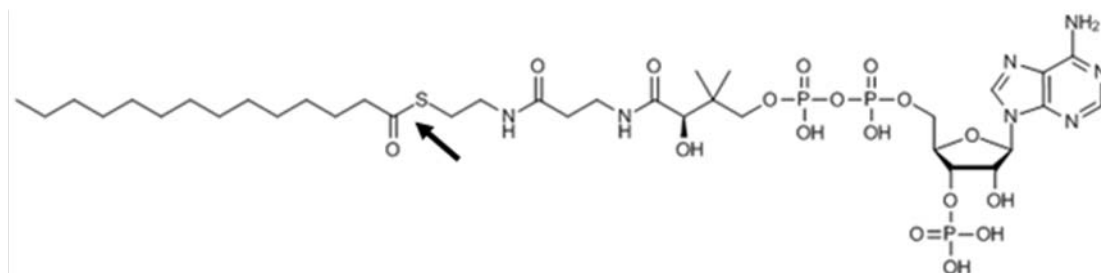
Figure 1.19: Structure of myristoyl-CoA. The arrow indicates the bond that is hydrolysed in the transfer of myristate to the protein subset.
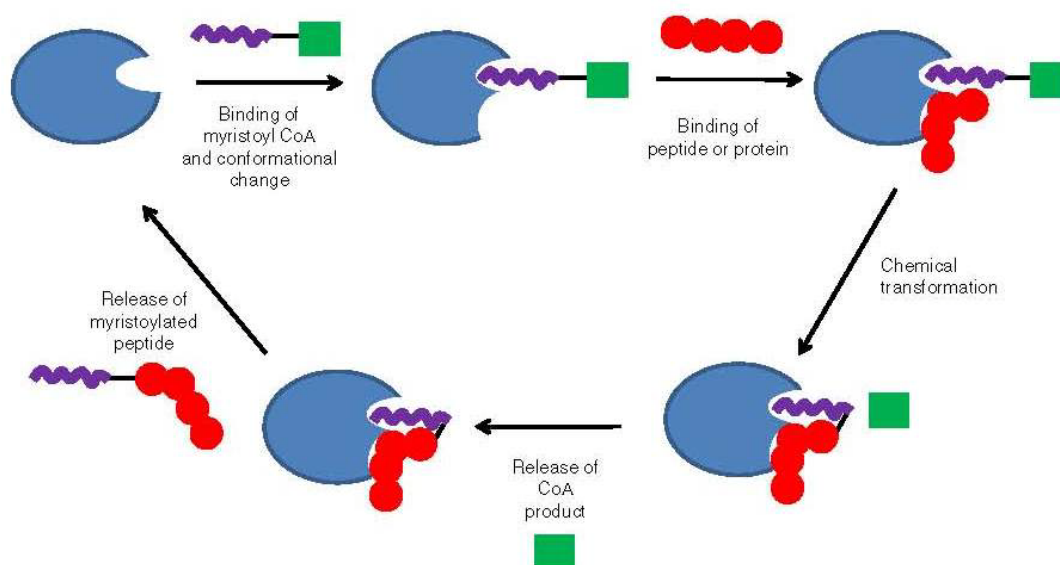
Figure 1.20: Catalytic mechanism followed by NMT enzyme (Adopted from Wright *et al.,* 2009).

1.26. NMT as a Potential Drug Target

Several eukaryotic human pathogens such as pathogenic fungi and parasites require NMT for their survival and some viruses and bacteria use the host NMT enzyme to myristoylate their own proteins. In addition, many myristoylated proteins are involved in signalling networks, and myristoylation may be important in in the

regulation of programmed cell death, apoptosis and reported to be involved in number of diseases. These many cellular functions of NMT have led to numerous investigations of this protein as a target for the development of antifungal, antiviral, anticancer, and antiparasitic compounds.

1.26.1. NMT as Antifungal Target

During the past two decades, the incidence of systemic fungal infections has increased dramatically due to the increasing number of immunocompromised hosts e.g., patients undergoing organ transplants or anticancer chemotherapy, and patients suffering from HIV infection (Fridkin & Jarvis, 1996). Clinically, *candidosis*, *cryptococcosis,* and *aspergillosis* have been identified three major opportunistic pathogens in the etiology of fungal infections (Latge, 1999; Steenbergen & Casadevall, 2000). The mortality rate associated with invasive *Candida* is approximately 40% (Edmond *et al*., 1999), whereas associated with invasive *Aspergillus* reaches 100% in solid organ transplant recipients (Minari *et al*., 2000). In contrast to the large number of antibacterial antibiotics, there are very few antifungal agents that can be used for life-threatening fungal infections. These drugs are amphotericin B (Gallis *et al*., 1990), azoles (such as fluconazole and itraconazole) (Sheehan *et al*., 1999), and echinocandins (such as caspofungin and micafungin) (Denning, 2002). However, current antifungal therapy is far from satisfactory, because currently available antifungal agents have several drawbacks, such as drug related toxicity, severe drug resistance, non-optimal pharmacokinetics, and serious drug-drug interactions (Cohen, 1998; Sanglard, 2002; Casalinuovo *et al*., 2004; Sheng *et al*., 2010; Malik *et al.,* 2012). Therefore, there is an emergent need to develop new antifungal drugs with higher efficiency, broader spectrum, lower toxicity, and improved pharmacodynamic profiles or with novel mode of action.

Genetic experiments have shown that NMT is a promising target enzyme for the development of novel fungicidal drugs having a broad antifungal spectrum (Georgopapadakou, 2002). Blockade of the myristoyl-CoA binding site was demonstrated by using acyl-CoA-based inhibitors obtaining K_i values in the micromolar range (Glover *et al*., 1991). Further potent inhibitors of this site include S-(2-oxo) pentadecyl-CoA and 2-substituted acyl-CoA analogues (Paige *et al*., 1989,

1990). However, comparative sequence and biochemical analyses as well as structural studies have demonstrated high conservation of the myristoyl-CoA binding sites across species, for example, between yeast and plant NMT (Heuckeroth *et al.*, 1988) and between yeast and human NMT (Kishore *et al.*, 1993). In contrast to the acyl-CoA binding site, the peptide binding site is not that highly conserved across species (Johnson, 1994). As a result, among numerous anti-fungal NMT drug discovery programmes, most have focussed on the differences in substrate specificity of *Sc*NMT or *Ca*NMT or *C. neoformans* and *Hs*NMT with selective inhibitors acting at the peptide-binding pocket of the NMTs of fungal species (Langner *et al.*, 1992; Lodge *et al.*, 1997, 1998; Georgopapadakou, 2002). Up to now, peptidomimetic inhibitors (Devadas *et al.*, 1995, 1997, 1998; Nagarajan *et al.*, 1997), myristic acid analogues (Paige *et al.*, 1990; Parang *et al.*, 1996), benzofuran inhibitors (Ebiike *et al.*, 2002; Kawasaki *et al*, 2003; Masubuchi *et al.*, 2001; Yamazaki *et al.*, 2005) and benzothiazole inhibitors (Yamazaki *et al.*, 2005; Ebara *et al.*, 2005) have been reported to possess NMT inhibitory activity. An inhibitor based on the N-terminus of ADP-ribosylation factor (ARF), the major myristoylated protein detected in [^{3}H]myristate-labeled *C. neoformans* (causative agent of chronic meningitis in AIDS patients), is competitive for peptide and non-competitive for myristoyl-CoA binding (Langner *et al.,* 1992), and a nonpeptide inhibitor based on this sequence is at least ~5-fold selective for *C. neoformans* against human NMT (Lodge *et al.*, 1998). Similarly, peptidomimetic inhibitors based on the N-terminal sequence of *C. albicans* ARF was shown to be 560-fold selective for the fungal NMT as compared to the human enzyme. A single dose of 200 µM of the inhibitor produced ~50% reduction in the myristoylation of *C. albicans* ARF, which is consistent with fungistatic activity as shown in Gel mobility assays (Lodge *et al.*, 1997; Price *et al.*, 2003).

1.26.2. NMT as Antiviral Target

Myristoylated viral proteins include both regulatory and structural proteins from a variety of virus families including human immunodeficiency viruses (HIV), the picorna and rotaviruses (Chow *et al.*, 1987; Gottlinger *et al.*, 1989). Interest in this modification of proteins has been stimulated by the finding that prevention of myristoylation, by mutation of the N-terminal glycine residue of the substrate protein,

for both HIV and polio virus, results in defective viral replication (Marc *et al*., 1989; Moscufo *et al*., 1991). The Gag and Nef proteins of HIV-1 virus require myristoylation by the host cell NMT to carry out their functions (Veronese *et al*., 1988). Gag is the polyprotein precursor for the four major structural components of the viral capsid and is targeted to the plasma membrane by myristoylation where it interacts through a myristate plus basic motif (Bryant & Ratner, 1990; Zhou *et al*., 1994). Myristoylation is required for the assembly and budding of newly formed virus particles (Bouamr *et al*., 2003; Gottlinger *et al*., 1989). NMT inhibitors block processing of the gag precursor at the cell membrane, and thus greatly reduce the infectivity of released virus particles (Furuishi *et al*., 1997).

Myristoylation of Nef also plays a critical role in HIV-1 pathogenesis. Nef is an HIV-1 accessory protein and one of the first retroviral proteins to accumulate in an infected cell and plays a role in down-regulating MHC class I molecules and cell-surface CD4 receptors (Garcia & Miller, 1992). Mutation of N-terminal glycine residue of Nef had been shown to affect its incorporation into HIV-1 particles and preventing it from downregulating the expression of MHC I and CD4 proteins (Welker *et al*., 1998; Peng & Robert-Guroff, 2001). Additionally, Nef has been shown to be important in viral infection, as wild-type viruses are six times more infectious than Nef deletion mutants (Miller *et al*., 1994). Harris, (1999) has shown that it enhances reverse transcription in infected cells.

As Gag and Nef are critical for high-titre virus replication and assembly *in vivo*, pharmacological inhibition of their myristoylation may be effective in blocking the progression of AIDS. However, since HIV-1 viruses use human NMT enzyme to myristoylate their own proteins, any anti-HIV inhibitor of myristoylation will target human NMT, raising the possibility of toxicity in uninfected cells. Seaton & Smith, (2008) reported that Nef being a higher-affinity substrate than Gag for both human isozymes, with NMT2 exhibiting a stronger preference both *in vitro* and *in vivo*. However, Takamune *et al*. (2008) suggested that NMT1 myristoylates Gag *in vivo* and that reducing NMT1 isoform activity, but not of NMT2, negatively affects HIV-1 production. A previous study had suggested that Gag binds to NMT2 only whereas Nef prefers NMT1 over NMT2 (Hill & Skowronski, 2005). These contradicting reports from different groups and lack of substantially proven specificity of anti-HIV inhibitor

against human NMT isozymes (NMT1 and NMT2) has left the potential of NMT inhibitors as anti-HIV agents remains as an open question so far.

1.26.3. NMT as Anticancer Target

It is known that several myristoylated proteins are involved in signalling processes e.g., that regulate cell proliferation and growth, stability, cellular location, and biological activity of the proteins. Felsted, Glover and Hartman proposed in 1995 that myristoylation should be considered as a chemotherapeutic target for cancer (Felsted *et al*., 1995). The Src family of tyrosine kinases (e.g., c-Src, Yes, and Fyn) is an example of oncogenic molecules that require N-myristoylation in order for them to function in cells. This was shown in studies of viral oncogene product v-Src that requires myristoylation for membrane binding and cellular transformation (Krueger *et al*., 1982; Cross *et al*., 1984; Kamps *et al*., 1986). Mutation of the N-terminal myristoylated glycine residue of Src-related tyrosine kinases blocks myristoylation and inhibits the ability of these proteins to transform cells without affecting their intrinsic kinase activity (Kamps *et al*., 1985, 1986). Increased Src activity is subsequently involved in the migration, proliferation, adhesion, and angiogenesis of the affected cells (Summy *et al*., 2003). Number of evidence has accumulated regarding the expression and activation of Src in several tumor types, including breast, colon, and ovarian tumors. Sharma and co-workers in particular showed that NMT is over-expressed in human colorectal tumours, adenocarcinomas and brain tumors (Rajala et *al*., 2000a, b; Magnuson *et al*., 1995; Raju *et al*., 1997; Lu *et al*., 2005). Shoji *et al*. (1990) showed that inhibiting myristoylation of Src in colon cancer cell lines prevents the localization of the kinase to the plasma membrane and results in decreased colony formation and cell proliferation. Ducker *et al.* (2005) showed that c-Src is a critical substrate of *Hs*NMT1 and NMT1 is involved in cellular proliferation. The ablation of NMT1 leads to loss of c-Src and reduced proliferation by around 27%. They also showed the simultaneous ablation of NMT1 while NMT2 was particularly lethal to the cells, producing 29-fold increase in the apoptosis. French *et al.* (2004) showed that compounds inhibiting NMT1 and NMT2 are cytotoxic to tumor cells. Over-expression of NMT has been shown relatively early in carcinogenesis, in colonic polyps and tumors of stage B1 in rat model (Magnuson *et al*., 1995). There is also evidence for the

overexpression of NMT in gallbladder and oral squamous cell carcinomas (Rajala *et al.*, 2000a; Shrivastav *et al.*, 2007) and brain tumours (Lu *et al.*, 2005). Because NMTs are overexpressed in carcinomas and tumors, treating with NMT inhibitors will be an important advance in cancer therapeutics.

1.26.4. NMT as Antiparasite Target

Parasitic diseases are a major global health problem, and many of the drugs used today to treat diseases such as malaria, leishmaniasis, and African sleeping sickness, affecting millions of people, were discovered and introduced decades ago and often display limited efficacy and many side effects (Renslo & McKerrow, 2006). Coupled with rising levels of resistance and a need for easily administered drugs with rapid effects, the need to identify and explore the potential new targets is essential. The development of new drugs targeted at parasite pathogens represents an important step towards decreasing the morbidity and mortality caused by these organisms, which affect hundreds of millions of people worldwide.

Several parasitic protozoa also possess NMTs essential for their survival in the host. Comparative biochemical studies of NMT from *P. falciparum* (a principal causative agent of malaria) with human NMT highlighted the potential of the enzyme as target for the development of antiparasitic compounds (Gunaratne *et al.*, 2000). Later on, genomic search and NMT protein sequence comparisons and gene targeting and RNAi (RNA interference) studies have proposed other protozoan parasitic NMTs such as those from *L. major*, *T. brucei*, and *T. cruzi* (causative vectors of cutaneous leishmaniasis, African sleeping sickness, and Chagas' disease, respectively), as good drug targets (Price *et al.*, 2003; Maurer-Stroh *et al*, 2002).

The myristoyl-CoA binding site in NMTs is very well conserved among species. Attention has therefore been directed toward the development of inhibitors that bind in the peptide binding site. Although, conservation of sequence in the peptide binding groove also reported; however, selective inhibitors of yeast and fungal NMTs that binds in this location, exploits the residue differences at a handful of positions and subtle differences in structure (Bhatnagar *et al.*, 1998; Sogabe *et al.*, 2002; Brannigan *et al.*, 2010). In the case of protozoan parasites, the "piggyback" approach, which involves screening of antifungal NMT inhibitors, has been used against recombinant *T.*

brucei NMT, *L. major* NMT, and *P. falciparum* NMT (Gelb *et al*., 2003). Initial screening of synthetic peptide substrates were found active against *T. brucei* NMT and *P. falciparum* NMT with micromolar affinities. Many more compounds were tested later on and the compounds that demonstrated inhibitory activity against *Tb*NMT and *Pf*NMT were also tested against cultured *T. brucei* and *P. falciparum* parasites *in vivo*. A number of these compounds showed potential as leads for future development as antiparasitic drugs for use in humans (Bowyer *et al*., 2007; Panethymitaki *et al*., 2006). In 2010, Frearson and co-workers showed *in situ* killing bloodstream form of *T. brucei* by acting on *Tb*NMT by using a model *Tb*NMT inhibitor DDD85646, probably due to inactivation of *Tb*ARF1 known to operate in protein trafficking and endocytic processes (Frearson *et al*., 2010).

The results coming out of studies from parasitic NMTs using different synthetic compounds shows that although the field is currently at early stages but compounds exist with inhibition characteristics and selectivity against human NMTs that are capable of leading to parasite death in culture and subsequently can be developed as potential drug molecule against the parasitic infections.

1.27. NMT in Leishmaniases

Leishmaniases is a spectrum of infectious diseases caused by a kinetoplastid parasite *Leishmania* species with more than 2 million new cases each year in 88 countries and 367 million people at risk (Chappuis *et al.*, 2007; Shaw, 2007). Cutaneous leishmaniasis, the most common form of leishmaniasis includes most prominently disfiguring skin lesions whereas visceral leishmaniasis (VL) is the most serious form, and is potentially fatal if untreated. In VL the parasites leave the inoculation site and proliferate in spleen, liver and bone marrow, resulting in host immune-suppression and ultimately death in the absence of treatment. These symptoms get aggravated in children and immune-compromised patients, such as those diagnosed as HIV positive. Some 90% of all reported cases of visceral leishmaniasis, also known as Kala azar, occur in just 6 countries: Bangladesh, Brazil, Ethiopia, India, Nepal and Sudan, mainly among the poorest people in rural areas (http://www.who.int/tdr/news/2012/vl_rapid_diagnostics/en/index.html).Though there is limited range of effective drugs in use including Pentostam, amphotericin B (in its

liposomal formulation, AmBisome) and miltefosine the only oral VL drug currently available which is teratogenic (Croft *et al.*, 2006) and a few others in clinical trials, developing resistance is an increasing problem, therefore there is a pressing need for new, preferably orally administered drugs against this important global health problem.

NMT from *Leishmania donovani* (*Ld*NMT), the principal causative agent of the most serious form of leishmaniasis, the visceral leishmaniasis has been shown to be essential for viability in extracellular parasites (Brannigan *et al.*, 2010). NMT from *L. donovani* strain LV9 is different from NMT of the closely related *L. major* Friedlin strain (*Lm*NMT) in only 11 amino acids. With the aim of exploiting *Ld*NMT for drug discovery, NMT from *L. donovani* strain LV9 was overproduced in *E. coli* and the recombinant protein was purified and its ligand binding properties and enzyme activity were checked. Additionally, *Ld*NMT has been co-crystallized with non- hydrolysable myristoyl CoA analogue S-(2-oxo) pentadecyl-CoA (PDB code: 2WUU, Brannigan *et al.*, 2010). The structure determined at a high resolution of 1.4 Å showed the well defined single molecule as asymmetric unit in the electron density maps except for the residues 1–10 at the amino terminus and residues 83–85 and 334–339, which are not defined by the maps and thus assumed to be disordered. The structure contained one molecule of S-(2-oxo) pentadecyl-CoA and well defined 378 water molecules.

*Ld*NMT has 41% sequence identity with NMT of *S. cerevisiae* and the structures were also similar except the electron density for first 26 residues in the yeast structure which were absent and considered to be disordered in the *Ld*NMT structure. The absence of density for some of the residues of "Ab" loop and their high B factors denotes the highly flexible nature of the "Ab" loops that facilitate the peptide binding into peptide binding pocket, following the binding of myristoyl-CoA. Like mentioned, NHM takes up a conformation resembling a question mark (Bhatnagar *et al.*, 1998), with the extended pantetheine and fatty acyl species folding around the adenine base. The conformation and interactions of the fatty acyl-CoA in *Ld*NMT are well conserved as in the structures of the yeast and fungal NMTs.

1.28. Scope of the thesis

Penicillin G acylase is an important enzyme with vital biotechnological and pharmaceutical roles. Thorough understanding of the mechanism in industrial applications and kinetics of PGA is valuable in maximizing the amount of 6-APA produced for the synthesis of semi-synthetic β-lactam antibiotics. With advancement in science and technology, the optimization of PGA molecular properties will continue to be studied. Detailed inspection of the scissile peptide bond and its environment in the mutated precursor molecule is an obvious step towards understanding the chemistry of autocatalytic processing in PGAs. There is a continued interest in probing the substrate specificity of penicillin acylases and engineering its specificity to expand its use not only in the semi-synthesis of new β-lactam antibiotics but also in other synthetic and hydrolytic reactions. More information about this enzyme and its structure will be useful. Knowledge of three-dimensional structure will be important for comparing the binding sites of enzymes to gain better understanding of their specificity towards substrate. Study of intra-molecular interactions responsible for the increase in stability towards pH or temperature will be essential for engineering proteins to achieve higher stability in these conditions. The structural studies on PGA are aimed at understanding and improving their molecular properties for applications.

NMT is one of the few proteins that have been comprehensively validated as a drug target for parasitic infections. The target product profile for a new drug to combat any parasitic infection requires that the drug is safe and efficacious against both the stage 1 disease, when parasites are present in the blood, lymph or interstitial fluids, and the stage 2 disease, when parasites are also present in the central nervous system (CNS). Drug should be selective over human NMTs as well as extreme potent against bloodstream form of parasite. The myristoyl-CoA binding site in NMTs is very well conserved among species, and the interactions of fatty acyl-CoA with the enzyme are conserved in the structures that have been determined. Attention has therefore been directed toward the development of inhibitors that bind in the peptide binding site. To achieve these objectives, a detailed understanding of the interaction between inhibitor bound at peptide binding pocket with the residues of the NMT enzyme is required. X-ray crystallography is a tool that can help to design better and more specific inhibitors which can be developed into effective drugs. The determination of the crystal structure

of the *Ld*NMT–NHM complex establishes a foundation both for the structural analysis of inhibitor complexes and for structure-assisted drug discovery against this protozoan parasite.

With the aim of exploring the application potential of PGA from *Kluyvera citrophila* (*Kc*PGA) and PGA from *Alcaligenes faecalis* (*Af*PGA) in pharmaceutical industry for the production of semi-synthetic penicillins and to exploit myristoyl-CoA–protein N-myristoyltransferase from *Leishmania donovani* (*Ld*NMT) as a potential drug target against leishmaniasis, the research carried out is presented in the thesis as details of the following objectives and procedures towards analysis and interpretation of data and the results:-

1.   Cloning, over-expression and purification of *Kc*PGA.

2.   Biochemical and biophysical characterization including crystallization and three-dimensional structure determination of purified *Kc*PGA.

3.   Three-dimensional structure determination of the thermostable *Af*PGA.

4.   Preparation of a new construct for Myristoyl-CoA–protein N-myristoyltransferase enzyme from *Leishmania donovani* (*Ld*NMT), its over-expression and purification.

5.   Co-crystallization of purified *Ld*NMT enzyme with its substrate analogue or inhibitor.

CHAPTER: 2

MATERIALS AND METHODS

2.1. INTRODUCTION

Much of the information about enzymes-what they are, how they do the what they do has been made possible because they can be isolated from cells and made to work in a test tube environment. This wealth of information regarding enzyme function can be extracted using the combined approaches of biochemical characterization, mutagenesis and structural analyses. This chapter describes the methods employed in the study of an industrially important enzyme penicillin G acylase (PGA) from two organisms named *Kluyvera citrophila* and *Alcaligenes faecalis* and a potential drug target Myristoyl-CoA:protein Nmyristoyltransferase (NMT) from a protozoal parasite *Leishmania donovani* combining the techniques of genetic engineering and X-ray crystallography.

Structure-based study is one component of multidisciplinary design cycles, as illustrated in **Figure 2.1**. In case of proteins the protein engineering cycle is of value. In this cycle, the protein is cloned, expressed and characterized kinetically, and the three-dimensional structure is explored, preferably as a complex with an inhibitor, ligand or a pseudo-substrate. The three-dimensional structure of a protein-ligand complex is used as the basis to elaborate the chemical mechanism of the reaction taking place at the active site of protein.

The entry to the cycle is usually by obtaining protein by identifying and cloning the gene followed by expression and purification of the recombinant protein. After the biochemical characterization of the protein, its three-dimensional structure can be determined using X-ray analysis provided the protein can be crystallised. The structure provides the platform to pursue more incisive and focused experiments, and at the same time allowing the interpretation of experimental observations. Once the structure has been determined, a more hypothetical and uncertain step is the introduction of a mutation to test hypotheses about parameters related to the crucial role of certain residues and the overall fold of protein in its function. The knowledge of the active site architecture certainly can provide a better starting point to identify possible candidates for mutagenic studies. The next step in the cycle is to test the design hypothesis using site-directed mutagenesis followed by expression, biochemical characterization and structure determination of the mutant protein. The ability to modify protein molecules

through site-directed mutagenesis plays a major part in addressing the many questions posed. The three-dimensional structure of a protein-ligand complex can help to unravel the key interactions and mechanism involved in enzyme catalysis and specificity. The methods involved in X-ray analysis and structure determination will be described in more detail later.
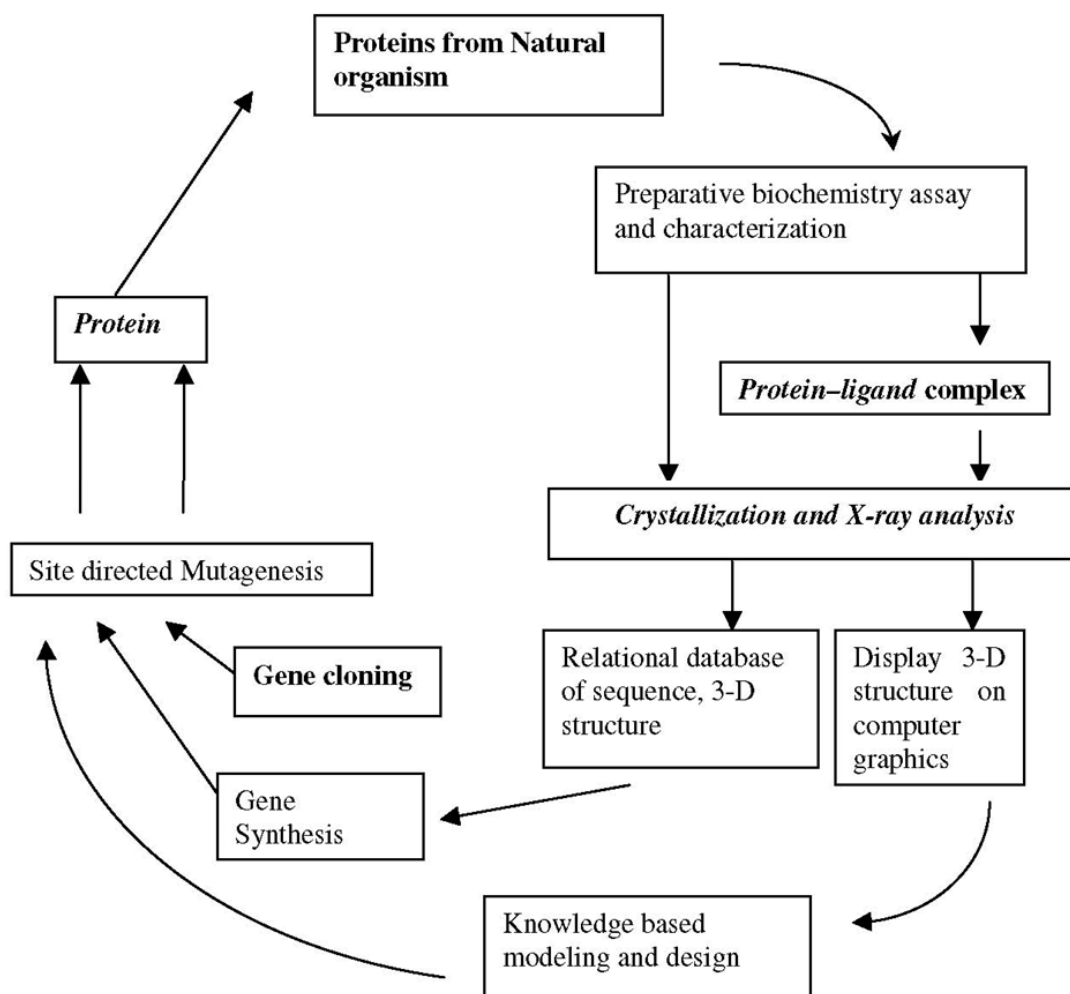
Figure 2.1: Design cycle to illustrate the structure based study of proteins. In most cases, we may have to go round the cycle several times to study the influence of a particular design (adapted from Blundell *et al*., 1989).

2.2. MATERIALS AND INSTRUMENTS USED IN THE STUDY

Media components such as Yeast Extract, Tryptone, Sodium Chloride (NaCl) for growing cells were obtained from HiMedia, India. Reagents used in molecular biology work were purchased from New England Biolabs (NEBs), Promega and Novagen. Oligonucleotides used in polymerase chain reaction (PCR) were obtained either from Sigma, USA or from Eurofins (MWG Operon, Ebersberg, Germany). Competent cells used for transforming deoxyribonucleotide (DNA) were purchased from Novagen directly or prepared from glycerol stock purchased from Novagen. Restriction endonucleases were purchased from either Promega or New England Biolabs (NEBs). Agarose, Kanamycin, Isopropyl thiogalactopyranoside (IPTG) were purchased from Sigma chemical company, St. Louis, USA. Chloramphenicol was purchased from HiMedia.

Potassium salt of Penicillin G was a gift from Hindustan Antibiotics, Pune. p-dimethylaminobenzaldehyde (PDAB) was obtained from Qualigens. PDAB reagent was prepared by dissolving 0.1 gm of PDAB in 17 ml of methanol containing 0.02% Hydroxy-quinone as stabilizer. All media were prepared in distilled water and all buffers were prepared in MilliQ water followed by filtration by 0.22 μm membrane. Chemicals used for the purification of proteins from crude extract of cells such as Trizma, di-Potassium hydrogen phosphate, Potassium dihydrogen Phosphate, Imidazole, Sodium acetate, Nickel chloride, Glycerol, β-mercapto-ethanol, Bromo-phenol-blue (BPB), Acrylamide, N,N'-methylene bisacrylamide, Sodium dodecyl sulfate (SDS), Acetic acid, Methanol, TEMED (N,N,N',N'-Tetramethylethylenediamine), Ammoniumpersulfate (APS) were either purchased from Sigma chemical company, St. Louis, USA or of high purity chemicals available commercially. Molecular weight marker for SDS-PAGE was purchased from BioRad (USA). All solvents of analytical as well as HPLC grade were purchased from Qualigens, India or Merck, Germany. Different types of matrix for protein purifications were purchased from Sigma, USA and GE Healthcare, USA.

Chemicals used in fluorescence studies: Acrylamide, potassium iodide, cesium chloride, 8-anilino-1-naphthalene sulfonic acid (ANS), adenine, guanidium hydrochloride were purchased from Sigma, USA. All other chemicals used were of analytical grade. Quencher solutions (acrylamide, potassium iodide, and cesium

chloride) were prepared as 5 M stocks in deionised water. Potassium iodide solution was made in 200 μM sodium thiosulfate solution to prevent the formation of triiodide (I^{-3}) ions.

Chemicals used extensively in the crystallization trials: Sodium cacodylate buffer, Tris, di-Potassium hydrogen phosphate, Potassium dihydrogen phosphate, Sodium acetate, HEPES-sodium salt, Ammonium sulphate, tri-sodium citrate tetrahydrate, Magnesium chloride hexahydrate ($MgCl_2.6H_2O$), MPD, Polyethylene glycols (PEG) of various sizes (400-20K), Lithium sulfate (Li_2SO_4), additives tried for improving diffraction quality of crystals like Isopropanol, Dioxan, Dimethyl formamide (DMF), Dimethyl Sulfoxide (DMSO), Glycerol, Ethelene Glycol etc were purchased from Sigma chemical company, St. Louis, USA. Commercial sparse matrix screens were purchased from Hampton Research (California, USA), Molecular Dimensions, (UK) and Emerald BioStructures (Washington, USA).

Sterile petridishes (85 mm diameter) were procured from Axygen, India. Disposable filter sterilization units (0.22 μM) were purchased from Millipore (USA). Microfuge tubes (1.5 mL and 2 mL capacity), microtips (10, 200, 1000 μL capacity) and PCR tubes (0.2 mL capacity, flat cap) were obtained from "Tarsons" and "Axygen", India. EIA/RIA medium binding flat bottom, 96 well plates were procured from "Costar" USA. 24 well tissue/cell culture plates used in Hanging drop vapour diffusion experiments for crystallization were obtained from TPP (Switzerland) and Griener bio-one (USA). Microbatch crystallization experiments were performed in MRC 96-well plates (Griener bio-one, USA). The slab gel electrophoresis units used to check the purity of the protein samples were purchased from BioRad and Houfer. The other specialized instruments used in the experiments are mentioned in the appropriate places.

Silicon graphics Octane and Linux based high performance workstations were used for graphics display, for running the crystallographic programs and for other related calculations. Programs from *CCP4* suite (Collaborative Computational Project, Number 4, 1994) and *PHENIX* 1.8.3 (Adams *et al.,* 2010) were used extensively for calculations. The program *QUANTA* (Accelrys) and *COOT* (Emsley & Cowton, 2004) were used for graphics display, visualization of protein structures and model building.

2.3. METHODS

2.3.1. Isolation, Cloning, Expression and Purification of "*pac*" gene from *Kluyvera citrophila* DMSZ 2660 (*KcPAC*)

2.3.1.1. Polymerase Chain Reaction (PCR)

PCR is a powerful technique to amplify a desired nucleotide (gene) sequence using sequence specific primers. This amplification may be either of and from a single template or of a template from a mixture of templates (Arnheim & Erlich, 1992; Mullis, 1990; Mullis & Faloona, 1987; Saiki *et al*., 1985). This technique has been successfully used for various purposes like fishing out of gene(s) from genomic DNA or from cDNA population, introducing restriction sites of interest in the amplified product for directional cloning (Scharf *et al.,* 1986), creating sequence mismatch/ deletion/ addition resulting in mutant version of a gene or nucleotide sequence.

In the present study, applications of PCR were exploited for amplifying a 2562 bp PCR fragment covering the region from 12 nucleotides upstream from the start codon of *K.citrophila* "*pac*" gene and 12 nucleotide downstream of the gene from chromosomal DNA of *K.citrophila* DMSZ 2660 (ATCC 21285) using gene specific primers designed according to the published coding sequence (Barbero *et al*, 1986, Accession No- M15418)

The PCR reaction mixture composition which contains components from KOD polymerase kit (Novagen/Merck) and generalized cycling conditions used were according to the manufacturer's recommendation.

2.3.1.2. Clean up of PCR Product

Amplified PCR product with desired restriction site near the ends were further purified to remove dNTPs and other PCR reaction components using PCR clean up kit (Promega, USA) as per manufacturer's recommendations.

2.3.1.3. Restriction Digestion and Purification of Insert DNA

Purified PCR product was digested with NdeI and XhoI (New England Biolabs) as per restriction enzymes manufacturer's recommendations at 37 °C**.** After restriction digestion the PCR product was loaded on to a 1.5% agarose gel and was recovered

from agarose gel by using Gel elution kit (Promega, USA) as per manual instructions. The product was stored at -20 °C and used for subsequent reactions.

2.3.1.4. Restriction Digestion, Dephosphorylation and Purification of Vector DNA

Vector pET26b (+) (**Figure 2.3**) was digested with same restriction enzymes (NdeI and XhoI) and dephosphorylated with Shrimp alkaline phosphatase (Fermentas Life Sciences, USA) as per manufacturer's recommendations. Digested and dephosphorylated vector was then purified with PCR purification kit (Promega, USA).
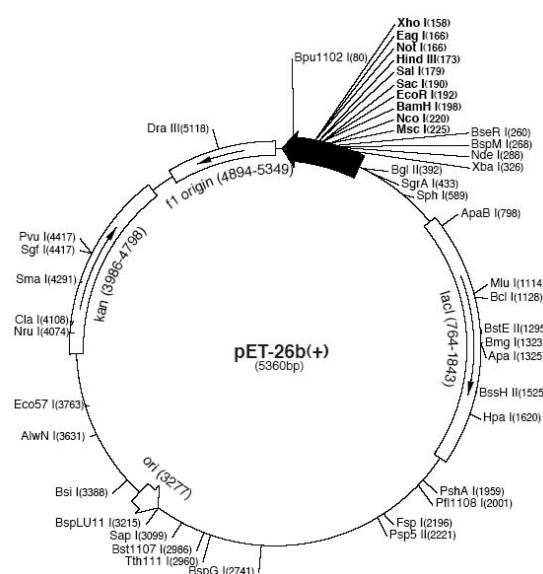
Figure 2.3: Vector map of pET26b (+)

2.3.1.5. Spectrophotometric Determination of Nucleic Acids Concentration

DNA concentration was determined by measurement of the absorption at 260 nm. Nanodrop was used to determine the concentration of DNA samples. Absorbance value (A_{260}) should fall between 0.1-1.0 to be accurate. Sample dilution was adjusted accordingly. An absorbance of 1.0 at 260 nm corresponds to 50 µg DNA/mL

2.3.1.6. Directional Ligation of insert and Vector

Purified vector and purified insert were ligated in 3:1 molar ratio of insert to vector with T4 DNA Ligase (NEB) with suitable buffer at 16 °C overnight.

2.3.1.7. Bacterial Transformation of ligation product

Ligation product was then transformed into NovaBlue competent cells (Novagen/Merck, Germany) and plated on LB + Kanamycin agar plate as per protocol recommended by manufacturers. In brief, competent cells were thawed on ice and 1μL of ligation reaction was added gently to the (50 μL) competent *E. coli* cells, mixed and kept on ice for 5 min. The cells were then incubated at 42 ℃ for exactly 30 sec (heat shock) and immediately kept back on ice for 5 min. 250 μL of SOC medium at room temperature was added and further incubated at 37 ℃ for 1 h at 200 rpm prior to plating 50 μL on LB agar medium containing 50 μg/ml Kanamycin and incubated overnight at 37 ℃.

2.3.1.8. Confirmation of Directional cloning

Cloning was confirmed by isolating plasmid from colonies chosen randomly, obtained on transformation plate following overnight incubation, using MiniPrep DNA purification kit (Promega, USA). Double digestion with NdeI and XhoI enzymes were set up with each plasmid DNA preparation in order to confirm which of the colonies contained the gene. Two gene specific primers were also designed to ensure the reading of full length gene and few of the isolated plasmids were sequenced using standard T7 forward, T7 reverse as well as gene specific primers to confirm their identity with reported sequences available at NCBI GenBank database. Sequence editing and analysis were carried out using Sequence Scanner Version: 1.0 (Applied Biosystems, USA). BLAST analysis was performed using full length gene sequence as well as translated amino acid sequence against NCBI database.

2.3.1.9. Site-Directed Mutagenesis of *KcPAC*

The mutagenesis is one of the techniques carried out to study the effect of a particular mutation on the function of the enzyme or its structure. Mutations can be introduced at specific places using site-directed mutagenesis or at random throughout the gene. During site-directed mutagenesis an amino acid can be changed into any other amino acids. The mutagenesis is generally performed *in vitro* in a cloned plasmid DNA gene using oligonucleotides with altered bases. Introduction of these oligos is normally PCR based and is a common tool in DNA manipulation. An oligonucleotide encoding

the desired mutation(s) is annealed to one strand of the DNA of interest and serves as a primer for initiation of DNA synthesis. In this manner, the mutagenic oligonucleotide is incorporated into the newly synthesized strand.

Site directed mutants of Penicillin G acylase from *Kluyvera citrophila* DMSZ 2660 have been prepared to study the mechanism of autoproteolysis and catalysis. The wild type "*pac*" gene from *K.citrophila* inserted into the plasmid pET26b (+) was used as template for the preparation of the mutants Serβ1Cys and Serβ1Gly. The QuickChange site-directed mutagenesis kit (Stratagene) was used for preparing mutants following manufacturer's recommended protocol.

"pac" gene with desired mutations were ligated with prepared nicked pET 26b (+) vector and transformed in NovaBlue competent cells as done for wild type gene. Transformed cells were plated on LB agar medium containing 50 μg/ml Kanamycin and incubated overnight at 37 ℃. Presence of desired mutations was confirmed by isolating plasmid from colonies chosen randomly followed by full length sequencing. Analyzed sequences were aligned with wild type "*pac*" gene using ClustalW programme (Thompson *et al.,* 1994).

2.3.1.10. Cryopreservation of Bacterial Culture

Bacterial culture harboring plasmid with target gene was preserved for further use in glycerol at ultra low temperature. In a microfuge tube (1.5 mL) culture, grown overnight in LB medium containing suitable antibiotics from sequenced positive colonies, were added with 20% sterile glycerol mixed thoroughly by pipeting and aliquots were frozen in liquid nitrogen and stored at -80 °C. The stocks were revived periodically (~3months) and fresh stocks were prepared.

2.3.1.11. Heterologous Expression in *E. coli*

Expression of the sequence confirmed plasmids for wild type were checked into three expression different hosts. BL21 (DE3), BL21 (DE3) pLysS and pLemo21 (DE3) were transformed with plasmid and the expression was checked after plating on LB agar containing desired antibiotics. A single isolated bacterial colony from overnight incubated plates was used to inoculate 5 mL liquid LB medium containing the desired antibiotic(s). Culture was grown overnight with shaking at 200 rpm at 37 °C. 1% of

overnight grown culture was used to inoculate 10 mL liquid LB containing antibiotic(s). Once the cultures reached OD_{600} 0.5 - 0.6, recombinant protein expression was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG), and the culture was grown for 16-18 h at 16 °C with shaking at 200 rpm. Expression of *Kc*PGA was checked after dissolving the cells in lysis buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% Triton X-100, 10 mM $MgCl_2$, 20 mM Imidazole, Lysozyme 100 μg/mL (added freshly), DNase from Bovine pancreas (added freshly), 1mM DTT) in periplasmic fraction, the cytoplasmic soluble fraction as well as in total cell fraction (soluble and insoluble). Equal amount of each of the three fractions were checked on SDS PAGE to verify heterologous expression of genes. Site directed mutants were transformed in BL21 (DE3) pLysS and expression was checked in both soluble as well as in total cell fractions.

2.3.1.12. Large Scale Production of Recombinant *Kc*PGA

After confirming the presence of soluble protein, the large scale production of *Kc*PGA was carried out to purify the recombinant protein for conducting biochemical and structural studies. LB broth containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol were inoculated with 1% of overnight grown culture containing *KcPAC* gene in BL21 (DE3) pLysS cells and incubated at 37 °C with shaking at 200 rpm. Once OD_{600} reaches 0.6-0.8, IPTG was added to a final concentration of 1.0 mM to induce protein expression and the culture was grown further for 16-18 h at 16 °C with shaking at 200 rpm. After incubation cells were harvested by centrifugation at 6,000 rpm for 10 min at 4 °C. Cells were washed with 0.05 M potassium phosphate buffer; pH 7.5. Cells were suspended in minimum volume of lysis buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% Triton X-100, 10 mM $MgCl_2$, 20 mM Imidazole, Lysozyme 100 μg/mL (added freshly), DNase from Bovine pancreas (added freshly), 1mM DTT) and disrupted by sonication on ice; 10 sec each for 3-5 min at 70% amplitude on an ultrasonic liquid processor, XL 2000 model (MISONIX). Cell debris was removed by centrifugation at 12,000 g for 30 min at 4 °C.

2.3.1.13. Purification of Recombinant Proteins using Immobilized Metal Ion Affinity Chromatography (IMAC)

Affinity chromatography serves as a powerful chromatography technique that separates proteins by use of a reversible interaction of the target protein with binding substance (ligand). The interaction can be biospecific, for example antibodies binding to specific Protein, or non-biospecific, such as hexahistidine-tagged proteins binding to metal ions in the process known as immobilized metal ion affinity chromatography (IMAC). The concept of this type of purification is rather simple. A gel bead is modified covalently so that it displays a chelator group for binding a heavy metal ion like Ni^{2+} or Zn^{2+}. When a protein with metal binding site finds the heavy metal, the protein will bind to the metal ion displayed on the chelator arm of the gel bead. For example a protein consisting of 6 or more histidine amino acid residues in a row can act as a metal binding site to the Ni^{2+} or Zn^{2+}. With high capacity and high selectivity, hence high resolution, for the protein(s) of interest, purification levels in the order of several thousand-fold with high recovery of active material are achievable with this simple procedure. The technique also has a concentrating effect, which enables large sample volumes to be handled. Samples are applied under suitable conditions that favour specific binding to the ligand and then target protein is eluted in a concentrated and much purified form by changing ionic strength, pH, or by introducing a competitive ligand like Imidazole.

Since wild type as well as mutant clones of *Kc*PGA contain histidine tag at C-terminal of the expressed protein, purification protocol was made easy by Affinity chromatography using Ni^{2+}-sepharose beads. This affinity matrix contains bound metal ion nickel, to which the poly histidine-tag binds with micro molar affinity. Clarified crude lysate were loaded onto packed column of Nickel-sepharose beads (Qiagen, Germany) connected to AKTA Explorer (GE Healthcare, USA) which were pre-equilibrated with equilibration buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM Imidazole, 1 mM DTT). The matrix was then washed with plenty of equilibration buffer until OD_{280} reaches to baseline. Non specific and weakly bound proteins were removed by washing with excess of washing buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM imidazole, 1 mM DTT). More than 90% pure *Kc*PGA protein was eluted with elution buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 150 mM imidazole, 1 mM

DTT). Protein elution was monitored by monitoring the absorbance at 280 nm of collected fractions. All the buffers used in this method of purification were free from chelating agent EDTA. The eluted protein was separated and analyzed by SDS polyacrylamide (SDS-PAGE) as described by Laemmli (1970). The gel was stained with Coomassie brilliant blue and protein concentration was determined by Bradford (Bio-Rad, USA) method using BSA as the standard.

2.3.1.14. Size Exclusion Chromatography

Size-exclusion chromatography (SEC) is a chromatographic method in which molecules in solution are separated by their size and in some cases molecular weight as they passed through porous material. Molecules that are smaller than the pore size can enter the particles and therefore have a longer path and longer passing time than larger molecules that cannot enter the particles. It is generally used to separate large molecules or macromolecular complexes such as proteins, and to determine molecular weights and molecular weight distributions of polymers. Typically, when an aqueous solution is used as a mobile phase, the technique is known as **gel-filtration chromatography**, whereas the name **gel permeation chromatography** is used when an organic solvent is used to transport the sample through the column.

The fractions containing *Kc*PGA protein after Nickel purification step were pooled and concentrated with Amicon centrifugal concentrator (Millipore, USA) with cutoff range of 30 kDa and passed through Gel filtration column (Sephacryl S-200) connected to AKTA Explorer and fractions were eluted with 10 mM potassium phosphate buffer pH 7.5 containing 150 mM NaCl.

The aliquots of the fractions were checked for the presence of PGA by enzyme assay. The purity and homogeneity of positive wild type enzyme fractions were checked using 12% (w/v) SDS-PAGE and those of the pooled fractions with 15% SDS-PAGE as well as with 7% native PAGE. The gels were stained using either Coomassie blue or by silver staining method and protein concentration was determined by Bradford (Bio-Rad, USA) method using BSA as the standard. The pooled fractions were concentrated with centrifugal concentrator and the protein was stored at –80 °C. More than 20 mg of pure protein was obtained from 1 L of culture.

Since the site directed mutations selected were to block the enzyme activity completely, the fractions could not be assayed and hence could be checked only by SDS-PAGE run.

2.3.2. Cloning, Expression and Purification of N-terminal Deletion Mutant of NMT gene from *Leishmania donovani* (Δ*N Ld*NMT)

2.3.2.1. Ligation Independent Cloning

Ligation-independent cloning (LIC) is an alternative technique of molecular cloning that is performed without the use of restriction endonucleases or DNA ligase. LIC vectors are created by treating a linearized backbone with T4 DNA polymerase adding only one dNTP. The 3' to 5' exonuclease activity of T4 DNA polymerase removes nucleotides until it encounters a residue corresponding to the single dNTP present in the reaction mix. At this point, the 5' to 3' polymerase activity of the enzyme counteracts the 3' to 5' exonuclease activity to effectively prevent further excision. Plasmid sequences adjacent to the site of linearization in the LIC vector are designed to produce specific noncomplementary 13- and 14-base single stranded overhangs. Target gene inserts with complementary overhangs are created using PCR technique by designing appropriate 5' base pair extensions into the primers. The PCR product is purified to remove dNTPs and original plasmid (if used as template) and then treated with T4 DNA polymerase in the presence of the appropriate dNTP to generate the specific vector-compatible overhangs. Cloning is very efficient because only the desired product is formed by annealing. The annealed LIC vector and insert are then transformed into competent *E. coli* cells. Covalent bond formation at the vector insert junctions occurs within the cell to yield circular plasmid (**Figure 2.4**) (Aslanidis & Dejong, 1990; Haun *et al*., 1992).

A new construct of *Ld*NMT gene, an N-terminal mutant with non-cleavable histidine tag (Δ*N Ld*NMT), was prepared using the technique of Ligation Independent Cloning (LIC cloning). *Ld*NMT gene was fished out from the plasmid DNA containing full length construct of *Ld*NMT gene cloned into modified pET28b (+) containing cleavable N-terminal hexahistidine tag (Brannigan *et al*., 2010). Primers were designed to generate LIC overhangs to original *Ld*NMT DNA to fuse it with pre digested LIC-

YSBL vector resulting in a formation of a new clone with N-terminal 12 amino acids from original clone substituted with non cleavable HisTag.

Primers were designed to generate LIC overhangs to original *Ld*NMT gene using PCR technique under conditions specified by manufacturer. PCR reaction mixture contains components from KOD polymerase kit (Novagen). The PCR product was cleaned up with DNA gel extraction kit (Promega) to remove the excess primers, dNTPs, other protein impurities and most importantly the original gene.
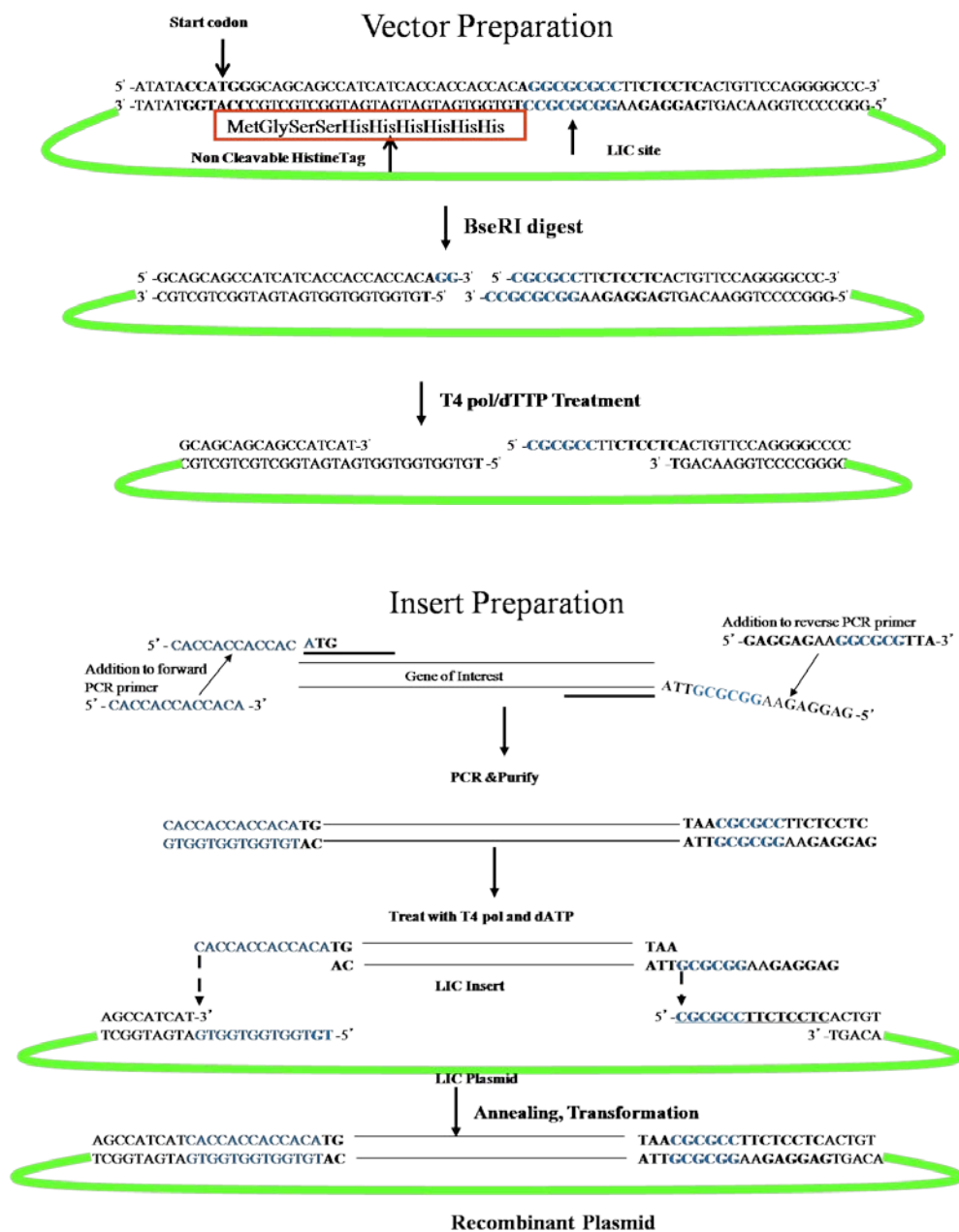
Figure 2.4: Ligation Independent Cloning strategy followed

2.3.2.2. Insert LIC T4 Polymerase Reaction

For the generation of the single stranded tails, the purified PCR product was treated with 1 unit of T4 DNA polymerase (Novagen/Merck, Germany) in a 20 µl volume with only dATP present in the reaction mixture. Contents were mixed and incubated at 22 ℃ for 30 min and reaction was stopped with further incubation of 75 ℃ for 20 min according to manufacturer's recommendation.

2.3.2.3. LIC Annealing

2 µl of T4 polymerase treated insert was added to 1 µl of LIC prepared vector (~30 ng/µl). The reaction mixture was incubated for 10 min at room temp (~20-22℃). Addition of 1 µl EDTA (25 mM) to give a final volume of 4 µl mixed with pipette tip and left at room temp for a further 10 min. LIC control reaction was identical, but substituted 2 µl H_2O for 2 µl of LIC reaction.

2.3.2.4. Transformation of LIC annealed product

LIC annealing reaction was transformed into NovaBlue singles competent cells according to manufacturer's protocol. Transformed cells were then plated out on plate of LB agar containing kanamycin at 50 µg /ml and incubated overnight at 37 ℃.

2.3.2.5. Checking for inserts from the colony

Following the overnight growth a number of colonies were picked up from the plate randomly and grown in LB liquid medium containing 50 µg /ml kanamycin and the plasmid DNA was isolated using Promega DNA purification kit (SV MiniPreps). As N-terminal deletion mutant cloned in LIC vector generated new restriction site of NcoI that was not present in original clone, a restriction digestion with NcoI and NdeI were set up with each plasmid DNA preparation in order to ascertain which colonies had the insert mutant DNA. Reaction mixture was incubated for 2 h at 37 ℃ and checked on 1% Agarose gel to confirm the presence of ΔN *Ld*NMT gene.

2.3.2.6. Expression studies for the cloned protein

The positive plasmids were transformed in pLEMO expression host after confirmation of sequencing results and screened for protein expression. One colony

each from control and sequenced confirmed plasmid, in pLEMO expression host was inoculated in 5 ml LB + kanamycin (50µg/ml) + chloramphenicol (34 µg/ml) liquid medium. Tubes were incubated at 37 °C overnight. 1% inoculums from overnight culture was inoculated in 10ml of LB + kanamycin + chloramphenicol and incubated at 37 °C for 2 h till OD_{600} reads to 0.5-0.7. Cultures were induced with 1 mM IPTG and kept for another 3 h at 37 °C. 1ml of culture were taken out from each tube before induction and kept on bench. Induced culture cells were harvested after 3 h and induced and uninduced cells dissolved in 250 µl of resuspension buffer (50 mM Tris-Cl, pH 8.0 + 100 mM NaCl + 1 mM DTT) for 1 ml culture harvested. Resuspended cells were sonicated for few seconds. Sonicated cells were centrifuged for 12,000 rpm for 10 min and clarified lysate as well as pellet were loaded on 12% SDS PAGE to check the expression.

2.3.2.7. Comparison of expression and Initial purification in LB vs. Auto Induction Medium (AIM)

500 ml of LB or Auto Induction Medium (Studier, 2005) were inoculated with 1% of overnight primary culture containing *Ld*NMT mutant in pLEMO cells. LB medium was induced with 1 mM IPTG after incubating at 37 °C for 2 h at OD ~0.6. AIM was kept at 37 °C for 7 h in total. Both flasks were shifted to 18 °C. LB culture was harvested at 5000 rpm for 15 min at 4 °C after 4 h of incubation while AIM was kept for overnight incubation at 18 °C and harvested.

Harvested cells for both culture broths were dissolved in minimum volume of lysis buffer (20 mM Tris-HCl pH 8.5, 10 mM Imidazole, 10 mM $MgCl_2$, 0.5 M NaCl, 0.1% Triton X-100) and pinch of DNaseI, lysozyme and a Protease Inhibitor cocktail tablet were added just before sonication. Cells were sonicated on ice for 3-5 min at 16 W with 10 sec pulse.

Unclarified sonicated cells were loaded on 1ml HisTrap crude column (GE life Sciences, USA) pre-equilibrated with binding buffer (20 mM Tris-HCl pH 8.5, 10 mM Imidazole and 0.5 M NaCl) independently on bench on two different columns. Flow through was collected for each column. Columns were washed with 10 CV with same buffer and protein was eluted with elution buffer (20 mM Tris-HCl pH 8.5, 500 mM

Imidazole and 0.5 M NaCl) till OD$_{280}$ drops very low. Eluted fractions were loaded on 12% SDS-PAGE gel to compare the expression levels and preliminary purification.

2.3.2.8. Purification of ΔN *Ld*NMT Enzyme

2.3.2.8.1. Affinity Chromatography

500 ml LB broth containing culture cells were dissolved in 50 ml of lysis buffer (20 mM Tris-HCl pH 8.5, 10 mM Imidazole, 10 mM MgCl$_2$, 0.25 M NaCl, 0.1% Triton X-100) and pinch of DNaseI, lysozyme and protease inhibitor cocktail tablet were added just before sonication. Unclarified lysate was loaded on 1 ml His-Trap crude column pre-equilibrated with 20 mM Tris-HCl pH 8.5, 10 mM imidazole, 0.25 M NaCl at flow rate of 0.5 ml/min on bench. Protein was eluted using AKTA explorer with 25 CV gradient of 0-100% buffer B (20 mM Tris-HCl pH 8.5, 500 mM imidazole, 0.25 M NaCl). Presence of the enzyme was checked using 12% SDS-PAGE.

2.3.2.8.2. Anion Exchange Chromatography

Fractions containing *Ld*NMT mutant enzyme purified from Affinity chromatography were pooled and diluted 12-15 times in dilution buffer (50 mM Tris-HCl pH 8.5, 1 mM DTT). Diluted protein was then loaded on bench on 1 ml QFF column pre-equilibrated with binding buffer (50 mM Tris-HCl pH 8.5, 30 mM NaCl, 1 mM DTT) with flow rate of 1 ml/min and eluted with 0-100% gradient of elution buffer (50 mM Tris-HCl pH 8.5, 500 mM NaCl, 1 mM DTT) in 35 CV and checked on 12% SDS-PAGE.

2.3.2.8.3. Size Exclusion Chromatography

Peak fractions showing relative purity were collected and concentrated to 2ml before loading on a HiLoad Superdex 75 16/60pg column (GE Life Sciences, USA) pre-equilibrated with 20 mM Tris-Cl pH 8.5, 250 mM NaCl, 1 mM DTT and protein was eluted with 0.5-0.75 ml/min flow rate.

Fractions containing pure protein were pooled and concentrated to 6 mg/ml in centricon (Millipore, USA) at 277 K. Protein tends to precipitate in the absence of non hydrolysable myristoyl CoA analogue while concentrating to higher concentration therefore 4 times molar concentration of non hydrolysable myristoyl CoA analogue

substrate (NHM) when concentrating protein. 50 µl aliquots of concentrated protein were prepared and stored at -80 °C after freezing in Liquid N_2.

2.3.3. SDS - Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The purity and homogeneity of protein preparations was checked routinely on a 12% sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE), according to procedure of Laemmli (1970). PAGE system is the widely used electrophoresis system for protein separations. The resolution in a Laemmli gel is excellent because the treated peptides are stacked in a stacking gel before entering the separating gel. The gel was prepared using the BioRad or Houfer SDS-PAGE apparatus with 1.5 mm spacers and the samples were electrophoresed alongside molecular weight markers applying voltage 150 V.

2.3.4. Staining using Coomassie blue

Protein bands were visualized by staining the gel with Coomassie blue. After running the gel it was transferred directly to a gel staining box containing 75-100 ml of Coomassie blue staining solution (0.2% coomassie-blue R 250 in 25% propane-2-ol and 10% Acetic acid) and was heated near to boiling in microwave oven. The gel box was kept shaking for 10-15 min at room temperature on dancing shaker. Staining solution was poured off and washed twice with de-ionized water. 100-150 ml of destaining solution (5% Propan-2-ol, 7% Glacial Acetic acid) was poured in. Gel box were then boiled for 3-4 min at maximum power in microwave keeping a stack of adsorbent tissue inside. Gel box was then kept on shaker for 5-10 min. Procedure was repeated 2-3 times with fresh destaining solution and new tissue paper every time till clear bands appeared.

2.3.5. Silver Staining of PAGE gel

Protein separated on SDS, was transferred to the fixer solution (40% Methanol, 10% Acetic acid) for 1 h. This was followed by twice washing in 50% ethanol (20 min each). The gel was transferred to 0.02% sodium thiosulfate solution for 1 min and rinsed thrice with de-ionized water (20 sec each). The gel was silver stained in silver solution (0.2% silver nitrate, 0.076% formaldehyde added to the solution just before use) for 20 min with intermittent shaking in dark. The gel was then rinsed thrice with

de-ionized water (20 sec each) and transferred to the developer solution (6% Na_2CO_3, 0.05% formaldehyde, 0.0004% (w/v) sodium thiosulfate, $Na_2S_2O_3$) till the bands developed. The gel was washed with de-ionized water and stored in fixer.

2.3.6. Biochemical and Biophysical Techniques for Protein Characterization

2.3.6.1. Protein estimation

Protein concentration in samples was estimated by the method of Lowry *et al.* (1951) or using Bradford (BioRad) method using bovine serum albumin (BSA) as the calibration standard. In this method different dilution of BSA solutions are prepared by using BSA stock solution (1 mg/ml), which range from 0.05-1 mg/ml. After addition of the copper sulphate-Folin Ciocalteau solution/Bradford, the reaction mixture was incubated and absorbance was recorded at 660 nm for Lowry and at 595 nm for Bradford method. The concentration of the proteins was determined by comparing with the standard curve for BSA.

2.3.6.2. Penicillin G acylase assay

Bomstein and Evans (1965) devised a specific assay for 6-APA, which depended on the reaction of the 6-amino group with p-dimethylaminobenzaldehyde to form a coloured Schiff base, which is estimated calorimetrically. The chromogen is stable only for a very short time (Bomstein & Evans, 1965).

Determination of acylase activity for PGA from *Kluyvera citrophila* was carried out at pH 7.5 and 50 ºC using 2% (w/v) Penicillin G. 6-aminopenicillanic acid (6-APA) formed was estimated using p-dimethyl aminobenzaldehyde (PDAB) spectroscopically at 415 nm, according to Bomstein and Evans (Bomstein and Evans, 1965), modified by Shewale (Shewale *et al.*, 1987).

For measuring enzyme activity for *Kc*PGA, enzyme sample contained in a final volume of 1 ml of 50 mM phosphate buffer pH 7.5 and 2% (w/v) penicillin G was incubated at 50 °C for 10 min. The reaction was quenched by the addition of 1 ml of 300 mM citric acid in 50 mM Phosphate buffer (CPB) of 2.5 pH. The 6-APA formed in the reaction mixture was estimated spectroscopically at wavelength 415 nm by the addition of 2 ml of a 0.6% (w/v) solution of PDAB in methanol. Appropriate dilution of the quenched reaction mixture in CPB was carried out to determine the absorbance in

the linear range of 0.6-0.7. A substrate concentration of 20 mg/ml (2%) was maintained to measure the enzyme activity in all the experiments except where stated otherwise. Standard graphs were drawn using pure 6-APA (Sigma).

The specific activity of the purified *Kc*PGA was determined using the standard curve plotted for 6-APA.One unit of enzyme activity is defined as the amount of enzyme required to produce 1 µmol of 6-APA per min under assay conditions.

2.3.6.3. Determination of Molecular weight of PGA from *Kluyvera citrophila*

2.3.6.3.1. Relative Molecular Mass (Mr) Determination by Matrix–assisted Laser Desorption Ionization/Time of flight-Time of flight Mass Spectrometry (MALDI-TOF TOF)

The mass spectrum was recorded by using AB SCIEX TOF/TOF^{TM} 5800 system (AB SCIEX, USA) with 1000 Hz high-repetition laser. 5 µl of 100 µg/ml of the purified native *Kc*PGA was mixed with equal volume of matrix solution in microcentrifuge tube. The matrix solution of 10 mg/ml sinnapic acid was prepared in 30% acetonitrile (ACN). About 5 µL of the mixture was applied to a stainless steel sample holder and introduced into the mass spectrometer after drying.

2.3.6.3.2. Molecular weight determination by Electrospray Ionization (ESI)

Molecular weight of the purified PGA was determined by Electrospray method on API QSTAR Pulsar I mass Spectrometer (Applied Biosystems, USA). Purified PGA was buffer exchanged with 2 mM Tris–HCl to remove any traces of salt and subjected to the mass spectrometer analysis.

2.3.6.3.3. Determination of molecular weight of subunits using SDS PAGE

Subunit molecular weight of the enzyme was determined by running SDS-PAGE slab gel electrophoresis (Laemmli, 1970) using BioRad Tetracell SDS-PAGE apparatus with 1.5 mm spacers and samples electrophoresed alongside lower range molecular weight marker proteins (Bangalore Genei/Merck, Germany) comprising phophorylase b (97,000 Da), bovine serum albumin (66,000 Da), ovalalbumin (43,000 Da), carbonic anhydrase (29,000 Da), soybean trypsin inhibitor (20,000 Da) and lysozyme (14,400 Da). The enzyme sample (20 µL) containing 10 µg of purified

*Kc*PGA was mixed with 5 μL SDS-PAGE denaturing dye. Enzyme sample was heated at 98-100 °C for 3 min and loaded in adjacent wells in a 15% SDS-PAGE gel. Protein bands were visualized by staining either with Coomassie blue or using silver staining method as already described.

2.3.6.4. Determination of Optimum Conditions for Penicillin acylase Activity

For determination of the optimum temperature of *Kc*PGA, activity was measured at different temperatures in the range of 20-90 °C at pH 7.5 in 0.05 M phosphate buffer for 10 min. For estimating thermostability, the enzyme was incubated at 25-70 °C up to 120 min in 0.05 M phosphate buffer pH 7.5 and residual activity was estimated at 30 min, at 60 min and at 120 min.

To ascertain the optimum pH of enzyme, activity was determined at 50 °C in the pH range 1.0-12.0 using 0.1 M buffers. The following buffers were used for these studies: Glycine-HCl for pH 1-3, acetate for pH 4-5, phosphate for pH 6-7.5, Tris- HCl for pH 8-9 and Glycine-NaOH for pH 10-12. Appropriate controls for pH studies were used that contained substrate at each respective pH without the presence of enzyme in the reaction mixture. For pH stability, the enzyme was incubated in 0.1 M buffers of pH range 1.0-12.0; using above mentioned buffers for 8 h at 25 °C. Residual activity was measured at different intervals by standard enzyme assay using 2% PenG and 2 min colour development with PDAB as described previously.

2.3.6.5. Substrate Specificity

The hydrolase activity was checked with different β-lactam antibiotics as substrates, purchased from Sigma, USA. penicillin V, penicillin G, ampicillin and amoxicillin were tested. A typical incubation mixture contained substrate at a concentration of 10 mM as free acid or sodium or potassium salt in 0.05 M potassium phosphate buffer pH 7.5. Reaction was initiated with the addition of enzyme to different substrates. After incubating at 50 °C for 10 min, the solutions containing the potential alternate substrates were quenched and the presence of 6-APA was determined using PDAB method. Each substrate was assayed in duplicates, along with appropriate controls from which the addition of enzyme was withheld.

2.3.6.6. Enzyme Kinetics

The purified *Kc*PGA preparations were characterized kinetically. Kinetic studies were carried out at pH 7.5 and 50 °C. The enzyme was incubated with various concentrations of the substrate (up to 50 mM) and the V*max*, K*m* values were calculated by fitting a linear regression curve to data points using Lineweaver-Burk plot. Catalytic turn over number *(*K$_{cat}$*)* value was calculated from the Equation (**2.1**)

$$V_{max} = K_{cat} \, [E]t. \qquad\qquad (2.1)$$

The K*cat* was expressed as V_{max}/μmole of the enzyme/min.

2.3.6.7. Effect of Modulators on Enzyme Activity

The enzyme (32 μg) was incubated with either metal ions (2 mM), 10 mM EDTA or detergents (1%) in 0.05 M potassium phosphate pH 7.5 (volume 0.5 ml), for 15 min at 28 °C. 10 μL of sample was withdrawn and relative catalytic activity of penicillin G acylase was measured using PDAB method as described above. Enzyme reaction without addition of modulator was considered as control. Chelating effect of 10 mM EDTA was checked after 24 h of incubation with enzyme as well.

2.3.6.8. Dynamic Light Scattering (DLS) Studies

Proteins consist of polypeptide chains that are sensitive to a wide range of parameters such as temperature and chemical environment. Preparation method, concentration of protein and/or buffer choice, storage conditions can all influence the size and quality of proteins in a sample.

Protein stability is a particularly relevant issue in the pharmaceutical field and continues to gain more importance as the number of therapeutic protein products in development increases. However, if a therapeutic protein cannot be stabilized adequately, its benefits to human health will not be realized. One major aspect of protein instability is self-association leading to aggregation. Aggregates can reduce the efficacy of protein drugs and can lead to immunological reactions and toxicity.

The presence of the aggregates and the size of the aggregates can be determined by Dynamic light scattering (DLS). DLS is a relatively fast method of characterizing the size of biomolecules in solution, taking only minutes for a measurement. DLS may

be used to distinguish between a homogenous monodisperse and an aggregated sample. For example, oligomer formation of proteins was detected by DLS as an increase of their hydrodynamic diameters. In addition, the polydispersity obtained from DLS analysis is often used as a parameter for prediction of formation of protein crystals. Empirical reports suggest that a protein preparation which is monodisperse (as judged by a polydispersity of less than 15% of the total particle size) is likely to crystallize while a sample that contains non-specific aggregates is unlikely to do so (Zulauf & D'Arcy, 1992; Ferré-D'Amaré & Burley, 1994).

A dynamic light scattering experiment was performed on a DynaPro instrument using protein solutions (Lakewood, USA). The measurements were made at 20 ℃ on the purified protein at 5 mg/ml in buffer solution of 20 mM Tris-HCL pH 8.5 containing 100 mM NaCl. Protein samples were either centrifuged at 10,000 *g* or sterile filtered to remove dust particles before analysis. The samples were then illuminated at 825 nm using a solidstate laser. Each sample was evaluated at least 20 times, with the quoted values being the mean of these independent evaluations. DLS experiments provide a direct determination of the diffusion coefficient (*D*) of the particles in solution. Using the diffusion coefficient, the hydrodynamic radius (*R*H) of the particles can be calculated from the Stokes-Einstein equation **(2.2)**

$$D = kT / 6hRH \qquad \textbf{(2.2)}$$

where *k* is the Boltzmann constant, *T* is the absolute temperature, and h is the solvent viscosity. Approximate molecular weights for each of the species were determined from RH calculated using the software package DYNAMICS supplied with the instrument.

2.3.6.9. Steady State Fluorescence Measurements

The aromatic amino acid residues in the protein, namely tryptophan, phenylalanine and tyrosine get excited on absorption of light energy (photons). Since the residue is unstable in their excited state, it returns to its ground state accompanied by the dissipation of excess energy in the form of fluorescence. Tryptophan has much stronger fluorescence and higher quantum yield as compared to tyrosine and phenylalanine. The intensity and the wavelength of maximum fluorescence emission of tryptophan are highly dependent on the polarity of the environment surrounding the

tryptophane residue. Hence by studying the tryptophan fluorescence, the conformational changes in the protein can be monitored.

Steady-state fluorescence measurements were performed at 28 °C on Perkin Elmer LS 55B luminescence spectrometer connected to a Julabo F20 water bath. The protein solution of 17-18 μg/ml (~0.2 μM) was excited at 295 nm and emission was recorded in the range of 310–400 nm. Slit widths of 7 nm each were set for excitation and emission monochromators and the spectra were recorded at 100 nm/min. To eliminate the background emission, the signal produced by buffer solution was subtracted.

2.3.6.9.1. Tryptophan Accessibility: Solute Quenching Studies

Interaction between a fluorophore and a molecule induces perturbation or modification in the fluorescence parameters like intensity, quantum yield and lifetime. Two types of fluorescence quenching can be observed when the interaction takes place between the fluorophore and the quencher molecule. Collisional quenching occurs when the fluorophore and another molecule diffuse in the solution and collide with each other. In this case, the two molecules do not form a complex. In static quenching, on the other hand, two molecules bind one to the other forming a complex (Albani, 2004).

The most commonly used quenching molecules are acrylamide (neutral), iodide (cationic) and cesium (anionic). Acrylamide, being a small uncharged molecule, can diffuse within a protein and can quench the fluorescence of buried tryptophans as well. Iodide and cesium ions, on the other hand, quench the fluorescence of tryptophans present at or near the surface of the protein. Being charged ions, their quenching efficiency also depends on the charge surrounding the tryptophan (Albani, 2004).

Fluorescence quenching experiments on *Kc*PGA were carried out at 28 °C. Sample containing 17-18 μg protein in 1 ml of 10 mM potassium phosphate buffer pH 7.5 containing 150 mM NaCl was titrated with small aliquots (5-10 μL) of 5 M quencher solution (acrylamide, potassium iodide or cesium chloride). Iodide stock solution contained 0.2 M sodium thiosulfate to prevent formation of triiodide (I^{-3}). The samples were excited at 295 nm and the emission spectra were recorded in the range of 310-400 nm as described above.

For quenching studies with denatured protein, the protein was incubated with 6 M Gdn-HCl prepared in same buffer for 16 h at room temperature. Fluorescence intensities were corrected for volume changes before further analysis of quenching data.

The steady-state fluorescence quenching data obtained with different quenchers were analyzed by Stern–Volmer (Equation **2.3**) and modified Stern–Volmer (Equation **2.4**) equations in order to obtain quantitative quenching parameters (Lehrer, 1971):

$$Fo/Fc = 1 + Ksv\,[Q] \qquad\qquad\qquad (2.3)$$

$$Fo/\Delta F = fa^{-1} + 1/\,[Kafa\,(Q)] \qquad\qquad (2.4)$$

where, Fo and Fc are the relative fluorescence intensities in the absence and presence of the quencher, respectively, (Q) is the quencher concentration, Ksv is Stern–Volmer quenching constant, $\Delta F = Fo - Fc$ is the change in fluorescence intensity at any point in the quenching titration, Ka is the quenching constant and fa is the fraction of the total fluorophore accessible to the quencher. Equation **2.4** shows that the slope of a plot of Fo/ΔF versus Q^{-1} (modified Stern–Volmer plot) gives the value of $(Kafa)^{-1}$ and its Y intercept gives the value of fa^{-1}.

2.3.6.9.2. Thermal denaturation

Effect of temperature on *Kc*PGA was studied using a thermostatic cuvette holder connected to an external constant temperature circulation water bath. The protein sample 23 μg (0.26 μM) was incubated for 10 min at specified temperature before taking the scan. For renaturation experiments the samples were cooled to 25 °C and left for 1 h before recording the spectra. Fluorescence spectra were recorded as described above.

2.3.6.9.3. Light scattering studies

Rayleigh light scattering experiments were carried out with the spectrofluorimeter to follow protein aggregation under different conditions. Both excitation and emission wavelength were set at 400 nm and the time dependent changes in scattering intensity were followed. The excitation and emission slit width were set at 12.5 nm and 2.5 nm, respectively. Scattering was observed for 30 sec.

2.3.6.9.4. Effect of pH

Protein samples (20µg) were incubated in an appropriate buffer for 4 h at 30 °C over a pH range of 1-12. The following buffers were used for these studies: 25 mM Glycine-HCl buffer for pH 1-3 and glycine-NaOH for pH 10-12. For other pH ranges, 25 mM acetate buffer (pH 4-5), 25 mM phosphate buffer (pH 6-7) and 25 mM Tris-HCl buffer (pH 8-9) were used. Fluorescence spectra were recorded as described above.

2.3.6.9.5. ANS binding studies

8-anilino-1-naphthalene sulfonic acid (ANS) is a fluorescent dye which can bind to the exposed hydrophobic patches on the protein surface. ANS alone shows a weak fluorescence at 515-520 nm, when excited at 375 nm. When this dye binds to the exposed hydrophobic patches of the protein, its fluorescence intensity increases considerably with a blue shift in the fluorescence maximum to 480-495 nm. This property is used to estimate the exposed hydrophobic patches on the protein under various conditions (Daniel & Weber, 1966; Cardamone & Puri, 1992). Since the hydrophobic patches are usually buried in the interior of the protein and do not get exposed until the protein is unfolded by thermal/pH denaturation, ANS binding studies provide valuable information about the unfolding behaviour of the protein, as well as about the intermediates formed during unfolding and refolding of the protein.

The intermediate states of *Kc*PGA incubated under denaturation conditions of pH and temperature were analyzed by the hydrophobic dye (ANS) binding. The final ANS concentration used was 40 µM, excitation wavelength was 375 nm and total fluorescence emission was monitored between 400 and 550 nm. Reference spectrum of ANS at each buffer of respective pH or at each temperature was subtracted from the spectrum of the sample.

2.3.7. Protein Preparation for Crystallographic studies

Purified and concentrated protein was spun at 12000 rpm at 277 K for 10 min to remove any particulate matter before setting up the crystallization trails. Clear supernatant was used to set up the crystallization plates.

2.3.8. Crystallographic Methods

In the 1930s, when biochemists began to realize that all chemical reactions in living cells are catalyzed by enzymes and the secret of life was widely believed to lie hidden in the structure of proteins, researchers has began the quest for finding the relationship between the protein structures and the function it performs. The most accurate experimental means of obtaining a detailed picture of a large molecule like protein molecule, allowing the resolution of individual atoms and hence its conformation and interactions, is by using X-ray crystallography.

Other methods like NMR, cryo-electron and atomic force microscopy are also useful tools to solve protein structures. Nuclear magnetic resonance (NMR) spectroscopy is unique among the methods available for three-dimensional structure determination of proteins and nucleic acids at atomic resolution, since the NMR data can be recorded on proteins in solution. Unlike X-ray crystallography, NMR studies do not require crystals, at the same time the dynamics of the structure can be analyzed. However, NMR can be used to determine structures of proteins having only molecular weights up to ~30 kDa. The Cryo-electron microscopy has been utilized in interpreting and visualizing unstained biological complexes of very large size (200 kDa or larger) such as ribosomes or viruses. Although these methods are useful in determining three-dimensional structure of the protein, these methods are still limited by the molecular size and restricted structural information that can be obtained. With more than 75,000 entries out of the total of more than 85,000 in the Protein Data Bank (PDB), are determined by diffraction methods, we could say that the method dominates the field of protein structural biology.

In general, determining the structure of a protein by X-ray crystallography entails growing high-quality crystals of the purified protein, measuring the directions and intensities of X-ray beams diffracted from the crystals, and using a computer to simulate the effects of an objective lens and thus produce an image of the crystal's contents followed by the structural analysis to give important details about the molecule. The methods employed starting from crystal to final model are graphically summarized in **Figure 2.5** and will be briefly discussed, emphasizing those methods specially used in research reported in this thesis, including crystallization, cryo-crystallography, data collection, molecular replacement and refinement.

Initial Screening

Crystal Improvement

Choosing cryoprotectant and method

X-ray
Diffraction Data

Data Reduction

Merged Data

| Experimental phasing | Molecular replacement |

Initial phases

Model in unit cell

Density improvement

Interpretable Map

Model Building

Refinement

Final model

Structural analysis

Analysis of Geometry and fit with data

Validation & Deposition

Figure 2.5: Graphical summary of the methods used in protein X-ray crystallography (Adapted and modified from CCP4 tutorial).

2.3.8.1. Crystallization and Crystal growth

The first step considered crucial in protein structure determination is the availability of diffraction quality crystals. It has been shown that two thirds of all proteins submitted to crystallization trials will fail to generate protein crystals (Dong *et al.,* 2007). In addition, of the proteins that do form crystals, only half may be optimized to form crystals of sufficient quality that will allow structure determination. Thus, in the course of going from a pure protein sample to its determined molecular structure, a success rate of merely 15% is estimated. Therefore, protein crystallization constitutes to be a major obstacle that deserves the utmost attention. A crystal may be defined as an orderly three-dimensional array of molecules held together by non-covalent interactions. The goal of crystallization is to produce a well-ordered crystal that is lacking in contaminants and large enough to provide a diffraction pattern when exposed to X-rays. This diffraction pattern can then be analyzed to discern the protein's three-dimensional structure. Protein crystallization, as people often say, is more like an art than science and is a significant bottleneck in protein crystallography. In the absence of any single concrete theory behind the mechanism of crystallization, protein crystallization is mainly trial and error procedure invoking experience and crystallization reports as guiding principles. Several factors are known to play important role in crystallization process and thus affect the quality of crystals obtained, such as temperature, pH, precipitating agents, ionic strength of the solution, presence of impurities and nucleating agents, additives and also several other unknown factors.

Protein crystallization is inherently difficult because of the fragile nature of protein crystals. Proteins have irregular surfaces, which results in forming large channels within any protein crystal. Therefore, the non-covalent bonds that hold together the lattice must often be formed through several layers of solvent molecules (Rhodes, 2000). Protein crystals are composed of approximately 50% solvent, though this may vary from 30−78% (Matthews, 1985). They are labile, fragile, and sensitive to external environments owing to their high solvent content, and the weak binding energies between protein molecules in the crystal (Littlechild, 1991). The only optimal conditions suitable for their growth are those that cause little or no perturbation in their molecular properties.

The crystallization of proteins from solution is a reversible equilibrium phenomenon. It can be considered as consisting of three stages: nucleation, growth and cessation of growth. **Nucleation** is necessarily the first step, influence the crystallization process decisively. Consequently, the ability to control it is of primary importance in crystallization experiments. In principle, crystals will form in any protein solution that is supersaturated i.e., when the protein concentration exceeds the solubility of the protein. The large supersaturation is required to overcome the activation energy barrier which exists when forming the crystal. This barrier represents the free energy required to create the small microscopic cluster of proteins known as a nucleus, from which the crystal will eventually grow. The formation of crystals is due to decreasing free energy of the system by the formation of many new interactions which outweighs the decreasing entropy of the system, allowing for a highly organized structure. Thus, the reduction in free energy by ~3-6 kcal/mole relative to the solution state is the thermodynamic driving force that causes ordering in crystals (McPherson, 1982, Drenth & Haas, 1998).

Second stage is **Crystal growth,** in which further molecules or ions are added to the nuclei in a regular manner resulting in a good crystal. The final stage is the termination or **cessation of growth** of the crystal (Weber, 1991). Crystals will grow only from a non-equilibrium supersaturated solution. At the equilibrium point, the number of protein molecules entering the solution is same as the number leaving the solution, which is referred to as the solubility limit of that protein. When the amount of protein in the solution is below this limit, the solution is undersaturated. If the amount of protein is equal to the solubility limit, the solution is in saturated state. Crystals can grow only in supersaturation state. Every protein has a unique solubility limit (**Figure 2.6**). Decreasing the solubility of the protein is the most effective way of creating supersaturation. Supersaturation can be achieved by different approaches including altering the precipitant concentration, buffer pH, temperature, protein concentration and/or dielectric constant of the medium.

Since there is an energy barrier, nucleation (the process of forming a nucleus) takes time. If the supersaturation is too less, the nucleation rate will be so slow that no crystals form after any amount of time. The corresponding area of the phase diagram is known as the "metastable zone." In the "labile" or "crystallization" zone, the

supersaturation is large enough that spontaneous nucleation is observable. If the supersaturation is too large, then disordered structures, such as aggregates or precipitates, may form. The "precipitation zone" is unfavourable for crystal formation, because the aggregates and precipitates form faster than the crystals (Asherie, 2004).
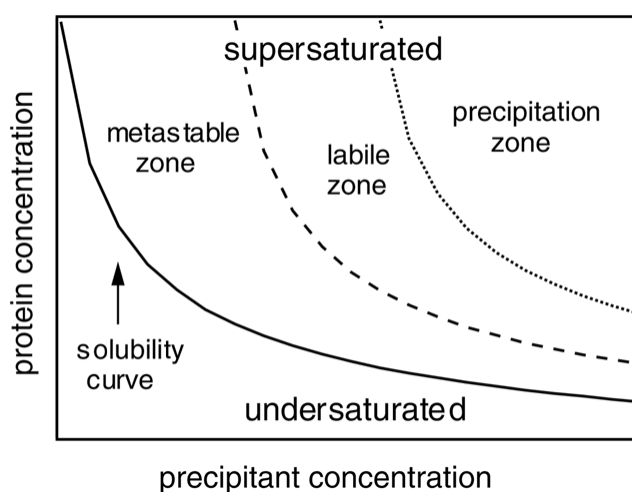
Figure 2.6: A schematic phase diagram showing the solubility of a protein in solution as a function of the concentration of the precipitant present (adapted from A. McPherson, 1999)

There are several methods available for protein crystallization that all involve bringing the protein to the point of crystallization in a slow, controlled manner. Purified protein is dissolved in an aqueous buffer containing a precipitant, such as salt or organic solvent, at a concentration just below that necessary to precipitate the protein. Then water is removed by controlled evaporation to produce precipitating conditions, which are maintained until crystal growth ceases. Several methods commonly used for crystallizing a protein are batch crystallization, liquid-liquid diffusion, vapour diffusion (hanging-drop and sitting-drop), dialysis etc., among which the most popular experimental method used for crystallization is the vapor-diffusion method, both hanging-drop and sitting-drop method. The hanging drop vapour diffusion method is widely adopted and has produced more protein crystals than all other methods combined (Chayen, 1998).

In this method, the protein and precipitant solutions are allowed to equilibrate in a closed container with a larger aqueous reservoir whose precipitant concentration is

optimal for producing crystals. This method is simple and it is easy to monitor the progress of crystallization. The advantage is that it requires only a small volume of droplet, which can be as low as 1 μL per experiment, so with a small amount of sample screening for many crystallization parameters and the optimization of the crystallization hits are possible (McPherson, 1999). The reason for the popularity of the hanging-drop method is the ease of performing the experiment, only a 24 well-plate (such as Linbro plate), grease and cover slips are required (**Figure 2.7**).

Hanging-drop vapour-diffusion method uses a 24-well plate that had a fine bead of vacuum grease applied around the edge of the well. Precipitating solutions (500 μL to 1ml) composed of precipitant, buffer solution, distilled water with or without additive are pipetted into the 24 well reservoirs of the crystallization plate. Then 1-5 μL of the concentrated protein sample is pipetted onto a siliconized coverslip, followed by equal amount of the well solution. The coverslip is then inverted over the well. This is then left undisturbed for at least 24 h to equilibrate. At the beginning of the experiment, the precipitant concentration in the drop is half of that present in the well. Equilibration then takes place via the vapour diffusion. Given the relatively large volume, concentration of the well effectively remains the same. The drop however loses water vapour to the well until the precipitant concentration equals that of the well. If the conditions are favourable, at some point during this process the protein will become supersaturated and will drive out of solution in the form of crystals.
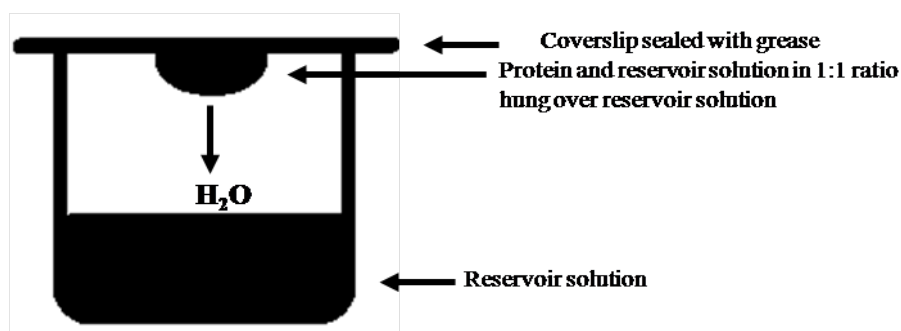
Figure 2.7: Hanging-drop vapor-diffusion method for protein crystallization

The sitting drop method involves a very similar procedure except that the protein containing drop is kept seated over the reservoir buffer instead of being suspended from the coverslip. The sitting-drop vapor-diffusion method is performed using 96-well Crystal Clear Strip plate. Usage of robotics has reduced the amount of

screen solution required where only 60-100 µL reservoir solution is used instead of 500 µL or 1ml as in case of hanging drop method. Protein as low as 150 nl and an equal amount of reservoir solution are mixed together at the top of the ledge, and finally the plates are sealed with sealing tape, placed on crystallization rack with temperature maintained at 20 ℃ or room temperature. Thus sitting drop vapour diffusion method can be used to screen thousands of crystallization conditions initially using very less amount of protein. Subsequently, crystallization hits can be optimized using bigger drop volume in hanging drop method to obtain better and larger crystals.

Seeding is a useful technique for growing crystals by producing a seed-stock solution in cases where spontaneous homogenous nucleation does not occur or where crystals are small and of poor diffraction quality. It is generally believed that the nucleation will be initiated at a higher level of supersaturation, the labile zone (Stura & Wilson, 1991). To use seeding techniques, good quality crystal seeds have to be selected and introduced into the metastable labile zone. After seeding, the seeded crystals will continue to grow. Seeding techniques can be classified into two categories based on the size of the seeds:

• Microseeding—transfer of submicroscopic seeds, too small to be distinguished individually.

• Macroseeding—transfer of a single crystal, usually 5–50 µm.

Seeding can be used in conjunction with other optimization measures such as fine-tuning the precipitant and protein concentrations, adjusting the pH, and screening for additives.

In this work, screening of crystallization condition for *Af*PGA was performed using hanging drop vapour diffusion method in 24 well plates using various commercial screens, whereas *Kc*PGA and *Ld*NMT mutant were screened for crystallization hits using 96 well microbatch experiments. Further optimization of the hits obtained were performed using 24 well hanging drop vapour diffusion method by changing various parameters like pH, temperature, precipitant concentration, protein concentration, drop size or by using various additives like salts, volatile solvents or detergents.

2.3.8.2. X-ray diffraction

After obtaining crystals of suitable size and quality the second step in crystal structure determination is obtaining diffraction pattern of the crystal by incidenting monochromatic X-ray beam. X-rays are produced when electrons generated from a heated filament accelerated by high voltage collide with atoms from a metal target such as copper. X-rays are electromagnetic radiation of wavelength about 0.1 – 10 nm (1 - 100 Å) (Blundell & Johnson, 1976). The amount of details or the resolution that can be viewed is limited by the wavelength of the electro-magnetic radiation used. For visualizing the atomic details of a protein molecule, the electromagnetic radiation required is within wavelength of around 0.1 nm or 1 Å, which corresponds to that of X-rays, hence these X-rays are the EM radiation of choice (Blow, 2002).

Since a single molecule is a very weak scatterer of X-rays. Most of the X-rays will pass through a single molecule without being diffracted, so the diffracted beams are too weak to be detected. Analyzing diffraction from crystals, rather than individual molecules, solves this problem. A crystal of a protein contains many ordered molecules in identical orientations, so each molecule diffracts equivalently, and the diffracted beams for all molecules augment each other to produce strong, detectable X-ray beams. If the crystal is well ordered, then diffraction will be measurable at high angles or high resolution and results in detailed information on protein structure. The X-rays are diffracted by the electrons in the structure and consequently the result of an X-ray diffraction experiment is a 3-dimensional map showing the distribution of electrons in the structure. Each reflection contains information on all atoms in the structure and conversely each atom contributes to the intensity of each reflection.

X-ray diffraction involves the set of experiments to measure intensities of Bragg reflections from the crystals. The materials needed are a crystal, an X-ray source, an area detector (Image plate or CCD), cryostream (if done at low temperature), and a goniometer to orient and rotate the crystal. 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.

Unwanted radiation is filtered using monochromator and the radiation is then focused using aligned mirrors. The most commonly used radiation from this source is the CuKα (λ = 1.5418Å) radiation. Synchrotrons pro duce 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. Substantial increase in resolution over those available with laboratory sources is another major advantage of the synchrotron facility. 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. An electron stream produced is then fed into a linear accelerator. Here the electrons assume closer to that of speed of light. The linear accelerator feeds electrons into the booster ring which uses magnetic fields to force the electrons to travel in a circular path. 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 as result of fast electron interacting with a magnetic field. Magnetic field will cause the electron to change direction by exerting a force on it perpendicular to the direction of electron's motion resulting in the acceleration of electron, causing it to radiate electromagnetic energy. If the electrons and the magnetic field have sufficient energy, the emitted radiation can be in the form of X-rays.

In a macromolecular X-ray diffraction experiment the protein crystal is either mounted in a low diffracting glass capillary (if data is to be collected at room temperature), or it is mounted on a loop made up of nylon fibres after brief soaking in the cryoprotectant (in case of low temperature data collection) as described in the next section. The capillary or loop is then mounted on a goniometer, which allows the crystal to be positioned accurately within the X-ray beam and rotated as required. The mounted crystal is then irradiated with a beam of monochromatic X-rays and the diffraction pattern in the form of spots from scattered X-rays is recorded on the detector placed at an appropriate distance.

During data acquisition the crystals were oscillated about an axis perpendicular to the X-ray beam, with a relatively small angle of oscillation chosen close to 0.2 - 0.3 degree per frame at the synchrotron or 0.5-1.0 degree per frame at the home source. Crystal to detector distance is chosen, based on the longest unit cell dimension, mosaic

spread etc., so that the diffraction spots are well resolved, and is often approximately equal to the longest crystal cell dimension. The exposure time depended on the quality of the crystal, the oscillation range, larger the oscillation range longer the exposure time required and the X-ray source, with synchrotron requiring much shorter exposure time compared to less intense beam of in-house source.

2.3.8.3. Cryocrystallography or Flash Cooling of Protein Crystals

When crystals are exposed to intense X-ray radiation, it can damage the crystals. This is due to the weak lattice forces within the crystal structure (McPherson, 1982) and the X-rays can produce sufficient free radicals to cause specific chemical changes on the protein molecules such as to break the disulfide bonds (Ravelli & Garman, 2006; Ravelli *et al*., 2003). This problem can be overcome to some extent by collecting the data at lower temperatures.

The crystals are prepared and rapidly flash cooled to prevent ice lattice formation in the aqueous solvent inside and on the surface of the crystal. In the presence of cryoprotectant, instead of ice, glass forms, leaving the crystal lattice with little or no damage. This provides crystals with a sort of 'immortality' in the X-ray beam and enable better quality data with the capability of long-term storage and reuse of crystals and often extending the limit of resolution obtainable (Gamblin & Rodgers, 1993; Garman & Schneider, 1997). In an effort to minimize lattice damage during cryogenic cooling, cryoprotective reagents or cryoprotectants are added to the solution to prevent the formation of crystalline ice in the internal and external solution as well as at the interface between crystal and solution. Organic solvents such as MPD, methanol, ethanol, isopropanol, ethylene glycol and glycerol are the commonly used cryoprotectants (Garman & Schneider, 1997). Glycerol is one of the most common cryoprotectants used in protein cryocrystallography. Lower molecular weight PEGs, glucose, sucrose and xylitol are also been widely used as cryoprotectant.

The intrinsic limitations with resolution due to diffraction quality had also been dramatically improved by data collection at cryogenic temperature. Many factors contribute to the improvement in data quality, including reduced thermal vibrations, decreased diffused scattering and reduced absorption due to absence of glass capillary and absence of excessive solvent surrounding crystal, enhancement of the signal-to-

noise ratio, reduced conformational disorder and potentially improved limiting resolution. Prevention of radiation damage has also resulted in complete data collection from a single crystal that in turn eliminates errors from merging and scaling multiple data sets from different crystals.

The introduction by Teng (1990) of a loop mounting technique for flash cooling of crystals was a major advancement in cryocrystallography. Loops are made of fine fibres, e.g., rayon, in which the crystal is scooped from the cryosolution and held within the loop suspended by a thin film of the solution. The loop is supported by a pin, which is attached to a steel base used for placing the assembly on a magnetic cap on the goniometer.

All of the above mentioned cryoprotectants have been tried to screen the cryo-conditions for *Af*PGA crystals, whereas Glycerol served the purpose in the cases of *Kc*PGA crystals and *Ld*NMT mutant crystals.

2.3.8.4. Data Collection and Data Processing

Data were collected on *Af*PGA crystals at room temperature (293K) from a crystal mounted inside a sealed thin glass capillary of 1.5 mm in diameter (Hampton Research) on the in-house X-ray diffraction facility at NCL, Pune, India using CuKα (λ = 1.5418Å) radiation using Raxis IV^{++} area detector mounted on a Rigaku rotating anode X-ray generator operating at 50 kV, 100 mA with Osmic focusing mirrors to monochromate and focus X-rays. The determination of the orientation of the crystals and initial processing of the diffraction data were performed using the *CRYSTALCLEAR* software supplied by Rigaku-MSC. After indexing the initial images the module "strategy" was used to get the range of the rotation angles to be used to acquire complete data. The best strategy during X-ray diffraction data collection depends on qualitative factors, such as crystal quality, type of X-ray source and detector, and quantitative ones, such as cell parameters, resolution limit, and crystal symmetry and its orientation in the X-ray beam, which can be determined by auto-indexing a preliminary image as input. On getting this information, the program determines the starting phi value and the total number of images required for collecting dataset with the highest completion possible. The crystal-to-detector distance was kept at 260 mm for orthorhombic crystals and 300 mm for tetragonal crystals for *Af*PGA

crystals. Successive frames were collected using a crystal oscillation of 0.5° about an axis perpendicular to the direction of the X-ray beam. The exposure time for data collection was chosen depending on the intensity of spots. For Orthorhombic data collection exposure time given was 2-3min, whereas for tetragonal crystals it was 5-6min.

Data from the initial crystals obtained from slow processing mutant of *Kluyvera citrophila* were collected at 100 K using 30% glycerol as cryoprotectant on Stanford Synchrotron Radiation Lightsource (SSRL) beam line BL12-2 at the SLAC National Accelerator Laboratory. Diffraction data were collected on DECTRIS PILATUS 6M detector with 0.2 sec exposure and 90% attenuation.

Crystals obtained for *Ld*NMT mutant were checked for diffraction at in-house X-ray diffraction facility at York Structural Biology Laboratory, University of York, York, UK using in-house X-ray facility at 100K using MAR345 detector mounted on Rigaku RUH3R generator equipped with Osmics multilayer optics. Subsequently, diffraction data from crystals were collected at Diamond Light Source synchrotron facility at Oxford, Didcot, UK using ADSC Q315r CCD detector. All data were collected at 100 K under liquid nitrogen with oscillation of 0.2° with exposure of 0.2 sec.

The analysis and reduction of single crystal diffraction data collected consists of the following steps:

1. Indexing of the diffraction pattern and determining of the crystal orientation

2. Refinement of the crystal and detector parameters

3. Integration of the diffraction intensities

4. Refining the relative scale factors between equivalent measurements

5. Precise refinement of crystal parameters using all the data

6. Merging and statistical analysis of the symmetry related reflections

First three steps are carried out by the program DENZO, while steps 4 to 6 are performed by the program, *SCALEPACK*. *DENZO* and SCALEPACK modules are the part of HKL suite (Otwinowsky & Minor, 1997). *DENZO* reads the diffraction data and outputs a file containing the intensity of spots, the estimated errors and the corrected

intensities. The scaling and the merging of the reflections and postrefinement are done using *SCALEPACK*, and it outputs the final intensities and the standard deviation and also the data quality statistics. *MOSFLM* can also be used for data processing. *MOSFLM* (Leslie & Powell, 2007) performs the functions similar to that of *DENZO*, while *SCALA* (Evans, 2006) of the *CCP4* suite is used to scale the reflections (Collaborative Computational Project No.4, 1994). *XDS* (Kabsch, 2010) and *xia2* (Winter, 2010) are another packages available for automated data reduction of the raw diffraction images.

Diffraction data obtained from *Af*PGA orthorhombic and tetragonal crystals were processed using *DENZO* and *SCALEPACK*. The diffraction data from slow processing mutant of *Kc*PGA and from monoclinc crystal of *Ld*NMT mutant were processed using automated *XDS* package, whereas triclinic data of *Ld*NMT mutant was processed using automated *xia2* package and scaled using *SCALA* programme in CCP4 suite.

2.3.8.5. Data Quality Statistics

The quality of the data collected and the resolution limit are decided based on the data quality statistics. The quality parameters output by *SCALEPACK* includes $\chi2$, R_{merge} and $<I/\sigma I>$. These parameters are defined by the equations given below (Blundell and Johnson, 1976, Otwinowsky & Minor, 1997) (Equations **2.5** and **2.6**).

$$R_{merge} = \frac{\sum_{h} \sum_{i} |I_i - \langle I \rangle|}{\sum_{h} \sum_{i} I_i}$$

(2.5)

$$\chi^2 = \Sigma \left[(I_i - <I_i>)^2 / (\sigma I^2 N / (N-1)) \right]$$ (2.6)

where I_i is the **i**th intensity measurement of reflection h, and $<I>$ is the average intensity from multiple observations.

One of the least defined criteria in data collection is the resolution limit of diffraction (Dauter, 1997, 1999). It is important to restrict the resolution to the point below which more than half the reflection intensities are higher than 2σ. This assumes that the estimate of the errors in measurement of intensities is correct. The R_{merge} (also called R_{sym} when only symmetry equivalents of a single dataset are used) is the most

widely accepted statistical parameter to indicate data quality in macromolecular crystallography. For a good data set, Rmerge will be less than 10%. The parameter σ (I) assigned to each intensity derives its validity from the χ2's (or the goodness-of-fit). A good refinement will have χ2 values closer to 1.0 in the different resolution bins. Very large values of χ2 may be an indication of serious problems with data processing. The R$_{merge}$ is less appropriate than the Chi-square, because of its dependency on the multiplicity of the data and on the symmetry of the crystal (Diederichs & Karplus, 1997). R$_{merge}$ alone is not sufficient to decide the high resolution cut-off because of its dependence on the redundancy of data (Weiss, 2001; Weiss and Hilgenfeld, 1997). The high resolution cut-off is decided such that the highest resolution bin has a R$_{merge}$ value less than 50% with at least 50% completion and <I/σI> value more than 2.0.

2.3.8.6. Matthews Number and Solvent Content

Once the space group and unit cell dimensions of the crystal are known, it is possible to estimate the number of molecules in the crystallographic asymmetric unit and the solvent content of the protein crystals based on knowledge of the molecular weight of protein. The following equations (**2.7** & **2.8**) are used (Matthews, 1968).

$$\textbf{Vm= Unit cell volume /Mol.Wt.* n * z} \qquad \textbf{(2.7)}$$

$$\textbf{V}_{\textbf{solv}} = \textbf{1} – \textbf{(1.23/ Vm)} \qquad \textbf{(2.8)}$$

where Vm is the Matthews number, n is the number of molecules per asymmetric unit and z is the Avagadro's number; V$_{solv}$ is the fraction of unit cell volume occupied by solvent. Cell Content Analysis module from the *CCP4* suite was used for the purpose. The Matthew's number for both the crystal forms of *Af*PGA were calculated by assuming 86500 Dalton as the molecular weight of a processed monomer whereas molecular weight of precursor PGA was taken as 93000 Dalton. Matthews number for *Ld*NMT mutant was calculating assuming molecular weight as 48700 Dalton.

2.3.8.7. Structure solution

The data collected and processed as described above from a diffraction experiment is defined as the reciprocal space representation of the original crystal lattice. The position of each diffraction 'spot' is depended on the size and shape of the

unit cell as well as the inherent crystal symmetry. The intensity of each diffraction 'spot' is measured, and structure amplitudes are directly obtained from reflection intensities as the square root of the intensities measured on the detector. Since X-rays, like ordinary visible light, are electromagnetic waves, they possess intensity as well as a phase. In order to work backwards, from diffraction patterns to crystal and molecular structures, it is necessary to measure, in addition to the intensities of the each reflection, their phases as well. However, only the intensities can be directly measured, but the phases cannot be directly measured in the ordinary diffraction experiment; they appear to be irretrievably lost. This is known as the "**phase problem**", the central problem of X-ray crystallography.

The first step in going from experimental X-ray data to protein structure is to estimate the phase information. Structure solution can be based on the many different techniques, depending on the data and information available. Multi-wavelength Anomalous Dispersion (MAD), Multiple Isomorphous Replacement (MIR) and Molecular Replacement (MR) are the three methods widely used in protein crystallography to solve protein structures. MAD technique is currently the most popular one owing to the availability of variable wavelength synchrotron radiation sources and sensitive detectors like Image plate and CCD to collect anomalous data on selenium derivative that can be prepared using molecular biology techniques for any protein (Beauchamp & Isaacs, 1999). Another advantage is MAD needs only one heavy atom derivative compared to MIR technique that requires more than one derivative. Molecular replacement technique is the simplest of all and can be used when a homologous protein with structural similarity is available in the database. The success of the MR method depends on the structural similarity between the search model and the unknown structure. Rossman & Blow (1962) first described the successful use MR methods. With the continuous increase in number of structures deposited in protein data bank, MR method is now routinely used in protein structure determinations.

2.3.8.8. Molecular replacement method

The crystal structures reported in this thesis were solved using Molecular Replacement (MR) technique. 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-dimensional problem.

The basic principle of molecular replacement is rooted in the Patterson function. Patterson methods (Rossmann & Blow, 1962) are used to calculate the rotation function, to obtain the orientation of the model in the new unit cell, and then the translation function calculated using model in correct orientation (Crowther & Blow, 1967) helps to place the model in the new unit cell. Commonly used programs for molecular replacement are *AMoRe* (Navaza, 1994), *PHASER* (McCoy *et al.*, 2007; Storoni *et al.*, 2004) and *MOLREP* (Vagin & Teplyakov, 1997).

AMoRe uses the fast rotation and translation functions for molecular replacement. It is reasonably fast and ranks the solution based on the R_{factor} R and correlation coefficient CC for a translation "t". Usually, R_{factor} less than 50% and correlation coefficient greater than 30% are indicators of a correct solution, but the final test is always the quality of the electron density map.

MOLREP (Vagin & Teplyakov, 1997) uses Crowther's fast rotation function, but differs from *AMoRe* in the translation function used. The modifications to the translation function results in an increase in the contrast. The incorporation of the packing function removes peaks of incorrect solutions with bad packing. It identifies correct solutions based on the contrast function.

PHASER (McCoy *et al.*, 2007) uses maximum likelihood based methods for performing the rotation and translation searches. It has been found to be useful where many other molecular replacement programs fail to give a solution. It has been especially successful in problems involving high symmetry space groups with large number of molecules in the unit cell and those involving low homology models. The success of a *PHASER* run is determined based on the parameters, log likelihood gain (LLG) and Z-score. The LLG is the difference between the likelihood of the model and the likelihood calculated from a Wilson distribution. It is a measure of how much better the data can be predicted with the model than with a random distribution of the same atoms. Z-score stands for the number of standard deviations over the mean value of the

peaks in the rotation or translation functions. A Z-score value greater than 8 usually indicates that a solution has been obtained (McCoy *et al*., 2007).

AutoMR from *PHENIX GUI* (McCoy *et al*., 2007; Adams *et al*., 2002) carries out automated likelihood-based molecular replacement using *PHASER* (McCoy *et al*., 2007). The procedure is highly automated, allows several copies of each of several components to be used as ensembles in a single run, and also tests different possible choices of space group. If there are alternative choices of model for a component, the molecular replacement calculation tries each of them in turn or combines them as a statistically weighted ensemble (Adams *et al*., 2010).

AutoMR programme was used to obtain an initial molecular-replacement solution for the dataset obtained for slow processing mutant of *Kluyvera citrophila* PGA. The structure of slow processing precursor of *E.coli* PGA (pdb code: 1E3A, Hewitt *et al*., 2000) was used a model. *PHASER* implemented in CCP4 was used as a molecular replacement package for the structure determination for both *Af*PGA datasets using structure of processed PGA from *E.coli* (pdb code: 1GK9, McVey *et al*., 2001). *MOLREP* from *CCP4* suite was used for the structure determination of $P2_1$ datasets and *PHASER* from *CCP4* suite was used as molecular replacement program for the P1 datasets in the case of *Ld*NMT mutant taking pdb coordinate of *L.major* NMT (pdb code: 2WSA, Frearson *et al.*, 2010) as a search model. The correct solution in each case was decided based on estimated values of the Z score and R factor.

2.3.8.9. Automated model building

The *AutoBuild* wizard (Terwilliger *et al.,* 2008) from *PHENIX* was applied to structure rebuilding of a model derived from molecular replacement for slow processing mutant of *Kc*PGA. AutoBuild combines density modification and chain tracing using *RESOLVE* (Terwilliger, 2003a; 2003b) with refinement using *phenix.refine* (Afonine *et al*., 2005, 2012) to generate a high-quality model. In the case of molecular replacement, the success of automated model building is a strong indicator of the correctness and completeness of the molecular replacement solution. Model building approach using model from *AutoMR* as a starting model was used in this case.

Initial model building was performed using *Buccaneer* (Cowtan, 2006) in the case of *Af*PGA structures. Buccaneer is a software tool implemented in *CCP4* suite to

trace protein structures in electron density maps by identifying connected alpha-carbon positions using a likelihood-based density target. The underlying principle of Buccaneer is the repeated application of an electron density likelihood function to identify probable oriented Cα positions in a noisy electron density map. The same likelihood is applied in several ways: to find candidate positions, using a six dimensional search, to grow a chain by adding new residues either side of an existing Cα position, and to refine Cα positions. The method is reasonably fast; taking minutes to an hour and can give a partial trace even at low resolutions (i.e. worse than 3.0Å) (Cowtan, 2006). However the method is dependent on the quality of the initial experimental phasing and phase improvement results. Since the resolution for the datasets obtained from the *Af*PGA crystals were no better than 3.3Å. Buccaneer program helped building 660 residues in the $C222_1$ crystal and 740 residues in the $P4_12_12$ crystal.

2.3.8.10. Structure refinement

The next step after getting the coordinates of a preliminary model using MR method is its refinement. Refinement is the process of fitting the parameters of the model to obtain a closer agreement between the calculated atomic model and observed diffraction data. This is achieved by minimizing the discrepancy between the calculated (Fcalc) and observed structure factors (Fobs). One of the problems in macromolecular refinement is the poor observation to parameter ratio. Hence refinement of the model is done by incrementing the positional parameters and the temperature factors of the atoms applying stereochemical restraints on the structure. Constraints are incorporated into the refinement in the form of stereochemical criteria deduced from data of small molecular structures of amino acids and peptides in which the bond lengths and angles have been determined to high precision (Engh & Huber, 1991). The stereochemical information can be applied in two different ways. For constrained refinement the geometry is considered as rigid which reduces the number of parameters to be refined. Whereas, when the specific parameter has restricted freedom around a standard value, the refinement is considered to be restrained. Restraints on bond lengths, bond angles, torsion angles, and van der Waals contacts have the apparent effect of increasing the number of observations.

2.3.8.11. Max-likelihood Refinement

Refinement of two of three structures studied in this thesis was carried out using restrained maximum likelihood refinement implemented in the program *REFMAC5* (Murshudov *et al.,* 1997). The program minimizes the coordinate parameters to satisfy a maximum-likelihood or least squares residuals. One of the expected advantages of using maximum likelihood refinement method is a decrease in refinement bias, as the calculated structure-factor amplitudes will not be forced to match the observed amplitudes (Read, 1997). Use of appropriate likelihood targets through the incorporation of the effect of measurement of error and the use of cross-validation data to estimate the values (error) are the key ingredients in the likelihood refinement. For protein structure refinement, maximum likelihood method is more than twice as effective compared to the least squares method, in improving the model (Pannu & Read, 1996). *REFMAC5* is a refinement module in which each cycle carries out two different steps:

(a) First is the estimation of the overall parameters of likelihood using the free set of reflections.

(b) And secondly it uses these parameters to build the likelihood function and refine the atomic parameters. To refine the atomic parameters only a working set of reflections are used.

Maximum likelihood method of refinement performs the calculation of the first derivative and makes an approximation of the second derivative of the likelihood function with respect to refinement parameters and then estimating the shifts to be added to the parameters. The values of standard deviations are calculated for the geometric restraints used during the refinement. The restraints used during refinement include bond distances, bond angles, peptide planarity, torsion angles, chirality, van der Waals radii and temperature factors. Symmetry between the molecules in the asymmetric unit is termed as noncrystallographic symmetry (NCS). The presence of NCS is helpful during the structure determination process. NCS information can also be used as restraints during initial refinement.

Initial model obtained after molecular replacement followed by model building for *Af*PGA crystals and molecular replacement model for the *Ld*NMT mutant crystals were refined using *REFMAC5* from the pre-release CCP4 v.6.1. *phenix.refine* (Afonine

et al., 2005) was used for the automated refinement of molecular replacement model improved by automated model building by *AutoBuild* for the dataset obtained for the slow processing mutant of *Kc*PGA.

Since *Kc*PGA and *Ld*NMT proteins had more than one molecule in their asymmetric units, their initial models were refined taking advantage of the presence of non-crystallographic symmetry.

2.3.8.12. Visualization and interpretation of electron density maps

*REFMAC*5 also produces an output file with extension MTZ (named after three of its progenitors, McLaughlin, Terry and Zelinka) containing weighted coefficients for weighted mFo-DFc and 2mFo-DFc maps. A 2Fo–Fc map gives more weightage to the observed structure factors, and hence reduces some of the model bias. It is normally viewed at a contour level of around 1.0σ. Fo–Fc map is examined at both positive and negative contours. Presence of a positive peak is an indication of unaccounted density, while a negative peak shows that the model contains features not conforms to the actual structure. After every refinement cycle the model was inspected and manual rebuilding done by inspecting (2Fo-Fc), positive (Fo-Fc) and negative (Fo-Fc) electron density maps. The programs *QUANTA* and *COOT* ((Emsley & Cowtan, 2004) were used for displaying and examining the electron density maps, for displaying atoms, interactive fitting and optimizing the geometry and also for adding water molecules. The addition of water molecules was done at the final stages of refinement of the model. The model that emerged after each cycle of model building and refinement was checked for unusual geometry at each residue, the peptide-flip values, and high temperature factors. The other parameters checked include values of torsion angles, including (Ramachandran angles φ and ψ), side chain torsion angles such as χl, χ2 etc. were verified between refinement cycles using the program *PROCHECK*.

2.3.8.13. Structure validation and evaluating the quality of structure

The cycle of rebuilding the model in *QUANTA* and *COOT* and restrained refinement in *REFMAC5* are followed by structure validation. The correctness and precision of the atomic parameters in a structure need to be assessed thoroughly, both during and at the end of the refinement. An initial indication of the reliability of a

protein crystal structure depends on the extent of resolution of the diffraction data used in refinement. The conventional R_{factor} (R_{cryst}) in equation (**2.9**) is an indicator of the general fit of the crystal structure with observed data. However, it is not a reliable independent validator, since it can be biased by the exclusion or overfitting of the data. A better assessment of the fit between observed and calculated structure factors is by calculating the free R factor (R_{free}) (Brunger, 1992).

$$\mathbf{R_{factor} = \Sigma \mid \mid Fobs \mid - \mid Fcalc \mid \mid / \Sigma \mid Fobs \mid} \qquad\qquad (2.9)$$

Despite stereochemical restraints, it is possible to overfit or misfit the diffraction data: an incorrect model can be refined to fairly good R value. To overcome this problem, the Rfree a statistical quantity was proposed. In cross validation, the data is randomly divided into working set and test set. The test set contains about 5-10% of the reflections that are not included in the refinement process. The rest of the data forms the working set and refinement is carried out using only these reflections. Rfree is calculated according to equation (**2.9**) using only the reflections from the test set. If the model is correct, the R_{factor} and R_{free} should be close, as the changes to the model should affect both of them similarly. Thus, monitoring of R_{free} serves as a check against over-fitting of the data. A difference between R and R_{free} of about 6% is accepted. The R-free value is also correlated with the accuracy of atomic models. In practice, this means that models with serious errors can be identified by a very high free R value (>0.40) irrespective of the value of the conventional R value, which may be very low (around 0.20). The advantage of the R_{free} value over knowledge based validation methods is that it can be applied to any type of model and does not depend on the availability of database-derived knowledge (Brunger, 1992).

2.3.8.14. Estimate of coordinate precision

Isotropic temperature factors are related to the thermal vibrations of atoms and therefore to some extent a reflection of the precision with which these atomic positions can be measured. It has been demonstrated that precision of atomic positions are strongly depended on magnitudes of B values (Daopin *et al.,* 1994). The program *PROCHECK* (Laskowski *et al.,* 1993) and *SFCHECK* (Vaguine *et al*., 1999) have been used to examine the stereochemistry of the final refined models. *PROCHECK* calculates the phi-psi angles and plots them on a two dimensional plot (Ramachandran

plot) (Ramachandran & Sasisekharan, 1968). It estimates the percentage of residues within or outside the allowed or partially allowed regions in Ramachandran plot. The program output a detailed list of the extend of deviations from the ideal geometry for residues along with corresponding plots of the parameters. A good model will have more than 90% of the residues in the allowed regions of the Ramachandran plot and none will be in the disallowed region.

2.3.8.15. Comparative Analysis of the Structures

The 3-D structures of the PGAs of *P. rettgeri* (PDB code: 1CP9) and *E. coli* (PDB codes: 1PNK and 1E3A) and NMT from *Leishmania donovani* (PDB code: 2WUU) and *Leishmania major* (PDB code: 2WSA) were extracted from the Brookhaven *PDB* (McVey *et al.,* 2001; Duggleby *et al.,* 1995; Hewitt *et al*., 2000; Brannigan *et al*., 2010; Frearson *et al*., 2010). The structures were superimposed in *CCP4MG* and a sequence alignment was derived from the structural equivalences using *SwissPdb* Viewer (http://www.expasy.org/spdbv/). The protein secondary structure motifs were computed using *ProMotif* programme (Hutchinson & Thornton, 2008) available at *PDBsum* server (http://www.ebi.ac.uk/pdbsum/). Protein Structure Classification was done using *CATH* (http://www.cathdb.info/) (Orengo *et al*., 1997).

2.3.9. Calculation of thermostability factors

The amino acid composition of *Af*PGA and its correlation with thermostability were also investigated. The protein sequence of the mature *Af*PGA, *Ec*PGA and *Pr*PGA were obtained using *UniProt* database (http://www.uniprot.org/) and were aligned by using the *ClustalW* program (Thompson *et al.,* 1994). Number of residues and relative percentage of each amino acid in the protein were calculated from the protein sequence. The salt bridges were calculated through the *WHAT IF* Web Interface (http://swift.cmbi.ru.nl/servers/html/index.html) (Vriend, 1990) using the pdb coordinates. A salt bridge is defined as a negative atom (side chain oxygens in Asp or Glu) and a positive atom (side chain nitrogens in Arg, Lys, or His) with an interatomic distance of less than 7.0 Å.

All the figures in the thesis have been generated using *CCP4MG* (McNicholas *et al*., 2011) or stated otherwise.

Chapter: 3

Cloning and Over-expression of *pac* gene from *Kluyvera citrophila* and Purification and Biochemical characterization of the expressed *Kc*PGA

3.1. INTRODUCTION

Almost 85% of 6-APA production for the manufacture of semi-synthetic penicillins is by using penicillin G acylase (PGA), the enzyme that hydrolyses penicillin G (benzyl penicillin), with the major share of the enzyme sourced from *E. coli*. However, for pharmacological applications when factors such as tolerance for environmental factors such as temperature, pH and nature of solvent, along with ease of immobilization are important, attention has turned to PGA enzymes from other sources such as *Kluyvera citrophila* (*Kc*PGA) (Alvaro *et al*., 1992; Fernandez-Lafuente *et al.,* 1991, 1996; Liu *et al.,* 2006).

Like many other members of Ntn hydrolase superfamily (Brannigan *et al.,* 1995), the *Kc*PGA is also produced as an inactive precursor (pre-pro-PGA). The precursor PGA is then processed by the autocatalytic removal of a 26 residue signal peptide and a 54 residue spacer peptide posttranslationally, to produce active mature enzyme in the periplasm which is a heterodimer of α and β chains of 209 and 557 amino acid residues, respectively. The cloning of this enzyme and overexpression in *E. coli* had been reported (Garcia *et al*., 1986; Barbero *et al*., 1986; Wen *et al*., 2004, 2005). Its nucleotide sequence is 80% similar to *E. coli* ATCC 11105 PGA (Schumacher *et al.,* 1986), indicating a common ancestor. It has been reported that altering amino acid residue Glyβ21 affects protein maturation (Prieto *et al*., 1992). The serine residue with a newly generated free α-amino group at the N-terminus of β-subunit acts both as nucleophile and as base in catalysis. The importance of the Arg α145 in this enzyme has been shown in our lab using chemical modification (Kumar *et al*., 2004). 82% and 78% inactivation was observed when the enzyme was chemically modified with arginine modifying reagents 20 mM phenylglyoxal and 50 mM 2, 3-butanedione respectively. Inactivation was prevented by protection with benzylpenicillin or phenylacetate at 50 mM concentration confirms that arginine residue is present at the active site of the enzyme and unavailable when substrate or inhibitor is bound (Kumar *et al*., 2004). Presence of Trpβ154 in active site of the enzyme has also been shown by Kumar *et al.* (2007), using chemical modification, fluorescence and docking studies. Kinetic studies confirm the involvement of this

residue in substrate binding. Sequence comparison with other PGAs showed the Trpβ154 is conserved in closely related PGAs that have better affinity for Pen G.

The most studied maturation pathways of PGA have been that of *E. coli* ATCC 11105. This process contains two essential steps: translocation of the precursor to the periplasmic membrane and processing of the precursor by an autocatalytic intramolecular peptide-bond cleavage (**Figure 3.1**). The exact mechanism of autocatalytic processing of PGA precursor is complex; however, several studies have resulted in accumulating sufficient evidence to suggest a plausible mechanism of activation. The first cleavage removes the 26 amino acid signal peptide and then, at least four subsequent proteolytic cleavages are necessary to generate the mature enzyme (Choi *et al*., 1992; Bock *et al*., 1983; Oh *et al*., 1987). One of these four cleavages includes the cleavage of the Thr289-Ser290 bond to expose the catalytic Ser290. Cleavage of the Thr289-Ser290 bond results in detaching C-terminal end of the linker peptide, unveiling the primary amine group of Serβ1 (Serβ290 in the precursor structure), forming the active centre of mature PGA. The rate-limiting step in the production of active enzyme is the intramolecular autoproteolytic processing of the precursor molecule, resulting in the removal of a linker peptide (Kasche *et al*., 1999; Hewitt *et al.,* 2000; Done *et al.,* 1998).
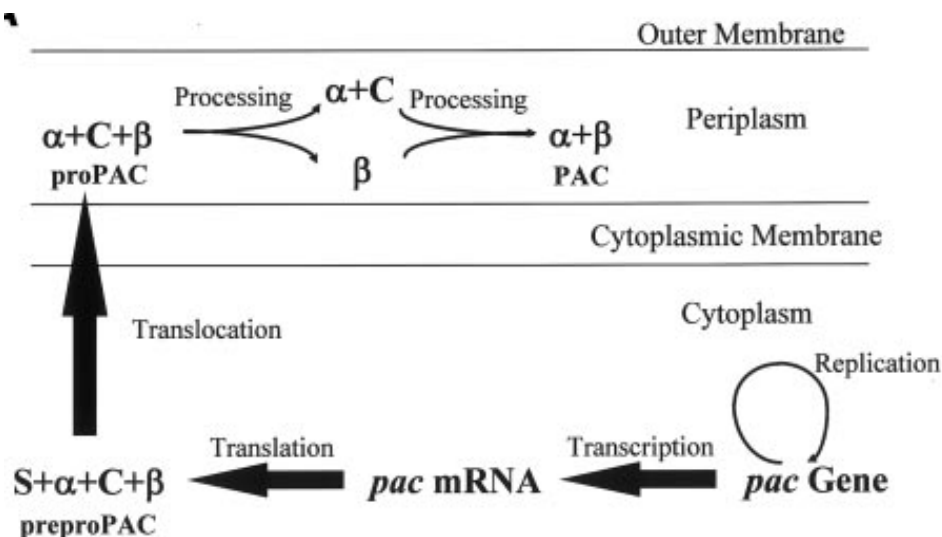
Figure 3.1: Showing maturation of Active PGA depicting autocatalytic processing of PGA from *E. coli* (Figure adopted from Xu *et al*., 2005)

Lee *et al*. (2000) showed that *in vitro* processing of the precursor exhibited the same behaviour as observed in the *in vivo* studies and was also pH dependent in the

same manner, with an optimum pH for processing from 6.4 to 7.0, physiological pH. A primary amino acid sequence alignment study of the precursors from four gram negative and two gram positive bacteria and the site directed mutagenesis study involving GST precursor PA fusion protein expression system conserved lysine residue (K299), which is near the N-terminal serine residue (S290), could be the most probable candidate responsible for the pH dependent activation. Therefore possible activation centre includes both Lys299 and Ser290 as critical residues for the autocatalytic processing of the PGA precursor (Lee *et al.*, 2000).

This chapter deals with the cloning of penicillin acylase (*pac*) gene from chromosomal DNA of *Kluyvera citrophila* DMSZ 2660 (ATCC 21285), over-expression, purification and biochemical characterization of expressed wild-type enzyme. This chapter also describes the site directed mutagenesis technique used to prepare two mutants of *Kc*PGA enzyme that resulted in the expression of either a slow processing mutant precursor enzyme or processed but catalytically inactive enzyme depending on the particular amino acid chosen to replace the nucleophile Ser290.

3.2. RESULTS

3.2.1. Isolation and Coning of "*pac*" gene from *Kluyvera citrophila* DMSZ 2660 (*KcPAC*)

      A 2562 bp PCR fragment covering the region 12 nucleotides upstream from the start codon of *K.citrophila* "*pac*" gene and 12 nucleotide downstream was amplified from chromosomal DNA of *K. citrophila* DMSZ 2660 (ATCC 21285) as a template, using gene specific primers designed according to the published coding sequence (Barbero *et al.*, 1986, Accession No- M15418)

*Pac*NdeF   5'-caagaggat**catatg**aaaaatagaaatcgtatgatcgtg-3' and

*Pac*XhoR   5'-gccgaa**ctcgag**gcgctgtacctgcagcactt-3'

(Eurofins MWG Operon, Ebersberg, Germany) with NdeI and XhoI restriction sites shown underlined.

      The gene was amplified using PCR under the following conditions. Initial denaturation was for 3 min at 98ºC. The denaturation step at 95 ºC was for 15 sec, annealing of primers at 61 ºC for 30 sec and extension at 72ºC for 50 sec. These steps were repeated for 30 cycles using KOD polymerase (**Figure 3.2**).
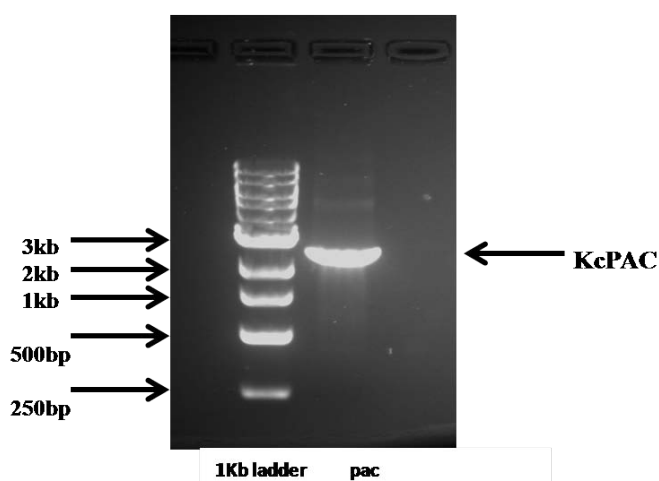
Figure 3.2: PCR amplification of *Kcpac* gene as visualized on 1% agarose gel

      The PCR products were digested with NdeI and XhoI and purified using gel extraction kit. The vector pET-26b (+) (**Figure 2.3**) was digested using same restriction enzymes, dephosphorylated with Shrimp alkaline phosphatase and purified with PCR

purification kit (Promega). Both restriction enzyme digestions were carried out at 37 ºC with suitable buffers. Purified vector and purified insert were ligated in 3:1 molar ratio of insert to vector with T4 DNA Ligase (NEB) with suitable buffer at 16 °C overnight and transformed into NovaBlue competent cells and confirmed by restriction analysis (**Figure 3.3**).
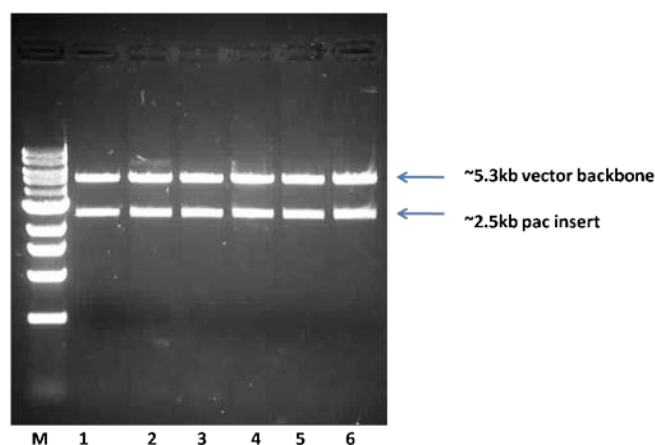
Figure 3.3: Clones of *Kcpac* in pET 26b (+) digested with NdeI and XhoI as visualized on 1% agarose gel; lanes 1-6 are digested products, lane M is 1kb ladder.

After confirmation by restriction digestion, 3-4 plasmids were sequenced using standard T7 forward, T7 reverse as well as gene specific primers. Sequencing result confirmed ORF of 2565 bp containing "*pac*" gene coding sequence of 2538bp encoding 855 amino acids shown as bold in **Figure 3.4** (rest coding for hexahistidine tag and stop codon).

Nucleotide Sequence

ATGAAAAATAGAAATCGTATGATCGTGAACGGTATTGTGACTTCCCTGATCTGTTGTTCTAGCCTGTCAGCGCTGGCG
GCAAGCCCGCCAACCGAGGTTAAGATCGTTCGCGATGAATACGGCATGCCGCATATTTACGCCGATGATACCTATCGA
CTGTTTTACGGCTATGGCTACGTGGTGGCGCAGGATCGCCTGTTCCAGATGGAAATGGCGCGCCGCAGTACTCAGGGG
ACCGTCTCCGAGGTGCTGGGCAAAGCATTCGTCAGTTTTGATAAAGATATTCGCCAGAACTACTGGCCGGATTCTATT
CGCGCGCAGATAGCTTCCCTCTCCGCTGAGGATAAATCCATTCTGCAGGGCTATGCCGATGGCATGAATGCGTGGATC
GATAAAGTGAACGCCAGCCCCGATAAGCTGTTACCCCAGCAGTTCTCCACCTTTGGTTTTAAACCCAAGCATTGGGAA
CCGTTTGATGTGGCGATGATTTTTGTCGGCACCATGGCGAACCGTTTCTCTGACAGCACCAGCGAAATTGATAACCTG
GCGCTGCTGACGGCGCTAAAAGATAAATACGGCAAGCAGCAGGGCATGGCGGTCTTTAACCAGCTGAAATGGCTGGTT
AATCCTTCCGCGCCAACCACCATTGCGGCGCGGGAAAGCGCCTATCCGCTGAAGTTTGATCTGCAAAACACGCAAACG
GCGGCGCTGCTGCCGCGCTACGACCAGCCGGCACCGATGCTCGACCGCCCGGCAAAAGGGACCGATGGCGCGCTGCTG
GCGCTGACCGCCGATCAGAACCGGGAAACTATCGCCGCGCAGTTCGCGCAAAGCGGCGCTAACGGCCTGGCTGGCTAC

```
CCGACCACTAGCAATATGTGGGTGATTGGCAAAAACAAAGCCCAGGATGCGAAGGCCATTATGGTCAATGGGCCGCAG
TTTGGTTGGTATGCGCCGGCGTACACCTACGGTATCGGCCTGCACGGCGCGGGCTATGACGTCACCGGCAATACGCCG
TTTGCCTATCCGGGCCTCGTTTTTGGTCACAACGGCACCATTTCATGGGGATCCACCGCCGGTTTTGGTGATGATGTC
GATATCTTTGCCGAAAAACTTTCCGCCGAGAAGCCGGGCTATTACCAGCATAACGGCGAGTGGGTGAAGATGTTGAGC
CGCAAGGAGACTATTGCGGTCAAAGACGGCCAGCCGGAGACCTTTACCGTTTGGCGCACGCTGCACGGCAACGTCATT
AAAACCGATACTGCGACGCAGACCGCCTATGCCAAAGCGCGCGCCTGGGATGGCAAAGAGGTGGCGTCCCTGCTGGCG
TGGACGCACCAGATGAAGGCCAAAAACTGGCCGGAGTGGACGCAGCAGGCGGCCAAACAGGCGCTGACCATTAACTGG
TACTACGCCGATGTGAACGGCAATATCGGCTATGTGCATACCGGCGCCTATCCGGATCGCCAGCCCGGCCACGACCCG
CGTTTGCCGGTTCCCGGCACTGGAAAATGGGACTGGAAAGGGTTGCTGTCGTTTGATTTGAATCCGAAAGTGTATAAC
CCGCAGTCGGGCTATATCGCCAACTGGAACAACTCGCCGCAAAAAGACTACCCGGCCTCTGATCTGTTCGCGTTCCTG
TGGGGCGGTGCGGATCGAGTTACTGAGATCGACACGATCCTCGATAAGCAACCGCGCTTCACCGCCGATCAGGCGTGG
GATGTGATCCGCCAAACCAGCCGTCGGGATCTCAACCTGCGGTTGTTCTTACCGGCGCTGAAGGACGCCACCGCGAAC
CTGGCGGAAAACGATCCGCGCCGCCAACTGGTGGATAAACTGGCGAGCTGGGACGGTGAAAACCTTGTCAACGATGAC
GGAAAAACCTATCAGCAACCGGGATCGGCGATTCTTAACGCCTGGCTGACCAGCATGCTCAAGCGCACGGTGGTTGCC
GCGGTCCCAGCGCCGTTTGGCAAGTGGTACAGCGCCAGTGGCTATGAAACCACCCAGGACGGGCCAACCGGCTCGCTG
AACATCAGCGTGGGGGCGAAAATCCTCTACGAAGCTCTGCAGGGTGATAAGTCGCCAATCCCGCAGGCGGTCGATCTG
TTTGGCGGGAAACCGCAGCAGGAAGTGATACTGGCGGCGCTGGACGACGCTTGGCAGACGCTGTCAAAACGCTACGGT
AACGACGTCACCGGCTGGAAAACCCCTGCCATGGCGCTTACCTTCCGGGCCAATAACTTCTTCGGCGTGCCGCAGGCG
GCAGCAAAAGAGGCGCGTCATCAGGCGGAGTACCAGAACCGCGGTACGGAAAACGACATGATTGTCTTCTCACCGACG
TCGGGTAACCGCCCGGTTCTTGCCTGGGATGTGGTGGCGCCGGGGCAAAGCGGTTTTATCGCGCCGGATGGCAAAGCC
GATAAGCACTATGACGATCAGCTGAAAATGTACGAGAGCTTTGGCCGTAAATCGCTGTGGTTAACGCCTCAGGACGTT
GACGAGCACAAAGAGTCTCAGGAAGTGCTGCAGGTACAGCGCCTCGAGCACCACCACCACCACCACTGA
```

Corresponding translated amino acid sequence

```
MKNRNRMIVNGIVTSLICCSSLSALAASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYVVAQDRLFQME
MARRSTQGTVSEVLGKAFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQGYADGMNAWIDKVNASPDKL
LPQQFSTFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEIDNLALLTALKDKYGKQQGMAVFNQLKWLVNP
SAPTTIAARESAYPLKFDLQNTQTAALLPRYDQPAPMLDRPAKGTDGALLALTADQNRETIAAQFAQSGA
NGLAGYPTTSNMWVIGKNKAQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTPFAYPGLVFGHNG
TISWGSTAGFGDDVDIFAEKLSAEKPGYYQHNGEWVKMLSRKETIAVKDGQPETFTVWRTLHGNVIKTDT
ATQTAYAKARAWDGKEVASLLAWTHQMKAKNWPEWTQQAAKQALTINWYYADVNGNIGYVHTGAYPDRQP
GHDPRLPVPGTGKWDWKGLLSFDLNPKVYNPQSGYIANWNNSPQKDYPASDLFAFLWGGADRVTEIDTIL
DKQPRFTADQAWDVIRQTSRRDLNLRLFLPALKDATANLAENDPRRQLVDKLASWDGENLVNDDGKTYQQ
PGSAILNAWLTSMLKRTVVAAVPAPFGKWYSASGYETTQDGPTGSLNISVGAKILYEALQGDKSPIPQAV
DLFGGKPQQEVILAALDDAWQTLSKRYGNDVTGWKTPAMALTFRANNFFGVPQAAAKEARHQAEYQNRGT
ENDMIVFSPTSGNRPVLAWDVVAPGQSGFIAPDGKADKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQE
VLQVQRLEHHHHHH-
```

Figure 3.4: Nucleotide and amino acid sequences of *Kcpac* gene (shown in bold) cloned in pET 26b (+)

BLAST analysis showed the "*pac*" gene from *K. citrophila* to have some changes both at gene as well as amino acid level from the sequence of "*pac*" gene reported by Barbero *et al.,* 1986 (**Figure 3.5 A, B**). It may be noted that some changes in the gene sequence were also reported by Rao *et al.* (1994). *Kcpac* from DSMZ 2660 cloned with pET 26b (+) showed an identity of 98% both at gene as well as amino acid level with reported sequences. Differences are highlighted with red colour in the alignment. However, the active site residues identified as directly involved in catalysis (Ser290, Glu312, Ala358 and Asn530, number corresponds to precursor) were found to be strictly conserved despite of the presence of changes in amino acid sequence and are highlighted in green (**Figure 3.5B**).

```
>gb|M15418.1|KCIPAC  K.citrophila pac gene encoding penicillin acylase, complete
cds, Length=2734
 Score = 4376 bits (4852), Expect = 0.0
 Identities = 2499/2542 (98%), Gaps = 14/2542 (1%) Strand=Plus/Plus

Query  1    ATGAAAAATAGAAATCGTATGATCGTGAACGGTATTGTGACTTCCCTGATCTGTTGTTCT  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  86   ATGAAAAATAGAAATCGTATGATCGTGAACGGTATTGTGACTTCCCTGATCTGTTGTTCT  145

Query  61   AGCCTGTCAGCGCTGGCGGCAAGCCCGCCAACCGAGGTTAAGATCGTTCGCGATGAATAC  120
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  146  AGCCTGTCAGCGCTCGCGGCAAGCCCGCCAACCGAGGTTAAGATCGTTCGCGATGAATAC  205

Query  121  GGCATGCCGCATATTTACGCCGATGATACCTATCGACTGTTTTACGGCTATGGCTACGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  206  GGCATGCCGCATATTTACGCCGATGATACCTATCGACTGTTTTACGGCTATGGCTACGTG  265

Query  181  GTGGCGCAGGATCGCCTGTTCCAGATGGAAATGGCGCGCCGCAGTACTCAGGGGACCGTC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  266  GTGGCGCAGGATCGCCTGTTCCAGATGGAAATGGCGCGCCGCAGTACTCAGGGGACCGTC  325

Query  241  TCCGAGGTGCTGGGCAAAGCATTCGTCAGTTTTGATAAAGATATTCGCCAGAACTACTGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  326  TCCGAGGTGCTGGGCAAAGCATTCGTCAGTTTTGATAAAGATATTCGCCAGAACTACTGG  385

Query  301  CCGGATTCTATTCGCGCGCAGATAGCTTCCCTCTCCGCTGAGGATAAATCCATTCTGCAG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  386  CCGGATTCTATTCGCGCGCAGATAGCTTCCCTCTCCGCTGAGGATAAATCCATTCTGCAG  445

Query  361  GGCTATGCCGATGGCATGAATGCGTGGATCGATAAAGTGAACGCCAGCCCCGATAAGCTG  420
            ||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct  446  GGCTATGCCGATGGAATGAATGCGTGGATCGATAAAGTGAACGCCAGCCCGGATAAGCTG  505

Query  421  TTACCCCAGCAGTTCTCCACCTTTGGTTTTAAACCCAAGCATTGGGAACCGTTTGATGTG  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  506  TTACCCCAGCAGTTCTCCACCTTTGGTTTTAAACCCAAGCATTGGGAACCGTTTGATGTG  565

Query  481  GCGATGATTTTTGTCGGCACCATGGCGAACCGTTTCTCTGACAGCACCAGCGAAATTGAT  540
            |||||||||||||||||||||||||||||||| || |||||||||||||||||||||||||
Sbjct  566  GCGATGATTTTTGTCGGCACCATGGCGAACCGGTTTTCTGACAGCACCAGCGAAATTGAT  625

Query  541  AACCTGGCGCTGCTGACGGCGCTAAAAGATAAATACGGCAAGCAGCAGGGCATGGCGGTC  600
            ||||||||||||||||||| |||||||||||||||||| |  | ||||||||||||||||
Sbjct  626  AACCTGGCGCTGCTGACGGCCGGTAAAAGATAAATACGGCAACGACGAGGGCATGGCGGTC  685

Query  601  TTTAACCAGCTGAAATGGCTGGTTAATCCTTCCGCGCCAACCACCATTGCGGCGCGGGAA  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  686  TTTAACCAGCTGAAATGGCTGGTTAATCCTTCCGCGCCAACCACCATTGCGGCGCGGGAA  745
```

```
Query  661   AGCGCCTATCCGCTGAAGTTTGATCTGCAAAACACGCAAACGGCGGCGCTGCTG---CCG   717
             ||| | ||||||||||||||||||||||||||||||||||||||||||||||||   |||
Sbjct  746   AGCTCGTATCCGCTGAAGTTTGATCTGCAAAACACGCAAACGGCGGCGCTGCTGGTCCCG   805

Query  718   CGCTACGACCAGCCGGCACCGATGCTCGACCGCCCGGCAAAAGGGACCGATGGCGCGCTG   777
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  806   CGCTACGACCAGCCGGCACCGATGCTCGACCGCCCGGCAAAAGGGACCGATGGCGCGCTG   865

Query  778   CTGGCGCTGACCGCCGATCA-GAACCGGGAAACTATCGCCGCGCAGTTCGCGCAAAGCGG   836
             ||||| | ||||| ||||| |||||||||||||||||||||||||||||   || |||
Sbjct  866   CTGGCCGTCACCGC-GATCAAGAACCGGGAAACTATCGCCGCGCAGTTCGC---AAACGG   921

Query  837   CGCTAACGGCCTGGCTGGCTACCCGACCACTAGCAATATGTGGGTGATTGGCAAAAACAA   896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  922   CGCTAACGGCCTGGCTGGCTACCCGACCACTAGCAATATGTGGGTGATTGGCAAAAACAA   981

Query  897   AGCCCAGGATGCGAAGGCCATTATGGTCAATGGGCCGCAGTTTGGTTGGTATGCGCCGGC   956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  982   AGCCCAGGATGCGAAGGCCATTATGGTCAATGGGCCGCAGTTTGGTTGGTATGCGCCGGC   1041

Query  957   GTACACCTACGGTATCGGCCTGCACGGCGCGGGCTATGACGTCACCGGCAATACGCCGTT   1016
             ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1042  GTACACTTACGGTATCGGCCTGCACGGCGCGGGCTATGACGTCACCGGCAATACGCCGTT   1101

Query  1017  TGCCTATCCGGGCCTCGTTTTTGGTCACAACGGCACCATTTCATGGGGATCCACCGCCGG   1076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1102  TGCCTATCCGGGCCTCGTTTTTGGTCACAACGGCACCATTTCATGGGGATCCACCGCCGG   1161

Query  1077  TTTTGGTGATGATGTCGATATCTTTGCCGAAAAACTTTCCGCCGAGAAGCCGGGCTATTA   1136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1162  TTTTGGTGATGATGTCGATATCTTTGCCGAAAAACTTTCCGCCGAGAAGCCGGGCTATTA   1221

Query  1137  CCAGCATAACGGCGAGTGGGTGAAGATGTTGAGCCGCAAGGAGACTATTGCGGTCAAAGA   1196
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1222  CCAGCATAACGGCGAGTGGGTGAAGATGTTGAGCCGCAAGGAGACTATTGCGGTCAAAGA   1281

Query  1197  CGGCCAGCCGGAGACCTTTACCGTTTGGCGCACGCTGCACGGCAACGTCATTAAAACCGA   1256
             |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Sbjct  1282  CGGCCAGCCGGAGACCTTTACCGTTTGGCGCACGCTCGACGGCAACGTCATTAAAACCGA   1341

Query  1257  TACTGCGACGCAGACCGCCTATGCCAAAGCGCGCGCCTGGGATGGCAAAGAGGTGGCGTC   1316
             |||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Sbjct  1342  TACTCGGACGCAGACCGCCTATGCCAAAGCGCGCGCCTGGGCTGGCAAAGAGGTGGCGTC   1401

Query  1317  CCTGCTGGCGTGGACGCACCAGATGAAGGCCAAAAACTGGCCGGAGTGGACGCAGCAGGC   1376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1402  CCTGCTGGCGTGGACGCACCAGATGAAGGCCAAAAACTGGCCGGAGTGGACGCAGCAGGC   1461

Query  1377  GGCCAAACAGGCGCTGACCATTAACTGGTACTACGCCGATGTGAACGGCAATATCGGCTA   1436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1462  GGCCAAACAGGCGCTGACCATTAACTGGTACTACGCCGATGTGAACGGCAATATCGGCTA   1521

Query  1437  TGTGCATACCGGCGCCTATCCGGATCGCCAGCCCGGCCACGACCCGCGTTTGCCGGTTCC   1496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1522  TGTGCATACCGGCGCCTATCCGGATCGCCAGCCCGGCCACGACCCGCGTTTGCCGGTTCC   1581
Query  1497  CGGCACTGGAAAATGGGACTGGAAAGGGTTGCTGTCGTTTGATTTGAATCCGAAAGTGTA   1556
             ||     |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1582  CGA---TGGAAAATGGGACTGGAAAGGGTTGCTGTCGTTTGATTTGAATCCGAAAGTGTA   1638

Query  1557  TAACCCGCAGTCGGGCTATATCGCCAACTGGAACAACTCGCCGCAAAAAGACTACCCGGC   1616
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1639  TAACCCGCAGTCGGGCTATATCGCCAACTGGAACAACTCGCCGCAAAAAGACTACCCGGC   1698

Query  1617  CTCTGATCTGTTCGCGTTCCTGTGGGGCGGTGCGGATCGAGTTACTGAGATCGACACGAT   1676
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1699  CTCTGATCTGTTCGCGTTCCTGTGGGGCGGTGCGGATCGAGTTACTGAGATCGACACGAT   1758

Query  1677  CCTCGATAAGCAACCGCGCTTCACCGCCGATCAGGCGTGGGATGTGATCCGCCAAACCAG   1736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1759  CCTCGATAAGCAACCGCGCTTCACCGCCGATCAGGCGTGGGATGTGATCCGCCAAACCAG   1818
```

```
Query  1737  CCGTCGGGATCTCAACCTGCGGTTGTTCTTACCGGCGCTGAAGGACGCCACCGCGAACCT  1796
             ||  ||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Sbjct  1819  CCTCCGGGATCTC---CTGCGGTTGTTCTTACCGGCGCTGAAGGACGCCACCGCGAACCT  1875

Query  1797  GGCGGAAAACGATCCGCGCCGCCAACTGGTGGATAAACTGGCGAGCTGGGACGGTGAAAA  1856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1876  GGCGGAAAACGATCCGCGCCGCCAACTGGTGGATAAACTGGCGAGCTGGGACGGTGAAAA  1935

Query  1857  CCTTGTCAACGATGACGGAAAAACCTATCAGCAACCGGGATCGGCGATTCTTAACGCCTG  1916
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1936  CCTTGTCAACGATGACGGAAAAACCTATCAGCAACCGGGATCGGCGATTCTTAACGCCTG  1995

Query  1917  GCTGACCAGCATGCTCAAGCGCACGGTGGTTGCCGCGGTCCCAGCGCCGTTTGGCAAGTG  1976
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1996  GCTGACCAGCATGCTCAAGCGCACGGTGGTTGCCGCGGTCCCAGCGCCGTTTGGCAAGTG  2055

Query  1977  GTACAGCGCCAGTGGCTATGAAACCACCCAGGACGGGCCAACCGGCTCGCTGAACATCAG  2036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2056  GTACAGCGCCAGTGGCTATGAAACCACCCAGGACGGGCCAACCGGCTCGCTGAACATCAG  2115

Query  2037  CGTGGGGGCGAAAATCCTCTACGAAGCTCTGCAGGGTGATAAGTCGCCAATCCCGCAGGC  2096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2116  CGTGGGGGCGAAAATCCTCTACGAAGCTCTGCAGGGTGATAAGTCGCCAATCCCGCAGGC  2175

Query  2097  GGTCGATCTGTTTGGCGGGAAACCGCAGCAGGAAGTGATACTGGCGGCGCTGGACGACGC  2156
             |||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Sbjct  2176  GGTCGATCTGTTTGGCGGGAAACCCGAGCAGGAAGTGATACTGGCGGCGCTGGACGACGC  2235

Query  2157  TTGGCAGACGCTGTCAAAACGCTACGGTAACGACGTCACCGGCTGGAAAACCCCTGCCAT  2216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2236  TTGGCAGACGCTGTCAAAACGCTACGGTAACGACGTCACCGGCTGGAAAACCCCTGCCAT  2295

Query  2217  GGCGCTTACCTTCCGGGCCAATAACTTCTTCGGCGTGCCGCAGGCGGCAGCAAAAGAGGC  2276
             |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Sbjct  2296  GGCGCTTACCTTCCGGGCCAATAACTTCTTCGGCGTGCCGCAGGCCGCAGCAAAAGAGGC  2355

Query  2277  GCGTCATCAGGCGGAGTACCAGAACCGCGGTACGGAAAACGACATGATTGTCTTCTCACC  2336
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2356  GCGTCATCAGGCGGAGTACCAGAACCGCGGTACGGAAAACGACATGATTGTCTTCTCACC  2415

Query  2337  GACGTCGGGTAACCGCCCGGTTCTTGCCTGGGATGTGGTGGCGCCGGGGCAAAGCGGTTT  2396
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2416  GACGTCGGGTAACCGCCCGGTTCTTGCCTGGGATGTGGTGGCGCCGGGGCAAAGCGGTTT  2475

Query  2397  TATCGCGCCGGATGGCAAAGCCGATAAGCACTATGACGATCAGCTGAAAATGTACGAGAG  2456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2476  TATCGCGCCGGATGGCAAAGCCGATAAGCACTATGACGATCAGCTGAAAATGTACGAGAG  2535

Query  2457  CTTTGGCCGTAAATCGCTGTGGTTAACGCCTCAGGACGTTGACGAGCACAAAGAGTCTCA  2516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2536  CTTTGGCCGTAAATCGCTGTGGTTAACGCCTCAGGACGTTGACGAGCACAAAGAGTCTCA  2595

Query  2517  GGAAGTGCTGCAGGTACAGCGC  2538
             ||||||||||||||||||||||
Sbjct  2596  GGAAGTGCTGCAGGTACAGCGC  2617
```

```
┌─────┐
│  A  │
└─────┘
```

>sp|P07941.1|*PAC_KLUCI*   RecName: Full=Penicillin   G   acylase;   AltName:
Full=Penicillin G amidase; AltName: Full=Penicillin amidohydrolase; Contains:
RecName: Full=Penicillin G acylase subunit alpha; Contains:
RecName: Full=Penicillin G acylase subunit beta; Flags: Precursor
 gb|AAA25047.1| penicillin   acylase   prepropeptide   (EC   3.5.1.11)   [Kluyvera
cryocrescens] Length=844

 Score = 1691 bits (4379), Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 828/847 (98%), Positives = 835/847 (99%), Gaps = 4/847 (0%)

```
Query  1    MKNRNRMIVNGIVTSLICCSSLSALAASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYV  60
            MKNRNRMIVNGIVTSLICCSSLSALAASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYV
Sbjct  1    MKNRNRMIVNGIVTSLICCSSLSALAASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYV  60
```

```
Query   61    VAQDRLFQMEMARRSTQGTVSEVLGKAFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQ    120
              VAQDRLFQMEMARRSTQGTVSEVLGKAFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQ
Sbjct   61    VAQDRLFQMEMARRSTQGTVSEVLGKAFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQ    120
Query   121   GYADGMNAWIDKVNASPDKLLPQQFSTFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEID   180
              GYADGMNAWIDKVNASPDKLLPQQFSTFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEID
Sbjct   121   GYADGMNAWIDKVNASPDKLLPQQFSTFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEID   180

Query   181   NLALLTALKDKYGKQQGMAVFNQLKWLVNPSAPTTIAARESAYPLKFDLQNTQTAALL-P   239
              NLALLTA+KDKYG  +GMAVFNQLKWLVNPSAPTTIAARES+YPLKFDLQNTQTAALL P
Sbjct   181   NLALLTAVKDKYGNDEGMAVFNQLKWLVNPSAPTTIAARESSYPLKFDLQNTQTAALLVP   240

Query   240   RYDQPAPMLDRPAKGTDGALLALTADQNRETIAAQFAQSGANGLAGYPTTSNMWVIGKNK   299
              RYDQPAPMLDRPAKGTDGALLA+TA +NRETIAAQFA +GANGLAGYPTTSNMWVIGKNK
Sbjct   241   RYDQPAPMLDRPAKGTDGALLAVTAIKNRETIAAQFA-NGANGLAGYPTTSNMWVIGKNK   299

Query   300   AQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTPFAYPGLVFGHNGTISWGSTAG   359
              AQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTPFAYPGLVFGHNGTISWGSTAG
Sbjct   300   AQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTPFAYPGLVFGHNGTISWGSTAG   359

Query   360   FGDDVDIFAEKLSAEKPGYYQHNGEWVKMLSRKETIAVKDGQPETFTVWRTLHGNVIKTD   419
              FGDDVDIFAEKLSAEKPGYYQHNGEWVKMLSRKETIAVKDGQPETFTVWRTL GNVIKTD
Sbjct   360   FGDDVDIFAEKLSAEKPGYYQHNGEWVKMLSRKETIAVKDGQPETFTVWRTLDGNVIKTD   419

Query   420   TATQTAYAKARAWDGKEVASLLAWTHQMKAKNWPEWTQQAAKQALTINWYYADVNGNIGY   479
              T TQTAYAKARAW GKEVASLLAWTHQMKAKNWPEWTQQAAKQALTINWYYADVNGNIGY
Sbjct   420   TRTQTAYAKARAWAGKEVASLLAWTHQMKAKNWPEWTQQAAKQALTINWYYADVNGNIGY   479

Query   480   VHTGAYPDRQPGHDPRLPVPGTGKWDWKGLLSFDLNPKVYNPQSGYIANWNNSPQKDYPA   539
              VHTGAYPDRQPGHDPRLPVP  GKWDWKGLLSFDLNPKVYNPQSGYIANWNNSPQKDYPA
Sbjct   480   VHTGAYPDRQPGHDPRLPVP-DGKWDWKGLLSFDLNPKVYNPQSGYIANWNNSPQKDYPA   538

Query   540   SDLFAFLWGGADRVTEIDTILDKQPRFTADQAWDVIRQTSRRDLNLRLFLPALKDATANL   599
              SDLFAFLWGGADRVTEIDTILDKQPRFTADQAWDVIRQTS RDL LRLFLPALKDATANL
Sbjct   539   SDLFAFLWGGADRVTEIDTILDKQPRFTADQAWDVIRQTSLRDL-LRLFLPALKDATANL   597

Query   600   AENDPRRQLVDKLASWDGENLVNDDGKTYQQPGSAILNAWLTSMLKRTVVAAVPAPFGKW   659
              AENDPRRQLVDKLASWDGENLVNDDGKTYQQPGSAILNAWLTSMLKRTVVAAVPAPFGKW
Sbjct   598   AENDPRRQLVDKLASWDGENLVNDDGKTYQQPGSAILNAWLTSMLKRTVVAAVPAPFGKW   657

Query   660   YSASGYETTQDGPTGSLNISVGAKILYEALQGDKSPIPQAVDLFGGKPQQEVILAALDDA   719
              YSASGYETTQDGPTGSLNISVGAKILYEALQGDKSPIPQAVDLFGGKP+QEVILAALDDA
Sbjct   658   YSASGYETTQDGPTGSLNISVGAKILYEALQGDKSPIPQAVDLFGGKPEQEVILAALDDA   717

Query   720   WQTLSKRYGNDVTGWKTPAMALTFRANNFFGVPQAAAKEARHQAEYQNRGTENDMIVFSP   779
              WQTLSKRYGNDVTGWKTPAMALTFRANNFFGVPQAAAKEARHQAEYQNRGTENDMIVFSP
Sbjct   718   WQTLSKRYGNDVTGWKTPAMALTFRANNFFGVPQAAAKEARHQAEYQNRGTENDMIVFSP   777

Query   780   TSGNRPVLAWDVVAPGQSGFIAPDGKADKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQ   839
              TSGNRPVLAWDVVAPGQSGFIAPDGKADKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQ
Sbjct   778   TSGNRPVLAWDVVAPGQSGFIAPDGKADKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQ   837

Query   840   EVLQVQR    846
              EVLQVQR
Sbjct   838   EVLQVQR    844
```

B

Figure 3.5: Nucleotide BLAST (A) and protein BLAST (B) results of *Kcpac* DSMZ 2660 showing difference from the reported sequences. Differences are highlighted in red. Active site residues are highlighted in Green.

Conserved domain database search on NCBI server showed non-specific hits with β-lactamase family that includes penicillin G acylases and Cephalosporin acylases

and specific hit to super family of Ntn hydrolase family based on the presence of the conserved domains (**Figure 3.6**).
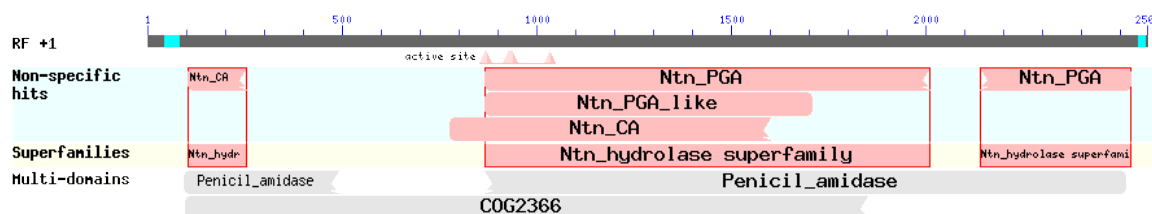
Figure 3.6: CDD search on NCBI database.

3.2.2. Amino Acid Sequence analysis

As mentioned above, the PGA precursor molecule undergoes autocatalytic processing event with the removal of signal and s*pac*er peptide resulting in the generation of active processed PGA which consists of a smaller α-chain and a larger β-chain of 209 and 557 amino acids, respectively intertwined with each other with non covalent forces. The calculated molecular mass using Expasy compute pI/Mw tool for α-chain and β-chain are 23.6 kDa and 62.9 kDa, respectively.

3.2.3. Preparing site-directed mutants

Choi *et al*. (1992) studied the effect of different mutations on the processing of the precursor PGA protein. Ser-290 is found critical for processing and also for the enzymatic activity of penicillin G acylase. In the mutant, in which Ser-290 is mutated to Cys, the precursor gets processed, but there is no detectable enzymatic activity, whereas mutations of this residue to Arg, Thr, Gly, or Ala have always yielded the unprocessed PGA precursor.

Site-directed mutants of *Kc*PGA DMSZ 2660 have been prepared to study the mechanism of autoproteolysis and catalysis. The wild type "*pac*" gene from *K.citrophila* inserted into the plasmid pET26b (+) was used as template for the preparation of the mutants Ser290Cys and Ser290Gly. The QuickChange site-directed mutagenesis kit (Stratagene) was used for preparing mutants following manufacturer's recommended protocol.

Primer pairs used to construct the mutant Ser290Cys were:

Forward primer 5'-CTACCCGACCACT**T**GCAATATGTGGGTG-3' and

Reverse primer 5'-CACCCACATATTGC**A**AGTGGTCGGGTAG-3'

Primer pairs used for Ser290Gly were:

Forward primer 5'- CTACCCGACCACT**G**GCAATATGTGGGTG-3' and

Reverse primer 5'-CACCCACATATTGC**C**AGTGGTCGGGTAG-3'

The mutation for codon change is shown underlined.

The constructs was cloned in *E. coli* NovaBlue cells. Presence of desired mutations was confirmed by isolating plasmid from colonies chosen randomly followed by full length sequencing.

3.2.4. Expression of *Kc*PGA Proteins

Expression of plasmids with confirmed sequence for wild type enzyme was checked in three different expression hosts. BL21 (DE3), BL21 (DE3) pLysS and Lemo21 (DE3) were transformed with plasmid and the expression from the 1 mM IPTG induced cultures expressed at 16 °C for 16-18 h using SDS-PAGE was checked. Expression of KcPGA was checked in periplasmic fraction, the cytoplasmic soluble fraction as well as in total cell fraction (**Figure 3.7**). Expression in BL21 (DE3) pLysS cells produced processed protein better than the other two hosts. Expression in BL21 (DE3) and BL21 pLEMO resulted in the production of unprocessed precursor PGA along with the expression of processed PGA protein. BL21 (DE3) pLysS contains a plasmid, pLysS, which carries the gene encoding T7 Lysozyme. T7 Lysozyme, a natural inhibitor of T7 RNA polymerase, lowers the background expression level of target genes thus increases the production of target gene protein in soluble fraction. BL21 (DE3) pLysS is also deficient in Proteases *lon* and *OmpT* that increases the stability of expressed protein. Thus BL21 (DE3) pLysS was chosen to be the expression host for expressing the *KcPAC* gene. Site directed mutants were transformed in BL21 (DE3) pLysS and expression was checked in both soluble as well as total cell fractions (soluble and insoluble) (**Figure 3.8**).
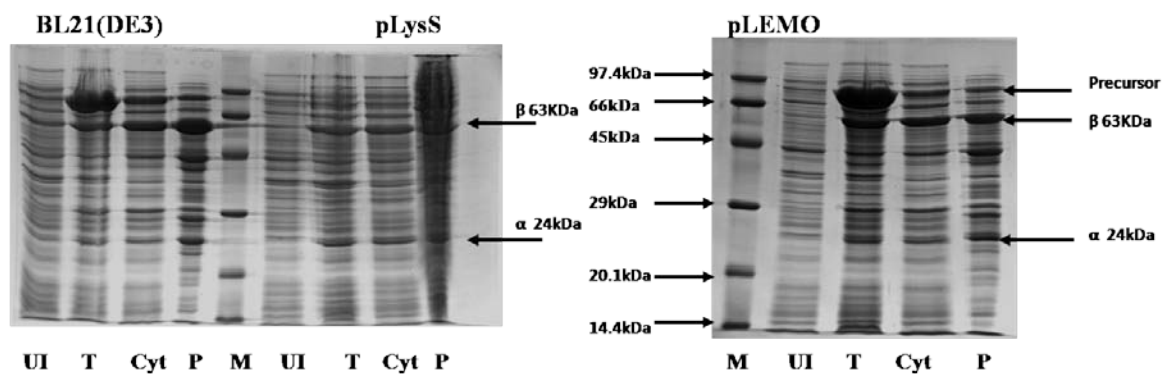
Figure 3.7: Expression of *KcPAC* cloned in pET 26b (+) in three different expression host, UI- Uninduced culture, T- Total cell fraction, Cyt- Clarified Cytoplasmic fraction, P- Periplasmic fraction, M- Low Molecular range protein marker, Proteins were separated on 12% SDS-PAGE and stained with Coomassie Blue R-250.
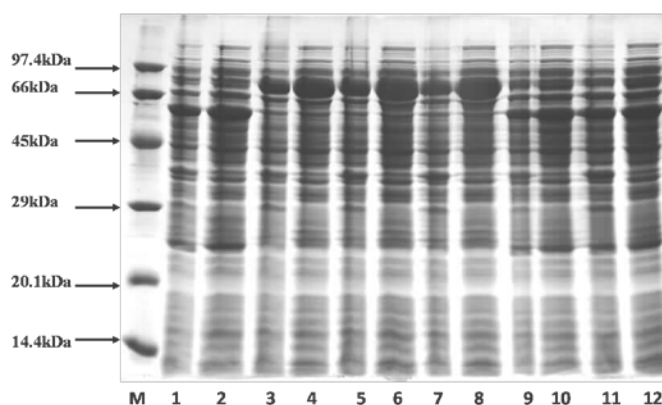
Figure 3.8: Expression of *Kc*PGA mutants in pLysS cells on 12% SDS PAGE stained with Coomassie Blue R-250. M- Low Range Molecular weight marker.

3.2.5. Purification of *Kc*PGA Proteins

The steps followed for purifying wild type and mutant forms of *Kc*PGA were described under materials methods in chapter 2. Briefly, the clarified sonicated fraction was loaded onto Nickel Sepharose column pre-equilibrated with binding buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 1 mM DTT) followed by washing and elution with washing buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM imidazole, 1 mM DTT) and elution buffer (25 mM Tris-HCl pH 8.0, 500 mM NaCl, 150 mM imidazole, 1 mM DTT), respectively. Fractions containing proteins, confirmed by activity assay for wild type and running 12% SDS-PAGE for mutant proteins, were

pooled and concentrated followed by final purification and polishing step of size-exclusion chromatography using Sephacryl S-200 using 10 mM potassium phosphate pH 7.5 and 150 mM NaCl. The elution profiles for the purification protocol are shown in **Figure 3.9 & 3.10**.

About 20 mg of the purified wild-type protein was obtained from 1 litre of the culture. The mutant proteins Ser290Cys and Ser290Gly were purified using the same procedure and the yield of pure protein was ~12 and ~15 mg/L, respectively. All the proteins were found to be highly soluble and could be concentrated to more than 50 mg/ml without any visible precipitation. The enzymes were electrophoretically homogenous as seen in native-PAGE and SDS-PAGE gels (**Figure 3.11 A, B and C**). Wild-type *Kc*PGA showed two bands of ~23.6 kDa and ~63 kDa corresponding to α- and β-chains, respectively. The mutant Ser290Cys also showed bands of similar size whereas Ser290Gly showed major band of precursor PGA protein with molecular weight ~93 kDa.
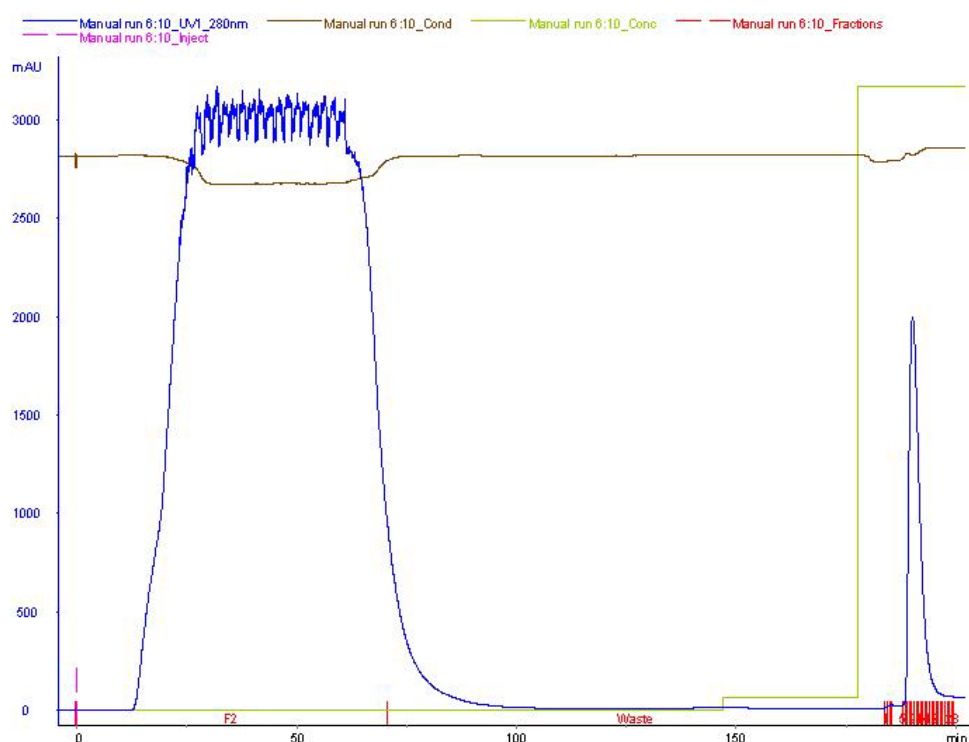
Figure 3.9: FPLC elution profile of Nickel affinity chromatography for *Kc*PGA wild-type. OD$_{280}$ is drawn in blue colour.
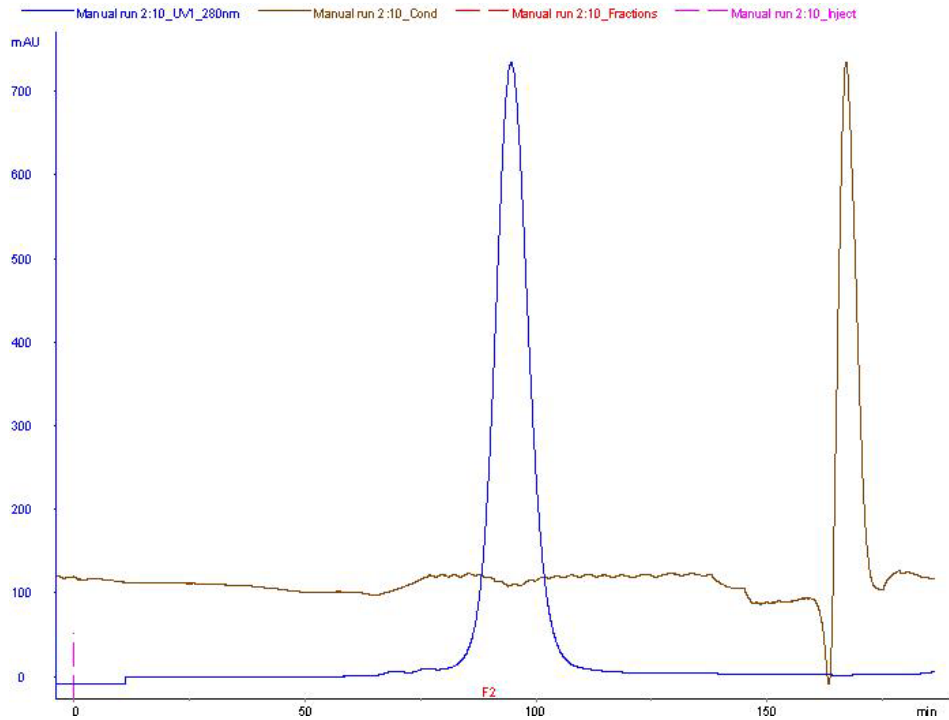
Figure 3.10: FPLC elution profile of Gel Filteration chromatography (Sephacryl S200) for *Kc*PGA wild-type protein. OD$_{280}$ is shown in blue colour.

Purification Protocol of KcPGA

Nickel Affinity Chromatography

↓

Sephacryl S-200

↓

Purified Protein
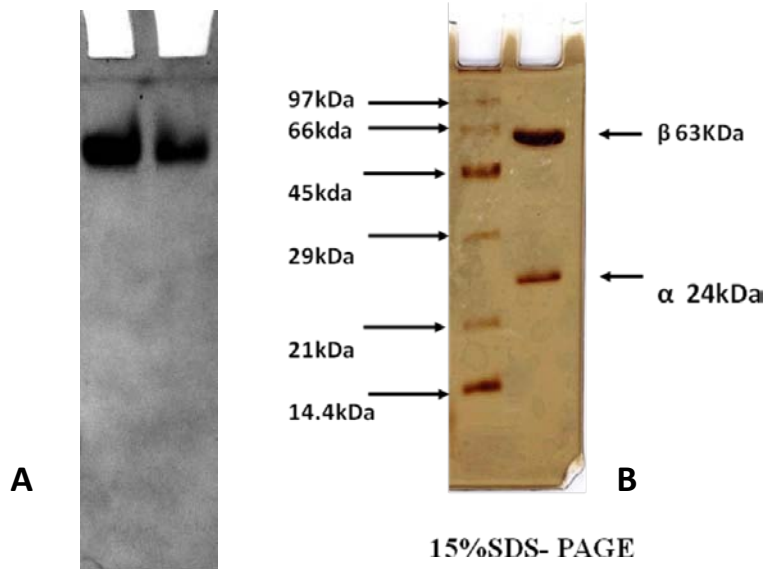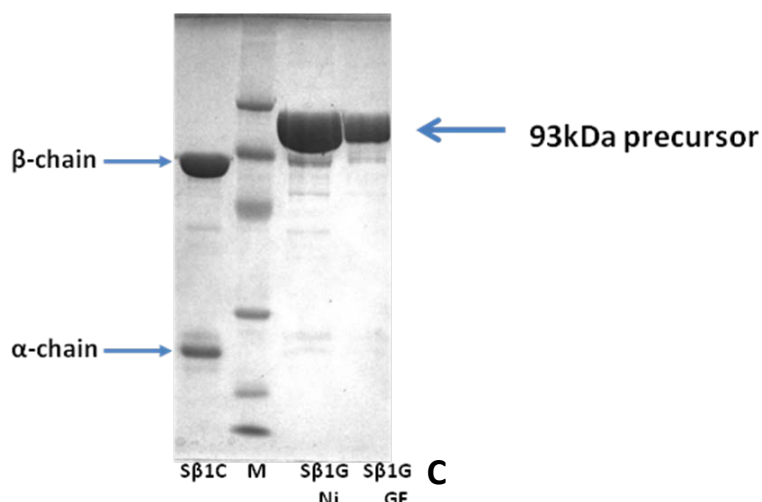
Figure 3.11: Results of PAGE runs showing purity and homogeneity of *Kc*PGA wild-type as well as mutant proteins, (A) 7.5% Native PAGE showing the purity of wild-type processed PGA, (B) Run of 15% SDS- PAGE showing the purity of wild-type processed *Kc*PGA stained with more sensitive silver stain, (C) Run of 12% SDS- PAGE showing the purity of Ser290Cys (Sβ1C) and precursor Ser290Gly (Sβ1G), Sβ1G Ni denotes the nickel purified precursor, Sβ1G GF denotes protein eluted from Gel filtration column, M : Low molecular Range protein Marker.

3.2.6. Preliminary Biochemical Characterization of *Kc*PGA Wild-type Protein

3.2.6.1. Molecular Weight Determination

Wild-type *Kc*PGA showed two bands of ~23.6 kDa and ~63 kDa corresponding to α- and β-chains, respectively on the SDS-PAGE (**Figure 3.11B**). The molecular weight of the wild-type enzyme was confirmed with the help of MALDI-TOF and electrospray ionization (**Figure 3.12 A, B**), which showed major peaks of α-chain and β-chain at 23,588 Da and 62,901 Da in MALDI-TOF and 23,607 Da and 62,920 Da in electrospray ionization spectrum.
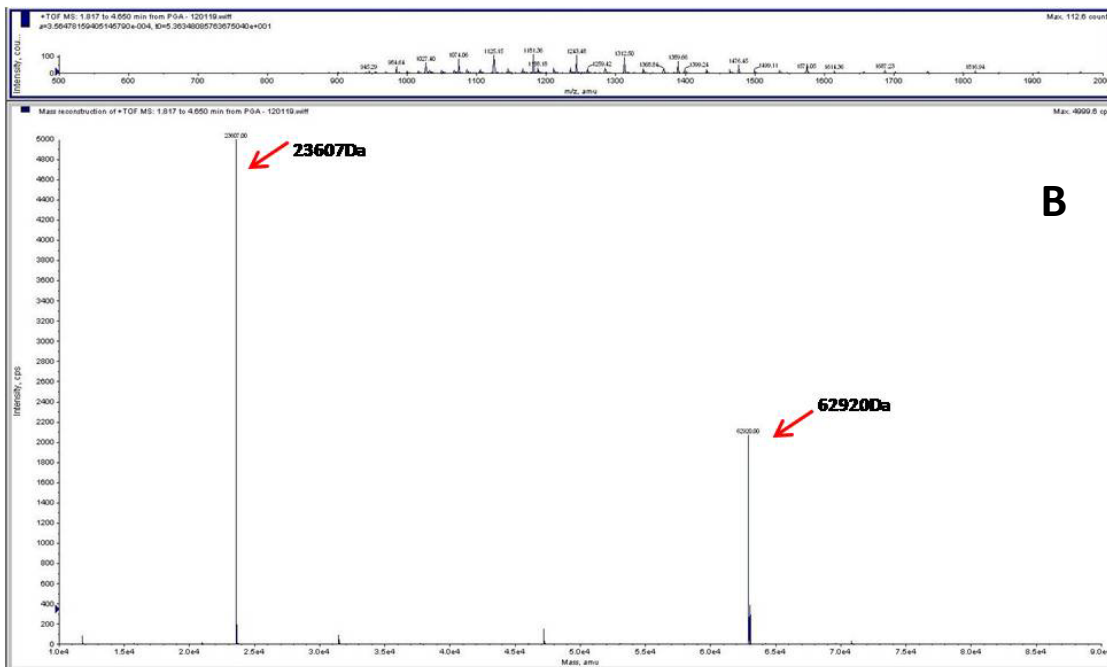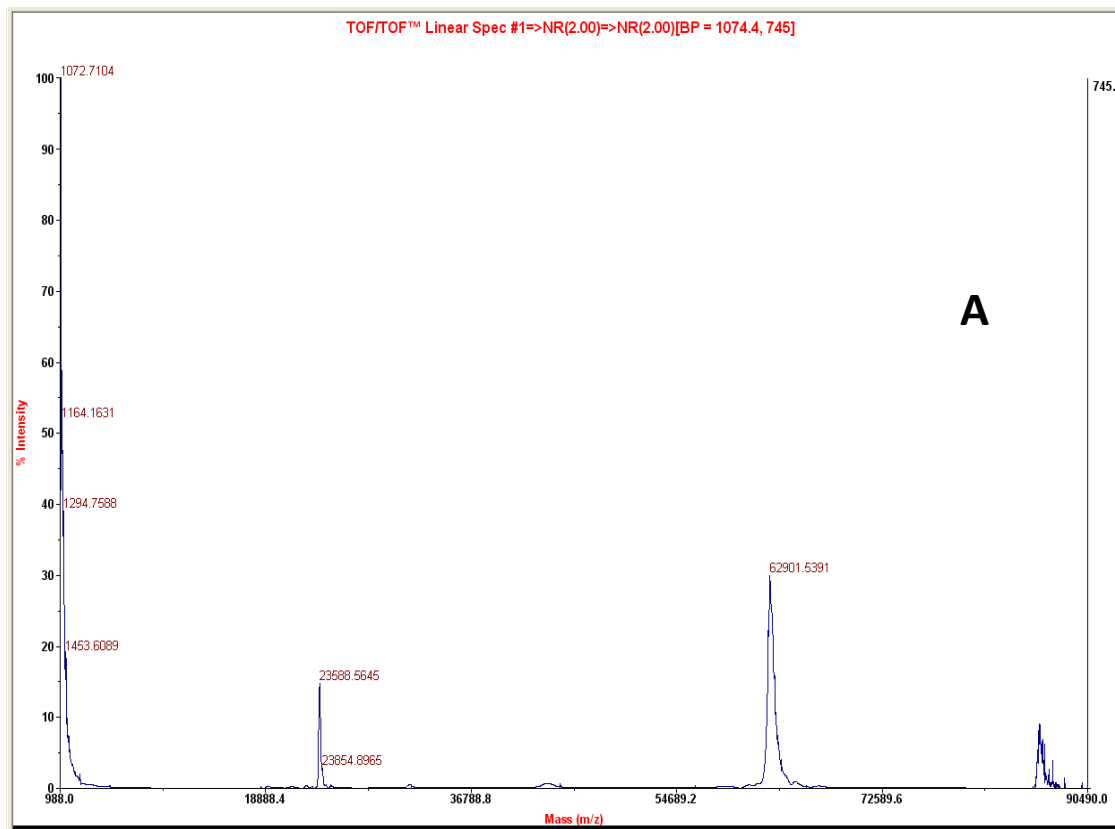
Figure 3.12: Matrix-assisted laser desorption ionization / time-of-flight time-of-flight mass spectrometry (MALDI-TOFTOF) (A) and Electrospray Ionization spectrum (B) of purified wild-type *Kc*PGA.

3.2.6.2. Dynamic Light Scattering

Dynamic Light Scattering experiment using 5 mg/ml as well as 25 mg/ml solutions of purified PGA showed protein to be highly monodispersed (**Figure 3.13**).
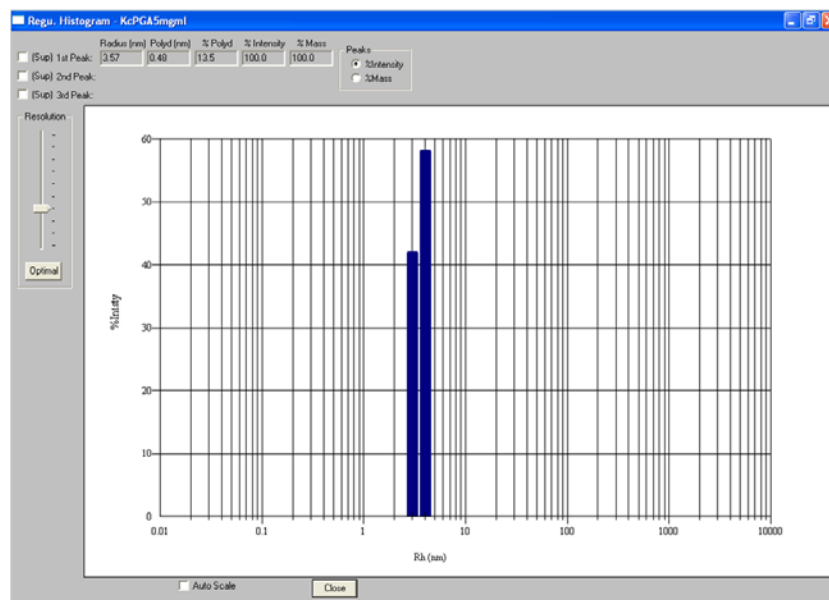
Figure 3.13: Dynamic Light Scattering profile showing high monodispersity of wild-type *Kc*PGA at a concentration of 5 mg/ml.
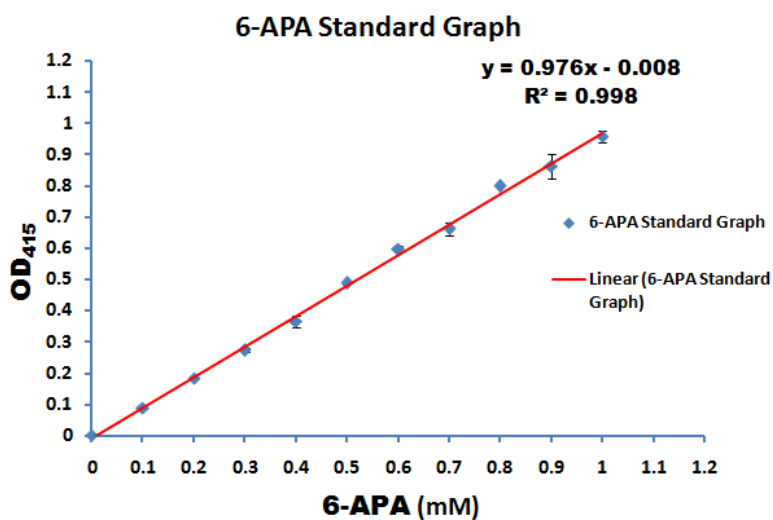
Figure 3.14: Standard graph of 6-APA detection. Different moles of pure 6-APA were reacted with PDAB for 2 min as described in materials and methods and the Schiff base formed was quantified by reading the resultant yellow colour at 415 nm wavelength.

3.2.6.3. Effect of Temperature on PGA Activity and Stability

When measuring the temperature effect, as expected in chemical reactions, the reaction velocity catalyzed by enzyme increased exponentially with increase in temperature over the rage where enzyme is stable. Enzymes are fragile and projected to operate only at physiological temperatures. When used at temperature above 37°C, they lose activity due to continuous denaturation of enzyme molecules with the increase in temperature or the increase in reaction time. **Figure 3.16A** shows the optimum temperature profile of *Kc*PGA. The optimum temperature was found to be 65 °C. The temperature stability of *Kc*PGA was studied from 25 to 70 °C in 5 degree increments for time interval of 120 min and it was found that the enzyme retains its activity over a range of temperatures. It retained almost 80% of its activity up to 50 °C when kept for 2 hours. Above this temperature, the activity reduces sharply on incubation for 30min. The activity is almost zero at 60 °C and above (**Figure 3.16B**).
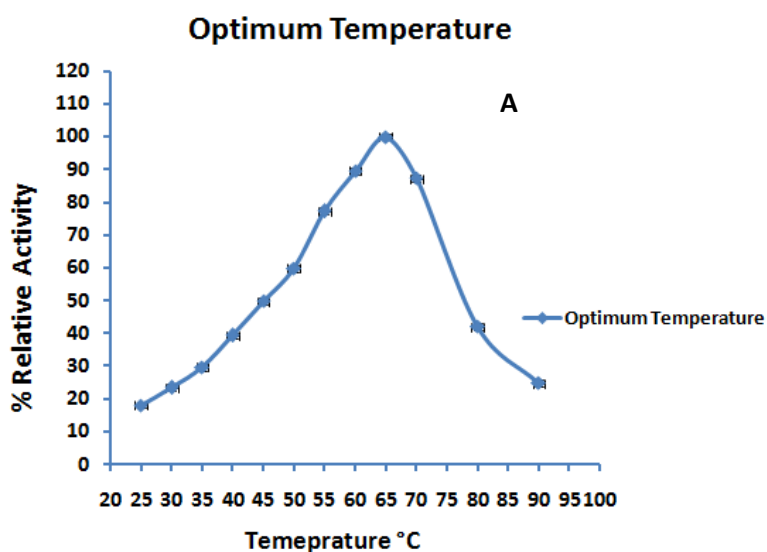
Figure 3.16A: The effect of temperature on *Kc*PGA activity. The activity was determined at different temperatures and pH 7.5 in 0.05 M potassium phosphate buffer using 7.2 μg of enzyme. The activity at 65°C was taken as 100 % (103.35 IU/mg).
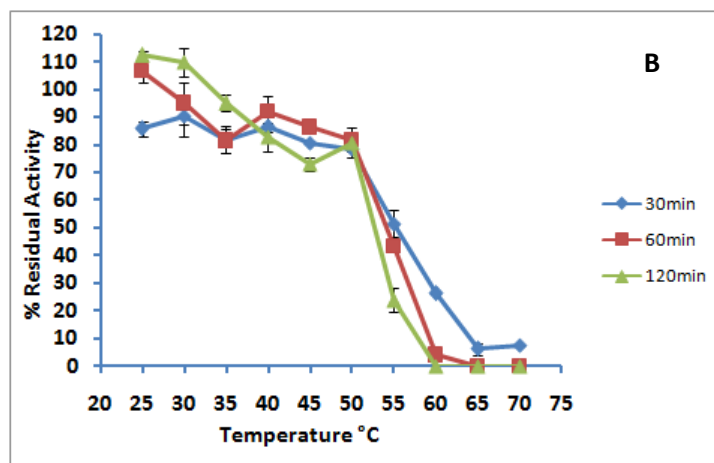
Figure3.16B: Thermostability of *Kc*PGA. Enzyme (32.5µg) was incubated at temperatures from 20-70°C up to 2 h in 0.05 M potassium phosphate buffer, pH 7.5. The activity of PGA incubated at 25°C at 0 min was set as 100 %.

3.2.6.4. Effect of pH on *Kc*PGA Activity and Stability

Many enzymes present characteristic pH value at which the activity will be a maximum. Away from this pH, the activity reduces. The enzyme's three-dimensional structure, responsible for its catalytic activity, is stabilized by hydrogen bonds, hydrophobic interactions and disulfide bonds. Altering the hydrogen ion concentration may modify the equilibrium of these forces, reversibly or irreversibly deactivating the enzyme. Other factors influencing enzymatic activity are related to pH dependence of enzyme's acid and alkaline behaviour and that of substrate. Generally only one of the ionic forms of a substrate is accepted by the enzyme and the concentration of ionic form (neutral or charged) depends on the pH. If an amino acid residue is directly involved in the catalysis, its charged state may have a major influence on the activity (de Souza *et al.,* 2005).

The ability of *Kc*PGA to hydrolyze substrate was measured in the range of pH 1-12. Studies on the effect of pH on enzyme catalysis revealed that the enzyme was active in a broad range of pH i.e., pH 5.0 to 10.0. The highest hydrolytic activity of PGA was observed at pH 7.5 (**Figure 3.17A**). The stability of *Kc*PGA was studied and it was found that the enzyme was quite stable over a wide range of pH (**Figure 3.17B**). Enzyme showed pH stability over the range 4.0 to 10.0. It retained more than 80% of its activity at pH 10.0 even after 8 hrs of incubation at room temperature. The higher stability at different pHs greatly improves handling or treatment of the enzyme during

preparation for industrial applications (e.g., purification, immobilization, chemical modification etc.), the higher stability at neutral pHs greatly improves the industrial hydrolysis of penicillin G, the higher stability at acidic pHs could be very useful in improving the performance of some synthetic processes (e.g., esterifications or synthesis of antibiotics).
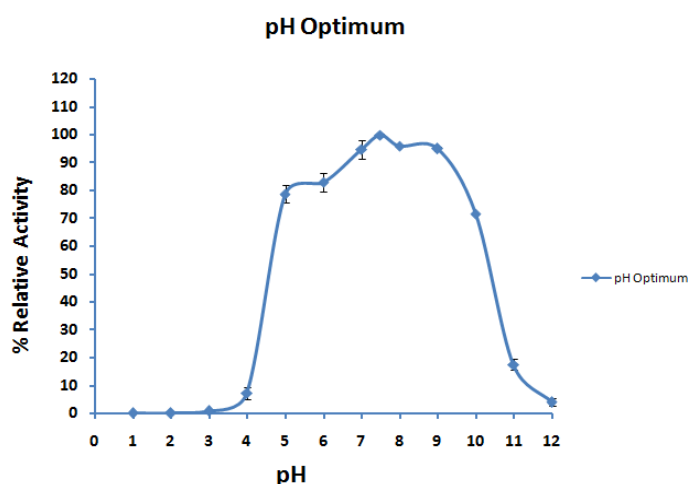
Figure 3.17A: Effect of pH on *Kc*PGA activity. The activity was determined at 50°C at different pH using different buffers at 0.1M concentration and ~7.5 µg of enzyme. The activity at pH 7.5 was set as 100% (63.59 IU/mg).
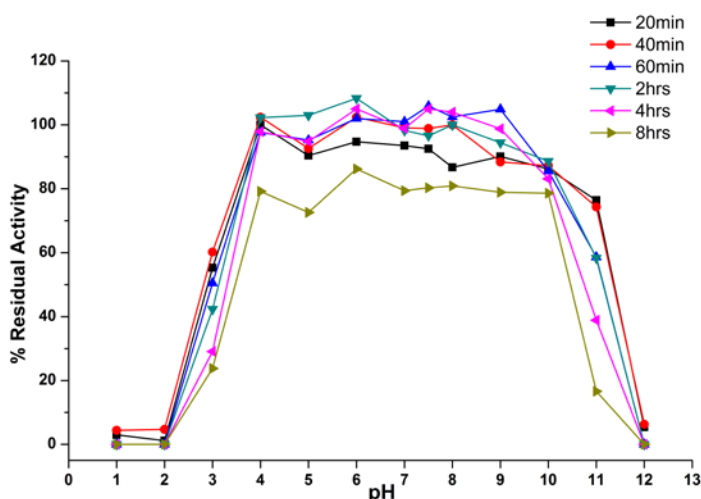
Figure 3.17B: Stability of *Kc*PGA at various pH and buffers. Enzyme (32.5µg) was incubated in different buffers of molarity 0.1M and different pH at 25 °C. The activity at pH 7.5 at 0 min was taken as 100%

Wen *et al*. (2004) showed that the pH stability of *Kc*PGA cloned in pET28b (+) and expressed in BL21 (DE3) cells had a plateau going from 5.0 to 10.0 when stored

for 30 min and degradation of enzyme was observed above 60 °C. Alvaro *et al*. (1992) also reported that *Kc*PGA is 10 fold more stable at pH 8.0 than *Ec*PGA when incubated at 58°C. Our results also showed that the *Kc*PGA was stable in a range of pH and showed almost 50% activity retention when stored at 55 °C for 30 min. These results points to its potential industrial importance.

3.2.6.5. Substrate Specificity

The purified *Kc*PGA was examined for its ability to hydrolyze natural and semi-synthetic β-lactam antibiotics under optimized conditions. The data is represented in **Figure 3.18.** Activity with Pen G was considered 100% and relative activities for other substrates were calculated. Results showed that *Kc*PGA had some affinity towards amoxicillin and Ampicillin. Hydrolyzing Pen V by *Kc*PGA is almost negligible.
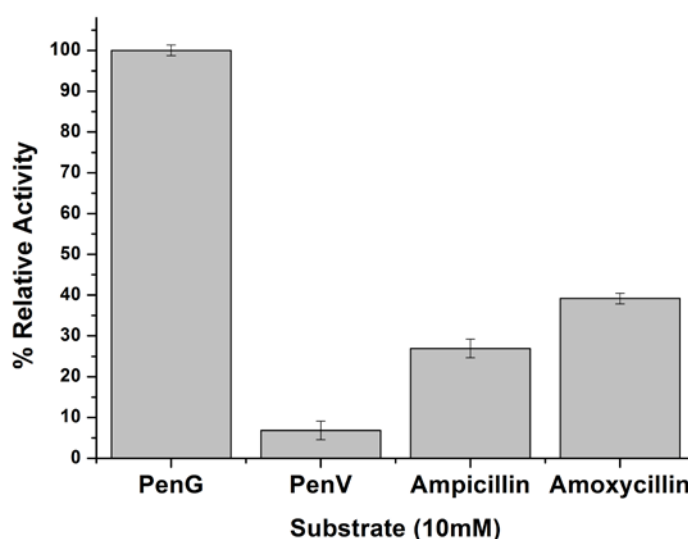
Figure 3.18: Substrate specificity profile of KcPGA.

3.2.6.6. Potential Modulators of Enzyme Activity

Effect of metal ions, non–ionic detergents and EDTA on PGA activity is shown in **Figure 3.19**. Amongst the metal ions tested, the PGA activity was inhibited by Hg^{2+}, Ag^{1+}, Fe^{2+}, Cu^{2+}, Zn^{2+}, Pb^{2+} and SDS by 41%, 10% 14%, 4%, 7%, 5% and 5%, respectively. No other metal ions or EDTA (10 mM) had any significant effect on the enzyme activity. Non ionic detergents like Triton X-100, Tween-80 and IGEPAL marginally improved the hydrolytic efficiency of *Kc*PGA.
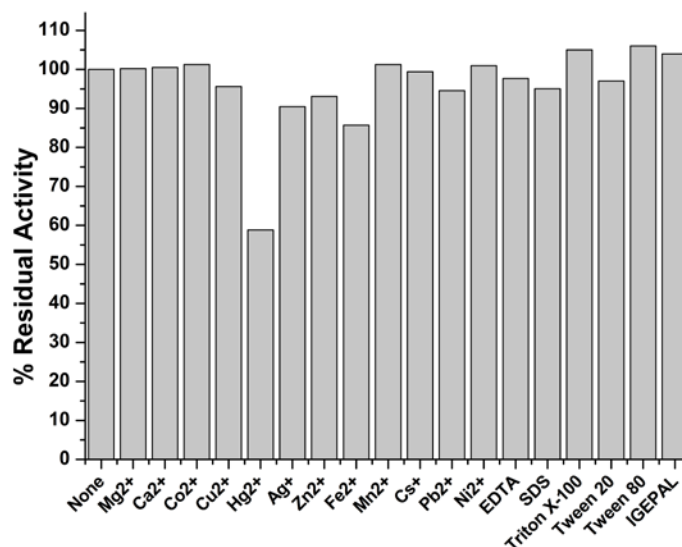
Figure 3.19: Effect of potential modulators on the PGA activity of *Kc*PGA.

3.2.6.7. Enzyme Kinetics

The purified *Kc*PGA preparations were characterized kinetically at pH 7.5 and 50 °C. The enzyme was incubated with various concentrations of the substrate from 0.1 mM to 50 mM. The effect of increasing the concentration of PenG on PGA activity of *Kc*PGA was shown in Michaelis–Menten plot (**Figure 3.20**). The V*max*, K*m* values were calculated by fitting a linear regression curve to data points using Lineweaver-Burk plot. Vmax and Km from Lineweaver Plot were found to be 62.5 μmole per min per mg and 2.68 mM respectively at the assay conditions of pH 7.5 and 50 °C (**Figure 3.21**). K_{cat} and K_{cat}/K_m were estimated are 90.11 sec^{-1} and 33.62 $mM^{-1}sec^{-1}$, respectively.
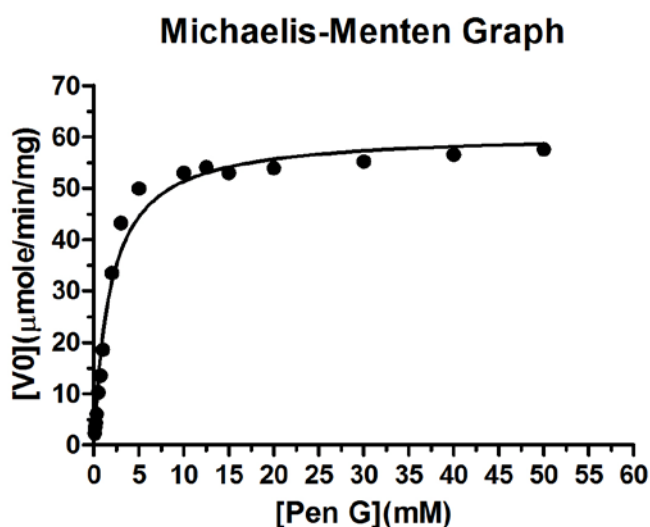
Figure 3.20: The Michaelis-Menten model adjusted to initial velocities of penG hydrolysis, determined for various initial substrate concentrations at pH 7.5 and 50 °C.
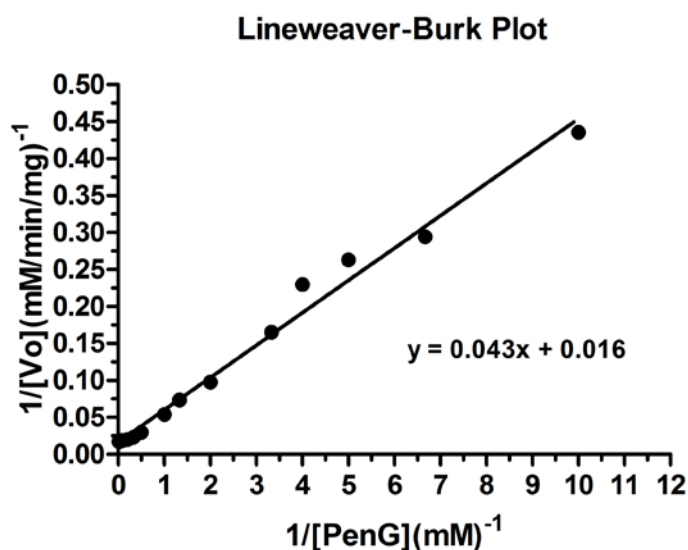
Figure 3.21: Lineweaver–Burk plot for the determination of Km and Vmax of *Kc*PGA

3.2.6.8. Effect of pH and Temperature on the *in vitro* Processing of *Kc*PGA Precursor

*Kc*PGA encoded by a single gene and is initially synthesized as a single polypeptide that undergoes posttranslational processing. It is reported that in *E.coli* PGA the Ser290 to Arg/Thr/Gly/Ala mutants lack auto-processing capability with the result that the spare peptide is retained in the mutant proteins (Choi *et al.*, 1992; McVey *et al.*, 1999). This observation raised the question whether the same phenomenon is

occurring in the case of *Kc*PGA also. From the expressed protein of Ser290Gly and Ser290Cys mutants in *Kc*PGA it is inferred that the processing is not completely blocked in the case of Ser290Gly but the mutation has resulted in the formation of slow processing precursor PGA. This mutant is similar to the Thr289Gly mutant reported by Kasche *et al*. (1999) where both respective mutants resulted in the formation of slow processing precursor proteins. *In vitro* processing of this slow processing mutant has been studied at different pH and temperature with respect to time (**Figure 3.22 A, B, C and D**).
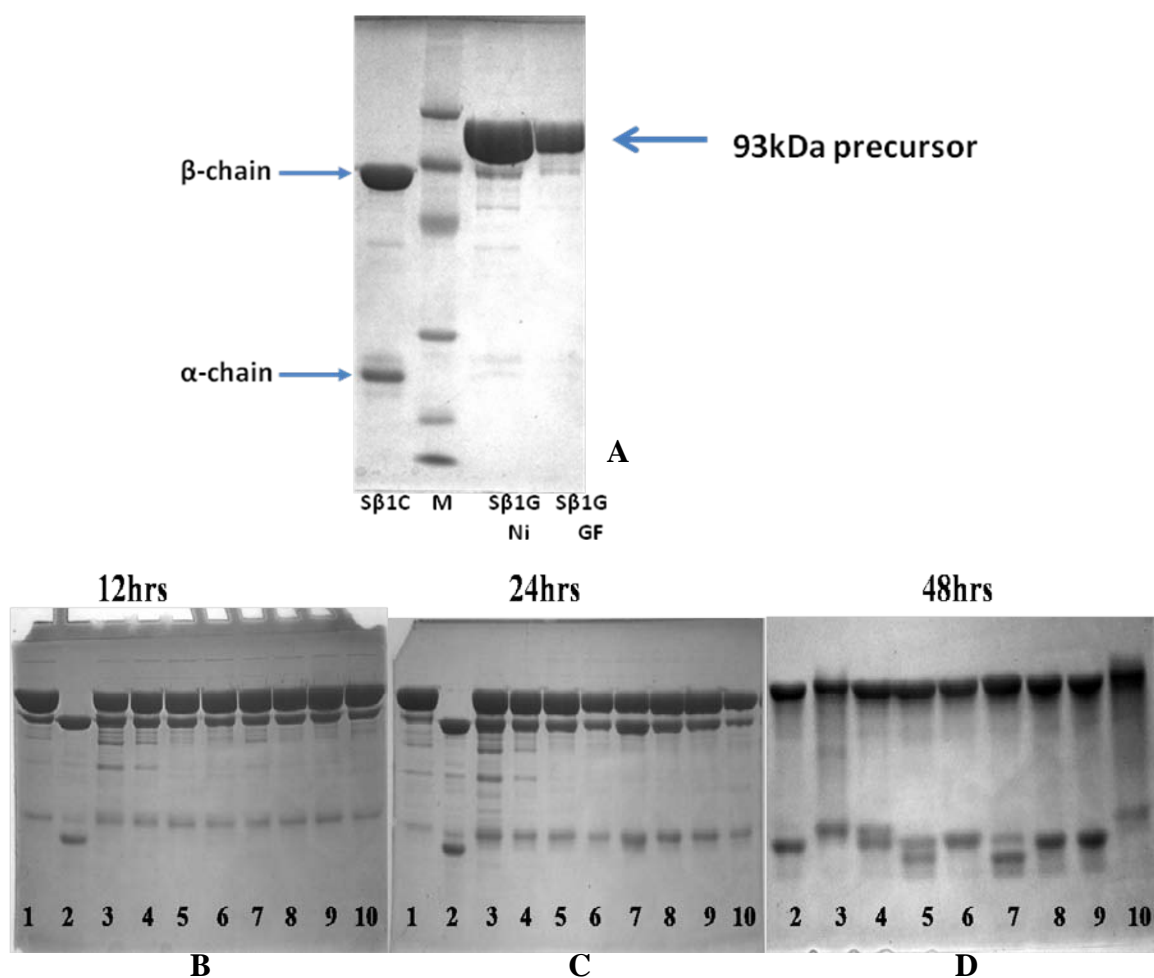
Figure 3.22: Series of SDS-PAGE showing the autocatalytic processing of slow processing precursor mutant (Ser290Gly) of KcPGA in vitro, lane1:precursor PGA stored at pH 7.5 in -20°C for a week, lane2:Wild-type processed KcPGA, lane3:precursor PGA incubated at pH 5.0 at room temperature, lane4:precursor PGA incubated at pH 6.0 at room temperature, lane5:precursor PGA incubated at pH 7.0 at room temperature, lane6:precursor PGA incubated at pH 7.5 at room temperature, lane7:precursor PGA incubated at pH 8.0 at room temperature, lane 8:precursor PGA incubated at pH 8.5 at room temperature, lane9:precursor PGA incubated at pH 9.0 at room temperature,

lane10: precursor PGA incubated at pH 7.5 on ice, (A) showing Ser290Gly precursor freshly purified along with Ser290Cys mutant protein.

Incubation of purified slow processing mutant (Ser290Gly) in 0.1M buffers from pH 5.0 to pH 9.0 showed the different rate of processing affected by pH of the buffer. From the **Figure 3.22D** it is evident that at lower pH (pH 5.0) and incubation at lower temperature of ice slows down the processing of the precursor PGA. Presence of different processing intermediates at different pH after 48 hrs supports the hypothesis of pH dependent autocatalytic processing of precursor PGA as proposed by Lee *et al.* (2000) in the case of *Ec*PGA. Previous studies in the autocatalytic process of LexA (Little, 1993), glycosylasparaginase (Guan *et al.*, 1998) and in some serine proteases (Keiler & Sauer, 1995; Vandijl, 1995) also shows the pH dependent precursor processing.

3.3. DISCUSSION

The profound impact of penicillin acylases in the manufacture of antibiotics is reason for continued interest in studying penicillin acylases from different sources. The most common antibiotics used are based on the derivatives of 6-APA and 7-ACA, which are manufactured from penicillin by the action of penicillin G acylases. Since the cost of producing 6-APA has a direct impact on profitability of semi-synthetic antibiotic production, an enzyme with enhanced property is always an economic boon. In this regard, PGA from *Kluyvera citrophila* (*Kc*PGA) has attracted attention due to its better and more suitable features for industrial applications. It is comparatively easier to immobilize and shows more stability towards extreme conditions of temperature, pH, and organic solvents as compared to PGA from *E.coli* (Alvaro *et al*., 1992; Fernandez-Lafuente *et al*., 1991, 1996; Liu *et al*., 2006). The stabilization of the enzyme by immobilization is also reported (Guisan *et al*., 1993). As one of the most prospective use in industrial application, overproduction of *Kc*PGA through the recombinant DNA technology is significant. However, the previous reports on *Kc*PGA mainly focused on its molecular characterization (Garcia *et al*., 1985) and substrate specificity (Roa *et al*., 1995). Few studies have been done with this enzyme regarding systematically genetic engineering including cloning and expression of recombinant enzyme (Barbero *et al*., 1986; Wen *et al*., 2004, 2005). Cloning and Overexpression of the enzyme will enable its manipulation to produce mutant proteins and greatly facilitate further investigation of this enzyme for industrial applications.

In this chapter, we have reported the successful cloning and overexpression of the recombinant *Kc*PGA containing a polyhistidine tag at the C-terminus in *E. coli*. High levels of soluble enzyme expression was achieved in *E.coli* pLysS expression cells and the enzyme was purified with simple protocol of affinity chromatography followed by polishing step of Gel Filtration chromatography to achieve monodispersity in the sample. MALDI TOF TOF, Electrospray and DLS studies indicate that the recombinant protein is highly soluble and monodisperse monomer and a fully active enzyme with expected molecular mass. The enzyme was found to be working optimally at pH range of 7.0-8.0 with best activity for Pen G at 7.5, however enzyme showed high stability in broad pH range of 4.0-10.0 as expected and shown also by Wen *et al*. (2004). However, they have showed the pH optimum of 8.5 and temperature optimum

of 55° C for the recombinant protein which is little different for our result of best pH optimum of 7.5 and 65° C as temperature optimum. However, the enzyme was found to be degrading above 50 °C and almost complete loss of the activity at 60° C within 30 min of incubation in both the cases. Enzyme showed best activity for its natural substrate penicillin G and little towards amoxicillin but nothing substantial against penicillin V or ampicillin. As shown by effect of potential modulators, most of the heavy and toxic metal ions reduced the enzyme activity after incubation. Published reports of K_m values for *K.citrophila* are far below than we have achieved (Barbero *et al.*, 1986; Prieto *et al.*, 1990; Rao *et al.*, 1994). V_{max} and K_m as calculated for penicillin G are deduced from Lineweaver Burk Plot and found to be 62.5 µmole per min per mg and 2.68 mM respectively, however are in the range for the other PGAs reported (Savidge & Cole 1975; de Souza, 2005; Chiang & Bennett, 1967; Ohashi *et al.*, 1989). This may be due to the use of crude penicillin G in this work.

The advantageous properties of this enzyme such as a broad pH, ease of immobilization and temperature stability are of industrial interest. However, it is necessary to re-examine the suitability and behaviour of the enzyme as a biocatalyst in an immobilized form since this is the most important criterion for evaluation of industrial applicability. Therefore, the high yield and high solubility of recombinant *Kc*PGA produced in *E.col*i represent an excellent basis for further development of recombinant *Kc*PGA as an immobilized biocatalyst for industrial applications.

Here we also tried to understand the pH dependent processing of the *Kc*PGA using the prepared site directed mutant. On the contrary to the report from Choi *et al.* (1992) on *E.coli* PGA, the Serβ1Gly mutation resulted in the expression of slow processing mutant in *Kc*PGA. Results from incubation of the mutant in different pH showed the pH dependent rate of processing of the precursor. Lower pH (pH 5.0) and low temperature incubation slows down the processing of the precursor PGA. The absence of Ser at β1 position and presence of different processing intermediates at different pH after 48 h hints at the possibility of pH dependent autocatalytic processing of precursor in *Kc*PGA with the involvement of another amino acid (Lysβ10 or Lys299 in pre-pro-enzyme) in the catalytic centre along with Serβ1 as suggested by Lee *et al.* (2000) in the case of *Ec*PGA. However, more studies need to be done in future to support and to generalise this observation in the case of other PGAs.

Chapter: 4

Structural and folding studies of wild-type enzyme and slow-processing mutant of *Kc*PGA

4.1. INTRODUCTION

Post-translational autoproteolysis is a mechanism that activates numerous proteins via self-catalyzed peptide bond rearrangements, which play an essential role in a wide range of biological processes. Intramolecular proteolysis, one such autoprocessing event, is utilized to cleave precursors and activate various proteins, including viral proteins (Douglass *et al.*, 1984) and in higher eukaryotic peptides like pre-pro-enzymes (Petrovan *et al.*, 1998; Stennicke & Salvesen, 1998). This maturation phenomenon was considered a hallmark of eukaryotes until a better understanding of posttranslational processing in prokaryotic systems became available (Trumpower, 1990; Uozumi *et al.,* 1989; Aronson *et al.,* 1989; Thony-Meyer *et al.,* 1992; Okada *et al.*, 2007).

A range of autocatalytic processing events proceed via protein splicing in many organisms, where the initial steps are proposed to be involve an N $\rightarrow$ S or N $\rightarrow$ O acyl rearrangement to generate a branched intermediate (Shao *et al.*, 1996). The Ntn hydrolase is one such family, where the members are expressed as precursors which undergo intramolecular autocatalytic cleavage to remove a spacer peptide to generate the active enzyme with N-terminal residue acting as a nucleophile. A snapshot of an active pre-autoproteolysis precursor molecule is essential to understand the mechanism of autoproteolytic reactions. The most accurate experimental means of obtaining a detailed picture of a protein molecule, allowing the resolution of individual atoms and hence its conformation and interactions, is by using X-ray crystallography.

Detailed 3D structural information is not available for most Ntn-hydrolase precursors, since they are processed rapidly. Indeed, the only Ntn-hydrolase precursor structures that have been determined thus far are mutant constructs that inhibit the speed of processing e.g., glycosylasparaginase (Xu *et al.*, 1999; Wang & Guo, 2010), the β-subunit of the proteasome (Ditzel *et al.*, 1998), the PGA precursor from *E.coli* (Hewitt *et al.*, 2000), a plant L-asparaginase (Michalska *et al.*, 2008b), glutaryl 7-aminocephalosporanic acid acylase (Kim *et al.*, 2003) and cephalosporin acylase (Kim *et al.*, 2002). A strained backbone conformation near autocleavage sites has been reported for many autoprocessing proteins. However, precursor activation mechanisms

are not clearly understood. In particular, the exact autoproteolytic potential is far from complete.

3D structure of a protein is held together by several factors like hydrogen bonds, ionic, hydrophobic and van der Waal's interactions as well as covalent interactions like disulfide linkages. The stability and the function of a protein depend on these interactions. The conditions that interfere with these interactions lead to loss of 3D structure as well as activity or function of the protein. Increase in temperature and changes in pH, denaturation by chemical agents like Gdn-HCl and urea, reduction of disulfide linkages by DTT are some of such conditions. Changes in the native conformation of a protein under different conditions can be followed by using steady state fluorescence spectroscopic technique. Tryptophan has much stronger fluorescence compared to tyrosine and phenylalanine when excited upon absorption of light energy (photon). Fluorescence of a tryptophan residue is influenced by its microenvironment; hence changes in tryptophan environment modify fluorescence properties (Glaudemans & Jolley, 1980). Hence by studying the tryptophan fluorescence, the conformational changes in the protein can be monitored. Understanding structure-function relationships also facilitate protein-engineering of the enzyme to enhance the potential for chemical, industrial and specific therapeutic applications.

Here we report the analysis of the X-ray diffraction data obtained from crystals of precursor PGA from *Kluyvera citrophila* (*Kc*PGA) (slow processing mutant Serβ1Gly). X-ray diffraction data for two crystal forms of this mutant *Kc*PGA belonging to space groups P1 and C2 were collected at resolutions 2.5 Å and 3.5 Å, respectively and the structure was determined using molecular replacement method using coordinates of *E. coli* precursor PGA (PDB code: 1E3A, Hewitt *et al.*, 2000) as input model. The structure was refined in space group P1 using data at 2.5 Å resolution.

This chapter also describes the structural changes accompanying denaturation by chaotropic agents such as guanidium hydrochloride (Gdn-HCl), temperature and pH and stability of wild-type processed *Kc*PGA were monitored using steady state fluorescence spectroscopy. The microenvironments of the Trp residues were also assessed by using various solute quenchers and change in intrinsic fluorescence.

4.2. RESULTS

Like other Ntn hydrolases, *Kc*PGA is also expressed as a single-chain 96 kDa precursor consisting of 844 amino acid residues in the cytoplasm. The autocatalytic processing removes a 26 residue signal peptide and a 54 residue linker peptide that results in the formation of active enzyme in the periplasm, which is a heterodimer of α-chain of 23.5 kDa and β-chain of 62 kDa with 209 and 557 amino acid residues, respectively (Alvaro *et al*., 1992; Barbero *et al*., 1986).

4.2.1. Structural Studies of *KcPGA* unprocessed Precursor Mutant

4.2.1.1 Preparation of Site Directed Mutants of *Kc*PGA

Site directed mutants of penicillin G acylase from *Kluyvera citrophila* DMSZ 2660 have been prepared to study the mechanism of autoproteolysis and catalysis. The wild-type "*pac*" gene from *K.citrophila* inserted into the plasmid pET26b (+) (Section **3.2**) was used as template for the preparation of two mutants Ser290Cys and Ser290Gly. The preparation and expression of site-directed mutants were described in sections **3.2.3** and **3.2.4.**

4.2.1.2. Purification of *Kc*PGA Mutants

Both the mutant PGAs of *Kluyvera citrophila,* Ser290Gly (Sβ1G) and Ser290Cys (Sβ1C) were purified with the same protocol as used for purifying wild-type (Section **3.2.5**). Briefly the mutant proteins were purified by using Nickel affinity chromatography followed by a size exclusion chromatography. The yield of purified mutant proteins Ser290Cys and Ser290Gly was ~12 and ~15 mg/L, respectively. Both the proteins were found to be highly soluble and could be concentrated to more than 50 mg/ml concentration without any visible precipitation. Ser290Cys mutant showed two bands of ~23.6 kDa and ~63 kDa correspond to α and β chains respectively, whereas Ser290Gly showed the presence of precursor PGA protein with molecular weight ~93 kDa as seen in the run of SDS-PAGE gel (**Figure 4.2**).
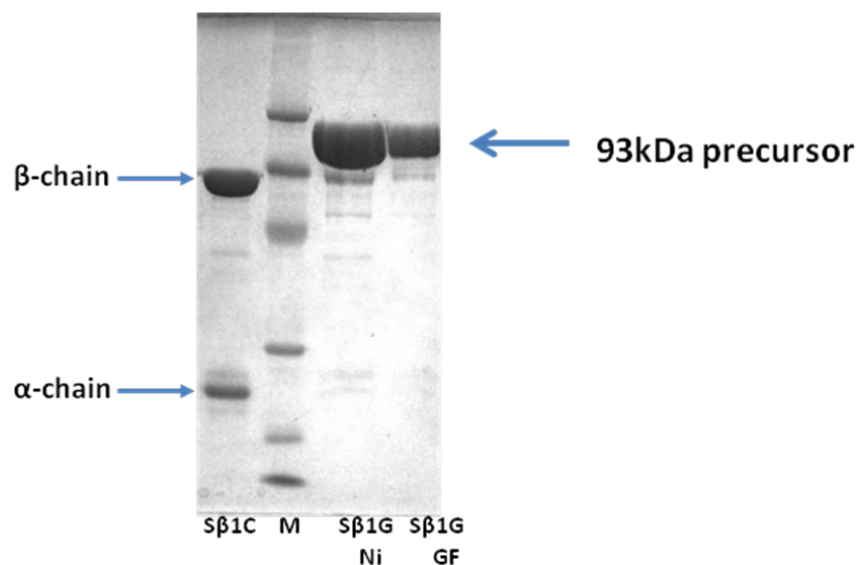
Figure 4.2: 12% SDS- PAGE to show the extent of purity of Ser290Cys (Sβ1C) and precursor Ser290Gly (Sβ1G), Sβ1GNi denotes the nickel purified precursor, Sβ1G GF denotes protein eluted from Gel filtration column, M– low molecular range protein marker.

4.2.1.3. Crystallization trials of *Kc*PGA Proteins

Wild-type *Kc*PGA and the mutants were screened for crystallization conditions. After purification, the proteins were buffer exchanged with 50 mM HEPES, 50 mM NaCl, 1 mM DTT and concentrated to 30 mg/ml. Proteins were then centrifuged at 10,000 rpm for 5 min at 277 K before keeping for crystallization. Since Ser290Gly (Serβ1Gly) mutation resulted in a slow processing mutant, crystallization was setup as soon as the protein was purified and buffer exchanged. Crystallization was performed at 293 K using the vapour-diffusion method with sitting-drops: 300 nl protein solution (23 mg ml^{-1}) was mixed with 300 nl reservoir solution and equilibrated against 100 µl reservoir solution. The screens were set up using the Mosquito crystallization robot (TTP LabTech Ltd., UK) as sitting-drop vapour-diffusion experiments in 96-well MRC plates (Hampton Research). Commercial crystallization kits from Hampton Research, Molecular Dimensions Ltd., Emerald BioSystems, Qiagen and some self prepared screens were employed in the screening experiments. However, presently the precursor *Kc*PGA (slow processing mutant, Ser290Gly) only has crystallized.

Initially crystals of Ser290Gly were observed in one of the self prepared matrix screens. Multiple thin plate like crystals were observed in solution containing 30% (w/v) PEG 4,000, 50 mM sodium cacodylate buffer pH 5.6 and 0.5 M KSCN in 3-4 days. Varying pH of the sodium cacodylate buffer (50 mM) and using similar concentrations of protein and precipitant, employing Gryphon crystallization robot (Art Robbins Instruments, USA) by mixing 500 nl protein solution with 500 nl of well solution, crystals were grown under a wide range of pH within a week. Crystals obtained at pH 4.6 and pH 6.5 had similar morphology but differed in diffraction quality (**Figure 4.3**).

4.2.1.4. Diffraction Data Collection on Ser290Gly Crystals

These crystals were transferred into a cryoprotectant solution composed of the reservoir solution containing 30% glycerol and were flash-cooled in a nitrogen stream at 100 K. Two datasets were collected at 100 K on Stanford Synchrotron Radiation Lightsource (SSRL) beam line BL12-2 at the SLAC National Accelerator Laboratory. Diffraction data were collected on DECTRIS PILATUS 6M detector with 0.2 sec exposure and 90% attenuation.

Diffraction data were integrated using *XDS* (Kabsch, 2010) and scaled with *SCALA* in *CCP4* suite (Winn, 2011). Indexing the diffraction pattern identified the crystals obtained from pH 4.6 and 6.5 as belonging to two different space groups. Crystals grown at pH 4.6 belonged to the space group P1 with cell dimensions $a = 54.00$Å, $b = 124.56$ Å, $c = 135.14$ Å, $\alpha = 104.05°$, $\beta = 101.37°$, $\gamma = 96.51°$ whereas crystals obtained at pH 6.5 belonged to space group C2 with unit cell dimensions $a = 265.14$ Å, $b = 53.96$ Å, $c = 249.19$ Å, $\beta = 104.44°$.

Datasets showed that the crystals encountered quick radiation damage, in spite of freezing with cryoprotectant. We could collect complete data from monoclinic crystals, but at a poor resolution of 3.5 Å. Data from triclinic crystals had better resolution of 2.5 Å spacing but the overall completion was only 76.5%. Diffraction images from both types of crystals are shown in **Figure 4.4**. The data collection and processing statistics for these crystals are summarized in **Table 4.1**.
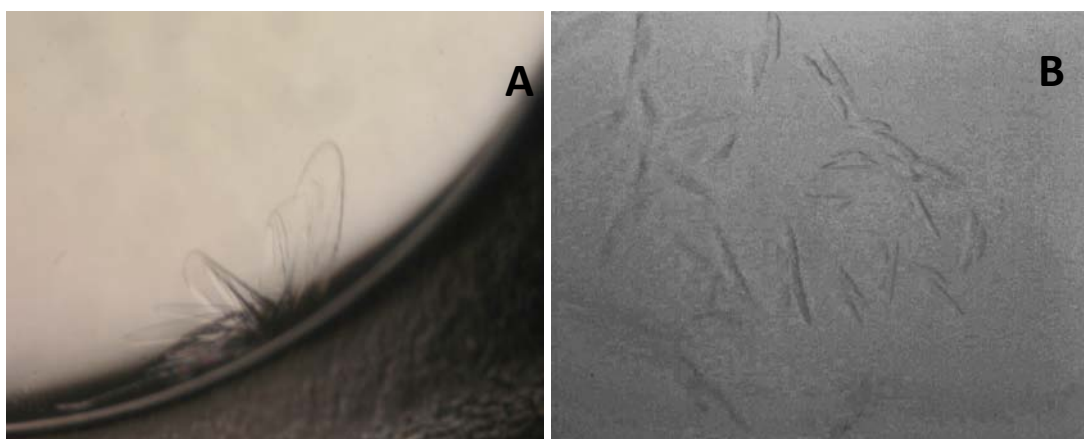
Figure 4.3: Crystals of slow processing mutant of *Kc*PGA (Serβ1Gly), which appeared within a week. (A) Crystals grown at low pH of 4.6 as clicked through a microscope, (B) Crystals grown at higher pH of 6.5 imaged using Rigaku Crystal Imager.
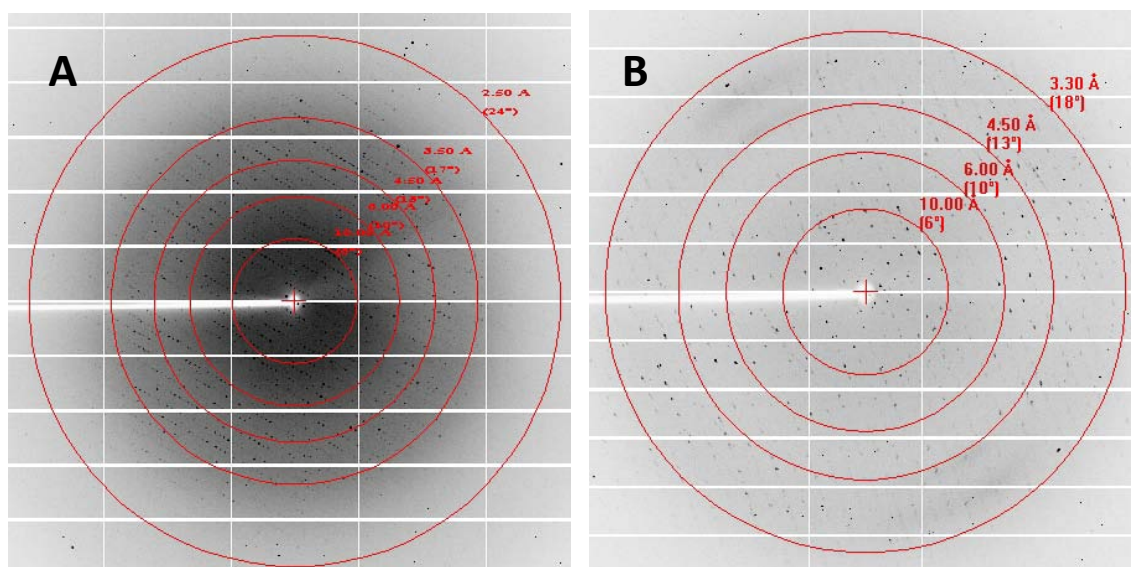
Figure 4.4: X-ray diffraction patterns obtained from *Kc*PGA precursor mutant crystals (A) triclinic P1 dataset, (B) monoclinic C2 dataset. Numbers denote the resolution of the rings.

4.2.1.5. Matthews coefficient and Solvent Content

Assuming molecular weight of precursor PGA as 92 kDa, the calculated Matthews coefficient (Matthews, 1968) for four molecules in the asymmetric unit of

triclinic unit cell P1 and monoclinic unit cell C2 are 2.48 $Å^3$ Da^{-1} and 2.51 $Å^3$ Da^{-1}, which corresponds to a solvent content of 50 % and 51%, respectively.

Table 4.1: Data collection statistics of *Kc*PGA slow processing precursor mutant crystals

Space Group	__P1__	__C2__
Temperature	100K	100K
X-ray source	BL12-2 SSRL	BL12-2 SSRL
Wavelength	0.9560	0.9560
Resolution range (Å)	38.59 - 2.5 (2.64-2.5)	38.76-3.5 (3.69-3.5)
Unit-cell parameter (Å)	$a = 54.00$, $b = 124.56$, $c = 135.14$, $\alpha = 104.05°$, $\beta = 101.37°$, $\gamma = 96.51°$	$a = 265.14$, $b = 53.96$, $c = 249.19$ $\alpha = 90.00°$, $\beta = 104.44°$, $\gamma = 90.00°$
Molecules per AU	4	4
Matthews coefficient ($Å^3$ Da^{-1})	2.48	2.51
Solvent content (%)	50	51
Total number of observations	148560	125434
Total number unique	87317	42189
Multiplicity	1.7 (1.6)	3.0(3.1)
Average I/σ(I)	6.1(1.2)	4.0 (2.8)
Remerge (%)	9.5 (55.1)	26.1 (40.5)
Completeness (%)	76.5 (80.6)	96.0 (96.5)

4.2.1.6. Structure determination

Reflections from both the datasets were phased using Molecular Replacement (MR) method. The precursor form of *Ec*PGA monomer (PDB code: 1E3A, Hewitt *et al.*, 2000) was used as search model for both the datasets. *AutoMR* programme from *PHENIX* (Adams *et al.*, 2010; McCoy, *et al.*, 2007) was used for MR calculations. A run of *AutoMR* wizard in default mode with 1E3A as template resulted in a clear single solution with an LLG gain of 9234.89, a rotation-function Z (RFZ) score of 10.1 and a translation-function Z (TFZ) score of 50.8 for P1 dataset. Similarly, molecular-replacement solution could be obtained for C2 dataset also. The LLG gain, RFZ and TFZ in this case were 2277.97, 17.1 and 15.9, respectively. TFZ score of above 8 usually indicate correct solution (McCoy *et al.*, 2007). A large non origin Patterson peak was found in the case of C2 dataset indicating presence of a non-crystallographic translational symmetry (NCS). The pseudo-translational vector identified has components 0.2451, 0.2576, 0.4973.

The initial phases obtained from molecular replacement calculation were sufficiently accurate for automatic tracing of the protein structure and preliminary model building. Automatic rebuilding was done using the *AutoBuild* wizard (Terwilliger *et al.,* 2008) from *PHENIX* rejecting option of adding water molecules. *AutoBuild* combines density modification and chain tracing using *RESOLVE* (Terwilliger, 2003a, b) with refinement using *phenix.refine* (Afonine *et al.*, 2005, 2012) to generate a high-quality model. 5 cycles of automated model building and refinement using *AutoBuild* built almost the complete models consisting of 3272 residues out of 3312 (including hexa-Histidine tag) in 8 fragments (2 fragments each for 4 molecules in asymmetric unit) with R_{cryst} and R_{free} values of 22.9% and 27.5% for P1 dataset and 3272 residues for C2 dataset with R_{cryst} and R_{free} values of 33.88% and 41.99%, respectively, yielding electron-density maps that were informative enough to evaluate the model.

4.2.1.7. Refinement of the Structure

Since the resolution of monoclinic crystal data was poor, structure refinement was attempted only for triclinic structure. In the absence of data at atomic resolution the crystallographic refinement here aims only at optimizing the agreement of atomic

model with both observed diffraction data and chemical restraints. The structure reported in this chapter for triclinic crystal form of slow processing mutant of *Kc*PGA was refined using *phenix.refine. phenix.refine* contains the options for group ADP refinement, rigid body refinement, restrained refinement (xyz, ADP: isotropic, anisotropic, mixed), automatic water picking, applying automatic NCS restraints, simulated annealing, occupancies (individual, group, constrains for alternative conformation), TLS refinement, refinement with twinned data. *phenix.refine* automatically detects and corrects flipped N/Q/H residues at each macrocycle and uses *MolProbity*/Reduce methodology (H-bonds, clashes) to determine correct orientation. Manual fitting of the molecular model and visualization of electron density maps were carried out using *COOT* (Emsley & Cowtan, 2004). Since both the structures contain four molecules in asymmetric unit, their models were refined taking advantage of the non-crystallographic symmetry present. The model building and positional refinements were iteratively carried out until all the residues in the model were mutated to side chains matching the electron density map (residue information were also obtained from sequence alignment).

4.1.2.8. Structure validation

During final cycles of refinement the stereochemistry and the geometry of the models were checked using *PROCHECK* (Laskowski *et al.,* 1993). The overall G factor considered to be a measure of stereochemical quality of the model output by *PROCHECK* was 0.04. This is within the limits expected for a correct structure that is refined at 2.5Å resolution. The Ramachandran (φ, ψ) plot (Ramachandran & Sasisekharan, 1968) showed that the residues were in the most favoured region or partially allowed region of the map (**Figure. 4.5**). The refinement statistics and the values of refined parameters are presented in **Table 4.2.** In addition to NCS constraint the poor completion of diffraction data also may be contributing to the comparatively higher values 23% and 27% for R-factors, R_{cryst} and R_{free}.
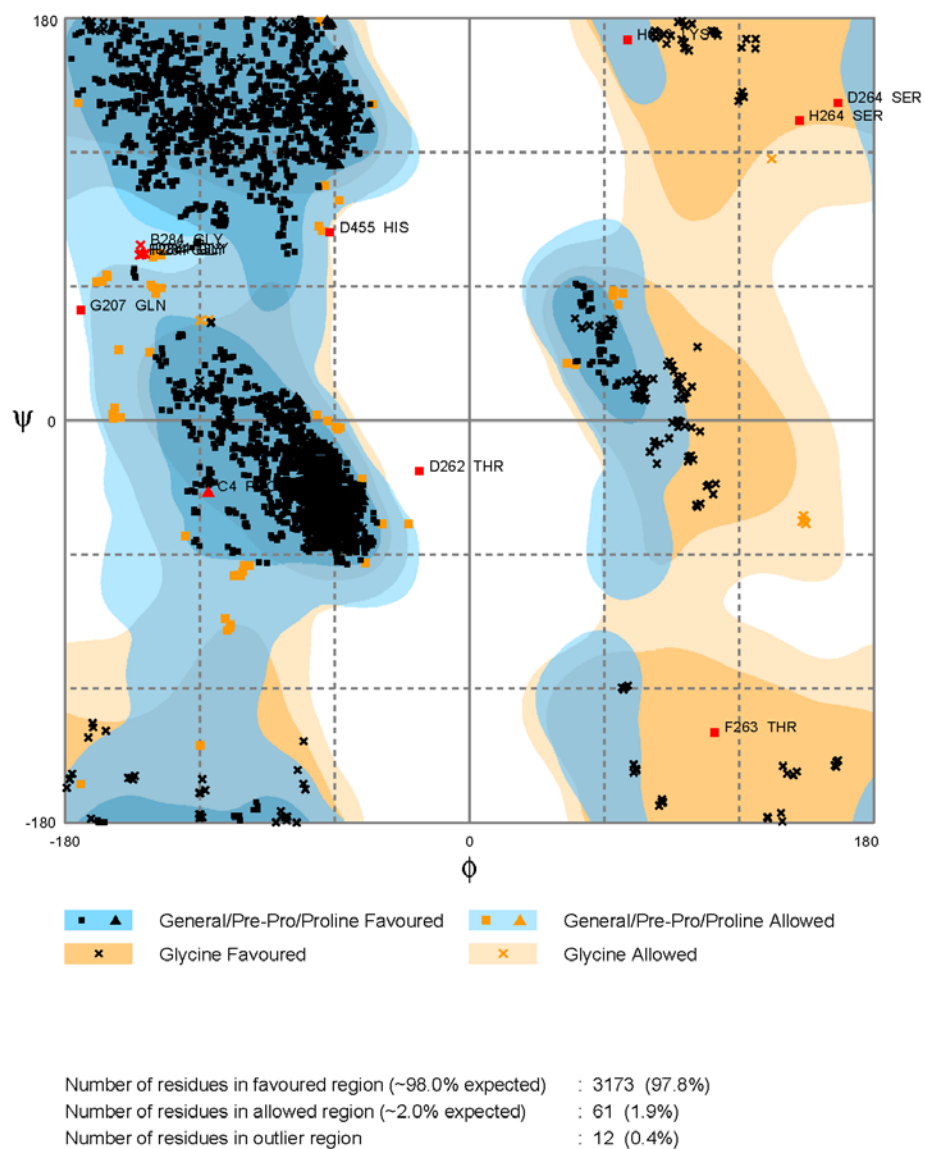
Figure 4.5: Ramachandran plot for triclinic form of *Kc*PGA precursor generated using PROCHECK.

Table 4.2: Refinement statistics for triclinic crystals of slow processing mutant of *Kc*PGA

Space Group	<u>P1</u>
Refinement statistics	
Resolution range for refinement (A°)	37.85-2.5
R_{cryst} (%)	22.97
R_{free} (%)	27.05
RMS Bond Length (Ǻ)	0.01
RMS Bond Angle (°)	1.2
Number of Protein atoms	25633
Number of Water atoms	428
Number of Heterogenous atoms	4
Overall B (Isotropic) from Wilson Plot	58.8
Ramachandran Plot	
Residues in most favoured Region	90.3%
Residues in additionally allowed Region	9.5%
Residues in generously allowed Region	0.2%

[a] $R_{cryst} = \Sigma\ |Fo\text{-}Fc\ |/\ \Sigma Fo$

[b] $R_{free} = \Sigma\ |Fo\text{-}Fc\ |/\ \Sigma Fo$ where the F values are test set amplitudes (5%) not used in refinement.

4.2.1.9. Description of the structure

In the course of structure refinement, the *Fo-Fc* difference Fourier map with spacer residues omitted for phase calculation, showed clear positive density for spacer sequence, confirming presence of residues 210-263 (numbering is according to pro-PGA enzyme ignoring signal sequence) using the unbiased map (**Figure 4.6**).

The structure of Serβ1Gly precursor *Kc*PGA is shown in ribbon and space fill diagrams (**Figure 4.7A and B**). The overall structure of *Kc*PGA precursor confirms the characteristic Ntn hydrolase fold comprising of a four-layered αββα core structure that is formed by two antiparallel β-sheets packed against each other, and these β-sheets are sandwiched between the layers of α-helices on either side. The short α-chain and the long β-chain together with the spacer peptide of the heterodimer are closely intertwined to form a pyramidal structure. The N-terminal end of the linker peptide extends away from the C-terminal end of α-chain, across to the active-site cavity where, after a β-hairpin turn, it forms a helix between residues 238 and 252. It then stretches directly down to the catalytic site of serine residue position in the mature PGA, which is deeply buried in the cavity. The secondary structure arrangement for *Kc*PGA precursor is shown in **Figure 4.8**.

Average B-factors of the main chain atoms for the entire protein calculated using Protein Server Analysis Package (PSAP) web server (http://iris.physics.iisc.ernet.in/psap/) was found to be 28.62 $\mathring{A}^2$. Listed in **Table 4.3** are the observed intra molecular hydrogen bonds of length in the range 2.25$\mathring{A}$-3.6$\mathring{A}$ for one of the chain calculated using PSAP web server.
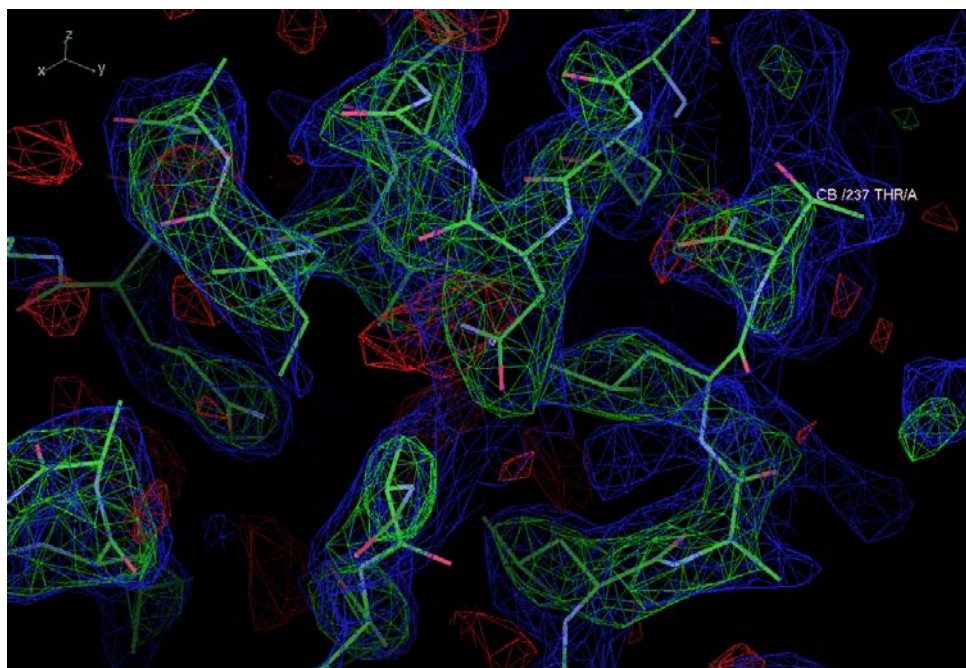
Figure 4.6: The difference electron density map showing the spacer residues superposed in the triclinic form of *Kc*PGA together with the 2Fo-Fc electron-density map contoured at 1 σ. The electron density corresponding to spacer peptide is that of a Fo-Fc map calculated in the absence of spacer and contoured at 3σ. The electron density for spacer is seen in the monoclinic form also (not shown). Figure is prepared using COOT.
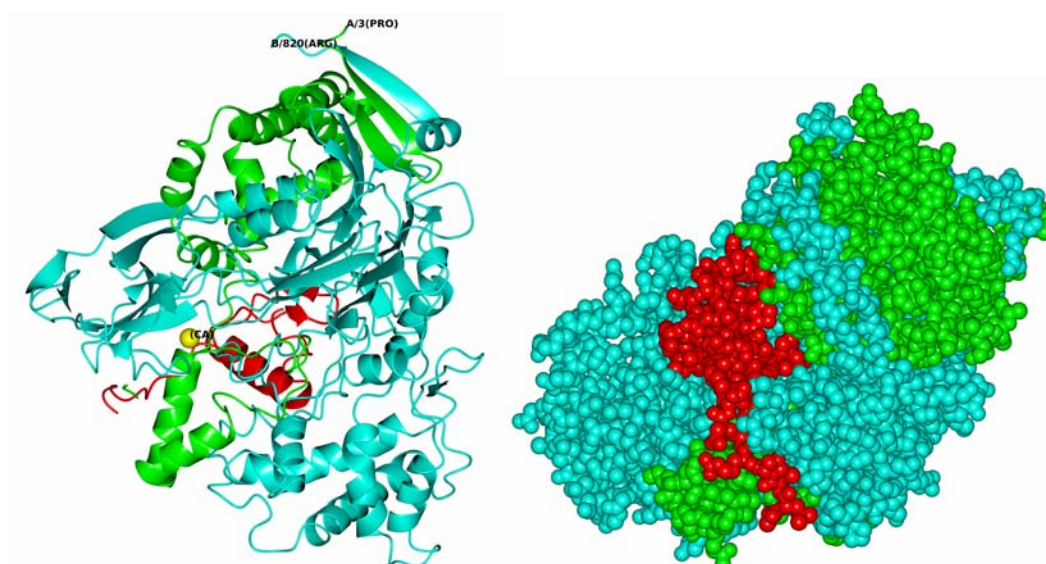
Figure 4.7: (A) Cartoon representation and (B) Space filling model of *Kc*PGA precursor showing the secondary structural elements: α-helices and β-sheets. The two chains α and β corresponding to processed form are shown in green and cyan while the linker peptide is in red. The calcium ion is shown as yellow sphere in (A). Figure was created with the program *CCP*4*MG* (McNichols *et al.*, 2011).
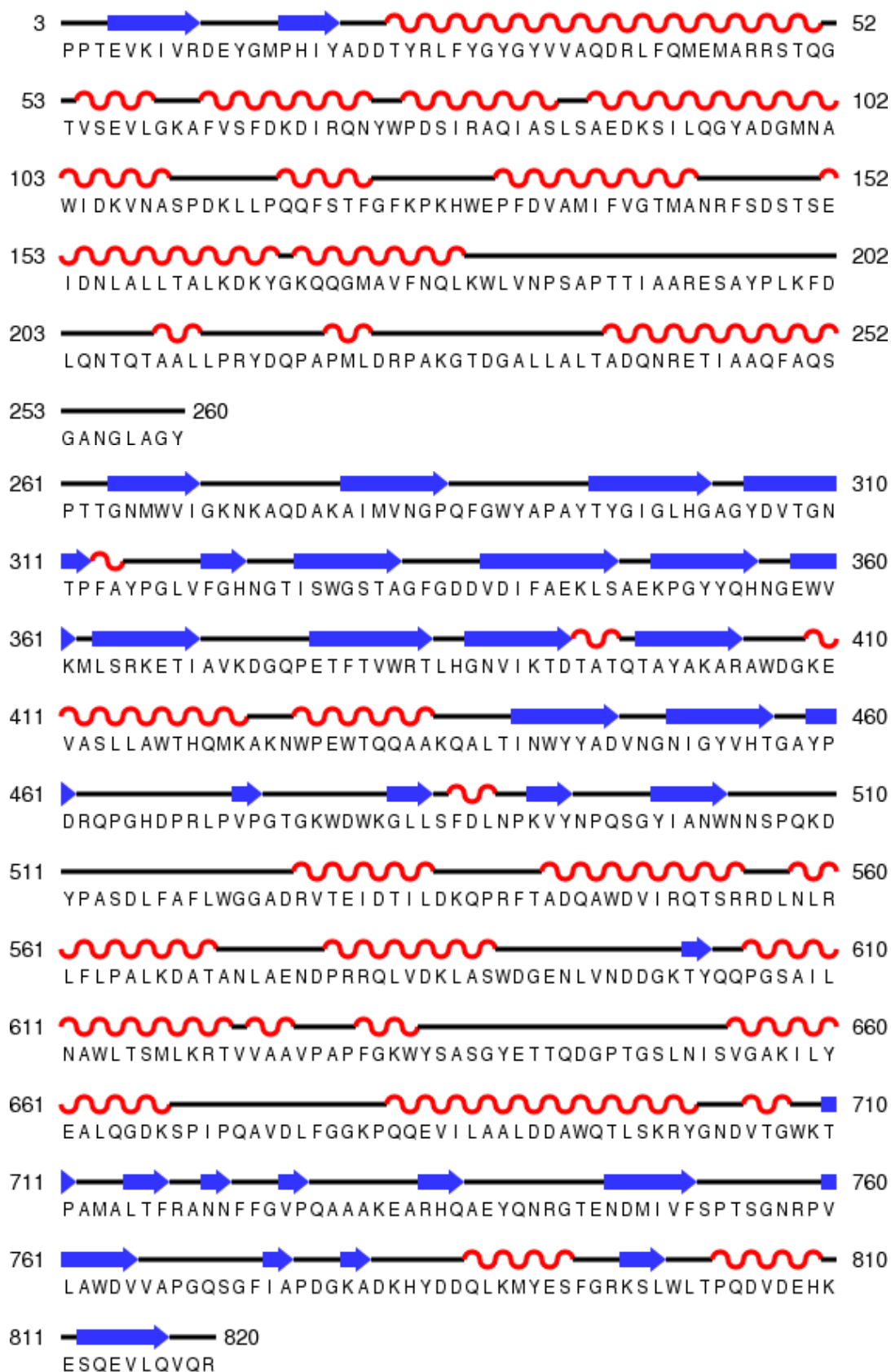
Figure 4.8: The secondary structure elements in the *Kc*PGA precursor structure marked along the amino acid sequence.

Table 4.3: The number of intra-subunit hydrogen bonds classified according to their types: main chain-main chain (Mc-Mc), main chain- side chain (Mc-Sc), side chain-side chain (Sc-Sc) and protein-water (Prot-Wat), observed for Triclinic crystal form of mutant precursor *Kc*PGA.

Interaction Chain	Mc-Mc	Mc-Sc	Sc-Sc	Prot-Wat
C	1531	734	137	217
D	3159	1844	537	368

4.2.1.10. Calcium Site

Like all other known PGA structures, precursor form of the *Kc*PGA also contains one molecule of Ca^{2+}. The residues 152Glu, 336Asp, 338Val, 339Asp, 468Pro, 515Asp coordinate with the calcium ion (**Figure 4.9**). Ca^{2+} coordinated by the side-chains of aspartyl residues 336, 339 and 515, and the main-chain carbonyl groups of Val338 and Pro468 from β-chain and the side chain of Glu156 from α chain. It is interesting to note that a bound calcium ion has been reported at the same position in the wild type PGA structures of *E.coli* and *P.rettgeri*. We have reported the presence of a bound calcium ion in PGA from *Alcaligenes faecalis* presented in this thesis (**Section 5.2.9**) suggesting that the calcium ion may have an important functional role in the PGAs.

4.2.1.11. Comparison of the precursor PGA structures of *Kc*PGA and *Ec*PGA

The structure of the precursor form of *Kc*PGA reported here is similar in structure to the Thr263Gly mutant PGA precursor from *E.coli* (PDB code: 1E3A). The root-mean-square deviation upon superposition of two precursor structures is 0.62 Å for 802 common Cα atoms (**Figure 4.10**). The amino acid sequence of precursor *Kc*PGA mutant deduced from the DNA sequence is found to be 87% identical to that of mutant *E.coli* precursor PGA (**Figure 4.11**).
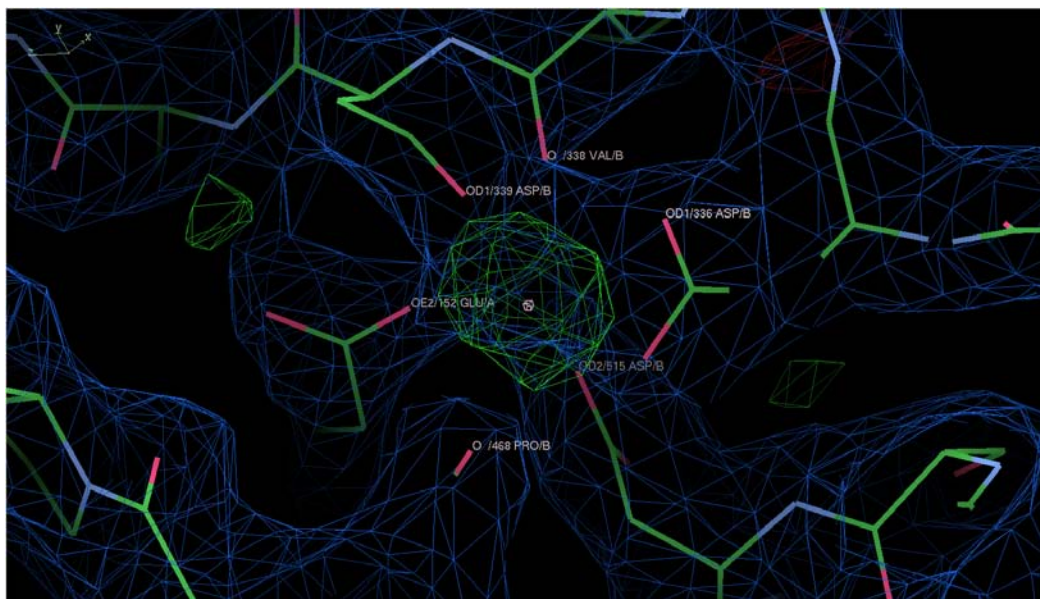
Figure 4.9: The calcium-binding site of *Kc*PGA showing the residues involved in calcium binding. The electron density surrounding Ca^{2+} is that of F$_o$ -F$_c$ map calculated in the absence of calcium and contoured at 3σ.

Figure 4.10: Superimposition of *Kc*PGA precursor structure on *Ec*PGA precursor (pdb code: 1E3A). The α-chain, spacer peptide, β-chain and calcium ion in *Kc*PGA precursor are shown in green, red, cyan, yellow whereas in *Ec*PGA precursor they are shown in magenta, blue, coral and black, respectively.

```
KcPGA        ASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYVVAQDRLFQMEMARRSTQGTVSEVLGK 60
EcPGA        EQSSSEIKIVRDEYGMPHIYANDTWHLFYGYGYVVAQDRLFQMEMARRSTQGTVAEVLGK 60
             ...:*:**************:**:***************************.*****

KcPGA        AFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQGYADGMNAWIDKVNASPDKLLPQQFS 120
EcPGA        DFVKFDKDIRRNYWPDAIRAQIAALSPEDMSILQGYADGMNAWIDKVNTNPETLLPKQFN 120
             **.******:*****:******:**.** *****************:.*:.***:**.

KcPGA        TFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEIDNLALLTALKDKYGKQQGMAVFNQLKW 180
EcPGA        TFGFTPKRWEPFDVAMIFVGTMANRFSDSTSEIDNLALLTALKDKYGVSQGMAVFNQLKW 180
             ****.**:********************************** .***********

KcPGA        LVNPSAPTTIAARESAYPLKFDLQNTQTAALLPRYDQPAPMLDRPAKGTDGALLALTADQ 240
EcPGA        LVNPSAPTTIAVQESNYPLKFNQQNSQTAALLPRYDLPAPMLDRPAKGADGALLALTAGK 240
             ***********.:** *****: **.********** ***********:*********.:

KcPGA        NRETIAAQFAQSGANGLAGYPTTGNMWVIGKNKAQDAKAIMVNGPQFGWYAPAYTYGIGL 300
EcPGA        NRETIAAQFAQGGANGLAGYPTGSNMWVIGKSKAQDAKAIMVNGPQFGWYAPAYTYGIGL 300
             ***********.**********   *******.***************************

KcPGA        HGAGYDVTGNTPFAYPGLVFGHNGTISWGSTAGFGDDVDIFAEKLSAEKPGYYQHNGEWV 360
EcPGA        HGAGYDVTGNTPFAYPGLVFGHNGVISWGSTAGFGDDVDIFAERLSAEKPGYYLHNGKWV 360
             ************************.*******************:********* ***.**

KcPGA        KMLSRKETIAVKDGQPETFTVWRTLHGNVIKTDTATQTAYAKARAWDGKEVASLLAWTHQ 420
EcPGA        KMLSREETITVKNGQAETFTVWRTVHGNILQTDQTTQTAYAKSRAWDGKEVASLLAWTHQ 420
             *****:***:**:**.********:***::.** :*******:*****************

KcPGA        MKAKNWPEWTQQAAKQALTINWYYADVNGNIGYVHTGAYPDRQPGHDPRLPVPGTGKWDW 480
EcPGA        MKAKNWQEWTQQAAKQALTINWYYADVNGNIGYVHTGAYPDRQSGHDPRLPVPGTGKWDW 480
             ****** ************************************.****************

KcPGA        KGLLSFDLNPKVYNPQSGYIANWNNSPQKDYPASDLFAFLWGGADRVTEIDTILDKQPRF 540
EcPGA        KGLLPFEMNPKVYNPQSGYIANWNNSPQKDYPASDLFAFLWGGADRVTEIDRLLEQKPRL 540
             ****.*::*********************************************** :*::.*:*:

KcPGA        TADQAWDVIRQTSRRDLNLRLFLPALKDATANLAENDPRRQLVDKLASWDGENLVNDDGK 600
EcPGA        TADQAWDVIRQTSRQDLNLRLFLPTLQAATSGLTQSDPRRQLVETLTRWDGINLLNDDGK 600
             **************:*********.*: **:.*::.*******.:*: *** **:*****

KcPGA        TYQQPGSAILNAWLTSMLKRTVVAAVPAPFGKWYSASGYETTQDGPTGSLNISVGAKILY 660
EcPGA        TWQQPGSAILNVWLTSMLKRTVVAAVPMPFDKWYSASGYETTQDGPTGSLNISVGAKILY 660
             *:*********.************** **.*****************************

KcPGA        EALQGDKSPIPQAVDLFGGKPQQEVILAALDDAWQTLSKRYGNDVTGWKTPAMALTFRAN 720
EcPGA        EAVQGDKSPIPQAVDLFAGKPQQEVVLAALEDTWETLSKRYGNNVSNWKTPAMALTFRAN 720
             **:**************.*******:*****:*.*:.*********::.:*************

KcPGA        NFFGVPQAAAKEARHQAEYQNRGTENDMIVFSPTSGNRPVLAWDVVAPGQSGFIAPDGKA 780
EcPGA        NFFGVPQAAAEETRHQAEYQNRGTENDMIVFSPTTSDRPVLAWDVVAPGQSGFIAPDGTV 780
             **********:*::*********************:.:.:*******************..

KcPGA        DKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQEVLQVQRLEHHHHHH 828
EcPGA        DKHYEDQLKMYENFGRKSLWLTKQDVEAHKESQEVLHVQR-------- 820
             ****:*******.********* ***: ********:***
```

Figure 4.11: Amino acid sequence of the mutant PGA precursors of *K. citrophila* and *E.coli*. Residues of the signal peptide are not included in the amino acid alignment. Mutated residue in each protein, to generate corresponding slow processing precursor enzyme, is shown underlined.

4.2.2. Steady State Fluorescence Studies of Wild-type processed *Kc*PGA

Wild type *Kc*PGA is purified as heterodimer of ~23.6 kDa and ~63 kDa correspond to α- and β-chains, respectively (**Section 3.6**). Purified protein was found to be fairly stable up to 50 °C (**Section 3.7.3**) and in the broad range of pH from pH 5.0 to 10.0 (**Section 3.7.4**). The conformational stability of this protein in various conditions has been studied using fluorescence spectroscopy. Protein fluorescence is chiefly due to tryptophan residues. Accessibility of tryptophan residue of the processed protein using various quenchers has also been studied.

Sequence of the wild type processed *Kc*PGA showed the presence of 26 Tryptophan residues in *Kc*PGA (**Figure 4.12**). The steady state fluorescence spectrum had a λ_{max} of 346 nm. The λ_{max} showed a red shift (361nm) on treatment with 6M Gdn-HCl indicating exposure of tryptophans to the solvent by opening up the structure.

Chain A
ASPPTEVKIVRDEYGMPHIYADDTYRLFYGYGYVVAQDRLFQMEMARRST
QGTVSEVLGKAFVSFDKDIRQNYWPDSIRAQIASLSAEDKSILQGYADGMN
AWIDKVNASPDKLLPQQFSTFGFKPKHWEPFDVAMIFVGTMANRFSDSTSEI
DNLALLTALKDKYGKQQGMAVFNQLKWLVNPSAPTTIAARESAYPLKFDLQ
NTQTA

Chain B
SNMWVIGKNKAQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTP
FAYPGLVFGHNGTISWGSTAGFGDDVDIFAEKLSAEKPGYYQHNGEWVKML
SRKETIAVKDGQPETFTVWRTLHGNVIKTDTATQTAYAKARAWDGKEVASL
LAWTHQMKAKNWPEWTQQAAKQALTINWYYADVNGNIGYVHTGAYPDR
QPGHDPRLPVPGTGKWDWKGLLSFDLNPKVYNPQSGYIANWNNSPQKDYP
ASDLFAFLWGGADRVTEIDTILDKQPRFTADQAWDVIRQTSRRDLNLRLFLP
ALKDATANLAENDPRRQLVDKLASWDGENLVNDDGKTYQQPGSAILNAWL
TSMLKRTVVAAVPAPFGKWYSASGYETTQDGPTGSLNISVGAKILYEALQGD
KSPIPQAVDLFGGKPQQEVILAALDDAWQTLSKRYGNDVTGWKTPAMALTF
RANNFFGVPQAAAKEARHQAEYQNRGTENDMIVFSPTSGNRPVLAWDVVA
PGQSGFIAPDGKADKHYDDQLKMYESFGRKSLWLTPQDVDEHKESQEVLQ
VQRLEHHHHHH

Figure 4.12: **Amino acid sequence of processed *Kc*PGA containing histidine tag cloned and purified from *E.coli* pLysS cells. The 26 tryptophan residues present are coloured in red.**

4.2.2.1. Thermal inactivation of the processed form of *Kc*PGA

Effects of temperature on the conformation and stability of the processed form of *Kc*PGA at pH 7.5 were studied by monitoring the changes in intrinsic fluorescence, ANS binding and Dynamic Light Scattering. Steady-state intrinsic fluorescence measurement of *Kc*PGA was performed as described in Material and Methods (Section **2.3.6.9**).

Fluorescence intensity of the native protein gradually decreased with increase in temperature in the range 30-55 °C at near neutral pH of 7.5, but no change in λ_{max} (346 nm) was observed (**Figure 4.13A**). However, at 60 °C the fluorescence intensity increased slightly, accompanied with a red shift in λ_{max}, presumably because of partial unfolding of the protein. However, the fluorescence intensity started dropping beyond 60 °C along with further red shift in λ_{max} to 354 nm, may be the result of exposure of Trp residues to solvent (**Figure 4.13B**). The decrease in fluorescence intensity must be due to the deactivation of the singlet-excited state by non-radiative processes that are in competition with fluorescence. Marginal increase in Trp emission at 60 °C indicates change in conformation of the protein. Lowering of temperature from 90 °C to 25 °C did not restore fluorescence intensity or λ_{max} to their original values. This is an indication of the irreversible nature of thermal denaturation. Protein seemed to aggregate close to 60 °C as monitored by Rayleigh light scattering. The experimental setup for Rayleigh light scattering was same as used in fluorescence studies and as described in Section **2.3.6.9.3**. Scattering was recorded for 30 seconds. Substantial change in the scattering intensity was observed beyond 55 °C, which confirmed the formation of insoluble aggregates at higher temperature (**Figure 4.13C**).

The hydrophobic dye ANS is widely used as a tool to identify and characterize partly folded states of proteins, including molten globule states. ANS binds to the hydrophobic regions of proteins with an increase in the fluorescence intensity and a blue shift of λmax, useful in probing exposed hydrophobic surface of molecule. When ANS binds to the native protein, shows a blue shift in the λmax from ~520nm towards 480 nm and an increase in the fluorescence intensity. Compared to ANS binding to native protein at 30 °C, there was almost three times increase in ANS binding at 60 °C accompanied with a blue shift of λ_{max} from 514nm to 484nm (**Figure 4.13 D**). ANS

binding was found to decrease beyond 60 °C as the temperature was further increased. Sudden increase in the ANS binding at 60 °C indicated partial unfolding and exposure of hydrophobic patches, which showed reduction with further increase in temperature. This could be due to shielding of the hydrophobic patches presumably an indication of the aggregation of protein.
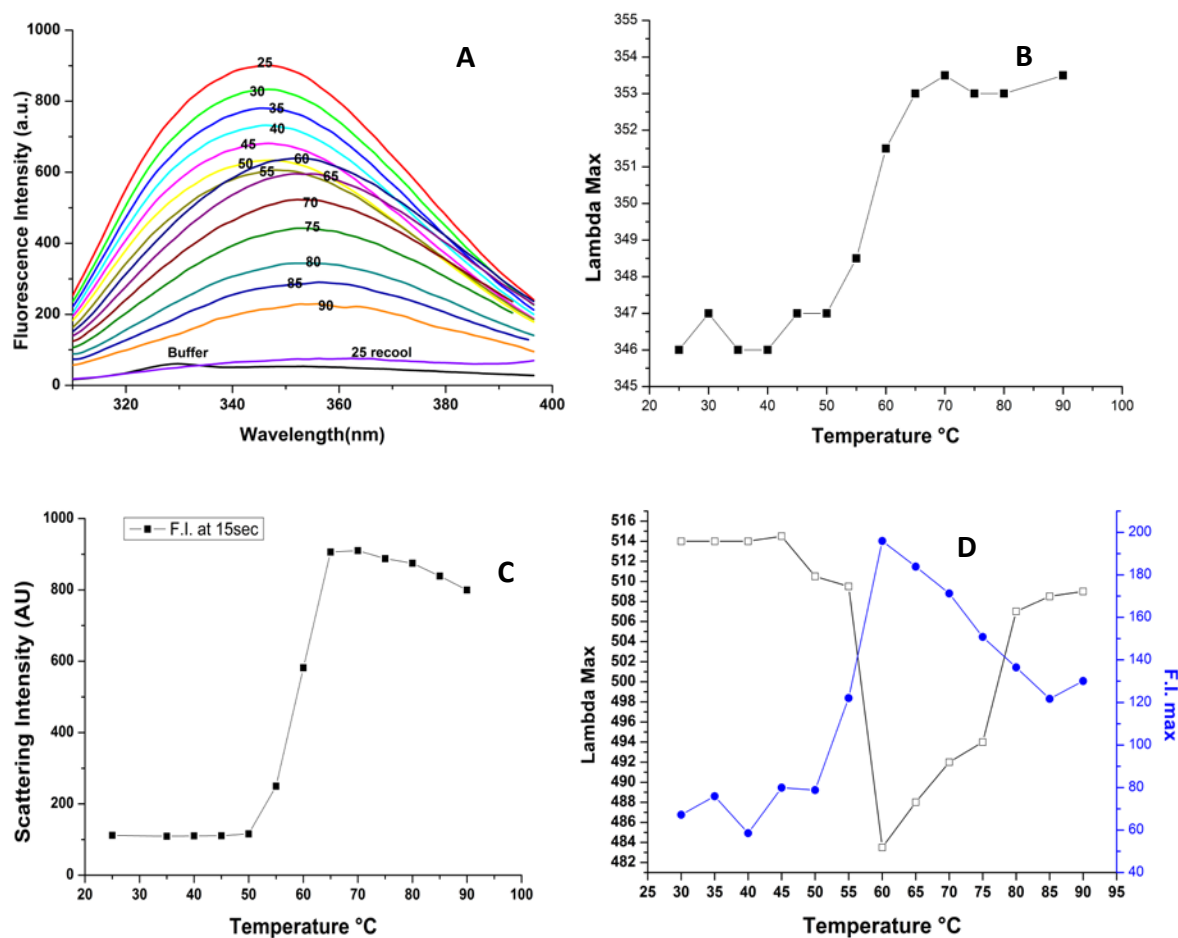
Figure 4.13: Monitoring the temperature stability of processed *Kc*PGA.

- A. **Fluorescence emission spectra of 23µg/ml (0.26µM) *Kc*PGA incubated at various temperatures for 10 minutes. The temperatures are indicated on the spectra.**

- B. **The change in λ_{max} of *Kc*PGA (23µg/ml) as a function of temperature**

- C. **Change in scattering intensity of *Kc*PGA (23µg/ml) with increase in temperature**

- D. **ANS fluorescence and binding at various temperatures (□) – F. I., (●) - λ_{max}.**

4.2.2.2. Effect of pH on stability of *Kc*PGA

Conformational stability of the processed form of *Kc*PGA at various pH was studied by measuring tryptophan fluorescence. Studies on the effect of pH variation on enzyme catalysis revealed that the enzyme was active in a broad range of pH i.e., pH 5.0 to 10.0 (**Section 3.7.4**). The fluorescence scans at different pH showed that the intensity increased with increase in pH till pH 10.0 and then started dropping beyond pH 10.0 without any change in the λ_{max} (**Figure 4.14A**). At extreme acidic and alkaline pH, the fluorescence emission maxima at 346 nm reduced considerably (**Figure 4.14A**). At pH 2.0 the red shift in λ_{max} was to 355 nm, whereas at pH 11.0 and pH 12.0 red shifts were to 348 nm and to 351 nm, respectively. This could be an indication of partial unfolding of the protein. Surprisingly, at pH 1.0, λ_{max} was seen to shift back to 350 nm. In strong alkaline conditions, deprotonation of amino group of amino acid residues occur. The ε-amino group of lysine, guanidino group of arginine and phenol group of tyrosine are known to quench the fluorescence of Trp residues in their deprotonated form (Bushueva *et al.,* 1975) which may explain the decrease in intensity at extreme alkaline pH. Deprotonation might induce charge destabilization in the local environment by inducing unfavourable changes in the native state like electrostatic repulsions, breakage of salt bridges or formation of isolated buried charges, leading to loss of function. The loss of PGA activity above pH 10.0 could also be due to unfolding of the protein. This unfolding at extremes of pH was further characterized by changes in hydrophobic dye binding.

Since the fluorescence intensity was very high and beyond the limit of recording, the protein concentration was reduced to one third for recording the ANS spectrum. ANS binding was observed only at extreme acidic pH; maximum at pH 2.0 with a blue shift in λmax from 520 to 482 nm and ANS fluorescence intensity increased seven times. At pH 1.0 there was six times increase in the ANS fluorescence intensity with λmax at 493 nm indicating possible rearrangement of structure as seen in intrinsic fluorescence. At pH 3.0 the ANS binding was only twice and λ_{max} 484 nm. The ANS binding to *Kc*PGA indicated the exposure of hydrophobic patches at extreme acidic pH. ANS fluorescence did not change in neutral and alkaline pH. No protein aggregation was detected in Rayleigh light scattering studies on excitation at 400 nm at different pH, maintaining the temperature at 30 °C.
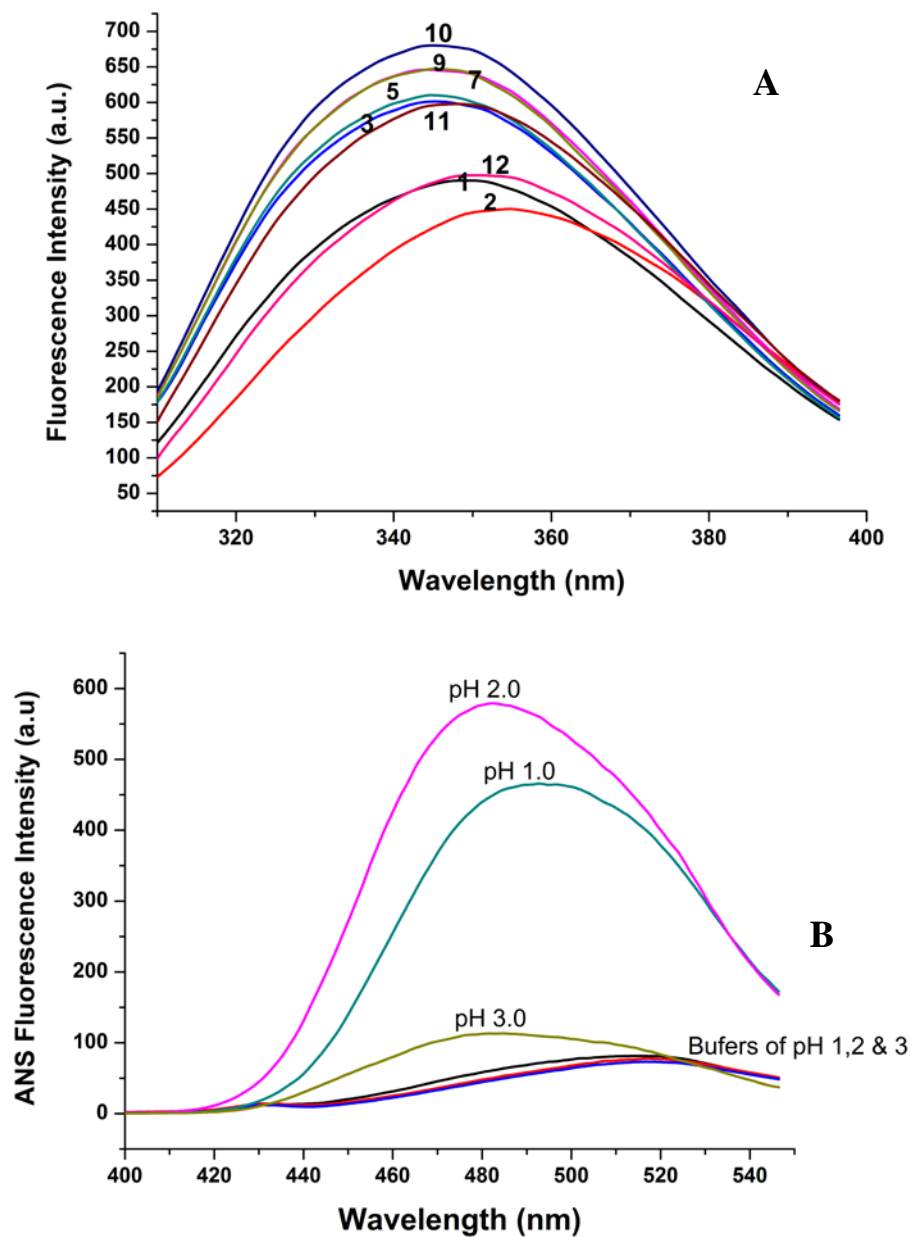
Figure 4.14: The pH dependent conformational changes of *Kc*PGA. The numbers marked on the spectra indicate pH of the sample.

    A. Fluorescence emission spectra of *Kc*PGA (18μg/ml) incubated at various pH

    B. Spectra indicating the binding of ANS to *Kc*PGA (5.5μg/ml) at different pH.

4.2.2.3. Solute Quenching Studies

The fluorescence characteristics of tryptophan residues depend strongly on the microenvironment and thus provide a sensitive parameter to investigate the conformational state of the protein. Fluorescence quenching of the proteins in solutions of quenchers have been widely used to study the degree of exposure and electronic environment of aromatic amino acid residues. Acrylamide is an efficient quencher of tryptophan fluorescence and can distinguish between buried and exposed side chains. In contrast, KI and CsCl are highly hydrated and charged molecules, whose quenching ability are limited to surface exposed tryptophans and also depends upon the neighbouring charged groups.

Analysis of the quenching data was carried out by using the Stern–Volmer equation) as well as by the modified Stern–Volmer equation described in Section **2.5.6.9.1**.

4.2.2.3.1. Quenching with Acrylamide

The fluorescence quenching profile, Stern–Volmer and modified Stern-Volmer plots, for quenching with the acrylamide of processed *Kc*PGA fluorescence at pH 7.5 are shown in **Figure. 4.15.** Under native condition, quenching of *Kc*PGA with acrylamide gave a linear Stern-Volmer plot with K_{sv} value as 4.58 M^{-1} (**Figure 4.15C**). Linear Stern-Volmer plot shows the fluorescence quenching by acrylamide is only collisional-type. 81% of the total fluorescence of *Kc*PGA was accessible to acrylamide under native condition as indicated by f_a of value 0.81 that gets increased to 1.1 on denaturation of the protein with 6M Gdn-HCl (**Figure 4.15 E and F**). Increase in accessibility of acrylamide showed that tryptophan residues were differentially exposed to the solvent and not fully accessible to the neutral quencher in the native condition but got fully exposed to solvent and consequent 100% quenching under denaturing condition (**Figure 4.15 E & F**). The fluorescence quenching of 6M Gdn-HCl treated protein by acrylamide showed an upward curvature of Stern-Volmer plot indicating static and collisional quenching, presumably as a result of altered conformations (**Figure 4.15D**).

Figure 4.15: Stern-Volmer plots for quenching of tryptophan fluorescence by acrylamide.
(A) & (B) Fluorescence emission spectra of *Kc*PGA (17-18 µg/ml) in native condition and
with 6 M Gdn-HCl, respectively in the presence and absence of quencher (acrylamide) are
shown. The quencher concentration was increased from 0 to 0.5 M, in the direction
indicated by the arrow. (C) & (D) Stern-Volmer plots for *Kc*PGA (17-18µg/ml) in the
presence of acrylamide in native condition and with 6 M Gdn-HCl, respectively. (E) & (F)
Modified Stern Volmer plots for *Kc*PGA (17-18 µg/ml) in acrylamide in native condition
and with 6M Gdn-HCl, respectively.

4.2.2.3.2. Quenching with Potassium Iodide (KI)

The fluorescence quenching profile, Stern–Volmer and modified Stern-Volmer plots for quenching with the KI for processed *Kc*PGA at pH 7.5 are shown in **Figure. 4.16.** The quenching profile of the native protein with KI showed a negative curvature (**Figure 4.16C**) indicating that certain tryptophans are selectively quenched before others. At low concentration of the quencher, the slope of the Stern–Volmer plots reflected largely the quenching of the more accessible residues. At higher concentrations, the easily quenched fluorescence depleted, and those tryptophans having lower quenching constants became dominant. Similar quenching patterns were reported for several multi-tryptophan proteins (Lehrer, 1971; Teale & Badley, 1970; Sultan & Swamy, 2005). From the slopes of the two linear components of the Stern–Volmer plots, collisional quenching constants, K_{sv1} and K_{sv2} were obtained for I^- and are listed in **Table 4.2**. K_{sv1} being greater than the K_{sv2} is a substantial evidence for selective tryptophan quenching by these quenchers. Linear modified Stern-Volmer plot under native condition estimates the fractional accessibility (f_a) for tryptophans by 0.45 indicating that 45% of the total tryptophans were in an electropositive environment (**Figure 4.16E**).

KI showed the linear Stern–Volmer plots for the denatured protein (**Figure 4.16D**) and increased quenching with I^-, making it clear that upon denaturation, even buried tryptophans become exposed to solvent and are more accessible for quenching. 100% of the quenching was seen in the case of KI under denatured condition (**Figure 4.16F**).
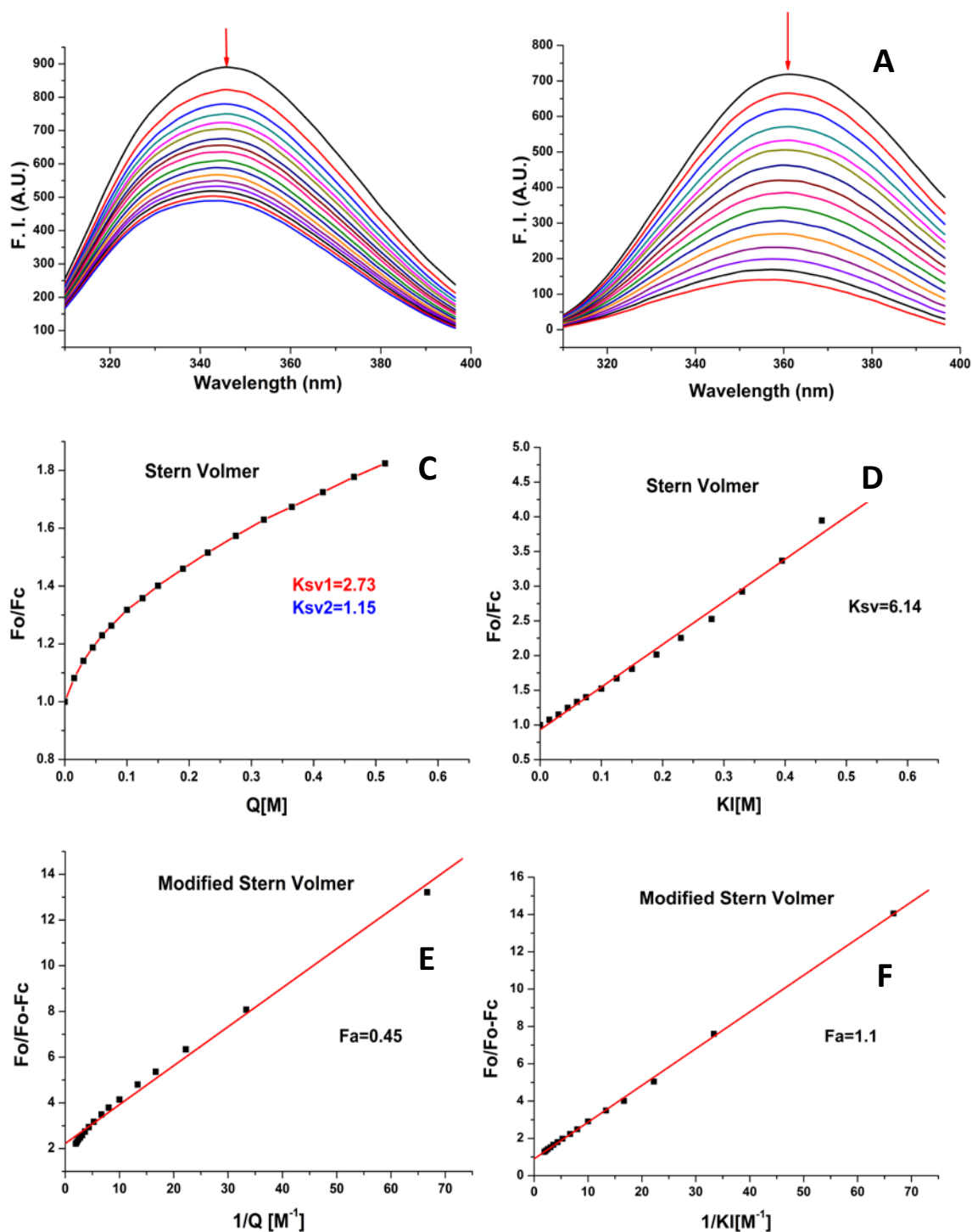
Figure 4.16: Stern-Volmer plots for quenching of tryptophanyl fluorescence by KI. (A) & (B) Fluorescence emission spectra of *Kc*PGA (17-18μg/ml) in native condition and with 6 M Gdn-HCl, respectively in the presence and absence of quencher (KI). The quencher concentration for spectra is increased from 0 to 0.5M, in the direction indicated by the arrow. (C) & (D) Stern-Volmer plots for *Kc*PGA (17-18μg/ml) with KI in native condition and with 6 M Gdn-HCl, respectively. (E) & (F) Modified Stern Volmer plots for *Kc*PGA (17-18μg/ml) with KI in native condition and with 6 M Gdn-HCl, respectively.

4.2.2.3.3. Quenching with Cesium Chloride (CsCl)

The fluorescence quenching profile, Stern–Volmer and modified Stern-Volmer plots, for quenching with CsCl of *Kc*PGA at pH 7.5 are shown in **Figure. 4.17.** The quencher CsCl, which cannot penetrate inside protein matrix but can access only surface exposed tryptophans, is found to quench *Kc*PGA with lowest efficiency indicating majority of the tryptophans are in an electropositive environment. The quenching profile of the native protein with CsCl also showed a negative curvature (**Figure 4.17C**) pointing to the possibility of certain tryptophans getting selectively quenched before others, already seen in the case of KI. Higher value of K_{sv1} ($0.44M^{-1}$) compared to K_{sv2} ($0.15M^{-1}$) showed selective tryptophan quenching by these quenchers. Modified Stern-Volmer plot for CsCl showed a negative curvature, slopes of which provided two values of F_a as 0.15 and 0.07(**Figure 4.17 E**).

Denaturation of *Kc*PGA slightly improved the accessibility to Cs^+. 54% of the total tryptophans were found to be quenched by Cs^+ after incubating with 6 M Gdn-HCl. Denatured protein showed linear plots for both Stern-Volmer as well as modified Stern-Volmer plots (**Figure 4.17 D and F**).
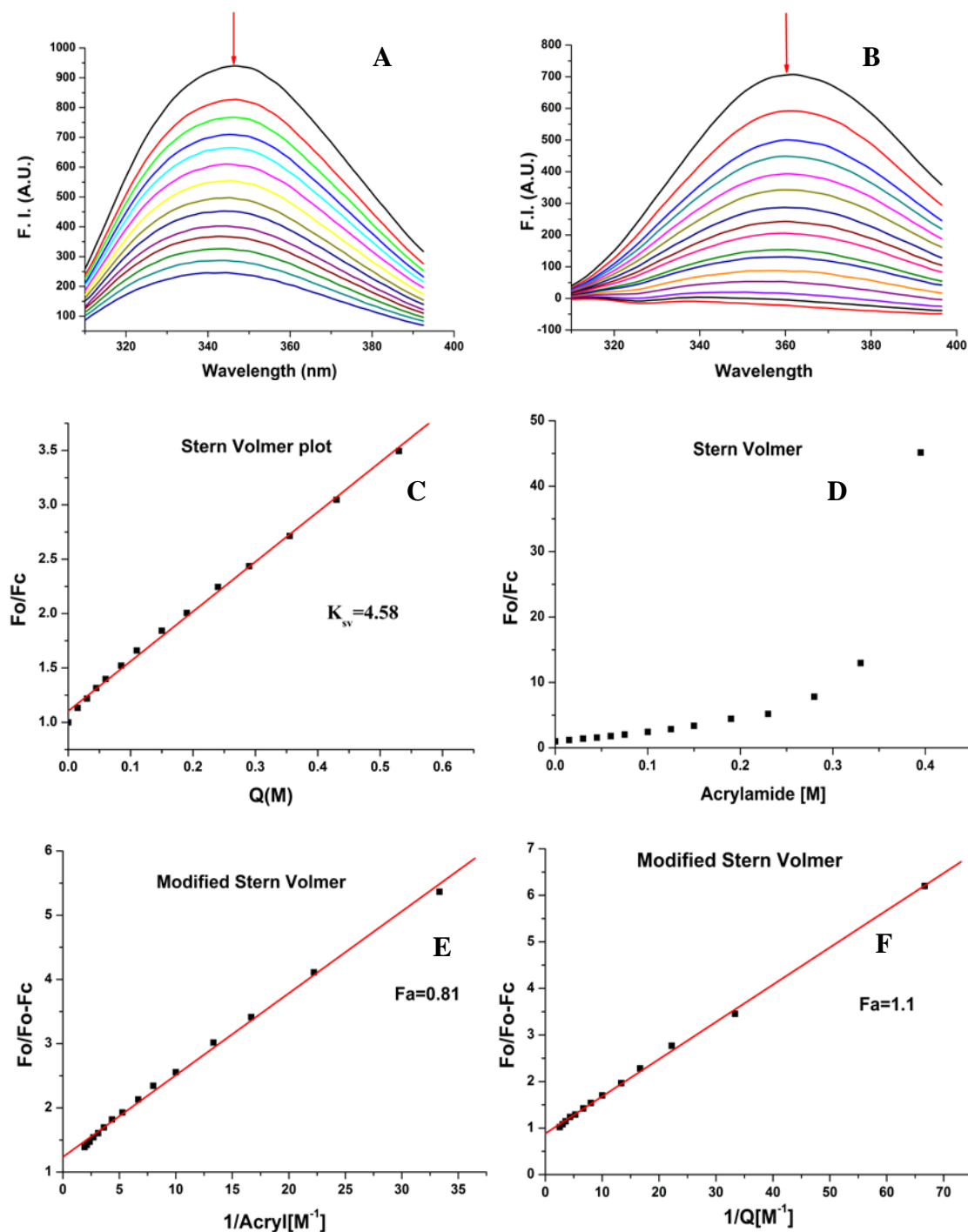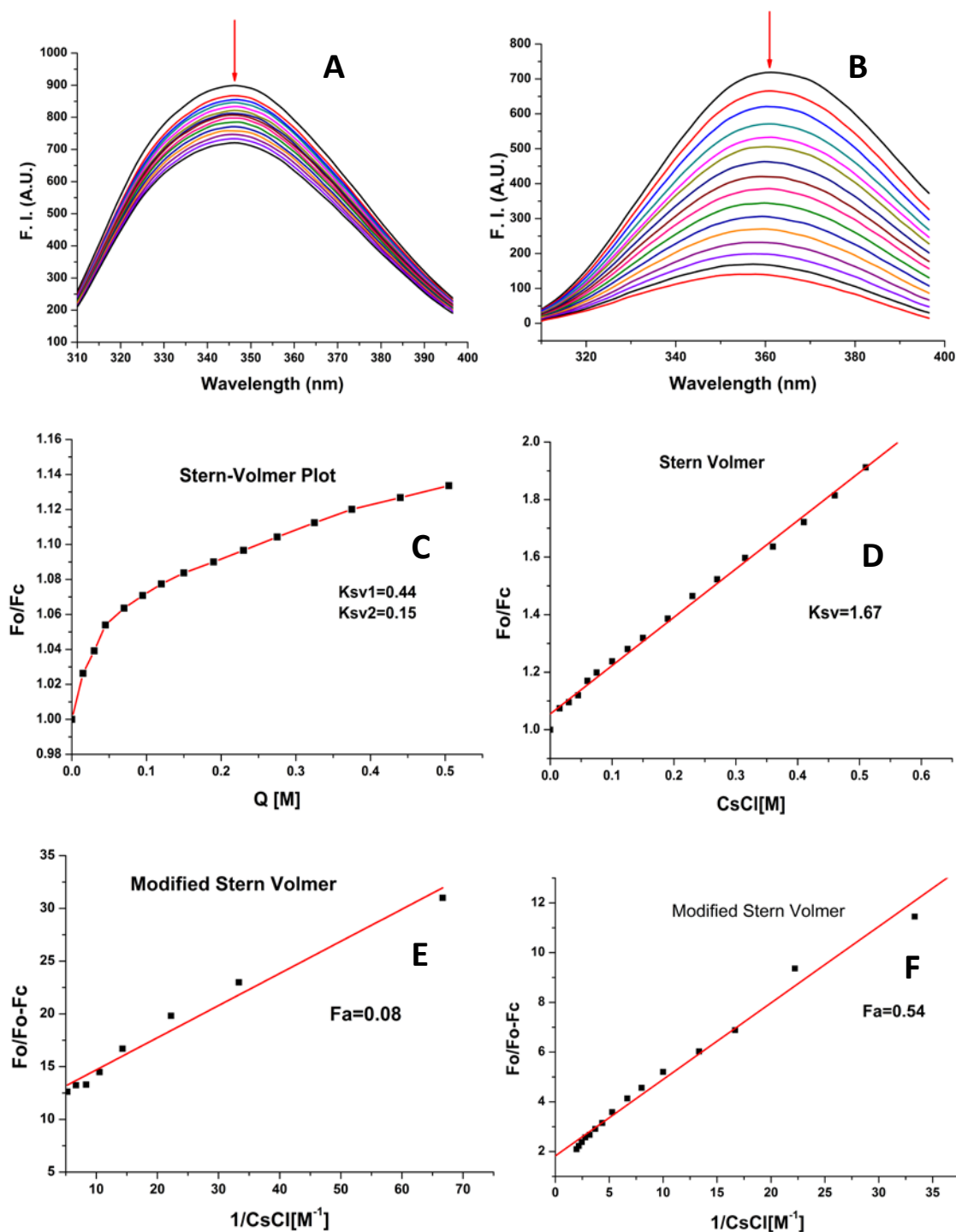
Figure 4.17: Stern-Volmer plots for quenching of tryptophanyl fluorescence by CsCl. (A) & (B) Fluorescence emission spectra of *Kc*PGA (17-18µg/ml) in native condition and with 6 M Gdn-HCl, respectively in the presence and absence of quencher (CsCl). The quencher concentration for spectra is increased from 0 to 0.5 M, in the direction indicated by the arrow. (C) & (D) Stern-Volmer plots for *Kc*PGA (17-18µg/ml) with CsCl in native condition and with 6 M Gdn-HCl, respectively. (E) & (F) Modified Stern-Volmer plots for *Kc*PGA (17-18µg/ml) with CsCl in native condition and with 6 M Gdn-HCl, respectively.

Table 4.2: Summary of parameters obtained from Stern-Volmer and modified Stern-Volmer plot analysis in the solute quenching studies of *Kc*PGA. nd - not determined

Type of Quencher and condition	K_{sv1} (M^{-1})	K_{sv2} (M^{-1})	f_a
Acrylamide			
Native	4.5771		0.81
Denatured (6M Gdn-HCl)	nd	nd	1.1
KI			
Native	2.733	1.15	0.45
Denatured (6M Gdn-HCl)	6.14		1.1
CsCl			
Native	0.44	0.15	0.15, 0.07
Denatured (6M Gdn-HCl)	1.67		0.54

4.3. DISCUSSION

Only in the early 1990s the protein editing events in prokaryotes was discovered (Thony-Meyer *et al.,* 1992). Specific peptidases was initially presumed to be performing the steps required for processing and activation of precursor proteins, but recently discovered mechanisms of protein splicing in many organisms, from prokaryotes to eukaryotes, has revealed a range of autocatalytic processing events (Perler *et al*., 1997). The autocatalytic activation of Ntn-hydrolases represents an interesting phenomenon. The residues that are important for activation have been identified from precursor structures of different Ntn hydrolases mutants (Ditzel et *al*., 1998; Xu *et al*., 1999; Hewitt *et al*., 2000; Kim *et al*., 2002, 2003; Yoon *et al*., 2004). There have been reports of many mutants that either abolishes or slows down the rate of autocatalytic processing in PGAs from *E.coli* (Kasche *et al*., 1999; Choi *et al*., 1992; Brannigan *et al*., 2000; McVey, 1999; Lee *et al*., 2000), *Alcaligenes faecalis* (Kasche *et al*., 2003) and *Kluyvera citrophila* (Prieto *et al*., 1992). *E. coli* PGA precursor was crystallized by Hewitt *et al*. (2000) where they reported the cleavage between Tyr260-Pro261 in the slow processing mutant (Thr263Gly). This mutation was unexpected as it was postulated that Thr263-Ser264 cleavage is the first step of the autocatalytic processing mechanism in the PGAs. Lee *et al.* (2000) proposed the possible role of another residue Lys299 in the autocatalytic processing of the PGAs along with the involvement of well known Ser264. Thus, the exact mechanism of the processing is still not entirely clearly defined in the case of PGAs. The understanding of mechanism of processing in Ntn-hydrolases cannot be complete without structural information on the first cleavage intermediates from different Ntn hydrolases.

Here, we tried to understand the autocatalytic mechanism of PGA from *Kluyvera citrophila*, which is expected to be similar to PGA from *E.coli* owing to their high percentage of similarity in the amino acid composition. Unlike, Ser264Gly mutation in *Ec*PGA that resulted in the expression of stable processing deficient single chain precursor PGA (Choi *et al*., 1992), Ser264Gly mutation in *Kluyvera citrophila* resulted in the expression of slow processing mutant. Enzyme crystallized in two different space groups of P1 and C2 and diffracted till 2.5 Å and 3.5 Å resolution respectively. Refinement of the 2.5 Å triclinic structure showed the good similarity with *E.coli* precursor structure (PDB code: 1E3A, Hewitt *et al.,* 2000) with r.m.s.

deviation of 0.62 Å and presence of intact space peptide and Ca^{2+} present at the similar position in the difference map calculated. Looking at the similarity in the structures and in the amino acid composition of both the precursor, an intramolecular autocatalytic processing following the same steps of cleavage might be the mechanism of activation of *Kc*PGA as well. However, the poor completion and resolution of the diffraction dataset was a limitation to comment on the exact point of cleavage.

The enzyme (*Kc*PGA) in the processed form has received more attention as compared to enzyme from *E.coli* (*Ec*PGA) due to its numerous industrial process friendly properties, namely increased resilience to harsh conditions and ease of immobilization (Alvaro *et al*., 1992; Fernandez-Lafuente *et al*., 1991, 1996; Liu *et al*., 2006). Amino acid composition showed the wild type processed PGA from *K.citrophila* as a multi-tryptophan protein. Measuring the tryptophan fluorescence intensity under various conditions is helpful in determining the effect of these conditions on the conformational stability of the protein. Fluorescence studies involving these multi tryptophans revealed that the enzyme is fairly stable till 55 °C at near neutral pH of 7.5 as there is no shift in the λ_{max} but tends to aggregate beyond 55 °C as revealed from light scattering and ANS binding studies.

ANS fluorescence is highly sensitive to hydrophobicity of the chemical environment and has been widely used to probe apolar binding sites in proteins. The maximum ANS binding was seen till 60 °C and thereafter protein starts to lose the hydrophobic patches presumably due to aggregation of the protein. Temperature mediated change is found to be irreversible for the processed *Kc*PGA. Processed *Kc*PGA enzyme was found to be stable in the broad pH range as studied by intrinsic fluorescence studies and ANS binding under various pH conditions as expected for this enzyme.

The fluorescence quenching profile for the processed enzyme was also studied with the various quenchers. Fluorescence spectra of the native and denatured *Kc*PGA recorded in the absence and in the presence of increasing concentrations of the quenchers displayed significant increase in the extent of quenching with denaturation, showing that the accessibility to all tryptophan residues is possible only on complete unfolding.

Under native conditions, among the three quenchers used, acrylamide was the most efficient one as it could penetrate into the interior of the protein. The Stern-Volmer plots obtained for quenching of the protein fluorescence with acrylamide was linear while that for cesium and iodide ions are non linear showing selective quenching by these ions. The lower quenching with charged quenchers (I^- and Cs^+) indicated most of the fluorescent tryptophan residues in the protein to be buried in the hydrophobic core of the protein. Additionally, lowest quenching observed with Cs^+ appears to be due to positively charged environment of the exposed tryptophan residues.

Since slow processing precursor enzyme mutant does not contain extra tryptophan in spacer region and because of its continuous processing *in vitro*, fluorescence studies on the precursor have not been attempted.

Chapter: 5

Structural studies of PGA from *Alcaligenes faecalis*

5.1. INTRODUCTION

$\mathbb{P}$GA from *Alcaligenes faecalis* (A*f*PGA) has received attention in the literature because of its higher thermostability and synthetic efficiency in enantioselective synthesis. This enzyme showed the high synthetic efficiency with ratio of the second-order rate constants for enzymatic hydrolysis of L- and D-enantiomers of N-phenylacetylarginine is more than 3,000 in enantioselective synthesis. The enzyme showed the highest specificity constant (k_{cat}/K_m) for hydrolysis of penicillin G compared to homologous enzymes from other bacterial sources. K_{cat}/K_m value of colorimetric substrate (NIPAB) hydrolysis was shown to be 42 and 27 times higher than the corresponding values for *E. coli* and *K. citrophila* enzymes, respectively (Svedas *et al*., 1997). The first cloning, sequencing and characterization of PGA from *Alcaligenes faecalis* was reported by Verhaert *et al*. in 1997. The sequence identity between A*f*PGA and the penicillin acylase from *E. coli* (*Ec*PGA) is 47%, and the residue similarity is even higher (67%). A*f*PGA showed clear advantage over other well-characterized PAs for industrial applications because of its higher thermostability, ascribed to the presence of two cysteine residues (absent in the *E. coli* enzyme), which form a disulfide bridge (Verhaert *et al*., 1997). The PGA from *E. coli* (*Ec*PGA) is reported to irreversibly lose more than 50% of its activity on incubation at 50 °C for 20 min, whereas A*f*PGA is not affected by the same treatment (Verhaert *et al*., 1997). The disulfide bridge also increased the conformational stability in wild-type in the temperature- induced unfolding (Yuryev *et al*., 2010).

The presence of a disulfide bond involving two cysteine residues at positions β492 and β525 in A*f*PGA enzyme was accounted on the basis of the following observations: (i) a comparison with the known 3D structure of the *E. coli* enzyme indicates that the two cysteine residues are in close proximity, (ii) the β-subunit of A*f*PGA shows an electrophoretic mobility that depends on its oxidation state, (iii) the A*f*PGA enzyme is significantly more thermostable in its oxidized state than reduced state, and (iv) A*f*PGA is more stable than the *E. coli* enzyme, which lack the disulfide bond (Verhaert *et al*., 1997). These four arguments led us to conclude that a disulfide bond could be one of the factors contributing for its higher stability compared to other PGAs.

A. faecalis PGA is formed *in vivo* via a complex post-translational processing from an inactive precursor to the catalytically active form, which also consists of α- and β-subunits. Interestingly the *A. faecalis* pro-enzyme is however transported through the membrane by the Sec-translocation system not by tat-pathway as proposed for other PAs (Ignatova, 2000, PhD Thesis). The processing undergoes through a multi-step excision of the small spacer peptide of 37 amino acids, which is located between Gln229 and Ala265 and much shorter than spacer present in *E.coli* and *K.citrophila* PGAs (numbering starts from the first amino acid of pro-PA) (Ignatova *et al.,* 1998). Intriguingly, this spacer peptide or parts of it were found to activate mature *A. faecalis* PGA when added in vitro (Kasche *et al*., 2003). Two subsequent mutations in the spacer peptide (T206G and S213G) markedly slow down the processing, so that the linker remains attached to the α-subunit.

Enzyme was shown to be irreversibly inactivated with micromolar concentration of PMSF within few minutes (Svedas *et al.,* 1997). Svedas and co-workers hypothesized, on the basis of kinetic studies that the active centres of the enzymes *Af*PGA and *Ec*PGA differ in the details of their organization, especially in the subsite responsible for the interaction with the leaving group of the substrate. However, the lack of crystal structure for this enzyme failed to support this hypothesis and the presence of the disulfide bridge. Structural interpretation and prediction of substrate selectivity and stereospecificity of *Af*PGA was done using homology modelling and docking studies (Braiuca *et al*., 2003). Substitution of the Pheβ71 found in *Ec*PGA for a Pro residue in *Af*PGA in aminic subsite enables the aminic subsite to accommodate bulkier nucleophiles.

Thermostability along with high stereo selectivity and highest specificity constant makes this enzyme a more attractive biocatalyst for both synthesis and modification of β-lactam antibiotics and industrial interest is demonstrated by the patent deposited in 1995 (De Vroom & der Mey, 1995). Since the rate of an enzymatic reaction increases generally with temperature, the application of an enzyme with enhanced thermal stability will improve the efficiency of many industrial processes. Since the cost of producing 6-APA has a direct impact on profitability of semi-synthetic antibiotic production, having an enzyme with enhanced property will always bring economic benefits. In industrial processes at elevated temperature, the

thermostable enzyme may provide some advantages like- higher reaction rate, a higher solubility of substrate, a lower viscosity of reaction medium and reduced risk of microbial contamination. Such characteristics make *Af*PGA a more attractive biocatalyst for both hydrolysis and synthetic conversions.

Approaches comparing the amino acid sequences and crystal structures of homologous mesophilic and thermophilic proteins have been used to determine structural features contributing to the thermostability of many proteins, and this knowledge will be essential for the rational design of highly thermostable protein using protein engineering techniques.

This chapter describes the crystallization, structure determination, refinement and analysis carried out on *Af*PGA in two crystal forms, which grow depending on the amount of detergent present in the crystallization condition. Although the resolution was modest we could identify features like the presence of a disulfide bridge in the structure that would be imparting comparatively high stability to *Af*PGA and the presence of a calcium binding site. Some of the other features presumed to be responsible for the thermostability of *Af*PGA were also studied by comparing the amino acid sequence and the three-dimensional structure of *Af*PGA with two other PGAs whose three-dimensional structures are available.

5.2. RESULTS

5.2.1. Crystallization of *Af*PGA

*Af*PGA shares 49% DNA-sequence identity and 39% protein sequence identity with the more elaborately studied *E. coli* penicillin acylase in the penicillin amidase family from the Gram-negative bacteria. Penicillin G acylase from *A. faecalis* (*Af*PGA) has two subunits and is expressed as a 92 kDa precursor polypeptide consisting of a

- signal peptide (26 amino acids) and

- α- and β-subunits (202 and 551 amino acids, respectively) separated by an endopeptide of 37 amino acids excised and removed during post-translational processing.

The purified protein sample was obtained from Dr. Zoya Ignatova (Department of Biochemistry, Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany). Details of the cloning, isolation and purification of *Af*PGA have been described by Ignatova *et al*. (1998) and Kasche *et al*. (2003). This sample was checked in the SDS-PAGE. The presence of β (~63 kDa) and α (~23 kDa) subunits were confirmed (**Figure 5.1**).
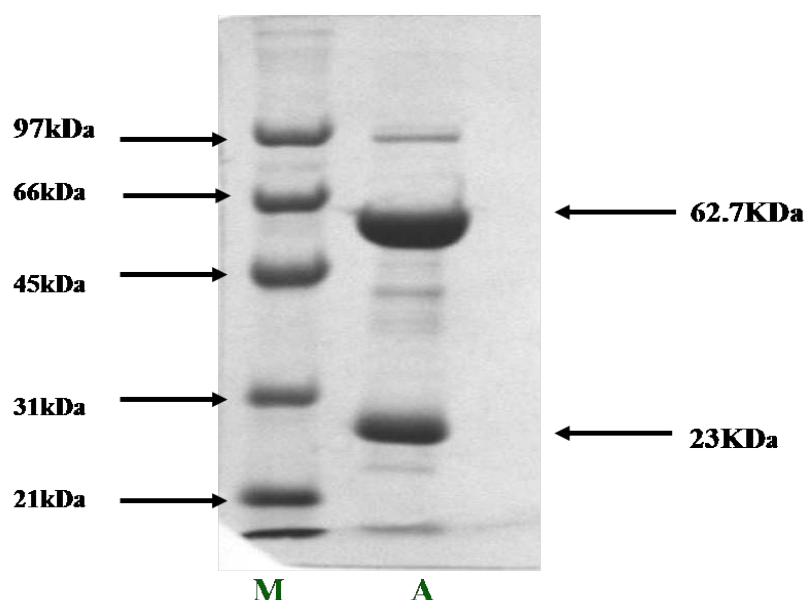
Figure 5.1: 12 % SDS-PAGE of purified *Af*PGA. The lanes contained A: *Af*PGA, M: Low Molecular Range Protein Marker

5.2.2. Protein Preparation

The protein sample was dialyzed overnight with two changes against 100 times volume of 10 mM potassium phosphate buffer pH 7.5 containing 1 mM DTT. Protein concentration was determined in accordance with the method of Lowry et al. (1951) with BSA as standard. The protein was concentrated to approximately 15 mg/mL in a centricon concentrator (Millipore) at 5,000 g and was stored at 253 K. The protein sample used for crystallization was thawed and spun at 10,000 g for 5 min before keeping for crystallization.

5.2.3. Crystallization of *Af*PGA

All crystallization experiments were carried out at 303 K using the hanging-drop vapour-diffusion method. The protein sample was thawed immediately prior to setting up crystallization and was centrifuged for 5 min at 10,000 rpm at 277 K to free the sample of particulate matter. In all experiments, 1 µl protein solution was mixed with 1 µl screen solution and was equilibrated over a reservoir containing 0.5 ml of optimized crystallization (Sureshkumar, 2006, Ph.D. Thesis) condition in 24-well plates (Axygen Biosciences, USA).

Reproducible diffraction-quality crystals were grown from optimized crystallization solution consisting of 15% (w/v) PEG 8,000, 0.1 M Tris–HCl pH 7.5 and 0.5% (w/v) β-octyl-glucopyranoside solution. The presence of 60 µl β-octyl-glucopyranoside in the well solution led to orthorhombic crystals (**Figure 5.2a**), whereas 80-100 µl β-octyl-glucopyranoside resulted in tetragonal crystals (**Figure 5.2c**). Varying the concentration of β-octyl-glucopyranoside resulted in crystals of varying morphology, but these were not better than the two forms analysed. For example, very thin plate-type crystals that were difficult to handle and that diffracted poorly were obtained in the presence of 70 µl of 0.5% (w/v) β-octyl-glucopyranoside solution (**Figure 5.2 b**).

a	**b**	**c**
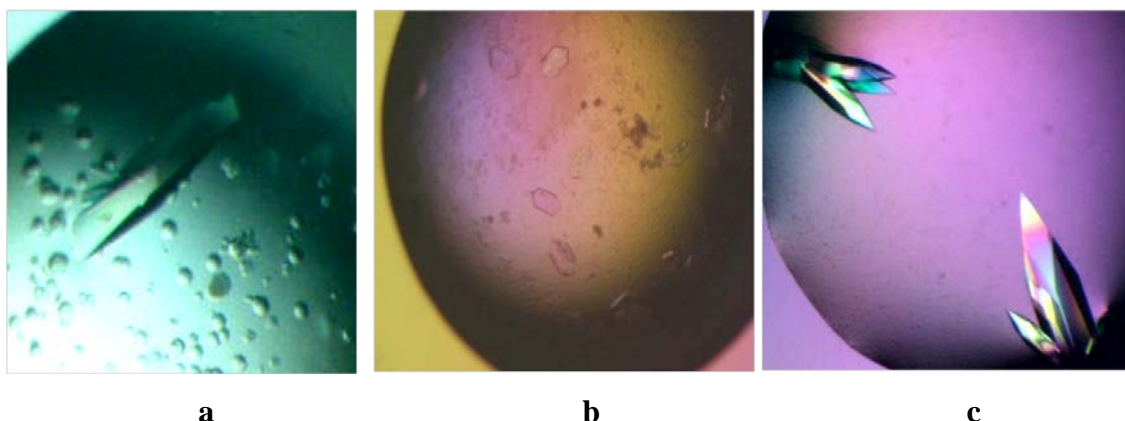
Figure 5.2: *Af*PGA **crystals obtained in the presence of β-octyl glucopyranoside; a) Orthorhombic crystals, b) plate type crystals, c) tetragonal crystals**

5.2.4. Data Collection and Processing

The collection of high-resolution data was limited by deterioration of diffraction quality on freezing the crystals under cryoconditions. Trials using a number of cryoprotectants such as PEG 400, glycerol, sucrose, PEG 8,000, ethylene glycol, 2-propanol and hexanetriol at concentrations of 20–30% in the crystallization solution, combined with strategies of slow and fast soaking in cryosolution, did not help. Soaking crystals in cryoprotectants invariably resulted in unmanageable mosaicity or in total elimination of diffraction. Therefore, data were collected on *Af*PGA crystals at room temperature (293 K) from a crystal mounted inside a sealed thin glass capillary of 1.0 or 1.5 mm in diameter on the in-house X-ray diffraction facility at NCL, Pune, India using CuKα (λ = 1.5418 Å) radiation using Raxis IV[++] area detector mounted on a Rigaku rotating anode X-ray generator operating at 50 kV, 100 mA. Typically, one single crystal was drawn along with mother liquor into the capillary. The crystal was taken out of the mother liquor and made to stick to the walls of the capillary by the surface tension of mother liquor surrounding the crystal, or otherwise the mother liquor on the crystal was dried out using thin strips of filter paper. A small amount of mother liquor was left in the capillary. The ends of the capillary were sealed with bee wax. The capillary was then mounted on a goniometer head with a small amount of plasticine. Because of one large unit cell dimension in both best diffracted orthorhombic and tetragonal crystals (260 Å and 298 Å, respectively) and not too excellent quality of the crystals we could collect data to an effective resolution of only 3.3 Å and 3.5 Å,

respectively using the laboratory X-ray source. The crystal-to-detector distance was kept at 260 mm for orthorhombic crystals and 300 mm for tetragonal crystals. The data collection control and initial Indexing were done using program *CRYSTALCLEAR* (Rigaku). Successive frames were collected using a crystal oscillation of 0.5°. Due to the fragile nature and poor diffraction quality of the crystals, data could not be collected on plate type crystals obtained with the addition of 70 µl of 0.5% β-octyl-glucopyranoside detergent in the well solution.

The orthorhombic data could be processed in either space group C222, with unit-cell parameters $a = 72.9$, $b = 86.0$, $c = 260.2$ Å, or in space group P2, with unit-cell parameters $a = c = 56.4$, $b = 260.2$ Å, $\beta = 99.4°$. Since tests for the presence of twinning were negative, the higher symmetry was chosen. These crystals diffracted to 3.3 Å resolution. Tetragonal-type crystals were indexed in space group P422, with unit-cell parameters $a = b = 85.6$, $c = 298.8$ Å, and diffracted to 3.5 Å resolution. Although diffraction extended slightly beyond these resolutions, the data could not be processed because of asymmetry of the diffraction and a sharp rise in R_{merge} values. Although several data sets were collected with varying quality, the statistics are summarized in **Table 5.1** for the sets which gave the best R_{merge} for this resolution. Assuming a molecular weight of 86 kDa for the protein and one molecule in the asymmetric unit of both crystal forms, the estimated Matthews coefficients (VM) were 2.37 $Å^3$ Da^{-1} for the orthorhombic form and 3.18 $Å^3$ Da^{-1} for the tetragonal form, which correspond to solvent contents of 48 and 61%, respectively. These values were within the normal ranges found in protein crystals (*Matthews*, 1968). The diffraction data were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL* suite (Otwinowsky & Minor, 1997).

Table 5.1: Data collection statistics of *Af*PGA crystals

Space Group	C222$_1$	P 4$_1$2$_1$2
Temperature	RT	RT
X-ray source	Cu-Kα	Cu-Kα
Wavelength	1.5418 Å	1.5418 Å
Resolution range (Å)	50-3.30 (3.42-3.3)	20-3.50 (3.56-3.50)
Unit-cell parameter (Å)	$a = 72.93$, $b = 86.01$, $c = 260.18$	$a = b = 85.56$, $c = 298.78$
Molecules per a.s.u	1	1
Matthews coefficient (Å^3 Da^{-1})	2.37	3.18
Solvent content (%)	48.13	61.31
No. of observed reflections	120675	99602
No. of unique reflections	12632	40387
Multiplicity	3.7 (3.3)	2.3 (2.0)
Average I/σ(I)	8.6 (3.8)	7.42 (3.08)
R$_{merge}$ (%) [†]	11.6 (28.8)	9.9 (23.9)
Completeness (%)	99.6 (97.9)	90.5 (95.4)

† **R$_{merge}$ = Σ$_{hkl}$Σ$_i$ |I$_j$ (hkl) - {I(hkl)}|/Σ$_{hkl}$Σ$_i$I$_i$(hkl), where I$_i$(hkl) is the intensity of an individual measurement of the reflection with Miller indices hkl and {I(hkl)} is the mean intensity of redundant measurements of that reflection.**

Numbers in the parentheses are for the last resolution shell

5.2.5. Structure determination

Trials for determination of the structure were performed using the molecular-replacement method with the structure of penicillin G acylase from E. coli in the processed form at 1.3Å resolution (PDB code: 1GK9; McVey *et al.,* 2001) as a template. Structures for both space groups were obtained using the program *PHASER* (McCoy *et al*., 2007; Storoni *et al.*, 2004) from the *CCP4* suite (Collaborative Computational Project, Number 4, 1994). The estimated values of the Z score and R$_{factor}$ were used as indicators of the choice of the correct solution. In the case of the orthorhombic form, space group C222$_1$ was chosen based on systematic absences and a lower R$_{merge}$ in diffraction data processing. Diffraction data from the orthorhombic

crystal form in the resolution range 20–3.3 Å were used in the molecular-replacement search and a unique solution with a correlation coefficient of 44.4% and an R_{factor} of 48.6% was selected.

For the tetragonal form, diffraction data in the resolution range 20–3.5 Å were used. Since the systematic absences were not clear all possible alternative space groups (P422, P42$_1$2, P4$_1$22, P4$_1$2$_1$2, P4$_2$22, P4$_2$2$_1$2, P4$_3$22 and P4$_3$2$_1$2) were tried in calculation of the translation function. The solution that corresponded to the highest Z score, in space group P4$_1$2$_1$2, was then subjected to packing-function calculations and it showed only permissible number of clashes, which in this case was fixed as zero. The correlation coefficient of the solution was 45.5% and the R_{factor} was 49.1% (Sureshkumar, 2006, PhD. Thesis).

5.2.6. Model Building and Refinement of the Structure

Initial model obtained after molecular replacement was built using Automated model building program *Buccaneer* (Cowtan, 2006), which built 660 residues in the C222$_1$ crystal and 740 residues in the P4$_1$2$_1$2 crystal. Initial models obtained after molecular replacement followed by model building for *Af*PGA crystals were refined using *REFMAC*5 (Murshudov *et al.,* 1997) from the pre-release *CCP4* v.6.1. Initial phases were improved by subsequent rigid body refinement. This was followed by several cycles of restrained refinement resulting in the improvement of both R_{factor} and R_{free}. Individual B-factors were refined in the final cycles. Electron density maps were calculated at this stage. Side chains were mutated according to the sequence of *Af*PGA and their conformations were selected based on the observed density and their correct positions were refined.

5.2.7. Structure Validation

During final cycles of refinement the stereochemistry and the geometry of the models were checked using *PROCHECK* (Laskowski *et al.,* 1993) and *SFCHECK* (Vaguine *et al*., 1999). The overall G factor considered to be a measure of stereochemical quality of the model output by *PROCHECK* was 0.13 for orthorhombic and 0.21 for tetragonal. This is within the limits expected for a correct structure that is refined at 3.3 Å and 3.5 Å resolution. The Ramachandran (φ, ψ) plot (Ramachandran &

Sasisekharan, 1968) showed that the most of the residues were placed in the most favoured region or partially allowed region of the map (**Figure 5.3 A, B**). The refinement statistics and the values of refined parameters are presented in **Table 5.2.** The R_{factor} reduced to not better than 27%, presumably due to poor quality of diffraction data. The coordinates of the final models and the merged structure factors have been deposited with the Protein Data Bank under the accession codes of 3K3W and 3ML0 for orthorhombic crystal form and tetragonal crystal form, respectively.

Table 5.2: Refinement Statistics for the *Af*PGA crystals

Crystal System/Space Group	Orthorhombic/ C222$_1$	Tetragonal/P4$_1$2$_1$2
Resolution range (Å)	50-3.30(3.42-3.30)	20-3.50(3.56-3.50)
Refinement Statistics		
R_{cryst} [a]	28.3	26.9
R_{free} [b]	31.5	28.3
Number of Proteins Atoms	6000	5995
Number of water molecules	10	0
Number of Heterogen atoms	1	1
Overall B (isotropic) from Wilson Plot ($Å^2$)	33.34	29.99
Root Mean Square Deviation		
Bond Length (Å)	0.0089	0.009
Bond Angle (°)	0.797	0.7082
Ramachandran Plot Analysis (% Residues)		
Most favoured region	87.4	87.5
Additionally allowed region	12.4	11.4
Generously allowed region	0.2	0.8
Disallowed region	0	0.3

[a] **$R_{cryst}= \Sigma$ |Fo-Fc |/ ΣFo,** [b] **$R_{free} = \Sigma$ |Fo-Fc |/ ΣFo where the F values are test set amplitudes (5 %) not used in refinement.**
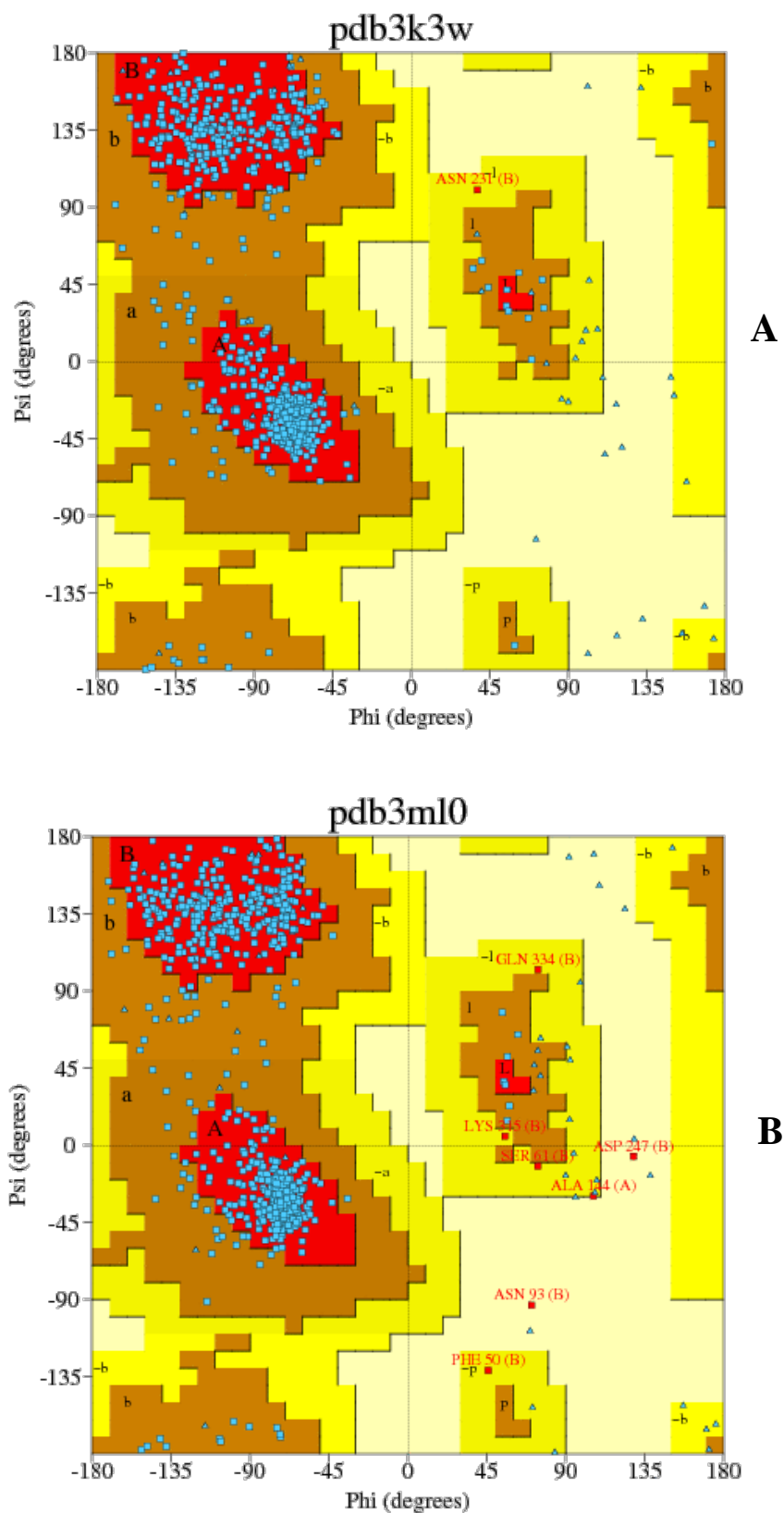
Figure 5.3: Ramachandran plot corresponding to A) Orthorhombic form B) Tetragonal form generated using *PROCHECK*. Triangles represent glycines.

5.2.8. Description of the Structure

The overall structure of PGA from *Alcaligenes faecalis* (*Af*PGA) confirms the characteristic Ntn-hydrolase fold comprising of a four-layered αββα core structure that is formed by two antiparallel β-sheets packed against each other, and these β-sheets are sandwiched between the layers of α-helices on either side. The smaller α-chain and bigger β-chain of the heterodimer are closely intertwined to form a pyramidal structure (**Figure 5.4**). *PROMOTIF* analysis using *PDBsum* (http://www.ebi.ac.uk/pdbsum/) (Hutchinson & Thornton, 1996) of the structure in the orthorhombic form showed seven β–sheets (A-G) and 25 helices. The α-chain consists of two strands (a part of sheet A) and ten helices (H1-H10); their connecting segments form one β-hairpin, two β-bulges, 30 β-turns and two γ-turns. Similarly, the β-chain consists of 26 strands (8 in sheet A, 6 in B, 4 in C and 2 each in sheet D, E, F, G) and 15 helices (H1-H15), and the connecting segments form 12 β-hairpins, two β-bulges, 67 β-turns and five γ-turns (**Figure 5.5A, B, Table 5.3**).

The structure of the tetragonal form of *Af*PGA showed eight β-sheets (A-G) and 27 helices. The α-chain consists of two strands (a part of sheet B) and 11 helices (H1-H11), and the connecting part of the polypeptide chain forms one β-hairpin, two β-bulges and 11 β-turns. Similarly, the β-chain consists of 27 strands distributed in various β-sheets (4 in sheet A, 9 in B, 6 in C, 5 in D, 4 in E and 2 each in sheet F, G, H) and 16 helices (H1-H16), and the connecting segment forms folds consisting of 12 β-hairpins, four β-bulges, 71 β-turns and six γ-turns (**Figure 5.6A, B**).

A stretch of 7 C-terminal amino acid residues in α chain and 2 C-terminal residues in β-chain in the case of orthorhombic structure and the N-terminal Glutamine and 7 C-terminal residues in α-chain in the case of tetragonal structure could not be modelled due to insufficient electron density, and are absent in the final model.

Protein Structure Classification using *CATH* (http://www.cathdb.info/) (Orengo *et al.*, 1997) classified orthorhombic form of *Af*PGA into four major domains. Smaller α-chain has been classified as Mainly Alpha class with orthogonal bundle architecture (CATH no-1.10.439.10), whereas the larger β-chain has been classified into three domains, Domain 1, 2 and 3 of 4-Layer sandwich architecture of alpha-beta class (CATH no-3.60.20.10), Roll architecture of class mainly beta (CATH no- 2.30.120.10)

and Orthogonal bundle of Mainly alpha class (CATH No- 1.10.1400.10), respectively (**Figure 5.7 A, B and C**).
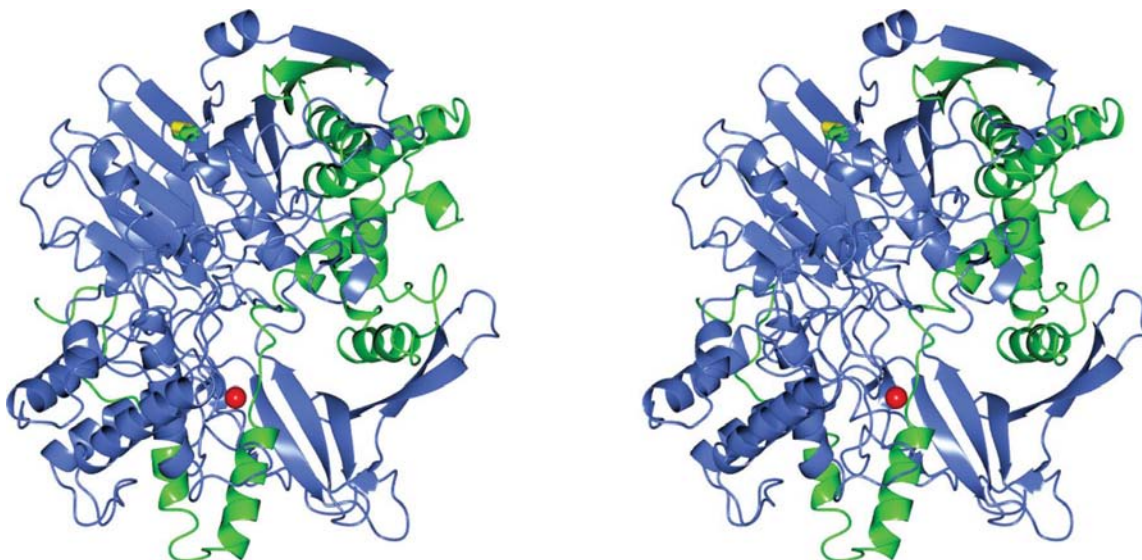
Figure 5.5: Stereoview of the three-dimensional structure of *Af*PGA showing the heterodimeric association (smaller α-chains are shown in green and larger β-chains in blue) and the αββα arrangement for the orthorhombic form (PDB code: 3K3W).
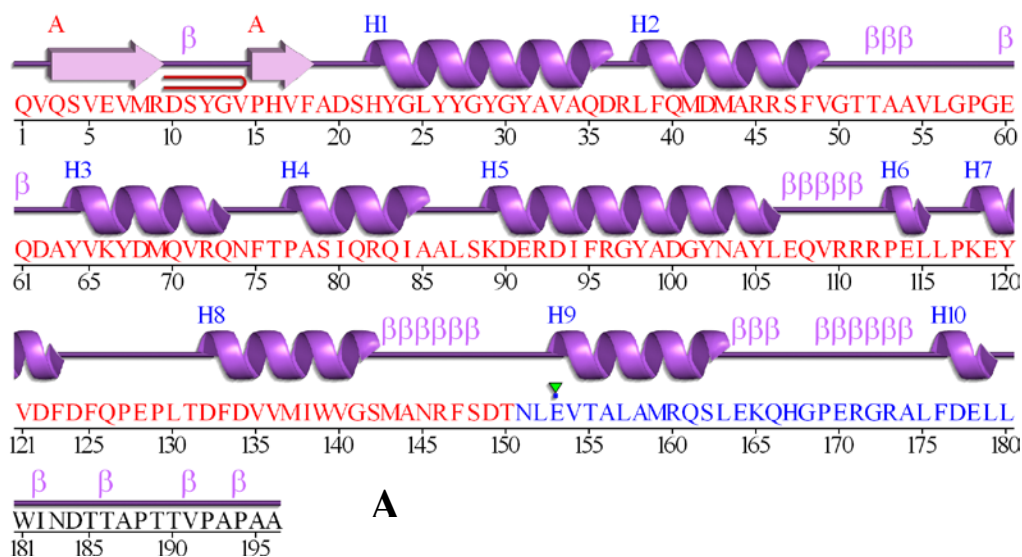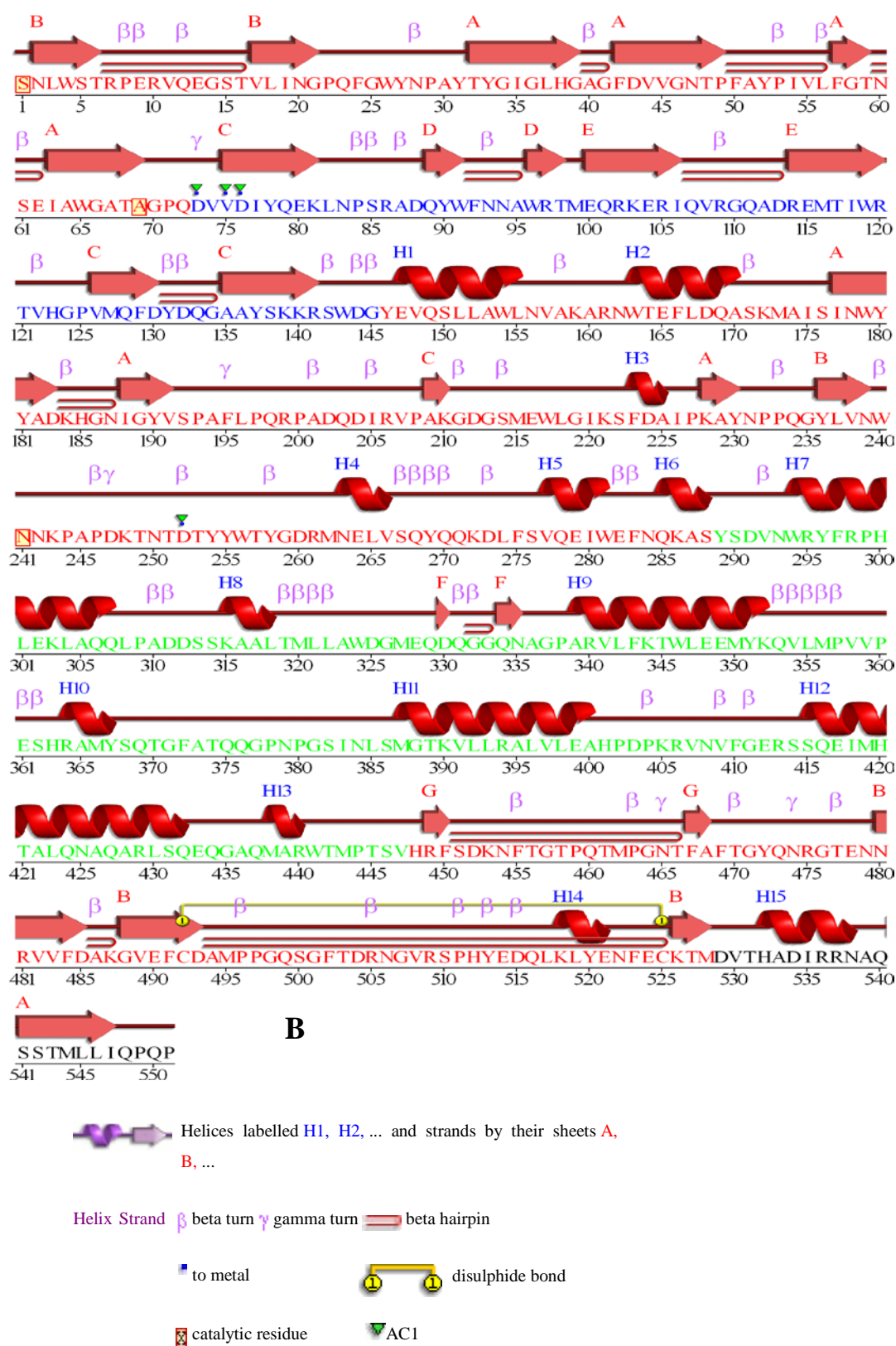
Figure 5.6: Arrangement of secondary structure elements in *Af*PGA orthorhombic form; A) α-Chain, B) β-Chain.

Table 5.3: List of Helices and Sheets in orthorhombic form of *Af*PGA

Strand number	Initial residue	Terminal residue
Sheet A (Mixed)		
Strand 1	Gln3α	Arg9α
Strand 2	Pro15α	Phe18α
Strand 3	Thr32β	Gly39β
Strand 4	Phe42β	Pro49β
Strand 5	Phe57β	Thr59β
Strand 6	Ile63β	Ala69β
Strand 7	Ile177β	Asp183β
Strand 8	Ile188β	Val191β
Strand 9	Lys228β	Tyr230β
Strand 10	Ser541β	Ile547β
Sheet B (Antiparallel)		
Strand 1	Asn2β	Thr6β
Strand 2	Val17β	Gly21β
Strand 3	Tyr236β	Asn239β
Strand 4	Asn480β	Asp485β
Strand 5	Gly488β	Asp493β
Strand 6	Lys526β	Met528β
Sheet C (Mixed)		
Strand 1	Val75β	Lys81β
Strand 2	Val126β	Asp130β
Strand 3	Ala135β	Arg141β
Strand 4	Ala209β	Lys210β
Sheet D (Antiparallel)		
Strand 1	Gln89β	Trp91β
Strand 2	Trp96β	Thr98β
Sheet E (Antiparallel)		
Strand 1	Glu100β	Ile106β
Strand 2	Arg114β	Arg120β
Sheet F (Antiparallel)		
Strand 1	Asp330β	Gln331β

Strand 2		Gln334β		Asn335β
Sheet G (Antiparallel)				
Strand 1		Arg449β		Phe450β
Strand 2		Phe467β		Ala468β
Helix number	**Helix ID**	**Initial residue**	**Terminal residue**	**Helix class**
α-Chain				
Helix 1	H1	His 22	Asp36	4-turn helix (α helix)
Helix 2	H2	Leu38	Val49	4-turn helix (α helix)
Helix 3	H3	Tyr64	Gln73	4-turn helix (α helix)
Helix 4	H4	Pro77	Ala85	4-turn helix (α helix)
Helix 5	H5	Lys89	Leu106	4-turn helix (α helix)
Helix 6	H6	Pro113	Leu115	3-turn helix (3_{10} helix)
Helix 7	H7	Lys118	Phe123	4-turn helix (α helix)
Helix 8	H8	Asp132	Ser142	4-turn helix (α helix)
Helix 9	H9	Glu153	Leu163	4-turn helix (α helix)
Helix 10	H10	Phe176	Leu179	4-turn helix (α helix)
β-Chain				
Helix 1	H1	Glu147	Leu155	4-turn helix (α helix)
Helix 2	H2	Trp 163	Ala170	4-turn helix (α helix)
Helix 3	H3	Phe223	Ala225	3-turn helix (3_{10} helix)
Helix 4	H4	Asn263	Val266	4-turn helix (α helix)
Helix 5	H5	Val277	Trp281	4-turn helix (α helix)
Helix 6	H6	Gln285	Ser288	4-turn helix (α helix)
Helix 7	H7	Trp294	Gln307	4-turn helix (α helix)
Helix 8	H8	Lys315	Leu318	4-turn helix (α helix)

Helix 9	H9	Ala339	Lys352	4-turn helix (α helix)
Helix 10	H10	Arg364	Tyr367	4-turn helix (α helix)
Helix 11	H11	Met387	Ala400	4-turn helix (α helix)
Helix 12	H12	Ser415	Gln432	4-turn helix (α helix)
Helix13	H13	Met438	Arg440	3-turn helix (3₁₀ helix)
Helix14	H14	Lys518	Glu521	4-turn helix (α helix)
Helix 15	H15	His532	Asn538	4-turn helix (α helix)
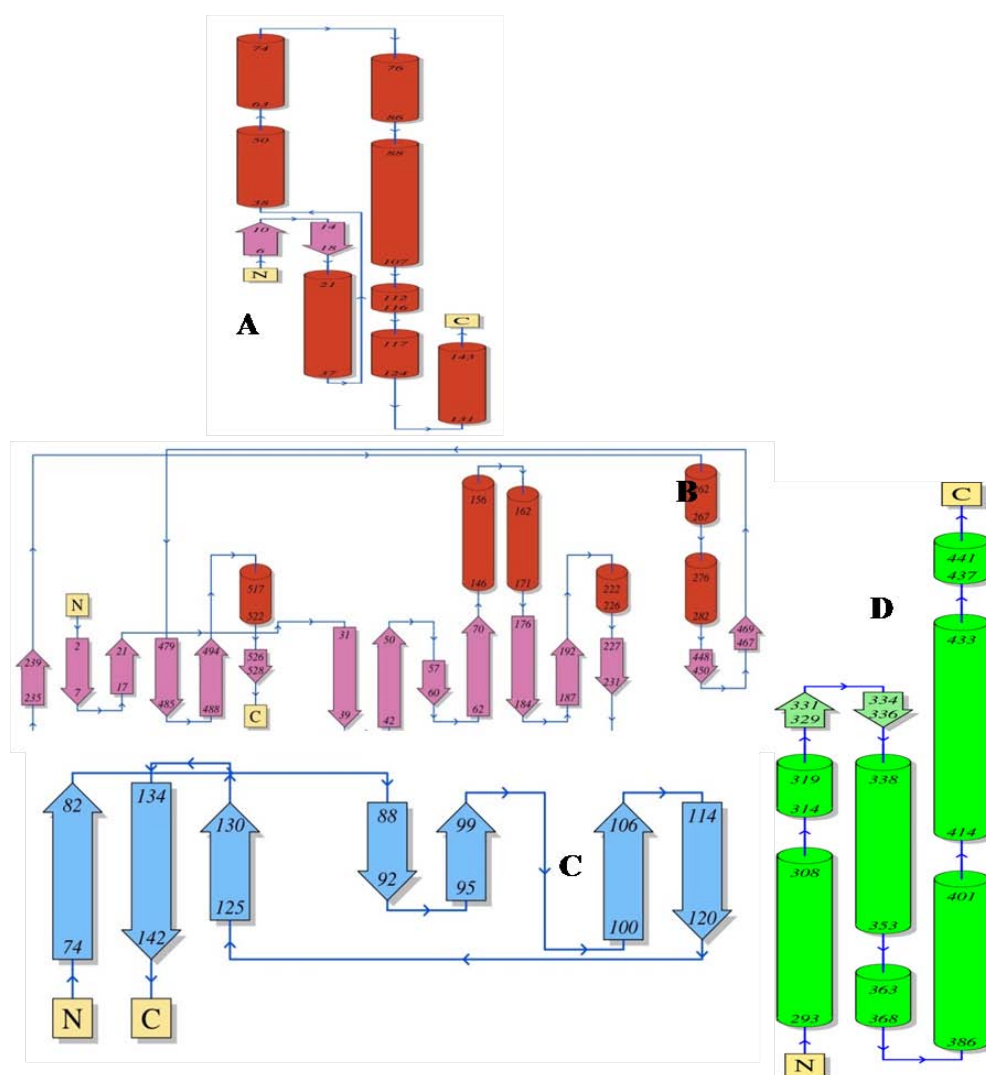
Figure 5.7: Domain Classification of *Af*PGA. The α-helices are represented by cylinders. The β-strands are represented by flat arrows. Residues forming the secondary structure are labelled as numbers. A) Domain represented by α-chain, B), C) and D); Domains of β-chain.

5.2.9. Calcium Site

A major positive density blob was located in the *Af*PGA structure during the initial stages of refinement. The positive density was identified as calcium ion as the highest peak present in the calculated difference Fourier map of both crystal forms and by comparison with the calcium-binding site in other PGAs. We could also find a major anomalous peak at the Ca^{2+} position in a DANO map. Interestingly, we did not use calcium in the crystallization solution. We therefore believe that the bound metal atom must have come from the protein sample itself, based on a report of the presence of a tightly bound calcium ion in purified *Af*PGA (Kasche *et al*., 2003). As previously mentioned, a bound calcium ion has also been reported at the same position in the *E.coli* and *P.rettgeri* PGA structures. Five of the six calcium coordinating residues in *E.coli* PGA (Glu156, Asp336, Val338, Asp339 and Asp515; numbering according to the *E.coli* precursor enzyme) are conserved in *Af*PGA. The conserved calcium-binding residues in *Af*PGA are Gluα153, Aspβ73, Valβ75, Aspβ76 and Aspβ252 (**Figure 5.8**). Ca^{2+} ion is found coordinating to side chain carboxyl oxygens of Glu153α, Asp73β, Asp76β and Asp252β. In the case of the valine residue the coordination is through the main-chain carbonyl O atom, as is the case for Pro468 in *Ec*PGA, which is replaced by Ileβ205 in *Af*PGA.

5.2.10. Disulphide Bond

The disulphide bond proposed to be one factor responsible for the comparatively high thermostability of this penicillin acylase by biochemical and gel retardation studies done by Verhaert *et al.* in 1997 was observed between the residues β492 – β525 (**Figure 5.9**). We could also find a minor peak for disulfide bond in DANO map. The stability of proteins is substantially increased by naturally occurring disulfide cross-links. A disulfide bond can contribute as much as 5 to 6 kcal/mole towards the stability of the folded protein at optimal temperature (Matsumura & Matthews 1991; Betz 1993).
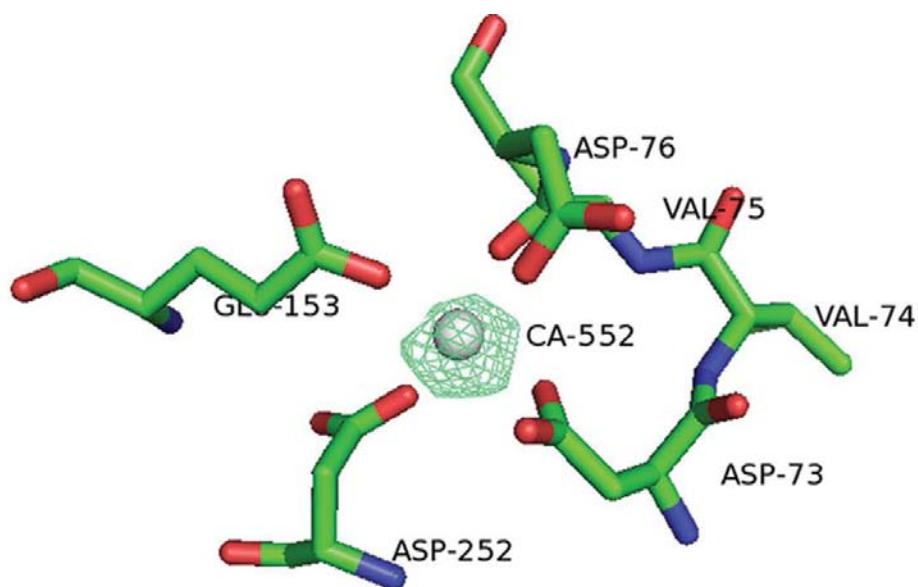
Figure 5.8: The calcium-binding site showing the residues (labelled as in *Af*PGA) involved in calcium binding. The electron density surrounding Ca^{2+} is that of a F$_o$-F$_c$ map calculated in the absence of calcium and contoured at 6σ. This figure and next figure are prepared in PyMol, Schrodinger, LLC.
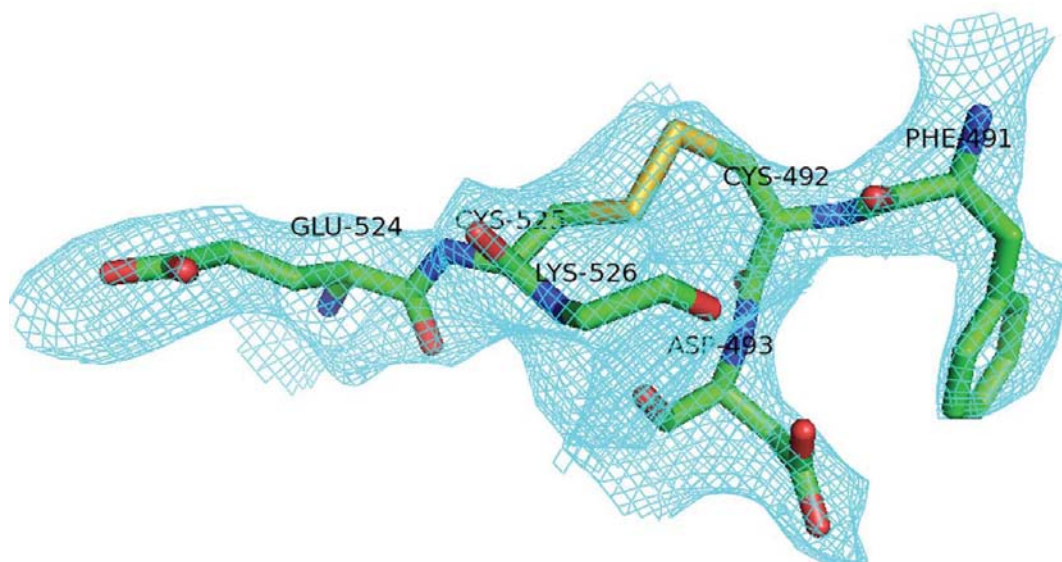
Figure 5.9: The disulfide bond (shown in yellow) formed by the two cysteine residues β492 and β 525 as observed in the orthorhombic form of *Af*PGA, together with the 2F$_o$-F$_c$ electron-density map contoured at 1σ. Two neighbouring residues of each cysteine are also shown. The same disulfide bond has also been identified in the tetragonal form.

5.2.11. Structural Comparison of *Af*PGA Structure with other PGAs

The structure of *Af*PGA reported here is similar to the reported structures of the enzyme from other sources. The r.m.s. deviations for superposition of Cα atoms of the orthorhombic form (PDB code: 3K3W) on other structures are as follows: tetragonal form (PDB code: 3ML0), 0.69 Å; *Ec*PGA (PDB code: 1PNK; Duggleby *et al*., 1995), 1.19 Å; Bro1 mutant *Pr*PGA (PDB code: 1CP9; McDonough *et al*., 1999), 1.21 Å (**Figure 5.10**).

Structural alignment of the *Af*PGA orthorhombic structure with *Ec*PGA and *Pr*PGA identified the key catalytic residues, the calcium binding residues, and residues that could be responsible for substrate specificity in three PGAs. The residues reported in the catalysis of the substrate Pen G including nucleophile Serine and the residues known to stabilize the oxyanion hole formed (Serβ1, Gluβ23, Alaβ69 and Asnβ241) are conserved in all three PGAs including *Af*PGA (**Figure 5.11**).

Structure of *Ec*PGA complexed with Phenylacetic acid (PAA) showed that PGA is a kidney shaped structure having a deep depression in the surface of the enzyme, at the base of which is a small hydrophobic pocket lined by many aromatic residues and hydrophobic side chains containing the serine residue that accommodates the phenyl ring of the substrate. This complementary fit is believed to confer specificity for the penicillin G phenylacetic R-group. The amide bond of the substrate is cleaved when the nucleophilic O^{γ} of Ser β1 attacks the carbonyl oxygen of the substrate forming an acyl-enzyme intermediate. Bond breakage is facilitated by positioning the substrate side-chain carbonyl group in oxyanion hole formed by the main-chain amide nitrogen of Ala β69 and N^{δ} of Asn β241 (Duggleby *et al.,* 1995). Similarly the conserved active site residues of *Af*PGA were also found to be present in the deep cavity (**Figure 5.12**).
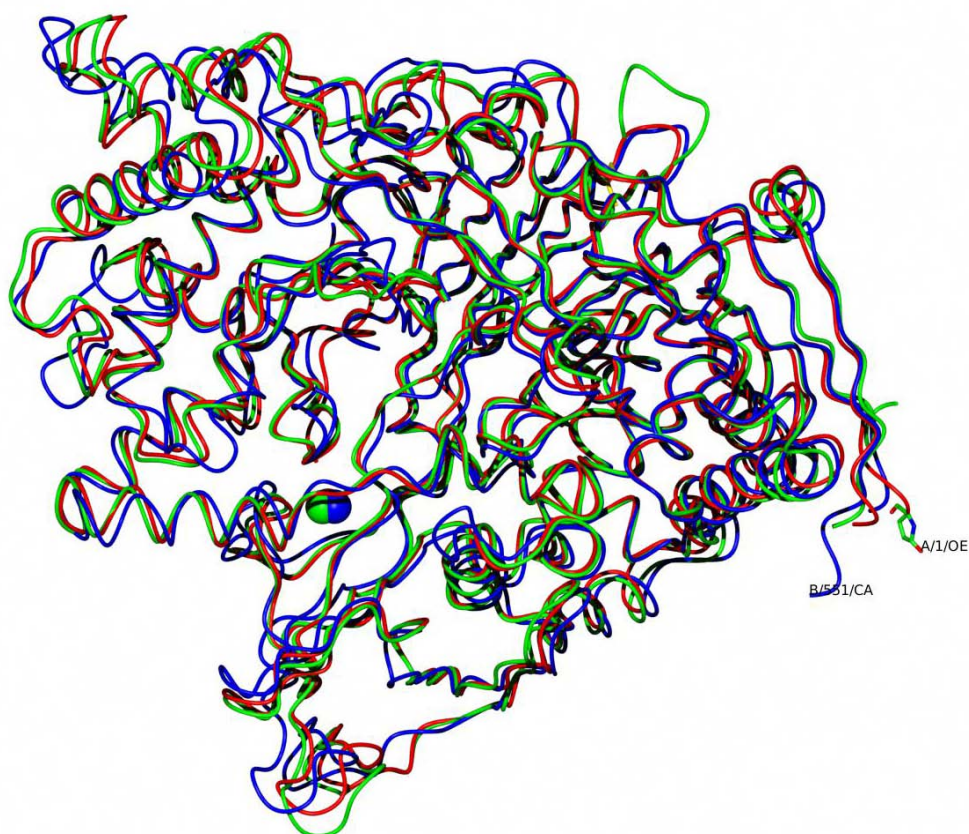
Figure 5.10: Superposition of Cα trace of the three PGA structures: *Af*PGA in Blue (PDB code: 3K3W), *Ec*PGA in Green (PDB code: 1PNK) and Bro mutant of *Pr*PGA in red (PDB code: 1CP9). Ca^{2+} present in all the three structures are represented as spheres using same colour code as the backbone. Superposition of the structures and figure was prepared using *CCP*4*MG.*

<u>α- Subunit</u>

```
1PNK_A     EQSSSEIKIVRDEYGMPHIYANDTWHLFYGYGYVVAQDRLFQMEMARRSTQGTVAEVLG- 59
1CP9_A     --ESTQIKIERDNYGVPHIYANDTYSLFYGYGYAVAQDRLFQMEMAKRSTQGTVSEVFG- 57
3K3W_A     --QVQSVEVMRDSYGVPHVFADSHYGLYYGYGYAVAQDRLFQMDMARRSFVGTTAAVLGP 58
             .   .::: **.**:**::*:. : *:*****.*********:**:**  **.: *:*

1PNK_A     ---KDFVKFDKDIRRNYWPDAIRAQIAALSPEDMSILQGYADGMNAWIDKVNTNPETLLP 116
1CP9_A     ---KDYISFDKEIRNNYWPDSIHKQINQLPSQEQDILRGYADGMNAWIKQINTKPDDLMP 114
3K3W_A     GEQDAYVKYDMQVRQNFTPASIQRQIAALSKDERDIFRGYADGYNAYLEQVRRRPE-LLP 117
             . ::.:* ::*.*: * :*: **  *. :: .*::***** **:.::. .*: *:*

1PNK_A     KQFNTFGFTPKRWEPFDVAMIFVGTMANRFSDSTSEIDNLALLTALKDKYGVSQGMAVFN 176
1CP9_A     KQFIDYDFLPSQWTSFDVAMIMVGTLANRFSDMNSEIDNLALLTALKDKYGEQLGVEFFN 174
3K3W_A     KEYVDFDFQPEPLTDFDVVMIWVGSMANRFSDTNLEVTALAMRQSLEKQHGPERGRALFD 177
           *::  :.* *.    ***.** **::****** . *: **: :*:.::* . *  .*:

1PNK_A     QLKWLVNPSAPTTIAVQESNYPLKFNQQNSQTA 209
1CP9_A     QINWLNNPNAPTTISSEEFTYSDSQKTKNIS-- 205
3K3W_A     ELLWINDTTAPTTVPAPAA-------------- 196
           :: *: :..****:.
```

β-Subunit

```
1PNK_B     SNMWVIGKSKAQDAKAIMVNGPQFGWYAPAYTYGIGLHGAGYDVTGNTPFAYPGLVFGHN 60
1CP9_B     SNVWLVGKTKASGAKAILLNGPQFGWFNPAYTYGIGLHGAGFNIVGNTPFAYPAILFGHN 60
3K3W_B     SNLWSTRPERVQEGSTVLINGPQFGWYNPAYTYGIGLHGAGFDVVGNTPFAYPIVLFGTN 60
           **:*       :.. ..:::******: ************:::.******* ::** *

1PNK_B     GVISWGSTAGFGDDVDIFAERLSAEKPGYYLHNGKWVKMLSREETITVKNGQAETFTVWR 120
1CP9_B     GHVSWGSTAGFGDGVDIFAEQVSPEDPNSYLHQGQWKKMLSRQETLNVKGEQPITFEIYR 120
3K3W_B     SEIAWGATAGPQDVVDIYQEKLNPSRADQYWFNNAWRTMEQRKERIQVRGQADREMTIWR 120
           . ::**:***  * ***: *::... .. * .:. * .* .*:* : *:.    : ::*

1PNK_B     TVHGNILQTDQTTQTAYAKSRAWDGKEVASLLAWTHQMKAKNWQEWTQQAAKQALTINWY 180
1CP9_B     TVHGNVVKRDKTTHTAYSKARAWDGKELTSLMAWVKQGQAQNWQQWLDQAQNQALTINWY 180
3K3W_B     TVHGPVMQFDYDQGAAYSKKRSWDGYEVQSLLAWLNVAKARNWTEFLDQASKMAISINWY 180
           **** ::;; *     :**:* *:*** *: **:** :  :*:** :: :** : *:;****

1PNK_B     YADVNGNIGYVHTGAYPDRQSGHDPRLPVPGTGKWDWKGLLPFEMNPKVYNPQSGYIANW 240
1CP9_B     YADKDGNIGYVHTGHYPDRQINHDPRLPVSGTGEWDWKGIQPFANNPKVYNPKSGYIANW 240
3K3W_B     YADKHGNIGYVSPAFLPQRPADQDIRVPAKGDGSMEWLGIKSFDAIPKAYNPPQGYLVNW 240
           *** .****** ..  *:*  .:* *:*. * *. :* *: .*   **.*** .**:.**

1PNK_B     NNSPQKDYPASDLFAFLWGGADRVTEIDRLLEQKPRLTADQAWDVIRQTSRQDLNLRLFL 300
1CP9_B     NNSPAKNYPASDLFAFLWGSADRVKEIDNRIEAYDKLTADDMWAILQQTSRVDLNHRLFT 300
3K3W_B     NNKPAPDKTNTDTYYWTYG--DRMNELVSQYQQKDLFSVQEIWEFNQKASYSDVNWRYFR 298
           **.* : . :* : : :* **:.*:   :   ::.:: * . :::* *:* * *

1PNK_B     PTLQAATSGLTQSDPRRQLVETLTRWDGINLLNDDGKTWQQPGSAILNVWLTSMLKRTVV 360
1CP9_B     PFLTQATQGLPSNDNSVKLVSMLQQWDGINQLSSDGKHYIHPGSAILDIWLKEMLKATLG 360
3K3W_B     PHLEKLAQQLPADDSSKAALTMLLAWDGMEQ--DQGGQNAGPARVLFKTWLEEMYKQVLM 356
           * *   :. *. .*     : *  ***::    .:*     *. .:;. ** .* * .:

1PNK_B     AAVPMPFDKWYSASGYETTQDGPTGSLNISVGAKILYEAVQGDKSPIPQAVDLFAGKPQQ 420
1CP9_B     QTVPAPFDKWYLASGYETTQEGPTGSLNISTGAKLLYESLLEDKSPISQSIDLFSGQPQN 420
3K3W_B     PVVPESHRAMYSQTGFATQQGPNPGSINLSMGTKVLLRALVLEAHPDPKRVNVFGERSSQ 416
            .** ..  *  :*: * *    .**:*:* *:*:* .::  :  * .: :::*. :..:

1PNK_B     EVVLAALEDTWETLSKRYGNNVSNWKTPAMALTFRANNFFGVPQAAAEETRHQAEYQNRG 480
1CP9_B     DVIRKTLNTTYQKMIEKYGDNPANWQTPATALTFRENNFFGIPQALPQENFHQNEYHNRG 480
3K3W_B     EIMHTALQNAQARLSQEQGAQMARWTMPTSVHRFSDKNFTGTPQTMPGNTFAFTGYQNRG 476
           :::  :*: :   : :. * : :.* *: . *  :** * **. . :.     *:***

1PNK_B     TENDMIVFSPTTSDRPVLAWDVVAPGQSGFIAPDGTVDKHYEDQLKMYENFGRKSLWLTK 540
1CP9_B     TENDLIVFT----EEGVSAWDVVAPGQSGFISPQGKPSPHYQDQLSLYQQFGKKPLWLNS 536
3K3W_B     TENNRVVFD----AKGVEFCDAMPPGQSGFTDRNGVRSPHYEDQLKLYENFECKTMDVTH 532
           ***: :**       . *    *.:.****** :* . **:***.:*::*  *.: :.

1PNK_B     QDVEAHKESQEVLHVQR-- 557
1CP9_B     EDVAPYIESTETLIIER-- 553
3K3W_B     ADIRRNAQSSTMLLIQPQP 551
            *:     :*   * ::
```

Active Site/Catalytic Residues Ca⁺² Binding Residues Disulfide bond forming Residues

Figure 5.11: Sequence alignment of *Af*PGA orthorhombic structure (PDB code: 3K3W) with *Ec*PGA (PDB code: 1PNK) and *Pr*PGA (PDB code: 1CP9), generated by structural superposition of their Cα atoms. Amino acids conserved in all three structures are marked with asterisks. Active site conserved residues are highlighted as Green; residues involved in Calcium binding site are highlighted as blue. Two cysteines present only in *Af*PGA involved in the formation of disulphide bond are highlighted as red.
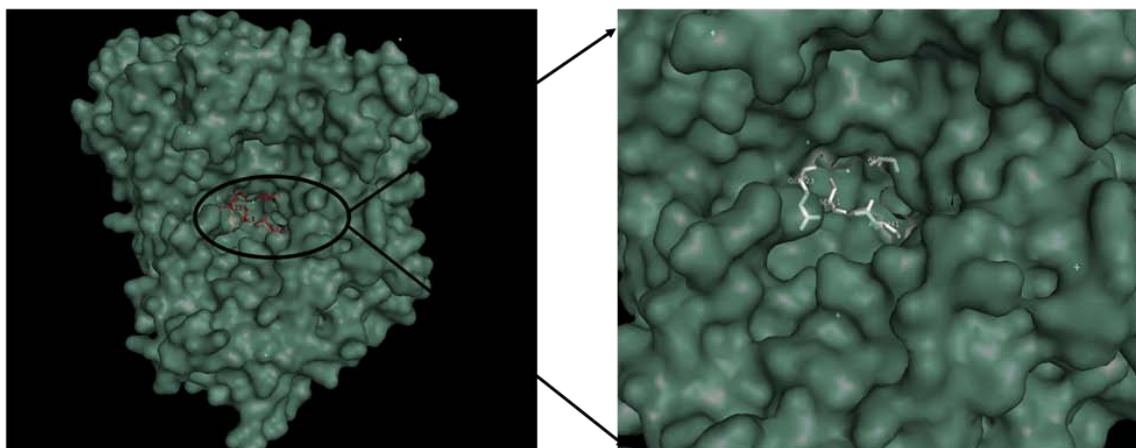
Figure 5.12: **Representation of deep cavity formed at the surface of *Af*PGA that accommodates the catalytic residues shown in ball and stick model.**

5.2.12. Thermostability Factors

Three-dimensional structure of *Af*PGA shows the presence of disulfide bond that was proposed to be a factor responsible for its enhanced thermostability. To further understand, the other possible structural characteristics associated with its thermostability were investigated comparing the amino acid sequence and the crystal structures of the *Af*PGA with other two PGAs whose three-dimensional structures are available.

Factors that affect thermostability of proteins have been a subject of numerous investigations (Ikai *et al.,* 1980; Merkler *et al*., 1981; Vihinen, 1987; Menendez-Ariaz & Argos, 1989; Matthews *et al*., 1987; Vetriani *et al.*, 1998; Das & Gerstein, 2000; Chakravarty & Varadarajan, 2000; Sterner & Liebl, 2001). Comparisons of structures from different thermophilic and mesophilic proteins are reported in various reports. *Af*PGA can be a good candidate to investigate the factors responsible for its thermostability in comparison with three-dimensional structures of PGAs from *E. coli* and *P. rettgeri.* However, a comparison based on the structural features of each of these three proteins is difficult because each protein structure was determined by a different group, refined by different approaches and at different resolutions. Fine details of each structure such as hydrogen bonds, salt bridges, etc, will depend on resolution and data quality. With such inadequacies in mind we have tried to analyze a limited set of

characteristics of the structures reported from those three organisms in an attempt to analyze the higher thermostability of *Af*PGA.

Salt bridges are calculated using molecular modelling package *WHAT IF* (http://swift.cmbi.ru.nl/servers/html/index.html, Vriend, 1990). The *WHAT IF* software atomic contact and salt bridges calculation results shows that the orthorhombic form of *Af*PGA has approximately 13% more salt bridges than that of *E.coli* and 44% more salt bridges than that of *P.rettgeri* (**Table 5.4**). A previous study has suggested that there is a strong correlation between the number of salt bridges and protein thermal stability (Das & Gerstein, 2000). Studies on protein three-dimensional structures have shown that in several ways the salt bridges can stabilize proteins. Ion pair networks, salt bridges buried in a hydrophobic core, helix-stabilizing salt bridges and surface salt bridges between two subunits are among the most frequently encountered types (Hennig *et al.*, 1995; Lebbink *et al.*, 1998; Salminen *et al.*, 1996; Sindelar *et al.*, 1998; Yip *et al.*, 1995). A single salt bridge can contribute 13 to 22 KJ/mol to the free energy of folding and salt bridges are relatively less affected at extremely high temperatures. They therefore may be responsible in providing the large increments of structural stability necessary to maintain function closer to 100 °C (Vetriani *et al*, 1998). Therefore the presence of extra salt bridges may be additional source of thermostability in *Af*PGA enzyme.

The amino acid composition of processed *Af*PGA and its correlation with thermostability were investigated as well. Results showed that the amino acid sequence of processed mature *Af*PGA has decreased content of Q+N+S+T than either of processed active *E.coli* PGA and *P.rettgeri* PGA which is consistent with the conclusion of the genomic studies conducted by Chakravarty & Varadarajan (2000) and Sterner & Liebl (2001) (**Table 5.4**). They found that mesophilic proteins normally have a decreased content of uncharged polar amino acids (i.e. Q, N, S, and T). As S and T can catalyze the deamination and backbone cleavage of Q and N residues, a reduction in the amount of all four of these residues would minimize deamination. Reduction of deamination in proteins confers stability on thermophilic proteins.

The number of Arginines, or rather, an Arg/ (Arg+ Lys) ratio, has been reported to be higher in thermostable proteins (Argos *et al.*, 1979; Merkler *et al.,* 1981; Kumar

et al., 2000). The higher Arg/ (Arg+ Lys) ratio in *Af*PGA may also account for its thermostability to some degree. This is due to both higher number of Arginines and lower number of Lysines present in *Af*PGA as compared to *Ec*PGA and *Pr* PGA (**Table 5.4**). Studies indicate that there is a correlation between protein stability and the number of arginines on the protein surface (Turunen *et al.*, 2002). The site-directed mutagenesis study has proved that Lys-Arg mutations known to increase the thermo tolerance of the enzymes because of the stronger hydrogen bonding of the large guanidium group of arginines with nearby polar groups (Cunningham & Wells, 1987; Mrabet *et al*., 1992; Turunen *et al.,* 2002). Because of its large side chains, arginines may be important in both local and long range interactions (Borders *et al.,* 1994; Kumar *et al.*, 2000). In addition, arginines residues may also provide advantage for thermostable proteins in the following ways: involvement in ion pair networks, more rigidity, better shielding of the hydrophobic hydrocarbon chain from water, and retaining charge at higher pH (Borders *et al.,* 1994; Cai *et al.,* 2004).

Another factor that may be related to the thermostability of the *Af*PGA is its notably greater amount of proline residues (**Table5.4**). Proline has a unique geometry, restricting the conformational flexibility of the folded polypeptide chain (Branden & Tooze, 1991); prolines also fit well into the first turn of an α-helix and consequently increases internal hydrophobicity, thus stabilizing the α-helix (Suzuki *et al.*, 1987). Proline can bend the polypeptide on itself so as to make the backbone much more easily hydrogen-bonding with the polar side chains of other turn formers (Lewis *et al*. 1973; Levitt, 1978). In addition, it is proposed that proline residues, in general reduces the entropy of unfolding and are therefore important in stabilizing the molecular structure (Vieille & Zeikus, 1996). It has been suggested that introduction of prolines at selected sites in target proteins increases its molecular stability (Herning *et al.*, 1992; Muslin *et al.*, 2002; Nakamura *et al.,* 1997; Zhu *et al*., 1999).

A comparison of genomes with respect to specific genes from mesophiles and thermophiles provides a data set that is sufficiently large enough to extract specific trends in amino acid usage. Compared to mesophiles, genomes of thermophiles encode higher levels of charged amino acids, primarily at the expense of uncharged polar residues (Jaenicke & Bohm, 1998). **Table 5.5** shows the amino acids distribution in each of the three PGAs and   **Table 5.6** shows that the processed *Af*PGA indeed has a

net increase of charged (D, E, K, R, H) amino acids as compared to that of *Ec*PGA and *Pr*PGA. Although we have not seen much decrease in the amount of uncharged polar residues in *Ec*PGA compared to *Af*PGA the decrease in the case of *Pr*PGA is significant.

Table 5.4: Putative factors responsible for the high thermostability of *Af*PGA

PGA Source	No. of Salt Bridges	%(N+Q+S+T)	%Arg	%Lys	%Arg/ (Arg +Lys)	%Proline
E.coli (1PNK) (Uniprot P06875)	310	24.54	4.18	5.35	0.4386	5.22
P.rettgeri (1CP9) (Uniprot Q7WZI9)	242	27.17	2.64	5.67	0.3176	5.41
Alcaligenes faecalis ((Uniprot Q8VQG6)	349	24.03	6.11	3.59	0.6305	6.10

Table 5.5: Amino acids composition of three PGAs whose three-dimensional structures are available

Amino Acid	*Af*PGA	*Ec*PGA	*Pr*PGA
A	62	69	50
C	2	0	0
D	48	48	48
E	39	35	36
F	33	30	33
G	51	57	56
H	14	13	17
I	23	30	46
K	27	41	43
L	48	56	56

M	26	18	14
N	40	42	54
P	46	40	41
Q	52	49	53
R	46	32	20
S	41	42	52
T	48	55	47
V	49	50	32
W	20	28	24
Y	38	31	36
Total residues	**753**	**766**	**758**

Table 5.6: Relative compositions of amino acids of three PGAs

Amino Acids	*Af*PGA	*Ec*PGA	*Pr*PGA
Charged Residues (DEKRH)	**23.11%**	**22.06%**	**21.63%**
Polar/uncharged (GSTNQYC)	**36.12%**	**36.03%**	**39.31%**
Hydrophobic residues (LMIVWP*AF*)	**40.77%**	**41.91%**	**39.06%**

5.3. DISCUSSION

Organisms tend to adapt their specific proteins to function efficiently within their normal environmental temperature. This generally implies that proteins have a limited temperature range within which structural integrity is maintained. Outside this thermal span, denaturation occurs with corresponding loss of function, such as enzyme activity. Since the rate of an enzymatic reaction generally increases with temperature, the application of an enzyme with greater temperature stability will improve the efficiency of many industrial processes such as detergent and antibiotics manufacturing, food and starch processing, PCR etc. Besides, using thermostable enzymes in industrial applications offers the extra benefits of higher substrate solubility, decreased media viscosity, longer enzyme shelf-lives at normal storage temperatures and lowered risk of microbial contamination when reactions are carried out at higher temperatures. Approaches comparing the amino acid sequences and crystal structures of homologous mesophilic and thermophilic proteins have been used to determine structural features contributing to the thermostability of many proteins, and such knowledge will be essential for the rational design of highly thermostable protein by protein engineering techniques.

This chapter deal with the three-dimensional structure determination of thermostable penicillin G acylase from *Alcaligenes faecalis* and the efforts to deduce the structural and molecular factors responsible for its thermostability. Penicillin G acylases share majority of the enzyme usage in the pharmaceutical industry for the production of semisynthetic β-lactam antibiotics via the key intermediate 6-aminopenicillanic acid. Pharmaceutical and biotechnological applications of all the penicillin acylases have emerged as a serious alternative to the traditional chemical procedures for the manufacture of β-lactam antibiotics owing to their environmental and economic benefits. The profound impact of penicillin acylases in pharmaceutical applications has sustained interest in identifying the structural determinants that are essential for catalysis by these enzymes. Efforts to enhance the catalytic rate by protein engineering have been reported (Gabor & Janssen, 2004). The PGA employed in semisynthetic penicillin manufacture is primarily sourced from *E. coli*. However, enzymes from other sources showing higher thermostability and catalytic efficiency will be more attractive biocatalysts. In industrial processes at elevated temperature, the

thermostable enzyme may provide some advantages Like- higher reaction rate, a higher solubility of substrate, a lower viscosity of reaction medium and reduced risk of microbial contamination. In this regard, a PGA produced by *Alcaligenes faecalis* (*Af*PGA) has clear industrial advantage over other well-characterized penicillin acylases in β-lactam conversion because of its reported higher thermostability and synthetic efficiency in enantioselective synthesis. The structure of the penicillin G acylase from *A. faecalis* reported here in two different crystal forms has similarity with the reported structures of the enzyme from two other sources with the r.m.s. deviations for superposition of Cα atoms of the orthorhombic form with others are lesser than 1.3 Å. The putative residues involved in catalysis (Ser β1, Glu β23, Ala β69 and Asn β241) are found to be conserved in all three PGAs. The structural similarity of *Af*PGA to other PGAs and its αββα architecture places it in the Ntn hydrolase superfamily and the free Nterminal residue Serβ1 has resulted from a post-translational autocatalytic proteolysis of the precursor enzyme after the removal of a linker peptide between the α- and β-subunits. The three-dimensional structure reported here, although at moderate resolution, has helped in carrying out a detailed analysis of all the factors that confer higher stability to this enzyme compared with other penicillin acylases. We have identified a disulfide bridge in the *Af*PGA structure between cysteine residues β492 and β525 which is absent in other penicillin acylases where three-dimensional structures are available (**Figure 5.9**). This could be one of the factors that determine the higher thermostability of this enzyme as proposed by Verhaert *et al*. (1997). The other structural and molecular features that could be considered as factors responsible for imparting thermostability to this enzyme are the increased number of salt bridges, decreased content of uncharged polar amino acids (N+Q+S+T), higher number of Arginines and a larger Arg/ (Arg+ Lys) ratio, larger number of Prolines, increase in the net content of ionic residues as compared to PGA from *E. coli and P. rettgeri.* Information on these stability factors will be useful when engineering penicillin acylases for industrial applications.

Chapter: 6

Cloning, Over-expression, Purification and Crystal Structure studies of N-myristoyltransferase from *Leishmania donovani*

6.1. INTRODUCTION

The leishmaniases, caused by species of the kinetoplastid parasite Leishmania, are a range of diseases associated with immune dysfunction that give rise to more than 2 million new cases each year in 88 countries, with 367 million people at risk (Chappuis *et al.,* 2007; Shaw, 2007; Brannigan *et al*., 2010). Cutaneous leishmaniasis, the most common form of leishmaniasis includes most prominently disfiguring skin lesions whereas visceral leishmaniasis (VL) is the most serious form, and is potentially fatal if untreated. In VL the parasites leave the inoculation site and proliferate in spleen, liver and bone marrow, resulting in host immunosuppression and ultimately death in the absence of treatment. These symptoms are severe in children and immunocompromised patients, such as those diagnosed as HIV positive. Though there is a limited range of effective drugs in use including Pentostam, amphotericin B and miltefosine, with a few potential ones in clinical trials (Croft *et al.,* 2006), delivery mode can be complex, while miltefosine, the only oral VL drug currently available, is teratogenic in nature. Developing resistance is an increasing problem, therefore a pressing need for new, preferably orally administered drugs, against this globally important but neglected health problem (Branigan *et al*., 2010).

Previous studies have identified Myristoyl-CoA:protein N-myristoyltransferase (NMT; **EC 2.3.1.97**) as a suitable candidate for drug development against protozoan parasitic infections, including *Plasmodium falciparum, Trypanosoma brucei*, *Leishmania major* and *Leishmania donovani* causative agents of human malaria, African sleeping sickness, cutaneous and visceral leishmaniasis, respectively (Bowyer *et al.,* 2007; Price *et al.,* 2003, 2010; Panethymitaki *et al.,* 2006; Brannigan *et al*., 2010).

NMT is ubiquitous in eukaryotic cells in which it catalyses the co- and post-translational addition of the C14:0 fatty acid, myristate, via an amide bond formation, to the N-terminal glycine residue of a subset of proteins. N-Myristoylation plays a role in facilitating protein–protein interactions, targeting proteins to membrane locations and stabilizing protein structures (Resh, 1999).

N-Myristoylation by NMT proceeds via an ordered Bi–Bi reaction mechanism; binding of myristoyl-CoA opens up a second pocket for docking of the substrate protein (Rudnick, *et al.,* 1991; Bhatnagar *et al.,* 1998). The myristate group is then transferred to the N-terminal glycine of the substrate in a nucleophilic addition–elimination reaction, followed by stepwise release of first the free CoA and then the N-myristoylated protein (Bhatnagar *et al.,* 1999).

NMT was first characterised in 1987 in *Saccharomyces cerevisiae* (Towler *et al.,* 1987), but subsequently has been isolated and characterised from number of pathogens, parasitic protozoans, plants and mammals. Comparative sequence and biochemical analyses as well as structural studies have demonstrated high conservation of the myristoyl-CoA binding sites in human NMT and parasites but divergent peptide-binding specificities (Johnson, 1994). Therefore, NMT has been the target of a number of antifungal drug development programmes, with the focus on development of selective inhibitors that act at the peptide-binding pocket. Peptide-based and peptidomimetic inhibitors have been developed that show selectivity against the NMTs of fungal species as compared to human NMT (Langner *et al.,* 1992; Lodge *et al.,* 1997, 1998).

NMT from *Leishmania donovani* (*Ld*NMT), the principal causative agent of the most serious form of leishmaniasis, visceral leishmaniasis has been proven to be essential for viability in extracellular parasites (Brannigan *et al.,* 2010). Cloning of the NMT gene from *Leishmania donovani* strain LV9 (*Ld*NMT) into a modified pET28 vector with a cleavable N-terminal hexahistidine tag, its purification and crystallization with non- hydrolysable myristoyl Co-A analogue S-(2-oxo) pentadecyl-CoA (NHM) has been reported in the literature (Brannigan *et al.,* 2010, PDB code: 2WUU). The structure contains one molecule of NHM and 378 well defined water molecules.

With the aim of exploiting *Ld*NMT for drug discovery, this chapter deals with the cloning of a new construct of the *Ld*NMT gene with deletion of a disordered N-terminal region, its overproduction in *E. coli* and the purification of recombinant protein. Since NMT from *Leishmania donovani* strain LV9 is different from Friedlin strain *Lm*NMT in only 11 amino acids, crystals obtained for *Ld*NMT in complex with

non hydrolysable myristoyl CoA analogue were soaked in reported inhibitor bound in *Lm*NMT peptide binding site (Frearson *et al.*, 2010). Co-crystallization by addition of a non-hydrolysable myristoyl CoA analogue and inhibitor after overnight incubation were also tried. Crystal structures in two different crystal forms were determined for the purified NMT enzyme bound with its non hydrolysable myristoyl CoA analogue. Co-crystallization of NMT with substrate and inhibitor resulted in poorly diffracting crystals.

6.2. RESULTS

6.2.1. Cloning of N-terminal mutant of the *Ld*NMT gene

A new construct of *Ld*NMT gene, an N-terminal deletant with a non-cleavable histidine tag (Δ*N LdNMT*), was prepared using the technique of Ligation Independent Cloning (LIC cloning) as described in **Section 2.3.2** and as shown in **Figure 2.4**. Plasmid DNA from *L. donovani* strain LV9 (Brannigan *et al*, 2010) was used to prepare the N-terminal deletion mutant.

The following primer pair was designed to add the LIC specific ends by PCR method using KOD polymerase (Novagen).

LdLicF   5'-CACCACCACCACATGGCATTCTGGAGCACACAGCCCGT-3'

LdLicR   5'GAGGAGAAGGCGCGTTACAACATCACCAAGGCAACCTGAGAG-3'

PCR includes one cycle of initial denaturation at 98 °C for 2 min followed by 25 cycles of denaturation at 98 °C for 15 sec, annealing at 55 °C for 2 sec and extension at 72 °C for 20 sec followed by final extension cycle at 72 °C for 3 min. A 1292 bp product was amplified for the *Ld*NMT gene with overhangs attached (**Figure 6.1**).

The amplified and gel purified *Ld*NMT gene fragment with LIC overhangs was then treated with 1 unit of T4 DNA polymerase (Novagen/Merck) in the presence of the 2.5 mM dATP to generate the single stranded vector-compatible overhangs. 0.2 pmol of the insert was used in the reaction mixture. The contents were mixed and incubated at 22 ºC for 30 min and the reaction was stopped by further incubation at 75 ºC for 20min.

The product with compatible single stranded overhangs was then ligated with linearized modified pET28 vector (LICYSBL) prepared using the same protocol as above, without in the use of T4 ligase and the annealed products were transformed into NovaBlue singles competent cells.

Confirmation of the mutant was done by double digestion of the plasmid preparation from random colonies. As N-terminal deletion mutant cloned in LIC vector generated new restriction site of NcoI that was not present in original clone, a restriction digestion with NcoI and NdeI were set up with each plasmid DNA

preparation in order to ascertain which colonies had the insert mutant DNA. The presence of an insert with size close to 1200bp confirms the presence of N-terminal mutant *L*dNMT gene which was confirmed by DNA sequencing of the isolated plasmids. Double digestion was done using 1 unit each of NcoI and NdeI.

1Kb   **amplified**

Ladder   **gene**
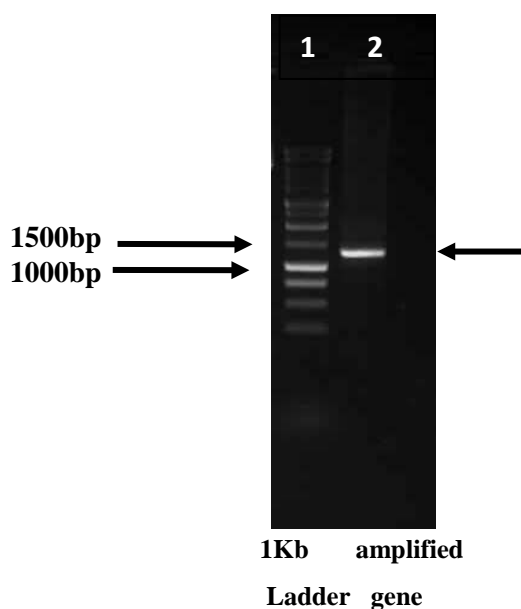
Figure 6.1: 1% agarose gel showing PCR amplified *L*dNMT gene with LIC overhangs, lane 1: 1Kb DNA ladder, lane 2: PCR product.

Sequencing of the double digestion confirmed plasmid prep confirms the cloning of N-terminal mutant of *L*dNMT gene as shown in the **Figure 6.2** and **6.3.**

M **G S S S H H H H H H M** A F W S T Q P V P Q T E D E T E K I V F A G P M D E P K T V A D I P E E P Y P I A S T F E W W T P N M E A A D D I H A I Y E L L R D N Y V E D D D S M F R F N Y S E E F L Q W A L C P P S Y I P D W H V A V R R K A D K K L L A F I A G V P V T L R M G T P K Y M K V K A Q E K G Q E E E A A K Y D A P R H I C E I N F L C V H K Q L R E K R L A P I L I K E V T R R V N R T N V W Q A V Y T A G V L L P T P Y A S G Q Y F H R S L N P E K L V E I R F S G I P A Q Y Q K F Q N P M A M L K R N Y Q L P N A P K N S G L R E M K P S D V P Q V R R I L M N Y L D N F D V G P V F S D A E I S H Y L L P R D G V V F T Y V V E N D K K V T D F F S F Y R I P S T V I G N S N Y N I L N A A Y V H Y Y A A T S M P L H Q L I L D L L I V A H S R G F D V C N M V E I L D N R S F V E Q L K F G A G D G H L R Y Y F Y N W A Y P K I K P S Q V A L V M L **Stop**

Figure 6.2: Translation product of the double digestion confirmed construct isolated from *E.coli* NovaBlue cells. Residues that were mutated are shown in bold and are underlined.

```
LdNMT         MSRNPSNSDAAHAFWSTQPVPQTEDETEKIVFAGPMDEPKTVADIPEEPYPIASTFEWWT 60

ΔN LdNMT      MGSSSHHHHHHMAFWSTQPVPQTEDETEKIVFAGPMDEPKTVADIPEEPYPIASTFEWWT 60

              *.  .. :  .      *******************************************

LdNMT         PNMEAADDIHAIYELLRDNYVEDDDSMFRFNYSEEFLQWALCPPSYIPDWHVAVRRKADK 120

ΔN LdNMT      PNMEAADDIHAIYELLRDNYVEDDDSMFRFNYSEEFLQWALCPPSYIPDWHVAVRRKADK 120

              ************************************************************

LdNMT         KLLAFIAGVPVTLRMGTPKYMKVKAQEKGQEEEAAKYDAPRHICEINFLCVHKQLREKRL 180

ΔN LdNMT      KLLAFIAGVPVTLRMGTPKYMKVKAQEKGQEEEAAKYDAPRHICEINFLCVHKQLREKRL 180

              ************************************************************

LdNMT         APILIKEVTRRVNRTNVWQAVYTAGVLLPTPYASGQYFHRSLNPEKLVEIRFSGIPAQYQ 240

ΔN LdNMT      APILIKEVTRRVNRTNVWQAVYTAGVLLPTPYASGQYFHRSLNPEKLVEIRFSGIPAQYQ 240

              ************************************************************

LdNMT         KFQNPMAMLKRNYQLPNAPKNSGLREMKPSDVPQVRRILMNYLDNFDVGPVFSDAEISHY 300

ΔN LdNMT      KFQNPMAMLKRNYQLPNAPKNSGLREMKPSDVPQVRRILMNYLDNFDVGPVFSDAEISHY 300

              ************************************************************

LdNMT         LLPRDGVVFTYVVENDKKVTDFFSFYRIPSTVIGNSNYNILNAAYVHYYAATSMPLHQLI 360

ΔN LdNMT      LLPRDGVVFTYVVENDKKVTDFFSFYRIPSTVIGNSNYNILNAAYVHYYAATSMPLHQLI 360

              ************************************************************

LdNMT         LDLLIVAHSRGFDVCNMVEILDNRSFVEQLKFGAGDGHLRYYFYNWAYPKIKPSQVALVM 420

ΔN LdNMT      LDLLIVAHSRGFDVCNMVEILDNRSFVEQLKFGAGDGHLRYYFYNWAYPKIKPSQVALVM 420

              ************************************************************

LdNMT         L 421

ΔN LdNMT      L 421

              *
```

Figure 6.3: Amino acid alignment of the N-terminal mutant with the original NMT construct (Brannigan *et al.*, 2010). Residues that were mutated are underlined.

6.2.2. Expression Studies for the Cloned Protein

The plasmid harbouring the coding sequence for the truncated NMT was transformed into *E.coli* pLEMO cells for the expression. 10 ml LB medium supplemented with kanamycin (50 µg/ml) + chloramphenicol (34 µg/ml) was inoculated with 100µl of overnight grown single colony and induced with 1 mM IPTG at OD_{600} of 0.5-0.7. Lysate from induced culture after 3 h at 37 °C of incubation showed the expression of the truncated protein mostly in insoluble fraction **(Figure 6.3.4)**.
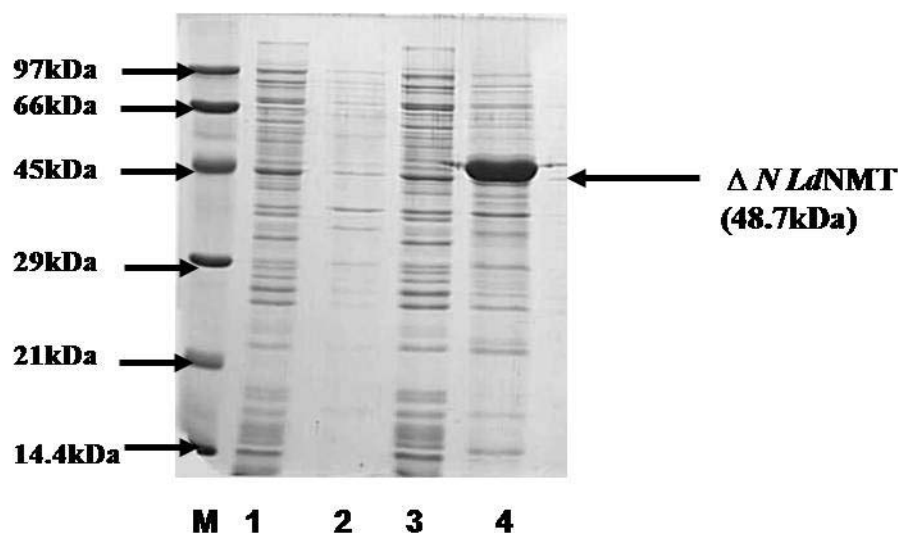
Figure 6.4: 12% SDS-PAGE showing expression profile for ΔN *Ld*NMT expressed in pLEMO cells after 3 h incubation at 37 °C on induction with 1mM IPTG. (M)- Low Molecular Range Protein Marker, lane 1- uninduced lysate, lane 2- uninduced pellet, lane 3- induced lysate, lane4- induced pellet.

6.2.3. Comparison of expression and initial purification of ΔN *Ld*NMT in LB and in Auto Induction Medium (AIM)

500 ml of LB and Auto Induction Medium each were inoculated with 5ml of overnight primary culture containing *Ld*NMT mutant in pLEMO cells. LB medium was induced with 1mM IPTG after incubating at 37 °C for 2 h. AIM was kept at 37 °C for 7 h. Both medium flasks were shifted to 18°C. LB culture was harvested at 5000 rpm for 15 min at 4 °C after 4 h of incubation while AIM was kept for overnight incubation at 18 °C. Harvested cells for both culture broths were resuspended in minimum volume of lysis on buffer. Unclarified sonicate cells were loaded onto a 1ml HisTrap crude column (GE life Sciences, USA) pre-equilibrated with binding buffer (20 mM Tris-HCl pH 8.5, 10 mM Imidazole and 0.5 M NaCl) and eluted with elution buffer containing 500 mM Imidazole. Eluted fractions along with flow through and lysate were loaded on a 12% SDS-PAGE gel to compare the expression levels and preliminary purification.

The auto-induction method of protein expression in *E. coli* is based on diauxic growth resulting from dynamic function of *lac* operon regulatory elements (*lacO* and *LacI*) in mixtures of glucose, glycerol, and lactose. During the initial growth period,

glucose is preferentially used as a carbon source and protein expression is low as a result of catabolite repression of alternative carbon utilization pathways (Inada *et al.,* 1996; Hogema *et al.,* 1998, 1999*)* and binding interactions between *lac* repressors (*LacI*) and *lac* operators (*lacO*). As glucose is depleted, catabolite repression is relieved, which leads to a shift in cellular metabolism toward the import and consumption of lactose and glycerol. Lactose import results in the production of allolactose from lactose by a promiscuous reaction of *β*-galactosidase. Allolactose then acts as the physiological inducer of the *lac* operon (Blommel *et al.,* 2007). AIM results in higher protein yields than IPTG induction. Furthermore, if a protein is very toxic or even slightly toxic, AIM has the added advantage of very low to no expression prior to the time of induction. Expression and initial purification of ΔN *Ld*NMT was found to be better in Auto Induction Media (AIM, Studier, 2005) as compared to LB media (**Figure 6.5**). As seen in the gels, there is more expression of the enzyme in the auto Induction medium and more of the binding to the affinity column. Almost all the protein expressed in LB medium was in the insoluble fraction that could not be available for the binding presumably due to unavailability of histidine tag.
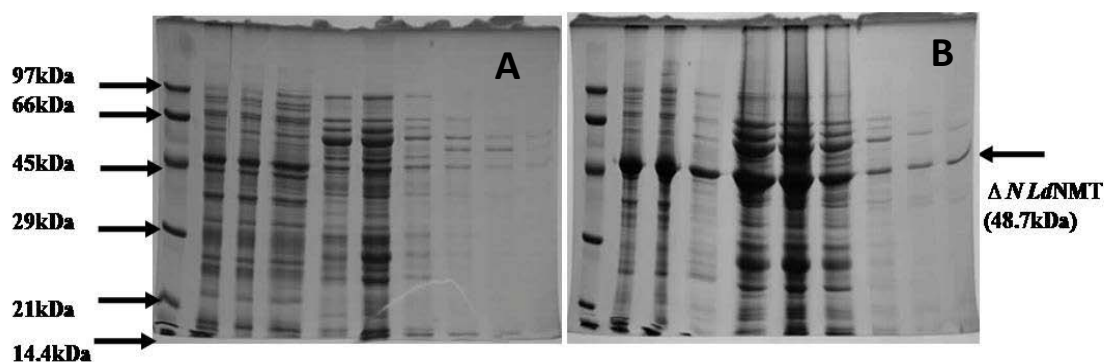
Figure 6.5: Comparison of expression and initial purification using affinity chromatography of ΔN *Ld*NMT expressed in pLEMO in (A) LB and (B) Auto Induction Medium; Lane 1-low range protein marker, lane 2-lysate 2μl, lane 3- FT 2μl, lane 4-washing 20 μl, lane 5-10-elution fraction 20 μl.

6.2.4. Purification of ΔN *Ld*NMT Enzyme

The steps ΔN *Ld*NMT enzyme purification were described under Methods in Chapter 2 under section **2.3.2.8.** Briefly the enzyme was purified from auto induction

media using a three step purification protocol. Since the enzyme was expressed containing N-terminal histidine tag, the affinity chromatography was employed as a first step of the protocol followed by anion exchange on Q-Sepharose and the polishing step of Size Exclusion chromatography. The elution profiles for the purification protocol are shown in **Figure 6.6**.
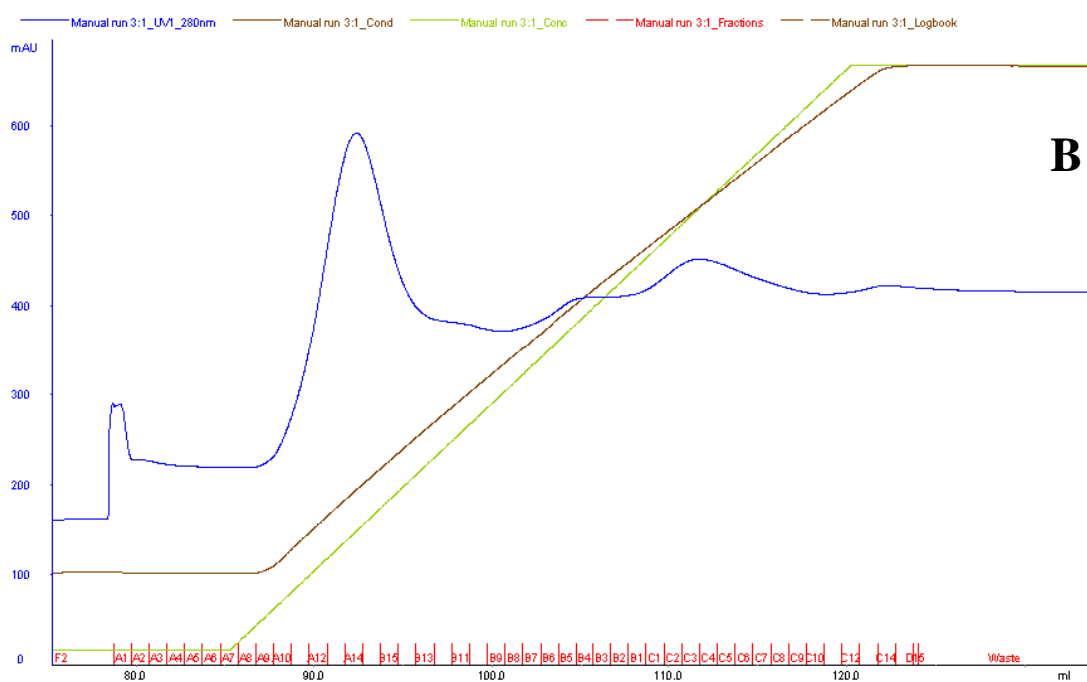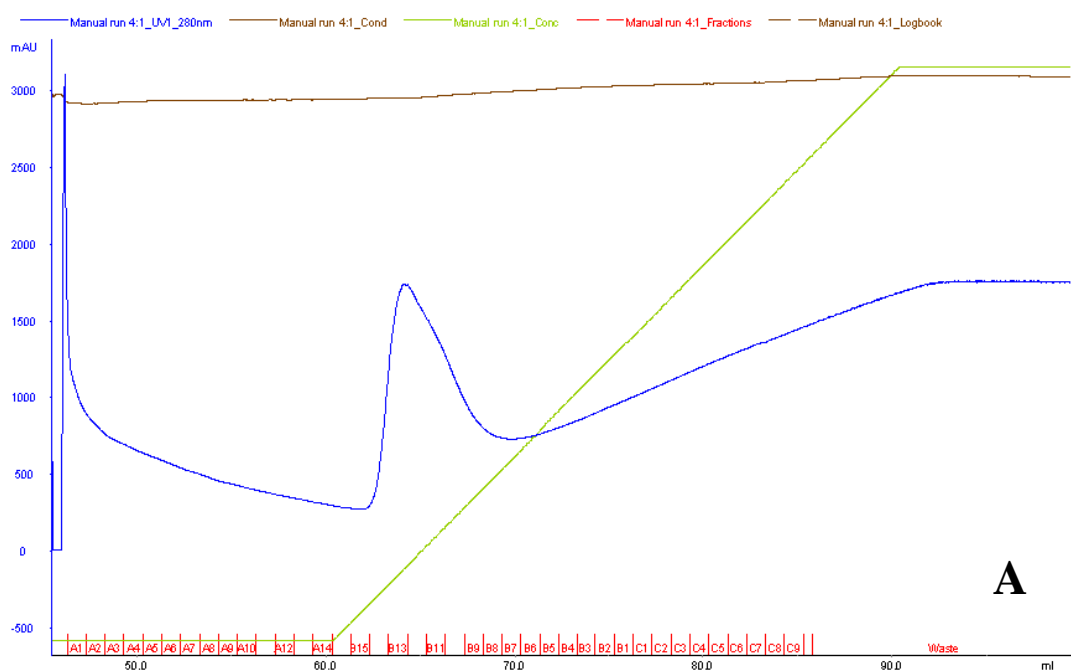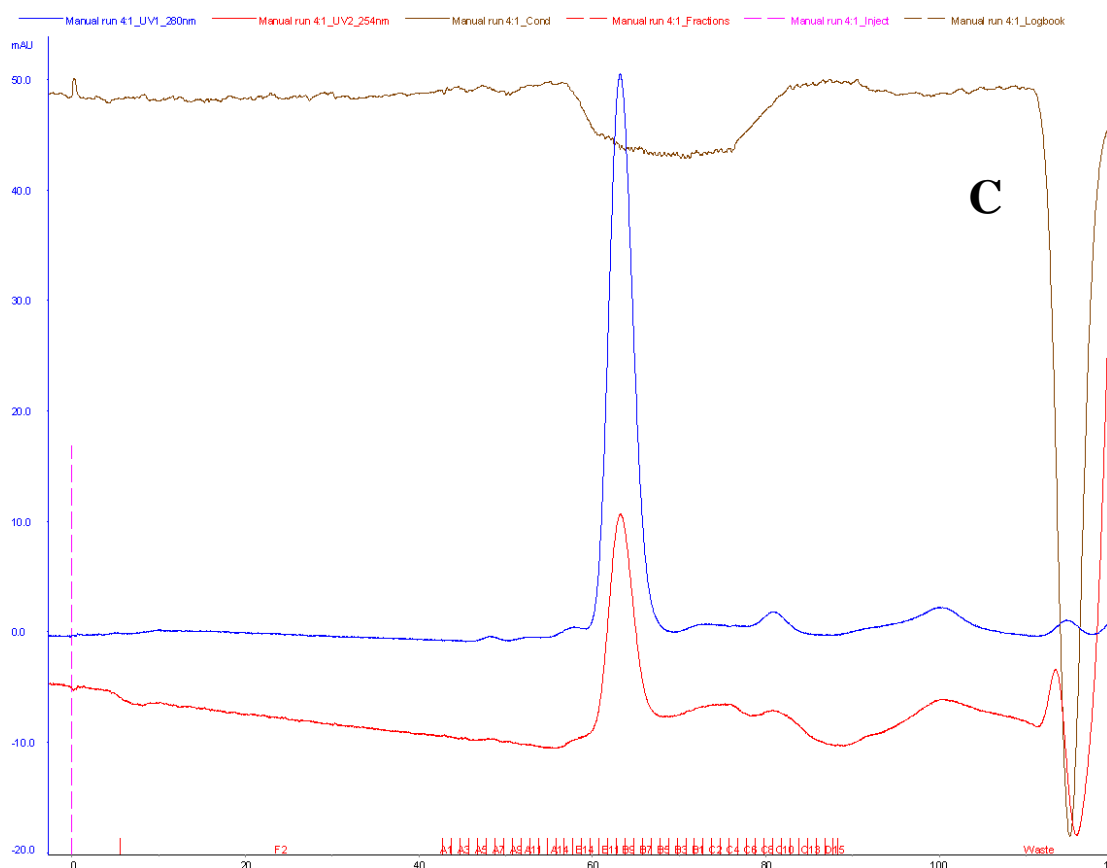
Figure 6.6: Purification protocol followed for the ΔN *Ld*NMT Enzyme; (A) Elution profile for the affinity chromatography, (B) Elution profile for Q-Sepharose, (C) Elution profile for Size exclusion Chromatography (HiLoad Superdex 75 16/60pg). OD$_{280}$ were shown in blue colour, OD$_{215}$ as red, buffer conductivity as brown and concentration of Elution buffer as green.

The typical yield for the enzyme was 3-4 mg/l of culture. Purified protein showed the presence of more than 90% pure protein of expected size on 12 % SDS-PAGE (**Figure 6.7**). Pure protein when concentrated tends to precipitate, therefore 4 times molar concentration of non-hydrolysable Myristoyl CoA analogue substrate (NHM) was added during concentration of protein. 50 µl aliquots of concentrated protein were made and stored in 193K after freezing in Liquid N$_2$.
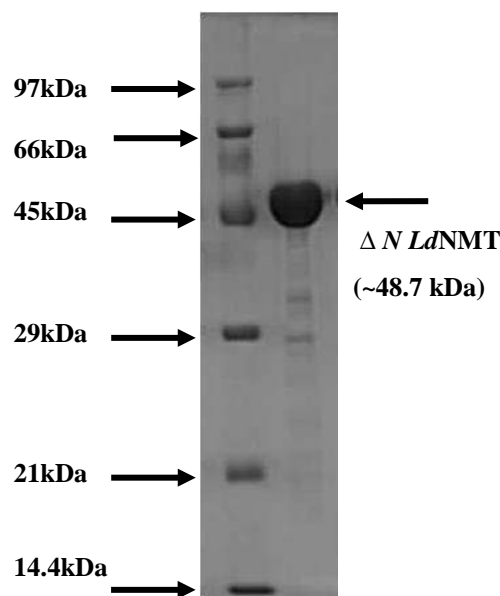
Figure 6.7: 12% SDS-PAGE showing final purified ΔN *Ld*NMT.

6.2.5. Crystallization of ΔN *Ld*NMT

ΔN *Ld*NMT protein was concentrated to 6 mg/ml in the presence of the non hydrolysable myristoyl CoA analogue (NHM or S-(2-oxo) pentadecyl-CoA). The final concentration of NHM was kept at 4 times the molar concentration of protein and the sample was kept overnight on ice. Crystallization was tried using available commercial sparse matrix screens from Hampton Research Ltd, Qiagen and Molecular Dimensions. Protein samples were centrifuged for 5 min at 10,000 rpm and 278 K immediately prior to the crystallization to ensure that samples were free from any particulate matter. Each experiment consisted of equilibrating a mixture of 150 nl protein solution and 150 nl screen solution over a reservoir of 60 nl screen solution in MRC 96-well plates (Griener bio-one, USA) using a Mosquito nano-litre high throughput robot manufactured by TTP labtech.

Initially, several different precipitants and buffers were used in crystallization attempts. Often, the crystals grew as multiple crystals or their diffraction quality was poor. Further optimization of hits by varying the concentration of precipitant and/or salt and drop size in 24 well hanging-drop vapour-diffusion experiments resulted in multiple crystals and/or single crystals (**Figure 6.8**). Single crystals used for the soaking experiment and subsequently for the data collection were grown by setting up

the hanging-drop vapour-diffusion method manually mixing 1.5 µl of protein at 6 mg/ml with 1 µl of well solution. Reproducible single crystals of diffraction quality were obtained primarily in two conditions of 2.0 M ammonium sulphate, 0.2 M sodium chloride in 1-2 weeks and 0.2 M tri-sodium citrate tetra hydrate, 25% PEG 3350, 0.1 M sodium acetate pH 4.6 in 3-4 weeks at 293K (**Figure 6.8**).
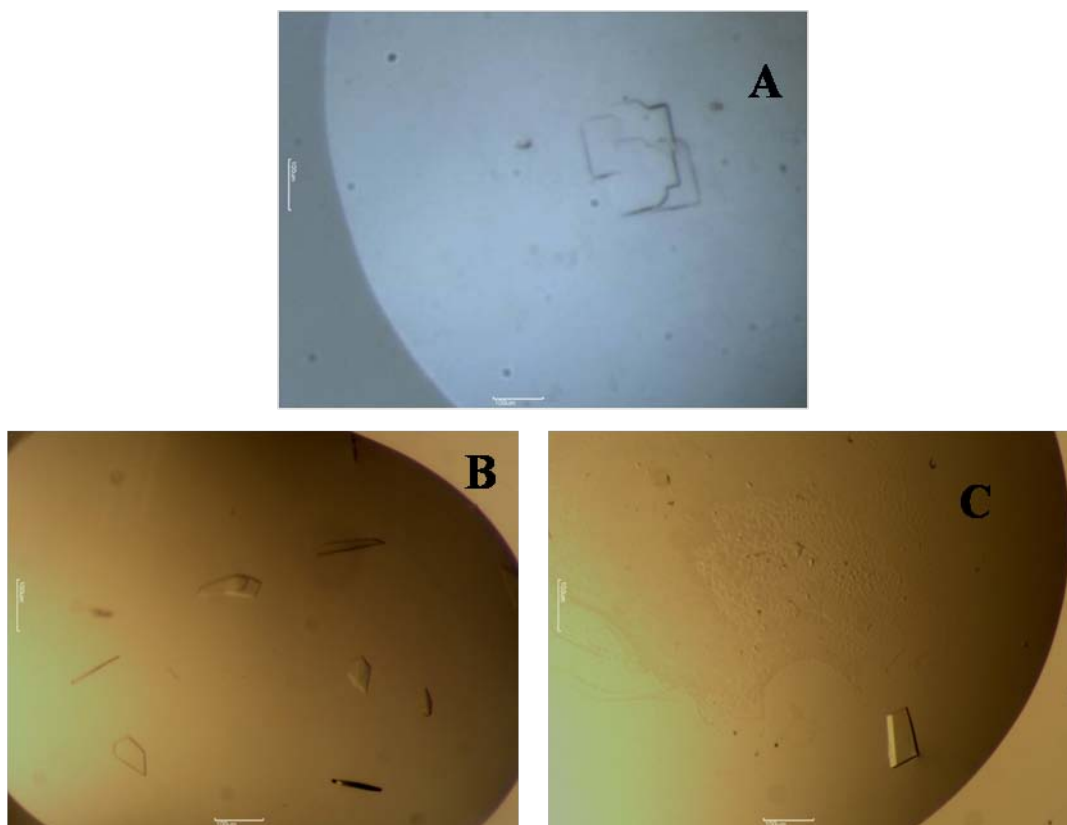
Figure 6.8: Crystals obtained in 24 well plate hanging drop-vapour diffusion in two different conditions for ΔN *Ld*NMT with its substrate incubated overnight; (A) Multiple crystals obtained for the mutant (B) Single crystals obtained in ammonium sulphate condition, (C) Single crystal obtained in PEG3350 condition.

6.2.6. Soaking of the Crystals in the Reported Inhibitor of *L.major* NMT and *T.brucei* NMT

DDD 85646 is a synthetic molecule containing a pyrazole sulphonamide scaffold known to inhibit the proliferation of the bloodstream form of *T. brucei* in culture (**Figure 6.9**) (Frearson *et al.,* 2010). Since *Lm*NMT has 74% overall sequence identity with *Tb*NMT and 94% identity within the peptide-binding site and ΔN *Ld*NMT is 98% identical to NMT from *Leishmania major*, crystals obtained for N-terminal

mutant of *Ld*NMT were soaked in the reported inhibitor of *Lm*NMT. This inhibitor (DDD 85646) is selective against peptide binding pocket for the *Lm*NMT and *Ld*NMT as well (*Ld*NMT IC$_{50}$= 3nM, Data not published). Crystals obtained from both the conditions were soaked in mother liquor containing 1-2 mM of DDD85646 for 30min and flash cooled in liquid nitrogen.
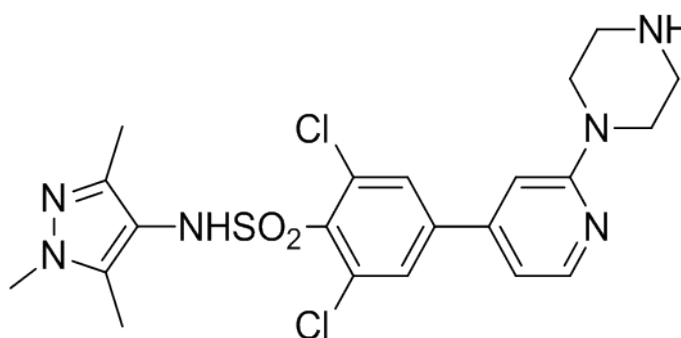
Figure 6.9: Representation of DDD85646 inhibitor

6.2.7. Data Collection for the Soaked Crystals

Diffraction data from crystals soaked with DDD85646 were collected at the Diamond Light Source synchrotron facility at Oxford, UK on beamline I02 (λ= 0.9795Å) using an ADSC Q315r CCD detector. Data for the crystals obtained from the ammonium sulphate condition were collected as 720 images with 0.25° oscillation and 0.25 sec exposure time per image and the crystals obtained from PEG3350 condition was collected as 400 images with 0.5° oscillation and 0.5 sec exposure per image (**Figure 6.10**). Data were found to be highly anisotropic from both the crystals and there were indications that the crystals were split. Moreover, diffraction showed overlapping spots at low resolution which mean possibility of twinning effect in the intensity distribution.

Data were automatically processed with xia2/XDS (Winter, 2010; Kabsch, 2010) and scaled using the SCALA programme in *CCP4* suite (Winn *et al.,* 2011). The crystal of *Ld*NMT mutant protein from ammonium sulphate condition is found to be triclinic P1 with unit cell *a* = 70.08Å, *b* = 79.79, *c* = 92.29Å, *α* =75.18° *β* = 73.07° *γ* = 72.68°, whereas crystals grown from PEG3350 crystallized in the monoclinic space group P2$_1$ with unit cell dimensions *a*= 92.58Å, *b*= 90.19Å, *c*= 121.71Å, *α*=*γ*= 90° *β*=

110.45°. The dataset obtained from P1 crystal was found to be twinned with twin fraction of 13%. The data collection and processing statistics are summarized in **Table 6.1**.
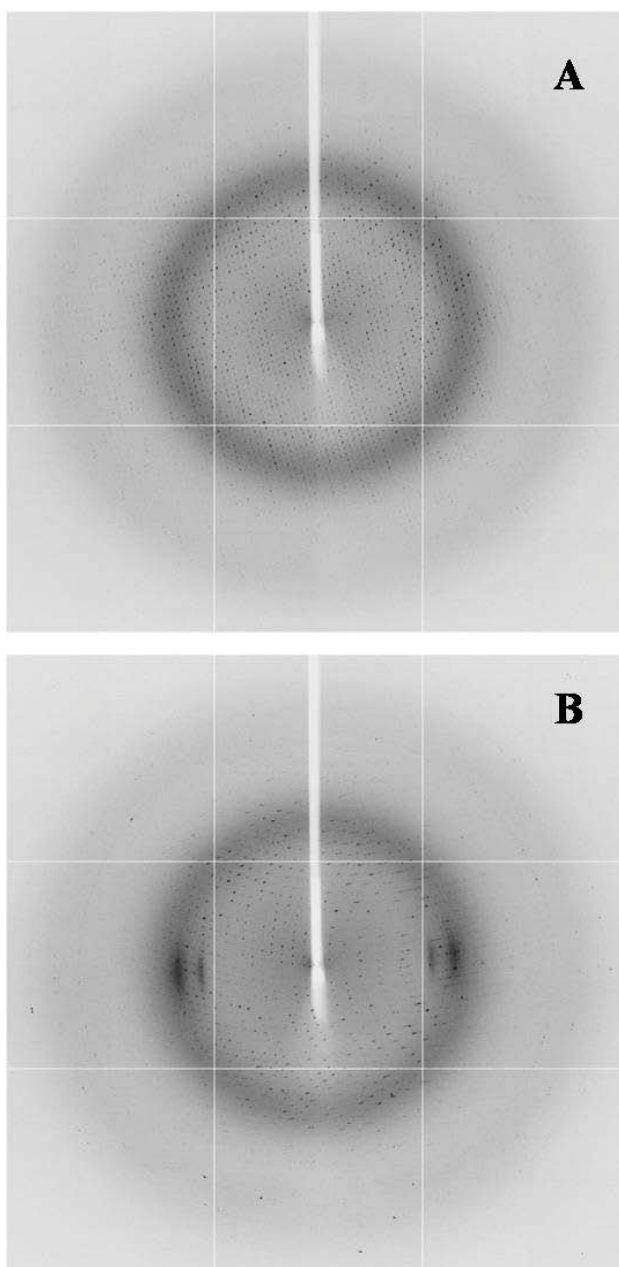
Figure 6.10: Diffraction images for two crystals obtained in two different conditions; (A) Diffraction image of the crystals obtained from the PEG3350 condition indexed in P2$_1$, (B) Diffraction image of the crystals obtained from the Ammonium sulphate condition indexed in P1.

Table 6.1: Data collection statistics for crystals obtained for ΔN *Ld*NMT

	Crystal form 1 with NHM	**Crystal form 2 with NHM**
Beam line	Diamond IO2	Diamond IO2
Wavelength	0.9795Å	0.9795Å
Space group	P2₁	P1
Unit cell dimensions (Å)	$a = 92.58$, $b = 90.19$, $c = 121.71$ $\alpha = \gamma = 90°$ $\beta = 110.45°$	$a = 70.08$, $b = 79.79$, $c = 92.29$ $\alpha = 75.18°$ $\beta = 73.07°$ $\gamma = 72.68°$
Resolution	1.8Å	2.23Å
% Solvent	49.7	49.7
Matthews Coefficient ($\mathring{A}^3$ Da^{-1})	2.44	2.45
Total Number of observations	695337 (94751)	170076 (24547)
Total number of unique observations	171280 (24396)	86979 (12571)
R_{merge} (%) [†]	12.3(87.4)	10.0(36.9)
Completeness (%)	99.0(97)	96.6(95.2)
Mean I/σ(I)	7.2(1.5)	6.3(2.1)
Multiplicity	4.1(3.9)	2.0(2.0)
Molecules in a.s.u	4	4

[†] $R_{merge} = \Sigma_{hkl}\Sigma_i |I_j (hkl) - \{I(hkl)\}|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, **where $I_i(hkl)$ is the intensity of an individual measurement of the reflection with Miller indices hkl and {I(hkl)} is the mean intensity of redundant measurements of that reflection.**

Numbers in the parentheses are for the last resolution shell

6.2.8. Solvent Content of Crystals

Considering molecular weight of the mutant as 48.7kDa, the calculated Matthews coefficient (Matthews, 1968) suggests four molecules in the asymmetric unit for both P1 and P2$_1$ datasets (2.45 $\mathring{A}^3$ Da^{-1} and 2.44 $\mathring{A}^3$ Da^{-1} corresponds to a solvent content of 49.7% respectively).

6.2.9. Structure Solution

Unique solutions were found for each of the two crystal forms taking coordinates of *L.major* NMT bound with substrate and inhibitor (DDD85646) using molecular replacement method in *CCP4* suite (PDB code: 2WSA, Frearson *et al.,* 2010). *MOLREP* (Vagin and Teplyakov, 1997) from *CCP4* suite using pdb coordinates of *Lm*NMT was used for the P2$_1$ dataset. *MOLREP* resulted in the detection of non-crystallographic translation vector of 0.000, -0.425, 0.500 (40% of the origin) for monoclinic dataset. *PHASER* (McCoy *et al*., 2007; Storoni *et al.*, 2004) from *CCP4* suite (Collaborative Computational Project, Number 4, 1994) was used as molecular replacement method in the case of triclinic dataset. As suggested by Matthews coefficient, both the resultant structures contained 4 molecules in the crystallographic asymmetric unit.

6.2.10. Refinement of the structures

The initial phases obtained from molecular replacement using the search model were improved by subsequent rigid body refinement keeping the sequence of the input model. Rigid body refinement using Refmac 5.6.0116 (Murshudov *et al.* 2011) gave the R$_{factor}$/ R$_{free}$ of ~0.42/ ~0.43 for P2$_1$ and ~36/35 for P1 structure. This was followed by several cycles of positional refinement using the data in the resolution shell of 48.67-1.8 $\mathring{A}$ (monoclinic), 86.76-2.3 $\mathring{A}$ (triclinic) resulting in the improvement of both R$_{factor}$ and R$_{free}$. Since both structures have more than one molecule in their asymmetric units, their initial models were refined taking advantage of the presence of non-crystallographic symmetry. Manual correction was done in *COOT* (Emsley & Cowtan, 2004). Individual B-factors were refined in the final cycles. Electron density maps were calculated at this stage. Side chains were mutated according to the sequence of

ΔN *Ld*NMT and their conformations were selected based on the observed density and their new positions were refined.

Water molecules were added using (Fo-Fc) map as reference. Automated water picking programme of the suite was employed for adding the water molecules during refinement. Water molecules were manually checked using following criteria.

(i) The electron density is observed at lσ level in the (2Fo-Fc) map and correspondingly in the (Fo-Fc) map at 5σ level initially and has been extended up to 3σ level during final cycles of refinement.

(ii) The distance between selected water molecule and interacting protein atoms or water molecules has to be between 1.8 to 3.2 Å. At least one hydrogen bond must be satisfied for selecting a water molecule. The appropriate geometry of hydrogen bonds around water molecule also has been considered.

ΔN *Ld*NMT was co-crystallised with NHM, a non hydrolysable analogue of myristoyl-CoA with a methylene group intervening between the sulphur and the carbonyl carbon in what would otherwise be the thioester linkage in myristoyl-CoA (**Figure 6.11**). A difference map calculated during refinement showed clear density for the NHM in both the structures (**Figure 6.12**). However conclusive density for the inhibitor was not seen in the difference map in any of the crystal forms (**Figure 6.13**).
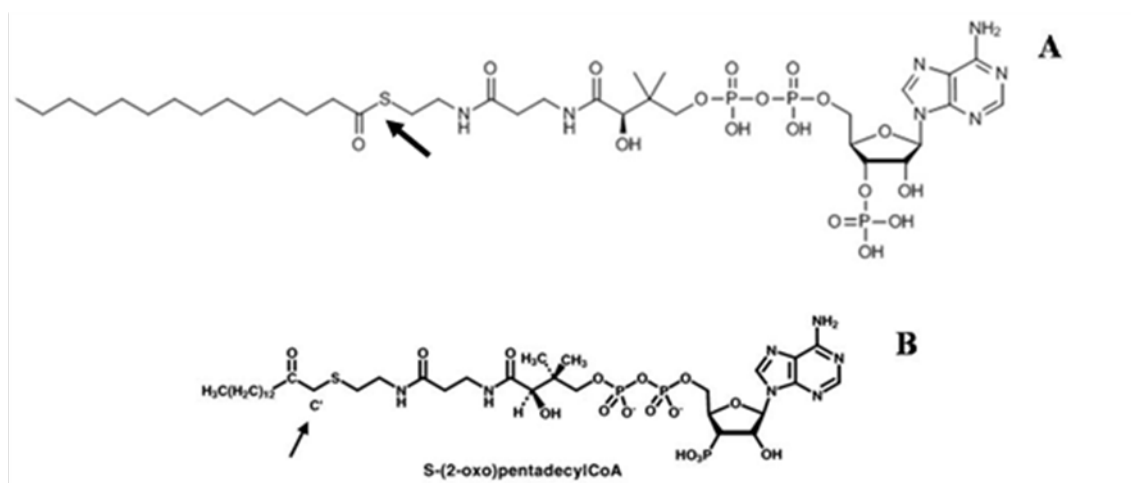
Figure 6.11: (A) Structure of Myristoyl-CoA. The arrow indicates the bond that is hydrolysed in the transfer of myristate to the protein subset. (B) Structure of S-(2-oxo) pentadecyl-CoA (NHM). The arrow shows the extra methylene group intervening between the sulphur and the carbonyl carbon in the Myristoyl-CoA.
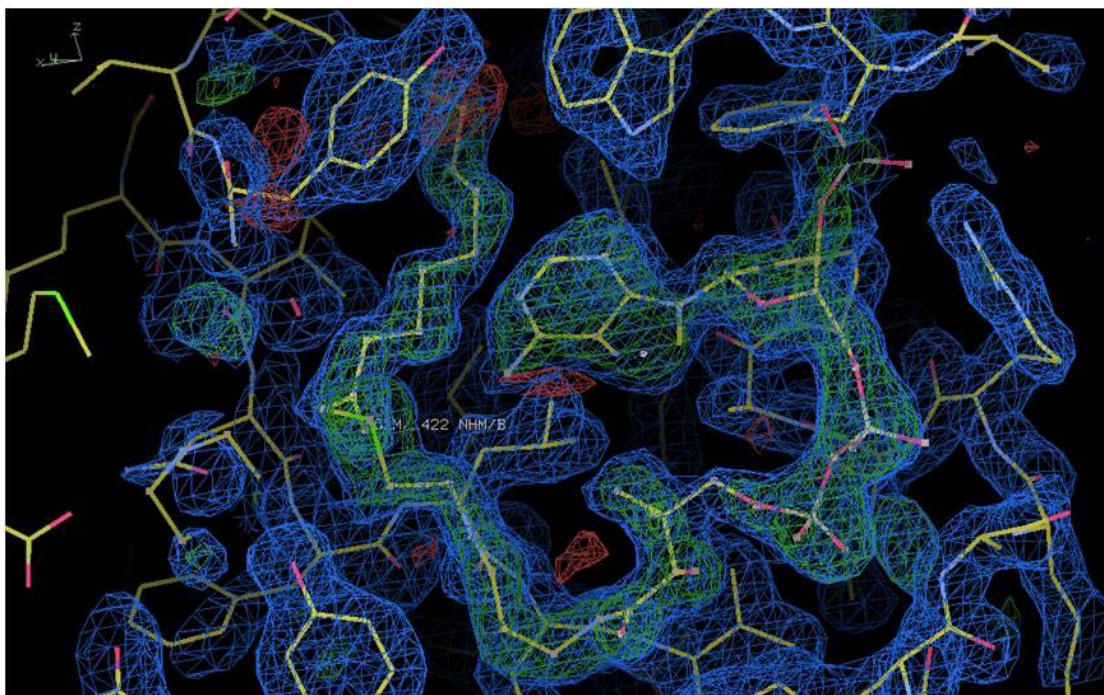
Figure 6.12: NHM density in Fo-Fc map at 3σ level showing presence of bound substrate in P2₁. Similar density was seen in all the four chains of monoclinic form and also in the triclinic form (not shown). The substrate was superimposed and added to the difference density obtained after refinement. Figure was prepared in COOT.
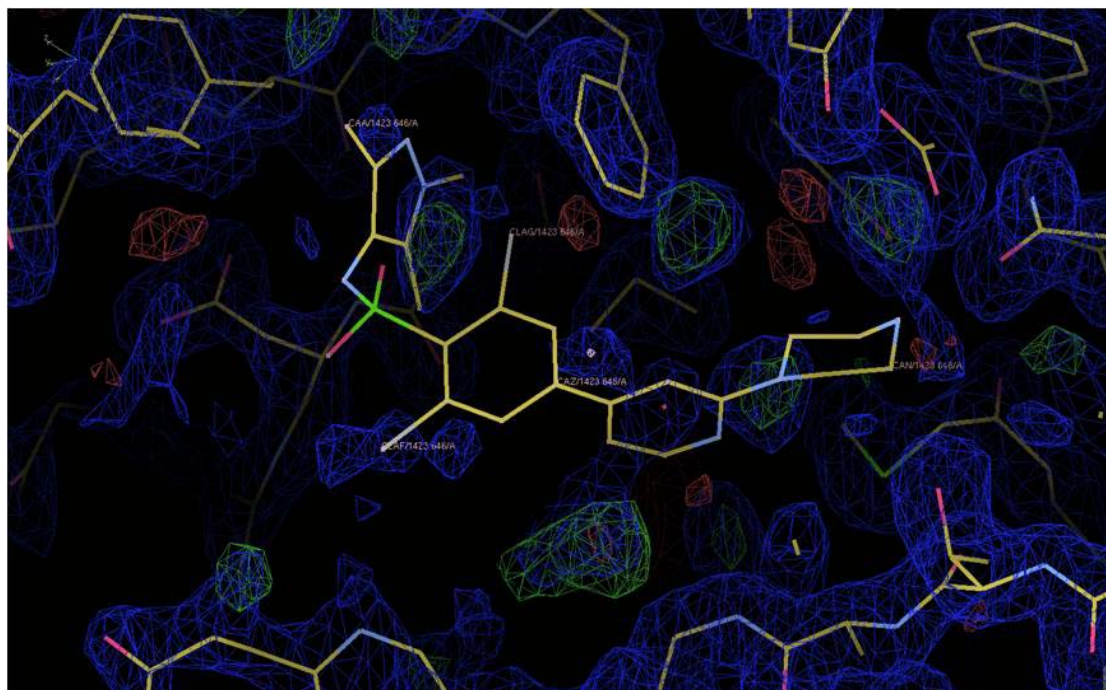
Figure 6.13: Inhibitor (DDD85646) superimposed on P2₁ electron density showing its absence in the crystal. The electron density for inhibitor is absent in all the four chains and also in the triclinic form (not shown). Figure was prepared in COOT.

6.2.11. Structure validation

During the final cycles of refinement the stereochemistry and the geometry of the models were checked using *PROCHECK* (Laskowski *et al.,* 1993). The overall G factors considered to be a measure of stereochemical quality of the model output by *PROCHECK* were -0.06 and -0.16 for the P2$_1$ and P1 crystal forms respectively. The values for G factors were within the limits expected for correct structures that are refined at 1.8Å and 2.23 Å resolution respectively. The refinement statistics and the values of refined parameters are presented in **Table 6.2**. The Ramachandran (φ, ψ) plots (Ramachandran & Sasisekharan, 1968) showed that most of the residues were placed in the most favoured or partially allowed regions of the map.

Table 6.2: Refinement Statistics for the ΔN *Ld*NMT crystals

Crystal System/Space Group	Monoclinic/P2$_1$	Triclinic/P1
Resolution range (Å)	48.72-1.80	86.76-2.23
Refinement Statistics		
R$_{cryst}$ [a]	0.254	0.293
R$_{free}$ [b]	0.296	0.327
Root Mean Square Deviation		
Bond Length (Å)	0.017	0.0133
Bond Angle (°)	1.8337	1.841
Chiral Volume	0.1155	0.327
Ramachandran Plot Analysis (% Residues)		
Most favoured region	90.1	86.6
Additionally allowed region	9.4	12.7
Generously allowed region	0.5	0.5
Disallowed region	0	0.2

[a] R$_{cryst}$= Σ |Fo-Fc |/ ΣFo

[b] R$_{free}$ = Σ |Fo-Fc |/ ΣFo where the F values are test set amplitudes (5 %) not used in refinement.

In addition to the anisotropy in both the datasets, the splitting of the crystals and non-crystallographic translation in monoclinic and possibility of twining in triclinic

crystal form also may be contributing to the comparatively higher values for R-factors, R_{cryst} and R_{free}.

6.2.12. Description of the Structure

In describing the structure, the secondary-structure nomenclature used for *C. albicans* NMT (Weston *et al.,* 1998), *S. cerevisiae* NMT (Bhatnagar *et al.,* 1998) and *L. donovani* wilt type structure (Brannigan *et al.,* 2010) has been adopted in which helices are denoted with uppercase letters and strands with lowercase letters (**Figure 6.13**). Secondary-structure elements additional to those observed in the *C. albicans* NMT (*Ca*NMT) are denoted by one or more primes (').
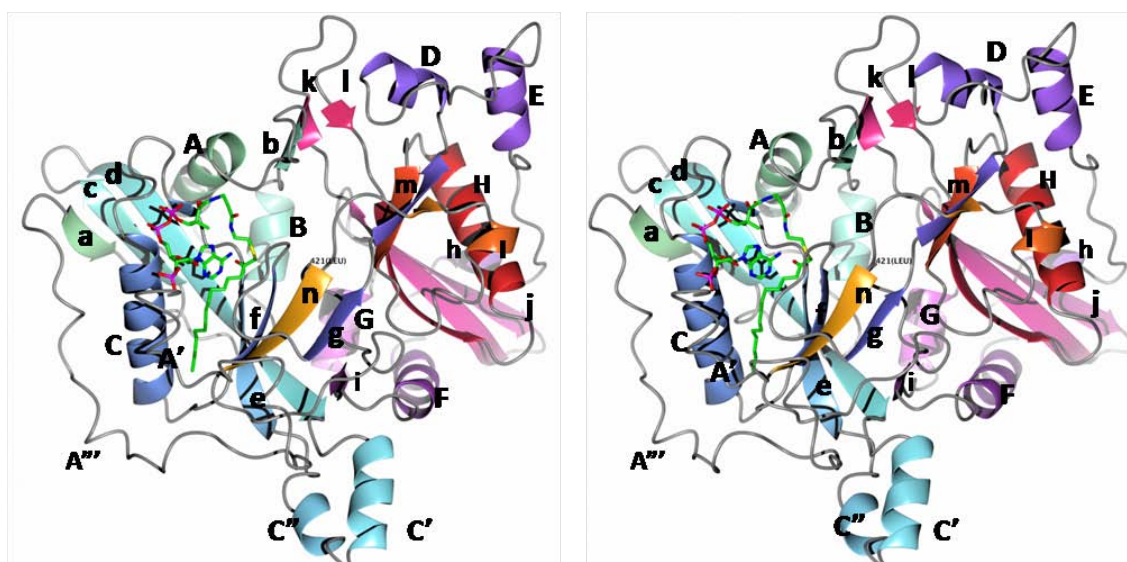
Figure 6.13: The three-dimensional structure of LdNMT. Stereo ribbon representation of LdNMT in its complex with NHM, which is shown in cylinder format and coloured by atom type: carbon, cyan; oxygen, red; nitrogen, blue, sulfur, yellow; phosphorus, magenta. The protein chain is colour ramped from residue 11 at the N-terminus (red) to residue 421 at the C-terminus (magenta). The secondary-structure elements are labelled with uppercase letters for α-helices and lowercase letters for β-strands. The ' and '' superscripts indicate elements additional to those observed in the *C. albicans* NMT (*Ca*NMT, Weston *et al.,* 1998).

Both the crystal forms are very similar to the wild type *Ld*NMT structure with rmsd of 0.96Å for 402 superposed Cα atoms (**Figure 6.14**) with volume occupied by the protein is 227669 $Å^3$. The NHM was also found to be bound with similar

conformation as in wild type *Ld*NMT. Like the wild type, the molecule showed two lobes, each of which has a multistranded β-sheet element at its core flanked by two α-helices on one face and a single α-helix on the other face (**Figure 6.13**). For the N-terminal lobe, the sheet is made up of strands "acdefng" with helices "A" and "B" packed onto one face and helix "C" onto the other. For the C-terminal lobe, the sheet comprises strands "gmlkjh" against one face of which helices "F" and "G" are packed, with helix "H" packed onto the other face.
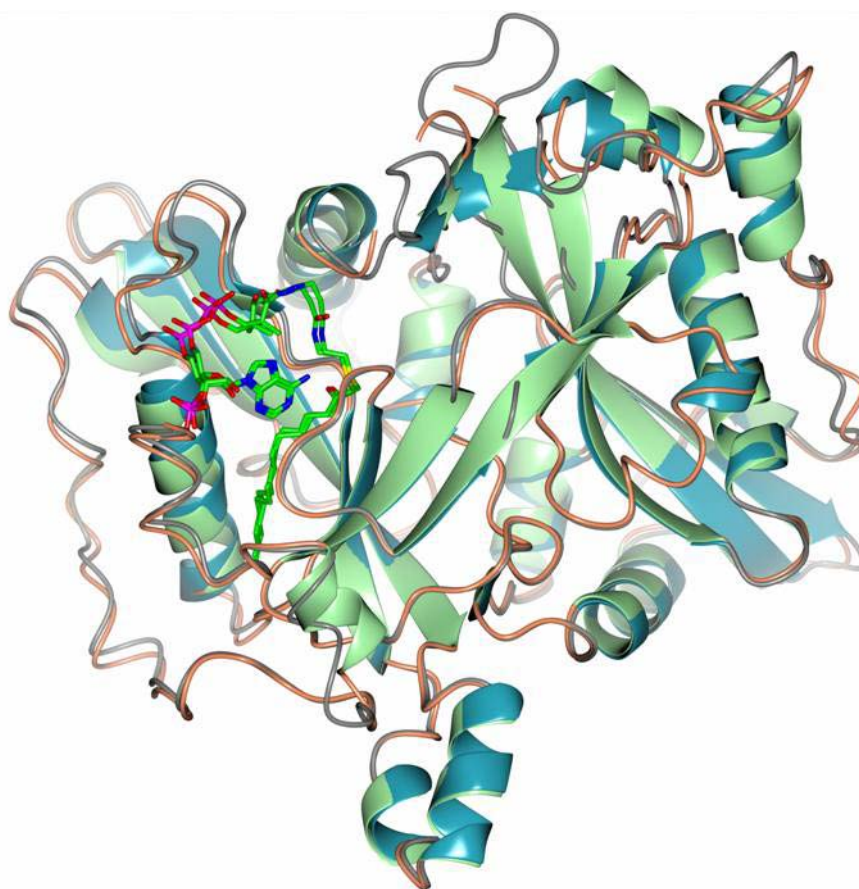
Figure 6.14: Comparison of the structures of *Ld*NMT and ∆N *Ld*NMT. The structures of *Ld*NMT–NHM (PDB: 2WUU, Brannigan *et al.,* 2010) and ∆N *Ld*NMT–NHM (P2₁ form) were superposed using the SSM superpose routine in CCP4MG. The structures are displayed as ribbons, *Ld*NMT (light green with turns coloured as coral) and ∆N *Ld*NMT (Light cyan with turns coloured grey), with the ligands represented in cylinder format and coloured by atom type: carbon, cyan; oxygen, red; nitrogen, blue, sulfur, yellow; phosphorus, magenta.

PROMOTIF analysis using *PDBsum* (http://www.ebi.ac.uk/pdbsum/) (Hutchinson & Thornton, 1996) showed the monoclinic form of the ΔN *Ld*NMT consist total of 2 βsheets with 16 strands and 18 α- helices. The secondary structure and topology diagram obtained from *PDBSum* are shown in **Figure 6.14** and **6.15**.
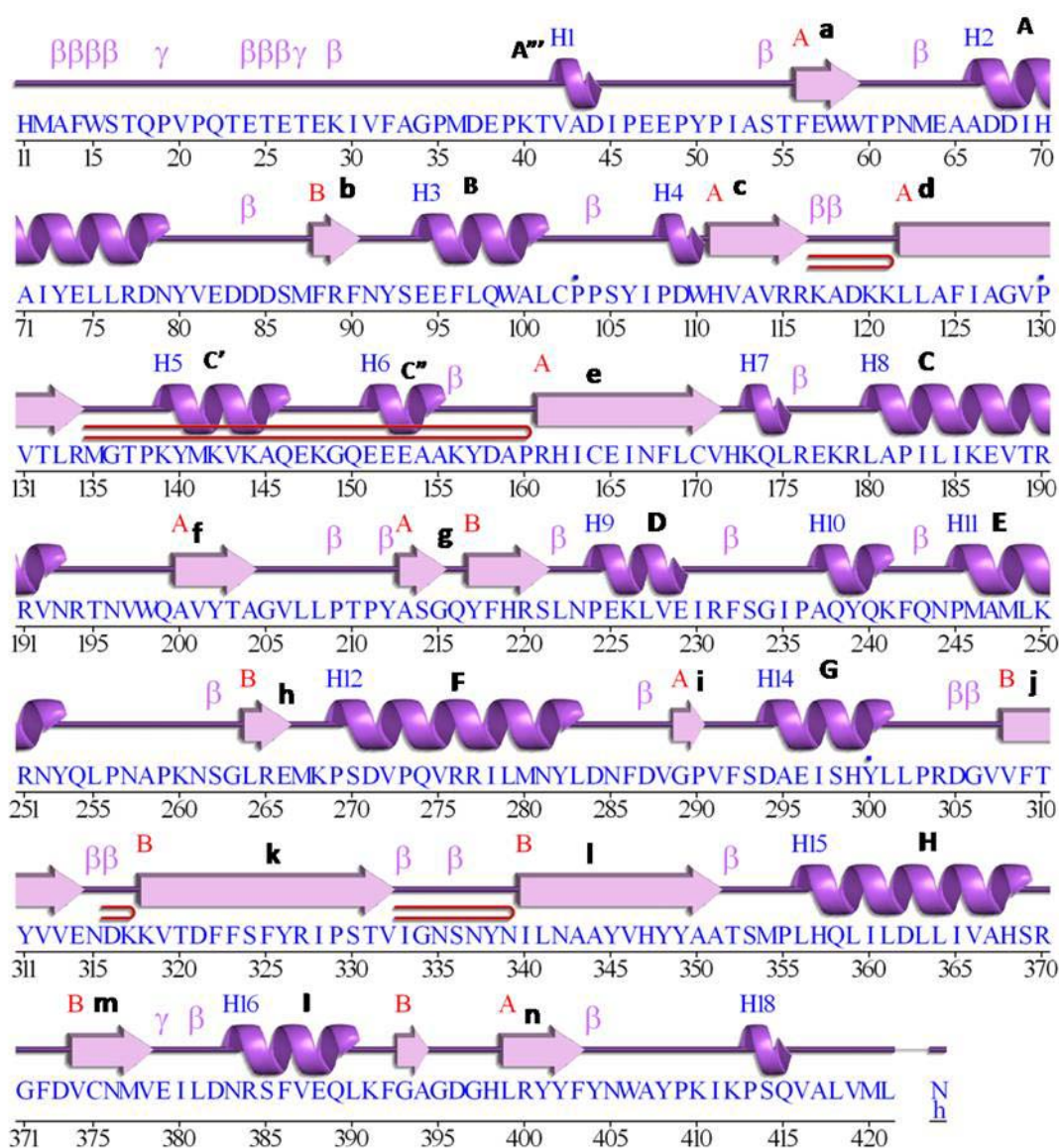
Figure 6.14: Arrangement of secondary structure elements in ΔN *Ld*NMT monoclinic form.
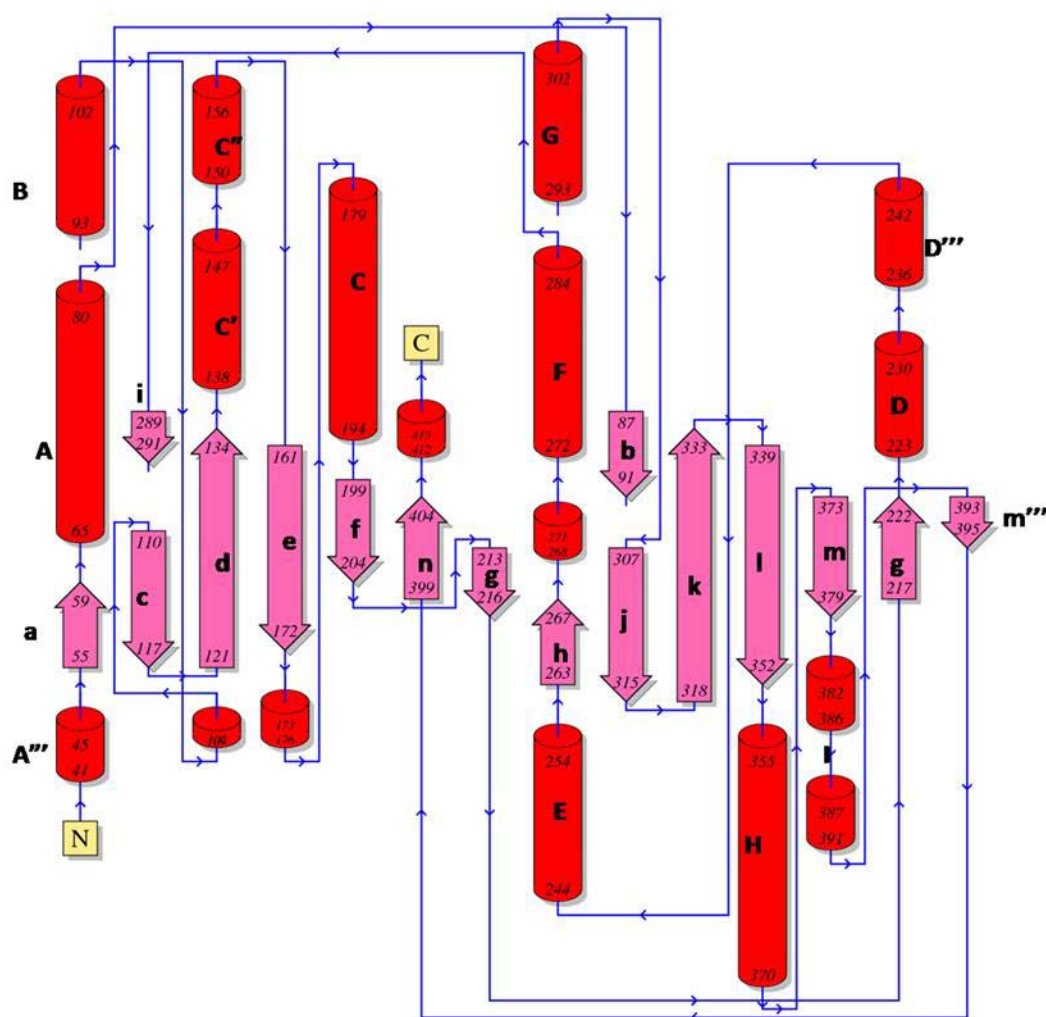
Figure 6.15: Topology diagram of ΔN *Ld*NMT with secondary-structure elements labelled and with α- helices represented as cylinders and β-strands as arrows

Most of the residues in the structure have mean atomic temperature factor for main-chain atoms as calculated from a web based suite for Protein Structure Analysis Package (PSAP, http://iris.physics.iisc.ernet.in/psap/) is low at 15-20 Å^2 with average of 23.1 Å^2 for entire protein. The regions corresponding to surface loops and loops "Ab" "DE" and "kl" have residues with significantly higher average main-chain temperature factors, suggesting that these regions of the structure have independent mobility.

Structurally deduced mechanism of the N-terminal Myristoylation of subset of proteins by NMT enzymes requires the displacement upon myristoyl-CoA binding of the "Ab" loop, which partially occludes the binding site for the N-terminal peptide of the substrate (Bhatnagar *et al.,* 1998; Farazi *et al.,* 2001; Brannigan *et al.,* 2010). This loop, which serves as a lid that opens and closes over the catalytic site, reported to be having variable conformations among the different structures. It partially occludes the myristoyl-CoA and peptide binding sites in the structure of the uncomplexed *Ca*NMT, while in the binary and ternary complexes of *Sc*NMT the loop has swung upward to different extents, opening up the active site. In this way, the binding of myristoyl-CoA assists peptide binding by ensuring the lid is open and the second substrate can bind unhindered. The different conformation of this "Ab" loop as reported in apo-enzyme and binary and ternary complexes has been depicted in the **Figure 6.16** with ΔN *Ld*NMT superimposed. ΔN *Ld*NMT structure showed the loop conformation in between the fully open conformation of *Sc*NMT ternary complex and closed conformation of apo-enzyme structure of *Ca*NMT.
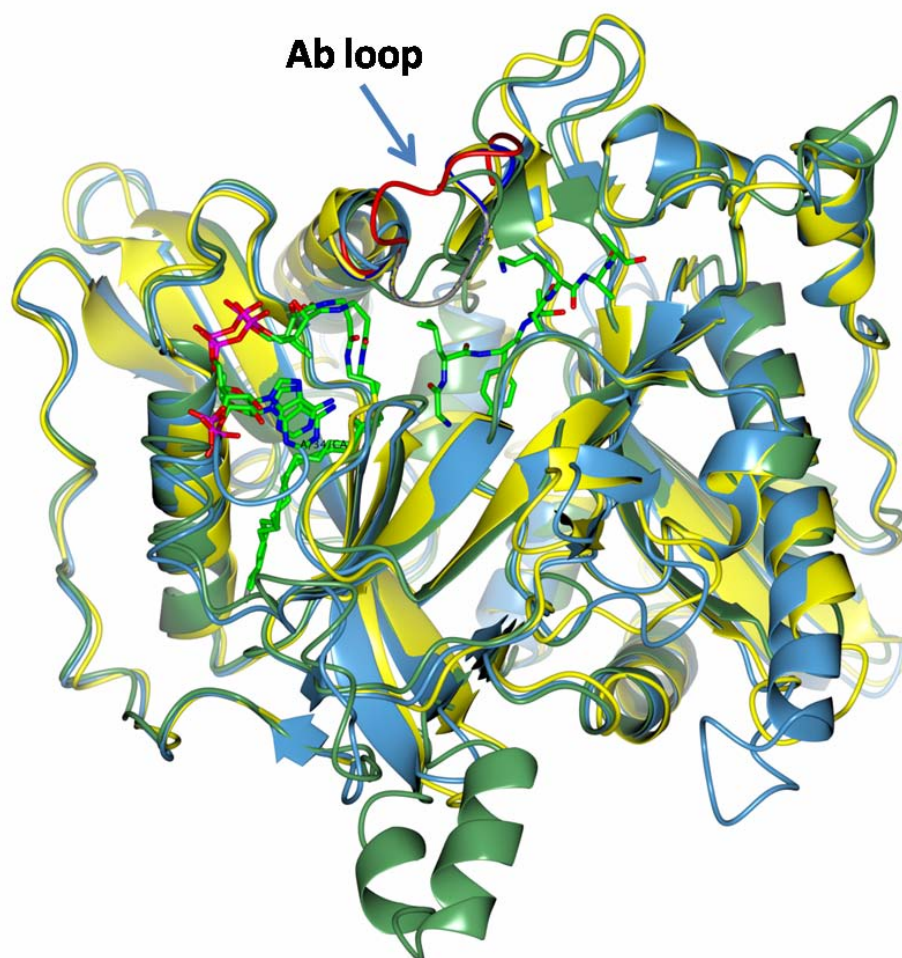
Figure 6.16: Different conformation of Ab loop as shown in apo-enzyme structure (*Ca*NMT PDB code: 1NMT, Weston *et al.*, 1998) and in the ternary complex (*Sc*NMT–NHM–GLYASKLA, PDB code: 1IID, Farazi *et al.*, 2001) superimposed with binary complex of ∆N *Ld*NMT-NHM superposed using the SSM superpose routine in CCP4MG. The structures are displayed as ribbons ∆N *Ld*NMT (green), *Ca*NMT (yellow) and *Sc*NMT (blue), with the ligands represented in cylinder format. The Ab loops from the superposed *Ca*NMT apo-enzyme, ∆N *Ld*NMT-binary and *Sc*NMT–NHM ternary complex are shown in grey, green and red repectively.

Out of 4 chains in ∆N *Ld*NMT structure chain A and chain B contains good electron density for the residues forming "Ab" loop but having much higher than average temperature factors, whereas, chain C and D shows no good density for most of the loop residues, suggesting intrinsic mobility of this segment of the structure. Taken together, it seems that the "Ab" loop is closed in the uncomplexed enzyme and

that following fatty acyl-CoA binding the "Ab" loop adopts an ensemble of structures that facilitate the binding of the N-terminal peptide of the substrate protein or inhibitor mimicking the N-terminal peptide.
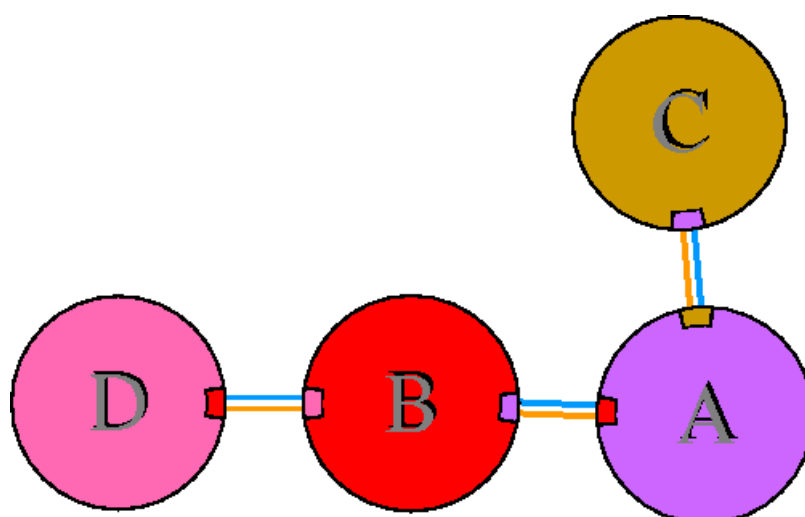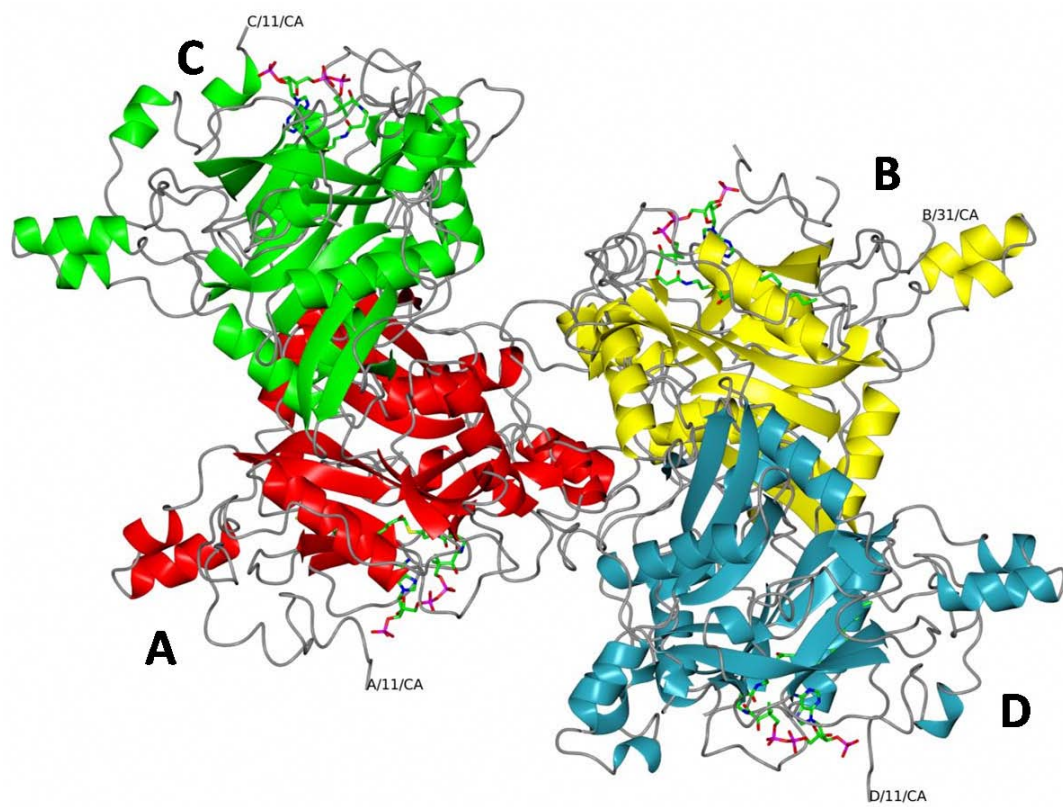
Figure 6.17: (A) Packing arrangement of the 4 chains in asymmetric unit of ΔN *Ld*NMT, (B) schematic diagram of interactions between protein chains of ΔN *Ld*NMT.

Table 6.3: Interface statistics for the four chains of ∆N*Ld*NMT

Chains	No. of interface residues	Interface area (Å2)	No. of salt bridges	No. of disulphide bonds	No. of hydrogen bonds	No. of non-bonded contacts
A : B	17:15	786:822	-	-	7	93
A : C	17:16	913:922	-	-	7	69
B : D	16:16	865:851	-	-	5	62

6.2.13. Co-Crystallization of ∆N *Ld*NMT with substrate and inhibitor

Various commercially available screens were tried for co-crystallization of ∆N *Ld*NMT with substrate and inhibitor. Condition G5 (20% PEG3350, 0.2 M sodium nitrate, 0.1 M Bis Tris propane pH 7.5) from PACT screen (Qiagen) gave small but single crystals (**Figure 6.18**). Fine screening of the condition varying either precipitant concentration or pH at a time gave single crystals in many conditions (**Figure 6.19**). However these crystals were found to be diffracting poorly to 12 Å even at synchrotron facility.

Figure 6.18: Original hit obtained for co crystallised ΔN *Ld*NMT with substrate and inhibitor added and incubated overnight.

21% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.1

19% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.5

19% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.3

18% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.3

19% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.6

19% PEG3350, 0.2M Sod. Nitrate, 0.1M Bis Tris Propane pH 7.3
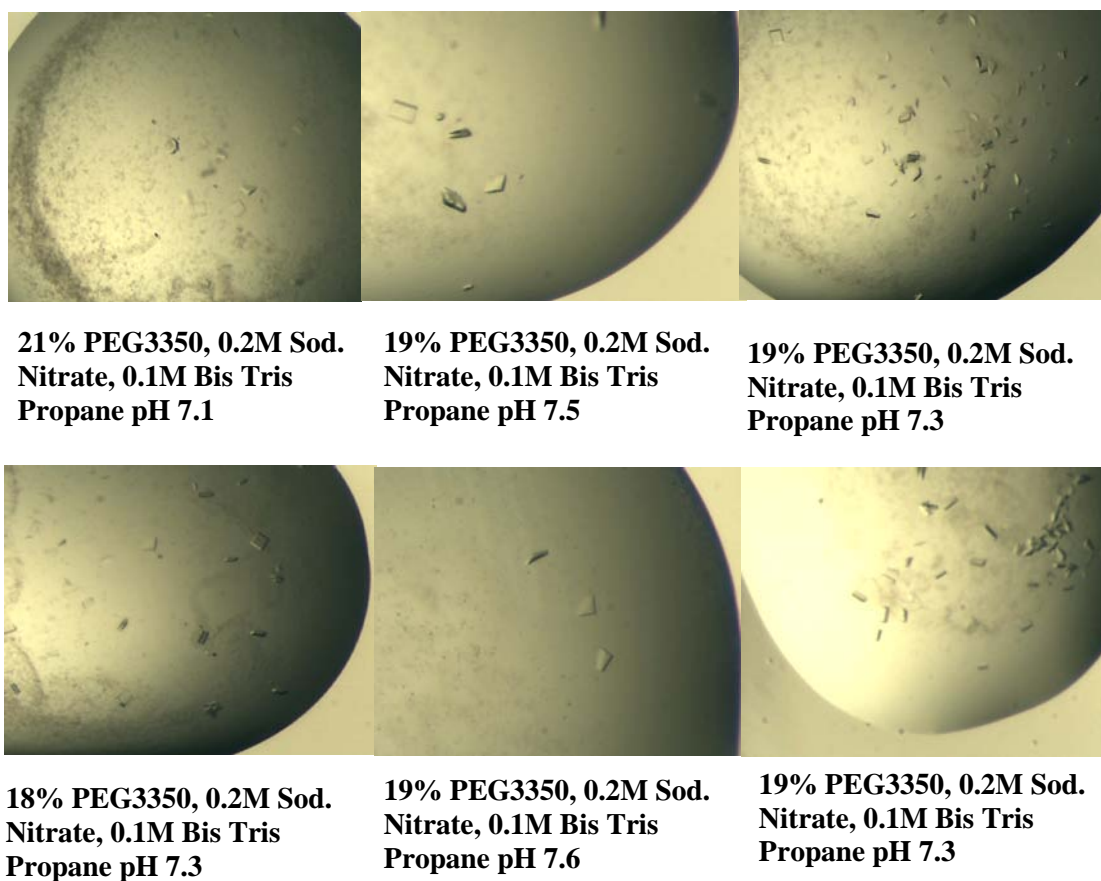
Figure 6.19: Various crystals obtained in fine screen after optimization of initial hit obtained for ΔN *Ld*NMT with substrate and inhibitor added.

6.3. DISCUSSION

NMT has been shown to be essential for the survival for many parasites making it a potential drug target against many parasitic infections. Targeted gene deletion methods showed NMT is likely to be an essential gene in the *L. donovani* life cycle, thus presenting NMT from *L. donovani* as a suitable target for drug discovery for deadly disease Visceral Leishmaniasis (Brannigan *et al.*, 2010). Industrial antifungal drug discovery programmes, high throughput screening and *in vivo* results have led to the identification of many highly potent selective inhibitors against various fungal and protozoan parasites. Comparative sequence and biochemical analyses as well as structural studies have demonstrated high conservation of the myristoyl-CoA binding sites across species, and the interactions of fatty acyl- CoA with the enzyme are conserved in the structures that have been determined. Attention has therefore been directed toward the development of inhibitors that bind in the peptide binding site. There is conservation of sequences in the peptide binding groove; nevertheless, selective inhibitors of yeast and fungal NMTs have been discovered that bind in this location, exploiting residue differences at a handful of positions and often subtle differences in structure (Bhatnagar *et al.,* 1998; Sogabe *et al.,* 2002). The determination of the crystal structure of the *Ld*NMT–NHM complex (Brannigan *et al.*, 2010) established a foundation both for the structural analysis of inhibitor complexes and for structure-assisted drug discovery. Comparison of *Ld*NMT-NHM bound structure with high-resolution structure of HsNMT-Myristoyl Co-A (PDB code: 3IU1; Structural Genomics Consortium, unpublished data) has shown differences in the loops (e.g., "Ab", "kl" and "DE") forming the upper surface of the peptide binding site. Non conservation of predicted residue in *Ld*NMT and *Hs*NMT, binding to peptide substrate or inhibitor and residues further to that may present an opportunity for selective inhibitor action. Also, the residues predicted to form the peptide binding site in *L. donovani* NMT is different in the NMTs of *T. brucei* and *T. cruzi* at two and four positions respectively (Brannigan *et al.*, 2010).

Recent availability of the structures of *Hs*NMT in complex with myristoyl-CoA and three different polycyclic inhibitor molecules (PDB codes: 3IU2, 3JTK and 3IWE, Structural Genomics Consortium, unpublished data), one of which crystallized with same DDD85646 inhibitor, can also present useful information determining the

differences in the conformations and interactions of these inhibitors in different pathogens and in human, hence useful in selective structure based drug discovery efforts.

However, failure to soak in the reported inhibitor (DDD85646) for *T. Brucei* (Frearson *et al.,* 2010) and *Ld*NMT (*Ld*NMT IC$_{50}$= 3nM, Data not published) to the new construct (ΔN *Ld*NMT) and poor diffraction of the co-crystallized crystals shows more considerable efforts to be made before gaining the insight to the binding mode and selective efficacy of this inhibitor with the *Ld*NMT.

Chapter: 7

Conclusions

7.1. INTRODUCTION

Three pharmaceutically and medicinally important enzymes, two of them bacterial penicillin G acylases (PGAs) and the third one N-myristoyltransferase from a protozoan *Leishmania donovani*, have been studied using various biophysical and biochemical techniques. The action of all three enzyme's concerns amide bond. PGAs hydrolyse the amide bond in the natural penicillins, useful in commercial production of the intermediate compound 6-APA (a β-lactam nucleus) used as a precursor molecule in the preparation of semi-synthetic antibiotics. Further, under a different condition the PGA condenses the appropriate D-amino acid derivative with 6-APA synthesising semi-synthetic antibiotics. The other enzyme, Myristoyl-CoA:protein N-myristoyltransferase (NMT), catalyses the co- and post-translational addition of the C14:0 fatty acid (myristate) to the N-terminal glycine residue of a subset of proteins via an amide-bond generation. PGA has been studied for its biotechnological and pharmaceutical application and NMT as a potential target for the development of antifungal, antiviral, anticancer, and antiparasitic compounds.

Two of these enzymes, PGA from *Kluyvera citrophila* and NMT from *Leishmania donovani*, were cloned, expressed and purified by the author. Whereas the three-dimensional structures of all three enzymes were determined and described here. Some of them were studied using various biochemical and biophysical techniques also. Main objective in carrying out these investigations was to understand the structure-function relation of these enzymes that can improve their application potential as catalyst in industry and as target in therapeutics.

7.2. CONCLUSIONS

From the view point of the pharmaceutical Industry, PGA from *Kluyvera citrophila* (*Kc*PGA) has been reported to be better than the currently used *Ec*PGA, due to its numerous industrial processes friendly properties, namely, ease of immobilization and endurable stability towards extreme conditions of temperature, pH, and presence of organic solvents (Alvaro *et al*., 1992; Fernandez-Lafuente *et al.,* 1991, 1996; Liu *et al.,* 2006). This enzyme was successfully cloned and expressed in high amount in *E.coli* pLysS cells. Recombinant protein was found to be autocatalytically processed with highly soluble monodisperse monomer and a fully active enzyme with expected

molecular mass of 23.6 kDa and 62.9 kDa correspond to α and β subunits, respectively on the SDS-PAGE. The processed recombinant enzyme was found to be highly active against penicillin G at 65 °C; however, the enzyme was not stable beyond 55 °C as inferred from activity assay as well as from fluorescence studies. Aggregation of the protein beyond 55 °C was confirmed by light scattering and ANS binding studies. Similar to other known PGAs the processed recombinant enzyme was stable in the pH range 4.0-10.0 as inferred from activity assays, intrinsic fluorescence and ANS binding studies. *Kc*PGA is found to be highly specific towards penicillin G compared to other natural penicillins. Retention of activity even after 24 h of incubation with 10 mM EDTA showed that metal ions in general had no role in enzyme activity; however, most of the heavy atom compounds inhibited the enzymatic activity. The enzyme followed the normal Michaelis–Menten kinetics, but the value of K_m was higher than that reported in the literature for penicillin G as substrate. The substrate available for the experiments was not completely pure. Under the native conditions, lower quenching with charged quenchers (I^- and Cs^+), compared to a neutral and penetrative quencher like acrylamide, indicate most of the tryptophan residues present in buried hydrophobic core in protein interior. Increase in the extent of quenching by quenchers with denaturation of the protein indicated increased accessibility of tryptophan residues to the quenchers when protein was unfolded. However, quenching by Cs^+ being low in both native and denatured state has implications to the presence of positively charged residues in the vicinity of the exposed (or partially exposed) tryptophan residues.

Contrary to report on *Ec*PGA (Choi *et al*., 1992), the *Kc*PGA Sβ1C mutant resulted in a slow processing precursor of size ~93 kDa as opposed to the stable precursor in the case of *Ec*PGA. Incubation in different pH for different intervals of time showed the pH dependence of the processing mechanism. The pH dependant autoprocessing in absence of nucleophile serine support the hypothesis that more than one residue may be involved in autocatalytic processing (Lee *et al*., 2000). More data may be required to further confirm this. We managed to grow the crystals of slow processing mutant in precursor form. The two crystal forms in space groups P1 and C2 diffracted up to 2.5 Å and 3.5 Å resolution, respectively. Three-dimensional structure obtained of the precursor mutant showed good resemblance with precursor structure of *Ec*PGA with r.m.s. deviations for Cα atoms of 0.62 Å along with the presence of intact

spacer peptide and Ca^{2+} at similar positions based on the difference electron density map. Looking at the similarity of structures and amino acid sequences a similar intramolecular autocatalytic processing mechanism can be expected for *Kc*PGA as well. However, presently, the poor completion and resolution of the diffraction data make the observation data insufficient to comment on the exact points of first and subsequent cleavages of peptide bonds in spacer.

The other pharmaceutically important enzyme studied in this thesis was a thermostable PGA from *Alcaligenes faecalis* (Verhaert *et al.*, 1997). This enzyme received attention because of the high synthetic efficiency for enzymatic hydrolysis in enantioselective synthesis. Enzyme also showed the highest specificity constant (k_{cat}/K_m) for hydrolysis of penicillin G compared to homologous enzymes from other bacterial sources (Svedas *et al.*, 1997). *Af*PGA has advantage over other well-characterized PGAs for industrial applications because of higher thermostability, by the presence of a disulphide bridge (absent in *E. coli* enzyme) as proposed by Verhaert *et al*. (1997) using biochemical and sequence comparison studies. However, the absence of structural information has prompted us to determine the three-dimensional structure of this enzyme. The enzyme crystallized in two different crystal forms decided by the amount of detergent added to the crystallization solution. The structure of the *Af*PGA reported here at 3.3 Å and 3.5 Å resolutions has similarity with enzymes from other two sources with r.m.s. Cα deviations less than 1.3 Å. The predicted disulfide bridge in the *Af*PGA structure between cysteine residues β492 and β525 has been located. Other structural and molecular features important for thermostability such as increased number of salt bridges, decreased content of uncharged polar amino acids (N+Q+S+T), higher number of arginines and a larger Arg/ (Arg+ Lys) ratio, larger number of prolines, increase in the net content of ionic residues were compared with PGA enzymes from *E. coli and P. rettgeri* to show that they really are responsible for its thermostability. Information on these stability factors will be useful for engineering penicillin acylases for industrial applications.

The third enzyme, therapeutically important, that was studied in this thesis was Myristoyl-CoA:protein N-myristoyltransferase (NMT) from a protozoan parasite *Leishmania donovani*. The enzyme shown to be coded by an essential gene in the *L. donovani* life cycle, is a suitable target for drug discovery for the deadly disease

Visceral Leishmaniasis caused by the parasite (Brannigan *et al*., 2010). Comparative sequence and biochemical analyses as well as structural studies have demonstrated high conservation of the myristoyl-CoA binding sites across species. Attention has therefore been directed toward the development of inhibitors that bind in the peptide binding site. Identification of highly potent inhibitor selective against the less conserved peptide binding site of the protozoan parasites and studying its selective toxicity toward the parasite through X-ray crystallography is an important step in the development of a drug candidate. With the aim of exploiting *Ld*NMT as a potent drug target, we successfully cloned an N-terminal mutant of the enzyme and expressed in the *E.coli* cells. The enzyme was crystallized in two different crystal forms of P21 and P1 from two different crystallization conditions in the presence of the non-hydrolysable myristoyl-CoA analogue (NHM). Crystals obtained were soaked into the reported inhibitor for NMT from *T. brucei* that has been shown to be binding in peptide binding site of *Leishmania major* NMT (Frearson *et al*., 2010). Analysis of X-ray diffraction data from the soaked crystals showed excellent density for the myristoyl-CoA analogue but no conclusive density for the inhibitor. The structures of the mutant solved in this thesis showed high similarity with the reported structure of wild type enzyme (PDB code: 2WUU, Brannigan *et al*., 2010) with high conservation of the interactions of NHM with the protein residues. The r.m.s. deviations of monoclinic form with wild type structure is 0.96 Å for 402 superposed Cα atoms. The trials of enzyme by co-crystallizing with both NHM and the inhibitor together resulted in poor diffracting crystals. Nevertheless, the successful crystallization of the mutant protein with non-hydrolysable myristoyl-CoA analogue alone as well as with added inhibitor, may be a starting point for the screening of inhibitors in the future.

Thus, the structure, stability, activity and structure-function relation of three proteins have been analysed using various biochemical and biophysical techniques that has provided information for further protein engineering to improve their application potential or for deigning better inhibitors for therapeutic purpose.

BIBLIOGRAPHY

❖ Abadjieva, A., Hilven, P., Pauwels, K. & Crabeel, M. (2000). The yeast ARG7 gene product is autoproteolyzed to two subunit peptides, yielding active ornithine acetyltransferase. *J. Biol. Chem.* **275**, 11361-11367.

❖ Abraham, E. P. (1981). The beta-lactam antibiotics. *Sci. Am.* **244**, 76-86.

❖ Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **66**, 213-221.

❖ Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P.D. (2005). *phenix.refine*. *CCP4 Newsl.* **42**, contribution 8.

❖ Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **68**, 352-367.

❖ Aitken, A., Cohen, P., Santikarn, S., Williams, D. H., Calder, A. G., Smith, A. & Klee, C. B. (1982). Identification of the NH_2 terminal blocking group of calcineurin B as myristic acid. *FEBS Lett.* **150**, 314-318.

❖ Albani, J. R. (2004). Structure and Dynamics of Macromolecules: Absorption and Fluorescence Studies. Elsevier Science.

❖ Alkema, W. B. L., Hensgens, C. M. H., Kroezinga, E. H., de Vries, E., Floris, R., van der Laan, J.-M., Dijkstra, B. W. & Janssen, D. B. (2000). Characterization of the beta-lactam binding site of penicillin acylase of *Escherichia coli* by structural and site-directed mutagenesis studies. *Protein Eng.* **13**, 857-863.

❖ Alkema, W. B. L., Hensgens, C. M. H., Snijder, H. J., Keizer, E., Dijkstra, B. W. & Janssen, D. B. (2004). Structural and kinetic studies on ligand binding in wild-type and active-site mutants of penicillin acylase. *Prot. Eng. Des. Sel.* **17**, 473-480.

❖ Alvaro, G., Fernandez-lafuente, R., Rosell, C. M., Blanco, R. M., Garcialopez, J. L. & Guisan, J. M. (1992). Penicillin G Acylase from *Kluyvera citrophila*- new choice as Industrial enzyme. *Biotechnol. Lett.* **14**, 285-290.

❖ Ambedkar, S. S., Deshpande, B. S., Sudhakaran, V. K. & Shewale, J. G. (1991). *Beijerinckia indica var penicillanicum* penicillin V acylase: enhanced enzyme production by catabolite

repression-resistant mutant and effect of solvents on enzyme activity. *J. Ind. Microbiol.* **7**, 209-214.

❖ Anon (1992). Technomarket survey report on enzymes (Genetically engineered and immobilized), New Delhi: TIFAC, DST, 125-147.

❖ Antignac, A., Boneca, I. G., Rousselle, J. C., Namane, A., Carlier, J. P., Vazquez, J. A., Fox, A., Alonso, J. M. & Taha, M. K. (2003). Correlation between alterations of the penicillin-binding protein 2 and modifications of the peptidoglycan structure in *Neisseria meningitidis* with reduced susceptibility to penicillin G. *J. Biol. Chem.* **278**, 31529-31535.

❖ Arbuzova, A., Murray, D. & McLaughlin, S. (1998). MARCKS, membranes, and calmodulin: kinetics of their interaction. *Biochimica Biophysica Acta,* **1376**, 369-379.

❖ Argos, P., Rossmann, M. G., Grau, U. M., Zuber, H., Frank, G. & Tratschin, J. D. (1979). Thermal stability and protein structure. *Biochemistry,* **18**, 5698-5703.

❖ Arnheim, N. & Erlich, H. (1992). Polymerase chain reaction strategy. *Annu. Rev. Biochem.* **61**, 131-156.

❖ Aronson, A. I., Song, H. Y. & Bourne, N. (1989). Gene structure and precursor processing of a novel *Bacillus subtilis* spore coat protein. *Mol. Microbiol.* **3**, 437-444.

❖ Arroyo, M., de la Mata, I., Acebal, C. & Castillon, M. P. (2003). Biotechnological applications of penicillin acylases: state-of-the-art. *Appl. Microbiol. Biotechnol.* **60**, 507-514.

❖ Asherie, N. (2004). Protein crystallization and phase diagrams. *Methods,* **34**, 266-272.

❖ Aslanidis, C. & Dejong, P. J. (1990). Ligation-independent cloning of PCR Products (LIC-PCR). *Nucleic Acids Res.* **18**, 6069-6074.

❖ Baldaro, E., Faiardi, D., Fuganti, C., Grasselli, P. & Lazzarini, A. (1988). Phenylacetyloxymethylene, a carboxyl protecting group removable with immobilized penicillin acylase, useful in benzyl penicillin chemistry. *Tetrahedron Lett.* **29**, 4623-4624.

❖ Ballio, A., Chain, E. B., Dentice Di Accadia, F., Mauri, M., Rauer, K., Schlesinger, M. J. & Schlesinger, S. (1961). Identification of a compound related to 6-aminopenicillanic acid, isolated from culture media of *Penicillium chrysogenum. Nature,* **191**, 909-910.

❖ Barbero, J. L., Buesa, J. M., Gonzalez de Buitrago, G., Mendez, E., Pez-Aranda, A. & Garcia, J. L. (1986). Complete nucleotide sequence of the penicillin acylase gene from *Kluyvera citrophila. Gene,* **49**, 69-80.

❖ Basso, A., Braiuca, P., De Martin, L., Ebert, C., Gardossi, L. & Linda, P. (2000). d-Phenylglycine and d-4-hydroxyphenylglycine methyl esters via penicillin G acylase catalysed resolution in organic solvents. *Tetrahedron: Asymm.* **11**, 1789-1796.

❖ Batchelor, F. R., Doyle, F. P., Nayler, J. H. C. & Rolinson, G. N. (1959). Syntheses of penicillin: 6-aminopenicillanic acid in penicillin fermentations. *Nature,* **183**, 257-258.

❖ Bauer K Fau - Kaufmann, W., Kaufmann W Fau - Ludwig, S. A. & Ludwig, S. A. (1971). A simplified determination of penicillin amidase from *E. coli. Hoppe Seylers Z Physiol. Chem.* **352**, 1723-1724.

❖ Beauchamp, J. C. & Isaacs, N. W. (1999). Methods for X-ray diffraction analysis of macromolecular structures. *Curr. Opin. Chem. Biol.* **3**, 525-529.

❖ Betz, S. F. (1993). Disulfide bonds and the stability of globular proteins. *Protein Sci.* **2**, 1551-1558.

❖ Bhatnagar, R. S., Futterer, K., Farazi, T. A., Korolev, S., Murray, C. L., Jackson-Machelski, E., Gokel, G. W., Gordon, J. I. & Waksman, G. (1998). Structure of N-myristoyltransferase with bound myristoyl CoA and peptide substrate analogs. *Nat. Struc. Biol.* **5**, 1091-1097.

❖ Bhatnagar, R. S., Futterer, K., Waksman, G. & Gordon, J. I. (1999). The structure of myristoyl-CoA:protein N-myristoyltransferase. *Biochim. Biophys. Acta,* **1441**, 162-172.

❖ Binda, C., Bossi, R. T., Wakatsuki, S., Arzt, S., Coda, A., Curti, B., Vanoni, M. A. & Mattevi, A. (2000). Cross-talk and ammonia channeling between active centers in the unexpected domain arrangement of glutamate synthase. *Structure,* **8**, 1299-1308.

❖ Blake, C. C., Koenig, D. F., Mair, G. A., North, A. C., Phillips, D. C. & Sarma, V. R. (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature,* **206**, 757-761.

❖ Blommel, P. G., Becker, K. J., Duvnjak, P. & Fox, B. G. (2007). Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. *Biotechnol. Prog.* **23**, 585-598.

❖ Blow, D. (2002). *Outline of Crystallography for Biologists*: Oxford University Press, Oxford.

❖ Blundell, T. L., Elliott, G., Gardner, S. P., Hubbard, T., Islam, S., Johnson, M., Mantafounis, D., Murray-Rust, P., Overington, J., Pitts, J. E., Sali, A., Sibanda, B. L., Singh, J., Sternberg, M. J. E., Sutcliffe, M. J., Thornton, J. M. & Travers, P. (1989). Protein engineering and design. *Philos. Trans. Royal Soc. London. B,* **324**, 447-460.

❖ Blundell, T. L. & Johnson, L. N. (1976). *Protein crystallography*: Academic Press.

❖ Boanca, G., Sand, A., Okada, T., Suzuki, H., Kumagai, H., Fukuyama, K. & Barycki, J. J. (2007). Autoprocessing of *Helicobacter pylori* gamma-glutamyltranspeptidase leads to the formation of a threonine-threonine catalytic dyad. *J. Biol. Chem.* **282**, 534-541.

❖ Bochtler, M., Ditzel, L., Groll, M. & Huber, R. (1997). Crystal structure of heat shock locus V (HslV) from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA,* **94**, 6070-6074.

❖ Bock, A., Wirth, R., Schmid, G., Schumacher, G., Lang, G. & Buckel, P. (1983). The two subunits of penicillin acylase are processed from a common precursor. *FEMS Microbiol. Lett.* **20**, 141-144.

❖ Boisson, B., Giglione, C. & Meinnel, T. (2003). Unexpected protein families including cell defense components feature in the N-myristoylome of a higher eukaryote. *J. Biol. Chem.* **278**, 43418-43429.

❖ Bokhove, M., Jimenez, P. N., Quax, W. J. & Dijkstra, B. W. (2010a). The quorum-quenching N-acyl homoserine lactone acylase PvdQ is an Ntn-hydrolase with an unusual substrate-binding pocket. *Proc. Natl. Acad. Sci. USA,* **107**, 686-691.

❖ Bokhove, M., Yoshida, H., Hensgens, C. M. H., Metske van der Laan, J., Sutherland, J. D. & Dijkstra, B. W. (2010b). Structures of an isopenicillin N converting Ntn-hydrolase reveal different catalytic roles for the active site residues of precursor and mature enzyme. *Structure,* **18**, 301-308.

❖ Bompard-Gilles, C., Villeret, V., Davies, G. J., Fanuel, L., Joris, B., Frere, J. M. & Van Beeumen, J. (2000). A new variant of the Ntn hydrolase fold revealed by the crystal structure of l-aminopeptidase d-Ala-esterase/amidase from *Ochrobactrum anthropi*. *Structure,* **8**, 153-162.

❖ Bomstein, J. & Evans, W. G. (1965). Automated colorimetric determination of 6-Aminopenicillanic acid in fermentation media. *Anal. Chem.* **37**, 576-578.

❖ Borders, C. L., Broadwater, J. A., Bekeny, P. A., Salmon, J. E., Lee, A. S., Eldridge, A. M. & Pett, V. B. (1994). A structural role for arginine in proteins: multiple hydrogen bonds to backbone carbonyl oxygens. *Protein Sci.* **3**, 541-548.

❖ Bossi, A., Cretich, M. & Righetti, P. G. (1998). Production of D-phenylglycine from racemic (D,L)-phenylglycine via isoelectrically-trapped penicillin G acylase. *Biotechnol. Bioeng.* **60**, 454-461.

❖ Bouamr, F., Scarlata, S. & Carter, C. (2003). Role of myristylation in HIV-1 Gag assembly. *Biochemistry,* **42**, 6408-6417.

❖ Boutin, J. A. (1997). Myristoylation. *Cell Signal,* **9**, 15-35.

❖ Bowyer, P. W., Gunaratne, R. S., Grainger, M., Withers-Martinez, C., Wickramsinghe, S. R., Tate, E. W., Leatherbarrow, R. J., Brown, K. A., Holder, A. A. & Smith, D. F. (2007). Molecules incorporating a benzothiazole core scaffold inhibit the N-myristoyltransferase of *Plasmodium falciparum. Biochem. J.* **408**, 173-180.

❖ Braam, B. & Verhaar, M. C. (2007). Understanding eNOS for pharmacological modulation of endothelial function: a translational view. *Curr. Pharm. Des.* **13**, 1727-1740.

❖ Braiuca, P., Ebert, C., Fischer, L., Gardossi, L. & Linda, P. (2003). A homology model of penicillin acylase from *Alcaligenes faecalis* and in silico evaluation of its selectivity. *Chembiochem,* **4**, 615-622.

❖ Branden, C. & Tooze, J. (1991). *Introduction to protein structure*: Garland Pub.

❖ Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J. L., Tomchick, D. R. & Murzin, A. G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature,* **378**, 644-644.

❖ Brannigan, J. A., Dodson, G. C., Done, S. H., Hewitt, L., McVey, C. E. & Wilson, K. S. (2000). Structural studies of penicillin acylase. *Appl. Biochem. Biotechnol.* **88**, 313-319.

❖ Brannigan, J. A., Smith, B. A., Yu, Z., Brzozowski, A. M., Hodgkinson, M. R., Maroof, A., Price, H. P., Meier, F., Leatherbarrow, R. J., Tate, E. W., Smith, D. F. & Wilkinson, A. J. (2010). N-myristoyltransferase from *Leishmania donovani*: structural and functional characterisation of a potential drug target for visceral leishmaniasis. *J. Mol. Biol.* **396**, 985-999.

❖ Bruggink, A., Roos, E. C. & de Vroom, E. (1998). Penicillin acylase in the industrial production of beta-lactam antibiotics. *Org. Process Res. Dev.* **2**, 128-133.

❖ Bruggink, A. & Roy, P. D. (2001). Industrial synthesis of semisynthetic antibiotics. In Bruggink, A. (ed.) *Synthesis of β-lactam antibiotics*. Kluwer academic publishers, Dordrecht, 12-56.

❖ Brunger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* **355**, 472-475.

❖ Bryant, M. & Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proc. Natl. Acad. Sci. USA,* **87**, 523-527.

❖ Buller, A. R., Freeman, M. F., Wright, N. T., Schildbach, J. F. & Townsend, C. A. (2012). Insights into cis-autoproteolysis reveal a reactive state formed through conformational rearrangement. *Proc. Natl. Acad. Sci. USA,* **109**, 2308-2313.

❖ Bushueva, T. L., Busel, E. P. & Burstein, E. A. (1975). Interaction of protein functional groups with indole chromophore.3. Amine, Amide, and Thiol-Groups. *Studia Biophysica,* **52**, 41-52.

❖ Bussey, H. (1988). Proteases and the processing of precursors to secreted proteins in yeast. *Yeast,* **4**, 17-26.

❖ Cai, G., Zhu, S. C., Yang, S., Zhao, G. P. & Jiang, W. H. (2004). Cloning, overexpression, and characterization of a novel thermostable penicillin G acylase from *Achromobacter xylosoxidans*: Probing the molecular basis for its high thermostability. *Appl. Environ. Biotechnol.* **70**, 2764-2770.

❖ Carr, S. A., Biemann, K., Shoji, S., Parmelee, D. C. & Titani, K. (1982). n-Tetradecanoyl is the NH$_2$-terminal blocking group of the catalytic subunit of cyclic AMP-dependent protein kinase from bovine cardiac muscle. *Proc. Natl. Acad. Sci. USA,* **79**, 6128-6131.

❖ Cardamone, M. & Puri, N. K. (1992). Spectrofluorimetric assessment of the surface hydrophobicity of proteins. *Biochem. J.* **282 (Pt 2)**, 589-593.

❖ Carrington Jc Fau - Freed, D. D., Freed Dd Fau - Sanders, T. C. & Sanders, T. C. (1989). Autocatalytic processing of the potyvirus helper component proteinase in *Escherichia coli* and in vitro. *J. Virol.* **63**, 4459-4463.

❖ Casalinuovo, I. A., Di Francesco, P. & Garaci, E. (2004). Fluconazole resistance in *Candida albicans*: a review of mechanisms. *Eur. Rev. Med. Pharmacol. Sci.* **8**, 69-77.

❖ CCP4. (1994). The CCP4 Suite—programs for protein crystallography. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **50**, 760–763.

❖ Chakravarty, S. & Varadarajan, R. (2000). Elucidation of determinants of protein stability through genome sequence analysis. *FEBS Lett.* **470**, 65-69.

❖ Chandra, P. M., Brannigan, J. A., Prabhune, A., Pundle, A., Turkenburg, J. P., Dodson, G. G. & Suresh, C. G. (2005). Cloning, preparation and preliminary crystallographic studies of penicillin V acylase autoproteolytic processing mutants. *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **61**, 124-127.

❖ Chappuis, F., Sundar, S., Hailu, A., Ghalib, H., Rijal, S., Peeling, R. W., Alvar, J. & Boelaert, M. (2007). Visceral leishmaniasis: What are the needs for diagnosis, treatment and control? *Nat. Rev. Microbiol.* **5**, 873-882.

❖ Chayen, N. E. (1998). Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **54**, 8-15.

❖ Chiang, C. & Bennett, R. E. (1967). Purification and Properties of Penicillin Amidase from *Bacillus megaterium. J. Bacteriol.* **93**, 302-308.

❖ Chisti, Y. & Moo-Young M. (1991). Fermentation technology, bioprocessing, scale-up and manufacture. In Moses, V. & Cape, R. E. (eds.) *Biotechnology: The Science and the Business*. New York: Harwood Academic Publishers, 167–209.

❖ Choi, K. S., Kim, J. A. & Kang, H. S. (1992). Effects of site-directed mutations on processing and activities of penicillin G acylase from *Escherichia coli* ATCC 11105. *J. Bacteriol.* **174**, 6270-6276.

❖ Chow, M., Newman, J. F., Filman, D., Hogle, J. M., Rowlands, D. J. & Brown, F. (1987). Myristylation of picornavirus capsid protein VP4 and its structural significance. *Nature,* **327**, 482-486.

❖ Claridge, C. A., Gourevitch, A. & Lein, J. (1960). Bacterial penicillin amidase. *Nature,* **187**, 237-238.

❖ Claridge, C. A., Luttinger, J. R. & Lein, J. (1963). Specificity of penicillin amidases. *Proc. Soc. Exp. Biol. Med.* **113**, 1008-1012.

❖ Cohen, B. E. (1998). Amphotericin B toxicity and lethality: a tale of two channels. *Int. J. Pharmaceut.* **162**, 95-106.

❖ Cole, M. (1964). Properties of the penicillin deacylase enzyme of *Escherichia coli*. *Nature*, **203**, 519-520.

❖ Cole, M. & Sutherland, R. (1966). The role of penicillin acylase in the resistance of gram-negative bacteria to penicillins. *J. Gen. Microbiol.* **42**, 345-356.

❖ Cole, M. (1967). Microbial synthesis of penicillin and 6-APA. *Process Biochem.* **2**, 35-43.

❖ Cole, M. (1969). Hydrolysis of penicillins and related compounds by the cell-bound penicillin acylase of *Escherichia coli*. *Biochem. J.* **115**, 733-739.

❖ Cordeddu, V., Di Schiavi, E., Pennacchio, L. A., Ma'ayan, A., Sarkozy, A., Fodale, V., Cecchetti, S., Cardinale, A., Martin, J., Schackwitz, W., Lipzen, A., Zampino, G., Mazzanti, L., Digilio, M. C., Martinelli, S., Flex, E., Lepri, F., Bartholdi, D., Kutsche, K., Ferrero, G. B., Anichini, C., Selicorni, A., Rossi, C., Tenconi, R., Zenker, M., Merlo, D., Dallapiccola, B., Iyengar, R., Bazzicalupo, P., Gelb, B. D. & Tartaglia, M. (2009). Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair. *Nat. Genetics,* **41**, 1022-U1095.

❖ Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **62**, 1002-1011.

❖ Croft, S. L., Seifert, K. & Yardley, V. (2006). Current scenario of drug development for leishmaniasis. *Ind. J. Med. Res.* **123**, 399-410.

❖ Cross, F. R., Garber, E. A., Pellman, D. & HanafUSA, H. (1984). A short sequence in the p60^{src} N-terminus is required for p60^{src} myristylation and membrane association and for cell transformation. *Mol. Cell. Biol.* **4**, 1834-1842.

❖ Crawford, L., Stepan, A. M., McAda, P. C., Rambosek, J. A., Conder, M. J., Vinci, V. A. & Reeves, C. D. (1995). Production of cephalosporin intermediates by feeding adipic acid to recombinant *Penicillium chrysogenum* strains expressing ring expansion activity. *Biotechnology,* **13**, 58-62.

❖ Crowfoot, D., Bunn, C. W., Rogers-Low, B. W. & Turner, A. J. (1949). *The Chemistry of Penicillin*: Princeton University Press.

❖ Crowther, R. A. & Blow, D. M. (1967). A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallogr.* **23**, 544-548.

❖ Cunningham, B. C. & Wells, J. A. (1987). Improvement in the alkaline stability of subtilisin using an efficient random mutagenesis and screening procedure. *Protein Eng.* **1**, 319-325.

❖ Daniel, E. & Weber, G. (1966). Cooperative effects in binding by bovine serum albumin. I. The binding of 1-anilino-8-naphthalenesulfonate. Fluorimetric titrations. *Biochemistry,* **5**, 1893-1900.

❖ Daopin, S., Davies, D. R., Schlunegger, M. P. & Grutter, M. G. (1994). Comparison of two crystal structures of TGF-beta2: the accuracy of refined protein structures. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **50**, 85-92.

❖ Das, R. & Gerstein, M. (2000). The stability of thermophilic proteins: a study based on comprehensive genome comparison. *Funct. Integr. Genomics,* **1**, 76-88.

❖ Daumy, G. O., Danley, D. & McColl, A. S. (1985). Role of protein subunits in *Proteus rettgeri* penicillin G acylase. *J. Bacteriol.* **163**, 1279-1281.

❖ Dauter, Z. (1997). Data collection strategy. In: Charles W. Carter, Jr. (ed.) *Methods Enzymol.*: Academic Press, 326-344.

❖ Dauter, Z. (1999). Data-collection strategies. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **55**, 1703-1717.

❖ Demain, A. L. (2000). Small bugs, big business: The economic power of the microbe. *Biotechnol. Adv.* **18**, 499-514.

❖ de Souza, V. R., Silva, A. C. G., Pinotti, L. M., Araujo, H. S. S. & Giordano, R. D. C. (2005). Characterization of the penicillin G acylase from *Bacillus megaterium* ATCC 14945. *Braz. Arch. Biol. Technol.* **48**, 105-111.

❖ De Koning, J. J., Kooreman, H. J., Tan, H. S. & Verweij, J. (1975). One-step, high yield conversion of penicillin sulfoxides to deacetoxycephalosporins. *J. Org. Chem.* **40**, 1346-1347.

❖ De Vroom, E. & van der Mey, M. (1995). New use of A*lcaligenes faecalis* penicillin G acylase. *International Patent* EP0638649A2

❖ Denning, D. W. (2002). Echinocandins: a new class of antifungal. *J. Antimicrob. Chemother.* **49**, 889-891.

❖ Deshpande, B. S., Ambedkar, S. S., Sudhakaran, V. K. & Shewale, J. G. (1994). Molecular biology of β-lactam acylases. *World J. Microbiol. Biotechnol.* **10**, 129-138.

❖ Devadas, B., Freeman, S. K., McWherter, C. A., Kishore, N. S., Lodge, J. K., Jackson-Machelski, E., Gordon, J. I. & Sikorski, J. A. (1998). Novel biologically active nonpeptidic inhibitors of myristoylCoA:protein N-myristoyltransferase. *J. Med. Chem.* **41**, 996-1000.

❖ Devadas, B., Freeman, S. K., Zupec, M. E., Lu, H. F., Nagarajan, S. R., Kishore, N. S., Lodge, J. K., Kuneman, D. W., McWherter, C. A., Vinjamoori, D. V., Getman, D. P., Gordon, J. I. & Sikorski, J. A. (1997). Design and synthesis of novel imidazole-substituted dipeptide amides as potent and selective inhibitors of *Candida albicans* myristoylCoA:protein N-myristoyltransferase and identification of related tripeptide inhibitors with mechanism-based antifungal activity. *J. Med. Chem.* **40**, 2609-2625.

❖ Devadas, B., Zupec, M. E., Freeman, S. K., Brown, D. L., Nagarajan, S., Sikorski, J. A., McWherter, C. A., Getman, D. P. & Gordon, J. I. (1995). Design and syntheses of potent and selective dipeptide inhibitors of *Candida albicans* myristoyl-CoA:protein N-myristoyltransferase. *J. Med. Chem.* **38**, 1837-1840.

❖ Diederichs, K. & Karplus, P. A. (1997). Improved *R*-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* **4**, 269-275.

❖ Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H. & Groll, M. (1998). Conformational constraints for protein self-cleavage in the proteasome. *J. Mol. Biol.* **279**, 1187-1191.

❖ Dodson, G. G. (2000). Catalysis in penicillin G amidase - a member of the Ntn (N-terminal nucleophile) hydrolase family. *Croatica Chemica. Acta.* **73**, 901-908.

❖ Done, S. H., Brannigan, J. A., Moody, P. C. E. & Hubbard, R. E. (1998). Ligand-induced conformational change in penicillin acylase. *J. Mol. Biol.* **284**, 463-475.

❖ Dong, A. P., Xu, X. H. & Edward, A. M. (2007). *In situ* proteolysis for protein crystallization and structure determination. *Nat. Methods,* **4**, 1019-1021.

❖ Dougherty, W. G. & Carrington, J. C. (1988). Expression and function of Potyviral gene products. *Annu. Rev. Phytopathol.* **26**, 123-143.

❖ Douglass, J., Civelli, O. & Herbert, E. (1984). Polyprotein gene expression: generation of diversity of neuroendocrine peptides. *Annu. Rev. Biochem.* **53**, 665-715.

❖ Drenth, J. & Haas, C. (1998). Nucleation in protein crystallization. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **54**, 867-872.

❖ Ducker, C. E., Upson, J. J., French, K. J. & Smith, C. D. (2005). Two N-myristoyltransferase isozymes play unique roles in protein myristoylation, proliferation, and apoptosis. *Mol. Cancer Res.* **3**, 463-476.

❖ Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, E. J., Dodson, G. & Moody, P. C. E. (1995). Penicillin acylase has a single-amino-acid catalytic center. *Nature,* **373**, 264-268.

❖ Duronio, R. J., Towler, D. A., Heuckeroth, R. O. & Gordon, J. I. (1989). Disruption of the yeast N-myristoyl transferase gene causes recessive lethality. *Science,* **243**, 796-800.

❖ Ebara, S., Naito, H., Nakazawa, K., Ishii, F. & Nakamura, M. (2005). FTR1335 is a novel synthetic inhibitor of *Candida albicans* N-myristoyltransferase with fungicidal activity. *Biol. Pharm. Bull.* **28**, 591-595.

❖ Ebiike, H., Masubuchi, M., Liu, P., Kawasaki, K., Morikami, K., Sogabe, S., Hayase, M., Fujii, T., Sakata, K., Shindoh, H., Shiratori, Y., Aoki, Y., Ohtsuka, T. & Shimma, N. (2002). Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 2. *Bioorg. Med. Chem. Lett.* **12**, 607-610.

❖ Edmond, M. B., Wallace, S. E., McClish, D. K., Pfaller, M. A., Jones, R. N. & Wenzel, R. P. (1999). Nosocomial bloodstream infections in United States hospitals: A three-year analysis. *Clinical Infec. Dis.* **29**, 239-244.

❖ Elander, R. P. (2003). Industrial production of beta-lactam antibiotics. *Appl. Microbiol. Biotechnol.* **61**, 385-392.

❖ Elkins, J. M., Kershaw, N. J. & Schofield, C. J. (2005). X-ray crystal structure of ornithine acetyltransferase from the clavulanic acid biosynthesis gene cluster. *Biochem. J.* **385**, 565-573.

❖ Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **60**, 2126-2132.

❖ Engh, R. A. & Huber, R. (1991). Accurate Bond and Angle Parameters for X-Ray Protein-Structure Refinement. *Acta Crystallogr. Sect. A: Foundn. Crystallogr.* **47**, 392-400.

❖ Essack, S. Y. (2001). The development of beta-lactam antibiotics in response to the evolution of beta-lactamases. *Pharm. Res.* **18**, 1391-1399.

❖ Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **62**, 72-82.

❖ Fadnavis, N. W., Sharfuddin, M., Vadivel, S. K. & Bhalerao, U. T. (1997). Efficient chemoenzymatic synthesis of (2S,3S)-3-hydroxyleucine mediated by immobilised penicillin G acylase. *J. Chem. Soc. Perkin Trans. 1*, 3577-3578.

❖ Farazi, T. A., Waksman, G. & Gordon, J. I. (2001). Structures of *Saccharomyces cerevisiae* N-myristoyltransferase with bound myristoylCoA and peptide provide insights about substrate recognition and catalysis. *Biochemistry,* **40**, 6335-6343.

❖ Felsted, R. L., Glover, C. J. & Hartman, K. (1995). Protein N-myristoylation as a chemotherapeutic target for cancer. *J. Natl. Cancer Inst.* **87**, 1571-1573.

❖ Fernandez-Lafuente, R., Rosell, C. M. & Guisan, J. M. (1991). Enzyme reaction engineering: synthesis of antibiotics catalysed by stabilized penicillin G acylase in the presence of organic cosolvents. *Enz. Microb. Technol.* **13**, 898-905.

❖ Fernandez-lafuente, R., Rosell, C. M. & Guisan, J. M. (1995). The use of stabilized penicillin acylase derivatives improves the design of kinetically controlled synthesis. *J. Mol. Cat. A,* **101**, 91-97.

❖ Fernandez-Lafuente, R., Rosell, C. M. & Guisan, J. M. (1996). Dynamic reaction design of enzymic biotransformations in organic media: equilibrium-controlled synthesis of antibiotics by penicillin G acylase. *Biotechnol. Appl. Biochem.* **24(2)**, 139-143.

❖ Ferré-D'Amaré, A. R. & Burley, S. K. (1994). Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. *Structure,* **2**, 357-359.

❖ Fleming, A. (1929). On the antibacterial action of cultures of a *penicillium*, with special reference to their use in the isolation of *B. Influenzae*. *Brit. J. Exp. Pathol.* **10**, 226-236.

❖ Fogg, M. J. & Wilkinson, A. J. (2008). Higher-throughput approaches to crystallization and crystal structure determination. *Biochem. Soc. Trans.* **36**, 771-775.

❖ Foster, W. & Raoult, A. (1974). Early descriptions of antibiosis. *J. R. Coll. Gen. Pract.* **24**, 889-894.

❖ Frearson, J. A., Brand, S., McElroy, S. P., Cleghorn, L. A. T., Smid, O., Stojanovski, L., Price, H. P., Guther, M. L. S., Torrie, L. S., Robinson, D. A., Hallyburton, I., Mpamhanga, C. P., Brannigan, J. A., Wilkinson, A. J., Hodgkinson, M., Hui, R., Qiu, W., Raimi, O. G., van Aalten, D. M. F., Brenk, R., Gilbert, I. H., Read, K. D., Fairlamb, A. H., Ferguson, M. A. J., Smith, D. F. & Wyatt, P. G. (2010). N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature,* **464**, 728-U100.

❖ French, K. J., Zhuang, Y., Schrecengost, R. S., Copper, J. E., Xia, Z. & Smith, C. D. (2004). Cyclohexyl-octahydro-pyrrolo[1,2-a]pyrazine-based inhibitors of human N-myristoyltransferase-1. *J. Pharmacol. Exp. Ther.* **309**, 340-347.

❖ Fridkin, S. K. & Jarvis, W. R. (1996). Epidemiology of nosocomial fungal infections. *Clin. Microbiol. Rev.* **9**, 499-511.

❖ Fritz-Wolf, K., Koller, K. P., Lange, G., Liesum, A., Sauber, K., Schreuder, H., Aretz, W. & Kabsch, W. (2002). Structure-based prediction of modifications in glutarylamidase to allow single-step enzymatic production of 7-aminocephalosporanic acid from cephalosporin C. *Protein Sci.* **11**, 92-103.

❖ Fuganti, C., Grasselli, P., Seneci, P. F., Servi, S. & Casati, P. (1986). Immobilized benzylpenicillin acylase : Application to the synthesis of optically active forms of carnitin and propranalol. *Tetrahedron Lett.* **27**, 2061-2062.

❖ Furuishi, K., Matsuoka, H., Takama, M., Takahashi, I., Misumi, S. & Shoji, S. (1997). Blockage of N-myristoylation of HIV-1 gag induces the production of impotent progeny virus. *Biochem. Biophys. Res. Commun.* **237**, 504-511.

❖ Gabor E. M., de Vries E. J. & Janssen D. B. (2005). A novel penicillin acylase from the environmental gene pool with improved synthetic properties. *Enz. Microb.Technol.* **36**, 182-190.

❖ Gallis, H. A., Drew, R. H. & Pickard, W. W. (1990). Amphotericin B: 30 years of clinical experience. *Rev. Infect. Dis.* **12**, 308-329.

❖ Gamblin, S. J., and Rodgers, D. W. (1993). Some Practical Details of Data Collection at 100 K. In: L. Sawyer, I., N. and Bailey, S. (ed.) *Proceedings of the CCP4 Study Weekend. In Data Collection and Processing,* 28-32.

❖ Garcia, J. L. & Buesa, J. M. (1986). An improved method to clone penicillin acylase genes - cloning and expression in *Escherichia coli* of penicillin G acylase from *Kluyvera citrophila*. *J. Biotechnol.* **3**, 187-195.

❖ Garcia, J. V. & Miller, A. D. (1992). Downregulation of cell surface CD4 by Nef. *Res. Virol.* **143**, 52-55.

❖ Garman, E. F. & Schneider, T. R. (1997). Macromolecular cryocrystallography. *J. Appl. Crystallogr.* **30**, 211-237.

❖ Gelb, M. H., Van Voorhis, W. C., Buckner, F. S., Yokoyama, K., Eastman, R., Carpenter, E. P., Panethymitaki, C., Brown, K. A. & Smith, D. F. (2003). Protein farnesyl and N-myristoyl transferases: piggy-back medicinal chemistry targets for the development of antitrypanosomatid and antimalarial therapeutics. *Mol. Biochem. Parasitol.* **126**, 155-163.

❖ Georgopapadakou, N. H. (2002). Antifungals targeted to protein modification: focus on protein N-myristoyltransferase. *Expert Opin. Invest. Drugs,* **11**, 1117-1125.

❖ Giang, D. K. & Cravatt, B. F. (1998). A second mammalian N-myristoyltransferase. *J. Biol. Chem.* **273**, 6595-6598.

❖ Glaudemans C. P. J. & Jolley, M. E. (1980). *Methods Carbohydr. Chem.* **8**, 145-149.

❖ Glover, C. J., Tellez, M. R., Guziec, F. S., Jr. & Felsted, R. L. (1991). Synthesis and characterization of inhibitors of myristoyl-CoA:protein N-myristoyltransferase. *Biochem. Pharmacol.* **41**, 1067-1074.

❖ Gordon, J. I., Duronio, R. J., Rudnick, D. A., Adams, S. P. & Gokel, G. W. (1991). Protein N-myristoylation. *J. Biol. Chem.* **266**, 8647-8650.

❖ Gottlinger, H. G., Sodroski, J. G. & Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA,* **86**, 5781-5785.

❖ Groll, M., Berkers, C. R., Ploegh, H. L. & Ovaa, H. (2006). Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure,* **14**, 451-456.

❖ Groll, M., Brandstetter, H., Bartunik, H., Bourenkow, G. & Huber, R. (2003). Investigations on the Maturation and Regulation of Archaebacterial Proteasomes. *J. Mol. Biol.* **327**, 75-83.

❖ Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature,* **386**, 463-471.

❖ Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A., Whitaker, R. & Paulus, H. (1998). Characterization and functional analysis of the cis-autoproteolysis active center of glycosylasparaginase. *J. Biol. Chem.* **273**, 9695-9702.

❖ Guisan, J. M., Alvaro, G., Fernandez-Lafuente, R., Rosell, C. M., Garcia, J. L. & Tagliani, A. (1993). Stabilization of heterodimeric enzyme by multipoint covalent immobilization: Penicillin G acylase from *Kluyvera citrophila. Biotechnol. Bioeng.* **42**, 455-464.

❖ Gunaratne, R. S., Sajid, M., Ling, I. T., Tripathi, R., Pachebat, J. A. & Holder, A. A. (2000). Characterization of N-myristoyltransferase from *Plasmodium falciparum. Biochem. J.* **348 (2)**, 459-463.

❖ Guranda, D. T., van Langen, L. M., van Rantwijk, F., Sheldon, R. A. & Svedas, V. K. (2001). Highly efficient and enantioselective enzymatic acylation of amines in aqueous medium. *Tetrahedron: Asymm.* **12**, 1645-1650.

❖ Hamilton-Miller, J. M. (1966). Penicillinacylase. *Bacteriol. Rev.* **30**, 761-771.

❖ Harris, M. (1999). HIV: a new role for Nef in the spread of HIV. *Curr. Biol.* **9**, R459-R461.

❖ Haun, R. S., Serventi, I. M. & Moss, J. (1992). Rapid, reliable ligation-independent cloning of PCR products using modified plasmid vectors. *Biotechniques,* **13**, 515-518.

❖ Heal, W. P., Wickramasinghe, S. R., Bowyer, P. W., Holder, A. A., Smith, D. F., Leatherbarrow, R. J. & Tate, E. W. (2008a). Site-specific N-terminal labelling of proteins in vitro and in vivo using N-myristoyl transferase and bioorthogonal ligation chemistry. *Chem. Comm.* 480-482.

❖ Heal, W. P., Wickramasinghe, S. R., Leatherbarrow, R. J. & Tate, E. W. (2008b). N-Myristoyl transferase-mediated protein labelling *in vivo*. *Org. Biomol. Chem.* **6**, 2308-2315.

❖ Hennig, M., Darimont, B., Sterner, R., Kirschner, K. & Jansonius, J. N. (1995). 2.0 A structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure,* **3**, 1295-1306.

❖ Herning, T., Yutani, K., Inaka, K., Kuroki, R., Matsushima, M. & Kikuchi, M. (1992). Role of proline residues in human lysozyme stability: a scanning calorimetric study combined with X-ray structure analysis of proline mutants. *Biochemistry,* **31**, 7077-7085.

❖ Heuckeroth, R. O., Towler, D. A., Adams, S. P., Glaser, L. & Gordon, J. I. (1988). 11-(Ethylthio)undecanoic acid. A myristic acid analogue of altered hydrophobicity which is functional for peptide N-myristoylation with wheat germ and yeast acyltransferase. *J. Biol. Chem.* **263**, 2127-2133.

❖ Hewitt, L., Kasche, V., Lummer, K., Lewis, R. J., Murshudov, G. N., Verma, C. S., Dodson, G. G. & Wilson, K. S. (2000). Structure of a slow processing precursor penicillin acylase from *Escherichia coli* reveals the linker peptide blocking the active site cleft. *J. Mol. Biol.* **302**, 887-898.

❖ Hill, B. T. & Skowronski, J. (2005). Human N-myristoyltransferases form stable complexes with lentiviral nef and other viral and cellular substrate proteins. *J. Virol.* **79**, 1133-1141.

❖ Hogema, B. M., Arents, J. C., Bader, R., Eijkemans, K., Inada, T., Aiba, H. & Postma, P. W. (1998). Inducer exclusion by glucose 6-phosphate in Escherichia coli. *Mol. Microbiol.* **28**, 755-765.

❖ Hogema, B. M., Arents, J. C., Bader, R. & Postma, P. W. (1999). Autoregulation of lactose uptake through the *LacY* permease by enzyme IIAGlc of the PTS in *Escherichia col*i K-12. *Mol. Microbiol.* **31**, 1825-1833.

❖ Holt, R. J. & Stewart, G. T. (1964). Production of Amidase and *β*-Lactamase by Bacteria. *J. Gen. Microbiol.* **36**, 203-213.

❖ Hu, G., Lin, G., Wang, M., Dick, L., Xu, R.-M., Nathan, C. & Li, H. (2006). Structure of the *Mycobacterium tuberculosi*s proteasome and mechanism of inhibition by a peptidyl boronate. *Mol. Microbiol.* **59**, 1417-1428.

❖ Huang, H. T., English, A. R., Seto, T. A., Shull, G. M. & Sobin, B. A. (1960). Enzymatic hydrolysis of the side chain of penicillins. *J. Am. Chem. Soc.* **82**, 3790-3791.

❖ Huang, H. T., Seto, T. A. & Shull, G. M. (1963). Distribution and substrate specificity of benzylpenicillin acylase. *Appl. Microbiol.* **11**, 1-6.

❖ Huber, Eva M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, Christopher J., Groettrup, M. & Groll, M. (2012). Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell,* **148**, 727-738.

❖ Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF-a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212-220.

❖ Ignatova, Z. (2000). PhD Thesis. Technische Universita¨t Hamburg-Harburg.

❖ Ignatova, Z., Stoeva, S., Galunsky, B., Hornle, C., Nurk, A., Piotraschke, E., Voelter, W. & Kasche, V. (1998). Proteolytic processing of penicillin amidase from *Alcaligenes faecalis* cloned in *E. coli* yields several active forms. *Biotechnol. Lett.* **20**, 977-982.

❖ Ignatova, Z., Wischnewski, F., Notbohm, H. & Kasche, V. (2005). Pro-sequence and Ca^{2+}-binding: Implications for folding and maturation of Ntn-hydrolase penicillin amidase from *E coli*. *J. Mol. Biol.* **348**, 999-1014.

❖ Ignatova, Z., Hornle, C., Nurk, A. & Kasche, V. (2002). Unusual signal peptide directs penicillin amidase from *Escherichia coli* to the Tat translocation machinery. *Biochem. Biophys. Res. Commun.* **291**, 146-149.

❖ Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J. Biochem.* **88**, 1895-1898.

❖ Inada, T., Kimata, K. & Aiba, H. (1996). Mechanism responsible for glucose–lactose diauxie in *Escherichia coli*: challenge to the cAMP model. *Genes to Cells,* **1**, 293-301.

❖ Ishii, Y., Saito, Y., Fujimura, T., Isogai, T., Kojo, H., Yamashita, M., Niwa, M. & Kohsaka, M. (1994). A novel 7-β-(4-carboxybutanamido)-cephalosporanic acid acylase isolated from Pseudomonas strain C427 and its high-level production in *Escherichia coli*. *J. Ferment. Bioeng.* **77**, 591-597.

❖ Ishiye, M. & Niwa, M. (1992). Nucleotide sequence and expression in *Escherichia coli* of the cephalosporin acylase gene of a *Pseudomonas* strain. *Biochim. Biophys. Acta,* **1132**, 233-239.

❖ Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denisot, M. A., Badet, B., Polikarpov, I., Littlechild, J. A. & Teplyakov, A. (1996). Substrate binding is required for assembly of the active conformation of the catalytic site in Ntn amidotransferases: evidence from the 1.8 Å crystal structure of the glutaminase domain of glucosamine 6-phosphate synthase. *Structure*, **4**, 801–810.

❖ Jacoby, G. A. & Munoz-Price, L. S. (2005). The new beta-lactamases. *N. Engl. J. Med.* **352**, 380-391.

❖ Jaenicke, R. & Bohm, G. (1998). The stability of proteins in extreme environments. *Cur. Op. Struct. Biol.* **8**, 738-748.

❖ Johnson, D. R., Bhatnagar, R. S., Knoll, L. J. & Gordon, J. I. (1994). Genetic and biochemical studies of protein N-myristoylation. *Annu. Rev. Biochem.* **63**, 869-914.

❖ Jose, L. M. H., Iliyna, A., Dominguez, M. L., Olga, S. C. & Dustet, M. J. C. (2003). Partial characterisation of penicillin acylase from fungi *Aspergillus fumigatus* and *Mucor gryseocianum. Moscow Univ. Chem. Bull.* **44**, 53-56.

❖ Kaasgaard, S. & Veitland, U. (1992). Process for the preparation of *β*-lactams. *International Patent* WO 92/01061 Novo-Nordisk.

❖ Kabsch, W. (2010). Xds. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **66**, 125-132.

❖ Kamps, M. P., Buss, J. E. & Sefton, B. M. (1985). Mutation of NH_2-terminal glycine of p60^{src} prevents both myristoylation and morphological transformation. *Proc. Natl. Acad. Sci. USA,* **82**, 4625-4628.

❖ Kamps, M. P., Buss, J. E. & Sefton, B. M. (1986). Rous sarcoma virus transforming protein lacking myristic acid phosphorylates known polypeptide substrates without inducing transformation. *Cell,* **45**, 105-112.

❖ Kasche, V. (1986). Mechanism and yields in enzyme catalysed equilibrium and kinetically controlled synthesis of *β*-lactam antibiotics, peptides and other condensation products. *Enz. Microb. Technol.* **8**, 4-16.

❖ Kasche, V., Lummer, K., Nurk, A., Piotraschke, E., Rieks, A., Stoeva, S. & Voelter, W. (1999). Intramolecular autoproteolysis initiates the maturation of penicillin amidase from *Escherichia coli. Biochim. Biophys. Acta,* **1433**, 76-86.

❖ Kasche, V., Galunsky, B. & Ignatova, Z. (2003). Fragments of pro-peptide activate mature penicillin amidase of *Alcaligenes faecalis. Eur. J. Biochem.* **270**, 4721-4728.

❖ Kato, K. (1980). *Agric. Biol. Chem.* **44**, 1083-1088

❖ Kaufmann, W. & Bauer, K. (1960). Enzymatic splitting and resynthese of penicillin. *Naturwissenschaften,* **47**, 474-475.

❖ Kaufmann, W. & Bauer, K. (1964). Variety of substrates for a bacterial benzyl penicillin splitting enzyme. *Nature*, **203**, 520.

❖ Kawasaki, K., Masubuchi, M., Morikami, K., Sogabe, S., Aoyama, T., Ebiike, H., Niizuma, S., Hayase, M., Fujii, T., Sakata, K., Shindoh, H., Shiratori, Y., Aoki, Y., Ohtsuka, T. & Shimma, N. (2003). Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. *Bioorg. Med. Chem. Lett.* **13**, 87-91.

❖ Keiler, K. C. & Sauer, R. T. (1995). Identification of active site residues of the Tsp protease. *J. Biol. Chem.* **270**, 28864-28868.

❖ Khan, J. A., Dunn, B. M. & Tong, L. (2005). Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure,* **13**, 1443-1452.

❖ Kim, D. J. & Byun, S. M. (1990). Purification and properties of ampicillin acylase from *Pseudomonas melanogenum. Biochim. Biophys. Acta,* **1040**, 12-18.

❖ Kim, J. H., Krahn, J. M., Tomchick, D. R., Smith, J. L. & Zalkin, H. (1996). Structure and function of the glutamine phosphoribosylpyrophosphate amidotransferase glutamine site and communication with the phosphoribosylpyrophosphate site. *J. Biol. Chem.* **271**, 15549-15557.

❖ Kim, D. W., Kang, S. M. & Yoon, K. H. (1999). Isolation of novel *Pseudomonas diminuta* KAC-1 strain producing glutaryl 7-aminocephalosporanic acid acylase. *J. Microbiol.* **37**, 200-205.

❖ Kim, Y., Yoon, K. H., Khang, Y., Turley, S. & Hol, W. G. J. (2000). The 2.0 Å crystal structure of cephalosporin acylase. *Structure,* **8**, 1059-1068.

❖ Kim, Y. & Hol, W. G. (2001). Structure of cephalosporin acylase in complex with glutaryl-7-aminocephalosporanic acid and glutarate: insight into the basis of its substrate specificity. *Chem. Biol.* **8**, 1253-1264.

❖ Kim, Y., Kim, S., Earnest, T. N. & Hol, W. G. (2002). Precursor structure of cephalosporin acylase. Insights into autoproteolytic activation in a new N-terminal hydrolase family. *J. Biol. Chem.* **277**, 2823-2829.

❖ Kim, J. K., Yang, I. S., Rhee, S., Dauter, Z., Lee, Y. S., Park, S. S. & Kim, K. H. (2003). Crystal structures of glutaryl 7-aminocephalosporanic acid acylase: Insight into autoproteolytic activation. *Biochemistry,* **42**, 4084-4093.

❖ Kim, J. K., Yang, I. S., Shin, H. J., Cho, K. J., Ryu, E. K., Kim, S. H., Park, S. S. & Kim, K. H. (2006). Insight into autoproteolytic activation from the structure of cephalosporin acylase: A protein with two proteolytic chemistries. *Proc. Natl. Acad. Sci. USA,* **103**, 1732-1737.

❖ Kinoshita, T., Tada, T., Saito, Y., Ishii, Y., Sato, A. & Murata, M. (2000). Crystallization and preliminary X-ray analysis of cephalosporin C acylase from *Pseudomonas* sp. strain N176. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **56**, 458-459.

❖ Kishore, N. S., Wood, D. C., Mehta, P. P., Wade, A. C., Lu, T., Gokel, G. W. & Gordon, J. I. (1993). Comparison of the acyl chain specificities of human myristoyl-CoA synthetase and human myristoyl-CoA:protein N-myristoyltransferase. *J. Biol. Chem.* **268**, 4889-4902.

❖ Klei, H. E., Daumy, G. O. & Kelly, J. A. (1995). Purification and preliminary crystallographic studies of penicillin G acylase from *Providencia rettgeri*. *Protein Sci.* **4**, 433-441.

❖ Konstantinovic, M., Marjanovic, N., Ljubijankic, G. & Glisin, V. (1994). The penicillin amidase of *Arthrobacter viscosus* (ATCC 15294). *Gene,* **143**, 79-83.

❖ Krueger, J. G., Garber, E. A., Goldberg, A. R. & HanafUSA, H. (1982). Changes in amino terminal sequences of p60[src] lead to decreased membrane association and decreased *in vivo* tumorigenicity. *Cell,* **28**, 889-896.

❖ Kumar, A., Prabhune, A., Suresh, C. G. & Pundle, A. (2008). Characterization of smallest active monomeric penicillin V acylase from new source: A yeast, *Rhodotorula aurantiaca* (NCIM 3425). *Process Biochem.* **43**, 961-967.

❖ Kumar, K. K., Sudhakaran, V., Deshpande, B. S., Ambedkar, S. S. & Shewale, J. G. (1993). Cephalosporin acylases: enzyme production, structure and application in the production of 7-ACA. *Hind. Antibiot. Bull.* **35**, 111-125.

❖ Kumar, R. S., Brannigan, J. A., Prabhune, A. A., Pundle, A. V., Dodson, G. G., Dodson, E. J. & Suresh, C. G. (2006). Structural and functional analysis of a conjugated bile salt hydrolase from *Bifidobacterium longum* reveals an evolutionary relationship with penicillin V acylase. *J. Biol. Chem.* **281**, 32516-32525.

❖ Kumar, R. S., Prabhune, A. A., Pundle, A. V., Karthikeyan, M. & Suresh, C. G. (2007). A tryptophan residue is identified in the substrate binding of penicillin G acylase from *Kluyvera citrophila. Enz. Microb. Technol.* **40**, 1389-1397.

❖ Kumar, R. S., Suresh, C. G., Pundle, A. & Prabhune, A. (2004). Evidence for the involvement of arginyl residue at the active site of penicillin G acylase from Kluyvera citrophila. *Biotechnol. Lett.* **26**, 1601-1606.

❖ Kumar, S., Tsai, C. J. & Nussinov, R. (2000). Factors enhancing protein thermostability. *Protein Eng.* **13**, 179-191.

❖ Kutzbach, C. & Rauenbusch, E. (1974). Preparation and general properties of crystalline penicillin acylase from *Escherichia coli* ATCC 11 105. *Hoppe Seylers Z. Physiol. Chem.* **354,** 45-53.

❖ Kwon, Y. D., Nagy, I., Adams, P. D., Baumeister, W. & Jap, B. K. (2004). Crystal structures of the *Rhodococcus* proteasome with and without its pro-peptides: implications for the role of the pro-peptide in proteasome assembly. *J. Mol. Biol.* **335**, 233-245.

❖ Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680-685.

❖ Lakomek, K., Dickmanns, A., Kettwig, M., Urlaub, H., Ficner, R. & Lubke, T. (2009). Initial insight into the function of the lysosomal 66.3 kDa protein from mouse by means of X-ray crystallography. *BMC Struct. Biol.* **9**, 56.

❖ Landsberg, H. (1949). Prelude to the discovery of penicillin. *Isis,* **40**.

❖ Langner, C. A., Lodge, J. K., Travis, S. J., Caldwell, J. E., Lu, T. B., Li, Q., Bryant, M. L., Devadas, B., Gokel, G. W., Kobayashi, G. S. & Gordon, J. I. (1992). 4-oxatetradecanoic acid

is fungicidal for *Cryptococcus neoformans* and inhibits replication of human immunodeficiency virus I. *J. Biol. Chem.* **267**, 17159-17169.

❖ Larsen, T. M., Boehlein, S. K., Schuster, S. M., Richards, N. G. J., Thoden, J. B., Holden, H. M. & Rayment, I. (1999). Three-dimensional structure of *Escherichia coli* asparagine synthetase B: a short journey from substrate to product. *Biochemistry,* **38**, 16146-16157.

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

❖ Latge, J. P. (1999). *Aspergillus fumigatus* and aspergillosis. *Clin. Microbiol. Rev.* **12**, 310-350.

❖ Lebbink, J. H. G., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. & de Vos, W. M. (1998). Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. I. Introduction of a six-residue ion-pair network in the hinge region. *J. Mol. Biol.* **280**, 287-296.

❖ Lee, H., Park, O. K. & Kang, H. S. (2000a). Identification of a new active site for autocatalytic processing of penicillin acylase precursor in *Escherichia coli* ATCC11105. *Biochem. Biophys. Res. Commun.* **272**, 199-204.

❖ Lee, Y. S., Kim, H. W. & Park, S. S. (2000b). The role of alpha-amino group of the N-terminal serine of beta subunit for enzyme catalysis and autoproteolytic activation of glutaryl 7-aminocephalosporanic acid acylase. *J. Biol. Chem.* **275**, 39200-39206.

❖ Lee, Y. S. & Park, S. S. (1998). Two-step autocatalytic processing of the glutaryl 7-aminocephalosporanic acid acylase from *Pseudomonas* sp. strain GK16. *J. Bacteriol.* **180**, 4576-4582.

❖ Lehrer, S. S. (1971). Solute perturbation of protein fluorescence. The quenching of the tryptophyl fluorescence of model compounds and of lysozyme by iodide ion. *Biochemistry,* **10**, 3254-3263.

❖ Leslie, A. G. W. & Powell, H. R. (2007). Processing diffraction data with *MOSFLM*. In Read, R. J. & Sussman, J. L. (eds.) *Evolving Methods for Macromolecular Crystallography*, **245,** Springer Press, 41-51

❖ Levitt, M. (1978). Conformational preferences of amino acids in globular proteins. *Biochemistry,* **17**, 4277-4284.

❖ Lewis, P. N., Momany, F. A. & Scheraga, H. A. (1973). Chain reversals in proteins. *Biochim. Biophys. Acta,* **303**, 211-229.

❖ Li, Y., Chen, J. F., Jiang, W. H., Mao, X., Zhao, G. P. & Wang, E. D. (1999). *In vivo* post-translational processing and subunit reconstitution of cephalosporin acylase from *Pseudomonas* sp 130. *Eur. J. Biochem.* **262**, 713-719.

❖ Little, J. W. (1993). LexA cleavage and other self processing reactions. *J. Bacteriol.* **175**, 4943-4950.

❖ Littlechild, J. A. (1991). Protein crystallization: magical or logical: can we establish some general rules? *J. Phys. D: Appl. Phys.* **24**, 111.

❖ Liu, S. L., Wei, D. Z., Song, Q. X., Zhang, Y. W. & Wang, X. D. (2006). Effect of organic cosolvent on kinetic resolution of tert-leucine by penicillin G acylase from *Kluyvera citrophila*. *Bioproc. Biosyst. Eng.* **28**, 285-289.

❖ Livermore, D. M. (2000). Antibiotic resistance in staphylococci. *Int. J. Antimicrob. Agents,* **16 (1)**, S3-10.

❖ Lodge, J. K., Jackson-Machelski, E., Devadas, B., Zupec, M. E., Getman, D. P., Kishore, N., Freeman, S. K., McWherter, C. A., Sikorski, J. A. & Gordon, J. I. (1997). N-myristoylation of Arf proteins in *Candida albicans*: an *in vivo* assay for evaluating antifungal inhibitors of myristoyl-CoA:protein N-myristoyltransferase. *Microbiology,* **143(2)**, 357-366.

❖ Lodge, J. K., Jackson-Machelski, E., Higgins, M., McWherter, C. A., Sikorski, J. A., Devadas, B. & Gordon, J. I. (1998). Genetic and biochemical studies establish that the fungicidal effect of a fully depeptidized inhibitor of *Cryptococcus neoformans* myristoyl-CoA : protein N-myristoyltransferase (Nmt) is Nmt-dependent. *J. Biol. Chem.* **273**, 12482-12491.

❖ Lodge, J. K., Jackson- Machelski, E., Toffaletti, D. L., Perfect, J. R. & Gordon, J. I. (1994). Targeted gene replacement demonstrates that myristoyl-Co A: protein N-myristoyltransferase is essential for viability of *Cryptococcus neoformans*. *Pro. Nat. Acad. Sci. USA,* **91**, 12008-12012.

❖ Lowe, D. A., Romanchik, G. & Elander, R. P., (1981). Penicillin acylases – a review of existing enzymes and the isolation of a new bacterial penicillin V acylase. *Dev. Ind. Microbiol.* **22**; 163-180.

❖ Lowe, J., Stock, D., Jap, R., Zwickl, P., Baumeister, W. & Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 angstrom resolution. *Science,* **268**, 533-539.

❖ Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265-275.

❖ Lu, Y., Selvakumar, P., Ali, K., Shrivastav, A., Bajaj, G., Resch, L., Griebel, R., Fourney, D., Meguro, K. & Sharma, R. K. (2005). Expression of N-myristoyltransferase in human brain tumors. *Neurochem. Res.* **30**, 9-13.

❖ Magnuson, B. A., Raja, R. V. S., Moyana, T. N. & Sharma, R. K. (1995). Increased N-myristoyltransferase activity observed in rat and human colonic tumors. *J. Nat. Cancer Inst.* **87**, 1630-1635.

❖ Malik, M. A., Al-Thabaiti, S. A. & Malik, M. A. (2012). Synthesis, structure optimization and antifungal screening of novel tetrazole ring bearing acyl-hydrazones. *Int. J. Mol. Sci.* **13**, 10880-10898.

❖ Mao, X., Wang, W. W., Jiang, W. H. & Zhao, G. P. (2004). His23 beta and Glu455 beta of the *Pseudomonas* sp 130 glutaryl-7-amino cephalosporanic acid acylase are crucially important for efficient autoproteolysis and enzymatic catalysis. *Protein J.* **23**, 197-204.

❖ Marc, D., Drugeon, G., Haenni, A. L., Girard, M. & van der Werf, S. (1989). Role of myristoylation of poliovirus capsid protein VP4 as determined by site-directed mutagenesis of its N-terminal sequence. *EMBO J.* **8**, 2661-2668.

❖ Margolin, A. L., Svedas, V. K. & Berezin, I. V. (1980). Substrate specificity of penicillin amidase from *E. coli*. *Biochim. Biophys. Acta.* **616**, 283-289.

❖ Martin, J., Slade, A., Aitken, A., Arche, R. & Virden, R. (1991). Chemical modification of serine at the active-site of penicillin acylase from *Kluyvera citrophila*. *Biochem. J.* **280**, 659-662.

❖ Martin, L., Prieto, M. A., Cortes, E. & Garcia, J. L. (1995). Cloning and sequencing of the pac gene encoding the penicillin G acylase of *Bacillus megaterium* ATCC 14945. *FEMS Microbiol. Lett.* **125**, 287-292.

❖ Martinez, A., Traverso, J. A., Valot, B., Ferro, M., Espagne, C., Ephritikhine, G., Zivy, M., Giglione, C. & Meinnel, T. (2008). Extent of N-terminal modifications in cytosolic proteins from eukaryotes. *Proteomics,* **8**, 2809-2831.

❖ Masubuchi, M., Kawasaki, K., Ebiike, H., Ikeda, Y., Tsujii, S., Sogabe, S., Fujii, T., Sakata, K., Shiratori, Y., Aoki, Y., Ohtsuka, T. & Shimma, N. (2001). Design and synthesis of novel benzofurans as a new class of antifungal agents targeting fungal N-myristoyltransferase. Part 1. *Bioorg. Med. Chem. Lett.* **11**, 1833-1837.

❖ Matsuda, A. & Komatsu, K. I. (1985). Molecular cloning and structure of the gene for 7-Beta-(4-Carboxybutanamido)cephalosporanic acid acylase from a *Pseudomonas* strain. *J. Bacteriol.* **163**, 1222-1228.

❖ Matsuda, A., Matsuyama, K., Yamamoto, K., Ichikawa, S. & Komatsu, K. (1987). Cloning and characterization of the genes for two distinct cephalosporin acylases from a *Pseudomonas* strain. *J. Bacteriol.* **169**, 5815-5820.

❖ Matsumoto K (1993) Production of 6-APA, 7-ACA and 7-ADCA by immobilized cephalosporin amidase. In Tanaka, A., Tosa, T. & Kobayashi, T. (eds.) *Industrial application of immobilized biocatalysis*. Marcel Dekker, New York, 67–88.

❖ Matsumura, M. & Matthews, B. W. (1991). Stabilization of functional proteins by introduction of multiple disulfide bonds. *Methods Enzymol.* **202**, 336-356.

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

❖ Matthews, B. W. (1985). Determination of protein molecular weight, hydration, and packing from crystal density. *Methods Enzymol.* **114**, 176-187.

❖ Matthews Bw Fau - Nicholson, H., Nicholson H Fau - Becktel, W. J. & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl. Acad. Sci. USA,* **84**, 6663-6667.

❖ Maurer-Stroh, S., Eisenhaber, B. & Eisenhaber, F. (2002). N-terminal N-myristoylation of proteins: prediction of substrate proteins from amino acid sequence. *J. Mol. Biol.* **317**, 541-557.

❖ Maurer-Stroh, S. & Eisenhaber, F. (2004). Myristoylation of viral and bacterial proteins. *Trends Microbiol.* **12**, 178-185.

❖ McCoy, A. (2007). Solving structures of protein complexes by molecular replacement with *Phaser*. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **63**, 32-41.

❖ McCullough, J. E. (1983). Gene Cloning in bacilli related to enhanced penicillin acylase production. *Nat. Biotechnol.* **1**, 879-885.

❖ McDonough, M. A., Klei, H. E. & Kelly, J. A. (1999). Crystal structure of penicillin G acylase from the Bro1 mutant strain of *Providencia rettgeri*. *Protein Sci.* **8**, 1971-1981.

❖ McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). Presenting your structures: the *CCP4mg* molecular-graphics software. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **67**, 386-394.

❖ McPherson, A. (1982). Preparation and analysis of protein crystals. John Wiley & Sons.

❖ McPherson, A. (1999). Crystallization of biological macromolecules. Cold Spring Harbor Laboratory Press.

❖ McVey, C. E. (1999). Substrate specificity of penicillin G acylase. PhD. Thesis. York: University of York, UK.

❖ McVey, C. E., Walsh, M. A., Dodson, G. G., Wilson, K. S. & Brannigan, J. A. (2001). Crystal structures of penicillin acylase enzyme-substrate complexes: Structural insights into the catalytic mechanism. *J. Mol. Biol.* **313**, 139-150.

❖ Menendez-Arias, L. & Argos, P. (1989). Engineering protein thermal stability. Sequence statistics point to residue substitutions in alpha-helices. *J. Mol. Biol.* **206**, 397-406.

❖ Merkler, D. J., Farrington, G. K. & Wedler, F. C. (1981). Protein thermostability. Correlations between calculated macroscopic parameters and growth temperature for closely related thermophilic and mesophilic bacilli. *Int. J. Pep. Prot. Res.* **18**, 430-442.

❖ Michalska, K., Borek, D., Hernandez-Santoyo, A. & Jaskolski, M. (2008a). Crystal packing of plant-type L-asparaginase from *Escherichia coli. Acta Crystallogr. Sect.D: Biol. Crystallogr.* **64**, 309-320.

❖ Michalska, K., Bujacz, G. & Jaskolski, M. (2006). Crystal structure of plant asparaginase. *J. Mol. Biol.* **360**, 105-116.

❖ Michalska, K., Hernandez-Santoyo, A. & Jaskolski, M. (2008b). The mechanism of autocatalytic activation of plant-type L-asparaginases. *J. Biol. Chem.* **283**, 13388-13397.

❖ Miller, M. T., Gerratana, B., Stapon, A., Townsend, C. A. & Rosenzweig, A. C. (2003). Crystal Structure of carbapenam synthetase (CarA). *J. Biol. Chem.* **278**, 40996-41002.

❖ Miller, M. D., Warmerdam, M. T., Gaston, I., Greene, W. C. & Feinberg, M. B. (1994). The human immunodeficiency virus-1 Nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages. *J. Exp. Med.* **179**, 101-113.

❖ Minari, A., Husni, R., Avery, R. K., Longworth, D. L., DeCamp, M., Bertin, M., Schilz, R., Smedira, N., Haug, M. T., Mehta, A. & Gordon, S. M. (2002). The incidence of invasive aspergillosis among solid organ transplant recipients and implications for prophylaxis in lung transplants. *Transpl. Infect. Dis.* **4**, 195-200.

❖ Mishkind, M. (2001). Morbid myristoylation. *Trends Cell Biol.* **11**, 191.

❖ Moreno, J., Nieto, J., Masina, S., Canavate, C., Cruz, I., Chicharro, C., Carrillo, E., Napp, S., Reymond, C., Kaye, P. M., Smith, D. F., Fasel, N. & Alvar, J. (2007). Immunization with H1, HASPB1 and MML Leishmania proteins in a vaccine trial against experimental canine leishmaniasis. *Vaccine,* **25**, 5290-5300.

❖ Moscufo, N., Simons, J. & Chow, M. (1991). Myristoylation is important at multiple stages in poliovirus assembly. *J. Virol.* **65**, 2372-2380.

❖ Mrabet, N. T., Vandenbroeck, A., Vandenbrande, I., Stanssens, P., Laroche, Y., Lambeir, A. M., Matthijssens, G., Jenkins, J., Chiadmi, M., Vantilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., Demaeyer, M. & Wodak, S. J. (1992). Arginine residues as stabilizing elements in proteins. *Biochemistry,* **31**, 2239-2253.

❖ Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. *Sci. Am.* **262**, 56-61, 64-55.

❖ Mullis, K. B. & Faloona, F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods Enzymol.* **155**, 335-350.

❖ Murao, S. (1955). Penicillin-amidase. III. Mechanism of penicillin-amidase on sodium penicillin G. *J. Agric. Chem. Soc. (Japan)*, **29**; 404-407.

❖ Murao, S., & Kishida, Y. (1961) Studies on penicillin-amidase. Part V Studies on the penicillin amidase producing microorganisms *J. Agric. Chem. Soc. (Japan)*, **35**, 610-615.

❖ Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **53**, 240-255.

❖ Muslin, E. H., Clark, S. E. & Henson, C. A. (2002). The effect of proline insertions on the thermostability of a barley alpha-glucosidase. *Protein Eng.* **15**, 29-33.

❖ Nagarajan, S. R., Devadas, B., Zupec, M. E., Freeman, S. K., Brown, D. L., Lu, H. F., Mehta, P. P., Kishore, N. S., McWherter, C. A., Getman, D. P., Gordon, J. I. & Sikorski, J. A. (1997). Conformationally constrained [p-(omega-aminoalkyl)phenacetyl]-L-seryl-L-lysyl dipeptide amides as potent peptidomimetic inhibitors of *Candida albicans* and human myristoyl-CoA:protein N-myristoyl transferase. *J. Med. Chem.* **40**, 1422-1438.

❖ Nakai, T., Hasegawa, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Ueki, T., Nanba, H., Ikenaka, Y., Takahashi, S., Sato, M. & Tsukihara, T. (2000). Crystal structure of N-carbamyl-D-amino acid amidohydrolase with a novel catalytic framework common to amidohydrolases. *Structure,* **8**, 729-737.

❖ Nakamura, S., Tanaka, T., Yada, R. Y. & Nakai, S. (1997). Improving the thermostability of *Bacillus stearothermophilus* neutral protease by introducing proline into the active site helix. *Protein Eng.* **10**, 1263-1269.

❖ Nara, T., Okachi, R. & Misawa, M. (1971a). Enzymatic synthesis of D(-)-alpha-aminobenzylpenicillin by *Kluyvera citrophila*. *J. Antibiot.* **24**, 321-323.

❖ Nara, T., Misawa, M., Okachi, R. & Yamamoto, M. (1971b). Enzymatic synthesis of D(-)-α-aminobenzylpenicillin. Part I. Selection of penicillin acylase-producing bacteria. *Agric. Biol. Chem.* **35**: 1676–1682

❖ Navaza, J. (1994). *AMoRe*: an automated package for molecular replacement. *Acta Crystallogr. Sect. A: Foundn. Crystallogr.* **50**, 157-163.

❖ Nayler, J. H. (1991a). Early discoveries in the penicillin series. *Trends Biochem. Sci.* **16**, 195-197.

❖ Nayler, J. H. (1991b). Semi-synthetic approaches to novel penicillins. *Trends Biochem. Sci.* **16**, 234-237.

❖ Negoro, S., Shibata, N., Tanaka, Y., Yasuhira, K., Shibata, H., Hashimoto, H., Lee, Y.-H., Oshima, S., Santa, R., Mochiji, K., Goto, Y., Ikegami, T., Nagai, K., Kato, D.-i., Takeo, M. & Higuchi, Y. (2012). Three-dimensional structure of nylon hydrolase and mechanism of nylon-6 hydrolysis. *J. Biol. Chem.* **287**, 5079-5090.

❖ Ntwasa, M., Aapies, S., Schiffmann, D. A. & Gay, N. J. (2001). Drosophila embryos lacking N-myristoyltransferase have multiple developmental defects. *Exp. Cell. Res.* **262**, 134-144.

❖ Ntwasa, M., Egerton, M. & Gay, N. J. (1997). Sequence and expression of Drosophila myristoyl-CoA:protein N-myristoyl transferase: evidence for proteolytic processing and membrane localisation. *J. Cell Sci.* **110 (2)**, 149-156.

❖ Oh, S. J., Kim, Y. C., Park, Y. W., Min, S. Y., Kim, I. S. & Kang, H. S. (1987). Complete nucleotide sequence of the penicillin G acylase gene and the flanking regions, and its expression in *Escherichia coli*. *Gene,* **56**, 87-97.

❖ Oh, B., Kim, M., Yoon, J., Chung, K., Shin, Y., Lee, D. & Kim, Y. (2003). Deacylation activity of cephalosporin acylase to cephalosporin C is improved by changing the side-chain conformations of active-site residues. *Biochem. Biophys. Res. Commun.* **310**, 19-27.

❖ Ohashi, H., Katsuta, Y., Hashizume, T., Abe, S. N., Kajiura, H., Hattori, H., Kamei, T. & Yano, M. (1988). Molecular cloning of the penicillin G acylase gene from *Arthrobacter viscosus*. *Appl. Environ. Microbiol.* **54**, 2603-2607.

❖ Ohashi, H., Katsuta, Y., Nagashima, M., Kamei, T. & Yano, M. (1989). Expression of the *Arthrobacter viscosus* penicillin G acylase gene in *Escherichia coli* and *Bacillus subtilis*. *Appl. Environ. Biotechnol.* **55**, 1351-1356.

❖ Okachi, R., Nara, T. & Misawa, M. (1972). Production of 6-aminopenicillanic acid by *Kluyvera citrophila* KY 3641. *Agric. Biol. Chem.* **36**, 925-930.

❖ Okachi, R., Kato, F., Miyamura, Y. & Nara, T. (1973). Selection of *Pseudomonas melanogenum* KY 3987 as a new ampicillin-producing bacteria. *Agric. Biol. Chem.* **37**, 1953-1957.

❖ Okada, T., Suzuki, H., Wada, K., Kumagai, H. & Fukuyama, K. (2006). Crystal structures of gamma-glutamyltranspeptidase from *Escherichia coli*, a key enzyme in glutathione metabolism, and its reaction intermediate. *Proc. Natl. Acad. Sci. USA,* **103**, 6471-6476.

❖ Okada, T., Suzuki, H., Wada, K., Kumagai, H. & Fukuyama, K. (2007). Crystal structure of the gamma-glutamyltranspeptidase precursor protein from *Escherichia coli.* Structural changes autocatalytic processing and implications for the maturation mechanism. *J. Biol. Chem.* **282**, 2433-2439.

❖ Oliver, G., Valle, F., Rosetti, F., Gomez-Pedrozo, M., Santamaría, P., Gosset, G. & Bolivar, F. (1985). A common precursor for the two subunits of the penicillin acylase from *Escherichia coli* ATCC11105. *Gene,* **40**, 9-14.

❖ Olsson, A., Hagstrom, T., Nilsson, B., Uhlen, M. & Gatenbeck, S. (1985). Molecular cloning of *Bacillus sphaericus* penicillin V amidase gene and its expression in *Escherichia coli* and *Bacillus subtilis*. *Appl. Environ. Microbiol.* **49**, 1084-1089.

❖ Olsson, A. & Uhlen, M. (1986). Sequencing and heterologous expression of the gene encoding penicillin V amidase from *Bacillus sphaericus*. *Gene,* **45**, 175-181.

❖ Oinonen, C., Tikkanen, R., Rouvinen, J. & Peltonen, L. (1995). Three-dimensional structure of human lysosomal aspartyiglucosaminidase. *Nat. Struct. Biol.* **2**, 1102-1108.

❖ Oinonen, C. & Rouvinen, J. (2000). Structural comparison of Ntn-hydrolases. *Protein Sci.* **9**, 2329–2337.

❖ Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). CATH--a hierarchic classification of protein domain structures. *Structure,* **5**, 1093-1108.

❖ Ospina, S., Barzana, E., Ramirez, O. T. & LopezMunguia, A. (1996). Effect of pH in the synthesis of ampicillin by penicillin acylase. *Enz. Microb. Technol.* **19**, 462-469.

❖ Otwinowski, Z. & Minor W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol.* **276**, Macromolecular Crystallography, part A, 307-326, Academic Press (New York).

❖ Paige, L. A., Zheng, G. Q., Defrees, S. A., Cassady, J. M. & Geahlen, R. L. (1989). S-(2-Oxopentadecyl)-CoA, a nonhydrolyzable analog of myristoyl-CoA, is a potent inhibitor of myristoyl-CoA protein: N-myristoyltransferase. *J. Med. Chem.* **32**, 1665-1667.

❖ Paige, L. A., Zheng, G. Q., Defrees, S. A., Cassady, J. M. & Geahlen, R. L. (1990). Metabolic activation of 2-substituted derivatives of myristic acid to form potent inhibitors of myristoyl CoA:protein N-myristoyltransferase. *Biochemistry,* **29**, 10566-10573.

❖ Panethymitaki, C., Bowyer, P. W., Price, H. P., Leatherbarrow, R. J., Brown, K. A. & Smith, D. F. (2006). Characterization and selective inhibition of myristoyl-CoA:protein N-myristoyltransferase from *Trypanosoma bruce*i and *Leishmania major. Biochem. J.* **396**, 277-285.

❖ Pannu, N. S. & Read, R. J. (1996). Improved Structure Refinement Through Maximum Likelihood. *Acta Crystallogr. Sect. A: Foundn. Crystallogr.* **52**, 659-668.

❖ Parang K Fau - Knaus, E. E., Knaus Ee Fau - Wiebe, L. I., Wiebe Li Fau - Sardari, S., Sardari S Fau - Daneshtalab, M., Daneshtalab M Fau - Csizmadia, F. & Csizmadia, F. (1996). Synthesis and antifungal activities of myristic acid analogs. *Arch. Pharm. (Weinheim)* **329**, 475-482.

❖ Parmar, A., Kumar, H., Marwaha, S. & Kennedy, J. F. (2000). Advances in enzymatic transformation of penicillins to 6-aminopenicillanic acid (6-APA). *Biotechnol. Adv.* **18**, 289-301.

❖ Pei, J. & Grishin, N. V. (2003). Peptidase family U34 belongs to the superfamily of N-terminal nucleophile hydrolases. *Protein Sci.* **12**, 1131-1135.

❖ Peng, B. & Robert-Guroff, M. (2001). Deletion of N-terminal myristoylation site of HIV Nef abrogates both MHC-1 and CD4 down-regulation. *Immunol. Lett.* **78**, 195–200.

❖ Perakyla, M. & Rouvinen, J. (1996). Ab initio quantum mechanical model calculations on the catalytic mechanism of aspartylglucosaminidase (AGA): A serine protease-like mechanism with an N-terminal threonine and substrate-assisted catalysis. *Chem. Eur. J.* **2**, 1548-1551.

❖ Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J. & Belfort, M. (1994). Protein splicing elements: inteins and exteins - a definition of terms and recommended nomenclature. *Nucleic Acids Res.* **22**, 1125-1127.

❖ Petrovan, R. J., Govers-Riemslag, J. W. P., Nowak, G., Hemker, H. C., Tans, G. & Rosing, J. (1998). Autocatalytic peptide bond cleavages in prothrombin and meizothrombin. *Biochemistry,* **37**, 1185-1191.

❖ Price, H. P., Guther, M. L. S., Ferguson, M. A. J. & Smith, D. F. (2010). Myristoyl-CoA:protein N-myristoyltransferase depletion in trypanosomes causes avirulence and endocytic defects. *Mol. Biochem. Parasitol.* **169**, 55-58.

❖ Price, H. P., Menon, M. R., Panethymitaki, C., Goulding, D., McKean, P. G. & Smith, D. F. (2003). Myristoyl-CoA : protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J. Biol. Chem.* **278**, 7206-7214.

❖ Prieto, I., Martin, J., Arche, R., Fernandez, P., Perez-Aranda, A. & Barbero, J. L. (1990). Penicillin acylase mutants with altered site-directed activity from *Kluyvera citrophila*. *Appl. Microbiol. Biotechnol.* **33**, 553-559.

❖ Prieto, I., Rodriguez, M. C., Marquez, G., Perez-Aranda, A. & Barbero, J. L. (1992). Changing glycine 21 for glutamic acid in the beta-subunit of penicillin G acylase from *Kluyvera citrophila* prevents protein maturation. *Appl. Microbiol. Biotechnol.* **36**, 659-662.

❖ Prieto, M. A., Perez-Aranda, A. & Garcia, J. L. (1993). Characterization of an *Escherichia coli* aromatic hydroxylase with a broad substrate range. *J. Bacteriol.* **175**, 2162-2167.

❖ Prieto, M. A., Diaz, E. & Garcia, J. L. (1996). Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: engineering a mobile aromatic degradative cluster. *J. Bacteriol.* **178**, 111-120.

❖ Provitera, P., El-Maghrabi, R. & Scarlata, S. (2006). The effect of HIV-1 Gag myristoylation on membrane binding. *Biophys. Chem.* **119**, 23-32.

❖ Qian, X., Guan, C. & Guo, H.-C. (2003). A dual role for an aspartic acid in glycosylasparaginase autoproteolysis. *Structure,* **11**, 997-1003.

❖ Rajala, R. V., Radhi, J. M., Kakkar, R., Datla, R. S. & Sharma, R. K. (2000a). Increased expression of N-myristoyltransferase in gallbladder carcinomas. *Cancer,* **88**, 1992-1999.

❖ Rajala, R. V. S., Dehm, S., Bi, X. G., Bonham, K. & Sharma, R. K. (2000b). Expression of N-myristoyltransferase inhibitor protein and its relationship to c-Src levels in human colon cancer cell lines. *Biochem. Biophys. Res. Commun.* **273**, 1116-1120.

❖ Rajendhran, J., Krishnakumar, V. & Gunasekaran, P. (2002). Optimization of a fermentation medium for the production of penicillin G acylase from *Bacillus sp. Lett. Appl. Microbiol.* **35**, 523-527.

❖ Rajendhran, J. & Gunasekaran, P. (2004). Recent biotechnological interventions for developing improved penicillin G acylases. *J. Biosci. Bioeng.* **97**, 1-13.

❖ Raju, R. V., Magnuson, B. A. & Sharma, R. K. (1995). Mammalian myristoyl CoA:protein N-myristoyltransferase. *Mol. Cell Biochem.* **149-150**, 191-202.

❖ Raju, R. V., Moyana, T. N. & Sharma, R. K. (1997). N-Myristoyltransferase overexpression in human colorectal adenocarcinomas. *Exp. Cell. Res.* **235**, 145-154.

❖ Raju, R. V. S., Datla, R. S. S. & Sharma, R. K. (1996). Overexpression of human N-myristoyltransferase utilizing a T7 polymerase gene expression system. *Protein Expr. Purif.* **7**, 431-437.

❖ Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283-438.

❖ Rathinaswamy, P., Pundle, A. V., Prabhune, A. A., SivaRaman, H., Brannigan, J. A., Dodson, G. G. & Suresh, C. G. (2005). Cloning, purification, crystallization and preliminary structural studies of penicillin V acylase from *Bacillus subtilis*. *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **61**, 680-683.

❖ Ravelli, R. B. & Garman, E. F. (2006). Radiation damage in macromolecular cryocrystallography. *Curr. Opin. Struct. Biol.* **16**, 624-629.

❖ Ravelli, R. B. G., Leiros, H. K. S., Pan, B. C., Caffrey, M. & McSweeney, S. (2003). Specific radiation damage can be used to solve macromolecular crystal structures. *Structure,* **11**, 217-224.

❖ Rawlings, N. D., Tolle, D. P. & Barrett, A. J. (2004). Evolutionary families of peptidase inhibitors. *Biochem. J.* **378**, 705-716.

❖ Read, R. J. (1997). Model phases: Probabilities and bias. In: *Methods in Enzymology* (Charles, W., Carter Jr, R. M. S., (ed.), **277**, Academic Press, p110-128.

❖ Read, T. D., Salzberg, S. L., Pop, M., Shumway, M., Umayam, L., Jiang, L., Holtzapple, E.,Busch, J. D., Smith, K. L., Schupp, J. M., Solomon, D., Keim, P. & Fraser, C. M. (2002). Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. *Science*, **296**, 2028-2033.

❖ Renslo, A. R. & McKerrow, J. H. (2006). Drug discovery and development for neglected parasitic diseases. *Nat. Chem. Biol.* **2**, 701-710.

❖ Resh, M. D. (1999). Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim. Biophys. Acta,* **1451**, 1-16.

❖ Resh, M. D. (2006). Trafficking and signaling by fatty-acylated and prenylated proteins. *Nat. Chem. Biol.* **2**, 584-590.

❖ Rhodes, G. (2000). Crystallography Made Crystal Clear. Academic Press.

❖ Rioux, V., Beauchamp, E., Pedrono, F., Daval, S., Molle, D., Catheline, D. & Legrand, P. (2006). Identification and characterization of recombinant and native rat myristoyl-CoA:protein N-myristoyltransferases. *Mol. Cell Biochem.* **286**, 161-170.

❖ Roa, A., Castillon, M. P., Goble, M. L., Virden, R. & Garcia, J. L. (1995). New insights on the specificity of penicillin acylase. *Biochem. Biophys. Res. Commun.* **206**, 629-636.

❖ Roche, D., Prasad, K. & Repic, O. (1999). Enantioselective acylation of beta-aminoesters using penicillin G acylase in organic solvents. *Tetrahedron Lett.* **40**, 3665-3668.

❖ Rolinson, G. N., Batchelor, F. R., Butterworth, D., Cameron-Wood, J., Cole, M., Eustace, G. C., Hart, M. V., Richards, M. & Chain, E. B. (1960). Formation of 6-aminopenicillanic acid from penicillin by enzymatic hydrolysis. *Nature,* **187**, 236-237.

❖ Rosell, C. M., Terreni, M., Fernandez-Lafuente, R. & Guisan, J. M. (1998). A criterion for the selection of monophasic solvents for enzymatic synthesis. *Enz. Microb. Technol.* **23**, 64-69.

❖ Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* **15**, 24-31.

❖ Rossocha, M., Schultz-Heienbrok, R., von Moeller, H., Coleman, J. P. & Saenger, W. (2005). Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product. *Biochemistry,* **44**, 5739-5748.

❖ Rudnick, D. A., Mcwherter, C. A., Rocque, W. J., Lennon, P. J., Getman, D. P. & Gordon, J. I. (1991). Kinetic and structural evidence for a sequential ordered Bi-Bi mechanism of catalysis by *Saccharomyces cerevisiae* myristoyl-CoA:protein N-myristoyltransferase. *J. Biol. Chem.* **266**, 9732-9739.

❖ Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* **239**, 487-491.

❖ Sakaguchi, K. & Murao, S. (1950). A preliminary report on a new enzyme, penicillinamidase. *J. Agric. Chem. Soc.(Japan),* **23**; 411.

❖ Salminen, T., Teplyakov, A., Kankare, J., Cooperman, B. S., Lahti, R. & Goldman, A. (1996). An unusual route to thermostability disclosed by the comparison of *Thermus thermophilus* and *Escherichia coli* inorganic pyrophosphatases. *Protein Sci.* **5**, 1014-1025.

❖ Sanglard, D. (2002). Resistance of human fungal pathogens to antifungal drugs. *Curr. Opin. Microbiol.* **5**, 379-385.

❖ Sankaranarayanan, R., Cherney, M. M., Garen, C., Garen, G., Niu, C., Yuan, M. & James, M. N. G. (2010). The molecular structure of ornithine acetyltransferase from *Mycobacterium tuberculosis* bound to ornithine, a competitive inhibitor. *J. Mol. Biol.* **397**, 979-990.

❖ Saridakis, V., Christendat, D., Thygesen, A., Arrowsmith, C. H., Edwards, A. M. & Pai, E. F. (2002). Crystal structure of *Methanobacterium thermoautotrophicum* conserved protein MTH1020 reveals an Ntn-hydrolase fold. *Proteins,* **48**, 141-143.

❖ Savidge, T. A. & Cole, M. (1975). Penicillin acylase (bacterial). *Methods Enzymol.* **43**, 705-721.

❖ Savidge, T. A. (1984). Enzymatic conversions used in the production of penicillins and cephalosporins. In Vandamme, E. (ed) *Biotechnology of Industrial Antibiotics*. Marcel Dekker, New York, p 171.

❖ Scharf, S. J., Horn, G. T. & Erlich, H. A. (1986). Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science,* **233**, 1076-1078.

❖ Schmidtke, G., Kraft, R., Kostka, S., Henklein, P., Frommel, C., Lowe, J., Huber, R., Kloetzel, P. M. & Schmidt, M. (1996). Analysis of mammalian 20S proteasome biogenesis: The maturation of beta-subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J.* **15**, 6887-6898.

❖ Schultz, A. M., Henderson, L. E., Oroszlan, S., Garber, E. A. & HanafUSA, H. (1985). Amino terminal myristylation of the protein kinase p60[src], a retroviral transforming protein. *Science,* **227**, 427-429.

❖ Schultz, A. M., Tsai, S. C., Kung, H. F., Oroszlan, S., Moss, J. & Vaughan, M. (1987). Hydroxylamine-stable covalent linkage of myristic acid in G0 alpha, a guanine nucleotide-binding protein of bovine brain. *Biochem. Biophys. Res. Commun.* **146**, 1234-1239.

❖ Schumacher, G., Sizmann, D., Haug, H., Buckel, P. & Bock, A. (1986). Penicillin acylase from *E. coli*: unique gene-protein relation. *Nucleic Acids Res.* **14**, 5713-5727.

❖ Seaton, K. E. & Smith, C. D. (2008). N-Myristoyltransferase isozymes exhibit differential specificity for human immunodeficiency virus type 1 Gag and Nef. *J. Gen. Virol.* **89**, 288-296.

❖ Selvakumar, P., Lakshmikuttyamma, A., Shrivastav, A., Das, S. B., Dimmock, J. R. & Sharma, R. K. (2007). Potential role of N-myristoyltransferase in cancer. *Prog. Lipid Res.* **46**, 1-36.

❖ Semba, K., Nishizawa, M., Miyajima, N., Yoshida, M. C., Sukegawa, J., Yamanashi, Y., Sasaki, M., Yamamoto, T. & Toyoshima, K. (1986). yes-related protooncogene, syn, belongs to the protein-tyrosine kinase family. *Proc. Natl. Acad. Sci. USA,* **83**, 5459-5463.

❖ Shao, Y., Xu, M. Q. & Paulus, H. (1996). Protein splicing: Evidence for an N-O acyl Rearrangement as the initial Step in the Splicing Process. *Biochemistry,* **35**, 3810-3815.

❖ Shaw, B. D., Momany, C. & Momany, M. (2002). *Aspergillus nidulans* swoF encodes an N-myristoyl transferase. *Eukaryot. Cell,* **1**, 241-248.

❖ Shaw, J. (2007). The leishmaniases - survival and expansion in a changing world. A mini-review. *Memorias Do Instituto Oswaldo Cruz.* **102**, 541-547.

❖ Sheehan, D. J., Hitchcock, C. A. & Sibley, C. M. (1999). Current and emerging azole antifungal agents. *Clin. Microbiol. Revi.* **12**, 40-79.

❖ Sheng, C., Xu, H., Wang, W., Cao, Y., Dong, G., Wang, S., Che, X., Ji, H., Miao, Z., Yao, J. & Zhang, W. (2010). Design, synthesis and antifungal activity of isosteric analogues of benzoheterocyclic N-myristoyltransferase inhibitors. *Eur. J. Med. Chem.* **45**, 3531-3540.

❖ Shewale, J. G., Deshpande, B. S., Sudhakaran, V. K. & Ambedkar, S. S. (1990). Penicillin acylases: applications and potentials. *Proces. Biochem.* **25**, 97-103.

❖ Shewale, G. J., Kumar, K. K., and Amberkar, G. R. (1987). Evaluation and determination of 6- aminopenicillinic acid by p-dimethylbenzaldehyde. *Biotechnol. Tech.* **1**, 69-72.

❖ Shewale, J. G. & Sivaraman, H. (1989). Penicillin acylase: Enzyme production and its application in the manufacture of 6-APA. *Proces. Biochem.* **24**, 146-154.

❖ Shewale, J. G. & Sudhakaran, V. K. (1997). Penicillin V acylase: Its potential in the production of 6-aminopenicillanic acid. *Enz. Microb. Technol.* **20**, 402-410.

❖ Shibuya, Y., Matsumoto, K. & Fujii, T. (1981). Isolation and properties of 7-β-(4-carboxybutanamido)cephalosporanic acid acylase-producing bacteria. *Agric. Biol. Chem.* **45**; 1561-1567.

❖ Shimizu, M., Okachi, R., Kimura, K. & Nara, T. (1975). Purification and properties of penicillin acylase from *Kluyvera citrophila. Agric. Biol. Chem.* **39**, 1655-1661.

❖ Sizmann, D., Keilmann, C. & Bock, A. (1990). Primary structure requirements for the maturation in vivo of penicillin acylase from *Escherichia coli* ATCC 11105. *Eur. J. Biochem.* **192**, 143-151.

❖ Shoji, S., Kurosawa, T., Inoue, H., Funakoshi, T. & Kubota, Y. (1990). Human cellular Src gene product - Identification of the myristoylated Pp60c-Src and blockage of its myristoyl acylation with N-fatty acyl compounds resulted in the suppression of colony formation. *Biochem. Biophys. Res. Commun.* **173**, 894-901.

❖ Shrivastav, A., Sharma, A. R., Bajaj, G., Charavaryamath, C., Ezzat, W., Spafford, P., Gore-Hickman, R., Singh, B., Copete, M. A. & Sharma, R. K. (2007). Elevated N-myristoyltransferase activity and expression in oral squamous cell carcinoma. *Oncol. Rep.* **18**, 93-97.

❖ Shtraizent, N., Eliyahu, E., Park, J. H., He, X., Shalgi, R. & Schuchman, E. H. (2008). Autoproteolytic cleavage and activation of human acid ceramidase. *J. Biol. Chem.* **283**, 11253-11259.

❖ Sindelar, C. V., Hendsch, Z. S. & Tidor, B. (1998). Effects of salt bridges on protein structure and design. *Protein Sci.* **7**, 1898-1914.

❖ Skrob, F., Becka, S., Plhackova, K., Fotopulosova, V. & Kyslik, P. (2003). Novel penicillin G acylase from *Achromobacter sp* CCM 4824. *Enz. Microb. Technol.* **32**, 738-744.

❖ Smith, J. L., Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H. & Satow, Y. (1994). Structure of the allosteric regulatory enzyme of purine biosynthesis. *Science,* **264**, 1427-1433.

❖ Sogabe, S., Masubuchi, M., Sakata, K., Fukami, T. A., Morikami, K., Shiratori, Y., Ebiike, H., Kawasaki, K., Aoki, Y., Shimma, N., D'Arcy, A., Winkler, F. K., Banner, D. W. & Ohtsuka, T. (2002). Crystal structures of *Candida albicans* N-myristoyltransferase with two distinct inhibitors. *Chem. Biol.* **9**, 1119-1128.

❖ Solorzano, C., Zhu, C., Battista, N., Astarita, G., Lodola, A., Rivara, S., Mor, M., Russo, R., Maccarrone, M., Antonietti, F., Duranti, A., Tontini, A., Cuzzocrea, S., Tarzia, G. & Piomelli, D. (2009). Selective N-acylethanolamine-hydrolyzing acid amidase inhibition reveals a key role for endogenous palmitoylethanolamide in inflammation. *Proc. Natl. Acad. Sci. USA,* **106**, 20966-20971.

❖ Sonawane, V. C. (2006). Enzymatic modifications of cephalosporins by cephalosporin acylase and other enzymes. *Crit. Rev. Biotechnol.* **26**, 95-120.

❖ Song, H. K., Bochtler, M., Azim, M. K., Hartmann, C., Huber, R. & Ramachandran, R. (2003). Isolation and characterization of the prokaryotic proteasome homolog HslVU (ClpQY) from *Thermotoga maritim*a and the crystal structure of HslV. *Biophys. Chem.* **100**, 437-452.

❖ Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S. & McKay, D. B. (2000). Crystal and solution structures of an HslUV protease–chaperone complex. *Cell,* **103**, 633-643.

❖ Stager, S., Smith, D. F. & Kaye, P. M. (2000). Immunization with a recombinant stage-regulated surface protein from *Leishmania donovani* induces protection against visceral leishmaniasis. *J. Immunol.* **165**, 7064-7071.

❖ Steenbergen, J. N. & Casadevall, A. (2000). Prevalence of *Cryptococcus neoformans var. neoformans* (serotype D) and *Cryptococcus neoformans var. grubii* (serotype A) isolates in New York city. *J. Clinical Microbiol.* **38**, 1974-1976.

❖ Stennicke, H. R. & Salvesen, G. S. (1998). Properties of the caspases. *Biochim. Biophys. Acta,* **1387**, 17-31.

❖ Sterner, R. & Liebl, W. (2001). Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* **36**, 39-106.

❖ Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **60**, 432-438.

❖ Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207-234.

❖ Stumpo, D. J., Graff, J. M., Albert, K. A., Greengard, P. & Blackshear, P. J. (1989). Molecular cloning, characterization, and expression of a cDNA encoding the "80- to 87-kDa" myristoylated alanine-rich C kinase substrate: a major cellular substrate for protein kinase C. *Proc. Natl. Acad. Sci. USA,* **86**, 4012-4016.

❖ Stura, E. A. & Wilson, I. A. (1990). Analytical and production seeding techniques. *Methods,* **1**, 38-49.

❖ Sudhakaran, V. K. & Borkar, P. S. (1985a). Microbial transformation of beta-lactam antibiotics: enzymes from bacteria, sources and study-a sum up. *Hind. Antibiot. Bull.* **27**, 63-119.

❖ Sudhakaran, V. K. & Borkar, P. S. (1985b). Phenoxymethyl penicillin acylases: sources and study-a sum up. *Hind. Antibiot. Bull.* **27**, 44-62.

❖ Sudhakaran, V. K. & Shewale, J. G. (1993). Enzymatic splitting of penicillin V for the production of 6-APA using immobilized penicillin V acylase. *World J. Microbiol. Biotechnol.* **9**, 630-634.

❖ Sudhakaran, V. K. & Shewale, J. G. (1995). Purification and characterization of extracellular penicillin V acylase from *Fusarium sp*. SKF 235. *Hind. Antibiot. Bull.* **37**, 9-15.

❖ Sukegawa, J., Semba, K., Yamanashi, Y., Nishizawa, M., Miyajima, N., Yamamoto, T. & Toyoshima, K. (1987). Characterization of cDNA clones for the human c-yes gene. *Mol. Cell Biol.* **7**, 41-47.

❖ Sultan, N. A. & Swamy, M. J. (2005). Fluorescence quenching and time-resolved fluorescence studies on *Trichosanthes dioica* seed lectin. *J. Photochem. Photobiol.* B**80**, 93-100.

❖ Summy, J. M. & Gallick, G. E. (2003). Src family kinases in tumor progression and metastasis. *Cancer Metastasis Rev.* **22**, 337-358.

❖ Sunder, A. V., Kumar, A., Naik, N. & Pundle, A. (2012). Characterization of a new *Bacillus cereus* ATUAVP1846 strain producing penicillin V acylase, and optimization of fermentation parameters. *Annals Microbiol.* **62**, 1287-1293.

❖ Suresh, C. G., Pundle, A. V., SivaRaman, H., Rao, K. N., Brannigan, J. A., McVey, C. E., Verma, C. S., Dauter, Z., Dodson, E. J. & Dodson, G. G. (1999). Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. *Nat. Struct. Biol.* **6**, 414-416.

❖ Sureshkumar, R. (2006). Investigation into the Structure and Activity of Conjugated Bile Salt Hydrolase and Related Penicillin Acylase. PhD. Thesis. Pune: University of Pune, Pune.

❖ Suzuki, Y., Oishi, K., Nakano, H. & Nagayama, T. (1987). A strong correlation between the increase in number of proline residues and the rise in thermostability of five Bacillus oligo-1,6-glucosidases. *Appl. Microbiol. Biotechnol.* **26**, 546-551.

❖ Svedas, V. K., Savchenko, M. V., Beltser, A. I. & Guranda, D. F. (1996). Enantioselective penicillin acylase-catalyzed reactions. Factors governing substrate and stereospecificity of the enzyme. *Annu. N.Y. Acad. Sci.*, **799**, 659-669.

❖ Svedas, V., Guranda, D., vanLangen, L., vanRantwijk, F. & Sheldon, R. (1997). Kinetic study of penicillin acylase from *Alcaligenes faecalis*. *FEBS Lett.* **417**, 414-418.

❖ Takamune, N., Gota, K., Misumi, S., Tanaka, K., Okinaka, S. & Shoji, S. (2008). HIV-1 production is specifically associated with human NMT1 long form in human NMT isozymes. *Microbes Infect.* **10**, 143-150.

❖ Teale, F. W. J. & Badley, R. A. (1970). Depolarization of intrinsic and extrinsic fluorescence of pepsinogen and pepsin. *Biochem. J.* **116**, 341-&.

❖ Teng, T. Y. (1990). Mounting of crystals for macromolecular crystallography in a free-standing thin film. *J. Appl. Crystallogr.* **23**, 387-391.

❖ Teplyakov, A., Obmolova, G., Badet, B. & Badet-Denisot, M.-A. (2001). Channeling of ammonia in glucosamine-6-phosphate synthase. *J. Mol. Biol.* **313**, 1093-1102.

❖ Terwilliger, T. (2003a). Automated side-chain model building and sequence assignment by template matching. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **59**, 45-49.

❖ Terwilliger, T. C. (2003b). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **59**, 38-44.

❖ Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J. & Adams, P. D. (2008). Iterative model building, structure refinement and density modification with the *PHENIX AutoBuild* wizard. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **64**, 61-69.

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

❖ Thony-Meyer, L., Bock, A. & Hennecke, H. (1992). Prokaryotic polyprotein precursors. *FEBS Lett.* **307**, 62-65.

❖ Topgi, R. S., Ng, J. S., Landis, B., Wang, P. & Behling, J. R. (1999). Use of enzyme penicillin acylase in selective amidation/amide hydrolysis to resolve ethyl 3-amino-4-pentynoate isomers. *Bioorg. Med. Chem.* **7**, 2221-2229.

❖ Torres-Bacete, J., Arroyo, M., Torres-Guzman, R., de la Mata, I., Castillon, M. P. & Acebal, C. (2000). Covalent immobilization of penicillin acylase from *Streptomyces lavendulae*. *Biotechnol. Lett.* **22**, 1011-1014.

❖ Torres-Guzman, R., de la Mata, I., Torres-Bacete, J., Arroyo, M., Castillon, M. P. & Acebal, C. (2002). Substrate specificity of penicillin acylase from *Streptomyces lavendulae*. *Biochem. Biophys. Res. Commun.* **291**, 593-597.

❖ Torres-Bacete, J., Hormigo, D., Stuart, M., Arroyo, M., Torres, P., Castillon, M. P., Acebal, C., Garcia, J. L. & de la Mata, I. (2007). Newly discovered penicillin acylase activity of Aculeacin A acylase from *Actinoplanes utahensis*. *Appl. Environ. Microbiol.* **73**, 5378-5381.

❖ Towler, D. A., Adams, S. P., Eubanks, S. R., Towery, D. S., Jackson-Machelski, E., Glaser, L. & Gordon, J. I. (1987). Purification and characterization of yeast myristoyl CoA:protein N-myristoyltransferase. *Proc. Natl. Acad. Sci. USA,* **84**, 2708-2712.

❖ Trumpower, B. L. (1990). Cytochrome bc1 complexes of microorganisms. *Microbiol. Rev.* **54**, 101-129.

❖ Tsuboi, K., Takezaki, N. & Ueda, N. (2007). The N-acylethanolamine-hydrolyzing acid amidase (NAAA). *Chem. Biodivers.* **4**, 1914-1925.

❖ Turunen, O., Vuorio, M., Fenel, F. & Leisola, M. (2002). Engineering of multiple arginines into the Ser/Thr surface of *Trichoderma reesei* endo-1,4-beta-xylanase II increases the thermotolerance and shifts the pH optimum towards alkaline pH. *Protein Eng.* **15**, 141-145.

❖ Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N. & Tsukihara, T. (2002). The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure,* **10**, 609-618.

❖ Uozumi, N., Sakurai, K., Sasaki, T., Takekawa, S., Yamagata, H., Tsukagoshi, N. & Udaka, S. (1989). A single gene directs synthesis of a precursor protein with beta- and alpha-amylase activities in *Bacillus polymyxa*. *J. Bacteriol.* **171**, 375-382.

❖ Vagin, A. & Teplyakov, A. (1997). *MOLREP*: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022-1025.

❖ Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *SFCHECK*: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **55**, 191-205.

❖ Valle, F., Balbas, P., Merino, E. & Bolivar, F. (1991). The role of penicillin amidases in nature and in industry. *Trends Biochem. Sci.* **16**, 36-40.

❖ van Langen, L. M., Oosthoek, N. H. P., Guranda, D. T., van Rantwijk, F., Svedas, V. K. & Sheldon, R. A. (2000). Penicillin acylase-catalyzed resolution of amines in aqueous organic solvents. *Tetrahedron: Asymm.* **11**, 4593-4600.

❖ Vandamme, E. J. & Voets, J. P. (1973). Some aspects of the penicillin V-acylase produced by *Rhodotorula glutinis var. glutinis*. *Z. Allg. Mikrobiol.* **13**; 701-710.

❖ Vandamme, E. J. & Voets, J. P. (1974). Microbial penicillin acylases. *Adv. Appl. Microbiol.* **17**, 311-369.

❖ Vandamme, E. J. & Voets, J. P. (1975). Properties of the purified penicillin V-acylase of *Erwinia aroideae*. *Experientia.* **31**, 140-143.

❖ Vandamme, E. J. (1988). Immobilized biocatalyst and antibiotic production. Biochemical, genetic, and biotechnical aspects. In: Moo-Young, M. (ed.) *Bioreactor, Immobilized Enzymes and Cells: Fundamentals and Applications*. Marcel Dekker, New York, p261-286.

❖ van den Heuvel, R. H. H., Ferrari, D., Bossi, R. T., Ravasio, S., Curti, B., Vanoni, M. A., Florencio, F. J. & Mattevi, A. (2002). Structural studies on the synchronization of catalytic centers in glutamate synthase. *J. Biol. Chem.* **277**, 24579-24583.

❖ Vandijl, J. M., Dejong, A., Venema, G. & Bron, S. (1995). Identification of the potential active site of the signal peptidase SipS of *Bacillus subtilis*. Structural and functional similarities with LexA-like proteases. *J. Biol. Chem.* **270**, 3611-3618.

❖ Varshney, N. K., Kumar, R. S., Ignatova, Z., Prabhune, A., Pundle, A., Dodson, E. & Suresh, C. G. (2012). Crystallization and X-ray structure analysis of a thermostable penicillin G

acylase from *Alcaligenes faecalis. Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* **68**, 273-277.

❖ Verhaert, R. M. D., Riemens, A. M., vanderLaan, J. M., vanDuin, J. & Quax, W. J. (1997). Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from *Alcaligenes faecalis. Appl. Environ. Biotechnol.* **63**, 3412-3418.

❖ Veronese, F. D., Copeland, T. D., Oroszlan, S., Gallo, R. C. & Sarngadharan, M. G. (1988). Biochemical and immunological analysis of human immunodeficiency virus gag gene products p17 and p24. *J. Virol.* **62**, 795-801.

❖ Verweij, J. & de Vroom, E. (1993). Industrial transformations of penicillins and cephalosporins. *Recueil des Travaux Chimiques des Pays-Bas.* **112**, 66-81.

❖ Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H. & Robb, F. T. (1998). Protein thermostability above 100°C: A key role for ionic interactions. *Proc. Natl. Acad. Sci.USA,* **95**, 12300-12305.

❖ Vieille, C. & Zeikus, J. G. (1996). Thermozymes: Identifying molecular determinants of protein structural and functional stability. *Trends in Biotechnol.* **14**, 183-190.

❖ Vihinen, M. (1987). Relationship of protein flexibility to thermostability. *Protein Eng.* **1**, 477-480.

❖ Vriend, G. (1990). *WHAT IF* - a Molecular modeling and drug design program. *J. Mol. Graphics,* **8**, 52-&.

❖ Wada, K., Irie, M., Suzuki, H. & Fukuyama, K. (2010). Crystal structure of the halotolerant γ-glutamyltranspeptidase from *Bacillus subtili*s in complex with glutamate reveals a unique architecture of the solvent-exposed catalytic pocket. *FEBS J.* **277**, 1000-1009.

❖ Waksman, S. A. (1947). What is an antibiotic or an antibiotic substance. *Mycologia,* **39**, 565-569.

❖ Wang, W.-C., Hsu, W.-H., Chien, F.-T. & Chen, C.-Y. (2001). Crystal structure and site-directed mutagenesis studies of N-carbamoyl-d-amino-acid amidohydrolase from *Agrobacterium radiobacter* reveals a homotetramer and insight into a catalytic cleft. *J. Mol. Biol.* **306**, 251-261.

❖ Wang, Y. & Guo, H. C. (2010). Crystallographic snapshot of glycosylasparaginase precursor poised for autoprocessing. *J. Mol. Biol.* **403**, 120-130.

❖ Wang, J., Rho, S.-H., Park, H. H. & Eom, S. H. (2005). Correction of X-ray intensities from an HslV-HslU co-crystal containing lattice-translocation defects. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **61**, 932-941.

❖ Weaver, S. S. & Bodey, G. P. (1980). CI-867, a new semisynthetic penicillin: in vitro studies. *Antimicrob. Agents Chemother.* **18**, 939-943.

❖ Weber, P. C. (1991). Physical principles of protein crystallization. *Adv. Protein Chem.* **41**, 1-36.

❖ Weinberg, R. A., McWherter, C. A., Freeman, S. K., Wood, D. C., Gordon, J. I. & Lee, S. C. (1995). Genetic studies reveal that myristoyl CoA:protein N-myristoyltransferase is an essential enzyme in *Candida albicans*. *Mol. Microbiol.* **16**, 241-250.

❖ Weiss, M. S. (2001). Global indicators of X-ray data quality. *J. Appl. Crystallogr.* **34**, 130-135.

❖ Weiss, M. S. & Hilgenfeld, R. (1997). On the use of the merging R factor as a quality indicator for X-ray data. *J. Appl. Crystallogr.* **30**, 203-205.

❖ Welker, R., Harris, M., Cardel, B. & Krausslich, H. G. (1998). Virion incorporation of human immunodeficiency virus type 1 Nef is mediated by a bipartite membrane-targeting signal: Analysis of its role in enhancement of viral infectivity. *J. Virol.* **72**, 8833-8840.

❖ Wen, Y., Feng, M. Q., Yuan, Z. Y. & Zhou, P. (2005). Expression and overproduction of recombinant penicillin G acylase from *Kluyvera citrophila* in *Escherichia coli*. *Enz. Microb. Technoly.* **37**, 233-237.

❖ Wen, Y., Shi, X. L., Yuan, Z. Y. & Zhou, P. (2004). Expression, purification, and characterization of His-tagged penicillin G acylase from *Kluyvera citrophila* in *Escherichia coli*. *Protein Expr. Purif.* **38**, 24-28.

❖ Weston, S. A., Camble, R., Colls, J., Rosenbrock, G., Taylor, I., Egerton, M., Tucker, A. D., Tunnicliffe, A., Mistry, A., Mancia, F., de la Fortelle, E., Irwin, J., Bricogne, G. & Pauptit, R. A. (1998). Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat. Struct. Biol.* **5**, 213-221.

❖ Wiegand, R. C., Carr, C., Minnerly, J. C., Pauley, A. M., Carron, C. P., Langner, C. A., Duronio, R. J. & Gordon, J. I. (1992). The *Candida albicans* myristoyl-CoA:protein N-myristoyltransferase gene. Isolation and expression in *Saccharomyces cerevisiae* and *Escherichia coli*. *J. Biol. Chem.* **267**, 8591-8598.

❖ Wilcox, C., Hu, J. S. & Olson, E. N. (1987). Acylation of proteins with myristic acid occurs cotranslationally. *Science,* **238**, 1275-1278.

❖ Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011).

Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect.D: Biol. Crystallogr.* **67**, 235-242.

❖ Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186-190.

❖ Wright, M. H., Heal, W. P., Mann, D. J. & Tate, E. W. (2010). Protein myristoylation in health and disease. *J. Chem. Biol.* **3**, 19-35.

❖ Wu, R., Richter, S., Zhang, R.-g., Anderson, V. J., Missiakas, D. & Joachimiak, A. (2009). Crystal structure of *Bacillus anthracis* transpeptidase enzyme CapD. *J. Biol. Chem.* **284**, 24406-24414.

❖ Wu, J., Tao, Y., Zhang, M., Howard, M. H., Gutteridge, S. & Ding, J. (2007). Crystal structures of *Saccharomyces cerevisiae* N-myristoyltransferase with bound myristoyl-CoA and inhibitors reveal the functional roles of the N-terminal region. *J. Biol. Chem.* **282**, 22185-22194.

❖ Xu, Q. A., Buckley, D., Guan, C. D. & Guo, H. C. (1999). Structural insights into the mechanism of intramolecular proteolysis. *Cell,* **98**, 651-661.

❖ Xuan, J., Tarentino, A. L., Grimwood, B. G., Plummer, T. H., Roey, P. V., Cui, T. & Guan, C. (1998). Crystal structure of glycosylasparaginase from *Flavobacterium meningosepticum. Protein Sci.* **7**, 774-781.

❖ Yamazaki, K., Kaneko, Y., Suwa, K., Ebara, S., Nakazawa, K. & Yasuno, K. (2005). Synthesis of potent and selective inhibitors of *Candida albicans* N-myristoyltransferase based on the benzothiazole structure. *Bioorg. Med. Chem.* **13**, 2509-2522.

❖ Yang, S. H., Shrivastav, A., Kosinski, C., Sharma, R. K., Chen, M. H., Berthiaume, L. G., Peters, L. L., Chuang, P. T., Young, S. G. & Bergo, M. O. (2005). N-myristoyltransferase 1 is essential in early mouse development. *J. Biol. Chem.* **280**, 18990-18995.

❖ Yang, Y. L., Yun, D. F., Guan, Y. Q., Peng, H. L., Chen, J. M., He, Y. S. & Jiao, R. C. (1991). Cloning of GL-7-ACA acylase gene from *Pseudomonas* sp. 130 and its expression in *Escherichia coli. Chin. J. Biotechnol.* **7,** 93-104.

❖ Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. & Rice, D. W. (1995). The Structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for Ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure,* **3**, 1147-1158.

❖ Yoon, J., Oh, B., Kim, K., Park, J., Han, D., Kim, K. K., Cha, S.-S., Lee, D. & Kim, Y. (2004). A bound water molecule is crucial in initiating autocatalytic precursor activation in an N-terminal hydrolase. *J. Biol. Chem.* **279**, 341-347.

❖ Yuryev, R., Kasche, V., Ignatova, Z. & Galunsky, B. (2010). Improved *A. faecalis* penicillin amidase mutant retains the thermodynamic and pH stability of the wild type enzyme. *Prot. J.* **29**, 181-187.

❖ Zeidan, Y. H., Jenkins, R. W., Korman, J. B., Liu, X., Obeid, L. M., Norris, J. S. & Hannun, Y. A. (2008). Molecular targeting of acid ceramidase: implications to cancer therapy. *Curr. Drug Targets,* **9**, 653-661.

❖ Zhang, F., Hu, M., Tian, G., Zhang, P., Finley, D., Jeffrey, P. D. & Shi, Y. (2009). Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Mol. Cell,* **34**, 473-484.

❖ Zhang, L., Jackson-Machelski, E. & Gordon, J. I. (1996). Biochemical studies of *Saccharomyces cerevisiae* myristoyl-coenzyme A:protein N-myristoyltransferase mutants. *J. Biol. Chem.* **271**, 33131-33140.

❖ Zhou, W. J., Parent, L. J., Wills, J. W. & Resh, M. D. (1994). Identification of a membrane-binding domain within the amino-terminal region of human immunodeficiency virus type 1 Gag protein which interacts with acidic phospholipids. *J. Virol.* **68**, 2556-2569.

❖ Zhu, G. P., Xu, C., Teng, M. K., Tao, L. M., Zhu, X. Y., Wu, C. J., Hang, J., Niu, L. W. & Wang, Y. Z. (1999). Increasing the thermostability of D-xylose isomerase by introduction of a proline into the turn of a random coil. *Protein Eng.* **12**, 635-638.

❖ Zmijewski, M. J., Briggs, B. S., Thompson, A. R. & Wright, I. G. (1991). Enantioselective acylation of a Beta-Lactam intermediate in the synthesis of Loracarbef using penicillin G amidase. *Tetrahedron Lett.* **32**, 1621-1622.

❖ Zulauf, M. & D'Arcy, A. (1992). Light scattering of proteins as a criterion for crystallization. *J. Cryst. Growth,* **122**, 102-106.

Curriculum vitae

NISHANT KUMAR VARSHNEY

C/O Dr. C. G. Suresh, Biochemical Sciences Division,

CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune-411008, Maharashtra, India. **+91-20-25902711, +91-8390564690**
nk.varshney@ncl.res.in, nikdna@gmail.com

PROFILE:

- Highly self motivated Ph.D. candidate with demonstrated research experience in the field of protein purification and structural biology

EDUCATION:

2006- 2013	**Doctor of Philosophy (Ph.D.) Biotechnology**
	Thesis Submitted to University of Pune, India
	CSIR-National Chemical Laboratory, Pune, India
2003-2005 (63.5%)	**Master of Science (M.Sc.) in Marine Biotechnology**
	Department of Biotechnology, Goa University, Goa, India
2000-2003 (67.6%)	**Bachelor of Science (B.Sc.) in Chemistry (Honors)**
	University of Delhi, Delhi, India

RESEARCH EXPERIENCE:

- **CSIR-National Chemical Laboratory, Pune, India**,

Ph.D. Candidate under the supervision of Dr. C. G. Suresh, **2006- 2013 (Thesis Submitted)**

"Structural investigations of selected microbial penicillin G acylases and a protozoal N-myristoyltransferase"

It is widely accepted that the amino acid sequence of a protein determines its three-dimensional structure and that structure is important for function. Understanding structure-function relationships may also unravel the mechanism of action and post-translational processing of proteins, which in turn facilitate protein-engineering to enhance the potential for chemical, industrial and specific therapeutic applications. With this objective, biochemical and biophysical characterization of three enzymes, two of them classified as Penicillin G acylases and the third one a N-myristoyltransferase from a protozoan

Leishmania donovani, has been carried out. Thesis provides an account of the study carried out on these three pharmaceutically and medicinally important enzymes, whose action concerns the amide bond. Main objective in carrying out these investigations was to understand the structure-function relation of these enzymes that can facilitate their applications. The enzymes were cloned, expressed and their structures were studied using various biochemical and biophysical techniques. Thesis summarizes how the above studies have provided insights into several chemical, physical and activity characteristics of the enzymes including three-dimensional structures that are important for improving their application potential.

- **Department of Biotechnology, Goa University, Goa, India**

M.Sc Dissertation Project under the guidance of Dr. Savita Kerkar**,** Lecturer, **July 2004-March 2005**

Population distribution in a Microbial Biofilm from Ribandar saltpan of Goa and studies on Sulphate Reducing Bacteria (SRB) localized to oxic zone of the mat

The main aim was to study the seasonal change on the biodiversity and the sulphate reduction and sulphide oxidation activities in the microbial mats present in hyper saline ponds of Goa and to study the presence of certain enzymes responsible for the oxygen tolerance in the selected strains of SRB growing in aerobic condition isolated from the same hyper saline pond.

- **Bioresources and Biotechnology Division, TERI, New Delhi, India**

Summer training Project under the guidance of Dr. Banwari Lal**,** Associate Director, **6th May - 14th July 2004.**

Enrichment and Isolation of Bacteria obtained from oil reservoir situated at Assam, India

The main aim was to isolate different bacteria present in the oil formation water and to characterize all the parameters for the present bacteria.

SKILLS/ TECHNIQUES KNOWN:

Proteins Crystallography: Crystallization of protein using different methods like Hanging Drop, Sitting Drop and Microbatch, Crystal mounting and Data collection of protein crystals using home source (Rigaku, Raxis IV^{++}) as well as at synchrotron, Data reduction and scaling in DENZO, SCALEPACK and MOSFLM, Use of CCP4 and PHENIX softwares for determining protein structure

Biochemical Techniques: Colorimetric and spectroscopic (UV, Visible) techniques, Native and SDS-Poly Acrylamide Gel Electrophoresis(PAGE), Purification of Proteins manually as well as on Automated purification system AKTA Explorer, Western Blot

Analysis, Fluorescence Spectroscopy, MALDI –TOF and other routine biochemical analysis

Molecular Biological Techniques: Isolation and estimation of DNA and RNA from bacteria, cloning by restriction ligation and ligation free methods, over-expression of gene of interest in bacteria and Yeast shake flask cultures

Bioinformatics: Protein sequence analysis,   PyMOL, ccp4mg,

Computer Skills: Linux/Windows, Microsoft Office, Internet etc.

PUBLICATIONS/PATENT:

❖ **Nishant Kumar Varshney**, R. Suresh Kumar, Zoya Ignatova, Asmita Prabhune, Archana Pundle, Eleanor Dodson and C. G. Suresh (2012).Crystallization and X-ray structure analysis of a thermostable penicillin G acylase from *Alcaligenes faecalis*, *Acta Cryst.* **F68**, 273–277.

❖ Sreejith, R.K., Yadav, V.N., **Varshney, N.K**., Berwal, S.K., Suresh, C.G., Gaikwad, S.M., Pal, J.K. (2009) Conformational characterization of human eukaryotic initiation factor 2alpha: a single tryptophan protein. *Biochemical and Biophysical Research Communications*, **390**, 273-279.

❖ **Nishant Kumar Varshney,** Sureshkumar Ramasamy, Jim A. Brannigan, Anthony J. Wilkinson, C. G. Suresh. Overexpression, crystallization and preliminary X-ray crystallographic analysis of slow-processing mutant of Penicillin G acylase from *Kluyvera citrophila* (ready to communicate).

❖ Priyabrata R Panigrahi, **Nishant Kumar Varshney**, C.G. Suresh. Application of molecular modelling and information to select genes of thermostable penicillin G acylases (ready to communicate).

❖ **Nishant Kumar Varshney**, Jim Brannigan, Eleanor Dodson, C. G. Suresh, Anthony J. Wilkinson. Cloning, Purification, Crystallization and comparative structural analysis of N-terminal mutant of N-myristoyltransferase from *Leishmania donovani* (manuscript under preparation).

❖ **Nishant Kumar Varshney** and C. G. Suresh. Structural and folding studies of wild-type enzyme and slow-processing mutant of *Kc*PGA (manuscript under preparation).

❖ Asmita, A. Prabhune, C. G. Suresh, **Nishant Kumar Varshney,** Priyabrata Panigrahi and Ruchira Mukherji. Quorum Quenching by *Kc*PGA, Patent Ref no: INV-2013-0004, Provisional filing number : 1232-DEL-2013

WORKSHOPS ATTENDED /POSTERS PRESENTED:

• **Crystallization and structure determination of thermostable penicillin G acylase from** *Alcaligenes faecalis* **poster** presented on December 9th , 2006 in Society of Biological Chemists (India) and 75th Annual Meeting held in Jawaharlal Nehru University, New Delhi, India,

- **Crystallographic analysis of penicillin G acylase from *Alcaligenes faecalis* and study of calcium influence in the post-translational processing of PGA from *Kluyvera citrophila* poster** presented in an International conference EMBO-WLC on "Recent Developments in Macromolecular Crystallography" held in NCL, Pune, India from 9-14th Nov. 2008,

- Attended **CCP4 workshop** held in Indian Institute of Science, Bangalore from 17th Feb. 2008 to 22nd Feb 2008 organized jointly by Indian Institute of Science, Bangalore, India and York Structural Biology Lab, University of York, York, UK,

- **Structural Analysis of thermostable Penicillin G Acylase from *Alcaligenes faecalis*. Oral presentation** at 38th National Seminar on Crystallography held from 11th to 13th Feb. 2009 in University of Mysore, Mysore, India,

- **Structural analysis of thermostable penicillin G acylase from *Alcaligenes faecalis* poster** presented in Science Day celebration held at NCL, Pune, February, 2009,

- **Crystal Structure of a Thermostable Penicillin G acylase from *Alcaligenes faecalis* – Structural Interpretation of Its Higher Stability poster** presented in Science Day celebration held at NCL, Pune, India February, 2010,

- **Structure of N-Myristoyltransferase from *Leishmania donovani* (LdNMT) in two crystal forms with bound myristoyl CoA analogue poster** presented in Science Day celebration held at NCL, Pune, India, during 27-28th February, 2012,

- Attended **18th CCP4 Northern Protein Structure Workshop** held at University of Cumbria, Carlisle, UK.

ACHIEVEMENTS AND HONOURS:

- Awarded Commonwealth Split-Site Scholarship to visit University of York, York, UK for one year during Ph.D. tenure (Feb,11-Jan,12),

- Awarded Junior Research Fellowship and Lecturership in Life Science in Dec 2005 by University Grant Commission and Council of Scientific & Industrial Research, Human Resource and Development Group, Government of India, New Delhi, India,

- Awarded DBT- Junior Research Fellowship (DBT-JRF) (2005) exam conducted by Department of Biotechnology, Government of India ,

- Awarded ICMR- Junior Research Fellowship (ICMR-JRF)(2005) exam conducted by Indian Council For Medical Research, New Delhi ,

- Stood 1st in all the three years of graduation in the college,

- Selected for M.Sc. Marine Biotechnology through All India Combined Entrance Examination conducted by Jawaharlal Nehru University, New Delhi, India,

- Selected for M.Sc. Environmental Sciences Course in Jawaharlal Nehru University, New Delhi through All India Combined Entrance Examination conducted by Jawaharlal Nehru University, New Delhi, India,

- Selected for M.Sc. Chemistry in Indian Institute of Technology, Roorkee, India,

- Awarded Scholarship given by Department of Biotechnology, Government of India in two years of M.Sc. (2003-2005)

WORK STYLE: - Willing to perform basic tasks and move on to solve complex problems

- Capable of learning new knowledge and adapt to new environment quickly

- Independent work style and excellent teamwork skills

- Organized, Enthusiastic and passionate about work

References:

1) Dr. C. G. Suresh,

Senior Principal Scientist,

Biochemical Sciences Division,

National Chemical Laboratory,

Dr. Homi Bhabha Road,

Pune-411008, Maharashtra, India.

Ph: - +91-20-25902236,

Fax: - +91-2025902648

Email: - cg.suresh@ncl.res.in

2) Prof. Anthony J. Wilkinson

York Structural Biology Laboratory,

Department of Chemistry,

University of York, York, YO10 5DD, UK

Ph: - +44(0)1904 328261,

Fax: - +44(0)1904 322516

Email: - tony.wilkinson@york.ac.uk

3) Dr. Sureshkumar Ramasamy

Scientist**,** Biochemical Sciences Division,

National Chemical Laboratory,

Dr. Homi Bhabha Road,

Pune-411008, Maharashtra, India.

Ph: - +91-7350850232

Email: - s.ramasamy@ncl.res.in