Structural and Biochemical Investigations of Novel Bacterial Choloylglycine Hydrolases

Thesis Submitted to AcSIR
For the Award of the Degree of
DOCTOR OF PHILOSOPHY
In
BIOLOGICAL SCIENCES

By
Philem Pushparani Devi
10BB11J26127

Under the guidance of
Dr. Dhanasekaran Shanmugam
(Research Supervisor)
Dr. Asmita Prabhune
(Research Co-Supervisor)

BIOCHEMICAL SCIENCES DIVISION
CSIR-NATIONAL CHEMICAL LABORATORY
PUNE-411008, INDIA
MARCH 2017

राष्ट्रीय रासायनिक प्रयोगशाला

(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)

डॉ. होमी भाभा मार्ग पुणे – 411 008. भारत

NATIONAL CHEMICAL LABORATORY

(Council of Scientific & Industrial Research)
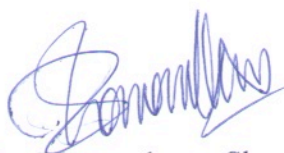
Dr. Homi Bhabha Road, Pune - 411 008. India.

NCL

CERTIFICATE

This is to certify that the work incorporated in this Ph.D. thesis entitled **"Structural and Biochemical Investigations of Novel Bacterial Choloylglycine Hydrolases "**submitted by Ms. **Philem Pushparani Devi** to Academy of Scientific and innovative Research (AcSIR) in fulfilment of the requirements for the award of the Degree of Doctor of Philosophy embodies original research work under our supervision/guidance. We further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text or illustration, table etc., used in the thesis from other sources, have been duly cited and acknowledged.

Philem Pushparani Devi

(Research Student)

Dr. Dhanasekaran Shanmugam

(Research Supervisor)

Dr. Asmita Prabhune

(Co-supervisor)

Date: 29th March, 2017

Place: Pune

DECLARATION BY RESEARCH SCHOLAR

I hereby declare that the thesis entitled "**Structural and Biochemical Investigations of Novel Bacterial Choloylglycine Hydrolases**" submitted by me for the Degree ofDoctor of Philosophy to Academy of Scientific & Innovative Research (AcSIR) is the record of workcarried out by me at Biochemical Sciences Division, CSIR- National Chemical Laboratory, Pune - 411008, India, under the supervision of Dr. Dhanasekaran Shanmugam (research guide) and Dr. Asmita Prabhune (Co-supervisor). The work is original and has not formed the basis for the award of any degree, diploma, associateship, and fellowship titles in this or any other university or other institute ofhigher learning. I further declare that the material obtained from other resources hasbeen duly acknowledged in the thesis.

Philem Pushparani Devi   **Dated: 29.3.2017**

Biochemical Sciences Division
CSIR-National Chemical Laboratory
Dr.HomiBhabha Road, Pashan
Pune – 411008
Maharashtra, India

Dedicated to

The way back home...

Table of Contents

Chapter 1: Introduction

Chapter 2: Cloning and characterization a choloylglycine hydrolase, *Sl*CGH1 from *Shewanella loihica* PV-4

Chapter 3: Structural features and mutational analysis of *Sl*CGH1

Chapter 4 : Characterization, molecular modeling of *Sl*CGH2 and phylogenetic analysis of *Shewanella loihica* PV-4 CGHs

Chapter 5: Native Penicillin V acylase of *Acinetobacter* sp. AP24, isolated from Loktak Lake - Purification and partial characterization

Acknowledgement

It has been an incredible and a truly transforming journey and I take a great pleasure to express my gratitude to everyone who has made this thesis possible. First and foremost thanks must go to my research supervisor, Dr. Asmita Prabhune for being my guru on so many levels. I am eternally grateful to her for giving me scientific insights, generous support at moments of great doubts and being so compassionate towards my needs. You have given me immense freedom to try out myriads of ideas and without you, this thesis would be lot less interesting.

I considered myself lucky for getting the chance to work with Dr. Sureshkumar Ramasamy. I am thankful to him for introducing me to the world of protein crystallography. You have not only supervise my work patiently but showered on me a great deal of trust and support which I will always remember. I will always carry a deep sense of gratitude for Dr. Dhanasekaran Shanmugam for his timely support, his invaluable suggestions and insights.

I am also grateful to my DAC members, Dr. Archana Pundle, Dr. Sushama Gaikwad, Dr. Pankaj Poddar and Dr. Jayant Khire, for periodically monitoring my work progress with genuine support and kindness. I would like to acknowledge Dr. Sreekumar Kurungot for allowing me to use his lab facility for my CSIR-800 project and Dr. Moneesha Fernandes (NCL) for fluorescence and CD measurements.

I owe a great deal to Deepak, Manu, Ameya and Deepanjan for patiently teaching me many techniques related to protein science as well as for carrying out X-ray diffraction experiments of my protein crystals. Special thanks to Yashpal, my partner in crime, for his valuable contributions in solving the protein structures, making of structure figures and for being such a fun company to work with. Heartfelt thanks to Isha for being such a strong support and for her contribution to the work in chapter 4.

List of abbreviations

6-APA	6-amino penicillanic acid
AHL	*N*-acylhomoserine lactone
Å	Angstrom unit
ATCC	American Type culture collection
*Bl*BSH	*Bifidobacterium longum* bile salt hydrolase
*Bt*BSH	*Bacteroides thetaiotaomicron* bile salt hydrolase
βME	β-mercapto ethanol
BSH	Bile salt hydrolase
*Bsp*PVA	*Bacillus sphaericus* penicillin V acylase
*Bsu*PVA	*Bacillus subtilis* penicillin V acylase
*Bt*BSH	*Bacteroides thetaiotamicron* bile salt hydrolase
CCD	Charge coupled device
CGH	Cholylglycine hydrolase
CPB	Citrate phosphate buffer
*Cp*BSH	*Clostridium perfringens* bile salt hydrolase
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
*Ec*PGA	*Escherichia coli* penicillin G acylase
ESRF	European Synchrotron Radiation Facility, France
EDTA	Ethylene diamine tetra-acetic acid
F(hkl)	Structure factor
GDCA	Glycocholic acid
HPLC	High pressure liquid chromatography
HSL	Homoserine lactone
IPTG	Isopropyl-β-D-thiogalactoside
*Kc*PGA	*Kluyvera citrophila* penicillin G acylase
k_{cat}	Turnover number
KDa	kilo Dalton
MALDI	Matrix-associated Laser Desorption and Ionization
MPD	2-methyl-2, 4-pentanediol

μg	Microgram
μL	Microlitre
μM	Micromolar
NCS	Non-crystallographic symmetry
Ntn	N-terminal nucleophile
*Pa*PVA	*Pectobacterium atrosepticum* penicillin V acylase
PAGE	Poly-acrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pDAB	para-dimethyl amino benzaldehyde
PEG	Polyethylene glycol
Pen G	Penicillin G
Pen V	Penicillin V
PGA	Penicillin G acylase
pI	Iso-electric pH
PVA	Penicillin V acylase
RMSD	Root mean square deviation
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
*Sl*CGH1	*Shewanella loihica* PV-4 first choloylglycine hydrolase
*Sl*CGH2	*Shewanella loihica* PV-4 second choloylglycine hydrolase
*Sl*PVA	*Streptomyces lavendulae* penicillin V acylase
*Sm*PVA	*Streptomyces mobarensis* penicillin V acylase
TDCA	Taurodeoxycholic acid
TCA	Taurocholic acid
Tris	Tris-hydroxymethyl amino methane
V_m	Matthews number
V_{max}	Maximum velocity

Abstract

The choloylglycine hydrolase (CGH) group of the Ntn hydrolase enzyme superfamily encompasses penicillin V acylase (PVA, E.C. 3.5.1.11) and bile hydrolase (BSH, EC 3.5.1.24). PVA is responsible for 15% of the annual production of 6-aminopenicillanic acid (6-APA), a key intermediate in manufacture of semi synthetic β-lactam antibiotics and BSH is the major player in probiotic preparations as the bile salt deconjugating enzyme. Both these enzymes exhibit αββα fold in their tertiary structure characteristic of Ntn hydrolases, housing N-terminal nucleophilic cysteine as the main catalytic residue. They cleave amide bonds present in their respective substrates and undergo autocatalytic post translational processing to give rise to the active enzyme.

Although they are industrially popular enzymes, the concrete role of CGHs in microbial physiology is not known to date. Interestingly, there is an emerging association of this group of enzymes recently with virulence phenomena. Adding onto this anonymity is the question of a widespread occurrence of uncharacterized CGH homologues than expected before, in single and multiple copies in genome sequences of different taxa spanning from bacteria, archaea to eukaryotes. Such a widespread evolutionary conservation of these genes suggests an important role of this enzyme family in the lifestyle of the host organism. All these findings stir new research interest towards exploring more appropriate substrates/activities to represent these enzymes and ultimately helping to decipher the *in vivo* roles.

The work described in the thesis revolves round the structural and biochemical characterization of two novel marine CGHs *Sl*CGH1 and *Sl*CGH2, from *Shewanella loihica* PV-4 that exhibit unique substrate specificity and might represent a new sub-class of CGH enzymes. Biochemical characterization of a PVA from fresh water *Acinetobacter* sp. AP24, isolated from Loktak Lake, was also performed. The thesis is organized into five chapters.

The *first chapter* is the general introduction to the thesis. It provides a brief review on literature on enzymes in general, their evolution and mechanistic aspects. The chapter also deals with the Ntn hydrolase super family members with special reference to CGHs, their structural features, catalysis behaviour and mechanism. The phenomenon of *N*-acylhomoserine lactone (AHL) based

quorum sensing (QS) and quenching (QQ) are also discussed as CGHs have been linked to AHL-regulated virulence and pathogenesis in Gram-negative bacteria.

The **second chapter** deals with the cloning of gene coding for *Sl*CGH1 from the Gram-negative marine bacterium *Shewanella loihica* PV-4. Expression of the enzyme in *E. coli*, purification and biochemical characterization has been discussed. Besides PenV and bile salts, other substrates including AHLs, casein and p-nitrophenylpalmitate were tested for possible AHL acylase, protease and lipase activities. *Sl*CGH1 was inactive on both CGH substrates, unlike previously reported functional PVA/BSHs. In addition, the enzyme showed good AHL acylase activity. Site-directed mutagenesis validated the role of N-terminal cysteine in the catalytic mechanism. Fluorescence and CD spectroscopy based stability analysis of the enzyme are elaborated.

The **third chapter** deals with the crystallization of *Sl*CGH1 and elucidation of its structural characteristics. The methodology used for crystallization and X-ray diffraction, and refinement of the structural parameters has been discussed. The structure of *Sl*CGH1 was compared with different CGHs, and AHL acylase to explain deviation in oligomeric interactions and active site organization accountable for the unique activity profile of *Sl*CGH1.

The **fourth chapter** describes the cloning and characterization of the second CGH, *Sl*CGH2, from *S. loihica* PV-4. *Sl*CGH2 exhibited a similar activity profile as *Sl*CGH1 with no activity on Pen V or bile salts; however, it could hydrolyze AHLs with longer acyl chains compared to *Sl*CGH1. Since attempts to crystallize *Sl*CGH2 were largely unsuccessful, the probable structure was constructed by homology modeling and structural characteristics analyzed. The ability of *Sl*CGH1 and *Sl*CGH2 to hydrolyze long chain AHLs involved in bacterial quorum sensing, and its significance in microbial physiology has also been explored. Moreover, detailed phylogenetic analysis of bacterial CGH sequences unraveled the segregation of *S. loihica* CGHs to a separate cluster distant from those previously described for their counterparts from Gram-positive and Gram-negative bacteria. Most marine CGH homologs also showed good similarity and conservation of critical residues in their binding site to *Sl*CGH1, confirming their unique nature and the probable existence of a separate CGH sub-class.

The **fifth chapter** describes isolation, identification of penicillin V acylase producing *Acinetobacter* sp. AP24 and partial characterization of the PVA. The substrate specificity was

explored and the enzyme was found highly specific for PenV. Stability of the enzyme in different pH and temperature has been studied. Optimum conditions for enzyme activity are also described.

A brief *summary* of the thesis work and scope for possible further research in these areas is presented at the end of the thesis.

INTRODUCTION

Structural and Biochemical Investigations of Novel Bacterial Choloylglycine Hydrolases

"The only way to make sense out of change is to plunge into it, move with it, and join the dance."

- Alan Watts

1.1. Enzyme-lore:

In 1897, a German chemist Eduard Buchner accidentally discovered that cell-free yeast extract could still convert sugar to ethanol and carbon dioxide, while working on a trial experiment for obtaining immunologically active proteins in his brother's lab (Buchner 1897). Little did the Buchner brothers knew that their observation had paved the way for the science of 'Biochemistry' to emerge. In 1907, Eduard Buchner received the Nobel Prize in Chemistry for the discovery of the first free enzyme which he named 'zymase'.

Although enzymes have been used for a long time in crude forms in brewing, baking processes etc, the mechanism of such conversions and importance of enzymes became a significant topic of research from the 19th century. Before Buchner, many eminent scientists such as Pasteur, Schwann, Liebig and Kuhne by virtue of their investigations on fermentation process, also contributed to this field. William Kuhne coined the term 'enzyme' in 1876 (Greek word 'zyme' meaning 'in yeast'). Around this time, the kinetic study of enzyme catalysis also gained popularity with the derivation of the MM equation (Michaelis and Menten 1913) which relates the rate of a reaction to its catalytic rate constant (k_{cat}) and the substrate concentration K_m, which reflects the affinity between enzyme and substrate.

The recognition of enzymatic activity as responsible for cell-free fermentation was controversial and provided the first glimpse of the possibility of life reaction outside living cell. Slowly scientists started to incorporate enzyme into their discussion to explain how life originated on earth. Eventually enzyme based theory gained popularity over the reigning theory of origin of life based on protoplasmic biology. Still the biochemical nature of enzyme was a mystery until James Sumner in 1926 demonstrated that enzymes are proteins by purifying and crystallizing the enzyme urease, followed by supporting discoveries of John Northrop and Wendell Stanley for isolating pure enzyme and viral protein. These three shared the 1946 Nobel Prize in Chemistry. The property of protein to form crystals led to the beginning of protein X-ray crystallography or structural biology and the glorious era of understanding enzyme mechanisms at the molecular level.

1.2. Enzyme catalysis:

The rate of a chemical reaction depends on the activation energy or the difference between the free energy between the substrate and the transition state, such that higher the activation energy, slower the rate of a reaction. Like all catalysts, enzymes increase the rate of biochemical reactions by lowering the activation energy and stabilizing the transition state (Fersht 1997). However, enzymes are more dynamic and complex than chemical catalysts. Unique to enzymes, during catalysis, one or more transition states are formed and the reaction could be divided into several smaller steps. They can achieve accelerations of rate to a factor of 10^7 - 10^{19} (Wolfenden and Snider 2001). Mostly globular protein and sometimes RNA in nature, enzymes first require binding to their substrate before catalysis, unlike their inorganic counterparts, giving enzymes the property of specificity towards their substrates. Moreover, upon binding to substrate, the enzyme undergoes small to large domain motions depending on different enzyme types (Gutteridge and Thornton 2004). Understanding the relation between protein dynamics and function still has to go a long way. New age enzymologists are coming up with new unconventional ways to explain enzyme catalysis based on phenomenon of quantum mechanics tunneling (Klinman and Kohen 2013).

The catalysis takes place in specialized pockets within the enzyme structure called the active site where the substrate forms a complex through interaction with various amino acids of the enzymes. Both covalent and non-covalent interactions are responsible for binding and catalysis. Especially non-covalent interactions such as hydrogen bond, hydrophobic, electrostatic, van der Waal, dipole, cation pi, and pi-pi interactions are indispensable as they provide the very much desired reversible interactions allowing stabilization of transient state and later release of the product. While the active site comprises of ~10 amino acids, only three or four of them are directly involved in catalysis (catalytic amino acids), while many of them help in orienting or polarizing residues for substrate binding and creating optimum environment for catalysis (Gutteridge and Thornton 2004) .

In more detail the sequence of chemical events taking place during enzyme catalysis has been categorized broadly into different groups – acid/base, electrostatic and covalent catalysis.

Acid/base catalysis the most common mechanism used by two-third of known enzymes (Findly 1961; Bartlett 2002),where transition state stabilization is accomplished by donating proton from substrate to enzyme by amino acids which are weak acid/base such as histidine. Electrostatic catalysis works on neutralization of transition intermediate by providing complementary charge coming from different sources such as charged amino acids, metal ions or even dipole of backbone amide group (Hooper 1994; Wilmouth et al. 2001). Lastly, covalent catalysis is characterized by the formation of a covalent bond between enzyme and substrate which begins by a nucleophilic attack on the substrate. Examples include the formation of electrophilic Schiff base in case of and tetrahedral intermediates in case of serine proteases (Willmouth et al. 2001).

1.3. Evolution of enzyme – A structure-function perspective:

Enzymes are catalysts of cellular metabolism and thus are the catalysts of life. For survival in changing environments, a living system has to adapt continuously in the direction of change by adjusting its metabolism to suit the new. Enzymes, under such evolutionary pressure, can evolve in different ways through their substrate specificity, stability, rate of catalysis or acquire a totally novel function. It is generally assumed that present day enzymes diverged from a small set of ancestral enzymes with broad substrate specificity or promiscuity and low effectiveness (Khersonsky and Tawfik 2010). With subsequent cycles of mutation events, the enzymes evolve to become more specialized and more efficient.

At the genetic level, mechanism of evolution of enzyme can be explained in term of point mutation and gene duplication events and a combination of both. Duplicated (paralogous) genes are considered the raw material for evolutionary novelties as they provide new material for mutation, genetic drift and selection to act upon (Lynch and Force 2000; Zhang 2003). Gene duplication may occur from either single gene duplication events such as unequal crossing over and retroposition or a genome wide/chromosomal duplication event. Three theoretical fates of such genes has been suggested – a) non-functionalization, where one copy stays functional and the second copy may simply become silenced by degenerative mutations, b) sub-functionlization, where both the copies of the gene share some aspect of the mother function while differing in some. Such mutation ensures stable maintenance of

both the copies as they provide their own individual advantage to the organism. c) neo-functionalization where one copy attains a totally novel function while the mother function remains with the other copy (Lynch and Force 2000; Zhang 2003). These events give rise to divergent evolution as the enzymes diverge from a common ancestor. On the contrary, enzymes of different origins can also evolve separately to converge to perform equivalent functions, termed as convergent evolution, as different species could face similar survival problems.
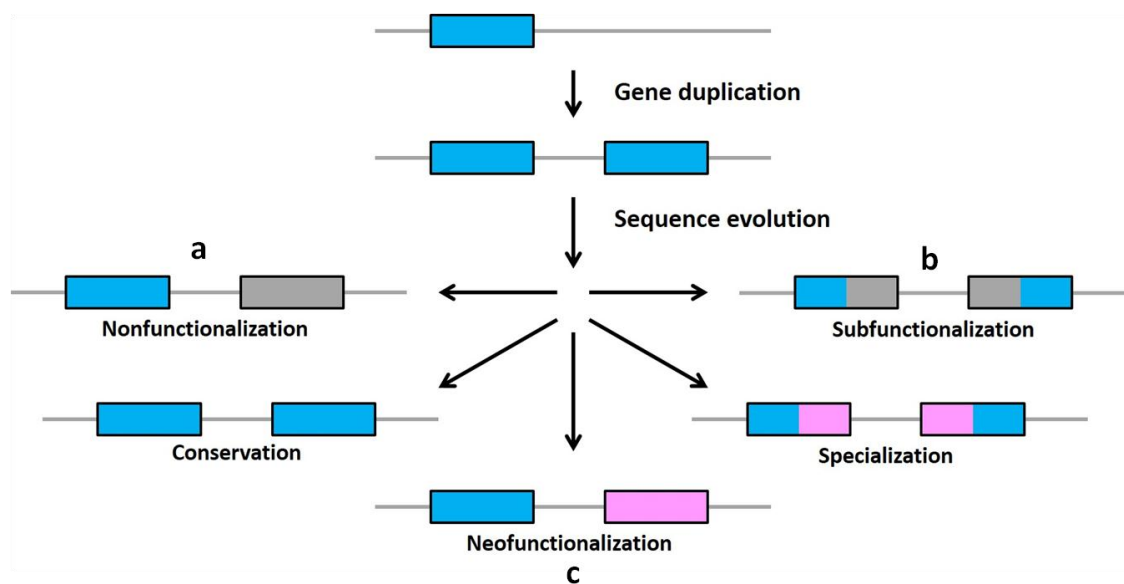
Fig. 1.1. Depiction of functional evolution of gene (Adapted from http://www.personal.psu.edu) a. Non-functionalization – loss of one copy of the duplicate genes as the most common fate of such genes, b. sub-functionlization – ancestral function gets sub divided between the copies of gene, c. neo-functionalization – one copy retains ancestral function while the other acquire new function.
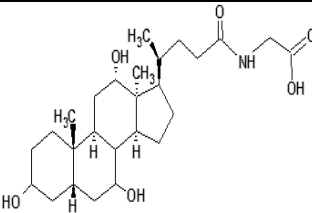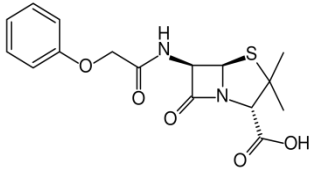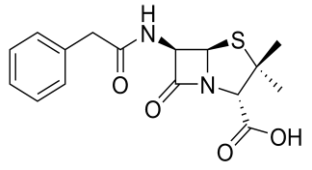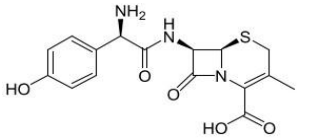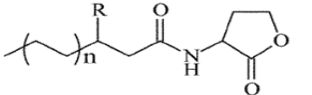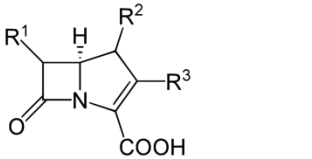
It is well established that enzymes are made up of two or more independent structure units called domains which can have either independent function or work incorporation with other domains in case of multidomain proteins (Doolittle 1995; Bashton and Chothia 2007). They are referred to as the evolutionary unit of protein as domain shuffling (non-homologous recombination) is the most common source of novel function. Besides domain shuffling, gene divergence can occur through by various mechanisms such as gene fusion, gene recruitment, gene transfer and posttranslational modification. Changes in sequence translate
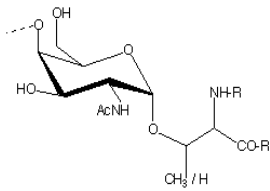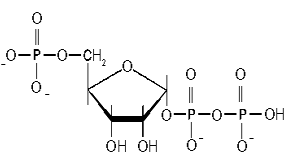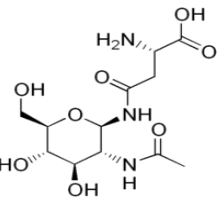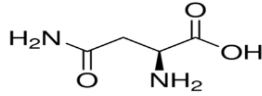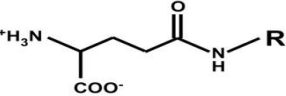
into modifications in three-dimensional structures of the enzyme which in turn reflects on function. However protein structure elements are usually far better conserved than sequence as a result of restrains imposed on tertiary structure of protein (Illergard et al. 2009). As a result, 3D structure becomes a better tool for detection of evolutionary relationship even when sequences become too diverged. Classification of proteins becomes possible based on collective homology and variation data observed at different level of protein structure as a result of divergent evolution. *Family* consists of proteins with significant sequence homology (greater than 30%) and catalyze with the same mechanism. However, orthologs, i.e., homologous gene present in different species, tend to have lower sequence identity and needs validation of their position from 3D structure data. *Superfamily* consists of families of proteins with common structural features and either catalyzes the same chemical reaction on different substrates or different overall chemical reactions which share certain stage of mechanism owing to conservation of active site residues (Gerlt and Babbitt 2001). Detailed entries of such classifications can be obtained from databases like Structural Classification of Proteins (SCOP; Hubbard et al. 1999) and Class Architecture Topology and Homology (CATH; Sillitoe et al. 2014).
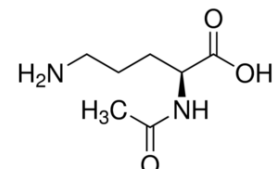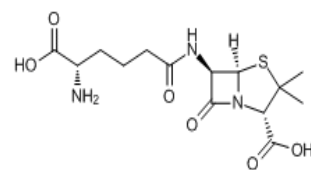
1.4. Ntn hydrolase superfamily:

Brannigan et al. (1995) suggested the name Ntn hydrolase to represent a newly found superfamily of enzymes, acting on non-protein amide bonds. Many members under this superfamily have now been documented with structural data. There are currently 7 representative families (SCOP database) under this superfamily - class II glutamine amidotransferase (GAT), penicillin G acylase, penicillin V acylase, proteasome subunits, glycosylasparaginase, γ-glutamyl transpeptidase (GGT) like and SPO2555-like. Ntn hydrolase superfamily is considered one of the most versatile groups of enzymes, with diverse substrate preferences and functions. Many Ntn hydrolase members have applications related to human health, including bile salt hydrolases (probiotics), penicillin and cephalosporin acylases (antibiotic industry), acid ceramidase and GGT (clinical diagnostic markers).

Table 1.1. List of structurally characterized Ntn hydrolases and their preferred substrates.

Enzyme	Organism	Nucleo-phile	Preferred substrate	PDB ID
Bile salt hydrolase	*a) Clostridium perfringens* (Rossocha et al. 2005) *b) Bifidobacterium longum* (Kumar et al. 2006)	Cys	 Bile salts	a) 2BJF b) 2HF0
Penicillin V acylase	*a) Bacillus sphaericus* (Suresh et al. 1999) *b) B. subtilis* (Rathinaswamy et al. 2005) *c)Pectobacterium atrosepticum* (Avinash et al. 2016)	Cys	 Penicillin V	a) 2PVA b) 2OQC c) 4WL2
Penicillin G acylase	*a) Escherichia coli* (Duggleby et al.1995) *b) Alcaligenes faecalis* (Varshney et al. 2012)	Ser	 Penicillin G	a) 1PNK b) 3K3W
Cephalosporin acylase	*Brevudimonas dimunuta* (Kim et al. 2006)	Ser	 Cephalosporin	1FM2
N-acyl homoserine lactone (AHL) acylase	*Pseudomonas aeruginosa* (Bokhove et al. 2010)	Ser	 *N*-acylhomoserine lactone	2WYE
Carbepenam synthase (*carA*)	*Pectobacterium carotovorum (*Miller et al. 2003*)*	Ser	 Carbapenam	1Q15

L-aminopeptidase – D-alaesterase/ amidase	*Ochrobactrum Anthropi (*Bompard-Gilles et al. 2000*)*	Ser	Peptides	1B65
Glycosyl asparaginase	*Flavobacterium Meningsepticum* (Guo et al. 1998)	Thr	O-linked ligosaccharide	1AYY
20S proteasome	*a) Thermoplasma acidophilum (*Lowe et al. 1995*)* *b) Saccharomyces cerivisiae (*Groll et al. 1997*)*	Thr	Ubiquitinylated proteins	a) 1PMA b) 1RYP
Glutamine phosphoribosyl pyrophosphate (PRPP) transferase	*a) B. subtilis* (Smith et al. 1995) *b)E. coli (*Kim et al. 1996*)*	Cys	PRPP	a) 1GPH b) 1ECG
Human aspartyl glucosaminidase	*Homo sapiens* (Oinonen et al. 1995)	Thr	Aspartyl glucosamine	1APY
L-asparaginase	*Lupinus luteus* (Michalska et al. 2005)	Thr	L-asparagine	2GEZ
γ-glutamyl transpeptidase	*E. coli* (Okada et al. 2006)	Thr	R=Cys−Gly γ-glutamyl compounds	2DBU

Ornithine acetyl transferase	*Streptomyces clavuligerus* (Elkins et al. 2005)	Thr	 N-acetyl L-ornithine	1VZ8
acyl coA: isopenicillin N acyltransferase (acyl coA: 6-APA transferase)	*Penicillium Chrysogenum* (Bokhove et al. 2010)	Cys	 Isopenicillin N	2X1C

Ntn hydrolases are named so because of the presence of an N-terminal nucleophile residue involved in catalysis which can be Ser, Cys or Thr. The most prominent structural feature of this superfamily is the conserved Ntn fold consisting two central antiparallel sheets sandwiched between 2 α-helical bundles, forming a four-layered $\alpha\beta\beta\alpha$-structural core (Brannigan et al. 1995; Oinonen and Rouvinen 2000). While the structural core remains conserved, variations occur mainly outside the core such as variation in length of loops connecting the secondary structures, addition of extra β strand in the central sheet and insertions in loops. An interesting exception to these general features is the archeal IMP cyclohydrolase PurO which lacks an N-terminal nucleophile, although it possesses the Ntn fold (Kang et al. 2007).

(a)   (b)
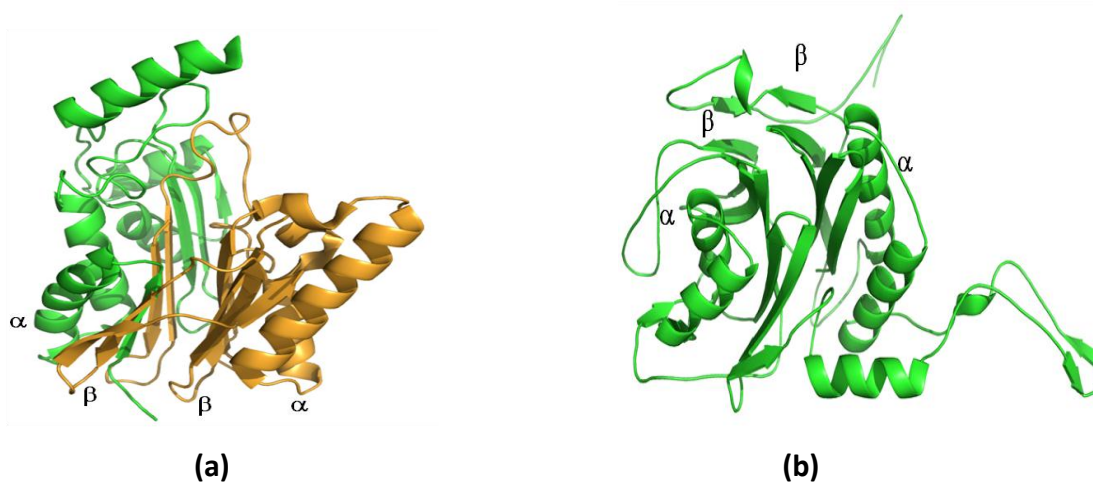
Fig. 1.2. Depiction of conserved $\alpha\beta\beta\alpha$ fold of Ntn hydrolase. (a) Glycosyl asparaginase (PDB ID: 2GL9) (b) Penicillin V acylase (PDB ID: 3PVA).
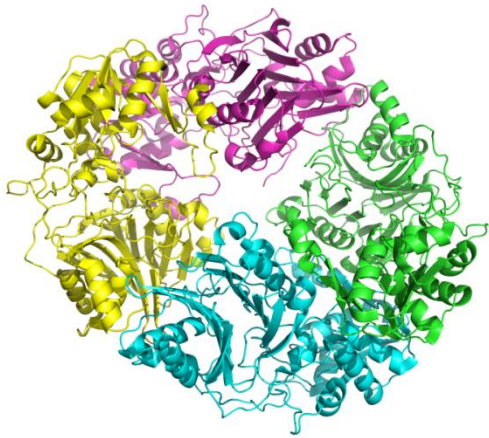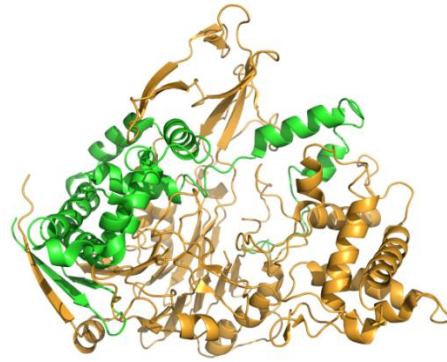
PRPP (PDB ID: 1GPH)   Penicillin G acylase (PDB ID: 3K3W)
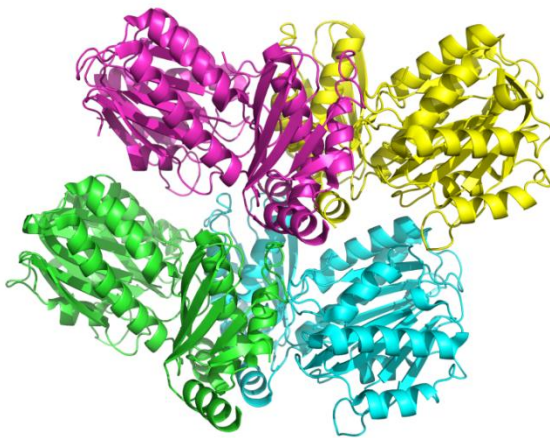
Ornithine acetyl transferase (PDB ID: IVZ8)   AHL acylase (PDB ID: 2WYE)
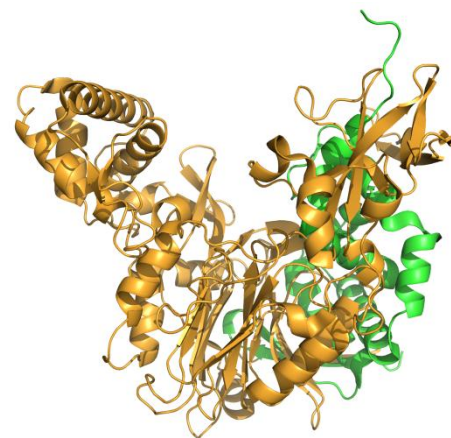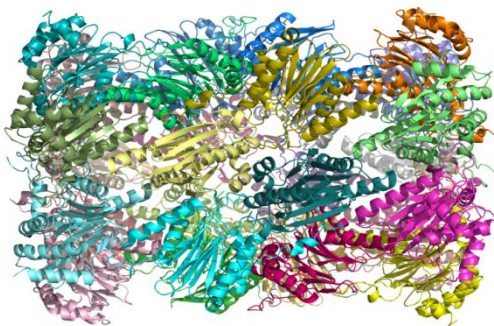
Proteaosome (PDB ID: 1RYP)   Glycosyl asparaginase (PDB ID: 2GL9)

Fig. 1.3. The versatile 3D structures of various Ntn-hydrolases.

All Ntn hydrolases are also synthesized as inactive precursors and undergo post translational processing to expose the catalytic N-terminal nucleophile and give rise to active enzyme. In

the precursor, the scissile peptide bond joining the pro-segment with the enzyme is concealed in the central β-sheet, inaccessible to external protease activity. Such observation indicated the activation step to be an autocatalytic or intermolecular cleavage event and it was demonstrated experimentally in penicillin G acylase and cephalosporin acylase (Kasche. 1999; Kim et al. 2006). The pro-segment may vary from a peptide segment as in penicillin V acylase to an N-terminal formyl-methionine as in case of bile salt hydrolase. The mechanism of autoproteolysis has been extensively studied (Kasche et al. 1999; Kim et al. 2006; Ditzel et al. 1998) and is highly diverse among different Ntn hydrolases.

1.5. Catalytic mechanism of Ntn-hydrolases:

The mechanism of catalysis in Ntn hydrolases resembles that of the serine protease and involves the formation of a catalytic triad. However a catalytic base like His is absent in Ntn hydrolases; it has been suggested that the α-amino group of the N-terminal nucleophile itself acts as the catalytic base for deprotonation of its own thiol or hydroxyl group. This unusual role of the N-terminal nucleophile to act as both the nucleophile and base was proved experimentally in cephalosporin acylase (Lee et al. 2000).

The first step of the reaction is the activation of the nucleophile by deprotonation of its OH or SH group by the α-amino group of the nucleophile itself. The proton abstraction initially was thought to occur directly between the two parts of the nucleophile, however Duggleby et al. (1995) observed the orientation between the α-amino group and the γ-O or γ-S proton was not ideal for the direct transfer. It was proposed that a water molecule would mediate the transfer acting like a 'virtual base' between the proton donor and the α-amino group. This is supported by the presence of water molecule near the nucleophile in all mature Ntn-hydrolase (Isupov et al. 1996; Suresh et al. 1999). However, in case of Cys-Ntn-hydrolase, it is suggested that the almost neutral side chain pKa of Cys might not require a catalytic base like water for nucleophile activation as required by other Ser/Thre-Ntn-hydrolases (Noren et al. 2000, Lodola et al. 2012). At the same time, in *Helicobacter pylori* γ-glutamyltranspeptidase (Williams et al. 2009) and *E. coli* asparaginase EcAIII (Michalska et al. 2005), it is proposed that side-chain hydroxyl can act as a virtual base instead of a water

molecule. After the activation of the nucleophile, it is stabilized by hydrogen bond interactions provided by nearby residues (**Fig 1.4**).
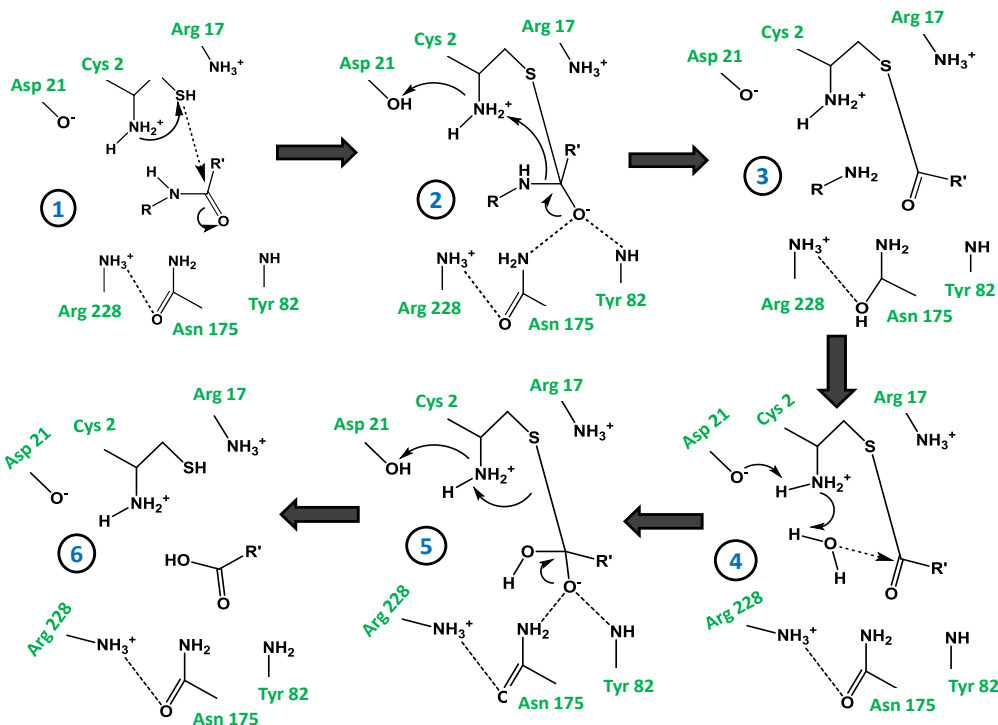
Fig. 1.4. Cys-Ntn hydrolase catalytic activity adapted from Avinash et al. (2016), depicting the role of important residues (Asp21, Arg 17) stabilizing the nucleophile Cys (1). Asn 175 and Tyr 82 are the oxyanion forming residues which form the tetrahedral intermediate with the substrate (2). Final release of the cleaved product is achieved by another nucleophilic attack with the help of a water molecule acting as a virtual base (4,5,6).

This is followed by the nucleophile attacking the carbonyl carbon of the substrate, generating the negatively charged tetrahedral intermediate. This intermediate is stabilized by the oxyanion hole formed by a backbone amide and the hydroxyl or amide group of an amino acid side chain. The oxyanion hole is the signature feature of catalysis among Ntn-hydrolases as well as serine protease and it is highly conserved. The nucleophile and the residues are separated by a distance of around 4.5–5.5 Å (Oinonen and Rouvinen 2000). This intermediate collapses as the step of acylation with the α-amino group donating proton to the nitrogen of the scissile amide bond of the substrate leading to the discharge of the amino part of the substrate. At this stage the acyl-enzyme complex is left behind. Deacylation takes

place with the help of a second water molecule which does a nucleophilic attack on the carbonyl carbon of the acyl-enzyme complex. The α-amino group receives the proton and a second tetrahedral intermediate is formed. The reaction becomes complete after α- amino group transfers the proton to the nucleophile (Duggleby et al. 1995; Oinonen and Rouvinen 2000). Although this mechanism is universal to all Ntn-hydrolase, Lodola et al. (2016) have suggested some unique features pertaining to cys-Ntn hydrolases such as a zwitter ionic form (S-/NH3+ ion pair) of the cysteine nucleophile is suggested.

Table 2.1 Variation in catalytic components such as oxyanion residues among different Ntn-hydrolases

Enzyme	Nucleophile	Nucleophile stabilizing residues	Base	Oxyanion Residue 1	Oxyanion Residue 2
PVA (Avinash et al. 2016)	Cysteine	Arg19, Asp22	None	Tyr82	Asn 183
BSH (Lodola et al. 2012)	Cysteine	Arg17, Asp21	None	Asn81	Asn172
PGA (McVey et al. 2001)	Serine	Gln 23	H_2O	AlaB69	AsnB241
AHL acylase (Bokhove et al. 20010)	Serine	His23, Arg297	H_2O	Val70	Asn269
Cephalosporin acylase (Kim et al. 2006)	Serine	His23	H_2O	Val70	Asn244
Proteasome (Lõwe et al. 1995)	Threonine	Lys33	H_2O	Gly130	Ser129
Glycosylasparaginase (Wang et al. 2007)	Threonine	Thr170	None	Gly204	Thr203
γ−Glutamyl transpeptidase (Okada et al. 2006)	Threonine	Thr-409	H_2O	Gly-483	Gly-484

1.6.   Choloylglycine hydrolases:

Two members of the Ntn hydrolase superfamily exhibiting a Cys nucleophile, penicillin V acylase (PVA) and bile salt hydrolase (BSH) make up the Choloylglycine hydrolase (CGH) family (Kumar et al. 2006). PVA and BSH share sequence homology and significant structural similarity. We will discuss CGH in terms of individual members.

1.6.1. Penicillin V acylase

Penicillin acylases (PA, E.C. 3.5.1.11) are industrially important enzymes that hydrolyze penicillin (Pen) to 6-aminopinicellanic acid (6-APA), the precursor for synthesis of semi synthetic β-lactam antibiotics (Shewale and Sudhakaran 1997). Based on different types of penicillins, PA are of several types. Two main types of PA are penicillin V acylase (PVA) and penicillin G acylase, which act on phenoxymethylpenicillin and benzylpenicillin respectively. PVA contributes to 15% of the annual production of 6-aminopenicillanic acid (6-APA), while the rest of the production is carried out by penicillin G acylase (PGA). Some reports emphasized that PVA offers several advantages over PGA such as higher conversion at higher substrate concentration and tolerance of broader pH range (Arroyo et al. 2003).
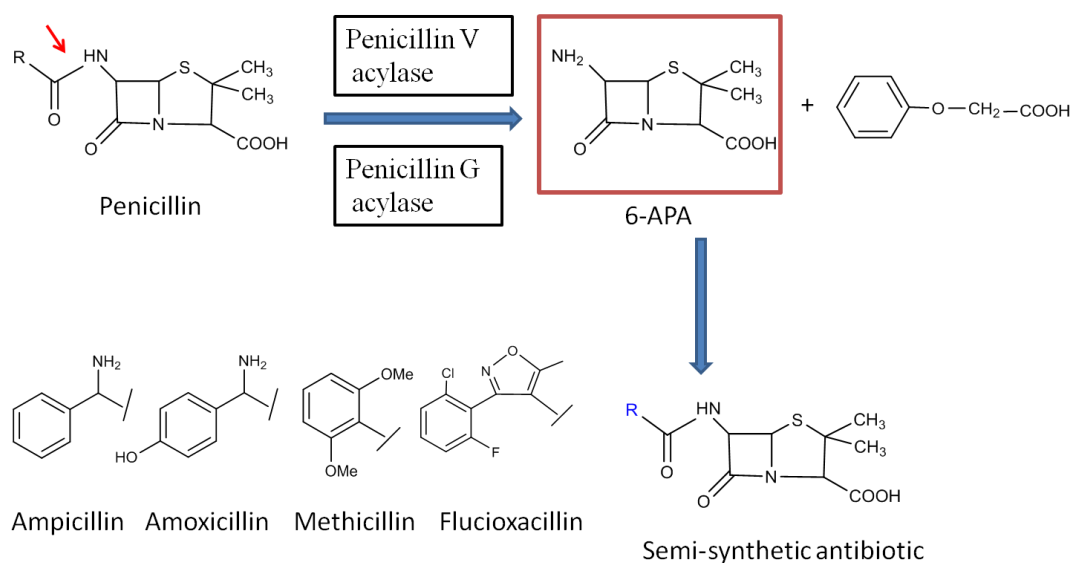
Fig 1.5. Hydrolysis of penicillin by penicillin acylases. Penicillin G acylase acts on benzyl penicillin and penicillin V acyase on phenoxy penicillin. Resulting hydrolysis product 6-amino penicillanic acid (6 APA) is used for the synthesis of semi-synthetic beta lactam antibiotics such as ampicillin, amoxicillin, methicilin etc.

PVAs are produced extracellularly by actinomycetes and fungi, intracellularly by bacteria (Sudhakaran and Borkar 1985). Gram-negative PVAs have been proposed to be periplasmic (Kovacikova et al. 2003) whereas PVAs from Gram positive bacteria are cytopasmic (Olsson et al. 1985, Rathinaswamy et al. 2005). PVA production is usually constitutive (Carlsen and Emborg 1981); but there are reports of enhancement of activity by the addition of

phenoxyacetic acid (Sudhakaran and Borkar 1985), sodium glutamate in the medium (Ambedkar et al. 1991, Philem et al. 2013). Catabolite repression by glucose has also been observed in *Streptomyces lavendulae* (Philem et al. 2013) and Chainia (Shewale and Sudhakaran 1997).

Structures of PVA solved using X-ray crystallography are now available from both Gram negative and positive bacteria – *Bacillus sphaericus* (Suresh et al. 1999), *Bacillus subtilis* (Rathinaswamy et al. 2005) and *Pectobacterium atrosepticum* (Avinash et al. 2016). PVAs are reported to exist as tetramers in crystal and in solution. Comparisons drawn from these reports show distinct structural differences between PVA homologs from Gram positive and negative bacteria including a reduction in tetrameric interface area. Gram negative PVAs show significant shortening in tetramer assembly motifs, (**Fig 1.6**) which in Gram positive PVAs are known to stabilize the tetramer by making a cross linking network between opposite monomers (Avinash et al. 2016). Moreover substitution of active side residues involved in substrate binding to Trp has been shown to be responsible for enhanced activity in PVAs from Gram-negative bacteria (Avinash et al. 2016).
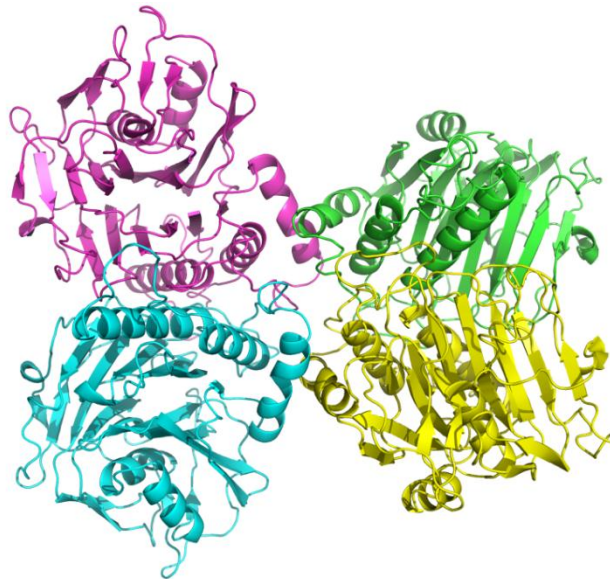
Fig 1. 6. 3D structure of tetrameric PVA from *Pectobacterium atrosepticum* (PDB ID 4WL2).

Although functionally similar and acting on related penicillin substrates, PVAs and PGAs do not show any significant sequence homology and also differ in subunit composition. The

catalytic mechanism of PVA differs from that of PGA. Cys1 allows direct transfer of proton from the water molecule during hydrolysis. The oxyanion hole forming residues in PVAs remain conserved as Tyr82 and Asn175 where Tyr82 at this position provides stacking interactions with the phenyl ring of Pen V. Variation in oxyanion hole forming residues among various Ntn hydrolases has been highlighted in **Table 2.1**. Residue Asp 274 in PVA has been suggested to interact with the extra oxygen in Pen V and confers preference over Pen G. Post translational processing in PVA involves the removal of terminal methionine to expose N-terminal cysteine. However the *Bacillus sphaericus* PVA features a Met-Leu-Gly tripeptide at the N-terminal and undergoes more complex processing (Suresh et al. 1999).

1.6.2 Bile salt hydrolase

Bile salt hydrolases are involved in the de-conjugation of bile salts and are usually present in intestinal microbiota. BSH activity has been exploited in application of probiotics supplements (Jones et al. 2008). Bile salts are amphipathic, end-metabolites of cholesterol and are important in vertebrates for intestinal absorption of lipids, digestion of dietary proteins, and exert potent antimicrobial activity in the small intestine (Hofmann and Hagey 2008). They also act as endocrine signals, regulating systemic functions including cholesterol and glucose homeostasis. BSHs allow gut microbes to tolerate bile salt toxicity by deconjugating the conjugated bile salt, giving out bile acid and the corresponding amino acid (taurine/glycine) (**Fig. 1.7**). BSH activity benefits the host by regulating the delicate balance between cholesterol and bile acid, thus probiotic brings health benefits such as lowering of serum cholesterol.

Three-dimensional structures of BSH are available from *Clostridium perfringens* (*Cp*BSH) (Rossocha et al. 2005), *Bifidobacterium longum* (*Bl*BSH) (Kumar et al. 2006) and *Lactobacillus salivarius* (*Ls*BSH) (Xu et al. 2016). Like PVAs, most bile BSHs are also known to exist as tetramer in both solution and crystal. However, recently crystal structure of BSH from *L. salivarius* has been reported to be a dimer, when in solution the enzyme is suggested to be stable as both dimer and tetramer (Xu et al. 2016). The inactive pro-enzyme form of BSH usually contains an initial methionine before the N-terminal cysteine which undergoes post translational cleavage by methionine aminopeptidase (MAP). An exception is

the BSH from *Bacteroides thetaiotamicron* and PVA homologues from Gram-negative bacteria which carry a long peptide directed to the periplasm (Panigrahi et al. 2014).
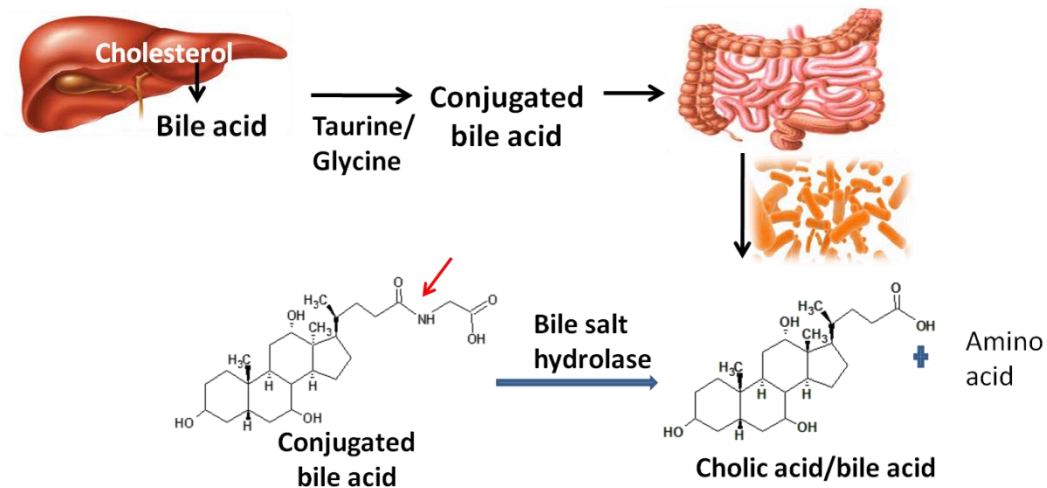
Fig. 1.7. Bile salt metabolism in vertebrate gut. Bile acid is synthesized from cholesterol in liver and conjugated with amino acid such as taurine/ glycine to form bile salts which gets transported to intestine. Intestinal bacteria producing BSH deconjugate the bile salt to release bile acid and its corresponding amino acid which microbes utilizes as carbon sources.
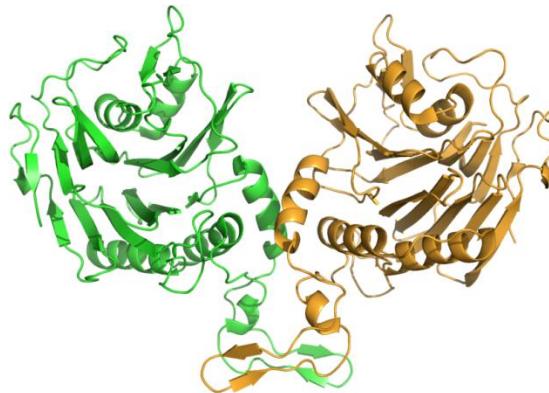
Fig 1.4. 3D structure of dimeric subunit of tetrameric BSH from *Bifidobacterium longum* (PDB ID 2HEZ).

It is suggested that BSH has diverged from PVA as adaptation to gut environment (Kumar et al. 2006). BSH and PVA share significant similarity in active site organization and catalysis mechanism, and have been proposed to share an evolutionary relationship. However, subtle

variations do exist in terms of loops surrounding the active site, and substitutions in substrate binding residues. The oxyanion hole forming residues in BSH are Asn81 and Asn175; in PVA Tyr 82 replaces Asn81. There are four conserved loops surrounding the active site in CGHs. These loops are comparatively longer in BSH than in PVAs, resulting in a higher active site volume and allowing entry of bulky steroid moiety bile salt. Moreover, loop 3 exhibits high flexibility and contains more polar residues to facilitate interaction with the cholate backbone of bile salts (Rossmann et al. 2008).

1.7.   Physiological role of CGH:

As a consequence of their pharmaceutical importance, researchs on CGHs have been focused on their applied aspects while the basic aspects such as physiological role and regulatory aspects are not still well understood. This is true especially in case of PVA. According to Valle et al. (1991), penicillin acylases may have a possible role in the utilization of alternative carbon sources by metabolizing aromatic compounds. Later, experimental finding of *E. coli* penicillin G acylase induction by phenylacetic acid (PAA) and repression by glucose supported such a relation. However, these studies are limited to PGAs and are not extendable to PVAs, as the two enzymes exhibit significant differences in sequence and subunit composition even though they act on related substrates.

In the last decade, there have been reports which link CGHs to metabolic pathways responsible for bacterial virulence and pathogenesis. For example, *pva* gene in *Vibrio cholerae* was down-regulated when virulence gene cascades are up-regulated by AphA, a transcription factor controlled by quorum sensing QS signaling pathway, indicating requirement of PVA in a non-virulent setting (Kovacikova et al. 2003). Many reports have raised the possibility of other roles for BSH besides bile tolerance. BSHs in *Listeria monocytogenes* (Dussurget et al. 2002) and *Brucella abortus* (Delpino et al. 2007) are known to act as virulence factors and influence the infectivity and colonization of the bacterial pathogens.   Certain strains of *Lactobacillus* also retain multiple copies of BSH; although for instance, only one product of the four bile salt hydrolase genes - bsh1, bsh2, bsh3, and bsh4 *Lactobacillus plantarum WCFS1* is active on bile salts (Lambert et al. 2008).   Another interesting report describes the ability of *Bifidobacterium* BSH to moonlight as one of its

plasminogen binding proteins as a virulence strategy (Candela et al. 2007). All these reports supported by the widespread occurrence of CGH genes in genome of different taxa spanning from bacteria, archaea to eukaryotes than expected before, point towards a diverse functionality of these enzymes. Thus exploring CGHs beyond their industrial or clinical applications would be a necessary step to answer questions on their role in bacterial metabolism and virulence.

1.8. Quorum sensing and quenching:

Quorum sensing is the social networking tool of the microbial world, a phrase first coined by Fuqua et al. (1994). Technically it is defined as the phenomenon of cell to cell signaling that enables the bacteria to collectively control gene expression via signaling molecules. It is a cell density dependent phenomenon; signaling molecules are produced constitutively and only upon reaching a threshold cell density (quorate), the signal molecule freely diffuses across the cell membrane and productively binds to its cognate receptor leading to gene expression of a number of phenotypes such as virulence, biofilm formation, sporulation etc (**Fig. 1.5**).
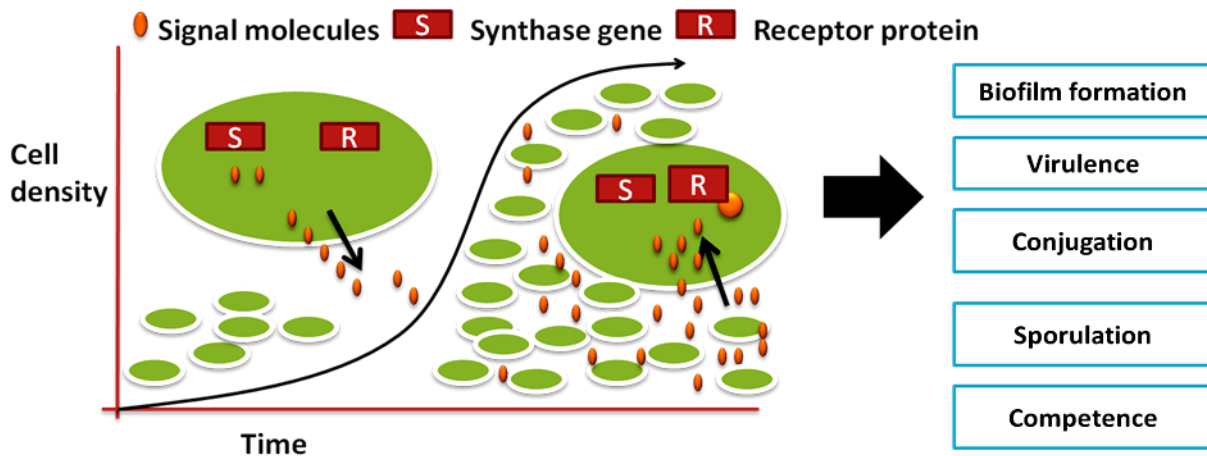
Fig.1.5. A simplified illustration of Quorum sensing. Signal molecules are synthesized by 'Synthase' gene (S) excreted into extracellular environment constitutively at low cell density signal. When population size increases to a threshold value, the signal molecules diffuses inside the cells and binds to receptor (R) to exert various phenotypic expressions.

Thus, QS controls a diverse range of cooperative behavior in microbes. Given the role in bacteria-host interaction, of both pathogenic and symbiotic nature, the disruption of the microbial QS system finds multiple applications in agriculture and medicine.

Many signaling molecules have been recognized so far, and are classified into three categories 1. N-acyl homoserine lactones (AHL) which are produced by Gram negative bacteria, 2. AutoInducing Peptides (AIP) in case of Gram positive bacteria and 3. AI-2, a furanosyl borate diester, which is found in both Gram positive and negative organisms, representing interspecies/universal signal molecule. In addition, there also exists inter-kingdom signaling as bacterial signal molecules, especially AHL, have shown to affect eukaryotic gene expression (William 2007; Diggle et al. 2007).

Interference with the quorum sensing system, termed quorum quenching (QQ) can be exploited to develop novel anti bacterials. QQ can be accomplished at different stages: 1) interfering with synthesis of signal molecule, 2) signal degradation and 3) receptor blocking. QQ agents can be of either biotic or abiotic in nature (Uroz et al. 2009). Abiotic factors capable of disrupting QS include temperature and pH. AHLs, particularly with short acyl chains such as C_4-HSL and C_6-HSL, are sensitive to inactivation at higher temperature and alkaline pH (Yates et al 2002; Uroze et al. 2009). Lactonolysis is the mechanism of AHL inactivation at alkaline pH and can be reversed upon acidification.

Enzymatic degradation of AHLs makes up the biotic side of QQ. Three QQ enzymatic activities are best recognized so far: 1. AHL acylases which hydrolyses amide bond in AHL, 2. AHL lactonase which causes lactone hydrolysis in AHL, 3. AHL oxidase and reductase which either reduce the keto group of 3-oxo-AHL or oxidize the w-1, w-2, and w-3 positions of long chain AHLs. These enzymes have been reported in both Gram positive and negative bacteria. Although most studies depict presence of a single type of QQ enzyme, few bacteria are also reported to possess both lactonase as well as acylase activities (*Rhodococcus erythropolis, Deinococcus radiodurans, Photorhabdus luminescens, Hyphomonas neptunium*) (Koul and Kalia 2011). It is also interesting that AHL lactonase and acylases

occurs in multiple copies in single organism in many examples elucidating the importance of these genes (Koul and Kalia 2011).
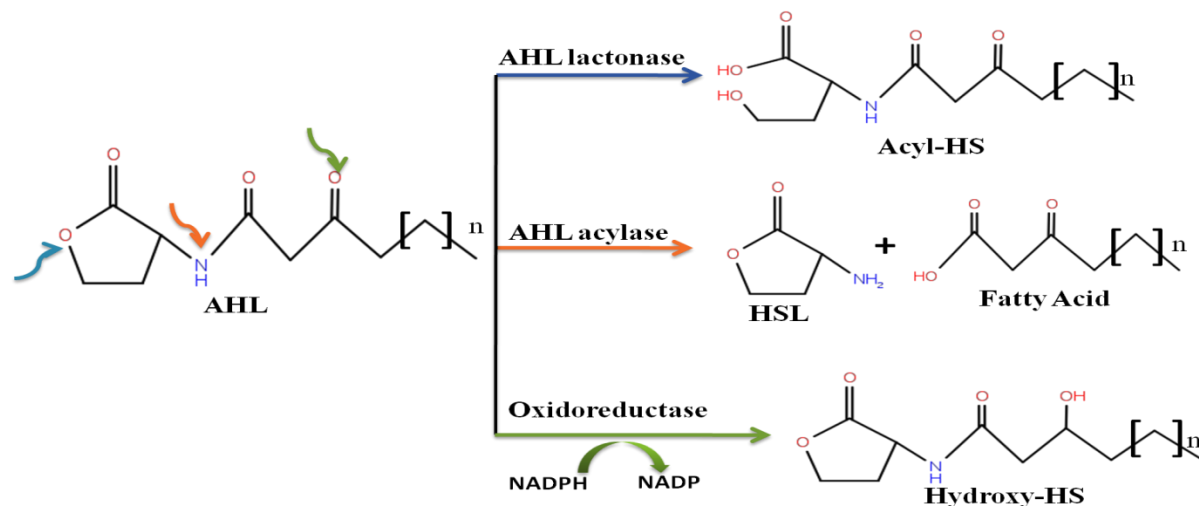
Fig. 1.6. **Three different QQ enzymatic activities on AHL. AHL lactonase cleaves the lactone ring of AHL ; acylase targets amide bond cleavage and oxidoreuctases reduce w-1, w-2, and w-3 positions of long chain AHLs.**

1.9.   AHLs and AHL acylases:

AHLs are the major QS signal molecules among Gram-negative bacteria which have been extensively studied in the last few decades. AHL signal molecules are synthesized by the LuxI family of acyl-HSL synthases that uses S-adenosylmethionine (SAM) as substrate and an intermediate of fatty acid biosynthesis, acyl-acyl carrier protein as the acyl chain donor (Fuqua et al. 2001). The LuxI-AHL synthase along with a second protein, LuxR-type protein play core role in the AHL mediated QS system. AHL interacts directly with the LuxR-type protein, after it reaches quorum concentration, and the resulting AHL-protein complex then binds at specific lux-box like gene promoter sequences affecting expression of QS target genes. AHLs can vary in the the length of their acyl side-chain, from C_4 to C_{14} atoms and the substitutions on the third carbon of the acyl chain. These different AHLs are generated by different LuxI homologs. Some of these homologues are TraR/TraI from *Agrobacterium tumefaciens* (Fuqua and Winans 1994) and LasR/LasI (Passador et al. 1993), RhlR/RhlI systems of *Pseudomonas aeruginosa* etc (Brint and Ohman 1995).
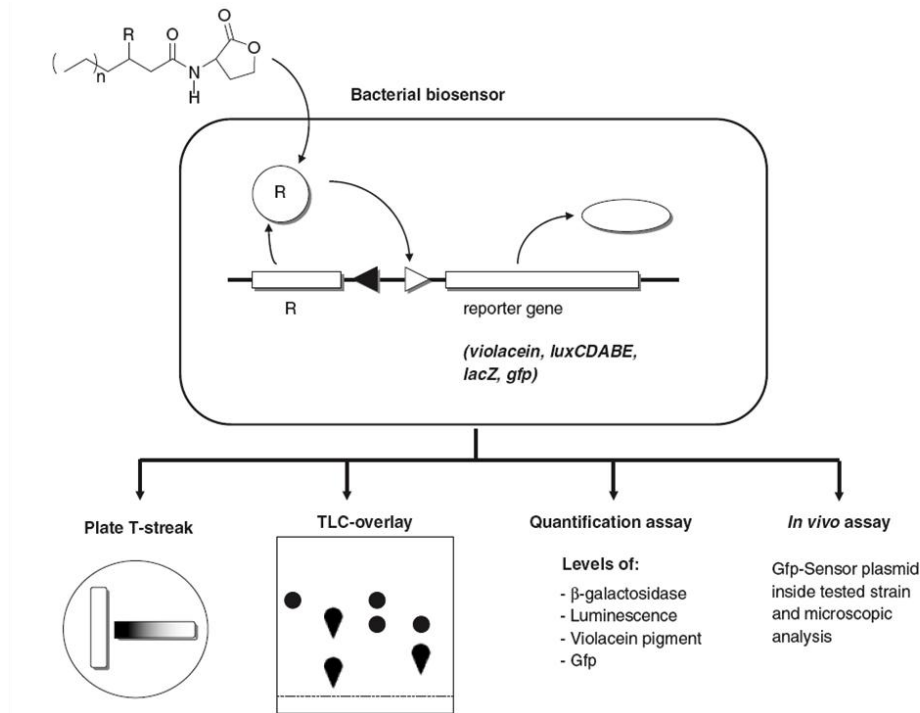
Fig. 1.7. Different methods of detection of quorum sensing and quenching using various bacterial biosensors, adapted from Steindler and Venturi 2007. LuxR family protein inside the bacterial biosensor (non-AHL producer) interacts with AHL which results in the transcription of a reporter gene(s) from a LuxR family-AHL regulated promoter.The plate 'T' streak assay is one the most common method for initial screening where reporter strain phenotype is expressed in a gradient near the test strain. TLC-overlay technique helps to separate AHL from a bacterial extract and visualize the reporter phenotype. Quantification assays are performed based on measurable reporter phenotypes such as bioluminescence, pigment production etc upon exposure to spent bacterial supernatants or bacterial organic extracts and finally *in vivo* assay are performed using gfp-based biosensors.

Over the years, a large number of bacterial biosensors have been developed for detection of the AHL-QS system and by extension AHL degradation as well. The simple AHL biosensors bacterial strains are deficient in AHL production and contain quorum sensing regulatory promoters fused to a reporter gene such as *lacZ* or *lux* operon etc which allow detection of various phenotypes such as bioluminescence, β-galactosidase, green-fluorescent protein and violacein pigment production etc. (**Fig. 1.7**) in the presence of broad range of AHLs (exogenously supplied) among Gram-negative bacteria.

AHL acylases catalyzes the irreversible cleavage of amide bond of AHL molecule releasing the homoserine lactone ring and the corresponding acyl chain. The fatty acyl chain is later used as carbon source by microbes. As a QQ agent, AHL acylase is a more promising enzyme choice as it provides more stable quenching unlike lactonase where the reaction could be reversed upon acidification. Moreover AHL acylase are more specific towards their substrate based on the acyl chain length, whereas lactonases usually exhibit broad substrate specificity. Majority of the AHL acylases are more active on long chain AHL with some lacking activity on AHLs shorter than 8 carbon atoms (Uroz et al. 2009)

AHL acylase also belongs to the Ntn hydrolase superfamily with Ser as the N-terminal nucleophile. They are heterodimeric and share closest sequence and structural relationship with penicillin G acylase (PGA) and cephalosporin acylase (CA). The 3D structure of AHL acylase, PvdQ from *P. aeruginosa* revealed that the active site pocket is unusually deep and hydrophobic to accommodate long chain AHLs, while the reaction mechanism and residues involved in catalytic mechanism were conserved as in PGA (Bokhove et al. 2010).

1.10. Cross reactivity and Promiscuity of CGH and related enzymes:

Classical enzyme studies considered specificity of reaction as one of the main characteristics of enzyme catalysis. However, many studies brought to light that some enzymes are capable of catalyzing more than one biochemical reaction (O'Brien and Herschlag 1999). Such a property where enzymes exhibit secondary activity other than the one they are evolved for is referred to as catalytic promiscuity or poly reactivity (James and Tawfik 2001). Promiscuity is considered a necessary arrangement for evolution of novel enzyme function (Copley 2003; Khersonsky and Tawfik 2006). It is reported that organisms inhabiting fluctuating environments are rich in promiscuous enzymes and such an enzymatic stock of new activities may help bacteria to face multiple ecological variables (Martínez-Núñez and Pérez-Rueda 2016). Among Ntn-hydrolases, PGA is considered a prominent example of promiscuous abilities - (1) Markovnikov addition reaction (Wu et al. 2005) which is an electrophile addition onto a double bond of a nucleophile; (2) transesterification reactions where an ester is transformed into another one through an interchange of the alkoxy moieties (Liu et al.

1999; Lindsay et al. 2002) and (3) Henry reaction known in organic chemistry to form carbon–carbon bond in the reaction of nucleophilic nitroalkane with an electrophilic aldehyde or ketone (Wang et al. 2010).

The two members of CGH family show promiscuous activity on each other's preferential substrates. Thus among previously reported CGHs with solved structures, there exists the 'pure' PVA exclusively specific for Pen V, with total absence of BSH activity, as in BsuPVA of *Bacillus subtilis* (Rathinaswamy et al. 2012) or *Pa*PVA from *Pectobacterium atrosepticum* (Avinash et al. 2015) and in a similar fashion there is the 'pure' BSH such as *Bl*BSH from *Bifidobacterium longum* (Kumar et al. 2006). Between these two types lies the cross reactive examples of CGH; BSH from *Clostridium perfingenes*, *Cp*BSH with low PVA activity; PVA from *B. sphaericus* (BspPVA) with low BSH activity (Kumar et al. 2006). A few BSHs including *Bacteroides, Lactobacillus salivarius* etc whose structures are known, have not yet been tested for PVA activity.

As mentioned in section 1.7, questions still abound regarding the possible physiological role of CGH. Keeping in mind the association of CGH with virulence related reports; it becomes rational to explore CGH activity in the virulence and pathogenesis circuit. One straightforward approach would be to screen quorum quenching activity of CGHs. Hence in the current project, we assess the QQ activity of selected novel CGHs from marine bacteria, *Shewanella loihica*-PV4 and Loktak Lake isolate, *Acinetobacter* AP24. Moreover PVAs and BSHs have been extensively studied from soil and gut inhabiting bacteria respectively; while CGHs from fungi, archea and marine bacteria with potential different /new properties remain largely unexplored.

2.0 Scope of the thesis:

The inspiration behind the thesis is to bridge the gaps in understanding CGHs from under explored sources such as marine environment which in turn would provide insights into the so far obscure in vivo role of these enzymes as well as their role in their new found association with virulence phenomenon. The work revolves around three CGHs of which two - *Sl*CGH1 and *Sl*CGH2, are from a single source, *Shewanella loihica*-PV4 which represent

marine CGHs and a third CGH from *Acinetobacter* APT13, a Gram negative bacterial isolate from water of Loktak lake, an Indo-Burma biodiversity hotspot. In **chapter 2** of the thesis, we describe biochemical characterization of SlCGH1 cloned and expressed in *E. coli* host. The substrate spectrum of the CGH and enzyme stability was explored using activity assays and fluorescence and CD spectroscopy methods. The results reveal that SlCGH1 is an unconventional CGH as it shows total absence of penicillin V acylase and Bile salt hydrolase activities and presence of AHL acylase activity. **Chapter 3** highlights the determination of the three dimensional structure of *Sl*CGH1 using X-ray crystallography and structural analysis of the unique substrate specificity of *Sl*CGH1. A comprehensive discussion on novel structural deviations from previously reported CGHs has also been given. **Chapter 4** focuses on cloning, enzymatic characterization and structural modeling of *Sl*CGH2. We found that SlCGH2 was also active on acyl homoserine lactones, however it shows more activity towards long chain AHLs when *Sl*CGH1 was more susceptible towards medium chain AHLs. Our results strongly suggested that *Sl*CGH1 and *Sl*CGH2 represent a potential new sub-class of CGH. **Chapter 5** represents work done on penicillin V acylase from *Acinetobacter* AP24, isolated from a water sample of Loktak Lake (Manipur, India) an Indo-Burma biodiversity hotspot. Optimization of fermentation parameters were done followed by enzyme purification and partial characterization. **Chapter 6** (Summary) integrates all the findings of the thesis and also highlights the future scope of the work.

Chapter 2

Cloning and characterization a choloylglycine hydrolase, *Sl*CGH1 from *Shewanella loihica* PV-4

2.1. Introduction:

At present, biological databases are flooded with protein sequences of unknown functions contributed by many genome and metagenomic sequencing programs. Annotation of such an enormous number of proteins sequences is an interesting research challenge. It is not a straightforward process as the presence of a gene doesn't guarantee an expected function; even a single mutation could be disruptive of the function. This is well illustrated in the case of CGHs; for instance, *Lactobacillus plantarum* WCFS1 carries four genes annotated as bile salt hydrolase - *bsh1, bsh2, bsh3, and bsh4*, however only one of these gene products was observed to be active on bile salts (Lambert et al. 2008). Moreover, enzyme moonlighting or promiscuity and experimental drawbacks to mimic the natural environment all adds to the complexity and magnitude of the problem.

Shewanella loihica PV-4 is a marine γ-proteobacterium, isolated from an iron-rich microbial mat at the active, deep-sea, hydrothermal vent of Loihi Seamount, Pacific Ocean (Gao et al. 2006). An examination of the whole genome sequence of *S. loihica* PV-4 available in NCBI database showed the presence of two putative choloylglycine hydrolase genes – *Shew_0681* and *Shew_2805*, predicted as penicillin V acylase like gene and choloylglycine hydrolase respectively. These two CGH genes showed low sequence homology (<20%) with other known CGHs and moreover, only 35% identity to each other. Further, *Shew_0681,* designated in this dissertation as *Sl*CGH1, also shows a high % similarity (70-64 %) to unannotated hypothetical proteins from marine bacteria (BLAST hits). The characteristics of CGH from marine bacteria have not been studied yet, and their low sequence homology with previously reported enzymes warrants an investigation. In addition, marine microbial mats also represent a potential source of unique ancestral enzyme homologues as it is a close replica of earth's earliest ecosystem. They represent extreme environments in terms of high metabolic rates, strong geochemical gradients inhabited by plethora of diverse microbial population enclosed in a biofilm matrix (Decho et al. 2009; Bolhuis et al. 2014).

The present study therefore ventured to characterize these enzymes to study their unique features in comparison to reported CGHs and gain a better understanding of their possible significance in the marine environment. To this end, we have cloned and expressed both the CGHs from *S.*

loihica in *E. coli* and characterized their biochemical and structural features. The biochemical characterization of *Sl*CGH1 has been described in this chapter. Extensive substrate screening was done against not only standard CGH substrates but also for related enzyme activities. Our results reveal a novel activity profile for a CGH.

2.2. Materials and methods:

2.2.1. Materials

All media components were procured from Himedia, India. Pen V (potassium salt), Bile salts, guanidine hydrochloride (Gdn-HCl), phenoxy acetic acid (POAA), ampicillin and HIS-Select matrix were procured from Sigma (USA). ENrich^{TM} SEC 650 column and molecular weight markers were from BioRad. DNA isolation and purification kits, cloning plasmids and *E. coli* strains were from Invitrogen (USA). All DNA manipulation enzymes and buffers were procured from New England Biolabs (NEB).

Shewanella loihica PV-4 ATCC BAA-1088 was procured from the American type culture collection (ATCC). Biosensor strains *Chromobacterium violaceum* CV026, *E. coli* pSB 401 and *E. coli* pSB1075 were kind gifts from Dr. Paul Williams, University of Nottingham and *Agrobacterium tumefaciens* NTL4(pZLR4) was kindly gifted by Dr. Stephen Farrand, University of Illinois, USA.

2.2.2. Cloning of *Sl*CGH1 gene from *S. loihica* PV-4

*Sl*CGH1 gene from the genomic DNA of *S. loihica* was PCR amplified using forward (5′ gcaatc*catatg*tgttctcgtctggtgacag 3′) and reverse (5′ caagat*ctcgag*catcttagaccaagtgatgg 3′) primers (restriction sites highlighted). PCR was carried out as follows: 95°C for 3 min, followed by 30 cycles of 95°C, 30 s; 58 °C, 30 s; 68°C, 60 s, and final extension at 68°C for 10 min. The PCR product was purified using NucleoSpin® Gel and PCR Clean-up kit (Macherey –Nagel) and inserted into pET22b(+) vector (Invitrogen) between NdeI and XhoI restriction sites. After ligation, the recombinant product was transformed into *E. coli* DH5α cells and selected on Luria-Bertani (LB) plates containing 100 μg ml^{-1} ampicillin. Positive clones were selected using colony PCR and sequence validation was done by sequencing the gene using T7 primers. Finally the recombinant vector was re-transformed into *E. coli* BL21 star (DE3) cells to express the enzyme with C-terminal His$_6$-tag.
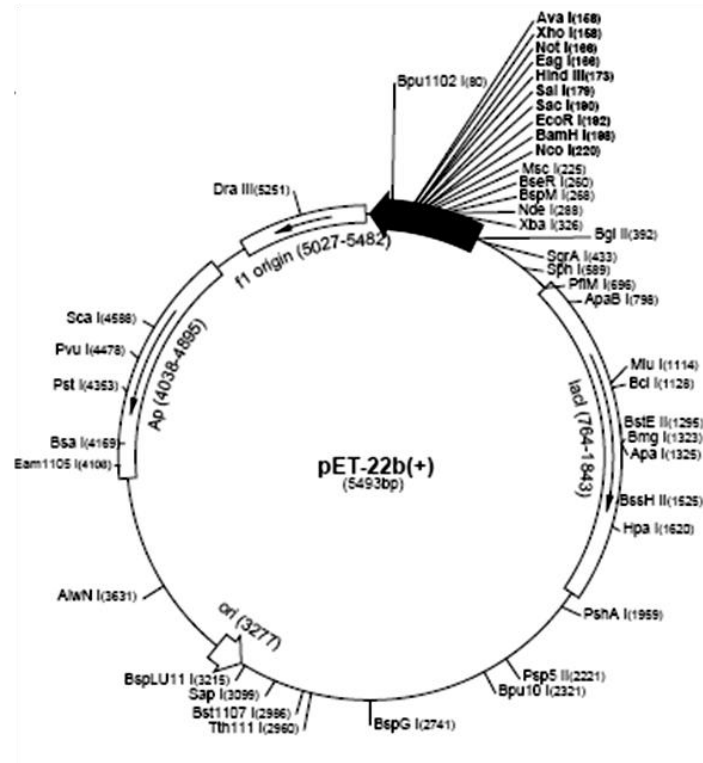
Fig. 2.1. Vector map of pET-22b(+) (Novagen:
http://www.helmholtzmuenchen.de/fileadmin/PEPF/pET_vectors/pET-22b_map.pdf)

2.2.3. Site directed Mutagenesis

To ascertain the role of cys1 as the N-terminal nucleophile, we mutated cys1 to ala (C1A), and to ser (C1S) using primers, 5' GAT ATA CAT ATG *GCG* TCT CGT CTG GTG ACA 3' and 5' GAT ATA CAT ATG *TCT* TCT CGT CTG GTG ACA 3' respectively (Mutated codon in italic). These two mutants C1A and C1S were constructed by site-directed mutagenesis, using a PCR following the above mentioned steps of cloning. The mutant PCR products were inserted into pET22b vector, cloned in *E. coli* DH5α and expressed in *E. coli* BL21 (DE3).

2.2.4. Expression and purification of *Sl*CGH1

Protein expression was induced by adding 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) to the culture of *E. coli* BL21 star (DE3) cells at OD_{600} = 0.5 (grown for 3-4 h at 37°C), followed by overnight incubation (16 h) at 16°C. The harvested cells were resuspended in lysis buffer (25mM Tris-Cl pH 7.0, 300mM NaCl, 10mM $MgCl_2$, 2mM β-mercaptoethanol and 0.1% Triton

X-100) and disrupted by sonication in ice bath for 15 min using a Branson sonifier. The supernatant was then applied to a HIS Select Ni^{2+} affinity column, pre-equilibrated with lysis buffer and unbound protein was washed with the same buffer. The bound protein was eluted with 250 mM imidazole. Protein fractions with high purity were dialyzed overnight in 100x volume of 20mM Tris buffer pH 7.2 containing 100mM NaCl and 1mM DTT at 4°C. The protein was further purified to homogeneity using a ENrich^{TM} 650 (BioRad) size exclusion column. Routine estimation of protein concentrations was done using Bradford method (Bradford 1976) and purity was confirmed by sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE), according to procedure of Laemmli (1970).

2.2.5. Molecular weight determination

Monomeric/subunit molecular weight was confirmed on a 12 % SDS-PAGE (Laemmli 1970). Electrophoresis was carried out at 25°C at 100 V (constant voltage) and the samples were electrophoresed alongside molecular weight markers till the bromophenol blue tracking dye reaches the end of the gel. The gel was stained by dipping in a solution of Coomassie brilliant blue solution (0.25% Coomassie brilliant blue R250 in 40% (v/v) methanol and 10% (v/v) glacial acetic acid) for 4 h or overnight. For visualization the stained gel was destained in a solution of 5% Propan-2-ol and 7% Glacial Acetic acid.

Oligomeric state and native molecular weight of *Sl*CGH1 were confirmed using size exclusion chromatography. 200 µl of the 20 mg/ml protein was loaded on ENrich^{TM} SEC 650 column (10X300 mm) fitted with a BioRad NGC^{TM} 10 Medium-pressure FPLC system. The column was initially equilibrated with 20mM Tris buffer pH 7.2 containing 100mM NaCl and 1mM DTT before protein loading and elution was carried out with the same buffer.

Matrix–assisted Laser Desorption Ionization/Time of flight-Time Mass Spectrometry (MALDI-TOF-MS) was also performed using a Voyager-De-STR MALDI-TOF mass spectrometer (Applied Biosystems) to obtain the monomer molecular weight. Sample was mixed with a matrix solution of 15 mg/mL sinapic acid in 30% acetonitrile in 1:5 ratios and was spotted on MALDI target plate. After drying the plate, it was introduced into the mass spectrometer. The spectrum was obtained using a pulsed N2 laser (337 nm), in linear mode with delayed ion extraction with an accelerating voltage of 25 kV.

2.2.6. Evaluation of enzyme activity

2.2.6.1. Penicillin V acylase activity

PVA activity was determined according to Bomstein and Evans method (1965), modified by Shewale et al. (1987), measuring the amount of 6-APA formed at 40 °C. A solution of Pen V, potassium salt (2% w/v) was prepared in 0.1 M sodium phosphate buffer pH 6.0. The 6-APA formed was estimated using 0.6% (w/v) p-dimethylaminobenzaldehyde in methanol. One unit (IU) of PVA activity is defined as the amount of enzyme that produces 1 μmol 6-APA per minute under the conditions defined.

2.2.6.2. Bile salt hydrolase activity

BSH activity was estimated by method described by Kumar et al (2006). The amount of free amino acids released upon incubation of the enzyme sample with 1 mM sodium glycodeoxycholate (GDCA) or sodium taurodeoxycholate (TDCA) at 30 ℃ in 10 mM sodium phosphate, pH 7.0 containing 1mM DTT was estimated by mixing an equal volume of 2% ninhydrin solution before boiling for 15 min. Absorption was measured at 570 nm and one unit of BSH activity was defined as the amount of enzyme that liberates 1 μmol of the amino acid from substrate per min.

2.2.6.3. Lipase and protease activity

Lipase assay was performed by spectrophotometric measurement of the p-nitrophenol released from hydrolysis of p-nitrophenylpalmitate (pNPP) (Winkler and Stuckmann 1979). Protease assay was done according to the method described by Sigma Aldrich for non-specific protease activity with slight modifications (Enyard 2008) measuring the amount of tyrosine with casein as substrate.

2.2.6.4. AHL degradation assay

Preliminary screening of AHL degradation activity was performed using biosensor strains *C. violaceum* CV026 for short to medium chain AHLs, C_4 to C_8 HSLs (McClean et al. 1997) and *A. tumefaciens* NTL4 (pZLR4) for long chains AHLs, unsubstituted and 3-oxo-C_8-C_{12} HSL (Farrand et al. 2002) . *Sl*CGH1 (40 μg) was incubated with 0.1 μl of AHLs (20 mM stock) for 2 h at 30 °C.

Medium chain AHL degradation bioassay using *C. violaceum* CV026: CV026 seed culture was grown overnight at 28°C in LB broth containing 100 μg ml^{-1} ampicillin and 30 μg ml^{-1} kanamycin. 0.2% (v/v) seed culture was added to 2 ml LB along with 20 μl of the enzyme-signal

mix and incubated overnight at 28°C. Short chain AHLs C_6-HSL and C_8-HSL were assayed using this method.

Long chain AHL degradation bioassay using *A. tumefaciens* NTL4 (pZLR4): *A. tumefaciens* NTL4 (pZLR4) culture was grown overnight in nutrient broth containing 30 μg ml^{-1} gentamicin. 10% innoculum of reportor culture was mixed with X-gal (60 μg ml^{-1}) and 20 μl enzyme-signal mix and grown overnight incubation at 28°C. Long chain AHLs C_{10}-HSL, C_{12}-HSL and 3-oxo-C_8-HSL were assayed using this method.

Lux-based bioluminescence assay for AHL degradation: Estimation of AHL degradation was carried out using lux-based bioluminesence assay based on two biosensor strains – 1) *E. coli* pSB401 for unsubstituted and 3-oxo-C_6 to C_8-HSL and 2) *E. coli* pSB1075 for long chain AHLs, unsubstituted and 3-oxo-C_{10} to C_{12}-HSL (Steindler and Venturi 2007). These reporter strains were grown in nutrient broth containing 10 μg ml^{-1} tetracycline for 12 h at 37°C. A protein concentration of 20 μg ml^{-1} was incubated with three different concentrations of each AHL substrates, 15, 25 and 100 μM from a 5 mM stock in DMSO, for 4 h at 30°C in a 96-well plate. The final reaction volume was made up to 100μl using 50 mM sodium phosphate buffer with 1mM DTT (Winson et al. 1998, Wahjudi et al. 2011). After incubation, 100 μl of 10 mM PBS (phosphate buffered saline) was added along with another 100 μl of the 1:100 diluted biosensors to appropriate AHL samples and incubated for 6 h at 30°C. All buffers used were filter sterilized. Heat-inactivated enzyme was used as control. Bioluminescence was measured in a 96-well white solid plate in a Varioskan® Flash spectral scanning multimode reader driven by SkanIt Software (Thermo Scientific). OD_{600} was also measured in a transparent 96-well plate using the same system, to monitor the bacterial growth. Results were expressed in terms of relative light unit (RLU; LU/OD_{600}) obtained from the ratio of the sample readings and the corresponding OD_{600}, in relation to the control (control RLU taken as 100%).

2.2.6.5. Analysis of AHL degradation by HR-MS spectroscopy

To elucidate and confirm AHL acylase activity, High resolution mass spectroscopy was carried out based on the method described by Huang et al. (2003) with slight modifications. Reaction products were analyzed using a C18 ultra-aqueous reverse-phase column fitted to a Q-Exactive system with Accela 1250 Pump and PDA detector from Thermo Scientific. Samples were run for

6 min and the mobile phase consisted of 50 % methanol (methanol-water with 1% acetic acid) for the initial 1 min and linearly increased to 80 % methanol. Samples were diluted in methanol before application.

2.2.7. Intrinsic fluorescence measurement

Fluorescence study of protein is based on the detection of conformational changes in proteins in different conditions using changes in the fluorescence of sensitive fluorophores like the aromatic amino acids Trp, Tyr and Phe. To get insights into stability of SlCGH1under different conditions such as pH, temperature and chemicals agents by assessing conformational changes, fluorescence measurements were made using a PerkinElmer Life Sciences LS50 fluorescence spectrophotometer connected to a Julabo F20 water bath. 60 µg enzyme in 2 ml of 10 mM potassium phosphate buffer pH 7.0 was excited at 295 nm and the emission was recorded from 310 to 400 nm wavelength. Blank readings were also taken by measuring the fluorescence of only buffer under similar conditions.

2.2.8. Circular dichroism measurement

Circular Dichroism (CD) spectra can provide an estimation of the secondary structure composition of proteins as they are primarily due to absorption by the peptide bond. Far UV CD (190-250 nm) spectra were recorded using a J815 spectro-polarimeter connected to a Peltier Type CD/FL Cell circulating water bath (Jasco, Tokyo, Japan). The protein sample was prepared in 10 mM potassium phosphate buffer pH 7.0 and filtered through a 0.22 µm membrane. Far-UV CD spectra were recorded in the range of 190-250 nm in a rectangular quartz cell of 1 mm path length at a scan speed of 100 nm/min with a slit width of 1 nm. Appropriate blank readings were subtracted from the sample readings and final spectra were expressed in terms of mean residue weight ellipticity (MRE in $degcm^2dmol^{-1}$).

2. 3. Results and Discussion:

Being pharmaceutically important enzymes, CGHs been exploited mainly for their applications in antibiotic production and probiotics while the basic aspects such as physiological role and regulatory aspects are not still well understood. This is true especially in case of PVA. Penicillin acylases are suggested to be 'scavenger enzymes' for aromatic carbon sources (Valle et al. 1991) particularly base on experimental finding of *E. coli* PGA induction by phenylacetic acid (PAA) and repression by glucose. Although PVA and PGA share similar substrates, they differ significantly in structure and sequence. BSHs, on the other hand, might be responsible for microbial bile tolerance and adhesion in mammalian gut as supported by a number of studies (Dambekodi and Gilliland 1998; Taranto et al. 1996; Taranto et al. 2003; Begley et al. 2005; Jones et al. 2008; Jarocki et al. 2014).

Interestingly, recent studies suggest that both PVA and BSH could be linked to bacterial signaling and virulence. For example, in *Vibrio cholerae,* PVA gene expression is down regulated when virulence gene cascades are up-regulated by AphA, a transcription factor controlled by quorum sensing QS pathway (Kovacikova et al. 2003). In *Bifidobacterium,* BSH moonlights as one of its plasminogen binding proteins as a virulence strategy (Candela et al. 2007). It is also noteworthy that there is widespread occurrence of uncharacterized CGH homologues in single and multiple copies in genomes of different taxa exhibiting sparse sequence similarity which suggests a more universal role of this enzyme family in the lifestyle of the host organisms.

All available enzyme structures of PVAs and BSHs are from free-living soil bacteria and gut inhabiting bacteria respectively; while CGHs from fungi, archea and marine bacteria with potential different or new properties remain largely unexplored. This is important in the context of finding insights into the physiological role of this enzyme family which is not yet distinctly defined. To get insights into the nature of CGHs of marine origin, the marine strain of *Shewanella loihica* PV-4 was selected for our current study. *S. loihica* PV-4 genome possesses two putative choloylglycine hydrolase genes, with low sequence homology of less than 20 % with known CGHs, while they are 35% identical to one another. Moreover, a BLAST search of

these two genes shows a majority of hits from marine CGHs, indicating the existence of a close group of marine CGHs which may be significantly different from CGHs of non-marine habitats.

Table 2.1. CGH sequence hits homologous to *Sl*CGH1 (https://blast.ncbi.nlm.nih.gov/).

Description	Max score	Total score	Query cover	E value	Identity	Accession
Penicillin V acylase-like protein (*Shewanella loihica* **PV-4**)	702	702	100%	0.0	100 %	WP_011864487.1
Hypothetical protein (*Vibrio fluvialis*)	525	525	99%	0.0	72 %	WP_032079919.1
Hypothetical protein (*Agarivorans gilvus*)	483	483	99%	$3e^{-169}$	66 %	WP_055731605.1
CGH (*Photobacterium swingsii*)	244	244	99%	$3e^{-75}$	40 %	WP_048897466.1
Hypothetical protein (*Agarivorans gilvus*)	230	230	99%	$6e^{-70}$	39 %	WP_055733775.1
Penicillin V acylase-like protein (*Vibrio tubiashii* **ATCC19109**)	228	228	100%	$2e^{-69}$	39 %	EGU54598.1
CGH (*Vibrio tubiashii*)	228	228	100%	$3e^{-69}$	39 %	WP_050809764.1
CGH (*Shewanella waksmanii*)	228	228	99 %	$4e^{-69}$	39 %	WP_028771424.1
CGH (*Vibrio ishigakensis*)	227	227	100 %	$8e^{-69}$	38 %	GAM56964.1
CGH (*Photobacterium jeanii*)	227	227	97 %	$1e^{-68}$	39 %	OAN13258.1
CGH (*Vibrio maritimus*)	226	226	97 %	$4e^{-68}$	38 %	WP_042471689.1
CGH (*Shewanella loihica* **PV-4**)	203	203	100 %	$2e^{-59}$	34 %	WP_011866602.1

In order to gain more insights into the unexplored characteristics of marine CGHs, the present study was aimed at performing biochemical and structural studies of these two genes. This chapter describes the cloning, expression of *Sl*CGH1from *S. loihica* PV-4 in *E. coli* and its biochemical characterization.

2.3.1. Cloning of *Sl*CGH1 gene from *S. loihica* PV-4

The 1080 bp gene coding for *Sl*CGH1was isolated by PCR amplification from the genomic DNA of *S. loihica* PV-4. *Sl*CGH1 gene also codes for a 26 amino acid signal sequence which helps to direct the protein to the periplasm. All Gram negative CGH are known to carry signal peptides of different length (Panigrahi et al. 2014). We amplified the *Sl*CGH1 gene without the signal peptide and added a methionine residue before the N-terminal cysteine (**Fig 2.3a**). Restriction digestion of the plasmid pET22b carrying *Sl*CGH1gene confirms the presence of the gene from the insert size (**Fig. 2.2**). Plasmid sequencing validated the absence of any mutation in the cloned gene. Colony PCR of the *E. coli* DH5α colony after transformation validated positive clones carrying intact *Sl*CGH1- pET22b plasmid.

A - 1 kb insert
B - Vector backbone
M - 1 kb DNA ladder

(a)   (b)
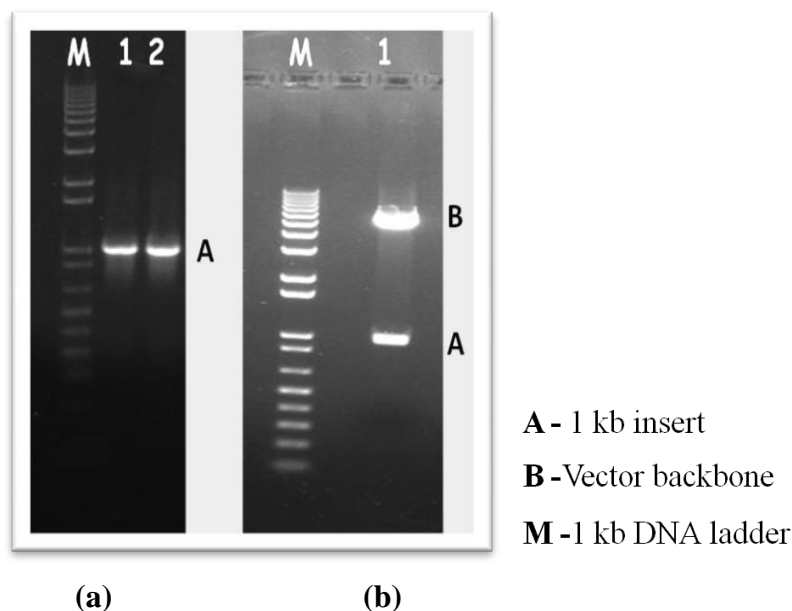
Fig 2.2. (a) PCR amplified band of *Sl*CGH1 on 1 % agarose gel (b) NdeI/XhoI restriction digested pET22b-*Sl*CGH1 showing 2 kb vector backbone (B) separated from 1 kb vector insert.

*Sl*CGH1 was expressed in *E. coli* BL21 as soluble cytoplasmic protein showing high level of expression, 200 – 250 mg/ l. This indicates post translational processing of *Sl*CGH1 by removal of N-terminal methionine in *E.coli*. Post translational modification for exposure of N-terminal nucleophilic residue is very important for Ntn-hydrolase activity. The importance of C1, R18 and D69 residues in the autocatalytic processing has been highlighted in the case of *Clostridium perfringens* BSH; mutants of these residues exhibited impaired post translation processing (Rossmann 2008).

a) **Nucleotide sequence**

```
   1 ATGAAAAAAG TACTACTAAG CACCACTATC GCAGCCCTAT TTGCCACTAC TATGGGCGCC
  61 GCCATAAACG CCGAAGCCTG TTCTCGTCTG GTGACAGAGA CCCAATACGG TACTATGTTA
 121 ATGCGTACCG CAGACTGGGT GAGCACAGCC CCCTTCGACG GTCACATGTC TGTCTTCCCT
 181 GTCGGCACCG AGCGCACCAT GCGTGGTCAG GTTGCCGAGT ACCAGCAGGC CATGACTAAG
 241 TGGCAGACTA AGTACCACAC CCTCTCTATC GAAGAGCATG GCGCCTTTGG CGGCCTGTCG
 301 GGTCAGACTT CTAACGAGAA GGGACTCAGC GTGATGGCGC TATCTCAGCA TGACAGCGAG
 361 CCTTATCTGA GCCAGCACAA AGATAATGGC GCGCCAGCGG TCAACACCGC CGACGTGGTG
 421 AGCTTTATCA CCGAGCGTTA TGCCACCACC GCCGAGGTGA AAGCGGCCCT GGATAATGGT
 481 GAGTTCCAAA TCGCCTGGGC CTCTGCGCCA AATGGTATGG AGCACGCTGC GCCGCTGCAC
 541 TACTCTGTGG TCGATGCCGA CGGTAACATC ATGCTTATCC AACTGGTGAA GGGCGGTGAG
 601 CAGAAGATCT ATCTGGGTGA TGCCGAATCG GATCTGCGCG TCAAGACTAA CGACCCGCTG
 661 CAGGAGAAGC ACAGAGAGTA TATGCAGCAG TTCGACCTTA AGGATCCTAG CGTTGCCACT
 721 AAAATGCCTT GGAGCATCGG CGGCCTGGAG CGTAACTCTC GCCTGCTGGC CATGTCGACT
 781 CACATGGACT TAGAAGGGCT GAGCTACACA GAGACAGTTG CTCGTCAGAA GGGCACCTTC
 841 GATGCGGCGG CGCTGGTGCC ATTCGGTGTG CAGGACCCTA AGACGGGCGA AGACTACCCA
 901 AGCTTCTTCA GCATGCAGTA CAACCTGGAT AACGGTGATA TCTGGTTCCG CAGCCTGATG
 961 AGTGGCAAGG AGATCAAGTT TAACCTGGAA GACACCAAGC AGTTCAAGAC GCCAATGCAT
1021 GCCGATATCA TGGCTCAGGT GGATAAAGGC GCTCAGACCA TCACTTGGTC TAAGATGTAA
```

b) **Amino acid sequence**

```
MKKVLLSTTIAALFATTMGAAINAEACSRLVTETQYGTMLMRTADWVSTAPFDGHMSVFPVGTERTMRGQVAEYQQAMTK
WQTKYHTLSIEEHGAFGGLSGQTSNEKGLSVMALSQHDSEPYLSQHKDNGAPAVNTADVVSFITERYATTAEVKAALDNG
EFQIAWASAPNGMEHAAPLHYSVVDADGNIMLIQLVKGGEQKIYLGDAESDLRVKTNDPLQEKHREYMQQFDLKDPSVAT
KMPWSIGGLERNSRLLAMSTHMDLEGLSYTETVARQKGTFDAAALVPFGVQDPKTGEDYPSFFSMQYNLDNGDIWFRSLM
SGKEIKFNLEDTKQFKTPMHADIMAQVDKGAQTITWSKMHHHHHH
```

Fig. 2.3. a) Nucleotide sequence of *Sl*CGH1, b) Corresponding amino acid sequence of *Sl*CGH1. Signal peptide sequence highlighted in red.

2.3.2. Purification and molecular weight determination of *Sl*CGH1

*Sl*CGH1was purified to homogeneity using Ni^{2+} affinity chromatography and purity was confirmed by the 37 kDa subunit molecular weight band of the protein on SDS-PAGE. The protein was stored at 4ºC at a concentration of 15 mg ml^{-1}.

From size exclusion chromatography, the native molecular weight of *Sl*CGH1 was deduced to be ~53 ± 4 KDa. Since a tetrameric protein of 37 KDa molecular weight would have a native mass of ~148 KDa, the probability of *Sl*CGH1 to be a tetrameric protein was considered very low and we suggest it could either be a monomer or a dimer. MALDI MS analysis of *Sl*CGH1 confirmed again a subunit molecular weight of 37772.4 Da and with an additional peak of 75494.4 Da which suggested that *Sl*CGH1 can be a dimer in solution (**Fig 2.3**).
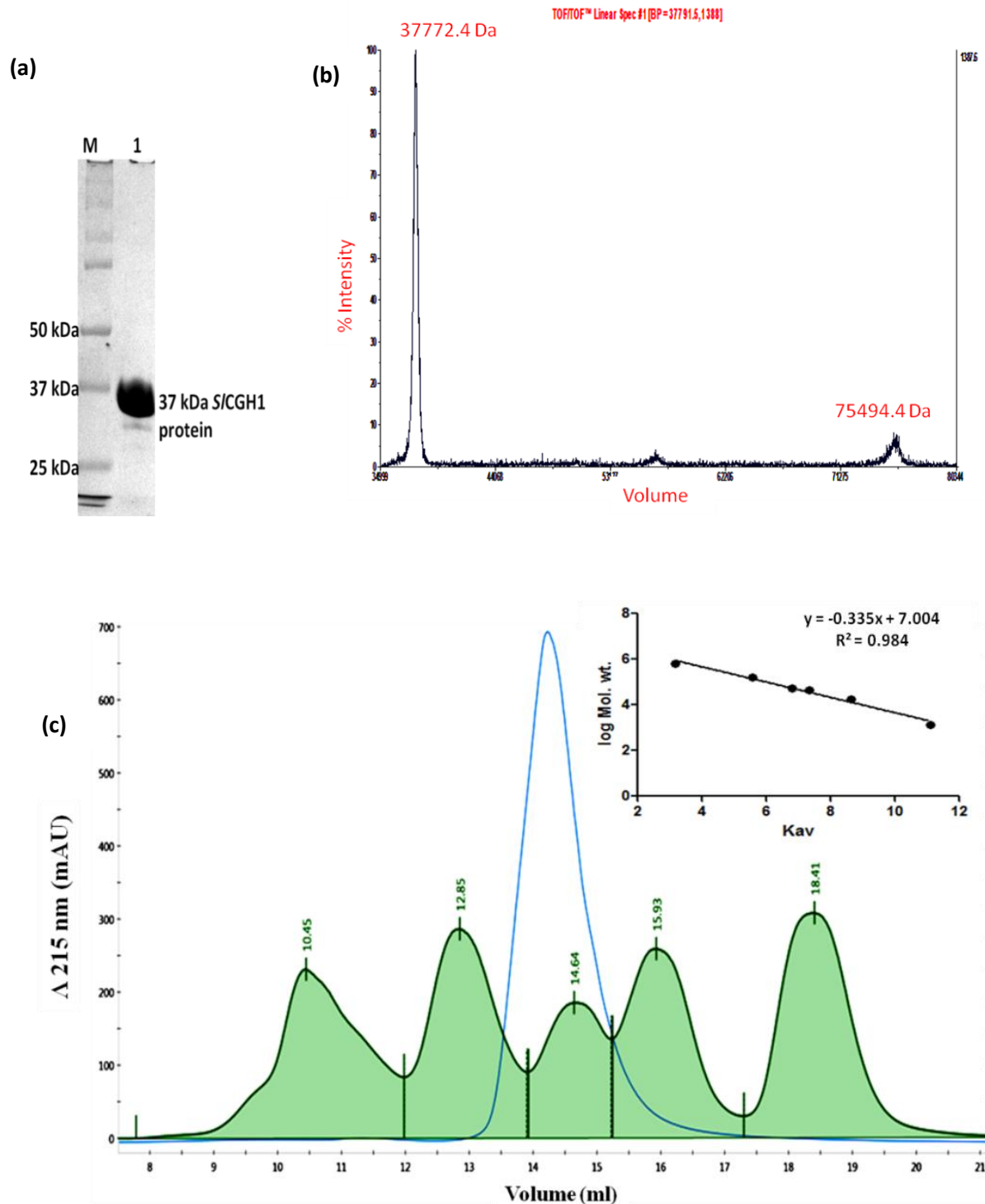
Fig 2.3. (a) Lane 1 - Purified *Sl*CGH1 of 37 KDa on 12% SDS-PAGE gel, M - BioRad protein molecular weight marker, MW in KDa. (b) MALDI MS spectrum showing subunit molecular weight of 37 KDa and a dimeric native molecular weight of 75 KDa. (c) Gel filtration chromatography profile of *Sl*CGH1.

2.3.3. Screening of enzyme activity

Based on sequence homology, *Sl*CGH1 has been annotated as a penicillin V acylase like enzyme in biological sequence databases (NCBI, UniProtKB). The substrate preference of *Sl*CGH1 was evaluated using standard CGH substrates - penicillins and bile salts. However, the enzyme was inactive on penicillins (Pen V and G) and bile salts. In previous report on *Lactobacillus plantarum* WCFS1, inactive BSH genes have been reported (Lambert et al. 2008). Such results reflect a gap in sequence based understanding of function of protein in general and a need to extensively explore functional perspectives of this enzyme family. We screened other amide containing substrates such as acyl homoserine lactones (AHL), p-nitrophenyl palmitate and casein for potential AHL acylase, lipase and protease activities respectively. The enzyme showed no lipase and protease activity; however, it was able to cleave the amide bond between lactone moiety and acyl chain in AHLs. We performed extensive bioassay on AHLs of various acyl chain lengths to gain further insights to the nature of degradation of this substrate by *Sl*CGH1.

Initial screening using biosensor strains *C. violaceum* CV026 and *A. tumefaciens* NTL4 (pZLR4) for C_8-HSL, C_{10}-HSL and 3-oxo-C_{10}-HSL degradation showed significant reduction in the production of extracellular pigments by these biosensors under AHL induction, indicating AHL degradation by the enzyme. The nature of AHL degradation by *Sl*CGH1 was investigated using High Resolution Mass spectroscopy (HR-MS) to detect the product formed after enzymatic treatment of 3-oxo-C_{10}-HSL (Huang et al.2003). The mass spectra of degradation products revealed peaks at m/z 124 and 102 corresponding to molecular formula of sodiated and protonated adducts of HSL, $C_4H_7O_2NNa$ and $C_4H_8O_2N$ respectively with error value of ±1.0-0.9 ppm (Fig 2). This result strongly supports that *Sl*CGH1 degrades AHLs via acylase activity. Similar peaks were detected in cases of C_8- HSL, 3-oxo-C_8-HSL, C_{10}-HSL, C_{12}-HSL and 3-oxo-C_{12}-HSL. The deviation in substrate specificity of *Sl*CGH1 could be attributed to the apparent low sequence homology of this enzyme (less than 20%) with the other well characterized CGHs.
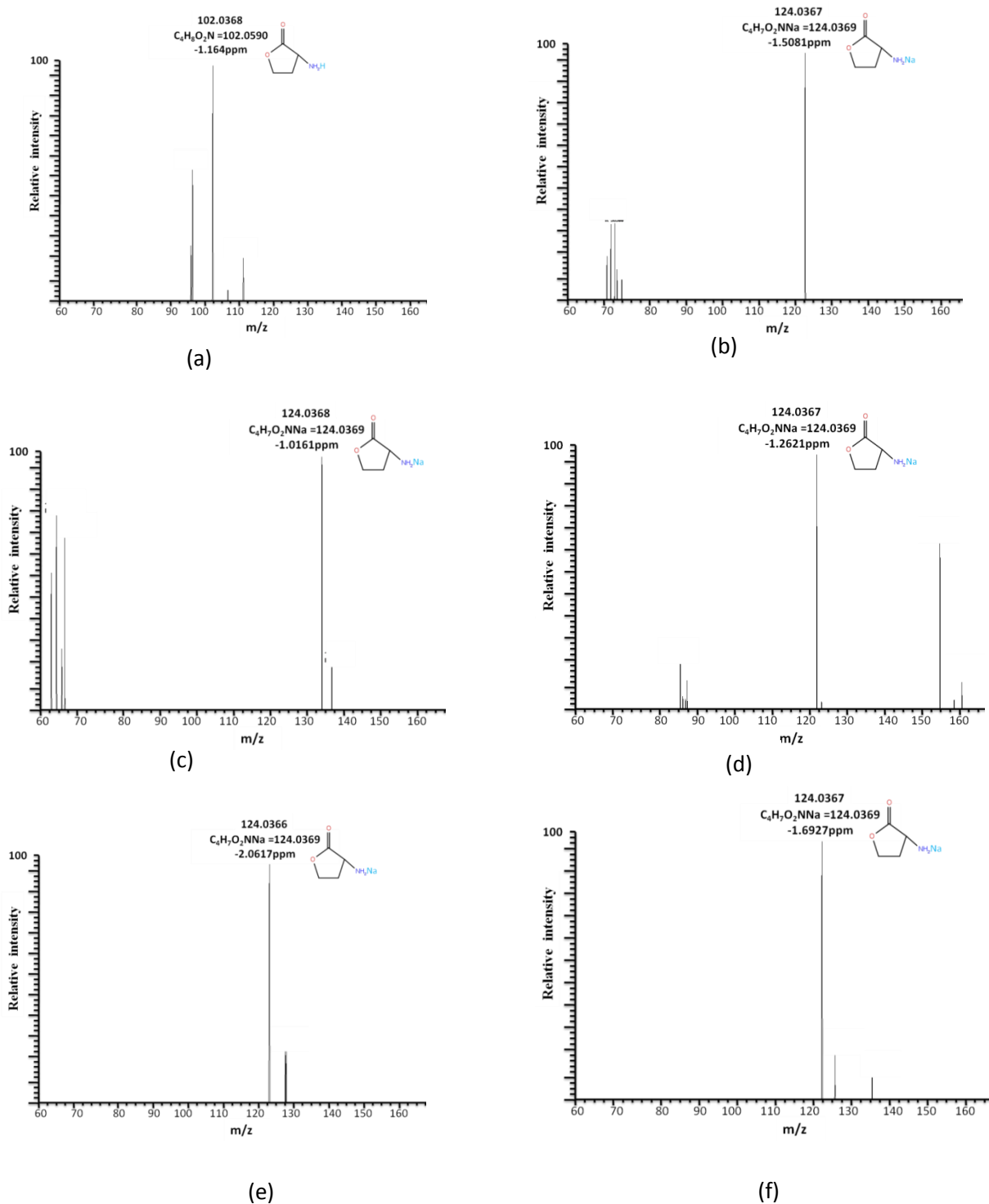
Fig 2.4. HR-MS spectrometry analysis of (a) C_8-HSL, (b) 3-oxo-C_8-HSL (c) C_{10}-HSL (d) 3-oxo-C_{10}-HSL (e) C_{12}-HSL and (f) 3-oxo-C_{12}-HSL, degradation by _Sl_CGH1. HSL peaks of m/z 102 (protonated) and 124 (sodiated) were present which confirm the acylase nature of _Sl_CGH1 activity. These results were within an error limit of 3 ppm.

During the course of this study, PVAs from Gram-negative bacteria, *Pectobacterium atrosepticum* and *Agrobacterium tumefaciens* have been reported to also degrade long chain AHLs (Avinash et al. 2017). However, *Sl*CGH1 differs from *Pa*PVA in that it shows activity exclusively on AHLs, and does not hydrolyze Pen V. The scarcity of reports on AHL degradation by both PVA and BSH is more precisely because AHL degradation assays have not usually been included in studies of this enzyme family.

Further quantification of AHL degradation was carried out using bioluminesence biosensors, *E. coli* pSB401 and *E. coli* pSB1075, covering C_6-HSL, C_8-HSL, 3-oxo-C_8 HSL, C_{10}-HSL, 3-oxo-C_{10}-HSL, C_{12}-HSL and 3-oxo-C_{12}-HSL as representative short, medium and long chain AHLs. The enzyme showed maximum activity against 3-oxo-C_{10}-HSL followed by C_{10}-HSL (**Table 2.2**) and was consistent in all three different AHL concentrations tested. At the lowest substrate concentration, *Sl*CGH1also showed good activity against C_8-HSL and 3-oxo-C_8-HSL. Overall the enzyme was most active on C_{10}-HSL, with weak yet confirmable activity towards AHLs with acyl chains shorter and longer than C_{10}. Among previously characterized AHL acylases, preference for medium and long chain AHLs (C_8 to C_{12}) has been more common. PvdQ and QuiP from *P. aeruginosa PAO1* (Huang et al. 2003; Huang et al. 2006), AhlM from *Streptomyces* sp. strain M664 (Park et al. 2005) and Aac from *Shewanella sp.* MIB015 (Morohoshi et al. 2008) showed preference for AHLs with acyl chain longer than 8 carbons. Unlike these examples, AiiC from Anabaena sp. PCC7120 (Romero et al. 2008) showed a wide range of specificity degrading short to long chains AHLs. Moreover, AiiD from a *Ralstonia* strains JX12B (Lin et al. 2003) has been reported to hydrolyze C_4-HSL as efficiently as it degrades 3-oxo-C_{12}-HSL, while it shows significantly low activity on 3-oxo-C_6-HSL.

Although *Sl*CGH1 activity was optimum against 3-oxo-C_{10}-HSL, the preference for 3-oxo substituted substrate was not reflected significantly in the rest of the AHLs tested. In the case of AhlM the preference for unsubstituted AHLs was reduced with increase in the number of acyl chain (Park et al. 2005). We attempted kinetic study of the AHL acylase activity using o-phthaldialdehyde (OPA) based detection of homoserine lactone (Park et al. 2005), but could not obtain consistent reading, probably due to inherent solubility problem associated AHLs.

Table 2.2. Result of AHL degradation by *Sl*CGH1 assayed using Lux-based biosensors (*E. coli* pSB401 and *E. coli* pSB1075). Residual RLU % = RLU of test/ RLU of control*100, expressed as mean± SD.

AHLs	Residual RLU %		
	15 µM	**25 µM**	**100 µM**
C_6 HSL	98.80 ± 4.51	107.50 ± 14	146.32 ± 1.67
C_8 HSL	40.09 ± 7	93.77 ± 10.15	111.50 ± 1.94
3-oxo C_8 HSL	66.77 ± 3.55	97.69 ± 9.8	107.91± 5.8
C_{10} HSL	17.32 ± 1.85	25.13 ± 6.8	95.10 ± 2.00
3-oxo C_{10} HSL	**0.42 ± 0.16**	**0.91 ± 0.2**	**33.06 ± 0.60**
C_{12} HSL	87.24 ± 5.21	95.80 ± 2.3	101.67 ± 0.34
3-oxo C_{12} HSL	93.79 ± 4.89	94.0 ± 5.5	100.56 ± 1.77

Majority of the AHL acylases reported so far, including AHL acylases from *P. aeruginosa*, *Ralstonia* sp.XJ12B (Lin et al 2003), *Streptomyces* sp. M664 (Park et al. 2005), *Shewanella* sp. MIB015, *Anabaena* sp. PCC7120 (Morohoshi et al. 2008) and *P. syringae* (Shepherd and Lindow 2009) belong to the Ntn- hydrolase superfamily. On the other hand, few other have also been shown to belong to α/β-hydrolase superfamily as in case of AHL acylases from *Ochrobactrum* sp. A44 (Czajkowski et al. 2011) and *Delftia* sp. VM4 (Maisuria and Nerurkar 2015). Such occurrence of similar function in members of different superfamilies may be an outcome of convergent evolution towards a common problem of bacterial signaling.

2.3.4. Site directed mutagenesis

The C1A and C1S mutants were expressed in *E. coli* BL21 as explained in case of native *Sl*CGH1. C1A showed a very poor expression level of 0.2-0.4 mg/ml whereas C1S showed a high level of expression similar to native *Sl*CGH1, 200 – 250 mg/ l, as soluble cytoplasmic protein. Bioluminescence assay results showed C1A could not degrade any of the AHLs, while C1S showed mild activity on long chain AHLs (**Table 2.3**). The dramatic lost of and reduction in activities strongly agree that Cys1 is the original nucleophile.

Table.2.3. Result of AHL degradation by C1A and C1S mutants assayed using Lux-based biosensors (*E. coli* pSB401 and *E. coli* pSB1075). Residual RLU % = RLU of test/ RLU of control*100, expressed as mean± SD.

AHLs	C1A	C1S
C_6-HSL	106 ± 6	129 ± 1
C_8-HSL	113 ± 8	130 ± 16
3-oxo-C_8-HSL	109 ± 6	92.3 ± 2.6
C_{10}-HSL	118 ± 8	89.3 ± 7
3-oxo-C_{10}-HSL	118 ± 4	89.1 ± 8
C_{12}-HSL	126 ± 4.5	85.2 ± 5
3-oxo-C_{12}-HSL	106 ± 5	84 ± 8.1

2.3.5. Fluorescence and Circular Dichroism studies

Conformational transitions and stability of *Sl*CGH1 were studied in thermal, chemical and pH mediated denaturation conditions using biophysical techniques such as fluorescence and CD spectroscopy. We drew comparisons between our results and that of previously reported tetrameric CGHs to understand any relationships between oligomeric state and stability of the protein. We could not perform functional stability at different thermal and pH conditions as AHLs are sensitive to alkaline pH and higher temperature (Yates et al. 2002).

The secondary structure composition for native *Sl*CGH1, calculated using CDPro software (http://sites.bmb.colostate.edu/sreeram/CDPro/) were 5.0% $\pm$ 2 % α helix; 38.4% $\pm$ 1.0 %, 22% $\pm$ 1.0 % turns and 33% $\pm$ 1.0 % unordered structures. However, in contrast to this calculated composition, experimental CD spectra shows a minimum at 208 nm and a shoulder at 218 nm indicating presence of prominent α and a small fraction of β structure in the enzyme. This variation in the calculated and experimental results might probably be due to the incompatibility of the data with the software used.

2.3.5.1. Temperature stability

The fluorescence intensity of *Sl*CGH1 showed emission maxima (λ_{max}) at 356 nm when excited at 295 nm. The protein showed a gradual decrease in fluorescence intensity with increase in temperature however without change in λmax (**Fig 2.5**). Each *Sl*CGH1 monomer contains 6 Trp and 11 Tyr residues. The exposure of such aromatic amino acids upon thermal induced denaturation of the protein may have resulted in quenching of the fluorescence intensity with increase in temperature. At 40ºC, the protein formed insoluble aggregates which are visually evident and also detected as an increase in Rayleigh scattering. A 50 % decrease in fluorescence intensity was observed at 45 ºC. Previous reports of fluorescence and CD studies of three CGHs, *Bl*BSH from *Bifidobacterium longum* (Kumar et al. 2006) *Pa*PVA from *P. atrosepticum* (Avinash et al. 2016) and *Ef*BSH, *Enterococcus faecalis* NCIM 2403 (Chand et al. 2015) suggested that tetrameric CGHs are stable upto 60ºC and aggregation reported upon thermal denaturation. Changes in *Sl*CGH1 secondary structure were correlated with the thermal denaturation by far-UV CD spectra (**Fig 2.4b**). Significant secondary structure changes were observed starting from 40ºC and the subsequent substantial change in minima (214 nm) suggested loss of β-sheet structure.

(a)   (b)
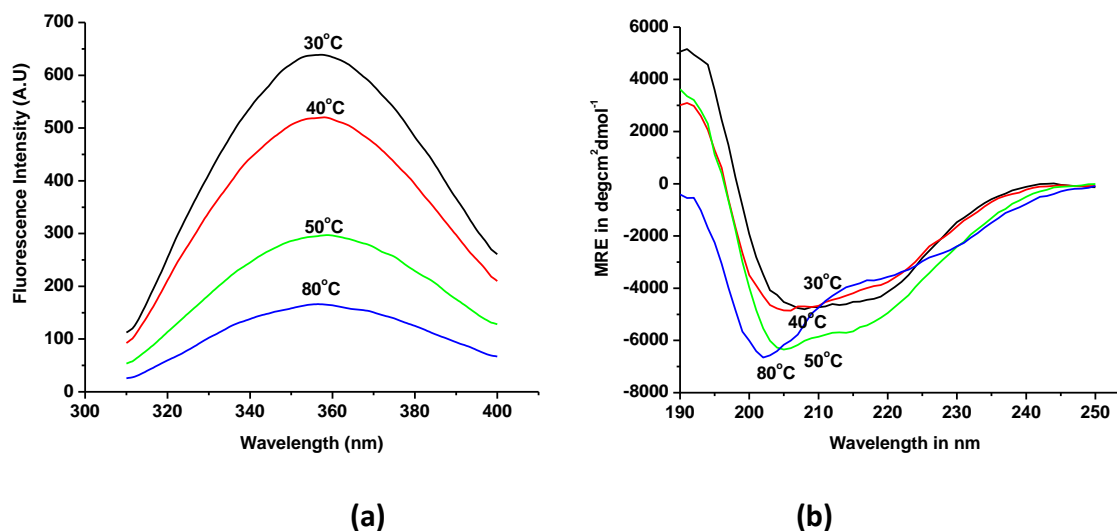
Fig. 2.4. (a) Fluorescence spectra of *Sl*CGH1 at different temperatures, (b) Far-UV CD spectra of *Sl*CGH1 at different temperatures, after 15 min incubation.

2.3.5.1. pH stability

Fluorescence intensities of *Sl*CGH1 decrease at extreme pH conditions, pH 1.0-2.0 and pH 9.0-10.0. Red shift was observed starting from pH 10 indicating increased unfolding of protein. The isoelectric point of *Sl*CGH1 is at pH 5 and the enzyme forms insoluble aggregate which is reflected in the absence of scattering in Rayleigh light scattering studies. At this pH the fluorescence intensity was found to be the lowest and the far UV spectrum also shows a total loss of secondary structures (**Fig. 2.5**). According to fluorescence and CD spectra, *Sl*CGH1 was stable between pH 6.0 to 9.0. BSHs, *Bl*BSH and *Ef*BSH were stable for the same range of pH and *Pa*PVA has been reported to be stable over a wide range of pH (3-11).

(a)   (b)
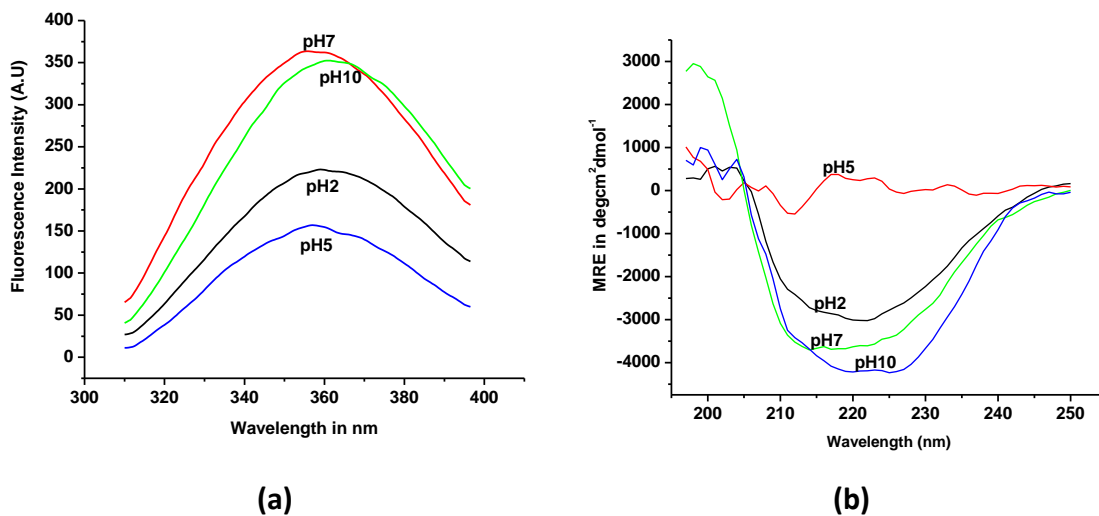
Fig. 2.5. (a) Fluorescence spectra of *Sl*CGH1 at different pHs, (b) Far-UV CD spectra of *Sl*CGH1 at different pH after 4 h incubation.

2.3.7. Chemical modification by Guanidine hydrochloride

Gdn-HCl at low concentrations has been known to confer effective charge shielding (Monera et al. 1994) and prevent aggregation in proteins (Hevehan and Clark 1997). When *Sl*CGH1was incubated with different concentrations of Gdn-HCl for 3 h, drastic secondary structure change and fluorescence red shift was observed from 1 M concentration of Gdn-HCl (**Fig. 2.6**). In previous reports of *Pa*PVA and in *Bl*BSH, the major secondary structure changes start after 1.5 M Gdn-HCl. We perform gel filtration chromatography of *Sl*CGH1 after incubation with 0.5 M

Gdn-HCl and found no dissociation of subunit as the enzyme was eluted in the same fraction as the native protein. The CD spectra of the Gdn-HCl treated protein when analyzed using the above programs did not show any significant change in the different structural elements compared to the native protein, while there was visible difference in the respective CD spectra.

(a)   **(b)**
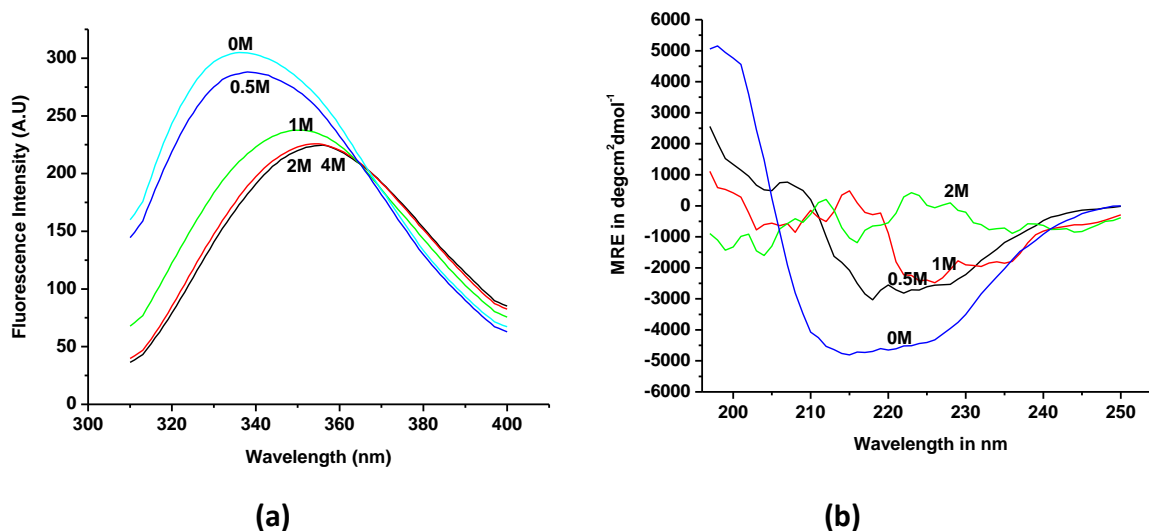
Fig. 2.6. (a) Fluorescence intensity at λ_{max} as a function of Gdn-HCl (incubation time 15 min)

(b) Far-UV CD spectra of *Pa*PVA after incubation at Gdn-HCl (0-2M) for 15 min.

2.4. Conclusion:

On the whole, biochemical characterization of *Sl*CGH1 from *Shewanella loihica* PV-4 reveals certain interesting features different from previously reported CGHs. The enzyme was a dimer and was found inactive on standard CGH substrates, PenV and bile salts, while hydrolyzing *N*-acylhomoserine lactones. *Sl*CGH1 showed best hydrolysis against the long chain AHL, 3-oxo-C_{10}-HSL followed by C_{10}-HSL while retaining mild activity on C_8 and C_{12}-HSLs. This unique activity profile suggests a possible adaptation to marine environment with special reference to microbial mat ecosystem where quorum sensing activity is considered highly active. While previously reported CGH of tetrameric nature are reported to be stable for a wide range of temperature and pH, *Sl*CGH1 was stable for a narrow range of temperature (till 35°C) and pH (6-9). This is the first report of a CGH with AHL acylase activity with total absence of conventional CGH activities. It would require further experimental studies to confirm whether AHLs are the true substrate or represent a promiscuous side of *Sl*CGH1.

Chapter 3

Structural features and mutational analysis of *Sl*CGH1

3.1. Introduction:

Protein molecules are tiny structures with an average size of ~2 nm. Three powerful techniques-X-ray diffraction, Nuclear magnetic resonance (NMR) and Cryo-Electron crystallography are currently employed for visualization of such small and sensitive biomolecule at the atomic level. X-ray crystallography has been the most important technique in the development of our understanding of enzyme structure and functions. Crystal structures with bound substrate and transition state analogs help to reveal the catalytic mechanisms of countless enzymes, thereby, assisting protein engineering, drug designing etc. The technique is based on the scattering of X-rays by the electrons in the molecules in a crystal. Enzyme molecules in the presence of some chemicals, probably a precipitant, arrange themselves in a crystal lattice where repeated structural motifs called unit cells are formed throughout the entire volume of the crystal in a periodic fashion. This enhances scattering of X-rays in selected directions amplifying the intensity of diffracted rays. A good crystal should have a high internal order and symmetry of the molecules.

Structures of many Ntn hydrolases and their ligand-bound complexes have been elucidated through X-ray crystallography. The 3D structural signature of Ntn hydrolases is the $\alpha\beta\beta\alpha$ structural fold common to all the members. The first structure to be solved among Ntn-hydrolases was that of penicillin G acylase from *E. coli* (Duggleby et al. 1995). So far, several structures of CGHs have been solved from both Gram positive and negative bacteria and there are distinct differences in structural features between the two groups (Avinash et al. 2017). Moreover, phylogenetic analysis revealed separate clusters for CGHs from Gram-negative and positive bacteria (Panigrahi et al. 2014). Structures of Gram positive CGH available are that of functional PVA from *Bacillus sphaericus* (*Bsp*PVA, Suresh et al. 1999), PVA from *B. subtilis* (*Bsu*PVA, Rathinaswamy et al. 2005), BSHs from *Clostridium perfringens* (*Cp*BSH, Rossocha et al. 2005) and *Bifidobacterium longum* (*Bl*BSH, Kumar et al. 2006). Gram negative CGH structure included BSH from *Bacteroides thetaiotaomicron* (*Bt*BSH, Stellwag and Hylemon 1976) followed by more recently solved PVAs from *Pectobacterium atrosepticum* (*Pa*PVA, Avinash et al. 2013) and from *Agrobacterium tumefaciens* (*At*PVA, Avinash et al. 2016).

All these above mentioned CGHs are from are soil (PVA) and gut (BSH) inhabiting bacteria, which leave a large number of other habitats to explore for novel CGHs with probable unique

features. Hence in an attempt to understand marine CGH, two CGH genes from marine isolate, *Shewanella loihica*-PV4, were selected the current study. This chapter describes the crystallization and determination of the three-dimensional structure of *Sl*CGH1. Unique structural features of *Sl*CGH1 as well as differences with respect to other choloylglycine hydrolases and AHL acylases has been discussed. Biological activity exhibited by *Sl*CGH1 has also been explained in terms of structural features.

3.2. Materials and methods:

3.2.1. Crystallization of *Sl*CGH1

Crystallization of a molecule is accomplished through trial and error experiments with different physico-chemical conditions to obtain the crystal and hence considered rather a 'non-scientific' approach. Protein crystals are usually sensitive and fragile as they are made up of 30-78 % solvent (Matthews, 1985) and the binding energies between protein molecules in the crystal is weak. The optimum conditions for growing a crystal are expected to give minimum or no disturbance to this molecular construct of the crystal.
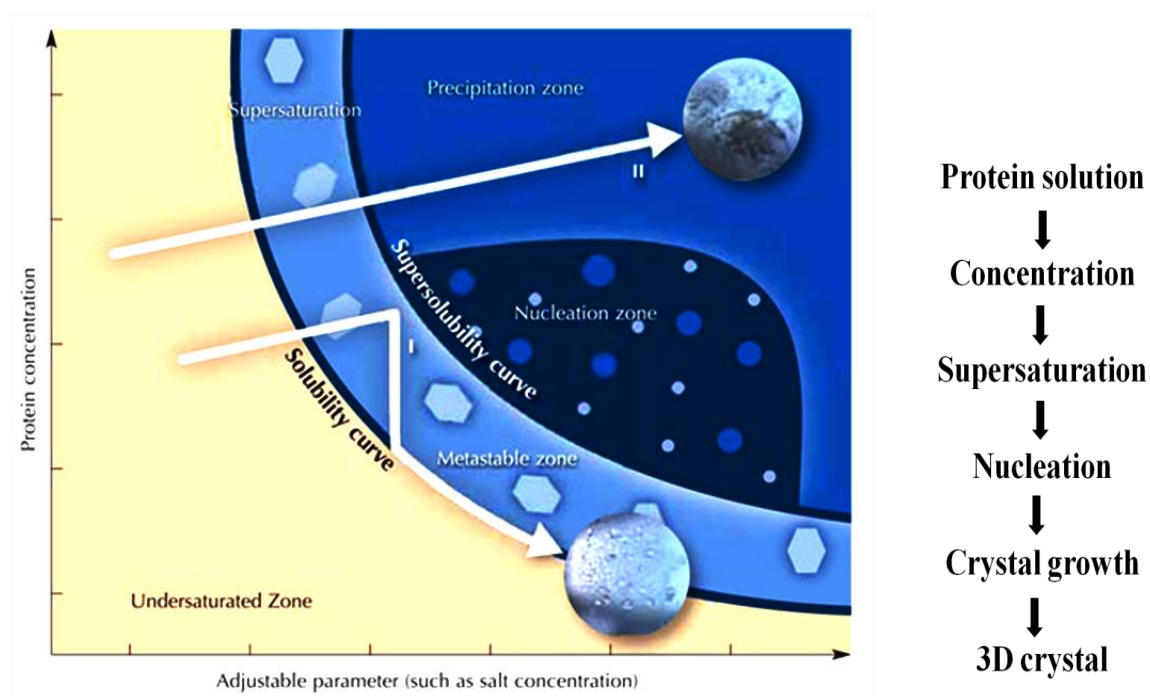
Fig.3.1. (Left) A classical illustration of solubility phase diagram of crystallization. The nucleation zone is where crystalline form occurs. The precipitation zone is characterized by

amorphous aggregates and metastable zone is the best zone for growth of large good quality crystals (Adapted from http://www.scienceinschool.org/2009/issue11/lysozyme); (Right) Stages of crystal formation.

Growing protein crystals is a reversible equilibrium phenomenon aiming at minimizing the free energy of the system (Weber 1991). At equilibrium conditions, the protein molecules in solution are fully solvated and upon introducing certain changes into the system which cause insufficiency of solvents to maintain full hydration of the molecules, the system equilibrium is disturbed. In this 'supersaturated' state, the system is thermodynamically driven toward a new equilibrium state and to minimize the free energy of the system. In order to do so, individual molecules lose rotational and translational freedom by forming many new stable, non-covalent chemical bonds. At this stage, the protein molecules appear either as amorphous precipitate or crystal nuclei in the solid phase. The overall process of crystallization occurs via three stages: nucleation, growth and cessation of growth. Nucleation is the first step where a 'nucleus' or a small protein aggregate starts the crystallization process and it is achieved at a supersaturation state. In the growth stage, further molecules or ions are added to the nuclei in a regular manner resulting in a good crystal. Thus the basic strategy of growing protein crystals is to attain the highest possible protein concentration without causing aggregation or precipitation. Supersaturation can be reached by different approaches including altering the buffer pH, temperature, protein concentration, dielectric constant of the medium, and precipitant concentration etc. If supersaturation is too high then too many nuclei appears resulting to the formation of numerous tiny crystals. In case of absence spontaneous nucleation, seeding strategy is used to grow the crystal where a small piece of single crystal called seed crystal is introduced to the system (Stura and Wilson SK, 1990).

Crystallization experiments can be set up using either vapor diffusion method or batch mode. Vapor diffusion is the more popular method requiring a simple set up and relatively small volume of protein. In this method, a mixture of protein solution and precipitating solution is placed in a sealed well, usually 24 or 96-well plates, which contains a reservoir of precipitating solution at higher concentration than that of the drop. Water vaporizes and diffuses from the drop to the reservoir solution, increasing the concentration of both protein solution and precipitating solution. Vapor diffusion setups are of two types - Hanging drop and Sitting drop methods (**Fig.3.2**).
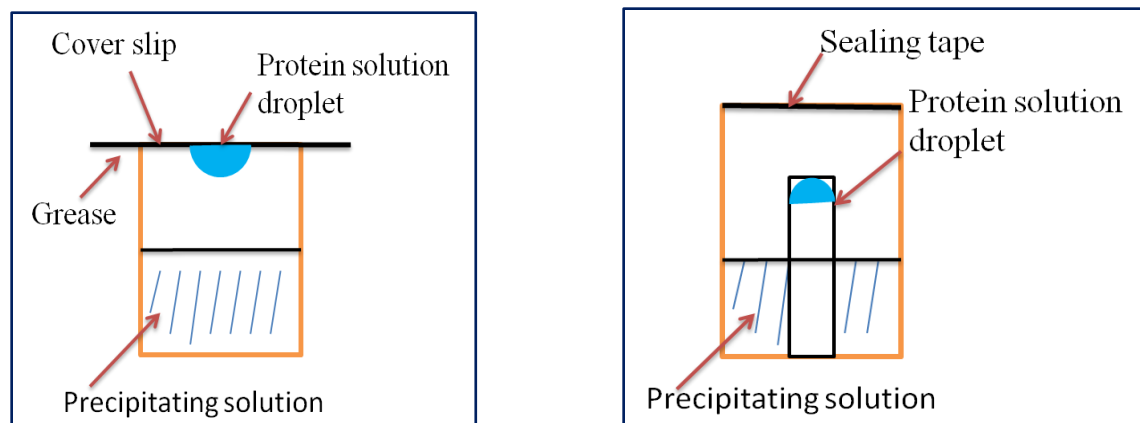
Fig: 3.2 a) Hanging Drop set up – protein solution drop faces or hangs up-side-down from the cover slip downward the bottom of the well. The system is sealed by using grease around the well's rim such that the cover slip sticks to the well. b) Sitting drop set up is a similar set up as hanging drop except the protein droplet is placed on a stage platform and sealing of system done using adhesive tapes.

Preliminary screening for identification of crystallization conditions for SlCGH1 was carried out using Mosquito Crystal Nanolitre protein crystallization robot (TTP Labtech, UK), by mixing 200 nl each of the protein and different precipitants on 96-well plates (sitting drop). Commercial crystallization screens used were Index, Crystal screen I and II and PEG-Ion screens (Hampton Research, CA, USA), Classic suite (Qiagen) and JCSG (Molecular Dimensions, UK). Manual optimization was carried out by vapor diffusion using the hanging drop approach. Purified SlCGH1 enzyme was concentrated to 15-20 mg/ml using Amicon ultra-15ml centrifugal filters (Millipore, USA). Protein was mixed with the precipitant solution in 1:1 ratio of protein and drop casted on the seliconized cover slip.

3.2.2   Cocrystallization of SlCGH1 and C2S mutant with substrates and inhibitors.

Structures of enzyme-substrate complex are valuable sources of information on catalysis mechanism, kinetics and regulation of enzyme activity etc. There are different strategies for obtaining crystals of protein–ligand complexes: (i) coexpression of the protein with the ligands of interest, (ii) use of the ligands during protein purification, (iii) cocrystallization and (iv) soaking the ligands into pre-existing crystals (Hassell et al. 2006). Cocrystallization is often the method of choice when the ligands have low solubility or the protein aggregates easily. Ligand is added to the protein to form a complex that is subsequently used in crystallization trials.

To understand the influence of structural changes on *Sl*CGH1 enzyme function, Native *Sl*CGH1 and C2S mutant was cocrystallized with different AHLs and PenV to capture an enzyme-substrate complex in all possible stages of the catalytic reaction. The protein was preincubated in solution with 1 to 2mM concentration of the substrates for 1 hour and kept for crystallization under the same crystallization condition as in case of Native *Sl*CGH1 and following the method explained in section 3.2.1.

3.2.3. X-ray diffraction and data collection

X-rays are electromagnetic radiation of wavelength about 0.1 – 10 nm (1 -100 Å) generated when electrons accelerated with high voltage collide with atoms of a metal target such as copper or molybdenum. Since this wavelength range falls in the same order of magnitude as the bond length between atoms (0.5 -1.6 Å) in the protein molecules, X-ray becomes the radiation of choice for protein crystal diffraction (Blow 2002). Although Individual atoms in a molecule can diffract X-rays, the diffraction pattern produced is too weak for analysis. Use of crystal solves this limitation by amplifying diffracted X-ray beam as it is composed of a number of molecules arranged in a regular or ordered manner (Rhode, 2000). A common laboratory source of X-ray is the rotating copper based anode generators producing X-rays of wavelength 1.542 Å. On the other hand, synchrotrons produce high intensity X-rays using particle accelerators and the diffraction pattern is recorded on charge coupled device (CCD) detectors. Synchrotron gives better resolution than laboratory sources, and data collection takes shorter time period (Drenth 1994).

Since X-ray can damage crystals through the generation of free radicals that cause chemical changes in the protein, crystals need cryoprotections and data collection is often carried out at low temperatures (-160°C). Commonly use cryoprotectants for proper freezing of the crystals includes glycerol, ethylene glycol, low molecular weight PEGs, MPD (2-methyl-2, 4-pentanediol), sucrose, xylitol etc. (Garman and Scheider 1997). The crystal is soaked in a cryoprotectant solution and mounted on a cryo loop made up of nylon fibre, followed by immediate flash-freezing in liquid nitrogen (Parkin and Hope 1998). After mounting the loop on the goniometer head, crystal is also kept exposed to a liquid nitrogen jet, to keep them frozen and aligned to the center of the X-ray beam. At synchrotron, an initial image is collected after an exposure of 2 seconds in order to determine the resolution limit, tentative space group and unit

cell parameters. The detector distance is adjusted based on the longest unit cell dimension, mosaic spread etc. The exposure time depends on the quality of the crystal and the oscillation range. Subsequent diffraction data are collected by exposing the crystal to the X-ray beam and rotating the crystal about an axis perpendicular to the X-ray beam over a given oscillation range varying from 0.2 - 0.3 degree per frame at the synchrotron or 0.5-1.0 degree per frame at the home source. The resulting diffraction pattern is recorded on a image-plate detector and the electron density of the molecule is calculated by inverse Fourier transform of phased reflection data.

The *Sl*CGH1 crystal was soaked in a cryoprotectant solution (30% glycerol) and mounted on a cryo loop, followed by immediate flash-freezing in liquid nitrogen. The diffraction data were collected at the Elettra synchrotron facility, Italy.

3.2.4. Data processing and Data Quality Statistics

Diffraction data of *Sl*CGH1 crystals were integrated and indexed using XDS (Kabsch 2010) and SCALA (Evans 2006). The important indicators of data quality are parameters such as resolution, I/σ (or signal-to-noise ratio), completeness, multiplicity and R_{merge}, overall and in the highest resolution shell. A diffraction experiment measures a large number of reflection intensities and some reflections have identical intensity owing to the symmetric nature of the crystal. The average number of measurements per individual, symmetrically unique reflection is called redundancy or multiplicity. High multiplicity (or redundancy) of measurements is desirable to assure accuracy of reflected intensity. The intensities of all repeatedly measured reflections have to be averaged after the application of all necessary corrections and appropriate scaling, a process known as 'scaling and merging', and the result expressed as a set of unique reflection intensities. Data completeness is the coverage of all theoretically possible unique reflections within the measured data set.

The merging R-factor ($R_{merge)}$ or symmetry R-factor (R_{sym}) is considered a global indicator of X-ray data quality and is given by the equation 3.1 (Blundell and Johnson, 1976):

$$R_{merge} = \frac{\sum\limits_{hkl} \sum\limits_{i=1}^{N} |\langle I_{hkl} \rangle - I_{hkl}(i)|}{\sum\limits_{hkl} \sum\limits_{i=1}^{N} I_{hkl}(i)}$$

.... (3.1)

where $I_{hkl}(i)$ is the ith measurement of the reflection with indices hkl and I_{hkl} is the mean value of the N equivalent reflections. For a good data set, R_{merge} will be less than 10%. Some authors considered R_{merge} as a limiting judge of data quality as it depends on the multiplicity of the data; higher the multiplicity, the higher R_{merge} becomes (Diederichs & Karplus 1997). It rather reflects data consistency better than quality of the reduced data.

3.2.5. Matthew's Number

Using the information of the crystal space group and unit cell dimensions, the number of molecules in the crystallographic asymmetric unit and the solvent content of the protein crystals can be estimated using the equations (**3.2 and 3.3**) by Matthew (1968) based on knowledge of the molecular weight of the protein:

Vm = unit cell volume * z / (MW * n) ----------- (3.2)

Vsolv = 1 – (1.23/ Vm)   ------------ (3.3)

where, *Vm* is the Matthew's number, n is the number of molecules per asymmetric unit and z is Avagadro's number; Vsolv is the fraction of unit cell volume occupied by solvent.

The Matthew's numbers for all the crystals of *Sl*CGH1 were estimated using the Matthews_coefficient module of the CCP4 suite by applying 37 KDa as the molecular weight of *Sl*CGH1 monomer.

3.2.6. Structure solution

The diffraction data collected and processed as described above is defined as the reciprocal space representation of the original crystal lattice. Each reflection is characterized by its amplitude and phase. The position of each diffraction 'spot' depends 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 amplitudes are directly calculated as the square root of the reflection intensities measured on the detector. However, information about reflection phases appears to be irretrievably lost as they cannot be directly provided by the diffraction experiment. This becomes the central problem of X-ray crystallography and known as the "phase problem".

Multi-wavelength Anomalous Dispersion (MAD), Isomorphous replacement (heavy atom method) and Molecular Replacement (MR) are widely used methods to solve the phase problem in protein structures. MAD method relies on data at different wavelengths from the same protein crystal, probably having a natural heavy atom or seleno methionine. Isomorphous replacement method makes use of heavy atom derivatives in a similar way. MR makes use of the fact that primary sequences of proteins determine their 3D structure and hence a known protein structure homologous to the proteins of interest can be used for phase determination. It is considered the easiest technique for structure solution (Rossman and Blow 1962). MR starts by placing the search molecule in the unit cell of the target crystal so as to account for the diffraction pattern. Proper orientation and precise positioning are achieved using two steps: rotation and translation. In the rotation step, the spatial orientation of the known and unknown molecule with respect to each other is determined, whereas in the next step, the translation needed to superimpose the now correctly oriented molecule onto the other molecule is calculated. The presence of non-crystallographic symmetry to improve phase information is also exploited in MR. Commonly used programs for molecular replacement are AMoRe (Navaza, 1994), MOLREP (Vagin & Teplyakov, 1997) and PHASER (McCoy et al., 2007; Storoni et al., 2004). The structure of SlCGH1was solved using Phaser ver 2.5.6 (McCoy et al. 2007) with the crystal structure of CGH from *Bacteroides thetaiotomicron* (PDB ID: 3HBC, 22 % sequence identity with SlCGH1) as the template.

3.2.7. Structure refinement and validation

The next step after getting the coordinates using a preliminary structure model is its refinement. Refinement fits the parameters of the model to achieve a closer agreement between observed and calculated structure factors and minimize errors in phase estimates. The model is refined 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 as well as temperature factors (Engh & Huber, 1991). At low resolution data, the hydrogen atom is omitted since they have only one electron and their influence on X-ray scattering is low. The residual or crystallographic factor provides an estimate of the correctness of the structure and is calculated using the formula below:

$$R - factor = \frac{\sum_{(h,k,l)} \| F_{obs}(h,k,l) | - | F_{calc}(h,k,l) \|}{\sum_{(h,k,l)} | F_{obs}(h,k,l) |}$$

............(**3.4**)

Where, F(h,k,l) = structure factor

However, R factor is not a reliable independent index of fit, as it is associated with the risk of exclusion or overfitting of the data. To overcome this problem, another a statistical parameter called the free R factor (R-free) was introduced (Brunger, 1992) for a better assessment of the fit between observed and calculated structure factors, given by the formula:

$$R\text{-free} = \frac{\sum_{hkl \in T} \| F_{obs} | - k | F_{calc} \|}{\sum_{hkl \in T} | F_{obs} |}$$

.... (**3.5**)

A set of reflections, commonly 5-10% of the observed data called 'test set T', are excluded from refinement. Refinement is carried out with the remaining reflections only, called 'working set W'. In practice, this means that models with serious errors can be indicated by a very high free R value (>0.40) irrespective of the value of the 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.

Model building of the *Sl*CGH1 structure was done using Coot and structure refinement by using Refmac5 software (CCP4 suite) using a restrained maximum likelihood refinement process.

3.2.8. Coordinate precision estimation and Structure analysis:

The estimate of the precision of atomic positions has been shown to depend on the magnitudes of B values or Isotropic temperature factors which represent the vibration of atoms in the lattice (Daopin et al., 1994). The program PROCHECK (Laskowski et al., 1993) was used to examine the stereochemistry of the final refined models, which calculates the phi-psi angles (Ramachandran plot, Ramachandran and Sasisekharan, 1968). Common coordinate problems which need fixing are unusual values of ω angles, eclipsed dihedral angles in side chains, abnormal van der Waals contacts and exceptionally high B-values as well as unpaired charged residues in the interior of the molecule.

All structural figures in this study were prepared using PyMol (The PyMOL Molecular Graphics System, Version 1.7.4, Schrödinger, LLC) or CCP4MG (McNicholas et al., 2011). Structural homology search was performed using the DALI server (Holm and Rosenstrom 2010). Analysis of oligomeric interactions was performed using PDBsum or PISA web servers.

3.2.9. Docking of 3-oxo-C_{10}-HSL, C_8-HSL and C_6-HSL to *Sl*CGH1 structure:

For the docking study, 3D conformation of 3-oxo-C_{10}-HSL, C_8-HSL and C_6-HSL were obtained from PubChem compound database (Bolton et al. 2008) with CID 5282982, CID 6914579, and CID 10058590 respectively. Docking these ligands was carried out using Grid based ligand docking program Glide (Freisner et al. 2006). Partial atomic charges of each ligand atom were determined from OPLS_2005 all-atom force field using LigPrep. A grid box of dimension 26x26x26 Å defining the binding site was generated and was centered on Cys1 residue. The ligands were docked flexibly using Glide's extra precision. The quality of docking in terms of free energy of binding is roughly estimated by using an empirical scoring function called GlideScore, which includes electrostatic, van der Waals interaction and other terms for rewarding or penalizing interactions that are known to influence ligand binding.

3.3. Results and Discussion:

3.3.1 Crystallization of *Sl*CGH1 and co-crystallization

Recombinant *Sl*CGH1was expressed and purified using techniques mentioned in Chapter 2. Preliminary trials for crystallization were performed using a high throughput crystallization robot on 96-well plates with different commercial standard screening kits containing pre-prepared solutions. The solutions for the grid screens were prepared by mixing of stock solutions of precipitant, salts and buffers and crystals were grown by vapour diffusion at 22°C.Thin needles like crystals appeared after 5 min in some conditions with PEG-salt combinations. Further manual optimization was done on 24-well plate using 20 mg/ml *Sl*CGH1 with three conditions selected from the preliminary trial screens (0.1 M Tris 8.5 + 25 % PEG1000; 0.2 M Sodium formate + 20 % PEG3350; 0.2 M Sodium sulphate + 20 % PEG3350). Diffraction quality crystals were obtained in an optimized condition consisting of 0.2 M sodium sulphate, 20 % PEG3350 and 5 % MPD, with a hanging drop containing 2 μl+2μl protein and the precipitant mix. The needle shaped crystal (**Fig.3.3 a**) diffracted at 1.8 Å resolution, and diffraction data was collected with 20% glycerol as cryoprotectant. The crystals belonged to $I12_1$ space group with a dimer per asymmetric unit (5494 amino acids, 172 water molecules).

(a)   (b)
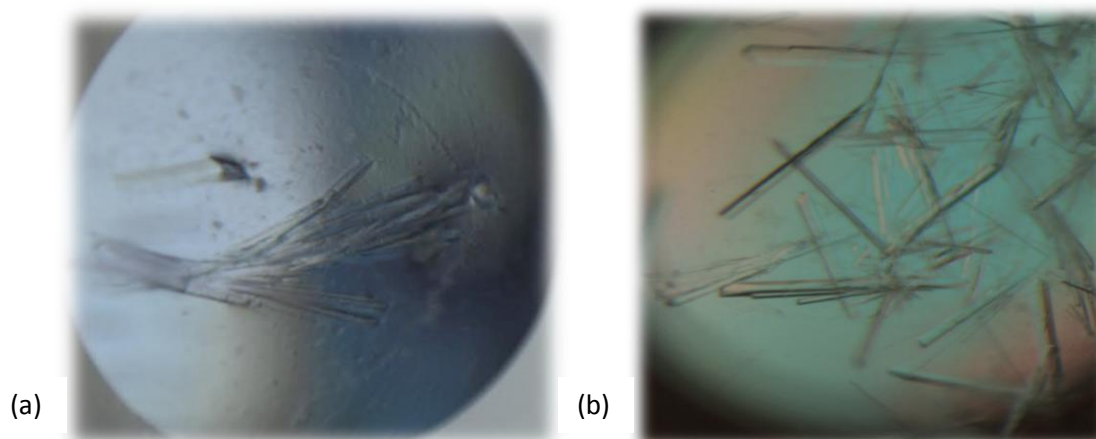
Fig.3.3. Crystals of (a) *Sl*CGH1 and (b) *Sl*CGH1-Cys1Ser grown under optimum crystallization conditions.

To gain insights into enzyme catalysis and function, we tried to crystallize the enzyme-substrate complex using both native *Sl*CGH1and site-directed (Cys1) mutants. We could not obtain

crystals of Cys1Ala owing to low protein concentration. Cys1Ser crystals were obtained (**Fig.3.3 b**) in similar condition as in case of native protein (0.2 M sodium sulphate, 20 % PEG3350 and 5 % MPD). Attempts were made to obtain complex structures with bound AHLs (C_6-HSL, C_8-HSL 3-oxo-C_8-HSL, C_{10}-HSL, 3-oxo-C_{10}-HSL and C_{12}-HSL), penV and bile salts employing both soaking and cocrystallization methods. Soaking *Sl*CGH1 crystals in any probable substrate solution resulted in cracking or dissolution of the crystals within seconds. Crystals prepared by cocrystallization of penV and bile salts with both native *Sl*CGH1and Cys1Ser, diffracted at low resolution 2.5 Å - 4 Å and showed high degree of mosiacity, which could not be solved. However, we were able to obtain good diffraction of Cys1Ser mutant cocrystalized with 3-oxo-C_8-HSL with a resolution of 1.5 Å.

Table 3.1: Data collection statistics of *Sl*CGH1 crystal and Mutant C1S cocrystalized with 3-oxo-C_8-HSL (Data in parantheses correspond to the outer shell).

X-ray diffraction and data collection	*Sl*CGH1	Mutant C1S
X-ray source	ELETTRA Synchrotron, ITALY	ESRF , France
Space group	I 12_1	C12_1
Temperature	100 K	100K
Resolution range	48.8-1.8 Å	50-1.5 Å
Unit cell parameters	a=99.4 , b= 52.34 , c= 140.07 ; α= γ = 90.00, β= 103.81	a=150.78 , b= 52.42, c= 99.20; α= γ=90.00, β= 115.96
Molecules per asymmetric unit	2	2
Matthews coefficient ($Å^3$ Da^{-1})	2.33	2.38
Solvent content (%)	48.58	48.39
Total no. of reflections	386710 (18943)	177994(15294)
No. of unique reflections	109332 (5258)	36355(3127)
Multiplicity	3.5 (3.6)	4.9(4.9)
Completeness (%)	97.6	100
Average I/σ(I)	13.8 (3.3)	12.2(5.1)
R-merge (all I+ and I-)	0.060 (0.0309)	0.084(0.276)

3.3.2. Structure solution, refinement and validation:

The structural model of *Sl*CGH1was built using MR approach of PHASER program, with the structure of CGH from *Bacteroides thetaiotamicron* (*Bt*BSH, PDB ID: 3HBC) as the search model. Further model building and refinement of the initial model of *Sl*CGH1 was accomplished using Coot and Refmac5 softwares (CCP4) respectively. In the last refinement cycle, solvent molecules were added at peaks of (Fo-Fc) density at 3σ, revealing sites with possible participation in hydrogen bonds with protein atoms. A total of 172 water molecules were fitted into the asymmetric unit.

Table 3.2 Summary of data collection and refinement statistics

Crystal system / Space group	*Sl*CGH1	Mutant C1S
Resolution range (Å)	1.8 Å	1.5 Å
Completeness (%)	97.5	100
Number of unique reflections	109332	36355
R$_{sym}$ (%)	6.0	8.4
Refinement statistics		
R$_{cryst}$	0.190	0.185
R$_{free}$	0.216	0.208
Number of protein atoms	5322	5445
Number of water molecules	172	175
Overall B value (Å^2)	15.487	20.656
Root mean square deviation		
Bond length (Å)	0.0253	0.219
Bond angle (º)	2.2400	2.0757
Dihedral angles	0.152	0.135
Ramachandran plot (% residues)		
Most favoured region	96.4 %	96.3 %
Allowed region	3.1 %	3.7 %
Outlier region	0.4 %	0

In the final cycles of refinement, the stereochemistry and the geometry of the models were checked using PROCHECK (CCP4). The final refined model has a calculated R-factor 0.190 and R-free 0.216. The geometric precision of the model was confirmed from Ramachandran plot (**Fig. 3.4**) which showed 97 % of the residues in the most favored or allowed regions of the plot.
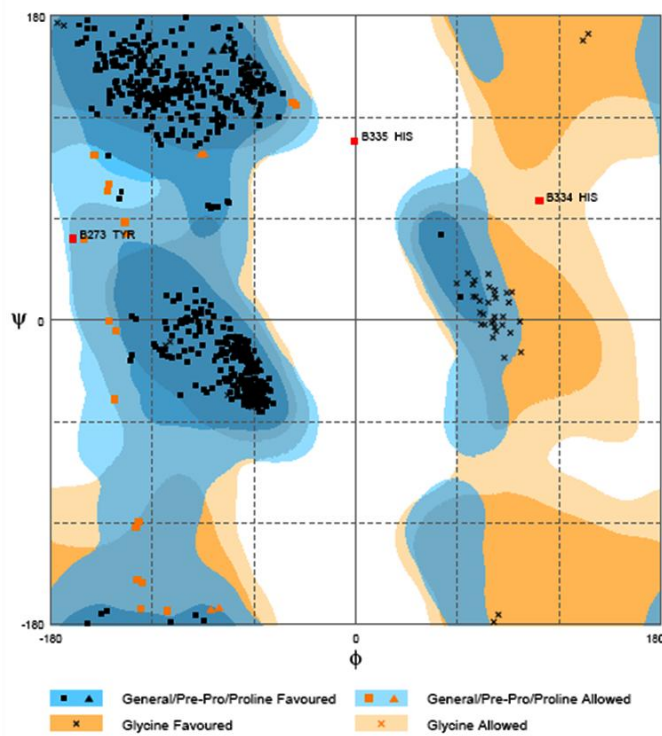
Fig.3.4. Ramachandran (phi-psi) plot of final refined *Sl*CGH1 model generated by PROCHECK. The His335 outlier is a part of C-terminal His-tag.

The structural model of Cys1Ser mutant was also built in similar way using MR approach with choloylglycine hydrolase from *Bacteroides thetaiotamicron* (*Bt*BSH, PDB ID: 3HBC) as the template. Refinement was carried out in similar manner as mentioned above. Total number of water molecules fitted into the asymmetric unit was 175. Calculated R-factor and R-free of the final refined model were 0.185 and 0.208 respectively. Ramachandran plot (**Fig. 3.5**) showed 100 % of the residues in the most favored region or partially allowed region of the plot indicating a high geometric precision of the model.
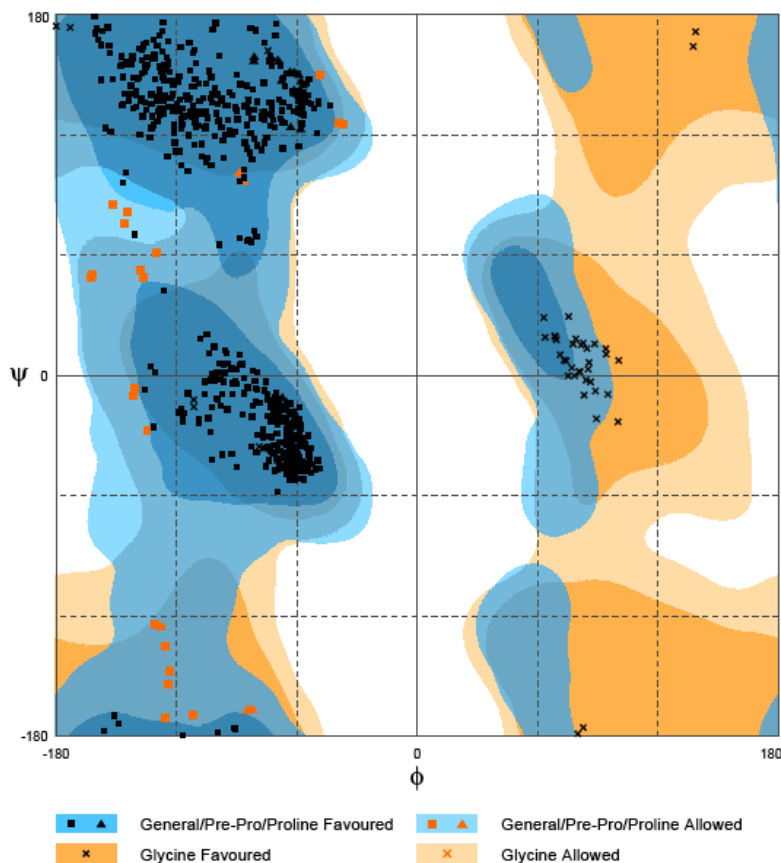
Fig. 3.5. Ramachandran plot of final refined C1S model.

3.3.3. Overall structure of apo- *Sl*CGH1

The apo-*Sl*CGH1 crystal structure reveals the presence of characteristic αββα-fold of the Ntn superfamily. The enzyme crystallized in a space group of I12$_1$ with two *Sl*CGH1 molecules per asymmetric unit. The two core β sheets of Ntn hydrolase host the active site and are spatially conserved. In *Sl*CGH1, β-sheet I consists of five β-strands while β-sheet II consists of nine β-strands (**Fig 3.8**). The topology of the strands in the first β-sheet is NH2-β1, β2, β12, β13 and β14; the second β-sheet consists of β11, β10, β3, β9, β8, β7, β6, β3 and β15. The angle between the strands of these two sheets is ~23.78º, previously reported CGH shows ~30° separation between the same (Kumar et al. 2006). The average B-factors calculated using the baverage module from CCP4 suit were 13.640 Å^2 for the main chain atoms and 17.288 Å^2 for side chains and water, respectively.
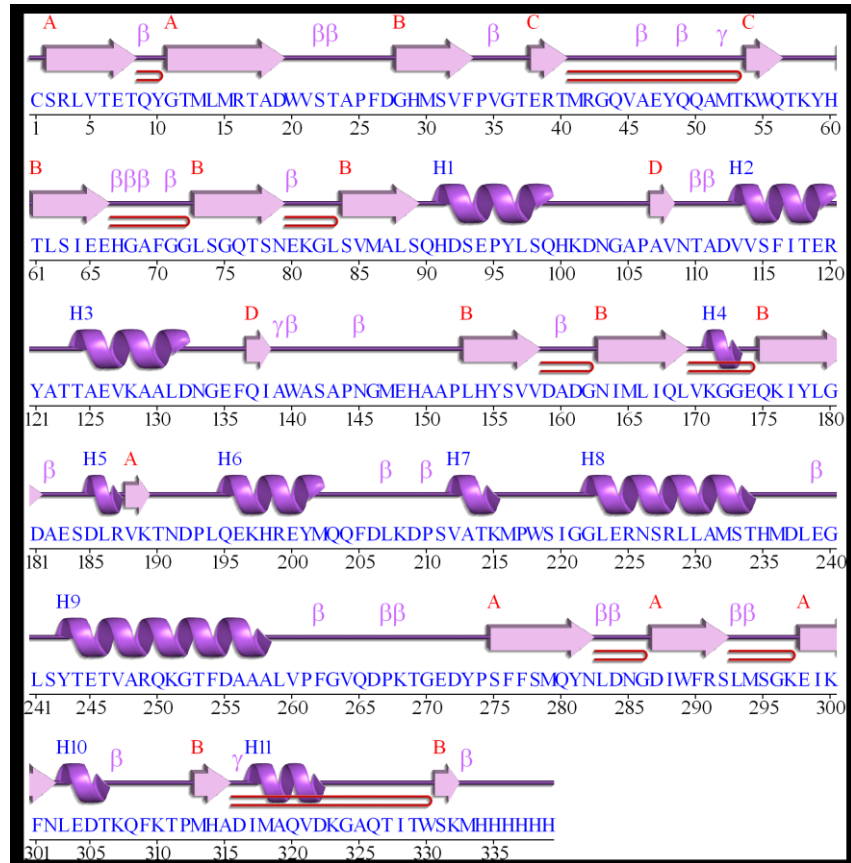
Fig. 3.6. Arrangement of secondary structure elements in *Sl*CGH1.
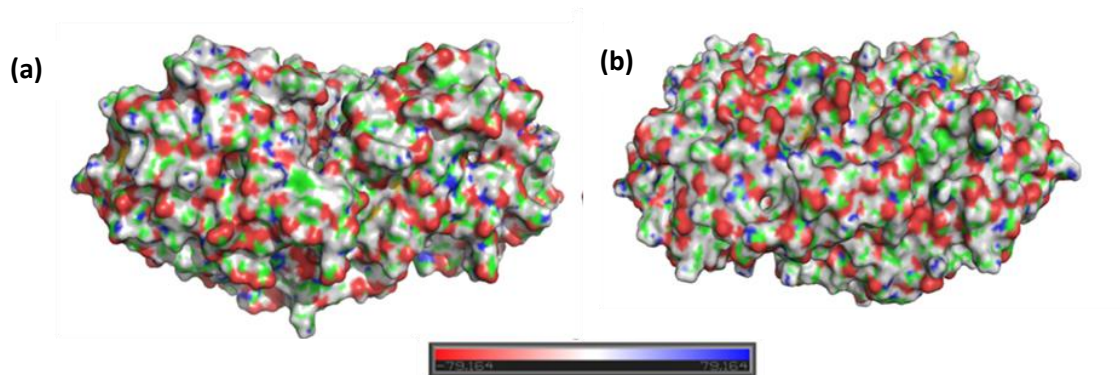
Fig. 3.7. Electrostatic surface representation of *Sl*CGH1 structure (a), rotation 180° (b).

(a)

(b)
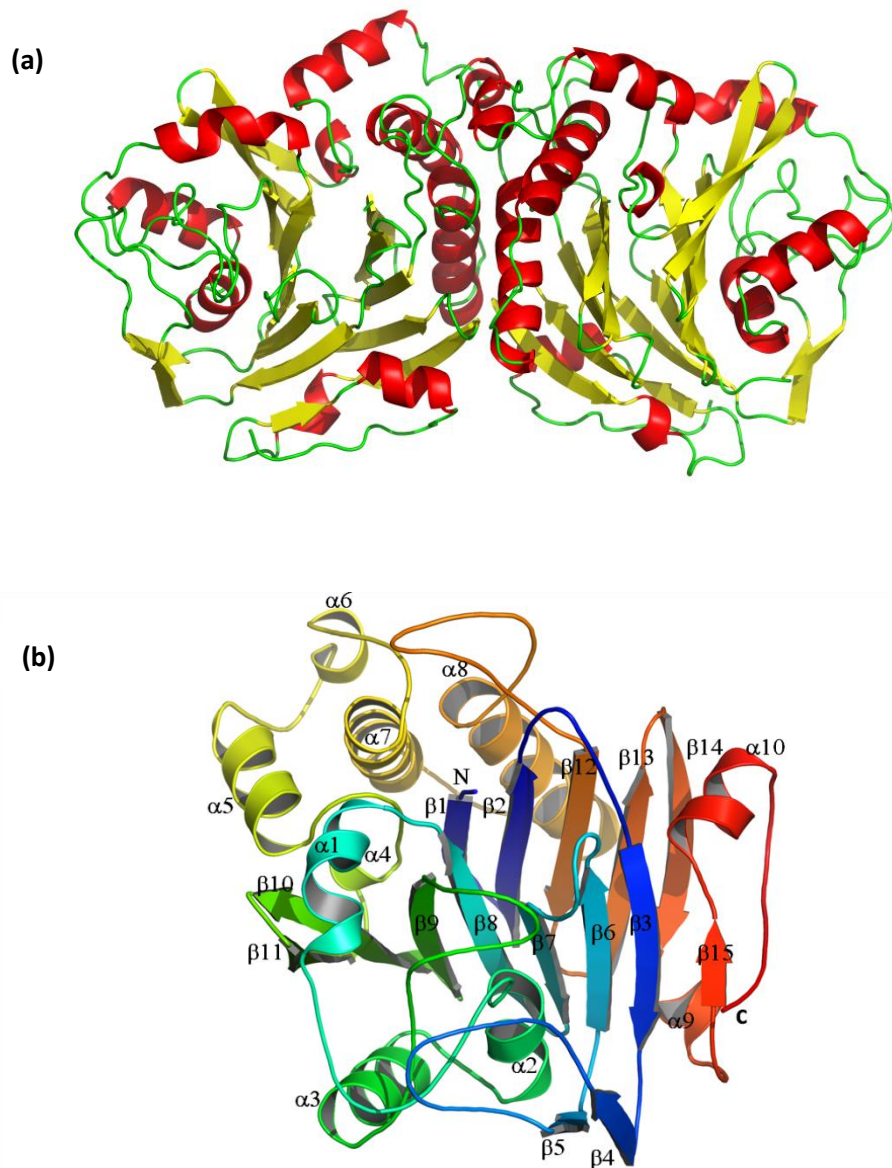
Fig 3.8. (a) Quaternary structure of *Sl*CGH1 dimer. (b) Cartoon representation of secondary structural elements in *Sl*CGH1 monomer color-ramped from N- to C-terminus (blue to red). The α-helices and the β-strands are labeled alphabetically.

Structure homology search using DALI server (Holm and Rosenstrom 2010) revealed similarity of *Sl*CGH1 to the CGHs from Gram negative bacteria, *Bacteroides* (PBD ID3HBC) and C chain of PVA of *P. atrosepticum* (PBD ID-4wl2) as with Cα RMSD values of 2.4Å and 2.6Å respectively. This was followed by CGHs from Gram-positive bacteria. The heterodimeric AHL acylase *P. aeruginosa*, PvdQ shares low structural similarity (7-8 %) with *Sl*CGH1.

Table 3.2. Structural alignment of SlCGH1 with other Ntn hydrolases using DALI server (Holm and Rosenstrom 2010).

PDB ID	Z-score	RMSD	L Ali	NRES	% identity	Protein
3HBC	33.8	2.4	295	309	20	CGH (*Bacteroides thetaiotamicron*)
4WL2	33.8	2.6	309	346	18	PVA (*Pectobacterium atrosepticum*)
2PVA	31.4	2.6	293	332	18	PVA (*Bacillus sphaericus*)
2OQC	31.1	2.7	293	317	19	PVA (*Bacillus subtilis*)
2BJF	31.1	2.7	291	328	16	BSH (*Clostridium perfingens*)
2HEZ	28.9	2.8	283	316	14	BSH (*Bifidobacterium longum*)
3FGT	16	3.3	231	344	7	Phospholipase B-like
1GK1	15.3	3.4	224	522	6	Cephalosporin acylase
4M1J	14.3	3.7	227	547	7	AHL acylase PvdQ
3M1O	13.9	3.7	224	551	10	Penicillin G acylase α subunit

3.3.4. Quaternary structure

*Sl*CGH1 adopts a stable dimeric quaternary structure in crystal form, which is also observed in solution. This finding is strikingly different as all characterized CGHs so far, both from Gram-positive and Gram-negative, are known to be homotetramer. PISA snalysis of protein interfaces suggested that the dimeric quaternary structure is stable in solution. The dimer interface in *Sl*CGH1 is formed by almost the same regions, (the alpha helices) of the Ntn fold as it occurs in other CGHs. Solvent accessible surface area of the dimer was 25577.2 Å^2 with a buried surface area of 2719.7 Å^2. The interface between the two monomers is spread across an area of 1359.8 Å ^{2} and is held by two hydrogen bonds along with 104 non-bonded contacts between 24 residues from both subunits (**Fig.3.9**). It is evident that the dimeric interface in *Sl*CGH1 is significantly reduced compared to CGHs from Gram-positive bacteria, and in fact even smaller than Gram-negative CGHs (*Pa*PVA).
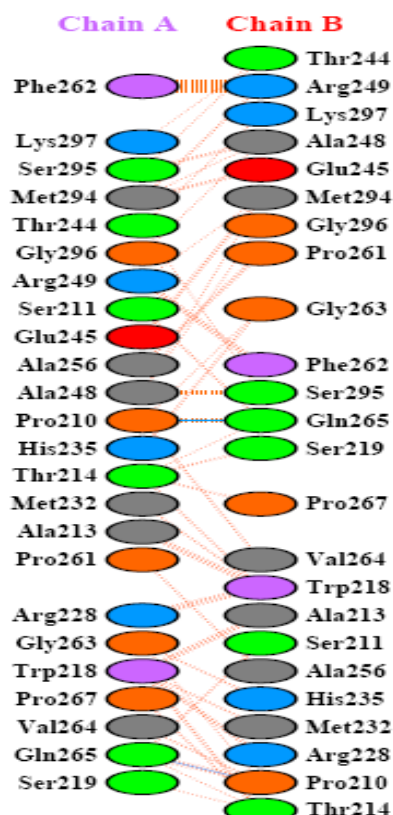
Fig 3.9. Interface residues interaction between two monomers of a *Sl*CGH1 dimer generated by PDBsum.

⎯⎯⎯ Hydrogen bond

⎯⎯⎯ Non-bond contact

In Gram positive CGHs, tetramerization is enabled and stabilized by an assembly motif of 20-25 residues that extends to the opposite monomer (Avinash et al. 2016, Panigrahi et al. 2015). In their Gram negative counterparts, the tetrameric loop is considerably shorten by deletion mutation reducing the interaction (**Fig. 3.10**). However, despite having a shorter tetramerization loop, Gram negative CGHs have been observed to retain their homotrameric state. To explain the undelaying molecular and structural features responsible for *Sl*CGH1 unique oligomerisation different from other Gram negative CGHs, we compared the *Sl*CGH1 and *Pa*PVA structures to look for deviations in sequence and structural features important for oligomerization. Two alpha helices upstream and downstream of the tetramerization loop in CGHs corresponding to, α5 (Gln195- Gln204) and α7 (Gly220-Thr234) in *Sl*CGH1 are important for interactions between the dimers. In *Pa*pva, two polar contacts (Tyr196:Asp214 and Gln199: Gln189) on either sides between these α-helices from adjacent subunits of the tetramer and non-polar contacts involving Val204 and Met205 from the tetramerization loop of opposite subunits constitute the main oligomeric interactions. Similar interactions were absent in *Sl*CGH1 as the above mentioned interacting residues were not positionally conserved. Another distinct feature in

SlCGH1 is the prominent occurrence of polar residues in the exposed α5 helix, while the corresponding region in other CGHs usually comprises a mixture of polar and non-polar residues. The change to a completely polar region may explain the feasibility of α5 to remain exposed to solvents thereby lessening the need to interact with another dimer for protection of non-polar residues. Also α5 is shorter in SlCGH1 by four residues, which increases the distance between potential interacting residues in a tetrameric alignment. Moreover, in addition to a short tetrameric loop, the presence of an additional short alpha helix (Val212-Lys215) in SlCGH1 formed in the middle of the loop reduces its flexibility. All these factors together explain the absence of tetramerization in SlCGH1. If we assume dimeric SlCGH1 to be a more ancient homolog, then the tetrameric forms might have evolved for more complex regulatory mechanisms and higher stability. This is supported by SlCGH1 being less stable than other tetrameric CGHs in CD and fluorescence studies in chapter2.

Fig. 3.10. Superposition of the monomers of SlCGH1 (green) with (a) PaPVA (red), (b) BlBSH (magenta) and (c) BspPVA (Orange). Locations of the assembly motif are encircled.

```
SlCGH1   CSRLVTETQYG---TMLMRTADWVSTAPFDGHMSVFPVGTERTMRGQVAEYQQAMTKWQTKYHTLSIEEHGAFG-71
PaPVA    CTRFVYLDPHNPDYPITARSMDWADDT--ETNLWIFPQ----ELKRSGGAGQYSLEWTSKYGSVIASAFDGRKGM69
BtBSH    CTRAVYLGPDR--MVVTGRTMDWKEDI--MSNIYVFPR----GMQRAGHNKEKTVNWTSKYGSVIATGYD-----62
BspPVA   CSSLSIRTTDD--KSLFARTMDFTMEP--DSKVIIVPR----NYGIRLLEKENVVINNSYAFVGMGSTDITS---64
BsuPVA   CTSLTLETADR--KHVLARTMDFAFQL--GTEVILYPR----RYSWN-SEADGRAHQTQYAFIGMG-RKLGN---62
BlBSH    CTGVRFSDDEG--NTYFGRNLDWSFSY--GETILVTPR----GYHYDTVFGAGGKAK-PNAVIGVGVVMADR---63
CpBSH    CTGLALETKDG--LHLFGRNMDIEYSF--NQSIIFIPR----NFKCVNKSNKKELTT-KYAVLGMGTIFDDY---63
LbBSH    CTAITLNGNS----NYFGRNLDLDFSY--GEEVIITPA----EYEFKFR-KEKAIKN-HKSLIGVGIVANDY---60
         *:            *. *                  :  . *                           :.

SlCGH1   GLSGQTSNEKGLSVMALSQHDSEPYLSQHKDNGAPAVNTADVVSFITERYATTAEVKAALDNGEFQIAWASAPNG146
PaPVA    ASTTDGVNEKGLAANVLWLAESEYPKTKPTA-KKPGLSVAAWAQYVLDNFATVDEAVKSLQQEKFILVTKQVE-G142
BtBSH    IGTCDGMNEKGLVASLLFLPESVY--SLPGD-TRPAMGISIWTQYVLDNFATVREAVDEMKKETFRIDAPRMP-N133
BspPVA   PVLYDGVNEKGLMGAMLYYATFATYADEPKK-GTTGINPVYVISQVLGNCVTVDDVIEKLTS--YTLLNEANI-I135
BsuPVA   ILFADGINESGLSCAALYFPGYAEYEKTIRE-DTVHIVPHEFVTWVLSVCQSLEDVKEKIRS--LTIVEKKLD-L123
BlBSH    PMYFDCANEHGLAIAGLNFPGYASFVHEPVE-GTENVATFEFPLWVARNFDSVDEVEETLRN--VTLVSQIVP-G134
CpBSH    PTFADGMNEKGLGCAGLNFPVYVSYSKEDIE-GKTNIPVYNFLLWVLANFSSVEEVKEALKN--ANIVDIPIS-E134
LbBSH    PLYFDAINEDGLGMAGLNFPGNAYYSDALEN-DKDNITPFEFIPWILGQCSDVNEARNLVEK--ININ-LSFS-E120
              :   ** **    *                   ::       :.   . :   .

SlCGH1   MEHAAPLHYSVVDADGNIMLIQLVKGGEQKIYLGDAESDLRVKTNDPLQEKHREYMQQFDLKD----------PS211
PaPVA    QKRLATLHLSLSDSSGDSAIIEYIDGKQVIHHS----KNYQVMTNSPTFDQQLTLNAYWDQIG---------GNV204
BtBSH    GGPESTLHMAITDETGNTAVIEYLDGKLSIHEG----KEYQVMTNSPRYELQLAVNDYWKEVG---------GLQ195
BspPVA   LGFAPPLHYTFTDASGESIVIEPDKTGITIHR-----KTIGVMTNSPGYEWHQTNLRAYIGVTPNPPQDIMMGDL205
BsuPVA   LDTVLPLHWILSDRTGRNLTIEPRADGLKVYD-----NQPGVMTNSPDFIWHVTNLQQYTGIRPKQLESKEMGGL193
BlBSH    QQ-ESLLHWFIGDGK-RSIVVEQMADGMHVHH-----DDVDVLTNQPTFDFHMENLRNYMCVSNEMAEPTSWGKA202
CpBSH    NIPNTTLHWMISDITGKSIVVEQTKEKLNVFD-----NNIGVLTNSPTFDWHVANLNQYVGLRYNQVPEFKLGDQ204
LbBSH    QLPLAGLHWLIADRE-KSIVVEVTKSGVHIYD-----NPIGILTNNPEFNYQMYNLNKYRNLSISTPQNTFSDSV189
             **  .*       ::            . : **.**       :

SlCGH1   VATKMPWSIGGLERN-------SRLLAMSTHMDLEGLSYTETVARQKGTFDAAALVPFGVQDPKTGEDYPSFFSM279
PaPVA    MLPGTNRAADRFVRASFYVKNVNPNKLIPGVAEKGKIEKDKADLATAFSIIRNASVPYGYSLPDMPNIASTRWRT279
BtBSH    MLPGTNRSSDRFVRASFYIH------AIPQTAD------AKIAVPSVLSVMRNVSVPFGINTPEKPHISSTRWRS258
BspPVA   DLTPFGQGAGGLGLPGDFTPSARFLRVAYWKKYTEKAKNETEGVTNLFHILSSVNIPKGVVLTNEGKTDYTIYTS280
BsuPVA   ALSAFGQGLGTVGLPGDYTPPSRFVRAVYLKEHLEPAADETKGVTAAFQILANMTIPKGAVITEEDEIHYTQYTS268
BlBSH    SLTAWGAGVGMHGIPGDVSSPSRFVRVAYTNAHYPQQNDEAANVSRLFHTLGSVQMVDGMAKMGDGQFERTLFTS277
CpBSH    SLTALGQGTGLVGLPGDFTPASRFIRVAFLRDAMIKNDKDSIDLIEFFHILNNVAMVRGSTRTVEEKSDLTQYTS279
LbBSH    DLKVDGTGFGGIGLPGDVSPESRFVRATFSKLNSSKG-MTVEEDITQFHILGTVEQIKGVNKTESGKEEYTVYSN263
            .  .       *         .: :

SlCGH1   QYNLDNGDIWFRSLMSGKEIKFNLEDTKQFKTPMHADIMAQVDKGAQTITWSKM--------------333
PaPVA    VVDHKSLQYFFESAVSPNIFWVDLKKINFAPRGGSAAKLDLGPNQSTIYSGQASGHFKPAQPFEFAGL347
BtBSH    VSDQKNKVYYFESTLTPNLFWLDLKKIDFSPKAG-VKKLSLT--KGEIYAGDAVKDLKDSQS------317
BspPVA   PVAAMCAQSKNYYFKLYDNSRISAVSLMAENLNSQDLITFEWDRKQDIKQLNQVNVMS---------338
BsuPVA   VMCNETGNYYFHHYDNRQIQKVNLFHEDLDCLEPKVFSAKAEESIHELN-----------------317
BlBSH    GYSSKTNTYYMNTYDDPAIRSYAMADYDMDSSELISVAR--------------------------316
CpBSH    CMCLEKGIYYYNTYENNQINAIDMNKENLDGNEIKTYKYNKTLSINHVN-----------------328
LbBSH    CYDLDNKTLYYTTYENRQIVAVTLNKD-KDGNRLVTYPFERKQIINKLN-----------------311
            ..:                :
```

Fig.3.11. Sequence alignment of CGH enzymes. *Sl*CGH1, PVAs from *Pectobacterium atrosepticum* (*Pa*PVA), *Bacillus subtilis* (*Bsu*PVA) and *B. sphaericus* (*Bsp*PVA); BSHs from *Bacteroides thetaiotamicron* (*Bt*BSH), *Bifidobacterium longum* (*Bl*BSH) and *Clostridium perfringens* (*Cp*BSH) and *Lactobacilus salivarus* (*Lb*BSH).The alignment highlights conserved active site residues (red and yellow), loops (grey) flanking the active site and assembly motif (Magenda), and α helix involved in oligomeric interactions between two monomeric subunits (cyan).

3.3.5. Active site of _Sl_CGH1 shows Conservation of CGH active site

The active site of _Sl_CGH1 was considered from two perspectives – the conserved catalytic residues and the loops surrounding the site. For comparative analysis, the _Sl_CGH1 structure was superimposed with PVAs from Gram negative _Pectobacterium atrosepticum_ (_Pa_PVA, Avinash et al. 2013) and Gram-positive _Bacillus sphaericus_ (_Bsp_PVA, Suresh et al. 1999), as well as BSHs from Gram-negative _Bacteroides thetaiotaomicron_ (BtBSH, Stellwag &Hylemon 1976) and Gram-positive _Bifidobacterium longum_ (_Bl_BSH, Kumar et al. 2006). _Sl_CGH1 share active site residues conservation with known CGHs with the exception of the oxyanion forming residue, Ser89 which corresponds to Tyr82 and Trp23 in PVA (Suresh et al. 1999; Avinash et al. 2016) and Asn81 in BSH (Rossocha et al. 2005) (**Fig. 3.12**). As it is the backbone amide that participates as electron acceptor during the formation of the intermediate oxyanion hole, this variation is considered of the least impact on the nature of active site (Rossocha et al. 2005). Other residues involved in catalytic mechanism or substrate binding, including Arg16, Asp19, Trp20, Asn191 and Arg225 are strictly conserved. The active site is dominantly hydrophobic due to the flanking aromatic residues including Trp20, Phe26, and Phe70 from the surrounding loops.
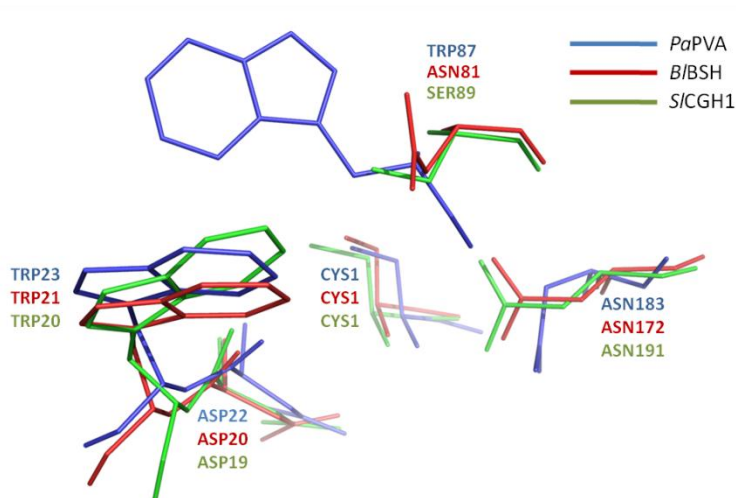
Fig. 3.12. Superimpose active sites of Gram negatives CGHs from _Sl_CGH1, _Pa_PVA and Gram positive _Bl_BSH.

In CGH, four conserved loops flanking the active site are important for substrate binding (Kumar et al. 2006; Panigrahi et al. 2014). It is also observed among known CGHs that the differential folding and conformations of these loops is responsible for variation in size and properties of their binding site pockets (Panigrahi et al. 2014). In *Sl*CGH1, these loops corresponds to W20 to F26 (loop1), H67-Q76 (loop 2), K128-P152 (loop3) and A257 to F276 (loop4). Although the core active site residues remain largely conserved in *Sl*CGH1, the active site loops, loop 2 in particular show significant deviations in position and length compared to other CGHs. Loop 2 displays an inward shift, moving closer to Cys1 by about 5 to 9 Å (**Fig.3.12**), consequently shrinking the active site to a much smaller volume (79 Å^3, SiteMap calculation) than in any other CGHs reported so far (153- 718 Å^3, Panigrahi et al. 2014). It is possible that the presence of four Gly residues in *Sl*CGH1 loop2 may have imparted extra flexibility to orient inward, while other CGHs possess only one or no glycine. In addition to reducing the active site volume, this also creates a narrow, close ended U-shaped architecture of the binding pocket as loop 2 and 3 come closer on one end of the pocket. We suggest the narrow active site might allow entry of substrates with a linear geometry such as AHL, while blocking entry of bulky substrates such as Pen V and bile salts and hence the inactivity on these substrates. The buried active site pocket showed a small globular hydrophobic patch which also might affect the substrate selectivity (**Fig. 3.13**)
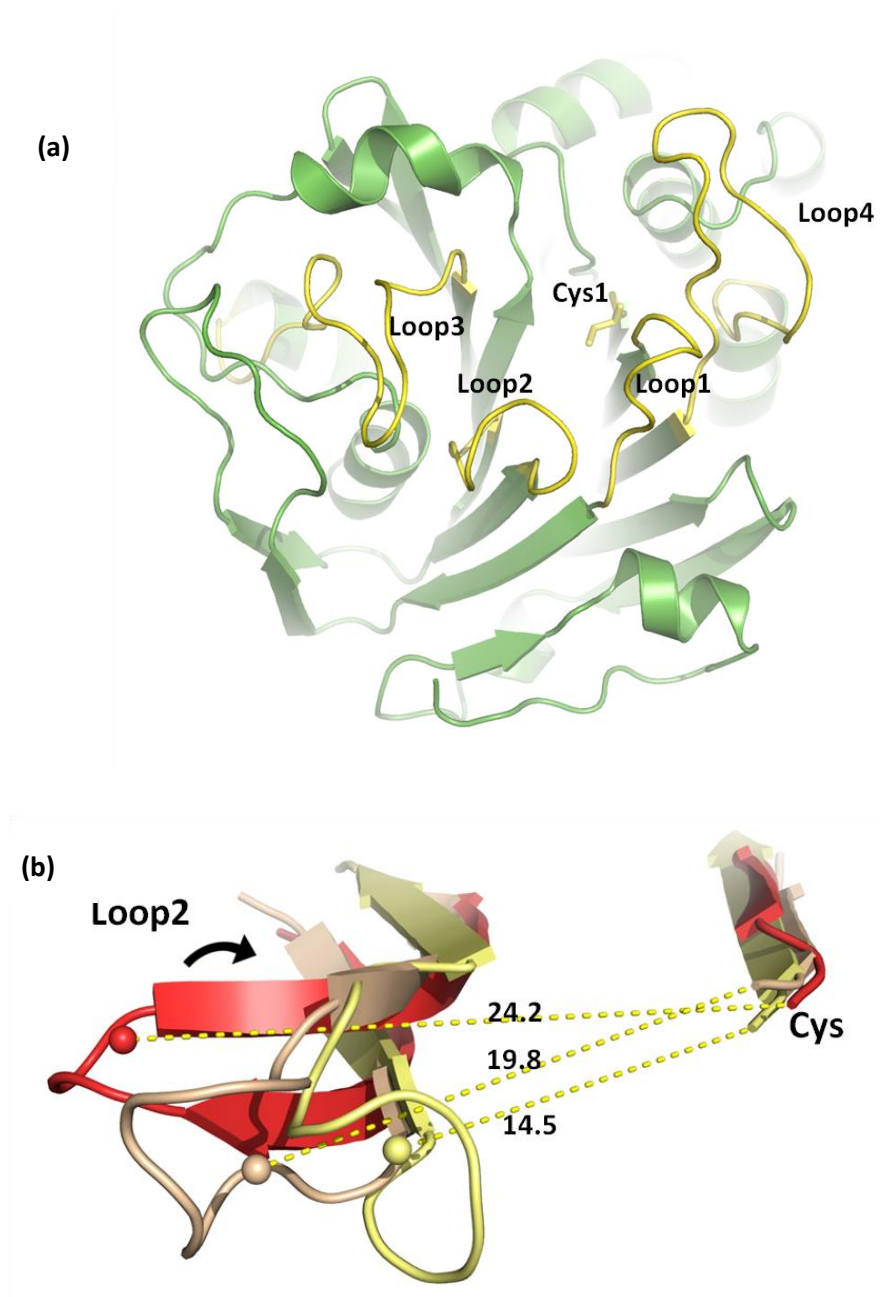
Fig. 3.12. (a) Four conserved active site loops (Yellow) surrounding Cys1. Loop 2 distance comparision between different CGHs . BSH from *Bifidobacterium longum*, PVA *Pectobacterium atrosepticum* and *Sl*CGH1.
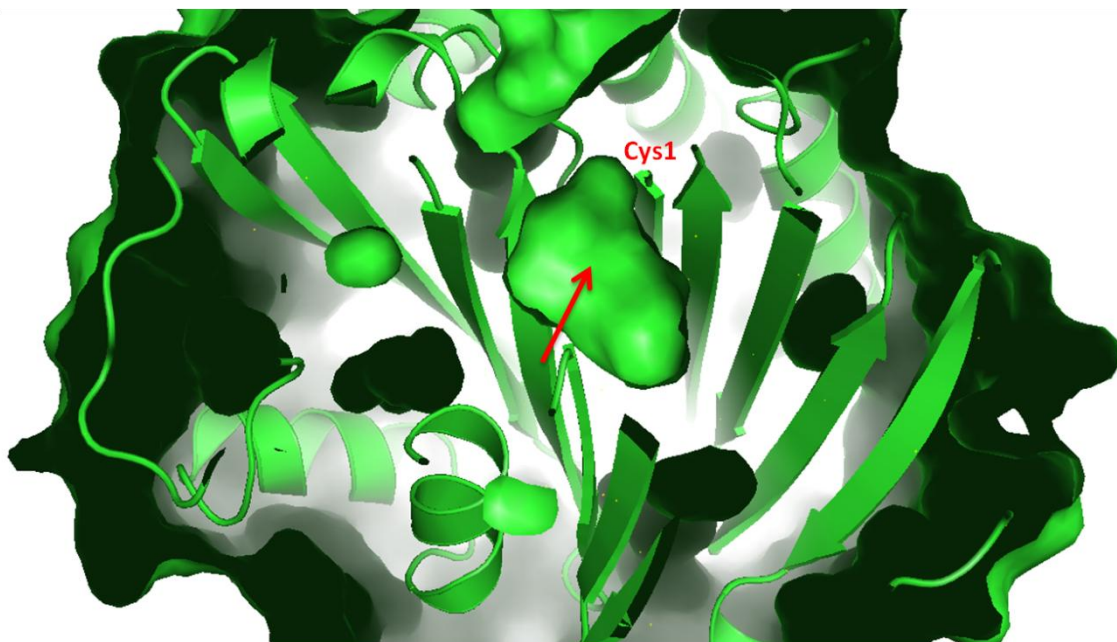
Fig. 3.13. Surface view of *Sl*CGH1 showing the hydrophobic patch of the central active site cavity.

In addition, loop3 was found to be extended by 4-14 residues than previously characterized CGHs, primarily due to the disappearance of a short β-strand. The loop 3 region also exhibits high average B factor with exceptionally highest value (31.2 $\mathring{A}^2$) centered on residue GluA148, which further increases in the ligand-bound structure (40.4 $\mathring{A}^2$). As B-factor represents real molecular motions, these values support the extra flexibility of loop3 as against the rest of the protein and Glu148 may act as gate allowing ligand entry to and exit from the buried active site pocket. Further studies could help elucidate the roles of loop 2 and 3 in deciding substrate specificity in CGHs.

3.3.6. Structure of Cys1Ser:3-oxooctanoic acid reveals a bent acyl chain conformation

In the product-bound structure, the 3-oxo-octanoic (3OAA) chain is bent in the middle between C4 and C5, imposed by the enclosed end of the U-shaped active site as well as the globular hydrophobic surrounding, on the hydrophilic acyl chain (**Fig 3.14**). We suggest that the presence of hydrophilic residues Glu65, Glu66 from the beta strand could also contribute to bending of the acyl chain. The Cys1Ser: 3-oxo-octanoic acid complex structure reveals the acyl chain trapped in the groove between β sheets I and II flanked by loop1, 2 and 3. Superposition of the ligand-

bound C1S variant on apo-*Sl*CGH1 structure reveals a low RMSD of 0.083 Å. I. There is a slight increase in the whole chain average B-factor of the complex structure (14.96 Å^2 to 20.09 Å^2) than that of the apo form. On the whole, it could therefore be assumed that the AHL binding in *Sl*CGH does not affect major conformational changes in the enzyme structure. The catalysis might basically follow Michaelis-Menten (MM) equation and do not show cooperativity between the monomers. MM kinetics is a common mode of catalysis in CGHs except for allosteric behavior reported in *Pa*PVA (Avinash et al. 2015). However another possibility might be that since only the hydrolyzed product was found bound to the enzyme, we might have captured the complex structure at a stage when the enzyme has return back to its native form.
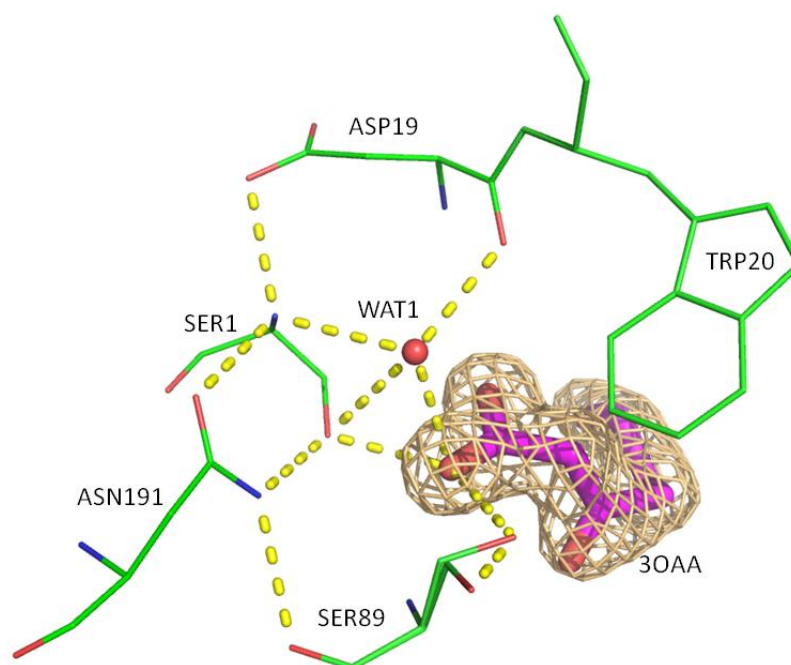
Fig. 3.14. 3-oxo-octanoic acid bound active site of C1S structure. A water molecule was found near the active site which participates in catalysis as the virtual base in the second nucleophilic attack of deacylation.

The complex structure further reveals that the hydrolyzed product (3OAA) was bound to the active site with one of its carbonyl oxygen in polar contact with a water molecule and the OH group of Ser1. At this stage the amino part of the substrate, here the lactone ring has left the active site. This depicts the second tetrahedral intermediate, formed after the cleavage of amide

bond, where in the water molecule acts as the virtual base to cause the second and final nucleophilic attack to release the acyl group (Duggleby et al. 1995; Oinonen and Rouvinen 2000). From structural superimposition with known CGHs, it was predicted that Ser89 and Asn191 would be the oxyanion hole forming residues in SlCGH1. However, in our complex the carbonyl oxygen atom of the substrate is hydrogen bonded with Ser89 and water, instead of the predicted pair of Ser89-Asn191. Instead of the substrate, Asn191 rather interacts with Ser1and Ser89 stabilizing their negative charges and by extension the tetrahedral intermediate. This observation would add a new possibility of catalysis for Cys-Ntn hydrolases where water plays the dual role as both the nucleophile and oxyanion forming residue. Some reports suggested that Cys might not require a catalytic base like water for nucleophile activation as required by other Ser/Thre-Ntn-hydrolases, owing to its almost neutral side chain pKa (Noren et al. 2000, Lodola et al. 2012). Hence, it is still not yet understood from our complex structure whether the first nucleophilic attack on the amide bond of the substrate is done in a similar manner involving the same water molecule in oxyanion hole formation.

3.3.7. Docking analysis of AHL binding to SlCGH1

Due to difficulty in obtaining the structure of intact AHL-bound complex, docking studies were applied to explore the enzyme-lactone interaction as well as to understand the difference in the binding modes of these different AHLs. 3-oxo- C_{10}-HSL, C_8-HSL and C_6-HSL were docked to C1S final model (**Fig. 3.16**). Docking of C_6-HSL and C_8 HSL showed lower Glidescores than that of 3-oxo-C_{10}-HSL. The distance of the carbonyl carbon of the amide bond from the SH of the cysteine nucleophile (nucleophilic attack distance, NAdist) is also favourably the lowest for 3-oxo- C_{10} HSL (**Table 3. 3**). These factors could explain the high activity of SlCGH1 on 3-oxo-C_{10}-HSL, as indicated by the bioluminence assay in chapter 2. Although Glide scores were similar for C_6-HSL and C_8-HSL, the farther Nadist for C_6-HSL might lead to unproductive binding and thus absence of activity on C_6-HSL, whereas moderate activity was shown on C_8-HSL at lower concentration. C_8-HSL and C_6-HSL were observed to bind in a similar orientation in the active site making a bent at the C3 of the chain.

Table 3.3. Properties of different AHL substrates and results of docking with _Sl_CGH1 structure (AlogP = hydrophobicity, SA = surface area, Nadist = Nucleophilic attack distance between SH group of cys1 and carbonyl carbon atom of AHL).

Ligand (AHL)	MW	Polar SA	AlogP	Rotatable bonds	H-bond donors	H-bond acceptors	Glidescore	Nadist (Å)
C_6 HSL	199.3	55.4	1.133	6	1	3	-2.84	8.65
C_8 HSL	227.3	55.4	2.6	7	1	3	-2.81	6.62
3-oxo- C_{10} HSL	255.4	55.4	2.958	10	1	3	-4.24	5.9

In 3-oxo- C_{10} HSL, there is a sharp bent at C4 atom of the acyl chain orienting the lactone ring closer to the cysteine nucleophile. The docked AHL ligand bound in a similar orientation as the acyl chain in the 3-oxo-octanoic acid bound C1S complex structure, although the acyl chain in the complex structure was captured closer to the nucleophile (3.3 Å away from Cys1). There were no polar contacts between the lactone ring of the AHL substrate and the active site residues.

In all the three docked AHL substrates, Arg16 interacts with the carbonyl oxygen of the lactone ring. Inside the active site cavity, shorter acyl chain ends in a surrounding dominated by hydrophilic residues, Glu65, Glu66, His67 and Gln76 etc, which may lead to weaker binding (**Fig.3.16. a,b**). Longer acyl chains are able to access a more hydrophobic part of the pocket, lined primarily by residues Ile64, Phe70, Leu73 and Met279. It can be inferred that the presence of extended acyl chain possibly leads to increased and better hydrophobic interactions with the residues in the hydrophobic pocket, thus allowing productive binding and better AHL degradation activity. Similar observation has been reported in _Pa_PVA docking studies with AHLs (Avinash et. al. 2013).
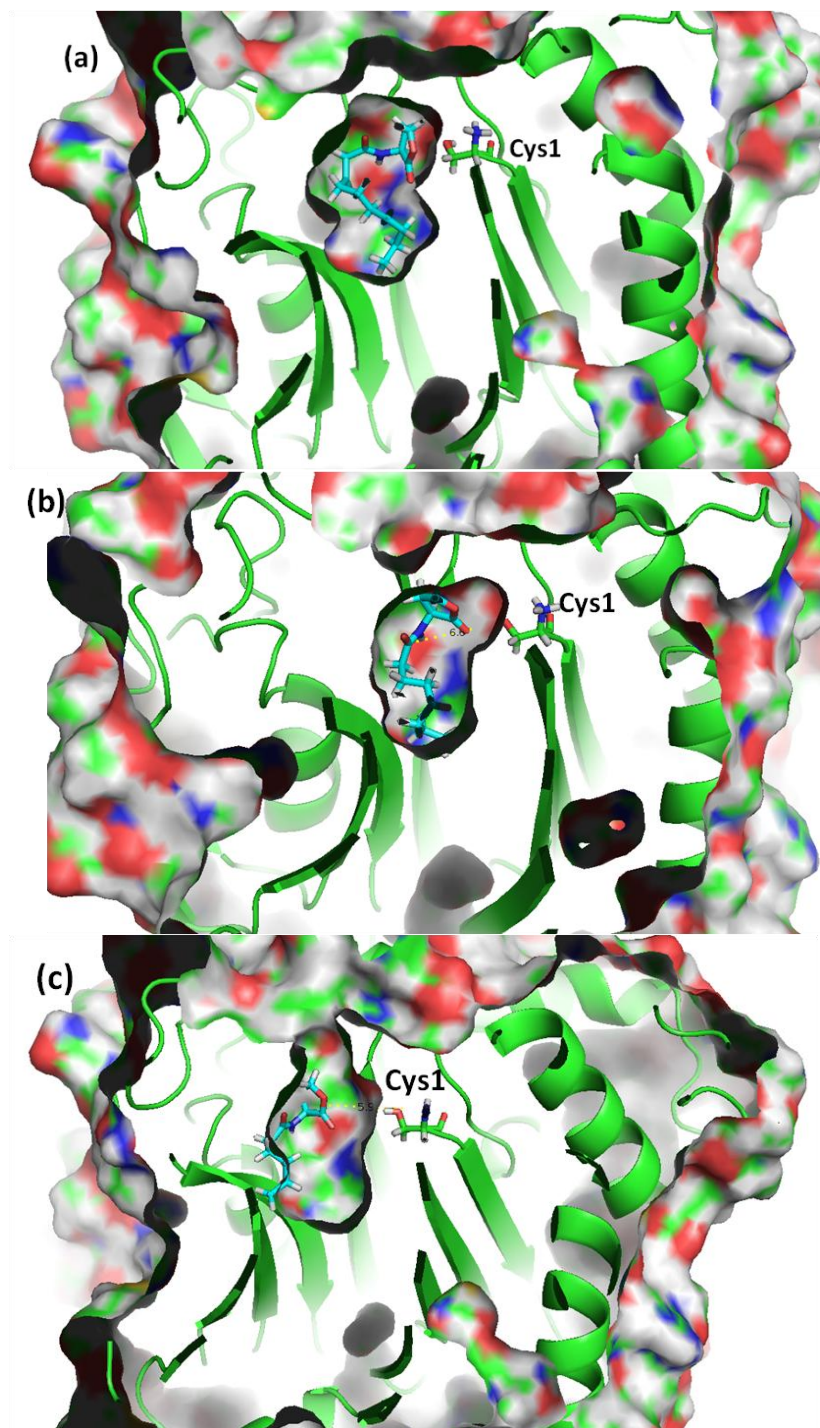
Fig. 3.16. Docked structures

(a) 3-oxo-C_{10}-HSL bound 5.9 Å away from Cys1

(b) C_8-HSL bound 6.6 Å away from Cys1

(c) C_6-HSL bound 8.6 Å away from Cys1
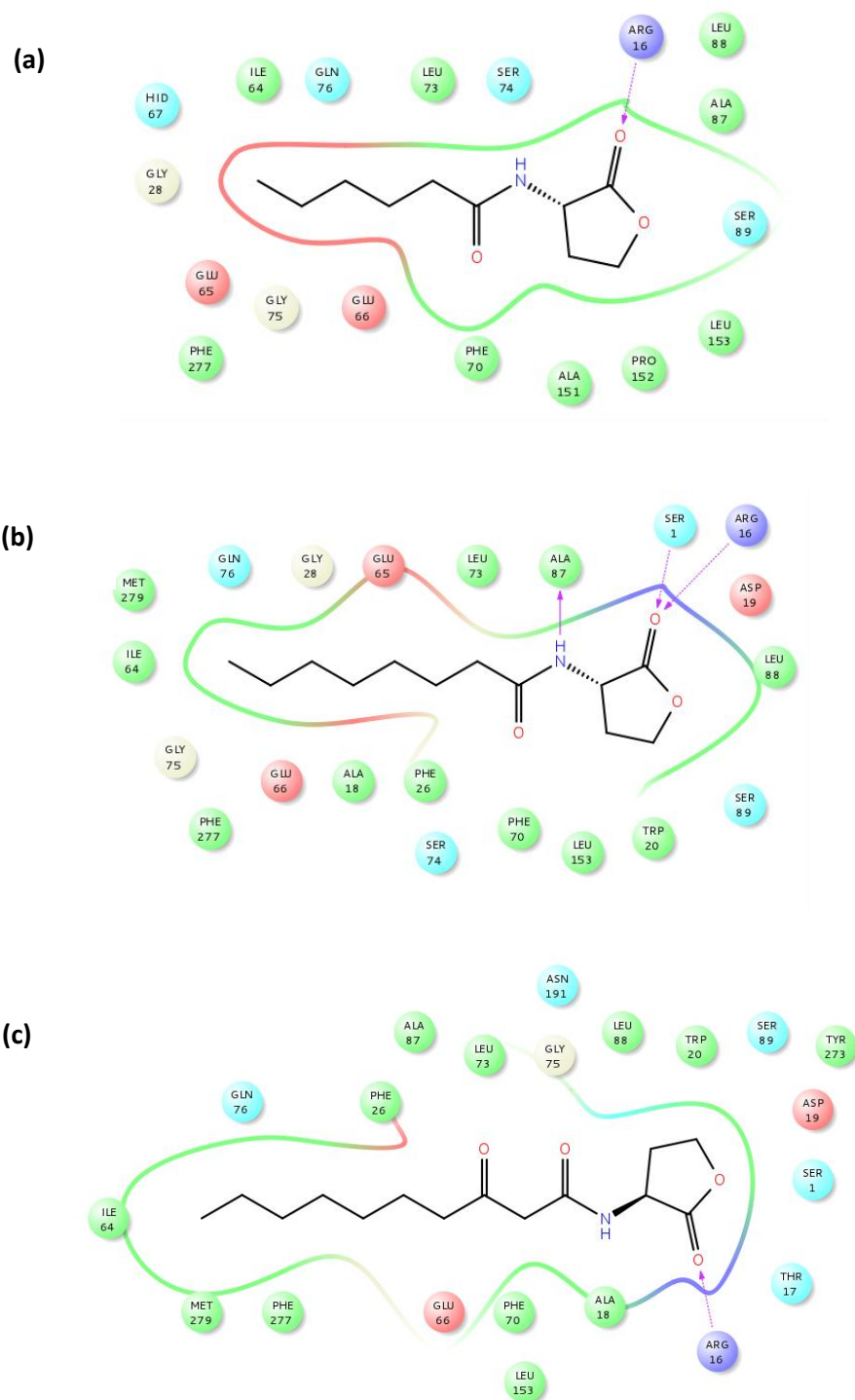
Fig. 3.17. Ligand interaction diagrams of (a) C$_6$-HSL, (b) C$_8$-HSL and (c) 3-oxo-C$_{10}$-HSL with PaPVA. Colour scheme: green – hydrophobic residues, red – acidic residues, purple – basic residues, cyan – polar residues, solid arrows – backbone hydrogen bonds, dotted arrows – side chain hydrogen bonds.

3.3.8. *Sl*CGH1 and conventional AHL acylase

Acylases active on AHLs described to date (with the exception of *Pa*PVA and *At*PVA) have been mostly heterodimeric Ser-Ntn hydrolases, exhibiting structural similarity with Pen G acylases and cephalosporin acylases. The AHL acylase PvdQ from *Pseudomonas aeruginosa* has been extensively studied and 3D structure has also been solved (Bukhove et al 2010). Structural comparison between *Sl*CGH1 and PvdQ was difficult as the two enzymes are oligomerically different. However, the core β sheets I and II that host the active site are highly conserved structurally (**Fig.3.18**). Superposition of only these β sheets revealed the conservation of core active site architecture, although with different residues (**Fig.3.19**).
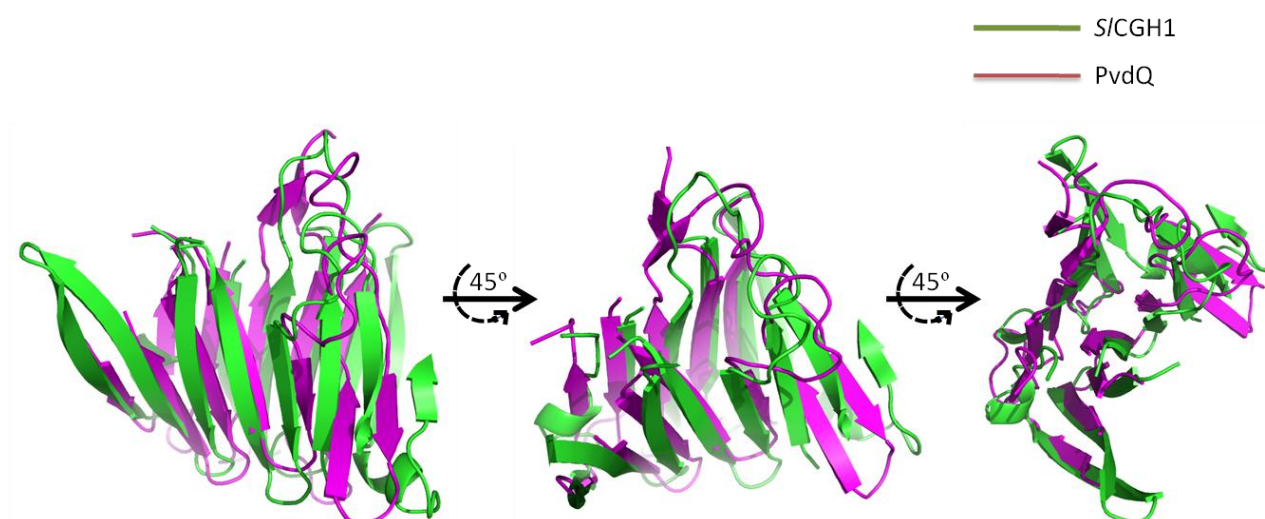
Fig.3.18. Superimposition of core β-sheets between *Sl*CGH1 and PvdQ shows conservation of β-sheets between the two structures.

The oxyanion forming residues in PvdQ, Val70 and Asn269 are replaced by Ser89 and Asn191 in *Sl*CGH1. Interestingly, the distances between carboxylate oxygen of the 3-oxo-octanoic acid and the backbone amide of Ser89 and Nδ of Asn191 are 2.8 and 4.1Å respectively which is almost a precise match with the corresponding values in PvdQ, 2.8/4.2Å. However, the active site of *Sl*CGH1 (79 Å^3) is smaller than that PvdQ (144 Å^3).

(a)

VAL70
SER89

———— SlCGH1
———— PvdQ

PHE24
TRP20

ASP269
ASP191

SER1
CYS1

HIS23
ASP19

(b)

———— SlCGH1
———— PvdQ

HIS23
ASP19

SER1
CYS1

3-oxo-octanoic acid

PHE24
TRP20

VAL70
SER89

WAT
WAT

3-oxo-dodecanoic acid

ASP269
ASP191
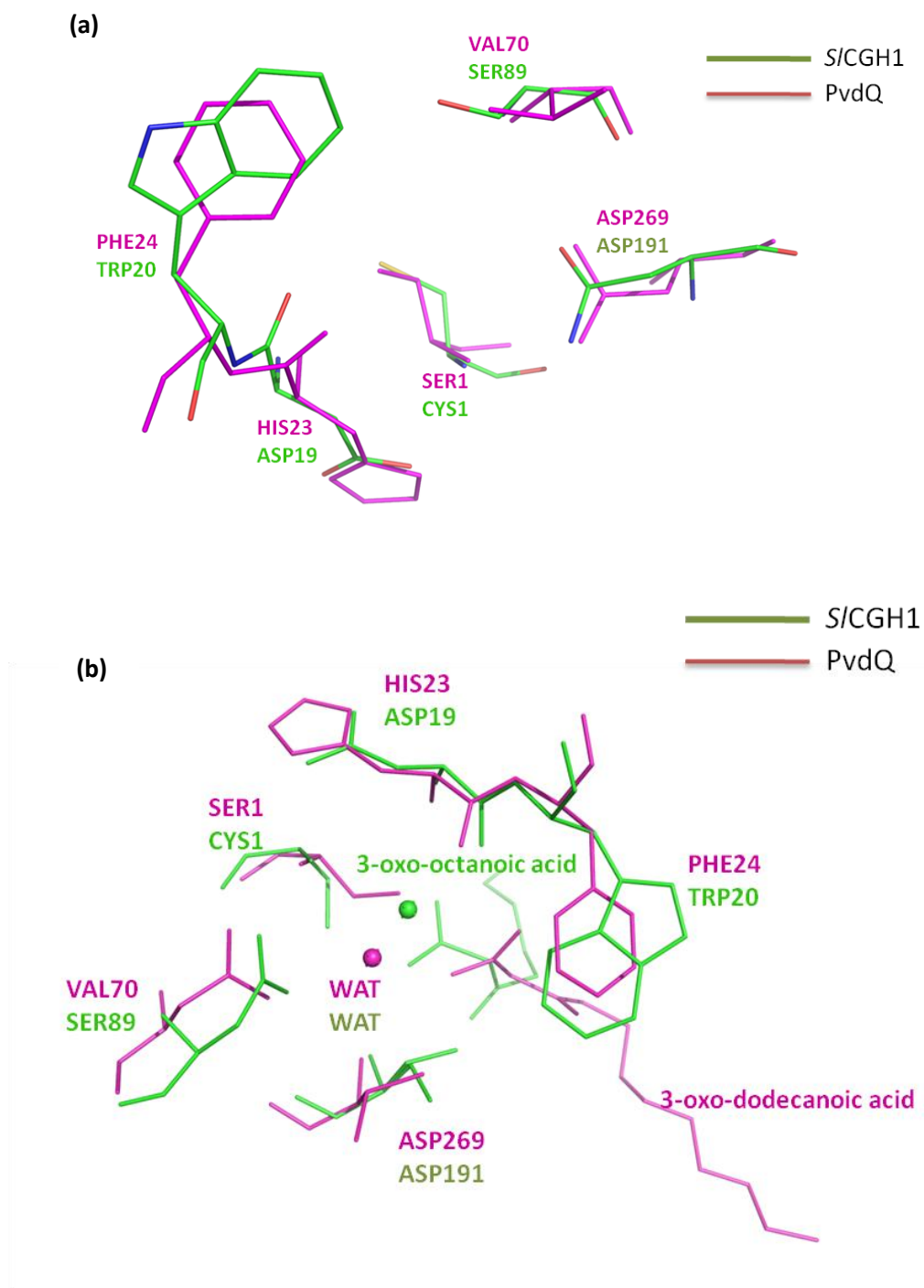
Fig. 3. 19.(a) Active site similarity shown between SlCGH1 and PvdQ. (b) 3-oxo-octanoic acid bound active site of C1S structure superposed with 3-oxo-dodecanoic acid bound PvdQ active site showing differences in acyl chain binding.

In contrast to the bent acyl chain in *Sl*CGH1, PvdQ accommodates a straight acyl chain conformation owing to a larger active site with deep hydrophobic region. The ligand-binding pocket in PvdQ has been suggested to have adapted for long acyl chain binding. PvdQ has been reported to undergo mild conformational changes to accommodate the long acyl chain substrate with Phe24 acting as the gate between the binding pocket and solvent. In *Sl*CGH1, the RMSD between apoenzyme and ligand-bound structure was minimal, but with an increased B factor, (13.6 Å^2 to 18.945 Å^2 for main chains) especially in loop3 centered on residues 147-149. This could indicate small changes in loop conformation in *Sl*CGH1 to allow substrate entry and binding. Moreover, no specific interactions were observed between the 3-oxo group of the AHL substrate and the enzyme, which is in agreement with the report on PvdQ. It has been suggested that the polar oxo group might ease the exit of the acyl chain out of the hydrophobic binding pocket (Bokhove et al 2010). Thus the two enzymes show similarity in active site architecture despite overall structure and oligomeric differences. Acyl chain binding mode is also different between the two enzymes. Although it seems both these enzymes are adapted for hydrolysis of long fatty acid chain substrates, further studies might be required to confirm whether AHLs are the true substrate or represent a promiscuous side of *Sl*CGH1 and to define the role *Sl*CGH1 like enzymes in the microbial communication circuit.

Chapter 4

Characterization, molecular modeling of *Sl*CGH2 and phylogenetic analysis of *Shewanella loihica* CGHs

4.1 Introduction:

The choloylglycine hydrolase (CGH) group of the Ntn hydrolase enzyme superfamily encompasses penicillin V acylase (PVA, E.C. 3.5.1.11) and bile salt hydrolase (BSH, EC 3.5.1.24). This project aims to study CGHs from the marine bacterium *Shewanella loihica* PV-4. As described in the previous chapters, *S. loihica* exhibits two CGHs in its genome with only 35% identity to each other. *Sl*CGH1 has been identified as a non-PVA, non-BSH enzyme, showing promiscuous hydrolyzing activity on N-acylhomoserine lactones (AHLs).

This chapter describes the biochemical and structural characteristics of *Sl*CGH2 to understand its substrate spectrum in comparison with *Sl*CGH1. In addition, analysis of bacterial CGH sequence phylogeny and the conservation of residues in their substrate binding sites have also been performed to reveal the uniqueness of marine CGH homologues.

4. 2 Materials and methods:

4.2.1 Cloning, expression and purification of *Sl*CGH2 enzyme

The gene coding for *Sl*CGH2 from *Shewanella loihica* PV-4 was PCR amplified using forward (5′ gaa tta *cat atg* tgc tct agc gtt tcg gcc ac 3′) and reverse (5′ cta gat *ctc gag* gct aca ctt gcc aaa ggt tg 3′) primers (restriction sites highlighted) and cloned into a pET22b(+) plasmid vector (Invitrogen) between NdeI and XhoI restriction sites. The recombinant plasmids were selected for ampicillin resistance and *Sl*CGH2 was expressed in *E. coli* BL21 star cells (Novogen) with a C-terminal His-tag. The enzyme was purified to homogeneity using a HIS Select Ni^{2+} affinity column and ENrichTM 650 (BioRad) size exclusion column, using a similar procedure as described for *Sl*CGH1 in Chapter 2. Purified *Sl*CGH2 was dialyzed against 10mM Tris HCl buffer pH 7.2 containing 100mM NaCl and 1mM DTT and stored at 4°C.

4.2.2. Screening of enzyme activity

Pen V hydrolysis activity was estimated by studying the formation of Schiff's conjugate with the product 6-APA and p-dimethyl amino benzaldehyde (Shewale et al. 1987). BSH assay was performed according to Kumar et al 2006 through a ninhydrin-based estimation of the amount of free amino acids released upon hydrolysis of bile salts. Lipase activity was estimated by the measuring the hydrolysis of pNPP at 410 nm (Winkler and Stuckmann, 1979).

AHL degradation was assayed by Lux based bioluminesence systems 1) *E. coli* pSB 401 for C_6-C_8 HSL and their oxo forms and 2) *E. coli* pSB1075 for long chain AHLs, C_{10}-C_{12} HSL and their oxo forms, as described in chapter 2 (Wahjudi et al. 2011). Results were expressed (mean of three independent experiments) in terms of residual relative light unit (RLU; LU/OD 600) obtained from the difference between the sample readings and that of the respective inactive enzyme controls.

4.2. 3. HR-MS spectroscopic analysis of AHL degradation

HR-MS was carried out to elucidate the nature of AHL degradation by activity of *Sl*CGH2, based on the method of Huang et al (2003) with slight modifications. Reaction products were analyzed using a C_{18} ultra-aqueous reverse-phase column fitted to a Q Exactive system with Accela 1250 Pump and PDA detector from Thermo Scientific. The mobile phase consisted of 50 % methanol (methanol-water with 1% acetic acid) for the initial 1 min and linearly increased to 80 % methanol. Samples were diluted in methanol before application and run time was 6 min.

4.2.4. Crystallization trials

Initial crystallization trials were set up employing the commercial screens - JCSG+ (Molecular Dimensions), PEG/Ion, Index and PEGRx (Hampton Research). The protein in 10mM Tris buffer pH 7.0, 0.1M NaCl, 1mM DTT, was concentrated to 11 mg/ml and the screens were set up using a Mosquito crystallization robot (TTP LabTech, UK) in 96-well Hampton Research MRC plates as sitting-drop vapour-diffusion experiments. Further manual optimization was performed using hanging drop technique, with 1μl protein mixed with 1μl or 2μl precipitant solution.

4.2.5. Homology Modeling of *Sl*CGH2 and Model Assessment

Homology modeling of *Sl*CGH2 was attempted by using two programs, MODELLER 9.17 (Sali 2016) and I-TASSER (Iterative Threading ASSEmbly Refinement) (Yang 2014). MODELLER performs protein structure modeling based on satisfaction of spatial restraints which include related protein structures (comparative modeling), NMR experiments (NMR refinement), fluorescence spectroscopy, rules of secondary structure packing (combinatorial modeling), cross-linking experiments, image reconstruction in electron microscopy, site-directed mutagenesis, intuition, etc. It derives these restraints in terms of distances, angles, dihedral angles, pairs of

dihedral angles etc from the known related structures and their alignment with the query sequence. The models generated by Modeller were scored on the basis of their DOPE score (Discrete Optimized Protein Energy) and the model with lowest DOPE score was selected for further studies.
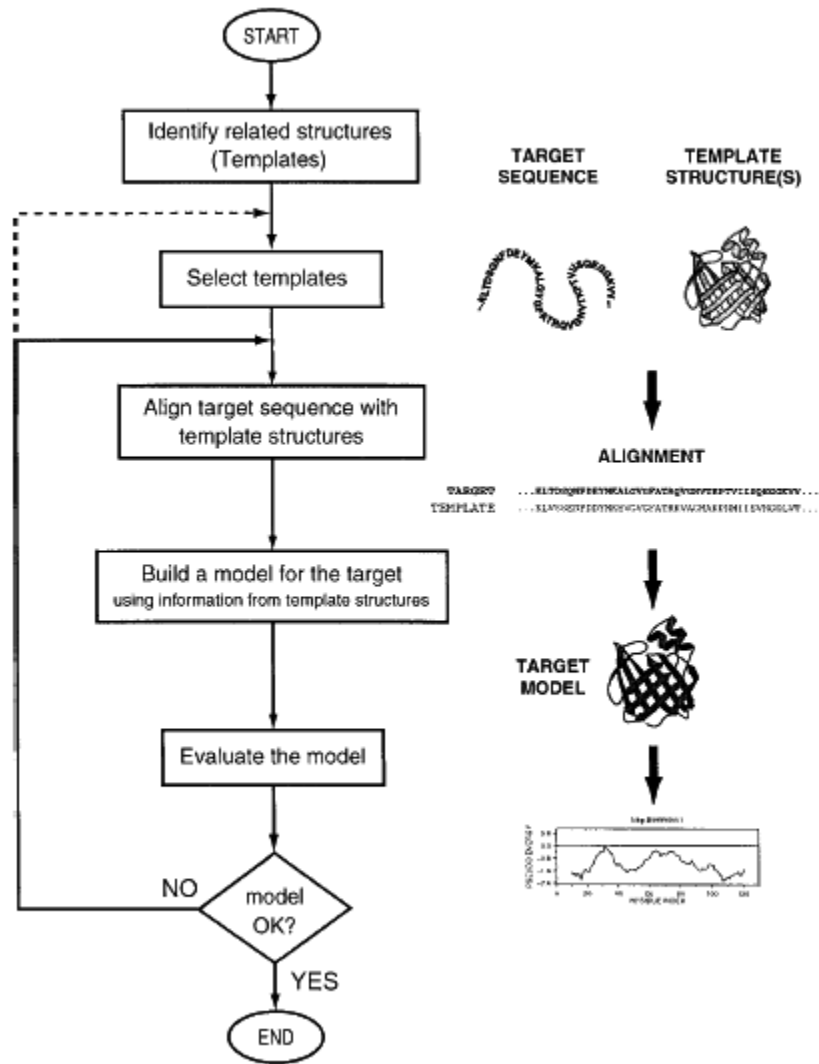
Fig.4.1. Flowchart elucidating the steps in protein homology modeling, adapted from Mart et al (2000).

I-TASSER, on the other hand, models 3D structures by iterative threading assembly simulations and also performs structure-based function annotation. I-TASSER is based on a combinatorial approach, employing all three conventional methods for structure modeling: comparative

modeling, threading, and *ab* initio modeling. The query sequence is first threaded through PDB library using LOMET, a meta-threading method containing eight fold-recognition programs (PPAS, Env-PPAS, wPPAS, dPPAS, dPPAS2, wdPPAS, MUSTER and wMUSTER) (Wu & Zhang, 2007). The LOMETS aligned sequence is divided into threading-unaligned regions and threading-aligned. The structure of unaligned regions is built from scratch by *ab initio* folding. The topology of full-length models is constructed by reassembling the continuously aligned fragments excised from templates by replica-exchange Monte Carlo simulations (Wu & Zhang, 2007) Figure. Finally, the structures of the lowest energy are selected, which are then refined for optimizing the hydrogen-binding network and removing steric clashes.
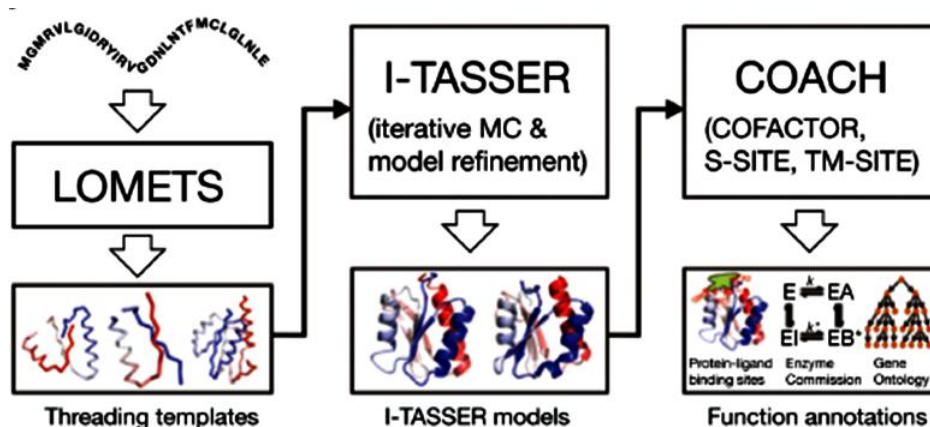
Fig.4.2. Flowchart taken from Zhang et al.2014, elucidating the schema for protein homology modelling used in I-TASSER. The I-TASSER Suite pipelines consist of LOMETS-based template identification, I-TASSER–based Monte Carlo (MC) structure assembly simulation and COACH-based function annotation.

4.2.6. Validation of structure models

The homology model was further validated by model validation programs like ProSA-web server, Verify3D and PROCHECK.

4.2.7. Phylogenetic analysis

Phylogenetic analysis of CGH family members was performed according to method described by Panigrahi et al. (2014) with modifications. Protein sequences of 7 structurally characterized CGHs, namely *Bacillus sphaericus* (BspPVA), *B. subtilis* (BsuPVA),*Pectobacterium atrosepticum* (PaPVA), *Bacteroides thetaiotamicron*(BtBSH), *Bifidobacterium longum* (BlBSH)

Clostridium perfringens (CpBSH) and *Lactobacillus salivaris* (LsBSH) along with *Sl*CGH1 sequence were used as 'templates' to retrieve CGH homologs sequences through a BLAST search. The search was filtered through a blast score cut off of 300 and an E- value of 0.008 (0.002 in case of *Sl*CGH1 due to limited number of hits). The cut-off limit was set at 60% redundancy (Jalview software) and sequences lacking N-terminal Cys residue were excluded. After multiple alignment using Clustal X , a phylogenetic tree was constructed in Mega6 (Tamura 2013) using the neighbor-joining method (Saitou 1987) with a bootstrap value of 1000. Further improvement in the graphical representation of the tree was done using Figtree, ver1.42 software (http://tree.bio.ed.ac.uk/software/figtree/).

4.2.8. Binding score similarity calculation

Panigrahi et al. 2014 developed the Binding Site Similarity Score (BSS score) which is a binding site profile-based scoring system to estimate quantitatively the BSS of each CGH family member, based on the binding site information from each of the six template CGH enzymes (*Bl*BSH, *Cp*BSH, *Bt*BSH, *Bsp*PVA, *Bsu*PVA and *Pa*PVA). The score of the query sequence with the *i*th template (*i*=1–6) was calculated as:

$$S(Query, Template\_i) = \sum_{j=1}^{N} S(AA\_query, AA\_Template\_i)j$$

Where, **S(AA_query,AA_Tempate_i)j** is the similarity score between amino acid residues of the query and the *i*th template sequence, at the *j*th position of the binding site profile.

4.3. Results and discussion:

4.3.1. Cloning and expression of *Sl*CGH2 gene from *Shewanella loihica* PV-4

The 1086 bp gene coding for *Sl*CGH2 was PCR amplified from the genomic DNA of *Shewanella loihica* PV-4. *Sl*CGH2 was expressed in *E. coli* BL21 as soluble cytoplasmic protein; however unlike *Sl*CGH1, the level of expression was very low, 0.2 - 0.5 g/ L. The protein was purified to homogeneity using Ni^{2+} NTA affinity chromatography and purity was confirmed by the 37 KDa subunit molecular weight band of the protein on a 12 % SDS PAGE gel (**Fig. 4.3a**).
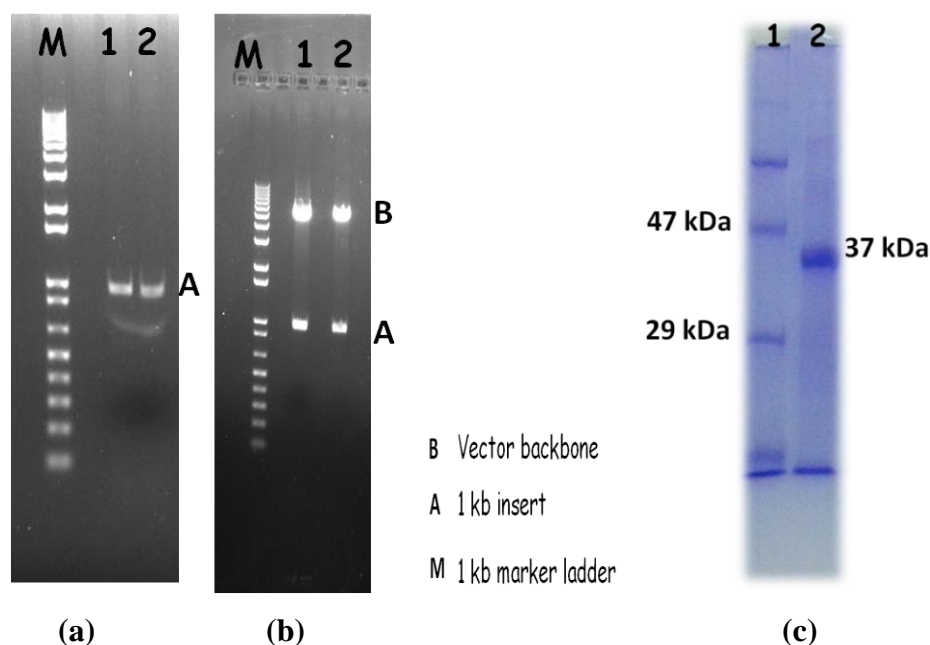
Fig 4.3 (a) PCR amplified band of *Sl*CGH2 on 1 % agarose gel (b) *Sl*CGH2- pET22b after NdeI/XhoI restriction digested showing 2 Kb vector backbone and 1 kb vector insert (c) *Sl*CGH2 on 12% SDS-PAGE gel (Lane 1 – BioRad protein molecular weight marker, MW in kDa, 2 – purified *Sl*CGH2).

4.3.2. Biochemical characterization of *Sl*CGH2

*Sl*CGH2 was also screened for various substrates such as penicillins, bile salts, AHLs, p-nitrophenylpalmitate and casein for potential penicillin acylase, BSH, AHL acylase, lipase and protease activities respectively. The lux-based bioluminesence assay using *E. coli* pSB 401 and *E. coli* pSB1075 reporter strains was used to detect the AHL activity spectrum of *Sl*CGH2. Similar to *Sl*CGH, the enzyme was inactive on all the above mentioned substrates except AHLs. However, the AHL acyl chain preference varied. While *Sl*CGH1 was predominantly active on C_{10}-HSL and 3-oxo-C_{10}-HSL, *Sl*CGH2 showed robust activity on AHLs with acyl chain longer than 10 carbons, with maximum activity on 3-oxo-C_{14}-HSL, followed by C_{14} HSL (**Table 4.1**). Almost 99% reduction in luminescence intensity was obtained in case of 3-oxo C_{14} HSL, followed by around 90 % reduction against C_{14} HSL and C_{12} HSL signals. There is mild activity on 3-oxo-C_8-HSL, C_{10}-HSL and 3-oxo-C_{10}-HSL and no activity on C_6 HSL. As in the case of *Sl*CGH1, no regular pattern could be discerned on the effect of 3-oxo substitution on AHL substrate preference.

The nature of AHL degradation by *Sl*CGH2 was analyzed using High Resolution Mass spectroscopy. The presence of peak corresponding to sodiated homoserine lactone (HSL), m/z of 124 with ppm value within ± 2.0, confirmed acylase activity of the enzyme (Fig 5). HSL peaks were confirmed in all the AHLs with positive bioluminesence activity.

*Sl*CGH2 has been annotated as a choloylglycine hydrolase in the GenBank database. This new activity profile could indicate the possibility of *Sl*CGH1, *Sl*CGH2 and their close homologs, which are predominantly marine CGH like enzymes might have preference for substrates different from the CGH of other habitats such as soil. This could also reflect a difference in the role of these enzymes in microbial physiology, evolved in response to the marine environment. Since *Sl*CGH1 and *Sl*CGH2 so far seem to share a common substrate, they might be paralogs evolved to share the same ancestral function in parts. They both might be adapted for long chain AHL like substrates with *Sl*CGH2 for longer substrates than *Sl*CGH1.

Table 4.1. Result of AHL degradation by *Sl*CGH 2 assayed using Lux-based biosensors (*E. coli* pSB401 and *E. coli* pSB1075). Residual RLU % = RLU of test/ RLU of control*100 (mean± SD).

AHLs	Residual RLU %
C_6-HSL	100.32 ± 4.3
C_8-HSL	114.86 ± 9.8
3-oxo-C_8-HSL	94.19 ± 3.9
C_{10} HSL	93.14 ± 2.8
3-oxo-C_{10}-HSL	95.93 ± 2.3
C_{12}-HSL	13.65 ± 6.8
3-oxo-C_{12}-HSL	39.36 ± 4.9
C_{14} HSL	9.41 ± 3.003
3-oxo-C_{14}-HSL	1.06 ± 0.67

Nature of AHL degradation by *Sl*CGH2 was analyzed using High Resolution Mass spectroscopy. The presence of peak corresponding to sodiated homoserine lactone (HSL), m/z of 124 with ppm value within ±2.0, confirmed acylase activity of the enzyme (Fig 5). HSL peaks were confirmed in all the AHLs with positive bioluminesence activity.
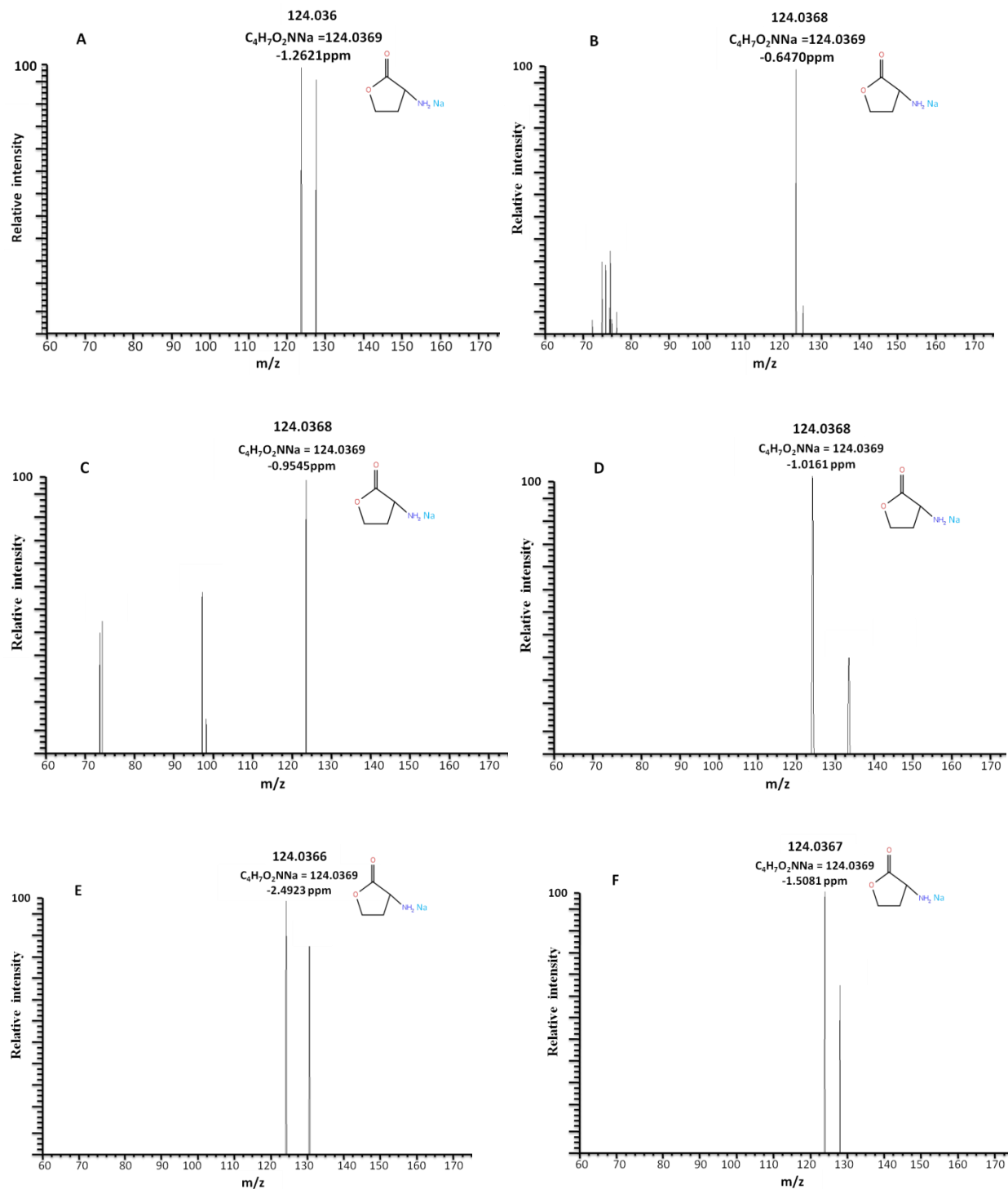
Fig. 4.4. HRMS spectrum showing sodiated HSL peaks of m/z of 124 with corresponding ppm values obtained from degradation of (A) C_{12} HSL (B) 3-oxo C_{12} HSL (C) C_{14} HSL (D) 3-oxo C_{14} HSL, (E) C_{10} HSL   (F)3-oxo C_{10} HSL.

We suggest *Sl*CGH1-*Sl*CGH2 pair is a case of subfunctionlization, where duplicate genes share in parts the ancestral gene function (Zhang 2003, Conant 2008). Together these two genes impart the capacity to impede signaling with a wide range of AHLs. Occurrence of multiple AHL acylase genes has also been reported in *P. aeruginosa PAO1,* where three AHL acylases PvdQ, Quip and Pa0305 has been characterized as long chain AHLs degraders (Huang et al. 2003; Huang et al. 2006). As discussed in chapter 2, preference for medium and long chain AHLs (C_8 to C_{12}) is also common among previously characterized AHLs (Park et al. 2005; Uroz et al. 2009). This abundance of medium and long chain degrading enzymes may be related to the less stable nature of short chain AHLs which bypass the need for 'extra' arrangements for recycling the molecule.

4.3.3. Crystallization trials

Preliminary crystallization screening using different commercial screens showed no hits except in two screens – PEG/Ion and Index1 (**Fig. 4.5**). Small cube-shaped crystals appeared in PEG/Ion screen under the composition, 0.2 M sodium fluoride and 20 % PEG 3350 (pH 7.3), within five minutes of robotic mixing. In the case of Index screen containing 0.1M Sodium acetate trihydrate pH 4.5 and 2.0 M Ammonium sulfate, short thin needles appeared after two days of application. Unfortunately, we couldn't able to reproduce the crystals despite several rounds of manual optimization of the above mentioned crystallization conditions. Therefore, homology modeling of *Sl*CGH2 was performed to glean structural information that could explain the substrate spectrum.
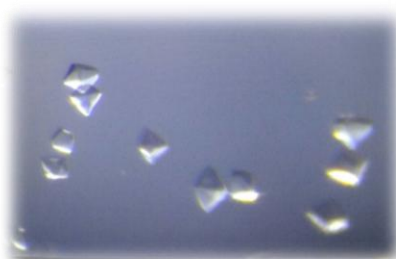
Fig. 4.5. Small crystals of *Sl*CGH2

4.3.4. Structure modeling of *Sl*CGH2

MODELLER is most frequently used for homology or comparative protein structure modeling. I-TASSER has also demonstrated to be particularly effective in *ab initio* modeling of small proteins with low primary sequence similarity (Wu et al., 2007). A search for the closest homolog with known tertiary structure was performed against PDB database for both the programs. Using MODELLER, the best templates identified were the crystal structures of cholylglycine hydrolases from *Bacteroides thetaiotaomicron* VPI (3hbcA), *Bifidobacterium longum* (2hf0A)*, Lactobacillus salivaris* (5hkeA), *Bacillus subtilis* (2oqcA) and *Bacillus sphaericus* (2quyA) and *Pectobacterium atrosepticum* (4wl2A). We selected 3hbcA as the template based on the relationship drawn between these sequences from a distance matrix, supported by % sequence identity and R-factor value. Five models were generated and were validated based on their DOPE score. The lowest DOPE score model was taken as the best model for further studies. Alternatively, we used the solved structure of *Sl*CGH1 (this study) as the template owing to high sequence homology (35 %), compared to the BLAST hits that showed <25% similarity. Two models selected each from these two templates were compared for better quality using different the model validation programs.

On the other hand, the I-TASSER server chooses 10 best templates from its LOMET threading programs. Five top ranking 3D models were generated by employing the top ten threading folds as the native structures. The models was compared based on their C-score (confidence score), TM-score (template modeling score) and RMSD (Root-mean-square deviation). The model with the highest C score (0.95) was selected as the final model (TM score = 0.84 ± 0.08, RMSD= 4.5 ± 3.0 Å). Another model was also constructed with *Sl*CGH1 as the template structure, showing a C-score of 1.07, TM score = 0.86 ± 0.07 and RMSD = 4.2 ± 2.8Å.

4.3.5. Validation of model quality

All models were validated by the model validation program such as Verify3D and Prosa. The models built using *Sl*CGH1 as the template were of better quality than those built on other CGH templates. Between *Sl*CGH1 based models of the two programs, MODELLER built model was a etter model than that of the I-tasser model as validated by Ramachandran plots (**Fig.4.6**).
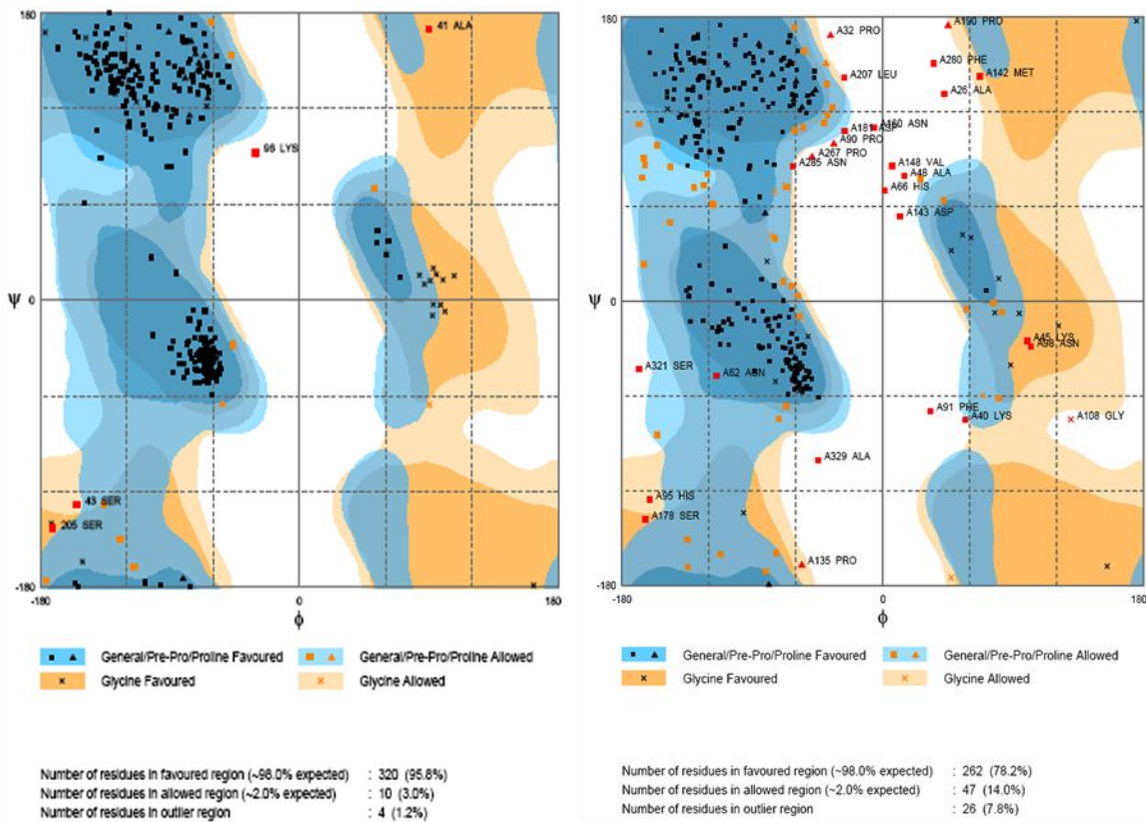
(a)   (b)
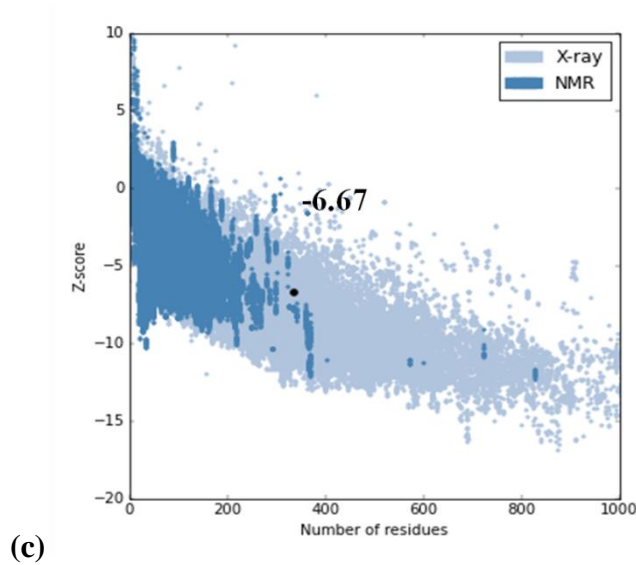
(c)

Fig.4.6. Validation reports: Ramachandran (phi-psi) plot of final models using (a) Modeller (b) I-tasser programmes. (c) Z-score of *Sl*CGH2 model from ProSA

ProSa displays overall model quality in terms of the Z-score and a plot of residue energy. Z-score measures the deviation of the total energy of the structure with respect to an energy distribution derived from experimentally determined protein chains in the current PDB. The final model showed an overall Z-score (-6.67), which was within the range of within the range of scores usually observed for native proteins of similar size. All residues had a Verify3D quality score greater than zero, indicating the absence of conformational error. Superimposed structures of *Sl*CGH2 and *Sl*CGH1 has a root mean square deviation (RMSD) of 1 Å of their backbone atoms (**Fig. 4.7**).
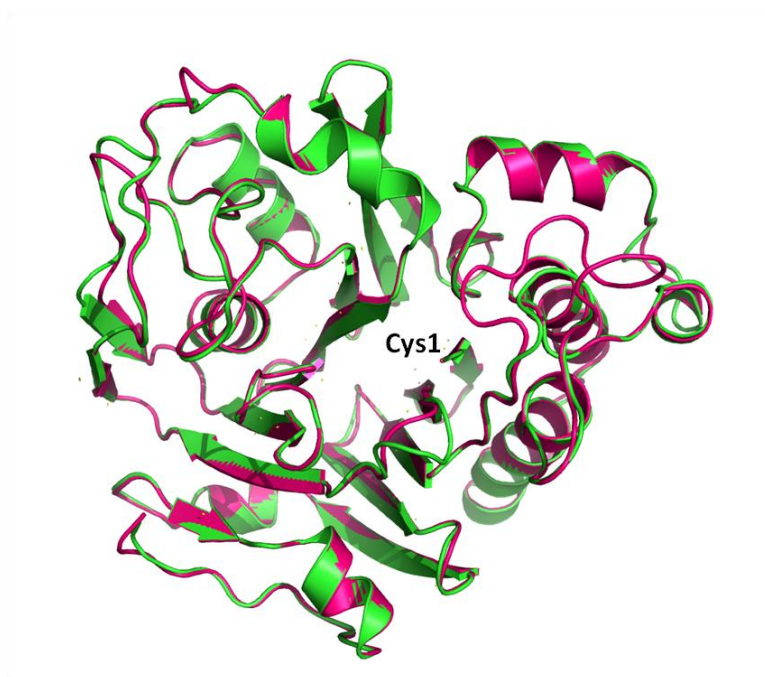
Fig. 4.7. Homology model of *Sl*CGH2 superposed on the crystal-derived structure of *Sl*CGH1.

4.3.6. Analysis of *Sl*CGH2 structure.

When the four conserved active site loops were compared between *Sl*CGH1 and *Sl*CGH2, loop 2 and 3 are found very similar in orientation and sequence length, while loop 1, 4 of *Sl*CGH2 are slightly longer (**Fig. 4.8**). However, since *in silico* modeling of loops may lack accuracy due to the dynamic nature of loops, difference in substrate specificity could not be explained with modeling data alone. We believe that the similar arrangement of loops especially in loop2 and 3, might be responsible for the inactivity of *Sl*CGH2 on PenV and bile salt in the same way it was seen in *Sl*CGH1. However, it could not convincingly explain the preference of *Sl*CGH2 for AHLs of longer chain (3-oxo C_{12} HSL, C_{12} HSL, C_{14} HSL, and 3-oxo C_{14} HSL) than that of *Sl*CGH1 (3-oxo-C_{10}- HSL, C_{10}-HSL), we look for significant differences between the two structures specially the active site.
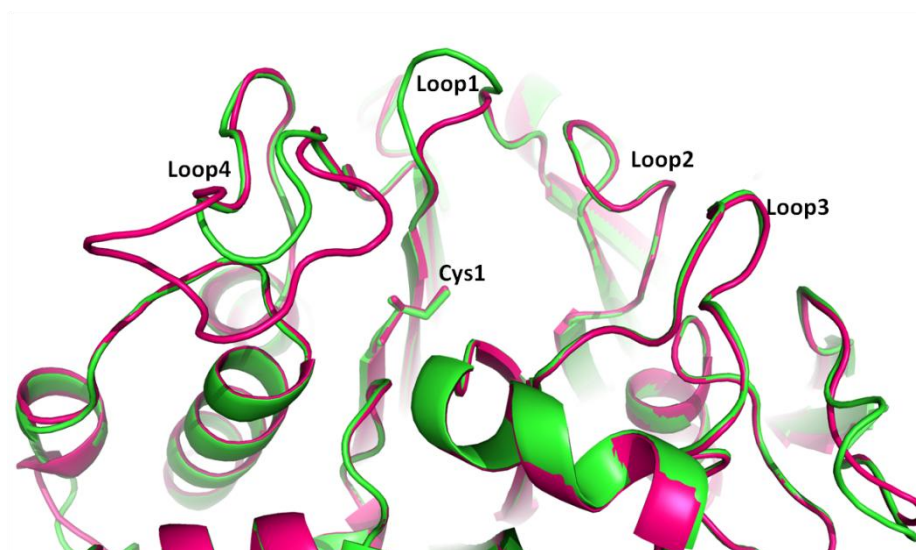
Fig. 4.8. Superposition of four conserved active site loops of *Sl*CGH1 and *Sl*CGH2. Loop 2 and 3 show minimum deviation between the two structures. Loops1 and 4 in *Sl*CGH2 active site are longer than that of *Sl*CGH1.

The active site of *Sl*CGH2 also shows conservation of basic CGH active site residues (**Fig. 4.9**). Highly conserved active site residues in various CGHs including *Sl*CGH1 are Arg16/17/19, Asp19/21/22, Trp20/21/23, Asn172/183/191 and Arg225. Also an important oxyanion residue which varies frequently in different CGHs is Ser89 in *Sl*CGH1, Tyr82/Tyr87 in PVA, Asn81 in BSH. In *Sl*CGH2, these conserved residues correspond to Arg16, Asp19, Trp20, Asn188,

Arg220 and Tyr84. Although one of the oxyanion forming residue varies in different CGHs, the oxyanion hole formation is usually not affected as it is the backbone NH- group that participates in the process.
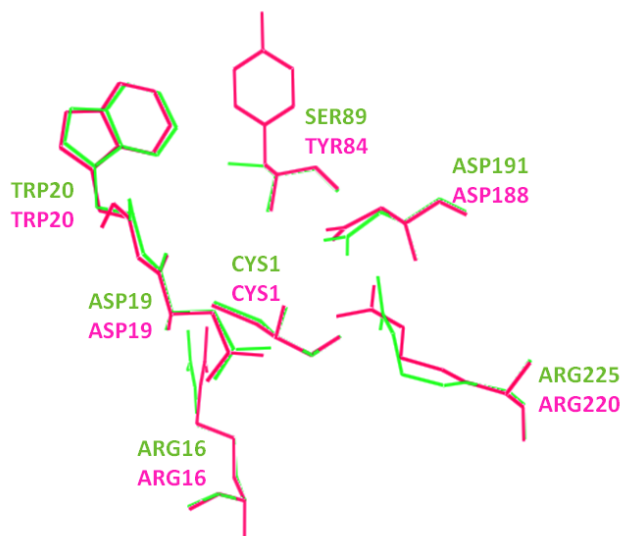
Fig.4.9. Superimposed active site of *Sl*CGH1 and *Sl*CGH2, showing conserved residues around the nucleophile Cys1.

However, it could be observed that *Sl*CGH2 exhibits a more pronounced and elongated hydrophobic patch than in *Sl*CGH1 (**Fig.4.10**). Leu18, Val 24, Ile64, Gly 82, Leu 110, Ala111, al148, Pro149, Gly150 and Phe152 etc., dominated the contributions to the apolar patch. The longer acyl chains could possibly be accommodated with more ease in *Sl*CGH2, in the elongated hydrophobic cavity. In *Sl*CGH1, the hydrophobic pocket housing the AHL acyl chain is smaller and circular, forcing the acyl chain to bend. The AHL acylase, PvdQ, reported to prefer long chain AHLs, exhibits a deep hydrophobic pocket housing the acyl chain in the active site (Bokhove et al. 2010). Similar binding geometry for the long chain AHLs could be expected in *Sl*CGH2, owing to the large hydrophobic cavity. Further analysis of AHL substrate binding to *Sl*CGH2 could provide a better understanding of the substrate binding and catalysis.
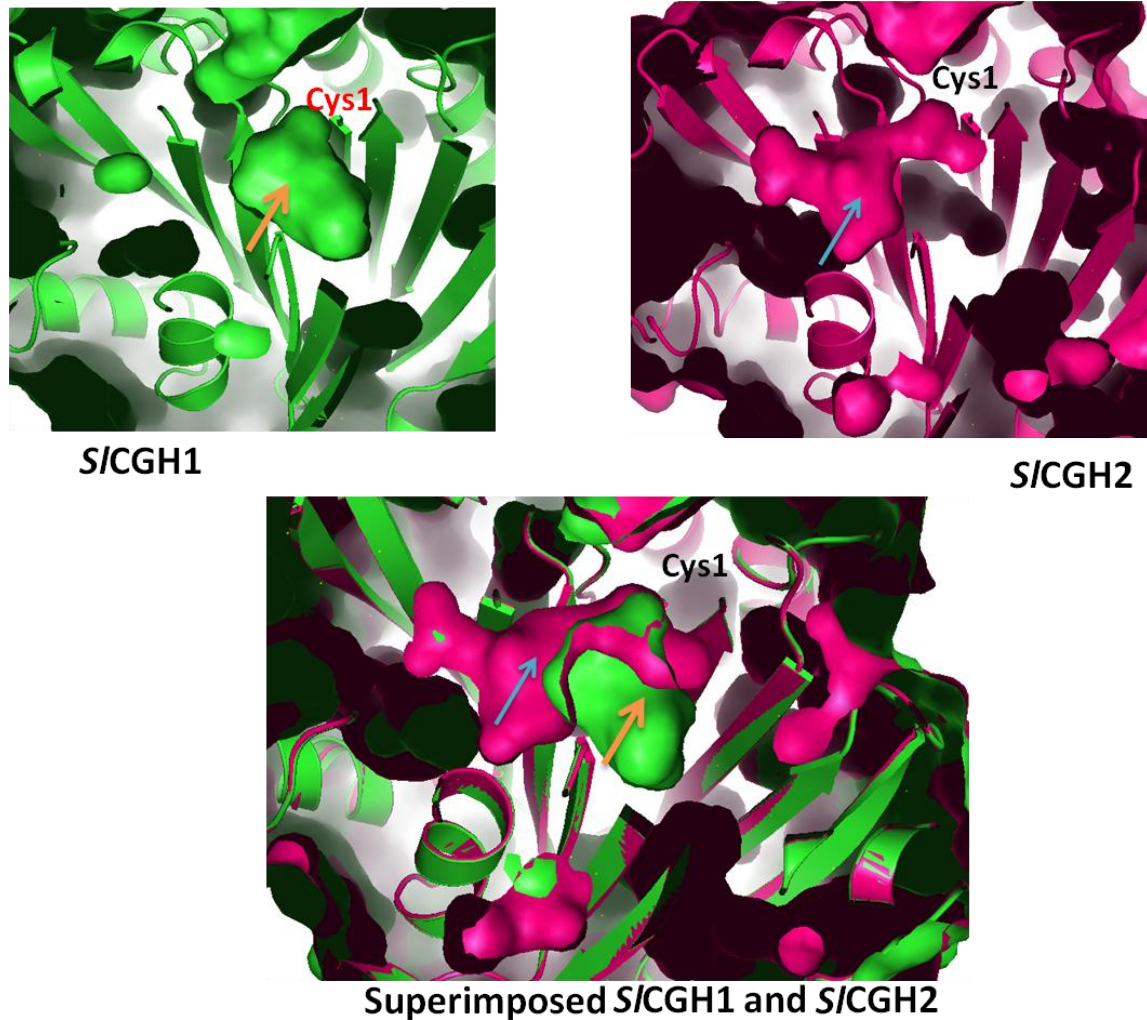
Fig.4. 10. Surface view of *Sl*CGH1 and *Sl*CGH2 superimposed active site, focusing on hydrophobic patch difference between the two. Hydrophobic patch at the active site are indicated by the arrow. Hydrophobic cavity is more elongated in *Sl*CGH2 which might accommodate long chain AHLs easily and hence the activity.

4.3.7. Phylogenetic analysis reveals a separate clade of *Sl*CGH1 homologs

Phylogenetic analysis of 198 CGH sequences by Panigrahi et al (2014) showed the existence of two separate clusters of CGHs where cluster1 included CGHs of Gram positive origin and cluster 2 was composed of CGHs of mostly Gram negative origins. Using the same methodology the homologs obtained from BLAST analysis of *Sl*CGH1 were incorporated to construct a more enhanced phylogenetic tree. The result revealed a distinctly segregated third clade/ cluster made

up of *Sl*CGH1 homologs, apart from the already shown clusters of gram positive and negative CGHs (**Fig. 4.11.a**). Almost all *Sl*CGH1 homologs (>99%) were sequences from marine dwelling bacteria. Unlike other CGHs, the total number of homologous hits was very less (around 100) for *Sl*CGH1. Moreover, it was also surprising that there were only 3 CGHs from *Shewanella* sp. in the list, including *Sl*CGH2. The first two best hits had an identity score of 72-66 % and were un-annotated hypothetical proteins from *Vibrio fluvialis* and *Agarivoran gilvus* respectively, while other hits had <40% identity. These findings may indicate an adaptation to a rarer niche within the marine environment. A large number of distant marine homologs could be identified with a relaxed E-score; however, these formed multiple smaller clusters on their own and were not included in the final phylogenetic tree to avoid complexity. Regardless, the phylogenetic information coupled with the unique activity of *Shewanella* CGHs on AHLs (while being inactive on Pen V or bile salts) provides intriguing information regarding the possible evolution of a new sub-class of CGHs in response to the marine environment. This becomes more evident through the differences in subunit composition of the *Shewanella* CGHs and the structural features in the active site/substrate binding pocket.

Panigrahi et al. (2014) have also employed the Binding site similarity (BSS) score as an improved method for correct annotation of and differentiation between PVA and BSH, based on sequence and available structural information (Panigrahi et al. 2014). BSS essentially calculates the conservation of residues critical for PVA/BSH activity and substrate binding. These included the amino acids in the active site involved in the catalytic mechanism, as well as other residues in the loops 1- 4 surrounding the active site which have been suggested to be essential for proper substrate binding. A strict conservation of these residues with any of the template sequences would essentially translate to a higher BSS, thereby providing conclusive information to categorize the CGH sequence as either a PVA or a BSH.

In the original scoring system, six templates which were either functional PVA or BSH and whose structures were available at the time were used to create the similarity score matrix. However, on employing this algorithm with both *S. loihica* CGHs, no significant similarity could be achieved with any of the six templates used. This observation confirms to the fact that the CGHs of marine origin, including those from *S. loihica*, show significantly lower sequence homology with known CGHs from other environments and cluster separately. Therefore, we
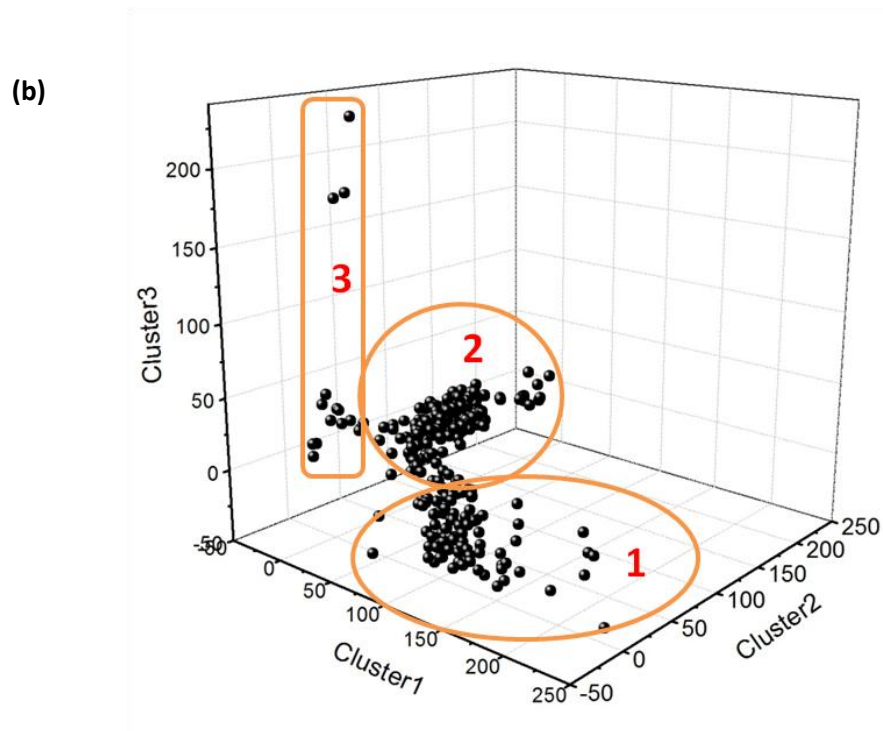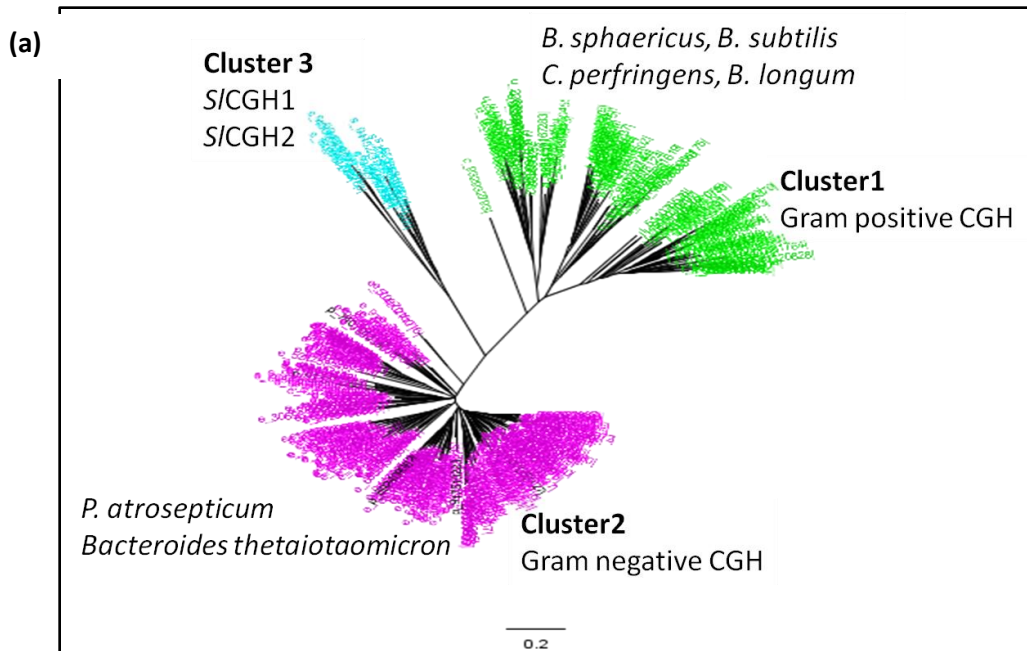
Fig. 4.11. (a) Phylogenetic study shows three different well segregated clusters. Cluster1 represents homologs of Gram positive CGHs; Cluster2 Gram negative CGH homologs and Cluster3 *Sl*CGH1 homologs including *Sl*CGH2. Cluster3 is also predominantly marine bacterial CGHs. (b) BSS score of CGH homologs presented as 3D graph, showing three clusters of enzymes.

added *Sl*CGH1 as another template to extend the scoring system. On recalculating the BSS scores, some of the marine homologs were observed to score higher with the *Sl*CGH1 template; these also showed better whole sequence similarity with *Sl*CGH1. A few sequences (from *Photobacterium swingsii, Vibrio tubiashii ATCC 19109, Shewanella waksmanii* and *Vibrio maritimus*) were ambiguous, showing BSS scores matching with Gram-negative templates *Pa*PVA and *Bt*BSH, in addition to *Sl*CGH1 (**Fig. 4.11.b**). It is possible that these enzymes might also show some activity with Pen V or bile salt. Previously reported CGHs are known to show substrate cross-reactivity. Between the pure PVAs active only on Pen V (*Bsu*PVA, Rathinaswamy et al., 2012 and *Pa*PVA, Avinash et al. 2015) and the pure BSH (*Bl*BSH, Kumar et al., 2006), *Cp*BSH (Coleman and Hudson1995) exhibits mild activity on Pen V and *Bsp*PVA exhibits low BSH activity (Kumar et al., 2006. Recently, two functional PVAs from Gram-negative bacteria (*Pa*PVA, *At*PVA) exhibiting high activity on Pen V have been shown to also degrade long chain AHLs (Avinash et al. 2016). Considering the activity profile of *Sl*CGHs, we suggest that AHLs could also be added to the list of potential substrates for CGH enzymes. Although it is not clear if AHL degrading activity specifically evolved in response to the marine environment, it is important to note that the marine CGH homologs represent a new sub-class of CGHs with unique substrate preference.

4.4. Are *Sl*CGH1 and *Sl*CGH2 AHL acylases ?

Biochemical studies reveal that although both *Sl*CGH1 and *Sl*CGH2 exhibit closer homology to CGHs, they are inactive on standard CGH substrates and instead exhibit AHL acylase activity. To evaluate the novelty of these two enzymes, their phylogenetic relationship with known AHL acylases (which are more homologous to PGA) from various bacteria was determined using the neighbor joining method. Ochiai et al. (2014) studied the phylogenetic relationship among different known AHL acylases and divided them into three families - Aculeacin A acylase protein family (EC 3.5.1.70), penicillin G acylase protein family (EC 3.5.1.11) and amidase family (EC 3.5.1.4). We performed a similar analysis with *Sl*CGH1, *Sl*CGH2 and known eleven AHL acylases - AiiD from *Ralstonia* sp. strain XJ12B, AhlM from *Streptomyces* sp. strain M664 , PvdQ from *P. aeruginosa* PAO1, and Aac from *Shewanella* sp. strain MIB015, AiiC from *Anabaena* sp. strain PCC 7120, AmiE from *Acinetobacter* sp. Ooi24, AmiE-Az from *Azospirillum* sp. B510, AmiE-Am from *A. orientalis* HCCB10007, SlPVA from *Streptomyces*

lavendulae, QuiP and HacB from *P. aeruginosa* PAO1. PVA from *Pectobacterium atrosepticum* (*Pa*PVA) and PGA from *Kluyvera citrophila* (KcPGA) were also included, as they have been reported to exhibit AHL acylase activity besides their PVA and PGA activities. The result of the phylogenetic analysis showed similar clustering of three AHL acylases families as reported before (Ochiai et al. 2014) and an additional fourth cluster was formed with the three CGHs - *Sl*CGH1, *Sl*CGH2 and *Pa*PVA. *Pa*PVA formed again a different clade within the family (**Fig.4.12**). This further confirms that *Sl*CGH1 and *Sl*CGH2 differ from known CGHs as well as AHL acylases and may represent a separate CGH sub-class with AHL hydrolysis activity.
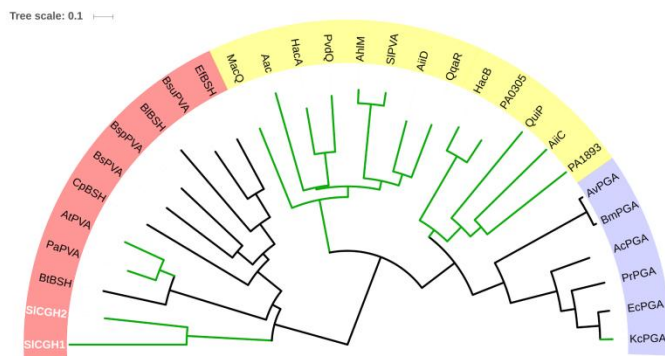
Fig. 4.12. Phylogenetic tree showing different families of AHL acylases based on sequence similarity.

It would be noteworthy, as we compare functionality, that previously characterized CGHs are from inhabitants of soil and gut environments while microbial mats (housing *S. lohica* PV-4) represent an extreme environment in terms of high metabolic rates, strong geochemical gradients inhabited by plethora of diverse microbial population enclosed in a biofilm matrix (Decho et al. 2009, Bolhuis et al. 2014). Decho *et al* (2009) reported abundant existence of a diverse range of AHL signaling molecules, C_4 to C_{14}, in microbial mat and AHLs with less than C_{10} acyl chain length undergo constant alkaline lactonolysis under a diel cycle influence. It is suggested that such fluctuation in the overall signal composition could well be a signal of the local environmental changes for microbes capable of detecting them. Ntn-hydrolases constitute a versatile group of enzymes and exhibition of AHL degrading capability in enzymes from such quorum sensing rich life styles is not surprising. As promiscuous enzyme activity is believed to be a common attribute of organisms of fluctuating environments to assist survival under multiple ecological variables (Martínez-Núñez and Pérez-Rueda, 2016), we need to investigate further

about the involvement of such enzyme in some metabolic pathways as well, to confer its physiological role. PvdQ is an excellent example of dual functionality, being an enzyme of the pyoverdine biosynthesis pathway and exhibiting AHL acylase activity. The *PvdQ* gene is a part of the pyoverdine locus; however in case of *Sl*CGH1, its surrounding genes do not provide an apparent relation as being part of an operon. Our results suggest that *Sl*CGH1 and *Sl*CGH2 have activities similar to the activities of known AHL acylases i.e, to degrade long chain AHLs (Fig 4.13). We assume that adaptation to marine environment provided selective pressure that propels the divergent evolution of CGHs.
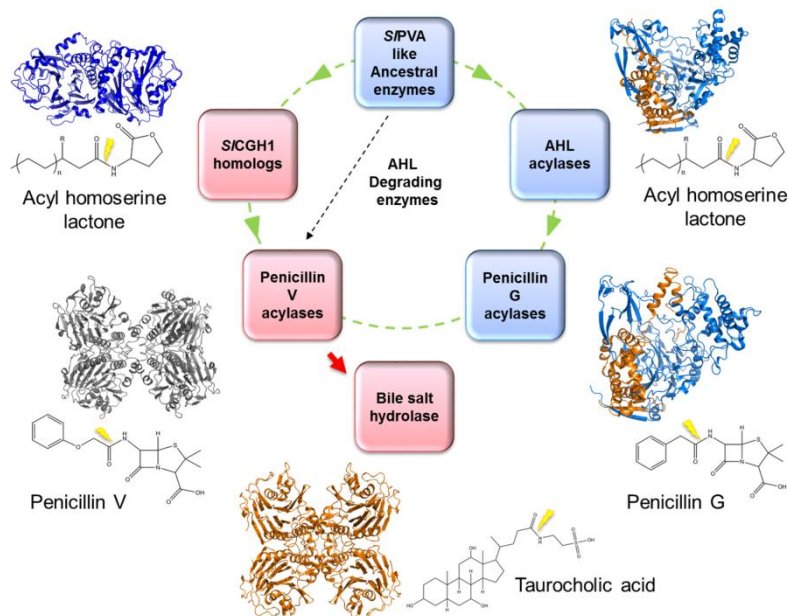
Fig. 4.13. A schematic representation of the probable evolutionary link between different AHL degrading enzyme (PvdQ, *Kc*PGA, *Pa*PVA and *Sl*CGH1) and ⍰-lactum degrading enzymes (*Kc*PGA ,*Pa*PVA, *Sl*PVA), based on activity profile and structural overlaps. A promiscious ancestral protein could be compared to the present day heterodimeric *Sl*PVA which shows a wider promiscuity acting on AHLs as well as both aromatic and aliphatic PenVs. Bile salt hydrolase (*Bl*BSH) might be a younger member evolved for adaptation in the mammalian gut for substrates other than AHLs. Here we assume heterodimers (AHL acylases andPGAs) evolving to homodimer and finally to more complex homotetramers (PVAs and BSHs). Considering the inactivity of *Sl*CGH1 on CGH substrate, independent evolution of PVAs and *Sl*CGH1 homologs remains another possibility.

Chapter 5

Native Penicillin V acylase of *Acinetobacter* sp. AP24, isolated from Loktak Lake - Purification and partial characterization

"Do you have the patience to wait until your mud settles and the water is clear?" - LaoTzu

5.1. Introduction:

Penicillin G acylase (PGA) and penicillin V acylase (PGA) are enzymes of immense pharmaceutical importance as they are responsible for the synthesis of semi synthetic antibiotics. They are produced by a number of bacteria, actinomycetes, yeasts and fungi (Sudhakaran and Borkar 1985). According to Valle et al. (1991) the possible role of these enzymes might be generation of alternative carbon sources by metabolizing aromatic compounds. This theory is in agreement with the experimental finding of *E.coli* penicillin G acylase induction by phenylacetic acid (PAA) and repression by glucose (Merino et al. 2006). However, PGA and PVA are significantly different from one another in terms of their sequence and structure. Recently, penicillin acylases have been implicated in virulence regulation (Kovacikova et al. 2003) and bacterial signalling networks (Candela et al. 2007). Based on these observations, the need for extensive exploration of penicillin acylases to understand more about its *in-vivo* role has been emphasized (Avinash et al. 2014).

In our attempt to search for novel penicillin acylases, water sample from the unexplored Loktak Lake, an Indo-Burma biodiversity hotspot, was screened for penicillin acylase producers. We isolated an intracellular penicillin V acylase producing bacterium, identified as *Acinetobacter* sp. AP24. Genome data search shows presence of penicillin acylase like genes from a number of *Acinetobacter* strains; however experimental activity has never been reported.

5.2. Materials and Methods:

All media components were procured from Hi-Media (India). Penicillin V (potassium salt) and different β-lactam substrates were kind gifts from KDL Biotech (Khopoli, India). Dimethyl aminobenzaldehyde (pDAB) and methanol (HPLC grade) were obtained from Qualigens, India. SP-Sepharose was purchased from Sigma. Protein molecular weight standards were obtained from Merck India. All other reagents and chemicals used were of analytical grade.

5.2.1. Screening for penicillin acylase production:

Water samples were collected from Loktak Lake, Manipur, India (24°33' N and 93°47' E). Bacteria were isolated using standard serial dilution and plating methods on nutrient agar plates. Colonies were purified by sub-culturing and maintained on slants as well as in glycerol stocks (-

80 °C). Primary screening of penicillin acylase producers was done by *Serratia marcescens* ATCC 27117 overlay technique (Meevootisom et al. 1983). *S. marcescens* ATCC 27117 is a penicillin resistant strain; however, it is sensitive to 6-APA and hence any culture that produces penicillin acylase inhibits the growth of *S. marcescens* ATCC 2711. The assay requires freshly prepared 0.8-1% NA and 24 h old culture of *S. marcescens* ATCC 27117 grown in milk agar. The test organism was spot inoculated on NA plate and kept for incubation at 30°C for 48-72 h. The NA plate was overlaid with 10ml of soft NA containing 100µl of the 24 h old *S. marcescens* ATCC 2711 and 20mg/ml PenV. The plates were then incubated overnight at 28°C and were observed for a zone of clearance in the overlaid mat of *S. marcescens* ATCC 27117.

5.2.2. Enzyme assay and protein estimation:

PVA activity was assayed according to the method by Shewale et al. 1987, measuring the amount of 6-APA formed at 40 °C from 2 % w/v solution of penicillin V (potassium salt) in 0.1 M phosphate buffer, pH 7.8. The reaction was stopped by adding 1.0 M citrate phosphate buffer (CPB) pH 2.5 and the 6-APA formed was estimated using 0.6 % (w/v) p-dimethyl-amino-benzaldehyde in methanol. One unit (IU) of PVA activity is defined as the amount of enzyme that produces 1µmol 6-APA per minute under the conditions defined. Protein estimation was done by method of Lowry et al. (1951) using crystalline bovine serum albumin as standard.

5.2.3. Identification of isolate:

Genomic DNA was extracted from the isolate *Acinetobacter* sp. AP24 by CTAB method as described by Wilson (1987). The 16S rRNA gene was amplified using the universal bacterial primers 16F27 (5'AGAGTTTGATCMTGGCTCAG 3') and 16R1525 (5' TTCTGCAGTCTAGAAGGAGGTGWTCCAGCC 3'). PCR was performed under the following conditions: an initial denaturation at 94 °C for 3 min, followed by 35 PCR cycles (94 °C, 30 sec ; 55 °C, 30 sec ; 72 °C 1.30 min) and a final elongation at 72 °C for 7 min. Amplified products were separated on 0.8 % agarose gel and purified using Antarctic phosphatase and exonuclease I treatment. Purified PCR product was sequenced by NCIM Resource Centre, National Chemical Laboratory, Pune, India. The 16S rRNA sequence obtained was compared with those in gene bank after blast searches and using EzTaxon server. The 16S rRNA gene sequence-based phylogenetic tree was constructed using MEGA 6 (Tamura et al. 2013)using the neighbour joining method. The genus *Bacillus* was used as outgroup.

5.2.4. Optimisation of PVA enzyme production:

Different physiochemical parameters for optimum PVA production were investigated. Fermentation media selected for the study were (g /L); (a) Minimal medium (Na_2HPO_4 – 12, K_2HPO_4 – 3.1, $(NH_4)_2SO_4$ – 3, $MgSO_4$ – 0.2, NaCl – 1, Glucose – 4), (b)Yeast malt extract (Yeast extract - 2, malt extract -5, glucose- 10), (c) Nutrient broth- (Peptone – 10, Beef extract – 5 , NaCl –5), (d) 0.5X Nutrient broth, (e) Luria Bertani broth (Tryptone - 10, yeast extract- 5, NaCl - 10) and (f) 0.5X Luria Bertani broth, (g) Starch-casein broth (SCB) (Starch – 10, KNO_3 – 2, K_2HPO_4 – 2, NaCl – 2, Casein - 0.3, $MgSO_4$ – 0.05, $FeSO_4$ -0.01, $CaCO_3$ – 0.02). Yeast malt extract (YME) showed optimum PVA production and was used as throughout the study as the basal medium. A 24 h old inoculum (10 %) was used as the seed culture. Enzyme activity was estimated every 12 h. Effect of aeration on enzyme production was studied by varying medium volumes in 250 ml Erlenmeyer flasks. Effect of incubation temperature and initial pH of the basal medium were also studied.

To study the effect of carbon and nitrogen sources, glucose and yeast in the basal medium (YME) were replaced by various carbon and nitrogen sources respectively. Effect of malt extract was studied by varying its concentration. Influence of various inducers of PVA production; skim milk (0.5 % w/v), corn steep liquor (1 % v/v), penicillin V (0.04 % w/v) and phenoxyacetic acid (0.04 % v/v) were also checked.

5.2.5. Purification of PVA from *Acinetobacter* sp. AP24

For enzyme purification, *Acinetobacter* sp. AP24 strain was grown in optimized medium for 24 h and cells were harvested by centrifugation. Wet cells were suspended in 20 mM phosphate buffer pH 7.5 (1 g in 3 ml buffer) containing 1 mM each of EDTA and DTT, and disrupted by sonication in ice bath for 15 min at 80 % amplitude using Branson Digital Sonifier. After centrifugation at 10,000 rpm for 30 min, the clear supernatant was brought to 0-30 %, then to 50 % saturation by gradual addition of fine powder of ammonium sulphate, while stirring. The suspension was allowed to stand for 6–8 h at 4 ℃. The precipitated protein collected from 30 to 50 % saturation was dialyzed against 10 mM, phosphate buffer pH 7.5, containing 1 mM DTT and 1 mM EDTA. The sample was loaded on SP-sepharose column (2.5 cm × 25 cm), pre-equilibrated with 10 mM phosphate buffer, pH 7.5. The column was washed with 10 mM

phosphate buffer pH 7.5 to remove unbound proteins. The enzyme was eluted subsequently with 0.15 M and 0.2 M NaCl at the same pH. Fractions containing PVA activity were pooled, concentrated in a SpeedVac system. Samples were analyzed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).

5.2.6. Size-Exclusion FPLC:

The molecular weight of the purified protein sample was determined by Gel Permeation chromatography. 100 µl of the sample was applied onto an analytical ENrich SEC 70 10 x 300 mm column fitted to an automated FPLC system (BioRad Labs). Column equilibration and elution were carried out with 10mM Tris-HCl, pH 7.5 containing 100 mM NaCl and 1mM DTT.

5.2.7. Matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF MS):

The mass spectrum was recorded on a Voyager-De-STR (Applied Biosystems) MALDI-TOF. The matrix solution of 15 mg/ ml sinapinic acid was prepared in 30 % acetonitrile (ACN). Enzyme mixed with the matrix solution in 1: 5 ratios was spotted on MALDI target plate and introduced into the mass spectrometer after drying. The spectrum was obtained using a pulsed N_2 laser (337 nm), in linear mode with delayed ion extraction with an accelerating voltage of 25 kV.

5.2.8. Substrate specificity:

For determination of substrate specificity, 5 mg/ml of various beta-lactam substrates containing amide bonds such as penicillin G, amoxicillin trihydrate, ampicillin, cephalosporin C, cephalosporin G, cephalexin, cefaclor and dicloxacillin were tested for activity measurement by standard enzyme assay.

5.2.9. Optimum pH and temperature:

Optimum temperature for maximal PVA activity was determined by measuring activity at different temperatures ranging from 30–70 °C at pH 7.5 in 0.1 M phosphate buffer. To find out optimum pH of enzyme, assay was carried out at 40 °C in the pH range of 4.0–9.0 using 0.1 M

buffers (pH 4.0 and 5.0, acetate buffer; pH 6.0 and 7.0, phosphate buffer; pH 8.0 and 9.0, Tris–HCl).

5.2.10. Thermal and pH stability of PVA:

For thermostability study, residual activity of enzyme incubated at 30, 40, and 50 °C up to 2 h in 0.1 M phosphate buffer pH 7.5 was estimated at intervals. pH stability was also estimated by measuring the residual activity of enzyme after incubation in 0.1 M buffers of pH range 1.0–11.0, using different buffers (pH 1.0 to 2.0, HCl-KCl buffer; pH 3.0, citrate, phosphate buffer; pH 4.0 to 9.0 as described above; pH 10 and 11, carbonate-bicarbonate buffer) for 2 h at 25 °C at intervals.

5. 3. Results and discussion:

5.3.1. Screening and identification of novel PVA producer

Loktak Lake, the largest fresh water lake in North-east India, was initially recognised as a wetland of international importance under the Ramsar Convention (Singh and Shyamananda 1994). As the lake faces various ecological problems, its once rich diversity of flora and fauna is declining. The lake is largely unexplored especially its microbial biodiversity. There are only few reports on study of anti-microbial properties of bacteria mostly *Streptomyces*, isolated from the lake sediment (Singh et al. 2006; Ningthoujam et al. 2009; Singh et al. 2009) and cyanobacterial diversity of the lake (Akoijam and Singh 2011; Singh and Dhar 2011). In the current study, we screened water samples from the lake for potent penicillin acylase producers. A bacterial isolate (AP24) showed PVA activity in the bioassay by *Serratia marcescens* overlay technique (Meevootisom et al. 1983), where zone of clearance was observed around the isolate against a mat of red coloured growth of *S. marcescens* ATCC 27117 with penicillin V (**Fig. 5.1**). Intracellular PVA activity was confirmed by standard enzyme assay.
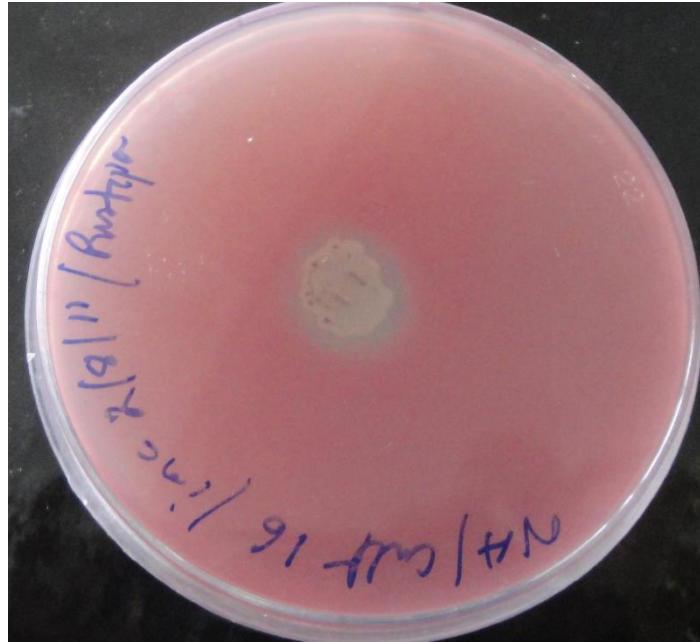
Fig. 5.1. Plate showing zone of inhibition around isolate AP24 as a result of 6-APA induced inhibition of growth of *Serratia marcescens* ATCC 27117.

Identification by16S rRNA gene sequence analysis confirmed the genus of the isolate AP24 as *Acinetobacter* (**Fig. 5.2**). Morphological and biochemical studies showed that the cell were Gram-negative cocco-bacillary rods, non-motile, catalase positive and oxidase negative, which are typical of the genus. The isolate was non-haemolytic when grown on blood agar, suggesting a non-pathogenic nature.

Our strain, AP24 shared closest similarity with *Acinetobacter venetianus* RAG-1 (98.11 %), *A. gyllenbergii* 1271 (98.11 %), *A. baumannii* ATCC19606 (97.82 %), *A. junni* CIP64.5 (97.54 %) and *A. beijerenckii* 58a (97.35 %). Phylogenetically, strain AP24 clustered with *A.venetianus*. The 16S rRNA sequence analyzed was deposited in NCBI GenBank database under the accession number KJ584607.

Bacteria of the *Acinetobacter* genus have been associated with a number of ecological and biotechnological applications; production of bioemulsifiers, lipases, proteases and biodegradation of numerous pollutants such as phenolic compounds, crude oil and heavy metals etc. (Abdel-El-Haleem 2003). Species like *A. baumannii* are notorious for their role in multidrug resistant (MDR) nosocomial infections (Fournier and Richet 2006). Moreover, Vallenet et al

(2008) reported the presence of putative penicillin acylase (*pac*) gene with unknown substrate specificity in the genome of *A. baumannii* AYE- a MDR human isolate. Genome databases also show the presence of penicillin acylase-like genes in many *Acinetobacter* strains. Considering the widespread presence of penicillin acylase-like gene sequences in this genus and its clinical importance, *Acinetobacter* can be considered a good model organism for the study of regulation of pencillin acylase.
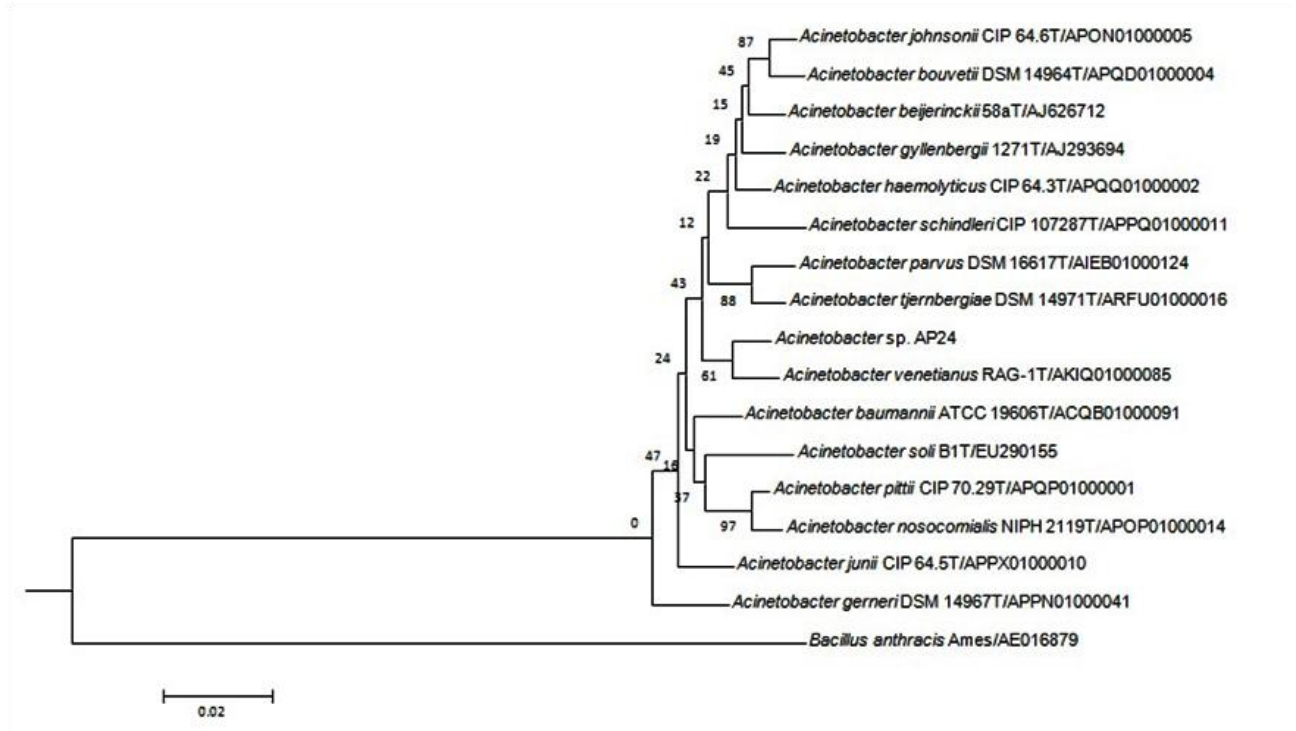
Fig. 5.2. Phylogenetic tree based on 16S rRNA of isolate, *Acinetobacter* sp. AP24.

5.3.2. Fermentation studies:

PVA production from *Acinetobacter* sp. AP24 was favoured by fermentation at 28 °C for 24 h under micro-aerophilic condition (100 ml of medium in 250 ml Erlenmeyer flask) with an initial medium pH of 7.0. Among the various carbon sources studied, glucose (1 %) gave maximum enzyme production (**Table 5.1a**) with no repression of enzyme production at higher concentration. Similar enhancement in PVA production by glucose has been reported in *Erwinia aroideae, Bacillus sphaericus* and *B. cereus* (Vandamme and Voets 1975; Carlsen and Emborg

1981; Sunder et al. 2012). On the hand, PVA production in *Streptomyces lavendulae* (Torres et al. 1999) and *Chainia* (Chauhan et al. 1998) was affected by catabolite repression in the presence of glucose.

Table 5.1. Effect of different (a) carbon sources and (b) nitrogen sources on PVA production.

(a) Carbon source	PVA activity (IU/g)	Cell weight (g/L)
Glucose	2.35	1.30
Fructose	1.55	2.42
Lactose	1.60	2.30
Starch	1.30	2.65
Glycerol	1.0	2.60
Xylose	1.20	3.50
(b) Nitrogen source	PVA activity (IU/g)	Cell weight (g/L)
Beef extract	0.77	1.15
Yeast extract	0.60	1.30
Peptone	0.55	1.25
Tryptone	0.83	1.05
Potassium nitrate	0.93	1.20
Sodium glutamate	1.85	1.32
Ammonium sulfate	0.54	1.28

Maximum PVA production was obtained with 0.2 % sodium glutamate as the nitrogen sources (**Table 5.1b**). Sodium glutamate enhanced PVA production has also been reported by Ambedkar et al. (1991) in *Beijerinckia indica var. penicillanicum*. Inducers known to increase PVA production in some bacteria, such as corn steep liquor (Pundle and Sivaraman 1997), phenoxyacetic acid (Deshpande et al. 1997), skim milk and penicillin V (Zhang et al. 2007)

showed no effect on PVA production in our study; hence the enzyme production might be considered constitutive. The PVA production was increased three fold to 2.3 IU/g after optimisation of various physical and nutritional parameters.

5.3.3. Purification of *Acinetobacter* sp. AP24 PVA

PVA from *Acinetobacter* sp. AP24 has been purified to electrophoretic homogeneity by a single chromatography step of cation exchange using SP-sepharose. Pure enzyme was eluted with 0.2 M NaCl. The purified enzyme had a specific activity of 11.9 IU/mg (**Table 5.2.**) which is higher than PVAs characterized in recent reports (Kumar, et al. 2008; Rathinaswamy et al. 2012). Purity of the enzyme was confirmed by SDS PAGE and MALDI-TOF MS analysis (**Fig. 5.3**). In MALDI-TOF, a prominent peak of 34 KDa molecular weight was observed, which was in agreement with the result seen in SDS PAGE. In previous reports, molecular weight of PVA monomers varies from 30-35 KDa (Torres et al. 1998). The molecular weight of the native enzyme obtained from gel permeation chromatography was 137,000±1000 Da, which suggested that PVA from *Acinetobacter* sp. AP24 is a homotetramer. *Bacillus sphaericus* PVA is also a homotetrameric protein of 138 kDa native and 35 kDa monomeric molecular weight (Pundle and Sivaraman, 1997).

Table 5.2. Purification table of PVA from *Acinetobacter* sp. AP24

Purification steps	Volume (ml)	Total protein (mg)	Total units (IU)	Specific activity (IU/mg)	Fold purification	Yield (%)
Cells free extract	34	284	21.75	0.074	1	100
Ammonium sulphate fractionation	4	54	17	0.1	1.35	78
SP-sepharose chromatography	4	1.14	13.57	11.9	119	62.4
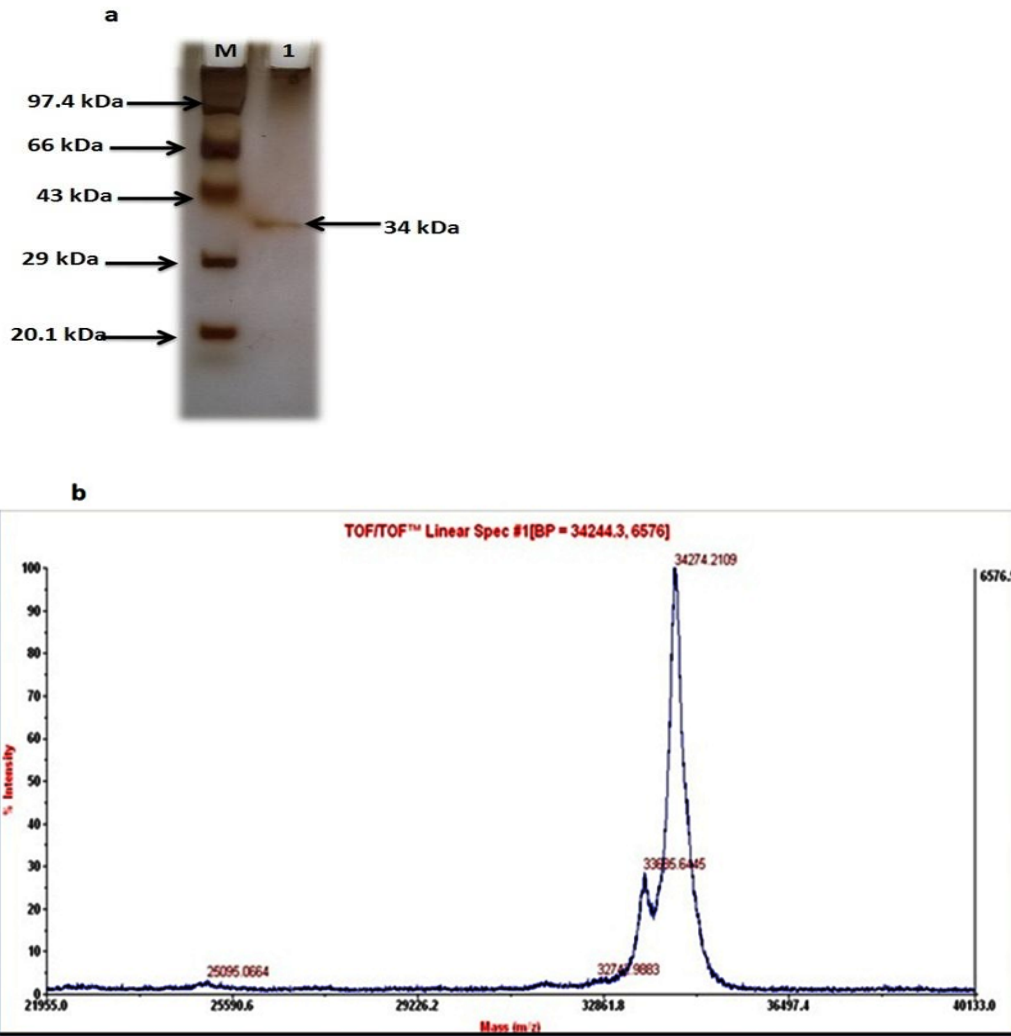
Fig.5.3. Molecular weight of a single subunit of PVA from *Acinetobacter* sp. AP24 was confirmed 34KDa, (a) SDS-PAGE on 12 % gel at 100 V and silver staining (Lane M, protein marker; lane 1, purified PVA) , (b) MALDI-TOF spectrum of the purified enzyme.

5.3.4. Partial characterization of *Acinetobacter* sp. AP24 PVA

Penicillin V was the best substrate for the enzyme. The enzyme had very little or no activity on ampicillin, amoxicillin, cephalosporin C, cefaclor, cloxacillin and dicloxacillin. Such high specificity towards penicillin V has been observed in PVA from *B. sphaericus* (Pundle and Sivaraman, 1997), *Erwinia aroideae* (Vandamme and Voets 1975) and *Pectobacterium atrosepticum* (Avinash et al. 2014). Whereas PVA from *S. mobaraensis* (Zhang et al. 2007) and

Rhodotorula aurantiaca (Kumar et al. 2008) have shown hydrolytic activity towards various other β-lactam antibiotics.

Table 5.3. Substrate specificity expressed as relative activity. Results shown as average of two duplicate assays.

Substrate	Relative Acitvity
Penicillin V	100
Penicillin G	0.84
Cephalosporin C	2.4
Cephalosporin G	0
Dicloxacillin	0
Ammoxycillin trihydrate	2.4
Ampicillin	0
Cephalexin	0
Cefaclor	0

Optimum conditions for PVA activity were 40 ° C (**Fig. 5.4a**) and at pH 7.5 (**Fig. 5.4b**). The temperature optimum is similar to that of PVA from *S. lavendulae* (Torres et al 1998). PVA from *Fusarium oxysporum* showed pH optima in a similar range of 7.4 to 7.6 (Sheng and Ye 1989). On the other hand, bacterial PVAs usually display maximum activity in acidic conditions (Shewale and Sudhakaran 1997). The enzyme retained 75 % activity for 1 h at 40 ° C (**Fig. 5.5**); Almost 80 % activity was retained from pH 5.0 to 9.0 for 2 h with 100 % activity retention at pH 7.0 (**Fig. 5.6**); however enzyme precipitated instantly in pH 1.0 and after an hour in pH 2.0. Such stability at low pH (< 7.0) is desirable during synthesis of semi-synthetic penicillin (Wegman et al. 2001). Further characterization of the enzyme would be necessary to reveal more information about its industrial applicability.
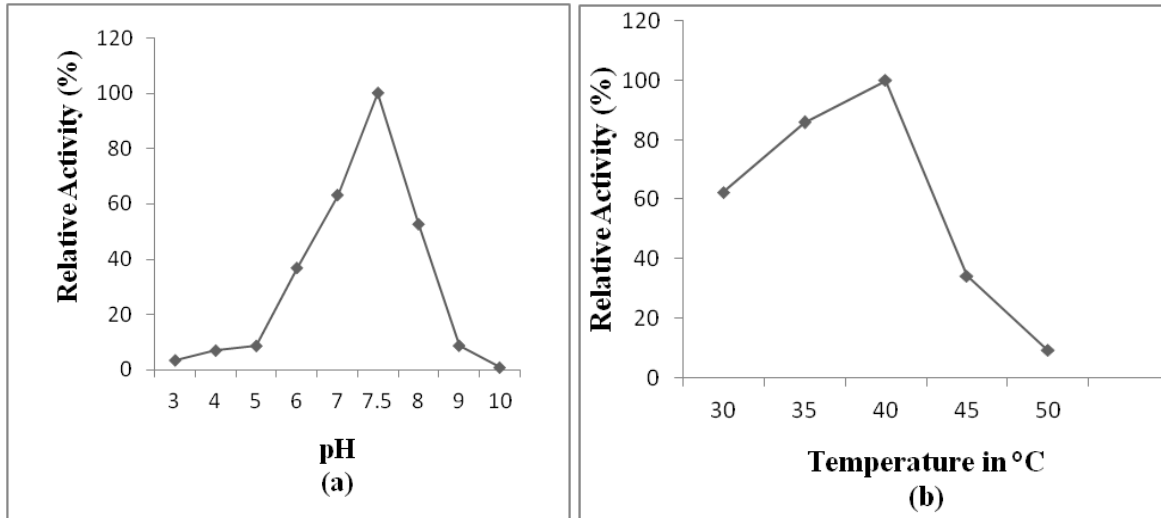
Fig. 5.4. Optimum temperature (a) of PVA activity studied at pH 7.5 in 0.1 M phosphate buffer at different temperatures and (b) optimum pH of PVA activity studied at 40 °C, at different pH. Results shown with ±SD.
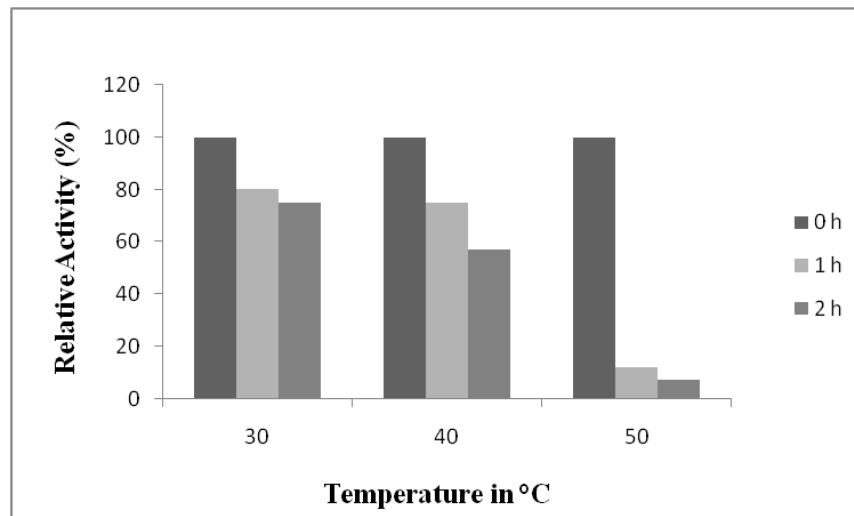
Fig. 5.5. Temperature stability profile. Residual activities measured at 40°C after incubation at 27, 40 and 50°C for 2h.
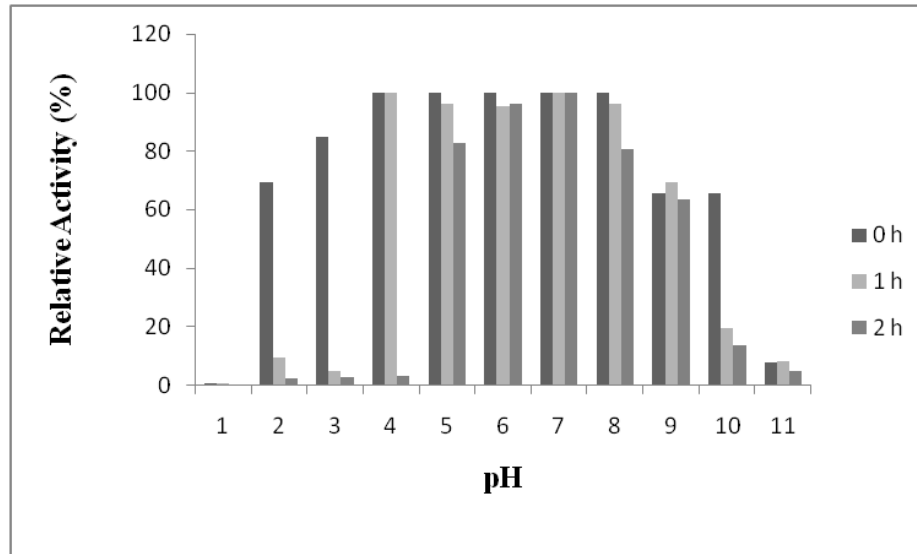
Fig. 5.6. pH stability profile. Residual activities measured at 40°C after incubation at different pH ranging from 1.0 to 11.0 for 2 h. Result shown with ±SD.

5.4. Conclusion:

Acinetobacter sp. AP-24 produces intracellular PVA in a constitutive manner. The PVA enzyme has been purified to homogeneity in a single chromatography step and is highly specific for penicillin V. Given the clinical and environmental importance of *Acinetobacter,* this enzyme holds great potential for research. Cloning of the *Acinetobacter* sp. AP-24 PVA is under progress and structural studies will be carried out for a deeper molecular understanding of the enzyme.

Summary and future prospects

The Ntn hydrolase enzyme superfamily is a consortium of structurally related enzymes that contain an N-terminal nucleophilic residue important for catalysis, and exhibit acylase/amidase activity on a wide range of substrates. The choloylglycine hydrolase (CGH) group includes two cys-Ntn hydrolases: penicillin V acylase (PVA, E.C. 3.5.1.11) and bile salt hydrolase (BSH, EC 3.5.1.24). Although CGHs occupy a significant place in the pharmaceutical/probiotic industry, their role in the natural environment has remained largely unexplored till recent years. Studies so far have pertained to soil and gut inhabiting bacterial CGHs, leaving a huge avenue for exploring CGHs from other sources or habitats. In this thesis, two marine CGHs and one from fresh water sample have been studied. The activity and substrate specificity of these enzymes have been explored using biochemical, structural and computational analyses.

The Gram negative, marine CGHs, *Sl*CGH1 and *Sl*CGH2, characterized in this work demonstrated unique activity profiles and active site structural features. Both the enzymes could not hydrolyze PenV and bile salts, i.e. neither PVA nor BSH activity. However, they showed good acylase activity on AHLs. Although there is a recent report of a functional/active PVA showing promiscuous AHL acylase activity, the present results from *S. loihica* CGHs are significant, as the enzymes have shown complete absence of BSH and PVA activities. The two enzymes differ in their preference of the type of AHLs, based on acyl chain length with *Sl*CGH1 showing peak activity on C_{10}-HSL and *Sl*CGH2 on C_{14}-HSL. We suggest that these could be paralogs evolved to adapt the marine environment or more specifically a fluctuating microbial mat which hosts them. Phylogenetic analysis of these enzymes among CGH sequences revealed that *Sl*CGH1 and *Sl*CGH2 homologs, which are predominantly of marine origin, segregated as a distinctly different clade. All these findings strongly supported the possibility of our enzymes to be a new sub-class of CGH with exclusively AHL acylase activity.

We have also reported first PVA activity from the genus *Acinetobacter* isolated from Loktak Lake, an Indo-Burma biodiversity hotspot. The PVA was a tetramer and demonstrated stability at low pH which is a trait desirable for industrial production of 6-APA.

Further exploration of these enzymes from *Shewanella* and other homologs could provide information on the probable place of marine CGHs as well as their role in the microbial physiology. Proteomics and metabolomic approaches for study of CGHs, which have not been reported before, would add new insights into the role of CGHs in the lifestyle of the host organisms.

References

- Abdel-El-Haleem D (2003) Acinetobacter: Environmental and biotechnological applications. *Afri J Biotechnol* 2:71–74.

- Akoijam C, Singh AK (2011) Molecular typing and distribution of filamentous heterocystous cyanobacteria isolated from two distinctly located regions in North-Eastern India. *World J Microbiol Biotechnol* 27:2187-2194.

- Ambedkar SS, Deshpande BS, Sudhakaran VK, Shewale JG (1991) *Beijerinckia indica* var *penicillanicum* penicillin V acylase - enhanced enzyme-production by catabolite repression-resistant mutant and effect of solvents on enzyme activity. *J Industr Microbiol* 7:209-214.

- Arroyo M, de la Mata I, Acebal C, Pilar Castillón M (2003) Biotechnological applications of penicillin acylases: state-of-the-art. *Appl Microbiol Biotechnol* 60:507–514.

- Avinash VS (2016) PhD thesis. Penicillin V Acylases From Gram-negative Bacteria: Biochemical And Structural Aspects. *University of Pune*, Pune.

- Avinash VS, Panigrahi P, Chand D, Pundle A, Suresh CG, Ramasamy S (2016) Structural analysis of a penicillin V acylase from *Pectobacterium atrosepticum* confirms the importance of two Trp residues for activity and specificity. *J Struct Biol* 193:85–94.

- Avinash VS, Pundle AV, Ramasamy S, Suresh CG (2016) Penicillin acylases revisited: importance beyond their industrial utility. *Crit Rev Biotechnol* 36:303–316.

- Avinash VS, Ramasamy S, Suresh CG, Pundle A (2015) Penicillin V acylase from *Pectobacterium atrosepticum* exhibits high specific activity and unique kinetics. *Int J Biol Macromol* 79:1–7.

- Bartlett GJ, Porter CT, Borkakoti N, Thornton JM (2002) Analysis of catalytic residues in enzyme active sites. *J Mol Biol* 324:105–121.

- Bashton M, Chothia C (2007) The generation of new protein functions by the combination of domains. *Structure* 15:85–99.

- Begley M, Sleator RD, Gahan CG, Hill C (2005) Contribution of the three bile-associated loci, *bsh*, *pva*, and *btlB*, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes*. *Infect Immun* 73:894–904.

- Blow D (2002) Outline of Crystallography for Biologists. *New York,Oxford University Press.*

- Blundell TL, Elliott G, Gardner SP, Hubbard T, Islam S (1989) Protein engineering and design.

Philos. Trans.R. Soc. London Set. B. 324-447.

- Bokhove M, et al. (2010) Structures of an Isopenicillin N Converting Ntn-Hydrolase Reveal Different Catalytic Roles for the Active Site Residues of Precursor and Mature Enzyme. *Structure* 18(3):301–308.

- Bokhove M, Jimenez PN, Quax WJ, Dijkstra BW (2010) The quorum-quenching *N*-acyl homoserine lactone acylase PvdQ is an Ntn-hydrolase with an unusual substrate-binding pocket. *Proc Natl Acad Sci* 107(2):686–691.

- Bolhuis H, Cretoiu MS, Stal LJ (2014) Molecular ecology of microbial mats. *FEMS Microbiol Ecol* 90:335–350.

- Bolton E, Wang Y, Thiessen PA, Bryant SH (2008) PubChem: Integrated platform of mall molecules and biological activities, in: *Annual Reports in Computational Chemistry*, Elsevier, Washington DC, 217–241.

- Bompard-Gilles C, Remaut H, Villeret V, Prangé T, Fanuel L, Delmarcelle M, Joris B, Frère J-M, Beeumen VJ (2000) Crystal structure of a D-aminopeptidase from *Ochrobactrum anthropi*, a new member of the "penicillin-recognizing enzyme" family. *Structure* 8:971–980.

- Bomstein J, Evans WG (1965) Automated colorimetric determination of 6-amino penicillanic acid in fermentation media. *Anal Chem* 37:576–578.

- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal Biochem* 72: 248-254.

- Brannigan JA, Dodson GG, Duggleby HJ, Moody PCE, Smith JL, Tomchick, DR and Murzin AG (1995) A protein catalytic framework with an N-terminal nu- cleophile is capable of self-activation. *Nature* 378:416–419.

- Brint JM, Ohman DE (1995) Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhlR-RhlI, another set of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family. *J Bacteriol* 177:7155–7163.

- Brunger AT (1992) Free R-value: a novel statistical quantity for assessing the accuracy of crystal-structures. *Nature* 355:472-475.

- Buchner E, Rapp R (1901) Alcoholic fermentation without yeast cells. *Berichte Der Dtsch Chem Gesellschaft* 34:1523–1530.

- Candela M, Bergmann S, Vici M, Vitali B, Turroni S, Eikmanns JB, Hammerschmidt S, Brigidi P (2007) Binding of Human Plasminogen to Bifidobacterium. *J Bacteriol* 189:5929–

5936.

- Carlsen F, Emborg C (1981) *Bacillus sphaericus* V-penicillin acylase .1. Fermentation. *Biotechnol Lett* 3:375-378.

- Chand D (2016) Structure-function study of Ntn hydrolase enzymes penicillin acylase and bile salt hydrolase. *Academy of Scientific and innovative Research (AcSIR)*, New Delhi.

- Chauhan S, Nichkawade A, Iyengar MRS, Chattoo BB (1998) *Chainia* penicillin V acylase: Strain characteristics, enzyme immobilization, and kinetic studies. *Current Microbiol* 37:186-190.

- Coleman JP, Hudson LL (1995) Cloning and characterization of a conjugated bile acid hydrolase gene from *Clostridium perfringens*. *Appl. Environ Microbiol* 61:2514-2520.

- Conant GC, Wolfe KH (2008) Turning a hobby into a job: how duplicated genes find new functions. *Nat Rev Genet* 9:938–50.

- Copley S (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr Opin Chem Biol* 7:265–272.

- Cupp-Enyard C (2008) Sigma's Non-specific Protease Activity Assay - Casein as a Substrate. *J Vis Exp* e899:1–3.

- Czajkowski R, Krzyzanowska D, Karczewska J, Atkinson S, Przysowa J, Lojkowska E, Williams P, Jafra S (2011) Inactivation of AHLs by *Ochrobactrum* sp. A44 depends on the activity of a novel class of AHL acylase. *Environ Microbiol Rep* 3:59–68

- Dambekodi PC, Gilliland SE (1998) Incorporation of cholesterol into the cellular membrane of *Bifidobacterium longum*. *J Dairy Sci* 81:1818-24.

- Daopin S, Davies DR, Schlunegger MP, and Grütter MG (1994) Comparison of two crystal structures of TGF-b 2: the accuracy of refined protein structures. *Acta Crystallogr D* 50:85-92.

- Decho AW, Norman RS, Visscher PT (2010) Quorum sensing in natural environments: emerging views from microbial mats. *Trends Microbiol* 18:73–80.

- Delpino MV, Marı́a IM, Silvia ME, Diego JC, Juliana C, Carlos AF, Pablo CB (2007) A bile salt hydrolase of *Brucella abortus* contributes to the establishment of a successful infection through the oral route in mice. *Infect Immun* 75:299–305.

- Deshpande BS, Ambedkar SS, Shewale JG (1996) Cephalosporin C acylase and penicillin V acylase formation by *Aeromonas* sp ACY 95. *World J Microbiol Biotechnol* 12:373-378.

- Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in

macromolecular crystallography. *Nat Struct Biol* 4:269-275.

- Diggle SP, Gardner A, West SA, Griffin AS (2007) Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philos Trans R Soc B-Biological Sci* 362:1241–1249.

- Ditzel L, Huber R, Mann K, Heinemeyer W, Wolf DH, Groll M. (1998) Conformational constraints for protein selfcleavage in the proteasome. *J Mol Biol* 279:1187–1191

- Doolittle RF (1995) The multiplicity Of domains in proteins. *Ann Rev Biochem* 64:287-114.

- Drenth J and Haas C (1998) Nucleation in protein crystallization. *Acta Crystallogr* D 54:867-872.

- Duggleby HJ, Tolley SP, Hill CP, Dodson EJ, Dodson G, Moody PCE (1995) Penicillin acylase has a single aminoacid catalytic center *Nature* 373:264–268.

- Dussurget O, Cabanes D, Dehoux P, Lecuit M, the European *Listeria* Genome Consortium, Buchrieser C, Glaser P and Cossart P (2002) *Listeria monocytogenes* bile salt hydrolase is a PrfA-regulated virulence factor involved in the intestinal and hepatic phases of listeriosis. *Mol Microbiol* 45:1095–1106.

- Elkins JM, Kershaw NJ, Schofield CJ (2005) X-ray crystal structure of ornithine acetyltransferase from the clavulanic acid biosynthesis gene cluster. *Biochem J* 385:565–573.

- Engh RA, Huber R (1991) Accurate bond and angle parameters for X-ray proteinstructure refinement. *Acta Crystallogr A* 47:392-400.

- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D* 62:72– 82.

- Farrand SK, Qin Y, Oger P (2002) Quorum-Sensing System of Agrobacterium Plasmids: Analysis and Utility. *Methods Enzymol* 358:452–484.

- Fersht A, Julet MR, Britch J (1997) Structure and mechanicsm in protein science - *A. Fersht (W H Freeman, 1999)*.

- Freisner RA, Murphy BR, Repasky PM, Frye LL, Jeremy GR, Halgren AT, Sanschagrin CP, Mainz TD (2006) Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J Med Chem* 6177–6196.

- Fuqua C, Parsek MR, Greenberg EP (2001) Regulation of Gene Expression by Cell-to-Cell Communication: Acyl-Homoserine Lactone Quorum Sensing. *Ann Rev Genet* 35:439–468.

- Fuqua WC, Winans SC (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J Bacteriol* 176:2796–2806.

- Gao H, Obraztova A, Stewart N, Popa R, Fredrickson KJ, Tiedje MJ, Nealson HK, Zhou J (2006) *Shewanella loihica* sp. nov., isolated from iron-rich microbial mats in the Pacific Ocean. *Int J Syst Evol Microbiol* 56(8):1911–1916.

- Garman EF, Schneider TR (1997) Macromolecular Cryocrystallography. *J Appl Cryst* 30: 211-237.

- Groll M, Ditzel L, Löwe J, Stock D, Bochtler M, Bartunik HD, Huber R (1997) Structure of 20S proteasome from yeast at 2.4Å resolution. *Nature* 386:463–471.

- Gutteridge A, Thornton J (2004) Conformational change in substrate binding, catalysis and product release: an open and shut case? *FEBS Lett* 567:67–73.

- Hassell AM, An G,Bledsoe R K, Bynum MJ, Carter H L III, Deng SJ, Gampe RT, Grisard TE,Madauss KP, Nolte TR, Rocque WJ, Wang L, Weaver KL,Williams SP, Wisely BG, Xua R, Shewchuk LM (2007) Crystallization of protein – ligand complexes. *Acta Crystallogr D* 4449:72–79.

- Hevehan D, Clark De BE (1997) Oxidative renaturation of lysozyme at high concentrations. *Biotechnol Bioeng* 54:221–230.

- Hofmann AF, Hagey LR (2008) Bile acids: Chemistry, pathochemistry, biology, pathobiology, and therapeutics. *Cell Mol Life Sci* 65:2461–2483.

- Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38(Web Server):W545–W549.

- Hooper NM (1994) Families of zinc metalloproteases. *FEBS Lett* 354: l-6

- Huang JJ, Han JI, Zhang LH, Leadbetter JR (2003) Utilization of acyl- homoserine lactone quorum signals for growth by a soil pseudomonad and *Pseudomonas aeruginosa* PAO1. *Appl Environ Microbiol* 69:5941-5949.

- Huang JJ, Petersen A, Whiteley M, Leadbetter JR (2006) Identification of QuiP, the product of gene PA as the second acyl-homoserine lactone acylase of *Pseudomonas aeruginosa* PAO1. *Appl Environ Microbiol* 72:1190-1197

- Huang JJ, Han JI, Zhang LH, Leadbetter JR (2003) Utilization of acyl-homoserine lactone quorum signals for growth by a soil pseudomonad and *Pseudomonas aeruginosa* PAO1. *Appl Environ Microbiol* 69:5941–5949.

- Hubbard TJP, Ailey B, Brenner SE, Murzin AG, Chothia C (1999) SCOP: a Structural Classification of Proteins database. *Nucleic Acids Res* 27:254–256.

- Guo HC, Xu Q, Buckley D, Guan C (1998) Crystal Structures of *Flavobacterium* glycosylasparaginase. 273:20205–20212

- Illergård K, Ardell DH, Elofsson A (2009) Structure is three to ten times more conserved than sequence - A study of structural response in protein cores. *Proteins Struct Funct Bioinforma* 77:499–508.

- Isupov NM, Obmolova G, Butterworth S, Badet-Denisot MA, Badet B, Polikarpov I, Littlechild AJ, 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.

- James LC, Tawfik DS (2001) Catalytic and binding poly-reactivities shared by two unrelated proteins: The potential role of promiscuity in enzyme evolution. *Protein Sci* 10:2600–2607.

- Jarocki P, Podleśny M, Glibowski P, Targoński Z (2014) A new insight into the physiological role of bile salt hydrolase among intestinal bacteria from the Genus *Bifidobacterium*. *PLoS One* 9:e114379.

- Jones BV, Begley M, Hill C, Gahan CGM, Marchesi JR (2008) Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc Natl Acad Sci* 105:13580–13585.

- Kabsch W (2010) XDS. *Acta Crystallogr D* 66:125–132.

- Kang Y, Tran A, White RH, Ealick SE (2007) A novel function for the N-terminal nucleophile hydrolase fold demonstrated by the structure of an archaeal Inosine monophosphate cyclohydrolase. *Biochem* 46:5050–5062.

- Kasche V, Lummer K, Nurk A, Piotraschke E, Rieks A, Stoeva S, Voelter W (1999) Intramolecular autoproteolysis initiates the maturation of penicillin amidase from *E. coli*. *Biochim Biophys Acta* 1433:76–86

- Khersonsky O, Tawfik DS (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem* 79:471–505.

- Kim JH, Krahn JM, Tomchick DR, Smith JL and Zalkin H (1996) Structure and function of the glutamine phosphoribosylpyrophosphate amidotransferase glutamine site and communication with the phosphoribosylpyrophosphate site. *J Biol Chem* 26:15549–15557

- Kim JK, Yang SI, Shin JH, Cho JK, Ryu KE, Kim HS, Park SS, Kim KH (2006) Insight into

autoproteolytic activation from the structure of cephalosporin acylase: a protein with two proteolytic chemistries. *Proc Natl Acad Sci* 103:1732–7.

- Klinman JP, Kohen A (2013) Hydrogen tunneling links protein dynamics to enzyme catalysis. *Annu Rev Biochem* 82:471–96.

- Koul S, Kalia VC (2017) Multiplicity of quorum quenching enzymes: A potential mechanism to limit quorum sensing bacterial population. *Indian J Microbiol* 57:100–108.

- Kovacikova G, Lin W, Skorupski K (2003) The virulence activator AphA links quorum sensing to pathogenesis and physiology in *Vibrio cholerae* by repressing the expression of a penicillin amidase gene on the small chromosome. *J Bacteriol* 185:4825–4836.

- Kumar A, Prabhune A, Suresh CG, 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 RS, Brannigan AJ, Prabhune AA, Pundle VA, Dodson GG, Dodson E, Suresh CG (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.

- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.

- Lambert JM, Bongers RS, De Vos WM, Kleerebezem M (2008) Functional analysis of four bile salt hydrolase and penicillin acylase family members in *Lactobacillus plantarum* WCFS1. *Appl Environ Microbiol* 74:4719–4726.

- Laskowski RA, McArthur MW, Moss DS, Thornton J (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:282-291.

- Lee SY, Kim WH, Parks SS (2000) The role of α-amino group of the N-terminal serine of α subunit for enzyme catalysis and autoproteolytic activation of glutaryl 7-aminocephalosporanic acid acylase. *J Biol Chem* 275:39200–39206.

- Lin YH, Xu JL, Hu J, Wang LH, Leong SO, Leadbetter RJ, Zhang LH (2003) Acyl-homoserine lactone acylase from *Ralstonia* strain XJ12B represents a novel and potent class of quorum-quenching enzymes. *Mol Microbiol* 47:849–860.

- Lin YH, Xu JL, Hu J, Wang LH, Ong SL, Leadbetter JR, Zhang LH (2003) Acyl-homoserine lactone acylase from *Ralstonia* strain XJ12B represents a novel and potent class of quorum-

quenching enzymes. *Mol Microbiol* 47:849–860.

- Lindsay JP, Clark DS, Dordick JS (2002) Penicillin amidase is activated for use in non-aqueous media by lyophilizing in the presence of potassium chloride. *Enzyme Microb Technol* 31:193–197.

- Liu YC, Chang WM, Lee CY (1999) Effect of oxygen enrichment aeration on penicillin G acylase production in high cell density culture of recombinant E. coli. *Bioprocess Eng* 21:227–230.

- Lodola A, Branduardi D, Capoferri MVL, Mor M, Piomelli D, Cavalli A (2012) A catalytic mechanism for cysteine N-terminal nucleophile hydrolases, as revealed by free energy simulations. *PLoS One* 7: e32397.

- Löwe J, Stock D, Jap B, Zwickl P, Baumeister W, Hubert R (1995) Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4Å resolution. *Science* 268:533–539.

- Lynch M, Force A (2000) The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154:459–473.

- Maisuria VB, Nerurkar AS (2015) Interference of Quorum Sensing by *Delftia* sp. VM4 Depends on the Activity of a Novel N-Acylhomoserine Lactone-Acylase. *PLoS One* 10:e0138034.

- Mart MA, Stuart AC, Roberto S, Melo F, Andrej S (2000) Comparative Protein Structuremodeling Of Genes And Genomes. *Annu Rev Biophys Biomol Struct* 29:291–325.

- Martínez-Núñez MA, Pérez-Rueda E (2016) Do lifestyles influence the presence of promiscuous enzymes in bacteria and Archaea metabolism? *Sustain Chem Process* 4:3

- Matthews BW (1968) The solvent content of protein crystals. *J Mol Biol* 33:491-497.

- McClean KH, Winson KM, Fish L, Taylor A, Chhabra RS, Camara M, Daykin M, Lamb JH. Swift S, Bycroft WB, Stewart GSAB and Williams P (1997) Quorum sensing and *Chromobacterium violaceum*: Exploitation of violacein production and inhibition for the detection of *N*-acylhomoserine lactones. *Microbiol* 143:3703–3711.

- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.

- McNicholas S, Potterton E, Wilson KS, Noble MEM (2011) Presenting your structures: the CCP4MG molecular-graphics software. *Acta Crystallogr D* 67: 386–394.

- Mcvey CE, Walsh MA, Dodson GG, Wilson KS, Brannigan JA (2001) Crystal structures of

penicillin acylase enzyme-substrate complexes : Structural insights into the catalytic mechanism. *J. Mol. Biol.* 313:139-150.

- Meevootisom V, Somsuk P, Prachaktam R, Flegel TW (1983) Simple screening method for isolation of penicillin acylase-producing bacteria. *Appl Environ Microbiol* 46: 1227-1229

- Merino E, Balbas P, Recillas F, Becerril B, Valle F (1992) Carbon regulation and the role in nature of the *Escherichia coli* penicillin acylase (*pac*) gene. *Mol Microbiol* 6: 2175-2182

- Michalska K, Brzezinski K, Jaskolski M (2005) Crystal structure of isoaspartyl aminopeptidase in complex with L-aspartate. *J Biol Chem* 280:28484–28491.

- Miller MT, Gerratana B, Stapon A, Townsend CA, Rosenzweig AC (2003) Crystal structure of carbapenam synthetase (CarA). *J Biol Chem* 278:40996–41002.

- Monera OD, Kay CM, Hodges RS (1994) Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Sci* 3:1984–1991.

- Morohoshi T, Kato M, Fukamachi K, Kato N, Ikeda T (2008) N- Acylhomoserine lactone regulates violacein production in *Chromobacterium violaceum* type strain ATCC 12472. *FEMS Microbiol Lett* 279:124– 130.

- Navaza J (1994) AMoRe - An automated package for molecular replacement. *Acta Crystallogr A* 50:157-163.

- Ningthoujam DS,Sanasam S, Tamreihao K, Nimaichand S (2009) Antagonistic activities of local actinomycete isolates against rice fungal pathogens. *Afri J Microbiol Res* 3:737-742.

- Noren C, Wang J, Perler F (2000) Dissecting the chemistry of protein splicing and its applications. *Angew Chem Int Ed Engl* 39:450–466.

- O'Brien PJ, Herschlag D (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem Biol* 6:R91–R105.

- Ochiai S, Yasumoto S, Morohoshi T, Ikeda T (2014) AmiE, a novel N-Acylhomoserine Lactone acylase belonging to the Amidase family, from the activated-sludge isolate *Acinetobacter* sp. strain Ooi24. *Appl Environ Microbiol* 80:6919–6925.

- Oinonen C, Rouvinen J (2000) Structural comparison of Ntn-hydrolases. *Prot Sci* 9:2329–2337.

- Oinonen C, Tikkanen R, Rouvinen J, Peltonen L. (1995) Three-dimensional structure of human lysosomal aspartylglucosaminidase. *Nature Struct Biol* 2:1102-1108.

- Okada T, Suzuki H, Wada K, Kumagai H, Fukuyama K (2006) Crystal structures of γ -

glutamyltranspeptidase from *Escherichia coli*, a key enzyme in glutathione metabolism, and its reaction intermediate. *Proc Natl Acad Sci* 103: 6471-6476.

- Panigrahi P, Sule M, Sharma R, Ramasamy S, Suresh CG (2014) An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholylglycine hydrolase family. *Microbiol Sgm* 160:1162-1174.

- Park SY, Kang HO, Jang HS, Lee JK, Koo BT, Yum DY (2005) Identification of extracellular *N*-acylhomoserine lactone acylase from a *Streptomyces* sp. and its application to quorum quenching. *Appl. Environ. Microbiol.* 71:2632–2641.

- Parkin S and Hope H (1998) Macromolecular cryocrystallography: cooling, mounting, storage and transportation of crystals. *J Appl Crystallogr* 31: 945-953.

- Passador L, Cook JM, Gambello MJ, Rust L, Iglewski BH (1993) Expression of *Pseudomonas aeruginosa* virulence genes requires cell-to-cell communication. *Science* 260:1127 LP-1130.

- Pundle A, Sivaraman H (1997) *Bacillus sphaericus* penicillin V acylase: Purification, substrate specificity, and active-site characterization. *Current Microbiol* 34:144-148.

- Pushparani DP, Godbole T, Prabhune A (2013) A novel alkaline intracellular penicillin V acylase from Streptomyces sp. APT13: Identification and effect of physiochemical conditions on enzyme production. *World J Pharm Pharm Sci* 2:6062–6075.

- Ramachandran GN, and Sasisekharan V (1968) Conformation of polypeptides and proteins. *Adv Protein Chem* 23:283-438.

- Rathinaswamy P, Pundle AV, Prabhune AA, SivaRaman H, Brannigan JA, Dodson GG, Suresh CG (2005) Cloning, purification, crystallization and preliminary structural studies of penicillin V acylase from *Bacillus subtilis*. *Acta Crystallogr F* 61:680–683.

- Rhode G (2000) Crystallography Made Crystal Clear. *Academic Press*.

- Romero M, Diggle SP, Heeb S, Camara M, Otero A (2008) Quorum quenching activity in *Anabaena* sp. PCC 7120: identification of AiiC, a novel AHL-acylase. *FEMS Microbiol Lett* 280:73–80.

- Rossmann M (2008) PhD thesis. Structural analysis of proteins of sphingolipid metabolism, Freie Universitat Berlin, Berlin.

- Rossmann MG, Blow DM (1962) The detection of subunits within the crystallographic asymmetric unit. *Acta Cryst* A15:24-31.

- Rossocha M, Schultz-heienbrok R, Moeller H Von, Coleman JP, 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. *Biochem* 2:5739–5748.

- Wilmouth RC, Edman K, Neutze R, Wright PA, Clifton IJ, Schneider TR, Schofield CJ and Hajdu J (2001) X-ray snapshots of serine protease catalysis reveal a tetrahedral intermediate. *Nat Struct Biol* 8:689-694.

- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.

- Sali A (2016) MODELLER A Program for Protein Structure Modeling Release 9.17, r10881. Available at: https://salilab.org/modeller/.

- Sheng G, Ye Y (1989) Some enzymological and kinetic properties of immobilize *Fusarium oxysporum*. *Shengwa Gonhvhrmg Xurbao* 5:129–34.

- Shepherd, RW, Lindow SE (2009) Two dissimilar N-acyl-homoserine lactone acylases of Pseudomonas syringae influence colony and biofilm morphology. *Appl Environ Microbiol* 75:45–53.

- Shewale JG, Kumar KK, Ambekar GR (1987) Evaluation of determination of 6-amino penicillanic acid by p-dimethylaminobenzaldehyde. *Biotechnol Tech* 1:69–72

- Shewale JG, Sudhakaran VK (1997) Penicillin V acylase: Its potential in the production of 6-aminopenicillanic acid. *Enzyme Microb Technol* 20(96):402–410.

- Sillitoe I, Lewis ET, Cuff A, Das S, Ashford P, Dawson LN, Furnham N, Laskowski AR, Lee D, Lees GJ, Lehtinen S, Studer RA, Thornton J, Christine AO (2014) CATH : comprehensive structural and functional annotations for genome sequences. *Nucleic Acids Res* 43:D376-81.

- Singh HT, Shyamananda RK (1994) Ramsar Sites of India, Loktak Lake, Manipur. *World Wide Fund for Nature* 96.

- Singh LL, Baruah I, Bora TC (2006) Actinomycetes of Loktak habitat: Isolation and screening for antimicrobial activities. *Biotechnol* 5:217-221

- Singh LS, Mazumder S, Bora TC (2009) Optimisation of process parameters for growth and bioactive metabolite produced by a salt-tolerant and alkaliphilic actinomycete, *Streptomyces tanashiensis* strain A2D. *Journal De Mycologie Medicale* 19:225-233

- Singh NK, Dhar DW (2011) Phylogenetic relatedness among *Spirulina* and related cyanobacterial genera. *World J Microbiol Biotechnol* 27:941-951.

- Smith JL (1995) Structures of glutamine amidotransferases from the purine biosynthetic

pathway. *Biochem Soc Trans* 23:894–898.

- Steindler L, Venturi V (2007) Detection of quorum-sensing N-acyl homoserine lactone signal molecules by bacterial biosensors. *FEMS Microbiol Lett* 266:1–9.

- Stellwag EJ, Hylemon PB (1976) Purification and characterization of bile-salt hydrolase from *Bacteroides fragilis* subsp *fragilis*. *Biochim Biophys Acta* 452:165-176.

- Stura EA and Wilson IA (1990) Analytical and production seeding techniques. *Methods* 1:38-49.

- Sudhakaran VK, Borkar PS (1985) Phenoxymethyl penicillin acylases: Sources and study-a summary. *Hind Antibiot Bull* 27:44-62

- Sunder AV, Kumar A, Naik N, Pundle AV (2012) Characterization of a new *Bacillus cereus* ATUAVP1846 strain producing penicillin V acylase, and optimization of fermentation parameters. *Ann Microbiol* 62:1287-1293.

- Suresh CG, Pundle A.V, SivaRaman H, Rao KN, Brannigan JA, McVey CE, Verma CS, Dauter Z, Dodson EJ, Dodson GG (1999) Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. *Nat Struct Biol* 6:414–416.

- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* 30:2725-2729.

- Taranto MP, De Llano DG, Rodriguez A, De Ruiz Holgado AP, Font de Valdez G (1996) Bile tolerance and cholesterol reduction by *Enterococcus faecium*, a candidate microorganism for the use as a dietary adjunct in milk products. *Milchwissenschaft* 51:383–385.

- Taranto MP, Fernandez Murga ML, Lorca G, de Valdez GF (2003) Bile salts and cholesterol induce changes in the lipid cell membrane of *Lactobacillus reuteri*. *J Appl Microbiol* 95 86-91.

- Torres R, Ramon F, Mata ID, Acebal C, Castillon MP (1999) Enhanced production of penicillin V acylase from *Streptomyces lavendulae*. *Appl Microbiol Biotechnol* 53:81-84.

- Uroz S, Dessaux Y, Oger P (2009) Quorum sensing and quorum quenching: the yin and yang of bacterial communication. *Chembiochem* 10:205–16.

- Vagin A and Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Cryst D* 56:1622-1624.

- 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.

- Vallenet D, Nordmann P, Barbe V, Poirel L, Mangenot S, Bataille E, Dossat C, Gas S,

Kreimeyer A, Lenoble P, Oztas S, Poulain J, Segurens B, Robert C, Abergel C, Claverie JM, Raoult D, Medigue C, Weissenbach J, Cruveiller S (2008) Comparative analysis of Acinetobacters: Three genomes for three lifestyles. *Plos One* 3:e1805.

- Vandamme EJ,Voets JP (1975) Properties of purified penicillin V acylase of *Erwinia aroideae*. *Experientia* 31:140–3.

- Varshney NK, Kumar RS, Ignatova Z, Prabhune A, Pundle A, Dodson E and Suresh CG (2012) Crystallization and X-ray structure analysis of a thermostable penicillin G acylase from Alcaligenes faecalis. *Acta Crystallogr F* 68:273–277.

- Wahjudi M, Papaioannou E, Hendrawati O, van Assen AHG, van Merkerk R, Cool RH, Poelarends GJ, Quax WJ (2011) PA0305 of *Pseudomonas aeruginosa* is a quorum quenching acylhomoserine lactone acylase belonging to the Ntn hydrolase superfamily. *Microbiol* 157:2042–55.

- Wang JL, Li X, Xie HY, Liu BK, Lin XF (2010) Hydrolase-catalyzed fast Henry reaction of nitroalkanes and aldehydes in organic media. *J Biotechnol* 145:240–243

- Wang Y, Guo H-C (2007) Crystallographic snapshot of a productive glycosylasparaginase–substrate complex. *J Mol Biol* 366:82–92.

- Weber PC (1991). Physical principles of protein crystallization. *Adv Prot Chem* 41:1- 36.

- Wegman MA, Janssen MHA, Van Rantwijk F, Sheldon RA (2001) Towards biocatalytic synthesis of β-lactam antibiotics. *Adv Synth Catal* 343:559–576.

- Williams K, Cullati S, Sand A, Biterova EI, Barycki JJ (2009) Crystal structure of acivicin-inhibited γ-glutamyltranspeptidase reveals critical roles for its C-terminus in autoprocessing and catalysis. *Biochem* 48:2459–2467.

- Williams P (2007) Quorum sensing, communication and cross-kingdom signalling in the bacterial world. *Microbiol* 153(12):3923–3938.

- Wilson K (1987) Preparation of genomic DNA from bacteria. In: Ausubel FM, Brent R, Kingston Re, Moore DD, Seidman JG, Smith JA, Struhl K, (eds) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 2.4.1-2.4.5

- Winkler UK, Stuckmann M (1979) Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. *J Bacteriol* 138:663–670.

- Winson MK, Swift S. Fish L, Throup JP, Jørgensen F, Chhabra SR, Bycroft BW, Williams P, Stewart GS (1998) Construction and analysis of luxCDABE -based plasmid sensors for

investigating N-acyl homoserine lactone-mediated quorum sensing. *FEMS Microbiol Lett* 163:185–192.

- Wolfenden R, Snider MJ (2001) The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc Chem Res* 34:938–945.

- Wu S, Zhang Y (2007) LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Res* 35:3375–3382.

- Wu W-B, Wang N, Xu J-M, Wu Q, Lin X-F (2005) Penicillin G acylase catalyzed Markovnikov addition of allopurinol to vinyl ester. *Chem Commun* 18:2348–50.

- Xu F, Guo F, Hu X-J, Lin J (2016) Crystal structure of bile salt hydrolase from *Lactobacillus salivarius*. *Acta Crystallogr F* 72(5):376–381.

- Yang J, Yan R, Roy A, Xu Dong, Poisson J, Zhang Y (2014) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12:7–8.

- Yates AE, Bodo P, Buckley C, Atkinson S, Chhabra SR, Sockett RE, Morris G, Dessaux Y, Camara M, Smith H, Williams P (2002) *N*-acylhomoserine lactones undergo lactonolysis in a pH, temperature, and acyl chain length-dependent manner during growth of *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa*. *Infect Immun* 70:5635–5646.

- Zhang D, Koreishi M, Imanaka H, Imamura K, Nakanishi K (2007) Cloning and characterization of penicillin V acylase from *Streptomyces mobaraensis*. *J Biotech* 128:788-800.

- Zhang J (2003) Evolution by gene duplication: An update. *Trends Ecol Evol* 18:292–298.

Web links:

1. http://tree.bio.ed.ac.uk/software/figtree/.

2. http://www.helmholtzmuenchen.de/fileadmin/PEPF/pET_vectors/pET-22b_map.pdf

List of Publications and Patents

1. Philem PD, Sonalkar VV, Dharne MS, Prabhune AA (2016) Purification and partial characterization of novel penicillin V acylase from Acinetobacter sp. AP24 isolated from Loktak Lake, an Indo-Burma biodiversity hotspot. Prep Biochem Biotechnol 46(5):524–530.

2. Philem PD, Yadav Y, Prabhune A, Ramasamy S (2017) A novel sub-class of cholylglycine hydrolases: Insights from structural and enzymatic characterization of cholylglycine hydrolases from *Shewanella loihica* PV-4 (Manuscript under preparation).

3. Pushparani Devi P, Godbole T, Prabhune A (2013) A Novel Alkaline Intracellular Penicillin V Acylase From Streptomyces Sp. APT13: Identification and Effect of Physiochemical Conditions on Enzyme Production. World J Pharm Pharm Sci 2(6):6062–6075.

4. D'britto V, Philem PD et al. (2012) Medicinal plant extracts used for blood sugar and obesity therapy shows excellent inhibition of invertase activity: synthesis of nanoparticles using this extract and its cytotoxic and genotoxic effects. Int J Life Sci Pharma Res 2(3):61–74.

Patent

1. Dubey P, Philem PD, Nisal AA, Prabhune AA (2017) Sophorolipid mediated accelerated generation of silk fibroin. US Patent 20,170,056,551.