

**Biochemical and structural analysis of Bile Salt
Hydrolases from gut microbiome and their *In vivo*
efficacy studies for biotherapeutic applications**

**Thesis Submitted to
AcSIR**

**For The Degree of
DOCTOR OF PHILOSOPHY**

**In
Biological Sciences**



**By
Yashpal Yadav
Registration Number: 10BB13A26032**

Under the guidance of

**Guide
Dr. Archana V. Pundle**

**Co-Guide
Dr. Sureshkumar Ramasamy**

**Biochemical Sciences Division
CSIR-National Chemical Laboratory
Pune-411008, India**

SEPTEMBER 2018

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 **Biochemical and structural analysis of Bile Salt Hydrolases from gut microbiome and their In vivo efficacy studies for biotherapeutic applications** submitted by Mr. **Yashpal Yadav** to Academy of Scientific and Innovative Research (AcSIR) in fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy**, embodies original research work carried out under my supervision. I further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table etc., used in the thesis from other sources, have been duly cited and acknowledged.

Dr. Archana V. Pundle
(Research Supervisor)

Yashpal Yadav
(Research Student)

Dr. Sureshkumar Ramasamy
(Research Co-Supervisor)

Biochemical Sciences Division
CSIR-National Chemical Laboratory
Pune-411008

Place:

Date:

Communication Channels

NCL Level DID : 2590
NCL Board No. : +91-20-25902000
EPABX : +91-20-25893300
: +91-20-25893400



FAX

Director's Office : +91-20-25902601
COA's Office : +91-20-25902660
SPO's Office : +91-20-25902664

WEBSITE

www.ncl-india.org

DECLARATION BY THE CANDIDATE

I hereby declare that the thesis entitled "**Biochemical and structural analysis of Bile Salt Hydrolases from gut microbiome and their *In vivo* efficacy studies for biotherapeutic applications**", submitted for the Degree of **Doctor of Philosophy in Biological Sciences** to Academy of Scientific and 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. Archana V. Pundle (research guide)** and **Dr. Sureshkumar Ramasamy (co-guide)**. The work is original and has not been submitted in part or full by me for any other degree or diploma to any other University. I further declared that the material obtained from other sources has been duly acknowledged in the thesis.



Yashpal Yadav

*Dedicated to my beloved family
and my teacher CKR....*

Acknowledgement

I would like to express my sincere gratitude towards my research guide Dr. Archana V. Pundle for her constant support and guidance during the present work. I would also like to thank my co-guide Dr. Sureshkumar Ramasamy for his motivation, valuable suggestions, constructive criticism and unconditional support throughout the course of this work. I take this opportunity to thank him for providing me the immense freedom to pursue my naïve ideas. His words of appreciation shall always be cherished! I am extremely thankful to Dr. C G Suresh for introducing me to Ntn hydrolases and the concern he showed regarding the progress of my work.

I would like to thank my research advisory committee: Dr. Dhanasekaran Shanmugam, Dr. Mahesh Dharne and Dr. H V Adikane, as well, for their insightful comments and encouragement.

My sincerest gratitude goes to Dr. Deepak Chand for his genuine interest, concern and constant inputs in regards to my work. I shall always cherish our association throughout the course of this work. In regards to the computational work, I enjoyed a lot working with Dr. Mrityunjay Tiwari. His critical comments and suggestions improved the quality of my work. I would also like to appreciate the help extended by Amit Kumawat and Vijay Rajput for the simulation studies performed in my thesis.

I take this opportunity to thank Dr. Kiran Kulkarni for allowing me to use in house structural biology facility. I am also thankful to Dr. Saikrishnan and Dr. Gayatri from IISER, Pune for allowing me to use the X-ray facility. I am also thankful to Dr. Radha Chauhan and Swasthik, Praveen and Ashiwini from NCCS for allowing me to use Mosquito facility and their help. I am grateful to Dr. Ravindra Makde, Dr. Ashwani Kumar and Dr. Biplab Ghosh from PXBL21 beamline, Indus-II, RRCAT, DAE, Indore, for their kind help & support in the crystal data collection.

I am also thankful to Dr. Ajit from CAMS, Venture Center, CSIR-NCL, who allowed me to use the Mass spectrometry facility and his valuable suggestions for data analysis.

I am thankful to Dr. Mahesh Kulkarni for allowing me to use the LC-MS facility and his students for helping me in sample preparation and data analysis. I am especially grateful to Shakuntala, for helping me with basic introductions to the nitty-gritties of MS, providing me insights related to data analysis and for healthy discussions regarding the use of mass spectrometry in structural biology.

It would be unfair not to mention of the immense help provided by Dr. Subhaschandraboze Chinnathambi and his group, particularly, Shweta and Nalini, during trying times and it is genuinely appreciated.

I am thankful to Dr. Sachin Agawane for his help and guidance in animal studies.

I am grateful to my past lab members Dr. Nishant, Dr. Manas, Dr. Priyabrata, Dr. Tulika, Dr. Ranu, Dr. Deepak, Dr. Ruby, Tejashri and present lab members, tagged as Dude's lab, Shiva, Manu, Ameya, Deepanjan, Vijay and Debjyoti for providing me an awesome work environment in the lab.

I am thankful to members of KK lab members Anand, Sneha, Ashwini, Zenia and Debopriya for being my media kitchen in SOS and fruitful discussion related to my work.

I shall cherish the countless moment I spent with my friends Manoj, Ravi, Deepak, Yuvraj, Shakuntala, Shahebaaz, Arun, Pinka, Brij, Dilip, Ulhaas, Parul, Preeti, Shamsaad, Priyanka, Amit, Anand, Sneha, Bhagyashri, Deepanjan, Vijay. Many thanks to Deepanjan who keeps capturing our very own "charminar" in IISER campus with surprising efficiency, so that we (Deepak Chand, Shakuntala, Ejaj, Deepanjan and I) can enjoy our tea in a peaceful night atmosphere. Thanks are due to my friends Priya, Tanmay, Pratima, Sonam, Raksha, Vishal, Pratap, Quazi Tousif, Akik, Utsav, Vishal bhujja, Nitya, Masuma and Sangita as well, who lent their support at various times during my Ph.D. tenure.

A special mention must nevertheless go to Shakuntala, for being my friend and confidant during my entire Ph.D. tenure and for constantly providing support, motivation and inspiration to further my work, in regards to Ph.D. and more.

My sincere thanks to Mrs. Chandrani Kundu Rahman for her motivation and help during my tough times.

I am also thankful to all divisional members and my friends for their support. I am also thankful to other staff of Biochemical Sciences Division, for their help. I am also grateful to the SAC and AcSIR office staff of CSIR-NCL especially Komal and Vaishali for the help they provided me time to time.

I gratefully acknowledge UGC, New Delhi India for Fellowship. I also thank Director, CSIR-NCL and Chair, Biochemical Sciences Division for their support and providing wonderful infrastructure.

Last but not the least, I would like to thank my family for supporting me unconditionally throughout my Ph.D.

Yashpal Yadav

Table of contents

Sr. No.	Title	Page No.
	List of tables	i
	List of figures	ii
	Abbreviations	vi
	Abstract	vii
Chapter I	Introduction	
1.1	Types and modification of Bile salts	2
1.2	Synthesis of bile acids	4
1.2.1	Classical pathway	4
1.2.2	Alternate pathway	6
1.3	Functions of Bile Acids	6
1.3.1	Digestion and related functions	6
1.3.2	Bile acid in FXR/TGR5 (GPBAR) signaling	8
1.3.3	Bile salt in regulation of cholesterol	9
1.3.3.1	Feedback regulation	9
1.3.3.2	Role of BSH in regulation of bile salts	9
1.4	Distribution of bile salt hydrolases in Gut microbiome	12
1.5	Hypothesis and objectives of present investigation	12
1.6	References	13
Chapter II	Dissection of Catalytic Site in Crucial Gut Microbiome Enzyme: Bile Salt Hydrolase	
2.1	Introduction	23
2.2	Materials	27
2.3	Methods	28
2.3.1	Heterologous expression and purification of <i>Ef</i> BSH and mutant proteins	28
2.3.2	Site directed mutagenesis	29
2.3.3	Measurement of Enzyme Activity and Kinetic parameter of <i>Ef</i> BSH mutants	29
2.3.4	Molecular Dynamics simulation	30
2.3.5	Crystallization, diffraction and data collection	30
2.3.6	X-ray diffraction and cryo-protection	31
2.3.7	Data collection and processing	31
2.3.8	Matthews coefficient	32
2.3.9	Structure determination and refinement	32
2.3.10	Structure validation	33
2.3.11	Computational Details	36
2.4	Results	38
2.4.1	Protein purification	38
2.4.2	Enzyme assay and steady state kinetics	40
2.4.3	Crystallization, data collection and refinement	42
2.4.4	Structure analysis	45

2.4.5	Molecular Dynamics simulation	56
2.4.6	Density Function Theory (DFT) calculations	58
2.5	Discussion	60
2.6	Summary and conclusions	64
2.7	References	65
Chapter III	Oligomeric Attributes of Cholyglycine Hydrolase (CGH) Family in Enzymatic Activity and Stability	
3.1	Introduction	72
3.2	Materials and Methods	74
3.2.1	Gene synthesis	74
3.2.2	Cloning, Heterologous expression and purification	79
3.2.3	Site directed mutagenesis	80
3.2.4	Homology modeling	81
3.3	Results	81
3.3.1	Gene synthesis and cloning	81
3.3.2	Protein purification	84
3.3.3	Kinetic characterization	92
3.3.4	Structural comparison of CGH family of proteins	94
3.3.5	Density Function Theory (DFT)	105
3.4	Discussion	106
3.5	Summary and conclusions	109
3.6	References	110
Chapter IV	Efficacy studies of encapsulated BSH enzyme on serum cholesterol level	
4.1	Introduction	114
4.1.1	Need for Enzyme formulation	115
4.1.2	Studies carried out-	115
4.2	Materials and methods	115
4.2.1	Enzyme preparation	115
4.2.2	Enzyme encapsulation	115
4.2.3	<i>In-vivo</i> study of Enzyme formulation	116
4.3	Results	118
4.3.1	Effect of cryo-protectant on lyophilized <i>Ef</i> BSH enzyme	118
4.3.2	Protein loading in liposomes	119
4.3.3	Hypocholesterolemic study of <i>Ef</i> BSH in animal model	122
4.4	Discussion	124
4.5	Conclusions	125
4.6	References	126
Chapter V	Summary and conclusions	131
	List of publications	133

List of tables

Table No.	Title	Page no.
2.1	List of primers for generating site directed mutagenesis in <i>Ef</i> BSH wild type	29
2.2	Protein component used for gel filtration standards	39
2.3	Percent residual activity profile of <i>Ef</i> BSH mutants as compared to wild type	41
2.4	X-ray diffraction and Data collection statistics for <i>Ef</i> BSH mutant structures viz. R207A and E269A	44
2.5	List of corresponding amino acids of Tyr20, Glu21, Glu269 and Arg207 of <i>Ef</i> BSH (PDB Id-4wl3) in <i>Ls</i> BSH, <i>Cp</i> BSH and <i>B</i> BSH	50
2.6	The influence of the non-cystine residues of the active site on the S-H bond cleavage in the pN for the geometries obtained after the constrained optimization at the M06-2X(water)/6-31G** level of theory.	59
2.7	Hydrogen bond distance (in Å) between Arg207 and Glu21 of WT <i>Ef</i> BSH and E269A mutant.	60
2.8	The interaction energy between acetate moiety of Glu21 and side chain of Tyr20 in E269A and R207A	60
3.1	Overlapping primers of <i>D</i> BSH	75
3.2	Overlapping primers of <i>Ms</i> BSH	76
3.3	PCR mixture for first PCR	78
3.4	List of primers used for site directed mutagenesis in <i>D</i> BSH	80
3.5	Retention volume and oligomer profile of wt <i>D</i> BSH and mutants using Sephacryl S200 column.	91
3.6	Energy calculation between tetramer interface residues using DFT	106
4.1	Composition of high fat diet for introducing hypercholesterolemia in rat experimental model	117
4.2	Enzyme activity comparison of cryoprotectant used during lyophilization studies of <i>Ef</i> BSH stability. Percent residual activity is shown in parenthesis.	119

List of figures

Figure No.	Title	Page no.
1.1	Anatomy of bile secretion system in human body	1
1.2	Structure of bile acids	3
1.3	Modifications of bile acids	4
1.4	Bile acid synthesis pathways	5
1.5	Detergent action of bile	7
1.6	Cascade of TGR5 signalling	8
1.7	Schematic representation of enterohepatic circulation and feedback control of interdependence of bile acids and cholesterol	10
1.8	Structural study of BSH from gut microbiome	11
2.1	Review of structural insight	24
2.2	Consurf image showing conservation of residues in core β -sheets and assembly loop	26
2.3	Dimeric form of wt <i>Ef</i> BSH revealing the Arg207 of chainB protrudes above the active site of chainA	26
2.4	Sequence alignment of Bile salt Hydrolases	27
2.5	Ramachandran map of <i>Ef</i> BSH E269A mutant generated using SFcheck (CCP4i)	34
2.6	Ramachandran map of <i>Ef</i> BSH R207A mutant generated using SFcheck (CCP4i)	35
2.7	The optimized geometry of chain A of a) wild type and b) E269A mutant type under constrained conditions at the M06-2X(water)/6-31G** level of theory	37
2.8	the optimized geometry of Glu21 and Tyr20 of a) E269A and b) R207A mutant type under constrained conditions at the M06-2X(water)/6-31G** level of theory	38
2.9	Purification of <i>Ef</i> BSH wild type and mutants	39
2.10	Size Exclusion Chromatography (SEC) of <i>Ef</i> BSH wild type and mutant and compared with protein standards.	40
2.11	Biochemical activity of <i>Ef</i> BSH wild type and mutants against glycol- and tauro- conjugated bile acids.	41
2.12	Difference in enzyme kinetic behavior of <i>Ef</i> BSH E269A and	42

	R207A mutant enzymes	
2.13	Diffraction protein crystal mounted under UV and visible light	43
2.14	X-ray diffraction image of <i>Ef</i> BSH E269A and R207A mutant	43
2.15	Analysis of Hydrogen bonding pattern in active site residues of four different chains of wild type <i>Ef</i> BSH	46
2.16	Dynamics of Active site	47
2.17	Analysis of loop1	49
2.18	Analysis of corresponding residues of Glu21, Glu269 and Arg207 of <i>Ef</i> BSH in other reported BSH structures	50
2.19	Structural superposition of all four chains of <i>Ef</i> BSH wild type and its mutants	51
2.20	Structural superposition of all four of <i>Ef</i> BSH R207A and <i>Ef</i> BSH E269A	51
2.21	Residue wise RMSD analysis of <i>Ef</i> BSH wild type	52
2.22	Residue wise RMSD analysis of <i>Ef</i> BSH E269A mutant	53
2.23	Residue wise RMSD analysis of <i>Ef</i> BSH R207A mutant	54
2.24	2Fo-Fc map of Glu21 and Tyr20 residues of all four chains from E269A and R207A mutant developed at 1σ	55
2.25	Docking of GCA in four chains of E269A mutant shown with respect to loop1 to loop4.	55
2.26	Comparison of electrostatic potential surface view of active site of <i>Ef</i> BSH wild type, R207A and E269A mutants	56
2.27	RMSD plot of MD simulation of <i>Ef</i> BSH wild type performed for 100 ns	57
2.28	Radius of gyration during MD simulation of <i>Ef</i> BSH showing no change in the quaternary arrangement and packing of molecule	58
2.29	Proposed mechanism of allostery in <i>Ef</i> BSH	65
3.1	Steps involved in optimized primer designing using DNAWorks server	74
3.2	Construct of <i>Df</i> BSH-pET22b(+)	79
3.3	Sequence alignment of <i>Ef</i> BSH and <i>Df</i> BSH	83
3.4	Gene synthesis and cloning	84
3.5	Comparison of expression profile of <i>Df</i> BSH at 37 °C and 16 °C	86

3.6	Thin Layer Chromatography (TLC) profile of <i>Dl</i> BSH(truncated) performed with whole cell	86
3.7	SDS-PAGE gel profile after Ni-NTA purification. <i>Dl</i> BSH fraction is present in unbound and wash, while absent in elution fraction	87
3.8	Sequencing result of cloned <i>Dl</i> BSH compared with the wild type <i>Dl</i> BSH	88
3.9	Ion exchange chromatography profile of truncated <i>Dl</i> BSH using anion exchanger Q-sepharose column	88
3.10	Purification of dimeric <i>Dl</i> BSH using SEC650 (Biorad) column	89
3.11	Qualitative estimation of the activity profile from purified dimeric form of truncated <i>Dl</i> BSH was performed using of TLC assay	89
3.12	Ni-NTA profile of <i>Dl</i> BSH and <i>Ms</i> BSH	90
3.13	Size exclusion chromatography profile of <i>Ms</i> BSH	90
3.14	Overlay analysis of Size exclusion chromatography profile of WT <i>Dl</i> BSH and its mutant	91
3.15	Biochemical characterization of <i>Dl</i> BSH	92
3.16	Kinetic characterization of <i>Dl</i> BSH	93
3.17	Preferential substrate activity profile with conjugated bile acids	94
3.18	Cartoon representation of tetramer form of <i>Ef</i> BSH (PDB ID: 4WL3) highlighting the tetramer interface	95
3.19	Tetramer interface of <i>Cp</i> BSH (Left; PDB ID-2RLC) and <i>Bi</i> BSH (Right; PDB ID-2HEZ)	95
3.20	Tetramer interface of <i>Dl</i> BSH (model) highlighting the presence of neutral amino acid i.e. Gln176 in interface minicluster	96
3.21	Superposed cartoon representation of <i>Ef</i> BSH and <i>Dl</i> BSH	97
3.22	Tetramer interface residue mapping in dimer (<i>Ls</i> BSH) and tetramer (<i>Ef</i> BSH) Bile Salt Hydrolase	97
3.23	Superposed cartoon representation of <i>Ef</i> BSH and <i>Ls</i> BSH showing Tyr205 and Asp206, respectively in stick form.	98
3.24	Interface analysis of <i>Ef</i> BSH (PDB ID – 4WL3)	99
3.25	Interface analysis of <i>Pa</i> PVA (PDB ID – 4WL2)	101

3.26	Cartoon representation of interface analysis in <i>Pa</i> PVA (4WL2)	102
3.27	Interface comparison of Tetramer and Dimer structure in gram negatives	103
3.28	Superposed cartoon representation of Gram positives BSH and PVA	104
3.29	Superposed cartoon representation of Gram negative PVA and BSH	104
3.30	Miniclusters formed by tetramer interface residue	105
4.1	Preparation of enzyme formulation for animal studies	116
4.2	Experimental design of animal model to study hypocholesterolemic effect	118
4.3	Qualitative analysis of Protein (<i>Ef</i> BSH) loading in liposomes by SDS-PAGE	120
4.4	Efficiency of <i>Ef</i> BSH protein loading in liposomes prepared at two different concentrations i.e. 1% and 5% by protein estimation in pellet and supernatant using Bradford reagent	121
4.5	Comparison of biochemical activity profile of <i>Ef</i> BSH using ninhydrin assay upon protein loading in liposome at two different concentration viz. 1% and 5% SL (SL-Sophorolipid)	121
4.6	Serum cholesterol profile of Group III rats fed on protein formulations	122
4.7	Secondary effects profile upon treatment with protein formulation in Group3 rats	123
4.8	Serum cholesterol profile of Group1, Group2 and Group3 rats fed on <i>Ef</i> BSH protein formulations after 42 days.	123

Abbreviations

aa	Amino acid
AHL	Acyl homoserine lactone
AU	Absorbance unit
<i>B</i> BSH	<i>Bifidobacterium longum</i> bile salt hydrolase
bp	base pair
BSH	Bile salt hydrolase
<i>Bsp</i> PVA	<i>Bacillus sphaericus</i> penicillin V acylase
<i>Bt</i> BSH	<i>Bacteroides thetaiotaomicron</i> VPI bile salt hydrolase
CCP4	Collaborative computational project no. 4
CGH	Cholyglycine hydrolase
CPB	Citrate phosphate buffer
<i>Cp</i> BSH	<i>Clostridium perfringens</i> bile salt hydrolase
DFT	Density function theory
DTT	Dithiothreitol
<i>D</i> BSH	<i>Dorea longicatena</i> bile salt hydrolase
<i>Ef</i> BSH	<i>Enterococcus faecalis</i> bile salt hydrolase
GCA	Glychocholic acid
GCDCA	Glychochenodeoxycholic acid
GDCA	Glychodeoxycholic acid
h	Hill coefficient
IPTG	Isopropyl β -D-1-thiogalactopyranoside
<i>Ls</i> BSH	<i>Lactobacillus salivarius</i> bile salt hydrolase
ns	Nanoseconds
Ntn	N-terminal nucleophile
OD	Optical density
<i>Pa</i> PVA	<i>Pectobacterium atrosepticum</i>
PCR	Polymerase chain reaction
PDB	Protein data bank
PVA	Penicillin V acylase
RMSD	Root mean square deviation
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
TCA	Taurocholic acid
TCDCa	Taurochenodeoxycholic acid
TDCA	Taurodeoxycholic acid

ABSTRACT

Abstract

Chapter 1: Introduction

Out of two catabolic pathways of cholesterol *i.e.* bile acids and steroids, conversion of cholesterol to bile acids is major route for cholesterol reduction. The bile acid pool regulates its own synthesis involving FXR signaling pathway. Any disturbances in the bile acid pool activate CYP7A1 in liver which starts conversion of cholesterol to bile acid. Moreover, Bile salt Hydrolase (BSH) in the human gut deconjugates the bile salts and reduces its affinity for enterohepatic circulation. Here, we reviewed the regulation of bile acids and involvement of gut microbiome in signaling as well as bile acid homeostasis. Role of bile salt hydrolase, present in gut microbiome, in removal of bile acid through feces and its role in cholesterol reduction are discussed.

Chapter 2: Dissection of Catalytic Site in Crucial Gut Microbiome Enzyme: Bile Salt Hydrolase

Bile Salt Hydrolase (BSH) is a crucial enzyme in bile acid metabolism in GI tract and present in many gut microbiomes. A gut inhabitant *Enterococcus faecalis* BSH (*Ef*BSH) displays substantially high activity compared to other reported BSHs. We have explored the mechanistic study of the highly active *Ef*BSH by looking into the dynamic hydrogen bonding interactions at substrate binding pocket. Apart from core substrate binding amino acids residues, R207 and E269 were have also been identified to be crucial in substrate binding and catalytic activity as the mutant R207A and E269A showed substantial modulation in biochemical activity and affinity towards the different bile acids. The mutant E269A showed a significant loss of in BSH activity and with increased affinity towards GCA. Whereas, R207 was found to be involved in allostery, which displayed the MM-Kinetics compared to native *Ef*BSH. Further, the structural investigations of R207A and E269A mutant revealed the molecular basis for activity, specificity and allostery in the *Ef*BSH. A comparison of electrostatic properties of the active sites of R207A suggests no major change in the in electrostatic potential surface of active site. However, E269A mutant showed more electropositive nature of active site as compared to the wild type. A comparative electrostatic force analysis of the active site of the E269A mutant and the wild type revealed that the long-range electrostatic force experienced by the S-H bond due the other amino acid residues at the active site of the two types are marginal. Thereby, provides a conformational basis of reduced enzymatic activity in the E269A mutant.

Chapter 3: Oligomeric Attributes of Cholylglycine Hydrolase (CGH) Family in Biochemical Activity and Stability

In this chapter, we have considered the oligomeric attribute of Bile Salt Hydrolase (BSH) in Cholylglycine Hydrolase (CGH) family (BSH, a member of Ntn hydrolase superfamily). N-terminal nucleophile (Ntn) hydrolase superfamily is one of the most diverse superfamily of hydrolytic enzymes. This superfamily includes many families of enzymes having physiological, clinical and pharmaceutical importance which harbours different homo and hetero oligomeric quaternary state. The cholylglycine hydrolase (CGH) family has been reported to consist of homo-oligomeric BSH and PVA.

Here we identified the crucial residues which is imperative to the oligomerization of CGH family of proteins and has implications on screening the highly active BSH/PVA using primary structure of protein from genomic databases and also predicts the oligomeric state of the protein. We propose, in general, there must be two fundamentally different pathway of oligomerization depending upon the intrinsic symmetry of the molecule. In cyclic symmetry, oligomeric proteins exist in two different forms i.e. active (monomer) or inactive (oligomer) or *vice-versa*. While in proteins having dihedral symmetry, can exist in multiple oligomeric state which perhaps, gives plasticity to the enzyme evolution and ability to metabolize or overcome the toxic effect of metabolites which leads to survival of the microorganisms.

Chapter 4: Efficacy studies of encapsulated BSH enzyme on serum cholesterol level

Here, we optimized the stability of *Ef*BSH enzyme in presence of cryoprotectant. Further, the *Ef*BSH was encapsulated in sophorolipid based liposome to develop an enzyme-based formulation. Animal studies were carried out to ascertain the hypocholesterolemic effect of enzyme based therapeutic formulation. We found that the level of cholesterol in *Ef*BSH enzyme-based formulation showed significant decrease in HDL and LDL cholesterol without influencing the other parameters such as glucose and creatine level, which was found to be normal throughout the study. The efficacy of enzyme based biotherapeutics was compared with statins, which is a widely used class of molecule which specifically inhibits HMG co-A reductase inhibitor in mevalonate pathway. Moreover, various blood parameters studied during the experimental phase showed less adverse effect of protein formulation as compared to statins. Further, cholesterol dependent pathways are not affected as we are targeting the by-product of cholesterol i.e. bile acids in gut.

Chapter 5: Summary and conclusions

The salient finding of the present work and conclusions drawn from the data are discussed in this chapter.

CHAPTER I

INTRODUCTION

1. Introduction

Bile is a greenish yellow fluid produced by hepatocytes which helps in emulsification of dietary lipids. It consists of 95 % water and endogenous component such as bile salt, cholesterol, amino acids, steroids etc. Bile is concentrated and stored in the gall bladder and part of it is secreted in the duodenum. Bile helps in excretion of major toxic lipophilic compounds which are not filtered by kidney. Bile salts in bile help in emulsification of dietary lipids. Bile salt exhibits detergent properties such as micelle formation which facilitate the lipid absorption and also reduces the detergent effect of bile salt in biliary epithelium. Bile also serves as a major route for the excretion of cholesterol [1,2].

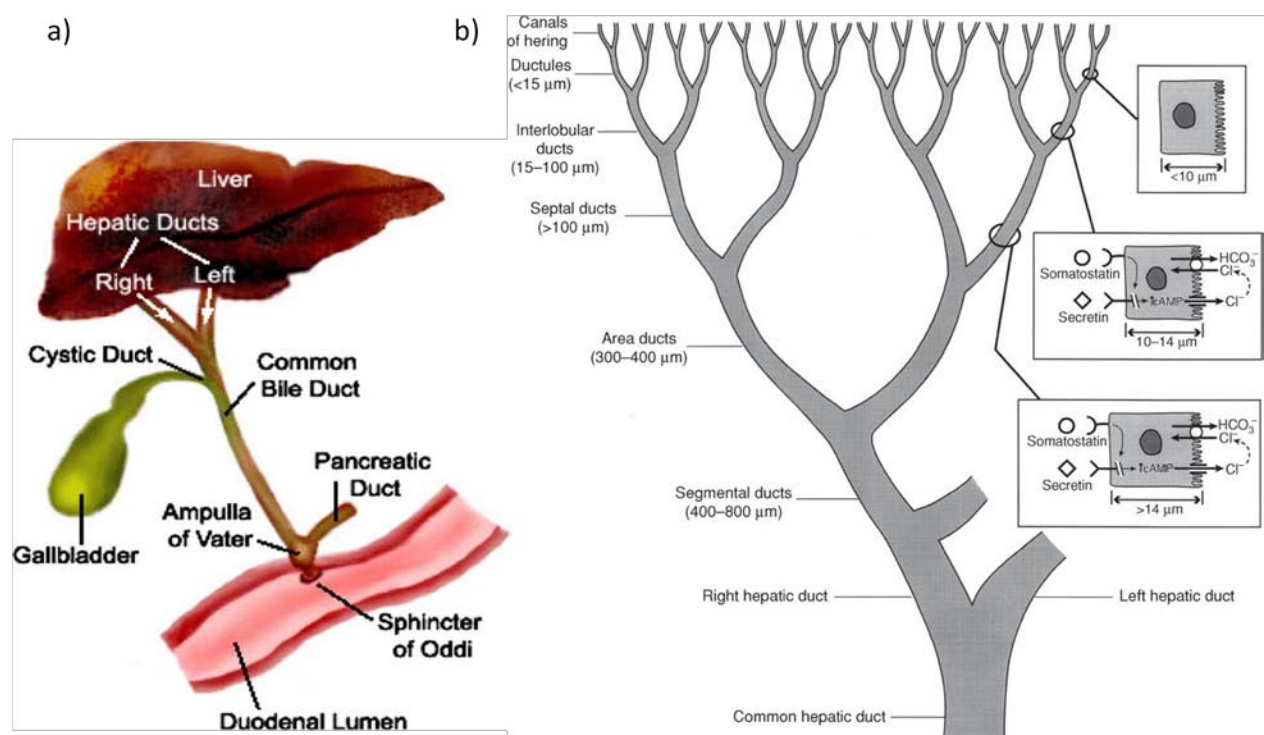


Figure 1.1: Anatomy of bile secretion system in human body: panel a) overview of biliary system [3] panel b) biliary tree showing heterogeneity in bile duct epithelial cells [4]

Rudolf schoenheimers in his bottle experiment reported the effect of diet on synthesis and destruction of cholesterol. Also, incorporation of Glycocholic acid (GCA) in the diet induces the synthesis of cholesterol [5]. This study was later confirmed in 1953 by using radioactive C^{14} -labeled acetate incorporation in cholesterol [6]. This feedback control was due to the presence of HMG co-A reductase, a rate limiting enzyme for cholesterol biosynthesis [7](Bucher and Lynen, 1959). Goldstein and Brown found that the activity of HMG co-A reductase is 10 fold higher in familial hypercholesterolemia (FH) patients and its activity is

controlled by lipoproteins, specifically low density lipoproteins [8,9] which eventually led to the discovery of low density lipoprotein receptor (LDL receptor) [10].

Here we reviewed the interplay of cholesterol and bile acids at cellular as well as tissue level and maladies associated with the mutation in associated genes involved in bile acid regulation.

1.1 Types and modification of Bile salts

Bile acids are categorized into primary bile acid and secondary bile acids. Primary bile acids are synthesized entirely in liver while secondary bile acids are generated upon bacterial modification in gut. Cholic acid (CA) and Chenodeoxycholic acid (CDCA) are the primary bile acids and rest others bile acids are derived from the CA and CDCA (**Fig.2**). Liver is the primary source for production of bile acid. Bile acids are conjugated with Glycine and taurine in the liver catalyzed by terminal enzyme, BAAT in bile acid synthesis pathway. The conjugation of these moieties to the bile acid enhances the hydrophilic property of the bile acid. Conjugation of cholic acid yields glycocholic acid and taurocholic acid while conjugation of chenodeoxycholic acid yields glycochenodeoxycholic acid and taurochenodeoxycholic acid. Bacterial modification of cholic acid and chenodeoxycholic acid leads to synthesis of deoxycholic acid and lithocholic acid respectively [1]. Conjugated, and to lesser extent unconjugated bile acids are reabsorbed from the intestine and modified by liver enzymes to produce tertiary bile acids [11]. This process is known as enterohepatic recirculation (**Fig.3**). SLC family of proteins play important role in active transportation of conjugated bile acid. SLC10A2 codes for apical sodium bile acid transporter (ASBT) also known as IBAT [12], which efficiently reabsorb >95 % of conjugated bile acids [13]. Since inhibition of ASBT transporters reduces the plasma cholesterol level, it was proposed to be a drug target for treatment of hypercholesterolemia [14,15]. Structural study from bacterial homologs showed the involvement of two domains which alternates the accessibility to inner and outer side of membrane [16,17]. In liver active transport of bile acids takes place through sodium transporter cotransporting polypeptide (NTCP) [18]. Bile salts also diffuses passively through upper small intestine and large intestine.

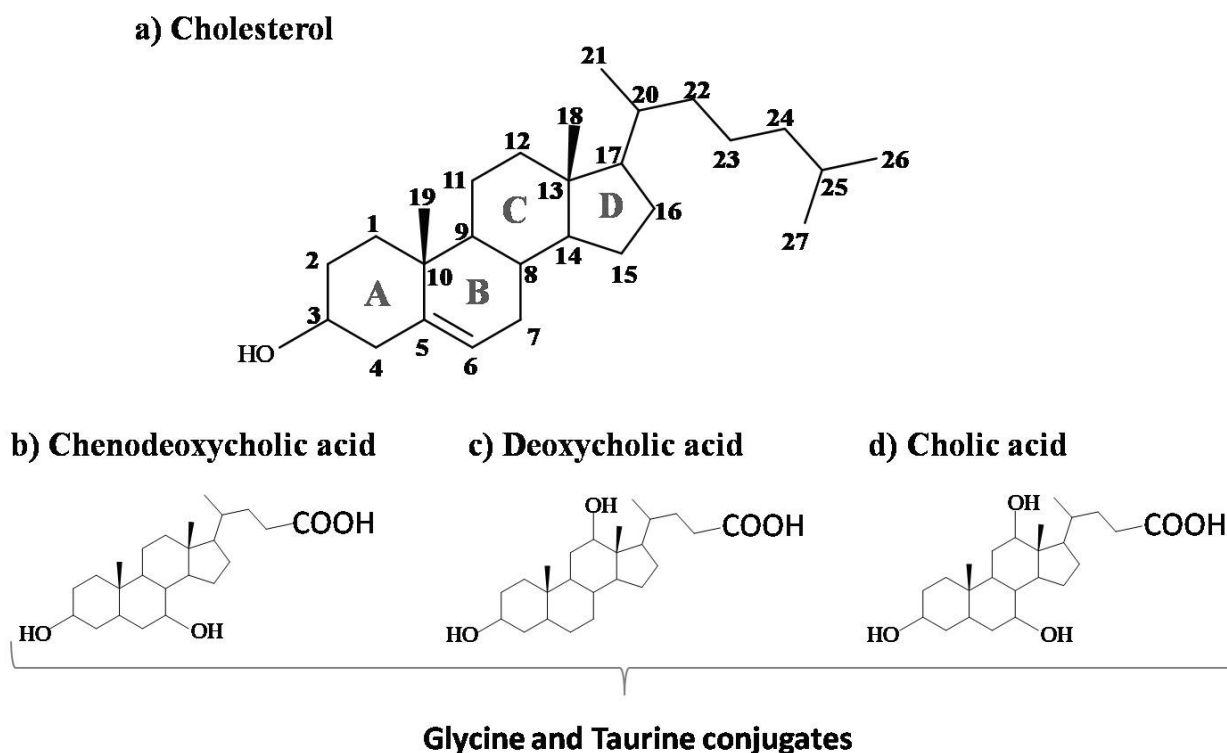


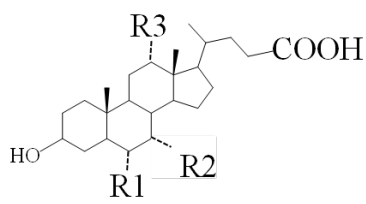
Figure 1.2: Structure of bile acids: panel **a)** describes the naming and numbering in ring system of cholesterol. **b)** & **d)** primary bile acids synthesized from cholesterol **c)** DCA, secondary bile acid synthesized from cholic acid upon bacterial modification in gut.

The effect of hydroxylation of bile acid makes huge difference in their solubility [19,20] (**Fig. 3a**). This hydroxylation occurs at different position of bile acid during synthesis which makes different bile acid pool (**Fig. 3b**) [21]. Ursocholic acid (UCA) and Cholic acid (CA) are more soluble as compared to other bile acids owing to the hydroxylation at two different sites (**Fig. 3c**). Although UCA and CA are more soluble, the extent of solubility of UCA is significantly higher owing to the axial conformation of hydroxyl group rather than planar as in CA [19,20].

a)

Bile acids	R1	R2	R3
CA	H	OH	OH
CDCA	H	OH	H
DCA	H	H	OH
LCA	H	H	H
MCA	OH	OH	H
UDCA	H	OH	H
UCA	H	OH	OH

b)



c)

Bile acids	Solubility
UCA	1670 μM
CA	273 μM
CDCA	28 μM
DCA	28 μM
LCA	0.05 μM

Hydrophilicity

Figure 1.3: Modifications of bile acids panel a) & b) shows the different modification of bile acid due to position of hydroxyl group in cholesterol moiety [21] Panel c) shows the effect of hydroxyl group on solubility of different bile acids [19,20].

1.2 Synthesis of bile acids

Synthesis of bile acid starts from liver which require 17 enzymes to complete the process. Most of these enzymes are synthesized in liver and belongs to cytochrome P450 superfamily which are mixed function oxidases. The CYP7A1 gene encodes first and the rate limiting enzyme, Cholesterol 7 α -hydroxylase for bile acid synthesis from cholesterol in liver. Bile acid synthesis takes place through two different route viz. classical pathway and alternate pathway. Classical pathways runs entirely in liver while alternate pathway works in other tissues.

1.2.1 Classical pathway

Both the pathways involves four different steps Initiation, sterol ring modification, side chain shortening and conjugation steps for bile acid synthesis independent of classical and alternate pathways. In classical pathway, cholesterol 7 α hydroxylase encoded by CYP7A1 add hydroxyl group at 7 α position in cholesterol whilst alternate pathway starts with side chain shortening which involve CYP27A1.

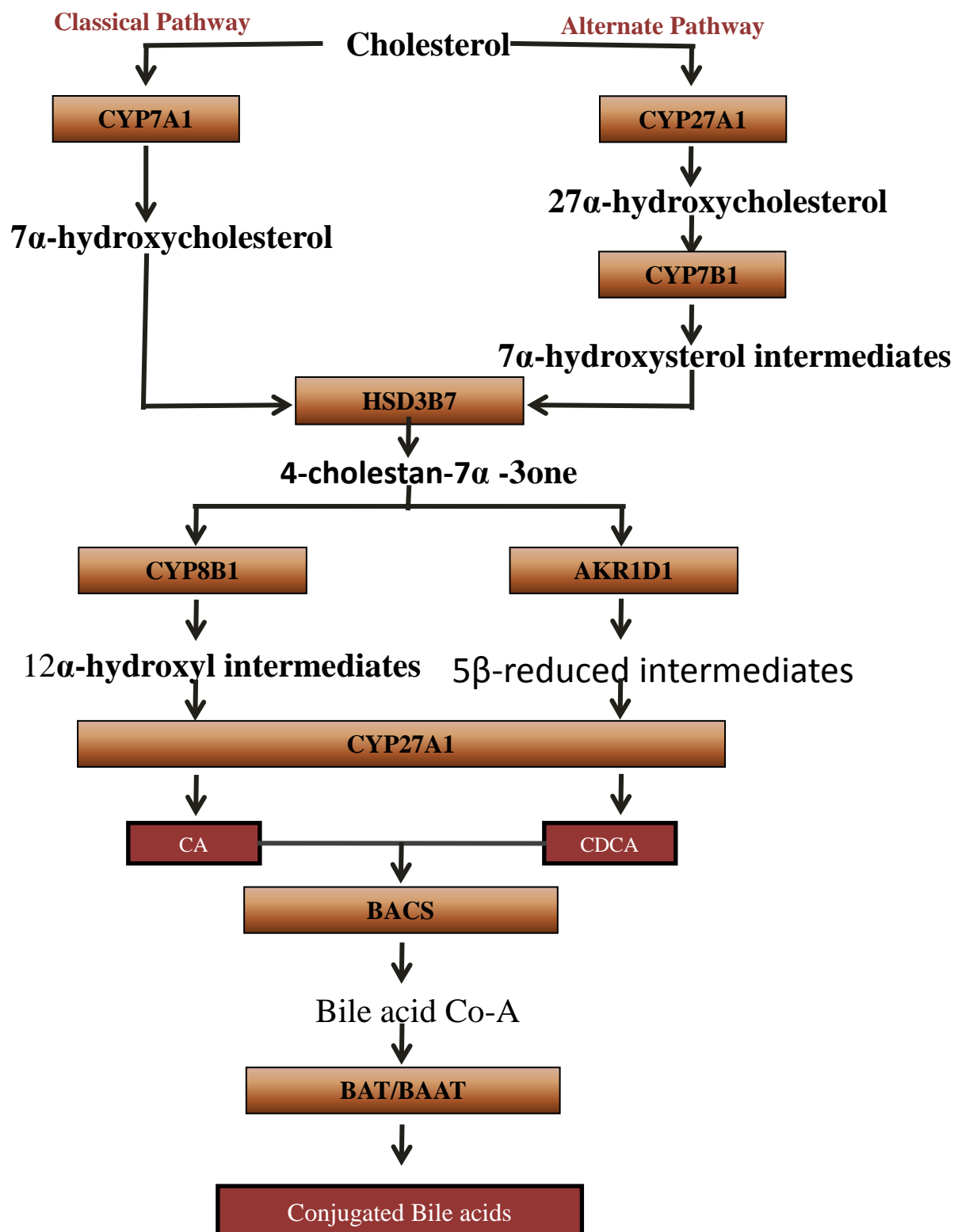


Figure 1.4: Bile acid synthesis pathways: Initiation, sterol ring modification, side chain shortening and conjugation are four different steps for bile acid synthesis independent of classical and alternate pathways. In classical pathway initiation step hydroxylates 7 α position in cholesterol whilst alternate pathway starts with side chain shortening which involve

CYP27A1. Later CYP7B1 hydroxylates 7α position of 27α hydroxy -cholesterol to form 7α hydroxysterol intermediates which then enters the classical pathway. Sterol ring modification involves HSD3B7, CYP8B1 and AKR1D1 which produce 12α and 5β intermediates. These intermediates are converted into respective bile acid by CYP27A1 viz. CA and CDCA. BAT is the terminal enzyme of this pathway leading to the conjugation of primary bile acids CA and CDCA to form GCA, TCA and GCDCA, TCDCA respectively.

Later CYP7B1 hydroxylates 7α position of 27α hydroxy -cholesterol to form 7α hydroxysterol intermediates which then enters the classical pathway. Sterol ring modification involves HSD3B7, CYP8B1 and AKR1D1 which produce 12α and 5β intermediates. These intermediates are converted into respective bile acid by CYP27A1 viz. CA and CDCA. BAAT is the terminal enzyme of this pathway leading to the conjugation of primary bile acids CA and CDCA to form GCA, TCA and GCDCA, TCDCA respectively.

1.2.2 Alternate pathway

Alternate pathway for bile acid synthesis plays crucial role in removal of cholesterol from tissues other than liver unlike classical pathway which operates entirely in liver. In this pathway Initiation step occurs after side chain shortening in which three different enzymes namely sterol 27-hydroxylase (cyp27) [22,23], cholesterol25-hydroxylase [24] and cholesterol 24-hydroxylase (cyp46) [25] participates depending upon different tissues which leads to the synthesis of oxysterols. The presence of oxysterols in blood serum prompted scientific community to look for metabolism of this cholesterol. Further study revealed that the pathway of oxysterol metabolism is known as Alternate pathway. CYP7B1 encodes 25-hydroxycholesterol 7-alpha-hydroxylase, which is the key enzyme in this pathway. Mutation reported in this enzyme showed phenotype such as severe cholestasis [26].

1.3 Functions of Bile Acids

1.3.1 Digestion and related functions

Being amphiphilic molecule, having hydrophobic and hydrophilic site, it is good in micellar solubilization of polar and non polar lipids. Bile acid forms two different types of micelles one merely bile acid micelle and other is mixed micelle consist of bile acid and polar lipids. It helps in the solubilization of dietary lipid and fat molecule. Bile salts are also reported to have antibacterial action in gut microflora which precludes the overgrowth in small intestine. Floch and coworker studied the antimicrobial action of conjugated and unconjugated bile

acid. They reported that unconjugated bile acids are more potent inhibitor to bacterial cell growth as compared to conjugated bile acids [27]. These findings suggest bile acid might show antimicrobial activity in synergism with fatty acid [28]. Later another group of scientists reported that in obstructive jaundice, bacterial cells overgrow in small intestine and translocate in neighboring organs [29,30]. Slocum and coworker confirmed whether the bacterial overgrowth is due to biliary obstruction or absence of bile in small intestine. They observed higher incidence of bacterial growth in case of bile duct ligated or diverted in rats as compared to sham ligated or control rats [31]. The first report on the effect of bile acid on innate immune response was appeared in 1994, when the histological study of mice deprived of bile for 48 hrs was carried out. They showed the outgrowth of coliforms and tissue damage which hints that multiple factors are involved in the regulation [32]. Lin et. al. identified the presence of 14 kDa bile acid binding protein (iBABP) in enterocytes of rat ileum. This 14 kDa bile acid binding protein was cloned and characterized in different tissues and found predominantly in ileal enterocytes [33]. The level of I-BABP increases in the presence of hydrophobic bile acids such as Chenodeoxycholic acid and deoxycholic acid as compared to more hydrophilic bile acids [34].

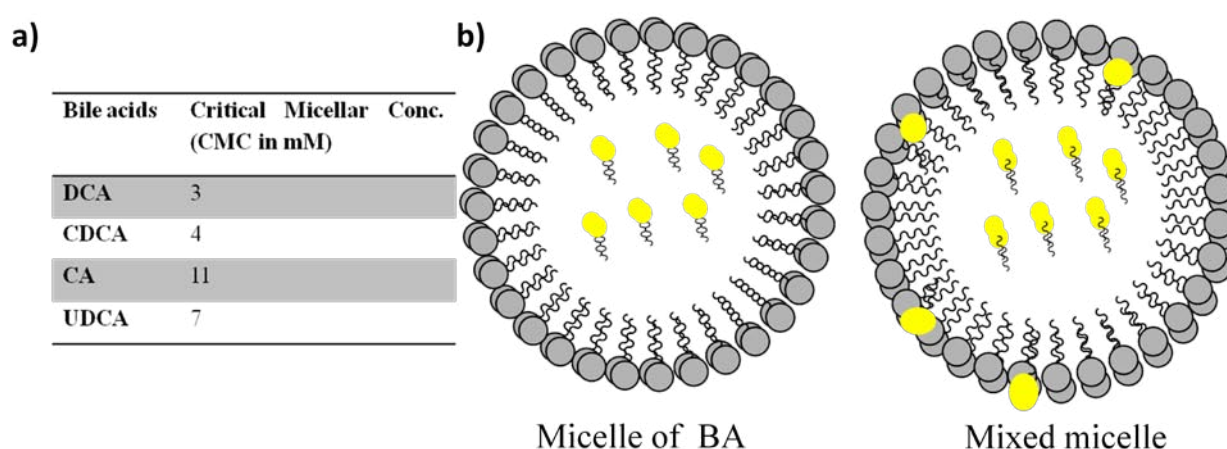


Figure 1.5: Detergent action of bile panel a) shows number of molecules required for formation of micelle, which is different for each type of bile acids. Panel b) shows the different types of micelles formed by bile acids.

At this point researchers established that bile acids play role in their regulation. To check whether this effect is mediated through hormone receptor family, Park and colleagues fused the ligand binding domain of orphan receptors and nucleotide binding domain of yeast transcription factors and found that CDCA selectively activates FXR [35]. Agonist of FXR

i.e. GW4064 has been identified to aid in the study of cholesterol and bile acid regulation [36]. Recently Inagaki et al. reported that activation of FXR blocks the bacterial overgrowth and prevent mucosal injury by synthesizing iNOS leads to the production of nitric oxide which is antibacterial [37]. However this result is debatable because absence of bacterial overgrowth in iNOS deficient mice.

1.3.2 Bile acid in FXR/TGR5 (GPBAR) signaling

Mayurama and Kawamata independently identified the G protein coupled receptor for bile acids, also known as TGR5/GPBAR [38,39]. Initially designated as Orphan receptor, TGR5 is now classified as a subclass of GPCR [40]. Later Watanabe et al., reported that the binding of bile acid activates thyroid stimulating hormone type 2 iodothyronine deiodinase (D2) in a cyclic AMP dependent manner, which is involved in energy expenditure process [41]. TGR5 is involved in multiple metabolic and immunomodulatory function. Its expression level is high in gall bladder as compared to other tissues. Also, it has been demonstrated that removal of Tgr5 gene prevents the gallstone formation in mice [42].

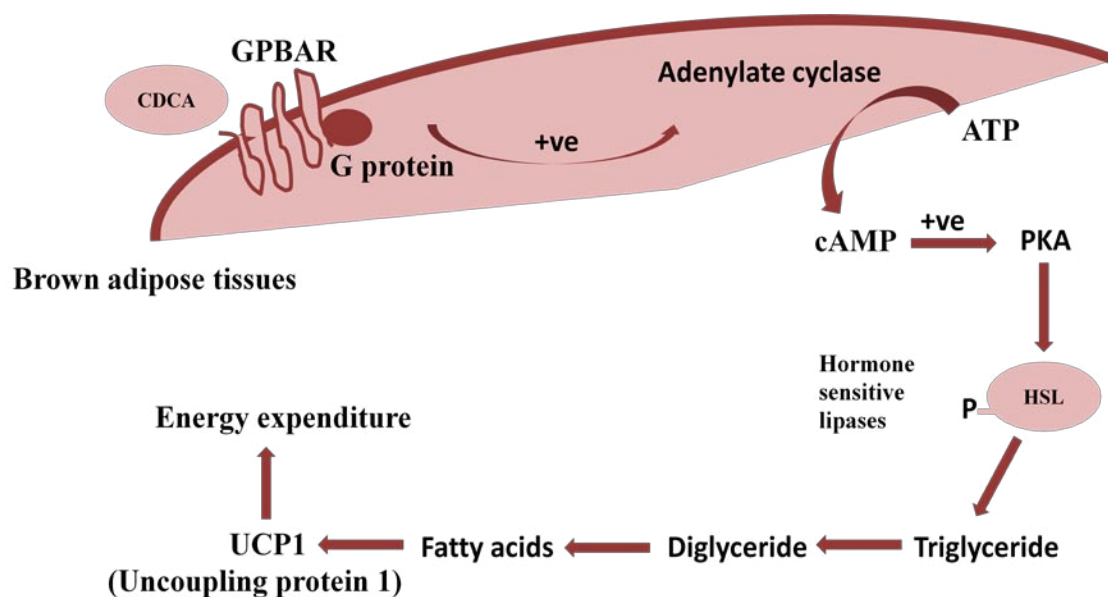


Figure 1.6: Cascade of TGR5 signalling.

1.3.3 Bile salt in regulation of cholesterol

1.3.3.1 Feedback regulation: Cholesterol homeostasis is maintained by de-novo synthesis via mevalonate pathway using acetate and from dietary cholesterol uptake from blood plasma through LDL receptor. The cholesterol is present in the esterified form in blood plasma with lipoproteins such as VLDL (very low density lipoprotein), IDL (intermediate density lipoprotein), LDL (low density lipoprotein), and HDL (high density lipoprotein) [43–45] which is transported through receptor mediated endocytosis to the lysosome where it is degraded. This dietary cholesterol when enters the cell, reduces the expression of HMG co-A reductase which is the rate limiting enzyme for cholesterol synthesis [46]. Moreover, excess of cholesterol also reduces the expression of LDL receptors which internalizes the cholesterol esters through lipoproteins [10,47,48]. This leads to inhibition of de-novo synthesis of cholesterol through mevalonate pathway. Further, the catabolism of cholesterol depends on the conversion of cholesterol to bile acids in liver [49,50] and steroid hormones where it serves as feedstock to the downstream pathway. This process depends upon the oxygenation of cholesterol which generates oxysterol which binds to LXR receptor. Targeted disruption of oxysterol nuclear receptor, LXR α in mice developed fatty liver starting from seven days [51]. Also, several fold increase in CYP7A1 mRNA was observed in LXR $^{-/-}$ mice when fed with high fat diet. This might be because of the presence of LXR response element (LXRE) upstream to CYP7A1 which is rate limiting enzyme in bile acid synthesis [52]. Later, Timothy et. al. delineate rationale behind the liver specific action of LXR α although expressed in different tissues. They found that activation of CYP7A1 by LXR α requires LRH1 and leads to synthesis of bile acids. Later, bile acid binds to Farnesoid X receptor (FXR) which synthesizes small heterodimer protein (SHP). SHP stops the bile acid biosynthesis by binding to LRH1 in a feedback manner and stops its own synthesis in a similar fashion [36]. Inagaki reported that SHP works synergistically with Fibroblast growth factor 15 (FGF15) [53]. The formation of bile acid is tightly regulated by the presence of four genes LRH1, FXR, RXR, SHP (Fig.7).

1.3.3.2 Role of BSH in regulation of bile salts: Decades after Schoenheimer reported the effect of bile acid in the regulation of cholesterol, researcher started looking into the regulatory aspect of bile acids. Subsequent report suggested that bile acids in the bile flow from liver to gall bladder are conjugated while the feces bile acids are deconjugated [54]. Later, Norman and Sjovald found that the labeled bile acids present in small intestine were conjugated; however in cecum conjugated bile acid reduces to only 20% and even less in fecal content [55]. Recent reports suggest that the binding affinity of Apical Sodium Bile acid

Transporter (ASBT) to unconjugated bile acids were significantly lower than conjugated bile acids [16]. Tettet in 1960 used high molecular weight polymer which binds to Bile acids in intestine and found lowering of serum cholesterol. He concluded that lowering of bile acids lowers serum cholesterol without influencing the distribution of cholesterol in tissues [56]. Shimada et. al. screened large number of anaerobic gut microbe for Bile Salt Hydrolase (BSH) activity [57]. Later, Kobashi et. al. reported that not all microorganisms are active against both tauro and glycol-conjugated bile acids, some are active against tauro conjugated while others are hydrolyzing only glyco- conjugated bile acids [58].

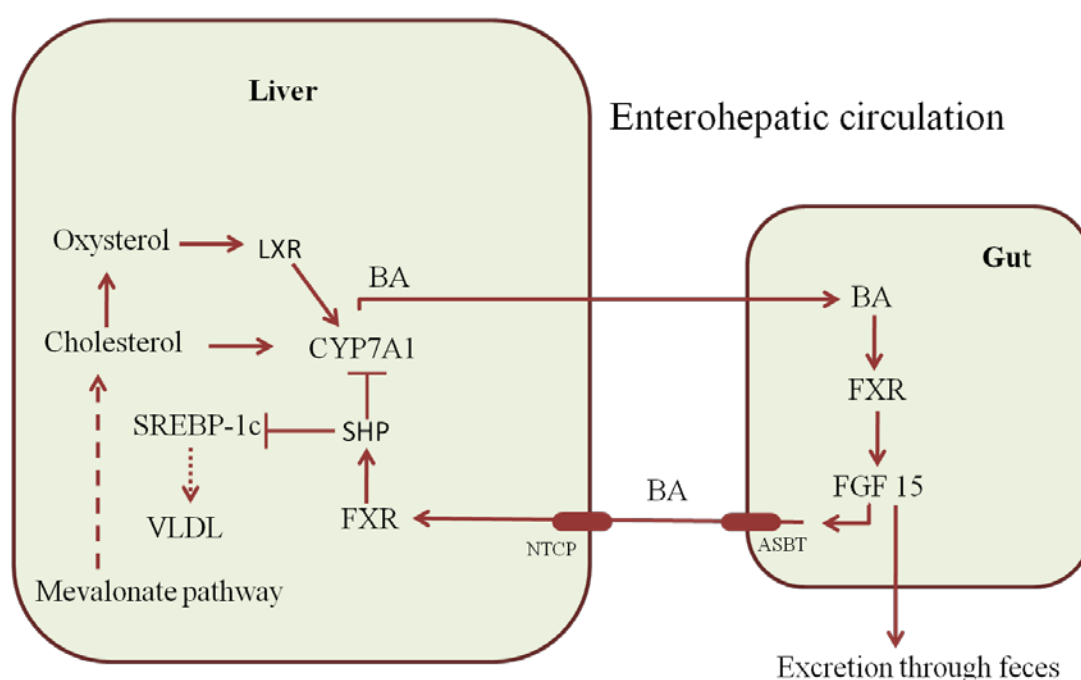


Figure 1.7: Schematic representation of enterohepatic circulation and feedback control of interdependence of bile acids and cholesterol.

Reports on animal studied showed that lactobacilli are majorly responsible for Bile salt hydrolase activity [59,60]. Tanaka et. al. compared the biochemical activity and found that BSH activity is much higher in bifidobacteria as compared to lactobacilli [61]. Begley in 2006 correlated the BSH activity as one of the probiotic selection criteria and a desirable trait for survival in human gut [62]. Studies on human gut microbiome using metagenomic approach revealed archeal species (*Methanobrevibacter smithii* and *Methanobrevibacter stadmanae*) and uncultured bacteria (*Dorea longicatena*) showing BSH activity [63]. This

study also suggests the growth of BSH positive organisms are better in presence of bile acids. The organisms having BSH activity play crucial role in bile acid homeostasis and maintenance of bile acid pool. Recently, Chand et. al. reported Bile Salt Hydrolase from *Enterococcus faecalis* which showed many fold higher specific activity than other reported BSHs. This enzyme acts on both tauro and glyco conjugated bile acids [64]. With few exceptions, gram positive as well as gram negative organism showed tetrameric form of BSH. Panigrahi et. al. reported that presence of indel of 20-22 amino acid in BSH leads to separate clustering of gram positive and gram negative organism [65]. Structural study of bile salt hydrolase revealed indel belongs to tetrameric loop region, which is absent in gram negative BSH [66,67]. Removal of the tetramerization loop in EfBSH changes the oligomerization state from tetramer to dimer and loss in activity. Consurf analysis showed the tetramer loop to be hypervariable, while the core β -sheets are highly conserved and helices in between (**Fig. 1.8**). The present scenario suggests that difference in BSH activity lies in variability of tetramer loop. Structural analysis showed that Arg207 in tetramer loop is protruding in active site of neighbouring chain which might influence the active site geometry. Hydrogen bonding network of Arg207 differ in different monomeric chain. The open conformation showed Arg 207 hydrogen bonded to Asp 21 which is a gatekeeper residue. While in closed conformation the hydrogen bonds were absent. This analysis correlates with the change in active site volume of different chains.

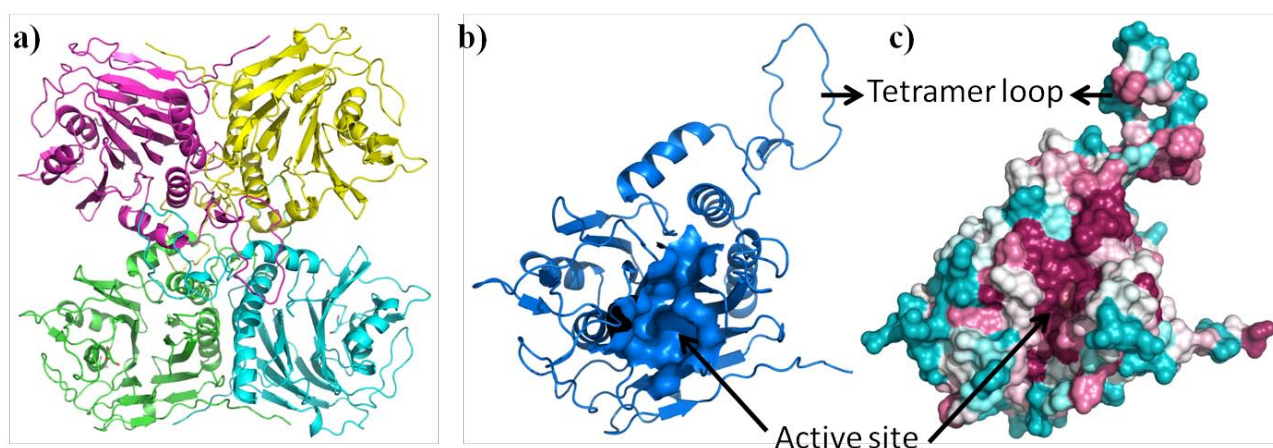


Figure 1.8: Structural study of BSH from gut microbiome: panel a) tetramer of *EfBSH* [68], panel b) CastP analysis showing binding groove of active site pocket of *EfBSH* [69]. Panel c) Consurf analysis of *EfBSH* color ramped from most conserved (red) to least conserved (cyan) [70].

1.4 Distribution of bile salt hydrolases in gut microbiome:

The human body is inhabited by a huge number of bacteria, archaea, viruses, and unicellular eukaryotes. The collection of microorganisms that live in coexistence with their hosts has been referred to as the microbiota. To remain viable in the gut environment bacteria need to be resistant to detergent action of bile acids. Both gram positive and gram negative organisms evolved differently to overcome the toxic effects of bile acids [3]. Metagenomic analysis of gut microbiome reported that BSH is widely distributed in gut microbiome which changes through different stages of life [63,71]. Recent report highlighted the role of bile salt hydrolases in altering host metabolism which has impacts on weight gain, circadian rhythm and immunological reaction in gut and liver [72]. The human microbiota consists of the 10–100 trillion symbiotic microbial cells harbored by each person, primarily bacteria in the gut. Studies of the diversity of the human microbiome started with Antonie van Leeuwenhoek, who, as early as the 1680s, had compared his oral and fecal microbiota and noted the striking differences in microbes between these two habitats. Human Microbiome Project (HMP) revealed the genome size of microbiome to be around 3.5 TB [73] which is many fold higher than the genome size of human and adds up the complexity of synergistic action in human gut. This gut microbiome involved in multifaceted role such as cholesterol regulation, bile acid signaling and glucose homeostasis [74]

1.5 Hypothesis and objectives of the present investigation

As reported by Tennet, removal of bile acids from intestine decreases the level of serum cholesterol by using polymer. Similar results were obtained by Kobayashi using cholestimide resin which binds to bile acids and found decrease in obesity in mice. And deconjugation of bile salt decreases the affinity to bind the epithelial transporter which helps in the enterohepatic circulation of conjugated bile salts. This deconjugation leads to removal of bile acids through feces and confer similar effect to that of synthetic polymer used by Tennet in 1960. Therefore a highly active *Enterococcus fecalis* BSH (*Ef*BSH) could be an alternative to reduce the serum cholesterol level. Using site directed mutagenesis approach, mapping the residue showing significant increase in activity will be of great importance. The contribution of individual residue in allostery may alleviate the need of conventional biochemical characterization and screening of plethora of microbes. In future, this study will complement the gut microbiome project and will be useful for designing the better probiotic drink. The rationale behind our study is to understand the effect of microbiome enzyme on host metabolism with respect to cholesterol regulation and their application in development of

biotherapeutics. Moreover, to screen novel Bile Salt Hydrolases (BSH) for higher activity at high alkaline pH and thermostability. The major objective of our work is to investigate the molecular characteristics for the high activity of *Enterococcus faecalis* Bile Salt Hydrolase (*EfBSH*) through mutational study to understand the mechanism of cooperativity. Further, identification and characterization of highly active microbial BSH from gut source which has implications in therapeutic formulation and development of enzyme based biotherapeutics.

References:

1. Russell, David W. "The enzymes, regulation, and genetics of bile acid synthesis." *Annual review of biochemistry* 72.1 (2003): 137-174.
2. Hofmann, Alan F. "The enterohepatic circulation of bile acids in man." *Clinics in gastroenterology* 6.1 (1977): 3-24.
3. Begley, Máire, Cormac GM Gahan, and Colin Hill. "The interaction between bacteria and bile." *FEMS microbiology reviews* 29.4 (2005): 625-651.
4. Boyer, James L. "Bile formation and secretion." *Comprehensive physiology* 3.3 (2013): 1035-1078.
5. Schoenheimer, Rudolf, and Fritz Breusch. "Synthesis and destruction of cholesterol in the organism." *J. biol. Chem* 103 (1933): 439-448.
6. Gould, R. Gordon, et al. "Cholesterol metabolism. 1. Effect of dietary cholesterol on the synthesis of cholesterol in dog tissue in vitro." *Journal of Biological Chemistry* 201 (1953): 519-528.
7. Bucher, N. L. R., P. Overrath, and F. Lynen. "Enzymes controlling cholesterol biosynthesis in livers of fasting rats." *FEDERATION PROCEEDINGS*. Vol. 18. No. 1. 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998: FEDERATION AMER SOC EXP BIOL, 1959.
8. Goldstein, Joseph L., and Michael S. Brown. "Familial hypercholesterolemia: identification of a defect in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity associated with overproduction of cholesterol." *Proceedings of the National Academy of Sciences* 70.10 (1973): 2804-2808.
9. Brown, Michael S., and Joseph L. Goldstein. "Suppression of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity and inhibition of growth of human

- fibroblasts by 7-ketocholesterol." *Journal of Biological Chemistry* 249.22 (1974): 7306-7314.
10. Brown, Michael S., and Joseph L. Goldstein. "A receptor-mediated pathway for cholesterol homeostasis." *Science* 232.4746 (1986): 34-47.
11. Hofmann, Alan F. "Enterohepatic circulation of bile acids." *Handbook of Physiology. The Gastrointestinal System* 4 (1989): 567-596.
12. Balakrishnan, Anand, and James E. Polli. "Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target." *Molecular pharmaceutics* 3.3 (2006): 223-230.
13. Dawson, Paul A. "Role of the intestinal bile acid transporters in bile acid and drug disposition." *Drug transporters*. Springer, Berlin, Heidelberg, 2011. 169-203.
14. West, Kristy L., et al. "1-[4-[4 [(4R, 5R)-3, 3-Dibutyl-7-(dimethylamino)-2, 3, 4, 5-tetrahydro-4-hydroxy-1, 1-dioxido-1-benzothiepin-5-yl] phenoxy] butyl]-4-aza-1-azoniabicyclo [2.2. 2] octane methanesulfonate (SC-435), an ileal apical sodium-codependent bile acid transporter inhibitor alters hepatic cholesterol metabolism and lowers plasma low-density lipoprotein-cholesterol concentrations in guinea pigs." *Journal of Pharmacology and Experimental Therapeutics* 303.1 (2002): 293-299.
15. Braun, Anne, et al. "Inhibition of intestinal absorption of cholesterol by ezetimibe or bile acids by SC-435 alters lipoprotein metabolism and extends the lifespan of SR-BI/apoE double knockout mice." *Atherosclerosis* 198.1 (2008): 77-84.
16. Zhou, Xiaoming, et al. "Structural basis of the alternating-access mechanism in a bile acid transporter." *Nature* 505.7484 (2014): 569.
17. Hu, Nien-Jen, et al. "Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT." *Nature* 478.7369 (2011): 408.

18. da Silva, Tatiana Claro, James E. Polli, and Peter W. Swaan. "The solute carrier family 10 (SLC10): beyond bile acid transport." *Molecular aspects of medicine* 34.2-3 (2013): 252-269.
19. Roda, Aldo, and Adamo Fini. "Effect of nuclear hydroxy substituents on aqueous solubility and acidic strength of bile acids." *Hepatology* 4.S2 (1984): 72S-76S.
20. Hofmann, Alan F., and Karol J. Mysels. "Bile acid solubility and precipitation in vitro and in vivo: the role of conjugation, pH, and Ca²⁺ ions." *Journal of lipid research* 33.5 (1992): 617-626.
21. Thomas, Charles, et al. "Targeting bile-acid signalling for metabolic diseases." *Nature reviews Drug discovery* 7.8 (2008): 678.
22. Cali, James J., and David W. Russell. "Characterization of human sterol 27-hydroxylase. A mitochondrial cytochrome P-450 that catalyzes multiple oxidation reaction in bile acid biosynthesis." *Journal of Biological Chemistry* 266.12 (1991): 7774-7778.
23. Cali, James J., et al. "Mutations in the bile acid biosynthetic enzyme sterol 27-hydroxylase underlie cerebrotendinous xanthomatosis." *Journal of Biological Chemistry* 266.12 (1991): 7779-7783.
24. Lund, Erik G., et al. "cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism." *Journal of Biological Chemistry* 273.51 (1998): 34316-34327.
25. Lund, Erik G., Joseph M. Guileyardo, and David W. Russell. "cDNA cloning of cholesterol 24-hydroxylase, a mediator of cholesterol homeostasis in the brain." *Proceedings of the National Academy of Sciences* 96.13 (1999): 7238-7243.

26. Sundaram, Shikha S., et al. "Mechanisms of disease: inborn errors of bile acid synthesis." *Nature Reviews Gastroenterology & Hepatology* 5.8 (2008): 456.
27. Floch, Martin H., et al. "The effect of bile acids on intestinal microflora." *The American journal of clinical nutrition* 25.12 (1972): 1418-1426.
28. Hofmann, Alan F., and Lars Eckmann. "How bile acids confer gut mucosal protection against bacteria." *Proceedings of the National Academy of Sciences* 103.12 (2006): 4333-4334.
29. Deitch, Edwin A., et al. "Obstructive jaundice promotes bacterial translocation from the gut." *The American Journal of Surgery* 159.1 (1990): 79-84.
30. Deitch, Edwin A. "Bacterial translocation of the gut flora." *Journal of Trauma and Acute Care Surgery* 30 (1990): 184.
31. Slocum, M. M., et al. "Absence of intestinal bile promotes bacterial translocation." *The American surgeon* 58.5 (1992): 305-310.
32. Kalambaheti, T., G. N. Cooper, and G. D. Jackson. "Role of bile in non-specific defence mechanisms of the gut." *Gut* 35.8 (1994): 1047-1052.
33. Gong, Yong-Zhong, et al. "Molecular cloning, tissue distribution, and expression of a 14-kDa bile acid-binding protein from rat ileal cytosol." *Proceedings of the National Academy of Sciences* 91.11 (1994): 4741-4745.
34. KANDA, Tatsuo, et al. "Regulation of expression of human intestinal bile acid-binding protein in Caco-2 cells." *Biochemical Journal* 330.1 (1998): 261-265.
35. Parks, Derek J., et al. "Bile acids: natural ligands for an orphan nuclear receptor." *Science* 284.5418 (1999): 1365-1368.
36. Maloney, Patrick R., et al. "Identification of a chemical tool for the orphan nuclear receptor FXR." *Journal of medicinal chemistry* 43.16 (2000): 2971-2974.

37. Inagaki, Takeshi, et al. "Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor." *Proceedings of the National Academy of Sciences* 103.10 (2006): 3920-3925.
38. Kawamata, Yuji, et al. "AG protein-coupled receptor responsive to bile acids." *Journal of Biological Chemistry* 278.11 (2003): 9435-9440.
39. Kawamata, Yuji, et al. "AG protein-coupled receptor responsive to bile acids." *Journal of Biological Chemistry* 278.11 (2003): 9435-9440.
40. Foord, Steven M., et al. "International Union of Pharmacology. XLVI. G protein-coupled receptor list." *Pharmacological reviews* 57.2 (2005): 279-288.
41. Watanabe, Mitsuhiro, et al. "Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation." *Nature* 439.7075 (2006): 484.
42. Vassileva, Galya, et al. "Targeted deletion of Gpbar1 protects mice from cholesterol gallstone formation." *Biochemical Journal* 398.3 (2006): 423-430.
43. Oncley, J. L., K. W. Walton, and D. G. Cornwell. "A Rapid Method for the Bulk Isolation of β -Lipoproteins from Human Plasma¹." *Journal of the American Chemical Society* 79.17 (1957): 4666-4671.
44. De Lalla, Oliver, and John W. Gofman. *Ultracentrifugal analysis of serum lipoproteins*. No. UCRL-2427. Radiation Lab., Univ. of Calif., Berkeley, 1953.
45. Fredrickson, Donald S. "Plasma lipoproteins and apolipoproteins." *Harvey lectures* 68 (1974): 185-237.
46. Luskey, K. L., et al. "Amplification of the gene for 3-hydroxy-3-methylglutaryl coenzyme A reductase, but not for the 53-kDa protein, in UT-1 cells." *Journal of Biological Chemistry* 258.13 (1983): 8462-8469.

47. Brown, Michael S., and Joseph L. Goldstein. "Lipoprotein metabolism in the macrophage: implications for cholesterol deposition in atherosclerosis." *Annual review of biochemistry* 52.1 (1983): 223-261.
48. Brown, Michael S., and Joseph L. Goldstein. "Regulation of the activity of the low density lipoprotein receptor in human fibroblasts." *Cell* 6.3 (1975): 307-316.
49. Russell, David W., and Kenneth DR Setchell. "Bile acid biosynthesis." *Biochemistry* 31.20 (1992): 4737-4749.
50. Princen, Hans MG, Sabine M. Post, and Jaap Twisk. "Regulation of bile acid biosynthesis." *Current Pharmaceutical Design* 3 (1997): 59-84.
51. Peet, Daniel J., et al. "Cholesterol and bile acid metabolism are impaired in mice lacking the nuclear oxysterol receptor LXR α ." *Cell* 93.5 (1998): 693-704.
52. Lehmann, Jürgen M., et al. "Activation of the nuclear receptor LXR by oxysterols defines a new hormone response pathway." *Journal of Biological Chemistry* 272.6 (1997): 3137-3140.
53. Inagaki, Takeshi, et al. "Fibroblast growth factor 15 functions as an enterohepatic signal to regulate bile acid homeostasis." *Cell metabolism* 2.4 (2005): 217-225.
54. Bergstrom, Sune, and Arne Norman. "Metabolic products of cholesterol in bile and feces of rat. Steroids and bile acids." *Proceedings of the Society for Experimental Biology and Medicine* 83.1 (1953): 71-74.
55. Norman, Arne, and Jan Sjövall. "On the Transformation and Enterohepatic Circulation of Cholic Acid in the Rat BILE ACIDS AND STEROIDS 68." *Journal of Biological Chemistry* 233.4 (1958): 872-885.
56. Tennent, David M., et al. "Plasma cholesterol lowering action of bile acid binding polymers in experimental animals." *Journal of lipid research* 1.5 (1960): 469-473.

-
57. Shimada, Kaoru, Kenneth S. Bricknell, and Sydney M. Finegold. "Deconjugation of bile acids by intestinal bacteria: review of literature and additional studies." *The Journal of infectious diseases* (1969): 273-281.
58. Kobashi, Kyoichi, et al. "A new hydrolase specific for taurine-conjugates of bile acids." *The Journal of Biochemistry* 84.2 (1978): 495-497.
59. Tannock, GERALD W., MICHAEL P. Dashkevicz, and SCOTT D. Feighner. "Lactobacilli and bile salt hydrolase in the murine intestinal tract." *Appl. Environ. Microbiol.* 55.7 (1989): 1848-1851.
60. De Smet, I., Patrick De Boever, and Willy Verstraete. "Cholesterol lowering in pigs through enhanced bacterial bile salt hydrolase activity." *British Journal of Nutrition* 79.2 (1998): 185-194.
61. Tanaka, H., et al. "Screening of lactic acid bacteria for bile salt hydrolase activity." *Journal of dairy science* 82.12 (1999): 2530-2535.
62. Begley, Máire, Colin Hill, and Cormac GM Gahan. "Bile salt hydrolase activity in probiotics." *Appl. Environ. Microbiol.* 72.3 (2006): 1729-1738.
63. Jones, Brian V., et al. "Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome." *Proceedings of the national academy of sciences* 105.36 (2008): 13580-13585.
64. Chand, Deepak, Sureshkumar Ramasamy, and C. G. Suresh. "A highly active bile salt hydrolase from *Enterococcus faecalis* shows positive cooperative kinetics." *Process Biochemistry* 51.2 (2016): 263-269.
65. Panigrahi, Priyabrata, et al. "An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholyglycine hydrolase family." *Microbiology* 160.6 (2014): 1162-1174.

-
66. Rossocha, Maksim, et al. "Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product." *Biochemistry* 44.15 (2005): 5739-5748.
 67. Kumar, R. Suresh, et al. "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* 281.43 (2006): 32516-32525.
 68. Chand, Deepak, et al. "Structure and function of a highly active Bile Salt Hydrolase (BSH) from *Enterococcus faecalis* and post-translational processing of BSH enzymes." *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1866.4 (2018): 507-518.
 69. Binkowski, T. Andrew, Shapor Naghibzadeh, and Jie Liang. "CASTp: computed atlas of surface topography of proteins." *Nucleic acids research* 31.13 (2003): 3352-3355.
 70. Glaser, Fabian, et al. "ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information." *Bioinformatics* 19.1 (2003): 163-164.
 71. Ottman, Noora, et al. "The function of our microbiota: who is out there and what do they do?." *Frontiers in cellular and infection microbiology* 2 (2012): 104.
 72. Yao, Lina, et al. "A selective gut bacterial bile salt hydrolase alters host metabolism." *Elife* 7 (2018): e37182.
 73. Gevers, Dirk, et al. "The Human Microbiome Project: a community resource for the healthy human microbiome." *PLoS biology* 10.8 (2012): e1001377.
 74. Chand, Deepak, et al. "Molecular features of bile salt hydrolases and relevance in human health." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1861.1 (2017): 2981-2991.

CHAPTER II

Dissection of Catalytic Site in Crucial Gut Microbiome Enzyme: Bile Salt Hydrolase

2.1 Introduction

Bile Salt Hydrolases (BSH's) is a class of enzyme that catalyzes the hydrolysis of amide bonds of Bile Acids (BA's). Due to presence of catalytic cysteine residue at N terminal, this enzyme is considered to be a member of Cholylglycine Hydrolase (CGH) family of enzymes, which is a member of Ntn hydrolase superfamily[1]. In CGH family, cysteine acts as nucleophile, which attacks the carbonyl carbon of peptide linkage present in bile acids. Upon hydrolysis of bile acids into cholesterol and corresponding amino acids, the solubility of cholesterol decreases and it precipitates out in the gut and excreted through feces[2]. The unconjugated bile acids have lower affinity for bile acid transporter such as ASBT and NTCP and are not recirculated in blood stream to liver[3–9]. Begley in 2006 correlated the BSH activity as one of the probiotic selection criteria and a desirable trait for survival of microflorain human gut[10]. Recent studies also suggests the growth of BSH positive organisms are better in presence of bile acids[11]. The organisms having BSH activity play crucial role in bile acid homeostasis and maintenance of bile acid pool. With few exceptions, gram positive as well as gram negative organism showed tetrameric form of BSH. Structural study of bile salt hydrolases revealed the active site pocket, which is surrounded with the four distinct loops (L1-L4) and a assembly loop or tetramerization (Figure 2.1) loop which holds the tetramer [12,13]. These tetrameric loop region is result of indel mutation which is evident in gram positive BSH. The presence of indel of 20-22 amino acid in BSH leads to separate clustering of gram positive and gram negative bacteria[14]. Moreover, loop1 and loop2 allow the entry of substrates into active site of enzyme and designated as 'Site A' as shown in *CpBSH* structure, whereas the site of glycine release is referred to as Site L (Figure 2.1) [13,14]. Recently, Bile Salt Hydrolase from *Enterococcus faecalis* (*EfBSH*) showed many fold higher specific activity than previously reported BSHs[15]. This enzyme acts on both *tauro-* and *glyco-* conjugated bile acids. Furthermore, structural studies of *EfBSH* reported that binding pocket of this enzyme has more hydrophobic residues (particularly aromatic residues) near site A consisting of loop1 and loop2, where choloyl moiety binds [14,16]. RMSD calculation of different chains of *EfBSH* exhibited dynamics owing to high RMSD, which indicates allosteric behavior of this enzyme[16].

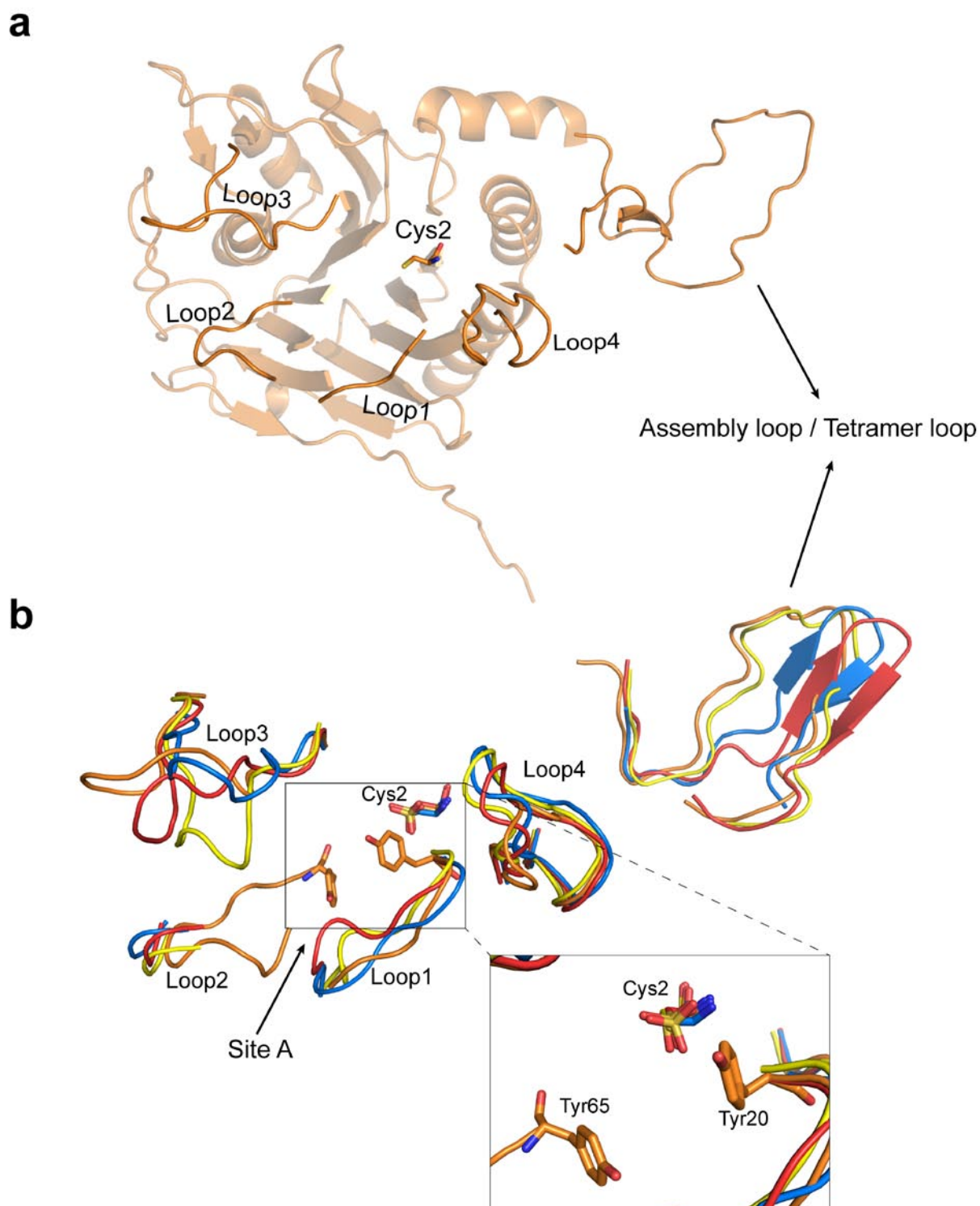


Figure 2.1: Review of structural insight: a) shows the arrangement of loop1 to loop 4 near active site and tetramer loop or assembly loop with reference to whole molecule of wild type *EfBSH*. b) Superposed image of reported BSH structure i.e. *EfBSH* (orange), 4wl3; *LsBSH* (yellow), 5hke; *BtBSH* (Red), 2hez; and *CpBSH* (Blue), 2rlc; showing the entry and exit site of substrates. Inset image shows two tyrosine residues reported to align substrates.

Moreover, disruption of loop2 is assumed to transmit allostery to different chains via tetramer loop (assembly loop) and residues viz. F18, Y20 and Y65 of which Y20 and Y65 belongs to loop 1 and loop2 respectively (Figure 2.1). These residues are proposed to align the substrate in an orientation, which has been projected to be the reason of higher specific activity of *Ef*BSH [16]. However, detailed molecular mechanism of enzyme's higher activity and its allosteric nature has not been explored so far. Owing to the pharmacological importance of BSH enzymes, it is of paramount importance to understand the structure and activity correlation of these enzymes. Further, pinpointing of residues responsible for this activity may help to tailor the molecule to obtain enhanced activity, and thus can be extrapolated to protein engineering of other industrially important enzymes.

Herein, we report that apart from loop1 and loop2, Loop4 also play crucial role in activity of *Ef*BSH. ConSurf analysis [17] showed the tetramer loop to be hypervariable, whereas the core β -sheets are highly conserved (Figure 2.2). However, helices showed intermediate behaviour. The present scenario suggests that difference in BSH activity also lies in variability of tetramer loop. Structural analysis showed that Arg207 in tetramer loop protrudes in active site of adjacent subunit, which might influence the active site geometry (Figure 2.3). Hydrogen bonding network showed Arg 207 interacts with two residues of adjacent subunit viz. Glu269 and Glu21 (Figure 2.3). However, levels of this interaction for each monomeric chain been found to be different. We also proposed here that the open and closed conformation of BSH active site owing to different hydrogen bonding pattern showed by Arg 207 with Asp 21 perhaps acts as a gatekeeper residue. We have used electrostatic analysis to understand the dynamics of residues involved in catalysis or turnover of enzyme. We prepared the mutants of *Ef*BSH viz. E269A and R207A to study the effect on catalytic efficiency and allostery exhibited by the *Ef*BSH molecule. Furthermore, in enzymes, formation and breakdown of hydrogen bond is very crucial to its activity. Understanding the dynamics of hydrogen bonding in active sites of Mutant and wild type will shed some light on the understanding of structure-function correlation. Recently proposed views in hydrogen bonding directionality and force have been exploited to analyze the structural difference in mutant and wild type structures.

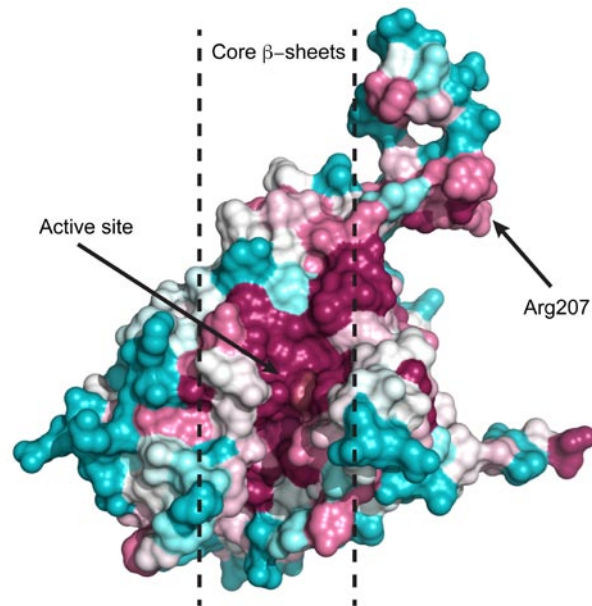


Figure 2.2: Consurf image showing conservation of residues in core β -sheets and assembly loop. The conservation score is color ramped from cyan (variable) to white (average) to red (conserved)

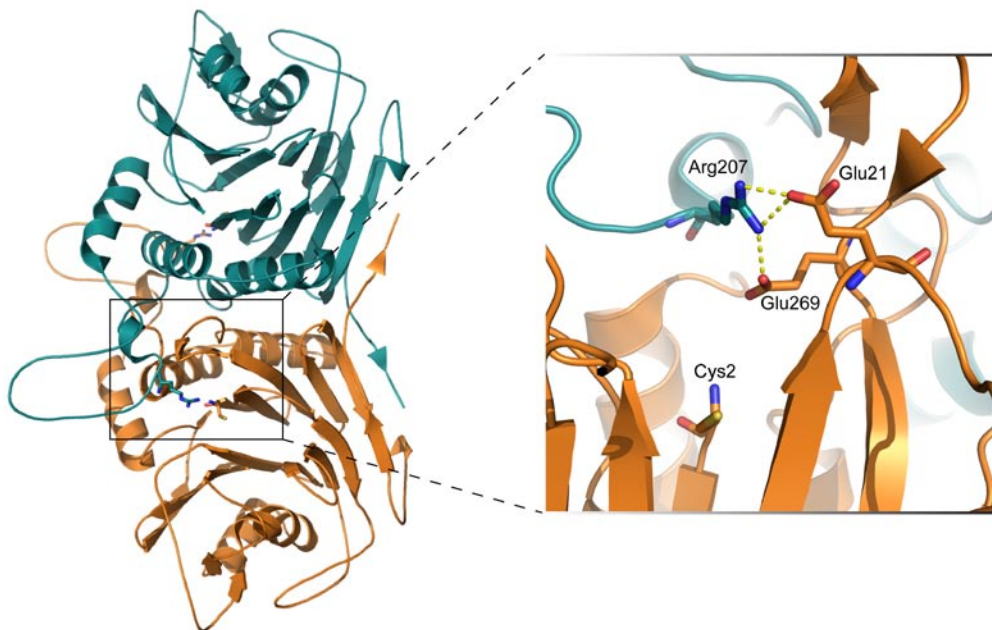


Figure 2.3: Dimeric form of wtEfBSH revealing the Arg207 of chain B protrudes above the active site of chain A. Inset image displays Arg207 from chain B protruding inside the active site pocket and hydrogen bonded to residues of chain A.

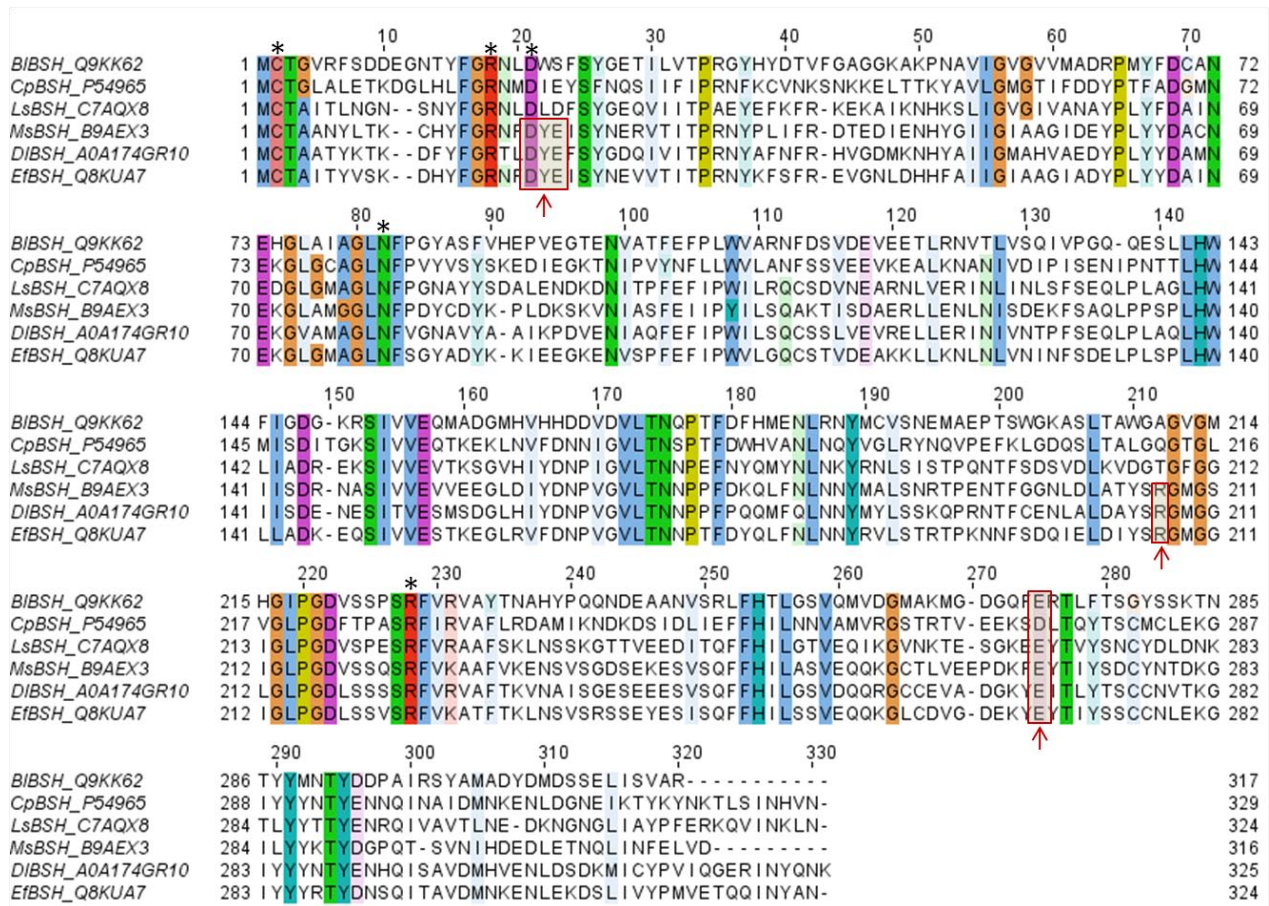


Figure 2.4: Sequence alignment of Bile salt Hydrolases: all the conserved residues are highlighted and catalytically active residues are marked with asterisks. Shown with red arrow highlighting the presence of motif DYE of loop1, Arg207 of Tetramer loop and Glu269 of loop4.

2.2 Materials

All chemicals were procured from Sigma (Sigma, USA), Merck (Merck, USA) and media components from Hi-Media, Thermal cycler (AB sciences, USA), plasmid isolation kits (Qiagen, Germany), Quick Change site-directed-mutagenesis kit (Stratagene, USA), restriction enzymes and ligases from New England Biolabs (NEB,UK) were used. Oligonucleotide DNA (Sigma, USA and Eurofin, India), pE T 22b(+) vector and Nova Blue competent cells were from Novagen (Novagen, USA). *EfBSH* wild type clone was gift from Deepak Chand[15].

Site-directed mutagenesis, Wild-type clone pET22b(+)-*EfBSH* was used as template for preparation of the *EfBSH_Y175A* and *EfBSH_Y175A* clones using Quick Change site-directed mutagenesis kit. Cloned plasmid was transformed into NovaBlue competent cells

and propagated. Recombinant plasmid was isolated and mutation was confirmed by sequencing. The details of experimental have been shown in following subsections.

For crystallization, Commercial screens were procured from Hampton research. Synchrotron radiation source was used for high resolution data, which were collected at INDUS-II, PX-BL21 beamline, RRCAT, Indore having MAR CCD detector. The crystals were flash cooled in a nitrogen stream produced by X-stream to collect diffraction data at low temperature, (Rigaku-MSD, USA). Different crystallographic softwares such as iMOSFLM for data integration, AIMLESS or SCALA for scaling the data during data processing. Molecular replacement was performed in phenix by using Phaser-MR module or by Molrep of CCP4i suite. Model building was done in Coot and refined in Refmac version 5.0.

2.3 Methods

2.3.1 Heterologous expression and purification of *EfBSH* and mutant proteins:

Recombinant plasmid of wild-type *EfBSH* and R207A, E269A, E269Q and E269D mutant was transformed into the expression host *E. coli* BL21 (DE3) cells. Bacterial cells were cultured in LB (Luria broth) media containing 0.1 mg/ml of Ampicillin antibiotics at 37 °C in shaking incubator (Steelmet, India) at 180 rpm. After OD_{600nm} reached 0.8 IPTG (isopropyl β-D-21-thiogalactopyranoside, Merck, USA) induction was given to the final concentration of 1 and 0.5 mM, respectively and culture was shifted to a lower temperature of 16 °C overnight (16–18 h). Bacterial cells were harvested by centrifugation at 5000 rpm at 4 °C for 15 min. Cell pellet was resuspended in the lysis buffer composed of 25mM Tris-Cl pH 8.0, 500mM NaCl, 0.1% (v/v) Triton X-100, 1 mM DTT and sonicated (E-Squire Biotech, India). Cell lysate was centrifuged at 12,000 rpm at 4 °C for 30 min. 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, 20mM imidazole) and washed with buffer A containing 60mM imidazole. Elution was performed using buffer A containing 500 mM imidazole. Eluted fractions were checked on SDS-PAGE. Fractions containing protein of interest were pooled, concentrated using 30K membrane cut-off (Millipore) and subjected to Size Exclusion Chromatography (SEC) using SephacrylS300 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 were pooled and concentrated to 10 mg/ml for crystal set-up and 1.2 mg/ml for kinetics study of mutants.

2.3.2 Site directed mutagenesis: Different mutants of wild-type *EfBSH* i.e. R207A, E269A, E269Q and E269D were prepared using WT *EfBSH* plasmid. PCR was performed at 95 °C for 3 min and 18 cycles of (94/ 30 sec, 55 °C for 30 sec, 72 °C for 6 min 30 sec.) and final ramping of 72 C for 10 min. PCR mixture was then treated with DpnI enzyme to degrade the methylated parental DNA and incubated at 37 °C for 3 hrs. 5ul of PCR mixture is further is transformed in *E.coli* DH5 α cells. Plasmids were purified and sequenced for confirmation of desired mutation. List of primers used for the desired mutations are listed in table 2.1

Table 2.1: List of primers for generating site directed mutagenesis in *EfBSH* wild type

Primer	5'-3' sequence
<i>EfBSH</i> R207AF	GAT ATT TAT AGT GCG GGA ATG GGT GGT
<i>EfBSH</i> R207AR	ACC ACC CAT TCC CGC ACT ATA AAT ATC
<i>EfBSH</i> E269A F	GGT GAT GAA AGA TAT GCG TAT ACG ATT TAT TCT
<i>EfBSH</i> E269A R	AGA ATA AAT CGT ATA CGC ATA TCT TTC ATC ACC
<i>EfBSH</i> E269D F	GGT GAT GAA AGA TAT GAT TAT ACG ATT TAT
<i>EfBSH</i> E269D R	ATA AAT CGT ATA ATC ATA TCT TTC ATC ACC
<i>EfBSH</i> E269Q F	GGT GAT GAA AGA TAT CAG TAT ACG ATT TAT
<i>EfBSH</i> E269Q R	ATA AAT CGT ATA CTG ATA TCT TTC ATC ACC

2.3.3 Measurement of Enzyme Activity and Kinetic parameter of *EfBSH* mutants: Bile Salt Hydrolase enzyme activity can be determined by ninhydrin assay [12,13,18]. 50 ul mixture of enzyme and GCA incubated at 37 °C for 10 min and the reaction was terminated using equal volume of 15% v/v TCA (Trichloroacetic Acid). Enzyme units are defined in terms of glycine release. One unit is defined as the amount of enzyme liberates 1umol of Glycine per minute. The kinetic constant, K_m and V_{max} , of *EfBSH* mutants were calculated using varying substrate concentration ranging from 0.5 mM to 60 mM in triplicates and kinetic constant were determined using Graphpad Prism version 5.0 (Graphpad Inc., USA) From previous study it is well established that GCA is preferred substrate over other glycol- and tauro-conjugated bile salts [15]. The enzyme concentration used for the kinetic study was

1.2 $\mu\text{g}/\mu\text{l}$ throughout the study. Kinetic parameters such as K_m , h and V_{max} were calculated using hill equation. As reported earlier, the presence of DTT does not contribute to allostery, all the kinetic studies were studied using 1mM DTT maintained in the buffer[15].

2.3.4 Molecular Dynamics simulation:

100 ns all atom MD simulation was performed for WT *Ef*B₅H using Amber forcefield[19] in Gromacs 4.5 [20]. All four chains of *Ef*B₅H were utilized for simulation. The *Ef*B₅H tetramer protein was solvated and the system was neutralized by addition of counter ions. Upon neutralization, the system was subjected to minimization and equilibration prior to production run.

2.3.5 Crystallization, diffraction and data collection: Crystallization of Protein is empirical process which is based on several round of trial and error and also, it is a multifactorial process. Factors contributing or sensitive to protein crystallization includes pH, temperature, ionic strength of salts, protein homogeneity and concentration (~ 10 mg/ml), atmospheric pressure, gravity, mixing, presence of substrates/ analogs and sometimes impurity. Temperature and pH play significant role in protein crystallization as it also affects the protein. With time different methods have been developed for crystallizing the proteins including, Sitting-drop, Hanging-drop, micro-batch method and Lipidic cubic phase (LCP)[21] majorly for membrane proteins. Initial screening for protein crystallization is generally performed with commercially available screens to get the conditions required for crystallization. Further optimization was performed by sitting-drop method is used widely to tailor the crystal condition for crystal improvement[22–25]. Moreover, the protein crystal quality can also be improved by seeding, a technique where crushed crystals were transferred to the fresh screen where it induces nucleation which allow the crystal to grow bigger and faster. The streak seeding uses a needle to transfer seeds of crystals to a drop of non-nucleated protein technique [26]. The needle can be introduced into a multiple pre-equilibrated drop which may support mainly growth. Seeding techniques are generally employed in a condition where crystals are very small and of poor quality which takes time to grow.

Here, we purified the *Ef*B₅H mutants proteins i.e. R207A, E269A, E269D, E269Q and wild type was purified to the homogeneity and concentrated upto 10 mg/ml using Amicon ultra-15 ml 30 kDa cutoff centrifugal filters (Millipore, USA) at 4 °C with 4000 rpm. *Ef*B₅H protein was screened against several commercially available crystallization screens including, Index (Hampton Research Corp.) Nextal PACT (Qiagen) and JCSG plus (Molecular Dimensions).

Screenings were performed using vapour diffusion method in 96 well MRC2 sitting drop plates (Hampton) having 200 nl of both protein and screening solution. The good diffractable crystals of *Ef*BSH E269A mutant obtained in JCSG plus Screen having 0.14 M CaCl₂, 0.07M CH₃COONa pH-4.6, 14% v/v isopropanol. While *Ef*BSH R207A mutant crystallized in 0.1M Sodium Citrate Tribasic Dihydrate (pH-5.5), 22% PEG1000 and the crystals were flash frozen in liquid nitrogen (100 K) without cryoprotectant and data were collected at PXBL-21 (Indus-2), RRCAT, Indore, India [27]

2.3.6 X-ray diffraction and cryo-protection

Upon crystallization, crystals are exposed to X-ray beam either in home source or in high energy x-ray wave in synchrotron radiation facility. X-rays are electromagnetic radiation whose wavelength ranges from 0.1 – 10 nm (1 -100 Å) (Blundell & Johnson, 1976). The wavelength of X-rays generated from home source using copper K- α is approx. 1.5418 Å. In a single crystal, X-rays are scattered by individual atoms in a unit cell. This scattered rays, on the basis of their path length, shows constructive or destructive interference. This type of scattering phenomenon is called Thomson scattering. Before exposing crystals to high energy X-rays, crystals are protected from radiation damage (breaks the forces between the lattices and generates free radical which degrades the quality of datasets) using cryoprotectant which is commonly used organic solvent such as ethylene glycol, glycerol, MPD, Isopropyl alcohol.

2.3.7 Data collection and processing

The confirmed protein crystals of *Ef*BSH mutants i.e. E269A and R207A were cryoprotected and flash frozen in liquid nitrogen. These crystals were diffracted at PX-BL21, INDUS-2 PX-BL2, RRCAT, Indore, India. Data were collected using CCD RAYONIX MX-225 detector, at 100K temperature. The mutant *Ef*BSH crystals diffracted well and 180 frames were collected with the exposure of 20 sec each having constant flux. These diffraction images were further integrated through iMOSFLM[28]. Later SCALA[29] module of the CCP4i suite is used to scale the reflections (Collaborative Computational Project No.4, 1994). During data processing, resolution limit were set on the basis of data statistics. SCALA output include χ^2 , Rmerge and $\langle I/\sigma I \rangle$. Mathematically these parameters are defined as (equations 2.1 & 2.2)[30].

$$R_{\text{merge}} = \frac{\sum hkl \sum_{j=1}^N |I_{hkl} - \langle I_{hkl} \rangle|}{\sum hkl \sum_{j=1}^N I_{hkl}} \dots \dots \dots (2.1)$$

$$\chi^2 = \frac{\sum [(I_i - \langle I_i \rangle)^2 / (\sigma I_i^2 N / (N-1))]}{\dots \dots \dots} (2.2)$$

Where, I_i is the intensity measurement of reflection hkl , and $\langle I \rangle$ is the average intensity of 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 (χ^2) is the measure of the good fit of the data and good refinement will have χ^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 [31]. Rmerge alone is not sufficient to decide the high resolution cut-off because of its dependence on the redundancy of data [32,33].

2.3.8 Matthews coefficient

It is an important parameter in the macromolecular crystallography, which helps in determining the number of macromolecules in asymmetric unit provided the unit cell parameter, space group and molecular weight of the macromolecule is known. Mathematically, Matthews coefficient (V_m) is expressed as

$$V_m = \frac{\text{volume of unit cell}}{\text{molecular weight of macromolecule} * Z * X}$$

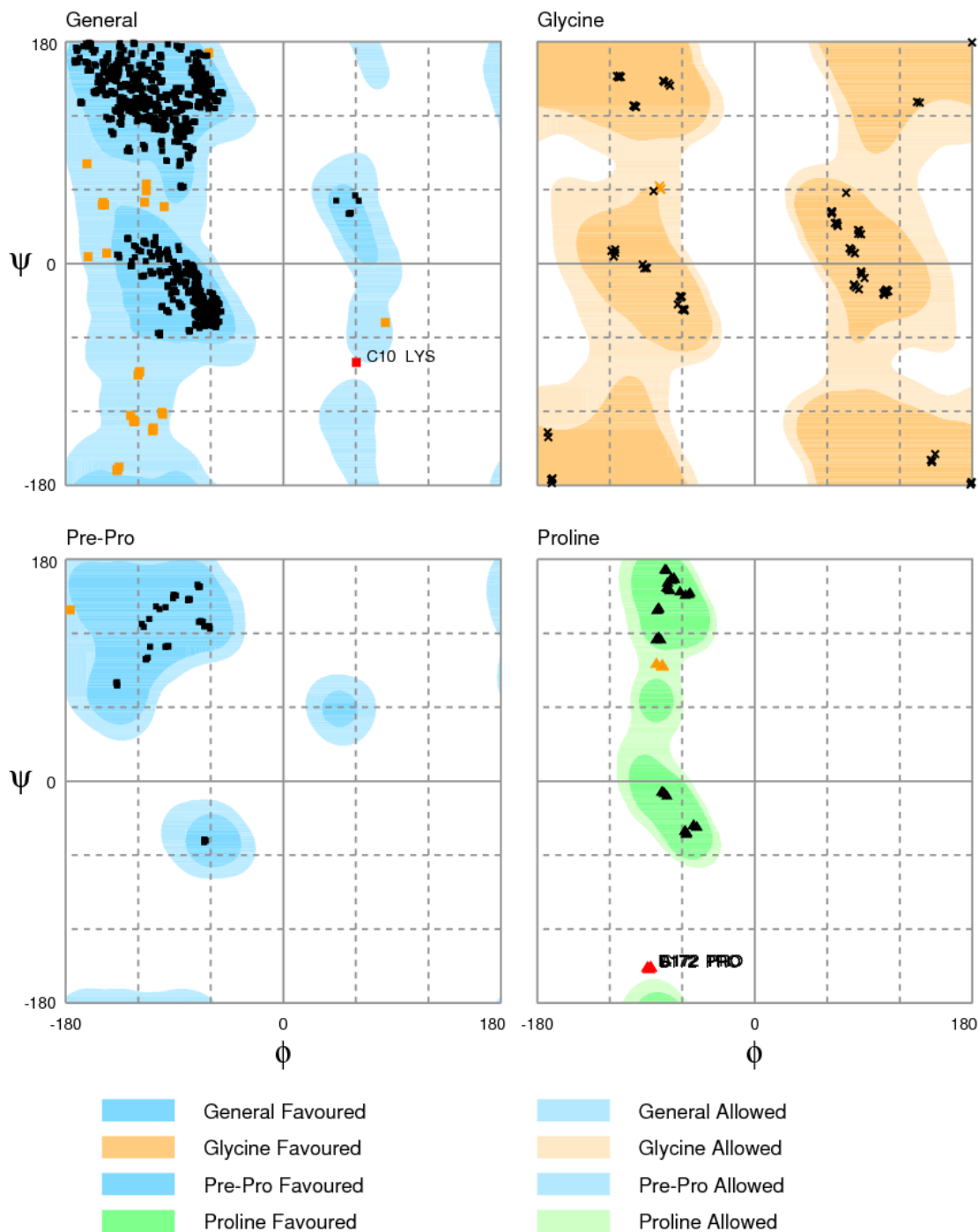
where, Z represents the number of asymmetric unit in unit cell, X represents the number of molecule in asymmetric unit [34]. Previously the difference in 'wet' and 'air dried' crystals were used to decipher the number of molecule in asymmetric unit but small sized crystal poses risk of wrong interpretation. Because of higher solvent content (nearly 40-60 %) [35], number of protein subunits in the asymmetric unit were difficult to elucidate. Matthews in 1968 plotted the graph of volume of asymmetric unit V_s molecular weight of protein and defined V_M as a ratio of the two parameters which sets the lower cut-off beyond that, the possibility of protein crystals would be exceptionally closed pack.

2.3.9 Structure determination and refinement: The data sets were recorded with a Rayonix MX225 CCD imaging plate detector system. The data were processed in iMosflm and scaled in the AIMLESS programs, respectively, and 5% of the data was reserved for cross-validation. Data collection statistics are detailed in the Table 1. The structure was determined using molecular replacement method using data between 50 and 2.0 Å in the Phaser-MR. The crystal structure of BSH from *Enterococcus faecalis* (PDB ID: 4wl3, chain A) was used as

template for search model. Model building and structure refinement were done using Coot and Refmac5 (CCP4 suite), respectively. Phaser-MR model was subjected to restrained refinement using REFMAC of CCP4i suite. Refinement using maximum likelihood method was carried out by multiple cycles in REFMAC5 followed by manual model building and fitting in the Coot 7.0. Water molecules were manually added in Coot at peaks of density above 4σ in the asymmetric unit. Finally, stereochemistry of the structure was validated by SFCHECK program. All structure related figures were prepared in Pymol.

2.3.10 Structure validation:

Upon phasing, the model is built using COOT and refined using REFMAC5[36] until crystallographic model is in agreement with the experimental X-ray diffraction data. The model of *Ef*B5H mutant *viz.* E269A and R207A were validated using SFcheck (Figure 2.5 & 2.6) and *thr*Rfactor and Rfree were in agreement. Rfactor is a structure validation parameter which measures the agreement between the model generated and the X-ray diffraction while Rfree help monitor the model over fitting. Before refinement, about 5-10 % of the experimental data are removed, called as test set, from the data set. Remaining 95% of experimental data are subject to refinement. The calculation of Rfree is based on how well the model predicts the remaining 5-10 % data in the test set, justifies the reason why Rfree is always higher than Rfactor.



Number of residues in favoured region (~98.0% expected) : 1241 (96.7%)
 Number of residues in allowed region (~2.0% expected) : 38 (3.0%)
 Number of residues in outlier region : 5 (0.4%)

Figure 2.5: Ramachandran map of *Ef*BSH E269A mutant generated using SFcheck (CCP4i).

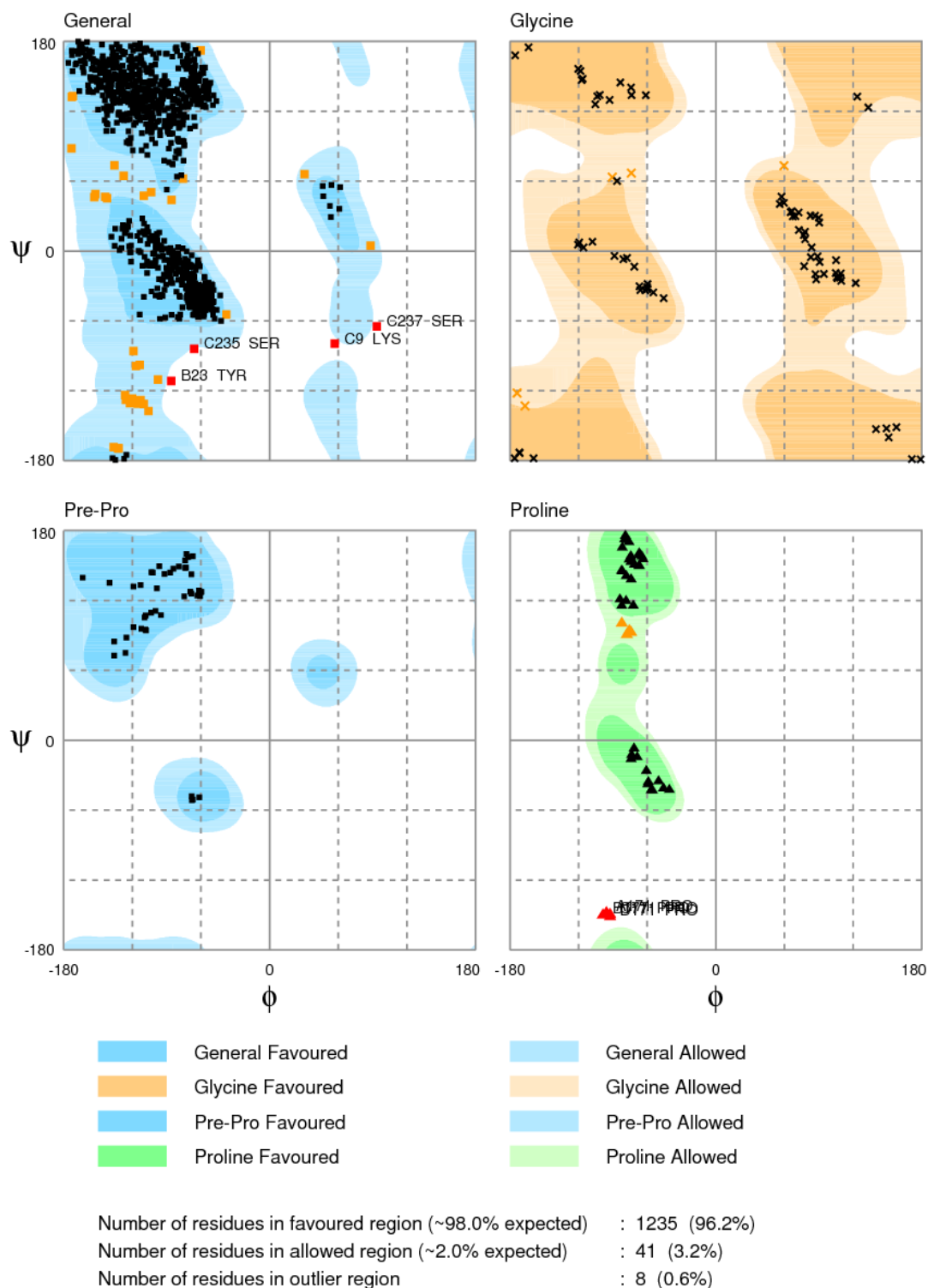


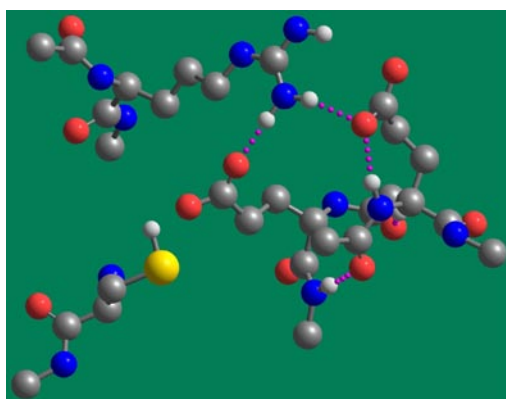
Figure 2.6: Ramachandran map of *EfBSH* R207A mutant generated using SFcheck (CCP4i)

2.3.11 Computational Details: All the DFT calculations, until unless mentioned specifically, were carried out at the M06-2X/6-31G** level of theory [37,38] using the Gaussian 09 suite of quantum-chemical programs [39]. The solvent effect has been introduced into the calculations employing Conductor like Polarization Continuum Model (CPCM) considering water as the solvent ($\epsilon = 78.3553$) [40,41]. The amino acids constituting active sites in all the four chains (A-D) of two cases (the wild type case and the E269A mutant type case) were chopped out from the crystal structure geometry, capped with the acyl (at N-end) and N-methyl (at C-end) groups (making all the amino acids as tripeptide except the cystine which is the N-terminal residue and resulted in a dipeptide after capping) followed by the addition of hydrogen atoms in order to complete the valences of all heavy atom. The geometries were optimized under the constrained condition where coordinates of all the heavy atoms (atoms other than hydrogen) were fixed at their crystal structure geometry and allowing only to optimize the externally added hydrogen atoms. Thus, the conformations of the active sites were preserved in the optimized geometries in order to mimic experimental conditions. Please see Figure 2.7 for the optimized geometries of the active sites of Chain A in two cases specified below.

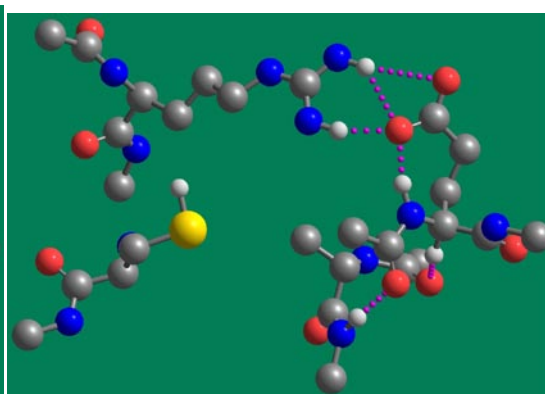
The forces were calculated employing Coulomb's law considering long range electrostatic influence of every atom on the chemistry of the enzyme is significant. Atoms in molecules have been considered as point charges for electrostatic force calculations. The charges have been obtained from the quantum chemical calculations. The NBO4 [42] and Mulliken5 [43] charges have been used to calculate the electrostatic forces. This approach has been shown to give reliable trend in our previous studies [44]. Since S-H bond cleavage (of cystine) has been reported to be one of the key steps in the catalytic process of the enzyme [45,46], the long range influence of amino acids (other than cystine) present at the active sites on the S-H bond cleavage has been studied to evaluate the long range electrostatic influence of the mutation. To calculate this effect, the cystine molecule (in dipeptidic form here) has been considered to be made up of two fragments: (i) the hydrogen atom (fragment 1) and (ii) rest of the part (fragment 2). The contribution of non-cystine atoms on the S-H bond cleavage has been calculated employing following procedure. Every non-cystine atom has been considered to be interacting with each atom of the cystine (from both the fragments). A component of each individual interaction forces experienced by the hydrogen atom in the fragment 1 due to each atom of non-cystine amino acids has been taken along a vector pointed from H to S. Then the magnitude of each component was added in order to get the net electrostatic force exerted by the non-cystine residues of the active site on fragment 1

(hydrogen fragment). The same procedure was repeated for the other fragment but the components of individual forces experienced by atoms in fragments 2 due to atoms in non-cysteine residues were calculated along a vector pointed towards H (from S atom). Then the sum of components of these forces has been taken to obtain the net force experienced by the fragment 2 due to non-cysteine atoms at the active site. Finally, the forces experienced by the two fragments had been added vectorily to obtain the net contributing force from the external environment of the active site. A schematic illustration of how these forces on each fragments have been calculated has been provided in figure 2.7 below. The magnitude of the individual interaction forces between have been calculated employing the dielectric constant of the water ($\epsilon = 78.3553$) as solvent as per the Coulomb's Law.

a)



b)

**Figure**

2.7. The optimized geometry of chain A of a) wild type and b) E269A mutant type under constrained conditions at the M06-2X(water)/6-31G** level of theory. Color representation: red - oxygen, grey - carbon, blue - nitrogen and white - hydrogen.

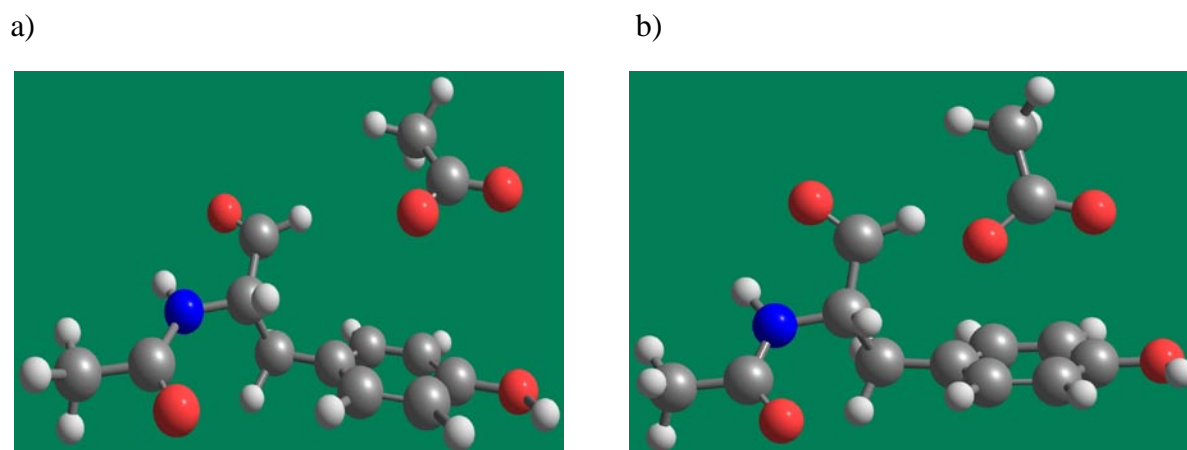


Figure 2.8. Shown here is the optimized geometry of Glu21 and Tyr20 of **a)** E269A and **b)** R207A mutant type under constrained conditions at the M06-2X(water)/6-31G** level of theory. Color representation: red – oxygen, grey – carbon, blue – nitrogen and white - hydrogen.

2.4 Results

2.4.1 Protein purification:

The purification of WT *Ef*BSH, R207A, E269A, E269D and E269Q were achieved as discussed in previous paper[15]. For kinetic study, Citrate-phosphate buffer (pH-5.0) was preferred as similar to optimum pH of wild type *Ef*BSH. The oligomeric profile of all the mutants was also checked in Size Exclusion Chromatography (SEC) using Sephacryl S300 column and compared with the biorad Gel Filtration standard listed in table 2.2 and were found to be tetramer (Figure 2.10). Peak 1 and peak 2 of standard correspond to protein aggregate and thyroglobulin (670 kDa) respectively. Peak 3 correspond to γ - globulin (bovine) of molecular weight 158 kDa, peak 4 correspond to ovalbumin and peak 5 correspond to myoglobin. All the *Ef*BSH wild type and mutant proteins are eluting at peak 3 of standard which corresponds to tetrameric form of BSH the purified fractions were run on SDS-PAGE and confirmed through western blotting using Anti His antibody (Figure 2.9)

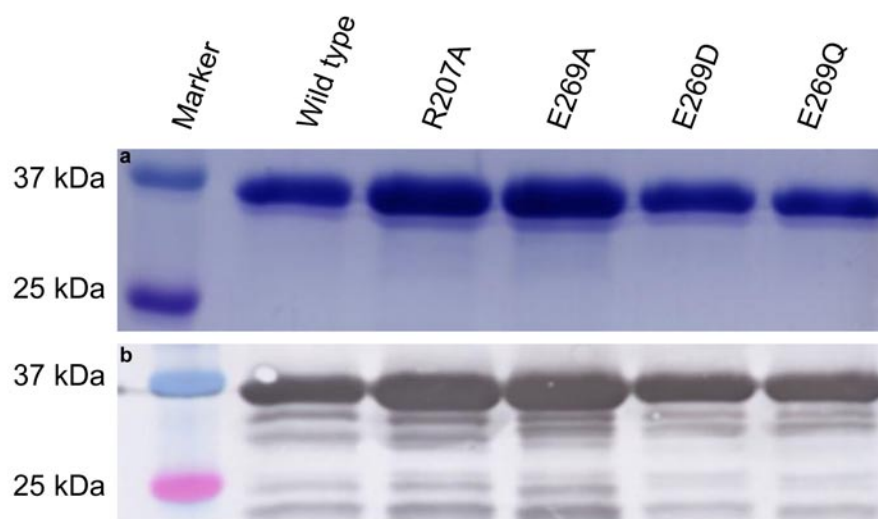


Figure 2.9: Purification of *EfBSH* wild type and mutants. a) SDS-PAGE image stained with Coomassie Brilliant Blue stain showing monomeric *EfBSH* protein bands corresponding to 37 kDa marker . b) confirmation of purified protein using western blot analysis showing band at 37 kDa.

Table 2.2: Protein component used for gel filtration standards

Sr. No.	Protein	Molecular weight
1	Thyroglobulin (bovine)	670,000
2	γ - globulin (bovine)	158000
3	Ovalbumin (chicken)	44,000
4	Myoglobin (horse)	17,000
5	Vitamin B 12	1,350

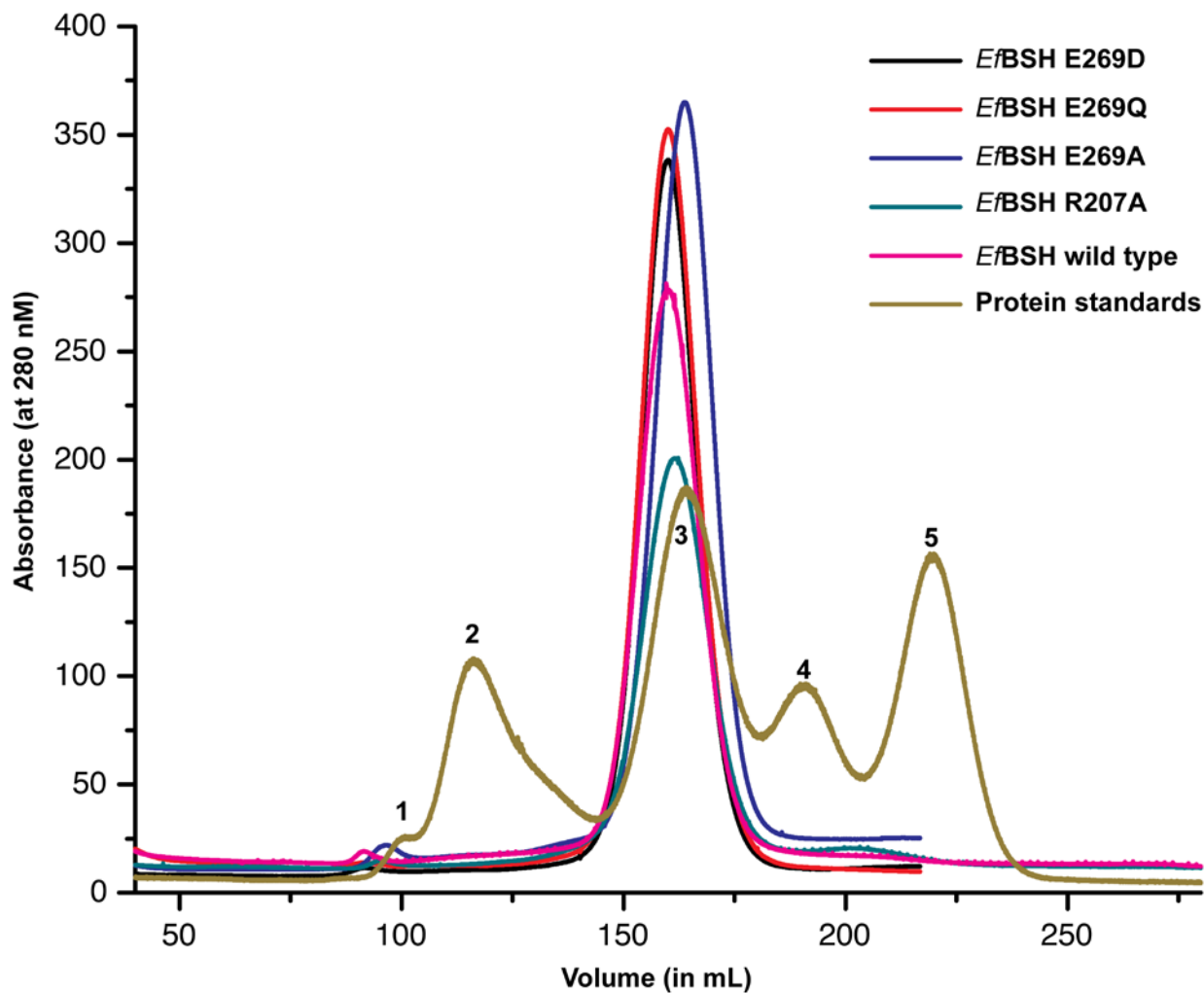


Figure 2.10: Size Exclusion Chromatography (SEC) of *EfBSH* wild type and mutant and compared with protein standards.

2.4.2 Enzyme assay and steady state kinetics:

Ninhydrin assay was used for the biochemical study to estimate the hydrolysed by-product glycine. The allosteric behavior of *EfBSH* was reported for different glyco- and tauro-conjugated bile salts in the previous study [15]. Upon mutation of residues such as R207A and E269A near active site, we investigated the change in kinetic constant of mutant *EfBSH* (Figure 2.12) and also calculated the percent residual activity of the mutants with comparison to wild type as 100% activity (Table 2.3). Both, tauro and glyco-conjugated bile acids were used as substrates for *EfBSH* mutants. Wild type *EfBSH* showed more activity towards Glyco-conjugated bile acids while R207A mutant showed increased activity towards tauro-conjugated bile acids (Figure 2.11). E269A and E269Q mutant showed reduction in biochemical activity against all substrates as compared to WT *EfBSH* and R207A mutant.

Surprisingly, mutating Glu269 to aspartic acid (E269D) completely abolishes the biochemical activity (Figure 2.11).

Table 2.3: Percent residual activity profile of EfBSH mutants as compared to wild type.

<i>EfBSH</i>	GCA	GDCA	TCA	TDCA
Wild type	100	100	100	100
E269A	22.10	27.07	19.80	19.95
E269D	0.18	0.76	0.19	0.41
E269Q	41.73	44.93	64.25	38.67
R207A	61.34	93.97	141.39	240.23

% Residual activity = $EA_{\text{mutant}}/EA_{\text{wild type}} * 100$; EA=Enzyme Activity

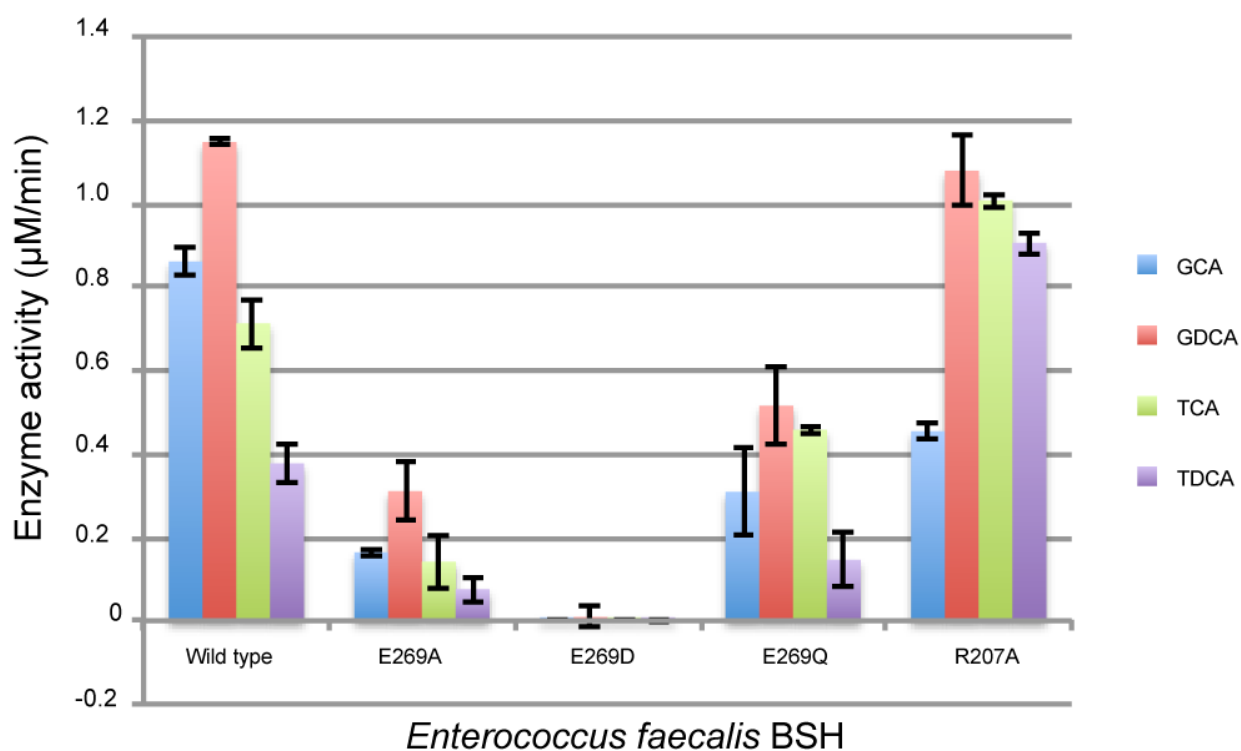


Fig. 2.11: Biochemical activity of EfBSH wild type and mutants against glycol- and tauro-conjugated bile acids.

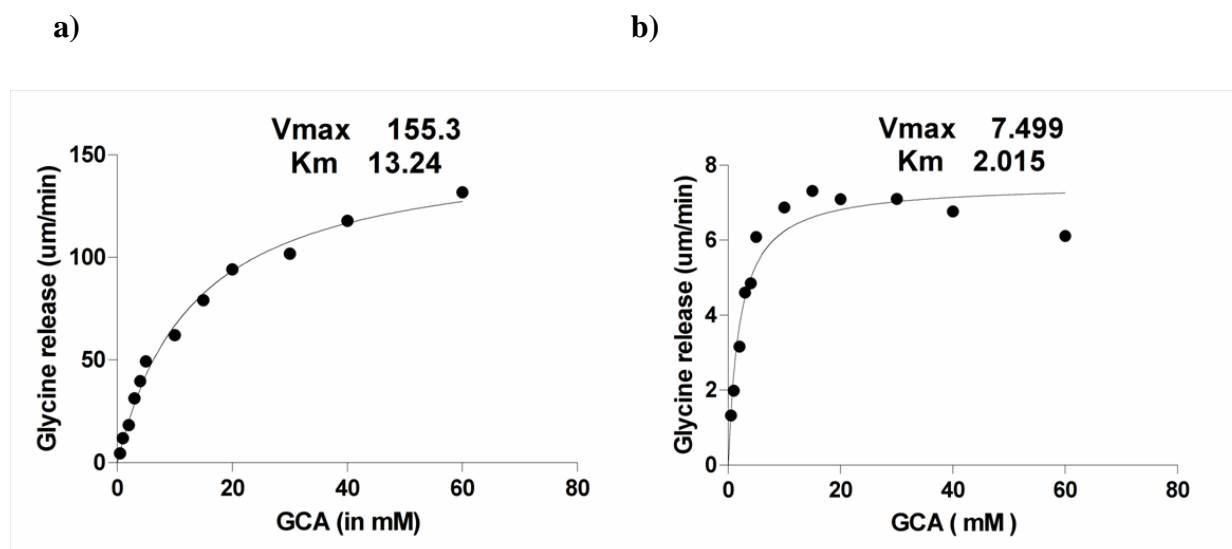


Figure 2.12: Difference in enzyme kinetic behavior of *EfBSH* E269A and R207A mutant enzymes: Left panel represents R207A mutant showing shift from Allostery(Middle panel) to MM kinetics[47]. Right panel represents E269A mutant showing loss in Activity as reflected by Vmax.

2.4.3 Crystallization, data collection and refinement:

The *EfBSH* mutant proteins were crystallized using commercial screens by sitting drop method. Good quality diffracting crystals were obtained in 0.1 M Lithium sulfate, 0.1 M Na citrate tribasic dehydrate pH-5.5 and 20% PEG1000 (Figure 2.13). The crystal of mutant *EfBSH* R207 and E269 diffracted at 2.0 Å and 1.5 Å respectively (Figure 2.14). The R207A and E269A mutant, both belongs to $P2_1$ space group having unit cell parameter $a=59.33\text{Å}$, $b=156.45\text{Å}$, $c=72.91\text{Å}$; $\alpha=\gamma=90$, $\beta=98.62$ and $a=59.84\text{Å}$, $b=156.32\text{Å}$, $c=73.15\text{Å}$; $\alpha=\gamma=90$, $\beta=98.8$, respectively. Details of data collection of *EfBSH* R207A and E269A are given in Table 2.4.

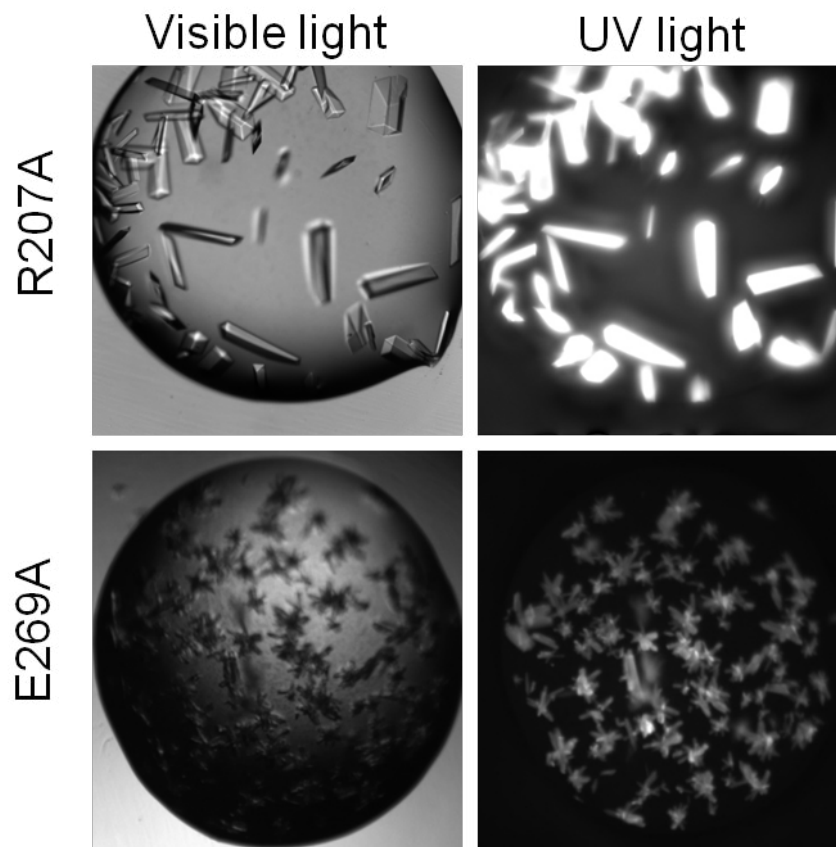


Figure 2.13: Diffraction protein crystal mounted under UV and visible light

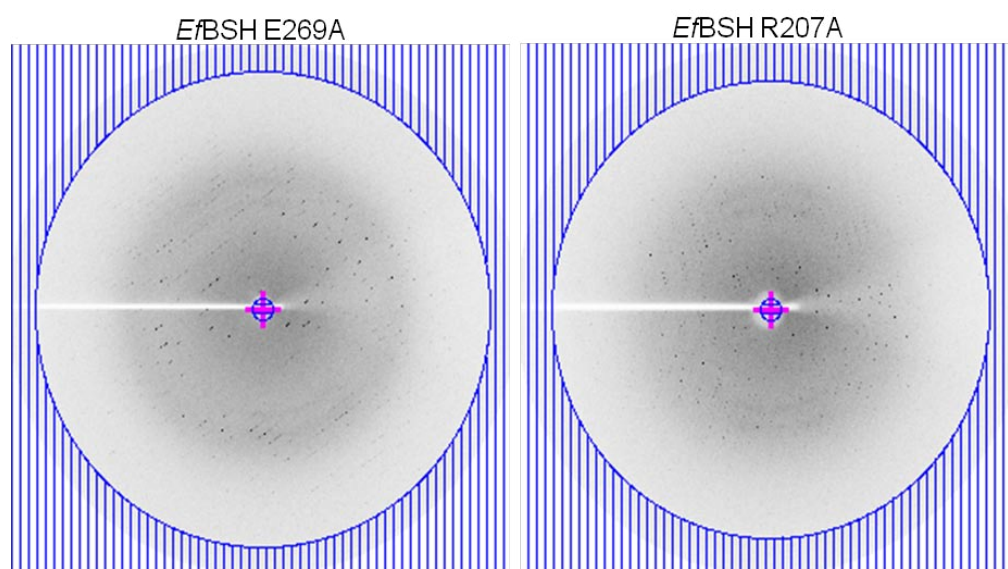


Figure 2.14: X-ray diffraction image of *EfBSH* E269A and R207A mutant

Table 2.4: X-ray diffraction and Data collection statistics for *Ef*BSH mutant structures viz. R207A and E269A. Numbers in parentheses correspond to highest resolution shell.

Data Statistics	R207A_mutant	E269A_mutant
Space group	P2 ₁	P2 ₁
X-ray source	INDUS-II, RRCAT	INDUS-II, RRCAT
Wavelength (Å)	0.97	0.97
Resolution (Å)	38.97–2.0	41.14-2.10
Unit cell parameter	a=59.33, b=156.45, c=72.91; $\alpha=\gamma=90$, $\beta=98.62$	a=59.84, b=156.32, c=73.15; $\alpha=\gamma=90$, $\beta=98.8$
Molecule per asymmetric unit	4	4
Matthews Coefficient (Å ³ Da ⁻¹)	2.21	2.20
Solvent content (%)	44.38	44.0
Total No. of reflection	335415(17397)	290807(17176)
No. of unique reflection	83294(4365)	77248(4592)
Multiplicity	4.0 (4.0)	12.0 (7.4)
Completeness (%)	98.7 (96.5)	100%(100%)
Average I/ σ (I)	5.0 (2.0)	3.8(3.7)
Rsym or Rmerge (%)	5.6 (0.41)	6.1 (0.74)
Refinement Statistics		
Rfactor (%)	20.73	20.35
Rfree (%)	23.62	24.8
RMS Bond Length (Å)	0.0133	0.0152
RMS Bond Angle (°)	1.6380	1.7650
Overall B (Isotropic) from Wilson Plot	13.4	13.3
Most favourable region (%)	96.6	96.2
Additional allowed Region (%)	3.0	3.2
Outlier region (%)	0.5	0.6
Disallowed region (%)	0	0

Data was collected using Image plate and the images were processed and integrated by imosflm. The merged mtz files were scaled in AIMLESS program of CCP4i suite, and 5% of the Data were kept Rfree for cross validation. Data collection and refinement statistics of both the *Ef*BSH mutants are given in table 2.4. In both the cases, the structure was phased using molecular replacement method by using wild type *Ef*BSH structure already deposited in PDB (PDB ID: 4WL3) as a search model template. Matthews coefficient[34] confirms the presence of four chains in asymmetric unit in both the mutant. The model generated was further subjected to model building and refinement using coot and Refmac5 suite of CCP4i software package. The R207A mutant was refined till R factor and R-free reaches to 20.73 and 23.62, respectively. The E269A mutant was refined till R factor and R-free reaches to 20.35 and 24.8, respectively. The geometry of the residues were further confirmed by Ramachandran plot and found no residues in disallowed region.

2.4.4 Structure analysis: *Ef*BSH mutants structures viz R207A and E269A were analyzed and compared with WT *Ef*BSH structure already deposited in PDB (4wl3) and previously reported structures i.e. *Bf*BSH (2hez) and *Cp*BSH (2rlc) and *Ls*BSH (5hke). Analysis of active site of wt *Ef*BSH showed similarity and differences in hydrogen bonding network in active site residues. Residues like Cys2, Asp19, Gln256 and Glu269 is in constant hydrogen bonded network (Figure 2.15) while Arg207, Glu269 and Glu21 showed dynamics in hydrogen bonding pattern in different chains of wt *Ef*BSH (Figure 2.15 & Figure 2.16). Crystal structure of Arg207 mutant showed the absence of any interaction between Glu21 and Glu269 in all different chains while in Glu269 mutant displayed presence of hydrogen bond between Glu21 and Arg207 throughout the different chains (Figure 2.16).

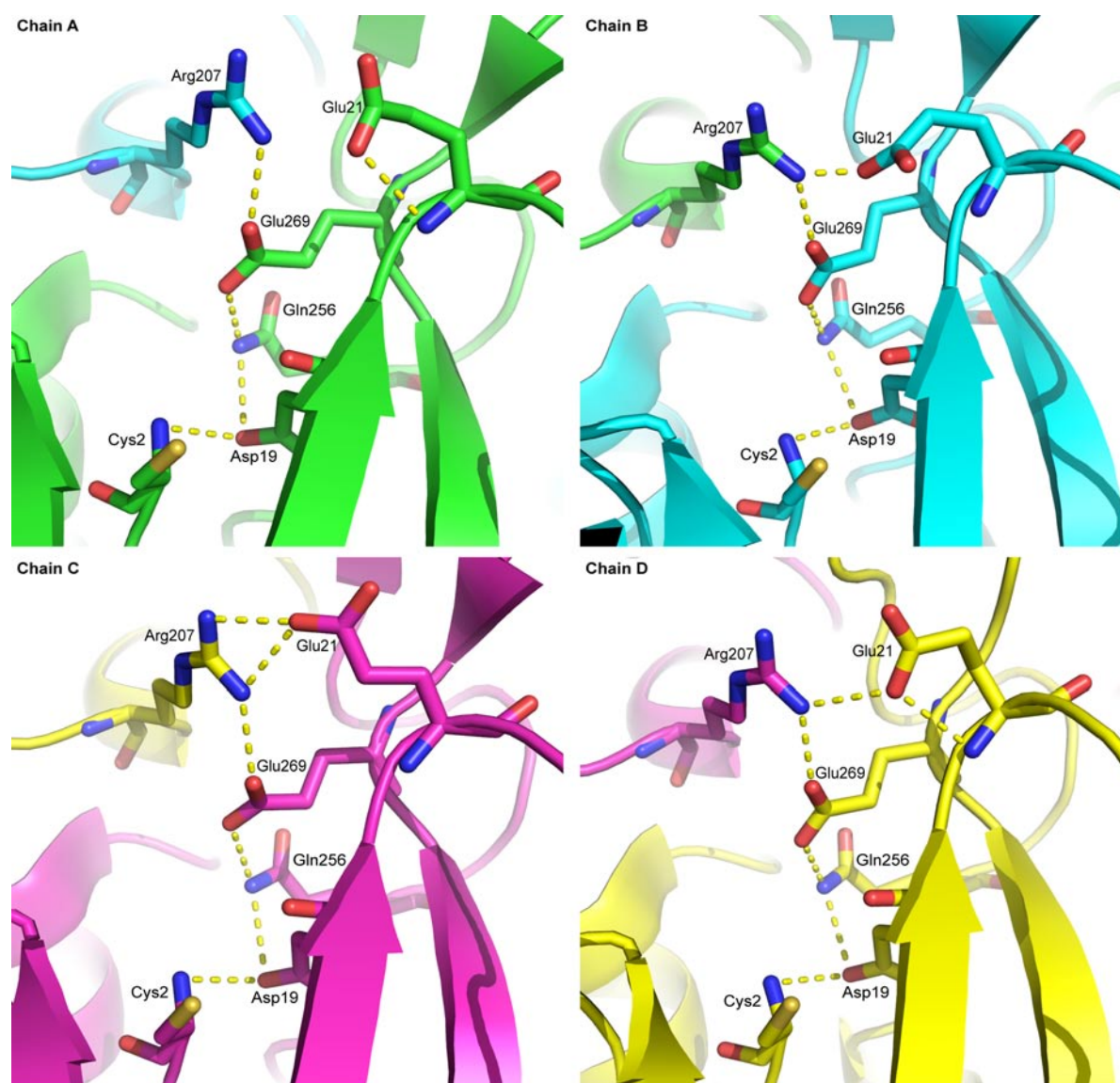


Fig 2.15: Analysis of Hydrogen bonding pattern in active site residues of four different chains of wild type *E/B5H*

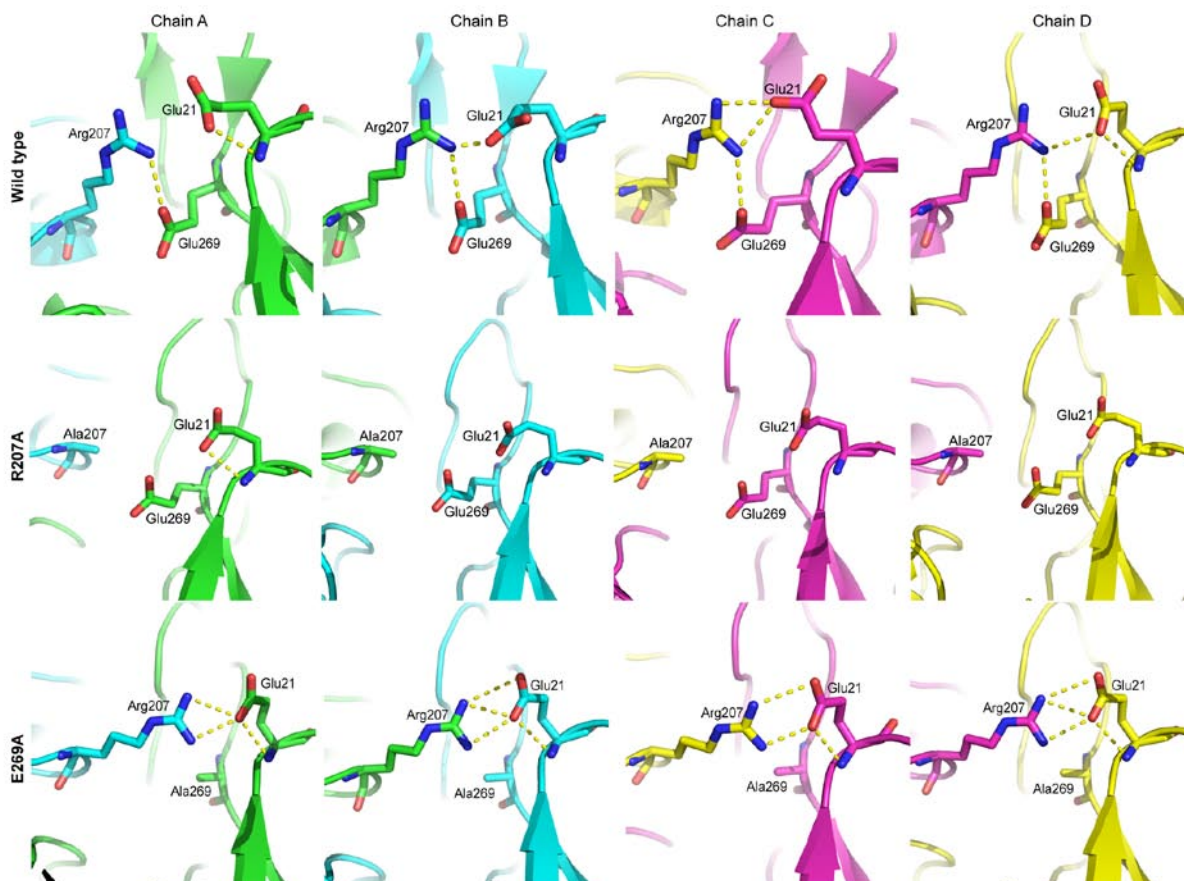


Fig. 2.16: Dynamics of Active site: Top panel describes the dynamics of active site in wild type *EfBSH* between Arg207, Glu21 and Glu269. Middle panel showed no interaction between the specified residues after mutating Arg207. Lower Panel describes the E269A mutant which shows the interaction in all the chains between R207 and Glu21.

The spatial position of Glu269 in *EfBSH* is found to be conserved in *BfBSH* (2HEZ), *LsBSH* (5HKE) and *EfBSH* (4WL3) while in *CpBSH* the spatial position Glu269 is replaced by Aspartic acid and Arg207 is replaced by Ala209 in *BfBSH*, Gln212 in *CpBSH* and Thr208 in *LsBSH* (Figure 2.18). Furthermore, the position of Glu21 in *EfBSH* is replaced by Ser22 in *BfBSH* and Asp21 in *LsBSH* while conserved in *CpBSH* i.e. Glu23.

Tyr20 in loop1 of *EfBSH* reported to align the substrate alongwith Tyr65 of loop2 which is why the biochemical activity is exceptionally higher than the reported BSH. We here, also sought the corresponding residues of Tyr20 in other reported BSH structures. Tyr20 is replaced by Trp21 in *BfBSH*, Ile22 in *CpBSH* and Leu20 in *LsBSH*. Side chain of tryptophan, being bulky in nature, further moving closer to Asn81, residue involved in oxyanion hole formation, and closing the active site which is evident in surface view of Asn81 and Trp21.

Similarly, surface view Leu20 and Asn79 of *Ls*BSH closes the active site. While in *Cp*BSH and *Ef*BSH, the surface view of corresponding amino acids showed open conformation (Figure 2.17). Upon superposition of all four chains of wild type *Ef*BSH showed approx. 2 Å inward movement of tyrosine residues, while in E269A and R207A mutant, such movement is restricted (Figure 2.19). In E269A, Arg207 is found to be hydrogen bonded in all the chains and hence restricting the dynamics of loop1 as compared to wild type (Figure 2.16). Due to the restriction imposed by Arg207, there is significant reduction in RMSD of loop1 and corresponding residues (Figure 2.22). While in R207A mutant, Glu269 and Glu21 do not interact thereby sets free the Glu21 of loop1 (Figure 2.16). Moreover superposition of all four chains of E269A mutant with all four chain of R207A mutant showed two extreme arrangement of Tyr20 (Figure 2.20). The calculated distance between the hydroxyl end of the Tyr20 of the superposed structure of E269A and R207A showed the movement of around 2.6 Å away from each other. Similar results were observed with RMSD plot of *Ef*BSH wild type, E269A and R207 mutant structure which is calculated between the chains (Figure 2.21, 2.22 & 2.23). The position of the Tyr20 and Glu21 of E269A and R207A mutant of *Ef*BSH residues involved in the analysis were further confirmed by 2Fo-Fc map. All the residues are found to be in the electron density developed at 1σ (Figure 2.24). Comparison of the electrostatic profile of active site of WT *Ef*BSH, E269A and R207A mutant showed similar electrostatic nature of wild type and R207A mutant. In case of E269A mutant, the active site becomes more electropositive as compared to WT *Ef*BSH and R207A mutant (Figure 2.25). Docking with GCA showed that none of the chains accepts the ligand from site A which is a binding site for cholyl moiety (Figure 2.26). All the conformation of GCA generated during Ligprep were superposed and showed in the figure 2.26

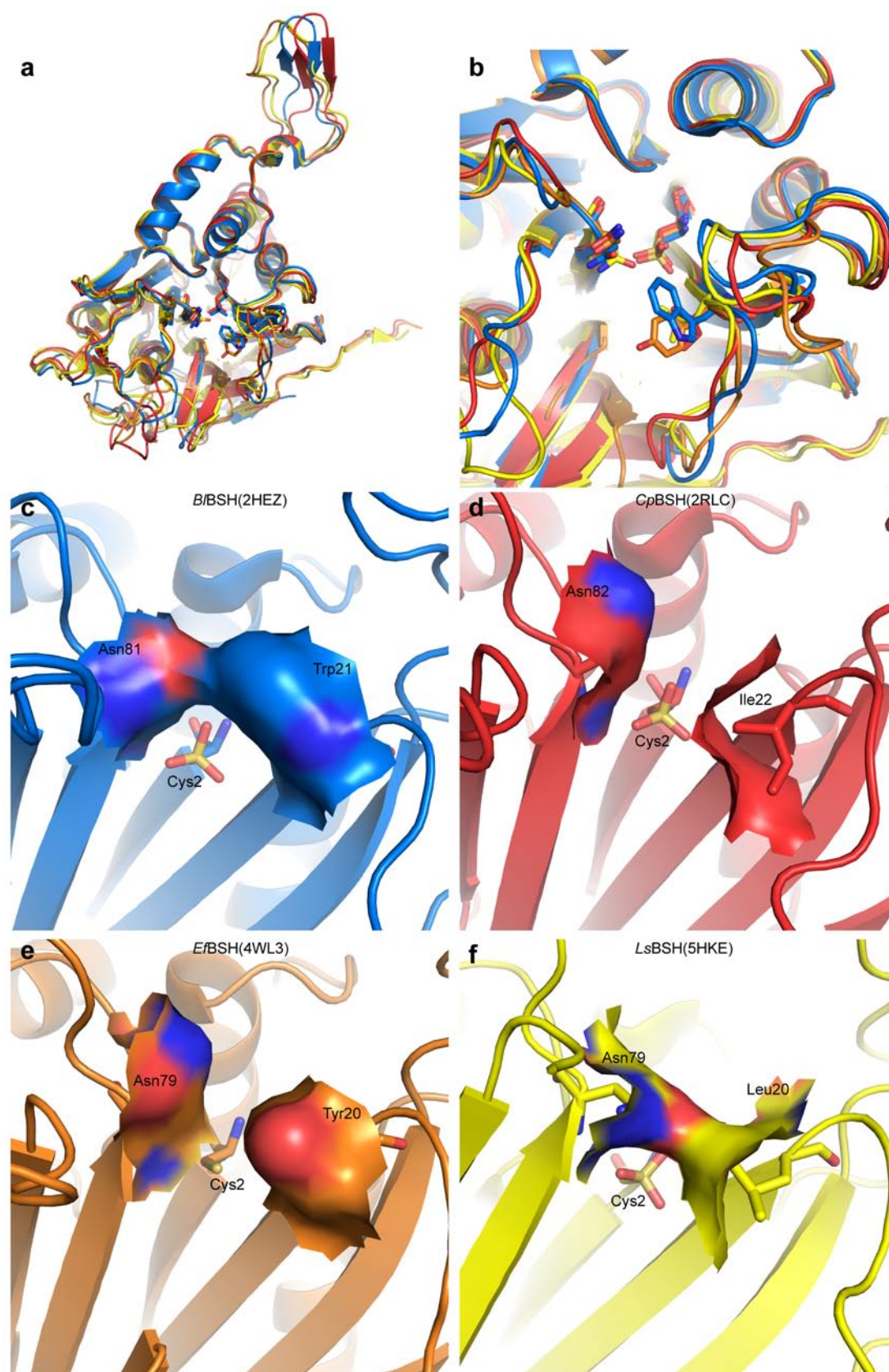


Figure 2.17: Analysis of loop1: a) shows the superposed image of Wt *EfBSH* (orange) PDB Id-4wl3, *B/BSH* (Blue) PDB Id-2hez, *CpBSH* (Red) PDB Id-2rlc, *LsBSH* (Yellow), PDB Id-

4wl3. b) Detailed insight of corresponding residues of Tyr20 in loop1 of EfBSH showed in stick form and individually shown in c-f panel. c) surface view of Trp21 and Asn81 residue of *B*/BSH. d) surface view of Ile22 and Asn82 residue of *Cp*BSH. e) surface view of Tyr20 and Asn79 residue of *Ef*BSH. f) surface view of Leu20 and Asn79 residue of *Ls*BSH.

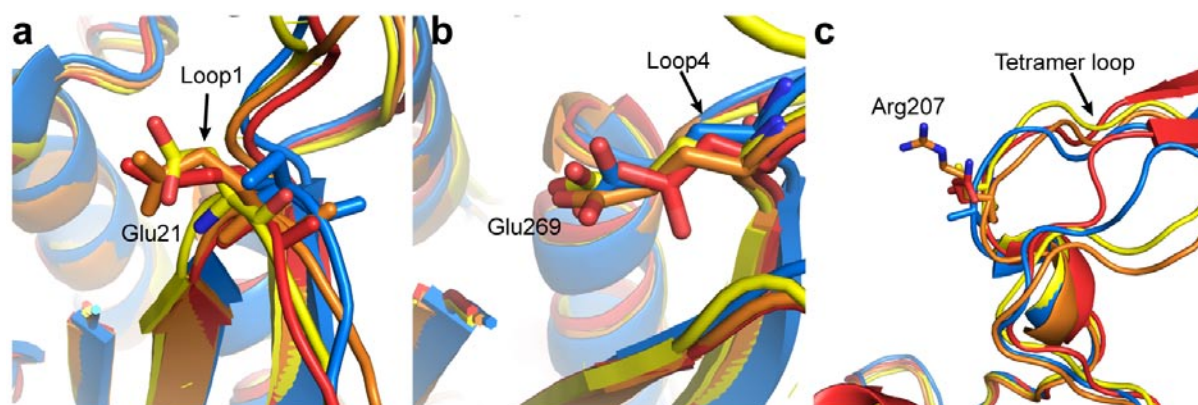


Figure 2.18: Analysis of corresponding residues of Glu21, Glu269 and Arg207 of *Ef*BSH in other reported BSH structures. Color coding of different PDBs is same as above. a) stick form of Glu21(*Ef*BSH) of loop1 and its corresponding residues. b) stick form of Glu269(*Ef*BSH) of loop4 and its corresponding residues. c) stick form of Arg207(*Ef*BSH) of assembly loop or tetramer loop and its corresponding residues in other reported BSHs.

Table 2.5: List of corresponding amino acids of Tyr20, Glu21, Glu269 and Arg207 of *Ef*BSH (PDB Id-4wl3) in *Ls*BSH, *Cp*BSH and *B*/BSH

4wl3	2hez	2rlc	5hke
Tyr20	Trp21	Ile22	Leu20
Glu21	Ser22	Glu23	Asp21
Glu269	Glu271	Asp274	Glu270
Arg207	Ala209	Gln212	Thr208

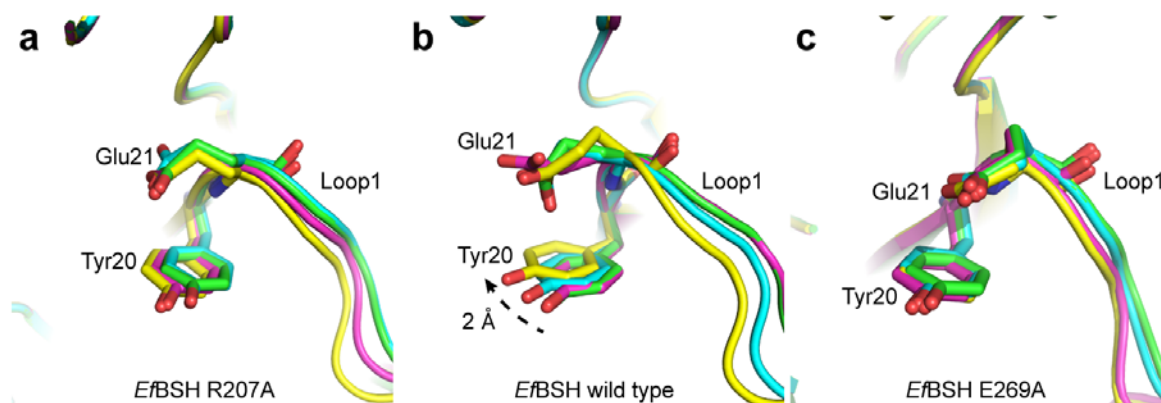


Fig 2.19: Structural superposition of all four chains (Green:Chain A; Cyan: Chain B; Magenta: Chain C; Yellow: Chain D) of *EfBSH* wild type and its mutants. fig a-c shows Stick form of Glu21 and Tyr20 in *EfBSH* R207A, *EfBSH* wild type and *EfBSH* E269A respectively.

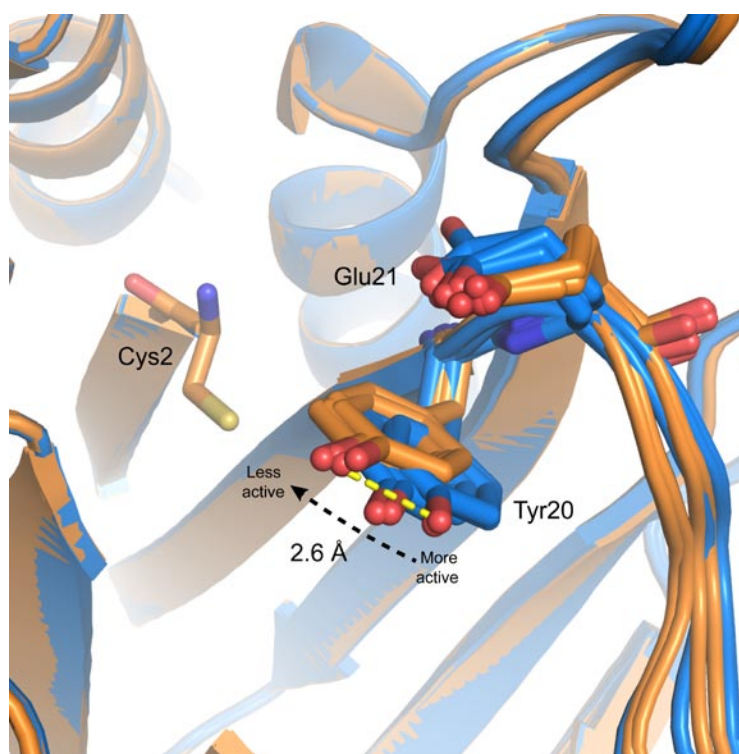


Figure 2.20: Structural superposition of all four chains (colour, orange: E269A and Blue:R207A) of *EfBSH* R207A and *EfBSH* E269A respectively. Glu21, Tyr20 and catalytic cys2 is shown in Stick form.

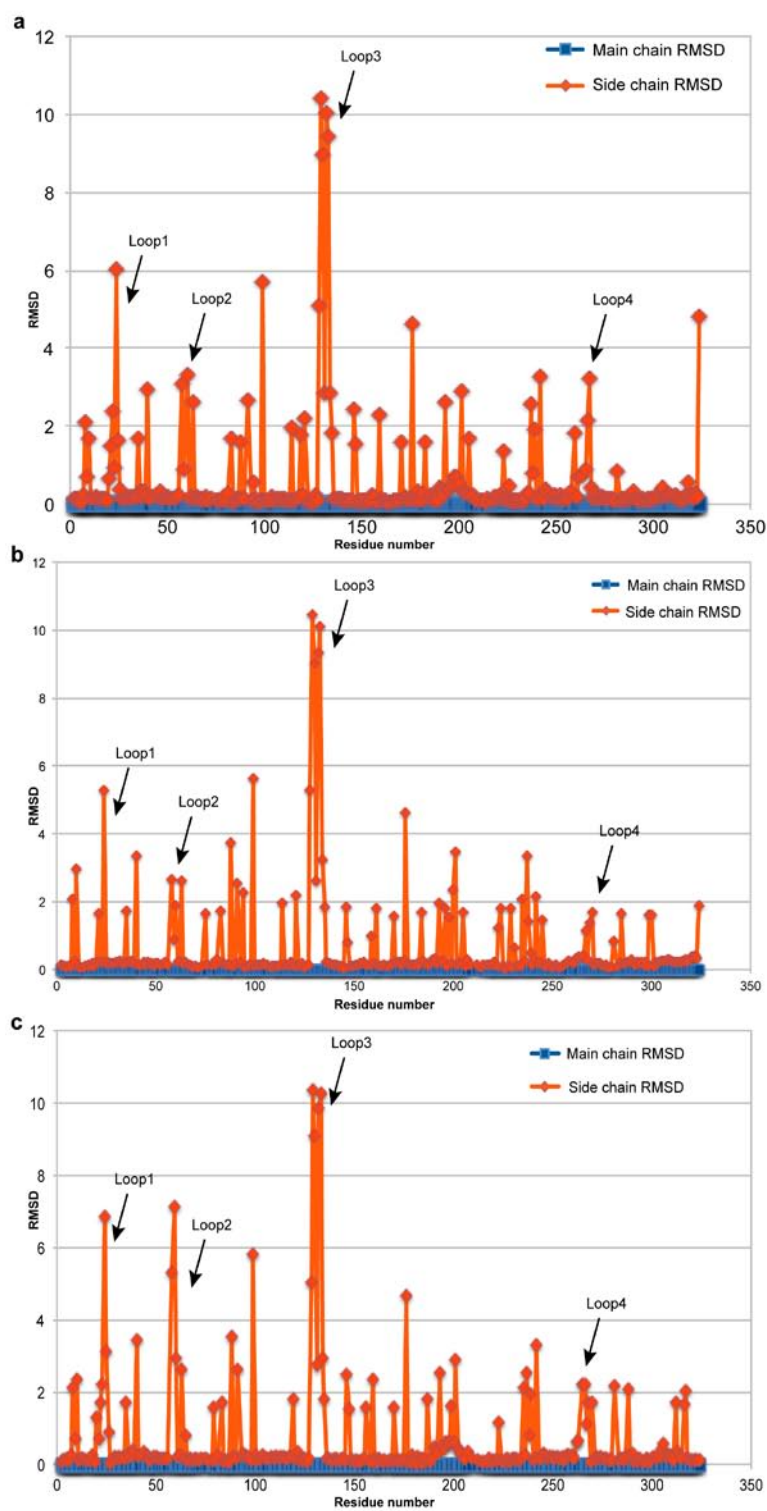


Figure 2.21: Residue wise RMSD analysis of *EfBSH* wild type a) chain A superposed with Chain B. b) Chain A superposed with Chain C. c) Chain A superposed with Chain D.

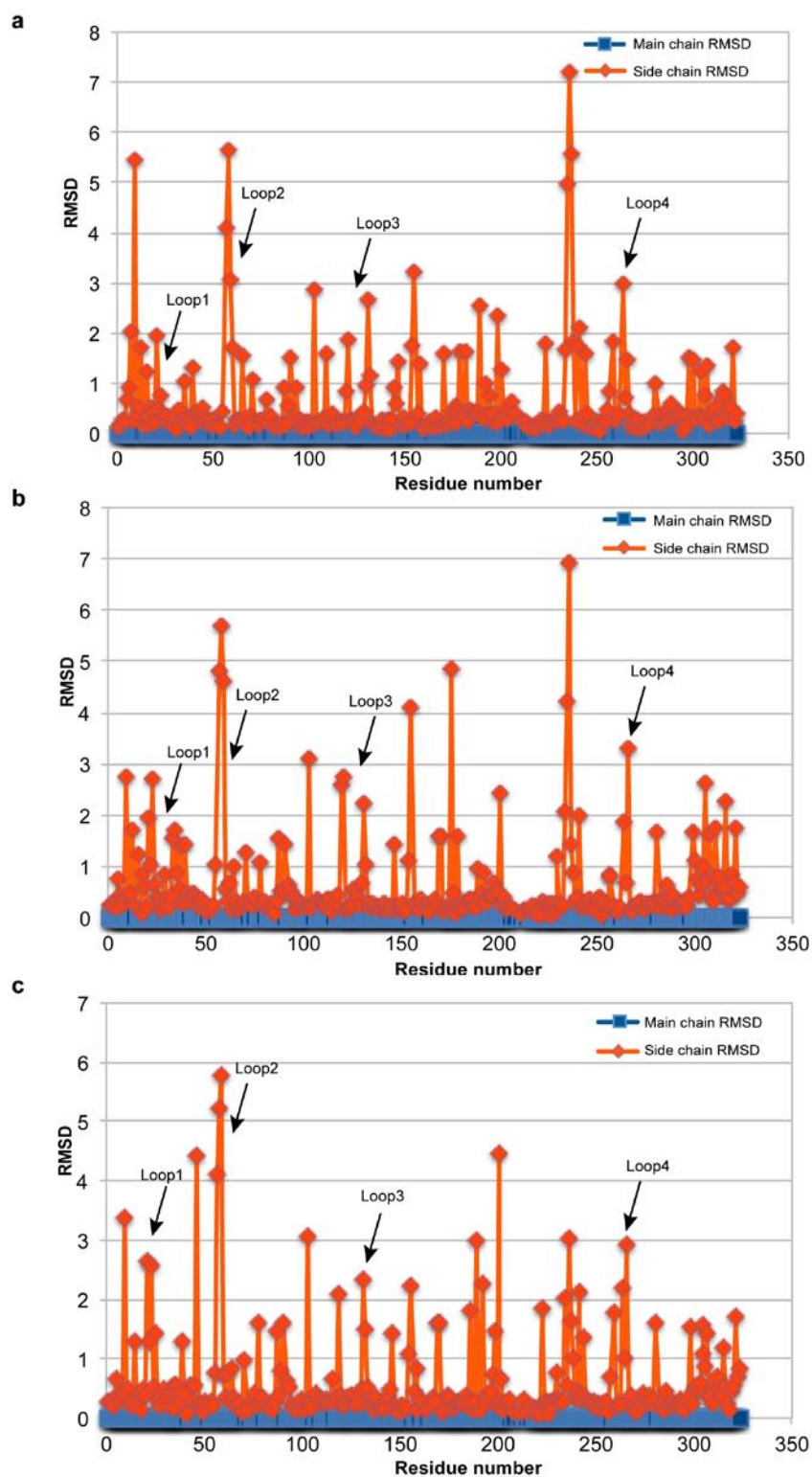


Figure 2.22: Residue wise RMSD analysis of *EfBSH* E269A mutant a) chain A superposed with Chain B. b) Chain A superposed with Chain C. c) Chain A superposed with Chain D.

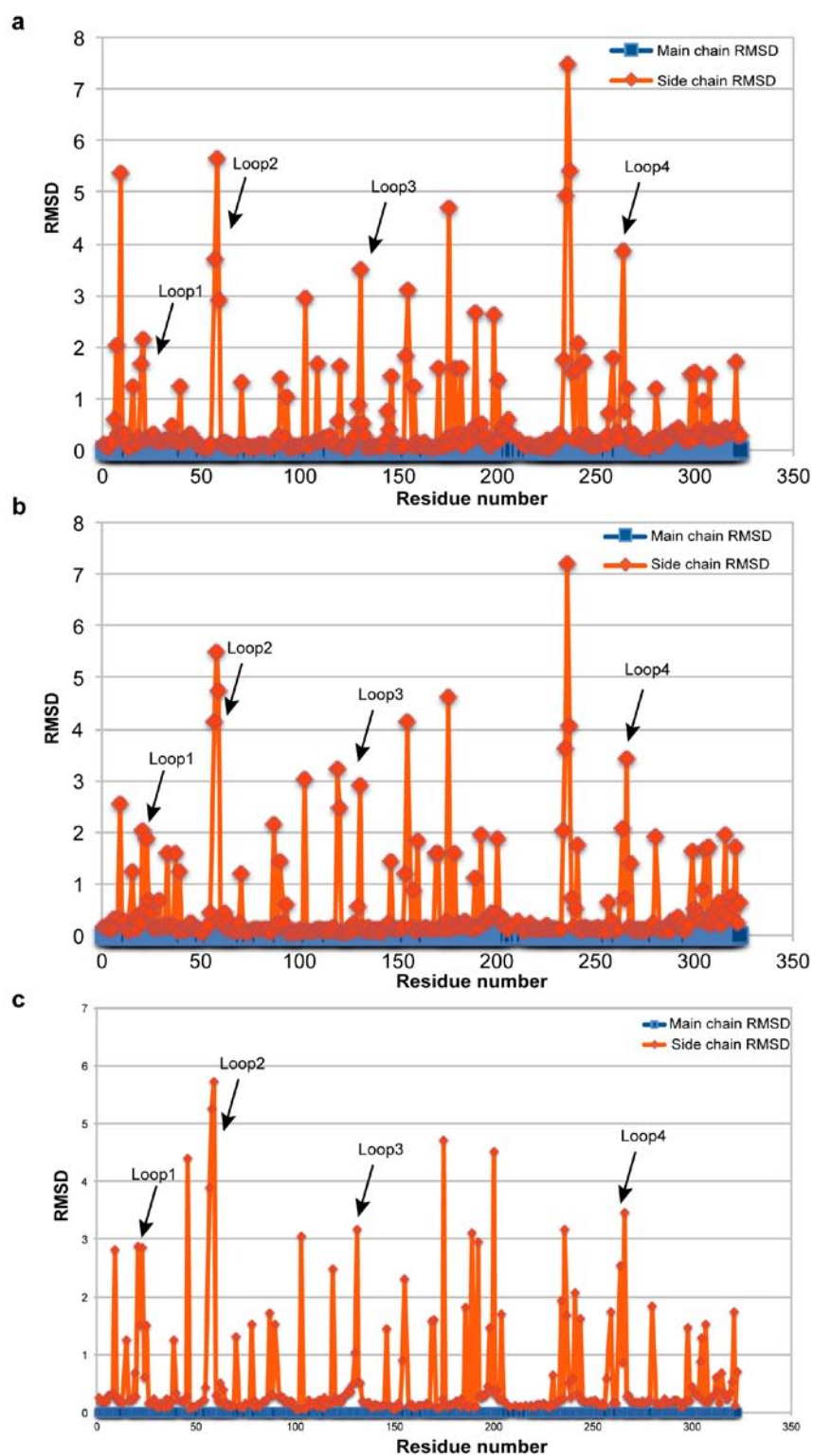


Figure 2.23: Residue wise RMSD analysis of *EfBSH* R207A mutant a) chain A superposed with Chain B. b) Chain A superposed with Chain C. c) Chain A superposed with Chain D.

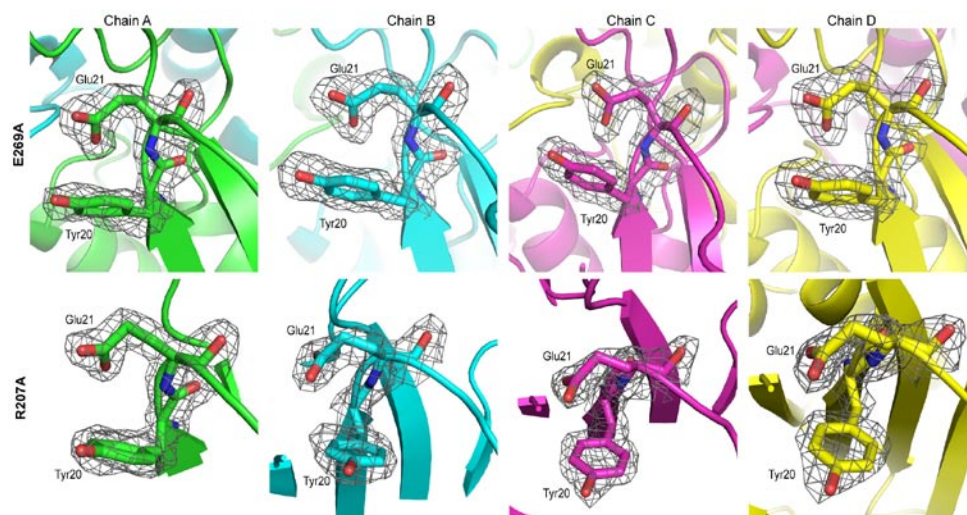


Figure 2.24: 2Fo-Fc map of Glu21 and Tyr20 residues of all four chains from E269A and R207A mutant developed at 1σ .

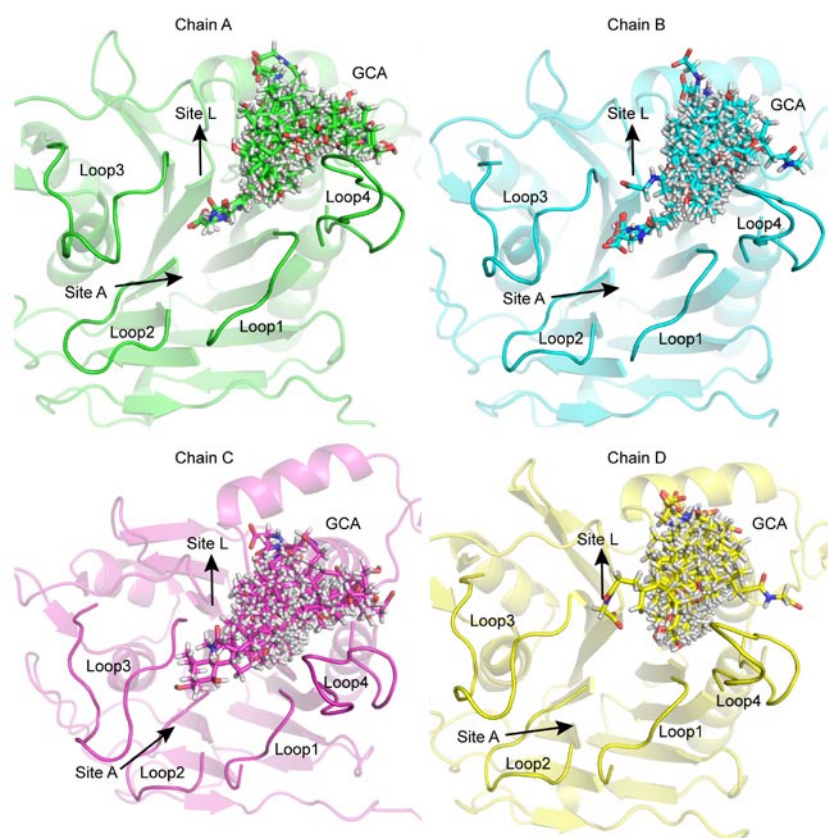


Figure 2.25: Docking of GCA in four chains of E269A mutant shown with respect to loop1 to loop4. Site A represents the cholyl moiety binding site and site L is designated as glycine release site.

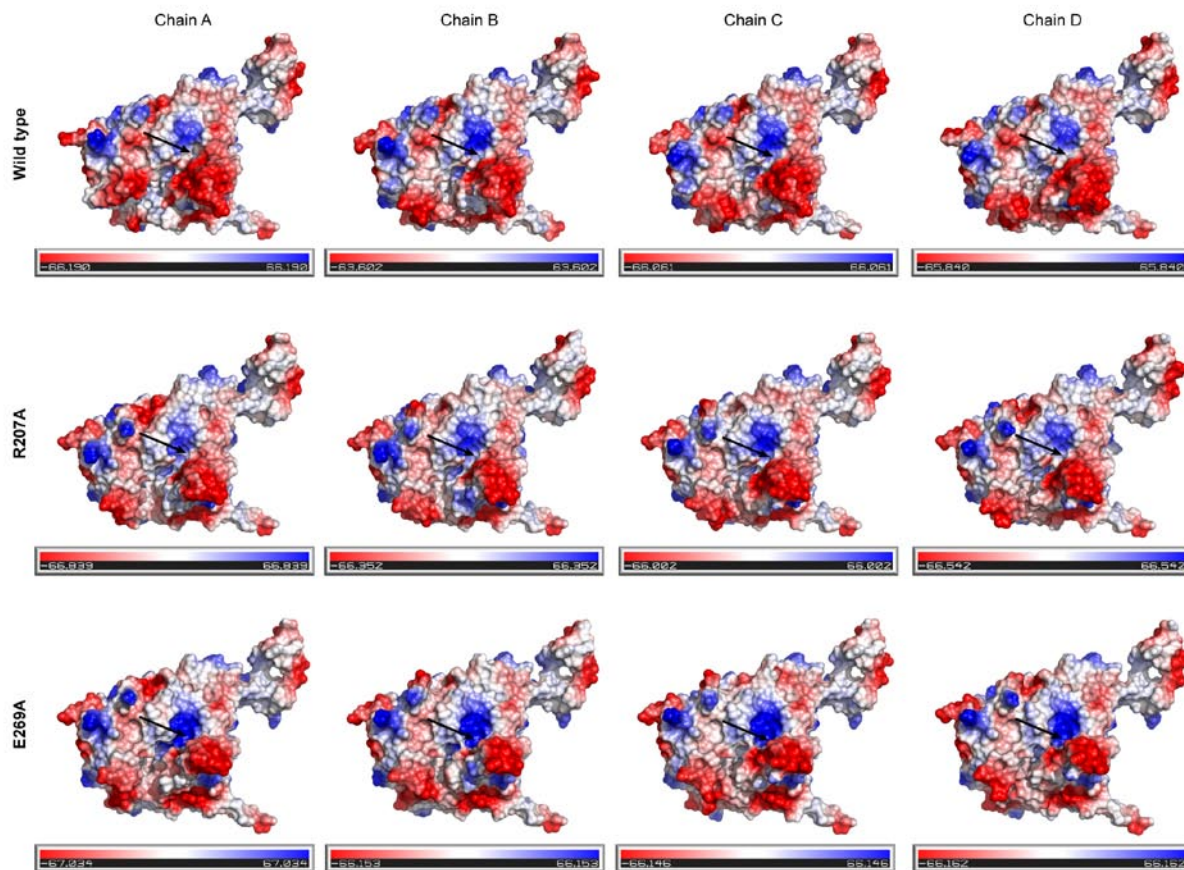


Figure 2.26: Comparison of electrostatic potential surface view of active site of EfBSH wild type, R207A and E269A mutants. All four chains are labeled accordingly. Difference in the electrostatics is marked by arrow pointing towards active site.

Electrostatic potential surface of wt *EfBSH* showed gradient of electropositive to electronegative and electroneutral in between near the active site involving loop1 and loop4. Electrostatic potential surface of R207A mutant was found to be similar to wild type *EfBSH*. Surprisingly, electrostatic potential surface of E269A mutant showed enhanced electropositive nature of active site and no demarcation of electroneutral in between electropositive and electronegative surface. This confirms the role of Glu269 in maintaining the dynamics in electrostatic potential of active site (Figure 2.26).

2.4.5 Molecular Dynamics simulation:

To establish the importance of Arg207 in allostery of WT *EfBSH* was performed for 100 ns and the trajectory was analysed using RMSD (Figure 2.27) and Rg - radius of gyration (Figure 2.28). The *EfBSH* tetramer was found to be stable during the entire simulation timescale. Radius of gyration showed no unfolding of protein in during production run.

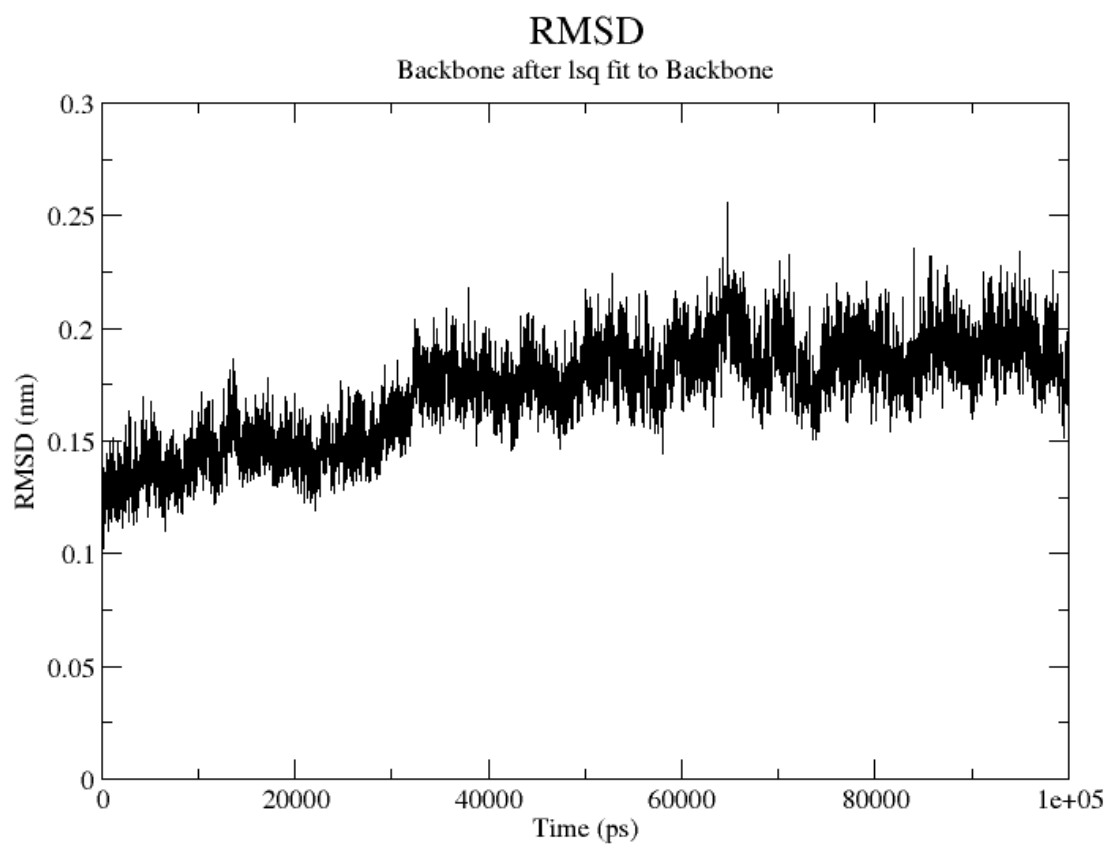


Figure 2.27: RMSD plot of MD simulation of *EfBSH* wild type performed for 100 ns.

Radius of gyration (total and around axes)

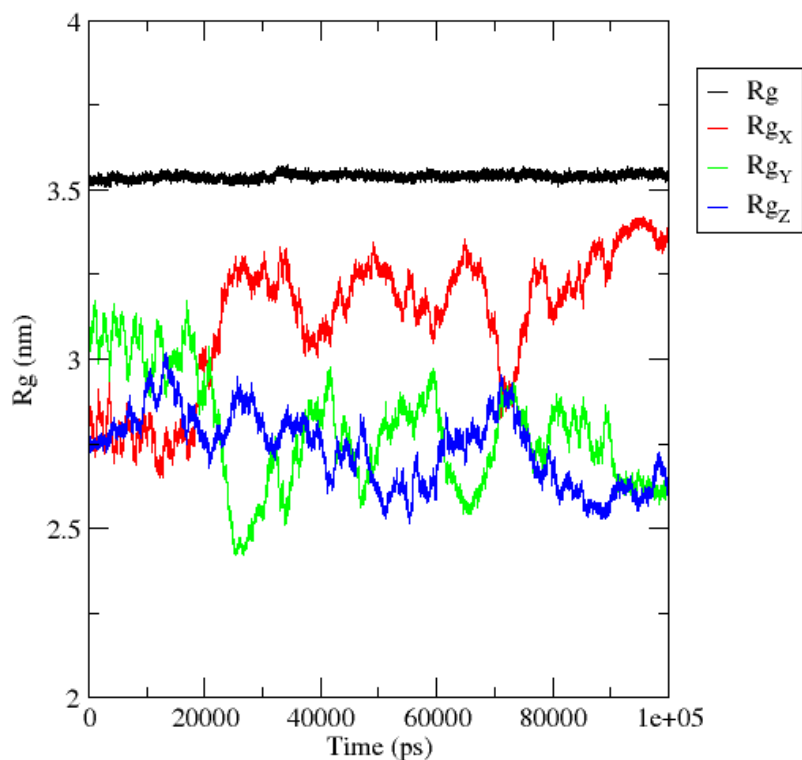


Figure 2.28: Radius of gyration during MD simulation of *Ef*BSH showing no change in the quaternary arrangement and packing of molecule.

2.4.6: Density Function Theory (DFT) calculations: In order to examine the electrostatic basis of reduced activity of the E269A mutant type with respect to the wt *Ef*BSH, electrostatic force analysis has been performed for the two cases. The effect of the non-cysteine amino acids of the active sites of all the four chains on the S-H bond stretching has been evaluated. The details of the procedure have been provided in the Computational Details section. The results of the force analysis have been summarized in the **Table 2.6** below.

Table 2.6:The influence of the non-cystine residues of the active site on the S-H bond cleavage in the pN for the geometries obtained after the constrained optimization at the M06-2X(water)/6-31G** level of theory.

Chain type	wt <i>Ef</i> BSH		S-H bond length (Å)	E269A mutant		S-H bond length (Å)	Difference in forces between wt <i>Ef</i> BSH and E269A mutant	
	Net binding force exerted by the non-cystine residues of the active site			Net binding force exerted by the non-cystine residues of the active site				
	Mulliken	NBO	Mulliken	NBO	Mulliken	NBO		
A	-2.0	-2.2	1.3403	1.0	1.0	1.3396	-3.0	-3.2
B	-2.7	-3.0	1.3403	1.3	1.2	1.3402	-4.0	-4.2
C	-2.0	-2.2	1.3402	1.0	1.0	1.3403	-3.0	-3.2
D	-1.9	-2.1	1.3402	1.4	1.4	1.3399	-3.3	-3.5

Mulliken = Forces obtained employing Mulliken charges

NBO= forces obtained employing NBO charges

The negative value of the force indicates stabilizing effect of the bond and positive force indicates destabilizing (cleaving) effect of S-H bond.

The interaction between side chain of Tyr20 and acetate group of Glu21 has been calculated using DFT in E269A and R207A. The interaction energy obtained for the two cases is provided in the Table 2.8. The interaction energy between acetate moiety of XX and the side chain of Tyr is found to be smaller in E269A in comparison to that in the R207A by almost 1.5 kcal/mol, which suggests that the acetate in the R207A is more strongly bound with the Tyr, and hence holds the structure tighter in this conformation. This result has been found to be consistent with different levels of theory involved for the calculation. The interaction energy has been calculated employing following formula.

$$E_i = E_{\text{comp}} - (E_{\text{frag1}} + E_{\text{frag2}}) \quad (2.3)$$

Where, E_{comp} is the energy of noncovalently bonded complex, and E_{frag1} and E_{frag2} are single point energies of two fragments being separated from an optimized complex. The optimized

complexes for two cases were obtained after constrained optimization. The starting point of each constrained geometry optimization were taken by chopping these structures out from the crystal structure in such as way that the only the noncovalent interactions between acetate and side chain of tyrosine persists. The constrained optimization was done to preserve the experimental geometry keeping position of all heavy atoms fixed while optimizing only the externally added hydrogen atoms. The optimized geometries at the ω B97XD (water)/6-311++g** level of theory have been shown in Figure 2.8.

Table 2.7: Hydrogen bond distance (in Å) between Arg207 and Glu21 of WT EfBSH and E269A mutant.

Chan ID	WT <i>EfBSH</i>		EfBSH E269A	
	H1	H2	H1	H2
A	2.241	2.753	1.941	2.619
B	2.118	2.770	2.018	2.58
C	1.949	1.649	1.729	2.419
D	2.002	2.927	2.206	2.442

*Distance cut off = 2.75 Å (sum of Van Der Waals radii of H and O).

Table 2.8:The intraction energy between acetate moiety of Glu21 and side chain of Tyr20 in E269A and R207A.

Level of theory	L ₁	L ₂	L ₃	L ₄
E269A	-0.3	-1.5	0.5	-0.3
R207A	-1.4	-2.7	-1.1	-1.8

L₁ = M06-2X(water)/6-311++g**

L₂ = M06-2X(water)/6-31g**

L₃ = B3LYP-D2(water)/6-311++g**

L₄ = ω B97XD (water)/6-311++g**

2.5 Discussion:

Recently, it was reported that *EfBSH* is a highly catalytic enzyme as compared to other known BSHs[15]. Since, the BSH enzyme exhibited multifaceted role in the gut environment [2], the structure of this highly active molecule was deciphered to study the molecular deviation from previously reported BSHs [16]. Detailed analysis of active site leads to

speculation that presence of Tyr20 in loop1 and Tyr65 in loop2 help align the substrates, being the reason of higher activity and the distorted loop2 transmit the allostery through the tetramer loop[16]. To begin with, we started our analysis with Cys2 (since Met1 is autocatalytically processed and removed) to map the hydrogen bonding network in the active site pocket in all chains of *Ef*BSH. The hydrogen bonding pattern near active site is uniform in all chains for Cys2, Gln256, and Asp19 but varies in Glu269 and Glu21 in different chains (Figure 2.15). These variation above the active site resulted from the involvement of Arg207 from other chain of dimer which protrudes into the active site to interact with Glu21 and Glu269 of loop1 and loop4 respectively (Figure 2.1 & 2.3). Mutating the Arg207 residue to alanine (R207A) releases the Glu269 and Glu21. While mutating Glu269 to alanine (E269A), Arg207 showed uniformity in hydrogen binding interaction with Glu21 in all chains of tetramer (Figure 2.16). Moreover, it is evident from the structural analysis that the Arg207 is asymmetrically placed in between Glu21 and Glu269 (Fig. S5). Hence, arginine's amino group, perhaps, experiences a similar electrostatic interaction but in opposite direction which results in the formation of weak but transient hydrogen bond. In case of E269A mutant, Arg207 forms triplet of hydrogen bond and forms stronger hydrogen bond network with Glu21[48]. This result is further validated using computational approach such as DFT studies. In E269A mutant, the distances between the residue Arg207 and Glu21 are reduced in all chains as compared to wild type (Table 2.7) which perhaps strengthens the hydrogen bond formed between Arg207 and Glu21. In case of E269A mutant distance of both the hydrogen bond is less than 2.75 Å in all chains. According to IUPAC, distance of hydrogen bond should be smaller than the sum of Van Der Waals radii of donor and acceptor atoms, which is around 2.75 Å. In WT *Ef*BSH, distance between the Arg207 and Glu21 is comparatively higher, which perhaps explains the weaker nature of hydrogen bond in the active site as compared to E269A mutant. This leads to reduction in active site dynamics of E269A mutant as evident in RMSD analysis of mutants and wt *Ef*BSH (Figure 2.22). Further, recent reports on *Ef*BSH highlighted the role of two tyrosine residues *viz.* Tyr20 and Tyr65 of loop1 and loop2 respectively [16], for higher biochemical activity. Analysis of Tyr20 in wt *Ef*BSH showed scattered arrangement of tyrosine side chain in different monomers (Figure 2.19). While similar analysis in Tyr20 of all four superposed chains showed congregation of tyrosine in identical orientation in E269A and R207A mutant respectively (Figure 2.19). Moreover, superposition of four chains of E269A with four chains of R207A showed two extreme conformations possible for Tyr20 and Tyr20 of wt *Ef*BSH drift between these two extremes (Figure 2.20). These conformations of Tyr20 resulted from the mutation of Arg207

and E269A which showed variable interaction with Glu21 of loop1. In order to decipher the effect of these mutants on Tyr20 which is mediated via Glu21, the interaction energy between side chain of tyrosine and acetate group of Glu21 has been calculated using DFT in E269A and R207A (Table 2.8). The interaction energy between acetate moiety of Glu21 and the side chain of Tyr20 is found to be smaller in E269A in comparison to that in the R207A by almost 1.5 kcal/mol, which suggests that the acetate in the R207A is more strongly bound with the Tyr20, and hence holds the structure tighter in this conformation. Structural superposition of wt *Ef*BSH with previously characterized BSHs structures showed corresponding amino acids of Arg207, Glu21, Glu269 and Tyr20 of *Ef*BSH (Fig. S9& S10). Tyr20 of *Ef*BSH is replaced by Trp21 in *Bf*BSH, Ile22 in *Cp*BSH and Leu20 in *Ls*BSH (Figure 2.18). Surface view of Tyr20 and its corresponding amino acids in other BSHs with conserved asparagine (crucial in anion hole formation) showed blockage of acceptor site in *Bf*BSH and *Ls*BSH (Figure 2.17). Alternatively, mutating the conserved Asn79 to tryptophan and tyrosine showed reduction in the biochemical activity to 17% and 88% respectively, as compared to wild type *Ef*BSH. This result indicates that increasing the bulkiness of side chains reduces the biochemical activity. As Asn79 is conserved in all BSHs (Figure 2.4), differences in biochemical activity can be attributed to Tyr20 in *Ef*BSH and corresponding amino acids of other BSHs. This, in part, explains the difference in biochemical nature of different Bile Salt Hydrolases (BSHs) depends on corresponding residues present in position of Tyr20 of loop1 in *Ef*BSH.

Insights from biochemical studies showed that the R207A mutant retains the activity towards GCA and GDCA and while showed enhanced activity towards TCA and TDCA (Figure 2.11). The E269A mutant showed reduced biochemical activity towards both Glyco- and Tauro- conjugated bile acids. Docking of different conformations of GCA, generated after ligprep, in active site of *Ef*BSH showed different binding mode (Figure 2.25) which restricted the entry of ligand from acceptor site[16]. The E269D mutant is inactive while E269Q is found to be active towards both glyco- and tauro- conjugated bile acids. Mutating Glu269 to neutral amino acids i.e. alanine and glutamine retains the feableactivity while reducing the length of side chain by mutating it to aspartic acid completely abolishes the biochemical activity towards both glyco- and tauro- conjugated bile acids (Figure 2.11). This confirms the electrostatic role of Glu269 of loop4 in the active site. The qualitative electrostatic potential surface of E269A mutant showed enhanced electropositive nature of active site. This indicates the role of Glu269 in maintaining the electrostatic texture of active site essential for the catalysis (Figure 2.26). It is to be noted that the different electrostatic potential

(viz. electrostatic field) have been reported to be crucial for the catalysis of different steps of the reaction in the catalytic process of an enzyme [49].

In order to examine the electrostatic basis of reduced activity of the E269A mutant with respect to the wt *Ef*B₅H. Electrostatic force analysis has been performed for the two cases which (Table 2.6) suggest that the non-cysteine residues in the wt *Ef*B₅H stabilizes the S-H bond of cysteine, which is contrary to the experimental finding. Since the S-H bond cleavage is one of the key steps in the catalytic cycle of the B₅H, any factor which stabilizes S-H bond is expected to increase the reaction barrier. However, S-H bond cleavage is coupled with the N-O-H bond formation in the transition state structure [46], which may or may not experience same effect. A complete electrostatic picture evaluating the long range electrostatic influence of non-cystine residues on the turn over frequency of the enzymatic activity of two proteins (wt *Ef*B₅H and E269A) is out of the scope of the current study. Moreover, electrostatic influence of the non-cystine residues of the active sites of each chain on the S-H bond cleavage is marginal (the same is expected for the N-H bond formation as well due to the large separation between the cystine and the Glu/Ala) as the difference in the electrostatic binding forces of S-H bond of each chain for two cases is very small. This result is further corroborated from the comparison of S-H covalent bond length, which is found to be almost same in the two cases (see column 4 and 7 of the Table 2.6). A slightly smaller S-H bond length for E269A suggests that the S-H bond is slightly activated in E269A, which is in accordance with our force analysis results as a very small positive values of the forces (which pushes hydrogen atom away to the rest part of the cystine) have been obtained for this case (columns 5 and 6, Table 2.6). Since previous reports suggest that the S-H bond cleavage is the rate determining step of the B₅H activity, (reference) the change in electrostatic due to mutation of Glu269 with Ala is suggested to have no influence on the activity of the E269 mutant type. Thereby, provides the conformational basis of reduced enzymatic activity in E269A mutant. The kinetic calculation showed increased affinity of GCA and reduced V_{max} towards E269A mutant and in R207A mutant showed MM kinetics (Figure 2.12).

Hence, it is evident from the above study that Arg207 interacts with Glu21, which directs the movement of Tyr20. In E269A mutant, interaction energy calculation suggest that Glu21 do not interact with Tyr20 and biochemical data showed very less activity as compared to wt *Ef*B₅H. Moreover, docking of GCA in active site of *Ef*B₅H restricted the entry of substrates from SiteA encompassing the loop1 to loop3. Therefore this conformation of Tyr20 can be designated as closed conformation. While in R207A mutant, the Glu21 showed

higher interaction energy with Tyr20 and showed increased activity towards Tauro-conjugated bile acids. Therefore, this conformation of Tyr20 can be designated as open conformation (Figure 4). This results shows that without Arg207 active site is open in all chains and involvement of Arg207 regulates the opening and closing of active site via Glu269 and Glu21. This result is further complemented by biochemical data which showed higher percent residual activity of R207A mutant with tauro- conjugated bile acids (Figure and Table). This is supported by the fact that previously reported BSH structures have different residues at corresponding location (figure suppl) which fails to form either of these combinations [12,13]. In conclusion, the conserved residues Asp19 and residues of loop1 i.e. Tyr20 and Glu21 forms a motif (DYE motif) which is responsible for higher activity and allostery in *Ef*BSH, provided Arg207 is present in assembly loop or tetramer loop (Figure 2.4).

2.6 Summary and conclusions:

In this chapter we dissected the active site residues to ascertain the high catalytic activity of *Ef*BSH. We found that Arg207 of tetramer loop protrudes inside the active site of other chain and influences the active site residues. Mutation of the Arg207 interacting residues i.e. Glu269 to alanine showed drastic decrease in biochemical activity. Glu269 when mutated to Aspartic acid loses the activity while when mutated to Glutamine, found to be active against bile acids. Electrostatic potential of active site of E269A mutant more electropositive nature hence electrostatics of active site could be one of the parameter which influences the activity. We here, observed that mutation of Glu269 to neutral amino acids retains the activity while when replaced with Aspartic acid completely shuts the biochemical activity against bile acids. Structural analysis showed that in Glu269 to alanine i.e. E269A mutant, Arg207 is hydrogen bonded to Glu21 of loop1 and constrain the movement or flexibility of loop in all the chains. Due to this, movement of Tyr20 of loop1, which aligns the substrates in the active sites, is also blocked. In R207A mutant, the Glu21 sets free in all different chains and do not interact with E269 of loop4. Superposition of Tyr20 of all four chain in E269A and R207 mutant showed very little deviation as compared to wild type. Further, superposition of four chains of E269A mutant with four chains of R207A mutant revealed two extreme arrangement possible for Tyr20. Moreover, in wild type, Tyr20 samples the conformation between the two extreme conformations possible as in E269A and R207A. Subsequently, the corresponding amino acids of other reported BSH structure were analyzed and we found that in *Bf*BSH (PDB Id: 2hez) tyrosine is replaced by tryptophan which closes the active site and

perhaps reduces the accessibility of substrates and hence activity. Therefore, Tyr20 occupies very strategic location near the active site and we found that Arg207 regulates the movement of this tyrosine by interacting with Glu21 and Glu269 (Figure 2.31).

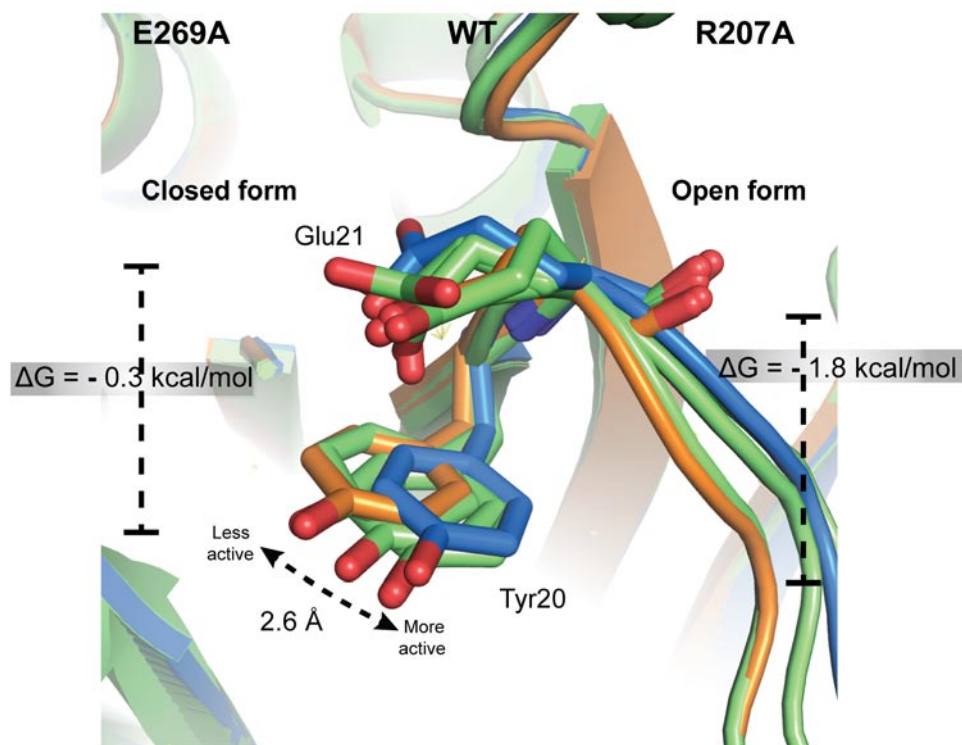


Figure 2.29: Proposed mechanism of allostery in *EfBSH*. Shown here is superposed cartoon image of E269A (orange), WT *EfBSH* (Green) and R207A (Blue). All four chains of WT *EfBSH* were used while ChainA of *EfBSH* and R207A were used for simplicity. Sticks of Glu21 and Tyr20 are shown for all chains.

Hence it is clear from above studies that Tyr20 acts as a gatekeeper residue and E269A mutant correspond to closed conformation and R207 correspond to open conformation. Also, kinetic calculation showed the reduction in hill coefficient which clearly indicates the significance of Arg207 in maintaining dynamic allostery near the active site.

References:

1. Duggleby, Helen J., et al. "Penicillin acylase has a single-amino-acid catalytic centre." *Nature* 373.6511 (1995): 264.
2. Chand, Deepak, et al. "Molecular features of bile salt hydrolases and relevance in human health." *Biochimica et Biophysica Acta (BBA)-General Subjects* 1861.1 (2017): 2981-2991.
3. Zhou, Xiaoming, et al. "Structural basis of the alternating-access mechanism in a bile acid transporter." *Nature* 505.7484 (2014): 569.
4. Hu, Nien-Jen, et al. "Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT." *Nature* 478.7369 (2011): 408.
5. Balakrishnan, Anand, and James E. Polli. "Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target." *Molecular pharmaceutics* 3.3 (2006): 223-230.
6. Hofmann, Alan F. "Enterohepatic circulation of bile acids." *Handbook of Physiology. The Gastrointestinal System* 4 (1989): 567-596.
7. Hofmann, Alan F. "The enterohepatic circulation of bile acids in man." *Clinics in gastroenterology* 6.1 (1977): 3-24.
8. Houten, Sander, and Johan Auwerx. "The enterohepatic nuclear receptors are major regulators of the enterohepatic circulation of bile salts." *Annals of medicine* 36.7 (2004): 482-491.
9. Norman, Arne, and Jan Sjövall. "On the Transformation and Enterohepatic Circulation of Cholic Acid in the Rat BILE ACIDS AND STEROIDS 68." *Journal of Biological Chemistry* 233.4 (1958): 872-885.
10. Begley, Máire, Colin Hill, and Cormac GM Gahan. "Bile salt hydrolase activity in probiotics." *Appl. Environ. Microbiol.* 72.3 (2006): 1729-1738.

11. Jones, Brian V., et al. "Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome." *Proceedings of the national academy of sciences* 105.36 (2008): 13580-13585.
12. Kumar, R. Suresh, et al. "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* 281.43 (2006): 32516-32525.
13. Rossocha, Maksim, et al. "Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product." *Biochemistry* 44.15 (2005): 5739-5748.
14. Panigrahi, Priyabrata, et al. "An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholylglycine hydrolase family." *Microbiology* 160.6 (2014): 1162-1174.
15. Chand, Deepak, Sureshkumar Ramasamy, and C. G. Suresh. "A highly active bile salt hydrolase from *Enterococcus faecalis* shows positive cooperative kinetics." *Process Biochemistry* 51.2 (2016): 263-269.
16. Chand, Deepak, et al. "Structure and function of a highly active Bile Salt Hydrolase (BSH) from *Enterococcus faecalis* and post-translational processing of BSH enzymes." *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1866.4 (2018): 507-518.
17. Ashkenazy, Haim, et al. "ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules." *Nucleic acids research* 44.W1 (2016): W344-W350.
18. Tanaka, H., et al. "Screening of lactic acid bacteria for bile salt hydrolase activity." *Journal of dairy science* 82.12 (1999): 2530-2535.

19. Salomon-Ferrer, Romelia, David A. Case, and Ross C. Walker. "An overview of the Amber biomolecular simulation package." *Wiley Interdisciplinary Reviews: Computational Molecular Science* 3.2 (2013): 198-210.
20. Pronk, Sander, et al. "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit." *Bioinformatics* 29.7 (2013): 845-854.
21. Gaisford, Wendy, Gebhard Schertler, and Pat Edwards. "mosquito® LCP: Making membrane protein crystallization accessible to the research scientist." *Nature Methods* 8.6 (2011).
22. Carter, Charles W. "Protein crystallization using incomplete factorial experiments." *Journal of Biological Chemistry* 254.23 (1979): 12219-12223.
23. Chayen, Naomi E. "Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques." *Acta Crystallographica Section D: Biological Crystallography* 54.1 (1998): 8-15.
24. Cudney, R., et al. "Screening and optimization strategies for macromolecular crystal growth." *Acta Crystallographica Section D: Biological Crystallography* 50.4 (1994): 414-423.
25. Jancarik, J. A. K. S., and S-H. Kim. "Sparse matrix sampling: a screening method for crystallization of proteins." *Journal of applied crystallography* 24.4 (1991): 409-411.
26. Stura, Enrico A., and Ian A. Wilson. "Analytical and production seeding techniques." *Methods* 1.1 (1990): 38-49.
27. Kumar, Ashwani, et al. "Protein crystallography beamline (PX-BL21) at Indus-2 synchrotron." *Journal of synchrotron radiation* 23.2 (2016): 629-634.
28. Battye, T. Geoff G., et al. "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM." *Acta Crystallographica Section D: Biological Crystallography* 67.4 (2011): 271-281.

-
29. Evans, Philip. "Scaling and assessment of data quality." *Acta Crystallographica Section D: Biological Crystallography* 62.1 (2006): 72-82.
 30. Blundell, Tom L., and Louise N. Johnson. "Protein crystallography." (1976).
 31. Diederichs, Kay, and P. Andrew Karplus. "Improved R-factors for diffraction data analysis in macromolecular crystallography." *Nature structural biology* 4.4 (1997): 269.
 32. Weiss, Manfred S. "Global indicators of X-ray data quality." *Journal of applied crystallography* 34.2 (2001): 130-135.
 33. Weiss, M. S., and R. Hilgenfeld. "On the use of the merging R factor as a quality indicator for X-ray data." *Journal of applied crystallography* 30.2 (1997): 203-205.
 34. Matthews, Brian W. "Solvent content of protein crystals." *Journal of molecular biology* 33.2 (1968): 491-497.
 35. Crick, Francis HC, and John C. Kendrew. "X-ray analysis and protein structure." *Advances in Protein Chemistry*. Vol. 12. Academic Press, 1957. 133-214.
 36. Murshudov, Garib N., et al. "REFMAC5 for the refinement of macromolecular crystal structures." *Acta Crystallographica Section D: Biological Crystallography* 67.4 (2011): 355-367.
 37. Zhao, Yan, and Donald G. Truhlar. "A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions." *The Journal of chemical physics* 125.19 (2006): 194101.
 38. Zhao, Yan, and Donald G. Truhlar. "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals." *Theoretical Chemistry Accounts* 120.1-3 (2008): 215-241.

-
39. Frisch, M. J., et al. "Gaussian, Inc., Wallingford CT." *Gaussian 03, Revision E 1* (2009).
 40. Barone, Vincenzo, and Maurizio Cossi. "Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model." *The Journal of Physical Chemistry A* 102.11 (1998): 1995-2001.
 41. Cossi, Maurizio, et al. "Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model." *Journal of computational chemistry* 24.6 (2003): 669-681.
 42. Reed, Alan E., Robert B. Weinstock, and Frank Weinhold. "Natural population analysis." *The Journal of Chemical Physics* 83.2 (1985): 735-746.
 43. Mulliken, Robert S. "Electronic population analysis on LCAO–MO molecular wave functions. I." *The Journal of Chemical Physics* 23.10 (1955): 1833-1840.
 44. Tiwari, Mrityunjay K., and Kumar Vanka. "Exploiting directional long range secondary forces for regulating electrostatics-dominated noncovalent interactions." *Chemical science* 8.2 (2017): 1378-1390.
 45. Lodola, Alessio, et al. "A catalytic mechanism for cysteine N-terminal nucleophile hydrolases, as revealed by free energy simulations." *PLoS One* 7.2 (2012): e32397.
 46. Oinonen, Carita, and Juha Rouvinen. "Structural comparison of Ntn-hydrolases." *Protein Science* 9.12 (2000): 2329-2337.
 47. Duggleby, Helen J., et al. "Penicillin acylase has a single-amino-acid catalytic centre." *Nature* 373.6511 (1995): 264.
 48. Friesner, Richard A., et al. "Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein– ligand complexes." *Journal of medicinal chemistry* 49.21 (2006): 6177-6196.

49. Lai, Wenzhen, et al. "External electric field can control the catalytic cycle of cytochrome P450cam: a QM/MM Study." *The Journal of Physical Chemistry Letters* 1.14 (2010): 2082-2087.

CHAPTER III

Oligomeric Attributes of Cholyglycine Hydrolase (CGH) Family in Enzymatic Activity and Stability

3.1 Introduction

Oligomerization is a key process in most of the cellular enzymes and significant proportion of cellular enzymes are oligomeric in nature. Moreover, the average oligomeric state of protein in cellular pool is also tetrameric[1]. Koshland and Bowden, in his classic paper discussed the thermodynamically possible arrangement viz. isologous and heterologous, of oligomers in respect to the stability of enzymes[2]. One of the early mechanisms proposed for oligomerization was domain swapping[3]. Appeared later another mechanism for oligomerization is “activating oligomerization” which includes most of the glycolytic enzymes, in which the enzyme switches from monomeric to oligomeric state and vice-versa. Yet another mechanism proposed by hashimoto is the indel of amino acids which enable the formation of oligomerization. These indels are frequently located on the interface of the oligomeric protein[4]. Indeed, there are numerous benefits for the protein to form oligomeric structure. It attains the bigger size without increasing the genome size and also increases the coding efficiency[5,6]. Moreover, the oligomeric proteins are more stable as compared to their monomeric counterpart[7,8]. Oligomerization also leads to allostery which improves the catalytic efficiency of enzymes[9]. In most of the cases, dedicated domain known as “oligomerization domain” is present which aids in the oligomerization of proteins. In p53, a tumor suppressor protein, removal of oligomerization domain leads to functional monomer[10]. The fragmented polypeptide of murine dihydrofolate reductase reassemble into functional protein owing to oligomerization domain[11].

Here we have considered the oligomeric attribute of Bile Salt Hydrolases (BSHs) of Cholyglycine Hydrolase (CGH) family (member of Ntn hydrolase superfamily). N-terminal nucleophile (Ntn) hydrolase superfamily, which is one of the most diverse superfamily of hydrolytic enzymes. This superfamily includes many families of enzymes having physiological, clinical and pharmaceutical importance which harbours different homo and hetero oligomeric quaternary state. The cholyglycine hydrolase (CGH) family has been reported to consist of homo-oligomeric BSH and PVA. BSH hydrolyzes the amide bond involved in the deconjugation of bile acids in human gut[12,13]. This reaction is very critical as it is involved in the regulation of cholesterol metabolism. Phylogenetic analysis revealed that evolution of BSH activity depends upon microbial habitat[12]. BSH from gram positive and gram negative organism form different cluster owing to difference in indel of 20 amino acids in gram positive bacteria[14]. These indels are known to form tetramerization loop which protrudes out from the core structure and helps in

tetramer formation[15,16]. However, dimeric BSH is reported from *Lactobacillus* even in the presence of tetramerization loop[17]. Theoretical prediction showed that absence of 20 amino acid indelin gram negative BSH reduces the interface area and stability of enzymes[14]. Hashimoto et al reported similar prediction after theoretical calculation of enabling and disabling loop in dimmers [4]. Recently, penicillin V-acylase was reported from Gram negative organism *Pectobacterium atrosepticum* (*PaPVA*) have been reported to be tetramer and thermostable without having tetramer loop[18,19]. Structural study of *PaPVA* showed non planar arrangement of dimer of dimer perhaps to compensate for the loss of tetramer loop[18] by increasing the surface area around tetramer interface. Moreover, BSH from *Lactobacillus salivarius* (*LsBSH*; PDB Id-5hke) is reported to be homodimer despite having tetramer loop and recently added member, Acyl Homoserine Lactone (AHL) acylase, from *Shewanella loihica* (*SICGH1*; PDB ID: 5x8z) showed homodimer (unpublished data). To understand the unifying principle of oligomerization in CGH family we analyzed the interface of previously reported structure and also performed structural and biochemical characterization of two recently added microorganism i.e. *Dorea longicatena* BSH (*DlBSH*) and *Methanobrevibacter smithii* BSH (*MsBSH*) from metagenomic analysis of human gut microbiome[12]. Oligomerization study was performed in greater detail with *DlBSH*, in which several mutations has been created using site directed mutagenesis approach. Furthermore, we identified a residue at tetramer interface, which perhaps regulates the BSH oligomerization. This result suggests that mere presence of tetramer loop cannot lead to formation of tetramer. Interface comparison with gram negative structure i.e. *PaPVA* revealed a unique minicluster, which holds and stabilizes the tetramer in gram negatives. In Gram positives, these miniclusters were supported by presence of aromatic residues such as Tyr205 in *DlBSH* and *EfBSH*, and in Gram negatives, dimer groove which has smaller tetramer loop makes very strong hydrophobic interaction with other dimer and stabilize the tetramer.

Also, we identified the crucial residues, which is imperative to the oligomerization of CGH family of proteins and has implications on screening the highly active BSH/PVA using primary structure of protein from genomic databases and also predicts the oligomeric state of the protein. We propose, in general, there must be two fundamentally different pathway of oligomerization depending upon the intrinsic symmetry of the molecule. In cyclic symmetry, oligomeric proteins exist in two different forms i.e. active (monomer) or inactive (oligomer) or

vice-versa. While in proteins having dihedral symmetry, can exist in multiple oligomeric states, which perhaps, give plasticity to the enzyme evolution and ability to metabolize or overcome the toxic effect of metabolites, which leads to survival of the microorganisms.

3.2 Materials and Methods

3.2.1 Gene synthesis

Genes of *Dorea longicatena* (uncultured bacteria) and *Methanobrevibacter smithii*(archeal) bile salt hydrolase (*D*IBSH and *M*sBSH) was characterized in Jones et. al. was optimized for *E.coli* classII and synthesised in lab by designing overlapping primers using DNAWorks software[12,20]. This software takes protein sequence as input parameters and execute through following steps:

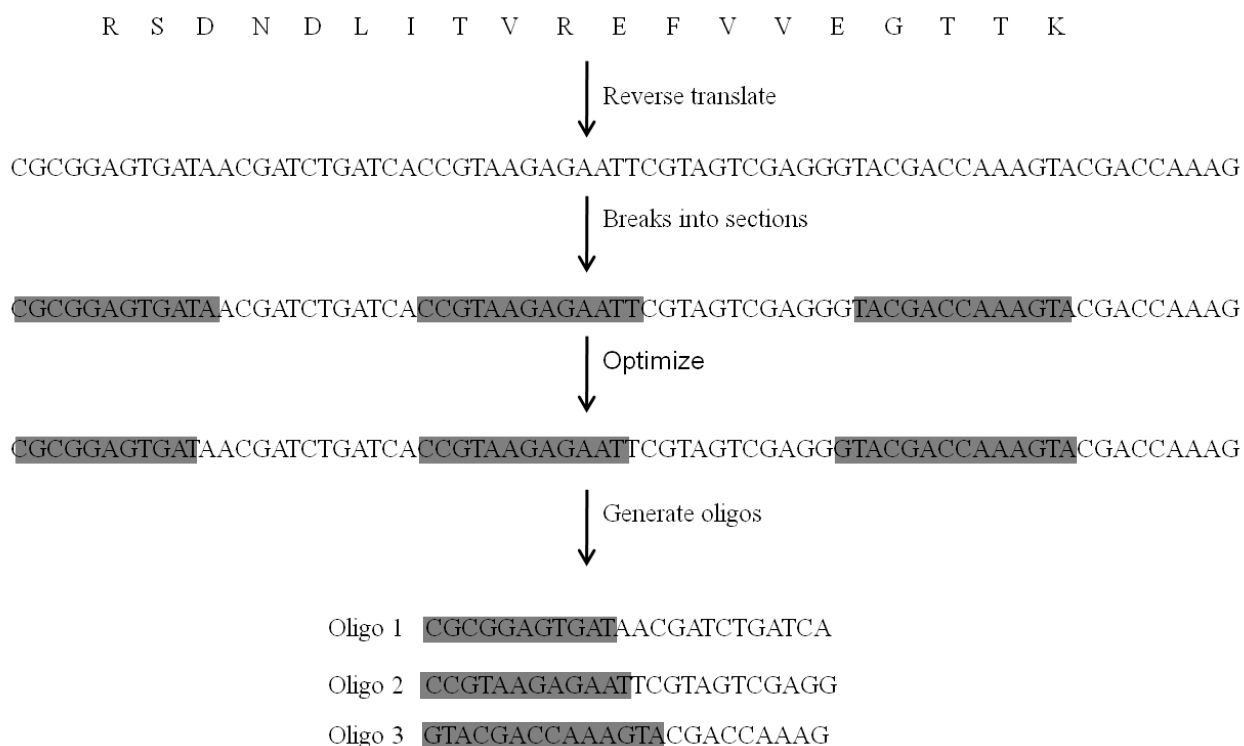


Figure3.1:Steps involved in optimized primer designing using DNAWorks server

*D*IBSH and *M*sBSH were broken into section and oligos were generated as explained above (Figure 3.1). The detailed information about the oligos is present in table 1 and 2 respectively.

Table 3.1: Overlapping primers of *D/BSH*

Sr. No.	Primer Sequence (5'-3')	Length (bp)
1	ATGTGCACCGCAGCGACCTACAAAACCAAAGACTTCTACTTTGGTC GTACCCTCGAC	57
2	CGCGGAGTGATAACGATCTGATCACCGTAAGAGAATTCGTAGTCGAG GGTACGACCAAAG	60
3	CAGATCGTTATCACTCCGCGTAACTATGCGTTTAACTTCCGTCACGTC GGTGACATGAAA	60
4	CTTCGGCAACGTGGGCCATACCGATGATGGCGTAGTGGTTTTTCATG TCACCGACGTGAC	60
5	GCCCACGTTGCCGAAGACTACCCGCTGTACTACGACGCGATGAACG AGAAGGGTGTTGCG	60
6	CGCCGCATAAACCGCATTACCAACGAAGTTCAGACCCGCCATCGCA ACACCCTTCTCGTT	60
7	ATGCGGTTTATGCGGCGATCAAACCGGACGTTGAAAACATCGCGCA GTTCGAATTCATCC	60
8	ACGAACTTCAACCAGGGAAGAGCACTGAGACAGGATCCACGGGAT GAATTCGAACTGCGC	60
9	CTTCCCTGGTTGAAGTTCGTGAACTGCTGGAGCGTATCAATATCGTC AATACCCCGTTTT	60
10	ATGATCCAATGCAGCTGAGCCAGCGGCAGCTGTTTCAGAAAACGGGG TATTGACGATATTG	60
11	GCTCAGCTGCATTGGATCATCTCCGACGAAAACGAATCTATTACCGT TGAATCCATGTCT	60
12	AACGCCTACCGGATTATCGTAGATGTGCAGACCGTCAGACATGGATT CAACGGTAATAGA	60
13	CGATAATCCGGTAGGCGTTCTGACCAACAATCCGCCGTTCCCGCAGC AGATGTTCCAAC	60
14	TGTTACGTGGCTGTTTCGGGGACAGATACATATAGTTGTTGAGTTGG	60

	AACATCTGCTGCG	
15	CCGAAACAGCCACGTAACACCTTCTGCGAAAACCTGGCGCTGGACG CTTATAGCCGTGGC	60
16	GGAGCTAGAGCTCAGGTCACCCGGCAGACCCAGACCACCCATGCCA CGGCTATAAGCGTC	60
17	TGACCTGAGCTCTAGCTCCCGCTTCGTTTCGTGTTGCCTTCACCAAGG TTAACGCGATCTC	60
18	GTGGAAGAAGTACGAGAAACAGATTCCTCTTCAGATTCACCAGAGATC GCGTTAACCTTGGT	60
19	AATCTGTTTCTCAGTTCTTCCACATCCTGGGTTCTGTTGACCAGCAG CGTGGTTGCTGCG	60
20	GCAAGAGGTGTACAGAGTGATTTTCGTATTTGCCGTCCGCAACCTCGC AGCAACCACGCTG	60
21	ATCACTCTGTACACCTCTTGCTGCAATGTTACCAAGGGTATTTACTAT TACAACACCTAT	60
22	GTGCATATCAACCGCAGAAATCTGATGGTTCTCATAGGTGTTGTAATA GTAAATACCCTT	60
23	ATTTCTGCGGTTGATATGCACGTGGAAAATCTGGACTCTGACAAAAT GATCTGCTACCCG	60
24	TTTGT TTTGGTAGTTGATGCGTTCACCCTGAATAACCGGGTAGCAGAT CATTTTGTC	57

Table 3.2: Overlapping primers of *MsBSH*

Sr.N o.	Primer sequence (5'-3')	Length (bp)
1	ATGTGCACCGCGGCAAACCTACCTGACGAAGTGCCACTACTTCGGCC GTAACCTT	53
2	GGTGATGGTAACACGTTTCGTTGTAAGAGATTTTCGTAGTCGAAGTTA CGGCCGAAGTAGTG	60
3	ACGAACGTGTTACCATCACCCCGCGTAACTACCCTCTGATCTTCCGT	60

	GACACCGAAGACA	
4	CGATGCCCGCAGCGATAACCGATGATAACCGTAGTGGTTTTTCGATGTCT TCGGTGTCACGGA	60
5	TCGCTGCGGGCATCGACGAATACCCGCTCTACTATGACGCGTGCAA CGAAAAAGGTCTGG	60
6	TTGTAGTCGCAATAGTCCGGGAAGTTCAGACCACCCATCGCCAGAC CTTTTTTCGTTGCAC	60
7	CCGGACTATTGCGACTACAAACCGCTGGACAAATCTAAAGTCAACA TCGCCTCTTTTGAA	60
8	AGAAATGGTTTTGGCCTGAGACAGGATGTACGGGATAATTTCAAAA GAGGCGATGTTGAC	60
9	TCTCAGGCCAAAACCATTTCTGACGCGGAACGTCTGCTGGAGAAC CTGAACATCTCCGAT	60
10	GTGGAGCGGAGACGGCGGCAGTTGCGCAGAAAACTTTTTCATCGGA GATGTTTCAGGTTCTC	60
11	GCCGTCTCCGCTCCACTGGATCATCTCTGACCGTAATGCGTCTATCG TTGTTGAAGTTGT	60
12	CACCAACCGGGTTGTCGTAGATGTCCAGGCCTTCTTCTACAACCTC AACAACGATAGACG	60
13	CGACAACCCGGTTGGTGTCTGACCAACAATCCGCCGTTTGACAAA CAGCTCTTCAATCT	60
14	TTCCGGGGTACGGTTAGACAGCGCCATGTAGTTATTCAGATTGAAG AGCTGTTTGTCAAA	60
15	GTCTAACCGTACCCCGGAAAACACCTTCGGCGGTAACCTGGACCTG GCGACCTACTCTCG	60
16	TGAGAAGAAACGTCACCCGGCAGACCGATAGAACCATAACCACGA GAGTAGGTCGCCAGG	60
17	CGGGTGACGTTTCTTCTCAGTCTCGTTTTTGTTAAAGCGGCTTTCGTT AAGGAAAATTCTG	60
18	AACTGGCTAACGGATTCTTTTTTCAGAGTCACCAGAAACAGAATTTT CCTTAACGAAAGCC	60

19	GAAAAAGAATCCGTTAGCCAGTTCTTCCACATCCTCGCGTCTGTTG AACAGCAGAAAGGC	60
20	TGGTGTATTCGAACTTGTCCGGTTCCTCAACCAGGGTGCAGCCTTT CTGCTGTTCAACAG	60
21	CGGACAAGTTCGAATACACCATCTACTCTGACTGCTACAACACCGA CAAAGGTATCCTGT	60
22	ATGTTAACGGAGGTCTGCGGACCGTCGTAGGTTTTGTAATACAGGA TACCTTTGTCGGTG	60
23	CGCAGACCTCCGTTAACATCCACGACGAAGACCTGGAAACTAACC AGCTGATCAACTTTG	60
24	GTCTACCAGCTCAAAGTTGATCAGCTGGTTAGTT	34

These oligos are mixed by aliquoting 5 μ l of each oligos into 500 μ l such that the final concentration of each primer is 1 μ M in mastermix. Whole gene synthesis took place in two step PCR protocol. In first PCR, mastermix of *Df*BSH and *Ef*BSH was mixed with the DNTPs (2.5 mM), Ex-Taq buffer(10x) and TaKaRa Ex Taq polymerase (5 units/ μ l) was diluted to 50 μ l (Table 3.3).

Table 3.3: PCR mixture for first PCR

PCR1 mix	<i>Df</i> BSH	<i>Ms</i> BSH
TaKaRa Ex Taq (5 units/ μl)	0.25 μ l	0.25 μ l
Ex-Taqbuffer(10x)	5 μ l	5 μ l
DNTPs (2.5 mM)	4 μ l	4 μ l
Primer mix	2.5 μ l	2.5 μ l
ddH₂O	38 μ l	38 μ l

Second PCR: This PCR was carried out using gene specific primer designed in Integrated DNA Technology web server consisting of NdeI/XhoI restriction sites on flanking regions by taking template from the product of first PCR.

3.2.2 Cloning, Heterologous expression and purification:

- Restriction Digestion:** *D/B*SH and *M_s*B*S*H were cloned in pET-22b(+) having NdeI/XhoI restriction sites. The pET-22b+ vector (Figure 3.2) and synthesized gene PCR product were digested with restriction endonucleases (NdeI/XhoI) and incubated at 37 °C for 4 hr. Upon restriction digestion, the reaction mixture was heated at 65 °C for denaturation of enzymes and followed by further purification with PCR clean up kit. The digested products were quantified by Nanodrop (Thermo-Scientific)

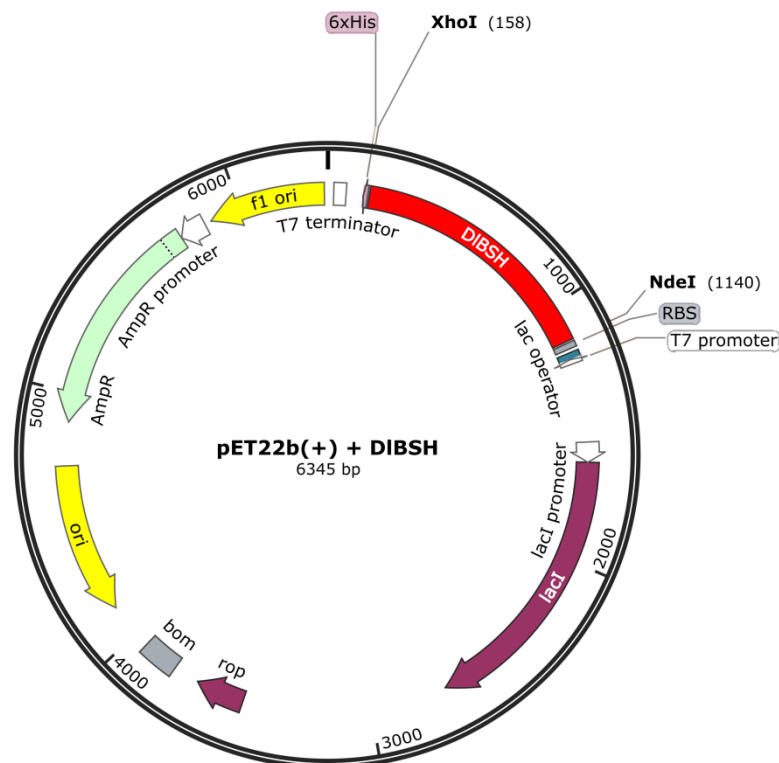


Figure 3.2: Construct of *D/B*SH-pET22b(+)

- Cloning:** 1:10 ratio mixture of vector to insert was incubated with T4 DNA Ligase (Invitrogen) with suitable buffer and two different temperatures were optimized for ligation strategy. i.e. at 16 °C overnight and at 45 °C for 5 min. Ligation mixture of plasmid pET22b+/bsh was then transformed into DH₅α cells and positive clones were selected on the basis of colony PCR. was transformed in the maintenance host DH₅α and plated in the LB agar containing 100 μg/ml Ampicillin and the plate was incubated at 37

°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.

- **Heterologous expression and purification:** Wild type *D/B*SH, *D/B*SH Y205D, *D/B*SH δ 189-209 and *M*sB*S*H were purified with affinity column using 6xHis affinity tag as explained in previous chapter. While *D/B*SH δ 189 (stop codon introduced after 189 aa) and *D/B*SH δ 207 (stop codon introduced after 207 aa) are purified using conventional method as it lacks C-terminal region along with 6xHis affinity tag. Upon cell homogenization, 40% (w/v) ammonium sulfate precipitation was carried out and kept overnight at 4 °C. The precipitated fraction further centrifuged at 10000 rpm and supernatant was discarded. The pellet was resuspended in 25mM Tris buffer pH 8 and 50 mMNaCl and passed through desalting column to remove excess of ammonium sulfate. Upon desalting, anion exchange chromatography (Q sepharose) was used to further purify the protein. In final step, the protein is passed through Sephacryl S-200 (GE Healthcare) gel filtration column and oligomeric profile was compared.

3.2.3 Site directed mutagenesis

Quick change of truncated clone of *D/B*SH was carried out using TaKaRa amplitaq followed by DpnI treatment for two hour and transformed into XL10 cells. Plasmid was isolated and transformed into *E.coli* BL-21 Gold (DE3). Transformed cells were checked for expression of proteins. Quick change of *D/B*SH δ 189 and *D/B*SH δ 207 were achieved by inserting stop codon after 189 aa and 207 aa, respectively which does not have 6xHis tag. List of primers used to create mutation is listed in table 3.4

Table 3.4: List of primers used for site directed mutagenesis in *D/B*SH. Nucleotide sequences code for stop codon is highlighted red.

Mutants	5'-3' sequence
<i>D/B</i> SH δ 189 F	CAGATGTTCCAACCTCTAACACCTTCTGCGA
<i>D/B</i> SH δ 189 R	TCGCAGAAGGTGTTAGAGTTGGAACATCTG
<i>D/B</i> SH δ 209F	GACGCTTATAGCCGTTAAGGCATGGGTGGTCTG
<i>D/B</i> SH δ 209R	CAGACCACCCATGCCTTAACGGCTATAAGCGTC
<i>D/B</i> SH δ 189-209F	TATATGTATCTGTCCGGTGGTCTGGGTCTG
<i>D/B</i> SH δ 189-209 R	CAGACCCAGACCACCGGACAGATACATATA

3.2.4: Homology modeling: Model of *Dorea longicatena* BSH and *Methanobrevibacter smithii* BSH are generated using *Ef*BSH as a template (PDB ID: 4WL3) in a Galaxy web server which generated the monomeric, dimeric and tetrameric form of BSH (<http://galaxy.seoklab.org/>). GalaxyGemini was used for protein structure prediction for homomeric protein based on template similarity and refined using GalaxyRefine.

Computational details of DFT analysis and biochemical experiments such as kinetic characterization is explained in previous chapter.

3.3 Results

3.3.1 Gene synthesis and cloning:

The protein sequence was retrieved from UniprotKb(<https://www.uniprot.org>) having uniprot ID A6BGZ5 of *Dorea longicatena* Bile Salt Hydrolase (*Dl*BSH) and B9AEX3 for *Methanobrevibacter smithii* Bile Salt Hydrolase (*Ms*BSH). The theoretical molecular weight and isoelectric point (pI) calculated using protparam tool is 36885.55 Da and 4.79, respectively for *Dl*BSH, and 35688.75 and 4.44 for *Ms*BSH (<https://web.expasy.org/cgi-bin/protparam>)

Gene ID: (*Dl*BSH)

```

ATGTGCACCGCAGCGACCTACAAAACCAAAGACTTCTACTTTGGTCGTACCCTCGAC
TAC GAATTCTCTTACGGTGATCAGATCGTTATCACTCCGCGTAACTATGCGTTTAAAC
TTCCGTCACGTCGGTGACATGAAAAACCACTACGCCATCATCGGTATGGCCACGTT
GCCGAAGACTACCCGCTGTACTACGACGCGATGAACGAGAAGGGTGTGCGATGGC
GGGTCTGAACTTCGTTGGTAATGCGGTTTATGCGGCGATCAAACCGGACGTTGAAAA
CATCGCGCAGTTCGAA TTCATCCCGTGGATCCTGTCTCAGTGCTCTTCCCTGGTTGA
AGTTCGTGAACTGCTGGAG CGTATCAATATCGTCAATACCCCGTTTTCTGAACAGCT
GCCGCTGGCTCAGCTGCATTGGATCATCTCCGACGAAAACGAATCTATTACCGTTGA
ATCCATGTCTGACGGTCTGCACATCTACGATAATCCGGTAGGCGTTCTGACCAACAA
TCCGCCGTTCCCGCAGCAGATGTTCCAACCTCAACAACCTATATGTATCTGTCCCCGAA
ACAGCCACGTAACACCTTCTGCGAAAACCTGGCGCTGGACGCTTATAGCCGTGGCAT
GGGTGGTCTGGGTCTGCCGGGTGACCTGAGCTCTAGCTCCCGCTTCGTTTCGTGTTGC
CTTCACCAAGGTTAACGCGATCTCTGGTGAATCTGAAGAGGAATCTGTTTCTCAGTT
CTTCCACATCCTGGGTCTGTTGACCAGCAGCGTGGTTGCTGCGAGGTTGCGGACGG
CAAATACGAAATCACTCTGTACACCTCTTGCTGCAATGTTACCAAGGGTATTTACTA

```


TTACAACACCTATGAGAACCATCAGATTTCTGCGGTTGATATGCACGTGGAAAATCT
GGACTCTGACAAAATGATCTGCTACCCGGTTATTCAGGGTGAACGCATCAACTACCA
AAACAAATAA

Uniprot ID: A6BGZ5

MCTAATYKTKDFYFGRTL DYEF SYGDQIVITPRNYAFNFRHVGDMKNHYAIIGMAHVA
EDYPLYDDAMNEKGVAMAGLNFVGNVYAAIKPDVENIAQFEFIPWILSQCSSLVEVRE
LLERINIVNTPFSEQLPLAQLHWIISDENESITVESMSDGLHIYDNPVGVLTNNPPFPQQMF
QLNNYMYLSPKQPRNTFCENLALDAYSRGMGGLGLPGDLSSSSRFVRVAFTKVNAISGE
SEESVSQFFHILGSVDQQRGCCEVADGKYEITLYTSCCNVTKGIYYNTYENHQISAVD
MHVENLSDKMICYPVIQGERINYQNK

Gene ID: BK798_RS07090 (*MsBSH*)

ATGTGTACTGCTGCAAATTATTTAACAAAATGCCATTATTTTGGCCGTAATTTTGACT
ATGAAATTTTCATATAATGAAAGAGTAACGATAACTCCTAGAACTATCCTTTAATAT
TCAGGGATACTGAGGACATTGAAAATCATTATGGGATTATTGGCATAGCTGCAGGTA
TTGATGAATATCCTTTGTATTATGATGCATGTAATGAGAAAGGATTAGCTATGGGGG
GATTAACCTTTCCGGATTACTGTGACTACAAACCACTAGATAAATCTAAAGTTAACA
TAGCTTCTTTTGAGATTATTCCATATATATTATCTCAAGCAAAAACCATCAGTGATGC
CGAAAAGTTATTGGAAAACCTAAATATTTTCAGATGAGAAATTTTCCGCCCAGTTGCC
TCCATCTCCACTTCATTGGATTATTTTCAGATAGGAATGCTTCAATTGTTGTAGAGGTT
GTAGAAGAAGGACTGGATATTTATGATAATCCTGTAGGAGTTTTAACAAACAACCCT
CCTTTTGATAAACAGCTATTTAATTTAAATAATTATATGGCATTATCAAACAGAACG
CCTGAAAATACATTTGGAGGCAATTTGGATTTGGCAACTTATAGTCGGGGGAATGGGT
TCAATTGGTCTTCCGGGGGATGTTTCTTCACAGTCCCGTTTTGTAAAAGCAGCTTTTG
TTAAAGAAAATTCGGTTTCCGGAGATTCTGAAAAAGAAAGTGTGTCTCAGTTTTTCC
ATATTCTGGCATCTGTTGAACAGCAAAAAGGATGTACGTTAGTGGAAGAACCTGAT
AAATTTGAATATACTATTTATTCGGACTGTTACAATACAGATAAGGGAATATTGTAT
TATAAACATATGATGGTCCTCAAACATCTGTTAATATACATGATGAGGATTTGGAA
ACCAATCAGTTAATTAATTTTGAGTTGGTTGATTAA

Uniprot ID: B9AEX3

MCTAANYLTKCHYFGRNFDYEISYNERVTITPRNYPLIFRDTEDIENHYGIIGIAAGIDEYP
 LYYDACNEKGLAMGGLNFPDYCDYKPLDKSKVNIASFEEIIPYILSQAKTISDAERLLENL
 NISDEKFSACLPPSPLHWIISDRNASIVVEVVEEGLDIYDNPVGVLTNNPPFDKQLFNLNN
 YMALSNRTPENTFGGNLDLATYSRGMGSIGLPGDVSSQSRFVKAAFVKENSVSGDSEKE
 SVSQFFHILASVEQQKGCTLVEEPDKFEYTIYSDCYNTDKGILYYKTYDGPQTSVNIHDE
 DLETNQLINFELVD

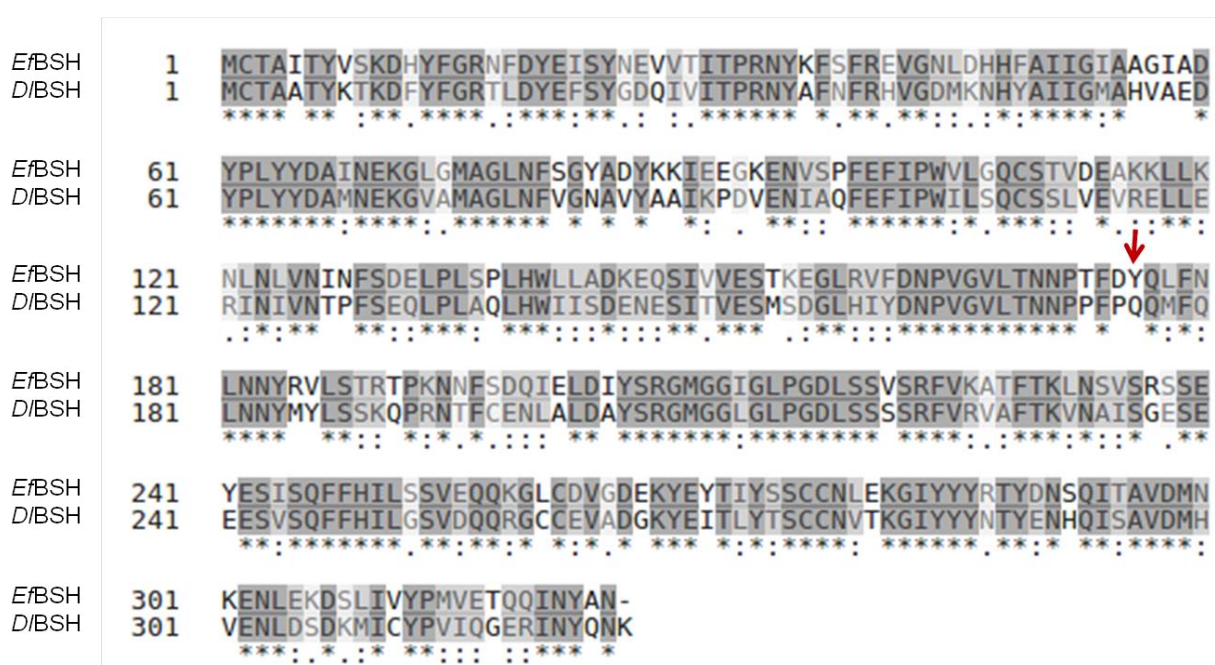


Figure 3.3: Sequence alignment of *Ef*BSH and *D*/BSH. Highlighted with red arrow correspond to tetramer interface which is involved in minicluster formation. Here, Tyr176 is replaced by Gln176 in case of *D*/BSH.

The reverse translated and optimized sequences were used for gene synthesis in two steps PCR. Upon second PCR, gene of *D*/BSH and *Ms*BSH correspond to 1Kb (Figure 3.4 a&b). Genes from both the species were further cloned in pET22b(+) vector (Figure 3.2) and transformed in *DH5 α* cells and plasmids were isolated from positive clones and sent for sequencing.

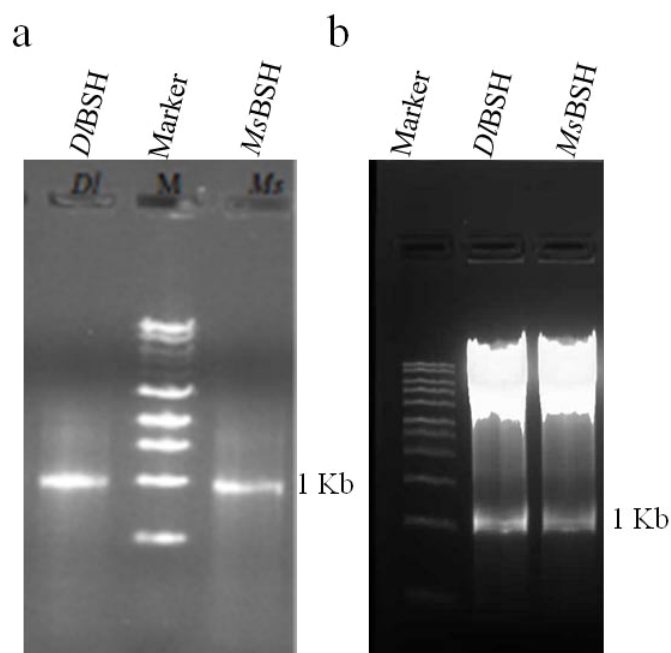


Figure 3.4: Gene synthesis and cloning a) shows the 1kb DNA after second step PCR of DNA synthesis. b) confirmation of clone using restriction digestion of cloned gene with NdeI/XhoI enzyme.

3.3.2 Protein purification

*D/B*SH protein samples were collected from lysis, supernatant and pellet and loaded on SDS PAGE for checking the expression of proteins at two different temperatures i.e. at 37 °C and at 16 °C and compared with uninduced expression profile. The SDS-PAGE gel profile shows inclusion body formation upon induction at 37 °C while soluble portion is increased upon induction at 16 °C (Figure 3.5). Furthermore, the size of *D/B*SH was not corresponding to 37 kDa, which is the expected size of monomer. To confirm the expression of *D/B*SH gene, we checked the hydrolase activity using GDCA as a substrate using whole cell with or without induction of 1mM IPTG (Figure 3.6). Since we performed the biochemical activity within bacterial cell, we incubated for two hours and checked the conversion of GDCA to DCA and compared with the standards TLC profile. No activity of *D/B*SH was detected even after two hours of incubation with uninduced cell while DCA was evident in IPTG induced cell (Figure 3.6). Purification of *D/B*SH using affinity chromatography showed protein in unbound and wash fraction and no protein band in elution fraction (Figure 3.7). Gene sequence analysis of the cloned gene when compared with the sequence of wt *D/B*SH showed insertion of cytosine base

which shifts the frame and code for stop codon (Figure 3.8). This confirms the absence of 6xHis tag and we tried to purify the heterologously expressed protein without affinity tag using conventional methods. Firstly, the protein was precipitated by 40% ammonium sulfate and desalted using desalting column (GE healthcare). Upon desalting, the protein was passed through Q-sepharose column and eluted with continuous gradient approach. The eluted fraction was loaded on SDS-PAGE gel and confirmed the presence of *D/B*SH protein after 70 % of buffer B (Figure 3.9). The *D/B*SH fractions were pooled and loaded on SEC650 (Biorad) gel filtration column and we found that truncated protein is eluting around 44 kDa and monomeric unit showed molecular weight around 22 kDa (Figure 3.10). This confirms that the oligomeric profile of the truncated *D/B*SH is dimeric. To check the activity of purified protein, the protein was incubated with GDCA and run on TLC plate for qualitative analysis of BSH activity. The purified dimer truncated *D/B*SH was found to be inactive towards GDCA (Figure 3.11). To fully understand the role of oligomerization and its effect on biochemical activity and stability, we generated two different truncated (before and after tetramer loop) *D/B*SH i.e. *D/B*SH δ 189 and *D/B*SH δ 209 by introducing stop codon after residue 189 and 209 respectively and purified using the same protocol as it lacks the 6X-His tag. Furthermore, *D/B*SH δ 189-209 mutant was also generated to ascertain the role of assembly loop or tetramer loop in maintaining the oligomeric profile of CGH family. The oligomeric profile of the entire mutant explained above was shifted to dimer which showed the importance of tetramer loop in oligomerization (Figure 3.14). The full length wt *D/B*SH and *M*sBSH was purified using the same protocol described above and both the proteins were found to be tetramer. Figure 3.12 shows the Ni-NTA profile of both *D/B*SH and *M*sBSH and figure 3.13 shows the Size exclusion profile of *M*sBSH. The SEC profile of *M*sBSH showed that it is tetramer, but unstable and forming soluble aggregate as evident in figure. Hence, further study was only focused on *D/B*SH, its mutant and comparison with other reported BSH structures.

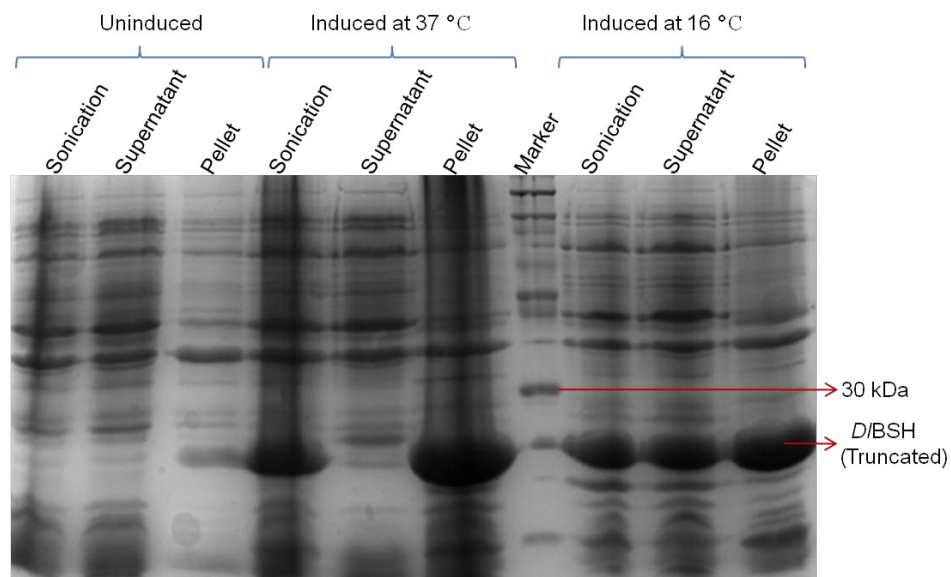


Figure3.5: Comparison of expression profile of *D/BSH* at 37 °C and 16 °C

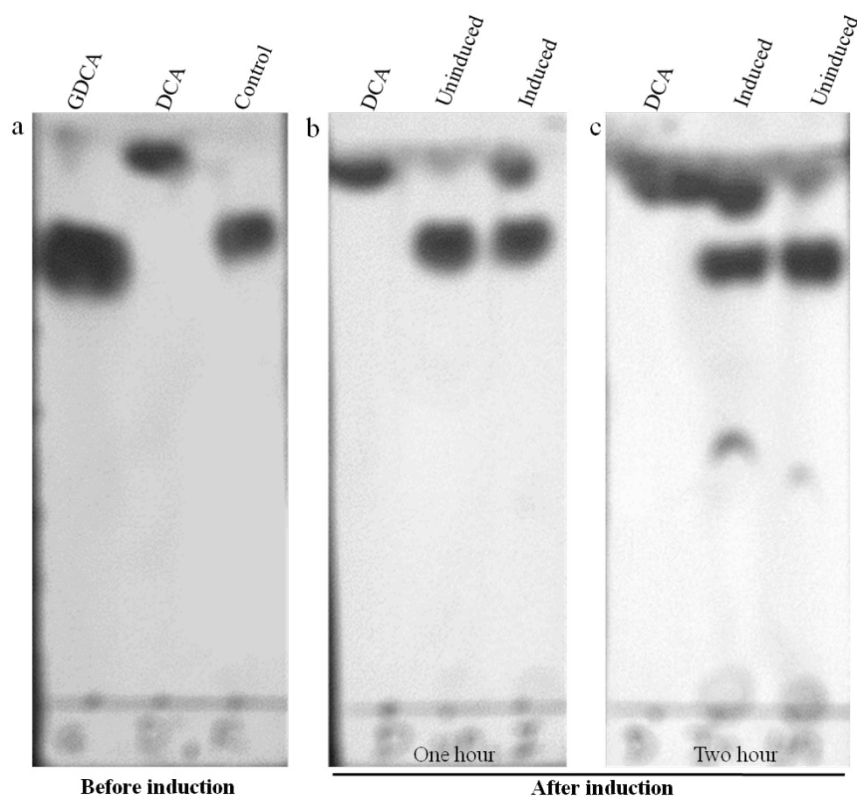


Figure3.6:Thin Layer Chromatography (TLC) profile of *D/BSH*(truncated) performed with whole cell.a) represents the run profile of standard GDCA, DCA and control (untransformed expression cells, i.e. *E.coli* Rosetta (DE3)) b) & c) represents the activity profile of truncated *D/BSH* in uninduced and induced cell with 1mM IPTG and compared with TLC profile of DCA as control after one hour and two hour, respectively.

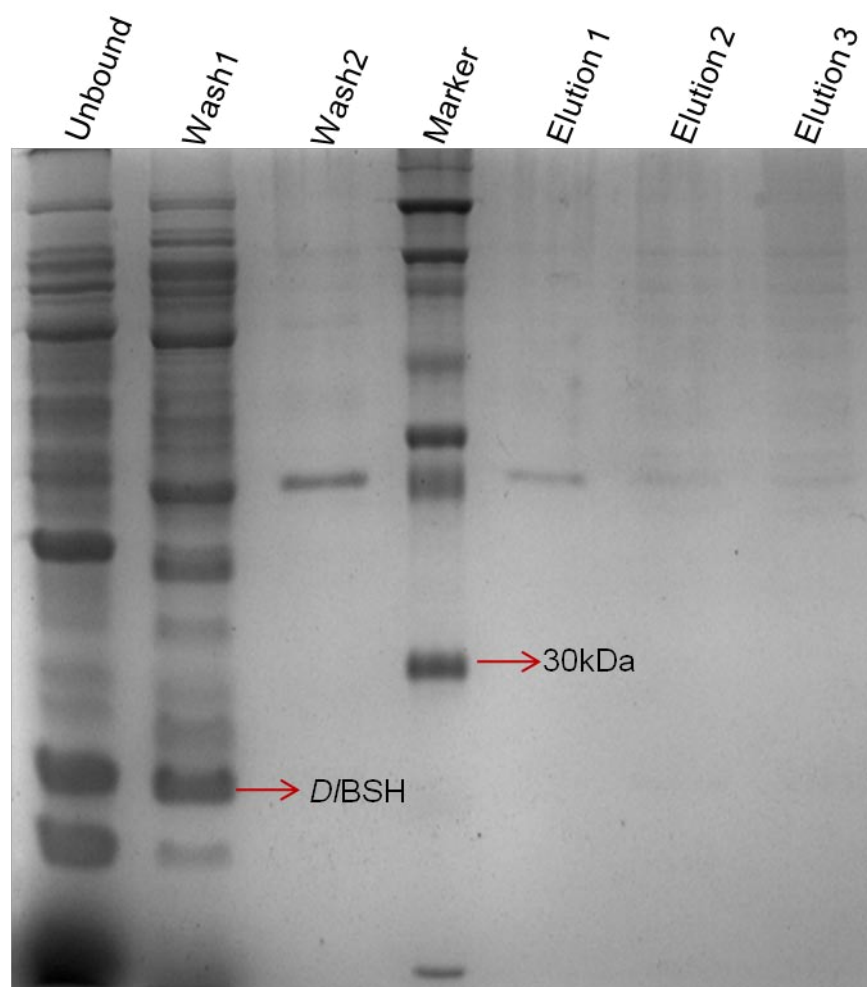


Figure3.7: SDS-PAGE gel profile after Ni-NTA purification. *D/BSH* fraction is present in unbound and wash, while absent in elution fraction.



Figure 3.8: Sequencing result of cloned *D*/BSH compared with the wild type *D*/BSH. Highlighted region shows the insertion of a base (Cytosine) leading to frameshift mutation which introduces stop codon (highlighted yellow).

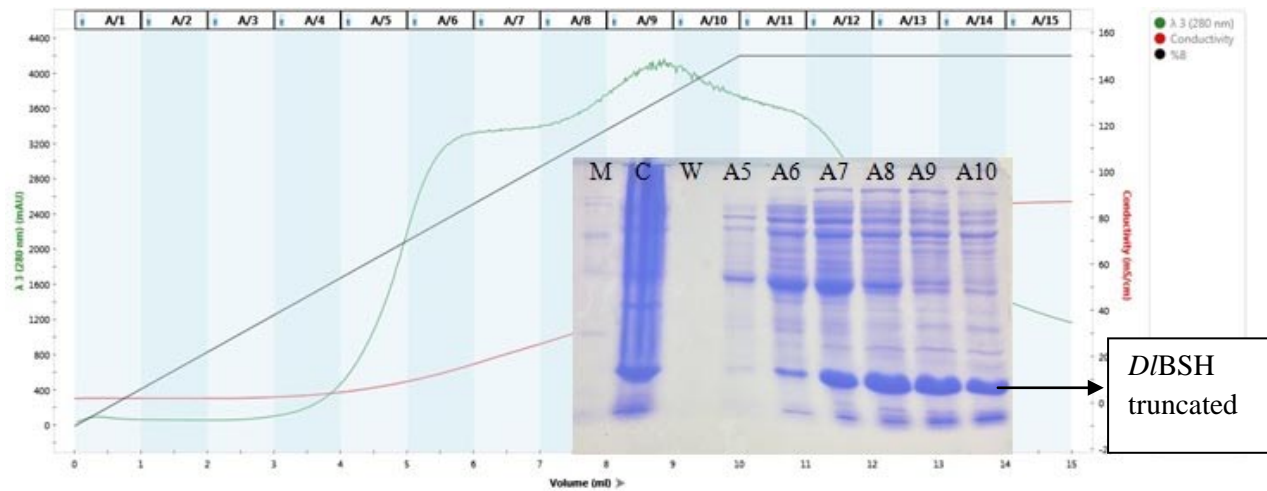


Figure 3.9: Ion exchange chromatography profile of truncated *D*/BSH using anion exchanger Q-sepharose column.

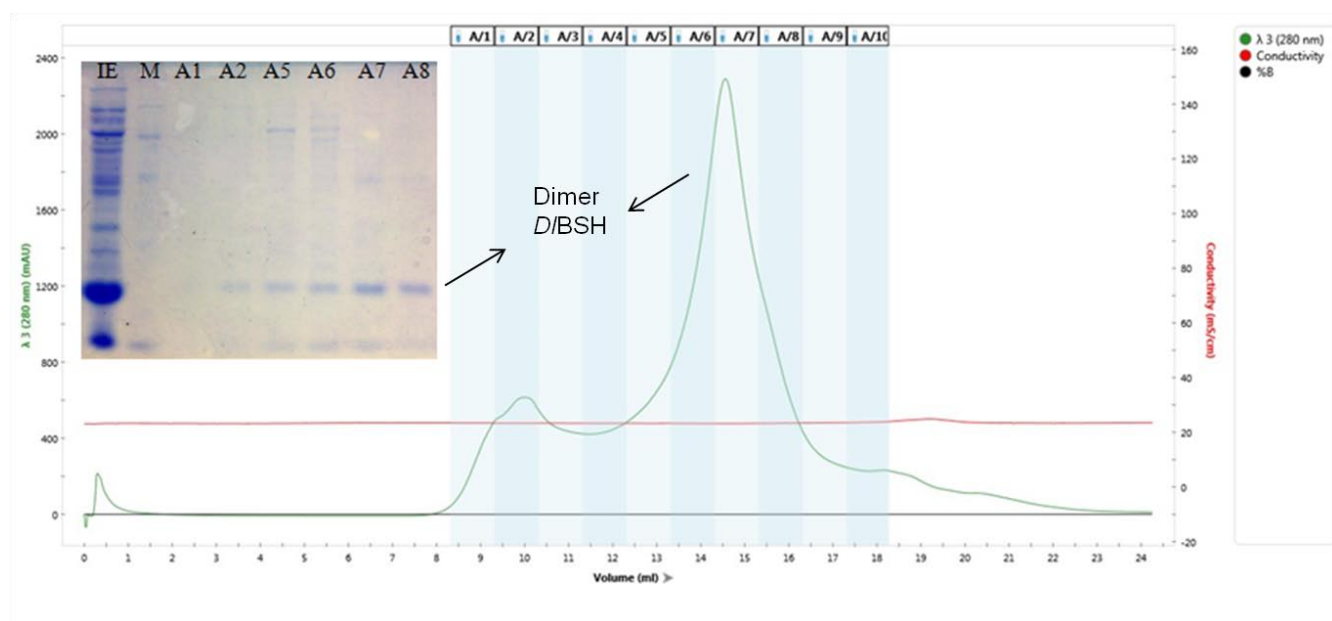


Figure 3.10: Purification of dimeric *D/BSH* using SEC650 (Biorad) column

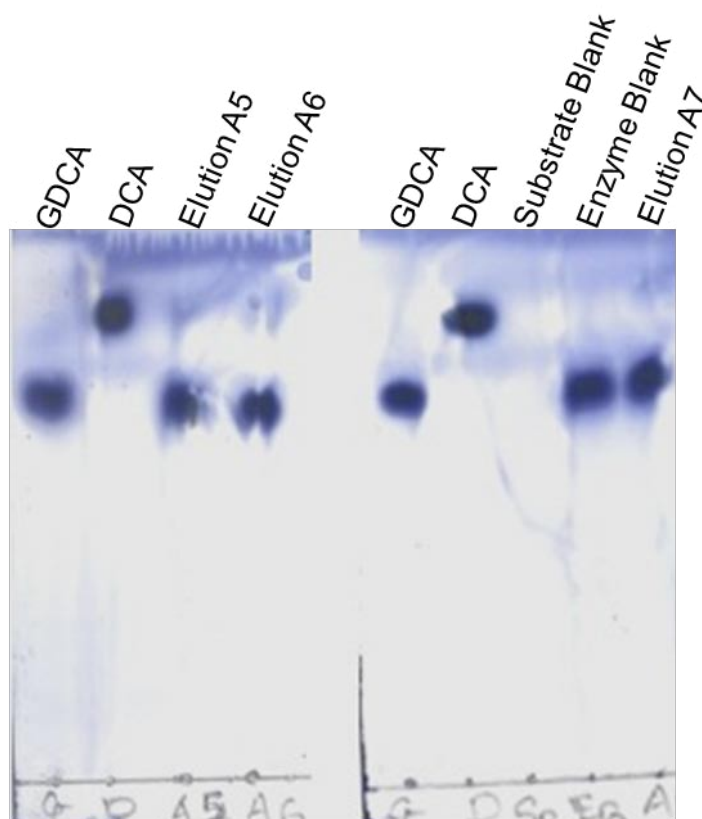


Figure 3.11: Qualitative estimation of the activity profile from purified dimeric form of truncated *D/BSH* was performed using of TLC assay.

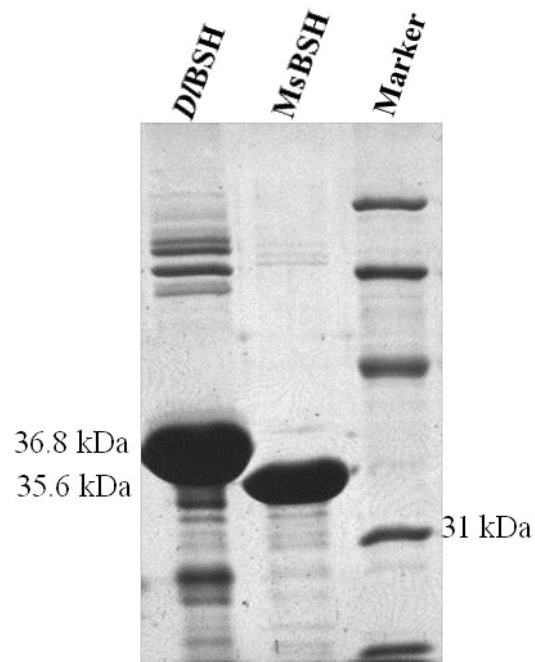


Figure3.12: Ni-NTA purification profile of *DIBSH* and *MsBSH* showing difference in molecular weight owing to reduced length of *MsBSH* (316 aa) as compared to *DIBSH* (324 aa).

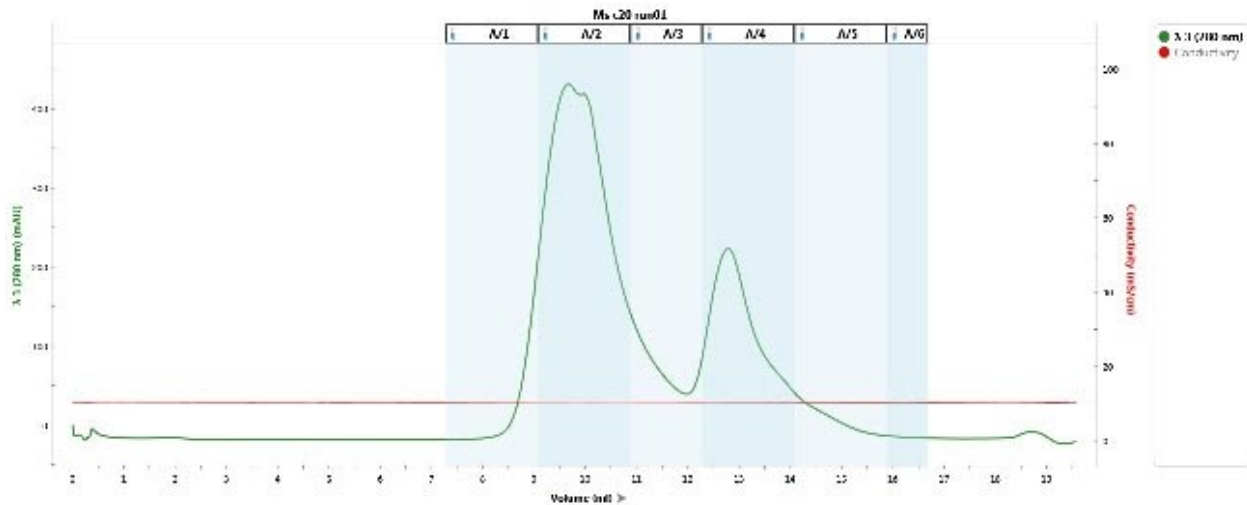


Figure3.13: Size exclusion chromatography profile of *MsBSH*. First peak is eluting at ~10 mL which is indicative of soluble aggregate. Soluble portion of *MsBSH* is eluting at ~12 mL.

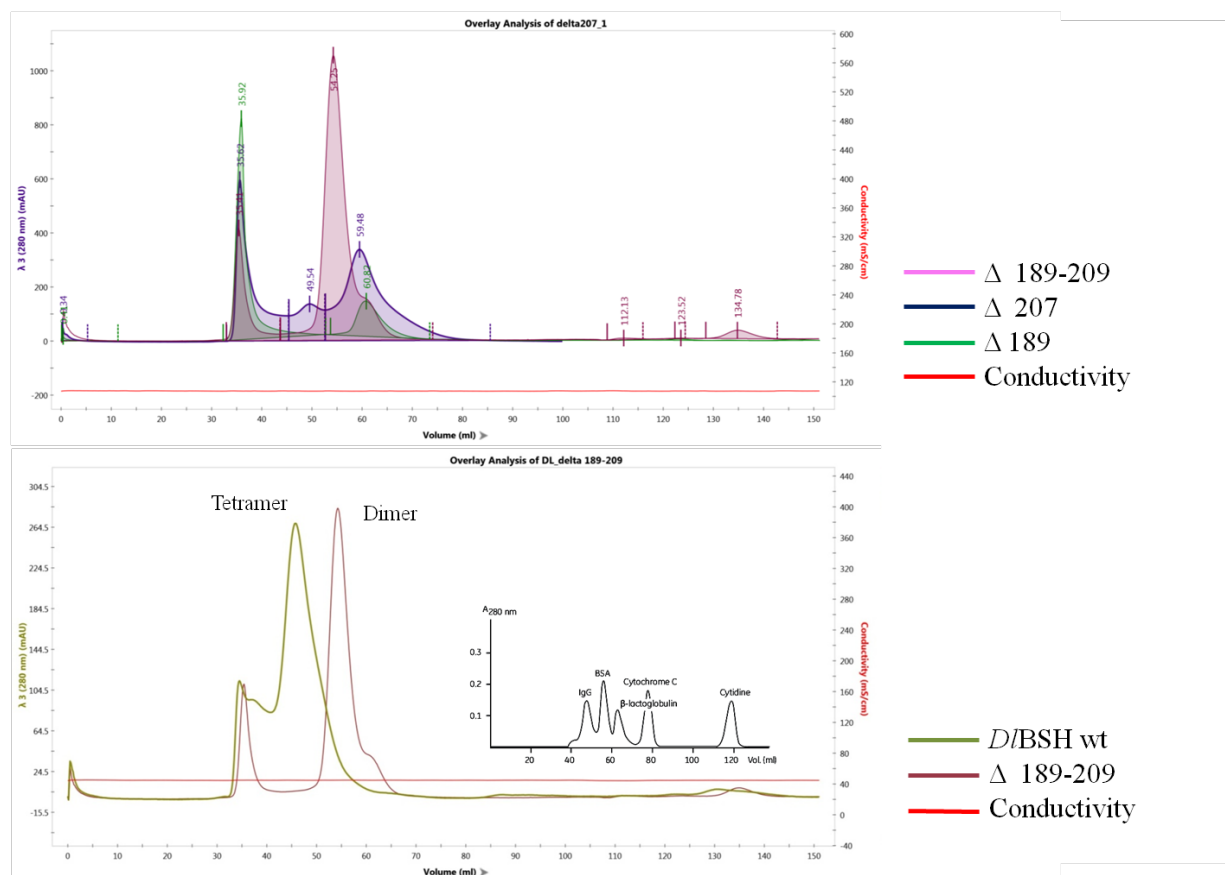


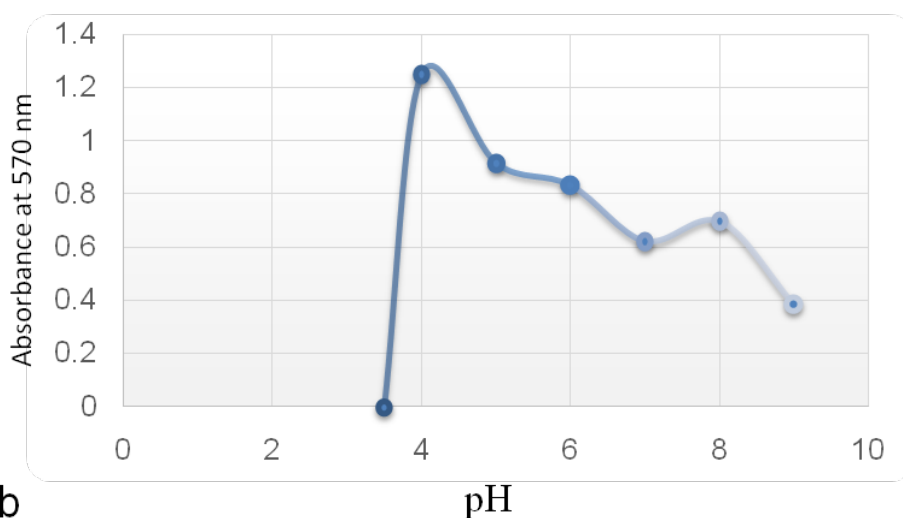
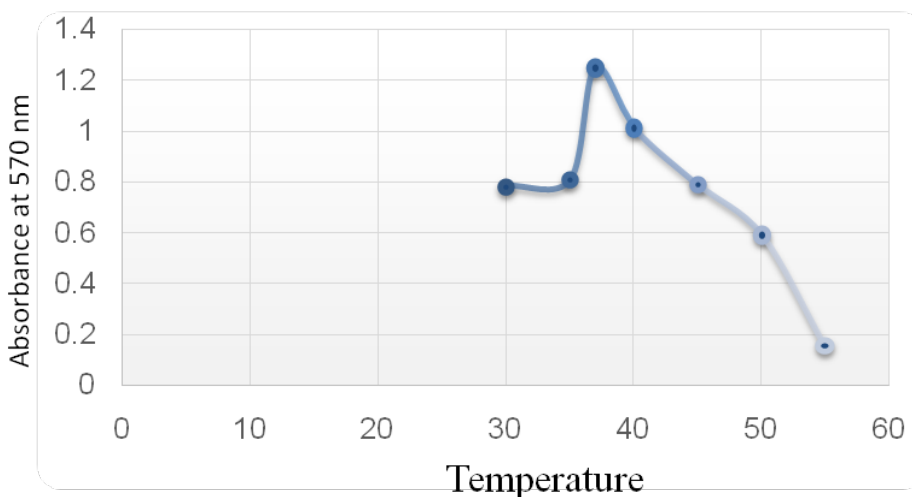
Figure 3.14: Overlay analysis of Size exclusion chromatography profile of WT *D/BSH* and its mutant. Top panel shows the overlay analysis of SEC profile of *D/BSH* deletion mutants from C-terminal end. Bottom panel represents the deletion of tetramer loop or assembly loop region and compared with WT *D/BSH*.

Table 3.5: Retention volume and oligomer profile of wt*D/BSH* and mutants using Sephacryl S200 column.

Protein	Retention volume (mL)	Oligomer state
<i>D/BSH</i> wt	45	Tetramer
<i>D/BSH</i> δ 189-209	54.25	Dimer
<i>D/BSH</i> δ 189	60.82	Dimer
<i>D/BSH</i> δ 209	59.48/49.54	Tetramer/Dimer

3.3.3 Kinetic characterization:

The purified *D*/BSH enzyme was assayed with GCA for determination of optimum pH and temperature. In case of *D*/BSH, the optimum pH was found to be around 4.0 and optimum temperature was 37 °C in citrate phosphate buffer (Figure 3.15). The mutation of interface residue i.e. Y175A in *Ef*BSH reduced the optimum temperature from 50 °C to 37 °C and kinetic calculation also showed drastic reduction in catalytic activity (Figure 3.16). The preferential substrate activity of *D*/BSH was studied using different substrates i.e. GCA, GDCA, TCA, TDCA, GCDCA. Only glycol- conjugated bile salts showed activity while tauro conjugate bile salts have very low activity as compared to Glyco-conjugated bile salts (Figure3.17).

a**b**

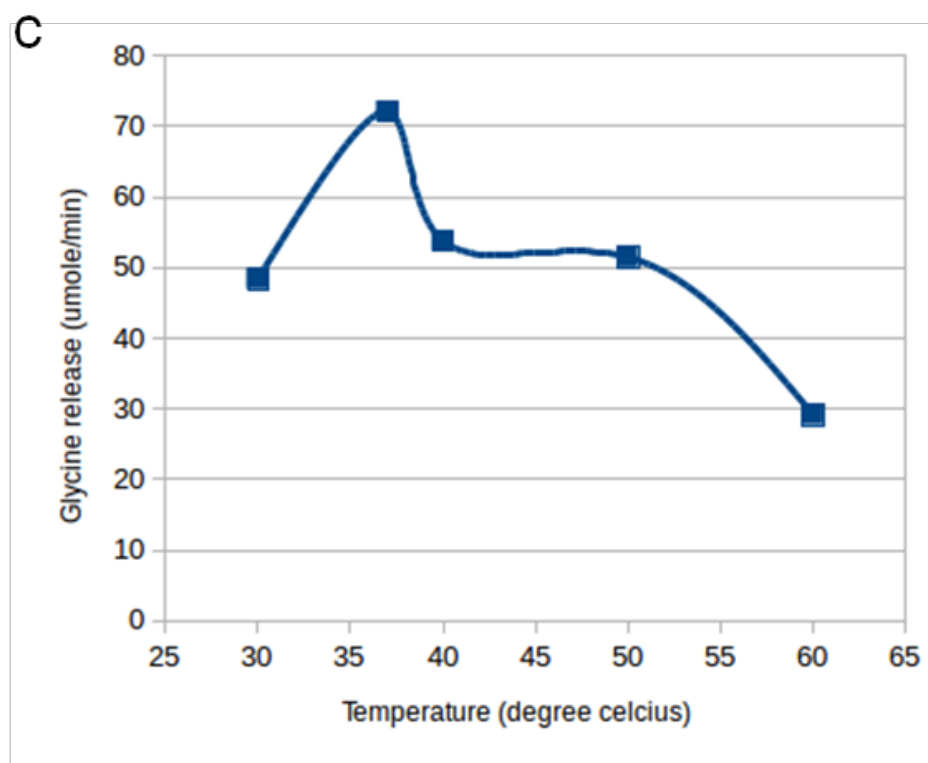


Figure 3.15: Biochemical characterization of *DfBSH* a) optimum pH for *DfBSH* b) &c) optimum temperature of *DfBSH* and *EfBSH* Y175A respectively.

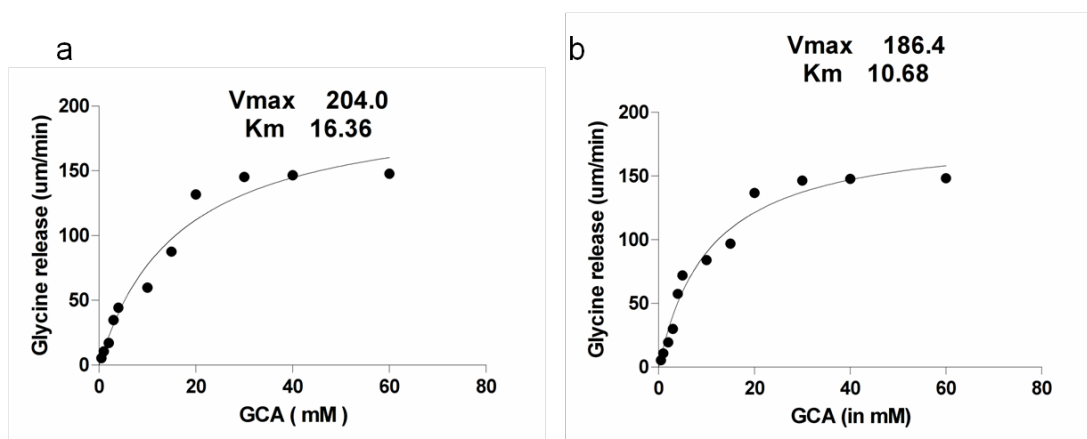


Figure 3.16: Kinetic characterization of *DfBSH* a) steady-state kinetics plot of *DfBSH*, and b) *EfBSH* Y175A mutant performed with GCA.

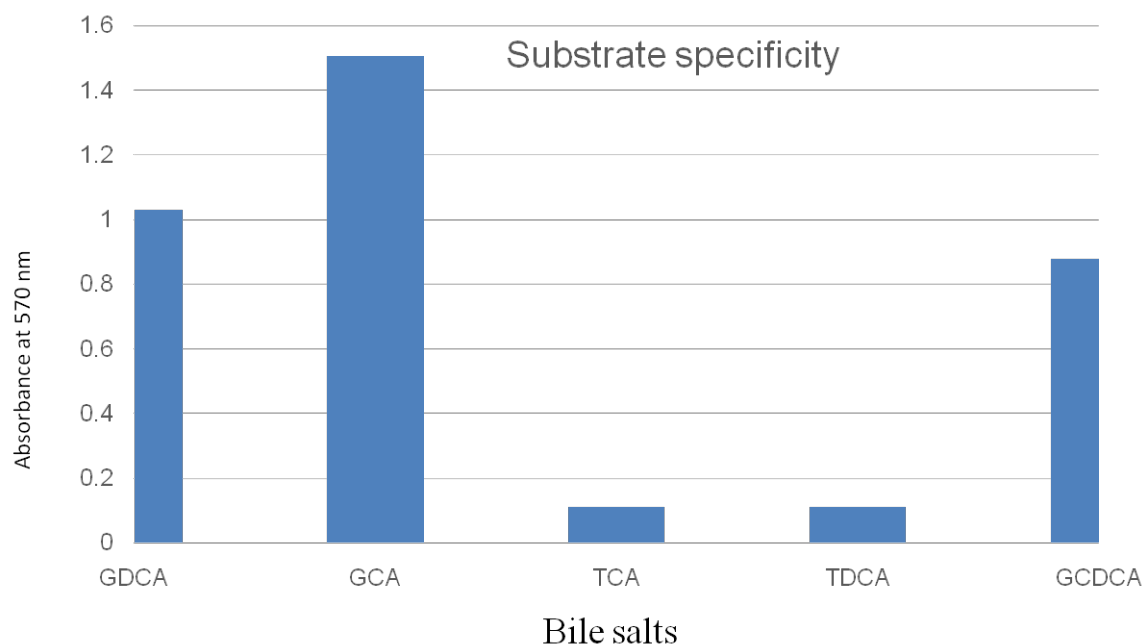


Figure 3.17: Preferential substrate activity profile of *D/B*SH with conjugated bile acids.

3.3.4 Structural comparison of CGH family of proteins

- **Structural analysis of Dimer/Tetramer in Gram positives**

Sequence alignment analysis showed more than 70 % identity in between *Ef*BSH and *D/B*SH and has different optimum temperature. The optimum temperature of *Ef*BSH is around 50 °C and of *D/B*SH is 37 °C. Hence, to find out the structural difference between the two proteins which showed different optimum temperature and therefore thermostability, we modeled the structure of *D/B*SH and compared with the other reported BSH and PVA structures. Interface residue analysis of *Ef*BSH showed Tyr176 faces each other and further stabilized by Phe179 (Figure 3.18). These residues form minicluster at the tetramer interface. However, in 2HEZ (*B/B*SH) and 2RLC (*Cp*BSH) these miniclusters are not possible as the interface residues are facing in opposite direction (Figure 3.19). In case of *D/B*SH, Gln176 is present instead of Tyr176 as in *Ef*BSH (Figure 3.20).

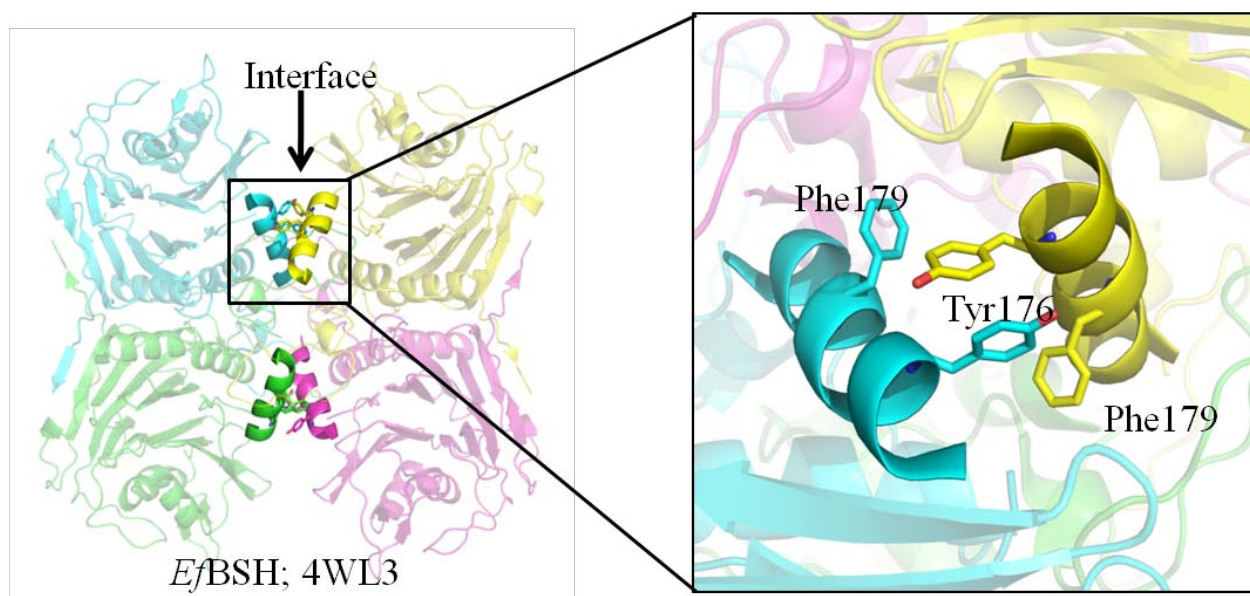


Figure 3.18: Cartoon representation of tetramer form of *EfBSH* (PDB ID: 4WL3) highlighting the tetramer interface. Inset shows the key residues of tetramer interface forming miniclusters.

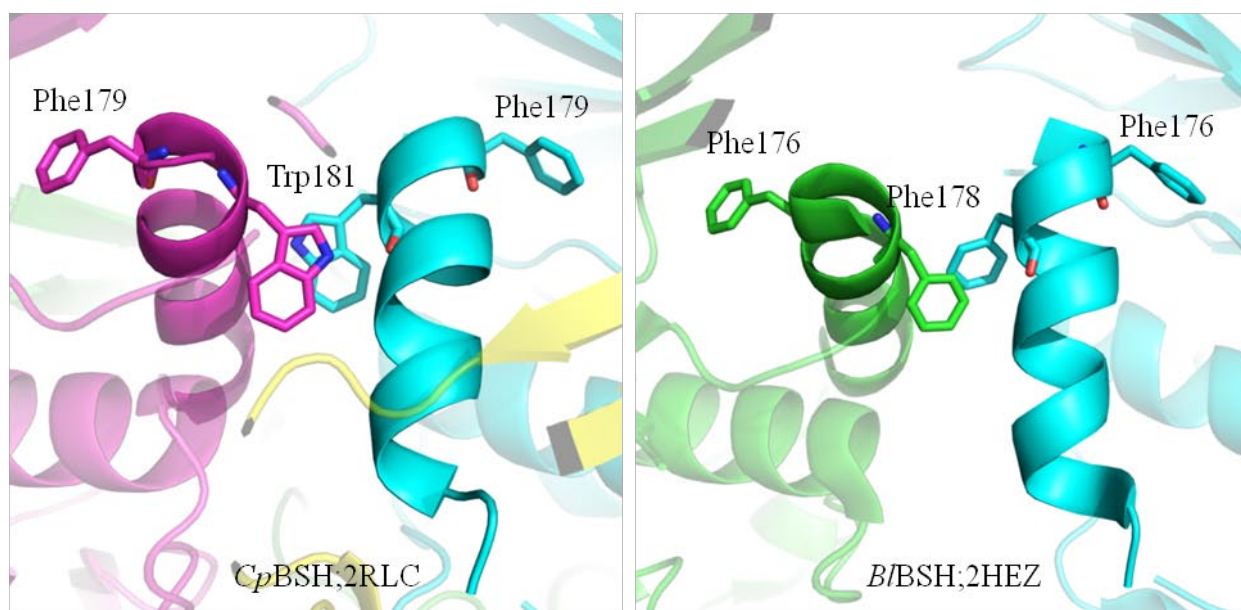


Figure 3.19: Tetramer interface of *CpBSH* (Left; PDB ID-2RLC) and *B/BSH* (Right; PDB ID-2HEZ).

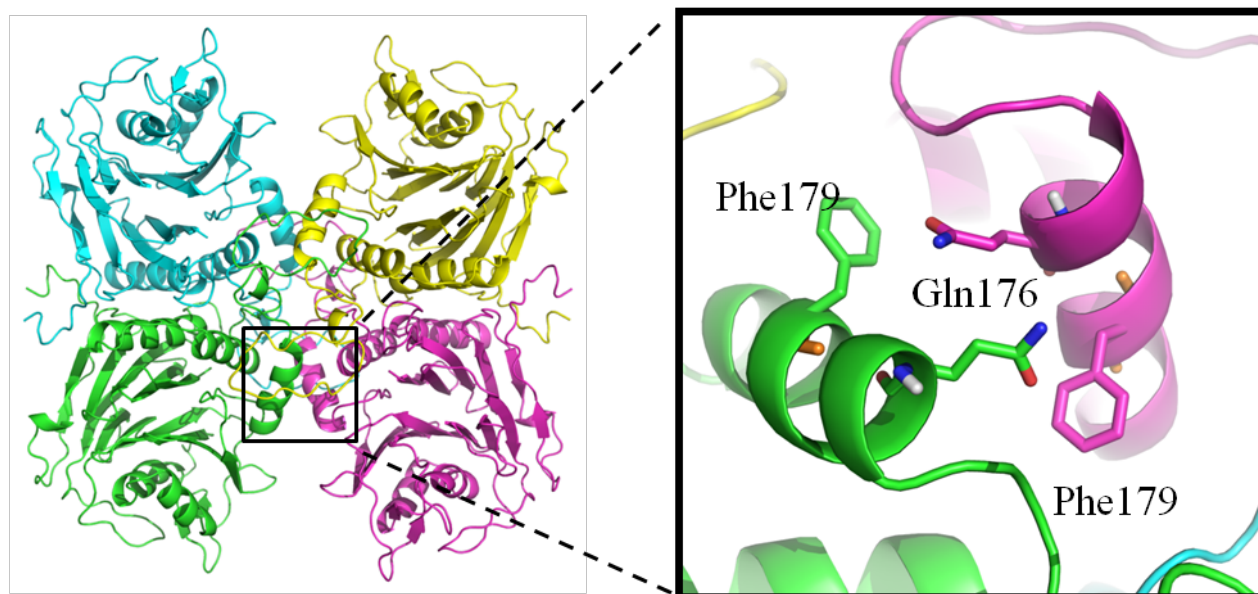


Figure 3.20: Tetramer interface of *DIBSH* (model) highlighting the presence of neutral amino acid i.e. Gln176 in interface minicluster.

The superposed structure of *DIBSH* and *EfBSH* showed that the spatial location of Gln176 and Tyr176 is conserved as well as, Phe179 of both chains are also superposing well (Figure 3.21). Removing the tetramer loop from both *EfBSH* and *DIBSH* leads to dimerization of protein. This result indicates the significance of tetramer loop in oligomerization, but recently *LsBSH* was reported to be dimer even with tetramer loop or assembly loop domain (Figure 3.22 & 3.23). This indicates rather than whole tetramer loop, perhaps, certain residues are taking part in maintaining oligomerization. Consurf analysis showed that the tetramer loop region is highly variable except Tyr205 to Arg207. Data from PDBsum showed that Tyr205 interact with Phe179 (Figure 3.24).

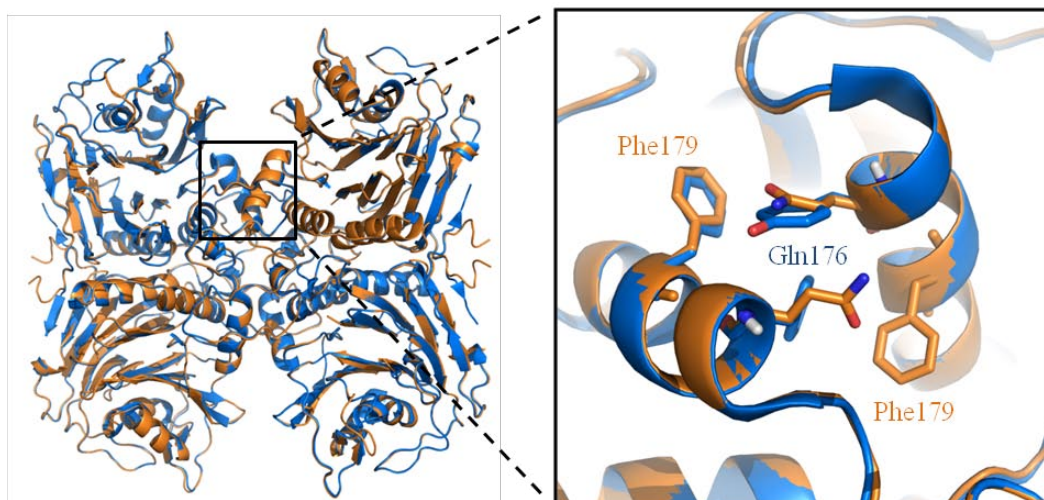


Figure 3.21: Superposed cartoon representation of *EfBSH* and *DfBSH*. Inset image shows the tetramer interface, pointing out the difference in residues forming miniclusters. Tyr176 in *EfBSH* (Blue) is replaced by Gln176 in *DfBSH* (Orange).

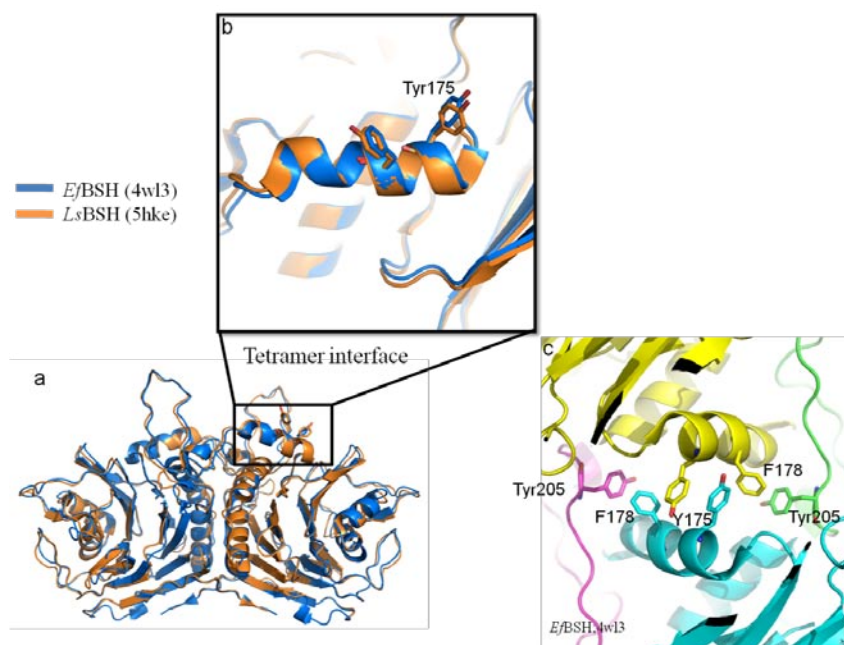


Figure 3.22: Tetramer interface residue mapping in dimer (*LsBSH*) and tetramer (*EfBSH*) Bile Salt Hydrolase a) superposed cartoon representation of *LsBSH* and *EfBSH*. b) Inset image shown in that the interface residues are conserved in *LsBSH* and *EfBSH*. c) Tetramer interface showing residue participation from all chains shown in different color. Tyr205 is coming from tetramer loop region.

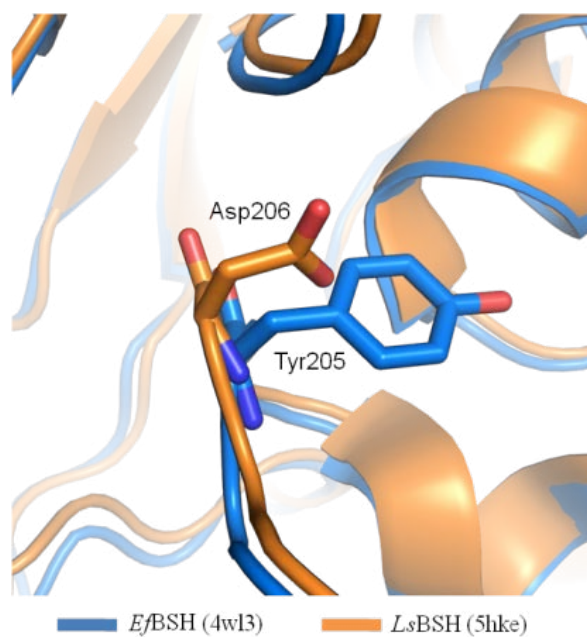


Figure 3.23: Superposed cartoon representation of *EfBsh* and *LsBsh* showing Tyr205 and Asp206, respectively in stick form.

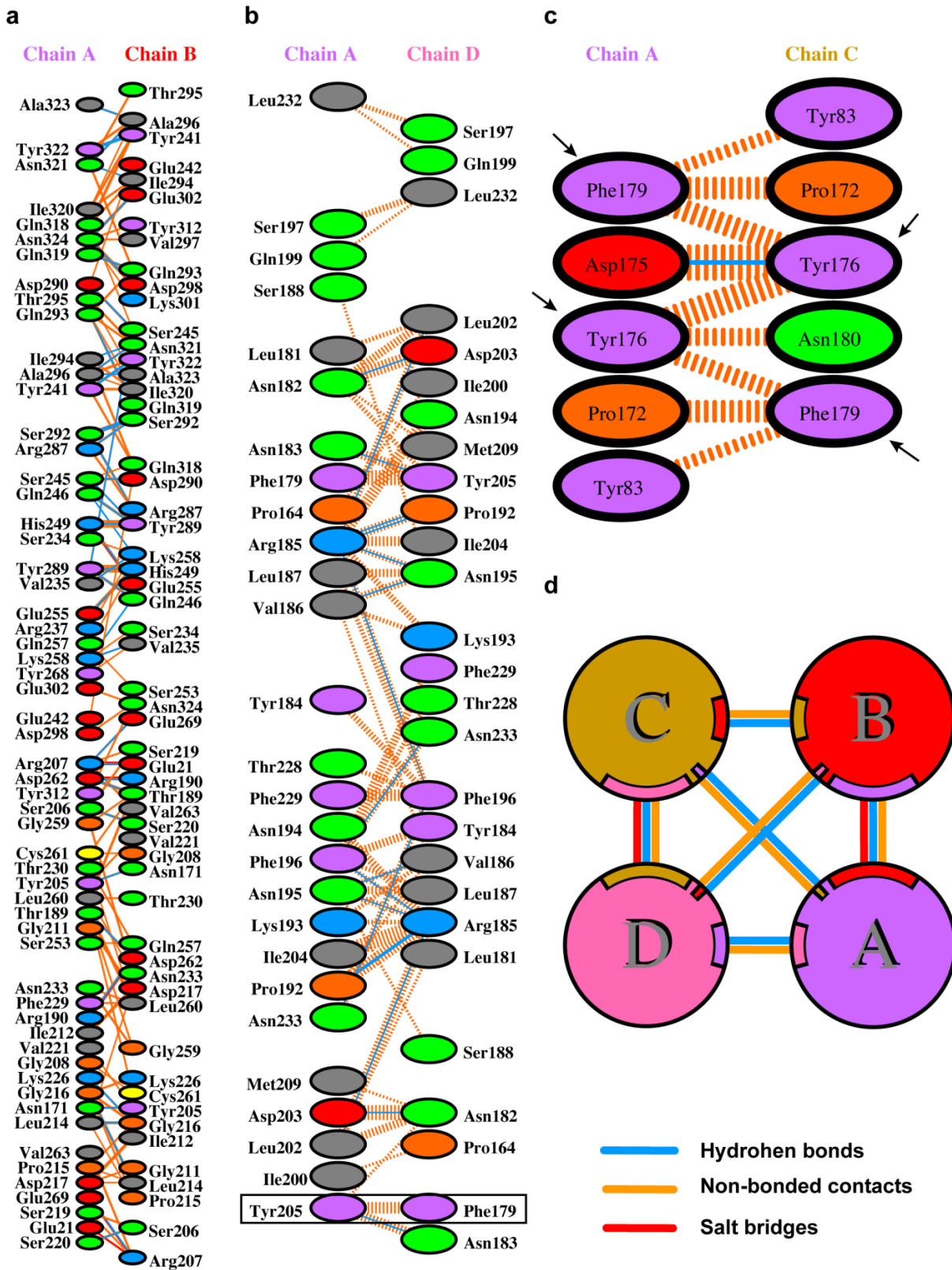


Figure 3.24: Interface analysis of *EfBSH* (PDB ID – 4WL3)

- **Structural analysis of Dimer/Tetramer in Gram negatives**

Although the tetramer loop is absent or shortened in gram negative organisms, *PaPVA* structure revealed stable tetramer. Analysis of interface using PDBsum showed the presence of Met205 which shows non-bonded contacts with Tyr196 (Figure 3.25). Recently, acyl homoserine lactone (AHL) acylases also incorporated in CGH family of Ntn-hydrolase superfamily (data not published). Structural studies showed that the protein exist in homodimeric state and has similar feature to PVA and hence annotated wrongly as penicillin acylases. Likewise, in gram positives, the structure of *Shewanella loihica* CGH (*SICGH1*) from Gram negatives is unique which is attributed to its dimeric oligomer profile. Hence, mapping the difference in sequence between tetramer and dimer of gram positives and Gram negatives will be of significance of predicting the oligomeric profile. Tetramer interface comparison of tetrameric and dimeric structure in gram negatives already reported in PDB showed major difference in interface residues. Superposed structure of dimer of *PaPVA* and *SICGH1* represented in cartoon form (Figure 3.26) with RMSD of 2.590 (2566 atoms). Residues of tetramer interface of *PaPVA* and *SICGH1* are compared and we found that hydrophobic amino acid is present in *PaPVA* is replaced by more hydrophilic amino acids on the tetramer interface (Figure 3.25 and 3.26) such as Met205 and Val204 in shortened tetramer loop region of *PaPVA* is replaced by more hydrophilic residues such as Lys215 and Thr214 in *SICGH1* (Figure 3.27b). Furthermore, corresponding position of Tyr196 and Trp197 which forms the unique minicluster in gram negatives (here *PaPVA*) is occupied by Gln204 and Phe205 in *SICGH1* (Figure 3.27a).

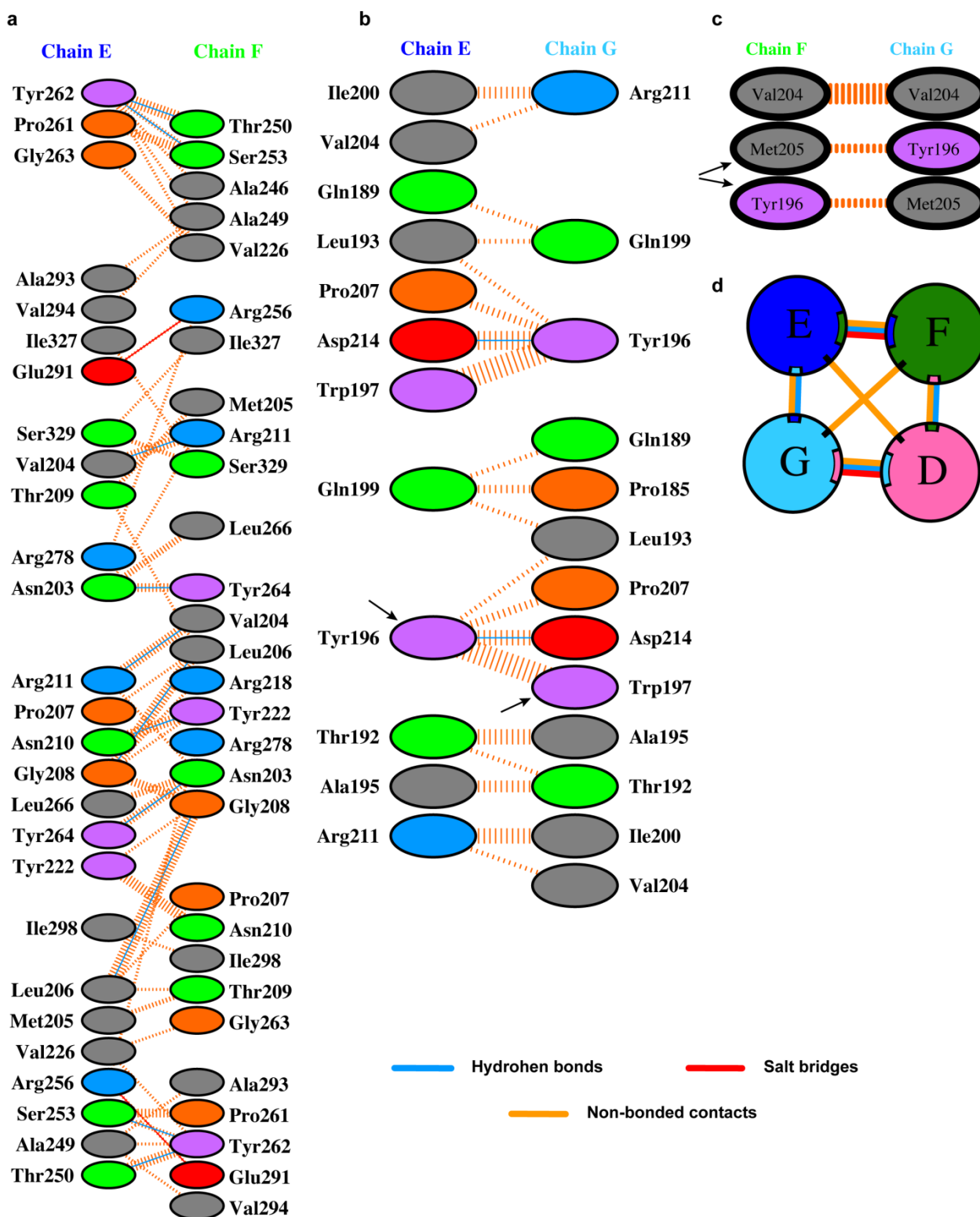


Figure 3.25: Interface analysis of *PaPVA* (PDB ID – 4WL2).

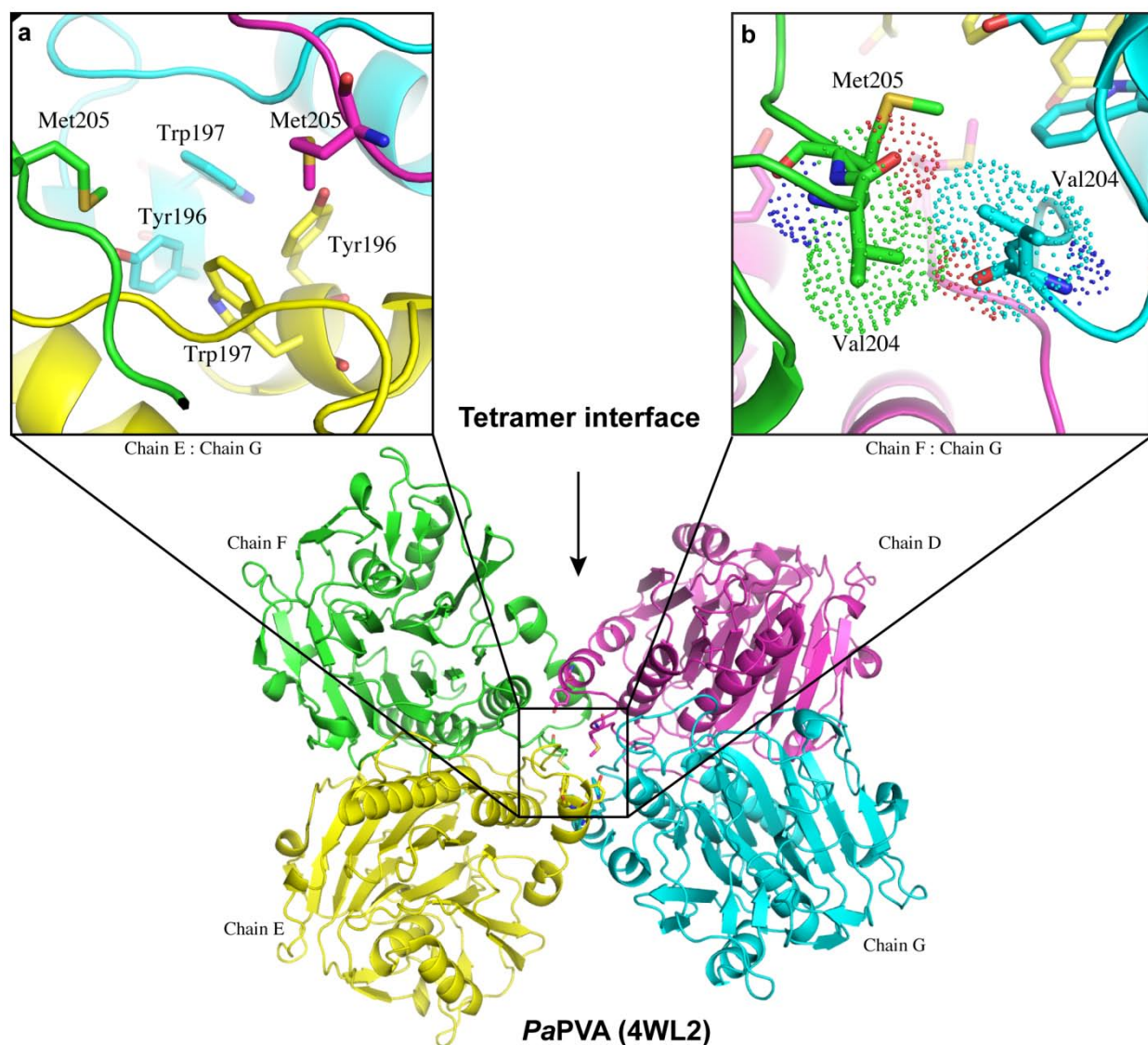


Figure 3.26: Cartoon representation of interface analysis in *PaPVA* (4WL2) Cartoon representation of tetramer of *PaPVA* with labeled chain ID as designated in PDB. Tetramer interface of *PaPVA* is indicated with arrow and further detail is shown in inset image. a) cartoon representation of residues which forms minicluster formed by Tyr196 and Trp197 of chain E and chain G and further stick form of Met205 is also shown which is analogous to the role of tetramer loop in Gram positive BSH. b) shows the presence of Val204 and Met205 in the chain F and chain G interface. Val204 forms hydrophobic interaction with Val204 of other chain of interface.

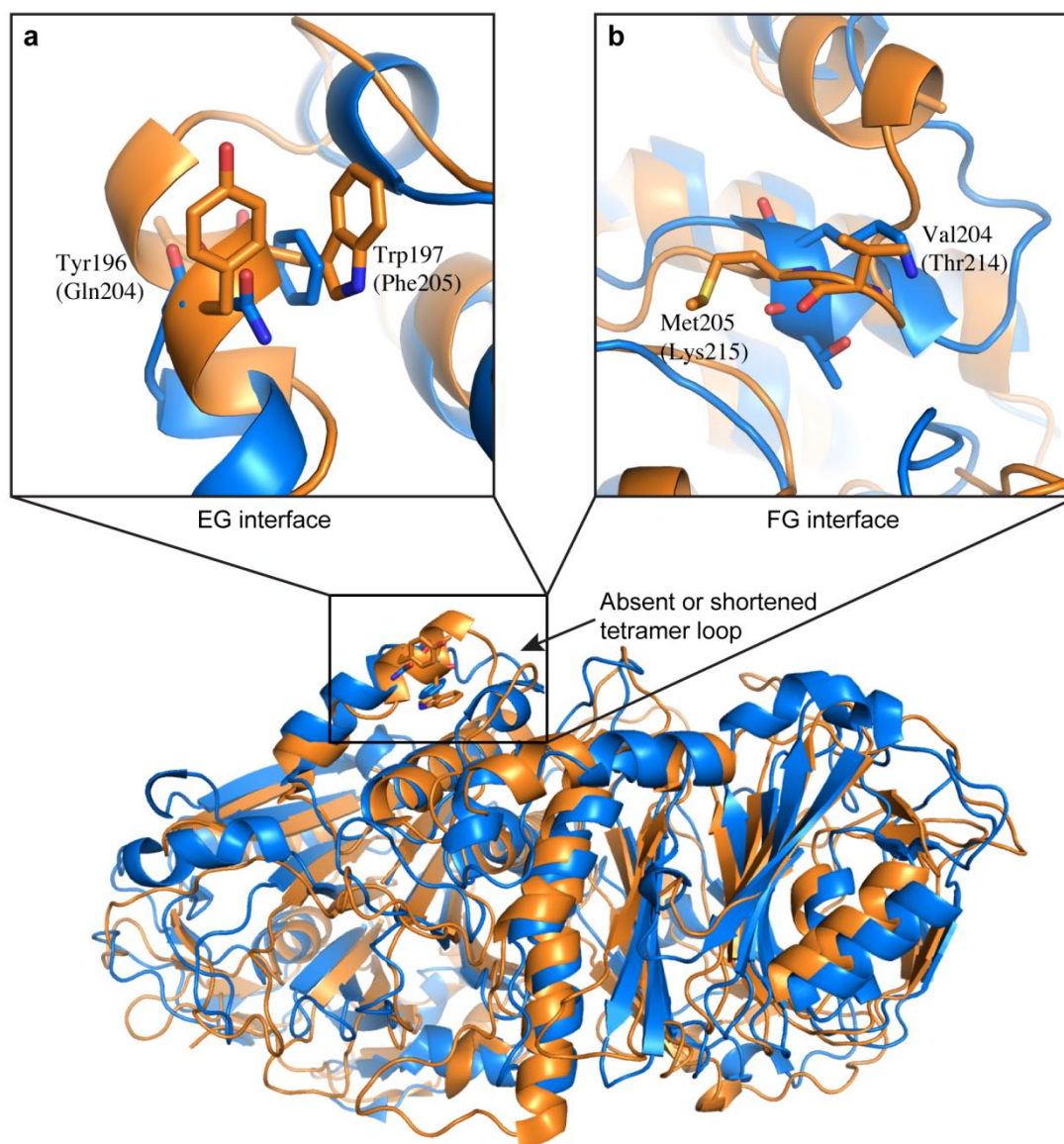


Figure3.27: Interface comparison of Tetramer and Dimer structure in gram negatives: superposed structure of dimer of *PaPVA* (orange) and *SICGH1* (Blue) shown in cartoon form with RMSD of 2.590 (2566 atoms). Residues of tetramer interface are compared in the inset image. a) stick representation of Tyr196 and Trp197 shows that the corresponding position in *SICGH1* is occupied by Gln204 and Phe205 in EG interface. b) Met205 and Val204 in tetramer loop region of *PaPVA* is replaced by more hydrophilic residues such as Lys215 and Thr214 in *SICGH1*.

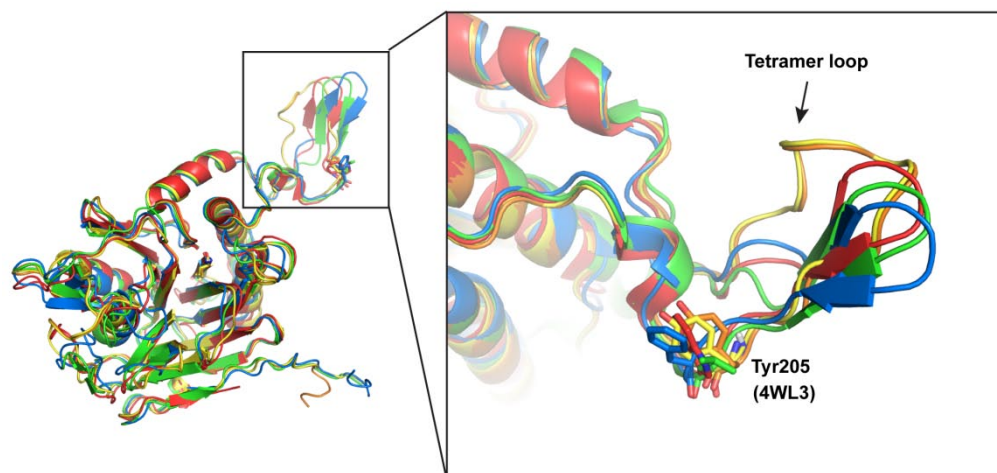


Figure 3.28: Superposed cartoon representation of Gram-positive BSH and PVA. BSH are from *Bt*BSH (2HEZ-Red), *Cp*BSH (2BJF-Green), *Ef*BSH (4WL3-Yellow), *Dt*BSH (Model-Orange) and *Bsp*PVA (3PVA-Blue) represents Gram positive PVA. Inset shows the corresponding amino acid of Tyr205 in *Ef*BSH.

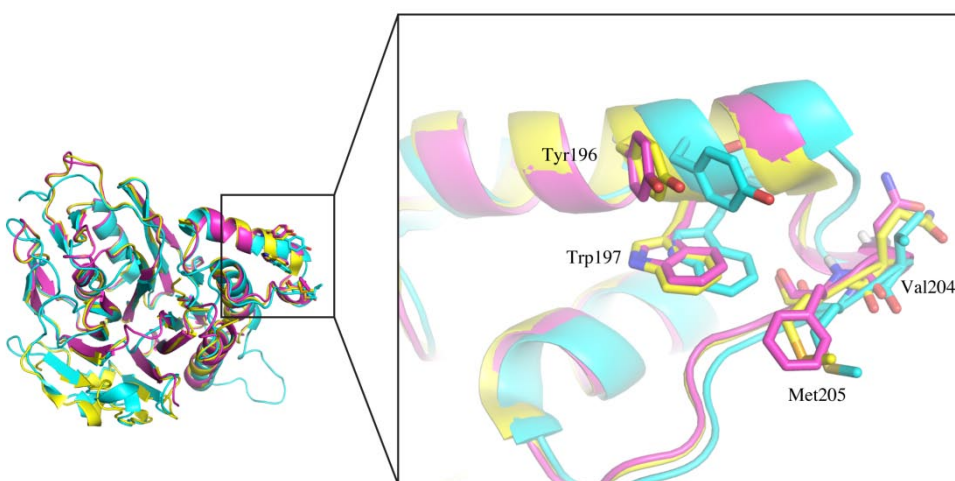


Figure 3.29: Superposed cartoon representation of Gram negative PVA and BSH. PVA are from *Bacteroides* (3HBC- Magenta), *Methanosarcina* (Model-Yellow) and *Pa*PVA (4WL2- Cyan). Inset picture shows the corresponding amino acid of Tyr196, Trp197, Met205 and Val204 in *Pa*PVA.

3.3.5 Density Function Theory (DFT):

We compared the energy of interface of highly active enzymes from both Gram positive as well as Gram negative interface i.e. 4WL3 (*Ef*BSH) and 4WL2 (*Pa*PVA). Both the enzyme showed unique miniclusters at the tetramer interface (Figure 3.30). We calculated the interaction energy between the dimers of dimer which, perhaps hold the tetramer. The calculated values using different theories have been listed in table 3.6.

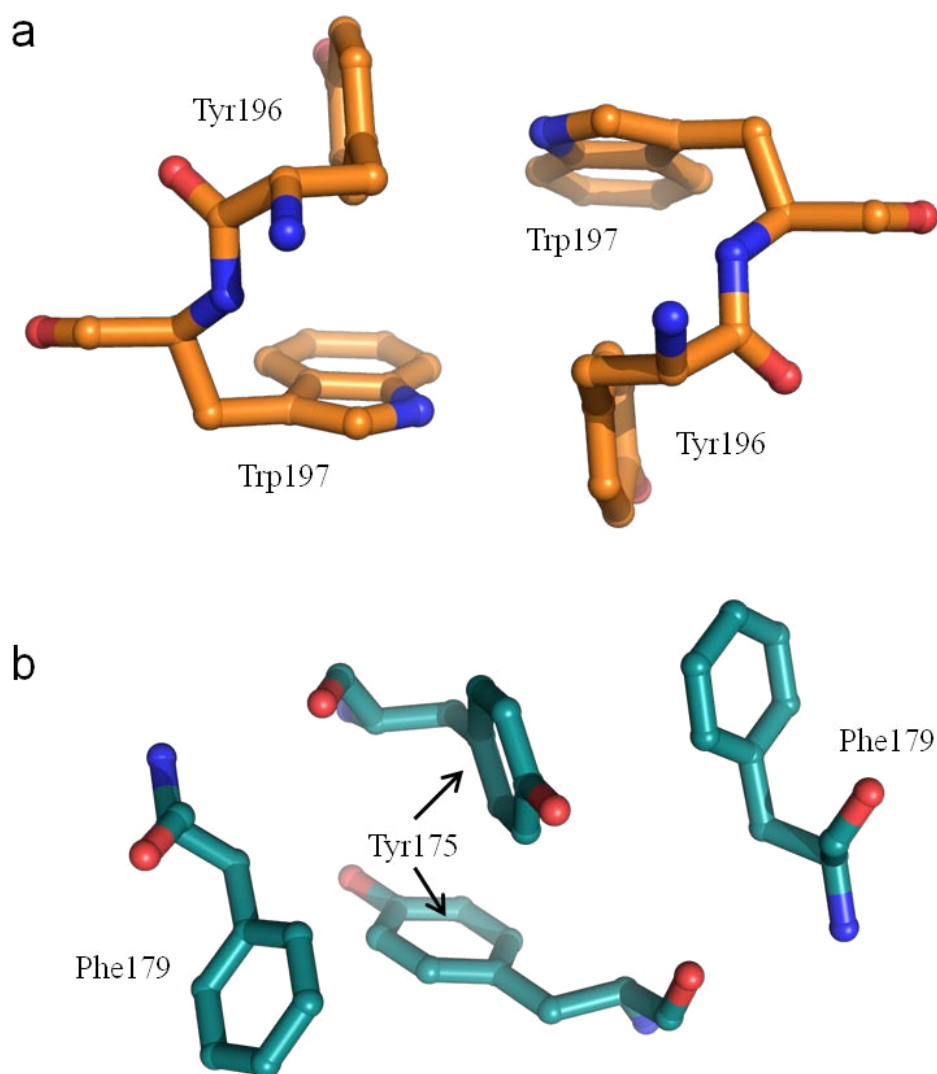


Figure 3.30: Miniclusters formed by tetramer interface residue of *Pa*PVA(a) and *Ef*BSH(b).

Table 3.6: Energy calculation between tetramerinterface residues using DFT.

Level of theory	E_i of 4w13 dimer (capped)	E_i of 4w12 dimer (capped)
PBE3(dis3)/COSMO (water)/TZVP	-9.7	-10.1
M06-2X/ CPCM (water) /6-311G**	-11.5	-9.3

E_i = interaction energy; Energies are in Kcal/mol

3.4 Discussion

Quaternary structure in Choloylglycine Hydrolase (CGH)

CGH subfamily of proteins is segregated further in two family viz. BSH and PVA. Recently, we suggested AHL acylase, on the basis of structure, biochemical function and oligomeric attribute, the third member of the family. Often in Ntn hydrolase superfamily, even vast sequence difference also shows signature $\alpha\beta\alpha$ fold. This makes difficult to classify the BSH from PVA and seldom wrongly annotated in databases. Panigrahi et. al., reported to specifically categorize the BSH/ PVA on the basis of Binding Score Similarity (BSS). Also, phylogenetic analysis showed two distinct clusters of Gram positive and Gram negative BSH/PVA on the basis of indel of 20-22 aa. These indel correspond to the tetramer loop region in gram positive organisms[16] and proposed to provide additional stability as compared to gram negative BSH/PVA[14]. Different homo-oligomeric state has been reported in CGH family of proteins[15–17,21]. Structural study of gram negative *Pa*PVA showed stable homotetramer without the presence of tetramer loop[18]. Recently, AHL acylase from gram negative *Sl*CGH1 was reported to be dimer in solution and crystal structure (PDB ID: 5X8Z and 5X9I). Similarly, crystal structure of *Ls*BSH(gram positive candidate) showed dimer (PDB ID: 5HKE) in spite of presence of tetramer loop region[17] while *B*BSH and *Cp*BSH were homotetramers[15,16]. Removing the tetramer loop (δ 189-209) region in *Dorea longicatena* BSH (*D*BSH) resulted into the homodimer (Figure 3.14). Truncating the C-terminal region i.e. δ 189 was purified to be dimer in solution while δ 207 showed two peaks which corresponds to tetramer and dimer on the basis of retention profile in

gel filtration (Figure 3.14). This confirms the role of tetramer loop in oligomerization. It is evident from the current and previous study that only specific residues take part in oligomerization rather than complete loop. In CGH family, BSH/PVA enzymes often show promiscuous activity. *L. plantarum*bsh1 shows feeble activity on 3-oxo-C6 and 3-oxo-C8 HSLs[22]; however, the inactive BSHs (bsh2-4) showed more activity on penicillins and AHLs than the bsh1 which is known to be active against bile acid. *SICGH* and *PaPVA* showed structural homology but differ in oligomeric state. Also, *SICGH* is active against C-10 HSL and *PaPVA* is known to hydrolyze Penicillin V. this difference in biochemical activity may be attributed to quaternary structure of proteins. Biochemical study of *DIBSH* showed preferential substrate activity towards glycol conjugated bile acids. In human, the ratio of glycol conjugated bile acid pool is more as compared to tauro-conjugated bile acids (3:1). Moreover, this study also suggests that the different strains of *Lactobacillus* showed difference in activity towards glycol- and tauro- conjugated bile acids which confirm the recognition of both glyceryl/tauryl and cholate moiety of substrates in the active sites. Kinetic data suggest that the *DIBSH* is less active as compared to previously reported structure (Figure 3.16 & 3.17). Mutation in Tyr276 in *EfBSH* to alanine (Y176A) showed shift in optimum temperature from 50 ° to 37 °C, which is similar to optimum temperature of *DIBSH* (Figure 3.15). Sequence analysis showed that Tyr176 in *EfBSH* is replaced by glutamine in *DIBSH* (Figure 3.3). DFT calculation of interaction energy of miniclusters formed between the chains resulted on average of -10kcal/mol in different level of theory (Table 3.6). Denaturation studies of *PaPVA* using 1M GdnHCl does not dissociate into dimer even without assembly loop which is contradictory to the study performed by panigrahi et al., which showed less thermodynamic stability of the tetramer in Gram-negative CGHs as compared to Gram-positive CGHs[14]. Tetramer interface of *PaPVA* also showed minicluster formed by tryptophan and tyrosine (Figure 3.30). It is evident from the above results that the aromatic residues near tetramer interface perhaps impart additional thermostability to the tetramer molecule.

Interface comparison of dimer and tetramer

Interface of wild type *EfBSH* showed unique arrangement of aromatic residues at the tetramer interface. Tyr175 of helix involved in Pi- Pi interaction with the same residue of adjacent chain. Furthermore, it also interacts with Phe179 which leads to clustering of four side chain, this

minicluster were further supported by Asp205 from assembly loop (Figure 3.24). Consurf analysis showed that Asp205, 206 and Arg207 are well conserved in all reported BSHs. The more appropriate role of these residues could be the stabilization of tetramers interface. Presence of aspartic acid in corresponding position of Tyr205 (Figure 3.23) revealed dimeric structure of *LsBSH*, as other corresponding mini-clusters in interface are conserved[17] and assembly loop or tetramer loop is also present. Hence the role of tyrosine 205 could be the stabilization of tetramer interface. In Gram negative *PaPVA* structure, similar unique arrangement of mini-cluster has been observed. Here the interface residues are Tyr and Trp, which forms tightly packed minicluster. These miniclusters are further supported by Met205 (Figure 3.25). This structural arrangement showed that two different type of mini-clusters are possible i.e. in between N and N+1 as in Gram negative PVA/BSH or N and N+3 in gram positive BSH/PVA (Figure 3.19). In gram negative organisms these tetramers are stabilized using such miniclusters as the interface area are very small as compared to Gram positives exhibiting assembly loop[14]. By looking at the interface residues it is well understood that gram negatives adopted modest approach of being stable without having assembly loop. This is in agreement with the previous study which showed that the diderms are more evolved as compared to monoderm[23]. The interaction energy calculated using DFT by applying two different theories showed very high interaction energy, nearly -10 Kcal/mol for one such minicluster in both *EfBSH* and *PaPVA* (Table 3.6). Interaction energy calculated between cationic – Pi interaction was around -5 Kcal/mol as shown by Dougherty et al. [24]. This shows that, for two such miniclusters in tetramer interface, the interaction energy would be sufficiently high to hold the tetramer structure. But removing the tetramer loop in *D/BSH* changes the oligomeric profile from tetramer to dimer, and dimer structure is reported for *LsBSH* even in the presence of tetramer loop. This indicates that presence of a residue rather than whole tetramer loop is essential for holding the tetrameric form of BSH/PVA. As reported by Avinash et al., these *PaPVA* structure are highly stable under higher temperature, the reason for which was not discussed in the previous paper[18]. Furthermore, BSH analysed from other species such as *D.longicatenash* showed variation in interface residues. Corresponding position of Tyr175 in *EfBSH* is replaced by Asn175 in *D/BSH* (Figure 3.20 & 3.21). Biochemical characterization showed that the optimum temperature of *D/BSH* is at 37 °C as compared to 50 °C in *EfBSH*, while optimum pH was nearly similar i.e. in acidic range of pH 4-5. From the calculation of interface energy, we can say that the arrangement

of such aromatic residues in interface may imparts additional thermostability to the protein molecule. This information can be extrapolated to improve the thermostability or tailoring the enzyme activity of industrially important enzymes. Interface analysis showed that both (Gram +ve and Gram -ve) the cluster have adopted a unique way to evolve in constrained environment. Both the mini-cluster have supporting residues viz. Met205 in *PaPVA* and Tyr205 in *EfBSH*. Superposed structures of Gram +ve and Gram -veBSHs/PVAs showed similar arrangement of corresponding residues (Figure 3.27& 3.28). Dimeric structure reported in PDB so far (5HKE and 5X8Z), showed more hydrophilic residues at the respective contact point which we discussed here as EG/FG interface in *PaPVA* (Figure 3.25) and AD/AC interface (Figure 3.24) in *EfBSH*. In *LsBSH* corresponding position of Tyr205 in *EfBSH* is replaced by Asp206 in *LsBSH* and similarly, in *SlCGH*, Met205 is replaced by Lys215. Although Val205 in *PaPVA* making strong hydrophobic contact (Figure 3.26) with Val204 of other chain, presence of aspartic acid in place of valine in *Methanosarcinae*(Model) and *Bacteroidesthetaiotaomicron* (3HBC) confirms no role in tetramer formation (Figure 3.29) as crystal structure of *Bacteroides* showed tetramer arrangement. This confirms the importance of similar role of Tyr205 in Gram +ve's and Met205 in Gram -ve's in maintaining tetrameric oligomer state.

3.5 Summary and conclusion

To understand the principle of oligomerization in CGH family, the interface of previously reported structure and also performed structural and biochemical characterization *Dorea longicatena* BSH (*DlBSH*). Oligomerization study was performed in greater detail with *DlBSH*, in which several mutations has been created using Site directed mutagenesis approach. Changes in oligomerization were observed upon deletion mutation in tetramer loop region. Removing the tetramer loop (δ 189-209) showed complete shift to dimer state. In gram negative organisms' tetramer BSH/PVA are stable without assembly loop while in Gram positive organisms removing the tetramer loop changes the oligomerization profile. Moreover, gram negative BSH/PVA structure showed non-planar tetramer. Analyzing the interface residue of gram positive and gram negative showed that both the classes evolved through unique mechanism of holding the tetramer. We identified, upon interface comparison of dimer and tetramer BSHs, Tyr205 in tetramer loop region to be essential in oligomerization. Similarly, in Gram negative *PaPVA* structure, showed presence of Met205 in tetramer interface present in shortened tetramer loop. These residues, perhaps crucial in holding the miniclusters present at tetramer

interface and also serve as contact point of residues from all four chains of tetramer. This in part explains why gram-negative structures displayed non-planar tetramer structure and perhaps the modest way of evolution. These residues have implications on screening the highly active BSH/PVA using primary structure of protein from genomic databases and also predicts the oligomeric state of the protein. We propose, in general, there must be two fundamentally different pathway of oligomerization depending upon the intrinsic symmetry of the molecule. In cyclic symmetry, oligomeric proteins exist in two different forms i.e. active (monomer) or inactive (oligomer) or *vice-versa*. While in proteins having dihedral symmetry, can exist in multiple oligomeric state which perhaps, gives plasticity to the enzyme evolution and ability to metabolize or overcome the toxic effect of metabolites which perhaps increases the survival rate of the microorganisms.

3.6 References:

1. Goodsell, David S. "Inside a living cell." *Trends in biochemical sciences* 16 (1991): 203-206.
2. Cornish-Bowden, Athel, and D. E. Koshland. "The Influence of Binding Domains on the Nature of Subunit Interactions in Oligomeric Proteins APPLICATION TO UNUSUAL KINETIC AND BINDING PATTERNS." *Journal of Biological Chemistry* 245.23 (1970): 6241-6250.
3. Bennett, M. J., S. Choe, and David Eisenberg. "Domain swapping: entangling alliances between proteins." *Proceedings of the National Academy of Sciences* 91.8 (1994): 3127-3131.
4. Hashimoto, Kosuke, and Anna R. Panchenko. "Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states." *Proceedings of the National Academy of Sciences* 107.47 (2010): 20352-20357.
5. Crick, Francis HC, and James D. Watson. "Virus structure: general principles." *The Nature of Viruses* 5 (1957): 5-18.
6. Goodsell, David S., and Arthur J. Olson. "Structural symmetry and protein function." *Annual review of biophysics and biomolecular structure* 29.1 (2000): 105-153.
7. Ullmann, A., F. Jacob, and J. Monod. "On the subunit structure of wild-type versus complemented β -galactosidase of *Escherichia coli*." *Journal of molecular biology* 32.1 (1968): 1-13.
8. Ullmann, Agnes, et al. "Determination of molecular weight of proteins and protein subunits in the presence of 6M guanidine hydrochloride." *Biochemistry* 7.1 (1968): 261-265.

9. Kundrot, Craig E., and Philip R. Evans. "Designing an allosterically locked phosphofructokinase." *Biochemistry* 30.6 (1991): 1478-1484.
10. Miller, Maria, et al. "The oligomerization domain of p53: crystal structure of the trigonal form." *FEBS letters* 399.1-2 (1996): 166-170.
11. Pelletier, Joelle N., F-X. Campbell-Valois, and Stephen W. Michnick. "Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments." *Proceedings of the National Academy of Sciences* 95.21 (1998): 12141-12146.
12. Jones, Brian V., et al. "Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome." *Proceedings of the national academy of sciences* 105.36 (2008): 13580-13585.
13. Begley, Máire, Colin Hill, and Cormac GM Gahan. "Bile salt hydrolase activity in probiotics." *Appl. Environ. Microbiol.* 72.3 (2006): 1729-1738.
14. Panigrahi, Priyabrata, et al. "An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholyglycine hydrolase family." *Microbiology* 160.6 (2014): 1162-1174.
15. Rossocha, Maksim, et al. "Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product." *Biochemistry* 44.15 (2005): 5739-5748.
16. Kumar, R. Suresh, et al. "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* 281.43 (2006): 32516-32525.

17. Xu, Fuzhou, et al. "Crystal structure of bile salt hydrolase from *Lactobacillus salivarius*." *Acta Crystallographica Section F: Structural Biology Communications* 72.5 (2016): 376-381.
18. Avinash, Vellore Sunder, et al. "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 (2016): 85-94.
19. Avinash, V. S., et al. "Penicillin V acylase from *Pectobacterium atrosepticum* exhibits high specific activity and unique kinetics." *International journal of biological macromolecules* 79 (2015): 1-7.
20. Hoover, David M., and Jacek Lubkowski. "DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis." *Nucleic acids research* 30.10 (2002): e43-e43.
21. Rossmann, Maxim. *Structural analysis of proteins of human sphingolipid metabolism*. Diss. 2008.
22. Lambert, Jolanda M., et al. "Functional analysis of four bile salt hydrolase and penicillin acylase family members in *Lactobacillus plantarum* WCFS1." *Appl. Environ. Microbiol.* 74.15 (2008): 4719-4726.
23. Gupta, Radhey S. "Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes." *Antonie Van Leeuwenhoek* 100.2 (2011): 171-182.
24. Dougherty, Dennis A. "Cation- π interactions involving aromatic amino acids." *The Journal of nutrition* 137.6 (2007): 1504S-1508S.

CHAPTER IV

Efficacy studies of encapsulated BSH enzyme on serum cholesterol level

4.1 Introduction

Bile helps in excretion of major toxic lipophilic compounds which are not filtered by kidney. One of the major components of bile is Bile salts which serves as a major route for the excretion of cholesterol, and also help in emulsification of dietary lipids. Bile salt exhibit detergent properties such as micelle formation which facilitate the lipid absorption and also reduces the detergent effect of bile salt in biliary epithelium[1,2]. Conjugated, and to lesser extent unconjugated bile acids are reabsorbed from the intestine and modified by liver enzymes to produce tertiary bile acids. This process is known as enterohepatic recirculation[3,4]. SLC family of proteins play important role in active transportation of conjugated bile acid. SLC10A2 codes for apical sodium bile acid transporter (ASBT) also known as IBAT, which efficiently reabsorb >95 % of conjugated bile acids[5,6]. Upon deconjugation the solubility of bile acid decreases and precipitates in intestine and excreted through feces[7]. Reduced concentration of bile acid is sensed by CYP7A1, a member of cytochrome P450 family of enzyme, and starts the conversion of cholesterol to bile acid in liver and thereby reducing the cholesterol level[8,9]. Bile salt Hydrolase (BSH), being active against Bile salts synthesized in liver secretion, helps reduces the cholesterol level by bacterial deconjugation through BSH active gut microbiota[10,11]. Hence the biochemical activity of BSH is of significance in terms of reducing the level of cholesterol. The more active the enzyme is, the more hydrolysis of bile salts will takes place in gut and reduces the reabsorption of deconjugated bile acids via enterohepatic recirculation[7]. This approach has an added advantage over the commercially used drug statins which acts on starting enzyme HMG-CoA reductase[12] of mevalonate pathway unlike *EfBSH* protein formulation which acts on bile acids where cholesterol lowering are viewed as secondary effect.

Previously, it was proposed that the hypocholesterolemic effects of *Lactobacillus acidophilus* and a consortium of *Lactobacillus casei* and *Bifidobacterium longuminipigs* were due to dehydroxylation of bile acid [13]. Further, upregulation of CYP7A1 activity upon feeding of *Lactobacillus plantarum* clearly indicates its effect on reducing bile acid pool [14]. Another study also showed the increase of bile acids pool in humans upon feeding of *Lactobacillus reuteri* within a week[15]. Comparison of the efficacy of different strains of *Lactobacillus* towards BSH activity showed affinity towards glyco- but tauro- conjugated bile acids[16]. Human trials of probiotic formulation in yogurt having *Lactobacillus reuteri* showed reduction in LDL cholesterol[17]. On contrary, the adsorption or colonization of bacterial consortia on the endothelial lining may vary from person to person in same

geographical location and having similar feeding habits[18,19]. In view to this, the cellular probiotic consortium for the development of probiotics, perhaps, limit its use on the basis of feeding habit and geographical location.

Here, we narrowed down our approach to study the enzyme-based formulations. Since BSH is actively involved in attributing probiotic character, we have studied extensively the highly active BSH from *Enterococcus faecalis* in animal model to study its hypocholesterolemic effect.

4.1.1 Need for Enzyme formulation

- ✚ Enzyme based formulation has an advantage over the probiotic formulation
- ✚ Direct action of the enzyme in the colon results quicker response time in cholesterol removal
- ✚ The enzyme-based therapeutics can be used across the globe, irrespective of food habits of an individual.
- ✚ Effectiveness dosage control

4.1.2 Studies carried out-

- ✚ Effect of absence/presence of a cryo-protectant on lyophilized BSH enzyme
- ✚ Comparison of effect of two different cryo-protectants on enzyme activity, stability and storage stability of lyophilized enzyme at 4°C (percent residual activity).
- ✚ Effect of the BSH enzyme-based formulation on Hypercholesterolemic rats

4.2 Materials and methods

*Ef*BSH (purified)[20], Wistar rats (arranged by PRADO), GCA, Ninhydrin reagent, Trichloro Acetic Acid (TCA), Statin Drug (Sigma). All animal related experiments were performed in PRADO Pvt. Ltd (*pradopreclinical.com*).

Methods:

4.2.1 Enzyme preparation: Efficacy of different methods was optimized in this study. Effect of various additives and storage studies was performed using Sucrose and Glycerol for powder-based formulation.

4.2.2 Enzyme encapsulation: Liposomes were prepared to study the efficiency of protein loading and its efficacy in biochemical activity. To prepare the liposomes, equimolar

concentration of Protein and sophorolipid were mixed together and sonicated at 4 °C for 40-50 min at 40% amplitude. Upon sonication, the mixture was centrifuged at 45000 rpm (Figure 4.1). Further, pellet were analyzed for protein loading using SDS-PAGE and biochemical activity using ninhydrin assay[21]

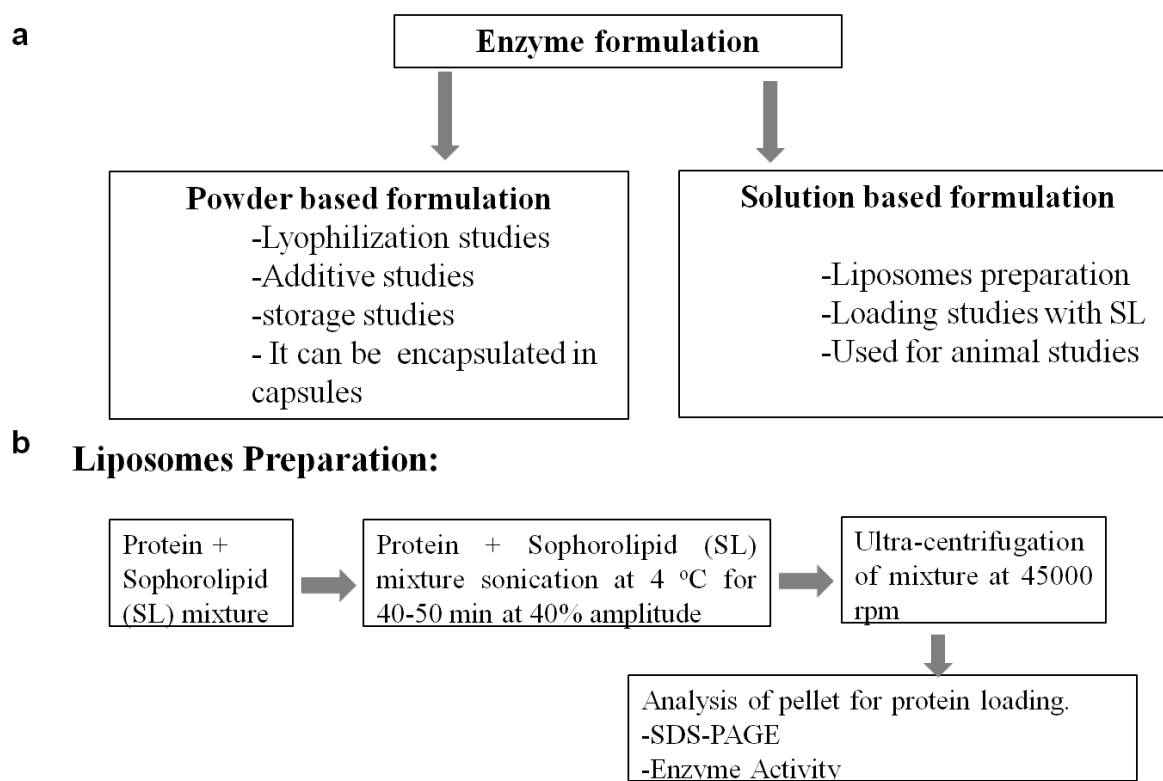


Figure4.1: Preparation of enzyme formulation for animal studies.

4.2.3 *In-vivo* study of enzyme formulation: The wistar rats were selected for the study and rats were fed on high fat diet to gain weight and induce hypercholesterolemia. These animals are further maintained in three different groups. Group1 consist of hypercholesterolemic diet control, Group2 consist of statin control (positive control) and group3 consist of hypercholesterolemic diet and fed with enzyme formulation (Figure 4.2). Each group consists of six rats that were placed in different cage. The animals were kept at room temperature and followed 12 hr day and 12 hr light cycle. Institutional Animal Ethics Committee (IAEC) approved animal experiments conducted as per CPCSEA guidelines. All groups receive solid conventional diet (rodent chow: 32% protein, 5% fat, 2% fibre 60% nitrogen-free extract). Detailed diet chart is shown in Table 4.1

Table 4.1: Composition of high fat diet for introducing hypercholesterolemia in rat experimental model[22].

Ingredients	composition	
	g	kcal
Casein	200	800
Corn starch	155.036	620
Sucrose	50	200
Dextrose	132	528
Cellulose	50	0
Soybean oil	25	225
Lard	175	1575
Mineral mixture	35	0
Vitamin mixture	10	40
TBHQ	0.014	0
DL-Methionine	0	0
L-cystine	3	12
Choline bitartarate	2.5	0
Total	838	4000
Cholesterol	20 g/kg	2%
Cholic acid	5g/kg	0.5%

The rats were forced-fed total 4ml (100ug/ml protein concentration) formulation, twice a day (2 ml morning and 2 ml evening session). Statin drug was used as a positive control (0.3mg/rat/day) and hypercholesterimicrats are used as a negative control. The blood samples were collected at day 0, day 14, day 21 and day 42. Various parameters such as cholesterol level, triglycerides, HDL, LDL, glucose and creatine were checked.

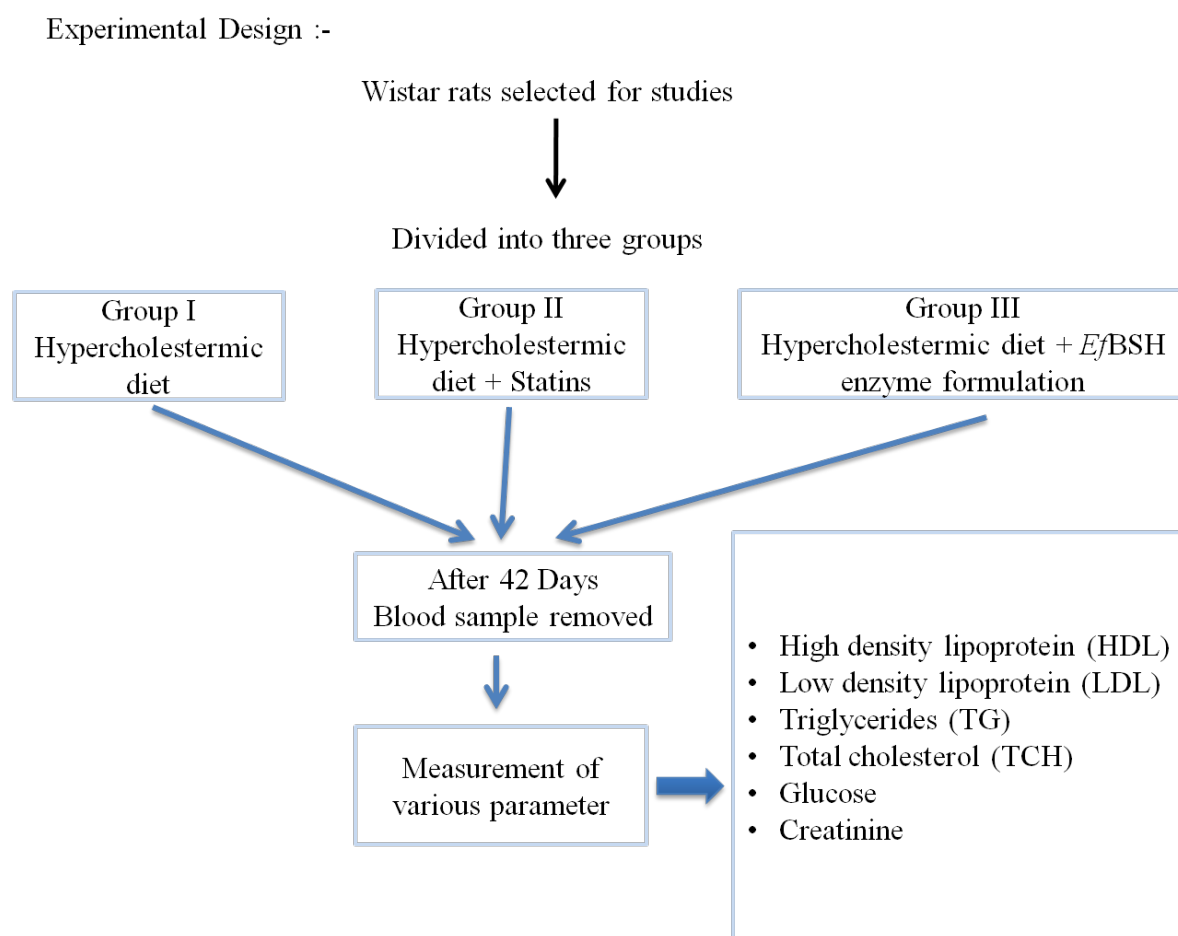


Figure4.2: Experimental design of animal model to study hypocholesterolemic effect.

4.3 Results

4.3.1 Effect of cryo-protectant on lyophilized *EfBSH* enzyme

Biochemical activity was compared to ascertain the good cryogenic agent, before and after lyophilization. Here, we used glycerol and sucrose as a cryogenic agent and checked the efficacy of cryogenic agent after calculating the residual activity at day1 and day7. Furthermore, the effect of lyophilization was also studied. We found that sucrose acts as a good cryogenic agent during lyophilization as it retains maximum activity as compared to enzyme incubated with glycerol, and enzyme alone. Percent residual activity calculated after seven days showed that sucrose retains nearly 81% activity whereas, enzyme alone retains only 2.5% activity and glycerol retains only 14% as compared to wild type (refer Table 4.2).

Table 4.2: Enzyme activity comparison of cryoprotectant used during lyophilization studies of *Ef*BSH stability. Percent residual activity is shown in parenthesis.

Enzyme Activity (IU/mg)	Only Enzyme (E)	E+ Sucrose	E+Glycerol
Before lyophilization			
Residual activity	390.26 (100)	390.26 (100)	390.26 (100)
After lyophilization			
Residual activity (Day1)	278.82(71.44)	350.98(89.93)	124.39(31.87)
Residual activity (Day 7)	9.672 (2.48)	317.18 (81.33)	53.781 (13.79)

E=*Enterococcus faecalis* BSH

4.3.2 Protein loading in liposomes

For animal studies, to keep the protein in solution form for oral administration of desired doses, we maintained the protein in liposomes prepared from sophorolipid (SL). Further qualitative and quantitative estimation was performed to check the protein loading capacity of liposomes. Out of two different ratios viz. 1:1 and 1:5, we found that protein loading is more in lane3 and lane5 which correspond to 1:1 and 1:5 ratio of SL supernatant after ultracentrifugation, respectively (Figure 4.3). The protein loading in liposomes got increased in 1:5 ratio of SL (lane6) as compared to 1:1(lane4) ratio of SL. This result was further confirmed by protein estimation using Bradford reagent. The protein concentration in supernatant is indeed higher than that of liposome pellet which confirms the limiting role of SL in protein loading. Upon increasing the SL concentration i.e.1:5 ratio of SL, the protein concentration increased in liposomes as compared to 1:1 ratio of SL (Figure 4.4). Further biochemical activity of whole enzyme, supernatant and liposome pellet after ultracentrifugation showed the reduction of enzyme activity in supernatant and increase in enzyme activity in liposomes pellet in 1:5 ratio as compared to 1:1 ratio (Figure 4.5).

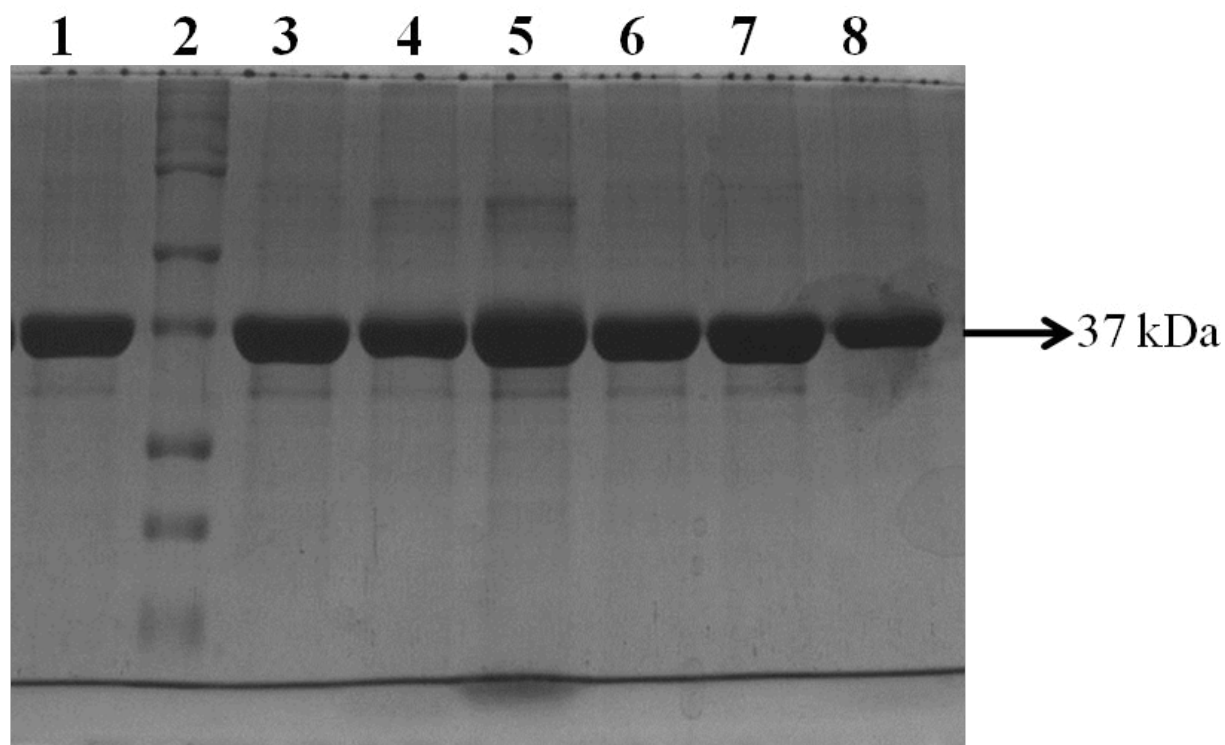


Figure 4.3:Qualitative analysis of Protein (*Ef*BSH) loading in liposomes by SDS-PAGE. Lane-1, BSH Enzyme (EB); 2, Marker; 3, E1(Enzyme +SL ,1:1 ratio) Supernatant after sonication and ultra-centrifuge; 4, E1(Enzyme +SL ,1:1 ratio) Pellet after sonication and ultra-centrifuge; 5, E2(Enzyme +SL ,1:5 ratio) Supernatant after sonication and ultra-centrifuge; 6, E2(Enzyme +SL ,1:5 ratio) Pellet after sonication and ultra-centrifuge; 7, Enzyme after sonication; 8, Enzyme after ultracentrifugation. Equal volume of each sample was loaded in respective well and the protein loading efficiency of sophorolipids liposomes were compared at two different ratios *viz.* 1:1 and 1:5

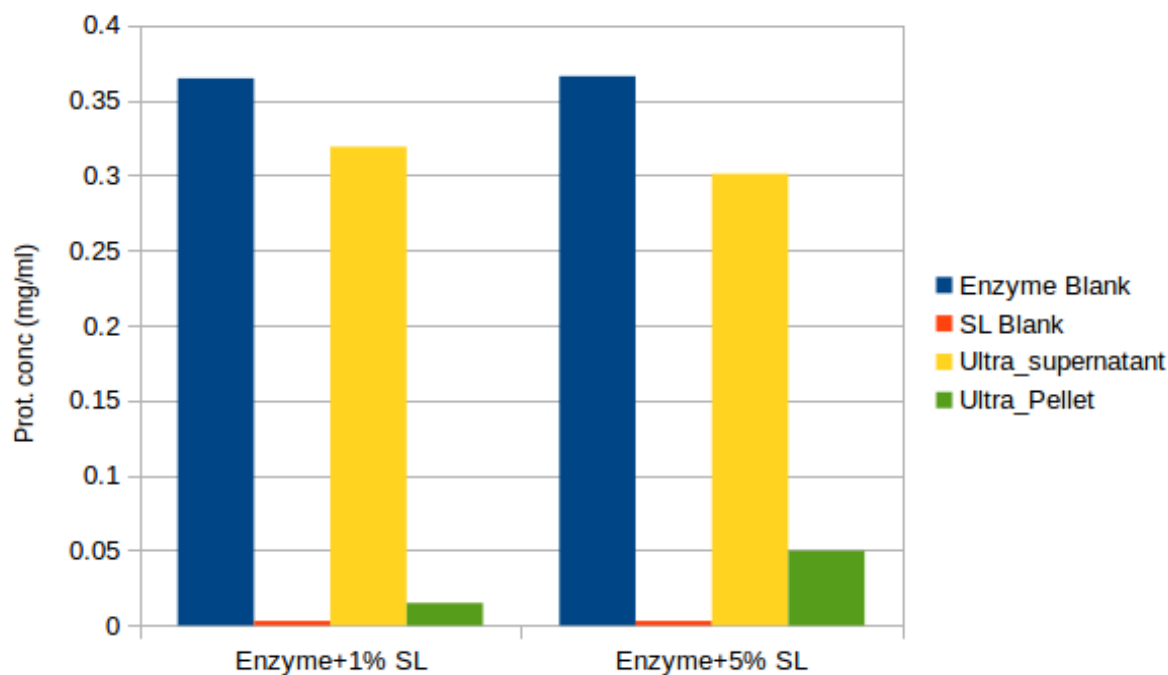


Figure 4.4: Efficiency of *Ef*BSH protein loading in liposomes prepared at two different concentrations i.e. 1% and 5% by protein estimation in pellet and supernatant using Bradford reagent.

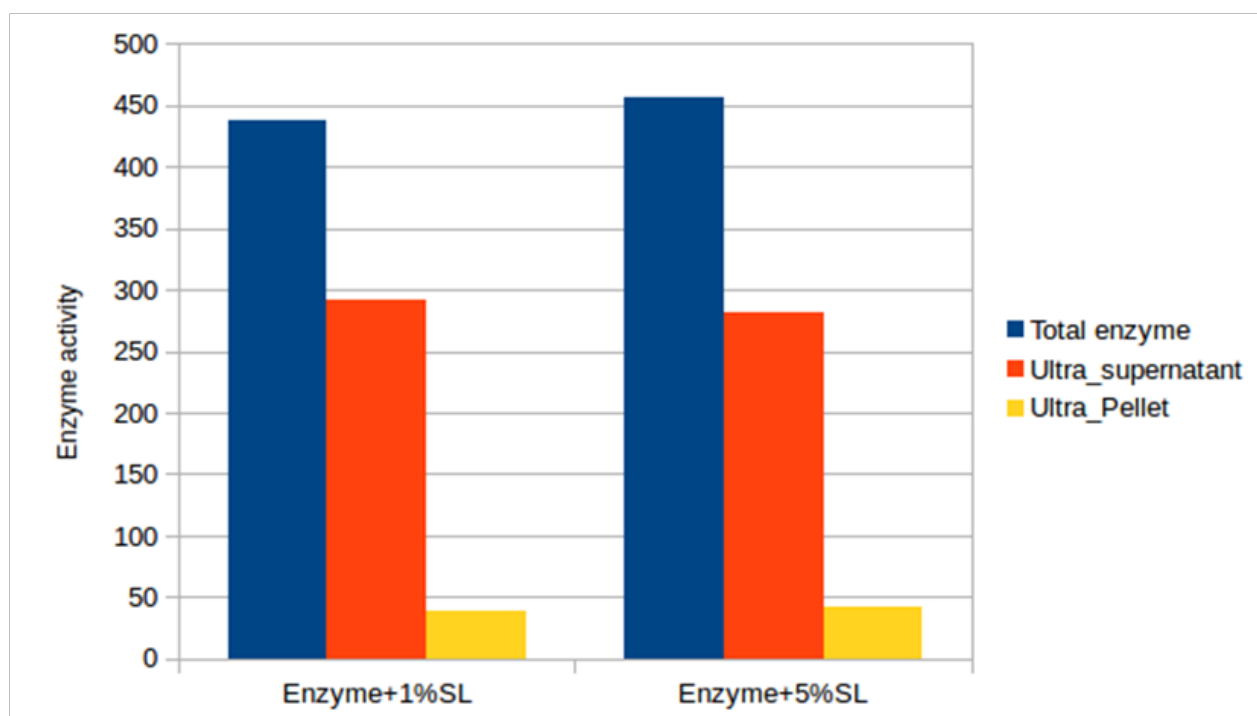


Figure 4.5: Comparison of biochemical activity profile of *Ef*BSH using ninhydrin assay upon protein loading in liposome at two different concentration viz. 1% and 5% SL (SL-Sophorolipid)

4.3.3 Hypocholesterolemic study of *Ef*BSH in animal model

The cholesterol level in all the groups were monitored and compared with statin drug which acts as a positive control and hypercholesterolemic rats were used as a control. The level of cholesterol found to be decreasing on the day 14, day 21, day 42 as compared to day 0 (Figure 4.6). The level of LDL showed decreasing pattern while level of HDL was found to be constant throughout the experiments. The level of glucose and creatine was found to be constant throughout and level of triglyceride decreased significantly on day 42 as compared to day 0 and day 14 (Figure 4.7).

The level of cholesterol and other parameter were measured on day 42 and compared with positive and negative control. The cholesterol level in *Ef*BSH protein formulation-based therapy was decreased and works better or equivalent to that of statin drug control (positive control). The protein-based formulation and the statin drug showed significant reduction in cholesterol as compared to hypercholesterolemic rats within 42 days (Figure 4.8). The level of triglycerides and LDL was reduced in group 2 (statin; positive control) and group 3 (Enzyme formulation) group 3 rats as compared to group 1 rats (Hypercholesterolemic rats; negative control). Glucose, creatine and HDL level was found to be normal and remain unchanged during the experimental period under consideration (Figure 4.7).

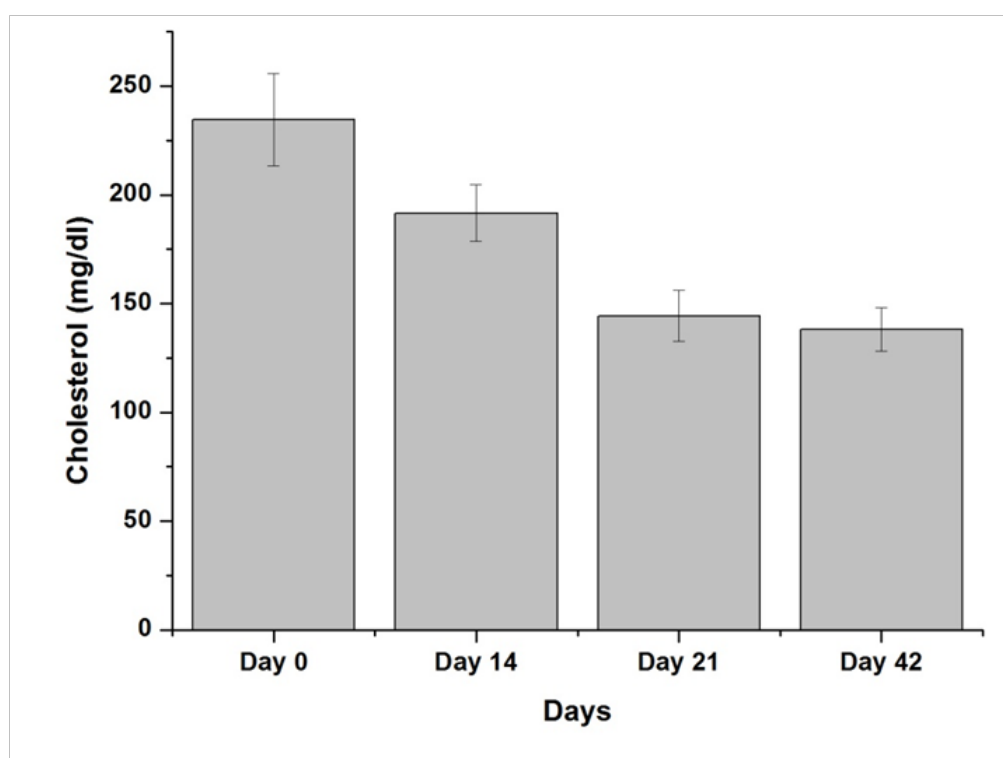


Figure 4.6: Serum cholesterol profile of Group III rats fed on protein formulations.

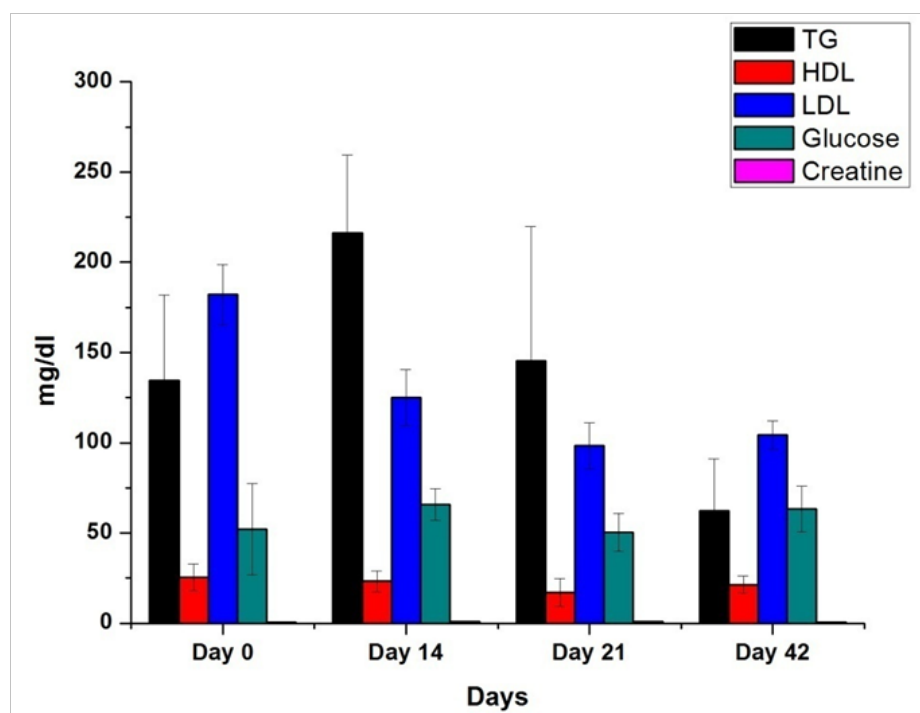


Figure 4.7: Secondary effects profile upon treatment with protein formulation in Group3 rats.

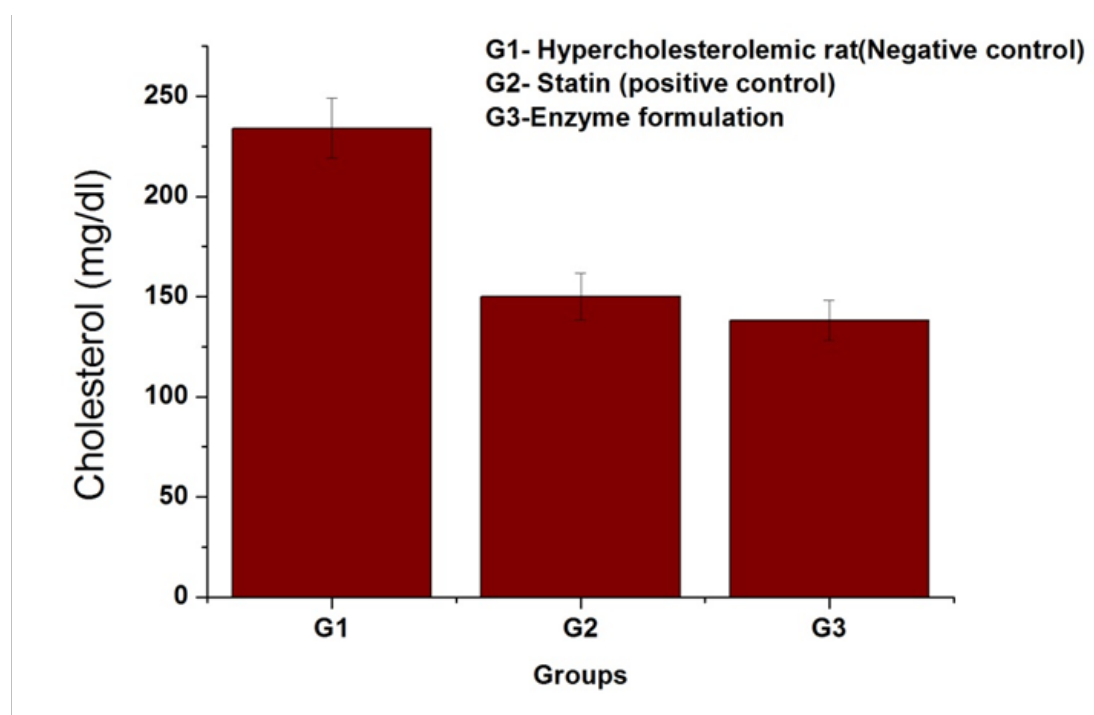


Figure 4.8: Serum cholesterol profile of Group1, Group2 and Group3 rats fed on *Ef*/BSH protein formulations after 42 days.

4.4 Discussion

Cryoprotection and protein loading in liposomes

To develop the *Ef*B₅H as an enzyme therapeutic for hypercholesterolemia, this enzyme needs to be stabilized and biochemically active for longer duration. Here, we studied the effect of two cryo-protectant i.e. sucrose and glycerol on lyophilization of *Ef*B₅H enzyme. Different studies suggest that sugars have been more commonly utilized as cryoprotectant. Use of sugars like low molecular weight malto-dextrin, lactose, sucrose and polymers like polyvinyl pyrrolidone (PVP) are common cryoprotectant added before lyophilization[23]. In our study also, we found that sucrose acts as a good cryoprotectant and compared to glycerol. Percent residual activity calculated after seven days showed that sucrose retains nearly 81% activity whereas, enzyme alone retains only 2.5% activity and glycerol retains only 14% as compared to wild type (Table 4.2). The solution-based formulation was studied to perform the loading of *Ef*B₅H in liposomes. The rationale of designing solution-based formulation was to keep the protein in solution form for oral administration of desired doses in animal model. We prepared liposome from sophorolipids (SLPs), a class of microbial bio-surfactant which is biodegradable and less toxic[24]. Among others, rhamnolipids and trehalolipids are also reported to be the principal amphiphilic compounds used in pharmaceutical preparations[25]. In our study, we used synthesized SLPs derived from jatropha oils, which showed enhanced detergent property with CMC of 9.5 ppm[26]. Two different ratios were used for protein loading studies and we found from the biochemical studies (Figure 4.5) and SDS-PAGE (Figure 4.3) analysis that protein loading is better in 1:5 ratio (Enzyme: SLPs). Determination of protein concentration using Bradford also showed higher concentration of protein in 5% SLPs (Figure 4.4).

In conclusion, we found sugar (sucrose) acts as a good cryoprotectant compared to glycerol. Also, although not efficient, sophorolipids showed loading of *Ef*B₅H. Since SLPs are biosurfactants which consist of hydrophobic fatty acid tail and a hydrophilic sugar head i.e. sophorose[27], different biosurfactant can be screened to give dual effect such as lyophilic attribute owing to sugar moiety and liposome formation because of amphiphilic nature of molecule.

Hypocholesterolemic study of *Ef*BSH in animal model

The wistar rats were selected for the study and divided into three groups i.e. Group1 (Hypercholesterolemic rats) which is negative control, Group2 (Hypercholesterolemic rats with statins) which is positive control and Group3 (Hypercholesterolemic rats with protein formulations) which is a test sample. The level of cholesterol and other parameter were measured on day 42 and compared with positive and negative control. The cholesterol level in *Ef*BSH protein formulation-based therapy showed decreased in serum cholesterol level (Figure 4.6 and 4.8) and worked equivalent to that of statin drug control (positive control) which is used commercially. Although report published by IMS health suggests atorvastatin is the bestselling drug and prescribed globally, these drugs are associated with adverse effects[28]. Statins inhibit HMG-CoA reductase which is the rate limiting enzyme for cholesterol synthesis in mevalonate pathway[12]. This results in the blockage of downstream products and other secondary pathways dependent on the product of mevalonate pathway. Since cholesterol is not the only product of this pathway, the level of other products such as coenzyme Q10, heme-A, and isoprenylated proteins also decreases simultaneously and poses potential risk[29]. Alternatively, the level of cholesterol is also regulated through bile acid synthesis in liver. A cytochrome P450 family of enzyme i.e. CYP7A1 (rate limiting enzyme) plays a key role in initiation of bile acid synthesis[1,2]. Also, the level of bile acids is regulated by enterohepatic recirculation of conjugated bile acids and maintains the bile acid pool[30]. Our approach to enhance the deconjugation activity of bile acids by bile salt hydrolase (BSH) enzymes, perhaps, poses less adverse effect as compared to statin. Various other parameters such as LDL, HDL and triglycerides level were monitored and found to be as effective as statins (Figure 4.7). Moreover, previous report suggests the incidence of diabetes mellitus exhibiting low glycaemic control and elevation in the level of creatine phosphokinase upon treatment with atorvastatin's[31]. In our study, glucose and creatine levels were found to be normal throughout the study (Figure 4.7) which showed that the lowering of cholesterol through deconjugation of bile acid can be the promising future of treatment of hypercholesterolemia and can be adopted as an alternative to well-known prescribed medicine "statins".

4.5 Conclusions

Here we showed the efficacy of *Ej*BSH protein formulation in cholesterol reduction in comparison to well-known marketed drug statins. Moreover, various blood parameter studied during the experimental phase showed less adverse effect of protein formulation as compared to statins. Further, no secondary pathways of sterol metabolism are affected as we are targeting the by-product of cholesterol i.e. bile acids in gut. In case of statins various secondary pathways which depends on cholesterol gets affected such as synthesis of steroid hormones such as sex steroids, corticosteroids and vitamin D which adversely affect the body physiology[32,33].

4.6 References

1. Russell, David W., and Kenneth DR Setchell. "Bile acid biosynthesis." *Biochemistry* 31.20 (1992): 4737-4749.
2. Russell, David W. "The enzymes, regulation, and genetics of bile acid synthesis." *Annual review of biochemistry* 72.1 (2003): 137-174.
3. Houten, Sander, and Johan Auwerx. "The enterohepatic nuclear receptors are major regulators of the enterohepatic circulation of bile salts." *Annals of medicine* 36.7 (2004): 482-491.
4. Hofmann, Alan F. "Enterohepatic circulation of bile acids." *Handbook of Physiology. The Gastrointestinal System* 4 (1989): 567-596.
5. Balakrishnan, Anand, and James E. Polli. "Apical sodium dependent bile acid transporter (ASBT, SLC10A2): a potential prodrug target." *Molecular pharmaceutics* 3.3 (2006): 223-230.
6. Hu, Nien-Jen, et al. "Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT." *Nature* 478.7369 (2011): 408.
7. Zhou, Xiaoming, et al. "Structural basis of the alternating-access mechanism in a bile acid transporter." *Nature* 505.7484 (2014): 569.
8. Ishibashi, Shun, et al. "Disruption of Cholesterol 7 α -Hydroxylase Gene in Mice I. POSTNATAL LETHALITY REVERSED BY BILE ACID AND VITAMIN SUPPLEMENTATION." *Journal of Biological Chemistry* 271.30 (1996): 18017-18023.
9. Schwarz, Margrit, et al. "Disruption of Cholesterol 7 α -Hydroxylase Gene in Mice II. Bile acid deficiency is overcome by induction of Oxysterol 7 α -Hydroxylase." *Journal of Biological Chemistry* 271.30 (1996): 18024-18031.

10. Jones, Brian V., et al. "Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome." *Proceedings of the national academy of sciences* 105.36 (2008): 13580-13585.
11. Begley, Máire, Colin Hill, and Cormac GM Gahan. "Bile salt hydrolase activity in probiotics." *Appl. Environ. Microbiol.* 72.3 (2006): 1729-1738.
12. Istvan, Eva S., and Johann Deisenhofer. "Structural mechanism for statin inhibition of HMG-CoA reductase." *Science* 292.5519 (2001): 1160-1164.
13. Park, Yoo Heon, et al. "Effects of *Lactobacillus acidophilus* 43121 and a mixture of *Lactobacillus casei* and *Bifidobacterium longum* on the serum cholesterol level and fecal sterol excretion in hypercholesterolemia-induced pigs." *Bioscience, biotechnology, and biochemistry* 72.2 (2008): 595-600.
14. Jeun, Jungae, et al. "Hypocholesterolemic effects of *Lactobacillus plantarum* KCTC3928 by increased bile acid excretion in C57BL/6 mice." *Nutrition* 26.3 (2010): 321-330.
15. Martoni, Christopher J., et al. "Changes in bile acids, FGF-19 and sterol absorption in response to bile salt hydrolase active *L. reuteri* NCIMB 30242." *Gut microbes* 6.1 (2015): 57-65.
16. Liong, M. T., and N. P. Shah. "Bile salt deconjugation ability, bile salt hydrolase activity and cholesterol co-precipitation ability of lactobacilli strains." *International Dairy Journal* 15.4 (2005): 391-398.
17. Jones, Mitchell L., et al. "Cholesterol-lowering efficacy of a microencapsulated bile salt hydrolase-active *Lactobacillus reuteri* NCIMB 30242 yoghurt formulation in hypercholesterolaemic adults." *British Journal of Nutrition* 107.10 (2012): 1505-1513.

18. Yatsunenکو, Tanya, et al. "Human gut microbiome viewed across age and geography." *nature* 486.7402 (2012): 222.
19. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome., *Nature*. 486 (2012) 207–14
20. Chand, Deepak, Sureshkumar Ramasamy, and C. G. Suresh. "A highly active bile salt hydrolase from *Enterococcus faecalis* shows positive cooperative kinetics." *Process Biochemistry* 51.2 (2016): 263-269.
21. Kumar, R. Suresh, et al. "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* 281.43 (2006): 32516-32525.
22. Lee, Do Kyung, et al. "Lactic acid bacteria affect serum cholesterol levels, harmful fecal enzyme activity, and fecal water content." *Lipids in Health and Disease* 8.1 (2009): 21.
23. Corveleyn, Sam, and Jean-Paul Remon. "Maltodextrins as lyoprotectants in the lyophilization of a model protein, LDH." *Pharmaceutical research* 13.1 (1996): 146-150.
24. de Oliveira, Marcos Roberto, et al. "sophorolipids a promising biosurfactant and it's applications." *Int J Adv Biotechnol Res* 6 (2015): 161-174.
25. Desai, Jitendra D., and Ibrahim M. Banat. "Microbial production of surfactants and their commercial potential." *Microbiol. Mol. Biol. Rev.* 61.1 (1997): 47-64.
26. Joshi-Navare, Kasturi, Poonam Khanvilkar, and Asmita Prabhune. "Jatropha oil derived sophorolipids: production and characterization as laundry detergent additive." *Biochemistry research international* 2013 (2013).

-
27. Davila, Anne-Marie, Rémy Marchal, and Jean-Paul Vandecasteele. "Sopporose lipid production from lipidic precursors: predictive evaluation of industrial substrates." *Journal of industrial microbiology* 13.4 (1994): 249-257.
 28. Golomb, Beatrice A., and Marcella A. Evans. "Statin adverse effects." *American Journal of Cardiovascular Drugs* 8.6 (2008): 373-418.
 29. Buhaescu, Irina, and Hassane Izzedine. "Mevalonate pathway: a review of clinical and therapeutical implications." *Clinical biochemistry* 40.9-10 (2007): 575-584.
 30. Thomas, Charles, et al. "Targeting bile-acid signalling for metabolic diseases." *Nature reviews Drug discovery* 7.8 (2008): 678.
 31. Sasaki, Jun, Mikio Iwashita, and Suminori Kono. "Statins: beneficial or adverse for glucose metabolism." *Journal of atherosclerosis and thrombosis* 13.3 (2006): 123-129.
 32. Mol, Marc JTM, et al. "Effects of inhibition of cholesterol synthesis by simvastatin on the production of adrenocortical steroid hormones and ACTH." *Clinical endocrinology* 31.6 (1989): 679-690.
 33. Hyyppä, Markku T., et al. "Does simvastatin affect mood and steroid hormone levels in hypercholesterolemic men? A randomized double-blind trial." *Psychoneuroendocrinology* 28.2 (2003): 181-194.

CHAPTER V

Summary and conclusion

Summary and conclusions

Bile salt hydrolases (BSH) are the enzymes that catalyzes the hydrolysis of glycine and taurine-conjugated bile salts into amino acid residues and free bile salts. Upon deconjugation the solubility of bile acids decreases and precipitates in the colon and excreted through faeces. Bile acids are synthesised from cholesterol and further bile acids pool maintained through enterohepatic recirculation. Deconjugating the bile acid in human gut reduces its affinity towards active transporters present in the gut i.e. Apical sodium Bile Acid (ASBT) transporter and have implication in cholesterol homeostasis. The rationale behind our study is to understand the effect of microbiome enzyme (BSH) on host metabolism with respect to cholesterol regulation and their application in development of biotherapeutics. Also, to screen novel Bile Salt Hydrolase (BSH) for understanding higher activity and thermostability. Here, we investigated the molecular characteristics for the high activity of *Enterococcus faecalis* Bile Salt Hydrolase (*Ef*BSH) through structural and biochemical analysis of *Ef*BSH mutants to understand the mechanism of cooperativity in *Ef*BSH.

We found that Arg207 of tetramer loop protrudes inside the active site of other chain and influences the active site residues. Mutation of the Arg207 interacting residues i.e. Glu269 to alanine showed drastic decrease in biochemical activity. Glu269 when mutated to Aspartic acid loses the activity while when mutated to Glutamine, found to be active against bile acids. Structural analysis showed that in Glu269 to alanine i.e. E269A mutant, Arg207 is hydrogen bonded to Glu21 of loop1 and constrain the movement or flexibility of loop in all the chains. Due to this, movement of Tyr20 of loop1, which aligns the substrates in the active sites, is also blocked. In R207A mutant, the Glu21 sets free in all different chains and do not interact with E269 of loop4. Superposition of Tyr20 of all four chain in E269A and R207 mutant showed very little deviation as compared to wild type. Further, superposition of four chains of E269A mutant with four chains of R207A mutant revealed two extreme arrangement possible for Tyr20. Moreover, in wild type, Tyr20 samples the conformation between the two extreme conformations possible as in E269A and R207A. Subsequently, the corresponding amino acids of other reported BSH structure were analyzed and we found that in *Bf*BSH (PDB Id: 2hez) tyrosine is replaced by tryptophan which closes the active site and perhaps reduces the accessibility of substrates and hence activity. Therefore, Tyr20 occupies very strategic location near the active site and we found that Arg207 regulates the movement of this tyrosine by interacting with Glu21 and Glu269.

To understand the principle of oligomerization in CGH family, the interface of previously reported structure and also performed structural and biochemical characterization *Dorea longicatena* BSH (*Dl*BSH). Oligomerization study was performed in greater detail with *Dl*BSH, in which several mutations has been created using Site directed mutagenesis approach. Changes in oligomerization were observed upon deletion mutation in tetramer loop region. Removing the tetramer loop (δ 189-209) showed complete shift to dimer state. In gram negative organisms tetramer BSH/PVA are stable without assembly loop while in Gram positive organisms removing the tetramer loop changes the oligomerization profile. Moreover, gram negative BSH/PVA structure showed non-planar tetramer. Analyzing the interface residue of gram positive and gram negative showed that both the classes evolved through unique mechanism of holding the tetramer. We identified, upon interface comparison of dimer and tetramer BSHs, Tyr205 in tetramer loop region to be essential in oligomerization. Similarly, in Gram negative *Pa*PVA structure, showed presence of Met205 in tetramer interface present in shortened tetramer loop. This residue, perhaps crucial in holding the miniclusters present at tetramer interface and also serves as contact point of residues from all four chains of tetramer. This in part explains why gram-negative structures displayed non-planar tetramer structure and perhaps the modest way of evolution. These residues have implications on screening the highly active BSH/PVA using primary structure of protein from genomic databases and also predicts the oligomeric state of the protein. We propose, in general, there must be two fundamentally different pathway of oligomerization depending upon the intrinsic symmetry of the molecule. In cyclic symmetry, oligomeric proteins exist in two different forms i.e. active (monomer) or inactive (oligomer) or *vice-versa*. While in proteins having dihedral symmetry, can exist in multiple oligomeric state which perhaps, gives plasticity to the enzyme evolution and ability to metabolize or overcome the toxic effect of metabolites which perhaps increases the survival rate of the microorganisms.

Here we showed the efficacy of EfBSH protein formulation in cholesterol reduction in comparison to well-known marketed drug statins. Moreover, various blood parameters studied during the experimental phase showed less adverse effect of protein formulation as compared to statins. Further, no secondary pathways are affected as we are targeting the by-product of cholesterol that is bile acids in gut. In case of statins various secondary pathways which depends on cholesterol gets affected such as synthesis of steroid hormones such as sex steroids, corticosteroids and vitamin D, which adversely affect the body physiology.

List of publications

1. **Yashpal Yadav**, Mrityunjay K Tiwari, Deepak Chand, Archana V Pundle and Sureshkumar Ramasamy. Dissection of Catalytic Site in Crucial Gut Microbiome Enzyme: Bile Salt Hydrolase. (Under Review)
2. **Yaspal Yadav** and Sureshkumar Ramasamy. Oligomeric attribute in cholyglycine hydrolase family. (manuscript under preparation)
3. Pushparani D Philem*, **Yashpal Yadav***, Avinash V Sunder, Deepanjan Ghosh, Asmita Prabhune, Sureshkumar Ramasamy. Expanding AHL Acylases Horizon - Insights from Structural Analysis of Choloylglycine Hydrolases from *Shewanella loihica* PV-4. (Communicated). *Equal contribution
4. **Yashpal Yadav** and Sureshkumar Ramasamy. Allostery mapping in Enterococcus faecalis bile salt hydrolase (BSH). **Acta Crystallographica Section A: Foundations and Advances**. 2017, 73:C274-C274.
5. BhideAmey, Sonal M. Channale, **Yashpal Yadav**, Kabita Bhattacharjee, Pankaj K. Pawar, V. L. Maheshwari, Vidya S. Gupta, Sureshkumar Ramasamy, and Ashok Giri. Genomic and functional characterization of coleopteran insect-specific α -amylase inhibitor gene from Amaranthus species. **Plant Molecular Biology**. 2017, 94: 1-14.
6. Chand Deepak, 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. **Biochimica et Biophysica Acta (BBA)-General Subjects**. 2017, 1861: 2981-2991.
7. Sonalchannale, AmeyBhide, **Yashpal Yadav**, Garima Kashyap, Pankaj K. Pawar, V.L. Maheshwari, Sureshkumar Ramasamy, Ashok Giri. Characterization of two coleopteran α -amylases and molecular insights into their differential inhibition by synthetic α -amylase inhibitor, acarbose. **Insect biochemistry and molecular biology** 2016, 74: 1-11.