Structure-Function Study of Ntn Hydrolase Enzymes Penicillin Acylase and Bile Salt Hydrolase

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

By

Deepak Chand
10BB11J26123

Under the guidance of
Dr. C. G. Suresh (Research Supervisor)

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

MAY 2016

Dedicated to.....

Maa, Papa

&

Family

सीएसआईआर - राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा मार्ग, पुणे - 411 008. भारत
CSIR - NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411 008, India

CERTIFICATE

This is to certify that the work incorporated in this Ph.D thesis entitled **"Structure-Function Study of Ntn Hydrolase Enzymes Penicillin Acylase and Bile Salt Hydrolase"** submitted by **Mr. Deepak Chand** to Academy of Scientific and Innovative Research (AcSIR) in fulfillment of the requirement for the award of the Degree of **Doctor of Philosophy**, embodies original research work under my 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 of diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table etc., used in the thesis from other sources have been duly cited and acknowledged.

Date: 26th May 2016

Place: Pune

Dr. C. G. Suresh
(Research Supervisor)
Chief Scientist

26-05-16
Deepak Chand
(Research Student)

Communication Channels
NCL Level DID : 2590
NCL Board No. : +91-20-2590 2000
EPABX : +91-20-2589 3300
: +91-20-2589 3400
FAX
Director's Office : +91-20-2590 2601
COA's Office : +91-20-2590 2660
COS&P's Office : +91-20-2590 2664
WEBSITE
www.ncl-india.org

DECLARATION BY RESEARCH SCHOLAR

I hereby declare that the thesis entitled "**Structure-Function Study of Ntn Hydrolase Enzymes Penicillin Acylase and Bile Salt Hydrolase**" submitted by me for the Degree of Doctor of Philosophy to Academy of Scientific & Innovative Research (AcSIR) is the record of work carried out by me at Biochemical Sciences Division, CSIR- National Chemical Laboratory, Pune - 411008, India, under the supervision of Dr. C.G.Suresh (research 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 of higher learning. I further declare that the material obtained from other resources has been duly acknowledged in the thesis.

26-05-16

Deepak Chand

26th

Date: May 2016

Structural Biology Group
Biochemical Sciences Division
CSIR-National Chemical Laboratory
Dr. Homi Bhabha Road, Pashan
Pune – 411008
Maharashtra, India

Table of Contents

Acknowledgment

This Thesis is the outcome of several people's contribution, without their support this wonderful journey, which is a blend of multitude of things, learning, patience, help, encouragement, aspiration, ecstasy and melancholy, would have not been possible. Hence it becomes my responsibility to acknowledge the people who had extended their support to me scientifically or morally to make a smooth running of my way. I, now take this opportunity to thank all who have helped me during these unforgettable years which is a crucial episode of my life. First and foremost, I owe to my parents and my family members specially my "Papa" who has walked through myriads of rough patches to pave my way and was always there for everything in my ups and down.

I am extremely grateful to Dr. C.G. Suresh, my research guide, for providing a great support and having faith in me for everything. This work would have not been possible without his guidance, critical inputs and freedom for executing my ideas as well. He has always supported me in every sense throughout these years. He has followed my progress from the first day and helped in a great deal in improving my speaking skill. I also appreciate his meticulous attention to the minute details. I appreciate all his effort, ideas and time to make this happen.

I am also grateful to Dr. Sureshkumar Ramasamy for his inevitable support and encouragement and due to whom my work took momentum moved to next step. I also acknowledge him for kindly replying my queries and giving suggestions out of his busy schedule for my work before he landed in India. He has been associated with my research work from beginning and has provided critical inputs and suggestions. I also appreciate his friendly nature in the lab.

I am thankful to my DAC chairman Dr. H.V. Thulasiram and DAC members Dr. Archana Pundle, Dr. Asmita Prabhune for their constant support and encouragement. I also acknowledge Director CSIR, NCL and head Biochemical Sciences Division, Dr Archana Pundle, and former head Dr. Vidya Gupta, for providing the lab facilities and infrastructure. I am thankful to Dr. Sushama Gaikwad and Dr Monisha for allowing me to use instruments. I am thankful to Dr. J.K. Pal from, Biotechnology, Pune University allowing me to work in his lab for six months. I am thankful to Dr.Kailash Datkhile under whom I have carried out my initial cloning work. I am thankful to Dr. Sunil, Dr. Sangeeta, Dr Radha, Dr. Shilpa, Ms. Varsha, Jemmy for help and friendly environment in the lab. I thank Dr. B.M Khan and Sakeel, Parth, Prasant, Somesh, Rishi for helping and using their instruments.

I am thankful to Dr. Saikrishna and Dr.Gayatri from IISER, Pune for allowing us to use the X-ray facility. I am also thankful to Dr. Radha Chauhan and Swasthik, Praveen and Ashiwini from NCCS for allowing us to use Mosquito and their help. I am grateful to Dr. Ravindra Makde, Mr. Ashwini and Mr. Biplab Ghosh from PXBL21 beamline, Indus-IIRRCAT, DAE, Indore , for their kind help & support in the crystal data collection. I am thankful to Dr. B. Gopal from MBU, IISc Bangalore, for allowing us to screen and collect the crystal data in MBU X-ray facility. I thank Mr. Babu for his kind help during the crystal screening and data collection in IISc. I am also thankful to Dr. Babu Manjashetty and Dr. Hassan from BM14 ESRF, France for their kind help during data collection at BM14 beam. I thank Abhishek (ICGEB), Ruby (JNU) and Dr. Shivendra from RCB, Delhi for their wonderful company in the ESRF, Grenoble.

I thank Dr. Vaibhav A. Mantri from Marine Algal Research Station (MARS) unit, Mandapam, Tamilnadu, for accepting our request to carry out my CSIR-800 project under his guidance. I also acknowledge Dr. Eshwaran and Dr. Ganesh for their valuable time and discussion on seaweed cultivation. I am thankful to Dr. Durga Prasad, Ashok, Barbara, Manikandan, Saminathan, Vinod and MARS staff people for their support.

I am grateful to my lab members Dr. Urvashi, Dr. Nishant, Manas, Dr. Tulika, Dr. Priyabrata, Dr. Ranu, Ruby, Prachi, Manu, Ameya, Shreedhar, Yashpal, Deepanjan, Vijay, Shiva, Debjyoti, Tejashri, Aditi, Parag for wonderful time and experiences in the lab. Apart from this I will also cherish the enjoyable moments spent with the "Dudes lab" members. A special thanks to Dr. Priyabrata for his help in the docking and simulation studies and Dr. Nishant for crystal data collection.

I will also cherish the countless moments spent with Manu, Rupa and Prasanjeet in all these years. I should also mention the GJ & HR-IV friends and my roommate for their help and quality time in NCL. I also enjoyed several games in the hostel, specially Table Tennis and thank to all TT players, had wonderful time with you guys. I thank Manoj and Anil for their help, support and healthy discussion. I also thank Ravi, Manoj, Anil (Night coffee group) and Hemendra who have voluntarily supported me during financial crunch. I also thank Pushpa, Avinash, Hrishikesh, Parul, Preeti and Amrita for their help and enjoyable time. I also thank Rakesh, Ameya, Priyanka and Saleem for their help.

A special thanks to my batchmates in NCL, Ashish, Ruchira, Sayali, and Sheon for great time spent with all of you.

I also thank, Sanjay, Nandu, Tushar, Shailesh and Santanu (Biochemistry group) who has always been there for me in every situation. I have always enjoyed their company and spent quality time with them since my M.Sc. to PhD. days. Also, thank Ruchi and Namita for wonderful time in Delhi.

I also thank my friends Anand Kumar and Devendra Negi who are my class 11th standard friends, have always been there for me in all my ups and downs. We all have a great bond of friendship and had spent countless memorable moments. Thanks to Varsha Tiwari and Naresh for their support and wonderful time. I also thank all whom I may have forgotten to mention here in this note.

26-11-16

Deepak Chand

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Full form
6-APA	6-amino penicillinic acid
aa	amino acid
*Af*PGA	*Alcaligenes faecalis* penicillin G acylase
AHL	Acyl Homoserine Lactone
AS	Asparagine synthetase
AU	Absorbance unit
BLAST	Basic Local Alignment Search Tool
*Bl*BSH	*Bifidobacterium longhum* Bile Salt Hydrolase
BME	beta-mercaptoethanol
bp	base pair
BSH	Bile Salt Hydrolase
*Bsp*PVA	*Bacillus Sphaericus* penicillin V acylase
*Bsu*PVA	*Bacillus subtilis* penicillin V acylase
*Bt*BSH	*Bacteroides thetaiotaomicron VPI* bile salt hydrolase
CA	Cephalosporin acylase
CCP4	Collaborative Computational Project No. 4
CD	Circular Dichroism
CGH	Cholylglycine hydrolase
CPB	Citrate phosphate buffer
*Cp*BSH	*Clostridium perfringens* bile salt hydrolase
CPC	Cephalosporin C
DTT	Dithiothreitol
*Ec*PGA	*E .coli* penicillin G acylase
EDTA	Ethylenediaminetetraacetic acid
*Ef*BSH	*Enterococcus faecalis* bile salt hydrolase
EG	Ethylene glycol
GAT	Glutamine amidotransferase
GCA	Glychocholic acid
GCDCA	Glychochenodeoxycholic acid
GFAT	Glucosamine-fructose-6-phosphate aminotransferase

GGT	Gamma-glutamyltransferase
GPATase	Glutamine phosphoribosylpyrophosphate amidotransferase
h	Hill coefficient
IPTG	Isopropyl β-D-1-thiogalactopyranoside
iRDP	*in silico* Rational Design of Proteins web server
$K_{0.5}$	Substrate concentration at half-maximal velocity
Kcat	Enzyme turnover rate
*Kc*PGA	*Kluyvera citrophila* penicillin G acylase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani media
LLG	Log Likelihood Gain
*Ls*BSH	*Lactobacillu salivarius* bile salt hydrolase
MALDI	Matrix-associated Laser Desorption and Ionization
MPD	2-Methyl-2, 4-pentanediol
MRE	Mean Residue Ellipticity
NAAA	N-acylethanolamine hydrolyzing acid amidase
Ni-NTA	Ni2+-nickel-nitrilotriacetic acid
ns	nanoseconds
NtCn	N-terminal cysteine nucleophile
Ntn	N-terminal nucleophile
NtSn	N-terminal serine nucleophile
NtTn	N-terminal threonine nucleophile
OD	Optical density
*Pa*PVA	*Pectobacterium atrosepticum*
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
Pen G	Penicillin G
Pen V	Penicillin V
PGA	Penicillin G acylase
PVA	Penicillin V acylase
R(hkl)	Structure factor
RMSD	Root mean square deviation
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate

TCA	Taurocholic acid
TCDCA	Taurochenodeoxycholic acid
TDCA	Taurodeoxycholic acid
V_{max}	Maximum rate
XP	Extra precision
TCA	Trichloro acetic acid
GdnHCl	Guanidine Hydrocholride

Abstract

We report here the extensive work carried out on two important enzymes bile salt hydrolase (BSH, E.C.3.5.1.24) from a Gram positive bacteria *Enterococcus faecalis* (*Ef*BSH) and penicillin G acylase (PGA, E.C. 3.5.1.11) from *Kluyvera citrophila* (*Kc*PGA). Both enzymes belong to the N-terminal nucleophile hydrolase superfamily or Ntn-hydrolases which are functionally identified as amidases. These enzymes are widely distributed in all three spheres of life; prokaryote, eukaryotes and archea. Based on the presence of N-terminal residue (either Cys, Ser or Thr), which acts as nucleophile and base, the superfamily can be classified into N-terminal cysteine nucleophile (NtCn) hydrolases, N-terminal serine nucleophile (NtSn) hydrolases and N-terminal threonine nucleophile (NtTn) hydrolases, respectively. Being members of Ntn-hydrolase superfamily BSHs and PGAs share a common (αββα) Ntn structural fold and similar catalytic mechanism. PGAs are heterodimers, whereas, BSHs form various oligomer assemblies, the most common form being tetramer.

BSH enzyme belongs to the NtCn-hydrolase family and has been distributed among bacteria (Gram-positive, Gram-negative) and archea. It catalyzes the de-conjugation of bile salt by hydrolyzing the amide bond in conjugated bile acids. This de-conjugation of bile salt by BSH has great contribution in lowering of cholesterol as well as in maintenance of bile acid pool homeostasis in mammalian gastrointestinal (GI) tract, which consequentially reduces further complication related to the hyper-cholesterol. In addition, it is also involved in the glucose homeostasis. Therefore, BSH activity has been a great consideration for selecting probiotic strains. Recently, a probable relation between BSH activity and quorum quenching has been reported.

Excessive and uncontrolled use of primary antibiotics resulted in the generation of antibiotic-resistant pathogenic bacteria. The existing antibiotics can be modified using semi-synthetic routes. Penicillin acylases have been long studied as enzymes important in the pharmaceutical industry. These enzymes have been characterized from a diverse range of bacteria and fungi. They catalyze the deacylation of natural penicillins to active pharmaceutical intermediate 6-aminopenicillanic acid (6-APA). Penicillin acylases are now widely used in the production of semi-synthetic antibiotics. They belong to NtSn-hydrolase family. PGA from *Kluyvera citrophila* (*Kc*PGA) undergoes post-translational processing to reach active form. *Kc*PGA is known for its resilience towards harsh conditions of pH, temperature and organic solvents, hence its importance in applications.

The detailed work presented in the thesis underpins the biochemical and structural characterization along with computational analysis of BSH from *Enterococcus faecalis* and and PGA from *Kluyvera citrophila*

The thesis is organized into six chapters:

Chapter 1: Introduction

Chapter 1 presents the general introduction and provides the review of the literature on BSH and PGA. This chapter also describes the applications of BSH and PGAs in industry and probable physiological role. It covers the biochemically and structurally characterized BSH and PGAs from various sources. Their probable role in the microbial and human physiology has also been emphasized.

Chapter 2: Cloning, expression, biochemical and biophysical characterization of BSH from *Enterococcus faecalis* (*Ef*BSH)

Chapter 2 presents the cloning, expression, purification, biochemical and biophysical characterization of *Ef*BSH. The stability of *Ef*BSH enzyme under various conditions (pH, temperature, GdnHCl) was assessed. *Ef*BSH was found to be highly active towards the bile acids compared to other BSH. Another unique feature of enzyme allosteric behaviour was also observed in the *Ef*BSH. Interestingly, BSH activity was not inhibited upon Pen V binding instead it had enhanced the BSH activity.

Chapter 3: Crystal structure determination and structural analysis of *Ef*BSH

Chapter 3 deals with the crystallization, data collection, 3-dimentional structure determination and detailed structural analysis. The *Ef*BSH structure was solved at 2Å resolution. The structure was compared with other BSHs and subtle changes in the active site and tetramerization loop were observed. These variations in the structure were correlated with the unique biochemical characteristics of *Ef*BSH.

Chapter 4: Substrate binding analysis, site-directed mutagenesis and autocatalytic processing of *Ef*BSH

Chapter 4 describes the substrate binding, site-directed mutagenesis to explore the important residues involved in substrate binding and activity and autocatalytic processing of *Ef*BSH. Docking and MD simulation studies was performed to understand the binding

of the substrate GCA and Pen V. GCA showed very high polar complementarity with *Ef*BSH binding pocket and aromatic residue interactions (F18, Y20, Y65) with hydrophobic surface of cholate ring which could probably the basis for its high activity and specificity. Autocatalysis processing in *Ef*BSH was also demonstrated where importance of Ser3-Ser4 residues after nucleophile residue in cholylglycine hydrolases was emphasized.

Chapter 5: Structure mediation in substrate binding and post-translational processing of penicillin acylases: Information from mutant structures of *Kluyvera citrophila* penicillin G acylase

Chapter 5 describes the comparison of the three-dimentioanl structure of unprocessed Serβ1Gly and Serβ1Cys mutant. The comparison of both the mutants demonstrated insights into the autocatalytic processing of spacer peptide in *Kc*PGA. The role of residue R145 and F146 of α-chain was highlighted in the autocatalytic processing and substrate discrimination. The docking of AHL molecule with *Kc*PGA and *Ec*PGA revealed that F146 of α-chain in the *Ec*PGA has moved toward (inward movement) the nucleophile residue compared to *Kc*PGA, which excludes the productive binding conformation of the AHL, hence *Ec*PGA has no AHL activity.

Chapter 6: Summary and Conclusion

This chapter highlights the important findings from this research and their future prospects.

Chapter-1

“Introduction”

1.1 Enzymes and their significance

Enzymes are the catalysts of life process. They accelerate biochemical reactions up to the rates at which biological processes take place in living organisms with extraordinary specificity and catalytic power. Enzymes are highly selective, accelerating rate of metabolic reactions, from the digestion to the synthesis of DNA, RNA and protein. Almost all metabolic processes in the cell require enzymes in all the metabolic pathways in order to occur at rates fast enough to sustain life. "Enzymology" documents research on enzymes and catalytic mechanisms. Most enzymes are proteins, although a few are catalytic RNA molecules. Like other biological macromolecules such as nucleic acids (DNA, RNA) and polysaccharides, proteins also form integral part of organisms. Proteins are large biopolymers consist of the amino acids residues which are linked by the peptide bond (CO-NH). All proteins consist of the standard natural 20 amino acid residues. Diversity of enzymes originates from different composition and arrangement of the 20 amino acids along the sequence and their specificity is due to unique three-dimensional structure and active site architecture.

The first enzyme to be discovered was amylase in 1833 (Payen & Persoz, 1833), which was isolated from barley and catalyses the conversion of starch into sugars. This discovery triggered the development of commercial bread making and brewing techniques as well as the development of fermentative enzymes at the beginning of the 20th century. Sumner isolated and crystallized urease in 1926, an enzyme that is involved in the hydrolysis of urea into ammonia and carbamate (Sumner, 1926). His original findings led to the demonstration that enzymes are actually proteins. Sumner was awarded Nobel Prize in Chemistry in 1946 for his work on enzymes. Almost forty years elapsed until the first three-dimensional structure of an enzyme was determined using X-ray diffraction techniques. In 1965, Phillips and coworkers obtained the structure of the hen egg-white lysozyme, which hydrolyses peptidoglycan in bacterial cell walls (Blake, et al., 1965). During the 1960s, studies by Anfinsen on ribonuclease shed light on the dynamic nature of protein structure and served as a model to explore protein folding.

In parallel, scientists started to capture enzymes performing function in the metabolic pathways. The idea of cataloguing and representing everything known about metabolism emerged from the original illustrations depicted by Krebs of the citric acid

cycle (Krebs, 1940), which later translated into the elaboration of comprehensive wall charts of metabolic pathways (Reitz et al., 2004). Current efforts to make metabolic data publicly available in online repositories such as KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/pathway.html) are prevalent (Kanehisa, et al., 2012). Apart from the molecular functioning of the enzymes in metabolic pathways of cells, they also share a pivotal role in multitude of industrial and commercial processes, from production of active pharmaceutical intermediates and medical treatment to bioremediation to clean up toxic environmental waste. In addition, structural understanding of the enzymes helps in elucidating the enzyme function and hence, could be specifically targeted in several biomedical applications.

In the last few decades recombinant DNA technology, where a particular gene for an enzyme can be cloned in the expression plasmid with the help of molecular scissors (Restriction enzymes), has revolutionized molecular biology by facilitating determination of three-dimensional structures of proteins of interest, which can be expressed in a heterologous host and the catalytic characteristics modified using targeted mutagenesis, facilitating improvement and production of commercially useful enzymes efficiently. Recently, structural genomics have been emerged which seeks to describe the three-dimensional structure of every protein encoded by a given genome. It is useful for the identification of novel protein folds and potential targets for drug discovery. There are several enzymes that show differential substrate activity and they can also hydrolyze other substrates apart from their original substrates. Such promiscuity of enzymes necessitates the study of the versatility of enzymes.

1.2 Evolution and promiscuity of enzymes

An enzyme's function is intrinsically linked to its three-dimensional structure. It follows folding routs from primary structure to secondary structure. Finally secondary structures are folded in the three-dimensional tertiary structure which is functionally active. Proteins also form quaternary structures with the association of combinations of subunits. Enzymes are very selective in their catalytic activity, which regulates all biochemical processes in the living cell. They are synthesized and transported to various destinations in the cell and each enzyme is involved in specific metabolic reaction. All enzymes are specific in their function and several million enzyme-mediated chemical reactions stir a living system. The number of unique genes in every living cell is far less

than the metabolic reactions occur in living cells. It is widely accepted that many of the enzymes have evolved from the existing parent enzymes through the process of gene duplication. In the process of genome evolution gene duplication is a very crucial event which turns out to be a major driving force in the gain of new biological functions. Natural evolution produces huge number of variants by mutation (point mutation, insertion or deletion) and subsequent selection of the suitable variant.

Sequence comparisons imply that these enzymes are derived from a common ancestral enzyme and have gathered mutations which allow them to adapt over millions of years. The natural evolution is complex: random noise and evolutionary selection pressures to evolve and sustain the life in various extreme conditions. Therefore functional novelties arise in the common scaffolds and eventually result in a new variant with new selectivity and properties such as stability (pH, temperature), solubility, catalytic efficiency and promiscuity. The concept of enzyme promiscuity, also known as the ability of enzymes to catalyze more than one biochemical reaction, has radically changed the way we understand enzymes and has implications across a broad range of scientific disciplines, from the evolution of enzyme function (Copley, 2003; Khersonsky & Tawfik, 2010; O'Brien & Herschlag, 1999) to the root of biotechnology and bio-catalysis (Hult & Berglund, 2007; Nobeli, et al., 2009). Enzyme activity is affected by temperature, chemical environment (pH), pressure, and substrate concentration. Activity can be affected by other molecules like, inhibitors which upon binding to protein may decrease or effect complete loss of enzyme activity, whereas activators enhance activity.

Microbial enzymes play a pivotal role in the biochemical investigation, diagnosis, and curing of many diseases. Microorganisms are an excellent source of therapeutic enzymes owing to their broad biochemical diversity and susceptibility to genetic manipulation. The results of many studies showed that nature has used the common binding site of the enzyme and mechanistic features to catalyse reactions involving similar substrates. Several enzyme superfamilies have been discovered using the sequences and structural homologies. The members of these superfamilies share similar catalytic residues and active site structural folds, which show the similar catalytic reaction mechanism despite sharing no sequence similarity. These superfamilies provide evidence for divergent evolution.

The Ntn-Hydrolase superfamily is one such superfamily, resulted from a divergent molecular evolution. This superfamily comprises of a vast range of enzymes and which displays the diversity in their substrate specificity and functioning. All Ntn-hydrolase superfamily members possess the conserved core structural Ntn-fold (Fig. 1.1) despite having low sequence similarity (~ <10% - 50%). All of them hydrolyze the amide bond of the substrate; hence they are also known as amidases. Amidases are widely distributed in all organisms including eukaryotes, prokaryotes and archea. Bacterial amidases are reported as enzymes of high therapeutic and industrial value.

1.3 Ntn (N-terminal nucleophile) hydrolase super-family

N-terminal nucleophile (Ntn) hydrolases form a superfamily of enzymes, which functionally belong to hydrolase class of enzymes but more specifically they are amidases (Artymiuk, 1995; Brannigan et al., 1995). Despite sharing low sequence similarity the members of Ntn-Hydrolase family possess a common Ntn fold (Fig.1.1), similar catalytic residues and hydrolytic mechanism is same (Fig. 1.2). Ntn stands for N-terminal nucleophile, where the N-terminal amino acid residue acts as nucleophile during catalysis. N-terminal residue can be either cysteine, serine or threonine. The side chain of Cys presents a sulfhydryl (-SH) group whereas those of Ser and Thr offer hydroxyl (-OH) group to act as nucleophile. Based on the N-terminal nucleophile residue Ntn-Hydrolases can be classified into three types: Nt**C**n-hydrolase and Nt**S**n-hydrolase and Nt**T**n-hydrolase.

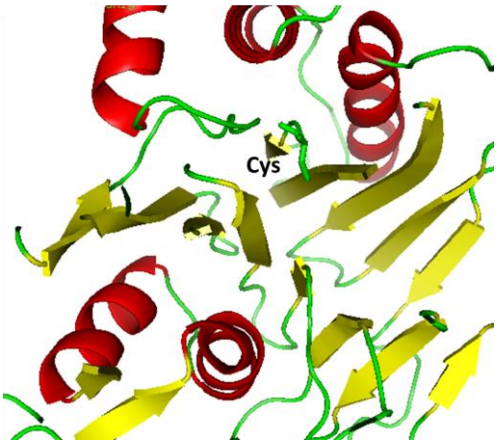


Figure1.1 Conserved structural core Ntn-fold comprised of αββα arrangement of the secondary structure in which β-sheets are sandwiched between α-helices. This fold is shared by all Ntn-hydrolase superfamily members.

One set of closely related Ntn hydrolases of interest here is the bile salt hydrolase (BSH) which belongs to the Cholylglycine hydrolase family or NtCn-hydrolases that includes penicillin V acylase (PVA) also (Kumar, et al., 2006). Another mechanistically related enzyme belonging to NtSn-hydrolases is Penicillin G acylase (PGA) which hydrolysis Pen G to produce 6-aminopenicillanic acid (6-APA) used in the synthesis of the semi-synthetic secondary antibiotics. Furthermore, PGA is also known to involve in quorum quenching by hydrolysing AHL (N-acyl Homoserine Lactone) a quorum sensing molecule in bacterium (Mukharji, et al., 2014). Interestingly, bile salt hydrolase (BSH), which catalyses the de-conjugation of six different bile acids in the GI- tract is also known to degrade AHL molecule (Mukharji et al., 2015). These enzymes show promiscuity towards various substrates.

1.4 Threonine-Ntn (NtTn) Hydrolase family

In the threonine-Ntn hydrolase, as name indicates the N-terminal threonine amino acid acts as nucleophile residue. These enzymes differ with respect to their subunit composition and substrate specificity. However, the catalytic site and overall structural Ntn-fold remain similar as in other Ntn-hydrolases. An example of NtTn-hydrolase is subunits of proteasome, the protein degradation machinery that degrades the ubiquitin-tagged proteins, plays a crucial role in non-lysosomal protein degradation and thus helps in protein turnover (Hershko & Ciechanover, 1992). The overall structures of the α- and β-subunits were observed to be quite similar (having αββα Ntn-hydrolase fold), they differ at their N-terminus (Groll, et al., 2003; Lowe et al., 1995; Huber, et al., 2012) (Fig. 1.0). Enzymes such as Glycosylasparaginase (GA) from *Flavobacterium meningosepticum* (Guan et al., 1998; Xu, et al., 1999), Ornithine acetyltransferase (OA) (Abadjieva, et al., 2000; Elkins, et al., 2005; Sankaranarayanan et al., 2010) also belong to NtTn-hydrolase family.

1.5 Cysteine-Ntn (NtCn) Hydrolases and Cholylglycine Hydrolase family

This superfamily includes Ntn-hydrolase enzymes which possesses the N-terminal amino acid residue cysteine residue that acts as nucleophile and base during catalysis. This family also shares the core conserved αββα-Ntn fold (Fig. 1.1). In the MEROPS database the NtCn superfamily represented as clan PB(C) which includes six structurally

or functionally characterized families of enzymes, family C59, C44, C45, C69, C89 and C95 (Dijkstra, et al., 2013). There is no substantial sequence homology found among enzymes of the families. Family C59 comprises of the bile salt hydrolase (BSH), penicillin V acylase (PVA), which are categorized under cholylglycine hydrolase family (Panigrahi, et al., 2014). The enzymes from C59 family have vital applications in the hypercholesterolemia and antibiotics industry.

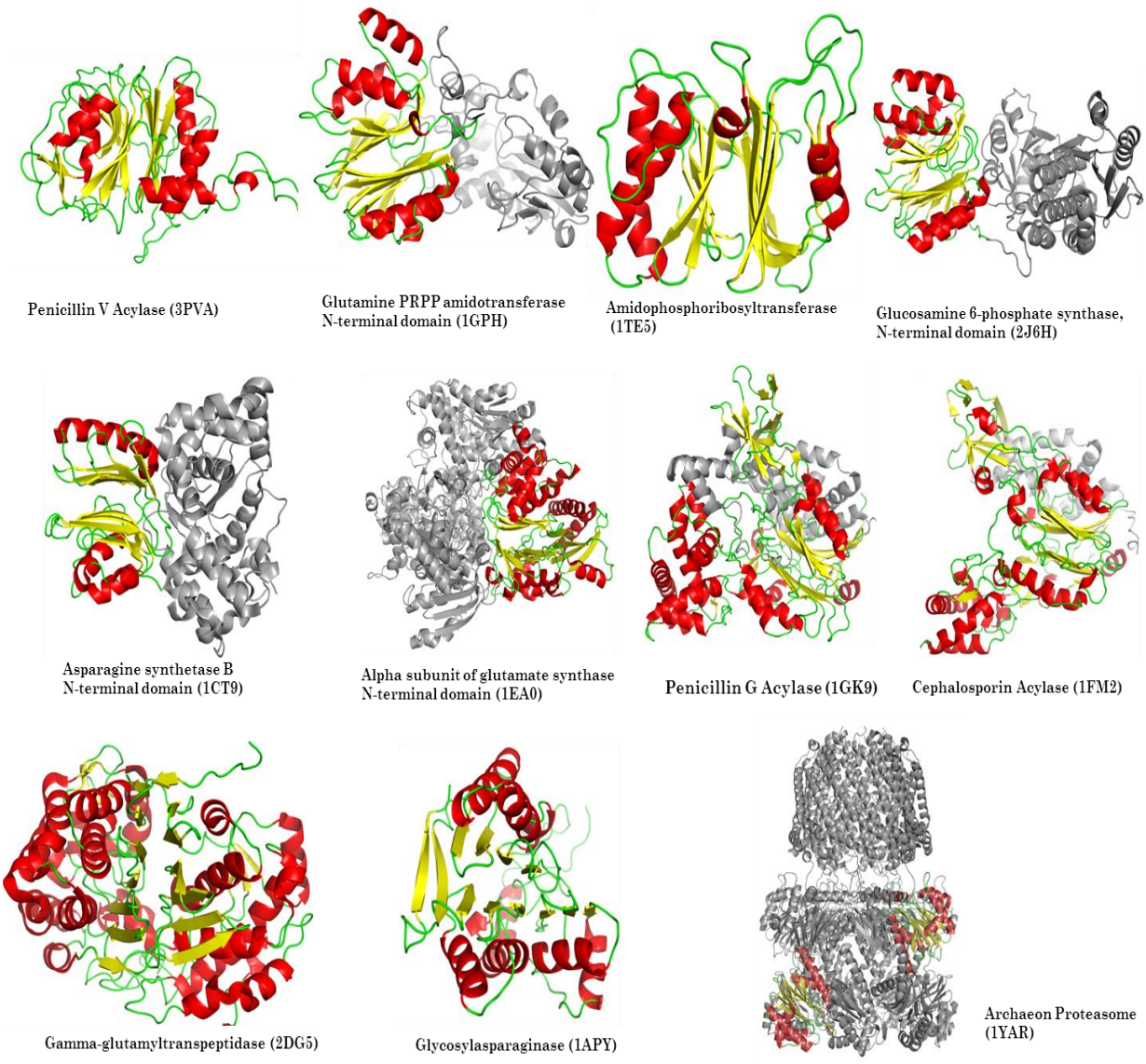


Figure 1.2 Members of the Ntn hydrolase super-family. The conserved core Ntn-fold is highlighted from various enzymes involved in the various pathways. They all share the similar catalytic mechanism even when sharing very less sequence similarity and show structural conservation.

1.5.1 Penicillin V acylases (PVAs)

Penicillin V acylases (PVA), belong to the cholylglycine hydrolase family of NtCn hydrolases. However, they catalyse substrates similar to that of penicillin G acylases (PGAs) belonging to NtSn-hydrolases. The PVAs have been characterized from Gram-positive bacteria, *Bacillus sphaericus* (Olsson, et al.1985; Suresh, et. al., 1999), *Bacillus subtilis* (Rathinaswamy, et al., 2009) and from a Gram-negative bacteria *Pectobacterium atrosepticum* (Avinash, et al., 2013). *Fusarium oxysporum a* fungus has also been found to produce PVA (Lowe, et al., 1986). An extracellular PVA from *Streptomyces lavendulae* has been biochemically characterized (Torres-Guzman, et al., 2002). A smallest monomeric PVA from *Rhodotorula aurantiaca* (NCIM 3425) has also been characterized (Atul, et al., 2007).

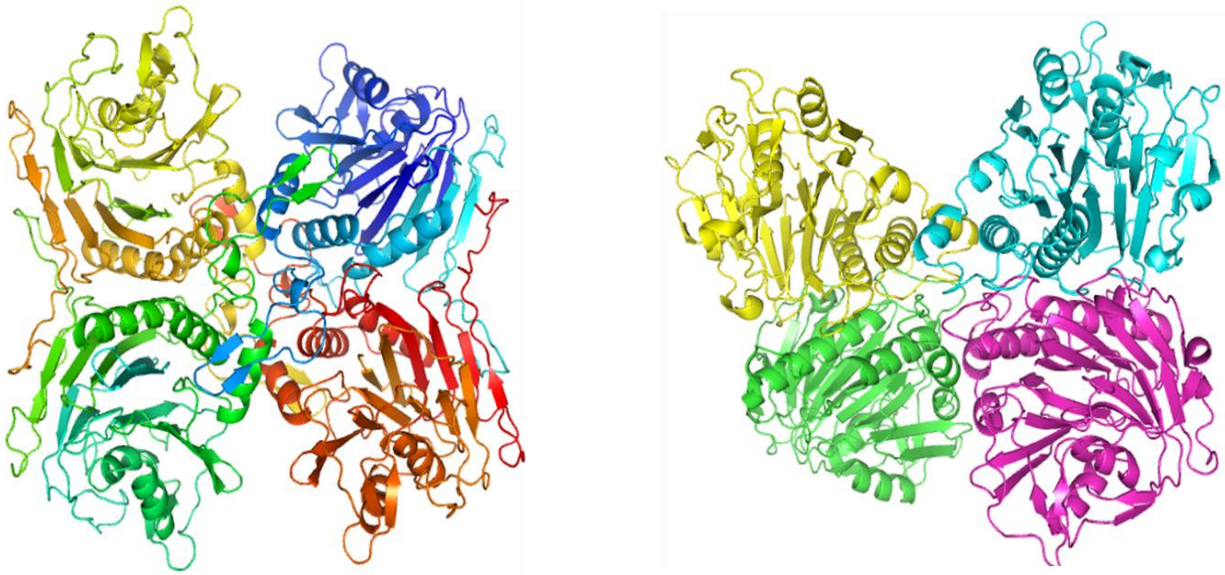


Figure1.3 [a] Crystal structure of the PVA from bacteria [a] A tetramer assembly of PVA from a Gram-positive bacteria *Bacillus sphaericus* (PDB ID: 3PVA, Suresh et al., 1999), [b] crystal structure of PVA from Gram-negative bacteria *Pectobacterium atrosepticum* (PDB ID: 4WL2, Avinash et al., 2015). In the Gram-positive bacterial structure assembly loop is shortened by 20 amino acid and arrangement of the subunits in tetramer is different compared to 3PVA.

As already mentioned PVAs catalyse the hydrolysis of Pen V substrate to produce 6-APA as a by-product (Lodola, et al., 2012; Avinash, et al., 2015). Although PGAs are used more in the production of the 6-APA at industrial scale there are some advantages of

using PVA over PGA. The optimum activity of PVA is towards the acidic pH (5.0-6.0) and PGA works in the range of pH 7.0-9.0. For industrial extraction PVA is a more suitable candidate for extraction at lower pH (Shewale and Sudhakaran, 1997).

The crystal structures of the PVAs have been elucidated from both Gram-negative and Gram -negative bacteria. The 3-dimentional structures of PVAs from Gram-positive bacteria *Bacillus sphaericus* (*Bsp*PVA, PDB ID: 3PVA, Suresh, et al., 1999) and *Bacillus subtilus* (*Bsu*PVA, PDB ID: 2OQC, Rathinaswamy, et el., 2008) have been determined using X-ray crystallography (Fig. 1.3a & 1.3b). Avinash, et al., (2015) has determined structure of a PVA from Gram-negative *Pectobacterium atrosepticum.* This PVA showed high activity and allosteric behaviour. In addition, the tetramerization assembly loop was found to be shortened by 20 amino acids, hence, reducing the tetrameric interactions across subunits. Also, the tetramer assembly arrangement was observed to be different from the reported PVA from Gram-positive bacteria. PVAs are evolutionary related to the bile salt hydrolase enzyme (Kumar, et al., 2006). The physiological role of the PVAs has been assumed in the aromatic compound degradation (Avinash, et al., 2014).

1.5.2 Bile salt hydrolases (BSHs)

The enzyme Bile Salt Hydrolase (E.C.3.5.1.24) commands an important position in the area of probiotics. It has been characterized from many bacteria constituting the indigenous mammalian gut microbiota. The presence of active BSH enzyme has long been considered an important criterion for selection of a bacterium as potential probiotic (Begely, et al., 2006).

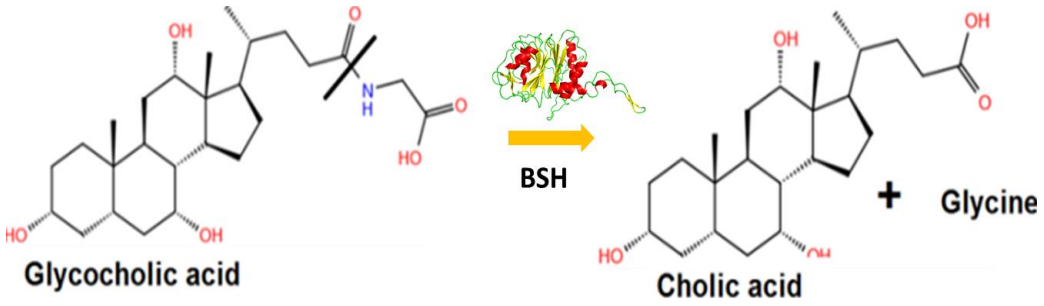


Figure1.4 Bile salt hydrolase catalysed reaction.

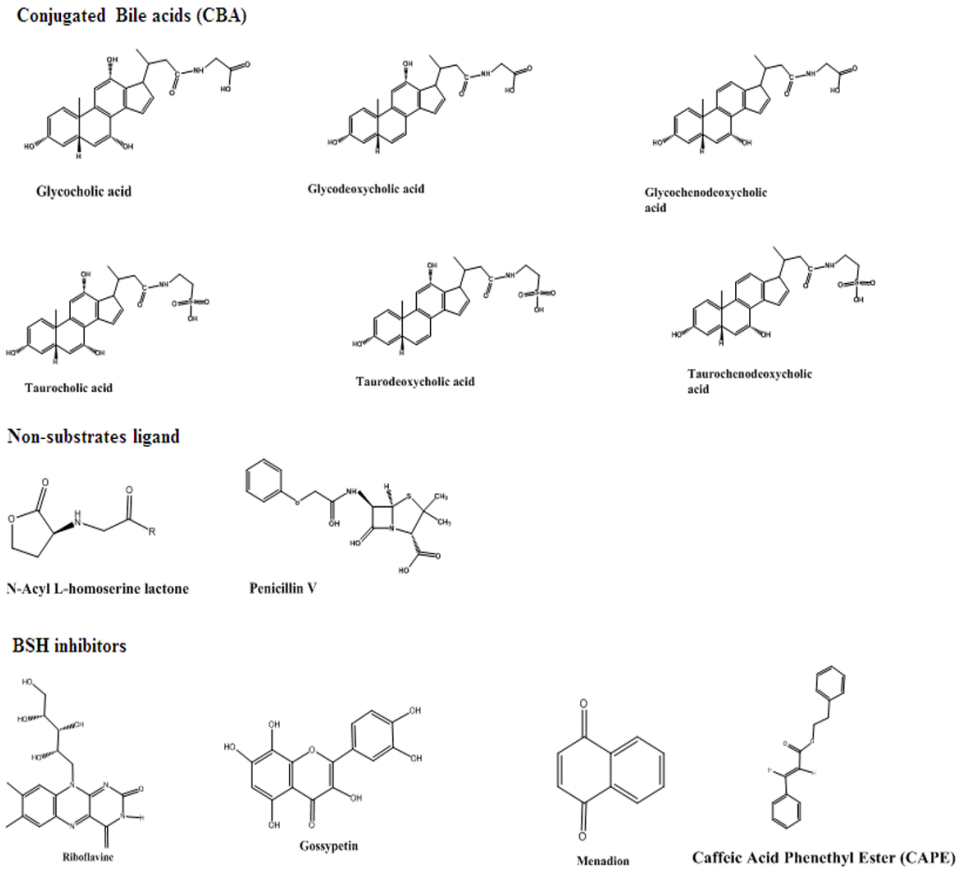


Figure1.5 The structures of substrates hydrolysed by bile salt hydrolase and some of the inhibitors of BSH.

This enzyme is responsible for the hydrolysis or metabolism of bile acids in healthy human GI tract (Fig. 1.4). Bile acids are synthesized in the liver from cholesterol as the precursor molecule, and are conjugated with either glycine or taurine (Russell, et al., 1992, Chiang, 1998, Begley, et al., 2005, Chiang, 2009) and secreted into the small intestine from the gall bladder. BSHs have been known to act on a wide range of bile acid conjugates, including all six major human conjugated bile acids (Fig. 1.5) - glycocholic acid (GCA), glycodeoxycholic acid (GDCA), and glycochenodeoxycholic acid (GCDCA), taurocholic acid (TCA), taurodeoxycholic acid (TDCA), taurochenodeoxycholic acid (TCDCA), releasing glycine- or taurine and free bile salts (Tanaka, et al., 2000, McAuliffe, et al., 2005; Begely, et al., 2006, Guo, et al., 2011).

There are several articles in literature highlighting the importance of BSH in probiotic selection and its importance in hydrolysis of bile salts leading to lowering of cholesterol levels. However, the BSH enzyme appears to be promiscuous and the full range of its functional spectrum which has not been studied in detail so far. BSH has now

been hypothesized to play multiple roles and has several applications, including the enhancement of bile resistance in intestinal microflora and the reduction of cholesterol levels (Tanaka et al., 2000, Moser, et al., 2000, Begely, et al., 2005, Begely, et al., 2006, Choi, et al., 2015). There is a fascinating interplay between the BSH of gut microbiota and the host metabolic processes including glucose homeostasis and cell signaling (Fig. 1.6). BSH has also been characterized from a few intestinal pathogenic bacteria, and is thought to be involved in the adhesion of organism to the intestinal membrane in both probiotic and pathogenic bacteria. Homologues of the *bsh* gene are present in various free-living soil and marine origin bacteria where the physiological role of BSH enzyme is not known. Recent reports have also found a probable link between BSH and related enzymes with the bacterial cell signaling system and quorum sensing. Such instances lend proof to the idea that BSHs might have evolved to play multiple roles in nature depending on the local environment of the bacteria that produce them (Begely, et al., 2006).

The biochemical and structural characteristics of BSH enzymes have been elucidated from different bacteria. BSHs from different classes of bacteria show remarkable variation in their sequence and substrate specificity. Certain bacteria also possess multiple BSH isoforms, although not all of them show bile de-conjugation activity (Jones, et al., 2008, Elkine, et al., 2001). The interaction of BSH with host physiology and the structural evolution of the enzyme in correlation with its substrate spectrum have been addressed here.

Bile salt hydrolase (BSH) enzyme is responsible for the de-conjugation of bile salts by commensal bacteria, thus playing a vital role in their colonization and survival in the mammalian intestine and decides their probiotic potential. Further, bile de-conjugation also leads to lowering of cholesterol, thus making BSH a clinically important enzyme. However, many recent observations seem to indicate that BSHs are involved in multifaceted roles, directly or indirectly in the host and microbial physiology. BSH activity is known to be linked to maintenance of glucose homeostasis in humans (Fig. 1.6).

Moreover, BSH paralogues have been found to exist in different microbes including free-living and pathogenic bacteria and Archaea. BSHs from various sources also show differential substrate activity and broad substrate spectrum. Certain bacteria are known to possess multiple genes for BSH enzymes, although only one usually de-

conjugates bile salts. These observations necessitate an intense study into the possible multi-faceted nature of BSH enzymes to better understand their role in regulating bacterial and host metabolism.

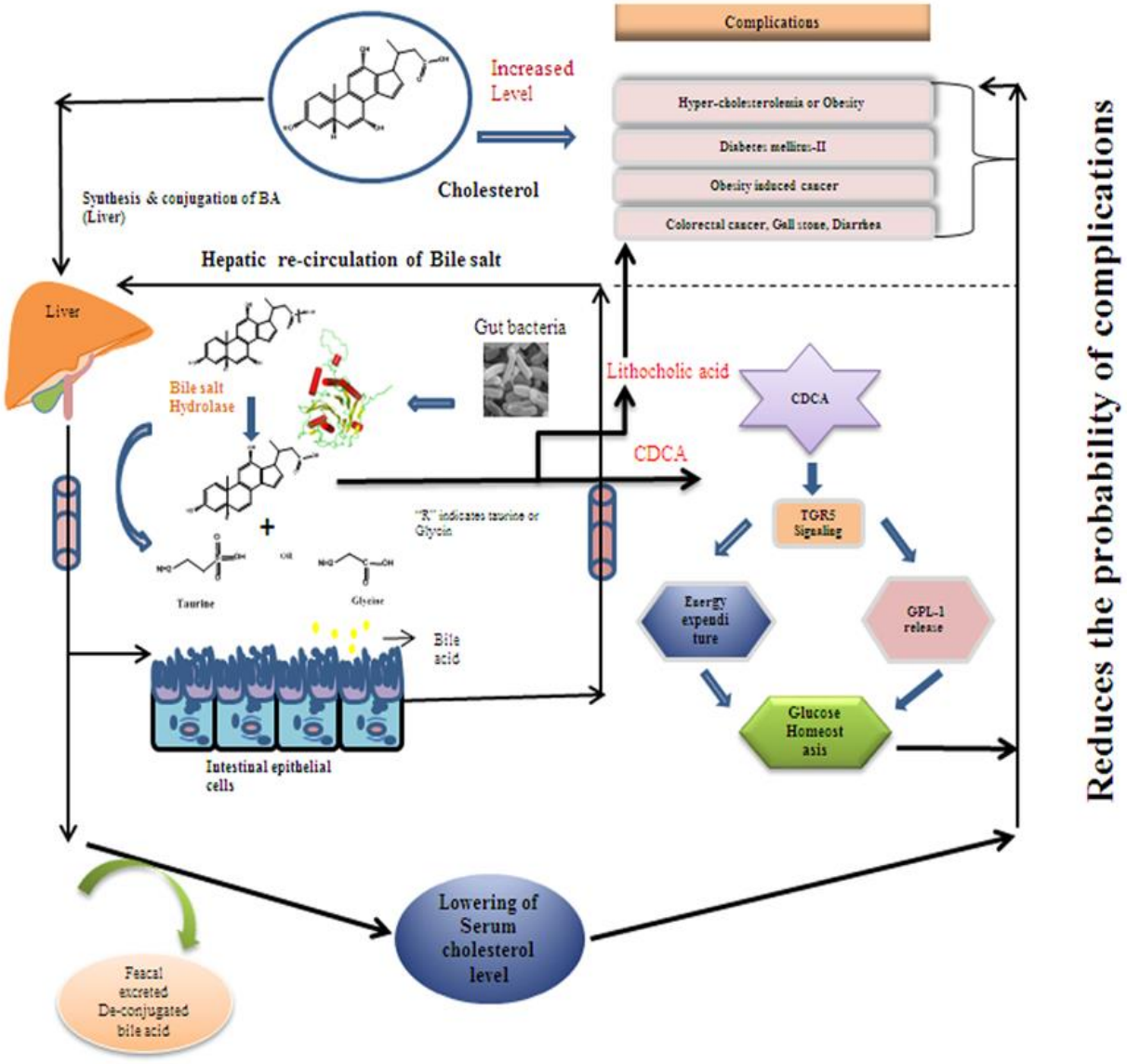


Figure 1.6 Multifaceted role of the BSH enzyme playing directly and or indirectly in the microorganism and host physiology.

A detailed exploration of BSH architecture and regulation could provide fascinating insights into its evolution and a better appreciation of the multiple functions of this enzyme in nature.

1.5.2.1Distribution of bile salt hydrolases in microorganisms

The ability to hydrolyze bile salts has been distributed among various Eubacteria, Archaea and fungi. BSHs have been isolated and characterized from several bacterial genera of the autochthonous gastrointestinal microbiota of animals including mice, rats, chickens, swine and human (Kim and Lee, 2005). BSH activity has been mainly studied in Lactobacillus, including *L. plantarum* (Christiaens, et al., 1992), *L. johnsonii* strain 100-100 (Lundeen and Savage, 1990), *L. acidophilus* (Corzo and Gilliland, 1999a), *L. casei* (Zhang, et al., 2009), *L. acidophilus* (Oh, et al., 2008), Other commensal probiotics such as *Bifidobacterium longum* (Grill, et al., 1995; Tanaka, et al., 2000) *and Pediococcus* sp. have also been reported to produce BSHs. Certain pathogens like *Clostridium perfringens* (Coleman and Hudson, 1995) , *Listeria monocytogenes* (Dussurget, et al., 2002) and *Brucella abortus* (Delpino, et al., 2007) are also BSH-active (Table 1). *Bacteriodes* sp. represents the only Gram-negative genera known to exhibit BSH activity so far (Kawamoto, et al., 1989; Stellwag and Hylemon 1976). Although *Bacteroides* and *Listeria* sp. are not commonly considered as members of the normal intestinal microbiota, they possess gene for BSH enzyme. Therefore, it has been suggested that these bacteria might be at the border between pathogenic and commensal microorganisms (Dussurget et al. 2002). *Enterococcus faecalis*, a Gram-positive bacteria, a normal inhabitant of the human gastrointestinal (GI) tract which can be an opportunistic pathogen, possesses a *bsh* homolog located within a pathogenicity island (Shankar, 2002. BSH from *Xanthomonas maltophilia,* a Gram-negative opportunistic nosocomial pathogen that causes a wide variety of diseases, is considerably different from other characterized BSH enzymes indicating a divergence of the BSH family. Even free-living bacteria isolated from marine sediments (*Methanosarcina acetovorans*) and Antartican lakes (*Planarus antarticus*) have been reported to possess BSH homologues (Reddy, et al., 2002; Jones, et al., 2008; Panigrahi, et al., 2014).

1.5.2.2 *Enterococcus faecalis*

Enterococcus faecalis is a Gram-positive commensal bacterium, which is classified as part of the group D *Streptococcus* system and inhabiting the gastrointestinal

tracts of humans and other mammal. *Enterococci* have gained notoriety over the past few decades as cause of multiple antibiotic resistance. *Enterococcus faecalis* is considered to prevent the Vancomycin resistance *Enterococci* and other pathogens colonization in the gut (Sakai, et al., 2006). Although it is an opportunistic pathogen but also play important role in human health by inhibiting the colonization of the pathogenic bacteria.

1.5.2.3 Biochemical characteristics and substrate specificity of BSH

Bile acids are conjugated in the liver with the amino acids taurine or glycine and secreted into small intestine. The cholate moiety is derived from cholesterol and serves as precursor molecule for the bile acid synthesis in the hepatocytes (Russell, et al., 1992, Chiang, et al., 2001). Deoxycholate and cholate are always hydroxylated on C_3 at α-position. Both glyco- and tauro- conjugated bile acids are anti-microbial in nature; the presence of BSH activity thus enables the persistence of bacteria in the intestinal tract through de-conjugation (Bateup, et al., 1995; Begley, et al., 2005). When the combinations between differently hydroxylated cholates and the two conjugated amino acids are taken into account, it is conceivable that BSHs have evolved to recognize bile acids on both the amino acid part and cholate groups. Moreover, the polar complementarity of hydroxyl group (-3α, -7α and -12α-OH) of bile acids are also estimated, which provides the probable basis for the substrate specificity in BSHs (Panigrahi, et al., 2015). BSH enzymes have been biochemically characterized from various bacteria and archea (Table 1). BSHs have been reported to mostly possess a homotetrameric composition with identical monomers of ~316-328 amino acids with an average molecular weight of ~37 kDa forming tetramers. The tetramer assembly is required for the biological activity of the enzyme (Suresh, et al., 2006). Gel filtration and X-ray crystallographic structures have shown that BSHs exist as homotetramers under physiological conditions (Gopal-Srivastava and Hylemon, 1988; Rossocha, et al., 2005; Suresh et al., 2006). Certain BSH homologues from *Lactobacillus* sp., however, are exceptions and exist in varied oligomeric states including dimeric or trimeric forms (Lundeen , et al., 1992) (Table 1.0). Sridevi, et al., (2009) have characterized a dimeric BSH from a thermophile *Brevibacillus species.*

Optimum bile salt hydrolysis by BSHs usually occurs in the acidic pH (~4.5-6.5) and temperatures ~40-50°C (Stellwag and Hylemon 1976; Gopal-srivastava and Hylemon 1988; Savage, et al., 1998; Corzo, et al., 1999 ; Grill, et al., 1995, Rossocha, et

al., 2005; Suresh, et al., 2006; Bi, et al., 2013). The *Xanthomonas maltophila* BSH showed better thermostability with complete recovery of activity after 75 min at 55°C, although it was deactivated rapidly at higher temperatures (Dean, et al., 2002). BSH enzymes vary widely in their catalytic efficiency and specificity. *Clostridum perfringens* BSH exhibits a low specific activity of 0.107 IU/mg, while BSHs from *L. acidophilus* and *B. longum* have higher activities of 3.4 and 11.4 IU/mg, respectively. BSH enzymes characterized from many bacteria including *C. perfringens* (Rossocha, et al., 2005) , *B. longum* (Tanaka, et al., 2000), *L.* and *acidophilus* (Corzo and Gilliland, 1999) are more specific for glyco- over tauro-conjugated bile salts, while *L. salivarius* (Fang, et al., 2009) shows preference for tauro-conjugated bile salts. Such differences probably hint at the scope and ability of the bacteria to colonize the gastro-intestinal environment. The intriguing spectrum of BSH activity necessitates the investigation of interactions between bile salt and binding enzyme, which would help better understand the significance of BSH in the gut microenvironment.

As already mentioned, the occurrence of multiple BSH homologues have been observed in certain cases, for instance four in *L. plantarum* and *L.sp. strain 100-100* (Lundeen, et al., 1992, Lambert, et al., 2008; Ren, et al., 2011). However, only one principal BSH shows good hydrolase activity, and the role of other BSH homologues present is not yet fully clear. A recent report has shown the role of BSH in the AHL molecule degradation that is involved in the quorum sensing and establishment of pathogenesis (Mukharjee, et al., 2015).

Table 1. Biochemically characterized bile salt hydrolase from various sources

S. No.	Domain	Source microorganism	Molecular weight (SDS-PAGE/Native)	Optimum pH	Km (mM)	Specific activity (U/mg)	Cellular location	Pathways involved	Oligomeric state	Reference
1	Gram-positive bacterial BSH	*Clostridium ATCC 19574*		5.6 ~5.8	-	-	Intracellular and extracellular	Bile acid Metabolism, bacterial defence system,		(Nair, et al. 1967: 7-11)
		C. perfringens PB 6 K		4.5 ~5.0	-	-	Intracellular	”		(Masuda 1981: 1-11)
		C. perfringens MCV 815	56.0 kDa ~250 kDa	5.8-6.4	-	GCA 0.107	Intracellular			(Gopal-Srivastava and Hylemon 1988: 1079-85)
		C. welchii		5-6	-	-	Intracellular			(Aries and Hill 1970: 526-34)
		Clostridium perfringens		5.5- 6.0	GCA 3.6 GDCA 1.2 GCDCA 14 TCA 37 TDCA 3.5 TCDCA 3	-	Intracellular	serum cholesterol lowering, Bile acid metabolism	Tetramer	(Rossocha, et al. 2005: 5739-48)
		B. longum BB536	~40.0 kDa-Monomer	5.5-6.5	-	-	Intracellular	-do-	Hexamer	(Grill, et al. 1995: 2577-82)

	B. longum SBT2928	37.3 kDa ~130 kDa	5.0-7.0	GDCA	GCA 11.84	intracellular	-do-		(Tanaka, et al. 2000: 2502-12)
	Bifidobacterium longum	~140 kDa, 37 kDa	5.0-6.5	Wild type GCA 0.22 TCA 0.32 Thr2Ala_mutant GCA 0.24 TCA 0.30	-	intracellular	-do-	Tetramer	(Kumar, et al. 2006: 32516-25)
	Bifidobacterium longum						Gel-formation, adhesion, cholesterol lowering	-	(Jarocki, et al. 2014: e114379)
	L. acidophilus O16	126 kDa	5.0-6.0	-	-	Intracellular	bacterial defence system, serum cholesterol lowering glucose homeostasis		(Corzo and Gilliland 1999: 466-71)
	L. acidophilus L1	126 kDa	3.5-4.5	-			-do-		(Corzo and Gilliland 1999: 466-71)
	L. acidophilus ATCC 43121	126 Kda	3.5-4.5	-	GCA 3.4	intracellular	-do-		(Corzo and Gilliland 1999: 466-71)
	L. brevis BCCM 18022						-do-		(Moser and Savage 2001: 3476-80)
	Lactobacillus salivarius BSH1			GCA 15.02 GDCA 14.12 GCDCA 15.48 TCA 17.16 TCDA 15.77 TCDCA 13.57			-do-	tetramer	(Bi, et al. 2013: 46-51)
	Lactobacillus			GCA 15.04 GDCA 11.88 GCDCA 15.05			-do-		

		salivarius BSH2			TCA 14.83 TCDA 18.17 TCDCA 15.92					(Bi, et al. 2013: 46-51)
		Lactobacilllus casei J57				GCA 44.91 GDCA 45.27 TCA 671.7 TDCA 61.57		-do-		(González-Vázquez, et al. 2015: 242-8)
		Lactobacillus sp. strain 100-100	BSH-A - 115 kDa BSH-B – 105 kDa BSH-C - 95 kDa BSH-D - 80 kDa	4.2-4.5	BSH-A - 0.76 BSH-B - 0.95 BSH-C - 0.45 BSH-D -0.37			Bile acid metabolism, cholesterol lowering, glucose homeostasis Role of prologues are unclear	Trimer	**(Lundeen and Savage 1992: 7217-20)**
		Enterococcus faecalis	~148 kDa, ~37 kDa	5.0-6.5	GCA 11.06 GDCA 9.94 GCDCA 16.97 TCA 13.16 TCDA 9.28 TCDCA 4.49	GCA 1390 GDCA 1289 GCDCA 970 TCA 1072 TDCA 1115 TCDCA 265.8		Bile acid Metabolism, cholesterol lowering, glucose homeostasis	tetramer	(Chand, et al. 2015)
2	Gram-negative Bacterial BSH	*Xanthomonas maltophilia*	52.0 kD ~100.0 kD	7.9-8.5	GDCA 0.15		intracellular	cell adhesion, protection against Bile acid		(Dean, et al. 2002: 3126-8)
		Brucella abortus 2308					extracellular	cell adhesion and contribution in pathogenesis		(Delpino, et al. 2007: 299-305)

		Bacteroids:					
		B. fragilis ssp. Fragilis ATCC 25285	250 kDa- Native 32 KDa - Monomer			Hexamer	(Stellwag and Hylemon 1976: 165-76)
		B. fragilis 2536					(Masuda 1981: 1-11)
		B. vulgatus		5.6 ~ 6.4	Intracellular	tetramer	(Kawamoto, et al. 1989: 1049-53)
3	Archeal BSH						
		Methanobrevibacter smithii.					(Jones, et al. 2008: 13580-5)
		Methanosphera stadmanae					(Jones, et al. 2008: 13580-5)

1.5.2.4 Genetic characterization and gene mobilization

Several bacteria which are exposed to bile salt stress express the *bsh* gene, allowing them to combat the lethal antimicrobial environment containing bile acids in the host. The presence of BSH enzyme is perceived as a positive impact on host health. In addition, BSH production has also been observed in pathogenic bacteria; the *bsh* gene in pathogens was plausibly acquired by horizontal gene transfer (HGT). HGT mechanisms occur by way of transposons, bacteriophages and other mobile elements, and are thought to play an important role in the acquisition of genes by the bacterial genome for evolution and better survivability (Dobrindt, et al., 2004). The presence and genetic variation of multiple *bsh* genes in *L. plantarum* and other Lactobacillus has been attributed to HGT (Bringel, et al., 2001; Molenaar, et al., 2005). Kumar, et al., (2012) have explored a natural transposon mediated insertional inactivation of *bsh* genes, which could help further understand the evolution of BSH and elucidate its functional roles in different environments.

The loci of *bsh* genes show different architectures in various bacteria. The transcript of *Lactocabicllus plantarum 80* is monocistronic (Christiaens, et al., 1992), however, *L. johnsonii strain 100-100* constitute a BSH operon, possessing functionally related genes *cbsT1, cbsT2* and *cbsHβ* genes, organized in tandem fashion (Elkins & Savage, 1998). Two separate loci on the genome of *L. johnsonii strain 100-100* for BSH were identified and characterized. Two genes *cbsT1, cbsT2* in operon belongs to the conjugated bile salt transporter (BST) family (Saier, et al., 1999). Notably, downstream of *cbsHβ* gene in the *L. Johnsonii strain 100-100* encodes for the maturase, which has the endonuclease and reverse transcriptase activity. This enzyme is known to involve in the cDNA splicing (Edgell, et al., 2000) suggesting that this locus in the *L.johnsonii strain 100-100* forms a genomic mobile element. Also, the *B. longum SBT2928 BSH* gene is coordinately regulated with glutamine synthetase adenylyltransferase (Tanaka, et al., 2000). Therefore gene acquisition by mobilization of the genetic material suggests adaption ability of the organism in a particular environment. The pathogenic bacteria in the gut might have acquired *bsh* gene to combat the antimicrobial activity of bile acid in the GI tract.

1.5.2.5 Phylogenetic analysis of cholylglycine hydrolase (CGH) family

Due to the obvious structural similarity and evolutionary relationship between BSHs and PVAs, the sequence annotation and functional characterization of CGHs has been difficult. Previously, biochemically and structurally well characterized BSH from *Clostridium perfringen* (*Cp*BSH) and *Bidifobacterium longum* (*Bl*BSH) (Rossecha, et al., 2005, Suresh, et al., 2006) were annotated in the Ntn-PVA-like family. Lambert, et al., (2008) have provided the correct annotation for the Gram-positive BSHs/PVAs, however Gram-negative CGH homologues were not taken into consideration in their study. Panigrahi, et al., (2014) have extended the analysis to Gram-negative bacteria and archea, and developed a new improved method for the annotation on the basis of phylogenetic analysis along with Binding Site Similarity based classification. CGH family members were classified into two distinct clusters of Gram-positive BSHs/PVAs and Gram-negative BSHs/PVA homologues, showing divergent evolution. In cluster 1, most of the enzymes annotated as BSH were gut-inhabiting bacteria (*firmicutes* and *Acenobactetia*) and Archea (*Methanobacteria smithii* and *Methanosphaera stadtmanae*). The enzymes annotated in cluster2 were widely distributed among gut-inhabiting and free living environmental bacteria. This highlights the probable evolution of BSHs with regard to the environment and functional differences. BSHs from different bacteria thus appear to be involved in promiscuous physiological roles; they are also involved in complex interactions with signaling and glucose homeostasis in their hosts. Although possessing BSH activity has been considered an essential part of probiotic bacteria, the significance of individual bile acid modifications by the enzyme in different metabolic pathways is still not clear. BSHs also appear to perform discrete functions in bacteria that adopt pathogenic or free living lifestyles. The varied functions of BSH enzymes in bacterial and host physiology therefore need to be explored in more detail.

1.5.2.6 Overall structure of bile salt hssydrolase:

The three-dimensional structures of bile salt hydrolases have so far been characterized from two Gram-positive bacteria *C. perfringens* (Rossocha, et al., 2005) and *B. longum* (Kumar, et al., 2006) and they exhibit a tetrameric subunit composition (Fig. 1.7a & b).

Monomers of both BSHs are made up of single globular domains with approximate dimensions of 40 Å × 50 Å × 55 Å and 75 Å x 38 Å x 44 Å, respectively (PDB ID:2RLC, Rossocha, ,et al., 2005; Kumar, et al., 2006, PDB ID:2HEZ). The tetrameric assembly can be considered a dimer of dimers as has been confirmed (Kumar, et al., 2006) by dissociation of *Bl*BSH with denaturant guanidine hydrochloride (GdnHCl). Tetramerization loops from each monomer (residues 185-220 in *Bl*BSH) cross over the symmetry center into the opposite subunit (reminiscent of domain swapping), leading to the formation of an extensively interlocked oligomer and stabilization of the quaternary structure (Fig. 1.7b). This oligomeric assembly motif is also shared by PVAs from Gram-positive bacteria (Fig. 1.3a). However, a significant reduction in the length of the assembly motif (Fig. 1.3b) and consequently the tetrameric interface area has been observed in CGHs from Gram-negative bacteria; these enzymes also phylogenetically separate into a different cluster (Panigrahi, et al., 2014; Avinash, et al., 2016).

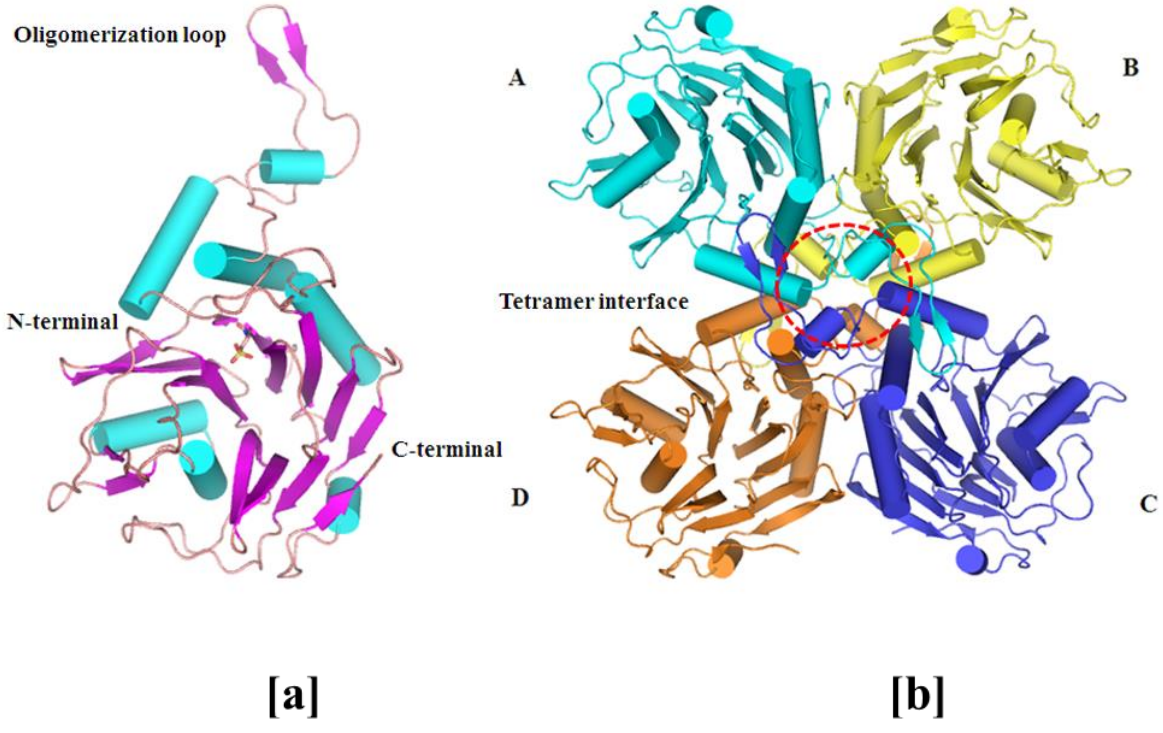


Figure 1.7 [a]. Structure of the monomer subunit of the *Bifidobacterium longum* BSH *(Bl*BSH*)* [3b]. Tetramer assembly of BSH is consisting of the dimer of two dimers. The assembly loop of the *Bl*BSH extends over 185-220 amino acid residues.

In the case of *Cp*BSH, the smallest interface contact between the opposite subunits A and B in the tetramer is formed between Trp181 and Asn185 (Rossmann , 2008). Trp181 also stacks hydrophobically with the Trp181 in the adjacent monomer. In other BSHs, Trp181 is sometimes replaced by phenylalanine or tyrosine which can form similar contacts. Many such reciprocal contacts also occur between residues from the tetramer interaction loops of opposite monomers. The subunits A and C that make up the dimers form a large interface comprising highly conserved residues (including Gly218-Asp222) making reciprocal contacts.

1.5.2.7 Catalytic mechanism, active site architecture and substrate binding specificity

The catalytic mechanism of Ntn hydrolases involves an initial nucleophilic attack on the amide bond by the N-terminal catalytic residue, followed by the formation of a tetrahedral intermediate which is stabilized by an oxyanion hole (Brannigan, et al., 1995; Oinonen and Rouvinen, 2000) (Fig. 1.8). Cysteine functions as the nucleophilic N-terminal residue in both BSH and PVAs (cholylglycine hydrolases). The importance of N-terminal cysteine in BSH activity has been proved through chemical modification and site-directed mutagenesis (Gopal-Srivastava and Hylemon, 1998; Kumar, et al., 2006). Many other catalytic residues in the active site of BSHs are conserved across Ntn-hydrolases (Duggleby, et al., 1995; Kim et al., 2000; Oinonen, et al., 1995; Tikkanen, et al., 1996).

The active site is located between the core β-strands (Rossocha, et al., 2005) in the αββα structural fold. Aside from the N-terminal cysteine, the active site residues such as Arg18, Asp21, Asn81, Asn175 and Arg228 (numbering according to *Cp*BSH sequence) are all strictly conserved in most BSHs.

The N-terminal cysteine Cys2 forms the catalytic Ntn-diad with Asp21 (Fig. 1.8). The α-amino group of Cys2 makes hydrogen bond to a water molecule in the vicinity that bridges to the nucleophilic cysteinyl sulphur. The sulfhydryl group of the N-terminal cysteine of BSH, stabilized by the conserved Arg18 residue, performs a nucleophilic attack on the amide bond of glycine or taurine conjugated bile salts and related substrates (Fig 1.8). The tetrahedral intermediate formed is stabilized by the oxyanion hole involving Asn175 and Asn81 in the loop (Rossmann, 2008).

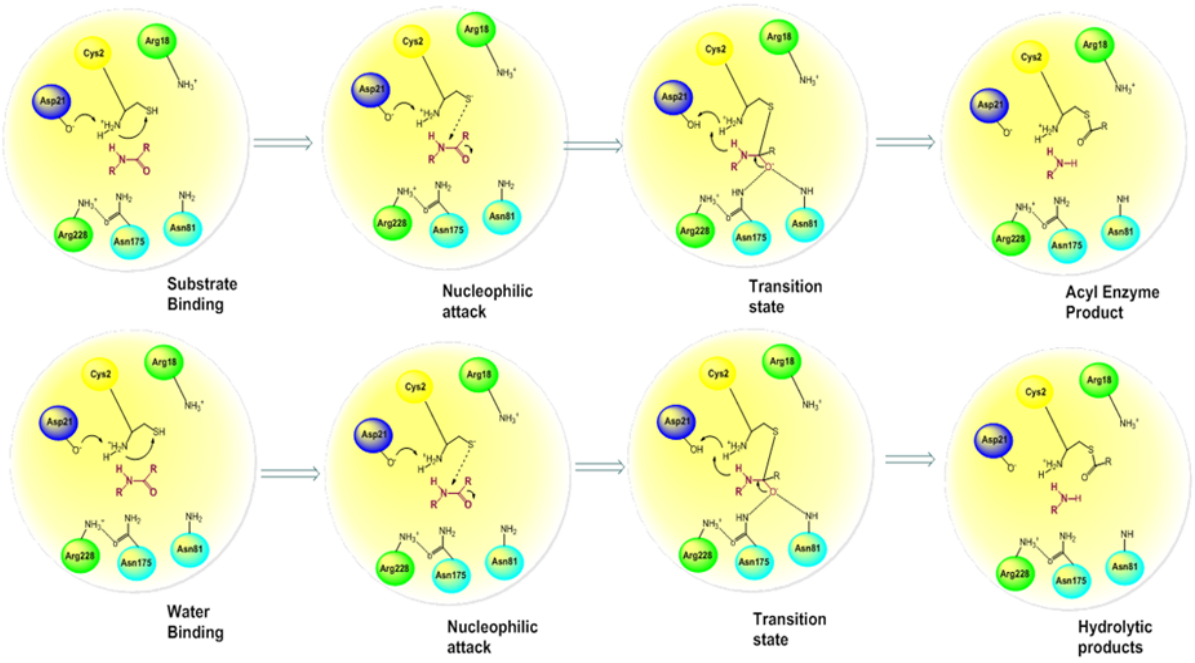


Figure 1.8 The steps in the substrate catalysis mechanism of bile salt hydrolase. Drawn in ChemBiodraw.

After the release of the leaving glycine or taurine, a second nucleophilic attack is carried out by a de-protonated water molecule on the thioester bond between Cys2 and the substrate. A similar tetrahedral intermediate is formed, eventually leading to the liberation of the cholate moiety and regeneration of free enzyme (Fig. 1.8). Recently, QM/MM calculations of the *Cp*BSH (Lodola, et al., 2012) have demonstrated that Cys2 participates in the catalysis in zwitterionic (Cys-S^-/ Cys-NH_3^+) form. This form being more stable (~15 kcal/mol) than the neutral Cys-SH/Cys-NH_2 was demonstrated. Simulation studies also, demonstrated the existence of the chair-like transition state, which might be one of the specific features of the catalytic cycle of Ntn-hydrolases. Similar finding were also demonstrated in the cysteine proteases cathepsin (Ma, et al., 2007) and papain (Mladevonic, et al., 2008).

Although the active site organization and catalytic mechanism of BSH is nearly identical to that of PVA, there are subtle variations in the nature of residues and structural elements involved in substrate binding. CGHs share four substrate binding loops flanking the active site (Loop1-4) (Fig. 1.9). The loops are generally shorter in BSHs to accommodate the bulky steroid nucleus in the active site. Loop3 extending from residues ~129-150 is known to be very dynamic and contains many hydrophilic residues in BSH; these residues complement the hydroxyl group of the bile acid substrates for proper

catalytic framework (Panigrahi, et al., 2014). In addition, the residue at 81 position whose main chain NH group is involved in the formation of oxyanion hole also varies between BSH and PVA; the respective side chain seems to strengthen interactions with the substrates. The amino acid at this position is almost always Asn in BSHs; while PVAs contain aromatic residues (Tyr in Gram-positive and Trp in Gram-negative homologues) that help in stacking interactions with the phenyl ring of substrate Pen V (Avinash, et al., 2016).

Residues of the active site are strictly conserved in bile acid hydrolases. On the other hand, the residues for substrate recognition are not particularly conserved, while most substitutions are conservative. The loop3 in BSHs is known to interact with the cholate backbone and contains more polar residues, while it is predominantly hydrophobic in PVAs. Mutation of the polar residues in loop3 of PVA enzymes could be employed as a strategy for engineering substrate specificity in CGHs (Panigrahi, et al., 2014, Avinash, et al., 2013). In spite of a similar tetrameric structure, subtle changes in the loop structure surrounding the binding pocket effect significant changes in catalytic specificity of these enzymes (Lambert, et al., 2008). PVA from *B. subtilis* (Rathinaswamy, et al., 2012) and *P. atrosepticum* (Avinash, et al., 2013) hydrolyze pen V exclusively, while PVA of *B. sphaericus* (Kumar, et al., 2006; Pundle, et al., 1997) can hydrolyse bile salts with a specific activity up to 30% of that shown on pen V as the substrate (Fig 1.1b). Similarly, *C. perfringens* BSH (Rossocha et al. 2005) can hydrolyze pen V to some extent, but the *B. longum* enzyme (Kumar, et al., 2006) exhibits pure BSH activity.

These subtle changes at the substrate binding pocket leads the structural changes which in turn imparts significant difference in the functional characteristics such as substrate specificity and catalytic efficiency. Hence, investigation of the BSH enzyme is important to understand the structural changes that lead the varied substrate activity of the enzyme. The binding of bile salts or Pen V by CGHs enzymes is known to show a directional preference for proper hydrolytic activity (Panigrahi, et al., 2014). The mode of substrate binding is dependent on the loops surrounding the active site, which also define the volume of the site (Fig. 1.9). In the case of GCA (glycocholic acid) binding to *Bl*BSH, the amide bond (CO-NH) was directed from site A (cholate adduct) towards site L (glycine leaving group). The PVA enzyme from the Gram-negative *P. atrosepticum* is highly specific for Pen V; the loops surrounding the active site are much longer and

restrict the volume available for proper binding of the bulky cholate moiety of bile salts. The GCA thus binds in reverse orientation, thus inhibiting the PVA activity even though the enzyme doesn't hydrolyze bile salts (Avinash, et al., 2013).

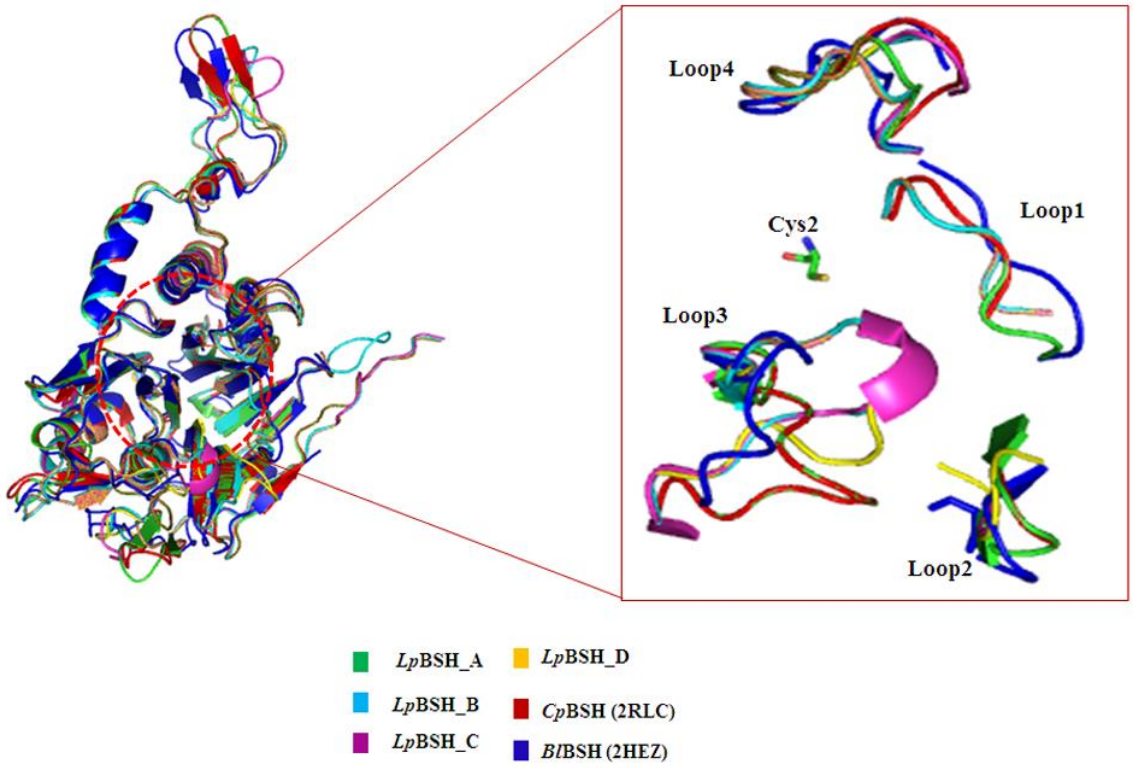


Figure 1.9 Binding pocket loop comparison of the homology modeled, BSH-A, BSH-B, BSH-C & BSH-D from *L. plantarum*, and corresponding loops in the structures of BSH from *C. perfringens* and *B. longum*.

In addition to the directional preference, polar complementarity towards the conjugated bile acids has also been proposed as an important facet of substrate specificity (Batta, et al., 1984). In *Bl*BSH, polar complementarity was found for all three hydroxyl groups of the GCA substrate along with good hydrogen bonding interaction. However, the polar complementarity and BSH activity of *Cp*BSH was observed to be lower than *Bl*BSH (Panigrahi, et al., 2014; Kumar, et al., 2006).

The substrate specificity of BSH has also been attributed to steric hindrance near the amide bond. It has been shown that bile salt hydrolysis decreases with increasing steric hindrance adjacent to the amide bond (Sjovall, et al., 2010). Glycine being the smallest amino acid provides least steric hindrance to rotation close to bonds (Bhagavan,

et al., 2002). However, taurine has a bulky sulphur atom, exhibiting higher steric hindrance, thus most BSHs show substrate preference toward the glyco-conjugated over taurine conjugated bile acids. It is also known that glyco-conjugated bile acids are the major components of the bile acid pool.

The occurrence of multiple BSH isoforms has also been observed in some *Lactobacillus* species. In *L. plantarum*, and *Lactobacillus plantarum WCFS1* which exhibits 4 different BSHs, only a principal BSH shows prominent bile salt hydrolysis activity (Begley, et al., 2006; Elkines, et a., 1997, Lambert, et al., 2008). It is thus probable that the other BSHs could be involved in other roles in the bacterium, given that they share less than 40% sequence similarity with each other. Homology modelling of all four *L. plantarum* BSHs (BSH-A, B, C, D) was performed using *Prime* (Schrodinger), using *Cp*BSH (PDB ID 2RLC, for BSH-A) and *Bl*BSH (PDB ID 2OQC, for BSH-B, C, D) as templates. Notably, comparison of superposed structures showed significant differences in the position of loop3 surrounding the binding pocket (Fig. 1.9). The loop 3 in BSH-A was similar to the template structures; however, it was displaced towards the N-terminal Cys residue in the other BSHs (Fig. 5). This probably reduces the size of the binding pocket, correlating with the loss of BSH activity and probably emergence of the differential substrate activity.

1.5.2.8 BSH in cholesterol lowering property of probiotics

The dynamic, diverse and complex population of intestinal microbiota contributes to the health of the host (Maria, et al., 2012). The ability of these microorganisms to de-conjugate bile salt via the production of bile salt hydrolase (BSH) has been widely associated with their cholesterol lowering potentials in the prevention of hypercholesterolemia (Kim, et al., 2005, Begley, et al., 2006, Ooi and Liong 2010, Homayouni, et al. 2012, Kumar, et al., 2012). De-conjugated bile salts are more hydrophobic than their conjugated counterparts, thus they are less reabsorbed and precipitate in intestines which results in higher excretion into the faeces (Kumar, et al., 2012, Lye, et al., 2010). As a result, the cholesterol flow is drawn towards the bile acid synthesis since cholesterol acts as precursor for this process, subsequently leading to decrease in serum cholesterol levels (Fig.1.6). Previous studies have also shown the relation between cholesterol removal and the deconjugation of bile salts by *L. acidophilus* and *L. casei* isolated from dairy products or from human origin (Liong and Shah 2005).

The use of BSH-producing probiotics has also given rise to some controversies, such as side-effects that have been attributed to the activities of de-conjugated bile acid products of the BSH reaction. De-conjugated bile acids have higher binding affinity towards the farsenoid X receptor (FXR), thus suppressing the transcription of the enzyme cholesterol 7-alpha hydroxylase (7AH) which is required in bile acid synthesis from cholesterol. Modification of bile acids by BSH is known to influence host bile acid synthesis through a feedback mechanism involving the regulation of hepatic cholesterol hydroxylase enzyme Cyp7A1 and Cyp8B1 (Schwarz, et al., 1996, Ishibashi, et al., 1996, Parks, et al., 1999, Chiang, et al., 2009, Evans, et al., 2009). Both the genes *CYP7A1*and *CYP8B1* are negatively regulated by bile acids (Russell, 2003). The increased de-conjugation of bile salts might lead to a few detrimental effects such as disruption of lipid metabolism, colon carcinogenesis, gallstone formation and altered gut microbial populations, which in turn contribute to other gut associated disorders (Jones, et al., 2008, Martin, et al., 2013, Ajouz, et al., 2014).

1.5.2.9 Role in the glucose homeostasis (Diabetes) and lipid metabolism

Recent studies have indicated that bile acid homeostasis is altered in diabetic patients, which suggests an important role for bile acids in type 2 diabetes mellitus (T2DM) (Bennion et al.,1997, Brufau, et al., 2010). T2DM is characterized by a state of insulin resistance and increased blood glucose levels (Edelstein, et al., 1997, Davies, et al., 2004, Prisant, et al., 2004, Padwal, et al., 2004). In peripheral tissues, insulin resistance leads to a decreased uptake of glucose from the blood, whereas hepatic insulin resistance results in enhanced de novo glucose production (Magnusson, et al., 1992, Wang, et al., 2012, Lu et al., 2013). Both phenomena contribute to elevated plasma glucose levels. T2DM has been a growing concern world-wide, but available therapy and preventions are insufficient. Bile acids are known to function as biological detergents to emulsify and solubilize the lipid. Apart from the fat solubilization they are considered as intriguing signaling molecules in the liver and intestine and control the bile acid, lipid and glucose metabolism. Bile acids have been shown to activate the TGR5 signaling pathway (Drucker, et al., 2006), which confers the ability to modulate the energy expenditure by controlling the type 2 deiodinase, in turn activating the thyroid hormone in the brown adipose tissue (BAT). Chenodeoxycholic acid (CDCA) and lithocholic aicd (LCA) have been reported to activate the TGR5 signalling pathway by acting as natural ligands for the TGR 5 (GPCR) receptor (Kwamata, et al., 2003) (Fig. 1.6)**.** These secondary bile acids

are strong agonists of the TGR5 receptor, which further activates the downstream signaling cascade and induce the secretion of the GLP1 in L cells of the intestine. GLP1 receptors are predominantly present in the pancreatic β-cells, and thereby responsible for the restoration of the glucose homeostasis by increasing the expression of the genes glucokinase, pdx-1 and Glut-2 (Wang, et al., 1999; Zhou, et al., 2002, Thomas, et al., 2009). BSH-mediated alteration in the liver bile acid pool is known to regulate (Joyce, et al., 2014) the transcription of genes *Pparγ* and *Angptl4* involved in the lipid metabolism, Abcg5/8 (cholesterol metabolism), *RegIIIγ and Dbp, Per1/2* involved in the gastrointestinal homeostasis and circadian rhythm, respectively. Regenerating islet-derived protein 3 gamma (REGIIIγ) which is an antibacterial protein produced by the host, targets Gram positive bacteria by modulating the local immune homeostasis in the gut (Joyce et al. 2014).

Moreover, the hydrolytic activity of BSH enzyme also leads to the accumulation of de-conjugated secondary bile acids (LCA, CDCA and DCA) that are less efficient than conjugated molecules in the emulsification of dietary lipids and the formation of micelles. This could potentially impair the normal lipid digestion and absorption of fatty acids and monoglycerides (Smet, et al., 1994; Kim and Lee 2005). Such malfunctioning in lipid metabolism has been reported to lead to short bowel syndrome (Bongaerts, et al., 2000). BSH activity has also been related to growth defects in chickens (Feighner 1987; Knarreborg, et al., 2002). In addition, increased LCA concentrations may induce colorectal cancer (McGarr, et al., 2005, Payne, et al., 2008). Accumulation of hydrophobic de-conjugated bile salts also strongly promotes cholesterol crystallization and formation of gall stone in the gall bladder (Begley, et al., 2006, Venneman and van Erpecum 2010).

1.5.2.10 Significance of BSH in microbial physiology

It has been proposed that probably the physiological role of the BSH in the bacteria producing this enzyme is to facilitate the incorporation of cholesterol or bile into bacterial membranes (Begely, et al., 2006, Lye, et al., 2010a). This may increase the membrane flexibility and strength due to change in their fluidity. In this context, it is significant that even certain pathogenic bacterial strains including *Brucella abortus* and *Listeria monocytogenes* possess BSH activity (Dussurget, et al., 2002; Delpino, et al., 2007). BSH activity could potentially offer protection against perturbation of the structure

and integrity of bacterial membranes by the immune system, and such resistance mechanisms may be important in establishing persistent infections (Kumar, et al., 2012).

In addition, it has also been suggested that bacteria that are able to de-conjugate bile salts may be able to use the amino acid released from hydrolysis as carbon, nitrogen, and energy sources, since glycine may be metabolized to ammonia, and carbon dioxide and taurine may be metabolized to ammonia, carbon dioxide, and sulfate (Gopal-Srivastava and Hylemon 1988; Begley, et al., 2006). Besides, BSH decreases the toxicity of conjugated bile acids for bacteria (Smet, et al., 1995, Kim, et al., 2005). Thus, it is considered that BSH activity in the bacteria might have evolved for the self-defence mechanism and persistent colonization in the gastrointestinal tract (Elkins, et al., 2001, Kim, et al., 2005, Ruiz, et al., 2013). Compared with their conjugated counterparts, de-conjugated bile acids have decreased solubility and diminished detergent activity and thus may be less toxic to bacteria in the intestine. The multifaceted role of the BSH related enzymes such as penicillin acylases in the production of the secondary semi-synthetic antibiotics, gene regulation to involvement in the signaling mechanism have been reviewed by the Avinash, et al., (2015).

1.5.2.11 Role of BSH in pathogenesis and quorum sensing

BSHs have also been characterized from pathogenic *L. monocytogenes*, considered as a member of the normal enteric flora (Dussurget, et al., 2002) and *Brucella abortus* (Delpino, et al., 2007), the causative agent of cow disease Brucellosis. Notably, it has been observed that in these bacteria, BSH also helps in the colonization and establishment of the pathogenic infection. BSH deletion mutants of *L. monocytogenes* showed substantial growth reduction in the presence of bile salts. In addition, along with *bsh*, *pva, btlB* and sigma b factor are also implicated in the bile acid resistance in *L.moncytogenes EGDe* (Begley, et al., 2005). The GC content of the *bsh* gene in *L. monocytogenes* is lower than that of the neighbouring genes, indicating that pathogenic strains could have acquired the *bsh* gene from the gut residing bacteria such as lactobacilli which are low in GC content (Dussurget, et al., 2002).

The Δ*cbh* deletion mutant was generated from *B. abortus* which was sensitive to BA, although it showed resistance towards the polymyxin B (cationic detergent) and faster growth in comparison to the wild-type. The pleiotropic effects due to Δ*cbh* reveal that CGH plays an important role in *B. abortus* physiology besides bile salt de-

conjugating activity. The Δ*cbh* mutation was also observed to influence the adhesive and internalization activity shown by *B. abortus*. Adaptive properties of Δ*cbh* in *Brucella abortus* are appreciated by differential expression of the cell envelope-associated proteins (Delpino, et al., 2007; Maria, et al., 2011). Moreover, Jorocki, et al., (2014) have reported the gel forming ability in the presence of bile salts of *B. pseudocatenulatum* and *B. longum suis,* exhibiting BSH activity, reinforcing the role of BSH in the efficient colonization and persistence of strains in the intestine.

A possible involvement of BSH and related enzymes in disrupting the bacterial signaling mechanism, quorum sensing, has also been suggested. Quorum sensing involves the perception of bacterial population size through the detection of auto-inducer molecules, consequently leading to the expression of selected genes. Many metabolic phenomena including virulence, bioluminescence, biofilm formation and antibiotic production are now known to be regulated through quorum sensing. Gram-negative bacteria use acyl homoserine lactones (AHLs) as their main signal, and the disruption of quorum sensing through enzymatic cleavage of AHLs has been put forward as a new anti-microbial strategy. AHL molecules are made of a homoserine lactone (HSL) moiety linked to acyl chain of variable length (C4-16). Apart from AHL acylases from *Pseudomonas* sp. (Bokhove, et al., 2010), the PGA enzyme from *Kluyvera citrophila* (Mukherji, et al., 2014) and PVA from *Streptomyces lavendulae* (Torres, et al., 2015) have also been observed to deacylate medium chain AHLs (C6-8 HSL). Considering that these enzymes belong to the same Ntn hydrolase superfamily as BSH, the recent identification of a *Streptococcus epiderdimis* (Mukherji and Prabhune 2015) having both BSH and AHL acylase activity could point to a possible role for BSHs in AHL degradation as well. Thus, judging by the various reports on BSHs from different environments, it appears that this enzyme is highly promiscuous and evolved to perform multiple functions in the bacteria as well as interacting with the host metabolism. Hill (2012) has suggested the use of the term "niche factor" for BSH as the enzyme is involved in both pathogenic and probiotic roles.

1.5.2.12 BSH inhibitors, an alternative to antibiotics growth promoters (AGP)

Sub-therapeutic levels of antibiotics have been used in the animal food industry for more than 50 years as food additives, in order to manipulate the microbial populations in the intestine and gain body weight (Frost and Woolcock, 1991). However, this has also

been linked to the emergence of antibiotic resistant bacteria (Wegener, 2003), which poses a threat to human health. BSH activity being a widely distributed function of the gut microbiota, has a profound impact on enhancing host lipid metabolism and energy harvest (Begley, et al., 2006). Joyce, et al., (2008) have demonstrated that in both gnotobiotic and conventionally raised mice, BSH activity has significant influence on host lipid metabolism. Moreover, AGPs are thought to act by targeting the BSH enzyme (Lin, et al., 2003). The AGP-mediated weight gain in animals can be correlated to the reduction in the levels of BSH enzyme and *Lactobacillus* sp. in the intestinal microbiota. These reports suggest that BSH can be a promising microbiome target for developing promising BSH inhibitors as novel alternatives to AGP (Smith, 2014). Approved feed additives like KIO_3, $NaHIO_3$, $NaIO_4$, $CuSO_4$, $CuCl_2$, $ZnSO_4$ and $ZnCl_2$ already used by the food industry could affect significant inhibition of *L. salivarius* BSH (Wang, et al., 2012). In addition, compounds such as menadione, riboflavin, gossypetin and caffeic acid phenethyl ester (CAPE) were found to be more potential inhibitors of BSH (Fig. 1) using a high throughput screening approach (Smith et al. 2014). Such inhibitors could be applied along with probiotics to develop new dietary supplements that regulate weight gain in animals.

1.6 Serine Ntn-Hydrolase family

The free α-amino group (-NH2) of the same N-terminal residue serves as *base* in order to generate the nucleophilic atom. In the NtSn-hydrolase family serine acts as nucleophile. NtSn-hydrolase enzymes also share the characteristic catalytic domain of four layered αββα core structural fold, known as the **Ntn-hydrolase fold** (Fig.1.1) (Brannigan, et al., 1995; Oinonen & Rouvinen, 2000). All Ntn-hydrolase members show spatially conserved active site topology and chemistry due to which the enzymes have similar reaction mechanisms (Duggleby, et al., 1995). However, due to variation in size and shape of the substrate binding pockets, enzymes display variation in their substrate specificity. This family possesses the cephalosporin acylase, AHL acylase and penicillin G acylase, which are very important enzymes in the production of the secondary semi-synthetic antibiotics and quorum quenching.

1.6.1 Cephalosporin acylase

Cephalosporin acylase (CA) is a commercially important enzyme which is involved in the production of 7-aminocephalosporanic acid (7-ACA) utilized in the synthesis of the semi-synthetic antibiotics.

Figure 1.10 Crystal structure of the Cephalosporin acylase from *Brevundimonas diminuta* (PDB ID: 1JW0 Kim, et al., 2001). Figures have been prepared using Pymol

Figure 1.11 Crystal structure of the Cephalosporin acylase from *Pseudomonas sp. SY-77-1* (PDB ID: 2ADV Kim, et al., 2006).

Semisynthetic cephalosporins are important antibacterials clinically. 7-ACA, a precursor for the industrial production of semisynthetic cephalosporins, obtained by deacylation of CPC (cephalosporin C) by use of the iminoether, nitrosyl chloride and methanol (Huber, et al., 1972).The structure of the cephalosporin acylase complex with glutaryl-7-aminocephalosporanic acid and glutarate was crystallized by Kim, et al., 2001 (Fig. 1.10). The CA is composed of the two non-identical α and β-subunit which are connected by the spacer peptide. The Precursor polypeptide of 74 kDa possesses the signal peptide, 16 kDa α-subunit, 54 kDa β-subunit and 9 amino acids spacer peptide. During the process of maturation spacer peptide is cleaved off similar to other Ntn hydrolases. CA from *Pseudomonas sp. strain SY-77-1 a* heterodimer (Fig. 1.11) of 166 aa and 55 kDa has been crystallized (Kim, et al., 2006).

1.6.2 Penicillin G acylase

The Penicillin G acylase (EC 3.5.1.11), belongs to the family of NtSn-hydrolases super family, also nomenclature as penicillin amidohydrolase which hydrolyses the carbon-nitrogen bonds (amide bond) other than peptide bonds. Enzymatic hydrolysis of penicillins to 6-APA was first reported by Sakaguchi & Murao (1950) with the enzymes obtained from *Penicillium chrysogenum Q176* and *Aspergillus oryzae*.

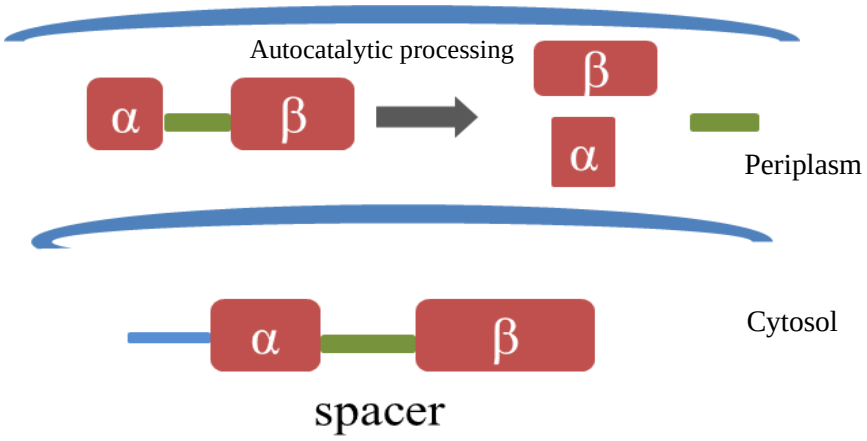


Figure 1.12 Schematic representation of the synthesis and maturation by autocatalysis processing of the penicillin G acylases. The enzyme is synthesized in the cytosol with a signal peptide which transports the inactive pre-pro-peptide to the periplasm where autocatalysis processing takes place.

The penicillin acylase are synthesized as pre-pro-peptide in the cytosol which remains in the inactive form where spacer peptide blocks the active site of the enzyme and transports to the periplasma where autocatalytic processing occurs (Fig. 1.12). The crystal structure of PGA from *E. coli* was reported to 1.9 Å resolution (Duggleby, et al., 1995).

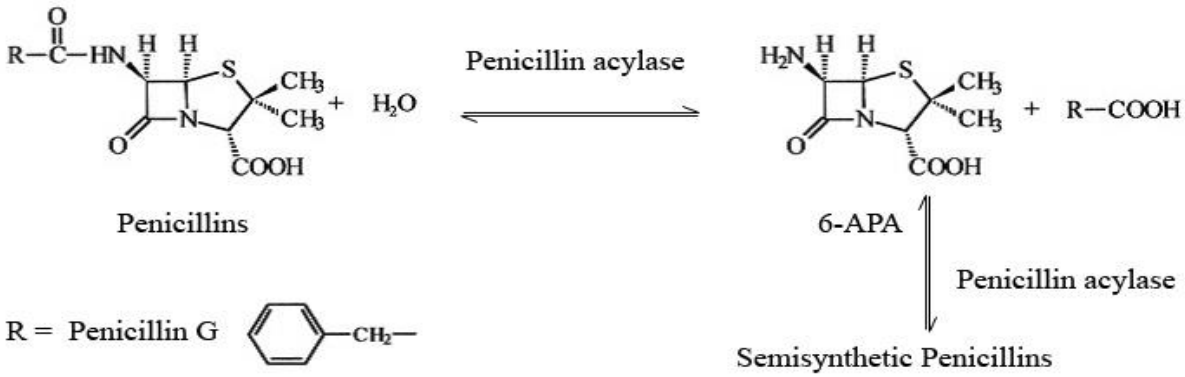


Figure 1.13 Hydrolysis of the penicillin by penicillin G acylases and production of by-product 6-APA (6-aminopenicillanic acid), a precursor for the secondary semi-synthetic antibiotics.

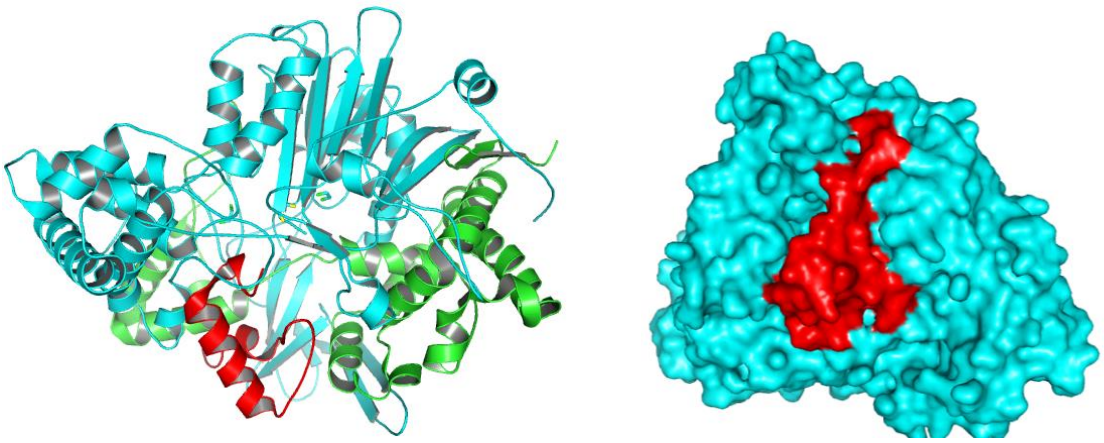


Figure 1.14 [a] Crystal structure of the a slow processing precursor penicillin acylase from *Escherichia coli* (PDB ID: 1E3A., Hewitt, et al., 2000). α- and β-subunits are shown in the green and cyan colour respectively. Spacer peptide is shown in the red colour, [b] Surface view of unprocessed PGA demonstrating spacer peptide (red) blocking active site groove.

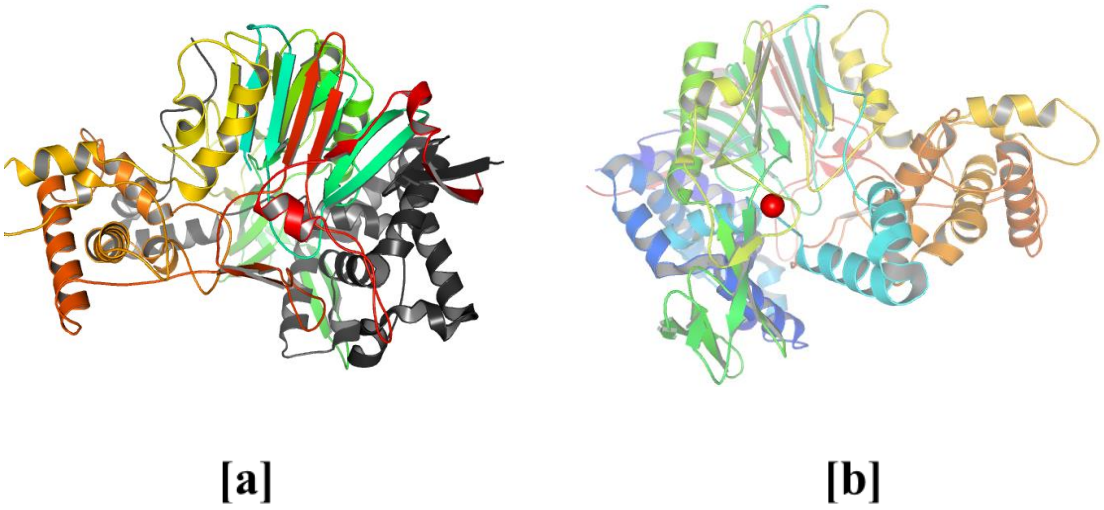


Figure 1.15 [a] Crystal structure of the penicillin G acylase from *Alcaligenes faecalis* in tetragonal form (PDB ID: 3ML0, Varshney, et al., 2012), [b]. Calcium ion bound to the PGA (red colour).

PGAs are widely distributed in the whole kingdom of prokaryotes and eukaryotes (Arroyo, et al., 2003). PGA has been studied from various bacteria *E. coli* (Schumacher, et al., 1986), *Arthrobacter viscosus* (Ohashi, et al.,1988), *Alcaligenes faecalis* (Verhaert, et al., 1997), *Kluyvera citrophila* (Barbero, et al., 1986; Martin, et al., 1991), *Proteus rettgeri* (Klei, et al., 1995), *Bacillus megaterium* (Chiang & Bennet, 1967; Martin et al., 1995), and *Achromobacter xylosoxidans* (Cai, et al., 2004). PGA has also been isolated from two fungal sources, *Aspergillus fumigatus H/6.17.3* and *Mucor griseocyanum H/55.1* (Jose, et al., 2003). PGAs are exploited extensively in the industrial application for the production of the semisynthetic antibiotics from 6-APA, which is obtained from the hydrolysis of Pen G substrate (Fig. 1.13). It is estimated that 10-30 million tons of immobilized PGA are utilized each year (deSouza, et al., 2005). The serine residue at N-terminal of the β-chain is present at the binding pocket and is involved in catalysis. Upon mutation of nucleophile serine, the catalytic base residue at the active site of *E. coli* PGA resulted in loss of the activity which led to the hypothesis that penicillin acylase has a single-amino-acid catalytic centre and classified in the Ntn-Hydrolase superfamily (Brannigan, et al., 1995, Duggleby, et al., 1995). Figure 1.14 has shown the crystal structure of the slow processing protein where spacer peptide has been captured in the crystal structure of the PGA from *E.coli.* shown in the red color in the surface view model (Fig. 1.14b) blocking the active site of the enzyme hence keeping it in the inactive form.

The important residues of the catalytic sites are Arg145 and Phe146 from the α-chain and Serβ1 Ala69 Asn241of the β-chains. The hydroxyl group of serine at the N-terminal of β-chain (Serβ1) acts as a nucleophile which is activated by its own α-amino group through bridging water molecule (Duggleby, et al., 1995; Dodson, 2000). However, the PGA crystal structure in complex with Pen G sulphoxide demonstrated presence of no catalytic water molecule suggesting that the auto-activation of the α-amino group of the nucleophile Serβ1of β-chain (McVey, et al., 2001). The transient tetrahedral intermediate is stabilized by the main chain amide of AlaB69 and the side chain nitrogen of AsnB241 of the β-chain PGA. The calcium ion is present at the interface between α and β-chains in the mature PGA enzyme structure from *E. coli* and *P. rettgeri* as well as in the precursor structure of *E.coli* which is coordinated by the residues, side chains of aspartyl residues 336, 339 and 515, and the main-chain carbonyl groups of Val338 and Pro468 from β-chain and the side-chain of Glu156 from α-chain in the case of *E.coli*.

The PGAs requires the Ca++ ion for their activity which is tightly coordinated by residues from both α- and β-subunits (Fig. 1.15b). PGA enzyme also shows the substrate activity towards the quorum sensing molecules, AHL (N-acyl homoserine lactone) apart from activity towards its true substrate Pen G (Mukharji, et al., 2014). Gram-negative bacteria use the *N*-acyl homoserine lactones (AHLs) as quorum sensing molecule.

Several Applications including peptide synthesis, where penicillin G acylase from *E. coli* proved to be an effective biocatalyst in the synthesis of the sweetener aspartame (Fuganti *et al*., 1986, resolution of racemic mixtures**,** such as amino acids (Fadnavis, et al., 1997; Bossi, et al., 1998), β-amino esters (Roche, et al., 1999) and secondary alcohols (Svedas, et al., 1996) in aqueous media and *β*-Lactam Antibiotics in which 6-APA and 7-ADCA are the central precursor molecules semi-synthetic secondary antibiotics worldwide.

1.7 Autocatalysis processing in Ntn-hydrolase

The mode of processing varies in different Ntn hydrolases. Autocatalytic processing in the Ntn-hydrolases occurs due to nucleophile residue (Ser, Thr, Cys) of enzyme which induces the removal of the spacer peptide auto-catalytically (Lee & Park, 1998; Li, et al., 1999; Hewitt, et al., 2000; Kim, et al., 2002). Most Ntn hydrolase enzymes are produced as inactive pro-enzymes. The pro-enzyme activates itself by a post translational intra-molecular autocatalytic peptide bond cleavage, which exposes the N-terminal nucleophile residue responsible for catalysis (Guane, et al., 1996; Seemuller, et

al., 1996; Xu, et al., 1999). On the basis of the autocatalytic activity, all the Ntn hydrolase pro-enzymes have been classified as peptidases in MEROPS database (Rawlings, et al., 2012).

PGAs are synthesized in cytosol with signal pre-peptide and a spacer peptide between the α and β subunits. They are transported to the bacterial periplasm by way of signal peptide, and further maturation takes place by autocatalytic post-translation modification to give the active heterodimeric enzyme (Brannigan, et al., 1995). Several mutation studies have been carried out to understand the autocatalytic processing. Site-directed mutagenesis of the residue at the enzyme's active site (Choi, et al., 1992), or at the processing site (Choi, et al., 1992; Hewitt, et al., 2000) affects the processing of the precursor, hence indicating that two processes use the same set of residues (Brannigan, et al., 2000). In the case of the *Ec*PGA, and *Af*PGA resulted the slow processing and inactive mutant of the PGA where spacer peptide remains bound in the catalytic groove (Hewitt, et al., 2000, Varshney, et al., 2012) (Fig. 1.15a) . In many cases, the N-terminal nucleophile residues could be substituted with one another, sometimes even without affecting the autocatalytic processing or activity of the enzyme. Cys can replace Ser in PGAs without affecting processing but this mutation inactivates the enzyme (Choi, et al., 1992). However, in proteasome where Thr is the nucleophile residue, Ser can be a suitable alternative for proteolytic activity, whereas Cys is effective for processing only (Seemuller, et al., 1996). Interestingly, substituting nucleophile Cys in *Bacillus sphaericus* PVA with Ser incapacitates processing and thus is inactive (Manish, et al., 2005, PDB: 2IWM).

However, BSHs have not been found to possess a pre-peptide, instead they only have the initial methionine before the nucleophilic cysteine residue. The N-terminal methionine in BSHs is cleaved by methionine aminopeptidase (MAP). Panigrahi, et al., (2014) have explored the distribution and length of the pre-peptide in cholylglycine hydrolases. While Gram-negative CGH homologues (including *Bt*BSH) contain a periplasm-directed long pre-peptide, many CGHs from Gram-positive bacteria possess only the methionine before the nucleophilic cysteine. *B. sphaericus* PVA has been reported as an exception (Suresh, et al., 1999) to process a tripeptide (Met-Leu-Gly) at the N-terminal end of the nucleophile cysteine; the *B. subtilis* PVA (Rathinaswamy, et al., 2005) and all BSHs possess only a methionine. Acid-base catalysis by conserved residues at active site and oxyanion hole are critical to facilitate the N›O or N›S acyl

shift or to stabilize the reaction intermediate to carry out the autocatalysis processing. Rossmann et al 2008, has hypothesised the role of Arg18 of *clostridium perfringens* BSH in autocatalysis processing. Apart from the nucleophile residue, the autocatalytic potential also originates partially from the extremely strained precursor and tight turn conformation at the scissile peptide bond. (Xu, et al., 1999, Ditzel, et al., 1998).

1.8 Scope and nature of the work

In this thesis our endeavour is to understand the structure-function relation of the protein and how the subtle changes which imparts significant differences to the functional characteristics of the enzyme. We have selected two enzymes bile salt hydrolase (*Ef*BSH) from *Enterococcus faecalis NCIM 2403,* belonging to NtCn-hydrolase family, and penicillin G acylase (*Kc*PGA) from *Kluyvera citrophila,* belonging to NtSn-hydrolase family. These enzymes have crucial role clinically and in industrial application.

BSH is a very important enzyme which degrades the bile acids in the GI tract, produced by the gut microbes and is responsible for the lowering of the cholesterol level in the blood. The BSH activity is one of the major requirements of a strain to be considered for the probiotic. BSH activity participates in crosstalk with host signaling pathways and has roles in lipid metabolism, glucose metabolism and circadian rhythm. The significance of bile acid de-conjugation by BSH enzyme in host metabolism and physiology underlines the potential for novel strategy to control the obesity (lowering of cholesterol) and associated metabolic syndrome.

The *Ef*BSH gene was cloned in pET22b+ and heterologous expression in *E. coli.* was performed. The enzyme was studied biochemically and biophysically to understand the optimum pH, temperature, stability and catalytic efficiency. The protein was expressed in very high levels (80mg/lit) in soluble form. It showed unique characteristics of the allosteric behaviour and very high catalytic efficiency as compared to the reported BSHs. In addition 3-dimentional structure of the protein was elucidated to understand the structural changes that have occurred in the *Ef*BSH which are imparting to the protein high catalysis and cooperativity kinetics. The mutational studies were also performed to understand the important residues in the substrate binding and catalysis which could provide insight to improve enzyme activity in protein engineering applications.

Moreover, autocaltalytic processing of the pre-peptide in the cholylglycine hydrolase family was studied by inserting the di-and tri-peptide before the nucleophile cysteine of *Ef*BSH.

Penicillin acylases have been long studied as enzymes important in the pharmaceutical industry. Penicillin G acylases have been characterized from a diverse range of bacteria and fungi. They catalyze the deacylation of natural penicillins to the active pharmaceutical intermediate 6-aminopenicillanic acid. Penicillin acylases are now widely used in the production of semisynthetic antibiotics. Penicillin G acylase from *Kluyvera citrophila* (*Kc*PGA) is known for its resilience towards harsh conditions of pH, temperature and organic solvents. *Kc*PGA undergoes post-translational processing to reach active form.

The *Kc*PGA_Serβ1Cys mutant was prepared, expressed and crystallized. This is a fully processed but inactive mutant of *Kc*PGA. The crystal diffraction data were collected and 3-dimentional structure was determined. To understand the autocatalytic processing of spacer peptide in *Kc*PGA processed Serβ1Cys mutant was compared to the unprocessed Serβ1Gly mutant structure and residues involved in autocatalysis were identified. In addition, *Kc*PGA shows activity toward AHL, apart from true substrate Pen G. However, *E. coli.* PGA even sharing 80% sequence similarity does not show AHL degradation activity. Hence, structural basis for the AHL hydrolysis was explored by comparing the *E.coli* PGA and *Kc*PGA structures.

Chapter-2

"Cloning, expression, biochemical and biophysical characterization of BSH from *Enterococcus faecalis* (*Ef*BSH)"

2.1 Introduction

BSH (cholylglycine hydrolase, E.C.3.5.1.24) is an enzyme involved in the crucial metabolic pathway of mammalian bile acid metabolism. The BSH enzyme belongs to the Ntn (N-terminal nucleophile) hydrolase structural superfamily and the cholylglycine hydrolase subfamily (Pfam: PF02275). Evolutionarily closely related BSH and penicillin V acylase (PVA) share all conserved active-site residues. Although distributed among bacteria (Gram positive, Gram negative) and the archaea kingdom, it is more widely found in gut microbes (Jones, et al., 2008). Several microbial species, including *Bifidobacterium*,(Tanaka, et al., 2000, Kim, et al., 2004, 2005, 2006, Kumar, et al., 2006, Öner et al., 2014, Jarocki, et al., 2014), *Lactobacillus*,(Lundeen et al., 1990, Chrishtiaens et al., 1992, Coleman et al.,1995, Anderson et al., 1999, McAuliffeet al., 2005, Chae et al., 2013, Gu et al., 2014) *Clostridium perfringens* (Gopal et al., 1988, Rossocha et al., 2005,) *Brevibacillus* (Sridevi et al., 2009) *Bacteroides* (Stellwag et al., 1976) and *Xanthomonas maltophilia*,(Dean et al., 2002) are reported to show BSH activity. In addition, some pathogenic bacteria such as *Listeria monocytogen* (Dussurget, et al., 2002, Sue, et al., 2003) and *Brucella abortus* (Delpino, et al., 2007) also showed BSH activity.

Bile salts are produced by liver hepatocytes to assist in dietary absorption of fat in the small intestine. Bile acids help in the solubilization and transport of cholesterol and fat across the intestinal epithelium (Eyssen, et al., 1973, Kim, et al., 2005, Begley, et al., 2005), these are the key signalling molecules that regulate host lipid metabolism and mucosal defence in the intestine.(Jones et al., 2008, Joyes et al., 2014). BSHs are present in most intestinal microbes, conferring resistance against the antimicrobial activity of bile. Two sets of molecules, the true substrates and substrate analogues, were tested for BSH activity. The substrates, the bile salts, showed variations in activity depending on the type of conjugated amino acid moiety (glycine or taurine) present and also depended on the position of the –OH group in the cholesterol ring. The reports which shows substrate specificity based on - OH group of GCA (Panigrahi, et al., 201, Bata, et al., 1984).

High blood cholesterol level is a major risk factor for occurrence of coronary heart disease (Pereira and Gibson, 2002). Pharmacological agents such as nicotinic acids, bile acid sequestrants and fibrates are also available for treatment of high cholesterol.

Although these drugs effectively reduce cholesterol levels, they are expensive and known to have side effects (Bliznakov, 2002; Willman, 2001).

BSHs from *Lactobacillus* have been studied in rats (Chikai, et al., 1987, Park et al., 1996) and humans (Larsen et al., 2000, Anderson et al., 2000). Hence, the role of BSH has been greatly attributed to lowering the serum cholesterol levels (Begley, et al., 2006) as well as maintaining bile acid pool homeostasis in the gastrointestinal (GI) tract, which consequentially reduces further complications related to hypercholesterol. Therefore, BSH activity is primarily considered in the study of probiotic strains.

The present chapter describes the cloning, over-expression and detailed biochemical and biophysical characterization of the Bile Salt Hydrolase from *Enterococcus faecalis* NCIM 2403 strain. The enzyme displayed high expression in soluble form, and demonstrated many-fold higher specific activity and distinctive enzyme kinetics as compared to other reported BSH enzymes. In addition, the non-substrate ligand Pen V exerted a positive modulating effect on *Ef*BSH activity.

2.2 Materials

Genomic DNA was isolated from *E. faecalis* strain NCIM 2403 obtained from NCIM of NCL. All enzymes used for DNA manipulation were procured from New England Biolabs (NEB). pET22b+ vector was used for cloning and *Escherichia coli.* DH5α cells were used as the maintenance host. The clone was transformed in the BL21* (DE3) expression host. Conjugated bile acids (GCA; GDCA; taurocholic acid, TCA; glycochenodeoxycholic acid, GCDCA; and taurochenodeoxycholic acid,TCDCA), guanidine hydrochloride (GdnHCl) and Pen V (potassium salt) were procured from Sigma (USA). The Sephadex200 (SEC-650 Bio-Rad) column for size exclusion chromatography and molecular weight markers were purchased from Bio-Rad. Ninhydrine was procured from Sigma (USA). Chemicals used in fermentation and protein expression were yeast extract, tryptone, sodium chloride (NaCl), ampicillin, kanamycin, isopropyl thiogalactopyranoside (IPTG) were purchased from Sigma chemical company, St. Louis, USA. Ammonium sulfate (AS), sodium acetate, nickel chloride, glycerol, β-mercapto-ethanol, bromo-phenol-blue (BPB), acrylamide, N,N'-methylenebisacrylamide, sodium dodecyl sulfate (SDS), acetic acid, methanol, TEMED (N,N,N',N'-Tetramethylethylenediamine), ammoniumpersulfate (APS), substrates and ligands such as

6-APA, PAA, POAA and compounds from bile salt kit, molecular weight marker kits for SDS-PAGE etc, were also purchased from Sigma chemical company, St. Louis, USA.

2.3 Methodology

2.3.1 Sequence analysis

The gene sequence of *Ef*BSH was retrieved from NCBI with GenBank ID EET97240.1 (http://www.ncbi.nlm.nih.gov). The global alignment of amino acid sequences and the percentage identity were calculated using the EMBOSS pairwise alignment algorithms (http://www.ebi.ac.uk/emboss/). The theoretical p*I* of and molecular weight of the protein was calculated from the ExPASy ProtParam tool server.

2.3.2 Construction of over-expression plasmid vector pET22b/*Efbsh*

Chromosomal DNA was isolated from *Enterococcus faecalis* NCIM strain 2403 by the bacterial DNA isolation protocol Sambrook et al. (1988). Nucleotide sequences for *Enterococcus faecalis* BSH was available at NCBI GenBank database. Gene was amplified by gene specific primers *Ef*BSH forward primer FP*_NdeI* 5'-GCGGCG**CATATG**TGTACAGCAATTACTTATGTATC-3' and reverse primer *Ef*BSH_Rev*_XhoI* 5'-CGCCTA **CTCGAG**ATTAGCATAGTTAATTTGTTGTG-3'. The restriction sites are highlighted. The PCR was carried out in a thermal cycler (AB-life technology, USA) under the following conditions: 1 cycle of pre-denaturation at 95°C (2 min) followed by 28 cycles of denaturation at 95°C (30 s), primer annealing at 55°C (45 s) and extension at 72°C (1 min), a final extension at 72°C (10 min), and then cooled to 4°C. The PCR amplified products were analysed on 1% agarose gel.

2.3.3 Restriction digestion of plasmid (pET22b+) vector and PCR product

The vector pET-22b+ (Fig. 2.1) and amplified PCR product were digested with appropriate restriction enzymes (*NdeI /XhoI)*. Both reactions were carried out at 37ºC with suitable buffers for 4 hr. After restriction digestion the products were loaded on to a 1.0 % agarose gel. The bands were carefully sliced from the gel. Digested vector and PCR products from the gel slices were purified from the gel using Invitrogen gel extraction kit. The digested product was visualized on the gel and quantified by Nanodrop (Thermo-scientific).

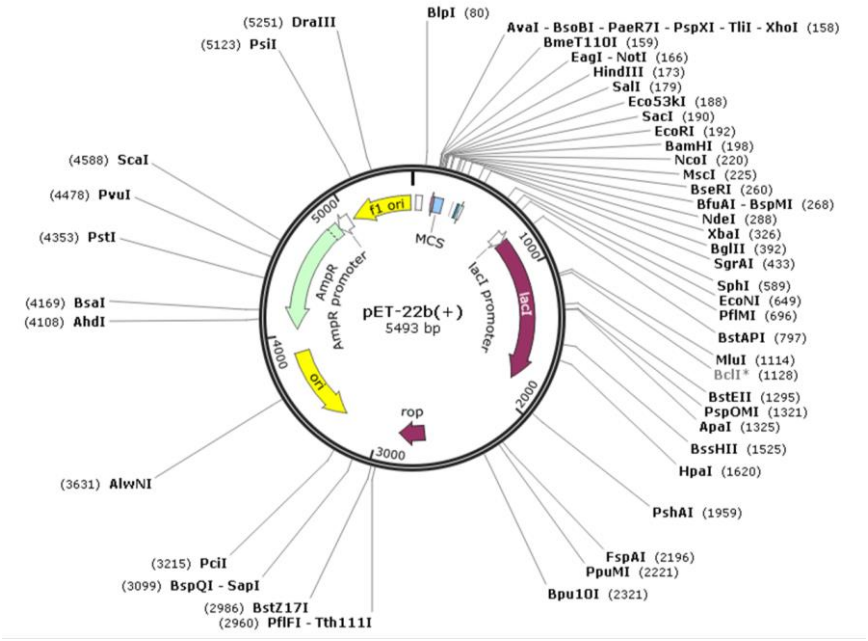


Fig. 2.1 Map of expression vector pET22b+

2.3.4 Ligation

Approximately 1:10 ratio of vector to insert was incubated with T4 DNA Ligase from Invitrogen with suitable buffer at 16ºC overnight.

2.3.5 Transformation of cloned plasmid (pET22b+/*bsh*) and colony PCR

Cloned plasmid pET22b+/*bsh* was transformed in the maintenance host DH5α and plated in the LB agar containing 100 mg/ml Ampicillin and the plate was incubated at 37^{o}C overnight. Single colony was picked up and colony PCR was performed using gene specific primers to confirm the *bsh* gene insert in the cloned plasmid.

2.3.6 Heterologous expression, cell extract preparation and protein purification

The recombinant plasmid (pET22b-*Ef*BSH) was mobilised into BL21* (DE) competent cells. A single colony of the transformed BL21* (DE) possessing the pET22b-*Ef*BSH construct was inoculated in LB media containing 100 µg/ml of ampicillin and incubated at 37°C until an optical density (OD_{600}) of 0.4–0.6 was attained. Further, isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to the media up to a concentration of 0.5 mM for induction followed by shaking incubation (180 rpm) at 16°C

overnight (16–18 h). Cells were centrifuged at 4000 rpm for 10 min at 4°C. They were then suspended in 25 ml of lysis buffer (25 mM Tris-HCl of pH 7.5, 20 mM imidazole and 500 mM NaCl), sonicated at 50% amplitude (E-squire Biotech Pvt Ltd) and centrifuged at 12,000 rpm for 30 min at 4°C. Purification of *Ef*BSH was performed by His-select Ni–NTA (Quiagen) followed by Sephadex200 size exclusion chromatography (SEC 650, Bio-rad). The purity of the protein was assessed on 12% sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE) by Coomassie brilliant blue (CBB) R-250 staining.

2.3.7 Determination of molecular weight

SDS-PAGE (12% gel) analysis determined the subunit molecular weight. The gel was stained with 0.25% Coomassie brilliant blue R250 in 40% (v/v) methanol and 10% (v/v) glacial acetic acid. Electrophoresis was conducted at 150 V (constant voltage). The apparent molecular weight was calculated by comparing the migration of the protein with Bio-Rad molecular weight markers.

2.3.8 Western blot analysis

Western blot was performed for the confirmation of the expression of the protein by electro-blotting on to the PVDF membrane. Blocking of the PVDF membrane was done with 5% bovine skim milk in 1X PBS for overnight followed by 3 washes of PBST (PBS+ Tween 20) buffer in the 5 minute interval. The PVDF membrane was incubated with primary anti-polyhistidine antibody at 1:3000 dilutions for 2 hours and subsequent three washes with PBST. Further membrane was incubated with secondary anti-mouse IgG (Fc-region specific) antibody at the 1:2000 dilutions for the 1 hour. Finally membrane was incubated with substrate (DAB) for visualizing the band.

2.3.9 Bile salt hydrolase Assay: TLC plate and ninhydrine assay

TLC plate assay is qualitative assay where 10 mM substrate GDCA was incubated with BSH enzyme at optimum temperature for 10 minutes. Supernatant of reaction mixture was taken out and loaded onto the silica TLC plate and run with glacial acetic acid (7 ml): Methanol (93 ml) solvent system. After running sample molybdenum phosphate was sprayed on TLC plate which binds with cholesterol ring and incubated in

oven at 90°C till blue band appeared. BSH enzyme activity is measured by estimating glycine released by de-conjugation

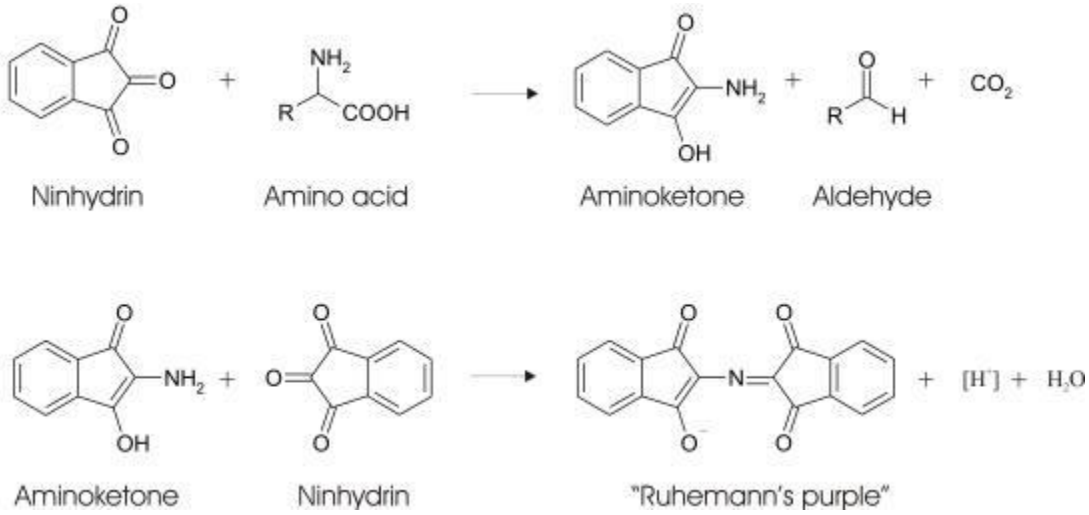


Figure 2.2 Detection of primary amines with ninhydrin. Hydrolyzed amino acid reacts with ninhydrin to turn into a deep blue or purple colour known as Ruhemann's purple that is quantified photometrically at 570 nm.

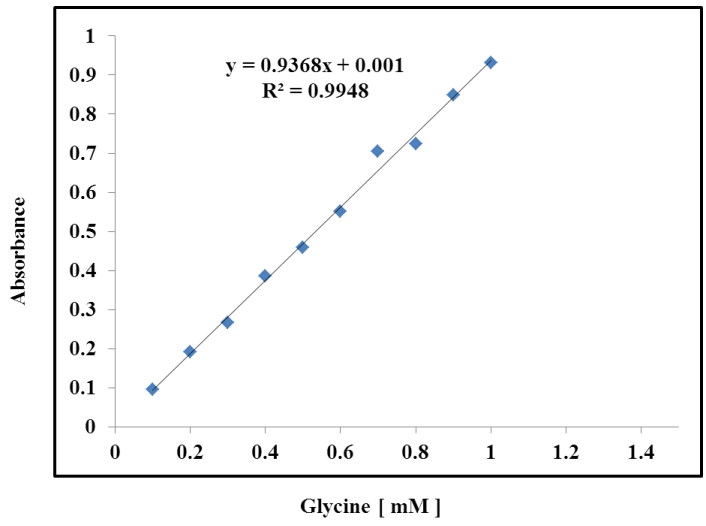


Figure 2.3 Standard graph of Glycine (mM) estimation.

activity on incubating the BSH enzyme with GDCA or TDCA at 50°C in 20 mM Citrate Phosphate buffer pH 5.0 (CPB). Reaction was quenched after 10 min incubation at 50°C using equal volume of 15% Trichloro acetic acid (TCA) and centrifuged at 10000 rpm for 5 min. 20 µl of the reaction mixture was withdrawn and equal volume of ninhydrine was

added and kept at 100^{o}C for 15 min. Product formed by the reaction gave blue colour (Fig.2.2) which was recorded at 570 nm. One unit of BSH activity is defined as the amount of enzyme that liberates 1 μmol of amino acid from the substrate per minute. Specific activity was defined as the number of units of activity per milligram of the pure protein. Protein concentration was measured by the Bradford method. Glycine standard graph was prepared to quantify the product released on substrate hydrolysis (Fig. 2.4).

2.3.10 Effect of pH and temperature on *Ef*BSH activity and stability

The BSH activity was assayed at different pH values in the range 3.0–11.0 and temperatures 25–80°C to determine the optimum conditions for enzyme activity. Enzyme activity and stability were also studied at different pH in the range 1.0–11.0 and temperatures 25–80°C. The effect of pH was studied by incubating the protein in 100 mM buffers of different pH (1.0–11.0), HCl-KCl (1-3 pH), citrate phosphate buffer (3-6), Tris-HCl (7-9) and carbonate-bicarbonate (10-11) for 4 hr and assaying the residual activity. Effect of the temperature on enzyme stability was studied at different temperatures from 25-80°C at different time intervals.

2.3.11 Effect of chemical denaturant GdnHCl on *Ef*BSH

The *Ef*BSH enzyme was also incubated with increasing concentrations of denaturant Gdn-HCl (0-6M) for 4h to study its unfolding effect and activity loss. The effect of chemical denaturant GdnHCl (2M) on oligomeric nature of *Ef*BSH was also evaluated using size exclusion chromatography (Bio-rad, SEC-650).

2.3.12 Steady state fluorescence measurement

Fluorescence of protein upon different treatment (different pH and GdnHCl concentration) was measured using a PerkinElmer Life Sciences LS50 fluorescence. Fluorescence emission spectra of *Ef*BSH were measured with slit width of 7 nm for both the monochromators. Spectra for the either buffer or buffer containing substrate were recorded in order to eliminate the background emission noise by correcting with observed fluorescence of the sample. Suitable amount of enzyme in 20 mM sodium citrate phosphate buffer, pH 5.0 buffer were filtered with the 0.22um fitters and protein was centrifuged at high speed 13000rpm in order to remove the particulate and aggregates in

the solution. 2 ml Samples were used throughout experiment at constant temperature (25°C ± 0.1°C) in a quartz cuvette to record the fluorescence spectra with the help of a Julabo F 25 circulating cryobath. Samples were excited at 295 nm and the emission spectra were recorded from 300 to 400 nm.

2.3.13 Circular dichroism (CD) spectroscopy

Circular dichroism (CD) is difference in absorption of right-handed circularly polarised light (R-CPL) and left-handed circularly polarised light (L-CPL). It occurs when a molecule contains the chiral chromophores. It is a spectroscopic technique where spectra of chiral molecules are measured over a range of wavelengths. This is employed extensively to study biological macro-molecules; protein and nucleic acid where it finds its most important applications to understand the structural conformational changes. Proteins are optically active due to their structural asymmetry and amino acid configuration. CD spectroscopy is employed to study the; conformational stability of a protein under stress induced by varying the pH, temperature and chemical denaturants, folding pattern of a protein and conformational changes upon protein-ligand interaction. The analysis of CD spectra recorded of proteins in the UV region is used to measure conformational changes that occurred in proteins secondary structure. In the case of Far-UV (190-250 nm) where CD spectra for protein is recorded provide an estimation of the secondary structure (alpha, β-sheets, loops) composition of total protein in that particular condition. The CD signal in the Far-UV is contributed by absorption primarily by the peptide bonds. However, in the Far-UV CD (250-300 nm) signal is contributed by the absorption by aromatic amino acids in the range of 390-310 nm wavelength which provides the tertiary structure estimation (Kelly et al 2005). The CD spectra of the enzyme were recorded on a J-715 Spectropolarimeter with a PTC 343 Peltier unit (Jasco, Tokyo, Japan) at 25°C in a quartz cuvette. CD spectra was recorded as a mean of three scan and and expressed in MRE.

All spectra were corrected for buffer contributions and observed values were converted to mean residue ellipticity (MRE) in deg cm^2 $dmol^{-1}$ defined as

$$MRE = M\theta\lambda/10dcr$$

Where, M is the molecular weight of the protein, θλ is CD in millidegree, d is the path length in cm, c is the protein concentration in mg/ ml and r is the average number of

amino acid residues in the protein. The relative content of secondary structure elements was calculated by using CD pro software. Low NRMSD values were observed for analysis with CONTINLL and SELCON.

2.3.14 Steady state kinetic parameters for *Ef*BSH

The kinetic constants $K_{0.5}$ and V_{max} of *Ef*BSH were estimated by assaying the enzyme activity with increasing concentrations of conjugated bile salts (Sigma) as substrates in the range 0.5–60 mM following the standard assay protocol. The assay was carried out in triplicates, and the kinetic constants ($K_{0.5}$ Vmax, hill coefficient 'h') were determined by non-linear regression by fitting the kinetic data in GraphPad Prism version 5.0

$$[Vmax \,.\, Sh] / [K0.5 + Sh + (S2h/Ki)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.1)$$

The effect of the non-substrate ligand Pen V (0.5–200 mM) on the conjugated bile acid (GDCA) activity of *Ef*BSH was also explored. Steady-state kinetics was performed by varying the GDCA concentration in the presence of 1 and 50 mM Pen V.

2.3.15 Effect of additives on *Ef*BSH

The effect of chemical modifiers and metal ions on the enzyme activity was also assessed. *Ef*BSH was dialysed against the buffer to remove any additives already present. The enzyme was then incubated with various modifiers at room temperature (25°C) for 30 min, and the activity was assayed at the aforementioned standard conditions. The reaction carried out without any additives was used as the control. All experiments were performed independently in triplicates.

2.4 Results & Discussion

Bile Salt Hydrolase has been distributed among bacteria (Gram positive, Gram negative), archea and fungal kingdom (Jones, et al., 2008) as mentioned earlier. BSH from both Gram positive and Gram negative have been studied. *Enterococcus faecalis* strain 2403 was obtained from the NCIM (National Collection of Industrial Microorganisms, NCL, Pune). Previously, two BSH enzymes from *Clostridium*

perfringens and *Bifidobacterium longhum* have been characterized biochemically and 3D structure elucidated (Rossocha, et al., 2005, Kumar, et al., 2006). Bile salt hydrolase from *E. faecalis* has shown very high activity towards the (glycol-/ tuaro-conjugated bile acids. This chapter has basically dealt with the cloning, expression, purification, biochemical and biophysical characterization of the *Ef*BSH (*Enterococcus faecalis* bile salt hydrolase).

2.4.1 Sequence analysis and cloning of *Efbsh* gene

The *Efbsh* gene was retrieved from NCBI with a GenBank ID of EET97240.1, which was annotated as cholylglycine hydrolase. Lambert et. al., (2008) and Panigarahi et. al., (2014) have conducted the sequence analysis and annotated this gene in the BSH cluster.

Uniprot id _EFBG_01849:

ATGTGTACAGCAATTACTTATGTATCAAAAGATCATTACTTTGGAAGGAATTT
TGATTATGAAATTTCTTATAATGAGGTGGTTACTATTACGCCGAGAAATTATA
AGTTTTCATTTCGAGAAGTTGGAAATTTAGATCATCATTTTGCAATAATTGGA
ATTGCTGCTGGGATTGCTGATTATCCGCTTTATTATGATGCAATAAATGAAAA
AGGATTAGGAATGGCTGGATTAAACTTTTCAGGCTATGCAGATTATAAAAAA
ATTGAAGAAGGAAAAGAAAATGTTTCTCCATTTGAGTTTATTCCTTGGGTATT
GGGCCAATGCTCTACTGTAGATGAAGCAAAAAAATTATTGAAGAATCTTAAT
TTAGTAAATATTAATTTTAGTGATGAACTTCCGTTATCCCCTCTCCATTGGCTG
TTGGCTGATAAAGAGCAATCTATTGTGGTTGAAAGCACGAAAGAAGGCTTAC
GTGTATTTGATAATCCTGTAGGCGTATTAACAAATAACCCAACATTTGATTAC
CAATTATTTAATTTAAACAATTATCGTGTACTTTCAACTAGAACTCCAAAAAA
TAATTTTTCAGATCAAATAGAGTTAGATATTTATAGTAGAGGAATGGGTGGTA
TTGGGTTGCCAGGAGATTTATCATCAGTATCTAGATTTGTAAAAGCAACTTTT
ACTAAGTTAAATTCTGTATCAAGAAGTTCAGAATATGAAAGTATTAGCCAATT
TTTTCATATTTTAAGTTCTGTCGAACAACAAAAAGGATTGTGTGATGTTGGTG
ATGAAAAATATGAGTATACGATTTATTCTTCATGTTGTAACCTGGAAAAGGGA
ATTTATTACTATCGTACGTATGACAATAGTCAAATTACTGCTGTGGATATGAA
TAAGGAAAATTTAGAGAAGGATAGCTTAATTGTTTATCCAATGGTGGAAACA
CAACAAATTAACTATGCTAATTAA

```
EfBSH  CTAITYVSKD--HYFGRNFDYEISYNEVVTITPRNYKFSFR-EVGNLDHHFAIIGIAAGI 57
LpBSH  CTAITYQSYN--NYFGRNFDYEISYNEMVTITPRKYPLVFR-KVENLDHHYAIIGITADV 57
CpBSH  CTGLALETKDGLHLFGRNMDIEYSFNQSIIFIPRNFKCVNKSNKKELTTKYAVLGMGTIF 60
BlBSH  CTGVRFSDDEGNTYFGRNLDWSFSYGETILVTPRGYHYDTVFGAGGKAKPNAVIGVGVVM 60
BtBSH  CTRAVYLGPDRMVVTGRTMDWKEDIMSNIYVFPRGMQRAGH-NKEKTVNWTSKYGSVIAT 59
       **       :     **.:* . .  . : . **                :  *
EfBSH  ADYPLYYDAINEKGLGMAGLNFSGYADYKKIE-EGKENVSPFEFIPWVLGQCSTVDEAKK 116
LpBSH  ESYPLYYDAMNEKGLCIAGLNFAGYADYKKYD-ADKVNITPFELIPWLLGQFSSVREVKK 116
CpBSH  DDYPTFADGMNEKGLGCAGLNFPVYVSYSKEDIEGKTNIPVYNFLLWVLANFSSVEEVKE 120
BlBSH  ADRPMYFDCANEHGLAIAGLNFPGYASFVHEPVEGTENVATFEFPLWVARNFDSVDEVEE 120
BtBSH  GYDIGTCDGMNEKGLVASLLFLP--ESVYSLPGDTRPAMGISIWTQYVLDNFATVREAVD 117
                *  **:**  : * :.   .          :       ::  :  :* *. .
EfBSH  LLKNLN--LVNINFSDELPLSPLHWLLADKE-QSIVVESTKEGLRVFDN-PVGVLTNNPT 172
LpBSH  NIQKLN--LVNINFSEQLPLSPLHWLVADKQ-ESIVIESVKEGLKIYDN-PVGVLTNNPN 172
CpBSH  ALKNAN--IVDIPISENIPNTTLHWMISDITGKSIVVEQTKEKLNVFDN-NIGVLTNSPT 177
BlBSH  TLRNVT--LVSQIVPGQQ-ESLLHWFIGDGK-RSIVVEQMADGMHVHHD-DVDVLTNQPT 175
BtBSH  EMKKETFRIDAPRMPNGGPESTLHMAITDETGNTAVIEYLDGKLSIHEGKEYQVMTNSPR 177
        ::: .  :    ..     : **  : *   .: *:*     : :...    *:**.*
EfBSH  FDYQLFNLNNYRVLS-TRTPKNNFSDQIELDIYSRGMGGIGLPGDLSSVSRFVKATFTKL 231
LpBSH  FDYQLFNLNNYRALS-NSTPQNSFSEKVDLDSYSRGMGGLGLPGDLSSMSRFVRAAFTKL 231
CpBSH  FDWHVANLNQYVGLRYNQVPEFKLGDQ-SLTALGQGTGLVGLPGDFTPASRFIRVAFLRD 236
BlBSH  FDFHMENLRNYMCVS-NEMAEPTSWGKASLTAWGAGVGMHGIPGDVSSPSRFVRVAYTNA 234
BtBSH  YELQLAVNDYWKEVG----------------------GLQMLPGTNRSSDRFVRASFYIH 215
       :: ::     :  :                        *   :**   . .**::.::
EfBSH  NSVSRSSEYESISQFFHILSSVEQQKGLCDVGDEKYEYTIYSSCCNLEKGIYYYRTYDNS 291
LpBSH  NSLPMQTESGSVSQFFHILGSVEQQKGLCEVTDGKYEYTIYSSCCDMDKGVYYYRTYDNS 291
CpBSH  AMIKNDKDSIDLIEFFHILNNVAMVRGSTRTVEEKSDLTQYTSCMCLEKGIYYYNTYENN 296
BlBSH  HYPQQNDEAANVSRLFHTLGSVQMVDGMAKMGDGQFERTLFTSGYSSKTNTYYMNTYDDP 294
BtBSH  AIPQTADAKIAVPSVLSVMRNVSVPFGINTPEKPHISSTRWRSVSDQKNKVYYFESTLTP 275
                 :  .:  : .*    *     . : . * : *    ..  ** .:
EfBSH  QITAVDMNKENLEKDSLIVYPMVETQQINYAN---------- 323
LpBSH  QINSVSLNHEHLDTTELISYPLRSEAQYYAVN---------- 323
CpBSH  QINAIDMNKENLDGNEIKTYKYNKTLSINHVN---------- 328
BlBSH  AIRSYAMADYDMDSSELISVAR-------------------- 316
BtBSH  NLFWLDLKKIDFSPKAGVKKLSLTKGEIYAGDAVKDLKDSQS 317
           :    : . .:.
```

Figure 2.4 Multiple sequence alignments of amino acid sequence of BSH from different bacteria. The amino acid residues conserved at the active site of BSH from diverse organisms are highlighted. Identical amino acids are marked with asterisk, conserved and semi-conserved amino acids are marked by two dots and single dots, respectively. Active site residues (Cys2, Arg16, Asp19, Asn79, Asn170 and Arg223) are highlighted in green colour. Sequences include those of *Enterococcus faecalis (Ef*BSH*), lactobacillus plantarum (Lp*BSH*), Clostridium perfringenes (Cp*BSH*), Bifidobacterium longum (Bl*BSH*), Bacteroid thetaiotaomicron (Bt*BSH*).*

The *Ef*BSH sequence contained 324 amino acid residues, and theoretical p*I* was calculated to be 4.9 (ExPASy ProtParam Server). The translated amino acid sequence of *Ef*BSH when compared with other characterised BSH sequences exhibited 69%, 44%, 37% and 19% identity with *Lactobacillus plantarum*, *C. perfringens*, *Bifidobacterium longum* and *Bacteroides thetaiotaomicron*, respectively. Active-site residues such as Cys2, Arg16, Asp19, Asn79, Asn170 and Arg223 are conserved (numbering is according to *Ef*BSH sequence) in all cholylglycine hydrolases, which are highlighted in the multiple sequence alignment and are important for the catalytic activity (Fig.2.4).

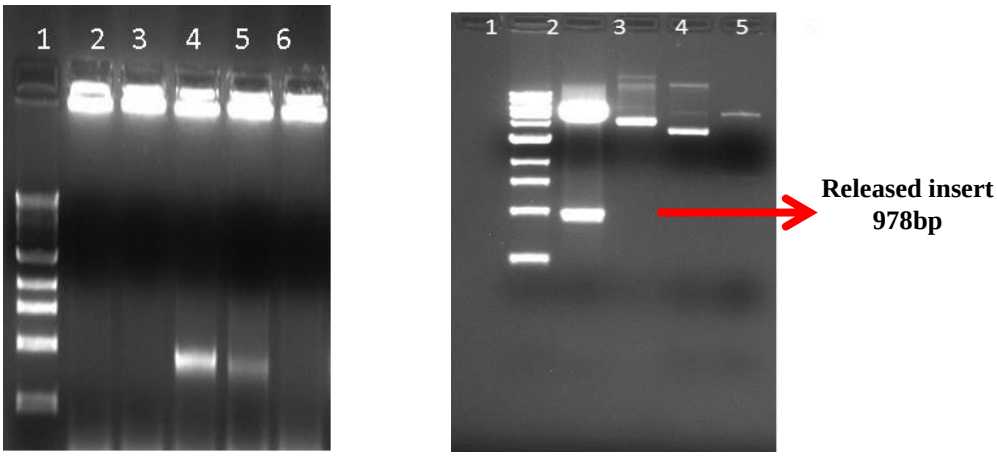


Figure. 2.5 (a). Colony PCR of positive clones. Lane 1-NEB 1Kb DNA ladder, Amplification of the positive clones by using gene specific primers in lane 4 & 5. (b). Lane 1-NEB 1Kb DNA ladder, lane 2-double digestion (*NdeI/XhoI*) of cloned plasmid, Lane 3-Undigested cloned Plasmid (*Ef*BSH/pET22b+), Lane 4-Undigested pET22b+, Lane 5-Digested pET22b+

The gene-specific primers were used for cloning the *Efbsh* gene and cloned in the pET22b+ vector (Fig. 2.1) described in the methodology section 2.3.2. Positive clones were subjected to colony PCR and double digestion to confirm the *Efbsh* gene insertion in the expression vector pET22b+. Colony PCR showed 1kb amplified PCR product corresponding to the *Efbsh* gene (Fig. 2.5a) and restriction digestion (*NdeI/XhoI*) has also shown the release of 1kb insert from positive clone pET22b+/*Efbsh* (Fig.2.5b). Positive clones were transformed in the maintenance host for propagation and plasmid was

isolated for sequencing. Sequencing results also showed that full length *bsh* gene has been cloned in correct frame having no mutation.

2.4.2 Preparation of cell extract, enzyme purification and molecular weight determination

The gene for *Ef*BSH was cloned in the pET22b+ expression vector, and the protein was expressed in *E. coli* BL21* (DE3) cells with a C-terminal 6X His tag. The protein expression was optimised and overexpressed by inducing culture with 0.5 mM IPTG and incubation at 16°C overnight to obtain the cytoplasmic soluble fraction.

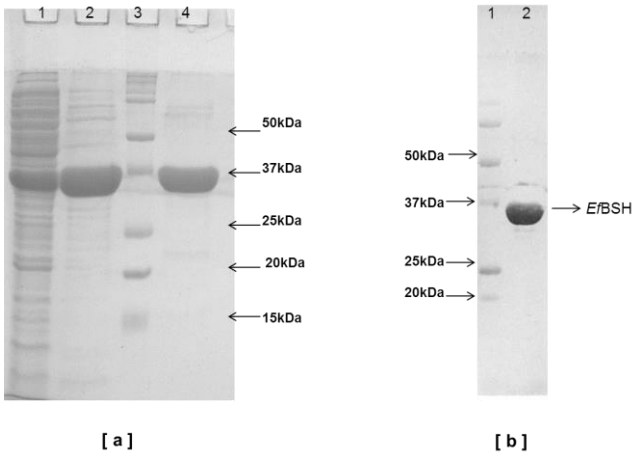
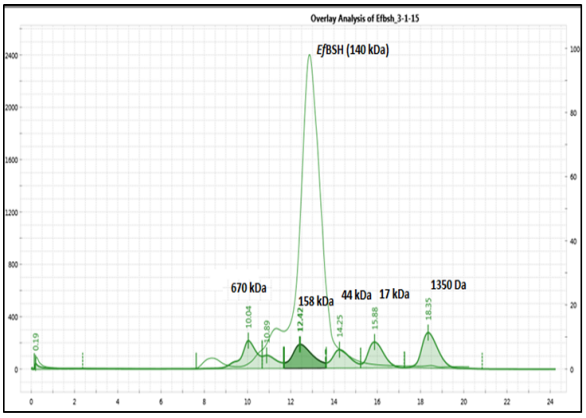
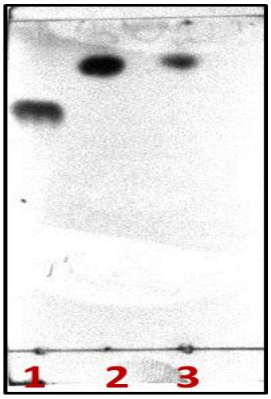


Figure 2.6 [a]. SDS-PAGE analysis of overexpressed recombinant *Ef*BSH protein at each purification step from *E.coli* BL21* (DE3). Lane 1: Supernatant of *Ef*BSH, Lane 2: Protein purified from Ni–NTA affinity chromatography, Lane 3: Precession Plus protein ™ dual colour marker, Lane 4: Purified protein fraction form size exclusion chromatography (Bio-rad SEC650 column). The ~37kDa molecular weight corresponds to monomer of *Ef*BSH. [b]. Western blot analysis of purified *Ef*BSH protein with anti-His antibodies; Lane 1: Precession Plus protein ™ dual colour marker, Lane2: *Ef*BSH protein.

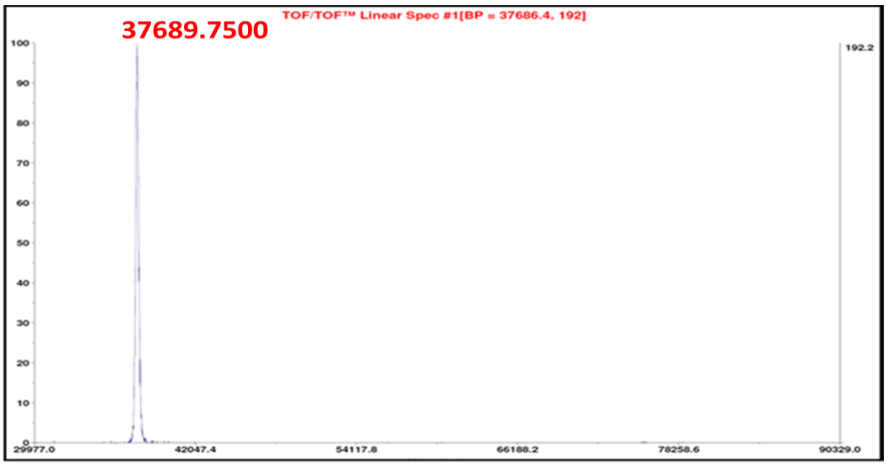
The overexpressed protein was purified to homogeneity using a Ni–NTA affinity column followed by gel filtration (column: SEC650-Bio-rad) to remove any contaminant protein and soluble aggregates present.

[a]





[b]



[c]

Figure 2.7 (a). Gel-filtration chromatogram to show that *Ef*BSH is a tetramer. Standards were run on the column and 158 kDa marker protein eluted at 12.42 ml volume. The *Ef*BSH homo-tetramer (~140kDa) was eluted at 12.88ml volume. (b). Qualitatively measuring the activity of the *Ef*BSH by detecting reaction products using TLC method, Lane 1. Glycodeoxycholoc acid (GDCA), Lane 2. Deoxycholic acid (DCA), Lane 3. *Ef*BSH hydrolysed substrate (GDCA). (b). Molecular mass determination of *Ef*BSH by MALDI-TOF-MS analysis. The image shows monomer of the ~37kDa protein.

The purity of the protein was assessed by 12% SDS-PAGE analysis where purified single bands corresponding to the ~37-kDa protein were observed (Fig.2.6a). The identity of the purified protein was confirmed by Western blot analysis (Fig. 2.6b). Protein yields were 80–100 mg/l of culture. Molecular weight of protein was measured by size exclusion chromatography and MALDI-TOF-MS. The peaks were obtained with the observed mass value of ~148 kDa (Fig.2.7a) (native, SEC650 chromatography peak) and 37 kDa from size exclusion chromatography and MALDI-TOF-MS, respectively

(Fig.2.7b). The gel filtration (SEC-650 column, Bio-rad) analysis of the purified protein confirmed the tetrameric form of *Ef*BSH (Fig. 2.7a). Qualitative analysis of *Ef*BSH activity was performed by identifying the reaction products using the TLC method (Fig. 2.7b).

2.4.3 Optimum pH and temperature

Most of the BSH enzymes from various Gram-positive bacterial sources show optimum pH in the acidic range of pH 4.0-6.5 (Rossocha, et al., 2005, Kumar, et al., 2006). However, Gram-negative bacteria *Xanthomonas maltophilia* shows optimum pH in the alkaline range 7.9-8.5 (Dean, et al., 2002). The purified enzyme *Ef*BSH was assayed for BSH activity with conjugated bile acids, and the reaction was carried out at pH 5.0 and temperature 50°C. One of the resultant by-products of the reaction, glycine or taurine, was estimated by ninhydrin assay by recording the colour developed at 570 nm. The *Ef*BSH enzyme displayed optimal activity at pH 5.0 (0.1 M citrate phosphate buffer) in the assay condition (Fig.2.8a). The optimum temperature for enzyme activity was found to be 50°C (Fig. 2.8b).

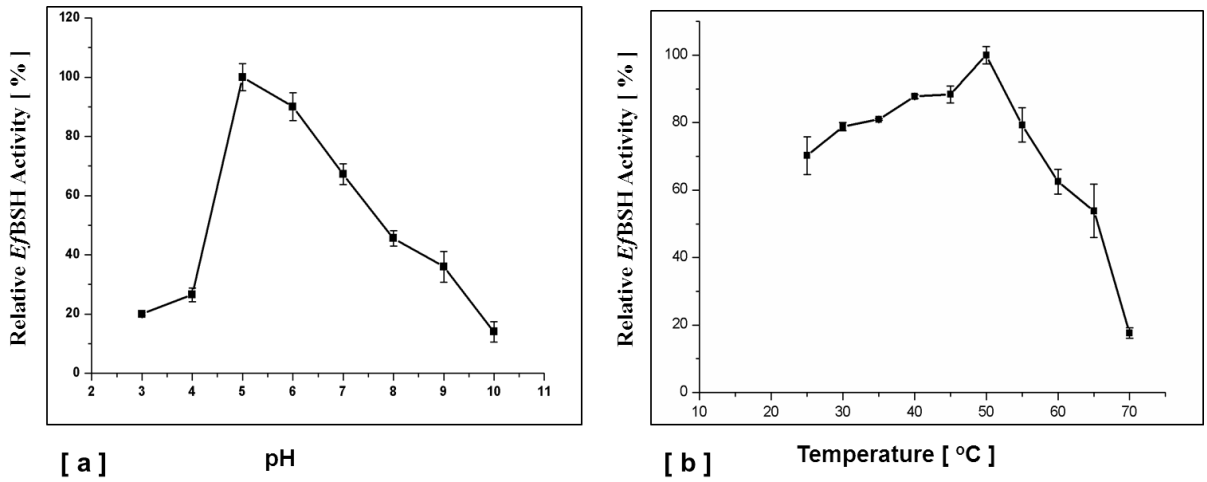


Figure 2.8 [a]. Optimum pH and [b] Optimum temperature determined for the recombinant *Ef*BSH.

2.4.4 Effect of pH on *Ef*BSH stability

The stability of the *Ef*BSH was assessed by incubating the enzyme in various pH buffers (1-11). The relative activity of the *Ef*BSH was measured and it was found that the enzyme is ~80 % stable in the pH range of 5.0-7.0. However, in the pH 8.0-9.0 the

enzyme could recover only ~50% residual activity (Fig. 2.9a). *Ef*BSH drastically lost activity in extreme protonated (pH-1.0-3.0) and deprotonated (pH-11.0) buffer (Fig. 2.9a). The intrinsic fluorescence emission maximum (λ_{max}) of the protein was recorded at 328 nm indicating the presence of tryptophan in a hydrophobic microenvironment (Fig. 2.9b).

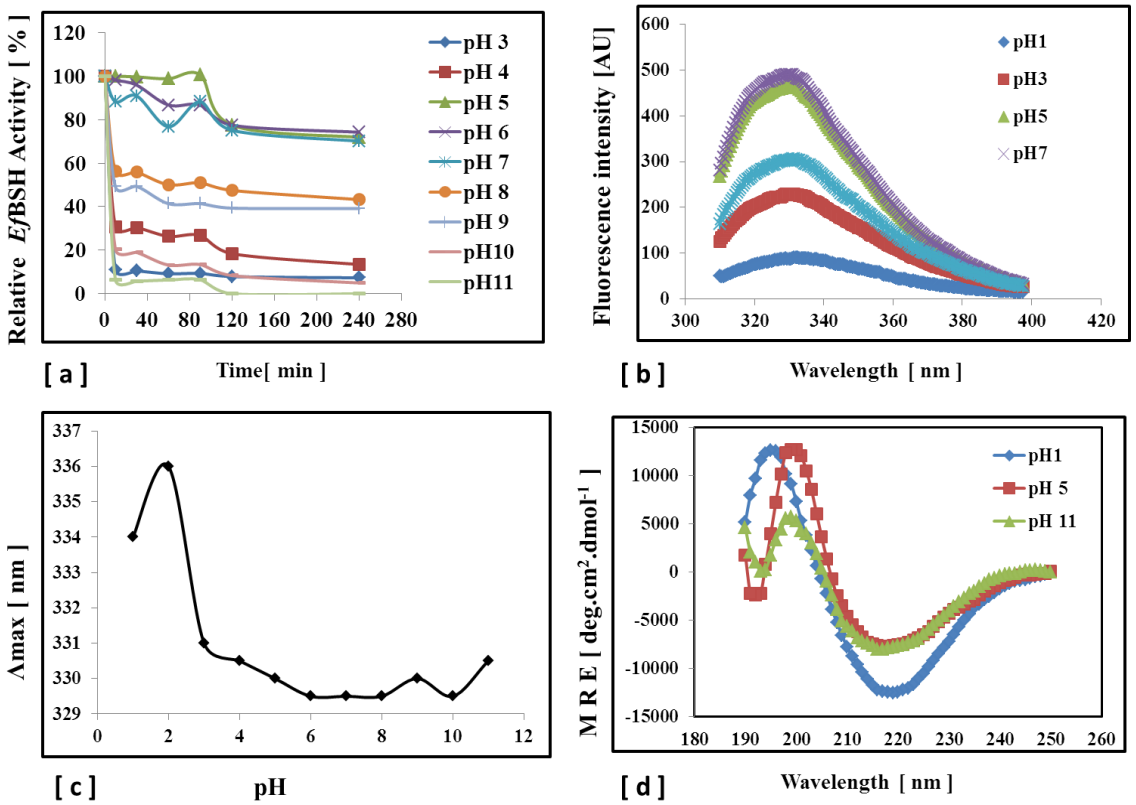


Figure 2.9 [a]. Relative *Ef*BSH activtiy (%) on incubating the protein in 0.1M buffer for 4 hr at respective pH [b] Fluorescence intensity measurement of the EfBSH at different pH (1-11) after 4 hr incubation of the protein in respective buffers of 0.1M. [c] Red shift (λmax shift) in protonated and deprotonated pH. [d]. Far-UV (MRE) CD spectra of *Ef*BSH protein at various pH.

Under highly protonated and deprotonated condition *Ef*BSH showed a decrease in the intrinsic fluorescence intensities, which further decreased in the protonated condition (Fig.2.9b). In the extreme acidic pH enzyme showed shift in λ_{max} to 336 nm which showed opening of the secondary structure (Fig. 2.9c). Far-UV CD spectrum showed a significant change in MRE at incubation at pH 1.0, indicating structural changes such as an increase in α-helix and a decrease in β-sheet fractions at alkaline pH 11.0 but the

reverse at pH1.0 (Fig. 2.9d.). The change in the secondary structure composition (helix, β-sheet, loop) of *Ef*BSH protein on treatment with various conditions, pH, temperature and chemical denaturant GdnHCL was estimated. All calculations were done using CDSSTR and Continll of the CDPro suite (Table 2.1). Upon incubating the protein in various conditions the disordered structure increased (table 2.1).

2.4.5 Thermal stability of *Ef*BSH

Thermal stability studies of *Ef*BSH were performed. The enzyme began to aggregate, and the activity reduced drastically at 60°C. The enzyme was fairly stable in the temperature range of 25-55°C and retained 60-80% of its activity. At 30°C the enzyme retained upto ~90% activity (Fig. 2.10a). However, the enzyme activity sharply reduced after 3 h of incubation at 50°C (65% activity), and the enzyme aggregated beyond 60°C and significantly lost its activity (Fig.2.10a). Most BSH enzymes are known to be stable upto 45°C (Rossocha, et al.,2005, Kumar, et al., 2006, Bi, et al., 2013), while dimeric BSH from Gram –ve bacteria *Xanthomonas maltophilia* has been reported thermally stable upto 70°C (Dean, et al.,2002). Far-UV CD spectra for *Ef*BSH enzyme were recorded and minima observed at 218 nm showed that the protein was rich in β-sheet structure (Fig. 2.10b). The enzyme was incubated at 100 ug/ml concentration at various temperatures (30-70°C) for 15 min and CD spectra recorded. The loss of secondary structure beyond 60°C (Fig.2.10c) correlated with the residual activity and increased disorder (Table 2.1). Near-UV CD spectra (MRE) of *Ef*BSH were also recorded to see the effect on tertiary structure. It was observed that at 60°C and beyond the tertiary structure collapsed. Change in Near-UV CD spectra was observed at 290-310 nm where aromatic residues absorbed (Fig.2.10d).

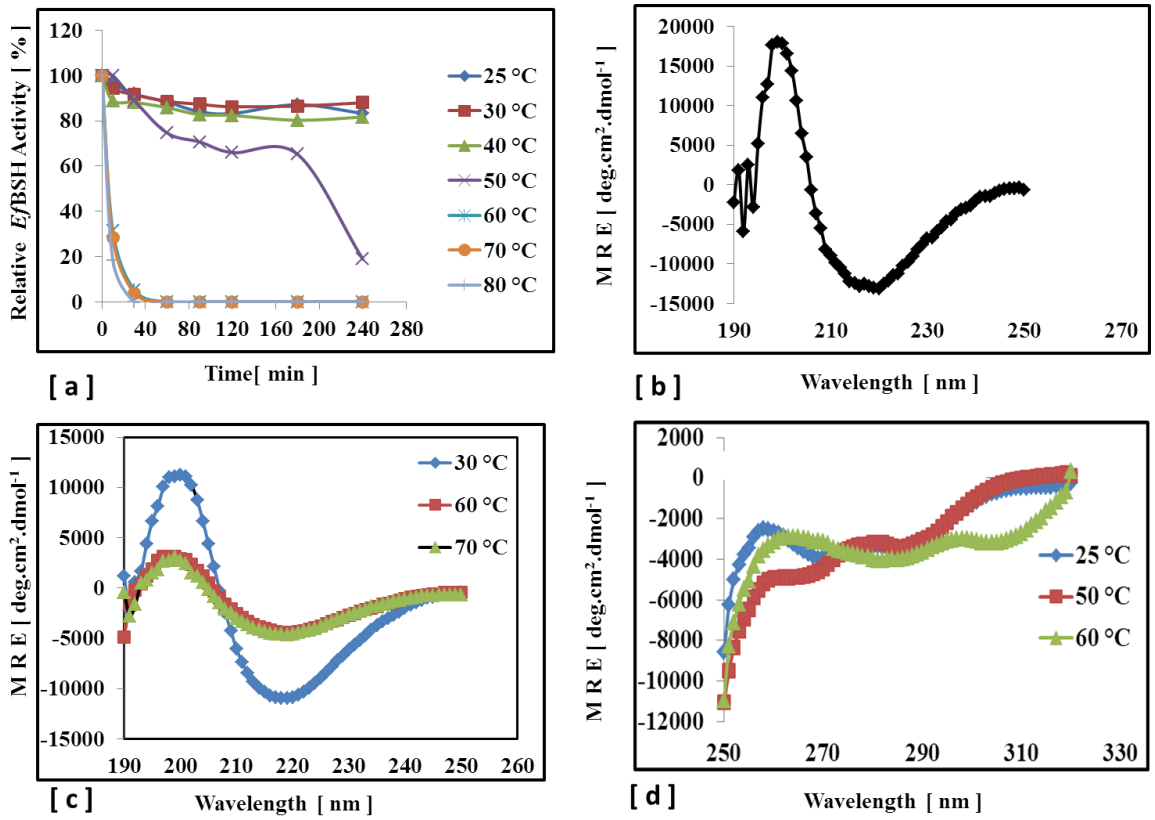


Fig.2.10 [a] Residual activtiy (%) of *Ef*BSH on incubating the protein in 50mM citrate phosphate buffer (CPB pH 5.0) at various temperatures (30-80°C) for different time intervals (min). [b] Far-UV (MRE) CD spectra of native-*Ef*BSH protein. [c] Far-UV(MRE) CD spectra of *Ef*BSH protein after incubating the protein at 30-70°C for 15 minutes. [d] Near-UV spectra (MRE) *Ef*BSH after incubation of 15 min at 25-70°C.

2.4.6 Effect of GdnHCl on *Ef*BSH stability

Chemical Denaturant Gdn-HCl at low concentrations is known to confer effective charge shielding (Monera, et al., 1994) and prevent aggregation in proteins (Hevehan, et al., 1997). *Ef*BSH showed decrease in fluorescence intensity with increase in the GdnHCl concentration above 1.5M, and shown drastic change in fluorescence beyond 2M GdnHCl concentration (Fig. 2.11a). Enzyme showed more fluorescence intensity (AU) in the presence of 0.25-0.5M GdnHCl than without the denaturant. Red shift in λmax (328 to 355 nm) was observed on increasing the concentration of GdnHCl showing opening up of secondary structure (Fig. 2.11b).

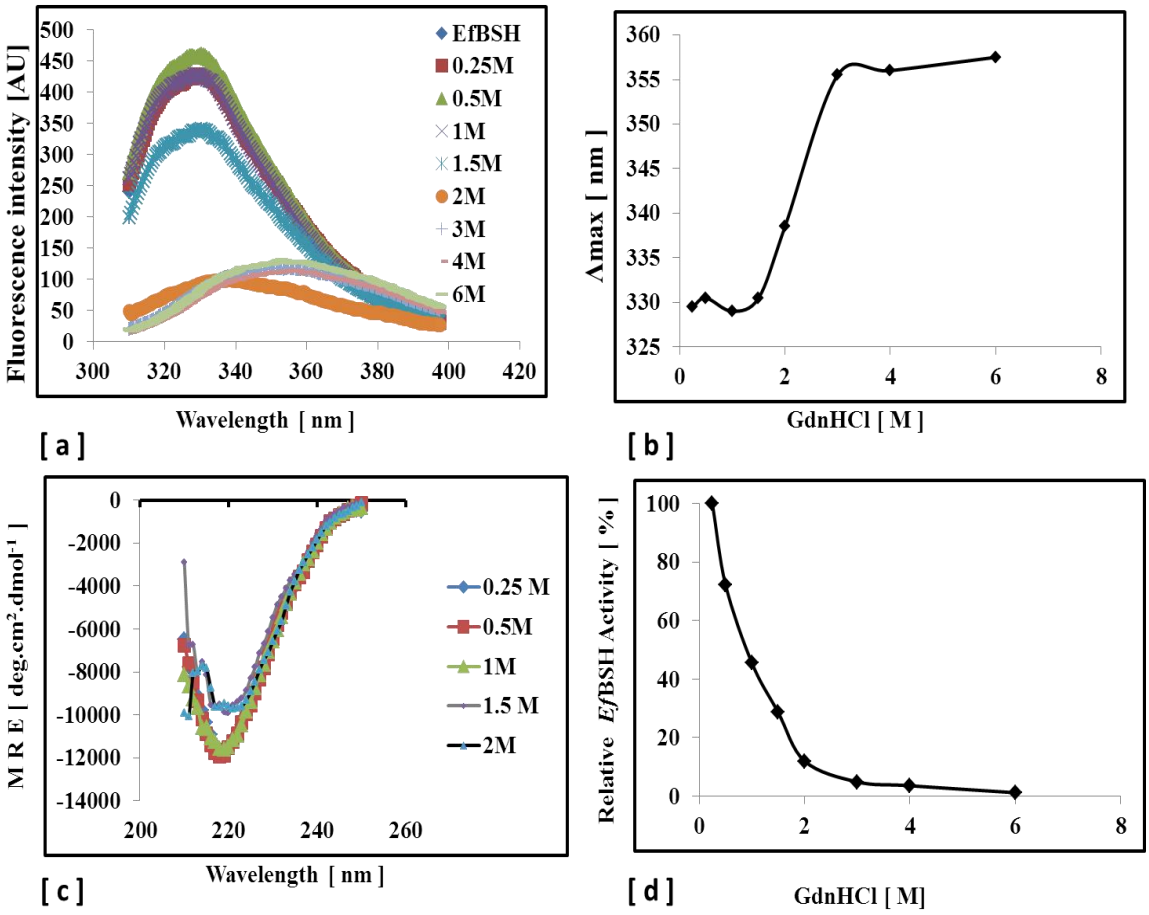


Fig.2 11 [a] Fluorescence spectra of the *Ef*BSH at different GdnHCl concentration (M) upon incubation of the protein for 4 hr in respective concentration. [b] Shift in λmax with increasing concentration of the GdnHCl. [c] Far-UV CD spectra of *Ef*BSH upon incubating with increasing conctration of GdnHCl (0-2M) for 4 hr. [d] Residaul activity of *Ef*BSH after incubating the eznyme for 4 hr in various GdnHCl concentrations (0-6.0M).

Minima for *Ef*BSH was observed to be around 218 nm. *Ef*BSH was incubated for 4 hr with given concentration (0.25 - 2 M) of GdnHCl and it was observed that after 1.5M concentration secondary structure was collapsing (Fig. 2.11c) (Table 2.1). This was also well correlated with the estimated residual activity of the *Ef*BSH with the GdnHCl (0.25-6M) concentration on incubating for 4 hr (Fig. 2.11d).

*Bl*BSH is a tetrameric protein composed of the dimer of dimer (Kumar, et al., 2006). *Ef*BSH enzyme when treated with 2M GdnHCl for 4 hr and on performing size exclusion chromatography, the absence of a dimer peak in the chromatogram was an

indication of the strong tetrameric association. However, *Bl*BSH tetrameric assembly have been reported to dissociate into dimers upon incubating the enzyme in 1.5M GdnHCl (Kumar, et al., 2006).

Table. 2.1 Change in secondary structure composition of *Ef*BSH protein with various denaturing conditions. All calculations were done using CDSSTR and Continll of the CDPro suite.

*Ef*BSH protein		% Helix	% Sheets	% Turn	% Unordered	NRMSD
1	pH					
	1	16.5	3	20.6	27.9	0.039
	5	6	38.2	19.5	36	0.069
	11	5	39.5	20.8	33.9	0.085
2	Temperature (^{o}C)					
	30	9	3	19.9	33.2	0.08
	60	4.84	34.1	24.3	36.8	0.08
	70	4	36.2	24.3	34.8	0.10
3	GdnHCl (M)					
	0.25	0	48.1	22.6	27.8	0.06
	0.5	0	37.1	22.6	33.6	0.10
	1	0	43.3	24.1	29.2	0.14
	1.5	2	46.2	18.2	30.7	0.11
	2	3	44.4	24.1	26.9	0.01

2.4.7 Effect of additives

The effect of divalent metal ions, chelating agent (ethylenediaminetetraacetic acid (EDTA) and detergent (Triton X-100) on the activity of *Ef*BSH are shown in Table 2.2 The *Ef*BSH activity was moderately enhanced by Mg^{2+}, Ca^{2+}, Co^{2+}, Ni^{2+} and Mn^{2+} metal ions. In the presence of Cu^{2+} ion ~28% inhibition occurred. EDTA had no effect on the enzyme activity. However, Triton X had marginally enhanced activity, which could be due to the dissociation into active form of soluble aggregated enzyme molecules.

Table 2.2 Effect of additives on *Ef*BSH activity.

S.No.	Additives	Relative activity (%)
1.	$MgSO_4$	105 ± 2
2.	$CaCl_2$	106 ± 7
3.	$HgCl_2$	0
4.	$AgNO_3$	0
5.	$CuSO_4$	28 ± 7
6.	$NiSO_4$	110 ± 3
7.	EDTA	102 ± 4
8.	$MnCl_2$	106 ± 1
9.	$CsCl_2$	108 ± 2
10.	$COCl_2$	109 ± 8
11.	Triton X	118 ± 7
12.	**SDS**	0

2.4.8 Steady-state kinetics of *Ef*BSH

The standard ninhydrin assay with some modification was used for the estimation of the hydrolyzed by-product, that is, glycine or taurine amino acid. BSH enzymes reported so far followed the Michaelis–Menten (MM) kinetics (Dean, et al., 2002, Rossecha et al., 2005, Kumar et al., 2006, Bi et al., 2013), while *Ef*BSH showed very distinct kinetic properties in comparison with other BSHs. Positive cooperativity and enhancement was noted for *Ef*BSH activity towards bile acids by the non-substrate ligand Pen V. The values of $K_{0.5}$, h and turnover number (k_{cat}) for different substrates GCA, GDCA, TCA, GCDCA and TCDCA were determined by incubating the enzyme sample with a range of concentrations of each substrate (0.5–60 mM) under the standard assay conditions. The concentration of the enzyme used was 1.2µg/reaction for all kinetic studies. Hill's equation (Hill, 1910) was used for computing the parameters $K_{0.5}$, h and V_{max}. *Ef*BSH showed Hill's coefficient <1, indicating apparent positive cooperativity for all the six bile salts studied (Fig.2.15).

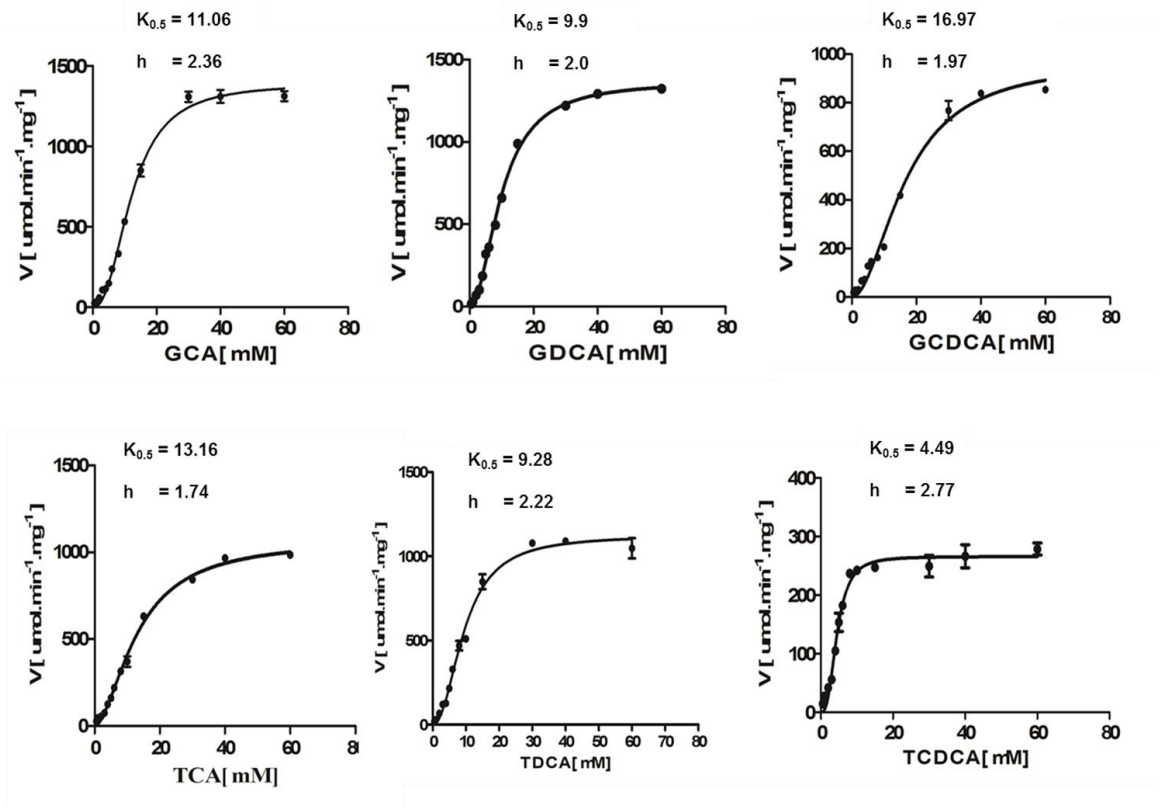


Figure 2.15 Steady-state kinetic graphs of recombinant *Ef*BSH with all six conjugated bile salts: [a] GCA, [b] GDCA, [c] GCDCA, [d] TCA, [e] TDCA and [f] TCDCA.

Cooperative behaviour has also been reported in *Lactobacillus salivarius* BSH (*Ls*BSH), when dithiothreitol (DTT) was removed by dialysis from the reaction. However, the enzyme showed MM kinetics in the presence of DTT (Bi, et al., 2013). Avinash, et al., 2015 have also observed the cooperative behaviour in one of the members of the cholylglycine hydrolase subfamily PVA from Gram-negative bacteria *Pectobacterium atrosepticum.* All the kinetic studies with *Ef*BSH were conducted in the presence of 1 mM DTT in the buffer. The presence of DTT did not modulate the allosteric behaviour of the enzyme. Removal of DTT (by PD10 column) from the enzyme solution showed a small decrease in activity presumably due to the oxidation of the thiol group of the active site cysteine (Rossocha, et al., 2005). To date, among the BSH enzymes, the k_{cat} value 3.013×10^3 S^{-1} and the catalytic efficiency ($k_{cat}/K_{0.5}$) 3.029×10^5 M^{-1} S^{-1} for GDCA determined for *Ef*BSH remain the highest (Gopal et al., 1988, Christiaens, et al., 1992, Tanaka, et al., 2000, Corzo et al., 1999, Pereira, et al., 2003,

Rossocha, et al., 2005, Kumar, et al., 2006). The kinetic parameters of *Ef*BSH calculated for all bile acid substrates are summarised in Table 2.3.

Table 2.3: steady state kinetic parameters:

Substrates	$K_{0.5}$	h	V_{max}	K_{cat} (S^{-1})	$K_{cat}/K_{0.5}$ (M^{-1} S^{-1})
GCA	11.065 ± 0.69	2.363 ± 0.14	1390 ± 56.56	3250.094 ± 132.26	2.937×10^5
GDCA	9.94 ± 0.004	1.99 ± 0.008	1289 ± 6.36	3.014×10^3	3.029×10^5
GCDCA	16.97 ± 0.736	1.929 ±0.112	970 ± 13.93	2.268×10^3	1.336×10^5
TCA	13.16 ± 0.915	1.745 ±0.088	1072 ± 23.33	2.506×10^3	1.903×10^5
TDCA	9.28 ± 0.405	2.22 ± 0.173	1115 ± 24.74	2.607×10^3	2.808×10^5
TCDCA	4.49 ± 0.177	2.77 ± 0.201	265.8 ±12.58	6.21×10^2	1.382×10^5

2.4.9 Pen V induced modulation of *Ef*BSH activity:

The activity-modulating effect of the non-substrate ligand Pen V (0.5–200 mM) for *Ef*BSH activity towards GDCA was studied.

The concentration of the bile acid (GDCA) was kept constant, while that of Pen V was increased up to 200 mM. It is worth noting that *Ef*BSH showed enhanced activity towards GDCA in the presence of Pen V and no inhibition was observed, contrary to anticipated results (Table 2.4). It may be noted that Pen V showed competitive inhibition in the case of *Bl*BSH (Kumar, et el., 2006).The steady-state kinetics was performed with varying concentrations of GDCA as the substrate (0.5–60 mM) in the presence of 1 mM and 50 mM Pen V concentrations (Fig. 2.16). The kinetic constants calculated indicated enhanced catalytic efficiency (k_{cat} / $k_{0.5}$) with values of 3.08×10^5 and 3.52×10^5 M^{-1} S^{-1},

respectively (Table 2.5), when compared to the catalytic efficiency towards GDCA in the absence of Pen V (Table 2.3).

Table 2.4 Enhanced catalytic activity of *Ef*BSH toward bile acids (GDCA) on addition of Pen V. GDCA concentration was kept constant (5 mM) in all reactions and Pen V concentration was varied in the range of 0.5–200 mM. *Ef*BSH activity without Pen V is considered as 100% activity.

S. No.	Pen V concentration	Relative activity (%) of *Ef*BSH
1	0.5	**104.9 ± 4.9**
2	1	**123.2 ± 4.2**
3	5	**127.6 ± 2.8**
4	10	**129.9 ± 3.5**
5	15	**137.8 ± 3.3**
6	30	**153.8 ± 2.5**
7	40	**152.2 ± 3.5**
8	50	**154.5 ± 1.4**
9	60	**145.9 ± 0.7**
10	80	**141.9 ± 1.4**
11	100	**139.7 ± 0.7**
12	**200**	**133.5 ± 2.1**

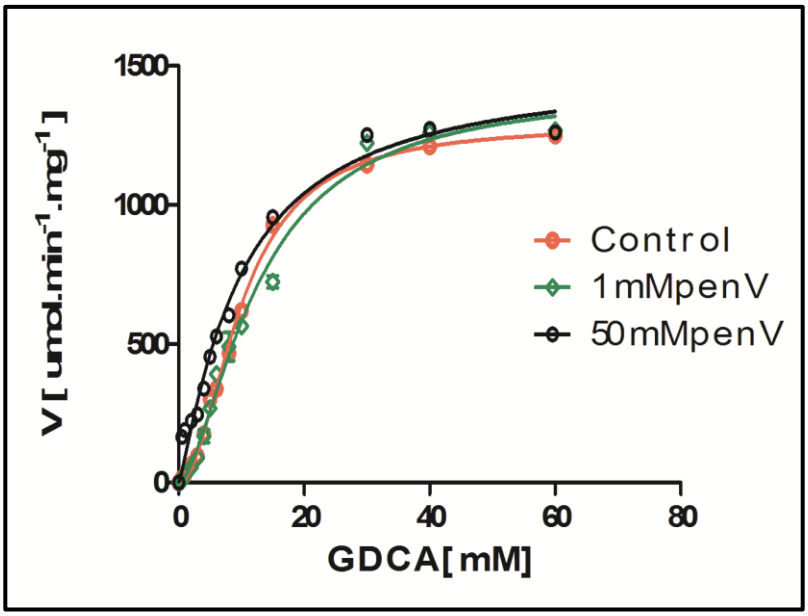


Figure 2.16 Steady-state kinetics of recombinant *Ef*BSH with GDCA in the presence of 1 and 50 mM penV. Kinetics with GDCA substrate without penV in reaction is shown as control (red), with 1 mM penV (green) and with 50 mM penV (black).

Table.2.5 Steady-state kinetic constants of *Ef*BSH activity calculated with varying GDCA concentrations in the presence of 1 and 50 mM Pen V.

S.No	GDCA kinetic + PenV	$K_{0.5}$ (mM)	*h*	V_{max}	K_{cat} (S^{-1})	$K_{cat}/k_{0.5}$ (M^{-1} S^{1})
1	GDCA kinetics with 1mM Pen V	10.8 ± 1.1	1.70 ± 0.06	1420 ±8	3320 ± 18	3.08 x 10^5
2	GDCA kinetics with 50mM Pen V	9.9 ±0.5	1.18 ±0.01	1493 ±13	3491 ± 31	3.52 x 10^5

*Ef*BSH activity was compared with the reported BSH activity from various sources and was found to be of highest catalytic efficiency in comparison to reported BSH till now.

Table 2.6 Comparison of *Ef*BSH activity with other reported BSH.

S.No.	Specific Activity (U/mg)[a]	Km (mM)	Kcat (S^{-1})	Substrate	Reference No.
1	0.107	0.5	-	GCA	Gopal, 1988
2	3.4	-	-	GCA	Corzo, et al.,1999
3	-	0.022	85	GCA	Kumar, et al., 2005
4	165.2	18.2	99	TDCA	Bi, et al., 2013
5	11.84	-	-	GDCA	Tanaka, et al., 2000
6	**1390**	**11.06**	**3250**	**GCA**	This Study

a. activity in umol/min/mg

2.5 Summary of the Chapter:

In summary, the present study has focussed on cloning, expression, purification and biophysical and detailed biochemical characterization of a BSH from *E. faecalis*.

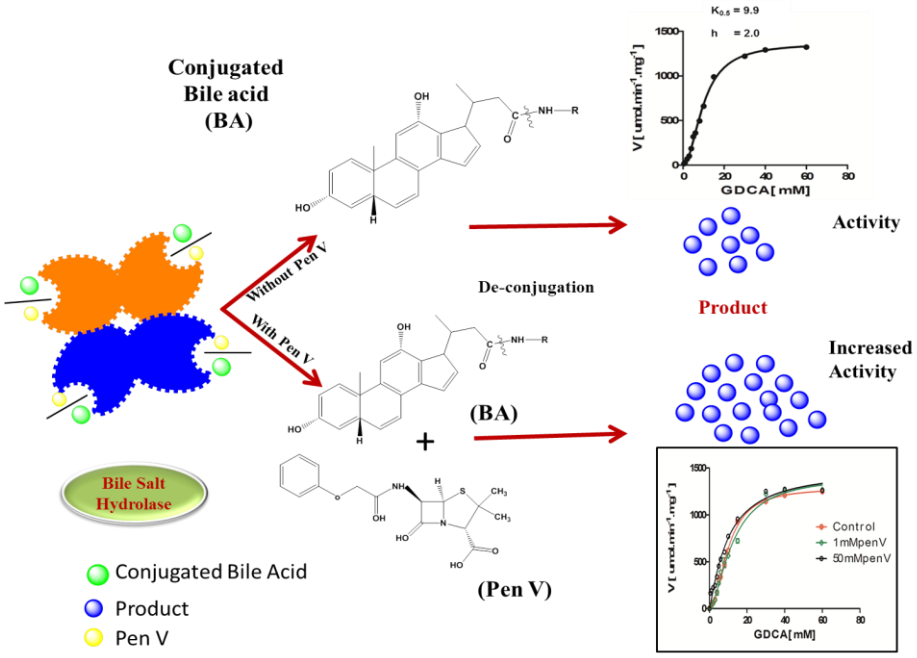


Figure 2.17 Schematic representation of the cooperativity in *Ef*BSH activity and Pen V modulation.

To the best of our knowledge, the catalytic ability of *Ef*BSH was found to be the highest for any BSH reported till date. The enzyme demonstrated unique kinetics and displayed apparent positive cooperativity. *Ef*BSH displayed enhanced activity in the presence of the non-substrate ligand Pen V instead of anticipated inhibition based on previous studies on other BSHs. Therefore, this enzyme is distinct and different from the previously reported BSHs. The present findings add further insight into the biochemical features and unique kinetic properties of *Ef*BSH. The overview of the chapter is summarized in the schematic representation shown above in Fig. 2.17.

Chapter-3

"Crystal structure determination and structural analysis of *Ef*BSH"

Macromolecular crystallography is one of major technique which has been exploited extensively to reveal the protein structure and function. Protein crystallography can help in understanding the structure-function relationship of proteins and supra-molecular assemblies like ribosome etc. Drug design is strongly reliant on structural information. Knowing the precise 3D-structure of protein by X-ray crystallography helps in identifying small molecules capable of blocking the active site of proteins (enzymes) that are involved in myriads of disease pathology and also in protein engineering to enhance the potential industrial and pharmaceutical application. Therefore, three-dimensional structure determination has a great impact in the biological sciences research in order to get insight into the molecular mechanism and associated functioning of molecules.

3.1 Introduction

Current studies have directed that gut microbiota is associated with host energy regulation and progression of obesity and other disorders in humans. Therefore, employing specific gut microbial functions may be one way to control obesity and associated chronic disorders (Joyce, et al., 2014, DiBaise et al., 2008; Tilg, et al., 2009, Kumar, et al., 2012). Bile salt hydrolase (BSH, Cholylglycine Hydrolase, E.C.3.5.1.24) activity is considered one of the important characteristics for selecting a strain to use as probiotic. Knowing the role of BSH in the host metabolism is noteworthy for selecting organisms for probiotic application. Joyce, et al., 2014 have shown the substantial evidence demonstrating that BSH activity alone can influence the host lipid metabolism significantly. BSHs have majorly studied in probiotics perspective, where it is known to play role in cholesterol lowering (Tanaka, et al., 2000, Moser et al., 2000, Begely, et al., 2005, Begely, et al., 2006, Choi, et al., 2015) by deconjugation of gut secreted bile acids. BSH is pharmaceutically an important enzyme which belongs to the Cysteine-Ntn hydrolase family of the Ntn-hydrolase superfamily. It is majorly distributed among the microbes of the gastrointestinal (GI) tract in mammals (Jones, et al., 2008). BSH catalyses, de-conjugation of the tauro- or glyco-conjugated bile acid by acting on the non-amide bond. It also, possess the core conserved Ntn-fold and catalyses the hydrolysis mechanism similar to other Ntn-hydrolase members (Suresh, et al., 1999, Kumar, et al., 2006).

To understand the structural basis of BSH activity or to manipulate the activity in BSH, knowledge of three-dimensional structure is essential. In this context, we present in this chapter the 2.0 Å resolution structure of the BSH from *Enterococcus faecalis* (*Ef*BSH). The gene encoding BSH was cloned, sequenced, and shown to be homologous to other known BSH enzymes and penicillin V acylases (PVAs). The overexpressed enzyme was purified and used for crystallization. The quaternary structure of *Ef*BSH is tetrameric and it crystallized in $P12_11$space group.

3.2 Materials

Commercial Hampton research screens. The circular coverslips (Φ = 19 mm) were of Blue Star make. Crystal screening were performed at home source (Cu-Kα radiation) Rigaku X-ray machine and RAXIS IV detector. The high resolution data were collected at Synchrotron, INDUS-II, BL21 beam, Indore with MAR CCD detector. To collect diffraction data at low temperature, the crystal was flash cooled in a nitrogen stream produced by X-stream (Rigaku-MSC, USA). Different crystallographic software such as XDS and *MOSFLM* to integrate the diffraction images and *AIMLESS* or *SCALA* to scale the data were used during data processing. Molecular replacement was done by Phaser-MR of CCP4i suite. Coot was used for visualizing the density and fitting the residues.

3.3 Methodology

3.3.1 Heterologous-expression and purification of *Ef*BSH

The *bsh* gene of *Enterococcus faecalis* was cloned into a pET22b+ plasmid vector using gene specific primers between *NdeI* and *XhoI* restriction sites. The *Ef*BSH enzyme was expressed in *Escherichia coli* BL21 star cells with a C-terminal His-tag, and was purified to homogeneity using a His-select Ni-NTA affinity column and desalting was done using PD10 column followed by ENrichTM 650 (BioRad) size exclusion column, as described in the previous chapter. The purity of the enzyme was checked by SDS–PAGE. Protein was concentrated up to 15mg/ml by 15ml centricon (Merk) and concentration was measured using Bradford method (Bradford, 1976).

3.3.2 Protein crystallization and crystal growth

Protein crystallization is principally a trial and error process.There are several factors considered to influence crystallization process which includes pH, temperature, ionic strength of salts present, homogeneity, pressure, gravity, mixing, presence of substrates/ analogs for enzymes and all precipitating agent. Temperature also affects protein solubility and hence crystallization process. There are several methods for crystallizing the proteins including, Sitting-drop method, Hanging-drop method micro-batch method and LCP (Wendy et al., 2011) for membrane proteins. For an initial screen with commercially available screens sitting-drop method is widely exploited by researchers (Carter and Carter, 1979; Chayen, 1998, Cudney et al., 1994; Jancarik and Kim, 1991).

Crystal improvement could be performed by seeding, a technique where crystals are transferred from nucleation conditions to those that will support growth. The technique of streak seeding as described by Stura and Wilson (Stura & Wilson, 1991) uses a needle to transfer seeds of crystals to a drop of non-nucleated protein. The needle is used to touch an existing crystal and dislodge seeds from it. Some of the seeds remain attached to the needle as it is drawn out of the solution, and these can be introduced into a pre-equilibrated drop which may support mainly growth. Seeding can be employed in situations where crystals are small and of poor quality, where they take too long to grow.

In this study *Ef*BSH protein was purified to the homogeneity and concentrated up to 15-20 mg/ml Amicon ultra-15 ml 30 kDa cutoff centrifugal filters (Millipore, USA) at 4 ^{o}C with 4000 rpm. *Ef*BSH protein was screened against several commercially available crystallization screens including, Crystal screens I & II, PEG/Ion HT, Index, (Hampton Research Corp.) Nextal PACT (Quiagen) and JCSG (Molecular Dimensions). Screening was performed with Vapor-diffusion sitting-drops containing 200 nano liter (nL) of protein and 200 nL of screen solution, were set up in 96-well Corning plates containing 50 µL screen solution in reservoir well using a Mosquito Crystal Nanolitre protein crystallization robot (TTP Labtech, UK). Small cuboidal and diamond shaped crystals appeared in several conditions after 2-3 days but best native-*Ef*BSH crystal was obtained in the 2M ammonium sulfate and 2% isopropyl alcohol. Further crystal quality

improvement was performed manually by hanging-drop method. The drop size was increased to 4uL (2uL protein + 2uL reservoir solution). Finally best crystals for diffraction were obtained in 2M ammonium sulfate and 5% isopropyl alcohol. The crystals were screened for best cryo-protectant among glycerol, MPD, isopropyl alcohol and Ethylene glycol. The crystals were freezed under flash-cooled liquid nitrogen (100 K) using 25% ethylene glycol as cryo-protectant. The crystal diffraction data were collected at synchrotron source PXBL21 beamline, (INDUS-II) at Indore, India.

3.3.3 X-Ray diffraction and cryo-protectant

Once crystals are grown in the crystallization condition of suitable size and quality afterward, crystals are exposed to high intensity X-ray beam. The X-rays are produced when electrons generated from a heated filament accelerated by high voltage collide with a metal target such as copper anode. X-rays are electromagnetic radiation of wavelength about 0.1 – 10 nm (1 -100 Å) (Blundell & Johnson, 1976). The X-rays from laboratory source used were α-copper radiation of wavelength 1.5418 Å. Individual atoms in a molecule can diffract X-rays; however, it is difficult to analyze an image from a single molecule. Each molecule within the crystal contributes to the diffraction equally (Rhode, 2000). Organic solvents such as MPD, ethanol, isopropanol, ethylene glycol and glycerol are the commonly used cryoprotectants (Garman & Schneider, 1997).The *Ef*BSH crystal was screened for good cryo-protectant. The cryo-protectant is prepared with the mother liquor of that particular crystallization condition. There are several cryo protectant namely, MPD (20-30 %), IP (Isopropyl alcohol, 20-30 %), 2,4-dexanediol (20-30 %), Glycerol, PEG 400 and Ethylene glycol (EG) in the range of 20-35%.

The *Ef*BSH crystals were screened at the home source. The crystals are exposed to intense X-ray radiation; it causes extensive damage to crystals. This is due to the weak lattice forces within the crystal structure (McPherson, 1982). The X-rays can produce sufficient free radicals to cause specific chemical changes on the protein molecules such as to break the disulfide bonds (Ravelli & Garman, 2006; Ravelli et al., 2003). This problem of damage and free radicals could be alleviated by collecting the data at lower temperature (100K).

3.3.4 Data collection and processing

The screening of the *Ef*BSH was performed using in-house X-ray diffraction facility at NCL, Pune, India using CuKα (λ = 1.5418Å) radiation and Raxis IV++ area detector mounted on a Rigaku rotating anode X-ray generator operating at 50 kV, 100 mA. The better diffracting crystals were stored and collected data at high brilliance beams. Wild-type *Ef*BSH crystals were diffracted at INDUS-2 synchrotron BEAMLINE PX-BL2, RRCAT, Indore, India. Data were collected using CCD RAYONIX MX-225 detector, at 100K temperature. For Cys2Ala mutant crystal data were collected at ELETRA, synchrotron, Italy. The wild-type *Ef*BSH diffracted well and successive frames were collected with the exposure of 20 sec.

MOSFLM (Leslie & Powell, 2007) performs the functions similar to that of XDS (Kabsch, 2010). SCALA (Evans, 2006) of the CCP4 suite is used to scale the reflections (Collaborative Computational Project No.4, 1994). Xia2 (Winter, 2010) is another package available for automated data reduction of the raw diffraction images. Both the wild-type and the mutant data were processed using SCALA. During the course of the data processing, data quality and resolution limit were decided based on the data quality statistics. The quality parameters output by SCALA include χ2, Rmerge and <I/σI>. These parameters are defined by the equations given below (equations 3.1 & 3.2) (Blundell and Johnson, 1976, Otwinowsky & Minor, 1997).

$$R_{merge} = \frac{\sum_{hkl}\sum_{j=1}^{N} | I_{hkl} - I_{hkl}(j) |}{\sum_{hkl}\sum_{j=1}^{N} | I_{hkl}(j) |} \qquad \text{…….. (3.1)}$$

$$\chi 2 = \Sigma [(Ii - < Ii>)2/ (\sigma I2N / (N-1))] \qquad \text{……. (3.2)}$$

where Ii is the intensity measurement of reflection hkl, and <I> is the average intensity from multiple observations. The R_{merge} is generally recognized for statistical data quality assessment parameter in macromolecular crystallography. It denotes the extent to which

observations of each unique reflection deviate from being equal to each other. A processed good data can have R_{merge} of less than 10%. The Chi square ($\chi 2$) is the measure of the good fit of the data and good refinement will have $\chi 2$ values closer to 1.0. The Rmerge is less appropriate than the Chi-square, because of its dependency on the multiplicity of the data and on the symmetry of the crystal (Diederichs & Karplus, 1997). Rmerge alone is not sufficient to decide the high resolution cut-off because of its dependence on the redundancy of data (Weiss, 2001; Weiss and Hilgenfeld, 1997).

3.3.5 Matthew's coefficient

To estimate the number of molecules in the crystallographic asymmetric unit and the solvent content of the protein crystals based on knowledge of the molecular weight of protein is quite possible once the space group and unit cell dimensions of the crystal are known,. The following equations (3.3 & 3.4) are used (Matthews, 1968 & 1985).

Vm= Unit cell volume /(Mol.Wt.in dalton x no.of protein molecules in unit cell)

……(3.3)

Vsolv = 1 – (1.23/ Vm) **…….(3.4)**

where Vm is the Matthews number, n is the number of molecules in the unit cell and z is the Avagadro's number; Vsolv is the fraction of unit cell volume occupied by solvent. The Matthew's number for crystal of *Ef*BSH was calculated by assuming 37000 Dalton as the molecular weight of a monomer

3.3.6 Structure determination

The data were collected and processed as described above. Molecular replacement (MR) was used for the solution of the phase problem. Patterson methods (Rossmann & Blow, 1962) are used to calculate the rotation function, to obtain the orientation of the model in the new unit cell, and then the translation function calculated using model in correct orientation (Crowther & Blow, 1967) helps to place the model in the new unit cell. Programs that are used for molecular replacement are AMoRe (Navaza, 1994), PHASER (McCoy et al., 2007; Storoni et al., 2004) and MOLREP (Vagin & Teplyakov, 1997). PHASER (McCoy et al., 2007) uses maximum likelihood based methods for performing the rotation and translation searches. The model obtained from the PHASER should give the good log likelihood gain (LLG) and Z-score. LLG is the likelihood calculated from the model and Wilson distribution. A Z-score value greater than 8 usually

indicates a good and reliable solution (McCoy, et al., 2007), but structures have been solved with Z-score of the 5-6. AutoBuild combines density modification with refinement using REFMAC5 (Murshudov et al., 1997) to generate a high-quality model.

The *Ef*BSH structure was determined using molecular replacement method using data between 2.0 - 43 Å in the Phaser-MR (McCoy, et al., 2007). The crystal structure of Cholylglycine hydrolase from *Clostridium perfringenes* (PDB ID: 2RLC, chain A) was used as template for search model to calculate the phases.

3.3.7 Structure refinement

The Phaser-MR model was subjected to restrained refinement using REFMAC5 of CCP4i suite. Refinement was carried out by multiple cycle of REFMAC5 in conjunction with manual model building and fitting in the Coot 7.0 and maximum likelihood refinement was used. Water molecules were manually added in the coot at peaks of density above 4σ in asymmetric unit. After every refinement cycle the model was inspected and manual rebuilding done by inspecting (2Fo-Fc), positive (Fo-Fc) and negative (Fo-Fc) electron density maps. The model that appeared after each cycle of model building and refinement was checked for unusual geometry at each residue, the peptide-flip values, and high values of temperature factors.

3.3.8 Structure validation and quality analysis of structure

Repetitive manual rebuilding of the structure in the COOT and refinement in the REFMAC5 (Murshudov et al. 1997) minimizes the errors in phase estimated and hence it comes closer to the observed and calculated structure factor. Rfactor is a measure of the agreement between the crystallographic model and the experimental X-ray diffraction data.

$$R_{factor} = \frac{\sum_{(h,k,l)} \| F_{obs}(h,k,l) | - | F_{calc}(h,k,l) \|}{\sum_{(h,k,l)} | F_{obs}(h,k,l) |} \quad \text{..... (3.5)}$$

where F is the structure factor and the sum extends over all reflections measured and their calculated counterparts respectively.

The resolution and R-factor are initial indicators of the reliability of the structure and general fit with the observed data. However, the R factor can be sometimes biased due to exclusion or over-modelling of the data (Brunger1992). A better assessment of the fit involves the use of the statistical parameter R_{free}. Crystallographers also use the R_{Free} to assess possible over-modelling of the data. R_{Free} is computed according to the same formula given above, but on a small, random sample of data that are set aside for the purpose and never included in the refinement. In this, about 5% of the experimental observations (test set) are removed from the data set before refinement. The other 95% data are subject to refinement, and R-free is calculated on how well the model predicts the remaining 5% data in the test set. The R_{free} is always higher than that of the R_{factor} and for good refined data the gap between both the parameters should not be more than 5%.

Finally, stereochemistry of the structure was validated by PROCKECK program. DALI server was used for the structural homology search using the three-dimensional structure against the data base. The program CONTACT (CCP4i suite) and PISA were used to calculate crystal contacts and interface interaction. All structure figures were prepared in the Pymol and CCP4MG (McNicholas et al., 2011).

3.4 Results & Discussion

3.4.1 Protein purification, crystallization and data collection

The *Ef*BSH purification protocol has been described in the chapter 2 section 2.3.6. Crystallization of native *Ef*BSH from *Enterococcus faecalis* was performed using commercial screens with sitting-drop method. The commercial crystallization screens showed that the protein crystallized with different morphologies under several conditions, whereas crystals appeared after 1-2 days at 22° C (Fig. 3.1 a, b & c).

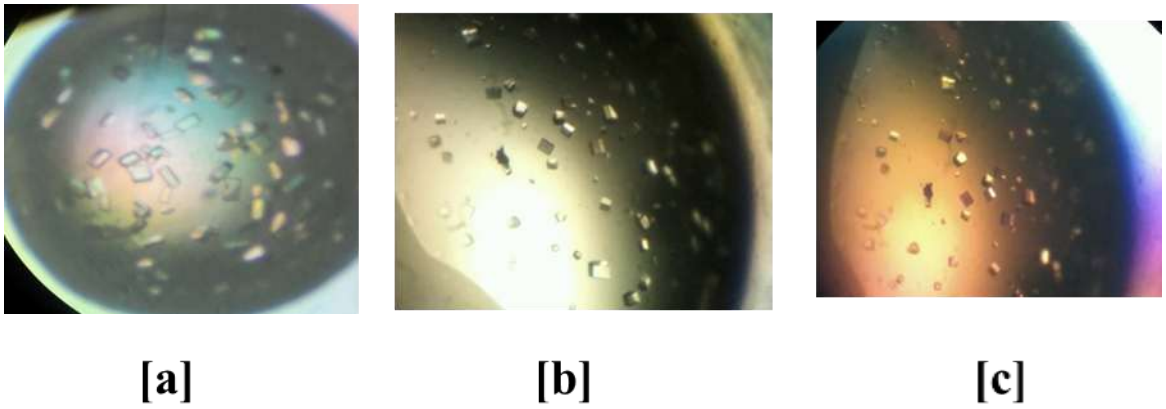


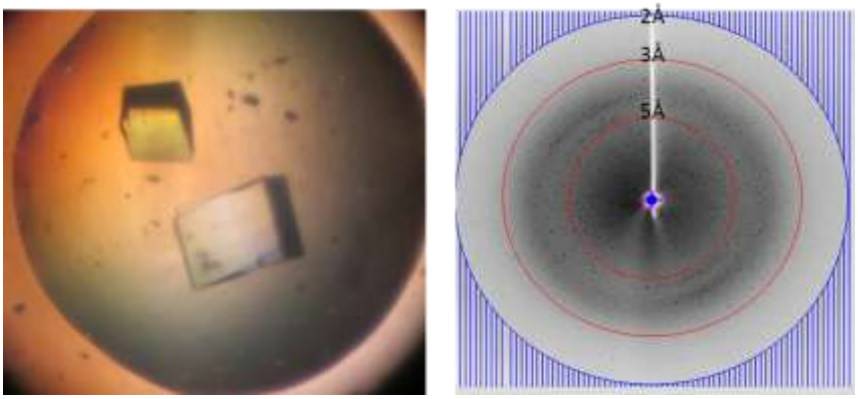
Figure 3.1 Optimization and improvement of the crystal diffraction quality, [a] 2M AS, 0.1M Na-Acetate, [b]. 1.5 M AS, 0.1M Na-Acetate, Isopropyl alcohol (IP), [c]. 2M ammonium sulfate, 2% IP

The best crystals of native *Ef*BSH were obtained in hanging-drop with 2.0 M ammonium sulfate, 5% isopropanol and grew to approximate dimensions of 0.3-0.5 mm and diffracted X-rays to 2.0 Å resolution (Fig. 3.2 a& b). The native *Ef*BSH crystal belong to space group $P2_1$ with unit-cell parameters of **a**= 66.20, **b**=131.62, **c**=86.72, **α**=90.00, **β**=94.48, **ϒ**=90.00. Detailed data collection statistics for the native *Ef*BSH dataset are given in Table.1.0

3.4.2 *Ef*BSH structure solution, refinement and validation

The data sets were recorded with a Rayonix MX225 CCD imaging plate detector system for native *Ef*BSH. The data was processed and scaled in the XDS (Kabsch, 2010) and AIMLES program respectively, and 5% of the data was reserved for cross-validation (Kasche et al., 2010). Data collection and refinement statistics for both native and mutants

structures are detailed in the table 3.1and 3.2. The Cys2Ala mutant was crystallized in two space groups: monoclinic and tetragonal. This mutant was prepared to obtain substrate-complex structure (Pen V and AHL). But no bound ligand density was found in the structure, instead density for the SO_4^- (Sulfate ion) was found near the nucleophile cysteine.



[a] **[b]**

Figure 3.2 [a] Diffraction quality crystal of *Ef*BSH in optimized condition 2M AS 5 % isopropanol and [b] diffraction pattern (PXBL21 beamline, INDUS-II synchrotron, RRCAT, Indore, India).

The structure of *Ef*BSH was determined using molecular replacement method using data between 2.0 - 41.0 Å in the Phaser-MR (McCoy et al., 2007). The crystal structure of cholylglycine hydrolase from *Clostridium perfringenes* (PDB ID: 2RLC, chain A) was used as template for search model. The best solution had a Z-score of 10 with 4 monomers (1 tetramer) in the asymmetric unit. The initial model of *Ef*BSH was subject to further model building and refinement, using Coot and Refmac5 software (CCP4) respectively. The calculated crystallographic R-factor for the final refined model was 19.7 and R-free 22.8. The geometry of the residues was confirmed from Ramachandran plot (Table 3.1 and 3.2, Fig. 3.3).

Table 3.1 X-ray diffraction and data collection of native *Ef*BSH (monoclinic). Values in parentheses represent outer shell.

Space group	**$P2_1$ Monoclinic (PDB ID: 4WL3)**
Temperature	100K
X-ray source	PXBL21 INDUS-II
Wavelength	1.0
Resolution (Å)	41.07-2.01
Unit cell parameter	a= 66.20 Å, b= 131.62 Å, c= 86.72 Å, β= 94.48˚,
Molecules per asymmetric unit	4 (one tetramer)
Matthews Coefficient (Å3 Da^{-1})	2.54
Solvent content (%)	51.7
Total No. of reflection	256366 (34168)
No. of unique reflection	97208 (13908)
Multiplicity	2.6 (2.5)
Completeness (%)	99.3 (97.8)
Average I/σ(I)	9.0 (3.5)
Rsym or Rmerge (%)	10.5 (38.7)

$$R_{symm} = \frac{\sum_{hkl}\sum_{i} | I_i(hkl) - < I(hkl) > |}{\sum_{hkl}\sum_{i}(hkl)}$$

Table 3.2 Refinement statistics of the native *Ef*BSH.

Parameters	wild type *Ef*BSH ($P2_1$) PDB ID:4WL3
Rfactor (%)	19.7
Rfree (%)	22.8
Root Mean Squire Deviation	
RMS Bond Length (Å)	0.018
RMS Bond Angle (°)	1.19
Overall B (Isotropic) from Wilson Plot	16.0
Ramachandran Plot (% residues)	
Most favourable region (%)	97
Allowed region (%)	2.8
Generously allowed region (%)	0.2
Disallowed region (%)	0

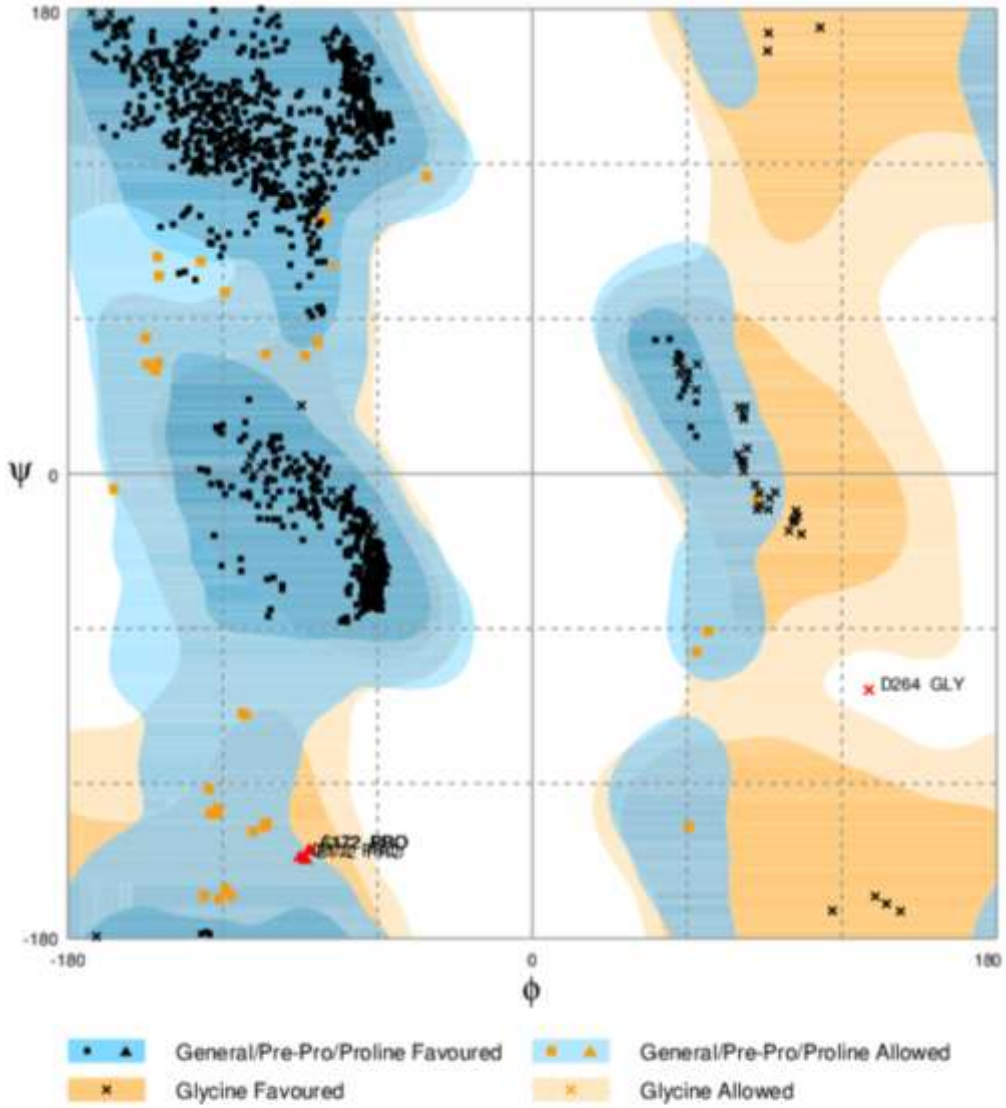
$$R_{cryst} = \frac{\sum_{hkl} | F_{obs}(hkl) - F_{cal}(hkl)|}{\sum_{hkl} F_{obs}(hkl)}$$

$$R_{free} = \frac{\sum_{hkl} | F_{obs}(hkl) - F_{cal}(hkl)|}{\sum_{hkl} F_{obs}(hkl)}$$

where the F values are test set amplitudes (5%) not used in refinement

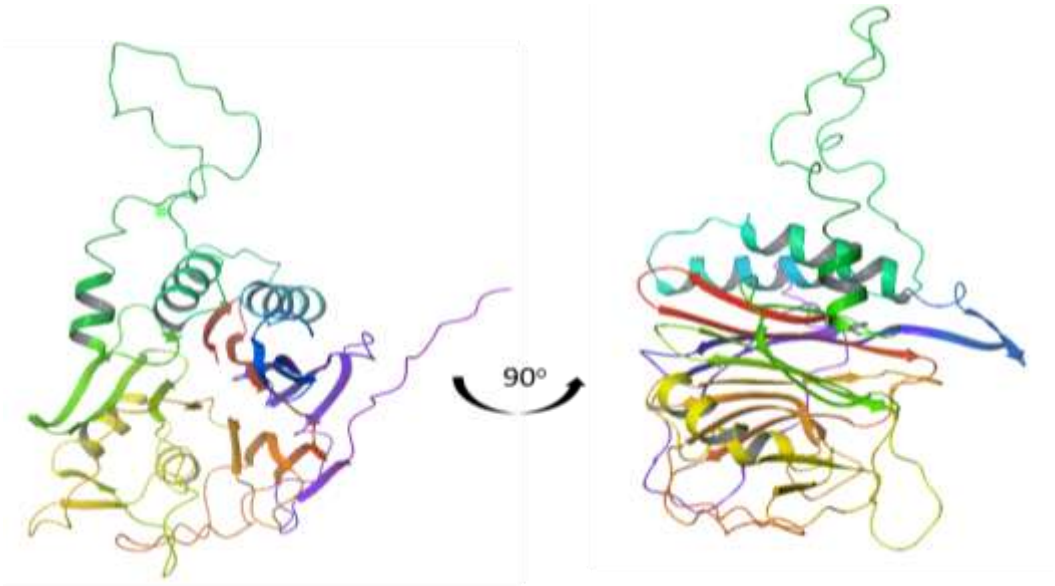
3.4.3 Overall structure of *Ef*BSH

Bile Salt Hydrolase (BSHs) belong to the Ntn hydrolase enzyme superfamily. This Ntn fold comprises of αββα arrangement of the secondary structures, wherein two β-sheets are sandwiched between α-helices, is a distinctive feature of Ntn hydrolases (Fig. 3.4a). Cholylglycine hydrolase family (CGH) comprises BSH and PVAs and exhibit noteworthy similarities in active site architecture and catalytic mechanism (Suresh et al. 1999; Kumar et al. 2006, Avinash et al., 2015). We have characterized the *bsh* locus from *Enterococcus faecalis*, a gut inhabitant, with substantially high de-conjugating activity and unique allosteric behaviour, observed for first time among characterized BSH reported. The monomer of *Ef*BSH is consist of the globular domain (~37 kDa) and forms a functional tetramer assembly of ~148 kDa molecular weight. *Ef*BSH showed 1.9 Å Cα root mean square deviations (RMSD) and 43% sequence identity with the closest BSH from *clostridium perfirngens*, calculated by the DALI server.

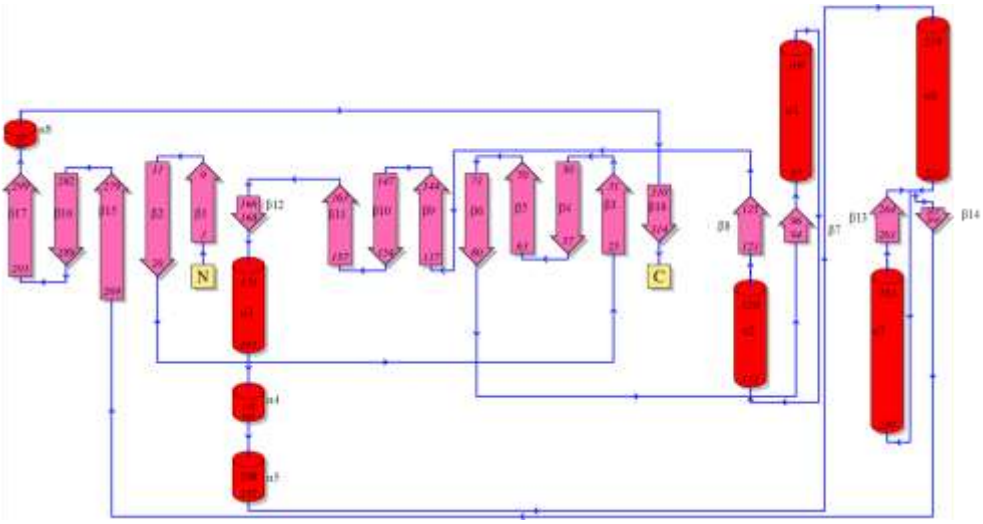


Number of residues in favoured region (~98.0% expected) : 1243 (96.8%)
Number of residues in allowed region (~2.0% expected) : 36 (2.8%)
Number of residues in outlier region : 5 (0.4%)

Figure 3.3 Ramachandran plot analysis of wild-type *Ef*BSH in PROCHECK



[a]



[b]

Figure 3.4 [a] Cartoon representation of *EfBSH* monomer from front and rotated 90°. [b]Topology diagram of the secondary structure content of *Ef*BSH (from PDBsum). The α-helices are represented by cylinders and the β-strands are represented by flat arrows.

The overall structure of the *Ef*BSH displays the core catalytic Ntn-fold of Ntn hydrolases superfamily (Oinonen and Rouvinen, 2000). The approximate dimensions for each monomer of *Ef*BSH were 40 Å x 45 Å x 50 Å, which forms a tetramer with total buried surface area (BSA) of 18910 $Å^2$. The*Ef*BSH structure possess 18 antiparallel and parallel β-strands and 8 α-helices which make up the 31 % β-strand, 16 % α-helices, 2.5% 3_{10} helix and 50.5 % other (Loops and turns)(PDBsum server) secondary structures.
The globular core of the *Ef*BSH is consists of two β-sheets one is present towards N-terminal and another towards the C-terminal (Fig. 3.4b, PDBsum server).

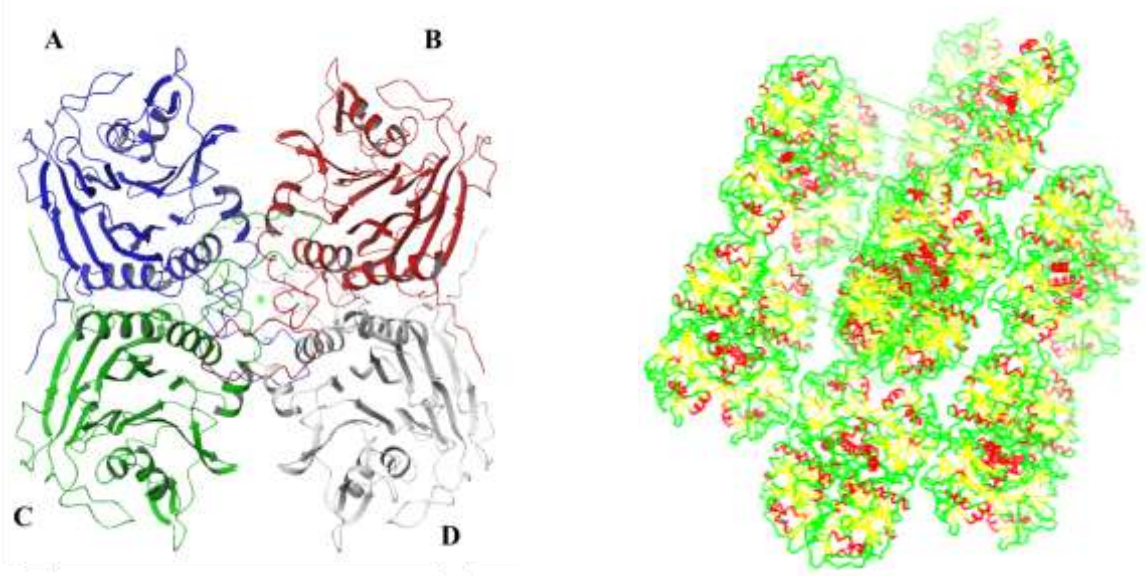


Figure 3.5 [a] Tetramer assembly of the native *Ef*BSH, [b] Native *Ef*BSH crystal packing diagram in monoclinic form.

The N-terminal β-sheets comprise of 5 antiparallel β-strands numbering β-1, β-2, β-15, β-16, β-17, and C-terminal β-sheets possesses the 9 antiparallel β-strands numbering β-3, β-4, β-5, β-6, β-9, β-10, β-11, β-12, β-18. The catalytic and substrate binding residues are distributed in the N-and C-terminal β-sheets, β-1, β-2, β-3, and β-6 strands including α 6-helice and loops (Fig. 3.4b). Cholylglycine hydrolases are known to

have four binding pocket loops in the active site binding pocket and provide catalytic framework for *Ef*BSH enzyme.

3.4.4 Tetramer assembly and interface analysis

Larger assembly formation is known to provide the structural stability and catalytic efficiency to the enzyme. *Ef*BSH forms a functional native homo-tetramer assembly in the solution confirmed by the gel filtration described in chapter 2, which is consisting of the dimer of dimer (Kumar et al., 2006)**.** The dimers are denoted as AC and BD (Fig. 3.5a). One tetramer of *Ef*BSH is present in an asymmetric unit (Fig.3.5b)**.** Monomer of the *Ef*BSH possesses an assembly motif of 25 amino acid residues extending from Arg185-Ser220 residues, inserted across to neighboring diagonally opposite monomer (AD, CB) in a criss-cross fashion as reported in other cholylglycine hydrolase members from Gram-positive bacteria (Suresh, et al., 1999, Rossocha, et al., 2005 Kumar, et al., 2006). In the tetramer, all four monomers interact with each other and maximum interface area was observed with the dimers AC /BD 3049.1 $Å^2$, however, AD/BC shares less interface area of 1417.0 $Å^2$ (Table 3.3).

The interface of the opposite subunits (AD /BC) has residues, Y83, P172, Y176, D17, F179 and N180 sharing the reciprocal interactions with 31 non-bonded contacts and 2 hydrogen bonds. However, adjacent dimer subunits (AC /BD) shares large interface with 84 amino acid residues, salt bridges, hydrogen and non-hydrogen bonds (Table 3.3). However, cholylglycine hydrolase representatives *Bt*BSH (*Bacteroides thetaiotaomicron* BSH) and *Pa*PVA (*Pectobacterium atrosepticum*) from the Gram –ve bacteria showed shortening of the assembly motif (Kim,et al., 2009, Avinash et al., 2015). The sequence analysis and annotation studies revealed an indel mutation of 13–19 amino acid residues in the assembly motif region of the members of the cholylglycine hydrolase belonging to the Gram negative bacteria (Panigrahi et al., 2014) indicating thermodynamically less stable tetrameric association. The total surface and buried surface area (BSA) for the *Ef*BSH tetramer assembly was also calculated which was found to be 45300 and 18910 $Å^2$ area, respectively, which was more than that of *Bl*BSH (43500 and15770 $Å^2$).

Table 3.3 Quantitative estimation of interface area, number of residues, hydrogen bonds, salt brides and the number of non-bonded interactions between individual subunits of *Bl*BSH, *Cp*BSH and *Ef*BSH in their quaternary structures (PISA web server).

	*Bl*BSH (2HEZ)			*Cp*BSH (2RLC)			*Ls*BSH (5HKE)			*Ef*BSH (4WL3)		
	AC/BD	AB /CD	AD/BC	AC/BD	AB /CD	AD/BC	AC/BD	AB /CD	AD/BC	AC/BD	AB/ CD	AD/BC
Interface area (Å2)	1964	310	1498	3089	355.6	686.1	2627	95.9	1407	3049	266.8	1417.0
ΔG (kcal/mol)	-27.8	-1.2	-16	-33	-1.2	-5.3	-23.7	-5.8	-12.6	-20.7	-4.2	-16.0
No. interface Residues	61	11	41	86	12	23	75	6	35	84	8	35
No. Hydrogen Bonds	20	4	14	38	8	4	38	5	23	49	2	20
No salt bridges	6	0	0	4	0	0	4	0	0	17	0	0
No. non bonded contacts	225	34	155	321	38	80	277	15	135	331	31	138
CSS	1.0	0.031	1.0	0.474	0.036	0.474	0.583	0.437	0.15	1.0	0.070	1.0

The interface of dimers AC/BD shared the 84 amino acid with the 48 H-bonds and 18 salt bridges which are more than any other BSH (Table 3.3). The detailed list of the residues involved in the H-bonding and salt bridges are given below in the table 3.4 and 3.5 respectively. Moreover, opposite dimers (AD/BC) interface has less hydrogen bonds (20) and no salt bridges were observed. However, *Ef*BSH shares a large interface between the dimers (AC/BD) in comparison to *Bl*BSH and *Cp*BSH.

The ΔG^{diss}, a measure of the free energy of assembly dissociation (with $\mathbf{\Delta G^{diss}>0}$ are thermodynamically stable) was calculated for *Ef*BSH, *Bl*BSH and *Cp*BSH corresponding to 43.7, 72.9 and 33.3 Kcal/mol (+ve values of ΔG^{diss} indicates that an external driving force required for dissociation). However, for *Bt*BSH G^{diss} value was estimated 4.9 kcal/mol, indicates the tetramer assembly is thermodynamic ally less stable (Panigrahi et al, 2014). The ΔG^{diss} value for adjacent dimer subunits (AC /BD) was estimated 31.1 which is less than for tetramer, indicating that existence of the dimer is thermodynamically less stable (Table 3.7).

Table 3.4 List of interface residues H-bonding between monomers of dimer (AC/ BD) in *Ef*BSH structure

S. No.	Monomer A	DISTANCE (Å)	Monomer C
1	A:ASN 171 [ND2]	2.9	C:TYR 205 [O]
2	A:ARG 190 [NH1]	2.9	C:ASP 262 [OD2]
3	A:ARG 207 [NH1]	2.9	C:GLU 269 [OE1]
4	A:ARG 207 [NH1]	2.1	C:GLU 21 [OE2]
5	A:GLY 216 [N]	2.8	C:LEU 214 [O]
6	A:ASP 217 [N]	3.7	C:LEU 214 [O]
7	A:SER 220 [N]	2.9	C:SER 206 [OG]
8	A:LYS 226 [NZ]	2.8	C:GLY 216 [O]
9	A: VAL 235 [N]	3.3	C:LYS 258 [O]
10	A:TYR 241 [OH]	2.7	C:TYR 322 [OH]
11	A:SER 245 [OG]	2.8	C:TYR 289 [O]
12	A:SER 245 [OG]	2.9	C:ASP 290 [OD1]
13	A:GLN 246 NE2]	3.1	C:LYS 258 [O]
14	A:HIS 249 [ND1]	2.6	C:GLU 255 [OE2]
15	A:HIS 249 [ND2]	2.8	C:TYR 289 [O]
16	A:ASP 262 [N]	2.7	C:THR 189 [O]
17	A:ARG 287 [NE]	2.8	C:SER 292 [OG]
18	A:TYR 289 [OH]	2.6	C:GLN 246 [OE1]
19	A:SER 292 [OG]	3.5	C:SER 292 [OG]
20	A:SER 292 [OG]	2.7	C:SER 292 [O]
21	A:GLN 293 [NE2]	2.9	C:GLN 319 [O]
22	A:ALA 296 [N]	2.8	C:ASN 321 [O]
23	A:TYR 312 [OH]	2.8	C:ASN 324 [OD1]
24	A:ASN 321 [N]	3.0	C:ILE 294 [O]
25	A:TYR 322 [OH]	2.6	C:TYR 241 [OH]
26	A:ALA 323 [N]	2.9	C:ALA 296 [O]
27	A:GLU 21 [OE2]	3.5	C:ARG 207 [NH2]
28	A:THR 189 [O]	2.8	C:ASP 262 [N]
29	A:TYR 205 [O]	2.9	C:ASN 171 [[ND1]
30	A:SER 206 [OG]	2.9	C:SER 220 [N]
31	A:LEU 214 [O]	2.7	C:GLY 216 [N]
32	A: LEU 214 [O]	3.8	C:ASP 217 [N]
33	A:GLN 246 [OE1]	2.6	C:TYR 289 [OH]
34	A:GLU 255 [OE2]	2.6	C:HIS 249 [ND1]
35	A:LYS 258 [O]	3.3	C:VAL 235 [N]
36	A:LYS 258 [O]	3.1	C:GLN 246 [NE2]
37	A:ASP 262 [OD2]	2.9	C:ARG 190 [NH1]
38	A:GLU 269 [OE1]	2.9	C:ARG 207 [NH1]
39	A:TYR 289 [O]	2.7	C:HIS 249 [NE2]

40	A:TYR 289 [O]	2.6	C:SER 245 [OG]
41	A:ASP 290 [OD1]	2.6	C:SER 245 [OG]
42	A:SER 292 [O]	2.9	C:SER 292 [OG]
43	A:SER 292 [OG]	2.9	C:ARG 287 [NE]
44	A:ILE 294 [O]	2.9	C:ASN 321 [N]
45	A:ALA 296 [O]	2.7	C:ALA 323 [N]
45	A:ASP 298 [O]	3.1	C:ASN 324 [ND2]
47	A:GLN 319 [O]	2.8	C:GLN 293 [NE2]
48	A:ASN 321 [O]	2.8	C:ALA 296 [N]

Table 3.5 Salt bridges of the dimer interfaces (AC/BD) of *Ef*BSH

S.No.	Monomer A	Distance (Å)	Monomer C
1	A:ARG 190 [NH1]	3.7	C:ASP 262 [OD1]
2	A:ARG 190 [NH1]	2.9	C:ASP 262 [OD2]
3	A:ARG207 [NE]	3.9	C:GLU 269 [OE1]
4	A:ARG 207 [NH1]	2.9	C:GLU 269 [OE1]
5	A:ARG 207 [NH1]	3.4	C:GLU 221 [OE1]
6	A:ARG 207 [NH1]	2.1	C:GLU 21 [OE2]
7	A:ARG 207 [NH1]	3.9	C:GLU 21 [OE2]
8	A:HIS 249 [ND1]	3.6	C:GLU 255 [OE1]
9	A:HIS 249 [ND1]	2.6	C:GLU 255 [OE2]
10	A:GLU 21 [OE1]	3.3	C:ARG 207 [NH1]
11	A:GLU 21 [OE2]	3.6	C:ARG 207 [NH1]
12	A:GLU 21 [OE2]	3.5	C:ARG 207 [NH2]
13	A:GLU 255 [OE1]	3.5	C:HIS 249 [ND1]
14	A:GLU 255 [OE2]	2.6	C:HIS 249 [ND1]
15	A:ASP 262 [OD1]	3.6	C:ARG 190 [ND1]
16	A:ASP 262 [OD2]	2.9	C:ARG 190 [ND1]
17	A:GLU 269 [OE1]	3.9	C:ARG 207 [NE]
18	A:GLU 269 [OE2]	2.9	C:ARG 207 [NH1]

Table 3.6 List of the interface H-bonding between opposite dimer interfaces (AD/BC).

S.No.	Monomer A	DISTANCE (Å)	Monomer D
1	A:ASN 183 [ND2]	2.6	D:TYR 205 [OH]
2	A:ARG 185 [NE]	2.7	D:PRO 192 [O]
3	A:ARG 185 [NH1]	2.9	D:LEU 202 [O]
4	A:ARG 185 [NH2]	2.9	D:PRO 192 [O]
5	A:ASN 195 [N]	2.8	D:ARG185 [O]
6	A:ASN 195 [ND2]	3.0	D:VAL 146 [O]
7	A:PHE 196 [N]	3.3	D:ARG 185[O]
8	A:GLN 199 [NE2]	3.0	D:LEU 202[O]
9	A:ASP 203 [N]	2.9	D:ASN 182[OD1]
10	A:TYR 205 [OH]	2.6	D:ASN 183 [OD1]
11	A:ASN 233 [ND2]	2.9	D:ASN 183 [OD1]
12	A:PRO 192 [O]	2.8	D:PHE 196 [O]
13	A:LEU 202 [O]	2.0	D:ARG 185[NE]
14	A:PRO 192 [O]	2.9	D:ARG 185[NH1]
15	A:ARG 185 [O]	2.8	D:ARG 185 [NH2]
16	A:VAL 186 [O]	3.1	D:ASN 195 [N]
17	A:ARG 185 [O]	3.0	D:PHE 196 [N]
18	A:LEU 185 [O]	3.5	D:GLN 199 [NE2]
19	A:ASN 182 [OD1]	2.9	D:ASP 203 [N]
20	A: PHE 196 [O]	3.0	D:ASN 233 [ND2]

Table 3.7 Assembly interface dissociation parameter analysis. ΔG^{diss} indicates the free energy of assembly dissociation, in kcal/M. The free energy of dissociation corresponds to the frenergy difference between dissociated and associated states. Positive values ofΔG^{diss} indicate that an external driving force should be applied in order to dissociate the assembly, therefore assemblies with ΔG^{diss}>0 are thermodynamically stable.

Structures	Assembly	ΔG^{diss}, kcal/mol
***Ef*BSH (4WL3)**	ABCD	43.7
	AB / CD	31.1
	AC / BD	11.2
***Bl*BSH (2HEZ)**	ABCD	72.9
	AB / CD	31.1
	AC / AD	23.6
***Bt*BSH (3HBC)**	ABCD	4.9
***Ls*BSH (5HKE)**	ABCD	36.1

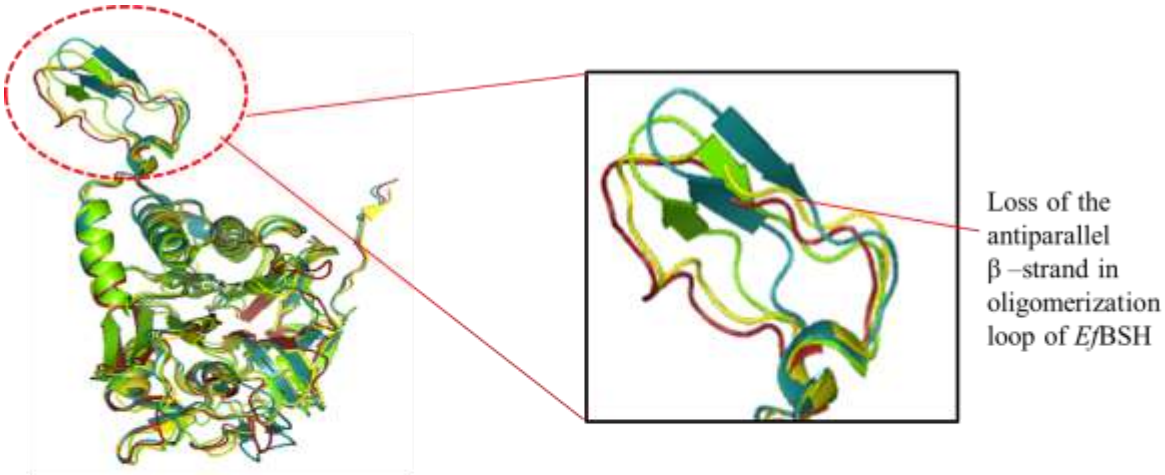
*Ef*BSH showed unique enzyme kinetics exhibiting the cooperativity. Superposition of all individual monomers of the *Ef*BSH tetramer assembly reveals significant conformational differences among monomers with overall Cα atom RMSD of 0.8-0.9 Å. Significant RMSD difference among monomers of tetramer assembly which could be due to cooperativity properties of enzyme and displays respectable agreement with kinetic data of *Ef*BSH enzymes which display the unique allosteric behavior (Chand, et al., 2015). However, the RMSD value for the *Cp*BSH monomers upon superposition was found less than 0.3 Å which showed the Michaelis-Menten kinetics (Rossocha et al., 2005) and no cooperativity (Table 3.8). Assembly motif in Cholylglycine hydrolase (BSH / PVA from Gram-positive) have an antiparallel β-strand, which was absent in the case of *Ef*BSH structure and was transformed in the loop (Arg185-Ser220 aa residues).

Table 3.8 The table shows the root mean square deviation (rmsd) of all four subunits upon superposition, of tetramer assembly from cholylglycine hydrolase family (BSH /PVA).

S.No.	PDB ID	RMSD among subunits of a tetramer assembly (Å)	Type of kinetics (MM/ sigmoidal)
1	2rlc (*Cp*BSH)	0.1-0.3	MM
2	2hez (*Bl*BSH)	0.1-0.3	MM
3	5hke (*Ls*BSH)	0.1-0.3	MM
4	4wl3 (*Ef*BSH)	**0.8-0.9**	**Allosteric behavior (Sigmoidal)**
5	2oqc (*Bs*PVA)	0.1-0.2	MM
6	3pva (*Bsp*PVA)	0.1	MM
	4wl2 (*Pa*PVA)	**0.6-0.7**	**Allosteric behavior (Sigmoidal)**

***MM corresponds to Michaelis-Menten**

This dynamic tetramerization loop (assembly motif) plausibly involved in the transduction of the conformational changes, upon substrate binding, from one subunit to diagonally opposite subunit (A-D subunit), resulting in the allosteric behavior. The loss of the antiparallel β-strands in the assembly motif probably has given the flexibility, which perhaps established better communication across subunits. In the *Bl*BSH assembly loop antiparallel β-strand were formed by the Thr-Ser and Ser-Leu residues which were replaced in *Ef*BSH by Asn-Phe and Glu-Leu pairs, resulting in loss of the antiparallel β-strand and formation of loop (Fig. 3.6). *Pa*PVA from Gram-negative bacteria had been reported loss of the assembly motif and small tetramerization loop remained as remnant of the assembly loop (Avinash et al., 2015). However, interface area and contacts among monomers of tetramer assembly were increased in comparison to *Bacillus shaericus* PVA (*Bsp*PVA).



S. No.	2RLC	2HEZ	5HKE	*Ef*BSH (4WL3)
1	EFK (β –strand 1)	TS (β –strand 1)	TF (Loop)	NF (Loop)
2	SLT (β –strand 2)	SL (β –strand 2)	DL (Loop)	EL (Loop)

Figure 3.6 Loss of anti-parallel β-strands of assembly loop in *Ef*BSH. This assembly or tetramerization loop extends across the subunit in criss-cross fashion and form the stable and functional tetramer assembly.

The feasible role of the increase or shortening of the tetramerization loop is still unclear. *Ef*BSH was treated with 2M GdnHCl for 4 hr and size exclusion chromatography was performed which demonstrated no dissociation of the tetramer into dimer. However *Bl*BSH is known to dissociate with 1.5M GdnHCl (Kumar et al., 2006).

3.4.5 Active site architecture and comparison to other cholylglycine hydrolases

The catalytic mechanism of *Ef*BSH is also similar to that of the other Ntn-hydrolases. Most of the catalytic residues at the active site which are involved in the hydrolysis of the substrate are conserved across the Ntn-hydrolase family members. In the cholylglycine hydrolase family conserved residues are Cys2 Arg16, Asp19, Asn79 Asn170 and Arg223 (numbering is according to *Ef*BSH) (Fig. 3.7). In the case of the PVAs Asn79 is replaced by the aromatic residues (Phe, Tyr, Trp).

The N-terminal catalytic cysteine residue is responsible for the nucleophilic attack on the amide bond followed by the stabilization of a tetrahedral intermediate by the formation of an oxyanion hole. The oxyanion hole is stabilized by the main chain NH2 group of Asn79 (Tyr, Phe, Trp in PVAs) and Asn170. The deacylation phase and release of the second product requires the involvement of a water molecule (Suresh et al. 1999).

The role of the Arg16 is known in the substrate hydrolysis by enhancing the nucleophilicity of the Cys2 –SH group. In addition, Arg16 has also been hypothesized in the autocatalysis processing of the pre-peptide (Rossman 2008). The distance of the conserved residues from the nucleophile cysteine was measured and it was found that Arg16 and Asp19 are 3.5 Å and 3.9 Å distant, respectively. The –NH2 of the Arg16 and –COOH group of the Asp19 form the hydrogen bonding to the amino group (-NH2) of the nucleophile cysteine (Fig. 3.8). However, the residues which form oxyanion hole Asn79 and Asn170 are 4 Å apart from the nucleophile cysteine. The active site catalytic framework of the BSH is formed by the four active site loops (Fig. 3.9, table 3.9). There was some deviation in the loops observed. The *Ef*BSH active site loops were slightly moved away from the nucleophile cysteine. However, *Cp*BSH loops were observed to be moved inward as compared to other BSH. *Cp*BSH shows less activity toward the conjugated bile acid compared to *Ef*BSH and also displays cross activity towards the Pen V substrate (Rossocha et al. 2005; Kumar et al. 2006).

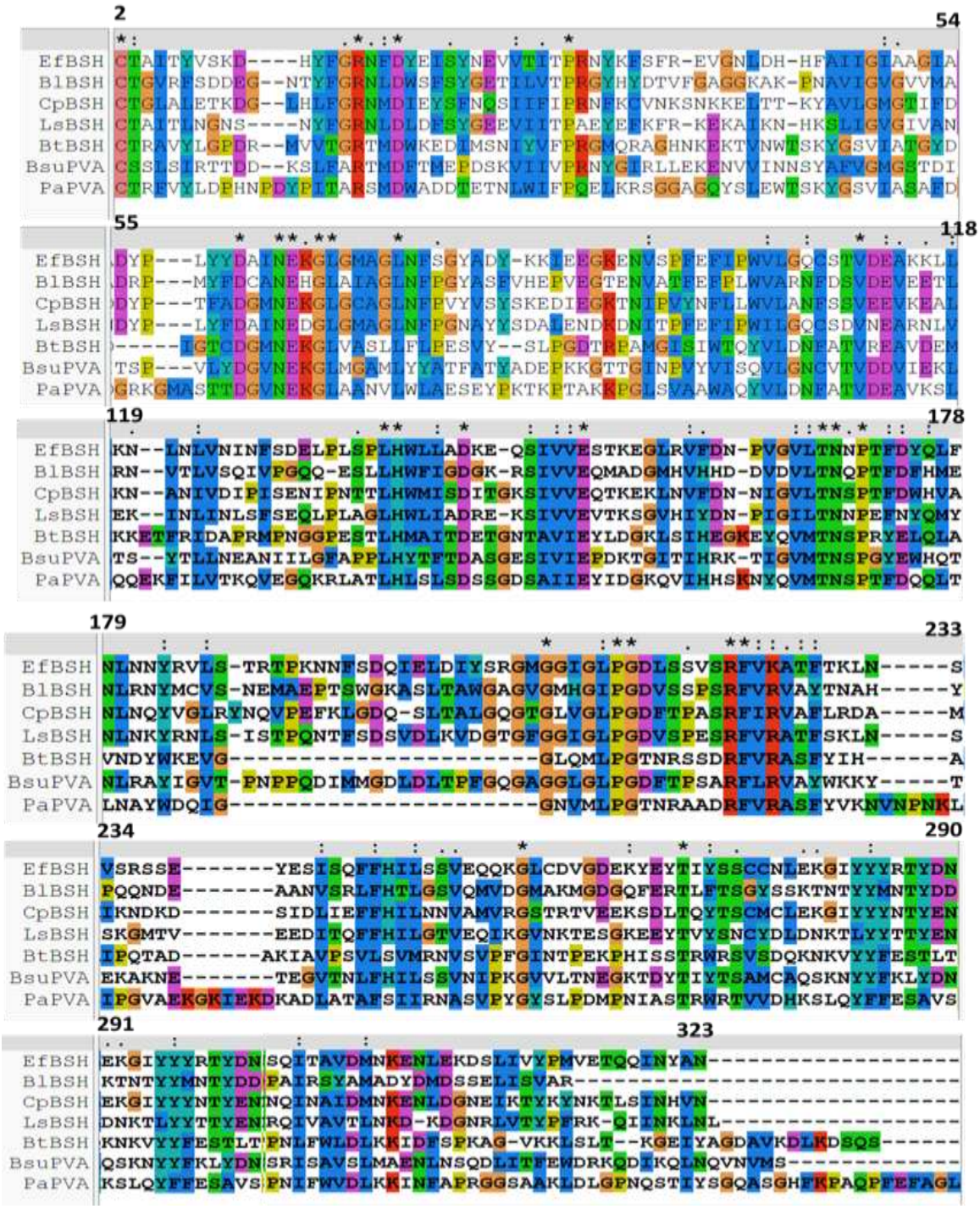
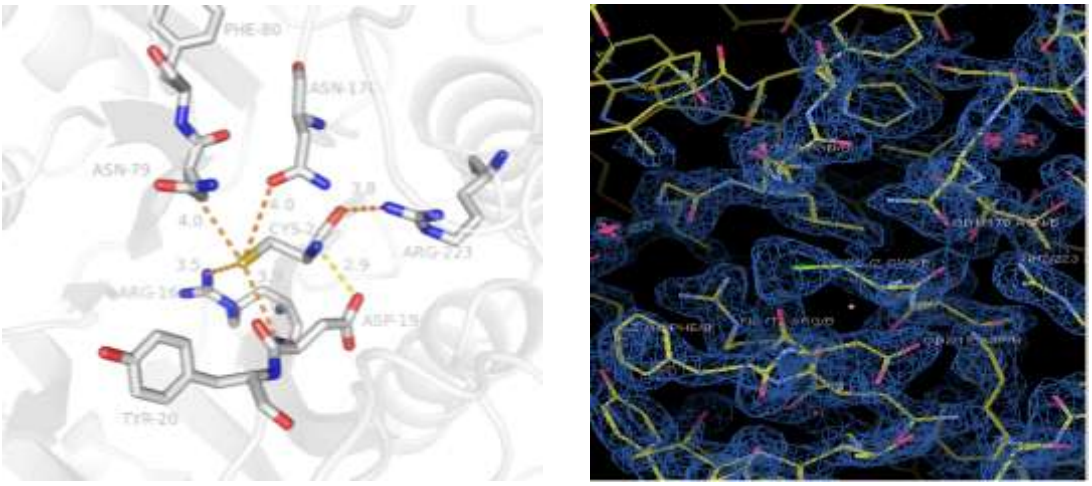


Figure 3.7 Sequence alignment of the all BSHs. *Ef*BSH (*Enterococcus faecalis*), *Bl*BSH (*Bifidobacterium longhum*), *Cp*BSH (*Clostridium perfringens*), *Ls*BSH (*Lactobacillus salivarius*) *Bt*BSH *(Bacteroides thetaiotaomicron VPI*), *Bsp*PVA (*B. Sphaericus*), *Pa*PVA(*Pectobacterium atrosepticum)*.Numbering is according to *Ef*BSH.



[a] **[b]**

Figure 3.8 [a] Active site architecture of the *Ef*BSH and distance (Å) of different active site residues from sulfhydryl group of the N-terminal nucleophile cysteine. [b] 2Fo-Fc map for the active site residues contoured at 1δ.

Probably inward movement of the loop resulted in small substrate binding pocket in which only Pen V could sit. *Ef*BSH shows large binding pocket as compared to other BSH in which loops are moved outward from nucleophile cysteine. Therefore, no Pen V activity was detected in *Ef*BSH since bile acids are bulky molecules in comparison to Pen V. Hence Pen V can't sit properly in the binding pocket.

Table 3.9 Substrate binding pocket loops of the *Ef*BSH and comparison with *Bl*BSH and *Cp*BSH

Loop No.	***Ef*BSH (4WL3)**	**2HEZ**	**2RLC**
Loop1	21-25 aa	22-27 aa	23-27 aa
Loop2	**57-64 (6 aa)**	61-62 aa	62-63 aa
Loop3	125-137	129-139 aa	130-141 aa
Loop4	255-269 aa	256-271 aa	259-274 aa

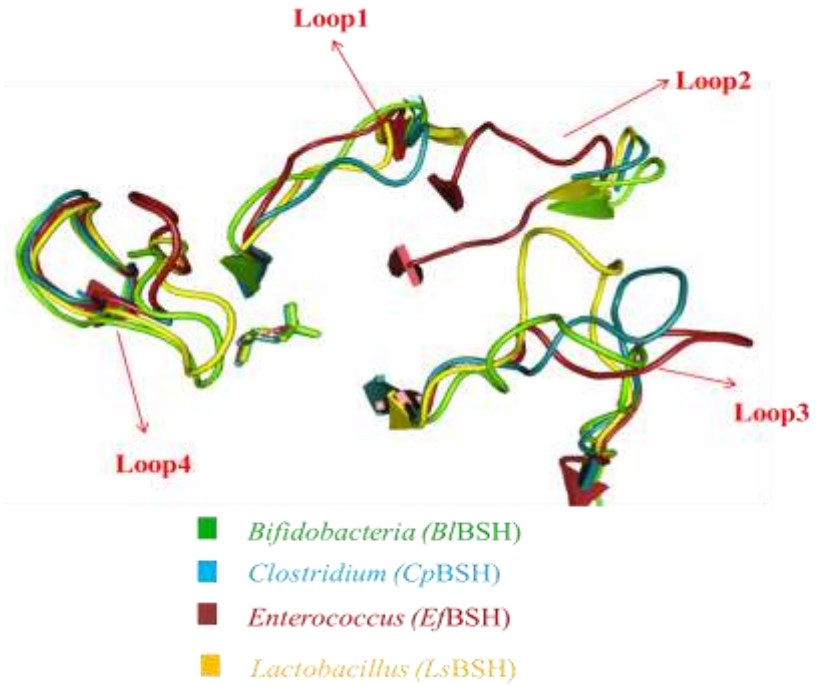


Figure 3.9 Binding pocket loop comparison of *Ef*BSH (Red color) with *Bl*BSH (green) and *Cp*BSH (Cyan). There were few insertions observed in Loop_2 and Loop_3. *aa corresponds to amino acids

Figure 3.10 Disruption of anti-parallel beta sheet which constitute the peculiar core αββα Ntn- fold. This disruption of 4 amino acid has resulted in the formation of the large loop_2. In other BSH loop_2 consist of two amino acid Ala and Asp (AD).

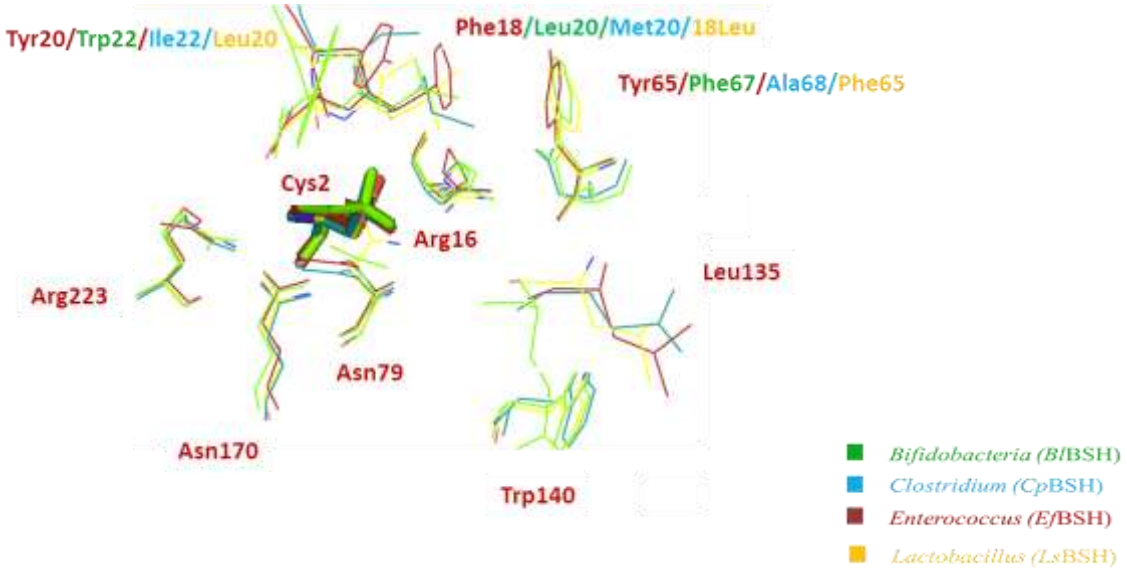


Figure 3.11 Comparison of the *Ef*BSH substrate binding pocket with *Bl*BSH and *Cp*BSH.

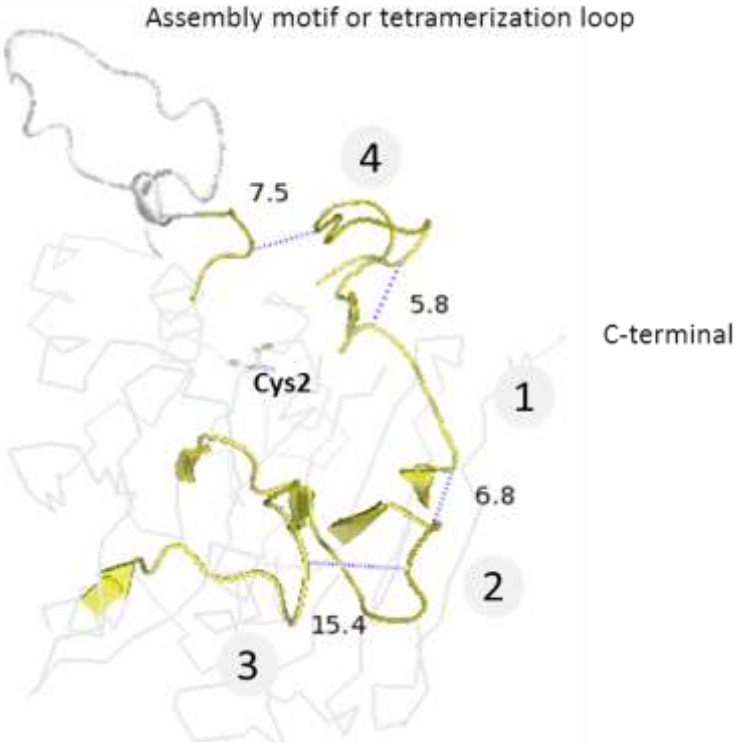


Figure 3.12 Binding pocket loops distance (Å) measured from each other (loops are numbered). The loops2 has insertion of 6 amino acids which is possibly assumed to be involved in the transmission of conformational dynamics to the assembly loop.

These four loops which enclose the active site in CGHs (Kumar et al. 2006) are comprised of residues 21-25 (loop1), 57-64 (loop2), 125-137 (loop3) and 255-269 aa (loop4) in case of *Ef*BSH as shown in Table 3.9. In loop2 of *Ef*BSH, insertion of 6 amino acid residues was found. This insertion has disrupted one of the core antiparallel β-strand of the core conserved αββα-Ntn-fold (Fig.3.10), in which half of the anti-parallel β-strand has resulted in the large dynamic loop. These loops are assumed to influence the catalytic efficiency and substrate specificity between BSH and PVA (Lambert et al. 2008, Panigrahi et al 2014).

*Ef*BSH was biochemically characterized and it was found that it showed very high BSH activity and unique allosteric behaviour (Chand et al., 2015). Major structural differences was observed such as 6 amino acid insertion in the core antiparallel β-strand of the core conserved αββα-Ntn-fold and loss of antiparallel β-strand in the assembly loop which plausibly has contributed high activity and allosteric nature to the *Ef*BSH enzyme.

The reduction of antiparallel β-strand of the core conserved αββα-Ntn-fold has resulted in the gain of the large loop2. Loop2 is known to involve in the substrate binding and it is at the opening of the substrate entry. This large loop2 has become dynamic in comparison to *Bl*BSH, *Cp*BSH and other cholylglycine hydrolase members. It probably, expedite the shuttling of the substrate and hence shows high catalytic efficiency. Moreover, active site of the *Ef*BSH was superposed on the other BSHs active site and some noteworthy differences were observed. The *Ef*BSH has the Phe18, Tyr20 and Tyr65 residues which might be interacting with the bile acid backbone cholesterol. However, in the case of the *Bl*BSH these residues are replaced by the Leu20, Trp22, and Phe67 respectively (Fig. 3.11). *Bl*BSH has very high affinity towards the bile acid but it shows less catalytic efficiency as compared to *Ef*BSH, suggesting tight binding of substrate in *Bl*BSH, which might be contributed by the strong pi-pi interaction of hydrophobic residues Trp22 and Phe67 with the cholesterol ring of the bile acids. Strong interaction of substrate basically leads the slow turnover rate for the enzyme. However, in the *Cp*BSH and *Ls*BSH similar positions are occupied by the Met20, Ile22, Ala68 and Leu18, Leu20, Phe65, respectively.

*Cp*BSH shows very less BSH activity firstly because loop3 has shown inward movement toward nucleophile Cysteine and reducing the binding pocket; secondly

affinity of *Cp*BSH is reduced probably because of the replacement of the aromatic residues (Trp22, Phe67, *Bl*BSH) to Ile22 and Ala68 residues (Fig 3.11). In addition, due to reduction in the binding pocket small Pen V molecule can sit and get hydrolysed. The Leu135 in the Loop3 is conserved in the cholylglycine hydrolase family members (Fig. 3.11) and Trp140 is conserved in the BHSs and in the PVAs it is replaced by the Met, Leu and Tyr residues (Fig. 3.11 & Fig 3.7). PVAs have another conserved residue, Met21, which in the case of the BSHs is replaced by either of Leu, Ile, or Phe residues. Although, *Cp*BSH possess Met21 and also shows the activity towards Pen V, hence the role of the Met21 could be important to differentiate the substrate Pen V and bile acids.

Table 3.10 Structural alignment of *Ef*BSH with other Ntn hydrolases (DALI server, Holm and Rosenstrom 2007)

PDBs	RMSD	% identity	Description
2rlc	1.9	43	Choloylglycine hydrolase, Clostridium perfringenes
2hez	1.9	37	Bile salt hydrolase;*Bifidobacterium longhum*
2oqc	1.8	32	Penicillin v acylase
3pva	2.2	32	Penicillin v acylase
3hbc	2.3	20	Cholylglycine hydrolase, gram negative
4wl2	2.7	23	Penicillin V acylase, gram negative
1gk1	3.5	10	Cephalosporin acylase
1ghd	3.4	11	Glutaryl-7-aminocephalosporanic acid acylase;
4m1j	3.5	10	Acyl-homoserine lactone acylase pvdq subunit beta
1pnk	3.2	10	Penicillin amidohydrolaseBeta subununit
5cz6	3.6	4	Proteasome subunit alpha type-2
5hke	1.1	54	*Lactobacillus salivarius* BSH

Apart from high efficiency *Ef*BSH display the allosteric behaviour. The insertion in loop2 which resulted in reduction of anti-parallel β-strand fraction of the core

conserved αββα-Ntn-fold into large Loop2 and loss of the anti-parallel β-strand in the assembly motif or tetramerization loop might be correlated with the cooperative nature of the enzyme. The distance between all four loops were measured; distance of Loop1-loop2 is 6.8 Å, loop2-loop3 15.4 Å, loop1-loop4 5.8 Å and loop4 to termini of the assembly loop 7.5 Å (Fig. 3.12). Due to the large loop2, it can transmit the conformational changes from loop4 to assembly loop (no anti-parallel β-strand) which further transmit to the neighbouring subunit.

*Ef*BSH structure is similar to the structure of *clostridium* (43% sequence similarity) and *Bifidobacterium* (37% sequence similarity) BSH. It shares very less sequence similarity with *Bacteroid thetaiotamicron* choloylglycine hydrolase, a Gram –ve BSH showing RMSD values for Cα atoms 2.3 Å (*Bt*BSH, 20%) which has emerged from a structural genomics project (Table 3.10) (DALI server, Holm and Rosenstrom, 2007). It shares more sequence similarity (32 %) with the Gram-positive and less with Gram-negative (23%) bacterial PVAs (Table 3.10).

3.5 Summary of the Chapter:

Crystal structure of the bile salt hydrolase from *Enterococcus faecalis* (*Ef*BSH) was solved at 2.0Å resolution. The structural comparison was performed with the existing BSHs structures from the *Bl*BSH, *Cp*BSH and *Ls*BSH. *Ef*BSH showed unique characteristics of very high BSH activity and allosteric behavior which is distinct from other BSHs. Structural comparison has demonstrated the probable reason for the unique characteristics of the *Ef*BSH. The structure of the *Ef*BSH has showed insertion of the 6 and 3 amino acid in the loop2 and loop3 respectively which resulted in the reduction of anti-parallel β-strand fraction of the core conserved αββα-Ntn-fold into large Loop2 due to which dynamicity of the substrate shuttling might have enhanced. In addition, active site residues Phe18, Tyr20 and Tyr65 may have role in the substrate affinity. Moreover, loss of beta strands in the assembly loop could plausibly be involved in the transmission of conformational changes upon substrate binding to another subunit.

Chapter-4

"Substrate binding analysis, site-directed mutagenesis and autocatalytic processing of *Ef*BSH"

4.1 Introduction

Protein-ligand docking is used as computational tool to study the protein and small- ligand interactions. The protein-protein docking can also provide insight into the intermolecular complex interactions formed between molecules. The protein-ligand docking process involves the screening of different ligand-conformations and orientations (the pose) within a given target receptor protein. These have been grouped into rigid-docking methods and flexible docking (flexible ligand, and flexible ligand - flexible receptor methods). Docking studies provide information that could be exploited to engineer the enzyme for various properties like high catalytic efficiency, modification of substrate spectrum and to design inhibitors for pharmaceutically important target enzyme. Therefore docking is a potential *in-silico* method for high throughput screening of the thousands of ligands or inhibitor molecules for pharmaceutical and industrial applications.

Docking gives the information of the ligand how it interacts with the receptor but the stability of the ligand can't be measured by this method. Therefore, molecular dynamics simulation studies enable the evaluation of backbone and side chain movement within protein-ligand docking and also measure stability of the ligand with time course. This enables identification of the best protein-ligand conformation pose which is more stable. The details of substrate binding also can be studied through a variety of experimental methods like site-directed mutagenesis. The site-directed mutagenesis approach allows specific residues in the substrate binding pocket of the enzyme to be mutated into desired amino acid and thus, explore the role of these residues in the activity and substrate binding. These mutants may result in enzyme acquiring unique properties which are different from the native enzyme, including catalytic efficiency, stability and substrate specificity. In protein engineering, similar method is exploited to modify the protein binding site in order to obtain desirable traits (stability, selectivity, high catalytic activity etc).

*Ef*BSH was observed to be highly active towards the bile acids, which also show unique allosteric kinetic behaviour (Chapter2) in comparison to other reported BSHs. To understand the structural changes that contribute towards the unique behaviour like high catalytic efficiency and cooperativity behaviour, we have determined the three-

dimensional structure of *Ef*BSH (Chapter3). The Cys2Ala a null mutant was also prepared to crystallize *Ef*BSH with the bile acid substrates (GCA, TCA, GDCA) and Pen V to understand the *Ef*BSH-substrate complex interaction at the binding pocket but no substrate bound complex was obtained. The Panigrahi et al., (2014) has reported polar complementarity in the CGHs members is a basis for the specificity and better catalytic activity. Therefore, to analyse the *Ef*BSH active interaction with substrate (GCA) and to estimate the polar complementarity, docking and MD simulation studies were carried out.

This chapter details the analysis on binding of GCA and Pen V by docking and molecular dynamics (MD) simulations. Potential binding sites on receptor was identified using SiteMap (Halgren, 2007). Grid based receptor (receptor) and ligand (flexible) docking program Glide (Friesner *et al.*, 2006) was used to predict the binding modes of ligand in the receptor binding site. Site-directed mutants were prepared to explore the importance of the active site residues involved in substrate binding and catalysis. Furthermore, to understand the autocatalytic processing of *Ef*BSH, di-and tri-pre-peptide (MG, MA, MLG & MLG_TA_SS mutant) were inserted before the nucleophile cysteine and some mutants also were prepared.

4.2 Materials

All chemicals TriCl buffer, NaCl, Pen V, imidazole DTT were used from Sigma (Sigma, USA), Merck (Merck, USA) and SRL (SRL, India). Enzymes used for site-directed mutagenesis were from Takkara (Ex-Taq polymerase) and NEB (*DpnI*). The 3-D conformers of the GCA and Pen V were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/search/). Schrodinger suite was used for Docking and Molecular Dynamics Simulation studies.

4.3 Methodology

4.3.1 Substrate specificity and Pen V modulation

The *Ef*BSH hydrolyses the conjugated bile acids. It displays specificity towards glyco-conjugated bile acids over tauro-conjugated chapter2. In addition *Ef*BSH has very high catalytic efficiency in comparison to other known BSH enzymes. Non-substrate ligand Pen V modulated the *Ef*BSH activity. Steady-state fluorescence measurement was performed for the native *Ef*BSH enzyme along with Pen V (2-16mM concentration). Change in fluorescence intensity after the binding of Pen V was monitored at λmax of the protein which has showed Pen V binding to *Ef*BSH.

4.3.2 Substrate binding mode analysis by docking

In silico methods to study ligand binding in proteins include docking and molecular dynamics simulations. For the molecular docking, the 3D conformers of Glycocholic acid (GCA) and non-substrate ligand Pen V were retrieved from the PubChem compound database (Bolton et al. 2008) with their CID 10140 and 6869, respectively. Grid based ligand docking program Glide (Freisner et al. 2006) was used for docking of these ligands in the binding site of receptor protein *Ef*BSH. The protein was first pre-processed, energy minimized and followed that by grid generation in the Glide module of the Schrodinger suite.

The receptor binding site was defined as a grid box of 20x20x20 Å dimension, by defining the active site residue of receptor protein using constrain residue Cys2 in which it will be the central residue. After grid generation ligand was prepared in the Ligprep generating several conformers of the ligand. The OPLS_2005 all-atom force field was used to determine the partial atomic charges of each ligand atom using LigPrep. Receptor grid generation was followed by ligand docking. The docking of ligand was performed using flexible docking and Glide's extra precision (XP) protocol. Free energy of receptor-ligand binding is estimated by using GlideScore, which includes van der Waals interaction, electrostatic and other terms for rewarding or penalizing interactions that are known to influence ligand binding. Best pose of the ligand is selected on the basis of the glide score and e-model score.

4.3.3 Site-directed mutagenesis (SDM)

Wild-type clone pET22b+*Ef*BSH was used as template for preparation of the site directed mutant clones. Primer pairs used to construct these mutants are, listed in the table 4.1. Mutation of active site residues C2, R16, D19, Y20, and N79 was done using ExTaq enzyme (TaKaRa Bio Inc.). In addition, stretch of the amino acids was also inserted in the *Ef*BSH clone before the nucleophile cysteine residues to prepare the clones MG, MA and MLG- peptide to understand the autocatalytic processing of the pre-peptide. All the mutants were generated by PCR (Polymerase chain reaction) using conditions: 94°C/2min denaturation and 25 cycles of (94°C/30s, 55°C/30s, 72°C/6.30 min) with final extension 72°C/10 min. *DpnI* treatment was given to the reaction and transformed into DH5α competent cells and propagated. Recombinant plasmid was isolated and mutations were confirmed by sequencing (Eurofin, Bangalore and First Base Laboratories, Malaysia).

Table 4.1 List of primers used to prepare the mutants.

S.No.	Mutants	Sequence (5 '- 3 ')
1	EFBSH_C2A_for	GAAGGAGATATA CAT ATG GCG ACAGCA ATT ACT TAT G
	EFBSH_C2A_Rev	CAT AAG TAA TTG CTG TCG CCA TAT GTA TAT CTC CTT C
2	EFBSH_R16A_for	GAT CAT TAC TTT GGA GCG AAT TTT GAT TAT GAA A
	EFBSH_R16A_Rev	TTT CAT AAT CAA AAT TCG CTC CAA AGT AAT GAT C
3	EfBSH_D19N_for	CTT TGG AAG GAA TTT TAA CTA TGA AAT TTC TTA TAA TG
	EfBSH_D19N_Rev	CAT TAT AAG AAA TTT CAT AGT TAA AAT TCC TTC CAA AG
4	EfBSH_Y20W_for	CTTTGGAAGGAATTTTGATTGGGAAATTTCTTATAATG
	EfBSH_Y20W_Rev	CATTATAAGAAATTTCCCAATCAAAATTCCTTCCAAAG
5	EfBSH_N79Y_for	GGA ATG GCT GGA TTA TAT TTT TCA GGC TAT GC
	EfBSH_N79Y_Rev	GCA TAG CCT GAA AAA TAT AAT CCA GCC ATT CC'
6	EfBSH_Y79W_for	GAT TAG GAA TGG CTG GAT TAT GGT TTT CAG GCT ATG CAG
	EfBSH_Y79W_Rev	CTG CAT AGC CTG AAA ACC ATA ATC CAG CCA TTC CTA ATC
7	MG_peptide_For	GAA GGA GAT ATA CAT ATG GGC TGT ACA GCA ATT AC
	MG_peptide_Rev	GTA ATT GCT GTA CAG CCC ATA TGT ATA TCT CCT TC
8	MA_peptide_For	GAA GGA GAT ATA CAT ATG GCG TGT ACA GCA ATT AC
	MA_peptide_Rev	GTA ATT GCT GTA CAC GCC ATA TGT ATA TCT CCT TC
9	EfBSH_MLG_For	GAA GGA GAT ATA CAT ATG CTG GGC TGT ACA GCA ATT AC
	EfBSH_MLG_Rev	GTA ATT GCT GTA CAG CCC AGC ATA TGT ATA TCT CCT TC
10	MLG_TA_SS_for	CAT ATG CTG GGC TGT AGC AGC ATT ACT TAT GTA TCA AAA G
	MLG_TA_SS_Rev	CTT TTG ATA CAT AAG TAA TGC TGC TAC AGC CCA GCA TAT G

4.3.4 Expression and purification of mutant proteins

The site-directed mutants were transformed into BL21* (DE3) competent cell and inoculated in LB media containing 100 µg/ml of ampicillin and incubated at 37°C until an optical density (OD_{600}) of 0.4–0.6 was attained and then IPTG was added to the concentration of 0.5 mM for induction followed by shaking incubation (180 rpm) at 16°C overnight (16–18 h). Purification for all mutants were performed by His-select Ni–NTA (Quiagen) followed by Sephadex200 size exclusion chromatography (SEC 650, Bio-rad) as similar to that carried out for wild-type *Ef*BSH described in the chapter 2 and section 2.3.6.

4.3.4 BSH enzyme assay

The BSH activity was measured by estimating the released glycine upon hydrolysis of the bile acids by ninhydrine at 570 nm wavelength (Rossocha, et al., 2005, Kumar, et al., 2006). One unit of BSH activity is defined as the amount of enzyme that liberates 1 µmol of amino acid from the substrate per minute.

4.3.2 Molecular dynamics simulation

Dynamics and stability of each receptor-ligand complex is assessed by molecular dynamics (MD) simulations over time scale using Amber force field in Gromacs 4.5 (Pronk et al., 2013). The MD simulation gives insight into the probable (theoretical information) conformational changes and molecular interaction in the receptor-ligand complex in a time dependent manner. MD simulations are extensively used in the ligand binding, enzyme reactions, drug designing and understanding the folding and stability of proteins, (Sabonmatsu and Tung 2007). Lodola, et al., 2012 have suggested the probable catalytic mechanism of the CGH members catalysed reactions.

*Ef*BSH crystal structure was used for docking the substrate GCA using Glide programme of the Schrodinger suit. The best pose based on Glide scoring obtained from docking was used for the MD simulation. The system was energy minimized using steepest descent followed by conjugated gradient minimization (maximum force for convergence: 10 kJ/mol/nm). Further, system was equilibrated with NPT ensemble for 100 ps at 1 atm pressure and respective temperatures, using Parrinello-Rahman pressure coupling. Simulation was carried out in The Gromacs 4.5 with amber99sb force field (Pronk et al. 2013).

4.4 Results & Discussion

4.4.1 Structural analysis and *Ef*BSH inactive mutant for substrate complex

There is significant variation observed in the binding pocket of *Ef*BSH as compared to reported BSHs. The CGHs (BSHs and PVAs) binding pocket consisted of four binding pocket loops (loop1-4) among which loop2 and loop3 had insertions resulting in longer binding pocket loops (Chapter3). The loop3 showed different conformations in the BSH and PVAs. In PVAs the loop3 moved inward closer to the nucleophile cysteine residue and thus reduced the binding pocket to preferably accommodate a smaller molecule Pen V than bile acids.

Table 4.2 SiteMap quantitative estimation of binding site properties of CGH enzymes

PDB	Binding pocket volume ($A^{\circ 3}$)	Hydrophobicity	Hydrophilic	#Hydrophobic/ Hydrophilic	Reference
***Cp*BSH (2RLC)**	485.35,	0.910	1.04	0.87	Panigrahi, et al., (2014)
***Bl*BSH (2HEZ)**	535.77	0.950	1.07	0.88	Panigrahi, et al., (2014)
***Ef*BSH (4WL3)**	378.3	0.638	1.066	0.59	This study

#Hydrophilic and hydrophobic are a measure of hydrophilic and hydrophobic nature of the binding site; the higher ratio of hydrophobic/ hydrophilic values indicates more hydrophobic active site.

The binding pocket volume of the *Ef*BSH was calculated by *SiteMap* tool which was found to be 378.3 $Å^3$ smaller than other BSHs (Table 4.2). *Ef*BSH displayed less hydrophobic binding pocket compared to that of *Cp*BSH and *Bl*BSH. This variation in the active site properties like binding pocket volume and hydrophobicity /hydrophilic ratio might influence the substrate binding. *Ef*BSH do not show any activity toward Pen V substrate. However, *Cp*BSH is known to display very less activity towards the Pen V apart from its original substrate bile acids. Such promiscuity of the enzymes might arise from slight changes in the binding pocket properties of the protein. The C2A mutant was prepared for to obtain the E*f*BSH-substrate complex but no density was observed for substrates (GCA, Pen V) in the C2A crystal structures. The data collection and refinement statistics are detailed in table 4.3 & 4.4.

Table 4.3 X-ray diffraction and data collection of C2A mutant of *Ef*BSH (monoclinic and tetragonal form). Values in parentheses represent outer shell.

Space group	C2A_P2$_1$ (mutant) Monoclinic	C2A_P4$_1$22 (mutant) tetragonal
Temperature	100K	100K
X-ray source	ELLETRA, Italy	ELLETRA, Italy
Wavelength	0.9	1.0
Resolution (Å)	47.24 - 1.5	45.29-1.5
Unit cell parameter	a= 66.08 Å, b= 131.78 Å, c= 86.77 Å, β= 94.48˚	a= b= 91.18 Å, c= 156.5 Å, α= β= γ= 90.0˚
Molecules per asymmetric unit	4 (one tetramer)	2 (one dimer)
Matthews Coefficient (Å3 Da^{-1})	2.55	2.20
Solvent content (%)	51.7	44.0
Total No. of reflection	645213 (91286)	1269344 (37354)
No. of unique reflection	232716 (33162)	105625 (5015)
Multiplicity	2.8 (2.8)	12.0 (7.4)
Completeness (%)	98.7 (96.5)	99.9 (97.5)
Average I/σ(I)	10.7 (2.3)	24.1 (2.9)
Rsym or Rmerge (%)	5.6 (41.0)	5.8 (68.0)

Table 4.4 Refinement statistics of the C2A mutants of *Ef*BSH.

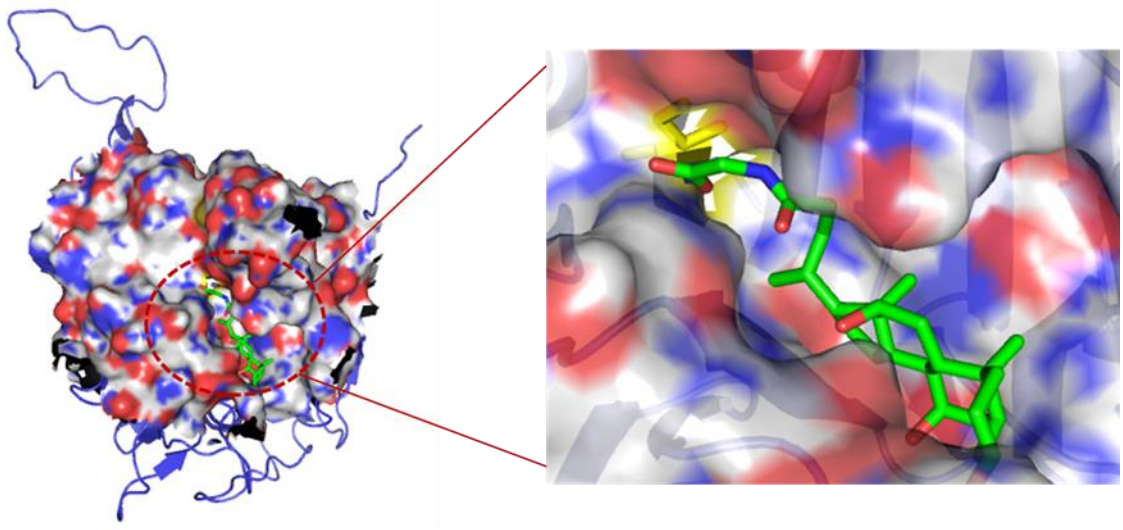
Parameters	C2A mutant (P2$_1$)	C2A mutant (P4$_1$22)
Rfactor (%)	24.0	20.05
Rfree (%)	26.3	23.35
Root Mean Squire Deviation (RMSD)		
RMS Bond Length (Å)	0.019	0.019
RMS Bond Angle (°)	1.903	1.963
Overall B (Isotropic) from Wilson Plot	23	18
Ramachandran Plot (% residues)		
Most favourable region (%)	96.3	97.4
Allowed region (%)	3.3	2.3
Generously allowed region	0.4	0.3
Disallowed region (%)	0	0

4.4.2 Substrate binding mode (GCA)

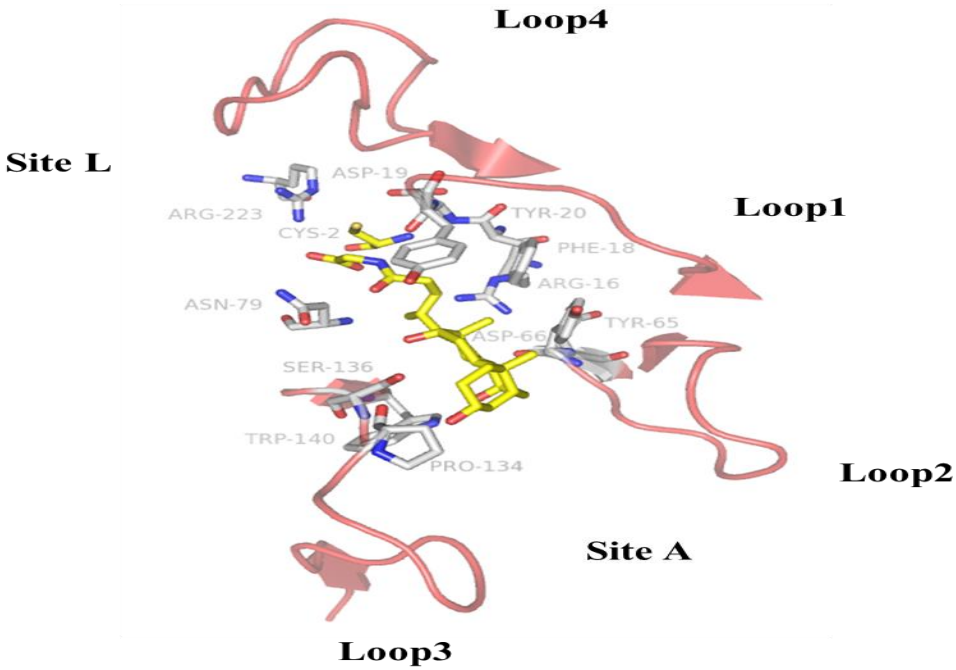
The Glychocholic acid (GCA) substrate molecule is amphipathic, synthesized in the liver and occurs as a sodium salt in the bile of mammals. The conjugation of the bile with glycine and taurine occurs in the liver in order to increase the solubility of the bile acid in the small intestine. The *Ef*BSH enzyme hydrolyses all six types of bile salts with various efficiencies. However, it shows preference for glyco-conjugated bile acids. The crystal structure of the *Cp*BSH (2RLC) with hydrolysed substrate has been reported by Rossocha, et al., 2005. Moreover, the probable catalysis mechanism for the CGHs members has been proposed by Lodola, et al., 2012, where nucleophilic attack by cysteine performed on sessile amide bond of the bile acid substrate releases the glycine as leaving group (Site*L*) and cholate as adduct group (Site*A*).

The chain A of the *Ef*BSH crystal structure was used for docking studies. The best pose for *Ef*BSH-GCA complex structure was obtained with free energy of binding (Glide score) of -7.59 kcal/mol (Fig.4.1a). The nucleophilic attack distance from nucleophile cysteine to sessile amide bond was calculated which was found to be 4.6 Å, however, for *Bl*BSH and *Cp*BSH this distance was observed to be 4.0 A˚ and 4.6 A˚, respectively (Panigrahi, et al., 2014), indicating that in *Bl*BSH, GCA sites are deep in the binding cavity toward the Nucleophile cysteine residue, compared to *Cp*BSH and *Ef*BSH. The detailed interactions of *Ef*BSH with the GCA are displayed in Fig.4.2. All binding loops (Loop1-4) constitute the binding pocket for GCA. GCA interacts with the binding pocket residues C1, R16, F18, D19, Y20, Y65, N79, P134, S136, N170 and R223 as shown in the Fig. 4.2.

GCA binding shows a directional preference for the amide bond orientation (CO–NH) in the direction from siteA to siteL. In the GCA after hydrolysis leaving group which leave first is Glycine and cholate forms the Enzyme-acyle adduct. As already described in the chapter 3, *Ef*BSH had aromatic residues (F18, Y20 and Y65) in binding pocket which were probably involved in substrate binding. When the *Ef*BSH-GCA docked complex was analysed it was observed that these aromatic residues interact with the cholate part of GCA. However, these residues were replaced in *Cp*BSH (M20, I22 and A68) and *Bl*BSH (L20, W22 and F67) as described in chapter3. *Bl*BSH shows high affinity towards GCA (Kumar, et al., 2006) which may be contributed by more hydrophobic residues W22 and F67 but has less catalytic efficiency compared to *Ef*BSH.

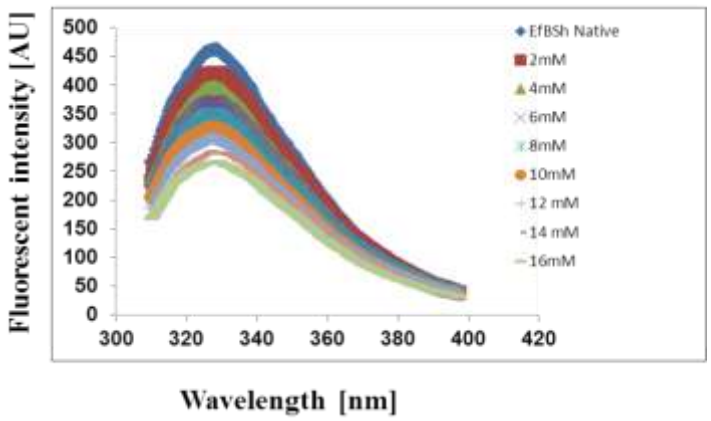


[a]

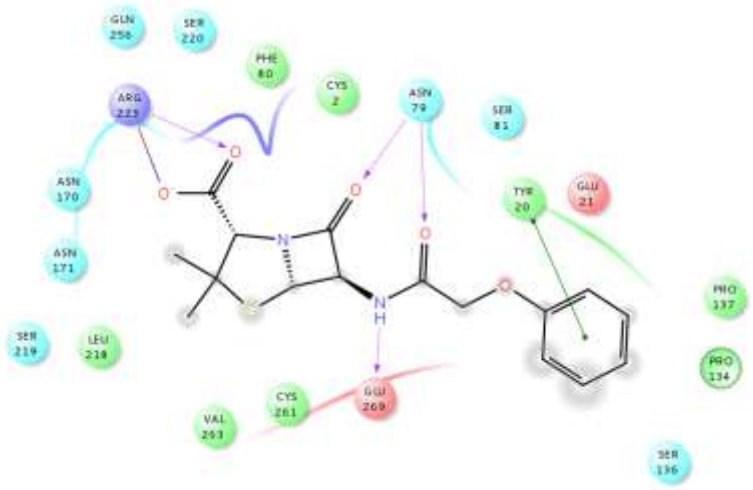


[b]

Figure 4.1 [a] Binding of the GCA substrate molecule in the *Ef*BSH binding pocket represented in surface view and stick model (ligand), [b] detailed interaction of the ligand (GCA) with the active site residues of the *Ef*BSH. The site of leaving group (siteL) and adduct group (siteA) are labelled.



[a]



[b]

Figure 4.2 [a]. Steady-state fluorescence intensity measurement upon Pen V binding. Decrease in fluorescence intensity (FI) upon treating *Ef*BSH with increasing Pen V concentration demonstrated the binding of Pen V. [b] Docked Pen V displaying the interactions with *Ef*BSH enzyme.

However, in the *Ef*BSH three aromatic residues at the binding pocket, interacting with the GCA probably forming good catalytic framework for the catalysis to happen (Fig. 4.2). We have prepared a mutation of Y20W assuming higher affinity similar to the *Bl*BSH and decreased turnover rate to demonstrate the role of the Y20 in *Ef*BSH. Interestingly, we have found that 25 % activity was reduced in Y20W mutant compared to wild-type *Ef*BSH, indicating Y20 has important role in the substrate binding which in turn influence the catalytic turnover rate of the enzyme.

Pen V substrate are hydrolysed by the PVAs enzyme, however, they are also known to inhibit their evolutionary related enzymes BSHs (Kumar, et al., 2006) in competitive fashion.

The Pen V binding was shown by the steady state fluorescence measurement in which increasing concentration of the Pen V demonstrated the reduction in fluorescence intensity (FI), indicating the binding of Pen V in the active site (Fig. 4.2a). Notably, *Ef*BSH do not showed any inhibition upon Pen V binding, instead it has enhanced the BSH activity described in chapter 2. Pen V was docked in the *Ef*BSH structure and best pose was obtained with Glide score of -6.6 Kcal/mol and 48.6 kcal/mol Glide e-model score, displayed in the ligand interaction diagram (Fig. 4.2b). It also has directional binding which is the reverse of the GCA in which Phenyl ring is oriented toward SiteA (Cholate in case of GCA) and 6-APA oriented toward the SiteL (Glycine in case of GCA). Tyr 20 has pi-pi interaction with the phenyl group of Pen V and residues N79, D169 and R223 have shown main chain interaction with Pen V molecule. Although it binds to the active site of *Ef*BSH but do not inhibit the BSH activity instead enhances the activity. The probable reason for this could be that Pen V molecule binds to one of the subunit of the native tetramer of *Ef*BSH and induce the binding of GCA substrate to another subunit, similar to allosteric mechanism.

4.4.3 Binding pocket mutations and their relative activity profile

The mutants of *Ef*BSH were prepared to understand the role of the important residues in the activity and stability of the protein given in table 4.5. The C2A mutation was introduced which is a null mutant to obtain the enzyme-substrate complex, but we couldn't get substrate bound crystal. Another important residue R16 (numbering according to *Ef*BSH) which is involved in the substrate catalysis mechanism (Rossocha, et al., 2005, Kumar et el., 2006) and also thought to be involved in autocatalysis processing of the pre-peptide or methionine in cholylglycine hydrolase family (Rossocha, et al., 2005). The R18A mutation in the *Bl*BSH and *Cp*BSH led complete loss of activity; however, *Ef*BSH_R16A mutant retains ~40 % of original activity (Table 4.5). Other mutants D19N and N79Y showed good expression of the protein and enzyme activity as well. In the PVAs aromatic residues (F, Y, W) at 81 positions are strictly conserved, however in the BSHs N (N79 in *Ef*BSH) is conserved. We tried to introduce the conserved residues of PVAs in the *Ef*BSH to gain the PVA activity. Therefore, we prepared N79Y, N79W and Y20W/N79W double mutants. The mutant N79Y showed

good expression but no PVA activity was detected, however 88 % BSH activity was retained. The N79W and Y20W/N79W mutants showed drastic decrease in expression and BSH activity and no PVA activity was detected. This indicates that there are other important residues apart from N79 (BSHs) and aromatic amino acids at position 81 (F, Y, W) in PVAs which are involved in distinguishing the substrates. The mutations highlighted were confirmed by DNA sequencing (Fig. 4.3).

Table 4.5 Relative % activity profile of the mutants. Wild-type *Ef*BSH activity was considered as 100%.

S. No.	Mutants	Relative % activity
1	C2A	Inactive
2	R16A	40
3	D19N	90.87
4	Y20W	74.54
5	N79W	17.74
6	N79Y	88.03
7	Y20W/N79W	Very unstable

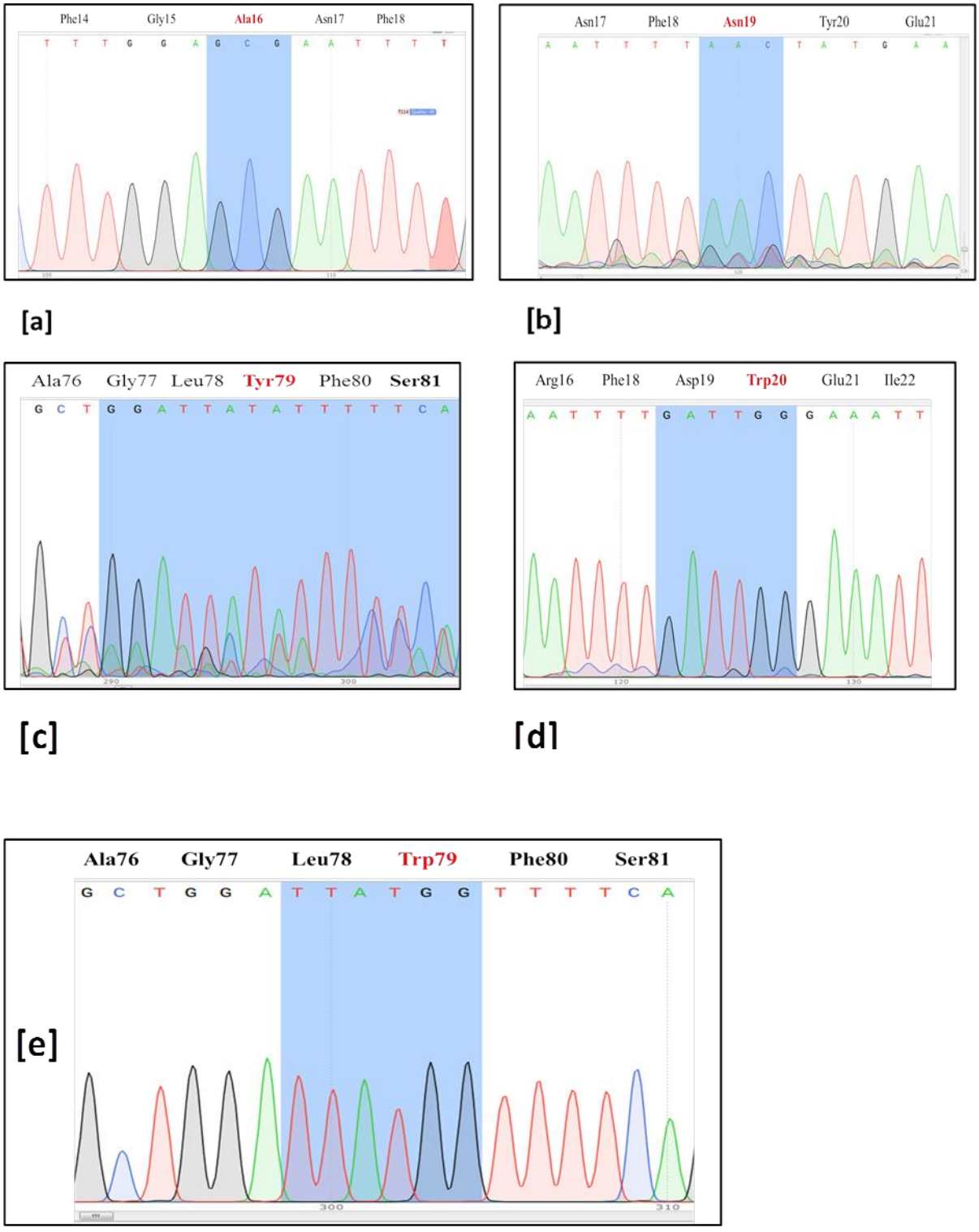
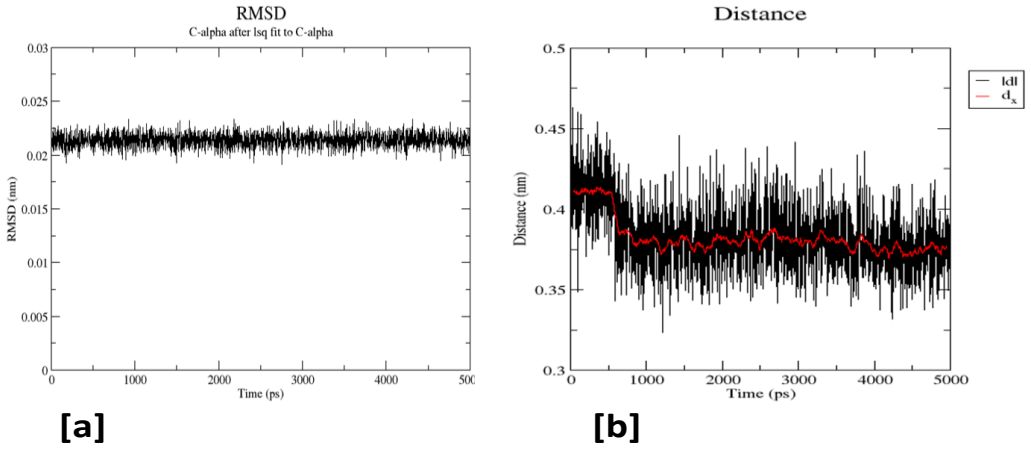
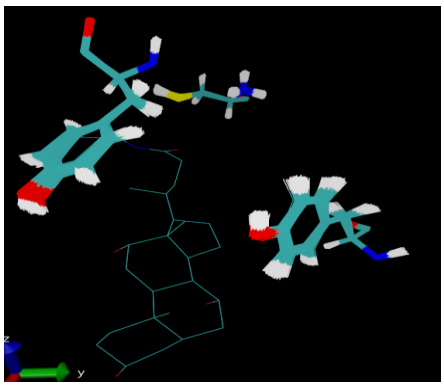


Figure 4.3 Sequencing chromatogram highlighting the mutations of active site mutants, [a] Arg16Ala mutant, [b] Asp18Asn mutant, [c] Asn79Tyr mutant, [d] Tyr20Trp mutant, [e] Asn79Trp mutant. All mutants were prepared to explore the importance of these residues in substrate binding and activity.

4.4.4 MD simulation and Polar complimentary for specificity and high catalytic efficiency

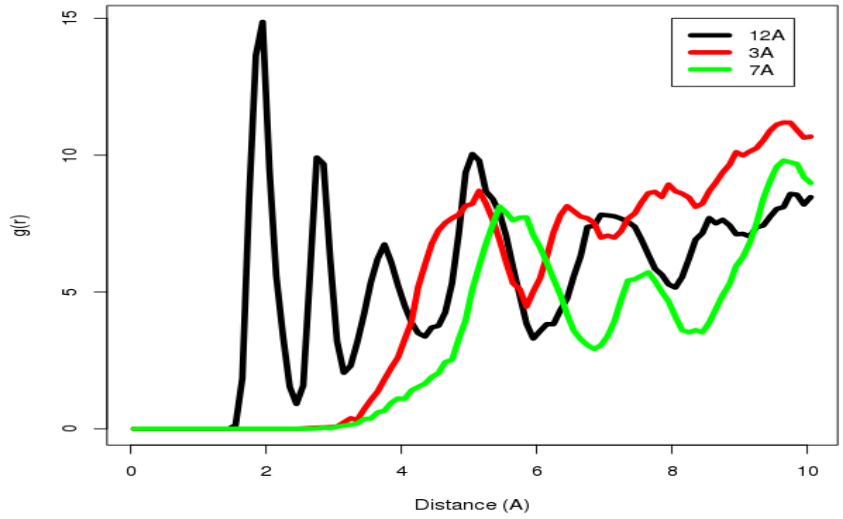
MD simulation was found to be stable which is measured on the basis of their Cα RMSD (average 0.05 Å; Fig.4.4a) values throughout the simulations (5 ns in case of *Ef*BSH). The substrate GCA was also found to be bound to the protein during the course of the simulation and which could be demonstrated by measuring the nucleophililc attack distance (average 0.05-0.1 Å, Fig.4.4b) from Cys2 to amide bond of the GCA. The fluctuations of Y20 and Y65 residues were also observed in the simulation but no significant fluctuation was observed (Fig.4.4c).



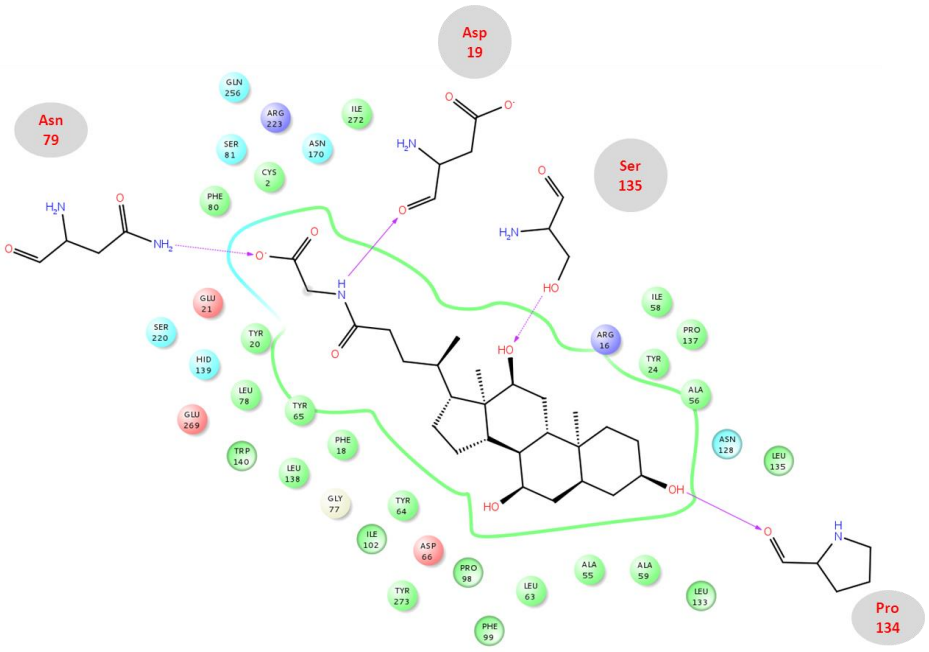


[c]

Figure 4.4 [a] RMSD of Cα atoms in wild-type *Ef*BSH in the course of MD simulation (5 ns), [b] Nucleophillic attack distance from nucleophile cysteine (Cys2) of *Ef*BSH to sessile amide bond of GCA substrate in the course of 5 ns time scale during MD simulation, [c] Interaction of Y20 and Y65 residues and their conformational changes during 5ns time scale with the GCA substrate in the active site.



[a]



[b]

Figure 4.5 [a] Radial distribution of receptor polar atoms around three hydroxyl groups of GCA (3a-, 7a- and 12a-OH) in *Ef*BSH. The y-axes correspond to the probability density of receptor polar atoms g(r) and the x-axis correspond to distance (in A^o). Polar complementarity at the respective hydroxyl group has been estimated at 5 A^o distance, [b] Ligand interaction diagram of the *Ef*BSH-GCA complex displaying the polar –OH groups interacting with various active site residues.

Polar complementarity at the binding pocket of CGHs has been assumed to provide specificity and better catalysis (Panigrahi, et al., 2014). As already mentioned GCA is an amphipathic planar molecule which has hydrophobic side having methyl group of the steroid ring and hydrophilic side consisting of 3-hydroxyl groups protruding out (3α-, 7α- and 12α-OH). The MD simulation studies provided the information regarding complementarity of these three hydroxyl groups. Radial distribution function was used to calculate the maximum probability density of receptor polar atoms g(r) within 5 Å of each hydroxyl group.

Table 4.6 Quantitative estimation of polar complementarities for the three hydroxyl groups of GCA molecule and percentage of times their involvement in hydrogen-bonding interactions (% H-bond) during molecular dynamics simulation of each enzyme–GCA complex.

S. No.	Property	Hydroxyl group	*Ef*BSH
1.	Maximum probability density within 5 A °(calculated by radial distribution function)	12-α-OH	14.86
		7-α-OH	3.96
		3-α-OH	8.15
2.	Hydrogen bonds (%)		
		12-α-OH	70.33 %
		7-α-OH	34 %
		3-α-OH	0.11 %

The values for the all three hydroxyl groups (3α-OH, 7α-OH, 12α-OH) in the *Ef*BSH-GCA complex were found to be very high compared to *Bl*BSH (3α-OH-1.62, 7α-OH-0.55 and 12α-OH 0.73, Panigrahi, et al., 2014), indicating high polar complementary of *Ef*BSH binding pocket (Table 4.6). The hydrogen bonds were observed maximum with 12-α-OH group (70%, Table 4.6). Significance of these hydroxyl groups in substrate binding is also supported by Batta et al., (1984). Furthermore, the hydrogen bonding interaction of these 3 hydroxyl groups with nearest receptor polar groups during dynamics was also assessed (Fig. 4.6 & Table 4.7). This analysis showed that hydrogen bonds existed in

*Ef*BSH-GCA complex including, ASN170ND2-GCA324O4, SER136OG-GCA324O1, ASN79ND2-GCA324N1 and GCA324O3-PRO134O were very stable during the course of the simulation (5 ns) (Fig. 4.6 & Table 4.7). Therefore, polar complementarity is an important factor which influences the GCA specificity and catalytic framework for the better hydrolysis among CGHs members.

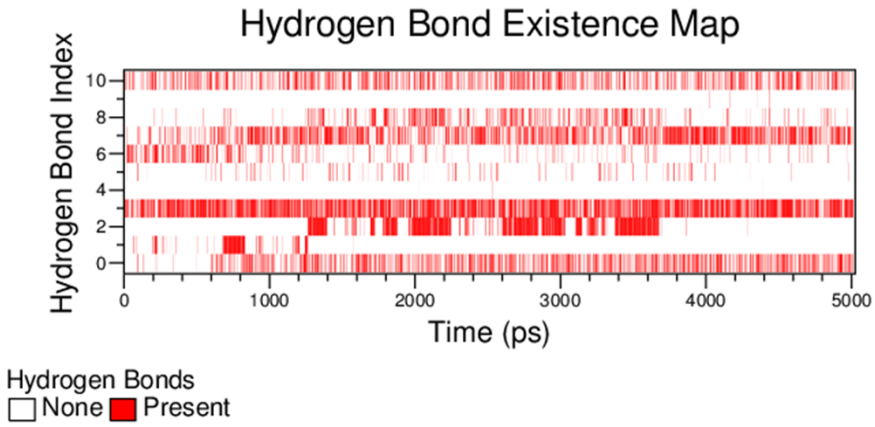


Figure 4.6 Hydrogen bonds index for the *Ef*BSH-GCA complex structure during the course of simulation for 5 ns time scale.

Table 4.7 Hydrogen bonds index and detailed interactions for the *Ef*BSH with GCA substrate in the 5 ns scale.

S.No.	Donor	Acceptor
0	ASN170ND2	GCA324O4
1	ASN170ND2	GCA324O5
2	ASN170ND2	GCA324O6
3	SER136OG	GCA324O1
4	SER136OG	GCA324O2
5	ASN79ND2	GCA324O4
6	ASN79ND2	GCA324O5
7	ASN79ND2	GCA324N1
8	ASN79ND2	GCA324O6
9	ASN79N	GCA324O4
10	GCA324O3	PRO134O

4.4.4 Autocatalytic processing of pre-peptide in CGH family

One of the major characteristics of the Ntn-hydrolase members is autocatalytic processing. CGH family, includes PVAs & BSHs which have "Cysteine" as N-terminal nucleophile residue as already described in chapter1. Phylogenetic analysis along with Binding Site Similarity (BSS) based classification CGH family members are categorized into 2 distinct groups: clusters-I (Gram-positive bacteria) and cluster-II (Gram-negative bacteria) showing divergent evolution (Panigrahi, et al., 2014). Both clusters differ with respect to the presence or absence of pre-peptides. Cluster-I possesses mostly methionine (80-90% sequences) present at N-terminal before nucleophile cysteine residue but very few organism also showed pre-peptide length in the range of 1-60 aa. However, in cluster-II pre-peptides are in the range of the 1-50 amino acids but most of the organisms have pre-peptide length in the range of 20-35 aa (Panigrahi, et al., 2014). Therefore, pre-peptide are prevalent in the Gram-negative compare to Gram-positive CGHs family members.

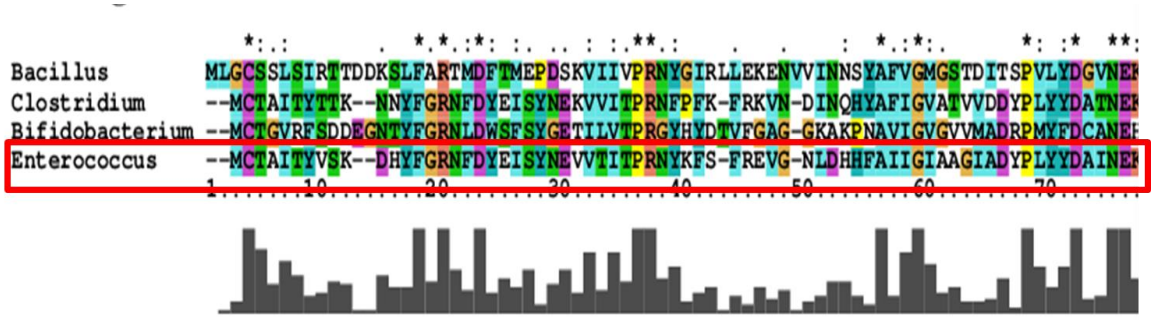


Figure 4.7 Sequence alignment showing the pre-peptide (MLG) in *Bacillus sphaericus* and 'methionine' in other BSHs (*Cp*BSH, *Bl*BSH, *Ef*BSH) before nucleophile cysteine residue.

Bacillus sphaericus, a Gram-positive bacterial PVA has been reported to have the "pre-peptide "MLG" (Fig.4.7)(Manish, et al., 2005) before nucleophile cysteine residue. The pre-peptide was reported to be auto catalytically processed and thus produced the mature and active form of the enzyme. However, BSHs from Gram-positive bacteria are reported to have the "Methionine" before nucleophile cysteine (Fig.4.7). So we tried to get some insight into autocatalytic processing of these enzymes using *Ef*BSH enzyme as example.

To understand the autocatalytic processing capability of *Ef*BSH we have prepared several mutants by inserting a pre-peptide before the nucleophile cysteine residue

(Fig.4.8). First we inserted the "MLG" pre-peptide in *Ef*BSH before the Cysteine2 and expressed in the BL21* (DE3) expression host cells. The protein preparation was performed but all protein had gone into the inclusion body (Fig.4.9).

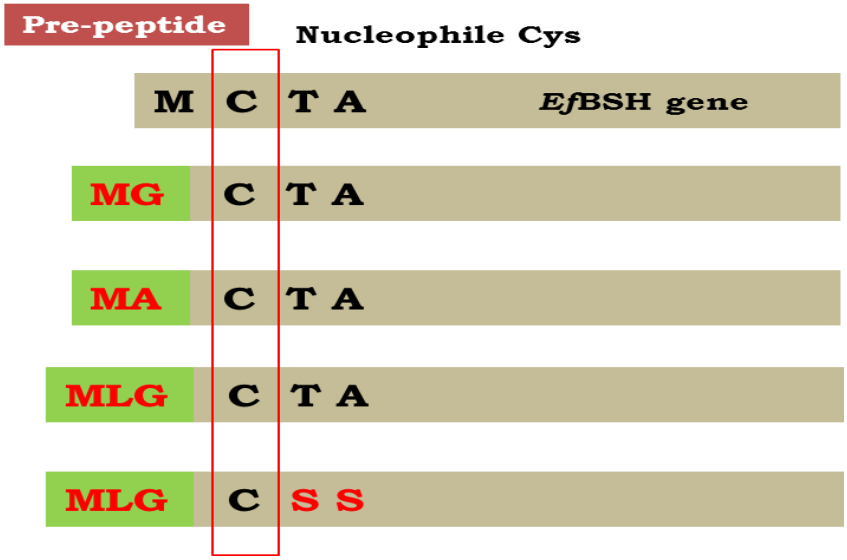


Figure 4.8 Schematic representations of pre-peptide insertion mutants of *Ef*BSH

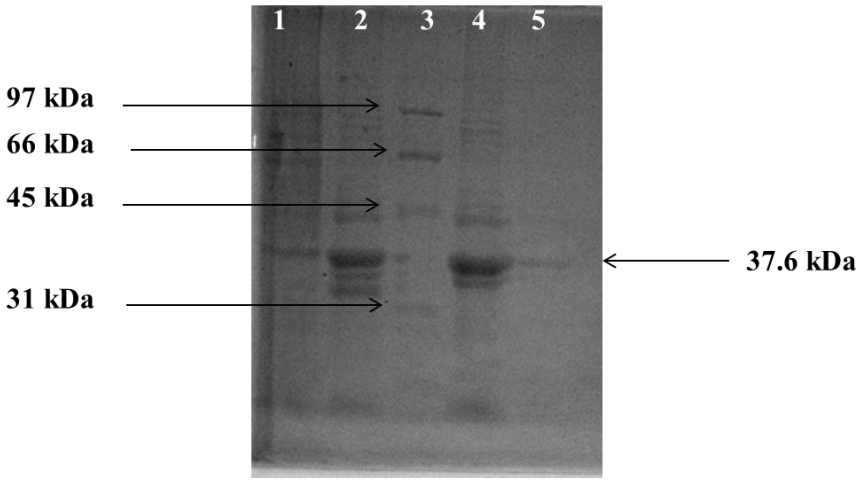
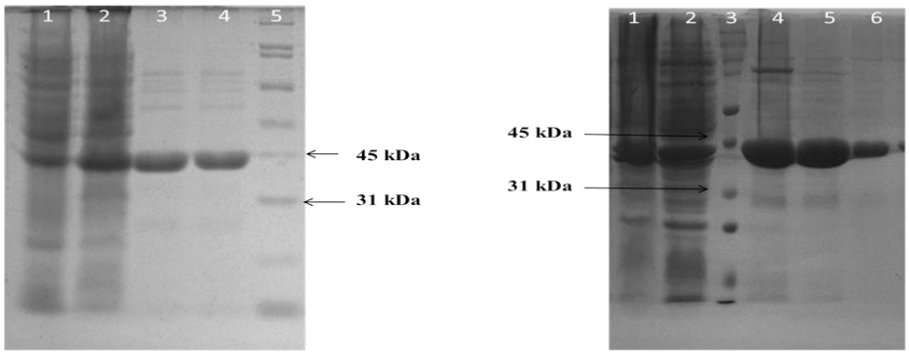


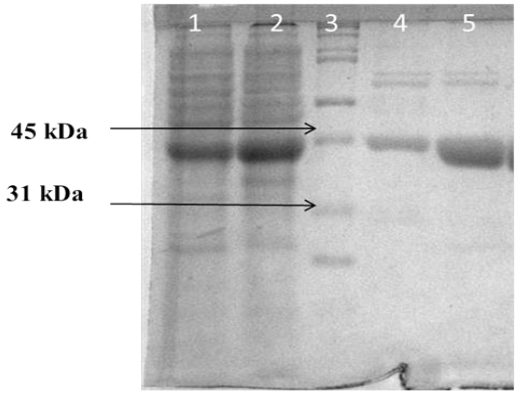
Figure 4.9 SDS-PAGE analysis of MLG-pre-peptide *Ef*BSH mutant purified from *E.coli* BL21 (DE3). Lane 1: Supernatant Lane 2: Pellet fraction or inclusion bodies, Lane 3: Bio-rad SDS-PAGE low range marker, Lane 4: flow through of on-column refolding (unfolded), Lane 5: Refolded protein (very less recovery and inactive).

On performing column refolding for the MLG-pre-peptide mutant, very less amount of the protein was obtained but no activity was detected indicating absence of processing of the MLG-pre-peptide. It probably could be also due to defect in the folding

of the enzyme due to presence of pre-peptide. Another mutant was prepared by reducing the length of pre-peptide (MLG to MG). This mutant showed good expression and was coming in soluble form and was fully active as native *Ef*BSH (Fig.4.10a).



[a] MG-Pre-peptide [b] MA-Pre-peptide



[c] MLG-TA-SS mutant

Figure 4.10 SDS-PAGE analysis of *Ef*BSH mutant purified from *E.coli* BL21 (DE3). [a] Lane 1: Supernatant, Lane 2: Pellet fraction or inclusion bodies, Lane 3: Bio-rad SDS-PAGE broad range marker, Lane 4-6: Ni-NTA eluted fractions (soluble and active). [b] Lane 1: Supernatant, Lane 2: Pellet fraction or inclusion bodies, Lane 3-4: Ni-NTA eluted fractions (soluble and active), Lane 5: Bio-rad SDS-PAGE broad range marker. [c] Lane 1: Supernatant, Lane 2: Pellet fraction or inclusion bodies, Lane 3: Bio-rad SDS-PAGE broad range marker, Lane 4-5: Ni-NTA eluted fractions (soluble and active).

To check the specificity at the position before nucleophile cysteine Ala was introduced in place of glycine (MG-pre-peptide) to obtain MA-pre-peptide mutant. Autocatalysis processing was observed in this MA-pre-peptide also like MG-pre-peptide

and produced fully active and mature *Ef*BSH (Fig.4.10b). When we looked back to the sequence of the *Bacillus Sphaericus* to understand why *Ef*BSH could not process the MLG-pre-peptide, we observed that after nucleophile cysteine residue it had Ser-Ser amino acids (MLG_CSS, *Bacillus Sphaericus*) which might give the flexibility for pre-peptide to accommodate it in the binding pocket and poise it to induce the autocatalytic processing. In the MLG-pre-peptide (MLG_CTA) mutant of *Ef*BSH Thr3-Ala4 to Ser3-Ser4 (Sequence No. is according to native *Ef*BSH) double mutation was introduced to generate the MLG_TA-SS mutant (MLG_CSS mutant of *Ef*BSH) (Fig.4.8). Interestingly, MLG_CSS mutant was expressed and protein purification was achieved (Fig.4.10c). The mutant produced fully matured and active *Ef*BSH protein, emphasizing the role of the Ser residue after nucleophile cysteine. The activity profile of the all the mutants is detailed in the Table 4.8. All the clones were confirmed by DNA sequencing (Fig. 4.11a-d) from 1st base sequencing company, Malaysia.

Table 4.8 Expression and activity profile of the pre-peptide mutants of *Ef*BSH.

S.No.	Pre-peptide mutants of *Ef*BSH	Activity	Expression Profile	Remarks
1	**MA-Peptide**	active	soluble	Processed and active
2	**MG_Peptide**	active	soluble	Processed and active
3	**MLG-Peptide**	inactive	Inclusion bodies	Un-processed and inactive
4	**MLG_TA/SS (MLGCSS)**	active	soluble	Processed and active

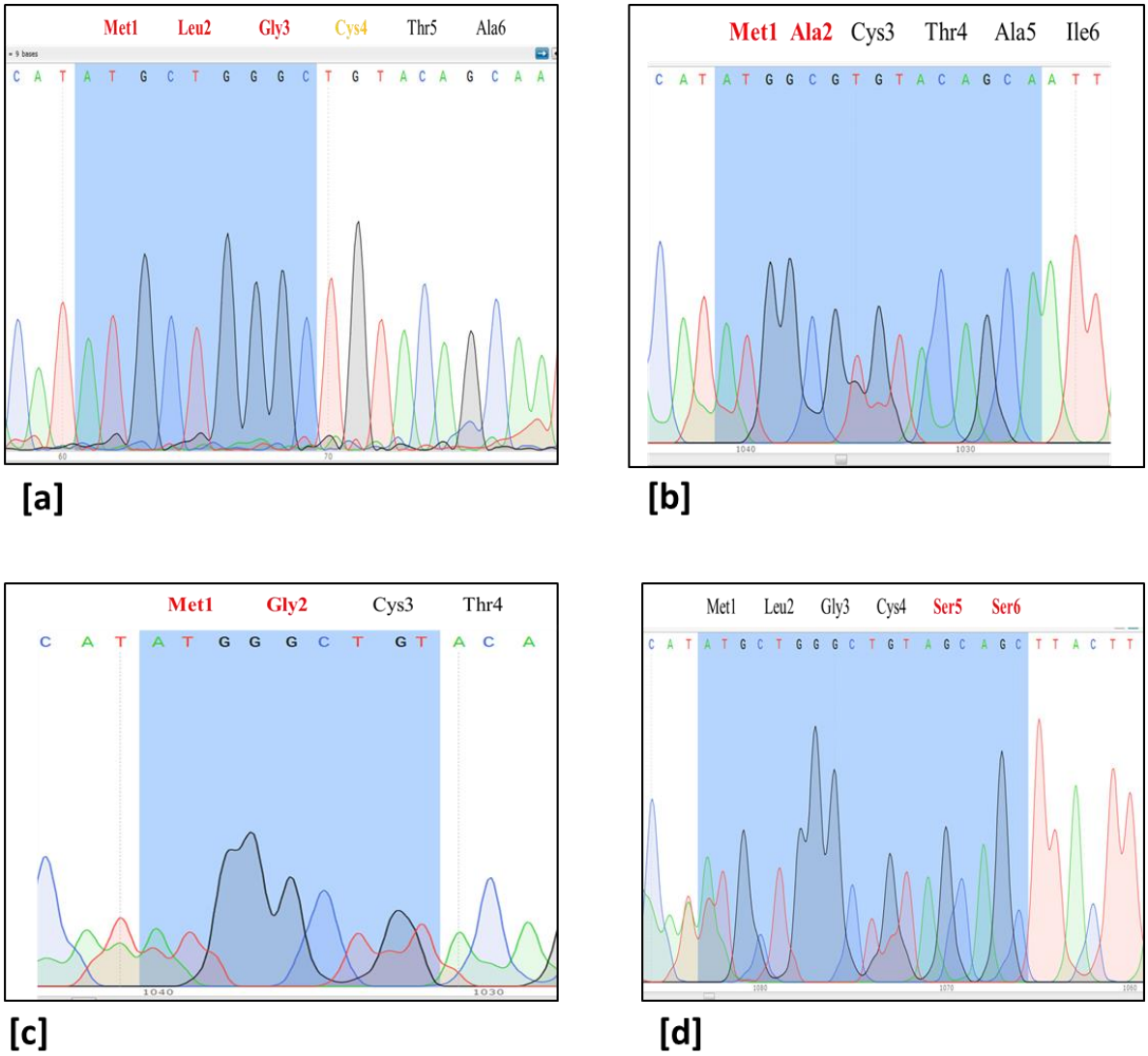


Figure 4.11 Sequencing chromatogram highlighting the pre-peptide insertion mutants of the *Ef*BSH. [a] MLG-pre-peptide mutant, [b] MA-pre-peptide mutant, [c] MG-pre-peptide mutant, [d] MLG-TA_SS (MLGCSS) pre-peptide mutant. All the sequencing were got done from 1st base sequencing company, Malaysia

4.5 Summary of the Chapter:

In this chapter we have presented the modes of substrate binding (GCA & Pen V) and probable basis for the specificity and better catalytic activity of *Ef*BSH towards bile acids (GCA) and no activity towards Pen V. The GCA binding analysis was carried out using the docking and MD simulation studies. Binding pocket analysis was done by SitMap program which had demonstrated that *Ef*BSH substrate binding pocket was more polar than other BSHs (*Bl*BSH and *Cp*BSH) which very well correlated with the high polar complementarity of all the hydroxyl groups (3α-OH, 7α-OH, 12α-OH) of GCA substrate in *Ef*BSH-GCA complex. The high polar complementarity among the CGHs members has thus been assumed one of the crucial factors which influence the substrate specificity and catalytic activity.

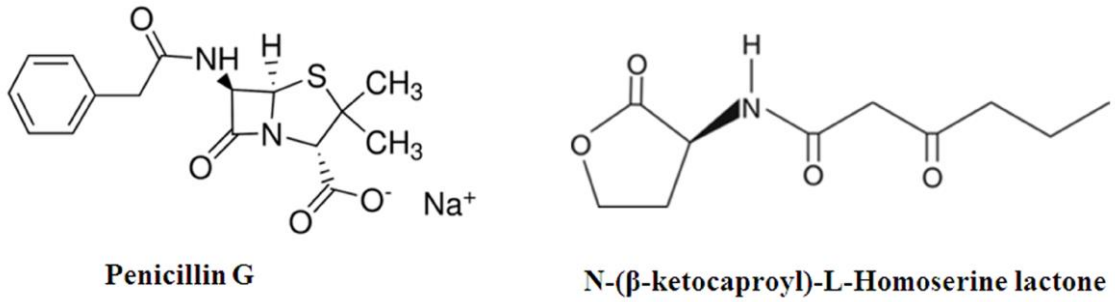
The active site mutants of the *Ef*BHS were prepared to explore the contribution of important residues in substrate binding and catalysis. In addition, autocatalytic processing of the pre-peptide was also studied in *Ef*BSH to understand the processing nuances in CGH family enzymes. Four mutants were prepared in which two or three amino acids were introduced at the N-terminus before the nucleophile residue cysteine and assessed the processing of the pre-peptide removal. MG and MA-pre-peptide mutants were processed fully and produced the mature and active form of the *Ef*BSH, however, MLG-pre-peptide expressed and went into the inclusion body. However the mutant MLG_TA-SS (MLG_CSS mutant of *Ef*BSH) in which Ser-Ser flexible linker was introduced in the place Thr-Ala demonstrated complete processing of the pre-peptide “MLG”, emphasizing the important role of the Ser-Ser residue in the autocatalytic processing of *Bsp*BSH and such enzymes of CGH family.

Chapter-5

"Structure mediation in substrate binding and post-translational processing of penicillin acylases: Information from mutant structures of *Kluyvera citrophila* penicillin G acylase"

5.1 Introduction

Penicillin acylases (PAs) are found in a variety of microorganisms, including bacteria, yeast and fungi (Vandamme & Voets, 1974) presumably to combat antibiotics and to metabolise aromatic compounds in the surrounding environment (Prieto et al., 1996). PAs are categorized into three sub-families viz, penicillin G acylase, penicillin V acylase and ampicillin amidase (Hamilton-Miller, 1966; Valle et al., 1991). Penicillin G acylase (PGA, E.C. 3.5.1.11) is an industrially important enzyme, which belongs to the Ntn-hydrolase (N-terminal nucleophile hydrolase) structural superfamily, sharing the αββα characteristic fold that forms the core structure of the family (Brannigan et al., 1995). PGA is widely employed in the pharmaceutical industry for the production of 6-APA (6-aminopenicillanic acid) from penicillin G substrate (Fig. 5.1), which is used as a precursor nucleus for manufacturing a range of semi-synthetic antibiotics (Bruggink et al., 1998; Nayler, 1991). Besides the production of semi-synthetic antibiotics, penicillin G acylases have also been deployed in peptide synthesis and in the resolution of racemic mixtures (Arroyo *et al.*, 2003; Waldmann, 1988). The physiological role of PGA remains speculative but the *pac* gene is located downstream of genes encoding proteins of the 4-hydroxyphenylacetic acid degradative pathway of *Escherichia coli* (Merino et al., 1992; Prieto et al., 1996). Therefore, it has been proposed that PGA may have a role in the assimilation of aromatic compounds as carbon sources and their degradation (Avinash *et al.*, 2014). PGA has a single amino acid catalytic centre, which is the N-terminal serine of the β-subunit, which also acts as a nucleophile in the cleavage of the Thr-Ser (TS) bond of the precursor to remove the spacer peptide and generate the active mature enzyme (Guan et al., 1998; Kasche et al., 1999). The fully processed mature PGA heterodimer comprises an α-chain of 23.2kDa and a β-chain of 63kDa (Duggleby et al., 1995; McVey et al., 2001). The crystal structures of PGA from *E. coli*. (*Ec*PGA) in processed, unprocessed and substrate bound forms have been reported at various resolutions (Alkema et al., 2000; Hewitt et al., 2000) .



Penicillin G **N-(β-ketocaproyl)-L-Homoserine lactone**

Figure 5.1 The drawing showing the two substrates pen G and AHL for which wild-type *Kc*PGA shows activity.

Penicillin G acylase from *Kluyvera citrophila* (*Kc*PGA) undergo post-translational processing to reach active form. *Kc*PGA is known for its resilience towards harsh conditions of pH, temperature and organic solvents. *Kc*PGA resembles *Ec*PGA in its post-translational processing and the two enzymes exhibit high sequence similarity, which is above 80%. *Kc*PGA is of interest as a consequence of its resilience towards harsh conditions like pH, temperature and organic solvents (Varshney et al., 2013) the property particularly important for pharmaceutical industries in semi-synthetic antibiotic synthesis. Another property that distinguishes it from *Ec*PGA is its capacity to hydrolyse N-Acyl Homoserine Lactone (Fig. 5.1) (Mukherji et al., 2014) .

This chapter details the 3D structure determination of the *K. citrophila* Serβ1Cys processed mutant and comparison with other PGA. The *K. citrophila pga* gene encodes a signal peptide, the α- subunit, a spacer peptide and the β-subunit. It is expressed as a ~93 kDa inactive precursor in the cytoplasm which is directed to the periplasmic space by the 26 amino acid residue signal peptide. Subsequently, the spacer peptide is cleaved off by autocatalytic processing to generate the mature and active *Kc*PGA. The sequence encoding *Kc*PGA was cloned into pET26b (+) vector and two mutants of *Kc*PGA were prepared by mutating the nucleophilic residue Serβ1 to Cys (Serβ1Cys) and Gly (Serβ1Gly). The Serβ1Gly mutant remained in the unprocessed precursor form (Varshney et al., 2013) while Serβ1Cys mutant, although fully processed, remained inactive. The crystal structure of the unprocessed *Kc*PGA Serβ1Gly mutant was reported previously (Varshney et al., 2013). In the present study we compare the refined structures of the processed inactive cysteine mutant Serβ1Cys and the unprocessed glycine mutant

Serβ1Gly at 2.8 Å and 2.5 Å, respectively. We have also compared these structures with processed, unprocessed and substrate bound structures of *Ec*PGA, in order to gain insights into the underlying mechanisms and dynamics of intra-molecular autocatalytic processing. Modelling studies helped to elucidate the role of structure in activity of *Kc*PGA and not *Ec*PGA towards homoserine lactone, a molecule identified in quorum sensing.

5.2 Materials

All chemicals Tri-Cl buffer, imidazole, ammonium sulphate, isopropanol, PEG-3350, Ni-NTA were procured from Sigma (Sigma, USA), Merck (Merck, USA). The materials such as media components (Hi-Media), Thermal cycler (AB sciences, USA), plasmid isolation kits (Qiagen, Germany), Quick Change site-directed-mutagenesis kit (Stratagene, USA), restriction enzymes and ligases (New England Biolabs, UK) were also used in the experiments. Oligonucleotide DNA (Eurofin, India), pET 26b+ vector and Nova Blue competent cells were from Novagen (Novagen, USA).

5.3 Methodology

5.3.1 Site-directed mutagenesis

Wild-type clone pET26b-*Kc*PGA was used as template for preparation of the *Kc*PGA_S1G (Serβ1Gly) and *Kc*PGA_S1C (Serβ1Cys) clones using Quick change site-directed mutagenesis kit (Varshney *et al.*, 2013). Primer pairs used to construct the Serβ1Cys mutant are, forward primer 5'-CTACCCGACCACT**T**GCAATATGTGGGTG-3, and reverse primer 5'-CACCCACATATTGC**A**AGTGGTCGGGTAG-3'. Cloned plasmid was transformed into Nova blue competent cells and propagated. Recombinant plasmid was isolated and mutation was confirmed by sequencing.

5.3.2 Overexpression and purification of mutant protein

Recombinant plasmid of wild-type *Kc*PGA and Serβ1Cys mutant was transformed into the expression host BL21 (DE3) pLys cells. Bacterial cells were cultured in LB (Luria broth) media containing 34 μg/ml kanamycin and chloramphenicol antibiotics at temperature 37^{o}C in shaking incubator (Steelmet, India) at 180 rpm, till OD_{600nm} reached

0.5-0.6. After 3 hour incubation at 37^{o}C IPTG (Isopropyl β-D-1-thiogalactopyranoside, Merck, USA) induction was given to the final concentration of 1 mM and 0.5 mM, respectively and culture was shifted to a lower temperature of 16^{o}C overnight (16-18 hrs). Bacterial cells were harvested by centrifugation at 5000 rpm at 4^{o}C for 15 min. Cells pellet was re-suspended in the lysis buffer composed of 25 mM Tris-Cl pH 8.0, 500 mM NaCl, 0.1% Triton X, 1 mM DTT, 20 mM imidazole, 10 mM $MgCl_2$ and sonicated (E-Squire Biotech, India). Cell lysate was centrifuged at 12000 rpm at 4^{o}C for 30 minutes. Clear supernatant was collected and allowed to bind on Ni-NTA affinity resin pre-equilibrated with buffer A (25 mM Tris-Cl pH 8.0, 500 mM NaCl, 20 mM imidazole) and washed with buffer A containing 60 mM imidazole. Elution was performed using buffer A containing 500 mM imidazole. Eluted fractions were checked on 12% SDS-PAGE. Fractions containing protein of interest were pooled and dialysed against 25 mM Tris-Cl pH 8.0, 50 mM NaCl buffer overnight at 10^{o}C. Dialysed sample was concentrated by centricon (Millipore, USA) and size exclusion chromatography (Sephacryl 200) was performed with 20 mM Tris-Cl pH 7.5 buffer containing 150 mM NaCl to purify the protein to homogeneity and purity was checked by SDS-PAGE. Fractions from gel filtration containing mutant protein were pooled and dialysed against 50 mM NaCl and 50 mM HEPES buffer pH 7.5 overnight at 10^{o}C. Expression and purification protocol for cloned Serβ1Gly was described in detail by Varshney et. al., (Varshney *et al.*, 2013).

5.3.3 Crystallization of *Kc*PGA_Serβ1Cys processed mutant

The Serβ1Cys mutant (processed form) was purified till homogeneity and buffer was exchanged with 50 mM HEPES buffer pH 7.5 and 50 mM NaCl. Protein was concentrated upto 19.2 mg/ml using Amicon Ultra-50 centrifugation filter device (30 kDa cutoff, Millipore, USA). Crystallization trials were conducted at temperature 20^{o}C with 1:1 ratio of protein to precipitant, exploiting commercially available screens from Hampton research (Hampton Research, USA) and Molecular dimension (MD, UK). Self-prepared additives were also screened to improve the crystal quality. Adding isopropyl alcohol in the crystallization solution improved the diffraction quality.

5.3.4 X-Ray diffraction and data collection for *Kc*PGA_Serβ1Cys processed mutant

Crystallization and data collection for Serβ1Gly mutant was reported by Varshney et. al., (2013). The Serβ1Cys mutant crystals were mounted on cryo-loop (Hampton research, USA) and frozen in liquid nitrogen using 30% glycerol for cryoprotection. Data were collected at 100K on Stanford Synchrotron Radiation Lightsource (SSRL, 12-2) using wavelength 0.9795 Å up to a resolution 2.8 Å.

5.3.5 Molecular replacement & structure Refinement

Automated model building was performed for Serβ1Cys with buccaneer programme of CCP4i suit, whereas *AutoMR* program from *PHENIX* (Adams et al., 2010; McCoy, 2007) was used in the case of Serβ1Gly mutant. Refinement was carried out using refmac5 programme of CCP4i (Collaborative, 1994). Cycles of model fitting and refinement as well as inputting of water molecules in difference electron density were carried out using Refmac5 in conjunction with COOT 7.2 graphics programme (Emsley & Cowtan, 2004) in CCP4 package.

5.3.6 Structural analysis

After final refinement, the stereochemistry and the geometry of the models were checked using PROCHECK programme (Laskowski et al., 1993). Receptor ligand docking was carried out using Glide (Version 5.8, Schrodinger, LLC, New York, NY, 2012). All structural figures were prepared using PyMol (The PyMOL Molecular Graphics System, Version 1.7.4, Schrödinger, LLC) or CCP4MG (McNicholas et al., 2011).

5.3.7 Molecular docking Studies

The 3D conformation of 3-oxo-hexanoyl homoserine lactone (AHL), a quorum signaling molecule with PubChem CID 688505 was used in the docking study (Bolton, 2008). LigPrep of Schrodinger suit was used for the partial atomic charges of each ligand atom which was determined from OPLS_2005 all-atom force field. The program Glide was used for docking the ligand in the binding site of receptor. Grid based ligand docking was performed in which center was considered from Cysβ1 residue and binding site

around the Cysβ1 was defined as a grid box. After receptor grid generation, ligand was docked flexibly using Glide's XP (extra precision). Free energy of ligand binding is roughly calculated by a scoring function called GlideScore.

5.4 Results & Discussion

5.4.1 Overexpression and purification of Serβ1Cys processed mutant

Overexpression of the fully processed Serβ1Cys mutant was achieved in *E.coli* strain pLys (DE3) by 1 mM and 0.5 mM IPTG induction for overnight (16-18 hour) at 16^{o}C. Both proteins were purified to homogeneity by Ni-NTA affinity column chromatography followed by gel filtration column chromatography (Sephacryl 200). The purity of the protein was assessed by SDS-PAGE analysis where the two bands corresponding to the α-subunit: ~23 kDa and the β-subunit: ~63 kDa, respectively were observed (Fig. 5.2).

For wild-type *Kc*PGA the fractions were assayed for PGA activity using 2% (w/v) penicillin G (Fig. 5.1) (Varshney, et al., 2013) and reaction was carried out at pH 7.5 and 50^{o}C. The product 6-aminopenicillanic acid (6-APA) was estimated using p-dimethyl aminobenzaldehyde (PDAB) and measuring OD at 415nm, a modified method described by Shewale et al., (Shewale et al., 1987), which was based on colorimetric assay proposed by Bombstein and Evans-(Bomstein & Evans, 1965). The enzyme was found to be active on Pen G substrate over a pH range 7.0 to 8.0 with maximum at pH 7.5. The enzyme showed stability over a broad pH range 4.0-10.0 (data not shown) and also reported by (Wen *et al.*, 2005). It retained more than 80% of its activity at pH 10.0 even after 8 hours of incubation at room temperature. The purified enzyme was inactivated above 50°C and completely lost activity at 60°C within 30 min of incubation. Values of V_{max} and K_M calculated for substrate pen G deduced from a Lineweaver-Burk plot were found to be 62.5 µmole/min/mg and 2.7 mM, respectively while K_M value for *E. coli* PGA reported was 7 µmole (Alkema et al., 2000).

5.4.2 Crystallization, crystal improvement and data collection

The purified protein of the Serβ1Cys mutant was subjected to crystallization using various commercial crystallization screens, PEG ions, Crystal screens I & II, JCSG and

Memfac. The protein concentration of 15-20 mg/ was used. Both micro-batch and hanging-drop method was used to obtain the diffraction quality crystals. However, no diffraction was observed with crystal obtained in the micro-batch method when they were tested at home source by capillary mounting method at room temperature (Fig. 5.3 a-d). After, rigorous screening, Serβ1Cys mutant was crystallized in a solution containing 50 mM sodium cacodylate buffer pH 5.6, 16% PEG 4000, 50 mM NaCl, 0.5 M NaSCN with 19.2 mg/ml concentration by hanging drop method (Fig.5.3e). The crystals appeared in 6-8 weeks but their diffraction was very feeble. Therefore, additives were tried for the improvement of crystal quality. Out of various additives screened 10% isopropyl alcohol was found suitable which had improved the diffraction quality of the crystal. The crystals diffracted up to 2.8 Å (Fig. 5.3f), data were integrated in the XDS (Kabsch, 2010).

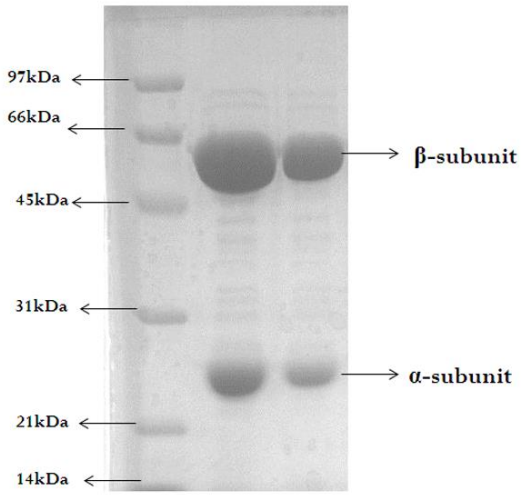


Figure 5.2 SDS-PAGE analysis of recombinant KcPGA_Serβ1Cys mutant purified from *E. coli* BL21 (DE3)/pLyse. Lane 1: Bio-rad SDS-PAGE low range marker. Lane 2: *Kc*PGA_Serβ1Cys purified form Ni-NTA affinity chromatography and Lane 3: purified protein after Ni-NTA followed by gel filtration (Sephacryl 200). Bands corresponds to α (23.6kDa) and β (63kDa) subunits.

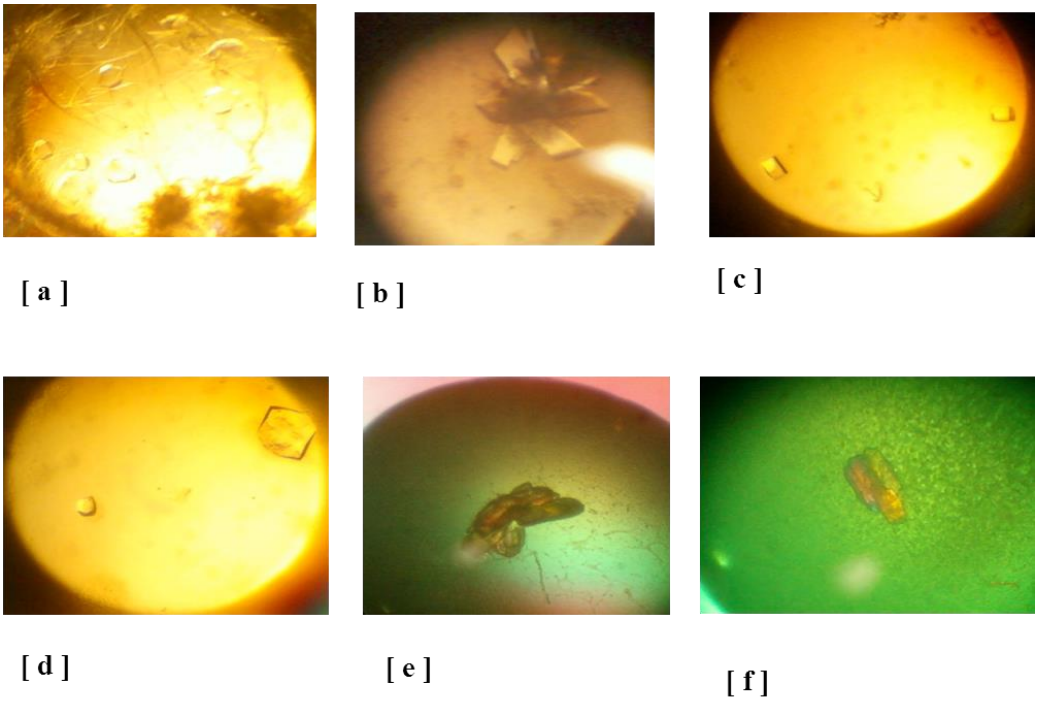


Figure 5.3 Screening of the best condition and improvement of the crystal quality. Crystal obtained in several conditions [a] 0.1 M HEPES pH 7.5, 0.1 M AS, 0.5 M sodium phosphate dibasic dehydrate, potassium phosphate dibasic (7.5 mg/ml) [b] PEG 4K (20%), 50 Mm sodium cacodylate pH 5.6, 0.5 M NaSCN (19.5 mg/ml), [c] 0.1 M ammonium phosphate dibasic, 0.1 M Tris-Cl pH 8.5, 0.5 M sodium phosphate dibasic dehydrate, potassium phosphate dibasic, (7.5 mg/ml), [d] 0.1 M Tris-Cl pH 8.5, 0.1 M sodium acetate trihydrate (7.5 mg/ml), [e] 50mM sodium cacodylate pH 5.6, 0.5 M NaSCN (19.5 mg/ml), 16% PEG 4K, [f] 50 mM sodium cacodylate pH 5.6, 0.5 M NaSCN (19.5 mg/ml), 20% PEG 4K, 10% Isopropyl alcohol.

The Serβ1Cys processed mutant crystallized in space group P1 with four molecules in the asymmetric unit. The estimated Matthew's coefficient was 2.3 $Å^3$ Da^{-1} with a solvent content of 46 % (Matthews, 1968). The summary of data collection is presented in Table 1.0 for Serβ1Cys mutant. However, details of crystallization conditions and data collection on Serβ1Gly mutant crystals were reported by Varshney et. al., (2013).

5.4.3 *Kc*PGA_Serβ1Cys processed mutant structure solution and refinement

The crystal structures of the processed mutant Serβ1Cys and the unprocessed mutant Serβ1Gly were refined at 2.8 Å and 2.5 Å resolutions, respectively. The data collection statistics for processed Serβ1Cys mutant is presented in the Table 5.1 and refinement statistics are in Table 5.2. However, details of crystallization conditions and data collection statistics for unprocessed Serβ1Gly mutant is reported by (Varshney et al., 2013). The refinement statistics for Serβ1Gly mutant are summarized in the Table 5.2 Initial phases of Serβ1Cys mutant reflections were assigned by using molecular replacement (Phaser-MR) using *Ec*PGA processed mutant (1PNK) as search model (Duggleby *et al.*, 1995) whereas a slow processing precursor of *Ec*PGA (1E3A) was used as model (Hewitt *et al.*, 2000) for phasing Serβ1Gly mutant.

The best model with highest Z-score value was obtained from Phaser-MR and used for the Autobuild to obtain the structure of the Serβ1Cys mutant structure. Structure was refined in Refmac5 in conjunction with Coot 7.0 program for several cycles till the desired R_{free} and R_{factor} could be achieved (Table 2.0). The Ramachandran plot analysis in the PROCHECK was also performed and no residues in the disallowed region were present (Table. 5.2 & Fig.5.4).

5.4.4 Overall Structure of *Kc*PGA_Serβ1Cys and *Kc*PGA_Serβ1Gly mutants

Penicillin G acylase belongs to Ser-Ntn-hydrolase family (Brannigan et al., 1995) of proteins which undergo post-translational maturation to generate the functionally active enzymes (Kasche et al., 1999). This auto-catalytic maturation process involves the removal of spacer peptide between the α and β-subunits of the pro-peptide remaining after the elimination of the signal peptide that directs the transport of the protein to the periplasm. In *Kc*PGA the N-terminal β1serine is a crucial catalytic residue in both enzyme catalysis and autocatalytic removal of spacer peptide. To evaluate the significance of the N-terminal serine residue in activity and to achieve insight into the autocatalytic processing mechanism of *Kc*PGA, two mutants Serβ1Cys and Serβ1Gly were prepared by site-directed mutagenesis and their structures were determined.

Table 5.1 Data collection statistics table for processed (Serβ1Cys, PDB Code: 4PEL) form of *Kc*PGA. Same for Serβ1Gly mutant was reported by Varshney et. al., (2013).

Space group	P1
Temperature	100K
X-ray source	SSRL (BL12-2)
Wavelength	0.9795
Resolution (Å)	75.75 - 2.8 (2.80)
Unit cell parameters (Å and °)	a=96.26, b=96.26, c=100.21, α=69.27, β=69.27, γ=76.77
No. of Molecules in Asymmetric unit	4
Matthews Coefficient ($Å^3$ Da^{-1})	2.3
Solvent content (%)	46
Total No. of reflections	143085 (21742)
No. of unique reflections	68030 (10055)
Multiplicity	2.1 (2.0)
Completeness (%)	88.4 (89)
Average I/σ(I)	6.6 (1.9)
Rsym or Rmerge (%)	(0. 14) (0.552)

$$R_{sym} = \sum \left| I-(I) \right| / \sum I$$

Table 5.2 Refinement statistics for processed (Serβ1Cys, PDB Code: 4PEL) and unprocessed (Serβ1Gly, PDB Code: 4PEM) mutants of *Kc*PGA.

Refinement Statics		
	Serβ1Cys (P1)	**Serβ1Gly (P1)**
Rcryst	0.212	0.253
Rfree	0.251	0.295
No. of Water Molecules	85	428
Root Mean Squire Deviation		
RMS Bond Length (Å)	0.009	0.010
RMS Bond Angle (°)	1.39	1.24
Overall B (Isotropic) from Wilson Plot	55.5	58.8
Ramachandran plot (% residues)		
Most favourable region (%)	96.8	97.3
Allowed Region (%)	2.9	2.1
Generously allowed region (%)	0.4	0.6
Disallowed region (%)	0	0

$\mathbf{R_{cryst} = \sum \left| F_o - F_c \right| / \sum F_o}$

$\mathbf{R_{free} = \sum \left| F_o - F_c \right| / \sum F_o}$ **where the F values are test set amplitudes (5%) not used in refinement**

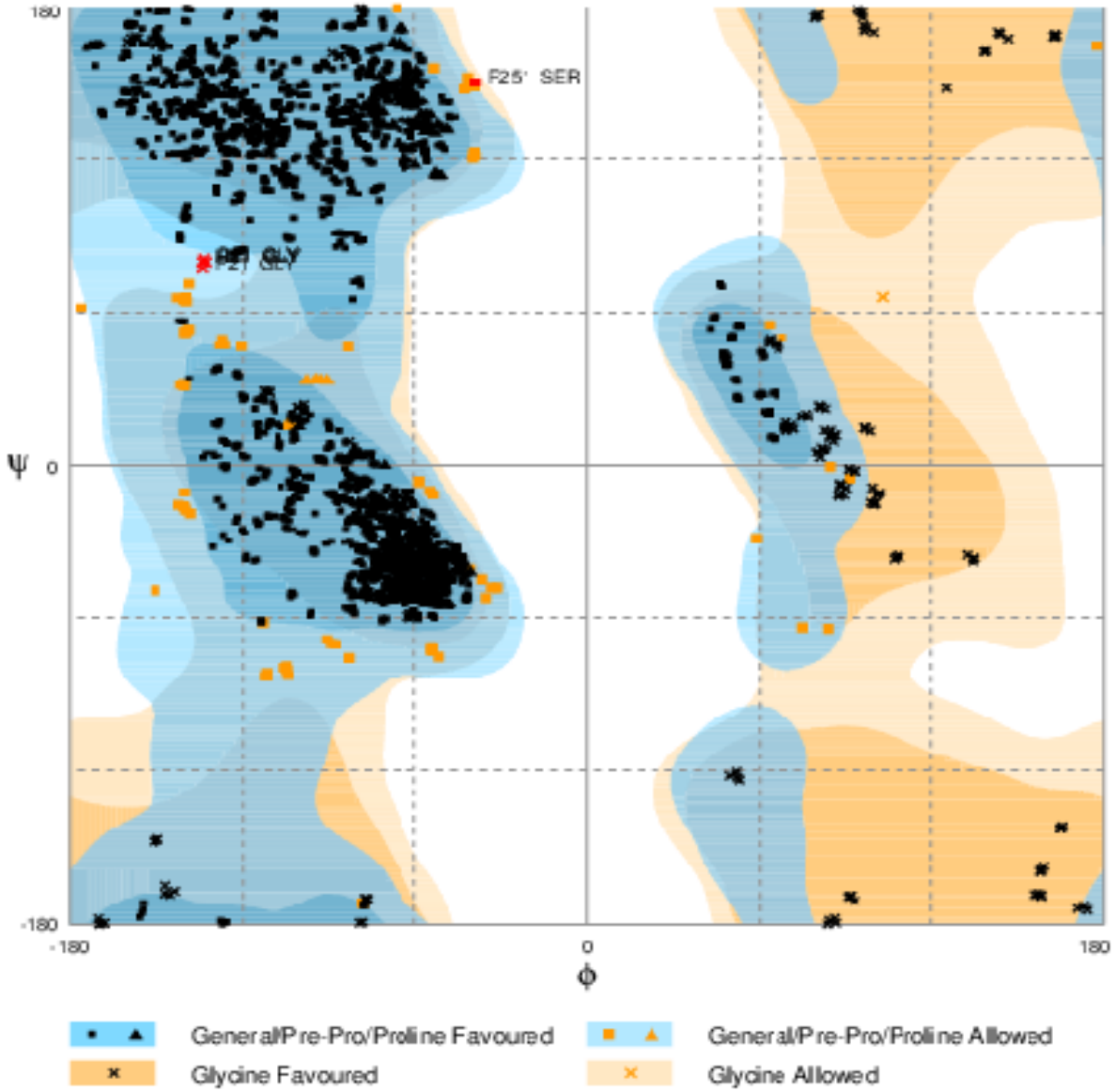


Figure 5.4 Ramachandran plot for *Kc*PGA_Serβ1Cys generated using PROCHECK program of CCP4i suit.

The structures of both the processed Serβ1Cys (PDB ID:4PEL) and unprocessed Serβ1Gly (PDB ID: 4PEM) *Kc*PGA mutants were refined to crystallographic $R_{factors}$ 21.2% and 25.3% and values of R_{free} 25.1% and 29.5%, respectively (Table 5.2). Both structures were validated by using PROCHECK (Table 5.2, Fig. 5.4). The overall structures of both processed and unprocessed mutants are comparable with that of *Ec*PGA. The root mean square deviation (RMSD) of Cα atoms superposition is given in Table 5.3. There is not much difference in Cα atoms RMSD of all the processed and unprocessed structure from *Kc*PGA and *Ec*PGA, since both shares ~80% sequence similarity (Varshney et al., 2013). The refined structures of both mutants Serβ1Cys and Serβ1Gly are presented as ribbon diagrams showing the characteristic Ntn-hydrolase αββα fold (Fig. 5.5 & Fig. 5.6).

The structure of processed Serβ1Cys comprises of two polypeptides as α and β-subunit which come together and form a deep-cavity like active site. For this mutant, the refined model consists of the α-subunit residues Pro3-Ala209 (207 residues) and β-subunit residues Cys1-Arg557 (557 residues) (Fig. 5.5). Whereas the unprocessed mutant Serβ1Gly structure comprises of a single polypeptide chain Pro3-Arg820 containing 818 amino acid residues (Fig. 5.6a). A break in the electron density was also observed at Pro260 which could be due to slow autocatalytic processing of spacer peptide in the mutant. The spacer peptide in the structure consists of 54 residues, Ala210-Thr263. Both processed (green) and unprocessed mutant (red) structures of *Kc*PGA were superposed which showed spacer peptide in the unprocessed form (Serβ1Gly), highlighted by yellow circle in the Fig. 5.6b and with a low C-alpha RMSD value of 0.513 (Table 5.3).

The α-subunit of the processed mutant Serβ1Cys consisting of 11 helices and 2 β-strands and β-subunit consist of 26 β-strands which make up the 6 sheets, 18 helices, 12 β-hairpins and 5 β-bulges (Fig. 5.7) (PDBsum, https://www.ebi.ac.uk/thornton-srv/databases/pdbsum). The average temperature factor for the Serβ1Cys mutant (PDB ID: 4PEL) and Serβ1Gly (PDB ID: 4PEM) structures was found to be 46.5 Å and 40.1 Å, respectively. The calculation was obtained from Protein Server Analysis Package (PSAP) web server (http://iris.physics.iisc.ernet.in/psap/). Matthews Coefficient for the Serβ1Cys was calculated 2.3 $Å^3$ Da^{-1} with 4 molecules in the asymmetric unit (Fig. 5.8).

Table 5.3 Comparison of overall Cα RMSD of *Kc*PGA and *Ec*PGA structures.

S. No.	Protein Structure	RMSD of C-alpha atoms				
		1	2	3	4	5
1	***Kc*PGA processed**	-	0.513	0.431	0.522	0.428
2	***Kc*PGA unprocessed**	-	-	0.579	0.677	0.636
3	***Ec*PGA processed**	-	-	-	0.520	0.386
4	***Ec*PGA unprocessed**	-	-	-	-	0.452
5	***Ec*PGA penG bound**	-	-	-	-	-

5.4.5 Mechanism of autocatalytic processing in the *Kc*PGA

To understand the mechanism of autocatalytic processing in *Kc*PGA two mutants, unprocessed Serβ1Gly and fully processed Serβ1Cys, were prepared. The Serβ1Cys mutant turned out to be PGA-inactive but fully processed, whereas the Serβ1Gly mutant was both PGA-inactive and unprocessed (Varshney et al., 2013). These observations demonstrate the indispensable role of a nucleophilic residue (Ser or Cys) at the N-terminus for the maturation of the enzyme through cleavage of the spacer-peptide. Site-directed mutagenesis of the β1 residue Ser to a non-nucleophilic Gly residue leads to a defect in processing. Interestingly, a serine residue as the nucleophile at the β1 position is essential for the catalytic activity of *Kc*PGA. The Cys at the β1 position supports processing, but is not capable of activating the enzyme. These observations evidently show the critical role of Ser residue at β1 position for both post-translational processing and enzyme activity, although Cys is adequate for post-translational processing.

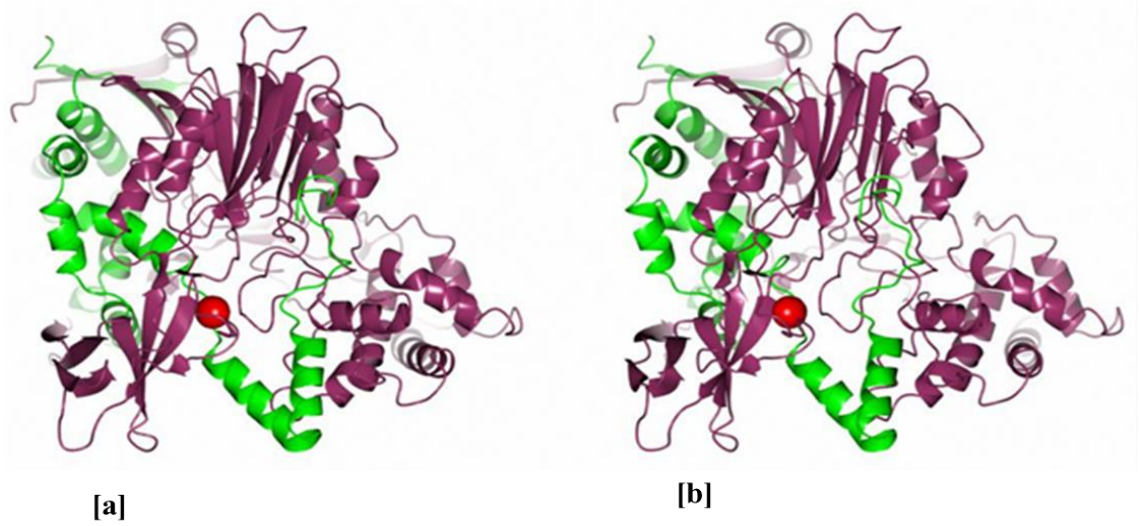


Figure 5.5 [a] Stereo view of the *Kc*PGA_Serβ1Cys mutant in cartoon representation (PDB Code: 4PEL, processed form) heterodimeric structure, α-subunit is shown in green and β-subunit in purple. The Ca2+ion is shown in red colour.

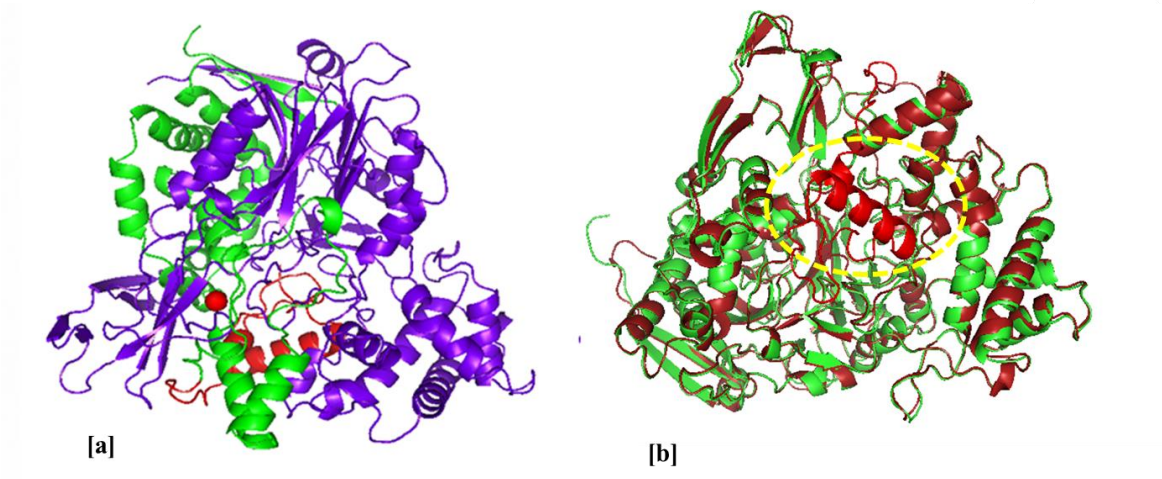


Figure 5.6 [a] *Kc*PGA-Serβ1Gly mutant (PDB Code: 4PEM, unprocessed form) showing spacer peptide (red colour), α-subunit in purple and β-subunit in green. [b] Superposed *Kc*PGA-Serβ1Cys (green) and *Kc*PGA_Serβ1Gly mutants (Red), spacer peptide in *Kc*PGA_Serβ1Gly unprocessed form highlighted by yellow circle.

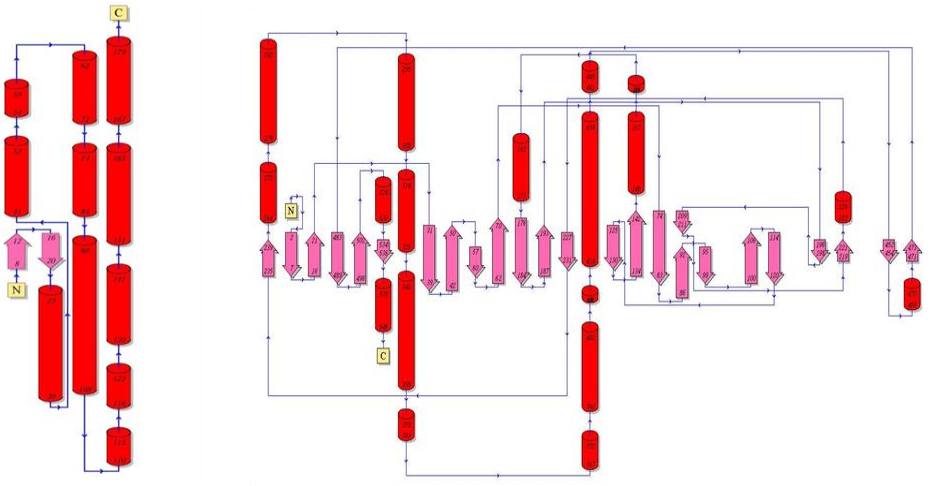


Figure 5.7 Topology diagram of the secondary structure content of heterodimer *Kc*PGA_Serβ1Cys mutant (from PDBsum, https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum). [a] Small α-subunit [~23 kDa] and [b] large β-subunit. The α-helices are represented by cylinders and the β-strands are represented by flat arrows, loops and coils are shown by lines.

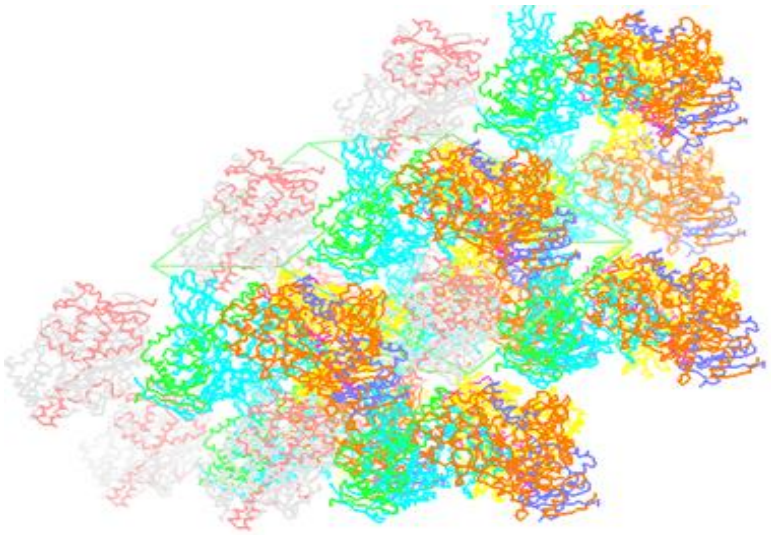


Figure 5.8 Crystal packing of the processed Serβ1Cys mutant structure.

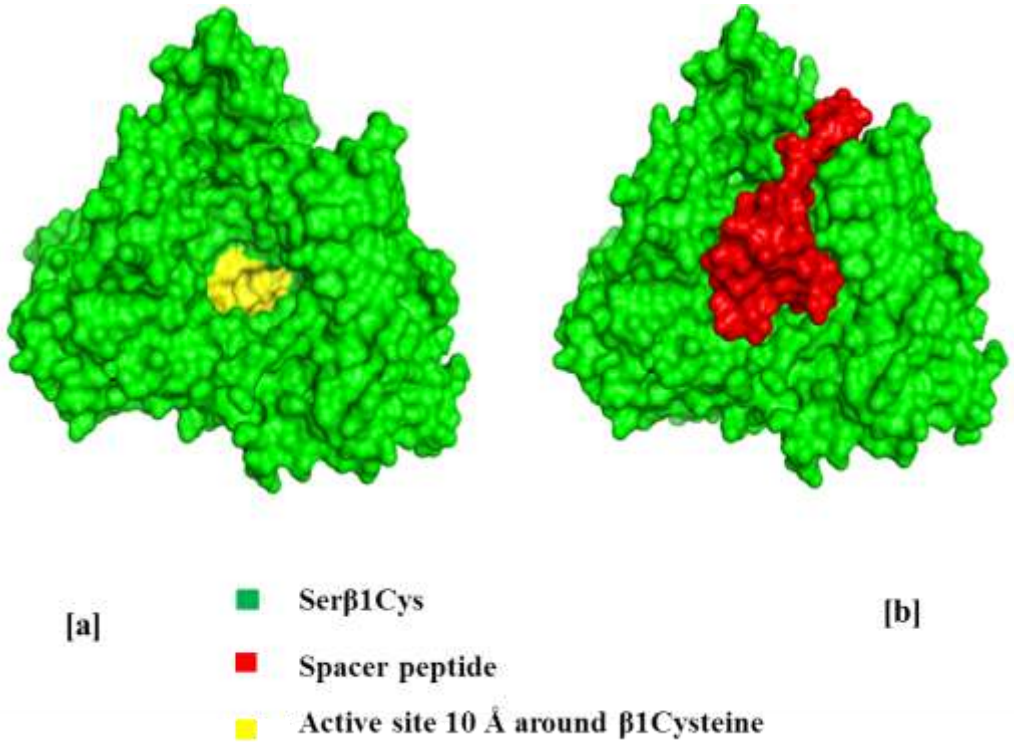


Figure 5.9 [a] Surface view of the Serβ1Cys mutant (green colour). Active site showing residues in 10 Å radius around β1Cys (yellow colour). [b] Surface view of the superposed Serβ1Cys mutant structure (green) and spacer peptide shown in red colour. Spacer peptide could be seen blocking the active site.

To get some structural insight into the autocatalytic processing in the *Kc*PGA both mutants structures were compared. Structural superposition of the Serβ1Gly and Serβ1Cys mutants reveals that the spacer peptide shown in red completely blocks the active site. Figure 5.9 a & b shows the surface view of the mutant structures in which the active site is shown as yellow surface (Fig. 5.9c). It can be clearly observed that the spacer peptide in Serβ1Gly mutant (shown as red surface in Fig. 5.9d) completely blocks the active site. It could be that the spacer peptide assumed a protective role of the active site in the periplasmic space. Figure 5.10 a & b shows RMS deviation and conformational changes of residues in the vicinity of the spacer peptide on comparing processed and unprocessed mutants of *Kc*PGA. As already reported, the structure contains a Ca^{2+} ion coordinated by Glu152 from the α-subunit and residues Val75, Asp73, Asp76, Pro 205 and Asp252 from the β-subunit (Fig. 5.10c). However, superposition of Ca^{2+} ion coordinating residues from both mutants showed no significant RMS deviation (Fig.

5.10d). Conformational differences examined in terms of RMSD of all atoms between the processed and unprocessed mutant structures showed more deviation in their α-subunits (Average 0.75 Å) compared to β-subunits (Average 0.69) (Fig. 5.10a).

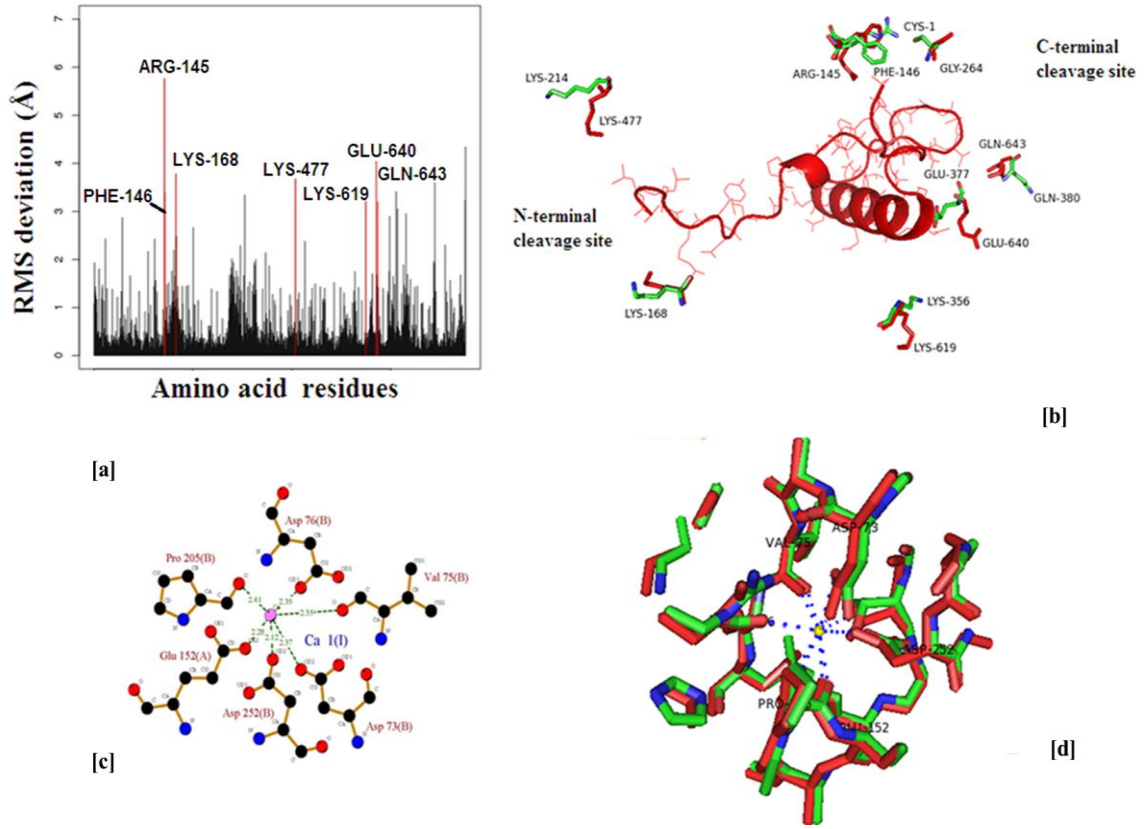


Figure 5.10 [a] All atoms Root Mean Square Deviation (RMSD) between the processed mutant (Serβ1Cys) and unprocessed mutant (Serβ1Gly) structures of *Kc*PGA is shown. Residues shown in the red colour are those that show high RMSD. [b] Figure shows the spacer peptide and deviation of the residues α145R, α146F, α168K, β477K, β619K, β640E and β643Q upon autocatalytic processing around the spacer peptide. Among them α168K and β477K were observed near to the N-terminal cleavage site of the spacer peptide. [c] Metal Ca^{2+} ion co-ordination in processed form of the Serβ1Cys mutant. [d] The figure shows the superposition of processed (Serβ1Cys, green colour) and unprocessed (Serβ1Gly, red colour) mutants of the *Kc*PGA. Ca^{2+} ion shows no significant RMSD.

The residues coordinating Ca^{2+} ion are from the β-subunit except for α152E (Fig. 5.10c) and the superposed structures show no significant Cα and side chain shift at coordinating residues (Fig. 5.10d) but show slight difference in the type of metal coordination, hexagonal bipyramid (equatorial) in processed form and pentagonal bipyramid (distorted) in unprocessed form. The residues such as α145R, α146F, α168K, β477K, β619K, β640E and β643Q in the vicinity of the spacer peptide (within 10 Å), showed significant conformational changes (all-atom RMSD > 3 Å; Fig. 3A). Their side chains moved in to partially fill the space left by spacer peptide. Among those residues α168K and β477K were found close to N-terminal cleavage site (209Ala-210Ala peptide bond) while the rest were near the C-terminal cleavage site of the spacer peptide (Fig. 5.10b). Therefore, α168K and β477K residues might have a role in the cleavage at N-terminal of the spacer peptide.

5.4.6 Comparison of the *Kc*PGA mutants (Serβ1Gly and Serβ1Cys) and *Ec*PGA structures

The active site pocket of *Kc*PGA consists of conserved Ser1, Asn23, Ala69, Asn241 and Arg263 β-chain amino acid residues. The Ser1 is the nucleophile residue which attacks adjacent Thr-Ser peptide bond (N → O acyl rearrangement) and cleaves off the spacer peptide. The residue Asn263 is involved in the stabilization of oxyanion hole formed during enzyme substrate complex intermediates. The processed form of *Kc*PGA (Serβ1Cys) was superposed on the precursor PGA (Serβ1Gly) around the 10 Å radius of the active site cysteine residue. Compared to the unprocessed *Kc*PGA mutant structure the residues α145R and α146F were found to flip towards the active site cysteine residue in the processed mutant and the values of all-atom RMSD for these residues were 5.76 Å and 3.38 Å, respectively (Fig. 5.10a & Fig 5.11).The above conformational changes in α145R and α146F residues led to the formation of a five-membered aromatic-aromatic interaction network involving residues α146F, β24F, β31Y, β57F, β71F (β71F interacts with α146F which interacts with β24F which interacts with β57F & β31Y) (Table 5.4) and a four-membered cation-pi interaction network involving α145R, α146F, β31Y, β24F (α146F interacts with α145R which in turn interacts with β24F & β31Y) (Fig. 5.12a) in the Serβ1Cys mutant structure which was absent in Serβ1Gly(Table 5.5).

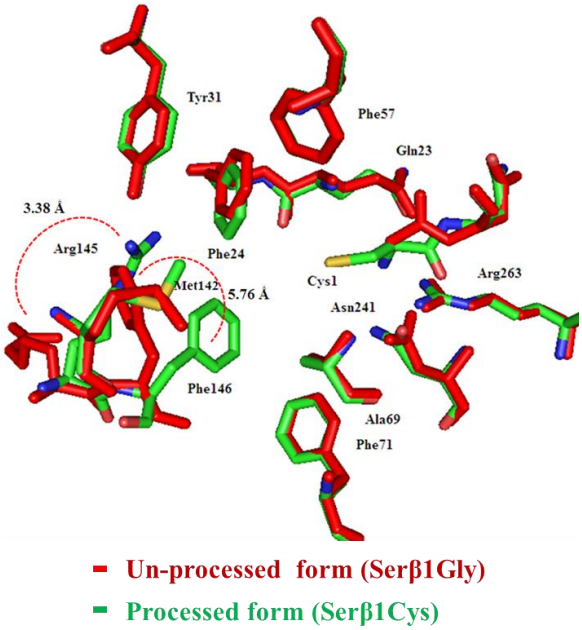


Figure 5.11 Figure shows the superposed active site residues of the processed (Serβ1Cys) and unprocessed (Serβ1Gly) mutants. The Arg145 and Phe146 amino acid residues showed large conformational changes and flipped toward the β1Cys in the processed form (Serβ1Cys, PDB Code: 4PEL) upon autocatalytic processing and removal of spacer peptide.

Table 5.4 Aromatic-Aromatic interactions around the 10 Å radius of *Kc*PGA_Serβ1Cys mutant active site.

Aromatic residue1			Aromatic residue2			Centroid Dist	Dihedral angle	ASA		Secstr	
Chain	Resno	Resid	Chain	Resno	Resid	(Å)	(Deg)	Aro1	Aro2	Aro1	Aro2
A	146	PHE	B	24	PHE	4.14	67.96	48.3	20.4	C	S
B	24	PHE	B	57	PHE	6.94	41.9	20.4	51.9	S	C
A	146	PHE	B	71	PHE	5.87	81.64	48.3	53.6	C	C
B	24	PHE	B	31	TYR	4.62	30.11	20.4	80.5	S	C
B	476	TYR	B	26	TRP	5.62	80	92.3	67.6	C	C
B	256	PHE	B	258	TRP	5.7	68.47	98	73.1	C	C

When Serβ1Gly mutant is compared with Serβ1Cys the centroid of α146F phenyl ring moved by 4.1 Å and the ring rotated by 68^{o} in the latter structure. Similar shifts in the α145R and α146F residues towards the active site and consequent formation of aromatic-aromatic (α146F interacts with β24F which interacts with β57F & β31Y) and

cation-pi (α146F interacts with α145R interacts with β24F & β31Y) interaction network was observed between the processed (PDB ID:1PNK) and unprocessed (PDB ID:1E3A) *Ec*PGA structures (Fig. 5.12b).

Table 5.5 Aromatic-Aromatic interactions around 10 Å radius of the *Kc*PGA_Serβ1Gly mutants active site.

Aromatic residue1			Aromatic residue2			Centroid Dist	Dihedral angle	ASA		Secstr	
Chain	**Resno**	**Resid**	**Chain**	**Resno**	**Resid**	**(Å)**	**(Deg)**	**Aro1**	**Aro2**	**Aro1**	**Aro2**
B	287	PHE	B	320	PHE	6.5	47.56	27.8	96.6	S	C
B	290	TYR	B	294	TYR	5.66	30.88	87.9	83.7	C	C
B	739	TYR	B	289	TRP	5.48	78.6	90.8	44.7	C	C
B	519	PHE	B	521	TRP	5.48	76.34	48.5	91.8	C	C

Hence, processing of spacer peptide has to be a crucial event in both these enzymes so as to form a well-structured catalytic framework involving aromatic residues at binding pocket to accommodate aromatic substrates. Interestingly, α-chain residues Arg145 and Phe146 in the substrate bound structure of *Ec*PGA (1FXV) flipped away upon substrate binding (Alkema et al., 2000) to position similar to those in unprocessed *Kc*PGA suggesting that they go through different conformational transition states (Fig. 5.12c) and substrate bound and unprocessed states have similarities of residue positioning for catalysis. There are three conformational states of the above residues observed in *Ec*PGA which includes those corresponding to processed, unprocessed and substrate bound states. Among them two states, unprocessed (PDB:1E3A) and substrate bound (PDB:1FXV) *Ec*PGA are in open conformational states, whereas, processed form without substrate remains in the closed state (PDB: 1PNK).

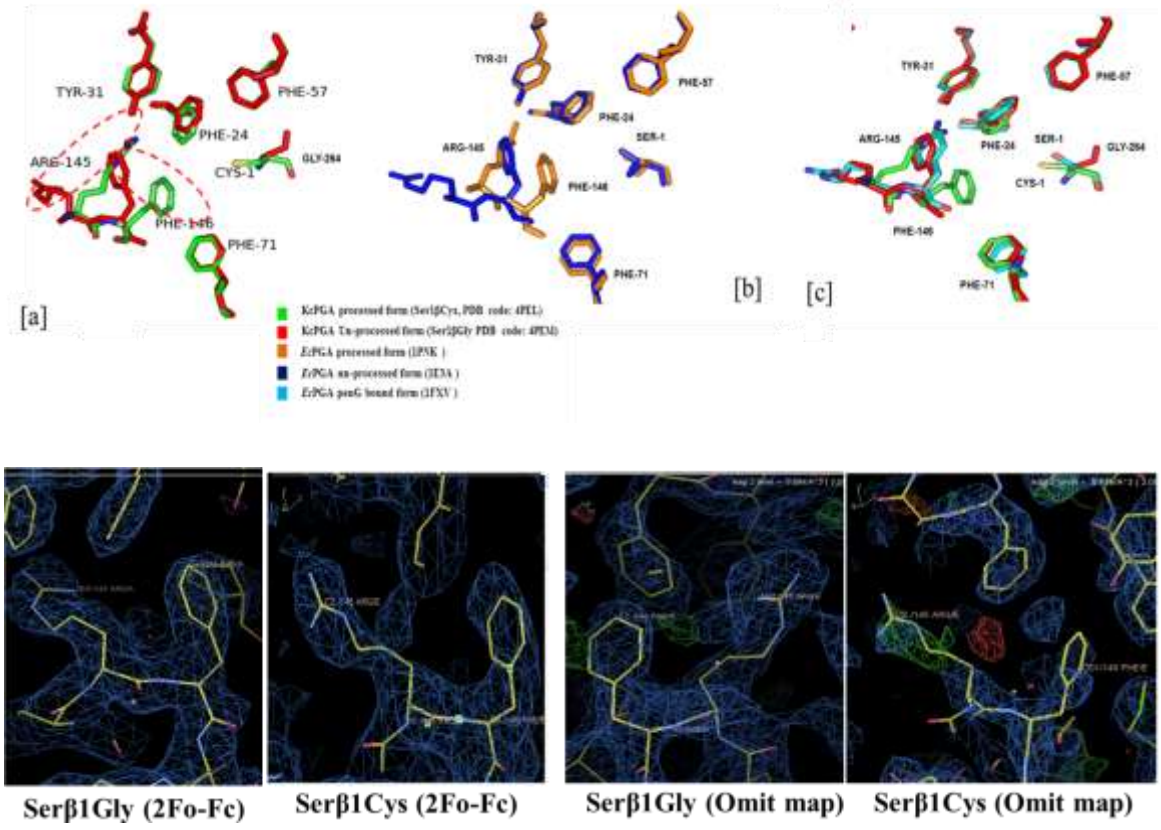


Figure 5.12 Figure shows the aromatic-aromatic and cation-pi interaction network. [a] Gain of aromatic-aromatic and cation-pi interaction network in the processed mutant of *Kc*PGA. The Arg145 and Phe146 amino acid residues of processed mutant (Serβ1Cys, Green colour) upon auto-catalytic processing of spacer peptide of unprocessed mutant (Serβ1Gly, red colour) lead to the formation of the 5-membered aromatic-aromatic and 4-membered cation-pi interaction network [b] The figure shows similar observation in the *Ec*PGA processed (1PNK, orange) and unprocessed (1E3A, blue) mutant structure. [c] The superposed image of *Kc*PGA processed form (green), unprocessed form (red) and *Ec*PGA substrate bound form, to show the conformational transition of Arg145 and Phe146. [d] 2Fo-Fc map for α-Arg145 and α-Phe146 residue in Serβ1Gly and Serβ1Cys mutants contoured at 1δ. [e] Omit map for both mutant contoured at 3δ.

Therefore, the Serβ1Gly mutant remains in open conformation whereas, Serβ1Cys mutant remains in closed conformation state. There are reports which say that open conformation might correspond to a low-energy conformation of the enzyme with no bound substrate (Done et al., 1998).

5.4.7 Homoserine lactone binding with processed *Kc*PGA and *Ec*PGA structure

The native *Kc*PGA hydrolyses acyl homoserine lactones (3-oxo-hexanoyl homoserine lactone) (Fig.1) whereas the Serβ1Cys mutant binds AHL but is unable to turnover either this substrate or pen G substrate. Previous reports have also shown that *Kc*PGA acts on AHL whereas *Ec*PGA does not (Mukherji et al., 2014). This may be due to differences in the conformations adopted by residues Phe146 and Phe24 in *Kc*PGA.

A comparison of the *Kc*PGA-Serβ1Cys mutant binding pocket with that of *Ec*PGA processed form (1PNK) in a 10 Å radius around the active site cysteine showed major deviations in main chain Cα and side chain atoms of Arg145 and Phe146 residues of α-subunit. A minor shift of phenyl ring of residue Phe24 of β-subunit was also observed in the Serβ1Cys structure when compared to Serβ1Gly (Fig. 5.12a). This was absent when *Ec*PGA processed and unprocessed forms were compared (Fig. 5.12b). In addition, upon superposition of the atoms in the active sites of Serβ1Cys and processed *Ec*PGA (1PNK) a slight displacement of the Phe24 amino acid of β-subunit in addition to Arg145 and Phe146 of α-subunit could be seen (Fig. 5.13).The movement of these residues at active site showed binding pocket differences in the *Kc*PGA processed form (Serβ1Cys) in comparison to processed *Ec*PGA (1PNK) binding pocket. Hence, the plasticity of Phe146 and Phe24 amino acid residues might provide the differential substrate specificity. Mutational studies of residues α146F, β24F and β71F have already been reported in *Ec*PGA (Alkema et al., 2004; Deaguero et al., 2012; Shapovalova et al., 2009). Various complexes of phenylacetic acid and (R)-α-methylphenylacetic acid with wild-type and mutants of *Ec*PGA have confirmed the role of α146F and β24F in substrate selectivity (Deaguero et al., 2012) and diastereoselectivity (Alkema et al., 2004). Similarly, it was demonstrated that replacing β71F selectively with other amino acids modified the catalytic properties and enantioselectivity of *Ec*PGA. To explore this possibility we performed docking studies of 3-oxo-hexanoyl homoserine lactone (3-oxo-hexanoyl HL) with the Serβ1Cys mutant of *Kc*PGA as well as processed substrate free (1PNK) and substrate bound *Ec*PGA (1FXV) structures. Substrate binding site and productive binding mode were identified from substrate bound *Ec*PGA structure, which was used for docking studies without the substrate.

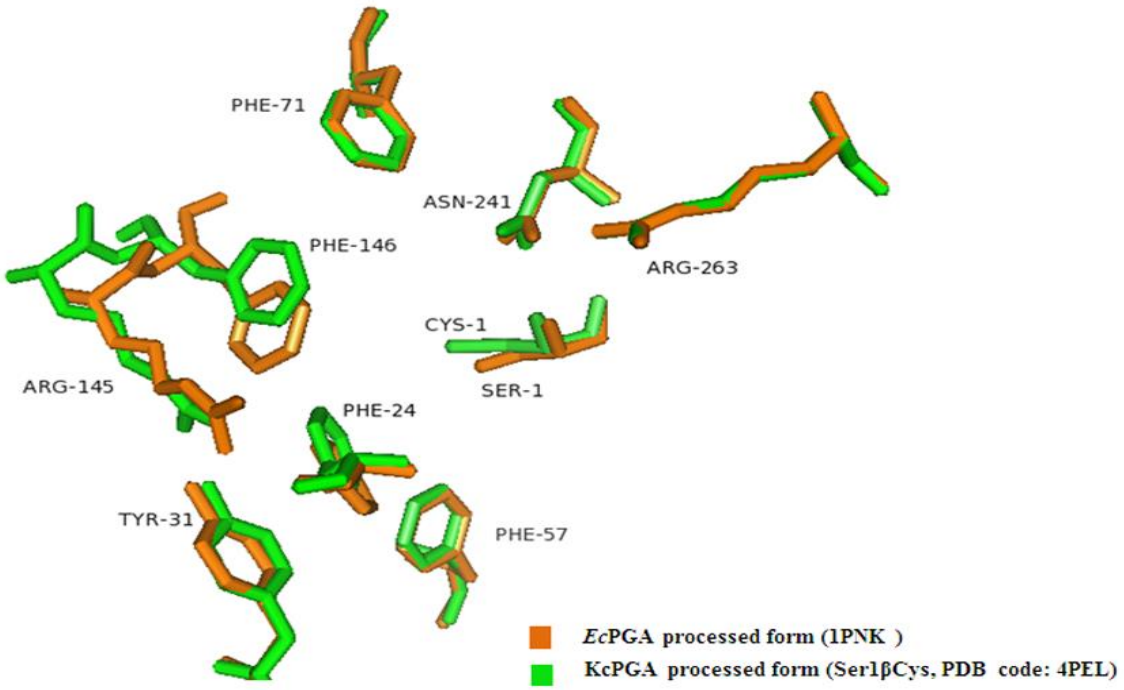


Figure 5.13 Comparison of the binding pockets of processed forms of *Kc*PGA (green) and *Ec*PGA 1PNK (orange). The binding pocket residues Arg145, Phe146 and Phe24 of *Kc*PGA processed form were found to flip away in comparison to processed *Ec*PGA (1PNK) which probably suggests the plausible expansion of binding pocket of *Kc*PGA to accommodate AHL.

Docking results showed non-productive binding of 3-oxo-hexanoyl HL, binding mode not conducive for activity due to steric clashes and distance from nucleophilic atom, with both processed (1PNK) and substrate bound (1FXV) forms of *Ec*PGA while productive binding, with no steric clashes and cleavage bond positioned near nucleophilic atom, was observed with the Serβ1Cys mutant of *Kc*PGA (Fig. 5.14a). During docking procedure receptor was treated as a rigid geometry while ligands were flexibly docked in the receptor binding site. The binding modes were evaluated based on spatial fit, interaction complementarity with binding site residues and the binding affinity values estimated as Glide scores. In the case of Serβ1Cys docked structure, the hydrophobic tail of the substrate was oriented towards the hydrophobic pocket and the carbonyl carbon of the substrate was positioned towards the nucleophilic cysteine residue (β1Cys) (Fig. 5.14b).

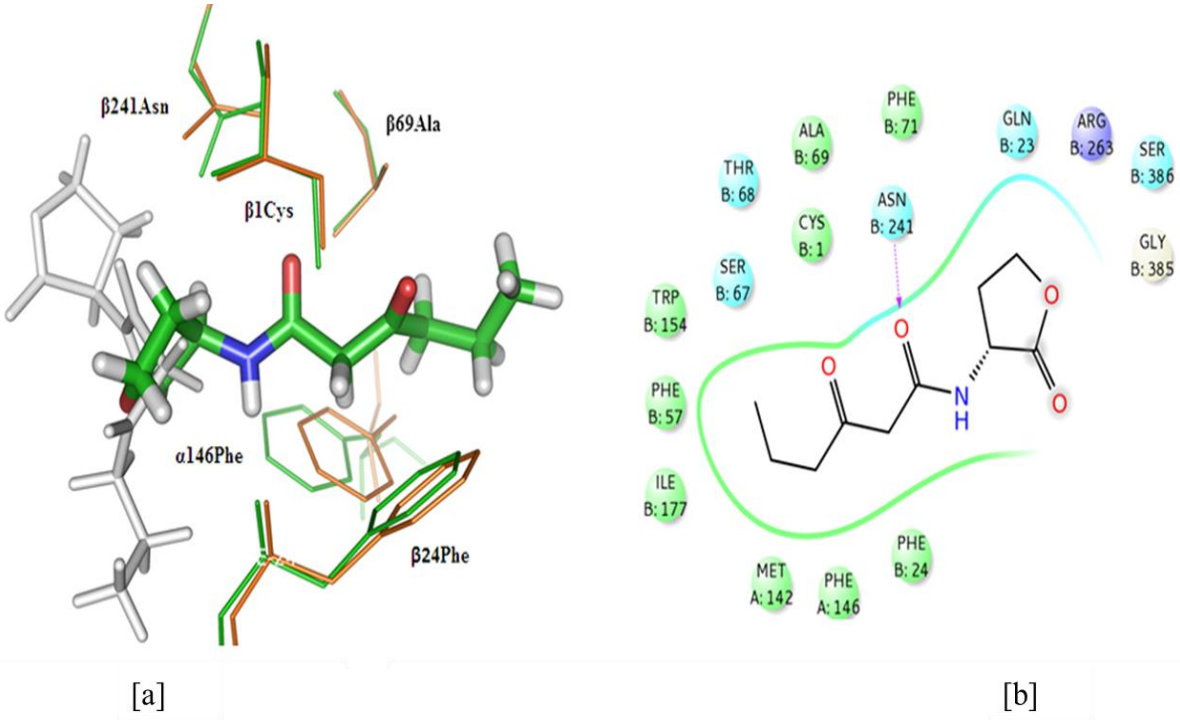


Figure 5.14 [a] Docking of 3-oxo-hexanoyl homoserine lactone with processed mutant of *Kc*PGA-Serβ1Cys (green) and processed *Ec*PGA (1PNK) (orange). Gray colour ligand is shown with processed *Ec*PGA (1PNK) docked structure showing non-productive binding. Green colour ligand shown with Serβ1Cys docked structure showed productive binding. [b] Productive binding conformation of 3-oxo-hexanoyl homoserine lactone bound at *Kc*PGA processed mutant Serβ1Cys binding pocket and interaction between enzyme and ligand.

The nucleophilic attack distance was measured at 3.2 Å and 5.4 Å from β1Cys of *Kc*PGA_Serβ1Cys and β1Ser of substrate free *Ec*PGA (1PNK) to the carbonyl carbon of 3-oxo-hexanoyl HL, respectively (Fig. 5.15a). In the substrate free *Ec*PGA (1PNK) docked structure the docked molecule adopted a totally different conformation while in substrate bound structure (1FXV) of *Ec*PGA it showed a mode of binding similar as Serβ1Cys docked structure (Fig. 5.15b). The residues α146Phe, β1Ser, β24Phe, β69Ala and β241Asn showed bad contacts or steric clashes on transferring docked 3-oxo-hexanoyl HL in the productive binding conformation of Serβ1Cys to the binding pocket of the processed *Ec*PGA (1PNK) and substrate bound *Ec*PGA (1FXV) structures, hence

leading to those becoming non-productive binding in both cases (Fig. 5.15c & d). The role of crucial residues α146Phe and β24Phe have been previously reported in diastereoselectivity (Deaguero et al., 2012, Andrew D. L. et al., 2012).

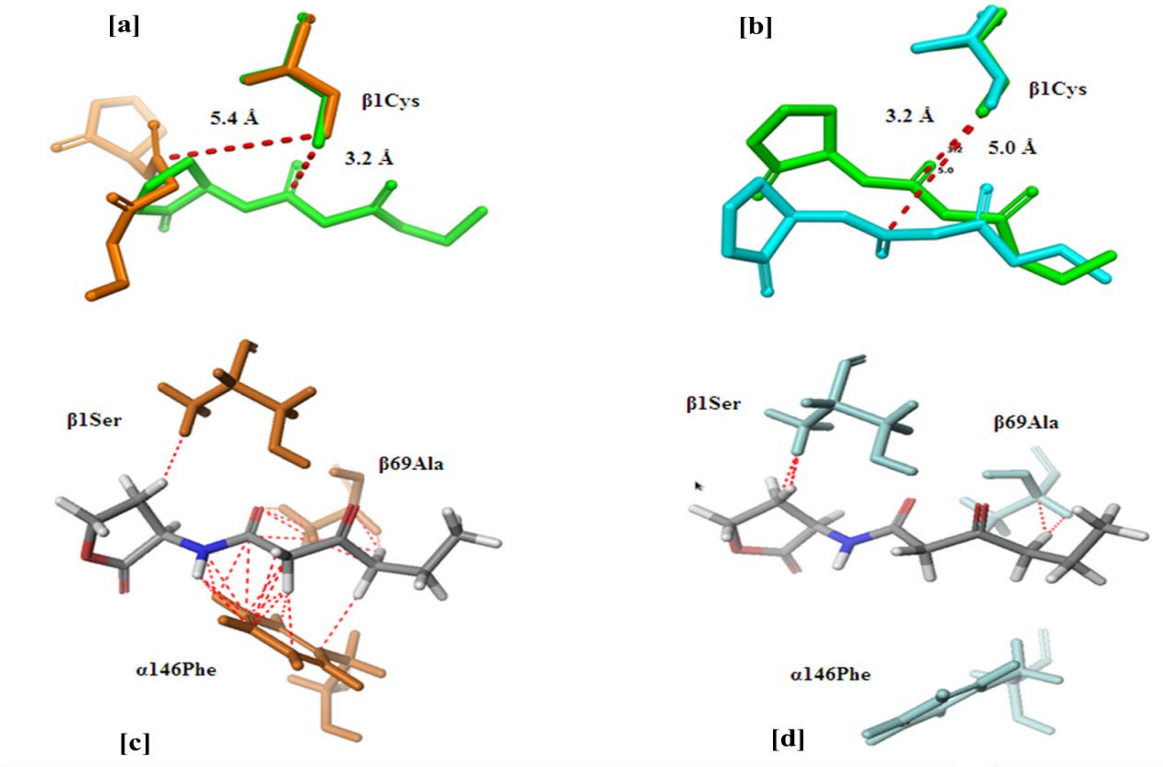


Figure 5.15 [a] Figure shows nucleophile distance of Serβ1Cys (green, 3.2 Å) and 1PNK (orange, 5.4 Å) docked structure to the carbonyl carbon of the 3-oxo-hexanoyl homoserine lactone [b]) Nucleophile distance of Serβ1Cys (green, 3.2 Å) and substrate bound *Ec*PGA (1FXV) (cyan, 5.0 Å) to the carbonyl carbon of the 3-oxo-hexanoyl homoserine lactone (cyan). [c] Figure shows steric clashes with ligand 3-oxo-hexanoyl homoserine lactone (gray) in its productive binding conformation with Serβ1Cys when placed in the binding pocket of processed *Ec*PGA (1PNK) (orange). [d] Steric clashes with ligand 3-oxo-hexanoyl homoserine lactone (gray) in its productive binding conformation with Serβ1Cys when placed in the binding pocket of substrate bound form of *Ec*PGA (1FXV) (cyan). In *Ec*PGA structure (1FXV) the residue α146Phe is moved away and hence has no bad contact but β1Ser and β69Ala showed steric clashes.

5.5 Summary of the Chapter

This chapter described about the crystallization, data collection and three-dimensional structure determination of the fully processed *Kc*PGA_Serβ1Cys mutant and comparison with *Kc*PGA_Serβ1Gly and *Ec*PGA structures. Structural analysis provides essential insights into autocatalytic processing of the spacer peptide. Structural comparison showed significant differences in the binding pocket of unprocessed (Serβ1Gly) and processed (Serβ1Cys) *Kc*PGA, where α145Arg and α146Phe were found to be greatly flipped toward the active cysteine in the case of processed form upon autocatalytic processing of the spacer peptide, indicating the crucial role of the both the residues in the processing dynamics and substrate binding as well.

In addition, structural and substrate (docking studies) binding studies shows differences in the binding mode of the substrate. Since *Ec*PGA enzyme does not show activity towards AHLs even when sharing 80% sequence similarity, it could be possible that the activity of *Kc*PGA towards AHL along with the other substrate Pen G is due to flexibility and different conformations of residues Phe146 of α-chain and Phe24 of β-chain as evidenced from our analysis and mutational studies reported on *Ec*PGA. Docking studies also supported the unproductive binding of AHL molecule due to clashes with Phe146 of α-chain which act as a gate keeper and occludes the productive conformation in the *Ec*PGA structure. However, our modeling studies show that β71F being away from bound AHL can hardly influence its binding. In conclusion the evidence is that the post-translational processing follows the same pathway involving similar residues in both enzymes whereas the changes in the dispositions of same residues in processed form due to conformational difference such that the enzymes can discriminate between certain molecules for activity.

Chapter-6

Summary

&

Conclusion

Bile salt hydrolase (BSH) and penicillin G acylase (PGA) which are clinically and pharmaceutically important enzymes, have been studied in this thesis work. Both the enzymes belong to Ntn-hydrolase superfamily, members of which are structurally related but display diversity in their primary structure and substrate spectrum. Main objective of our investigations was to understand the structure-function relation of these enzymes that can substantially enhance their potential applications. The study involved biochemical, biophysical, structural and computational analysis to establish the structure-function relation of these enzymes.

BSH is an important enzyme which de-conjugates the bile acid in the gut and intervenes in the bile acid metabolism pathway. The gene coding for *Ef*BSH from *Enterococcus faecalis* (a Gram-positive bacteria) was fished out using gene-specific primers and cloned, expressed and purified from BL21*(DE3) expression host cells. *Ef*BSH enzyme was noted to exhibit very high activity towards bile acid specifically glyco-conjugated bile acids, compared to other BSHs reported. It also displayed the unique characteristic of allosteric behavior which was observed in a BSH enzyme for the first time. Pen V substrate which is hydrolyzed by PVA, the enzyme closely related to BSH, both belonging to CGH family, is known to bind at BSH active site and shows competitive inhibition. However, *Ef*BSH is not inhibited by Pen V, instead it enhanced the BSH activity.

The three-dimentional structure of *Ef*BSH was determined at 2 Å resolution. The structural analysis and comparison with other BSHs showed variation in active site binding loops (loop1-4) and the tetramerization or assembly loop. Insertions were observed in the loop2 and loop3. Loop2 had large insertion which resulted in the disruption and reduction of the core anti-parallel β-strand of Ntn-fold. Due to this insertion, loop2 became longer compared to other BSHs. In addition, loss of anti-parallel β-sheet was also observed in the tetramerization loop which extended to the diagonally opposite subunit to form the tetramer. We hypothesize that the allosteric behavior of the *Ef*BSH is the result of insertion in the loop2 and loss of the anti-parallel β-sheet which might be transmitting the conformational changes to the neighbouring subunit. Docking and MD simulation studies were performed and active site of the *Ef*BSH was analyzed. It was observed that residues F18, Y20 and Y65 interact with the hydrophobic part of the

GCA substrate. These residues were found replaced in other BSHs such as *Bl*BSH (L20, W22, F67) *Cp*BSH (M20, I22, A68) and *Ls*BSH (L18, L20, F65). Moreover, polar complementarity (basis for the better substrate activity and specificity) of the three –OH groups (3α-OH, 7α-OH & 12α-OH) of GCA with the active site of *Ef*BSH was also calculated by radial distribution function. Polar complementarity of these –OH groups was found to be very high, co-ordinated by residues D19, N79, P135 and S136. All these structural variations might have changed the character of the binding pocket of *Ef*BSH which probably contributed to its unique functions of high catalytic activity and cooperativity behavior. Selected active site mutants were prepared to explore the role of mutated residues in the substrate binding and activity.

Autocatalytic processing is one of the key features of Ntn-hydrolases. Autocatalytic processing in *Bacillus sphaericus* PVA (*Bsp*PVA) is reported. Here, we tried to explore the capacity for autocatalytic processing by *Ef*BSH when pre-peptide of *Bsp*PVA was inserted at its N-terminal. Four mutants (MG, MA, MLG and MLG_TA/SS (MLG-CSS) pre-peptide mutants) were prepared in which di and tri-pre-peptides were introduced at the N-terminus and autocatalytic processing was assessed. MA and MG pre-peptides were found to process fully and produced active form of the *Ef*BSH. However, MLG pr-peptide mutant was expressed into inclusion body. We analyzed the sequence alignment with *Bsp*PVA which was known to process and remove the MLG pre-peptide. We noticed the difference of Ser-Ser being present after nucleophile cysteine (MLG_CSS) in *Bsp*PVA instead of Thr-Ala of *Ef*BSH. The mutant S-S replacing T-A (MLG_CSS) in *Ef*BSH was prepared. The double mutant MLG_TA/SS (MLG_CSS of *Ef*BSH) produced the processed and fully active form of *Ef*BSH, indicating the importance of Ser-Ser residues in that position of the sequence for autocatalytic processing.

PGAs are industrially important enzymes which hydrolyse the amide bond in the natural penicillins and produce 6-APA (β-lactam nucleus) a precursor molecule in the preparation of semi-synthetic antibiotics. *Kluyvera citrophila* PGA (*Kc*PGA) is synthesized in the cytosol and transported to the periplasm. *Kc*PGA is expressed as pro-peptide of ~89 kDa (unprocessed and inactive) and it undergoes autocatalytic processing in the periplasm and produces mature fully active enzyme consisting of α-subunit (~23kDa) and β-subunit (~63kDa). *Kc*PGA exhibits promiscuity by hydrolyzing the

quorum sensing molecule N-Acyl homoserine lactone (AHL) in addition to its known substrate Pen G.

*Kc*PGA mutant Serβ1Cys (the N-terminal nucleophile Ser replaced by Cys), which underwent autocatalytic processing, was prepared, purified, crystallized and three-dimentional structure elucidated. The Structural comparison of the Serβ1Cys processed mutant and Serβ1Gly unprocessed mutant of *Kc*PGA was performed which provided insights into autocatalytic processing and removal of the spacer peptide. Structural comparison showed significant differences in the binding pocket of unprocessed (Serβ1Gly) and processed (Serβ1Cys) *Kc*PGA, where α145Arg and α146Phe residues showed large conformational changes. *Kc*PGA shares 80% sequence similarity with *E.coli* PGA (*Ec*PGA) but *Ec*PGA do not show activity towards AHL. Structural comparison and docking studies revealed that *Kc*PGA activity towards AHL along with the original substrate Pen G was due to flexibility and different conformations of residues α146Phe and β24Phe. In *Ec*PGA the residue Phe146 of α-chain was found to exclude productive binding of AHL molecule and showed steric clashes when coordinates of the *Kc*PGA docked AHL (productive binding) was transferred to the *Ec*PGA structure, indicating the crucial role of the different conformations of Phe146 of α-chain in substrate discrimination.

In conclusion, our research reported here has identified a highly active BSH from a Gram-positive bacteria *Enterococcus faecalis.* The cloning of the gene, expressing and characterizing the enzyme using various biochemical and biophysical techniques, along with computational studies have established the role of structure in its high specific activity, allosteric behavior and promiscuity. Mutational studies have shown the importance of two serine residues next to N-terminal cysteine in autocatalytic processing of cholylglycine hydrolases. Similarly, studies on PGAs showed the role of various residues and their conformations in autocatalytic processing and discriminating between substrates. Apart from offering a highly active BSH for therapeutic applications these findings will also help in protein engineering to produce related enzymes with modified desirable properties.

Appendix

Introduction

Copper is a heavy metal which plays critical roles in biological metabolic processes, growth and physiology of organism and excess amount results in cell death (Gupta, et al., 1995). Copper maintains the cellular osmotic balance and enzymes utilize as cofactors and protein structures as stabilizers. The protein possessing the Cu^{++} ions are widely distributed in organisms from prokaryote to eukaryotes and involved in several physiological processes, oxidative stress protection, iron transport, respiration and pigmentation (Puig, S. and Thiele, 2002). Cu^{++} ions have been known to exist in two states, an oxidized state as Cu^{++} and a reduced state as Cu^{+} (Rensing and Grass, 2003). Copper deficiency and copper sensitivity due to genetic alteration in the copper metabolism may result in disorders like Wilson's and Menkes' disease Alzheimer and Creutzfeld-Jacob syndrome in humans (Bull, and Cox, 1994).

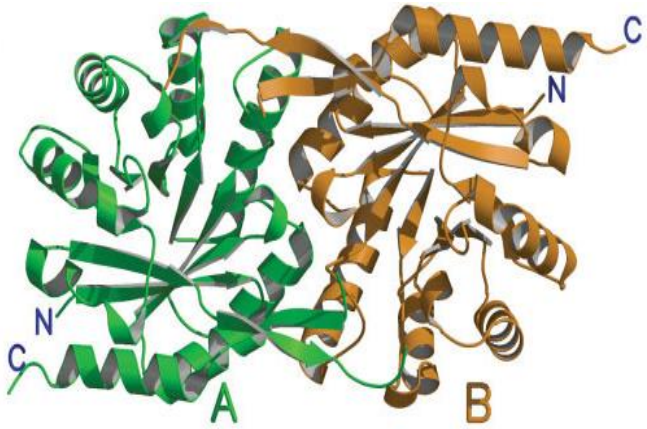


Figure 1.0 Crystal structure of the CutCm from *Shigella flexineri* at 1.7Å resolution.

Organisms have evolved copper homeostasis mechanism to reduce copper toxicity in the cell. In bacteria two families are involved in the copper homeostasis: cop gene family and Cut gene family. Cop gene family is a well understood system of active transport efflux pump. However, Cut gene family, constituting of six members *CutA, CutB, CutC, CutD CutE* and *CutF* , is not fully studied. The first crystal structure of CutCm from *Shigella flexneri 2a str. 301* was determined using multiple-wavelength anomalous dispersion (MAD) method at 1.7 Å resolution (Zhu, et al., 2005). But CutC

from *E.coli* has been assigned the function of transporting reduced form of Cu^{+} in the cytosol.

Primarily our aim was to clone the *bsh* gene using the forward primers 5'GCGGCGCATATGTGTACAGCAATTACTTATGTATC 3' and reverse primer 5'TTA ATT AGC ATA GTT AAT TTG TTG TGT TTC CAC 3'. But we have resulted in the amplification of ~500bp length product. Amplified product was sequenced and it was identified as CutC gene instead of bsh. Then, specific primers were designed for the Cut C gene and cloned. In, the present work we have cloned the CutC gene, expressed and purification was achieved. The purified CutC protein was subjected to extensive crystallization trials by various high throughput screening methods using the commercially available crystallization screens.

Materials and Methodology

Genomic DNA isolation and Cut C gene cloning

The genomic DNA was isolated from *Lactobacillus plantarum JDM1* using DNA isolation Kit (Quagen). Gene specific primers were designed for CutC gene and was amplified and cloned into the pET28a+ vector with *NdeI/HindIII* restriction sites.

Primers:

Forward primer (*NdeI*): 5'-GCCACA TATGTTATTAAAAGAAGTCTGTG-3' Tm = 61°C

Reverse primer (*HindIII*) : 5' GTA AGC TT A ATA ATC TTG GTC CCA TG 3' Tm = 60.1°C

The *CutC* gene was amplified in the PCR condition; denaturation at 94°C/2min, [94°C/30sec, 50-58°C/30sec, 72/1min] x 25 cycles and final extension at 72°C for 10 min. The ligation mixture was transformed in the DH5α competent cells and incubated overnight at 37°C. Colonies were screened for positive clones by using the colony PCR. In the colony PCR single colonies were picked and mixed with the autoclaved milliQ water and heated for 10 min to disrupt the cells. After that 1ul of the lysate was taken and used in the PCR as a template. Positive clones gave the amplification of the inserted gene. In addition, to confirm the gene insertion in the cloned plasmid, double digestion was performed (*NdeI/XhoI*) in which the inserted gene was released.

Expression, purification & crystallization of CutC protein

The positive clones of the CutC were expressed in the BL21 rossetta (DE3) cells. The cells were grown in the LB media and induced with IPTG for the overexpression of the protein and incubated at 16^{o}C overnight. The cells were harvested and lysed with the lysis buffer (Tris-Cl pH 7.5, 500mM NaCl, and 1mM DTT). Protein was purified by the Ni-NTA affinity followed by the Sephacryl-200 size exclusion chromatography. The molecular weight of protein was estimated by SDS-PAGE and MALDI-TOF analysis. The purified CutC was subjected for the crystallization with several commercially available screens.

Results

Copper homeostasis gene (Cut C) cloning *L. plantarum JDM1*

The *CutC* gene was fished out from the genomic DNA of *Lactobacillus plantarum JDM1*and was cloned in the pET28a expression vector. The gene-specific primers were used to amplify the *CutC* gene and amplified product of 639 bp was observed in the Agarose gel shown in the Fig.1

CutC gene:

ATGTTATTAAAAGAAGTCTGTGTTGAAAATTACACTAACATTCCCGCAGCCAT
TGCGGCCGGTGCCAATCGGATCGAACTCAACGATAACCTATCCGTTGGCGGT
ACAACGGTCAGTCGTGGTGTAATGGCTGAGGCAGCCAAGTATACCAGTGAAC
ATCACGTGCCCCTAGTGACCATGATTCGCCCCCGTGGTGGCAATTTCGTTTAT
AATGATACGGAATTAAAAATCATGGAAGCGGACCTGCTACAGGCTCAATCAT
TAGGTGTTGATGGTGTCGCATTTGGTGCTTTGACTGCGGATAATCAGCTTGAC
GAAGAAGCCTTAGCGTTACTAATTGGTGCAGCCGGTGGCATGTCGATCACCTT
CCACATGGCTTTTGACGCCATCCCTGAAAATCAACAAGCCGCGGCAATTGACT
GGTTAGTAGACCATGACGTCGATCGCATTTTGACACATGGCGGCCCACTTGAT
CAGCCGATTGCTGACTGTGTCCCCCACTTGCAAACGACGATCAAACAGGCCG
CTGGCCGAATTCAGATCCTACCTGGTGGTGGCATTACCACCGCCAATGTCGCC
GACATCACCACAACGCTCGGCGTCAAACAGGCCCATGGGACCAAGATTATTA
ATTACTAA

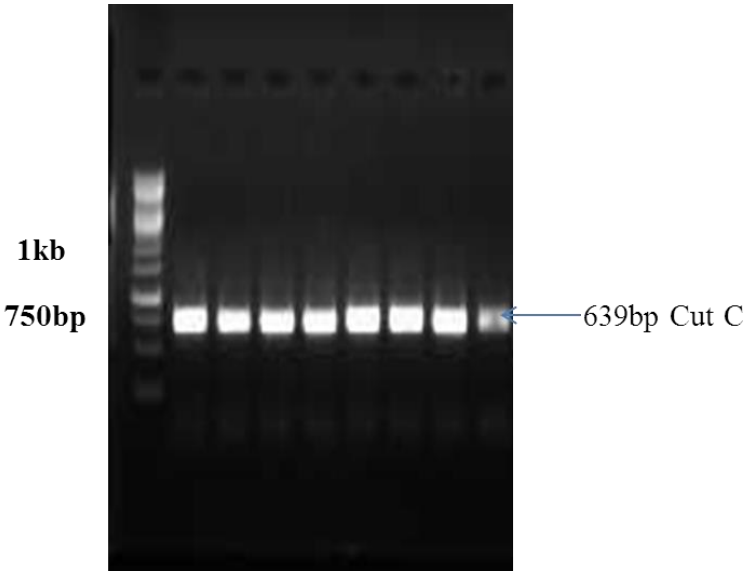


Figure 1 CutC gene amplification (gradient) from genomic DNA by CutC specific primers and loaded into 1 % agarose gel Lane 1: Marker Lane 2-9: PCR amplification of CutC gene (50-58°C).

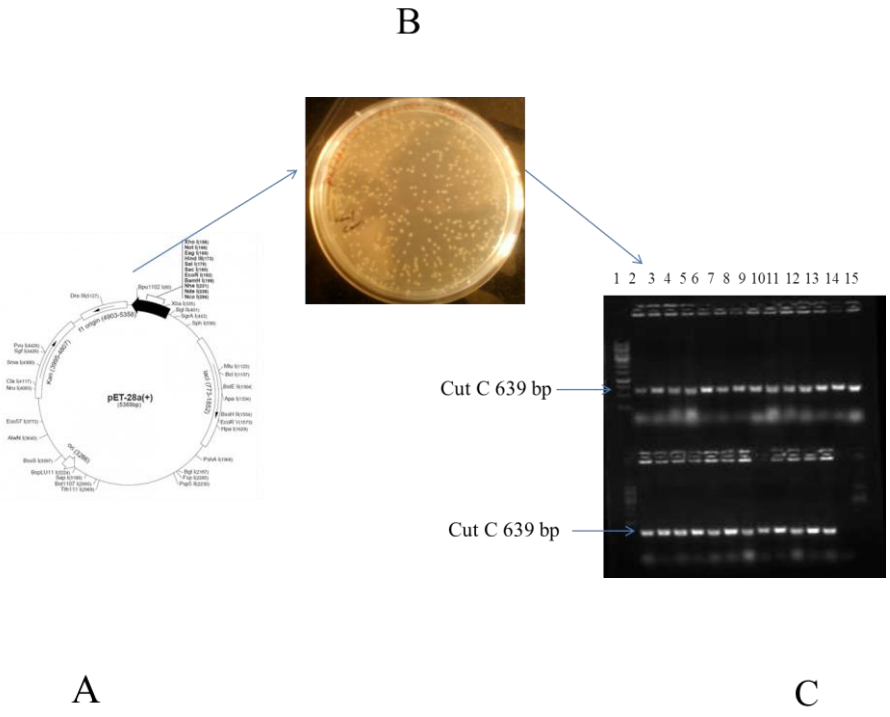


Figure 2 Cloning and confirmation of pET28a+/CutC Clones by colony PCR: A. map of pET28a+ vector B. Transformed clones (34µg/ml, Kan/Cam) C. Colony PCR of clones with Cut C-specific primers, lane 1- Marker, Lane 2-15 - Cut C positive clones, 1% Agarose gel.

The amplified product and pET28a vector were digested with *NdeI/HindIII* restriction enzyme for 4 hr at 37^{o}C. Linear vector was purified by the gel elution method. The purified and *NdeI/HindIII* vector and PCR products were used for the ligation. The ligation mixture was transformed in the DH5α competent cells. Many colonies were obtained and colony PCR was performed to confirm the insertion of the *CutC* gene in the vector. All colonies were found to be positive clones (Fig.2). Further plasmid was isolated from positive clones and sequencing was done.

Expression, purification and crystallization of CutC protein

The positive clone of *CutC* was expressed in Bl21 rosseta (DE3) expression host. The protein was expressing in the inclusion bodies and very less amount of protein was observed in the soluble form. To produce CutC protein in the soluble form various parameters like, IPTG concentration, temperature (10^{o}C, 16^{o}C, 22^{o}C and 37^{o}C) and incubation time were standardized (Fig.3a).

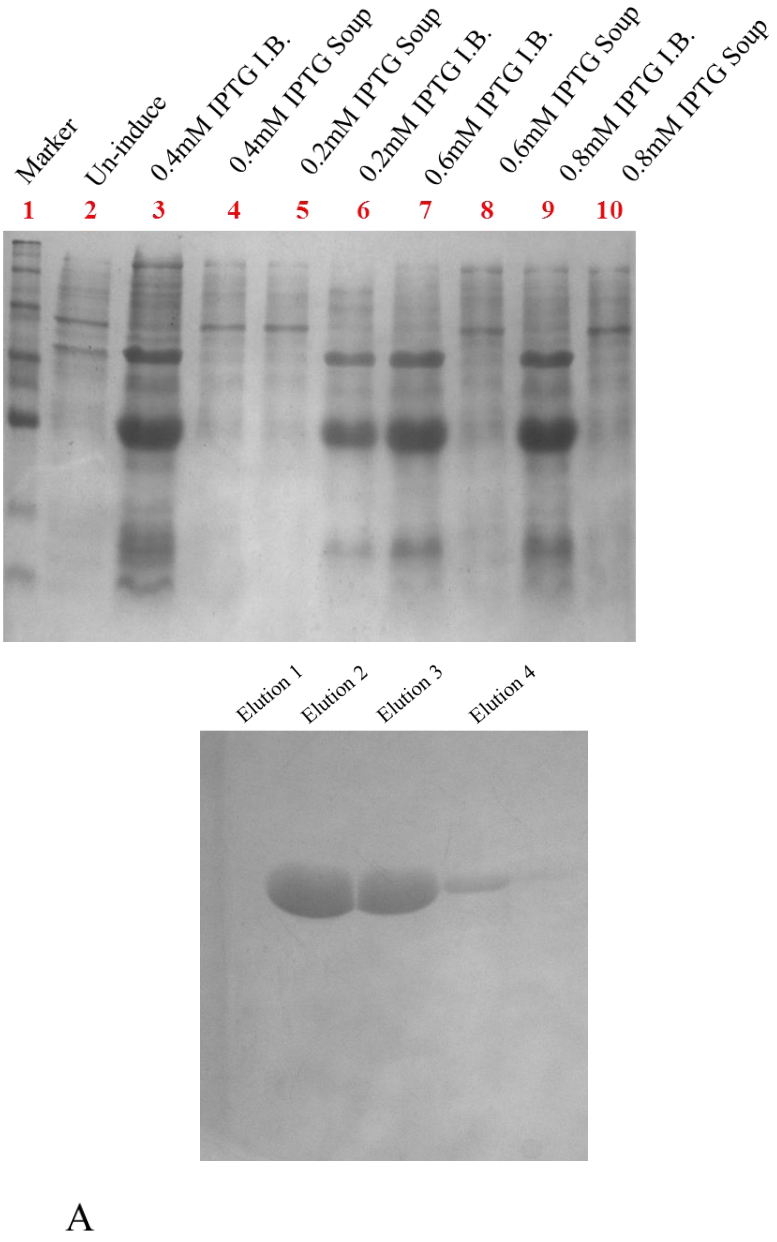


Figure 3 Standardization and purification of the CutC protein. [a] **SDS PAGE Analysis of CutC at different concentration of IPTG induction (0.2, 0.4, 0.6 and 0.8 mM) at 37 ^{0}C. [b] Purified CutC protein after Ni-NTA followed by gel filtration.**

The best condition selected was incubation of the cell culture at 37^{o}C until OD_{600} reaches around 1 to have more cell density. Then the culture was induced with 0.2 mM IPTG concentration and incubated at 10^{o}C for 16-18 hr. The protein was obtained in the soluble fraction and purified by the Ni-NTA affinity column followed by the gel filtration column. The purity of the CutC protein was analysed by the SDS-PAGE analysis

(Fig.3b). The molecular weight was also estimated to be ~25 kDa in the SDS-PAGE (Fig.3b) and MALDI-TOF analysis (Fig.4).

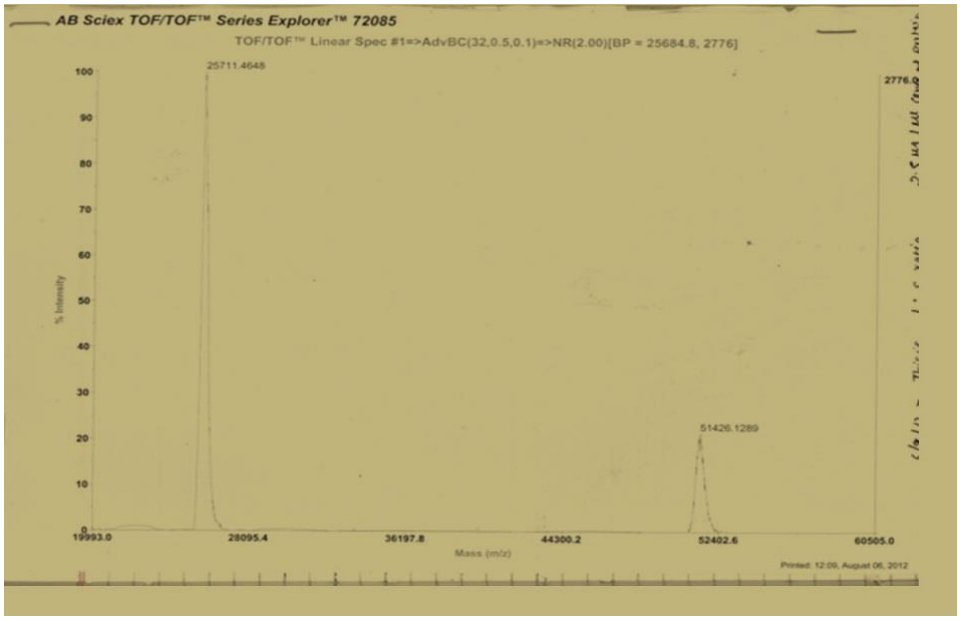


Figure 4 MALDI-TOF analysis for molecular weight estimation.

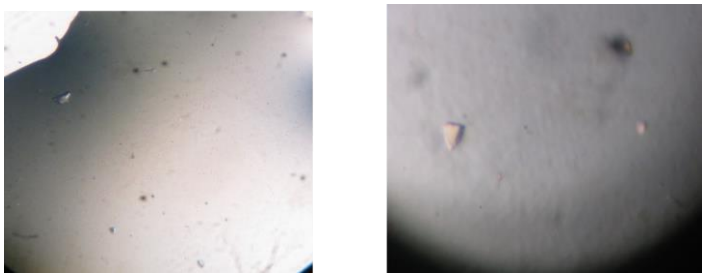


Figure 5 Crystallization of the CutC protein in the 50mM sodium acetate and 20 %PEG3350

The protein was concentrated upto 10-12 mg/ml and subjected to screening of crystallization condition. Few hits were obtained and carried forward for the improvement of the crystal.

The crystallization condition 50 mM sodium acetate and 20% PEG3350 was found to be good in which triangle shaped crystals were obtained (Fig.5). When the crystals were exposed to X-rays no diffraction was observed. Several batches of proteins were prepared and tried for crystallization. Even after extensive efforts no diffraction quality crystals could be obtained.

Bibliography

Bibliography

- Abadjieva, A., Hilven, P., Pauwels, K. & Crabeel, M. (2000). The yeast ARG7 gene product is autoproteolyzed to two subunit peptides, yielding active ornithine acetyltransferase. J. Biol. Chem. 275, 11361-11367.
- Adams, P. D., Afonine, P. V., Bunkóczi, G. & other authors (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D: Biological Crystallography 66, 213-221.
- Ajouz, H., Mukherji, D. & Shamseddine, A. (2014). Secondary bile acids: an underrecognized cause of colon cancer. World Journal of Surgical Oncology 12, (1) 1.
- Alkema, W. B., Hensgens, C. M., Kroezinga, E. H., de Vries, E., Floris, R., van der Laan, J.-M., Dijkstra, B. W. & Janssen, D. B. (2000). Characterization of the β-lactam binding site of penicillin acylase of Escherichia coli by structural and site-directed mutagenesis studies. Protein engineering 13, 857-863.
- Alkema, W. B., Hensgens, C. M., Snijder, H. J., Keizer, E., Dijkstra, B. W. & Janssen, D. B. (2004). Structural and kinetic studies on ligand binding in wild-type and active-site mutants of penicillin acylase. Protein Engineering Design and Selection 17, 473-480.
- Anderson J. W. & Gilliland S. E. (1999). Effect of fermented milk (yogurt) containing Lactobacillus acidophilus L1 on serum cholesterol in hypercholesterolemic humans. Journal of the American College of Nutritio;18:43-50.
- Aries, V. & Hill, M. (1970). Degradation of steroids by intestinal bacteria I. Deconjugation of bile salts. Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism 202, 526-534.
- Arroyo, M., De, M, I., Acebal, C. & Castillon, M. P. (2003). Biotechnological applications of penicillin acylases: state-of-the-art. Applied microbiology and biotechnology 60, 507-514.
- Atul K. Prabhune, A. A. Suresh C. G. and A. V. Pundle (2008) Characterization of smallest active monomeric penicillin V acylase from new source: a yeast, Rhodotorula aurantiaca (NCIM 3425) Process Biochemistry, 43, 961-967

- Avinash V.S., Ramasamy S, Suresh, C. G., Pundle A.V. (2015). Penicillin V acylase from Pectobacterium atrosepticum exhibits high specific activity and unique kinetics. International journal of biological macromolecules, 79:1-7.
- Avinash, V. S., Panigrahi, P., Chand, D., Pundle, A.V., Suresh, C. G., & Ramasamy, S. (2016). Structural analysis of a penicillin V acylase from Pectobacterium atrosepticum confirms the importance of two Trp residues for activity and specificity. Journal of structural biology, 193(2), 85-94.
- Avinash, V. S., Panigrahi, P., Suresh, C. G., Pundle, A. V., & Ramasamy, S. (2013). Structural modelling of substrate binding and inhibition in penicillin V acylase from Pectobacterium atrosepticum. Biochemical and biophysical research communications, 437(4), 538-543.
- Avinash, V. S., Pundle, A. V., Ramasamy, S. & Suresh, C. G. (2014). Penicillin acylases revisited: importance beyond their industrial utility. Critical reviews in biotechnology, 1-14.
- Avinash, V. S., Pundle, A. V., Ramasamy, S. & Suresh, C. G. (2016). Penicillin acylases revisited: importance beyond their industrial utility. Critical reviews in biotechnology 36, 303-316.
- Barbero, J. L., Buesa, J. M., Gonzalez de Buitrago, G., Mendez, E., Pez-Aranda, A. & Garcia, J. L. (1986). Complete nucleotide sequence of the penicillin acylase gene from Kluyvera citrophila. Gene, 49, 69-80.
- Bassler, B. L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. Current opinion in microbiology 2, 582-587.
- Bateup, J. M., McConnell, M. A., Jenkinson, H. F. & Tannock, G. W. (1995). Comparison of Lactobacillus strains with respect to bile salt hydrolase activity, colonization of the gastrointestinal tract, and growth rate of the murine host. Applied and environmental microbiology 61, 1147-1149.
- Batta, A., Salen, G. & Shefer, S. (1984). Substrate specificity of cholylglycine hydrolase for the hydrolysis of bile acid conjugates. Journal of Biological Chemistry 259, 15035-15039.
- Begley M, Gahan C. G, Hill C. (2005). The interaction between bacteria and bile. FEMS microbiology reviews, 29:625-651.
- Begley M, Hill C, Gahan C. G. (2006). Bile salt hydrolase activity in probiotics. Applied and environmental microbiology, 72:1729-1738.

- Bennion, L. J. & Grundy, S. M. (1977). Effects of diabetes mellitus on cholesterol metabolism in man. New England Journal of Medicine 296, 1365-1371.
- Bhagavan, N. V. (2002). Medical biochemistry: Academic press.
- Bi J, Fang F., Lu S., Du G., Chen J. (2013) New insight into the catalytic properties of bile salt hydrolase. Journal of Molecular Catalysis B: Enzymatic, 96:46-51.
- Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1965). Structure of hen egg-white lysozyme: a three-dimensional Fourier synthesis at 2 Å resolution. Nature, 206(4986), 757-761.
- Blundell, T. L. & Johnson, L. N. (1976). Protein crystallography: Academic Press.
- Bokhove, M., Jimenez, P. N., Quax, W. J. & Dijkstra, B. W. (2010). The quorum-quenching N-acyl homoserine lactone acylase PvdQ is an Ntn-hydrolase with an unusual substrate-binding pocket. Proceedings of the National Academy of Sciences 107, 686-691.
- Bomstein, J. & Evans, W. G. (1965). Automated Colorimetric Determination of 6-Aminopenicillanic Acid in Fermentation Media. Analytical chemistry 37, 576-578.
- Bongaerts, G. P., Severijnen, R. S., Tangerman, A., Verrips, A. & Tolboom, J. J. (2000). Bile acid deconjugation by Lactobacilli and its effects in patients with a short small bowel. Journal of gastroenterology 35, 801-804.
- Bossi, A., Cretich, M. & Righetti, P. G. (1998). Production of D-phenylglycine from racemic (D,L)-phenylglycine via isoelectrically-trapped penicillin G acylase. Biotechnol. Bioeng. 60, 454-461.
- Bradford M. M.(1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical biochemistry, 72:248-254.
- Brange, J. (1987). Insulin preparations. In Galenics of Insulin. Springer Berlin Heidelberg. (pp. 17-73).
- Brannigan, J., Dodson, G., Duggleby, H., Moody, P., Smith, J., Tomchick, D. & Murzin, A. (1995). A Protein Catalytic Framework with an N-Terminal Nucleophile Is Capable of Self-Activation (Vol 378, Pg 416, 1995). Nature 378, 644-644.

- Bringel, F., Quénée, P. & Tailliez, P. (2001). Polyphasic investigation of the diversity within Lactobacillus plantarum related strains revealed two Lactobacillus plantarum subgroups. Systematic and applied microbiology 24, 561-571.
- Brufau, G., Bahr, M. J., Staels, B. & other authors (2010). Plasma bile acids are not associated with energy metabolism in humans. Nutrition & metabolism 7, 1.
- Bruggink, A., Roos, E. C. & de Vroom, E. (1998). Penicillin acylase in the industrial production of β-lactam antibiotics. Organic Process Research & Development 2, 128-133.
- Brunger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature, 355, 472-475.
- Bull, P. C., & Cox, D. W. (1994). Wilson disease and Menkes disease: new handles on heavy-metal transport. Trends in Genetics, 10(7), 246-252.
- Cai, G., Zhu, S., Yang, S., Zhao, G., & Jiang, W. (2004). Cloning, overexpression, and characterization of a novel thermostable penicillin G acylase from Achromobacter xylosoxidans: probing the molecular basis for its high thermostability. Applied and environmental microbiology, 70(5), 2764-2770.
- Carnes E. C, Lopez DM, Donegan N. P, Cheung A, Gresham H, et al. (2010) Confinement-induced quorum sensing of individual Staphylococcus aureus bacteria. Nat Chem Biol 6: 41–45. doi: 10.1038/nchembio.264
- Carter Jr, C. W., & Carter, C. W. (1979). Protein crystallization using incomplete factorial experiments. J. Biol. Chem, 254(23), 12219-12223
- Castillo, A. (2015).How Bacteria Use Quorum Sensing to Communicate. Nature Education 8 (2): 4
- Chae J, Valeriano V, Kim G. B, Kang D. K. Molecular cloning, characterization and comparison of bile salt hydrolases from Lactobacillus johnsonii PF01. Journal of applied microbiology 2013;114:121-133.
- Chand, D., Ramasamy, S. & Suresh, C. G. (2016). A highly active bile salt hydrolase from Enterococcus faecalis shows positive cooperative kinetics. Process Biochemistry. 51:263–269
- Chand, D., Varshney, N., Ramasamy, S., Panigrahi, P., Brannigan, J. A., Wilkinson, A. J. & Suresh, C. (2015b). Structure mediation in substrate binding and post-translational processing of penicillin acylases: Information from mutant structures of Kluyvera citrophila penicillin G acylase. Protein Science 24, 1660-1670.

- Chandra, P. M., Brannigan, J. A., Prabhune, A., Pundle, A., Turkenburg, J. P., Dodson, G. G. & Suresh, C. (2005). Cloning, preparation and preliminary crystallographic studies of penicillin V acylase autoproteolytic processing mutants. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 61, 124-127.
- Chayen, N. E. (1998). Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques. Acta Crystallogr. Sect.D: Biol. Crystallogr. 54, 8-15.
- Chiang, C. & Bennett, R. E. (1967). Purification and Properties of Penicillin Amidase from Bacillus megaterium. J. Bacteriol. 93, 302-308.
- Chiang, J. Y. (2002). Bile acid regulation of gene expression: roles of nuclear hormone receptors. Endocrine reviews 23, 443-463.
- Chiang, J. Y. (2009). Bile acids: regulation of synthesis. Journal of lipid research 50, 1955-1966.
- Chikai T, Nakao H, Uchida K. (1987). Deconjugation of bile acids by human intestinal bacteria implanted in germ-free rats. Lipids, 22:669-671.
- Choi, K. S., Kim, J. A. & Kang, H. S. (1992). Effects of site-directed mutations on processing and activities of penicillin G acylase from Escherichia coli ATCC 11105. Journal of bacteriology 174, 6270-6276.
- Choi, S. B., Lew, L. C., Yeo, S. K., Nair, P. S. & Liong, M. T. (2015). Probiotics and the BSH-related cholesterol lowering mechanism: a Jekyll and Hyde scenario. Critical reviews in biotechnology 35, 392-401.
- Christiaens, H., Leer, R., Pouwels, P. & Verstraete, W. (1992). Cloning and expression of a conjugated bile acid hydrolase gene from Lactobacillus plantarum by using a direct plate assay. Applied and environmental microbiology 58, 3792-3798.
- Coleman, J. P. & Hudson, L. L. (1995). Cloning and characterization of a conjugated bile acid hydrolase gene from Clostridium perfringens. Applied and environmental microbiology 61, 2514-2520.
- Collaborative, C. P. (1994). The CCP4 suite: programs for protein crystallography. Acta crystallographica Section D, Biological crystallography 50, 760.
- Compare, D., Coccoli, P., Rocco, A., Nardone, O., De Maria, S., Cartenì, M. & Nardone, G. (2012). Gut–liver axis: the impact of gut microbiota on non alcoholic

fatty liver disease. Nutrition, Metabolism and Cardiovascular Diseases 22, 471-476.

- Copley, S. D. (2003). Enzymes with extra talents: moonlighting functions and catalytic promiscuity. Current opinion in chemical biology, 7(2), 265-272.
- Corzo G, Gilliland S. (1999). Bile salt hydrolase activity of three strains of Lactobacillus acidophilus. Journal of Dairy Science, 82:472-480.
- Corzo, G. & Gilliland, S. (1999). Measurement of bile salt hydrolase activity from Lactobacillus acidophilus based on disappearance of conjugated bile salts. Journal of dairy science 82, 466-471.
- Crowther, R. A. & Blow, D. M. (1967). A method of positioning a known molecule in an unknown crystal structure. Acta Crystallogr. 23, 544-548.
- Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. V. O. N. N. E., & McPherson, A. (1994). Screening and optimization strategies for macromolecular crystal growth. Acta Crystallographica Section D: Biological Crystallography, 50(4), 414-423.
- Davies, M. J., Tringham, J., Troughton, J. & Khunti, K. (2004). Prevention of Type 2 diabetes mellitus. A review of the evidence and its application in a UK setting. Diabetic Medicine 21, 403-414.
- de Souza, V. R., Silva, A. C. G., Pinotti, L. M., Araujo, H. S. S. & Giordano, R. D. C. (2005). Characterization of the penicillin G acylase from Bacillus megaterium ATCC 14945. Braz. Arch. Biol. Technol. 48, 105-111.
- Deaguero, A. L., Blum, J. K. & Bommarius, A. S. (2012). Improving the diastereoselectivity of penicillin G acylase for ampicillin synthesis from racemic substrates. Protein Engineering Design and Selection 25, 135-144.
- Dean, M., Cervellati, C., Casanova, E., Squerzanti, M., Lanzara, V., Medici, A., de Laureto, P. P. & Bergamini, C. M. (2002). Characterization of cholylglycine hydrolase from a bile-adapted strain of Xanthomonas maltophilia and its application for quantitative hydrolysis of conjugated bile salts. Applied and environmental microbiology 68, 3126-3128.
- Delpino, M. V., Marchesini, M. I., Estein, S. M., Comerci, D. J., Cassataro, J., Fossati, C. A. & Baldi, P. C. (2007). A bile salt hydrolase of Brucella abortus contributes to the establishment of a successful infection through the oral route in mice. Infection and immunity 75, 299-305.

- DiBaise, J. K., Zhang, H., Crowell, M. D., Krajmalnik-Brown, R., Decker, G. A., & Rittmann, B. E. (2008). Gut microbiota and its possible relationship with obesity. In Mayo Clinic Proceedings (Vol. 83, No. 4, pp. 460-469).
- Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat. Struct. Biol. 4, 269-275.
- Dijkstra, K. D. B., Bechly, G., Bybee, S. M., Dow, R. A., Dumont, H. J., Fleck, G., & May, M. L. (2013). The classification and diversity of dragonflies and damselflies (Odonata). In: Zhang, Z.-Q.(Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013). Zootaxa, 3703(1), 36-45.
- Ditzel, L., Huber, R., Mann, K., Heinemeyer, W., Wolf, D. H. & Groll, M. (1998). Conformational constraints for protein self-cleavage in the proteasome. Journal of molecular biology 279, 1187-1191.
- Dobrindt, U., Hochhut, B., Hentschel, U. & Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. Nature Reviews Microbiology 2, 414-424.
- Dodson, G. G. (2000). Catalysis in penicillin G amidase - a member of the Ntn (N-terminal nucleophile) hydrolase family. Croatica Chemica. Acta. 73, 901-908.
- Drenth, J. & Haas, C. (1998). Nucleation in protein crystallization. Acta Crystallogr. Sect. D: Biol. Crystallogr. 54, 867-872.
- Drucker, D. J. (2006). The biology of incretin hormones. Cell metabolism 3, 153-165.
- Duggleby, Helen. J., Shirley P. Tolley, Christopher P. H, Eleanor J. D, Guy D. and Peter C. E. M. (1995). "Penicillin acylase has a single-amino-acid catalytic centre." 264-268.
- Dussurget, O., Cabanes, D., Dehoux, P., Lecuit, M., Buchrieser, C., Glaser, P. & Cossart, P. (2002). Listeria monocytogenes bile salt hydrolase is a PrfA-regulated virulence factor involved in the intestinal and hepatic phases of listeriosis. Molecular microbiology 45, 1095-1106.
- E. Bolton, Y. Wang, P.A. Thiessen, S.H. Bryant, PubChem: Integrated platform of mall molecules and biological activities, in: Annual Reports in Computational Chemistry, Elsevier, Washington DC, 2008, pp. 217–241

- Edelstein, S. L., Knowler, W. C., Bain, R. P. & other authors (1997). Predictors of progression from impaired glucose tolerance to NIDDM: an analysis of six prospective studies. Diabetes 46, 701-710.
- Edgell, D. R., Belfort, M. & Shub, D. A. (2000). Barriers to intron promiscuity in bacteria. Journal of Bacteriology 182, 5281-5289.
- Elkins, C. A. & Savage, D. C. (1998). Identification of genes encoding conjugated bile salt hydrolase and transport in Lactobacillus johnsonii100-100. Journal of bacteriology 180, 4344-4349.
- Elkins, C. A., Moser, S. A. & Savage, D. C. (2001). Genes encoding bile salt hydrolases and conjugated bile salt transporters in Lactobacillus johnsonii 100-100 and other Lactobacillus species. Microbiology 147, 3403-3412.
- Elkins, J. M., Kershaw, N. J. & Schofield, C. J. (2005). X-ray crystal structure of ornithine acetyltransferase from the clavulanic acid biosynthesis gene cluster. Biochem. J. 385, 565-573.
- Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography 60, 2126-2132.
- Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. Sect.D: Biol. Crystallogr. 62, 72-82.
- Eyssen, H. (1973). Role of the gut microflora in metabolism of lipids and sterols. Proceedings of the Nutrition Society, 32(02), 59-63.F.
- Huber, B. Jackson, (1972). Preparative methods for 7-aminocephalosporanic acid and 6-aminopenicillanic acid in cephalosporins and penicillins, in: E. Flynn (Ed.), Chemistry and Biology, Academic press, New York, pp. 27–48.
- Fadnavis, N. W., Sharfuddin, M., Vadivel, S. K. & Bhalerao, U. T. (1997). Efficient chemoenzymatic synthesis of (2S,3S)-3-hydroxyleucine mediated by immobilised penicillin G acylase. J. Chem. Soc. Perkin Trans. 1, 3577-3578.
- Fang, F., Li, Y., Bumann, M., Raftis, E. J., Casey, P. G., Cooney, J. C., Walsh, M. A. & O'Toole, P. W. (2009). Allelic variation of bile salt hydrolase genes in Lactobacillus salivarius does not determine bile resistance levels. Journal of bacteriology 191, 5743-5757.
- Feighner, S. D. & Dashkevicz, M. P. (1987). Subtherapeutic levels of antibiotics in poultry feeds and their effects on weight gain, feed efficiency, and bacterial

cholyltaurine hydrolase activity. Applied and Environmental Microbiology 53, 331-336.

- Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., & Mainz, D. T. (2006). Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. Journal of medicinal chemistry, 49(21), 6177-6196.
- Woolcock, J. B. (1991). Microbiology of animals and animal products. World Animal Science, A6 (No. 6).
- Fuqua, C. & Greenberg, E. P. (1998). Self-perception in bacteria: quorum sensing with acylated homoserine lactones. Current opinion in microbiology 1, 183-189.
- Fuqua, W. C., Winans, S. C. & Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. Journal of bacteriology 176, 269.
- Gaisford, Wendy, G. F. Schertler, and Pat Edwards. "mosquito LCP: Making membrane protein crystallization accesible to the research scientist." Nature Methods 8 (2011).
- Garman, E. F. & Schneider, T. R. (1997). Macromolecular cryocrystallography. J. Appl. Crystallogr. 30, 211-237.
- Glide, version 5.8, Schrodinger, LLC, New York, NY, 2012.
- González-Vázquez, R., Azaola-Espinosa, A., Mayorga-Reyes, L., Reyes-Nava, L., Shah, N. & Rivera-Espinoza, Y. (2015). Isolation, Identification and Partial Characterization of a Lactobacillus casei Strain with Bile Salt Hydrolase Activity from Pulque. Probiotics and antimicrobial proteins 7, 242-248.
- Gopal-Srivastava, R. & Hylemon, P. B. (1988). Purification and characterization of bile salt hydrolase from Clostridium perfringens. Journal of lipid research 29, 1079-1085.
- Grill, J., Schneider, F., Crociani, J. & Ballongue, J. (1995). Purification and Characterization of Conjugated Bile Salt Hydrolase from Bifidobacterium longum BB536. Applied and environmental microbiology 61, 2577-2582.
- Groll, M., & Huber, R. (2003). Substrate access and processing by the 20S proteasome core particle. The international journal of biochemistry & cell biology, 35(5), 606-616.

- Gu X. C, Luo X. G, Wang C. X, Ma D. Y, Wang Y, He Y.Y, Li W. Zhou H., Zhang T. C. (2014). Cloning and analysis of bile salt hydrolase genes from Lactobacillus plantarum CGMCC No. 8198. Biotechnology letters, 36:975-983.
- Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L. & Comb, D. (1996). Activation of Glycosylasparaginase formation of active n-terminal threonine by intramolecular autoproteolysis. Journal of Biological Chemistry 271, 1732-1737.
- Guan, C., Liu, Y., Shao, Y., Cui, T., Liao, W., Ewel, A. & Paulus, H. (1998). Characterization and functional analysis of the cis-autoproteolysis active center of glycosylasparaginase. Journal of Biological Chemistry, 273(16), 9695-9702.
- Guo, C., Zhang, L., Han, X., Li, J., Du, M., Yi, H., Feng, Z., Zhang, Y. & Xu, X. (2011). Short communication: A sensitive method for qualitative screening of bile salt hydrolase-active lactobacilli based on thin-layer chromatography. Journal of dairy science 94, 1732-1737.
- Gupta, S. D., Lee, B. T., Camakaris, J., & Wu, H. C. (1995). Identification of cutC and cutF (nlpE) genes involved in copper tolerance in Escherichia coli. Journal of bacteriology, 177(15), 4207-4215.
- Halgren, T. (2007). New method for fast and accurate binding-site identification and analysis. Chem Biol Drug Des 69, 146-148
- Hamilton-Miller, J. (1966). Penicillinacylase. Bacteriological reviews 30, 761.
- Hammer, B. K. & Bassler, B. L. (2003). Quorum sensing controls biofilm formation in Vibrio cholerae. Molecular microbiology 50, 101-104.
- Hershko, A., & Ciechanover, A. (1992). The ubiquitin system for protein degradation. Annual review of biochemistry, 61(1), 761-807.
- Hevehan, D. L., & De Bernardez Clark, E. (1997). Oxidative renaturation of lysozyme at high concentrations. Biotechnology and bioengineering, 54(3), 221-230.
- Hewitt, L., Kasche, V., Lummer, K., Lewis, R., Murshudov, G., Verma, C., Dodson, G. & Wilson, K. (2000). Structure of a slow processing precursor penicillin acylase from Escherichia coli reveals the linker peptide blocking the active-site cleft. Journal of molecular biology 302, 887-898.
- Hill A. V. (1910).The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol (Lond) ;40:4-7.
- Hill, C. (2012). Virulence or niche factors: what's in a name? Journal of bacteriology 194, 5725-5727.

- Holm, L., & Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic acids research, 38(suppl 2), W545-W549.
- Homayouni, A., Payahoo, L. & Azizi, A. (2012). Effects of probiotics on lipid profile: a review. Am J Food Technol, 7 (5).
- Huber, Eva M., Basler, M., Schwab, R., Heinemeyer, W., Kirk, Christopher J., Groettrup, M. & Groll, M. (2012). Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. Cell, 148, 727-738.
- Huber, F., Chauvette, R. R., & Jackson, B. G. (1972). Preparative methods for 7-aminocephalosporanic acid and 6-aminopenicillanic acid.Cephalosporins and penicillins, chemistry and biology. Academic, New York, 27-48.
- Hult, K., & Berglund, P. (2007). Enzyme promiscuity: mechanism and applications. Trends in biotechnology, 25(5), 231-238.
- Ishibashi, S., Schwarz, M., Frykman, P. K., Herz, J. & Russell, D. W. (1996). Disruption of Cholesterol 7α-Hydroxylase Gene in Mice. Journal of Biological Chemistry 271, 18017-18023.
- Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denisot, M.-A., Badet, B., Polikarpov, I., Littlechild, J. A. & Teplyakov, A. (1996). Substrate binding is required for assembly of the active conformation of the catalytic site in Ntn amidotransferases: evidence from the 1.8 Å crystal structure of the glutaminase domain of glucosamine 6-phosphate synthase. Structure 4, 801-810.
- Jacobsen C, Garside J, Hoare M (1997) Nucleation and growth of microbial lipase crystals from clarified concentrated fermentation broths. Biote chnol Bioeng 57: 666–675.
- Jancarik, J. A. K. S., & Kim, S. H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. Journal of applied crystallography, 24(4), 409-411.
- 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, B. V., Begley, M., Hill, C., Gahan, C. G. & Marchesi, J. R. (2008). Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. Proceedings of the national academy of sciences 105, 13580-13585.

- Jose, L. M. H., Iliyna, A., Dominguez, M. L., Olga, S. C. & Dustet, M. J. C. (2003). Partial characterisation of penicillin acylase from fungi Aspergillus fumigatus and Mucor gryseocianum. Moscow Univ. Chem. Bull. 44, 53-56.
- Joyce, S. A., MacSharry, J., Casey, P. G., Kinsella, M., Murphy, E. F., Shanahan, F., Hill, C. & Gahan, C. G. (2014). Regulation of host weight gain and lipid metabolism by bacterial bile acid modification in the gut. Proceedings of the National Academy of Sciences 111, 7421-7426.
- Judge RA, Johns MR, White ET (1995) Protein purification: The recovery of ovalbumin. Biotechnol Bioeng 48: 316–323.
- Kabsch, W. (2010). Xds. Acta Crystallographica Section D: Biological Crystallography 66, 125-132.
- Kasche, V., Lummer, K., Nurk, A., Piotraschke, E., Rieks, A., Stoeva, S. & Voelter, W. (1999). Intramolecular autoproteolysis initiates the maturation of penicillin amidase from Escherichia coli. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 1433, 76-86.
- Kawamata, Y., Fujii, R., Hosoya, M. & other authors (2003). AG protein-coupled receptor responsive to bile acids. Journal of Biological Chemistry 278, 9435-9440.
- Kawamoto, K., Horibe, I. & Uchida, K. (1989). Purification and characterization of a new hydrolase for conjugated bile acids, chenodeoxycholyltaurine hydrolase, from Bacteroides vulgatus. Journal of biochemistry 106, 1049-1053.
- Kim G-B, Brochet M, Lee BH. (2005). Cloning and characterization of a bile salt hydrolase (bsh) from. Biotechnology letters, 27:817-822.
- Kim GB, Lee B. (2008). Genetic analysis of a bile salt hydrolase in Bifidobacterium animalis subsp. lactis KL612 . Journal of applied microbiology, 105:778-790.
- Kim, G. & Lee, B. H. (2005). Biochemical and molecular insights into bile salt hydrolase in the gastrointestinal microflora-a review. Asian Australasian Journal of Animal Sciences 18, 1505.
- Kim, G. B., Yi, S. H. & Lee, B. (2004). Purification and characterization of three different types of bile salt hydrolases from Bifidobacterium strains. Journal of dairy science 87, 258-266.
- Kim, J. K., Yang, I. S., Shin, H. J., Cho, K. J., Ryu, E. K., Kim, S. H. & Kim, K. H. (2006). Insight into autoproteolytic activation from the structure of cephalosporin acylase: a protein with two proteolytic chemistries. Proceedings of

the National Academy of Sciences of the United States of America, 103(6), 1732-1737.

- Kim, J. K., Yang, I. S., Shin, H. J., Cho, K. J., Ryu, E. K., Kim, S. H., Park, S. S., Kim, K. H. (2006) Insight into autoproteolytic activation from the structure of cephalosporin acylase: a protein with two proteolytic chemistries. Proc.Natl.Acad.Sci.USA 103: 1732-1737
- Kim, Y., & Hol, W. G. (2001). Structure of cephalosporin acylase in complex with glutaryl-7-aminocephalosporanic acid and glutarate: insight into the basis of its substrate specificity. Chemistry & biology, 8(12), 1253-1264.
- Kim, Y., Yoon, K.-H., Khang, Y., Turley, S. & Hol, W. G. (2000). The 2.0 Å crystal structure of cephalosporin acylase. Structure 8, 1059-1068.
- Klei, H. E., Daumy, G. O. & Kelly, J. A. (1995). Purification and preliminary crystallographic studies of penicillin G acylase from Providencia rettgeri. Protein Sci. 4, 433-441.
- Klyushnichenko V (2003) Protein crystallization: From HTS to kilogram-scale. Curr Opin Drug Discovery Dev 6: 848–854.
- Knarreborg, A., Engberg, R. M., Jensen, S. K. & Jensen, B. B. (2002). Quantitative determination of bile salt hydrolase activity in bacteria isolated from the small intestine of chickens. Applied and Environmental Microbiology 68, 6425-6428.
- Kotera, M., Hirakawa, M., Tokimatsu, T., Goto, S., & Kanehisa, M. (2012). The KEGG databases and tools facilitating omics analysis: latest developments involving human diseases and pharmaceuticals. Next Generation Microarray Bioinformatics: Methods and Protocols, 19-39.
- Kumar, M., Nagpal, R., Kumar, R. & other authors (2012). Cholesterol-lowering probiotics as potential biotherapeutics for metabolic diseases. Experimental diabetes research 2012.
- Kumar, R. S., Brannigan, J. A., Prabhune, A. A., Pundle, A. V., Dodson, G. G., Dodson, E. J. & Suresh, C. G. (2006). Structural and functional analysis of a conjugated bile salt hydrolase from Bifidobacterium longum reveals an evolutionary relationship with penicillin V acylase. Journal of Biological Chemistry 281, 32516-32525.

- Lambert, J. M., Siezen, R. J., de Vos, W. M. & Kleerebezem, M. (2008). Improved annotation of conjugated bile acid hydrolase superfamily members in Gram-positive bacteria. Microbiology 154, 2492-2500.
- Larsen C, T., (2000). Blood pressure level and relation to other cardiovascular risk factors in male hypertensive patients without clinical evidence of ischemic heart disease. Blood pressure, 9:91-97.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26, 283-291.
- Lee, Y. S. & Park, S. S. (1998). Two-step autocatalytic processing of the glutaryl 7-aminocephalosporanic acid acylase from Pseudomonas sp. strain GK16. J. Bacteriol. 180, 4576-4582.
- Leslie, A. G. W. & Powell, H. R. (2007). Processing diffraction data with MOSFLM. In Read, R. J. & Sussman, J. L. (eds.) Evolving Methods for Macromolecular Crystallography, 245, Springer Press, 41-51
- Li, Y., Chen, J. F., Jiang, W. H., Mao, X., Zhao, G. P. & Wang, E. D. (1999). In vivo post-translational processing and subunit reconstitution of cephalosporin acylase from Pseudomonas sp 130. Eur. J. Biochem. 262, 713-719.
- LigPrep, version 2.5, Schrodinger, LLC, New York, NY, 2012.
- Lin, J. (2007). Antibiotic growth promoters enhance animal production by targeting intestinal bile salt hydrolase and its producers. Low-dose antibiotics: current status and outlook for the future, 129.
- Liong, M. & Shah, N. (2005). Bile salt deconjugation ability, bile salt hydrolase activity and cholesterol co-precipitation ability of lactobacilli strains. International Dairy Journal 15, 391-398.
- Lodola, A., Branduardi, D., De V. M., Capoferri, L., 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.
- Lowe, A. A., Santamaria, J. D., Fleetham, J. A., & Price, C. (1986). Facial morphology and obstructive sleep apnea. American Journal of Orthodontics and Dentofacial Orthopedics, 90(6), 484-491.
- Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., & Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. Science, 268(5210), 533-539.

- Lu, Y., Xiong, X., Wang, X. & other authors (2013). Yin Yang 1 promotes hepatic gluconeogenesis through upregulation of glucocorticoid receptor. Diabetes 62, 1064-1073.
- Lundeen, S. G. & Savage, D. C. (1990). Characterization and purification of bile salt hydrolase from Lactobacillus sp. strain 100-100. Journal of bacteriology 172, 4171-4177.
- Lundeen, S. G. & Savage, D. C. (1992). Multiple forms of bile salt hydrolase from Lactobacillus sp. strain 100-100. Journal of bacteriology 174, 7217-7220.
- Lye, H. S., Rahmat A. G. R. & Liong, M. T. (2010). Mechanisms of cholesterol removal by lactobacilli under conditions that mimic the human gastrointestinal tract. International Dairy Journal 20, 169-175.
- Ma, S., Devi, K., L. S. & Gao, J. (2007). Molecular dynamics simulations of the catalytic pathway of a cysteine protease: a combined QM/MM study of human cathepsin K. Journal of the American Chemical Society 129, 13633-13645.
- Magnusson, I., Rothman, D., Katz, L., Shulman, R. & Shulman, G. (1992). Increased rate of gluconeogenesis in type II diabetes mellitus. A 13C nuclear magnetic resonance study. Journal of Clinical Investigation 90, 1323.
- Marchesini, M. I., Connolly, J., Delpino, M. V., Baldi, P. C., Mujer, C. V., DelVecchio, V. G. & Comerci, D. J. (2011). Brucella abortus choloylglycine hydrolase affects cell envelope composition and host cell internalization. PloS one 6, e28480.
- Martin, J., Slade, A., Aitken, A., Arche, R. & Virden, R. (1991). Chemical modification of serine at the active-site of penicillin acylase from Kluyvera citrophila. Biochem. J. 280, 659-662.
- Martin, L., Prieto, M. A., Cortes, E. & Garcia, J. L. (1995). Cloning and sequencing of the pac gene encoding the penicillin G acylase of Bacillus megaterium ATCC 14945. FEMS Microbiol. Lett. 125, 287-292.
- Martín, R., Miquel, S., Ulmer, J., Kechaou, N., Langella, P. & Bermúdez-Humarán, L. G. (2013). Role of commensal and probiotic bacteria in human health: a focus on inflammatory bowel disease. Microbial cell factories 12, 1.
- Masuda, N. (1981). Deconjugation of bile salts by Bacteroides and Clostridium. Microbiology and immunology 25, 1-11.
- Matthews, B. W. (1968). Solvent content of protein crystals. J. Mol. Biol. 33, 491-497.

- Matthews, B. W. (1968). Solvent content of protein crystals. Journal of molecular biology 33, 491-497.
- Matthews, B. W. (1985). Determination of protein molecular weight, hydration, and packing from crystal density. Methods Enzymol. 114, 176-187.
- McAuliffe, O., Cano, R. J. & Klaenhammer, T. R. (2005). Genetic analysis of two bile salt hydrolase activities in Lactobacillus acidophilus NCFM. Applied and environmental microbiology 71, 4925-4929.
- McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallographica Section D: Biological Crystallography 63, 32-41.
- McGarr, S. E., Ridlon, J. M. & Hylemon, P. B. (2005). Diet, anaerobic bacterial metabolism, and colon cancer: a review of the literature. Journal of clinical gastroenterology 39, 98-109.
- McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr. Sect.D: Biol. Crystallogr. 67, 386-394.
- McPherson, A. (1982). Preparation and analysis of protein crystals. John Wiley & Sons.
- McPherson, A. (1999). Crystallization of biological macromolecules. Cold Spring Harbor Laboratory Press.
- McVey, C. E., Walsh, M. A., Dodson, G. G., Wilson, K. S. & Brannigan, J. A. (2001). Crystal structures of penicillin acylase enzyme-substrate complexes: structural insights into the catalytic mechanism. Journal of molecular biology 313, 139-150.
- Merino, E., Balbás, P., Recillas, F., Becerril, B., Valle, F. & Bolivar, F. (1992). Carbon regulation and the role in nature of the Escherichia coli penicillin acylase (pac) gene. Molecular microbiology 6, 2175-2182.
- Miller, M. B. & Bassler, B. L. (2001). Quorum sensing in bacteria. Annual Reviews in Microbiology 55, 165-199.
- Mladenovic, M., Fink, R. F., Thiel, W., Schirmeister, T. & Engels, B. (2008). On the origin of the stabilization of the zwitterionic resting state of cysteine proteases: a theoretical study. Journal of the American Chemical Society 130, 8696-8705.

- Molenaar, D., Bringel, F., Schuren, F. H., De V. W. M., Siezen, R. J. & Kleerebezem, M. (2005). Exploring Lactobacillus plantarum genome diversity by using microarrays. Journal of Bacteriology 187, 6119-6127.
- Monera, O. D., Kay, C. M., Hodges, R. S. (1994) Protein denaturation with guanidinehydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. Protein Sci. 3, 1984–91.
- Moser, S. A. & Savage, D. C. (2001). Bile salt hydrolase activity and resistance to toxicity of conjugated bile salts are unrelated properties in lactobacilli. Applied and Environmental Microbiology 67, 3476-3480.
- Mukherji, R. & Prabhune, A. (2015). Possible Correlation Between Bile Salt Hydrolysis and AHL Deamidation: Staphylococcus epidermidis RM1, a Potent Quorum Quencher and Bile Salt Hydrolase Producer. Applied biochemistry and biotechnology 176, 140-150.
- Mukherji, R., Varshney, N. K., Panigrahi, P., Suresh, C. G. & Prabhune, A. (2014). A new role for penicillin acylases: degradation of acyl homoserine lactone quorum sensing signals by Kluyvera citrophila penicillin G acylase. Enzyme and microbial technology 56, 1-7.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect.D: Biol. Crystallogr. 53, 240-255.
- Murzin, A. G. (1996). Structural classification of proteins: new superfamilies. Current opinion in structural biology 6, 386-394.
- Myerson A. S. & Toyokura K. (1990). Crystallization as a separation process. ACS Symposium Series 438. American Chemical Society, Washington,
- Nair, P., Gordon, M. & Reback, J. (1967). The enzymatic cleavage of the carbon-nitrogen bond in 3α, 7α, 12α-trihydroxy-5β-cholan-24-oylglycine. Journal of Biological Chemistry 242, 7-11.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. Sect. A: Foundn. Crystallogr. 50, 157-163.
- Nayler, J. H. (1991). Early discoveries in the penicillin series. Trends in biochemical sciences 16, 195-197.

- Ng W. L. & Bassler B. L. (2009). Bacterial quorum-sensing network architectures. Annu Rev Genet 43: 197–222. doi: 10.1146/annurev-genet-102108-134304
- Nobeli, I., Favia, A. D., & Thornton, J. M. (2009). Protein promiscuity and its implications for biotechnology. Nature biotechnology, 27(2), 157-167..
- O'Brien, P. J., & Herschlag, D. (1999). Catalytic promiscuity and the evolution of new enzymatic activities. Chemistry & biology, 6(4), R91-R105.
- Oh, H. K., Lee, J. Y., Lim, S. J., Kim, M. J., Kim, G. B., Kim, J. H., Hong, S. K. & Kang, D.-K. (2008). Molecular cloning and characterization of a bile salt hydrolase from Lactobacillus acidophilus PF01. J Microbiol Biotechnol 18, 449-456.
- Ohashi, H., Katsuta, Y., Hashizume, T., Abe, S. N., Kajiura, H., Hattori, H., Kamei, T. & Yano, M. (1988). Molecular cloning of the penicillin G acylase gene from Arthrobacter viscosus. Appl. Environ. Microbiol. 54, 2603-2607.
- Ohashi, H., Katsuta, Y., Nagashima, M., Kamei, T. & Yano, M. (1989). Expression of the Arthrobacter viscosus penicillin G acylase gene in Escherichia coli and Bacillus subtilis. Appl. Environ. Biotechnol. 55, 1351-1356.
- Oinonen, C. & Rouvinen, J. (2000). Structural comparison of Ntn-hydrolases. Protein Science 9, 2329-2337.
- Olsson, A., Hagstrom, T., Nilsson, B., Uhlen, M. & Gatenbeck, S. (1985). Molecular cloning of Bacillus sphaericus penicillin V amidase gene and its expression in Escherichia coli and Bacillus subtilis. Appl. Environ. Microbiol. 49, 1084-1089.
- Öner Ö, Aslim B, & Aydaş S. B. (2014). Mechanisms of cholesterol-lowering effects of lactobacilli and bifidobacteria strains as potential probiotics with their bsh gene analysis. Journal of molecular microbiology and biotechnology, 24:12-18.
- Ooi, L. G. & Liong, M. T. (2010). Cholesterol-lowering effects of probiotics and prebiotics: a review of in vivo and in vitro findings. International journal of molecular sciences 11, 2499-2522.
- Otwinowski, Z. & Minor W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods Enzymol. 276, Macromolecular Crystallography, part A, 307-326, Academic Press (New York).

- Padwal, R. & Laupacis, A. (2004). Antihypertensive Therapy and Incidence of Type 2 Diabetes A systematic review. Diabetes care 27, 247-255.
- Panigrahi, P., Sule, M., Sharma, R., Ramasamy, S. & Suresh, C. G. (2014). An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholylglycine hydrolase family. Microbiology 160, 1162-1174.
- Park S, Chung H, J. G, Ko Y, Jeong H, Yang J, & Kim Y. (1996). Effect of various lactic acid bacteria on the serum cholesterol levels in rats and resistance to acid, bile and antibiotics. Korean Journal of Applied Microbiology and Biotechnology (Korea Republic).
- Payen , A. , & Persoz , J. F. (1833) . Memory on the diastase , the main products of its reactions, and their applications to industrial arts. Ann . chem. Phys , 53 (2) , 73-92
- Payne, C. M., Bernstein, C., Dvorak, K. & Bernstein, H. (2008). Hydrophobic bile acids, genomic instability, Darwinian selection, and colon carcinogenesis. Clin Exp Gastroenterol 1, 19-47.
- Pereira D. I, McCartney A. L, Gibson G. R. (2003). An in vitro study of the probiotic potential of a bile-salt-hydrolyzing Lactobacillus fermentum strain, and determination of its cholesterol-lowering properties. Applied and environmental microbiology, 69:4743-4752.
- Pesci, E. C., Milbank, J. B., Pearson, J. P., McKnight, S., Kende, A. S., Greenberg, E. P. & Iglewski, B. H. (1999). Quinolone signaling in the cell-to-cell communication system of Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences 96, 11229-11234.
- Peters J, Minuth T, & Schroeder W (2005) Implementation of a crystallization step into the purification process of a recombinant protein. Protein Expression Purif 39: 43–53.
- Platt TG, & Fuqua C. (2010). The semantics of quorum sensing. Trends Microbiol 18: 383–387.
- Prieto, M. A., Diaz, E. & García, J. L. (1996). Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of Escherichia coli W: engineering a mobile aromatic degradative cluster. Journal of bacteriology 178, 111-120.
- Prisant, L. M. (2004). Preventing type II diabetes mellitus. The Journal of Clinical Pharmacology 44, 406-413.

- Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R. & Hess, B. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics, 29(7), 845-854.
- Puig, S., & Thiele, D. J. (2002). Molecular mechanisms of copper uptake and distribution. Current opinion in chemical biology, 6(2), 171-180.
- Pundle, A. & SivaRaman, H. (1997). Bacillus sphaericus penicillin V acylase: purification, substrate specificity, and active-site characterization. Current microbiology 34, 144-148.
- Rathinaswamy, P., Gaikwad, S. M., Suresh, C. G., Prabhune, A. A., Brannigan, J. A., Dodson, G. G. & Pundle, A. V. (2012). Purification and characterization of YxeI, a penicillin acylase from Bacillus subtilis. Int J Biol Macromol 50, 25–30.
- Rathinaswamy, P., Pundle, A. V., Prabhune, A. A., SivaRaman, H., Brannigan, J. A., Dodson, G. G. & Suresh, C. G.(2005). Cloning, purification, crystallization and preliminary structural studies of penicillin V acylase from Bacillus subtilis. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 61, 680-683.
- Ravelli, R. B. & Garman, E. F. (2006). Radiation damage in macromolecular cryocrystallography. Curr. Opin. Struct. Biol. 16, 624-629.
- Ravelli, R. B. G., Leiros, H. K. S., Pan, B. C., Caffrey, M. & McSweeney, S. (2003). Specific radiation damage can be used to solve macromolecular crystal structures. Structure, 11, 217-224.
- Rawlings, N. D., Barrett, A. J. & Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic acids research 40, 343-350.
- Reddy, G., Prakash, J., Vairamani, M., Prabhakar, S., Matsumoto, G. & Shivaji, S. (2002). Planococcus antarcticus and Planococcus psychrophilus spp. nov. isolated from cyanobacterial mat samples collected from ponds in Antarctica. Extremophiles 6, 253-261.
- Reitz, M., Sacher, O., Tarkhov, A., Trümbach, D., & Gasteiger, J. (2004). Enabling the exploration of biochemical pathways. Organic & biomolecular chemistry, 2(22), 3226-3237.
- Ren, J., Sun, K., Wu, Z., Yao, J. & Guo, B. (2011). All 4 Bile Salt Hydrolase Proteins Are Responsible for the Hydrolysis Activity in Lactobacillus plantarum ST-III. Journal of food science 76, M622-M628.

- Rensing, C., & Grass, G. (2003). Escherichia coli mechanisms of copper homeostasis in a changing environment. FEMS microbiology reviews, 27(2-3), 197-213.
- Rhodes, G. (2000). Crystallography Made Crystal Clear. Academic Press.
- Roche, D., Prasad, K. & Repic, O. (1999). Enantioselective acylation of beta-aminoesters using penicillin G acylase in organic solvents. Tetrahedron Lett. 40, 3665-3668.
- Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. Acta Crystallogr. 15, 24-31.
- Rossocha, M., Schultz-Heienbrok, R., Von Moeller, H., Coleman, J. P. & Saenger, W. (2005). Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product. Biochemistry 44, 5739-5748.
- Ruiz, L., Margolles, A. & Sánchez, B. (2014). Bile resistance mechanisms in Lactobacillus and Bifidobacterium. Microbial mechanisms of tolerance to weak acids: an overview, 65.
- Russell, D. W. & Setchell, K. D. (1992). Bile acid biosynthesis. Biochemistry 31, 4737-4749.
- Russell, D. W. (2003). The enzymes, regulation, and genetics of bile acid synthesis. Annual review of biochemistry 72, 137-174.
- Saarela, J., Oinonen, C., Jalanko, A., Rouvinen, J. & Peltonen, L. (2004). Autoproteolytic activation of human aspartylglucosaminidase. Biochemical journal 378, 363-371.
- Sanbonmatsu, K. Y., & Tung, C. S. (2007). High performance computing in biology: multimillion atom simulations of nanoscale systems. Journal of structural biology, 157(3), 470-480.
- Saier Jr, M. H., Beatty, J. T., Goffeau, A. & other authors (1999). The major facilitator superfamily. J Mol Microbiol Biotechnol 1, 257-279.
- Sakaguchi, K. & Murao, S. (1950). A preliminary report on a new enzyme, penicillinamidase. J. Agric. Chem. Soc.(Japan), 23; 411.
- Sakai, Y., Tsukahara, T., Bukawa, W., Matsubara, N., & Ushida, K. (2006). Cell preparation of Enterococcus faecalis strain EC-12 prevents vancomycin-resistant enterococci colonization in the cecum of newly hatched chicks.Poultry science, 85(2), 273-277.

- Sankaranarayanan, R., Cherney, M. M., Garen, C., Garen, G., Niu, C., Yuan, M. & James, M. N. G. (2010). The molecular structure of ornithine acetyltransferase from Mycobacterium tuberculosis bound to ornithine, a competitive inhibitor. J. Mol. Biol. 397, 979-990.
- Sanschagrin, P. C. & Mainz, D. T. (2006). Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49, 6177-6196.
- Schumacher, G., Sizmann, D., Haug, H., Buckel, P. & Bock, A. (1986). Penicillin acylase from E. coli: unique gene-protein relation. Nucleic Acids Res. 14, 5713-5727.
- Schwarz, M., Lund, E. G., Setchell, K. D., Kayden, H. J., Zerwekh, J. E., Björkhem, I., Herz, J. & Russell, D. W. (1996). Disruption of Cholesterol 7α-Hydroxylase Gene in Mice II. Bile acid deficiency is overcome by induction of Oxysterol 7α-Hydroxylase. Journal of Biological Chemistry 271, 18024-18031.
- Seemüller, E., Lupas, A. & Baumeister, W. (1996). Autocatalytic processing of the 20S proteasome. Nature 382, 468-470.
- Shankar, N., Baghdayan, A. S. & Gilmore, M. S. (2002). Modulation of virulence within a pathogenicity island in vancomycin-resistant Enterococcus faecalis. Nature 417, 746-750.
- Shapovalova, I., Alkema, W., Jamskova, O., de Vries, E., Guranda, D., Janssen, D. & Švedas, D. (2009). Mutation of Residue βF71 of Escherichia coli Penicillin Acylase Results in Enhanced Enantioselectivity and Improved Catalytic Properties. Acta naturae 1, 94.
- Shewale, J. & SivaRaman, H. (1989). Penicillin acylase: enzyme production and its application in the manufacture of 6-APA. Process Biochemistry 24, 146-154.
- Shewale, J. G., & Sudhakaran, V. K. (1997). Penicillin V acylase: its potential in the production of 6-aminopenicillanic acid. Enzyme and Microbial Technology, 20(6), 402-410.
- Shewale, J. G., Kumar, K. K. & Ambekar, G. R. (1987). Evaluation of determination of 6-aminopenicillanic acid by p-dimethyl aminobenzaldedyde. Biotechnology Techniques 1, 69-72.
- Sjövall, J., Griffiths, W. J., Setchell, K. D., Mano, N. & Goto, J. (2010). Analysis of bile acids. In Steroid Analysis, pp. 837-966: Springer.

- Smet, I. D., Hoorde, L. V., Saeyer, N. D., Woestyne, M. V. & Verstraete, W. (1994). In vitro study of bile salt hydrolase (BSH) activity of BSH isogenic Lactobacillus plantarum 80 strains and estimation of cholesterol lowering through enhanced BSH activity. Microbial Ecology in Health and Disease 7, 315-329.
- Smet, I., Hoorde, L., Woestyne, M., Christiaens, H. & Verstraete, W. (1995). Significance of bile salt hydrolytic activities of lactobacilli. Journal of Applied Bacteriology 79, 292-301.
- Smith, K., Zeng, X. & Lin, J. (2014). Discovery of bile salt hydrolase inhibitors using an efficient high-throughput screening system. PloS one 9, e85344.
- Sridevi, N., Srivastava, S., Khan, B. M. & Prabhune, A. A. (2009). Characterization of the smallest dimeric bile salt hydrolase from a thermophile Brevibacillus sp. Extremophiles 13, 363-370.
- Stellwag, E. & Hylemon, P. (1976). Purification and characterization of bile salt hydrolase from Bacteroides fragilis subsp. fragilis. Biochimica et Biophysica Acta (BBA)-Enzymology 452, 165-176.
- Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. Acta Crystallogr. Sect.D: Biol. Crystallogr. 60, 432-438.
- Stura, E. A. & Wilson, I. A. (1990). Analytical and production seeding techniques. Methods, 1, 38-49.
- Sue D, Boor KJ, Wiedmann M. σB-dependent expression patterns of compatible solute transporter genes opuCA and lmo1421 and the conjugated bile salt hydrolase gene bsh in Listeria monocytogenes. Microbiology 2003;149:3247-3256.
- Sumner, James B. "The isolation and crystallization of the enzyme urease preliminary paper." Journal of Biological Chemistry 69.2 (1926): 435-441.
- Suresh, C. G. Pundle, A., SivaRaman, H. & other authors (1999). Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. Nature Structural & Molecular Biology 6, 414-416.
- Suresh, C. G., Pundle, A. V., SivaRaman, H., Rao, K. N., Brannigan, J. A., McVey, C. E. & Dodson, G. G. (1999). Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. Nature Structural & Molecular Biology, 6(5), 414-416
- Svedas, V. K., Savchenko, M. V., Beltser, A. I. & Guranda, D. F. (1996). Enantioselective penicillin acylase-catalyzed reactions. Factors governing

substrate and stereospecificity of the enzyme. Annu. N.Y. Acad. Sci., 799, 659-669.

- Tanaka, H., Hashiba, H., Kok, J. & Mierau, I. (2000). Bile salt hydrolase of Bifidobacterium longum—biochemical and genetic characterization. Applied and environmental microbiology 66, 2502-2512.
- Tawfik, O. K. & Dan S. (2010). Enzyme promiscuity: a mechanistic and evolutionary perspective. Annual review of biochemistry, 79, 471-505.
- Thomas, C., Gioiello, A., Noriega, L. & other authors (2009). TGR5-mediated bile acid sensing controls glucose homeostasis. Cell metabolism 10, 167-177.
- Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, R. & Peltonen, L. (1996). Functional analyses of active site residues of human lysosomal aspartylglucosaminidase: implications for catalytic mechanism and autocatalytic activation. The EMBO journal 15, 2954.
- Tilg, H., Moschen, A. R. & Kaser, A. (2009). Gastroenterology, 136, 1476–1483.
- Torres-Bacete, J., Hormigo, D., Torres-Gúzman, R., Arroyo, M., Castillón, M. P., García, J. L., Acebal, C. & de la Mata, I. (2015). Overexpression of penicillin V acylase from streptomyces lavendulae and elucidation of its catalytic residues. Applied and environmental microbiology 81, 1225-1233.
- Torres G. R., De L. M., I., Torres, B. J., Arroyo, M., Castillón, M. P., & Acebal, C. (2002). Substrate specificity of penicillin acylase from Streptomyces lavendulae. Biochemical and biophysical research communications, 291(3), 593-597.
- Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022-1025.
- Valle, F., Balba, P., Merino, E. & Bollvar, F. (1991). The role of penicillin amidases in nature and in industry. Trends in biochemical sciences 16, 36-40.
- Vandamme, E. & Voets, P. (1974). Microbiol penicillin acylases. Adv Appl Microbiol 17, 311.
- Varshney, N. K., Ramasamy, S., Brannigan, J. A., Wilkinson, A. J. & Suresh, C.G. (2013). Cloning, overexpression, crystallization and preliminary X-ray crystallographic analysis of a slow-processing mutant of penicillin G acylase from Kluyvera citrophila. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 69, 925-929.

- Venneman, N. G. & van Erpecum, K. J. (2010). Pathogenesis of gallstones. Gastroenterology clinics of North America 39, 171-183.
- Verhaert, R. M. D., Riemens, A. M., vanderLaan, J. M., vanDuin, J. & Quax, W. J. (1997). Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from Alcaligenes faecalis. Appl. Environ. Biotechnol. 63, 3412-3418.
- Waldmann, H. (1988). The use of penicillin acylase for selective N-terminal deprotection in peptide synthesis. Tetrahedron letters 29, 1131-1134.
- Wang, X., Cahill, C. M., Piñeyro, M. A., Zhou, J., Doyle, M. E. & Egan, J. M. (1999). Glucagon-like peptide-1 regulates the beta cell transcription factor, PDX-1, in insulinoma cells. Endocrinology 140, 4904-4907.
- Wang, Y., L. G., Goode, J., P. J. C., Ouyang, K., Screaton, R. & Montminy, M. (2012). Inositol-1, 4, 5-trisphosphate receptor regulates hepatic gluconeogenesis in fasting and diabetes. Nature, 485(7396), 128-132.
- Wang, Z., Zeng, X., Mo, Y., Smith, K., Guo, Y. & Lin, J. (2012b). Identification and characterization of a bile salt hydrolase from Lactobacillus salivarius for development of novel alternatives to antibiotic growth promoters. Applied and environmental microbiology 78, 8795-8802.
- Waters, C. M. & Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol 21, 319-346.
- Weber, P. C. (1991). Physical principles of protein crystallization. Adv. Protein Chem. 41, 1-36.
- Wegener, H. C. (2003). Antibiotics in animal feed and their role in resistance development. Current opinion in microbiology 6, 439-445.
- Weiss, M. S. & Hilgenfeld, R. (1997). On the use of the merging R factor as a quality indicator for X-ray data. J. Appl. Crystallogr. 30, 203-205.
- Weiss, M. S. (2001). Global indicators of X-ray data quality. J. Appl. Crystallogr. 34, 130-135.
- Wen, Y., Feng, M., Yuan, Z. & Zhou, P. (2005). Expression and overproduction of recombinant penicillin G acylase from Kluyvera citrophila in Escherichia coli. Enzyme and microbial technology 37, 233-237.
- Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. J. Appl. Crystallogr. 43, 186-190.

- Xu, Q. A., Buckley, D., Guan, C. D. & Guo, H. C. (1999). Structural insights into the mechanism of intramolecular proteolysis. Cell, 98, 651-661.
- Zhang, W. Y., Wu, R. N., Sun, Z. H., Sun, T. S., Meng, H. & Zhang, H. P. (2009). Molecular cloning and characterization of bile salt hydrolase inLactobacillus casei Zhang. Annals of microbiology 59, 721-726.
- Zhou, J., Pineyro, M. A., Wang, X., Doyle, M. E. & Egan, J. M. (2002). Exendin-4 differentiation of a human pancreatic duct cell line into endocrine cells: Involvement of PDX-1 and HNF3β transcription factors. Journal of cellular physiology 192, 304-314.
- Zhu, D. Y., Zhu, Y. Q., Huang, R. H., Xiang, Y., Yang, N., Lu, H. X. & Wang, D. C. (2005). Crystal structure of the copper homeostasis protein (CutCm) from Shigella flexneri at 1.7 Å resolution: The first structure of a new sequence family of TIM barrels. PROTEINS: Structure, Function, and Bioinformatics, 58(3), 764-768.

List of Publications:

- **Deepak Chand,** Nishant Kumar Varshney, Sureshkumar Ramasamy, Priyabrata Panigrahi, James A Brannigan , Anthony J Wilkinson and C G Suresh (2015). Structure mediation in substrate binding and post-translational processing of penicillin acylases: Information from mutant structures of *Kluyvera citrophila* penicillin G acylase. Protein Science. 24(10):1660-70.

- **Deepak Chand,** Sureshkumar Ramasamy and C.G. Suresh (2015). A Highly Active Bile Salt Hydrolase from *Enterococcus faecalis* Shows Positive Cooperative Kinetics. Process Biochemistry, 51 (2)- 263–269

- **Deepak Chand**, Vellore Sunder Avinash, Yashpal Yadav, Archana Vishnu Pundle, Cheravakattu Gopalan Suresh and Sureshkumar Ramasamy. Molecular features of bile salt hydrolases and relevance in human health (Communicated)

- **Deepak Chand,** Priyabrata Panigrahi, Sureshkumar Ramasamy and C. G. Suresh. Subtle structural variations impart unique characteristics to Bile Salt Hydrolase from *Enterococcus faecalis*. (under preparation)

- Priyabrata Panigrahi, **Deepak Chand,** Ruchira Mukherji , Sureshkumar Ramasamy , C. G. Suresh (2015) Sequence and structure-based comparative analysis to assess, identify and improve the thermostability of penicillin G acylases J Ind Microbiol Biotechnol. Nov; 42(11):1493-506

- Avinash, V. S., Panigrahi, P., **Chand, D**., Pundle, A., Suresh, C. G., & Ramasamy, S. (2016). Structural analysis of a penicillin V acylase from Pectobacterium atrosepticum confirms the importance of two Trp residues for activity and specificity. *Journal of structural biology*, *193*(2), 85-94.