The comparative structural and functional studies of acid ceramidase and acid ceramidase-like proteins using biophysical and computational techniques

Thesis Submitted to AcSIR
For The Degree of
DOCTOR OF PHILOSOPHY

In
Biological Sciences

By
Rajput Vijay Bhagchand
Registration Number: 10BB13A26036

Under the guidance of

Guide
Dr. M. Karthikeyan

Co-Guide
Dr. Sureshkumar Ramasamy

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

JULY 2019

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 **The comparative structural and functional studies of acid ceramidase and acid ceramidase-like protein using biophysical and computational techniques** submitted by Mr. **Rajput Vijay Bhagchand** to Academy of Scientific and Innovative Research (AcSIR) in fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy**, embodies original research work under my/our supervision/guidance. We further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table etc., used in the thesis from other sources, have been duly cited and acknowledged.

It is also certified that this work done by the student under my supervision, is plagiarism free.

Rajput Vijay Bhagchand
(Research Student)

Dr. M. Karthikeyan
(Research Supervisor)

Dr. Sureshkumar Ramasamy
(Research Co-Supervisor)

Place: Pune
Date: 01 JULY 2019

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 **"The comparative structural and functional studies of acid ceramidase-like proteins using biophysical and computational techniques"**, 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. M. Karthikeyan (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.

Rajput

Rajput Vijay Bhagchand

Dedicated to my beloved daughter Vedika…

Acknowledgement

I have been inspired and encouraged by numerous people who have helped me in innumerable ways to reach at this stage. There is no doubt that this list is incomplete.

First and foremost, I would like to express my sincere gratitude towards my present research guide Dr. M. Karthikeyan for his constant support and guidance during this work. He has been my inspiration for the in silico work included in my thesis.

I would like to thank my supervisor, Dr. C. G. Suresh and Dr. Archana V. Pundle for the concern they showed regarding the progress of my work. I am certain that the ethics and moral values which I learnt from them will go a long way in making me a better human being.

I take this opportunity to gratefully acknowledge my co-guide Dr. Sureshkumar Ramasamy for his motivation, valuable suggestions, constructive criticism and never-ending support and wholehearted interest during the course of the investigation. I am extremely thankful to him for his invaluable guidance and everlasting encouragement. His enthusiastic attitude, innovative ideas, and scientific knowledge have inspired me profoundly. I truly feel privileged to have joined his research group, and I shall cherish my learning experience under him.

I am grateful to the members of my doctoral advisory committee, Dr. A. K. Bhattacharya, Dr. Dhanasekaran Shanmugam and Dr. Subhaschandrabose Chinnathambi, for their regular assessment of my work, insightful comments and constant encouragement.

I also thank Dr Kavita Babu from IISER Mohali for providing C. elegans and Dr. Mahendra Sonawane from TIFR, Mumbai for providing cDNA of zebrafish.

I am thankful for the continuous support and help by Dr. Subhaschandrabose Chinnathambi and his group, particularly, Shweta, Nalini, Abhishek and Tushar.

I feel indebted to Dr. Sushama Gaikwad and Sanskruti for their help and guidance in CD data analysis. I am also grateful to Dr. V. Koteswara Rao for his help.

I am grateful to my seniors and present lab members Dr. Nishant, Dr. Manas, Dr. Tulika, Dr. Ranu, Dr. Priyabrata, Dr. Ruby, Dr. Deepak, Dr. Manu, Dr. Ameya, Dr. Yashpal, Deepanjan, Shridhar, Tejashri, Aditi Shiva, and Debjyoti for providing me an awesome work environment in the lab.

I would like to thank Deepanjan and Debjyoti for their useful contribution in shaping all manuscripts and this thesis in its entirety. Special credits should go to Dr. Yashpal, Dr. Deepak, Dr. Priyabrata and Dr. Nishant for their fruitful discussions and inputs regarding the Ntn Hydrolase superfamily. I would also like to acknowledge Dr. Ruby, Shridhar and Teju for their support during tough times. The list wouldn't be complete without thanking Dr. Manu and Shiva who diligently helped me for protein expression in E.coli initially.

I am also grateful to Dr. Sandeep who helped me in maintaining C. elegans cultures, Dr. Prabhakar, Dr. Atul, Dr. Krithika, Dr. Avinash , Dr. Dipesh, Dr. Ameya , Dr. Rubina, Dr. Yogesh, Dr. Gajanan, Dr. Ejaj, Dr. Ekta, Vijashree, Uttara, Yugendra, Monika, Veena, Nidhi, Amit, Atish,Vishwa and Amol.

A special thanks to Deepanjan Ghosh for helping me with almost each and every problem, I faced during my Ph.D. tenure.

Special thanks to my dear hostelmates, Veer, Pravin, Sandeep, Deepak, Roby, Salim, Indrapal and Dilip for being just themselves. Not to menion, my roommates Amol, Sachin, Nitin, Shivaji and Krishna and my batchmates Anand, Shakuntala, Bhagyashree, Sneha, Zenia, Shrikant and Deepak, who have been with me through thick and thin during my Ph.D. tenure.

My sincere thanks to Dr. Arun Kharat and Dr. Jitendra Kulkarani for their motivation and support.

I want to thank my former colleagues Dr. Savita, Dr. Yoginee, Dr. Suchitra, Rohinee, Anupriya, Fara, Areeb and Suraj.

I also want to thank my friends Jitu, Pankaj, Swapnil, Prshant, Siddheshwar, Sudharshan, Aannata, Raghu, Ram, Vaibhav, Sagar, Mayur, Sandeep, Rahul and Kiran.

I am thankful to all the members of Biochemical Sciences Division for their support. I am also grateful to the SAC and AcSIR office staff of CSIR-NCL.

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

I would like to thank my entire family for supporting me unconditionally throughout my PhD. Finally, I would like to acknowledge my life partner Manisha Rajput. She has been a constant source of strength and inspiration throughout this journey.

Rajput Vijay

Table of Contents

Abbreviations

aa	Amino acid
AC	Acid ceramidase
ASAH1	N-acylsphingosine amidohydrolase
AU	Absorbance unit
BMGY	Buffered Glycerol-complex medium
BMMY	Buffered Methanol-complex medium
bp	base pair
BSH	Bile salt hydrolase
CD	Circular dichroism
CGH	Cholylglycine hydrolase
CPB	Citrate phosphate buffer
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
***Ef*BSH**	*Enterococcus faecalis* bile salt hydrolase
FAAH	Fatty acid amide hydrolase
FD	Farber disease
h	Hill coefficient
hAC	Human acid ceramidase
hNAAA	Human N-acylethanolamine-hydrolyzing acid amidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MRE	Mean residue ellipticity
NAAA	N-acylethanolamine-hydrolyzing acid amidase
Ni-NTA	Ni^{2+}-nikel-nitrilotriacetic acid
ns	Nanoseconds
Ntn	N-terminal nucleophile
OD	Optical density
PCR	Polymerase chain reaction
PDB	Protein data bank
PGA	Penicilin G acylase
PME	Progressive myoclonic epilepsy
ps	Picosecond
PVA	Penicillin V acylase
RMSD	Root mean square deviation
RMSF	Root mean square fluctations
ROG	Radius of gyration
rpm	Revolutions per minute
SASA	Solvent-accessible surface
SDS	Sodium dodecyl sulfate
SMA	Spinal muscular atrophy
WT	Wild type

List of Tables

List of Figures

Abstract

We report here the extensive work carried out on two essential lysosomal enzymes acid ceramidase (AC, EC 3.5.1.23) from zebrafish (*Danio rerio*), *Caenorhabditis elegans* and acid ceramidase-like protein (N-acylethanolamine hydrolyzing acid amidase (NAAA), EC 3.5.1.4) from mouse (*Mus musculus*) and *Micromonas commoda (strain RCC299) (Picoplanktonic green alga.* These enzymes belong to the N-terminal nucleophile hydrolase superfamily, which is functionally identified as amidases. Based on the presence of N-terminal residue Cys, which acts as a nucleophile, the superfamily classified into N-terminal cysteine nucleophile (NtCn) hydrolases. Being members of Ntn-hydrolase superfamily ACs and NAAAs share a common (αββα) Ntn structural fold and similar catalytic mechanism. Both enzymes are heterodimers and form homodimer assemblies of heterodimer assemblies. AC catalyzes the degradation of ceramide by hydrolyzing the amide bond in ceramide into sphingosine and free fatty acid. NAAA catalyzes the degradation of N-Acylethanolamine (NAE) by hydrolyzing the amide bond in NAE into ethanolamine and free fatty acid. AC and NAAA are glycoprotein and undergo post-translational processing to mature as an active form.

The detailed work presented in the thesis underpins the biochemical and structural characterization along with computational analysis of AC and NAAA.
The thesis is organized into five chapters:

Chapter 1: Introduction

The general introduction and provides a review of the literature and present status of the work on AC and NAAA. This chapter also discusses the comparative account of AC and NAAA with respect to function and structure. It covers the biochemically and structurally characterized AC and NAAA from various sources. Their probable role in human physiology has also been emphasized.

Chapter 2: Characterization of acid ceramidase (AC)

Chapter 2 presents the cloning, expression, purification, biochemical and biophysical characterization of AC. Zebrafish asah1b and *C. elegans.* Asah1 gene was cloned in pPICZα-A and heterologously expressed in *Pichia pastoris*. AC was purified through affinity and Size

exclusion chromatography. Glycosylation analysis was done by a very reliable prediction algorithm. The confirmation of glycosylation was based on glycoprotein staining and deglycosylation of AC. Demonstrated insights into the autocatalytic processing of AC. Zebrafish AC was found to be active towards the C12-NBD-Ceramide, another unique feature of enzyme reverse activity was also observed. The optimum pH, temperature, time and DTT concentration of zebrafish AC has been determined. Kinetic parameters of zebrafish AC catalyzed reaction were also determined. Site-directed mutagenesis of perceived mutations of SMA-PME has been designed to explore the important residues involved in stability and activity and autocatalytic processing of zebrafish AC. Biophysical techniques circular dichroism was used for the structural transitions studies of zebrafish AC by monitored by subjecting the enzyme to thermal denaturation.

Chapter 3: Characterization of acid ceramidase-like protein (NAAA)

Chapter 3 presents the cloning, expression, purification, biochemical and biophysical characterization of NAAA. Mouse and *M. commoda* NAAA gene was cloned in pPICZα-A and heterologously expressed in *Pichia pastoris*. The expression was confirmed by western blot. AC was purified through affinity, Size exclusion chromatography and Concanavalin A-sepharose chromatography. Glycosylation analysis was done by prediction and confirmation based on glycoprotein staining and deglycosylation of NAAA. Western blotting was also used to insights into the autocatalytic processing of NAAA. Mouse NAAA was found to be active towards the PEA, another unique feature of enzyme ceramidase hydrolase activity was also observed. The optimum pH, temperature, time and DTT concentration of mouse NAAA determined. Kinetic parameters of mouse NAAA catalyzed reaction were also determined. Biophysical techniques like fluorescence, differential scanning fluorimetry and circular dichroism were used for the studies. Through the enzyme thermal denaturation studies, the structural transition of zebrafish AC was studied.

Chapter 4: In silico study of acid ceramidase (AC) and acid ceramidase-like protein (NAAA)

In this chapter, the homology models of AC and NAAA generated and zebrafish AC docked with substrates ceramide. The prediction study was most damaging nsSNPs done by analysis deleterious and destabilizing nsSNPs. *In silico* mutagenesis were also generated to explore important of the structural effect of Farber and SMA-PME mutation on function. MD simulations studies of WT and disease mutant models were executed to understand the

molecular mechanism of Farber and SMA-PME disease. Based on the observations made from the above studies, 7 mutations (**L182V, R226P, G235R, G235D, R254G, R333H and P362R)** were identified as highly deleterious from 30 disease causing nsSNPs. The four substrate binding pocket loops mapped on the structure of AC and NAAA. AC and NAAA separated based on the difference in length of loop1 by phylogenetic analysis of the sequences.

Chapter 5: Summary and conclusions

This chapter highlights the key findings from the present work and conclusions drawn from the data are discussed in this chapter.

Chapter 1: Introduction

1.1 Enzyme and their importance for life

Metabolism is a fundamental process of life in which organic compounds are synthesized (anabolism) and breakdown (catabolism) by stepwise reaction using the help of the enzymes. Many organisms break down carbohydrates, lipids and proteins to obtain energy and precursor for biosynthesis, while autotrophs satisfy their needs completely with inorganic materials. The catalysts of metabolism are enzymes and nearly all enzymes are proteins but some are nonprotein in nature (catalytic RNA). The discovery of the first enzyme amylase (Payen, A. & Persoz, 1833) triggered the development of fermentation technology in the 20th century. Enzyme urease isolated and crystallized proves that enzyme nature is protein (Sumner, 1926). During the 1960s, a series of experiments conducted by Anfinsen on ribonuclease shed light that the primary structure of the protein determines the final conformation of the protein.

The function of the enzyme generally depends on the structure of the enzyme. Proteins are macromolecules made up from standard 20 amino acids residues which are linked together by a peptide bond. Protein diversity arises from different compositions and arrangement of the 20 amino acids. The protein diversity information is stored in the genetic material in the form of nucleotides. The primary structure of the enzyme is linear sequences of amino acid. The regular periodic secondary structural element (α-helices, β-sheet and loop) formation depends on the pattern of hydrogen bonds between the amino group and the carbonyl group of amino acid residues. The secondary structure further folds and form a three-dimension structure called a tertiary structure which is functionally active. Sometimes enzymes are made up of more than one polypeptide chain, which assembles in the form of quaternary structure.

Enzymes are highly specific and enhance the rate of reaction by the formation of the transition state (Pauling, 1946). Enzymes in fine quantities are adequate enough to convert the bulk amount of substrates to products. Some enzymes require the presences of cofactors for activity. Such cofactors can be metal ions, organic molecules (coenzymes). Endoenzymes are synthesized by living cells and remain inside the cells. On the other hand, exoenzymes leave the cells and function outside. Every enzyme has a specific optimum pH and optimum

temperature. Any shift from optimum pH and temperature results in a decrease in enzyme activity because it denatures the enzyme molecule.

Enzymes being large molecules consist of many hundreds of amino acids but the functional regions of an enzyme (generally the active site) consists about ten to twelve amino acids (Gutteridge and Thornton, 2005). The amino acids histidine, cysteine and aspartate are more frequently involved in catalytic activities, whereas aliphatic amino acids such as alanine, leucine and glycine are hardly ever involved (Holliday, Mitchell and Thornton, 2009). The active site residues can identify by chemical labeling (Aktories, 1997) and site-directed mutagenesis (Morrison and Weiss, 2001). *In silico* prediction of function of an enzyme from its sequence and structure, is possible with the help of structural genomics (Burley *et al.*, 1999; Shapiro and Harris, 2000). The comparative study of sequences and structures of the enzyme are helpful to extract the evolutionary relationship between organisms (Gabaldón and Koonin, 2013). Sometimes function prediction from sequences fails due to the fact that even proteins with high sequence similarity have been seen to perform quite different functions (Karp, 1998).

1.2 Ceramidase

Ceramidase (EC 3.5.1.23, acylsphingosine deacylase, glycosphingolipid ceramide deacylase) are amidohydrolase enzyme that catalyzes the cleavage of ceramide into free fatty acids and sphingosine (SPH). Ceramides are fatty acid derivatives that are amide-linked to a sphingosine moiety. They can differ in the lengths of their acyl chains and even in degree of unsaturation. Ceramidases are classified into three groups, acid (ASAH1), neutral (ASAH2, ASAH2B and ASAH2C) and alkaline (ACER1, ACER2 and ACER3) according to their optimum pH (Gangoiti *et al.*, 2010).

Acid ceramidases (AC; N-acylsphingosine deacylase, EC 3.5.1.23) are autoproteolytic heterodimeric glycoproteins which localizes in lysosomes. AC activity was first identified in rat brain (GATT, 1963), but was purified and characterized from human urine (Bernardo *et al.*, 1995). The first antibody produced against AC using urine AC was further used for the study of sphingolipid metabolism in skin fibroblast (Chen, Moser and Moser, 1981). AC is a key enzyme in sphingolipid metabolism and degrades ceramide into sphingosine and free fatty acid (GATT, 1963; Zeidan and Hannun, 2007). Deficiency or imbalance of enzyme causes defects in levels of sphingolipid metabolite which interferes in the sphingolipid-mediated signal transduction, responsible for diseases (Park and Schuchman, 2006). Since

AC is overexpressed in cancers, it is a potential target of treatment in cancer disease (Zeidan *et al.*, 2008). AC is functionally similar to other ceramidases but structurally more similar to NAAA (Figure 1.1). AC has different characteristic features compare to other ceramidases such as optimum pH, substrates preferences, primary structure and cellular location.

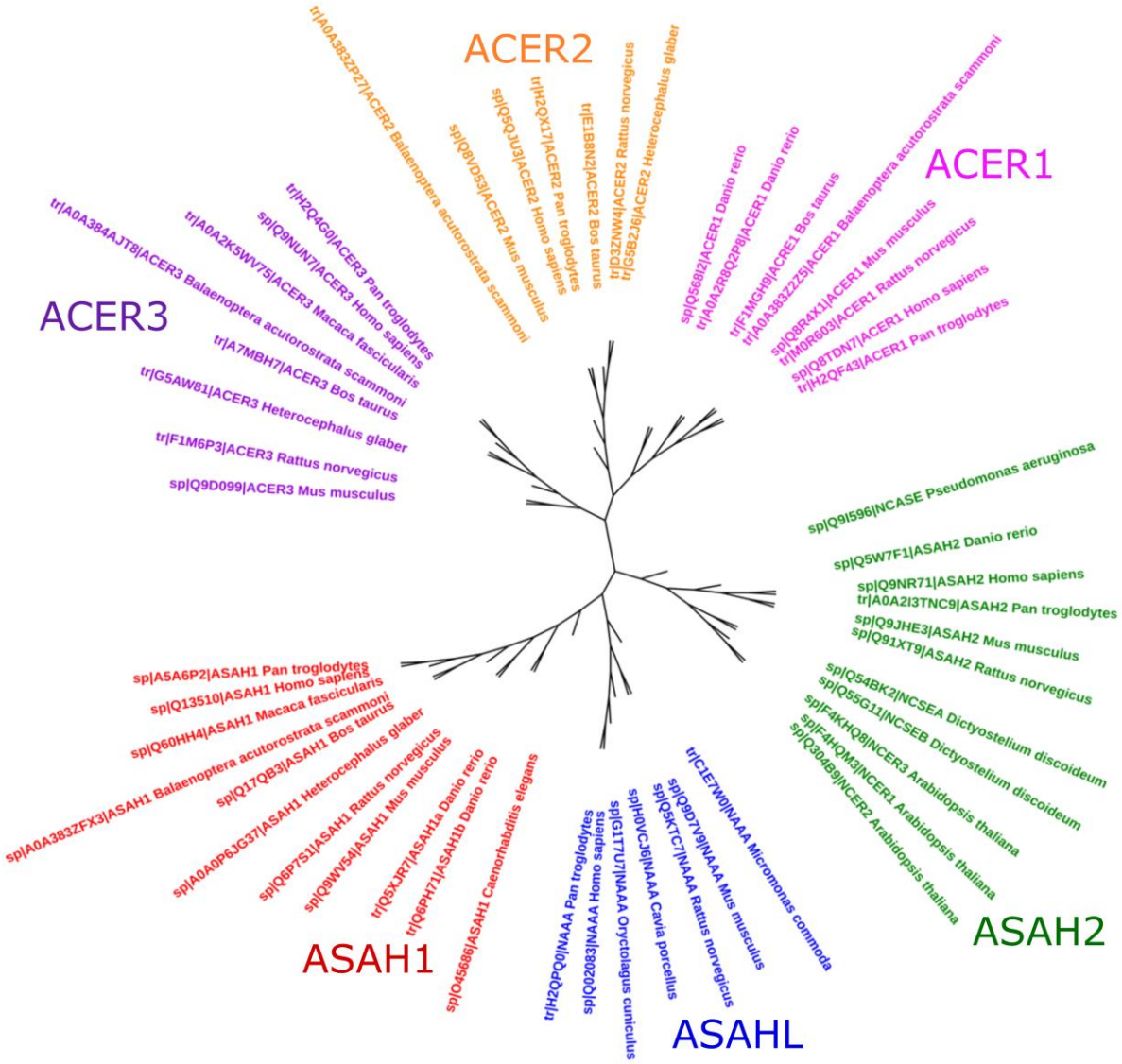


Figure 1.1: Phylogenetic analysis of ceramidase. The red branches indicate acid ceramidases, blue branches indicate acid ceramidase-like proteins, green branches indicate neutral ceramidases, magenta, orangeand violet branches indicate alkaline ceramidases (ACER1, ACER2 and ACER3). The tree was prepared in MEGA 6 and plotted using iTOL.

Acid ceramidase-like protein (N-acylethanolamine-hydrolyzing acid amidase; EC 3.5.1.4) gene was first cloned from human placenta having homology (33-35% amino acid identity) to AC but showed no ceramidase activity (Hong *et al.*, 1999). The term acid

ceramidase-like protein (ASAHL) was coined by Hong et al. because of the sequences similarity with AC. Later, a similar protein was purified from rat lung and it was found that similar genes were present in human and mouse having sequence similarity with ASAHL (Tsuboi *et al.*, 2005). The naaa cloned from human megakaryoblastic (CMK) cells has 99.8% identity to the previously cloned acid ceramidase-like protein from human placenta. Tsuboi et al., changed the protein name to N-Acylethanolamine-hydrolyzing acid amidase (NAAA) because of its functional similarity with fatty acid amide hydrolase (FAAH; EC 3.5.1.99). Bioactive NAE is hydrolyzed to fatty acids and ethanolamine by the catalysis of FAAH (Ueda *et al.*, 2000). NAEs are fatty acid derivatives that are acyl chains bound to ethanolamine by an amide bond. Variation is observed in acyl chain length and the occurrence of double bonds. FAAH is a member of the amidase signature superfamily constituting of a Ser-Ser-Lys triad at the catalytic site (Bracey, 2002). NAAA catalyzes the hydrolysis of many NAE into ethanolamine and free fatty acid but has more activity against N-palmitoylethanolamine (PEA). NAAA catalyzes the hydrolysis of PEA into ethanolamine and palmitic acid. The NAAA is a potential target of treatment for inflammatory disorders due to the anti-inflammatory nature of its substrate, PEA. NAAA is functionally similar to FAAA but structurally more similar to AC.

1.3 Ntn (N-terminal nucleophile)

AC and NAAA sequences revealed high homology with the bile salt hydrolase (BSH) and penicillin V acylase (PVA), suggesting that AC and NAAA is a member of the chologlycin hydrolase family (CGH) (PFAM PF02275). The CGH is a part of the superfamily N-terminal nucleophile (Ntn) hydrolase. (Rossocha *et al.*, 2005; Tsuboi, Takezaki and Ueda, 2007). Ntn hydrolases functionally belong to hydrolase class of enzymes but most of them are amidases (Brannigan *et al.*, 1995). Ntn hydrolase superfamily shows insignificant homology in primary structure but shows high similarity in their αββα sandwich structure (**Ntn-hydrolase fold**), as well as the cleavage of the amide bond and autoproteolytic processing. (Oinonen and Rouvinen, 2000). In Ntn-hydrolase fold, two central anti-parallel β-sheets are sandwiched between layers of α-helices on either side (Figure 1.2). All Ntn hydrolase have conserved active site topology and similar reaction mechanisms (Duggleby *et al.*, 1995). Ntn hydrolase is active after autocatalytic processing which creates new N-terminal residues (Cys/Thr/Ser), which is responsible for the hydrolysis of an amide bond in Nt**C**n-hydrolase and Nt**T**n-hydrolase and Nt**S**n-hydrolase (Figure 1.3) (Ekici, Paetzel and Dalbey, 2008).The sulfhydryl (-SH) group of Cys and hydroxyl (-OH) group of Ser and Thr

act as a nucleophile in Ntn hydrolase. The zymogen of AC and NAAA are autoproteolytically processed into α- and β-subunit and this subunit forms an active heterodimer. N-terminal Cys residue of β-subunit act as nucleophile and form a catalytic triad with Arg and Asp residues.

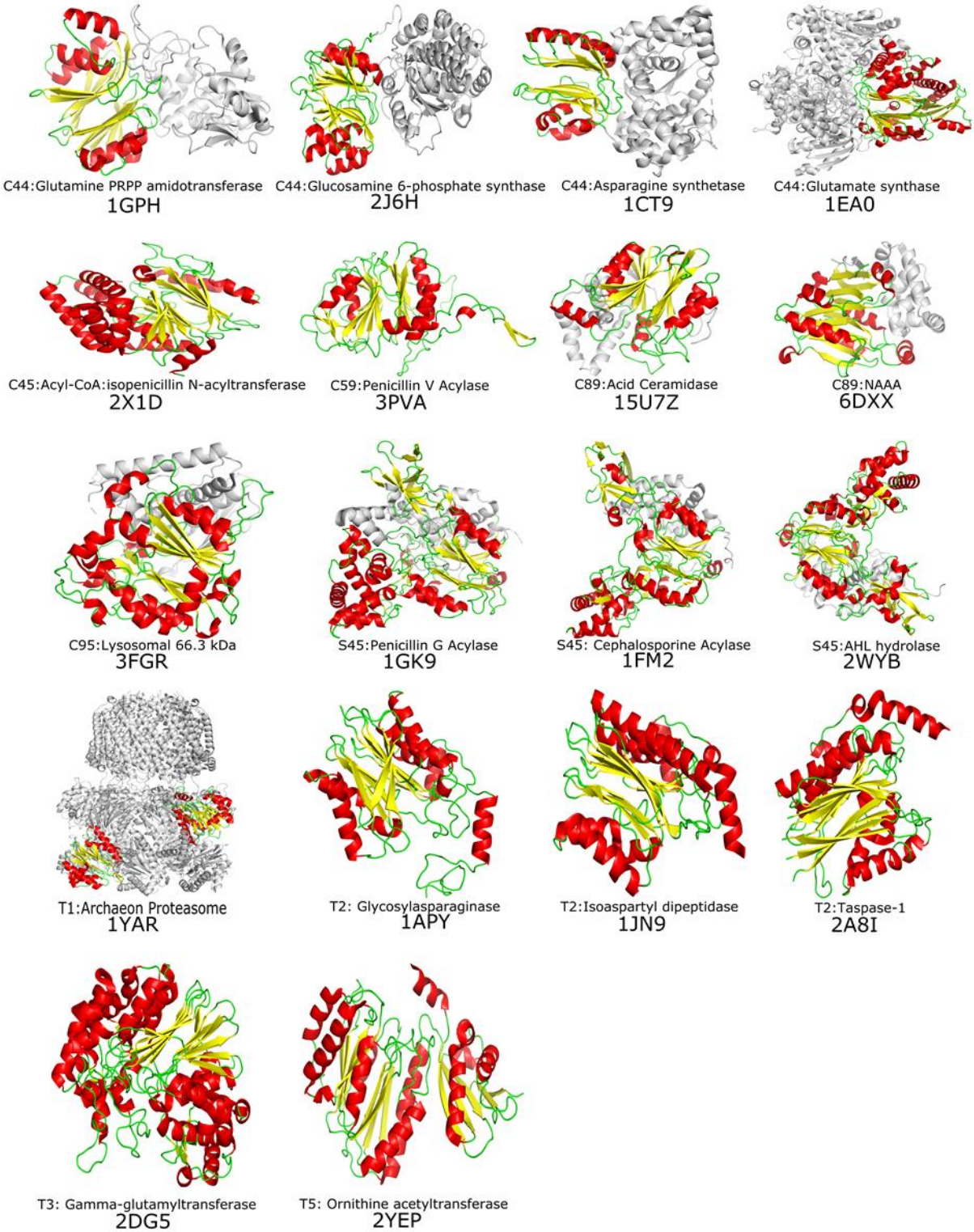


Figure 1. 2: Members of the Ntn hydrolase superfamily. The conserved core Ntn-fold is highlighted in the structure of proteins.

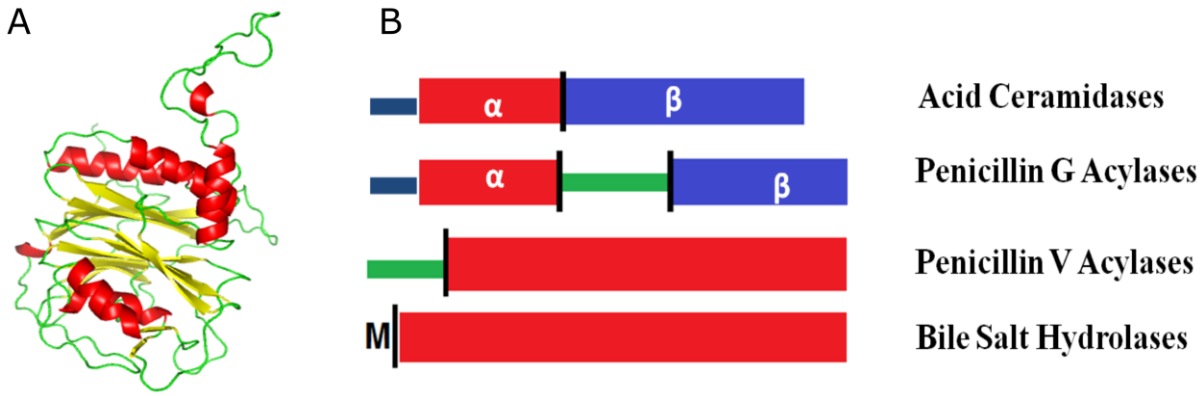


Figure 1.3: Modes of autoprocessing in Ntn hydrolase. A: Ntn hydrolase fold. B: Autoprocessing in Ntn hydrolase. The signal peptide, α- and β-subunit of AC and PGA showed in dark blue, red and light blue color, respectively. Spacer peptide of PGA and PVA showed in green color. M, initial methionine of BSH.

Based on the N-terminal nucleophile residue, Ntn-hydrolase enzymes (Clan PB) are classified into NtCn-hydrolase (subclan PB(C)) and NtSn-hydrolase (subclan PB(S)) and NtTn-hydrolase (subclan PB(T)) in MEROPS database (Rawlings et al., 2014). One set of closely related Ntn hydrolases of interest here is AC and NAAA which belongs to the Cholylglycine hydrolase family or NtCn-hydrolases (subclan PB).

1.3.1 Cysteine-Ntn (NtCn) Hydrolases (subclan PB) and Cholylglycine Hydrolase family

This clan PB(C) includes six structurally and functionally characterized families of self-processing enzymes, family C59, C44, C45, C69, C89 and C95. These family members are related to each other through structural homology, not sequence homology (Panigrahi et al., 2014). Family C89 comprises of the enzymes AC and NAAA. Other enzymes such as Cholylglycine hydrolase (BSH and PVA), Glutamine phosphoribosylpyrophosphate amidotransferase (GPATase), Glucosamine-fructose-6-phosphate aminotransferase (GFAT), Asparagine synthetase (AS), Glutamate synthase (GS), Dipeptidase DA, lysosomal 66.3 kDa protein and Acyl-CoA:isopenicillin N-acyltransferase (IPAT) also belong to NtTn-hydrolase family.

1.3.2 Serine Ntn-Hydrolase family

This clan PB(S) includes family S45 (cephalosporin acylase, AHL acylase and penicillin G acylase (PGA)). These enzymes are targeted to the periplasmic space by the signal peptide. They are synthesized as inactive premature precursor polypeptide (signal

peptide, α-chain, spacer peptide, β-chain), which is subsequently processed into mature active heterodimer consisting of α- and β-chains (Figure 1.3 and 1.6).

1.3.3 Threonine-Ntn (NtTn) Hydrolase family

This clan PB (T) includes four family T1, T2, T3 and T4. These enzymes differ with respect to their subunit composition and substrate specificity but the catalytic site and overall structural Ntn-fold remain similar. Enzymes such as Glycosylasparaginase (GA), Ornithine acetyltransferase (OA) and subunits of proteasome belong to NtTn-hydrolase family.

1.4 Distribution (evolution) of AC and NAAA

Distribution of AC and NAAA are found in bacteria, protozoa, animals and viruses but are absent in archaea and fungi (Rawlings et al., 2014). In bacterial and viral homologs, protein predicted to be AC and NAAA do not have consrvrd active site residues. In plants this protein is absent but Micromonas sp. Has NAAA protein predicated to be AC and NAAA do not have conserved active site residues. In plants this protein is absent but Micromonas sp. has NAAA protein. The reason behind the lack of both protein in plants may be due to the lack of lysosomes in most plants. The function of AC and NAAA in plants may be carried out by other ceramidase and FAAH, respectively. The homologous amino acid sequences of AC and NAAA could not be detected in Drosophila but neutral ceramidase was present; however, they were found to have activity in neutral pH as well as in acidic pH (Yoshimura et al., 2002). NAAA has high homology (33-35% amino acid identity) to AC with respect to amino acid sequence; however, more similarity is found in β-subunit amino acid sequences and both AC & NAAA have conserved active site residues (Figure 1.4 and 1.5). Similarities are also observed in their structure, cellular location, post-translation modification and autoproteolytic cleavage. AC and NAAA have.

1.5 The molecular biology of AC and NAAA

AC is encoded by the ASAH1 (N-acylsphingosine amidohydrolase) gene. Skin fibroblast and pituitary tissue cDNA was used to partially amplify the fragments and form the full length, human AC (hAC) (Koch et al., 1996). The cDNA encoding murine AC has also been cloned (Li et al., 1998). The locus of ASAH1 gene is on chromosome 8 (8p21.3-22) was confirmed by in situ hybridization and FISH analyses (Hong et al., 1999).

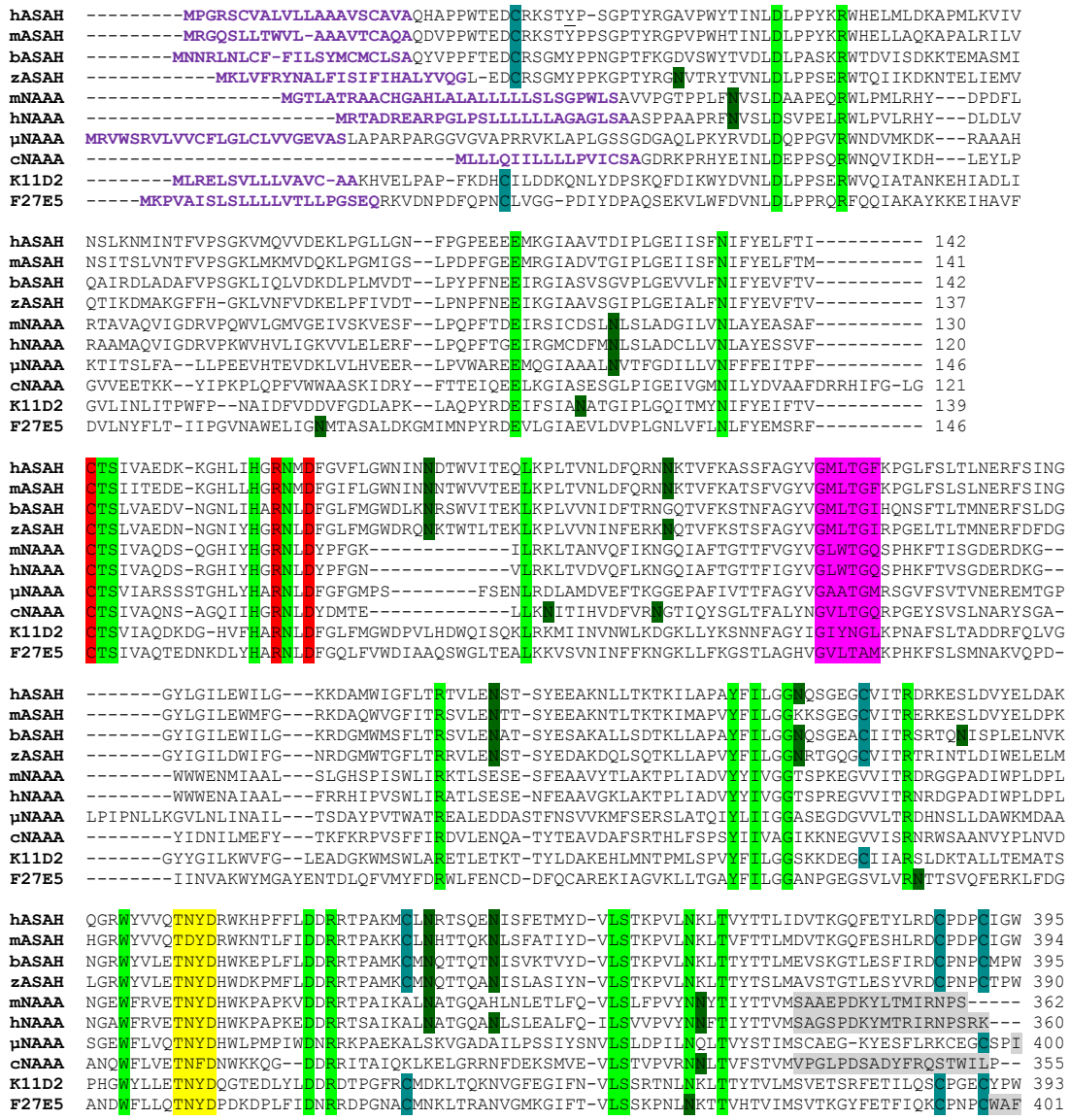


Figure 1.4: Multiple sequences alignment of AC and NAAA.

Multiple sequences alignment of AC and NAAA from different species analyzes using Clustal W. The red color amino acids are catalytic residues of AC & NAAA. Magenta and yellow color indicates respectively N-myristoylation and casein kinase II phosphorylation sites. N-glycosylation site shown in green dark color predicated using NetNGlyc 1.0 server.
hASAH- acid ceramidase precursor (*Homo sapiens*).
mASAH-acidceramidase precursor (*Mus musculus*).
bASAH-acidceramidase precursor (*Brachydanio rerio*).
zASAH-acidceramidase precursor (*Brachydanio rerio*).
mNAAA-N-acylethanolamine acid amidase precursor (*Mus musculus*).
hNAAA-N-acylethanolamine acid amidase precursor (*Homo sapiens*).
µNAAA-N-acylethanolamine acid amidase precursor (*Micromonas commoda*).
cNAAA-N-acylethanolamine acid amidase precursor (*Caenorhabditis elegans*).
K11D2- acid ceramidase precursor (*Caenorhabditis elegans*).
F27E5-(*Caenorhabditis elegans*).

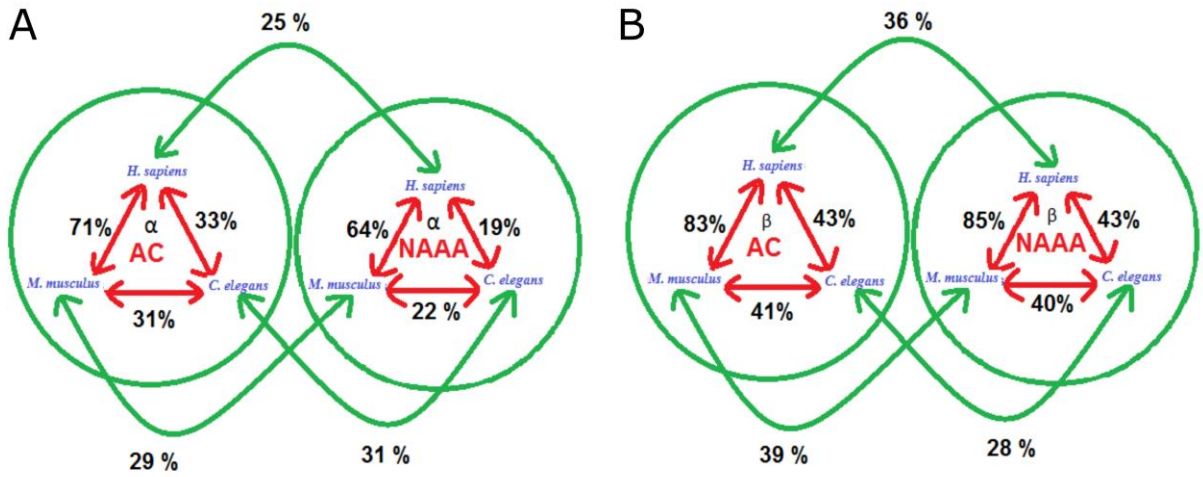


Figure 1. 5: The comparison of homology between AC and NAAA. A: sequence similarity between α-subunits of AC & NAAA of *H. sapiens, M. musculus* & *C. elegans.* B: sequence similarity between β-subunits of AC & NAAA of *H. sapiens, M. musculus* & *C. elegans.*

ASAH1 promoter region has SP1, AP2, NF-ķb, KLF6 and CACC box sequences in the promoter region (confirmed by luciferase assay using fibroblast cell line NIH3T3) (Li *et al.*, 1999). KLF6 expression positively correlates with *ASAH1* mRNA expression, suggesting that KLF6 transcription factor positively regulates the *ASAH1* gene expression at the transcription level (Park and Schuchman, 2006). The cAMP-response element binding protein (CREB) also regulates *ASAH1* gene expression at the transcription level (Lucki *et al.*, 2012). The *ASAH1* expression is found to be the highest in heart and kidney compared to placenta and lungs (Sugita *et al.*, 1975). AC negatively regulates gene expression of steroidogenic factor 1 dependent gene by binding to its promoter (Lucki *et al.*, 2012). USP2 positively regulate AC protein expression (Mizutani *et al.*, 2015). *NAAA* gene is located on the long arm of chromosome 4 (4q21.1), confirmed by FISH analyses (Hong *et al.*, 1999).

The human recombinant AC was overexpressed in Chinese hamster ovary (CHO) (He *et al.*, 2003) and Sf9 insect cell line by secretion into the culture media (Schulze, Schepers and Sandhoff, 2007) and purification by sequential chromatography. Recombinant AC is also found to be heterodimeric like human AC purified from a natural source (urine), has both α- and β-subunits and follows normal Michaelis-Menten kinetics. Heterologous expression of recombinant human NAAA was obtained in HEK293 cells (Tsuboi et al., 2005) and Sf9 insect cell line. Ni-NTA chromatography was used to purify secreted protein from the culture media (Gorelik et al., 2018), followed by size exclusion chromatography.

1.6 The biogenesis of AC and NAAA

AC and NAAA are targeting to the lumen of the endoplasmic reticulum (ER) by signal peptide (Figure 1.7) (Ferlinz *et al.*, 2001; Zhao *et al.*, 2007). It is synthesized as inactive premature precursor polypeptide (signal peptide, α- and β-subunit), which is subsequently processed into mature active heterodimer consisting of α- and β-subunit in lysosomes (Figure 1.6) (Ferlinz *et al.*, 2001; Zhao *et al.*, 2007). Maturation of the pre mature protein results in formation of a new N-terminal cysteine residue in the β-subunit (βcys1) which acts as a nucleophile during hydrolysis of the substrate (Zhao et al., 2007; Shtraizent et al., 2008). Amino acid sequencing of AC expressed in CHO cell line suggested the carboxyl-terminal processing of β-subunit (He et al., 2003). Both subunits of AC heterodimer is linked with disulfide bridge, confirmed by SDS-PAGE, mass spectrometry (MS) and the crystal structure of recombinant human AC (Schulze, Schepersand Sandhoff, 2007; Gebai et al., 2018). Mature NAAA is a heterodimer of the α- and β-subunits linked by non-covalent interactions (Zhao et al., 2007; West et al., 2012).

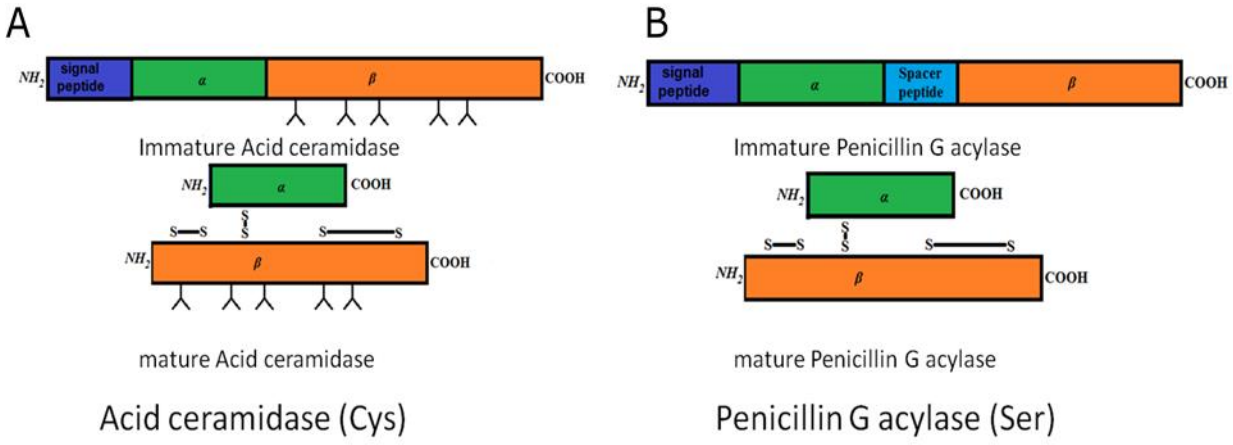


Figure 1.6: Schematic illustration of AC and PGA processing. The primary structure of mature and immature acid ceramidase compared with bacterial Ntn hydrolase enzyme PGA.

AC and NAAA are N-glycosylated in the ER and Golgi body (Figure 1.7). Peptide-N-glycanase (PNGase F) and Endo-H reduce the molecular weight of β-subunit of hAC suggesting that there is glycosylation in the hAC, however the actual site of glycosylation was confirmed by site-directed mutagenesis of AC cDNA expressed in CHO cells (Al et al., 1989; Ferlinz et al., 2001; He et al., 2003). MALDI-TOF MS analysis of human NAAA (hNAAA) treated with PNGase F, purified from expression in HEK293 cell line suggested the presence of glycosylation in hNAAA as well (West et al., 2012). AC glycosylation is necessary for the formation of its heterodimer, proper cellular targeting, enzyme stability and

activity. However recent studies have shown that glycosylation in hNAAA is not necessary for its activity (Pavlopoulos *et al.*, 2018). The hAC is localized in the lysosome and lysosomal targeting depends on mannose-6-phosphate receptor experimentally proven using I cell disease fibroblasts and NH4Cl (lysosomotropic agent) treated fibroblasts (Ferlinz et al., 2001). It is also observed that a small amount of unprocessed hAC is secreted outside the cell but its function is not characterized. An isoform 2 is also found to be localized in the cytoplasm and nucleus (Lucki et al., 2012). AC and NAAA are auto-processed into α and β-subunit in the lysosomal compartment (Ferlinz et al., 2001; Zhao et al., 2007).

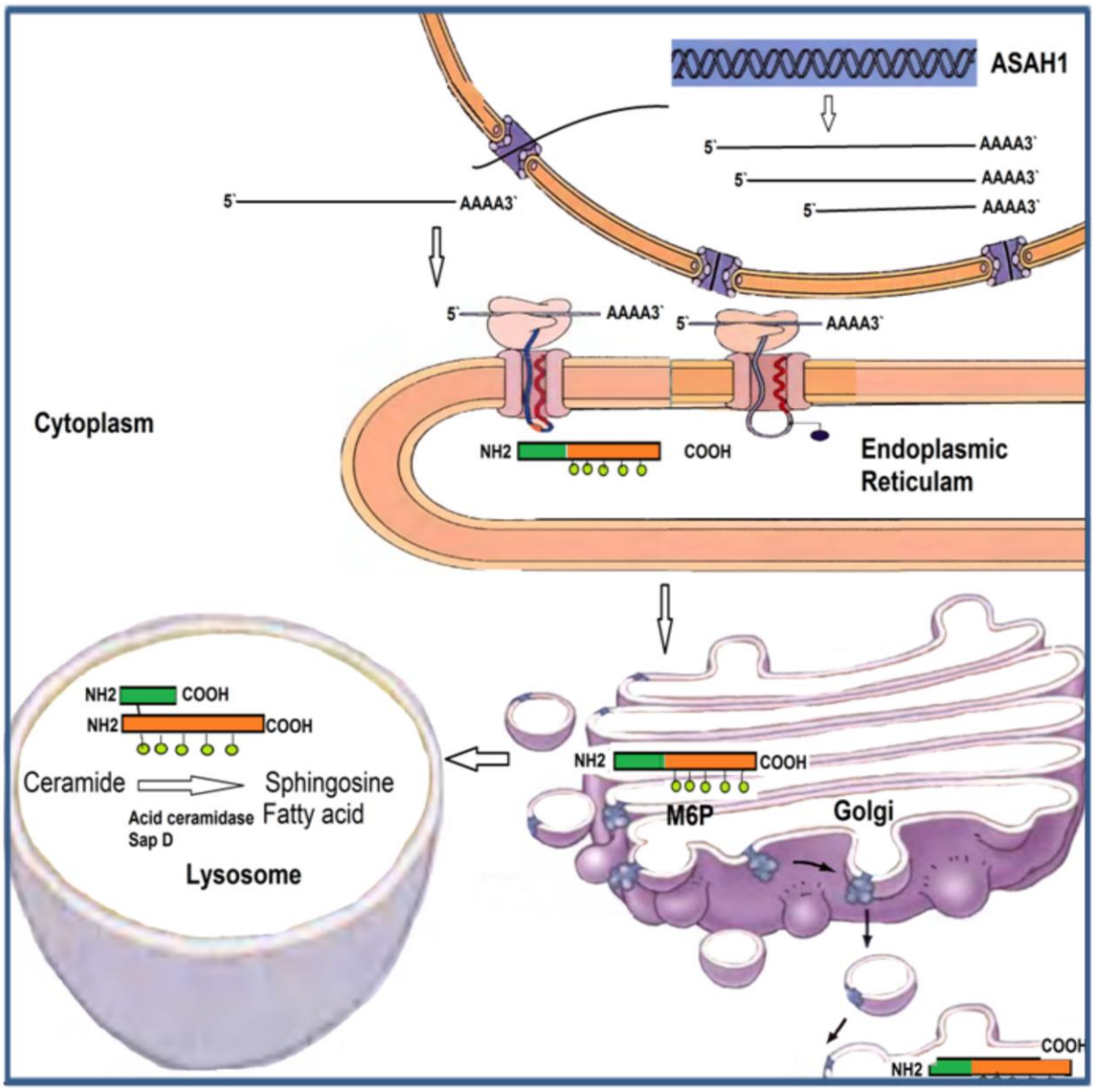


Figure 1. 7: The biogenesis of AC. The gene of AC produces different transcript by alternative splicing. The AC is initially targeted to ER lumen via signal peptide, N- and O-glycosylated in the ER and Golgi, respectivelyand finally targeted to the lysosome.

The β-subunit has N-terminal Cys and conserved Arg and Asp residues located further apart, which play an important role in processing and activity, proved by site-directed mutagenesis and cysteine protease inhibitor (Zhao et al., 2007; Shtraizent et al., 2008; Wang et al., 2008). Asp forms an H bond with Cys and then Arg act as proton acceptor as a result of which the thioether bond between Ile and Cys is the site of AC cleavage (Zhao et al., 2007; Shtraizent et al., 2008; Wang et al., 2008). Finally, water molecules catalyze the hydrolysis of the peptide bond (Figure 1.8) (Zhao et al., 2007; Shtraizent et al., 2008; Wang et al., 2008). The α-subunit is also essential for activity, confirmed by expression of β-subunit alone in COS -1 cell line (Shtraizent et al., 2008) and Farber disease mutation mostly reported in the α-subunit (Park and Schuchman, 2006). Maturation of recombinant AC and NAAA in vitro required acidic condition and physiological temperature (Shtraizent et al., 2008; Wang et al., 2008).

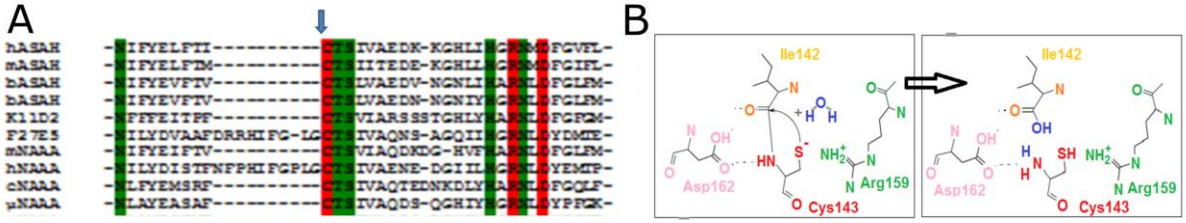


Figure 1.8 : The autoproteolytic cleavage mechanism of AC. Adapted from (Shtraizent *et al.*, 2008). A: The conserved autoproteolytic cleavage sites are shown by an arrow. B: Proposed mechanism of AC processing. Asp-162 forms a hydrogen bond with the amino group of Cys-143. The formation of this bond stabilizes a new conformation that is favorable for the amino group of Arg-159, to accept protons from the side chain sulfydril. The latter becomes a polar nucleophile and can form a thioether intermediate bond with the carbon atom of the carboxyl group of Ile-142. The hydrolysis of the thioether bond is catalyzed by a water molecule, resulting in cleavage of the peptide bond between Cys-143 and Ile-142.

1.7 Biochemistry of AC and NAAA

AC catalyzes the hydrolysis of ceramide into sphingosine (SPH) and free fatty acid (GATT, 1963) (Figure 1.9). Amide bond hydrolysis of ceramide by AC in enhanced in the presence of the sphingolipid activator protein, Sap-D. (Kishimoto, Hiraiwa and O'Brien, 1992; Klein *et al.*, 1994; Linke *et al.*, 2001). Sap-D stimulates activity by binding to negatively charged lipids and solubilizing the lipids (Azuma *et al.*, 1994). NAAA hydrolyze NAE into free fatty acid and ethanolamine (Figure 1.9) (Ueda, Yamanaka and Yamamoto, 2001). FAAH also hydrolyze NAE into free fatty acid and ethanolamine. Both NAAA and FAAH hydrolyze NAE with different substrates preferences. FAAH is active in a wide range

of pH but NAAA is only active at acidic pH (Ueda, Yamanaka and Yamamoto, 2001; Tsuboi *et al.*, 2005). In the presence of non-ionic detergents such as NP-40 and Triton X- 100 NAAA hydrolyzes PEA much faster than other NAEs but it was hardly able to hydrolyze 2-AG (Ueda, Yamanaka and Yamamoto, 2001; Tsuboi *et al.*, 2005). Moreover, dithiothreitol (DTT) is found to dose-dependently activate NAAA but does not stimulate FAAH (Ueda *et al.*, 1999; Ueda, Yamanaka and Yamamoto, 2001; Tsuboi *et al.*, 2005).

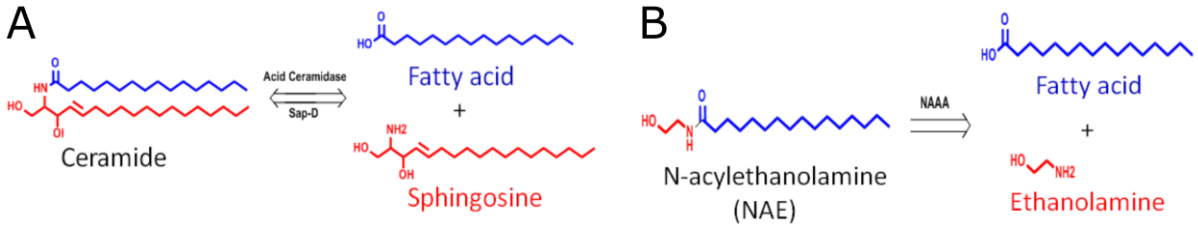


Figure 1.9: Reaction catalyzed by AC and NAAA. A: AC hydrolyzes ceramide into free fatty acid and sphingosine. B: The NAAA hydrolyze NAE into free fatty acid and ethanolamine.

AC also synthesizes ceramide *in vitro* and *in vivo* from sphingosine and free fatty acid (He *et al.*, 2003). Residues important for catalytic activity in hAC are Cys143 (Gebai *et al.*, 2018; Dementiev *et al.*, 2019), Asp162 (Gebai *et al.*, 2018), Asn320 (Gebai *et al.*, 2018) and Arg 333 (Gebai *et al.*, 2018).

AC inhibitors (Liu, Beckman and Foroozesh, 2013) are methyl methanethiosulfonate (Shtraizent *et al.*, 2008), benzoxazolone carboxamides (hAC, IC_{50}64 nm), (Bach *et al.*, 2015; Pizzirani *et al.*, 2015), carmfur (rat AC, IC_{50} 0.029 µM) (Pizzirani *et al.*, 2013; Realini *et al.*, 2013), tamoxifen (Morad *et al.*, 2013), toremifene (Morad *et al.*, 2013), B-13 (hNAAA, IC_{50} 10 µM) (Proksch, Klein and Arenz, 2011), N-oleoylethanolamine (hNAAA, *Ki*500 μM) (Proksch, Klein and Arenz, 2011), DM102 (Gouazé-Andersson *et al.*, 2011), desipramine (Elojeimy *et al.*, 2006), 2-substituted aminoethanol amides (Bedia *et al.*, 2008) and cystatin SA (Eliyahu *et al.*, 2011).

NAAA inhibitors are cyclopentyl hexadecanoate (hNAAA, IC50 10µM) (Saturnino et al., 2010) pyrrolidine derivative 16 (rNAAA, IC502 µM) (Li et al., 2012), isothiocyanates AM9023 (hNAAA, IC50 0.600 µM) (West et al., 2012), AM9053 (hNAAA, Ki 27 µM) (Alhouayek et al., 2017), β-lactones(S)-OOPP (rNAAA, IC500.420µM) (Solorzano et al., 2009) (Solorzano et al., 2009), ARN077 (hNAAA, IC500.007 µM) (Ponzano et al., 2013), (rNAAA, IC50 0.01–0.130 µM) (Duranti et al., 2012; Ponzano et al., 2013; Sasso et al.,

2013), β-lactams ARN726 (hNAAA, IC500.027–0.073 µM) (Ribeiro et al., 2015; Nuzzi et al., 2016), (rNAAA,IC50 0.063 µM) (Ribeiro et al., 2015), N-O-aryl-substituted derivative 37 (hNAAA, IC500.006 µM) (Petracca et al., 2017), oxazolidonesF96 (rNAAA, IC50 0.270 µM) (Yang et al., 2015; Ren et al., 2017), F215 (rNAAA, IC50 0.009 µM) (Li et al., 2017), EPT4900 (diacerein) (hNAAA, IC500.700 µM) (Petrosino et al., 2015) and benzothiazolepiperazine derivative 8 (hNAAA, IC500.230 µM) (Migliore et al., 2016).

Recombinant hAC was co-immunoprecipitated with acid sphingomyelinase and β-galactosidase enzyme in CHO cell line supernatant (He et al., 2003). Sphingomyelinase activity was increased by 20% compared to normal CHO cell line. It was also observed in the case of Farber disease (FD) cell when AC overexpressed. These studies suggest that AC interacts with sphingomyelinase and β-galactosidase (Koch et al., 1996).

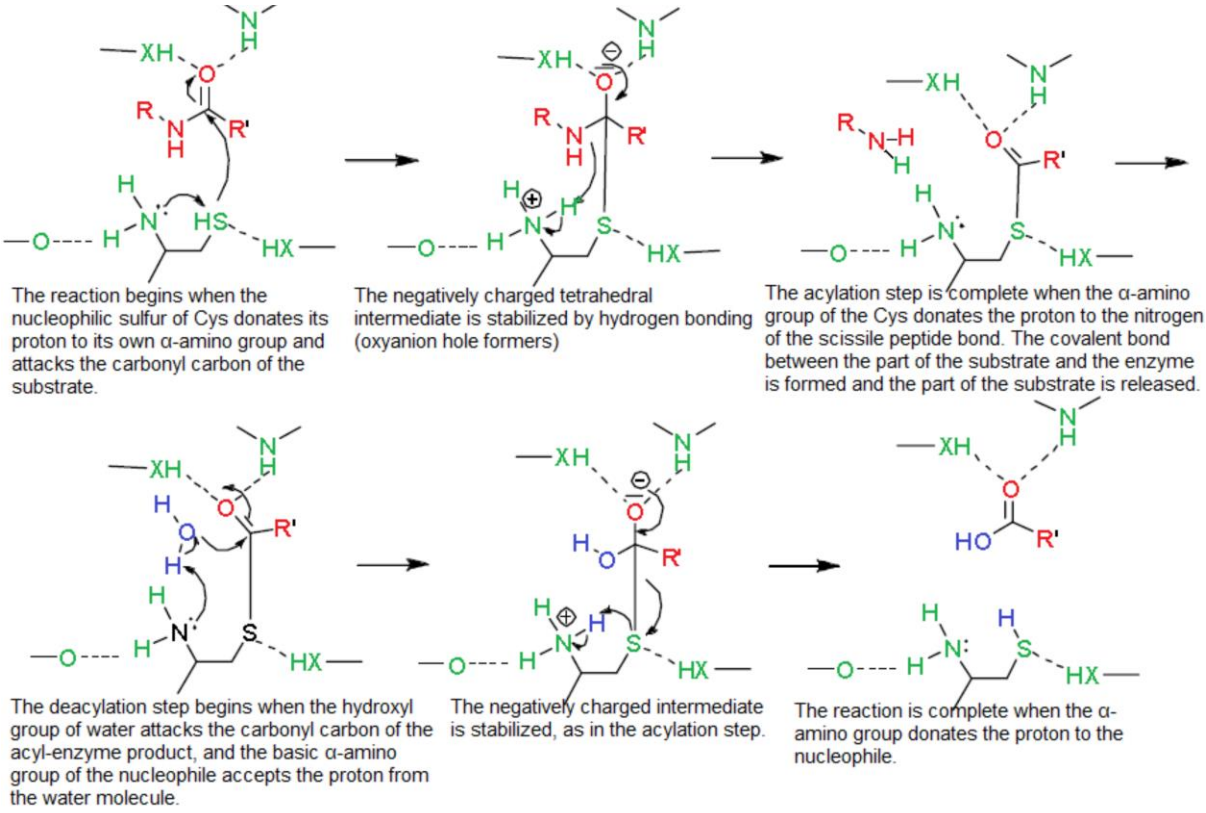


Figure 1.10 : The catalytic mechanism of acid ceramidase. Modified from (Oinonen and Rouvinen, 2000).

1.8 Sphingolipid metabolism and signaling

Ceramide generated by SMase pathway by hydrolysis of sphingomyelin, de novo synthesis from palmitate and serine and via salvage pathway from sphingosine and fatty acid (Figure 1.11). Ceramide can be hydrolyzed by acid, neutral and alkaline ceramidases into

fatty acids and sphingosine (Liu, Beckman and Foroozesh, 2013). Ceramide promotes apoptosis, cell cycle and autophagic response (Liu, Beckman and Foroozesh, 2013). Activation of PP2A, PKCζ, ASK1-P38 by ceramide inactivates the AKT (anti-apoptotic kinase) through dephosphorylation (Bourbon, Sandirasegarane and Kester, 2002; Wang *et al.*, 2005; Li *et al.*, 2018). Inactivated AKT decreases the phosphorylation of BCL-2. BCL-2 phosphorylation is inhibited by PP2A, it also promotes p53/BCL-2 complex formation (Deng, Gao and May, 2009). Cathepsin is also activated by lysosomal ceramide (Heinrich *et al.*, 2004) and is known to induce outer membrane permeability (von Haefen *et al.*, 2002).

Ceramide-1-phosphate (C1P) is generated by phosphorylation of ceramide by ceramide kinase. Ceramide promotes cell proliferation, migration, invasion and angiogenesis (Liu, Beckman and Foroozesh, 2013). Ceramide-1-phosphate blocks apoptosis by inhibition of SMase and serine palmitoyltransferase (Gómez-Muñoz et al., 2004; Granado et al., 2009). JNK, PI3/AKT and ERK1/2 are activated by C1P simulating cell proliferation (Gangoiti et al., 2008, 2012).

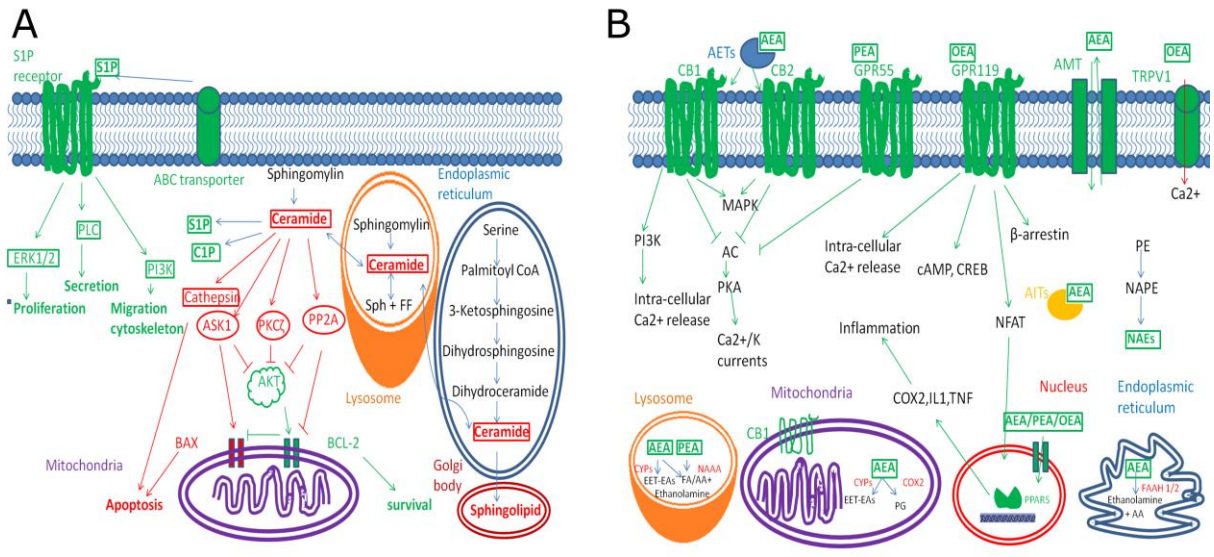


Figure 1.11: The metabolism and signal transduction pathway of ceramide and NAE. A: Ceramide metabolism and signal transduction pathway. B: NAE metabolism and signal transduction pathway.

sphingosine-1-phosphate (S1P) is generated by phosphorylation of sphingosine by a sphingosine kinase (Maceyka *et al.*, 2002). Sphingosine-1-phosphate promotes survival, proliferation, migration, invasion and angiogenesis (Liu, Beckman and Foroozesh, 2013). In the intracellular signaling pathway, S1P is known to upregulate the anti-apoptotic protein (BCL-2), which increases SK1 levels (Bektas et al., 2005; Pyne and Pyne, 2010) and

downregulate the pro-apoptotic protein (BAX) (Sauer *et al.*, 2005; Avery *et al.*, 2008). In the S1P receptor-dependent mechanism, ABC transporters export intracellular S1P outside the cells where they bind to the 7 trans membranes S1P receptor (Mitra *et al.*, 2006).

1.9 NAE metabolism and signaling

NAEs (palmitoylethanolamide (PEA), oleoylethanolamide (OEA) and arachidonoylethanolamide (anandamide, AEA) are generated by transacylation-phosphodiesterase pathway (Figure 1.11). N-acylphosphatidylethanolamines (NAPEs) are generated by N-acyltransferase (NAT) from membrane glycerophospholipids (Ogura *et al.*, 2016). The acyl donors are phosphatidylcholine (PC), 1-acyl-lyso-PC, cardiolipin and phosphatidylethanolamine (PE). The acceptors are PE, alkylacyl-PE and plasmenylethanolamine. NAPEs are hydrolyzed by NAPE- hydrolyzing phospholipase D (NAPE-PLD) into NAEs. NAEs are further hydrolyzed by FAAH and NAAA into fatty acids and ethanolamine (Hussain *et al.*, 2017).

Cytochrome P450 enzymes, Cyclooxygenase-2 and lipoxygenases metabolize NAE into epoxyeicosatrienoic acids, ethanolamide derivatives (prostaglandins) and ethanolamides of hydroxyeicosatetraenoic acids, respectively (Urquhart, Nicolaou and Woodward, 2015). N-oleoyethanolamide binds to the ion channel transient receptor potential cation channel subfamily V1, G-protein-coupled receptor GPR119 and, peroxisome proliferator-activated receptor α (PPARα) (Ahern, 2003; Lauffer, Iakoubov and Brubaker, 2009; O'Sullivan, 2016). PEA is mostly a PPARα ligand. N-docosahexaenoylethanolamine (DHEA), N-eicosapentaenoylethanolamine and AEA activate cannabinoid receptors CB1 and CB2 which also activates PPARγ (Alhouayek and Muccioli, 2014).

1.10 Structure of AC and NAAA

Recently AC and NAAA crystal structure in both states (processed and unprocessed) was solved (Gebai *et al.*, 2018; Gorelik *et al.*, 2018). The processed and unprocessed have overall structural architectures based on the basis of their subunit organization. The α-subunit of AC and NAAA consists of five α-helices and exhibits the highest structural variability. The β-subunit of AC and NAAA share the typical Ntn-hydrolase. The β-subunit of AC and NAAA constitutes of two central anti-parallel β-strands flanked on either side by six and two α-helices respectively. In human NAAA contains one disulfide bond within the α-subunit, whereas in AC two disulfide bonds are present. The α-subunit and β-subunit of AC are

connected by a disulfide and a turn is stabilized in the β-subunit by the second disulfide bond. Activation of AC and NAAA require autocleavage by Cys, breaking the peptide bond between it and its preceding residue by intramolecular cleavage. Either a bound water molecule or Arg causes Cys deprotonation. The overall structure of the enzyme does not change significantly after autocleavage. In the α-subunit the last two residues of the C-terminus are disordered exposing the active site. Cys residues carry out both autocleavage and substrate hydrolysis but the oxyanion hole is formed by different residues. The substrate binding channel in AC comprised hydrophobic residues and the surrounding region of substrate binding channel is also hydrophobic, composed of 39 residues from both the α- and β-subunits. NAAA hydrophobic helices α3 and α6 possibly interact with the membrane as a result of which a hollow space is generated between the two subunits, which accommodates the substrate.

1.11 Role of AC in human health and disease

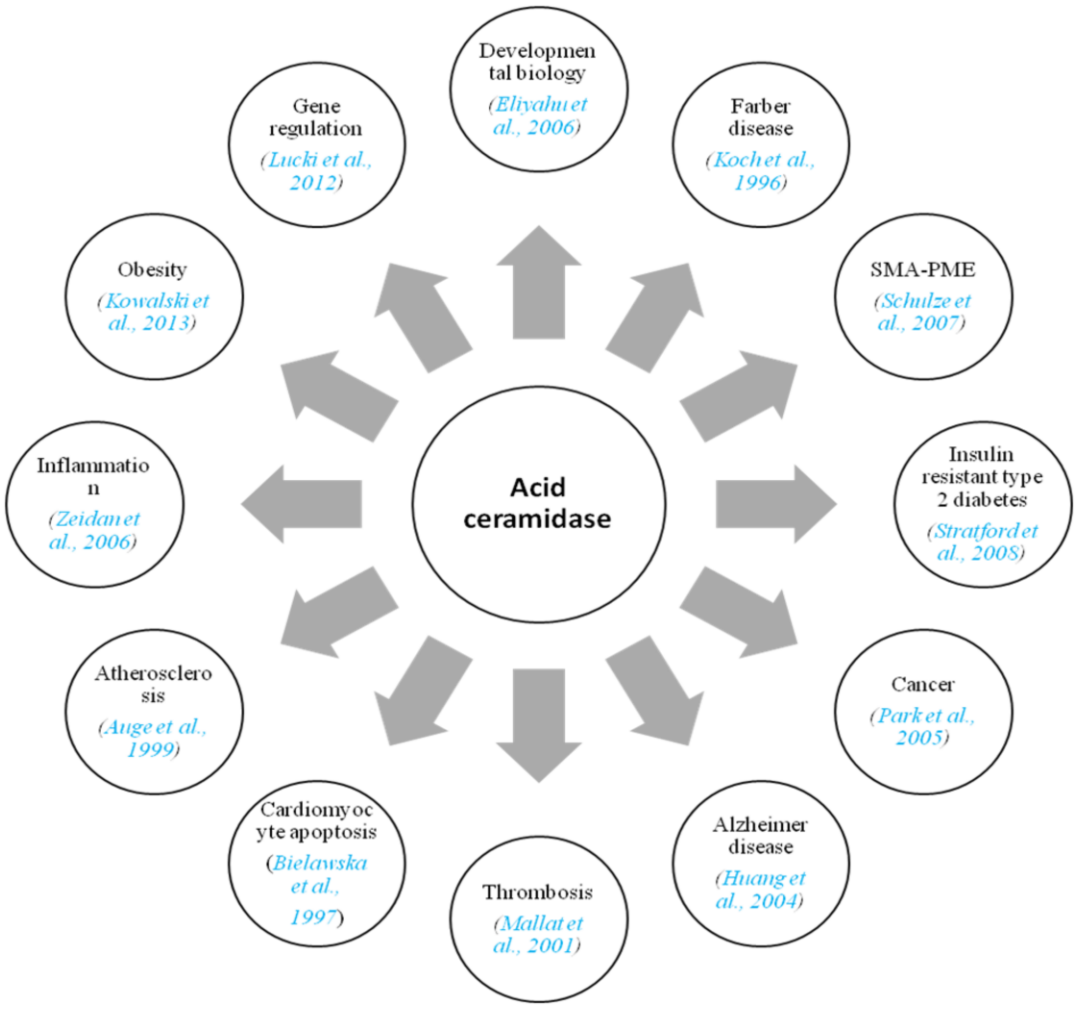


Figure 1.12: AC associated with human disease.

1.11.1 Farber disease (FD)

Farber lipogranulomatosis (having autosomal recessive inheritance) was described and characterized by Sidney Farber (FARBER, COHEN and UZMAN, 1957). Accumulation of ceramide in subcutaneous nodule and absent of ceramidase activity in kidney and cerebellum from a patient with FD suggested the deficiency of ceramidase in the body (Sugita et al., 1975). Characteristic features of this disease are subcutaneous nodules, painful and progressively deformed joints and hoarseness (Alves et al., 2013) (Figure 4. 13). HPLC, mass spectrometry, immunofluorescence microscopy, transmission electron microscopy and thin layer chromatography based assay used for diagnosis of FD (He et al., 1999; Bedia et al., 2010). N-laurylsphingosine can be used as a substrate for diagnosis and carrier detection of FD (Ben-Yoseph Y, Gagné R, Parvathy MR, Mitchell DA, 1989). Deficiencies of prosaposin in FD was also reported (Schnabel et al., 1992; Chatelut et al., 1997). AC gene therapy for FD is presently available by using lentivirus (Ramsubir et al., 2008).

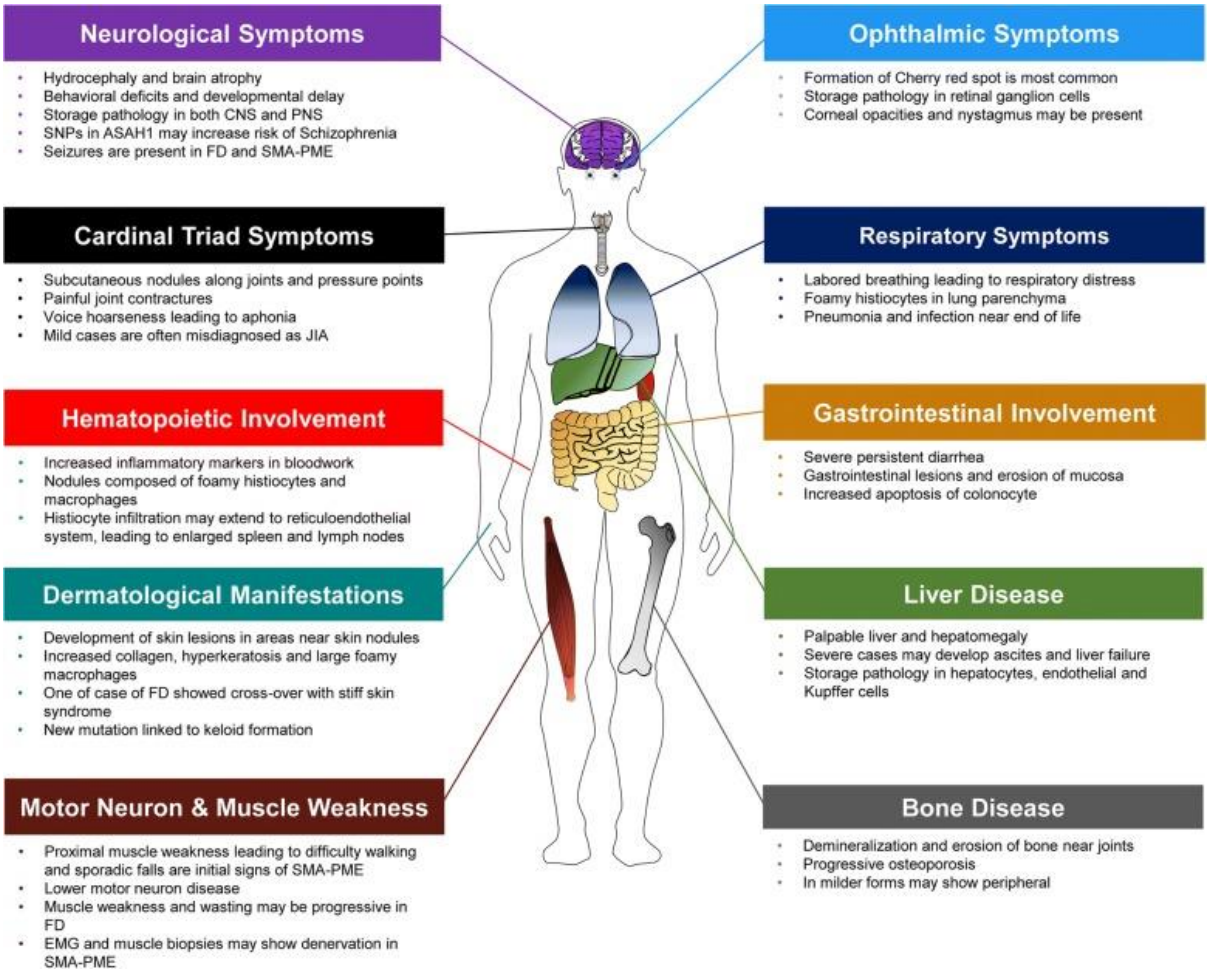


Figure 1.13: The typical clinical manifestations by organ type that have been reported in cases of FD and SMA-PME. Adapted from (Yu *et al.*, 2018).

1.11.2 Spinal muscular atrophy with progressive myoclonic epilepsy (SMA-PME)

Spinal muscular atrophy (SMA) is a neuromuscular disorder that causes degeneration of lower motor neurons of the spinal cord resulting in paralysis (Lefebvre et al., 1995) (Figure 4.13). Progressive myoclonic epilepsy (PME) is causing generalized and myoclonic seizures with progressive degeneration of neurons (Jankovic and Rivera, 1979; Haliloglu et al., 2002). A homozygous missense mutation (C125T) (Thr42Met) in *ASAH1* was identified by genomic-wide SNP genotyping and exome sequencing of families affected by childhood SMA-PME (Zhou et al., 2012).

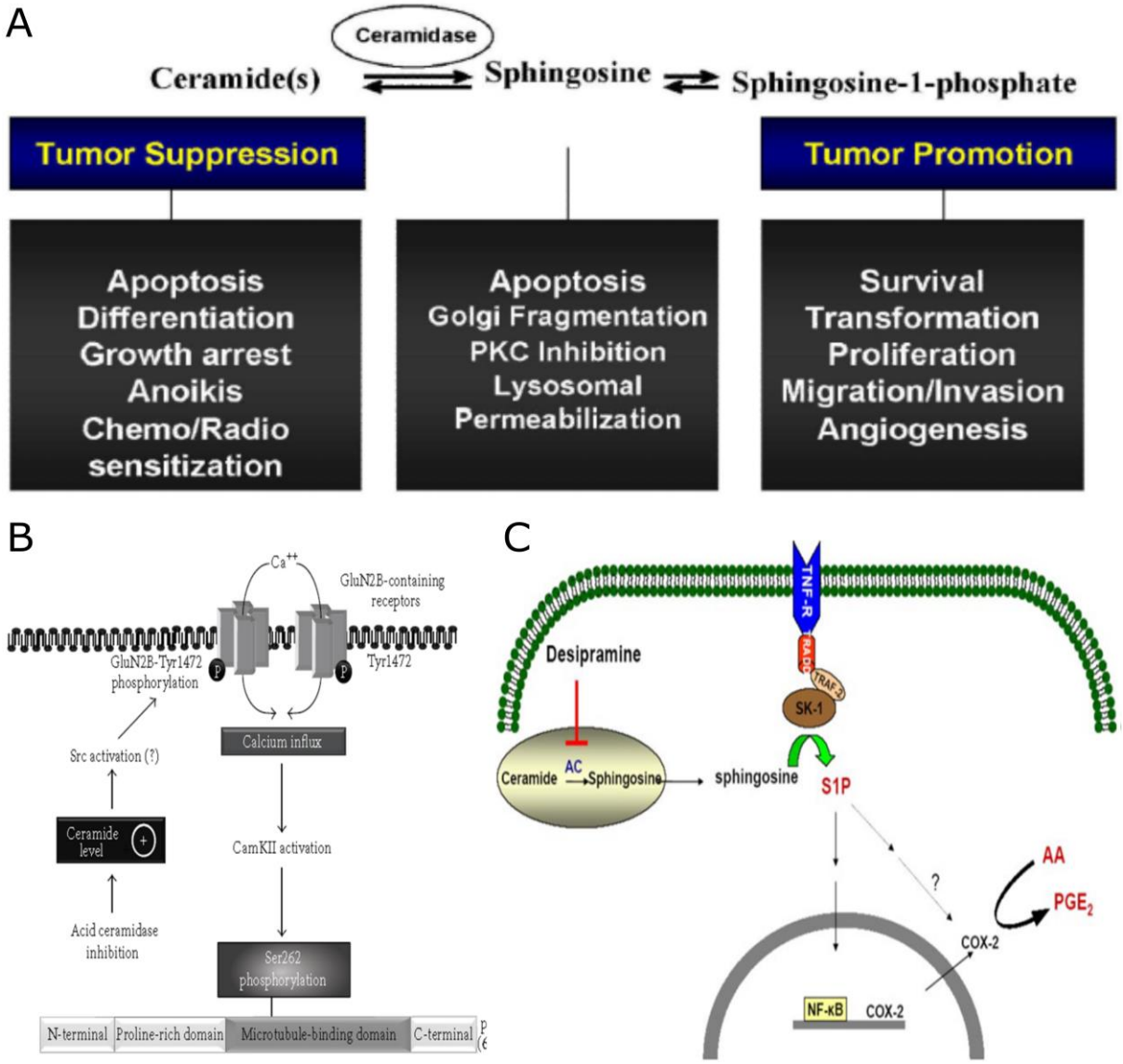


Figure 1.14: AC role in cancer Alzheimer and inflammation. A: The AC role in cancer biology. Adapted from (Zeidan *et al.*, 2008). B: The accumulation ceramide responsible for hyperphosphorylation tau. Adapted from (Laurier-Laurin *et al.*, 2014). C: The sphingosine-1-phosphate up-regulated cyclooxygenase two responsible for inflammation. Adapted from (Zeidan *et al.*, 2006).

1.11.3 Developmental biology

First AC "knock-down" mouse model was constructed, but the homozygote mutant was notably absent after developmental day E8.5 (Li et al., 2002). Knockdown mice undergo apoptotic death at blastula (Eliyahu et al., 2007). AC suppresses ceramide-mediated apoptosis essential for embryo survival (Morita et al., 2000). AC also improves the quality of oocytes and embryos (Butler et al., 2002).

1.11.4 Alzheimer's disease

The level and activity of AC was found to increase in the brain in case of Alzheimer's disease and may be involved in neurofibrillary degeneration (Huang et al., 2004). AC inactivation induces tau hyperphosphorylation, which impacts on neuronal survival due to protein localization and aggregation in cells (Laurier-Laurin et al., 2014) (Figure 4. 14 B).

1.11.5 Insulin resistant type 2 diabetes

The ceramide accumulation inhibits glucose regulation by inhibition of Akt/protein kinase B of insulin signaling (Stratford *et al.*, 2004) and overexpression of AC prevents this effect (Chavez *et al.*, 2005).

1.11.6 Cancer

AC balances cell cycle and apoptosis by regulating many signaling molecules and plays a central role in sphingolipid metabolism (Figure 4. 14A). The overexpression of AC suppressed sindbis virus (SV) (Jan, Chatterjee and Griffin, 2000) and tumor necrosis factor alpha (TNFα) induced apoptosis (Strelow *et al.*, 2000). Suppressed AC activity by N-oleoylethanoamide enhances apoptosis in L929 cells (Strelow *et al.*, 2000) TNFα/IFNα induced apoptosis in primary placental trophoblasts (Payne, Brindley and Guilbert, 1999). Suppressed AC activity by other inhibitors like B13 and NGF, induces apoptosis in prostate cancer cells (Selzner *et al.*, 2001; Samsel *et al.*, 2004) and hippocampal neurons (Brann *et al.*, 2002) respectively. Apoptosis was increased in primary keratinocytes as well, by suppressed AC activity (Raisova *et al.*, 2002). Lysosomal ceramide does not involve in stress-induced apoptosis (Ségui *et al.*, 2000; Burek *et al.*, 2001).

AC is overexpressed in most prostate cancers and causes export PTEN tumor suppression in the nucleus leading to oncogenesis (Beckham *et al.*, 2013). . AC is overexpressed in various prostate, head, and squamous cells cancers. AC contributes to

cellular proliferation by inhibiting the amassing ceramide and generating growth simulating sphingolipid signaling moities, predominantly S1P which bypasses the cell arrest signal. (Monick *et al.*, 2004). The concentration of S1P increases due to the overexpression of AC, which promotes the up-regulation of cathepsin B by Ets 1 transcription factor (Beckham *et al.*, 2012). AC also induces cell invasion through AKT2 (Berndt *et al.*, 2013).

1.11.7 Complex diseases

Atherosclerosis, thrombosis, cardiomyocyte apoptosis and inflammation. Atherosclerosis (Augé *et al.*, 1999), thrombosis (Mallat and Tedgui, 2001), cardiomyocyte apoptosis (Bielawska AE, Shapiro JP, Jiang L, Melkonyan HS, Piot C, Wolfe CL, Tomei LD, Hannun YA, 1997) and inflammation (Figure 4. 13 C) (Zeidan *et al.*, 2006) have been associated to defective ceramide metabolism.

1.12 Role of NAAA in human disease

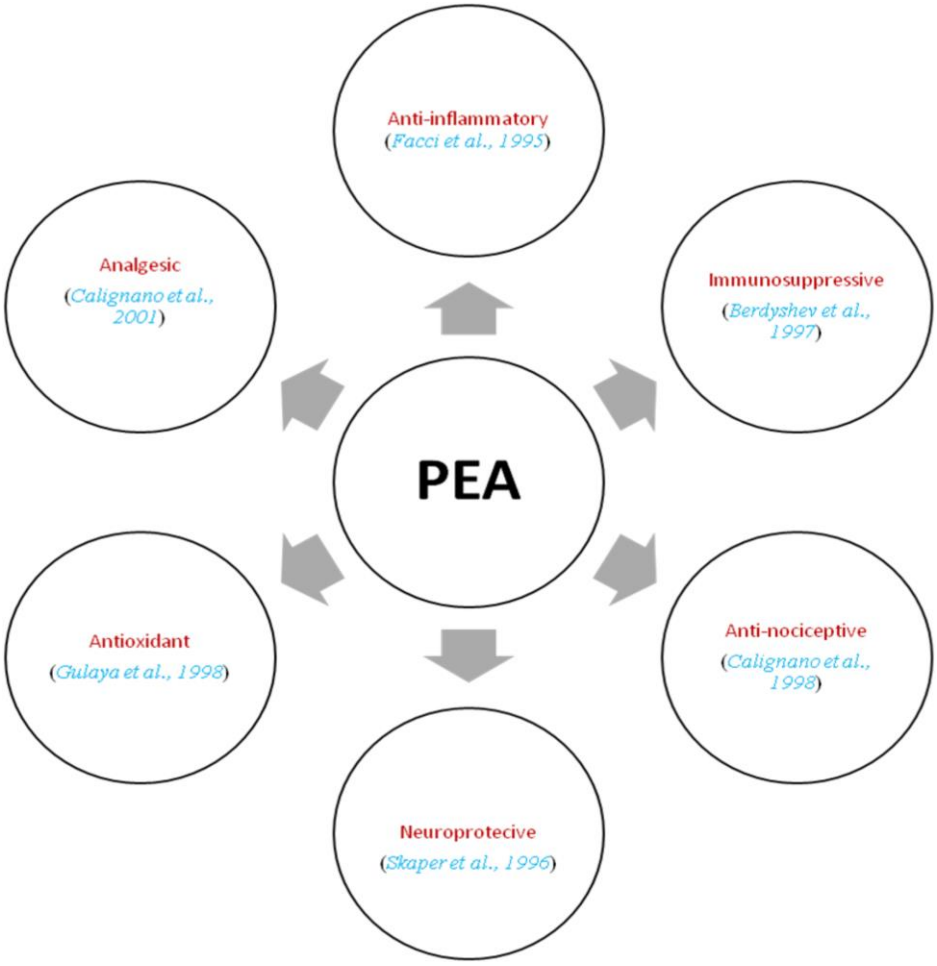


Figure 1. 15: NAAA associated with human disease.

The substrate of NAAA, PEA is anti-inflammatory (Facci et al., 1995; Mazzari et al., 1996), immunosuppressive (Berdyshev et al., 1997), anti-nociceptive (Calignano, La Rana and Piomelli, 2001), neuroprotective (Skaper et al., 1996), antioxidant (Gulaya et al., 1998) and analgesic (Calignano, La Rana and Piomelli, 2001) in nature. Lung inflammation (Ribeiro et al., 2015; Li et al., 2018), arthritis (Bonezzi et al., 2016), allergic contact dermatitis (Sasso et al., 2018), spinal cord trauma (Solorzano et al., 2009) and multiple sclerosis (Migliore et al., 2016) are some of the diseases which are influenced by levels of NAAA in the body (Figure 4. 15).

1.13 Statement of problem

In this doctoral research work, we intended to understand the comparative structure-function relationship between AC and NAAA. We have selected ACs from zebrafish and *C. elegans* and NAAAs from mouse and *M. commoda*. Both AC and NAAA enzymes belong to Ntn-hydrolase family. These enzymes have a crucial role in human physiology. Zebrafish and *C. elegans* are good model organisms for the study of neuronal disease (spinal muscular atrophy with progressive myoclonic epilepsy) and lysosomal storage disease (FD), respectively. *M. commoda* NAAA was chosen because; in Plantae, AC and NAAA homologs are absent. In *Micromonas sp.* lysosome organelle is absent but it's speculated that in *Micromonas sp.* this protein located in vacuoles.

Many reports have indicated that the human AC is overexpressed in several types of human cancers including prostate, breast, head and neck cancer and melanoma. These findings have implied that human AC can be a potential clinical target to treat cancer. The NAAA is a potential target for treatments of inflammatory disorders. The recent three-dimensional structures have been reported for AC and NAAA, which were produced from complicated eukaryotic expression system in insect cells. Our study optimized the expression of AC and NAAA in the simple eukaryotic system for structural studies, which demand an active and relatively large amount of highly pure protein. Initially, heterologous expression of AC and NAAA were tested in different strains of E. coli such as BL21 (DE3), Rosetta, C41 (DE3), C43 (DE3) and Lemo21 (DE3) but no expression was observed. The possible reason may be the inherent problem of the bacterial expression systems, which lack post-translational modifications like glycosylation and thereby for eukaryotic protein expression, it spells disaster. These modifications are generally crucial for eukaryotic protein expression, folding and stability. In the case of AC and NAAA, the post-translational modifications are

necessary for its autoproteolytic processing, localization and function. To overcome this challenge, we have chosen the *Pichia pastoris* expression system for high-level expression of recombinant AC and NAAA. The AC and NAAA genes were cloned in pPICZα-A and heterologous expression in *P. Pastoris* was performed. The enzymes were studied biochemically and biophysically to understand the optimum pH, temperature, DTT concentration, substrate specificity, stability and catalytic efficiency. The post-translation modification such as glycosylation, disulfide bond and autoproteolytic cleavage were comparatively studied in AC and NAAA. The mutational studies of zebrafish AC were also performed to understand the SMA-PME at the protein level. Thirty known disease-causing missense mutations were analyzed *in silico,* to understand Farber and SMA-PME disease at molecular levels. The seven mutations (L182V, R226P, G235R, G235D, R254G, R333H and P362R) were identified as highly deleterious from 30 diseases causing nsSNPs and were further investigated for the effects of mutations on protein structure using MD simulation approach.

Chapter 2: Characterization of acid ceramidase (AC)

2.1 Introduction

Acid ceramidase (AC) is a heterodimeric glycoprotein and localizes in the lysosome (Bernardo *et al.*, 1995; Ferlinz *et al.*, 2001). It catalyzes the hydrolysis ceramide and this reaction enhanced by the presence of the Sap-D (Klein *et al.*, 1994; Linke *et al.*, 2001). AC belongs to the Cholylglycine hydrolase family, a subfamily of Ntn (N-terminal nucleophile) hydrolase structural superfamily (Pfam: PF02275) (Oinonen and Rouvinen, 2000; Rossocha *et al.*, 2005; Tsuboi, Takezaki and Ueda, 2007). AC is also known for the synthesis of ceramide using free fatty acid and sphingosine as a precursor *in vitro* and *in vivo* (Okino *et al.*, 2003). AC activity was first identified in rat brain (GATT, 1963), but purified and characterized from human urine (Bernardo *et al.*, 1995). The full length, human AC cDNA was isolated, assembled and cloned from partial cDNA of pituitary tissue and skin fibroblast (Koch *et al.*, 1996). The cDNA encoding murine AC has also been cloned (Li *et al.*, 1998). The recombinant hAC was overexpressed in Chinese hamster ovary (CHO) cells (He *et al.*, 2003) and Sf21 insect cell line (Schulze, Schepers and Sandhoff, 2007), where it got secreted into the culture media and was purified by stepwise chromatography.

Mutation in ASAH1 has been connected with a rare lysosomal storage disorder known as FD (Koch *et al.*, 1996) and the rare neurodegenerative condition is known as SMA-PME (Zhou *et al.*, 2012). AC is synthesized as inactive premature precursor polypeptide (signal peptide, α-subunit, β-subunit) targeted to the lumen of the endoplasmic reticulum by an ER-specific signal peptide, which is subsequently processed into mature active heterodimer (consisting of α- and β-subunits) in lysosomes (Ferlinz *et al.*, 2001). Maturation of pre mature enzyme results in the formation of a new N-terminal Cys residue in the β-subunit (βCys1) of hAC which acts as a nucleophile during hydrolysis of ceramide (Shtraizent *et al.*, 2008). The catalytic mechanism of AC is similar to other well-established members of the Ntn hydrolase superfamily. The active site of the enzyme contains three amino acid residues, (Cys, Arg and Asp) which form the catalytic center of the enzyme, out of which the cysteine residue acts as a nucleophile (Shtraizent *et al.*, 2008), which is located at N-terminal of the β-subunits. Amino acid sequencing of the AC expressed in CHO cell line suggested the carboxyl-terminal processing of the β-subunit (He *et al.*, 2003). AC

glycosylation is necessary for precise cellular targeting, the formation of heterodimer, stability and activity of the enzymes (Ferlinz *et al.*, 2001). Morpholino knockdown of the asah1b in zebrafish suggests that zebrafish AC is required for motor-neuron axonal branching (Zhou *et al.*, 2012).

In the present chapter, we have reported for the first time, an efficient expression system for the heterologous overexpression of biologically active AC of zebrafish (*Danio rerio*) and *Caenorhabditis elegans*. The eukaryotic expression system *Pichia pastoris* was used to express the protein successfully. This expression system is easier and well suited for expressing recombinant AC with post-translational modifications such as proteolytic processing, disulfide-bonds and glycosylation compared to prokaryotic E. coli and animal cell lines. This is an important consideration for the production of functionally active recombinant forms of naturally glycosylated proteins such as zebrafish AC and *C. elegans* AC.

2.2 Materials and Methods

2.2.1 Cloning of asah1b and asah1

2.2.1.1 Sequences analysis

The cDNA sequences of zebrafish asah1b and *C. elegans* asah1 and signal peptide length were retrieved from GenBank (GenBank Accession No. NM_200577 and No. NM_493173). Autoproteolytic cleavage site, α-and β-subunit sequences were obtained from the MEROPS database. N-glycosylation and O-glycosylation sites were predicated using NetNGlyc 1.0 server and NetOGlyc 4.0 server, respectively. The theoretical p*I* and molecular weight of the protein was calculated from the ExPASy ProtParam tool server.

2.2.1.2 RNA isolation, cDNA preparation, primer design and PCR amplification

N2 strain of *C.elegan*s was maintained on NGM OP50 plates at 15°C. Total RNA was isolated from nematode using Trizol® Reagent (Life Technology, Cat#10296-010) as per the manufacturer`s instruction. Purified RNA samples were analyzed by agarose gel electrophoresis and concentration was spectrophotometrically determined using Nanodrop (Thermo Scientific, USA). One µg of purified RNA was used for the preparation of cDNA using the SuperScript™ III First Strand Synthesis System (Life Technology, Cat#18080-051).

2.2.1.3 Construction of plasmid vector (pPICZ*α*-A) and gene (asah1b and asah1)

Zebrafish cDNA was used as a template for PCR amplification of the gene encoding zebrafish AC (asah 1b) using primers asah1b forward, (5′-CGCGCGGAATTCCAATATGTACCACCGTTCAC-3) and asah1b reverse, (5′-ACTAGTGTCGACCCAGGGCATGCAGGGATTTG-3). *C. elegans* cDNA was used as a template for PCR amplification of the gene encoding *C. elegans* AC (asah 1) using primers asah1 forward, (5′-ACTAGTGAATTCAAGCATGTGGAGCTTCCGGC -3) and asah1 reverse, (5′-ACTAGTGTCGACCCATGGATAGCATTCTCCCG -3). The underlined sequences correspond to the restriction sites of EcoRI and SalI, respectively. These sites were selected to insert the gene into the pPICZα-A vector, in order to ensure that the recombinant DNA fragment had an alpha-factor signal peptide preceding its N-terminal, followed by a C-terminus 6× His extension under the AOX promoter. The gene was amplified using Ex Taq DNA polymerase (TaKaRa) under the following conditions: 1 cycle of pre-denaturation at 94°C (1 min) followed by 28 cycles of denaturation at 94°C (30 s), primer annealing at 58°C (30 s) and extension at 72°C (1:30 min), a final extension at 72°C (10 min) and then cooled to 4°C. The PCR amplified products were analyzed on 1% agarose gel. The amplicon of the desired size was gel extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704).

Table 2.1: PCR reaction mixture for amplification of AC gene PCR.

Sr. No.	PCR mix	asah 1b	asah1
1	TaKaRa Ex Taq (5 units/ µl)	0.25 µl	0.25 µl
2	Ex-Taq buffer(10x)	5 µl	5 µl
3	dNTP (2.5 mM)	4 µl	4 µl
4	Forward Primer	1 µl	1 µl
5	Reverse Primer	1 µl	1 µl
6	ddH_2O	38.75 µl	38.75 µl

2.2.1.4 Restriction digestion and ligation

The asah1b and asah1were digested with EcoRI and SalI at 37°C with suitable buffers for 4 hours and ligated into pPICZα-A vector. Approximately 1:3 ratio of vector to insert was incubated with T4 DNA ligase from Invitrogen with suitable buffer at 16°C overnight.

2.2.1.5 Transformation of cloned plasmid asah1b-pPICZα-A and asah1-pPICZα-A

The asah1b-pPICZα-A and asah1-pPICZα-A constructs were transformed into competent *E. coli* DH5α. Positive colonies were selected on low salt LB zeocin plates (1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl and 2% w/v agar and zeocin at 25

µg/ml of media) and gene insertion was confirmed by colony PCR using 5′AOX1 vector specific forward primer and insert specific reverse primer. The clone was further confirmed by RE double digestion and sequencing.

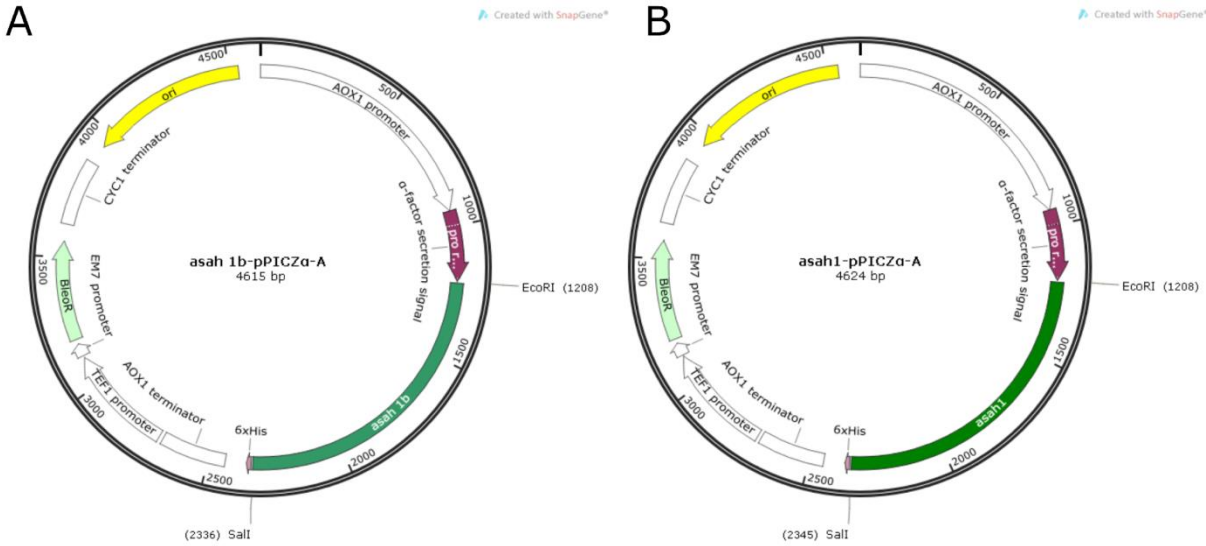


Figure 2.1: Map of vectors A: asah1 b-pPICZα-A. B: asah1-pPICZα-A.

2.2.2 Site-directed mutagenesis

Different mutants of wild-type zebrafish AC i.e. Y36C, T42M and N152K were prepared using WT zebrafish AC plasmid (asah1b-pPICZα-A). PCR was performed at 95 ºC for 3 min and 18 cycles of (94/ 30 sec, 55 ºC for 30 sec, 72 ºC 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 ºC for 3 hrs. 5 ul of PCR mixture is further is transformed in *E.coli* DH5α cells. Plasmids were purified and sequenced for confirmation of the desired mutation. List of primers used for the desired mutations are listed in table 2.2

Table 2.2: List of primers for generating site-directed mutagenesis in wild type zebrafish AC.

Primer	5'-3' sequence
Zebrafish AC Y36C FP	TGC AGA AGT GGC ATG TGT CCT CCA AAT GGC CCC
Zebrafish AC Y36C RP	GGG GCC ATT TGG AGG ACA CAT GCC ACT TCT GCA
Zebrafish AC T42M FP	CCT CCA AAT GGC CCC ATG TTT AAA GGT GAT GTG
Zebrafish AC T42M RP	CAC ATC ACC TTT AAA CAT GGG GCC ATT TGG AGG
Zebrafish AC K152N FP	GTT GCT GAG GAC GTC AAG GGC AAT CTC ATC CAT
Zebrafish AC K152N RP	ATG GAT GAG ATT GCC CTT GAC GTC CTC AGC AAC

2.2.3 Expression asah1b and asah1 in *Pichia pastoris*

The asah1b-pPICZα-A and asah 1-pPICZα-A constructs were linearized using PmeI restriction enzyme and transformed into competent *Pichia pastoris* (GS115 strain) by electroporation (4 kΩ, 50 μF, 400 V) (Bio-Rad Gene Pulser). The transformants were selected on YPDS-zeocin plates (1% w/v yeast extract, 2% w/v peptone, 2% w/v dextrose, 1M sorbitol, 2% w/v agar and zeocin at 1000 μg/ml of media) and colony PCR using the 5′ α-factor and 3′AOX1 primers. The selected clone was cultured into 25 ml of BMGY medium and grown at 28 °*C* till OD_{600} reaches to 2-6. These cells were re-suspended in 300 ml of BMMY medium and incubated at 28°*C*. This secondary culture was induced by the addition of 0.5% (v/v) methanol at 24 h intervals. The cell-free supernatant was concentrated by lyophilization and then was dialyzed through Millipore 12-kDa cut-off membrane to exclude ions and salts.

2.2.4 Confirmation of expression

The purity, molecular weight and autoproteolytic processing of the protein were assessed on SDS-PAGE by coomassie brilliant blue (CBB) R-250 and silver staining. The molecular weight was estimated by using a prestained dual color protein ladder (10-250 kDa, Bio-Rad) as standard. Western blotting was achieved using anti-His mouse monoclonal antibodies (Qiagen) and anti-ASAH1 rabbit polyclonal antibodies directed against human recombinant ASAH1 (Sigma). The glycosylation staining was done by glycoprotein staining kit (Pierce). Deglycosylation was done with PNGase F enzyme (New England Biolabs) and protein deglycosylation Mix II (β1-4 Galactosidase S, O-Glycosidase, β-N-Acethylhexosaminidase, Neuraminidase A, PNGase F) (New England Biolabs) according to manufacturer's protocol.

2.2.5 Protein purification

The C-terminal His-tagged zebrafish AC and *C.elegans* AC protein were purified using Ni-NTA affinity chromatography. The supernatant with Ni^{2+}- agarose beads were incubated for overnight for binding at 4°C with shaking. After incubation supernatant was loaded on to a column. The matrix was then washed with equilibration buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM DTT, pH 8) followed by removal of non-specific and weakly bound proteins by washing with wash buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM DTT and 10 mM imidazole, pH 8). The matrix-bound zebrafish AC enzyme was eluted by passing the elution buffer (50 mM Tris-HCl, 300 mM NaCl, 10% v/v glycerol, 2 mM DTT and 250 mM imidazole, pH 8). The high concentration of imidazole in elution fraction was removed by

buffer exchange using a 10-kDa cut-off ultrafiltration membrane (Millipore). The protein solution obtained was finally subjected to size exclusion chromatography using sephacryl S200 (16/60) column, pre-equilibrated with Tris buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 2 mM DTT and 10% v/v glycerol). Fractions containing zebrafish AC activity were pooled and analyzed by running 12% SDS-PAGE and western blot. Gel filtration standard (Bio-Rad) was used as a standard for estimation of the native molecular mass of protein.

2.2.6 Autoproteolytic cleavage assay

Zebrafish AC was incubated at 37°C in autoproteolytic assay buffer for 96 hours (100 mM sodium acetate buffer, pH 4, 2 mM DTT and 0.1% v/v NP-40 substitute) (Shtraizent *et al.*, 2008). Sampling was done at 24-hour intervals and visualized in 10% SDS-PAGE.

2.2.7 Qualitative enzyme assay

In vitro, AC activity assay was modified and performed as previously described (Mizutani *et al.*, 2015). The 50 µg of zebrafish AC and 50 µM of fluorescent C12-NBD-Ceramide (Sigma) substrate were mixed in 50µl Assay buffer I (25mM acetate buffer pH 5.6, 0.5 % v/v Triton X-100). The reaction was performed at 37°C for 90 min and stopped by addition of 125 µl methanol/chloroform (1:2) (v/v). The organic phase was collected and dried. Lipids were dissolved in 5µl methanol/chloroform (1:2) then spotted on TLC Silica gel 60 plates (Merck). The substrate and product were separated by developing the plates in chloroform/methanol/25% v/v ammonium hydroxide (90:30:0.5) solvent system. The TLC plate was observed under UV and RF value was calculated. AC activity assay was also performed using substrates C6-NBD-Ceramide (Avanti#:810209) and Omega-NBD-18:0 Ceramide (Avanti#:810210).

2.2.8 Reverse AC assay

In vitro, reverse AC activity assay was modified and performed as previously described (Okino *et al.*, 2003). The 50 µg of zebrafish AC and 50 µM of fluorescent N-(NBD-aminolauroyl) fatty acid (Sigma) and Sphingosine (Avanti#: 860490) substrate were mixed in 50µl Assay buffer II (200 mM citrate-phosphate buffer pH 6.0, 300mM NaCl, 0.05 % Triton X-100). The reaction was performed at 37°C for 90 min and spotted on TLC plate and separated.

2.2.9 Quantitative enzyme assay

AC enzyme activity is measured by estimating D-Sphingosine released by hydrolysis activity on incubating the AC enzyme with ceramide (N-Lauroyl-D-sphingosine) at 37°C in assay buffer III (100 mM acetate buffer pH 4, 0.1 % v/v NP-40 substitute). 500 µl of the reaction mixture was withdrawn and an equal volume of OPA was added and kept at RT for 2 min. The product formed by the reaction was recorded at 340 nm. One unit of AC activity is defined as the amount of enzyme that liberates 1 µmol of D-Sphingosine from the substrate per minute. Specific activity was defined as the number of units of activity per milligram of the pure protein. Protein concentration was measured by the Bradford method. D-Sphingosine (Sigma) standard graph was prepared to quantify the product released on substrate hydrolysis (Figure. 2.3).

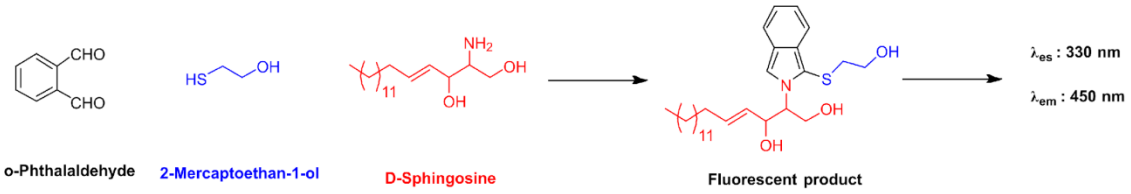


Figure 2.2: Detection of primary amines of D-Sphingosine with OPA. Hydrolyzed D-Sphingosine reacts with OPA in the presence of a nucleophile (2-mercaptoethanol) and under alkaline conditions yields highly fluorescent product that is quantified photometrically at 340 nm.

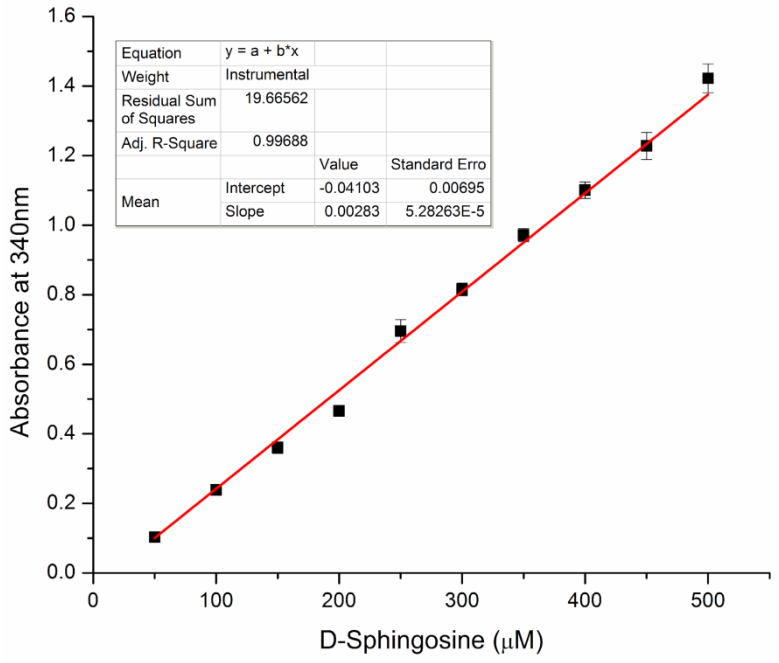


Figure 2.3: Standard graph of D-Sphingosine (µM) estimation.

2.2.10 Effect of time, pH, temperature and DTT on AC activity and stability

The AC activity was assayed at different time value in the range 1-36 hours, DTT 0-10 mM, pH values in the range 3.0-10 and temperatures 10-70°C to determine the optimum conditions for enzyme activity.

2.2.11 Circular dichroism (CD) spectroscopy

The CD spectra of the enzyme were recorded on a Jasco J-815-150S spectropolarimeter (Jasco, Tokyo, Japan) with a PTC 343 Peltier unit at 25°C in a quartz cuvette. Far CD spectra were recorded as a mean of three scans and expressed in MRE. The purified enzyme at a concentration of 0.2012 mg ml^{-1} was used for the entire sample. All spectra were corrected for buffer contributions and observed values were converted to mean residue ellipticity (MRE) in deg cm2 dmol-1 defined as

MRE = Mθλ/10dcr

Where M is the molecular weight of the protein, θλ is CD in millidegree, d is the path length in cm, c is the protein concentration in mg/ ml and r is the average number of amino acid residues in the protein. The relative content of secondary structure elements was calculated by using CD pro software (Sreerama and Woody, 2000). Low NRMSD values were observed for analysis with CONTINLL and SELCON.

2.2.12 Steady-state kinetic parameters for zebrafish AC

The kinetic constants $K_{0.5}$ and V_{max} of zebrafish AC were estimated by assaying the enzyme activity with increasing concentrations of N-Lauroyl-D-Sphingosine (Sigma) as substrates in the range 100–1600 µM following the standard assay protocol. The assay was carried out in triplicates and the kinetic constants (K0.5 V_{max}, hill coefficient 'h') were determined by non-linear regression by fitting the kinetic data in Origin version 8. The *k*cat value for N-Lauroyl-D Sphingosine calculator by non-linear regression by fitting the kinetic data(real velocity in µM/s) in GraphPad Prism version 5.0 with constraints enzyme concentration (0.04 µM) and Km (273.6544 µM).

$[V_{max} .Sh] / [K0.5 + Sh + (S2h/Ki)]$ ……..... (2.1)

2.3 Results

2.3.1 Cloning of asah 1b gene of zebrafish AC and asah1 of *C.elegans* AC

RNA was isolated from the N2 strain of *C. elegans* at a concentration of 1153 ng ul^{-1} (Figure 2.5). The cDNA library prepared was utilized to amplify the asah1 gene of size 1131 bp with the primers. We also amplify the asah2 and naaa to check the presences of acid ceramidase-like protein (NAAA) and isoform of AC.

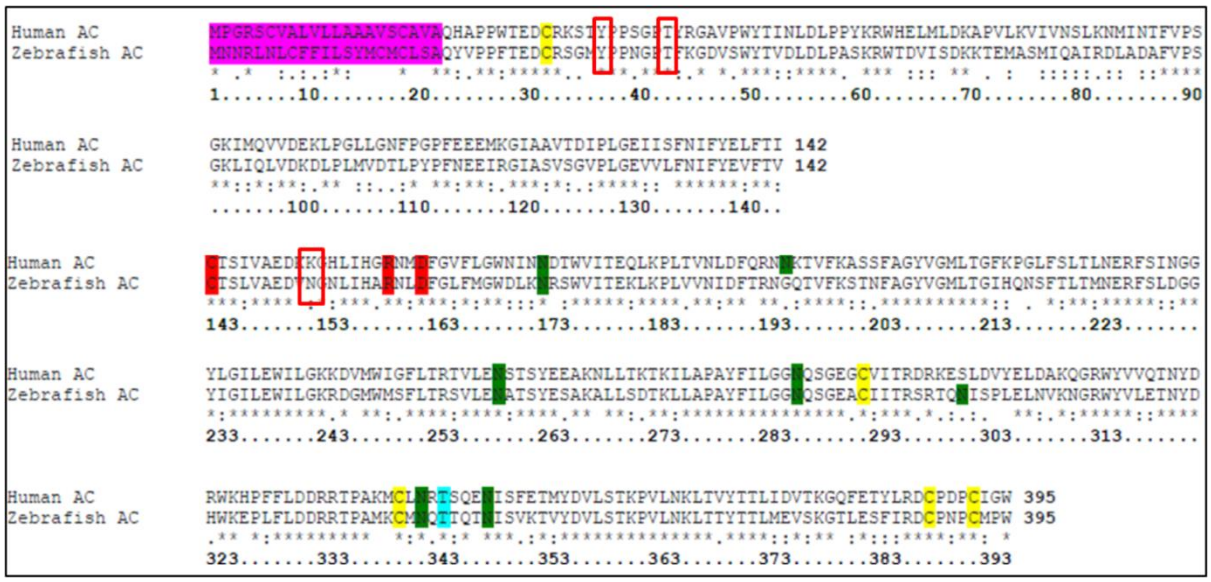


Figure 2.4: Sequences alignment of acid ceramidase of human and zebrafish analyzed using clustal W. The red color amino acids are catalytic residues of acid ceramidase. N-glycosylation site shown in green color predicated using NetNGlyc 1.0 server. The conserved cysteine residues involved in disulfide bonds shown in yellow color.

The C. elegans AC consists of 393 amino acid residues with a 15 amino acid at N-terminal that, functions as a signal peptide. The catalytic triad of C. elegans AC consists of Cys 140, Arg 156 and Asp 159 amino acid residues. C. elegans asah1 was cloned into vector pPICZα-A between N-terminal the yeast α-factor secretion signal and c-terminal 6X His tag. C. elegans AC translates into a 388 aa protein, theoretical molecular weight of 44 kDa having pI at pH 5.34 (http://web.expasy.org/protparam/).

Zebrafish AC consists of 395 aa residues having a 21 aa at N-terminal signal peptide. Active site residues consist of Cys 143, Arg 159 and Asp 162; these were predicted via sequence conservation analysis of acid ceramidase protein family. The asah1b gene without the signal peptide sequence was fished out from zebrafish cDNA using asah1bF and asah1bR primers. Zebrafish asah1b was cloned into the vector pPICZα-A between N-terminal the

yeast α-factor secretion signal and c-terminal 6X His tag. The resulting plasmid was sequenced and named as asah1b-PICZα-A. Zebrafish AC translates into a 384 aa protein, theoretical molecular weight of 44 kDa having pI at pH 5.22.

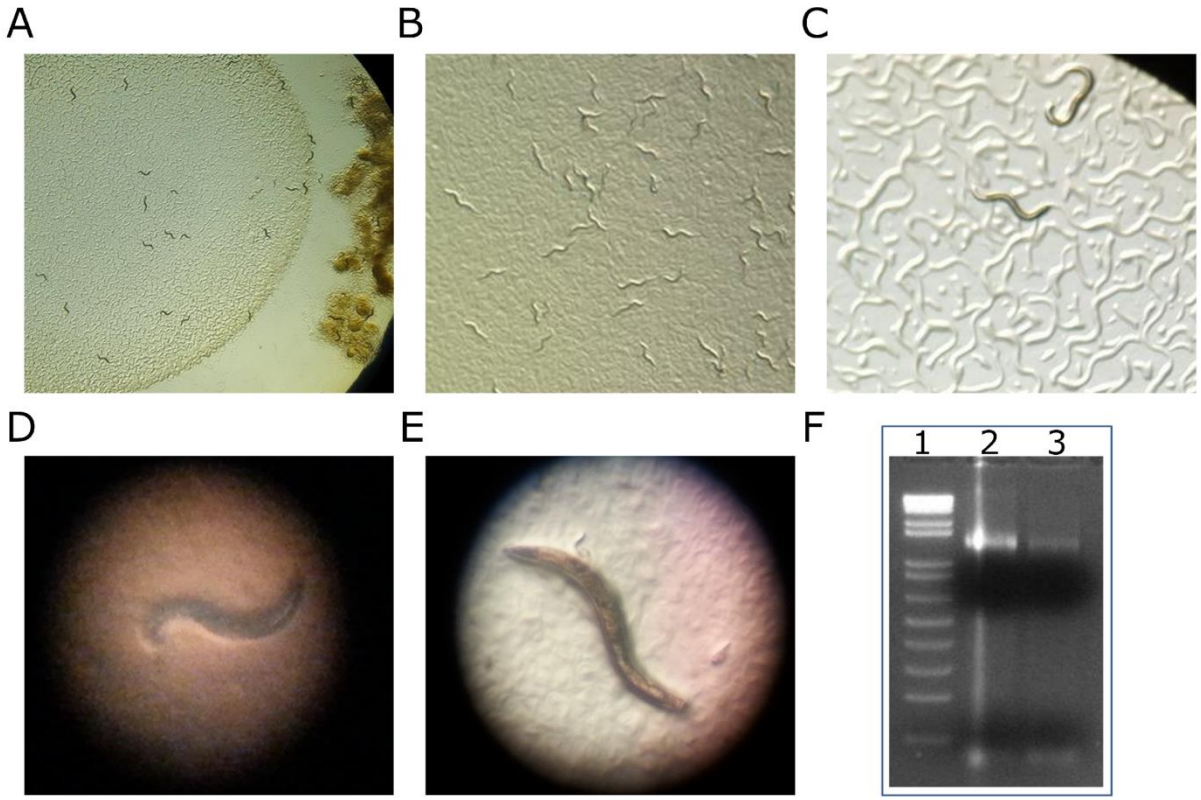


Figure 2.5: Total RNA isolation from *C. elegans.* A: Egg and L1 larva. B: L2 larva. C: L3 larva. D: L4 larva. E: Adult C. elegans. F: Lane 1, NEB 1 Kb plus DNA ladder; Lanes 2 and 3, *C. elegans* total RNA, electrophoresed on 2% agarose gel.

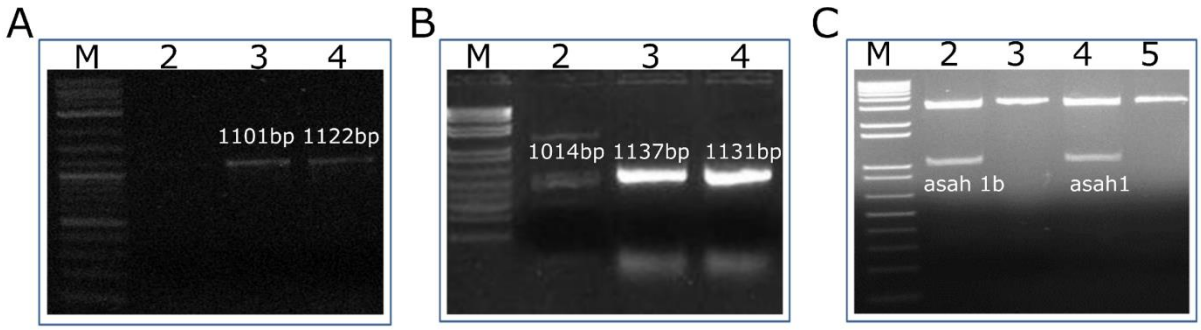


Figure 2.6: Cloning of asah1b and asah1. A: 1% agarose gel showing PCR product of *zebrafish* AC gene. Lane 1, NEB 1Kb plus DNA ladder; Lane 2, PCR blank as negative control; Lane 3, 1101 bp PCR product of asah 1a; Lane, 1122 bp PCR product of asah 1b. B: 1% agarose gel showing PCR product of *C. elegans* AC and NAAA gene. Lane 1, NEB 1Kb DNA ladder; Lane 2, 1014 bp PCR product of naaa; Lane 3, 1137 bp PCR product of asah 2; Lane 4, 1131 bp PCR product of asah 1. C: Restriction digestion. Lane 1, NEB 1Kb plus DNA ladder; Lane 2, Double digestion (*EcoRI/SalI*) of cloned plasmid (asah1b-pPICZα-A); Lane 3, Undigested cloned plasmid (asah1b-pPICZα-A); Lane 4, Double digestion (*EcoRI/SalI*) of cloned plasmid (asah1-pPICZα-A); Lane 5, Undigested cloned plasmid (asah1-pPICZα-A).

2.3.2 Heterologous expression of zebrafish AC

The recombinant P. pastoris GS115 strain having asah-1b gene was inoculated in a 500 ml shake flask with 25 ml BMGY and grown for 48 h at 28 °C, 200 rpm. Cell pellets were harvested by centrifugation and resuspended in a 2 L shake flask with 300 ml BMMY. Methanol induction was given every 24 h for asah-1b gene expression. The culture was grown in BMMY medium for 144 h at 28 °C, 200 rpm.

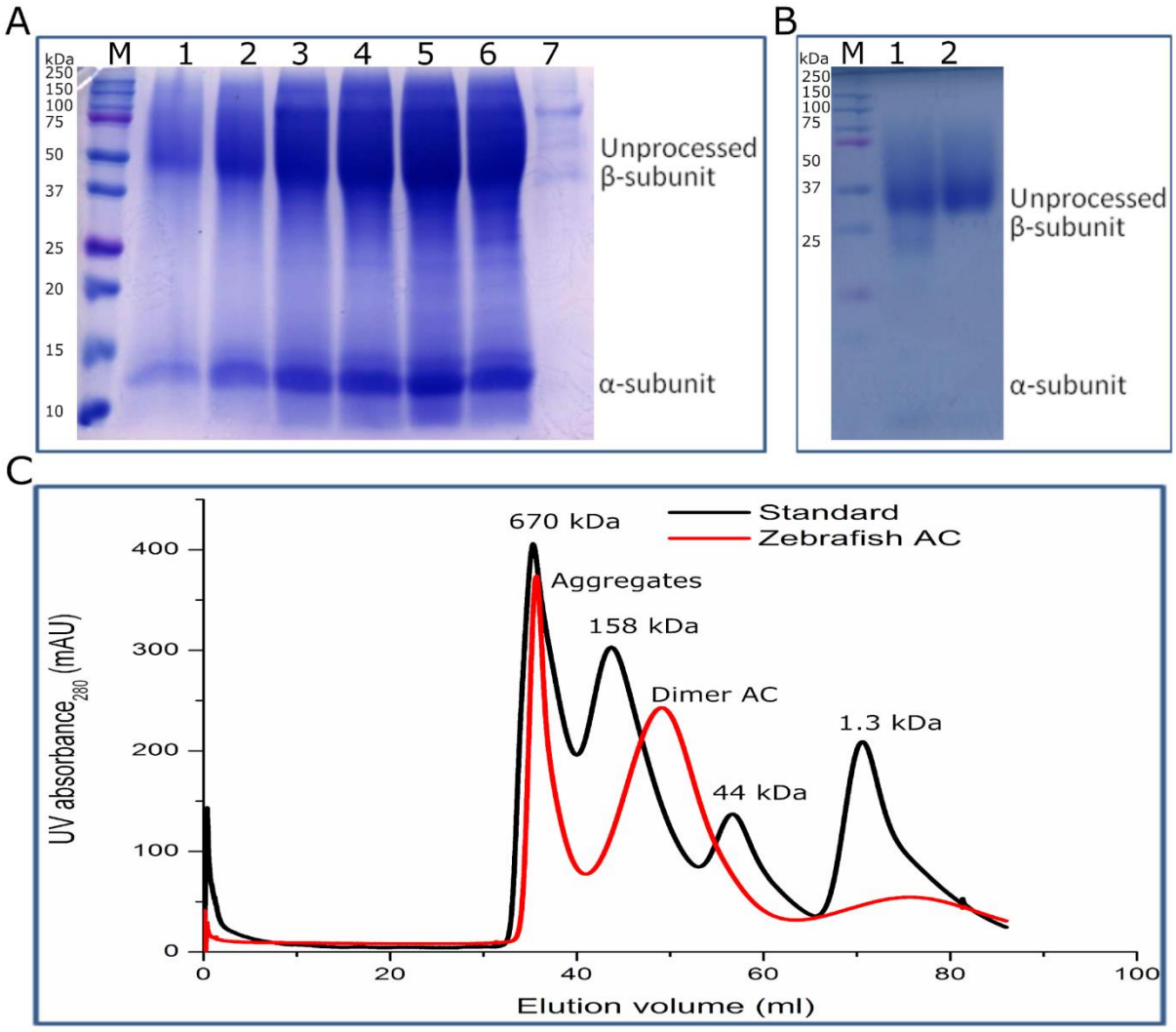


Figure 2.7: Purification of zebrafish AC. A) 12 % SDS-PAGE analysis shows expression of zebrafish AC protein in *pichia*. M, Prestained protein ladder; Lanes 1-6, 20 µl of the concentrated supernatant after being induced by methanol for 24, 48, 72, 96, 120 and 144 h, respectively; Lane 7, 20 µl of the concentrated supernatant obtained from GS115 cells containing empty pPICZα-A plasmid induced for 120 h. B: 12% W/V SDS-PAGE analysis shows purification of zebrafish AC protein by sequential chromatography. M, Prestained protein ladder; Lane 1, Zebrafish AC protein was purified using Ni-NTA affinity chromatography; Lane 2, Zebrafish AC protein was purified by AKTA purifier with size exclusion chromatography. C: Chromatogram of size exclusion chromatography using sephacryl S200 (16/60) column for purification of zebrafish AC protein.

Since the inserted gene is present downstream to the secretory signal sequences, it was expected that the expressed gene product would be secreted out in the medium. The culture medium was duly collected and cell debris was removed. The cell-free supernatant was concentrated and desalted by dialysis. The supernatant was subsequently analyzed on 10% w/v SDS-PAGE. Protein bands of unprocessed AC approximately 50 kDa were observed along with a significant amount of processed form of AC approximately 30-40 and 14 kDa, α and β-subunits respectively (Figure 2.9). As expected, the same protein bands were not detected for pPICZα-A transformed cells used as the control. Expression of zebrafish AC gradually increased over time until the 5th day. Expression was further confirmed using western blot and enzyme assay.

Table 2.3: Protein component used as gel filtration standards.

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

2.3.3 Purification and confirmation of zebrafish AC

Zebrafish AC was purified from approximately 300 ml of cell-free medium, 72 hours after induction. A two-step purification protocol, using Ni-NTA followed by size exclusion chromatography was employed for zebrafish AC purification. As shown in the chromatogram, three peaks were obtained (Figure 2.7.C). These were further analyzed by SDS-PAGE and checked for activity. Only one peak has activity and bands of interest on SDS-PAGE (Figure 2.7 B: lane 2). Majority of the purified enzyme was obtained in the unprocessed form; however, small amounts of α- and β-subunits were also obtained (Figure 2.7 B: lane 1 and 2). The broad peak in size exclusion chromatography and SDS-PAGE suggests that the purified enzyme fraction is heterogeneous due to glycosylation. Western blotting using anti-His antibodies detected unprocessed AC band along with the β-subunit. (Figure 2.8 A: lane 1 and 2). The unprocessed AC band merged with the β-subunits band on SDS-PAGE may be due to glycosylation but clearly resolved in western blotting. The western blot results showed that β-subunit have 4 detectable different bands; these bands were further investigated for glycosylation study.

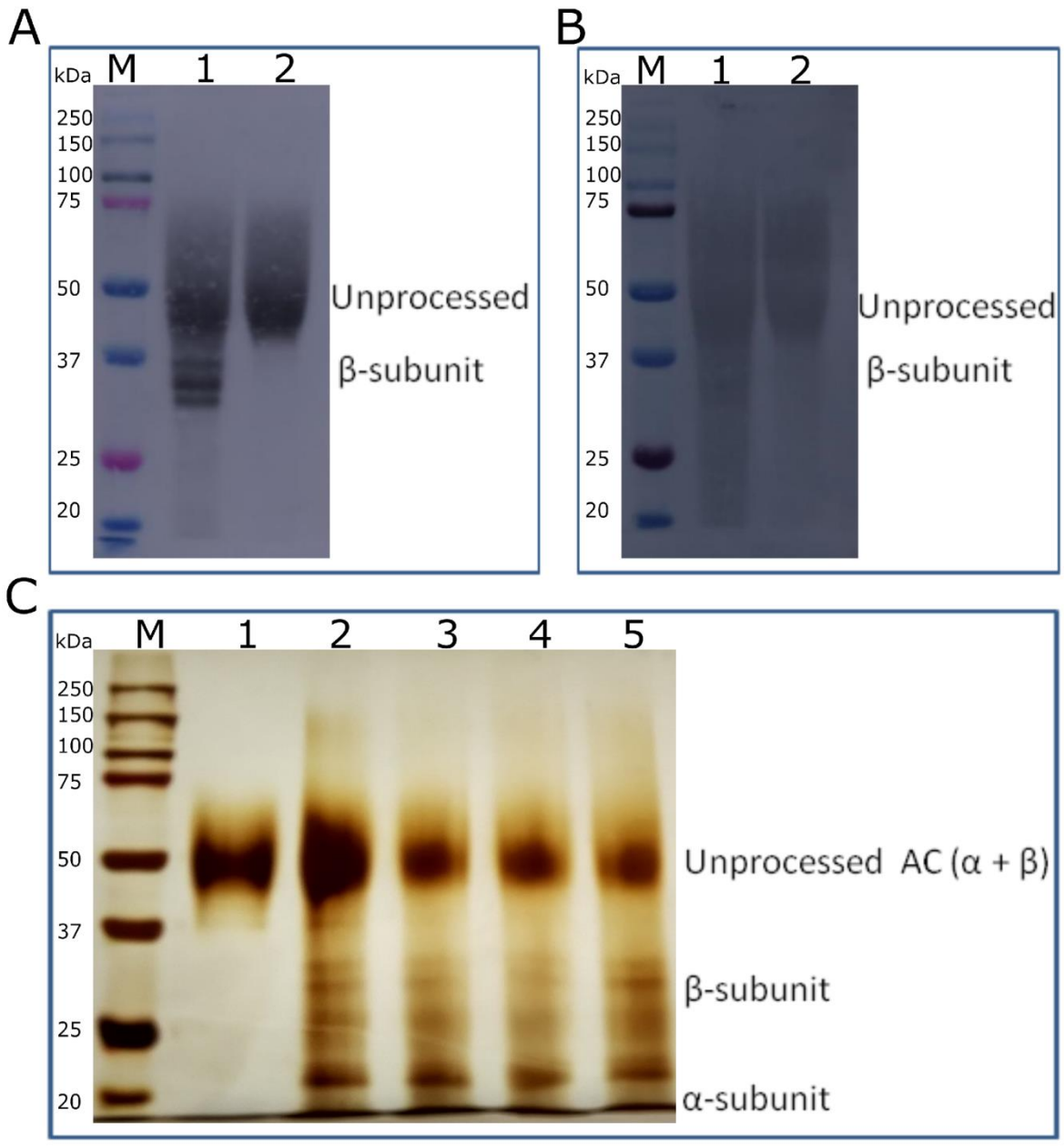


Figure 2.8: Confirmation of zebrafish AC. A: Western blotting analysis confirms zebrafish AC protein expression with monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 1, Zebrafish AC protein was purified using Ni-NTA affinity chromatography; Lane 2, Zebrafish AC protein was purified by AKTA purifier with size exclusion chromatography. B: Western blotting analysis confirms zebrafish AC protein expression with polyclonal anti-human AC antibody. M, Prestained protein ladder; Lane 1, Zebrafish AC protein was purified using Ni-NTA affinity chromatography; Lane 2, Zebrafish AC protein was purified by AKTA purifier with size exclusion chromatography. C: Self-processing of purified recombinant zebrafish AC. M, Prestained protein ladder; Lanes 1-5, Zebrafish AC of incubated at 37 ^{o}C in autoproteolytic assay buffer analyzed on 10% SDS-PAGE with time interval such as 0, 24, 48, 72 and 96 hrs, respectively.

Polyclonal anti-human AC antibody showed cross-reactivity against zebrafish AC (Figure 2.8 B: lane 1 and 2). The purified zebrafish AC fraction showed activity towards the substrate NBD-C12-ceramide as substrate. Protein yield obtained was ~10 mg/l of culture is more compared to human AC expressed in CHO, Sf21 and SF9 cells ~8 mg/l, 1 mg/l and 80 µg/l, respectively (He *et al.*, 2003; Schulze, Schepers and Sandhoff, 2007; Chien, 2009). This study illustrates that the P. pastoris expression system is efficient for the heterologous expression of proteins undergoing autoproteolytic processing such as zebrafish AC.

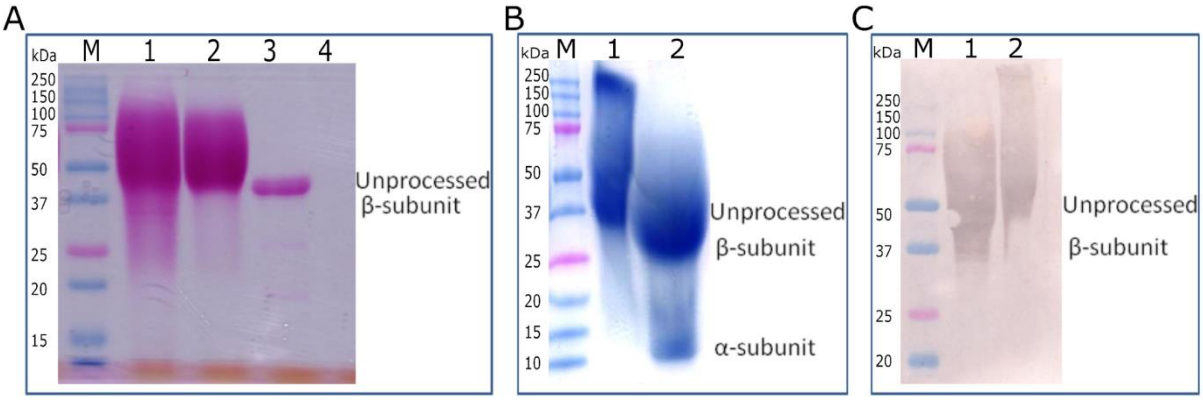


Figure 2.9: Glycosylation analysis and electrophoretic mobility shift assay of zebrafish AC. A: Glycosylation analysis of zebrafish AC protein by glycosylation staining kit. M, Prestained protein ladder; Lane 1, Zebrafish AC protein was purified using Ni-NTA affinity chromatography; Lane 2, Zebrafish AC protein was purified by AKTA purifier with size exclusion chromatography; Lane 3, Horseradish peroxidase as a positive control; Lane 4, Soybean trypsin inhibitor as a negative control. B: 8-16% SDS-PAGE analysis showing that the electrophoretic mobility of the cleaved form of zebrafish AC changes in reducing and non-reducing conditions. M, Prestained protein ladder; Lane 1, Zebrafish AC protein without β-mercaptoethanol (BME) and heat; Lane 2, Zebrafish AC protein with 5% BME and heat. B: Western blotting analysis showing that the electrophoretic mobility of the cleaved form of zebrafish AC changes in reducing and non-reducing conditions. M, Prestained protein ladder; Lane 1, Zebrafish AC protein with 5% BME and heat; Lane 2, Zebrafish AC protein without β-mercaptoethanol (BME) and heat.

2.3.4 Molecular mass determination of zebrafish AC

The native molecular mass of zebrafish AC was found to be ~100 kDa, as estimated by size-exclusion chromatography using a sephacryl-S200 column in FPLC system when compared with molecular weight standards to determine the mass (Figure 2.7 C). The molecular mass corresponding to ~50, ~30-40 and 14 kDa was α-, β-subunits and unprocessed zebrafish AC respectively estimated on SDS-PAGE (Figure 2.7 B). The difference between masses on size-exclusion chromatography, SDS-PAGE suggested that zebrafish AC exists as a homodimer of the heterodimer (α + β) having two α-subunit and two

β-subunits. Our findings resemble previous findings of the human AC as homotrimers of the heterodimer in urine (He et al., 2003).

2.3.5 Autoproteolytic activation of zebrafish AC

Zebrafish AC like human AC is also subjected to proteolytic cleavage during maturation. With respect to the similarity between the primary structure of zebrafish AC and human AC, zebrafish AC Cys 143 corresponds to Cys 143 of hAC, the N-terminal residue of the β-subunit. Western blotting confirmed zebrafish AC cleavage, the anti-his antibody binds to only the C-terminal 6X His-tagged zebrafish AC. Based on the subunit molecular masses obtained from SDS-PAGE, Cys 143 was predicted to be the cleavage site. The difference in electrophoretic mobility between mature and immature zebrafish AC on SDS-PAGE in the presence and absence of β-mercapto ethanol suggests (Figure 2.9 B and C), zebrafish AC forms disulfide bond linked αβ heterodimer (Bernardo *et al.*, 1995). The ratio of immature protein to mature α and β-subunit in the purified varied to a moderate extent from batch to batch. This may be due to, the zebrafish AC is secreted as an unprocessed non-numeric precursor (immature form) and then slowly gets converted into matures heterodimer during expression and purification. The 15 µg of zebrafish AC was incubated for 96 h in autoproteolytic clevage assay buffer (pH = 4) at 37 °C for autoproteolytic processing. The processing of the zebrafish AC was assessed at 24hour time intervals from 0 hours till 96 hours, by running on 10% SDS-PAGE and visualizing α and β-subunits by silver staining. (Figure 2.8 C). The intensity of α and β-subunit slightly increased with one more new band at 22 kDa. The unprocessed form slowly converts into mature processed in autoproteolytic clevage assay buffer (pH = 4) at 37 °C after 24 h. The highest activity observed for more than 24 hrs as compared to unincubated protein. The unprocessed form did not entirely convert into the mature form may be due to the oxidation of the active site cysteine.

2.3.6 Glycoprotein analysis of zebrafish AC

On the basis of gel electrophoretic results, the molecular masses corresponding to the bands ~50, ~30-40 and 14 kDa were observed with slight variation in the molecular weight than the predicted *in silico* amino acid sequence 44, 30 and 14 kDa for unprocessed AC, β- and α- subunit, respectively. Six putative N-glycosylation sites (Asn-Xaa-Thr/ser) and one potential O-glycosylation site were predicted in the zebrafish AC on sequence analysis on NetNGlyc 1.0 (Gupta, Jung and Brunak, 2016) and NetOGlyc 4.0 Server (Steentoft et al., 2013), respectively. Six N-glycosylation sites at amino acid positions 173, 259, 286, 301, 342

and 348 and a single O-glycosylation site at positions 344 present on the β-subunit (Figure 2.10).

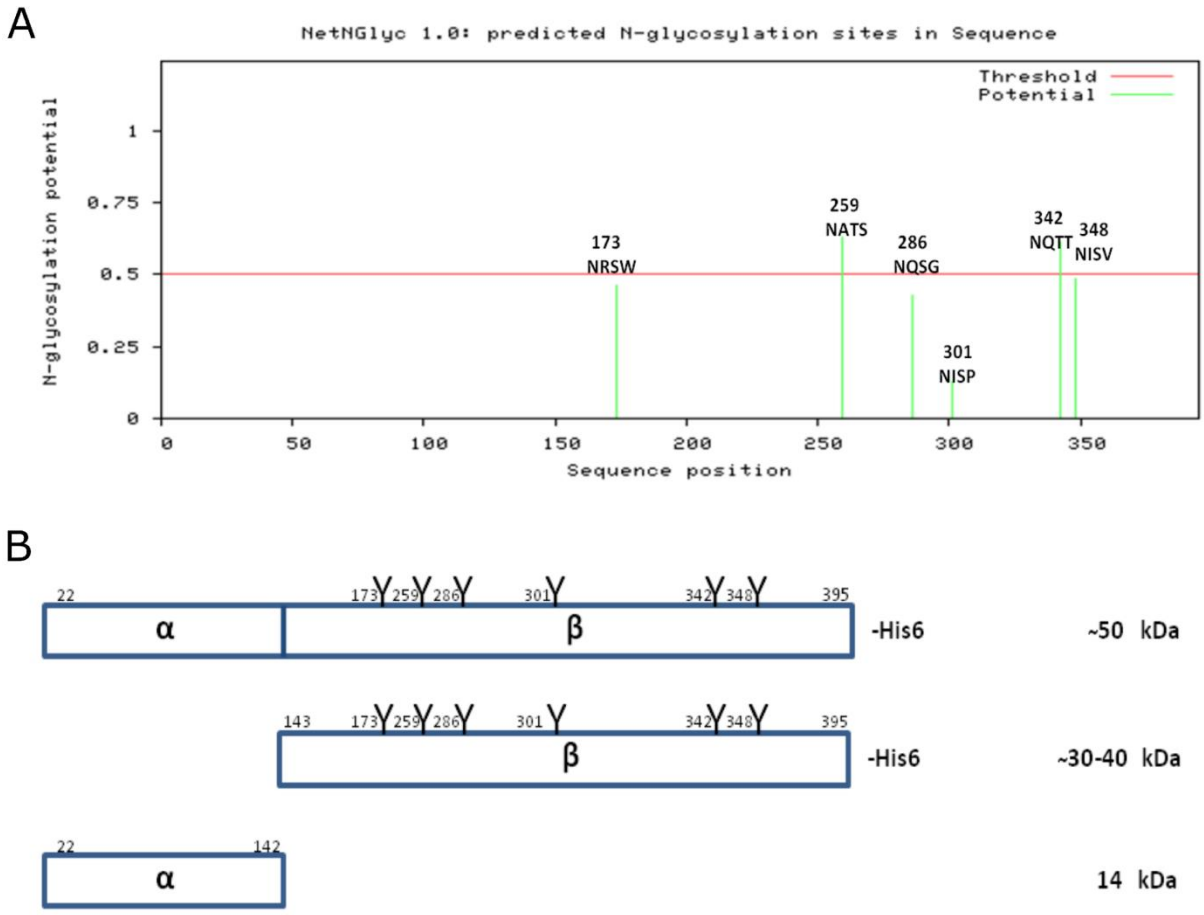


Figure 2.10: The predicted N-glycosylation site and observed molecular weight of zebrafish AC. A: Predicted N-glycosylation site in zebrafish AC using NetNGlyc 1.0 server B: The schematic diagram of recombinant zebrafish AC showed that different subunit with sequence number, glycosylation number and molecular weight.

The α-subunit does not have any putative glycosylation site. The purified zebrafish AC is known to have heterogeneity due to post-translational processes such as glycosylation and autoproteolytic cleavage. Kit based Glycoprotein staining on zebrafishAC confirmed glycosylation. The ~50 kDa and ~30-40 kDa band were stained magenta, whereas the 14 kDa band did not, implicating glycosylation of ~50 kDa and ~30-40 kDa bands which were the unprocessed and the β-subunit of zebrafish AC respectively (Figure 2.7 B and 2.9 A). These results also indicate that α-subunit having 14 kDa bands is nonglycosylated similar to human AC (Figure 2.7 B) (Bernardo *et al.*, 1995; Ferlinz *et al.*, 2001). Deglycosylation using enzyme PNGaseF appeared to decrease the protein size to about ~44, ~30 and 14 kDa (Figure 2.11 A, B and C). The β-subunit had four bands in western blot but after deglycosylation by

PNGase F, the β-subunit had only two bands in western blot, suggesting that there were two different poly peptides, as a result of carboxy-terminal processing or which may reflect different glycosylation status, similar to human AC (Figure 2.8 A and 2.11 B) (He *et al.*, 2003).

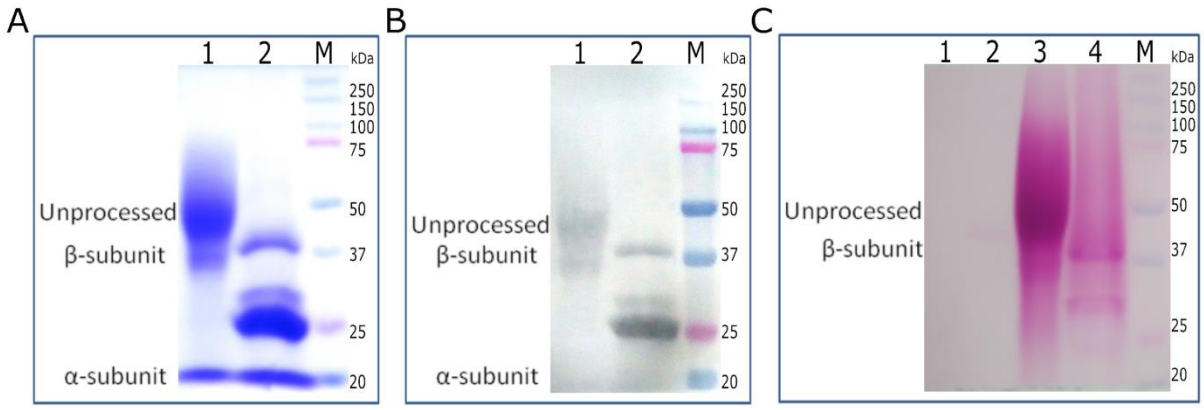


Figure 2.11: N-glycosylation analysis of zebrafish AC. A: 100 µg of zebrafish AC protein were treated with PNGase F for 1 h, separated by 10% SDS-PAGE. Lane 1, Zebrafish AC protein; Lane 2, Zebrafish AC protein was treated with PNGase F; M, Prestained protein ladder. B: 100 µg of zebrafish AC protein were treated with PNGase for 1 h, separated by SDS–PAGE and analyzed by western blotting using with monoclonal anti-His antibodies. Lane 1, Zebrafish AC protein; Lane 2, Zebrafish AC protein was treated with PNGase F; M, Prestained protein ladder. C: 100 µg of zebrafish AC protein were treated with PNGase F for 1 h, separated by SDS-PAGE and analyzed by glycosylation staining kit. Lane 1, Soybean trypsin inhibitor as negative control; Lane 2, Horseradish peroxidase as a positive control; Lane 3, Zebrafish AC protein; Lane4, Zebrafish AC protein was treated with PNGase F; M, Prestained protein ladder.

In order to make sure, the exact nature of the two forms of the β-subunit, we carried out glycosylation staining, where the "prominent" band remained unstained, proving that it was non-glycosylated, while the other one got stained and thereby can be concluded to consist of glycosylated β-subunit. However, we were curious about the nature of the glycosylation present in the relatively fainter band and we decided to treat the same with Protein Deglycosylation Mix II (PDM II) to check if there was any O-glycosylation present. We found that the fainter band disappeared on treatment with PDM II, indicating that it contained O-glycosylated moieties (Figure 2.11 A, B and C). Thereby we can infer that the variation in β-subunit is entirely due to N- and O-glycosylation and not due to carboxy-terminal truncation, unlike human AC (He *et al.*, 2003). After deglycosylation, the intensity of the β-subunit band was found to be more than the intensity of the unprocessed AC band, which indicated that the β-subunit band merges with the unprocessed AC in glycosylated

condition. The glycosylation content of the recombinant enzyme clearly confirmed by deglycosylation using PNGase F and PDM II, glycoprotein staining kit but the structure of oligosaccharides have not been investigated.

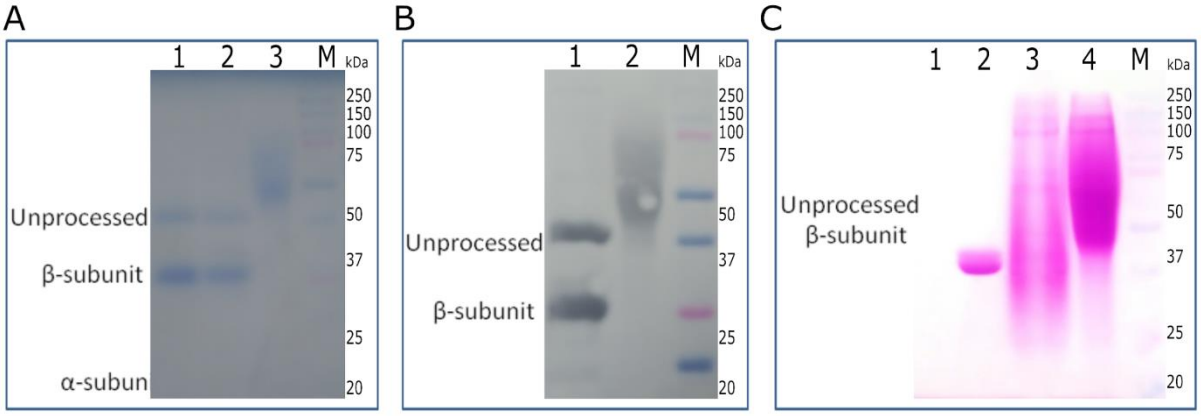


Figure 2.12: O-glycosylation analysis of zebrafish AC. A: 100 µg of zebrafish AC protein were treated with deglycosylation mix, separated by 10% SDS-PAGE. Lane 1, Zebrafish AC protein was treated with deglycosylation mix for 16 h; Lane 2, Zebrafish AC protein was treated with deglycosylation mix for 24 h; Lane 3, Zebrafish AC protein; M, Prestained protein ladder. B: 100 µg of zebrafish AC protein were treated with deglycosylation mix for 16 h, separated by SDS-PAGE and analyzed by western blotting using with monoclonal anti-His antibodies. Lane 1, Zebrafish AC protein was treated with deglycosylation mix; Lane 2, Zebrafish AC protein; M, Prestained protein ladder. C: 100 µg of zebrafish AC protein were treated with deglycosylation mix for 16 h, separated by SDS-PAGE and analyzed by glycosylation staining kit. Lane 1, Soybean trypsin inhibitor as a negative control. Lane 2, Horseradish peroxidase as a positive control; Lane 3, Zebrafish AC protein was treated with deglycosylation mix; Lane 4, Zebrafish AC protein; M, Prestained protein ladder.

2.3.7 Enzymatic activity of zebrafish AC

The purified recombinant zebrafish AC was used for enzyme activity assays with NBD-C12-ceramide as substrate. Thin-layer chromatography (TLC) used to separate NBD-dodecanoic acid (N-(NBD-Aminolauroyl) fatty acid) product from parent substrate NBD-C12-ceramide. TLC showed that some ceramide converted into NBD-Fatty acid (Figure 2.13). The product NBD- dodecanoic fatty acid compared with NBD-dodecanoic acid. RF value of NBD-C12-ceramide and NBD dodecanoic acid respectively 0.92 and 0.71. Acid ceramidase also hydrolyzes NBD-C6-ceramide and Omega-NBD-18:0 ceramide (Figure 2.14 A). RF value of NBD-C6-ceramide, Omega-NBD-18:0 ceramide, (N-(NBD-Aminohexanoyl) fatty acid) and (N-(NBD-Aminostearoyl) fatty acid) respectively 0.82, 0.88, 0.34 and 0.33. This enzyme not only hydrolyzes ceramide but also synthesize ceramide. It is synthesized C12-NBD-ceramide from using substrates NBD-dodecanoic acid and sphingosine (Figure 2.14 B).

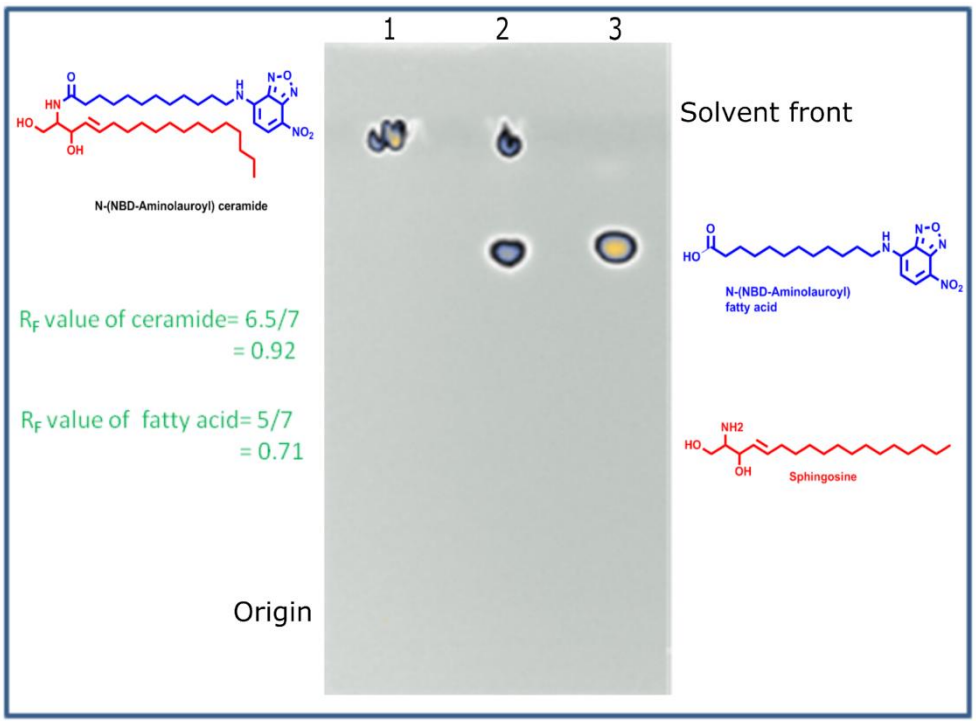


Figure 2.13: Detections of fluorescent fatty acid released from labeled ceramides by the action of a zebrafish AC. Lanes 1, C12-NBD-Ceramide as a negative control; Lanes 2, C12-NBD-Ceramide extracted after treated with zebrafish AC; Lane 3, NBD-Dodecanoic acid as a positive control.

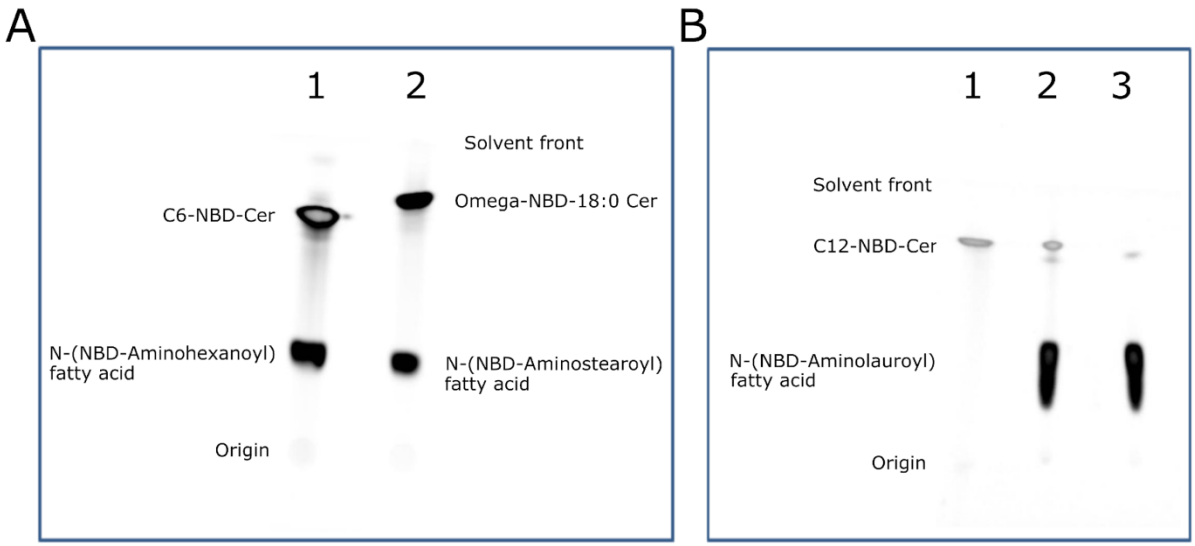


Figure 2.14: Forward and reverse activity of zebrafish AC. A: Detections of fluorescent fatty acid released from labeled ceramides by the action of a zebrafish AC. Lane 2, C6-NBD-Ceramide extracted after treated with zebrafish AC; Lanes 2, Omega-NBD-18:0 Ceramide extracted after treated with zebrafish AC. B: Detections of fluorescent C12-NBD-Ceramide synthesized from D-Sphingosine and labeled fatty acid by the action of a zebrafish AC. Lane 1, C12-NBD-Ceramide as a positive control; Lane 2, D-Sphingosine and NBD-Dodecanoic acid extracted after treated with zebrafish AC for 1hrs; Lane 3, D-Sphingosine and NBD-Dodecanoic acid extracted after treated with zebrafish AC for 24 hrs.

2.3.8 Enzyme assay and steady-state kinetics of zebrafish AC

The o-Pthaladehyd assay has used the estimation of the hydrolyzed by-product sphingosine by recording absorbance at 340 nm. The purified enzyme zebrafish AC was assayed for ceramide hydrolase activity with l-laurylsphingosine as substrate and the reaction was carried out at pH 4 and 37°C. Optimal enzyme activity was at pH 4 (sodium acetate buffer) in the assay condition (Figure 2.15 C). 40°C was determined to be the optimum temperature for enzyme activity (Figure 2.15 D). The optimum time and DTT concentration also observe 24 hrs and 0.5 mM, respectively (Figure 2.15 A and B). Optimum pH of hAC is 3.8-5.0 for the hydrolysis of N-dodecanoylsphing-4-enine (Bernardo et al., 1995; He et al., 2003) and 5.5-6.5 for the synthesis of N-dodecanoylsphing-4-enine (He et al., 2003).

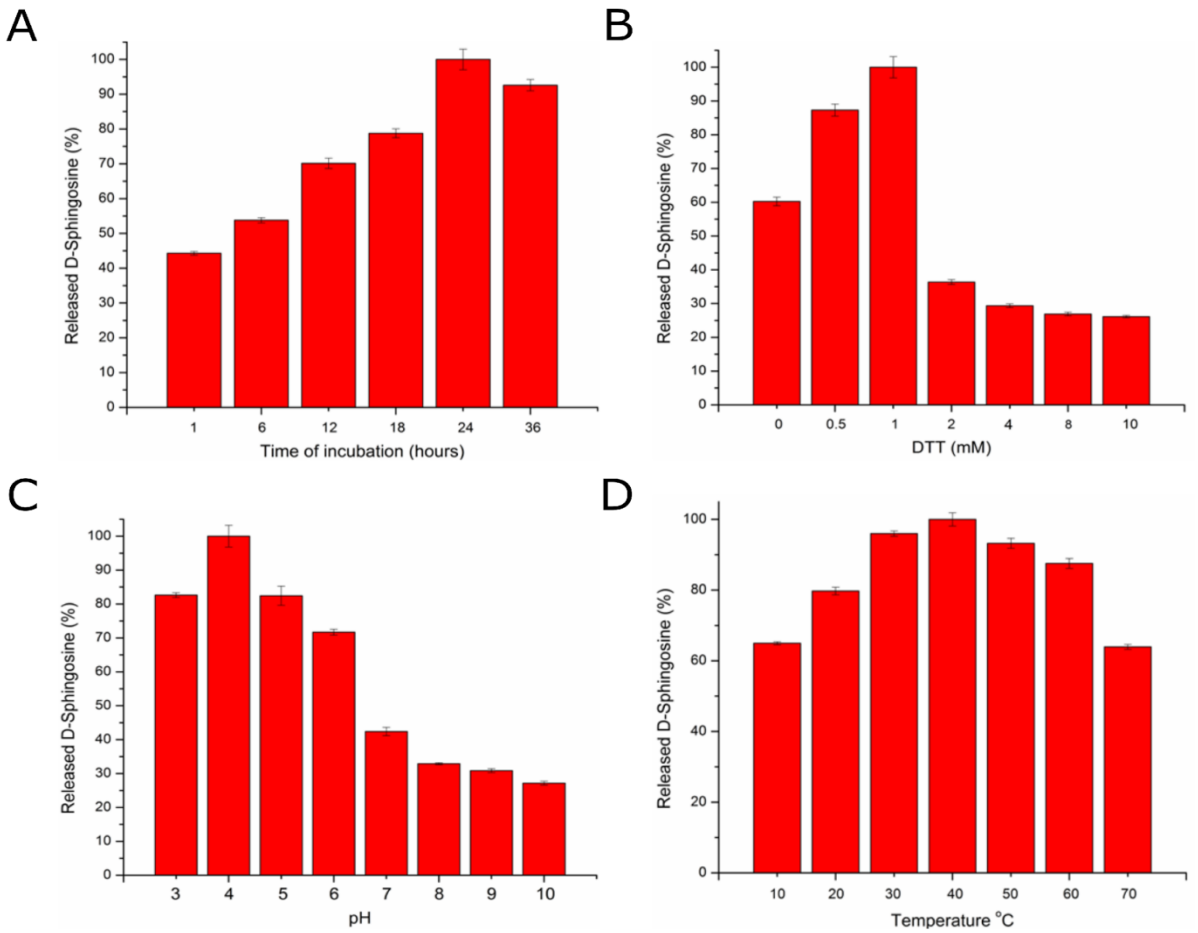


Figure 2.15: Optimum assay condition of zebrafish AC. A: Optimum time, B: DTT concentration, C: Optimum pH and D: Optimum temperature determined for the recombinant zebrafish AC.

Zebrafish enzyme followed the Michaelis-Menten (MM) kinetics (Figure 2.16). The values of K0.5, h and turnover number (kcat) for substrate L-lauroyl-D-sphingosine was determined by incubating the zebrafish AC with a range of concentrations of L-lauroyl-D-

sphingosine (100-1600 µM) under the standard assay conditions. The concentration of the zebrafish AC used was 1µg/reaction for all kinetic studies (The final concentration of enzyme in 500 µl assay reaction is 40 nM).

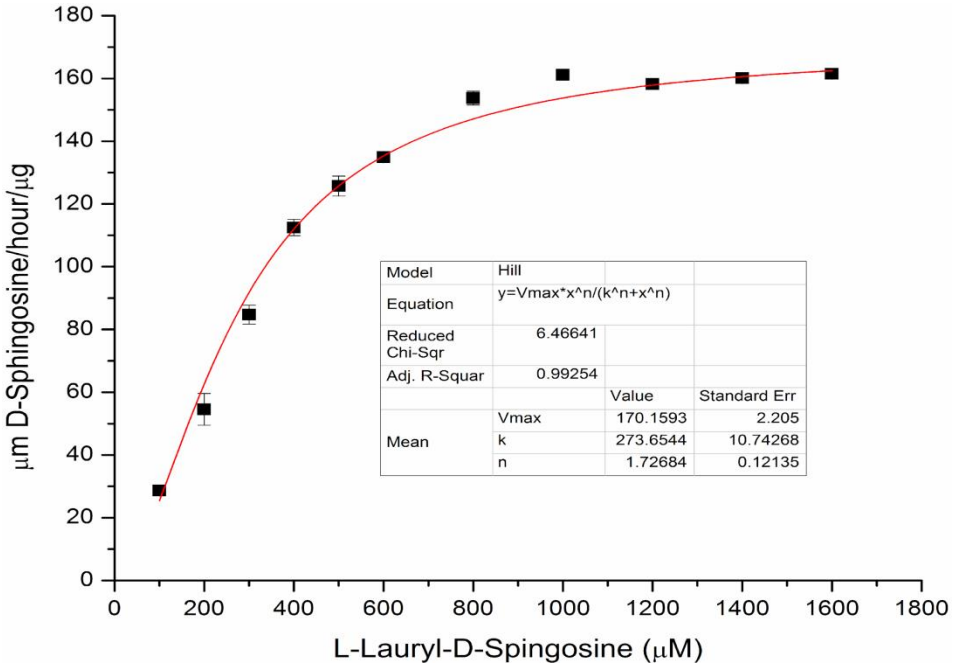


Figure 2.16: Steady-state kinetic graphs of recombinant zebrafish AC with N-Lauroyl-D Sphingosine.

The kinetic parameters K0.5, h and Vmax were calculated using Hill's equation The kinetic parameters of zebrafish AC calculated for substrate N-lauroyl-D-sphingosine are summarised in Table 2.4.

Table 2.4: Steady state kinetic parameters of Zebrafish AC

Substrate	$K_{0.5}$	h	V_{max}	Kcat (S^{-1})	Kcat/K0.5 (M^{-1} S^{-1})
N-Lauroyl-D-sphingosine	273.6544±10.74268	1.72684±0.12135	170.1593±2.205	39.66± 0.6856	0.1449

Table 2.5: Comparison of zebrafish AC activity with other reported hAC

Sr. No.	Substrate	$K_{0.5}$ (µM)	V_{max}	pH	Temp	Reference No.
1	N-dodecanoylsphing-4-enine	149	136 nmol/h/mg	**4.2**	**37**	(Bernardo *et al.*, 1995)
2	N-dodecanoylsphing-4-enine	389	28 nmol/h/mg	**4.5**	**37**	(He *et al.*, 2003)
3	Sphingosine	24	208 pmol/h/µg	**6.0**	**37**	(Okino *et al.*, 2003)
4	Dodecanoate	74	233 pmol/h/µg	**6.0**	**37**	(Okino *et al.*, 2003)
5	N-dodecanoylethanolamine	55		**4.5**	**37**	(Tsuboi *et al.*, 2005)
6	N-Lauroyl-D Sphingosine	273	170 µmol/h/ µg	**4.0**	**37**	This Study

2.3.9 Reported mutation of SMA-PME

The mutation Y36C, T42M and K152N of asah1 were reported associated with SMA-PME and knockdown of asah 1b increased apoptosis in the spinal cord (Zhou *et al.*, 2012). Mutant residue Asn is conserved in asah 1b at 152 position (Figure 2.4). Site-directed mutagenesis was used to engineer the mutant (Figure 2.17). All mutant expressed, purified and characterized to understand the SMA-PME at protein level (activity, processing and stability) (Table 2.6).

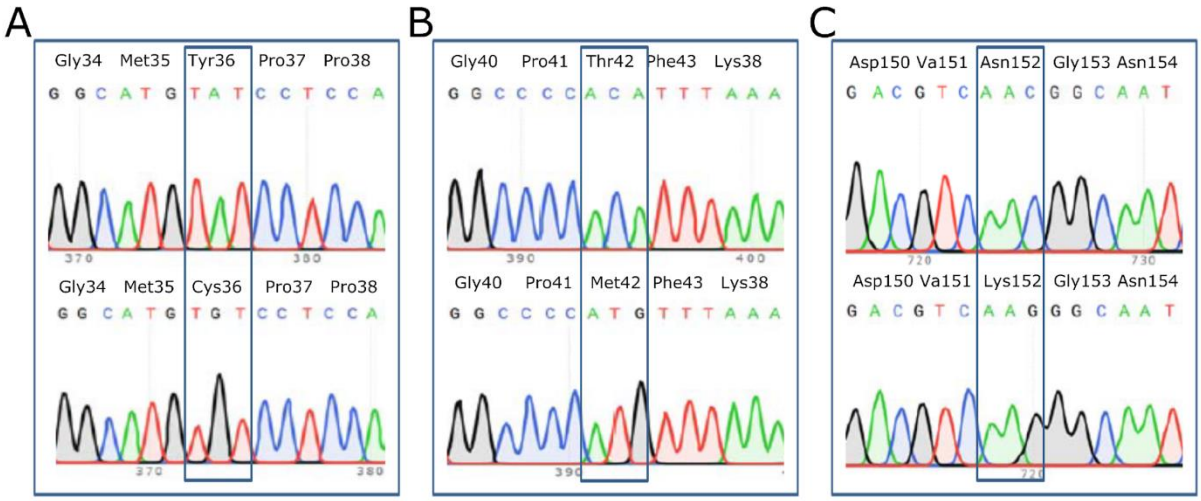


Figure 2.17: Sequences electrogram of site-directed mutagenesis. The upper panel is the WT sequences and the lower panel is the mutant sequence. The blue box highlights the mutation, A: Y36C, B: T42M, C: N152K.

Table 2.6: Expression and activity profile of the SMA-PME mutants of zebrafish AC

Sr. No.	SMA-PME mutants of *AC*	Activity	Expression Profile	Remarks
1	Tyr36→Cys	Active	Soluble but more prone to aggregates	Processed and less active
2	Thr42→Met	Active	Soluble but more prone to aggregates	Processed and less active
3	Asn152→Lys	Active	Soluble	Processed and active

2.3.10 Circular dichroism of zebrafish AC

The far UV CD spectrum of zebrafish AC had positive ellipticity maxima at 197 nm and negative minima at 220nm (Figure 2.18). The spectrum was analyzed with CD Pro to estimate the secondary structure contents as 34% α-helix and 7% β-sheet, suggested that zebrafish AC is rich in α-helix with slight β-sheet. The ordered CD profile also indicates that the recombinant protein was in a properly folded form.

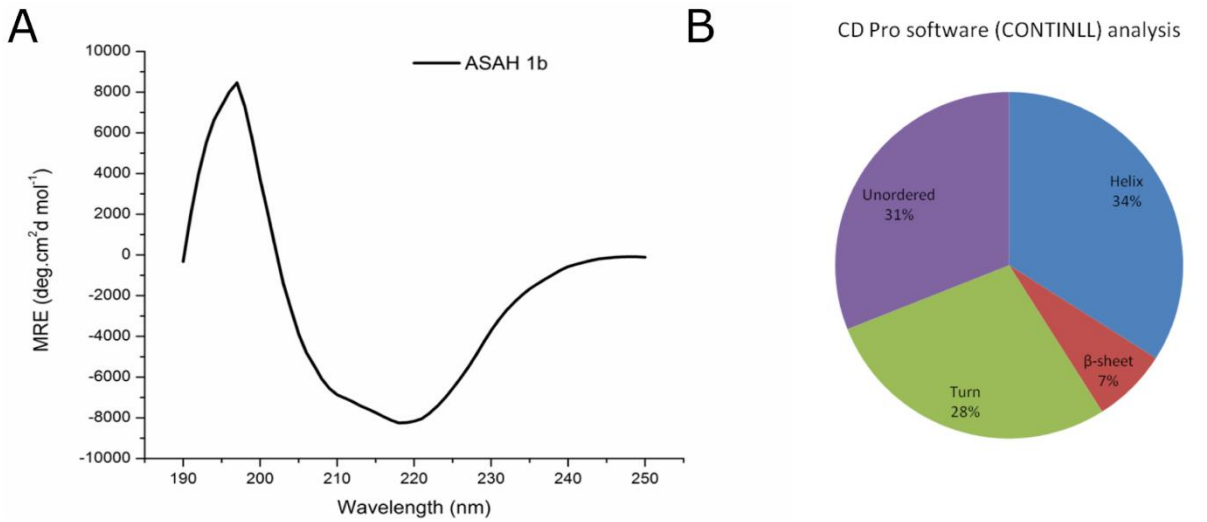


Figure 2.18: Far-UV CD spectra of zebrafish AC. A: Far-UV CD scan of zebrafish AC at pH 7.0 (0.2012 mg ml^{-1}). B: CD Pro analysis

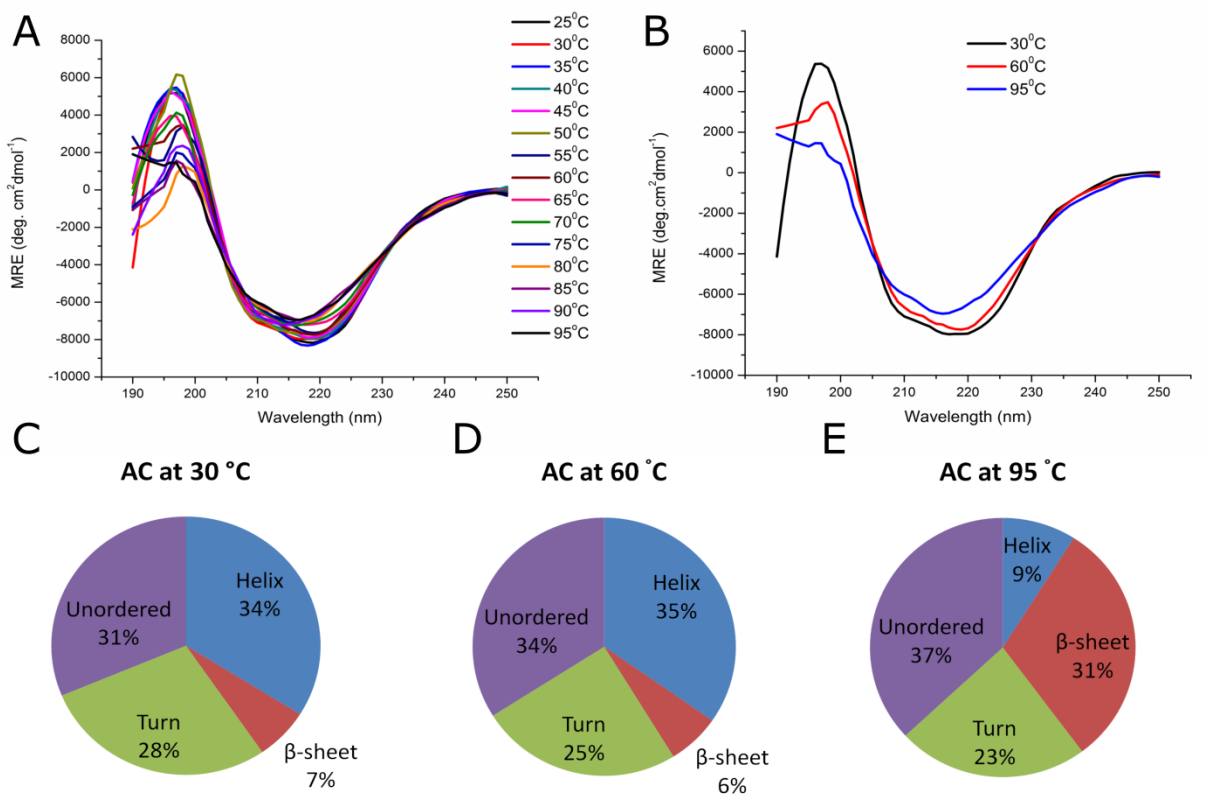


Figure 2.19: Thermal denaturation of zebrafish AC using the CD. A and B: Far-UV CD spectra of zebrafish AC incubated at different temperatures for 5 min, 0.178 mg ml^{-1} of protein was used, each sample was incubated for 5 min at the respective temperature and then the scans were recorded. C, D and E: CD Pro analysis

Thermal stability of the protein provides information about the non-covalent and covalent forces involved in stabilizing the enzyme structure. Secondary structure monitoring during thermal denaturation CD spectroscopy revealed almost unperturbed CD profile up to 95°C (Figure 2.19). CONTINLL analysis indicates that the α-helix content of the protein decreased significantly above 55°C with a simultaneous increase in β-sheet and unordered structure.

2.3.11 Characterizations of *C. elegans* AC

C. elegans AC was purified using Ni-NTA affinity chromatography. Majority of the purified enzyme was obtained in the unprocessed form; however, small amounts of β-subunits were also obtained (Figure 2.21 A). The α-subunit is difficult to observe in SDS-PAGE may be reason less concentration due to less processing of *C. elegans* AC. The unprocessed AC band was detected by western blot with monoclonal anti-His antibodies (Figure 2.21 D).

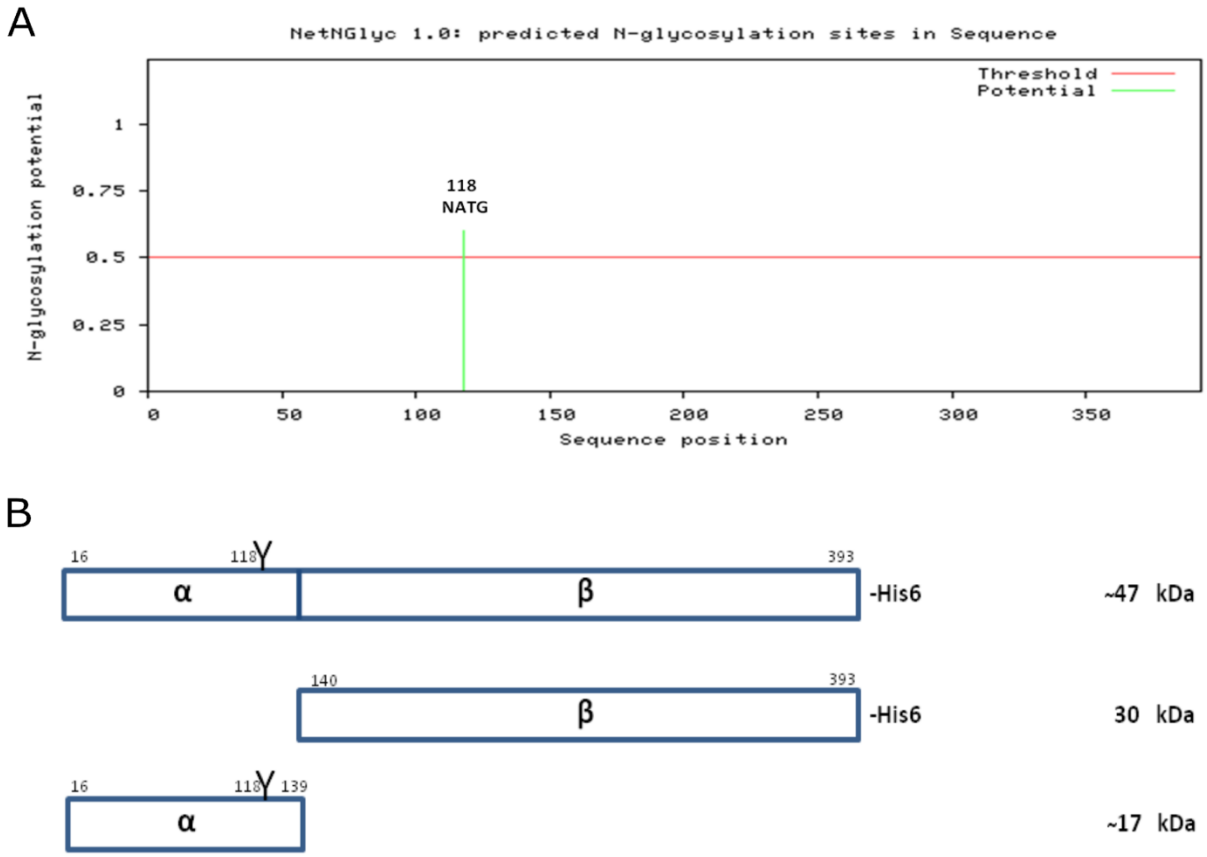


Figure 2.20: The predicted N-glycosylation site and observed molecular weight of *C. elegans* AC. A: Predicted N-glycosylation site in *C. elegans* AC using NetNGlyc 1.0 server B: The schematic diagram of recombinant *C. elegans* AC showed that different subunit with sequence number, glycosylation numberand molecular weight.

The β-subunit observed in SDS-PAGE but not in western blot, which was probably the result of carboxy-terminal processing, similar to human AC (Figure 2.21 A and C) (He et al., 2003). On the basis of gel electrophoretic results, the molecular masses corresponding to the bands ~47, 30and 17 kDa were predicted with slight variation in the molecular weight than the predicted in silico amino acid sequence 44, 30 and 14 kDa for unprocessed AC, β- and α- subunit, respectively (Figure 2.20 A and B). One putative N-glycosylation sites at amino acid positions 118 and no potential O-glycosylation site were predicted in the C. elegans AC when the amino acid sequence was analyzed using NetNGlyc 1.0 Server (Gupta, Jung and Brunak, 2016) and NetOGlyc4.0 Server, respectively (Steentoft et al., 2013).

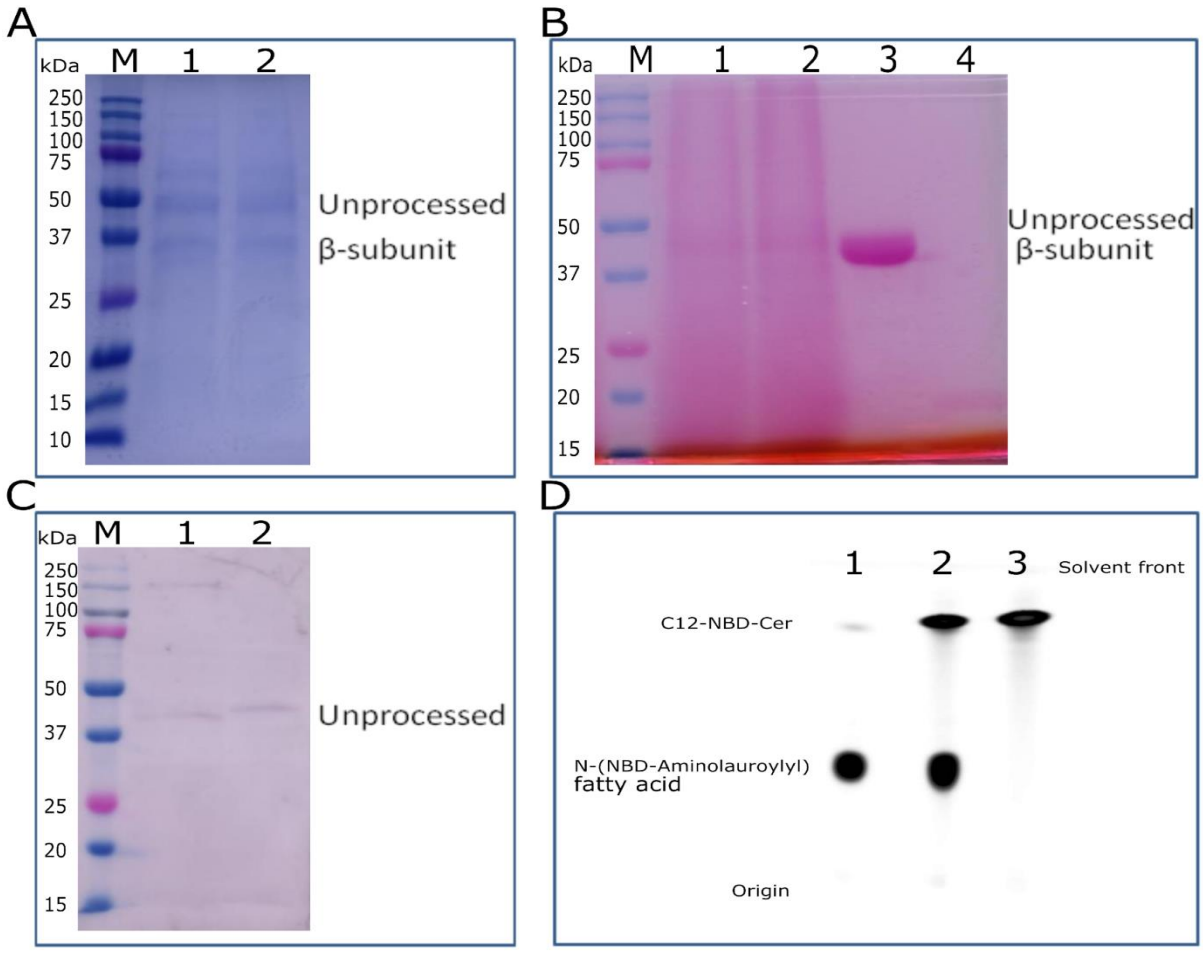
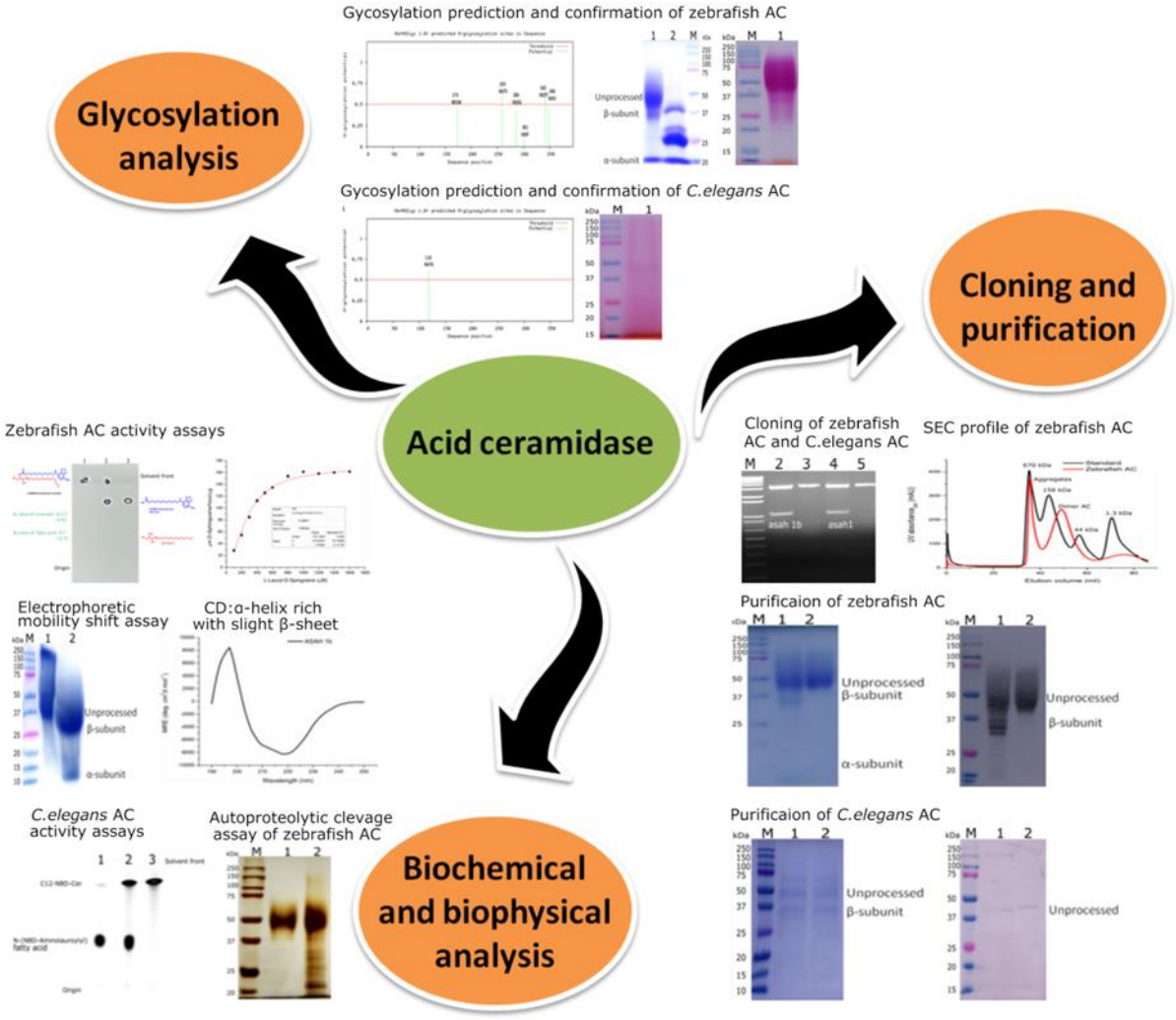


Figure 2.21: Purification and characterization of *C. elegans* AC. A: 8-16% SDS-PAGE analysis shows purification of *C. elegans* AC protein. M, Prestained protein ladder; Lane 1 and Lane 2, *C. elegans* AC protein was purified using Ni-NTA affinity chromatography. B: Glycosylation analysis of *C. elegans* AC protein by glycosylation staining kit. M, Prestained protein ladder; Lane 1 and Lane 2, *C. elegans* AC protein was purified using Ni-NTA affinity chromatography; Lane 3, Horseradish peroxidase as a positive control; Lane 4, Soybean trypsin inhibitor as a negative control. C: 100 µg of *C. elegans* AC protein were treated with deglycosylation mix for 24 h, separated by SDS–PAGE and analyzed by western blotting using with monoclonal anti-His antibodies. M, Prestained protein ladder. Lane 1, *C. elegans* AC protein was treated with deglycosylation mix. Lane 2, *C. elegans* AC protein.

The β-subunit does not have any putative glycosylation site. Kit based glycoprotein staining on zebrafish AC confirmed glycosylation. The 47 kDa band was stained magenta, whereas the 30 kDa band did not, implicating glycosylation of 47 kDa band which were the unprocessed AC (Figure 2.20 A and 2.21 A, B and C). These results also indicate that β-subunit having 44 kDa bands is nonglycosylated. The enzymatic deglycosylation using PDM II appeared to decrease the protein size to about 47, kDa to about 44, kDa (Figure 2.21C: lane 1 and 2). These enzymes hydrolyze C12-NBD-Ceramide (Figure 2.20 D: lane 2) and also synthesize C12-NBD-Ceramide (Figure 2.21 D: lane 1). This forward and reverse reaction is dependent on pH, the concentration of the product, salt and detergent.

2.4 Graphical summary

Schematic illustration of AC characterization



Chapter 3: Characterization of acid ceramidase-like protein (NAAA)

3.1 Introduction

Acid ceramidase-like protein (ASAHL) is also called as N-acylethanolamine hydrolyzing acid amidase (NAAA). NAAA catalyzes the hydrolysis of PEA into ethanolamine and palmitic acid (Ueda, Yamanaka and Yamamoto, 2001). Inhibiting enzymatic degradation of N-acylethanolamines (NAEs) like PEA is a potential strategy for pain treatment. The NAAA is a potential target of treatments for inflammatory disorders (Ueda, Yamanaka and Yamamoto, 2001). NAAA is functionally similar to FAAA but structurally more similar to AC. Like AC, NAAA also belongs to the Cholylglycine hydrolase family, a subfamily of Ntn (N-terminal nucleophile) hydrolase structural superfamily (**Pfam: PF02275**) (Oinonen and Rouvinen, 2000; Rossocha *et al.*, 2005; Tsuboi, Takezaki and Ueda, 2007). *ASAHL* gene first cloned and Purified from human placenta (Hong *et al.*, 1999) but NAE-hydrolyzing activity identifies by N-terminal sequencing of amidohydrolase purified from rat lung (Ueda, Yamanaka and Yamamoto, 2001).

NAAA is synthesized as inactive premature precursor polypeptide (signal peptide, α-subunit, β-subunit) targeted to the lumen of the endoplasmic reticulum (ER) by an ER-specific signal peptide, which is subsequently processed into mature active heterodimer (consisting of α- and β-subunits) in lysosomes (Zhao *et al.*, 2007). Maturation of the pre mature enzyme results in the formation of a new N–terminal cysteine residue in the β-subunit (βCys1) which acts as a nucleophile during hydrolysis of the NAE (Zhao *et al.*, 2007; Wang *et al.*, 2008). The catalytic mechanism of NAAA is like to other well-established members of the Ntn hydrolase. The active site of the enzyme contains three amino acid residues, (Cys, Arg and Asp) which form the catalytic center of the enzyme, out of which the cysteine residue acts as a nucleophile (Zhao *et al.*, 2007; Wang *et al.*, 2008). NAAA is glycoprotein that undergoes removal of an N-terminal signal peptide after targeting to ER, finally transported by the mannose-6-phosphate pathway to the lysosomes via Golgi bodies. It is N-glycosylated and O-glycosylated glycosylated in ER and Golgi, respectively (Zhao *et al.*, 2007).

In the present study, we have reported the first time, an efficient expression system for the heterologous overexpression of biologically active NAAA from mouse (Mus musculus) and *Micromonas commoda* (strain RCC299) (Picoplanktonic green alga). The expression

system used was a eukaryotic expression system, Pichia pastoris. This expression system is easier and well suited for expressing recombinant NAAA with post-translational modifications to prokaryotic E. coli and animal cell lines. This is an important consideration for the production of functionally active recombinant forms of naturally glycosylated proteins such as NAAA.

3.2 Materials and Methods

3.2.1 Cloning of mouse and *M. commoda* naaa

3.2.1.1 Sequences analysis

The cDNA sequences and signal peptide length of mouse and *M. commoda* naaa were retrieved from GenBank (GenBank Accession No. NM_025972 and No. 8244433). Autoproteolytic cleavage site, αand β subunit sequences were obtained from the MEROPS database. N-glycosylation and O-glycosylation sites were predicated using NetNGlyc 1.0 server and NetOGlyc 4.0 server, respectively. The theoretical pI and molecular weight of the protein was calculated from the ExPASy ProtParam tool server.

3.2.1.2 Construction of plasmid vector (pPICZα-A) and gene (naaa)

Mouse cDNA was used as a template for PCR amplification of the gene encoding mouse NAAA using primers forward, (5`-ACTAGTGAATTCGCTGTCGTCCCCGGGACTCC-3) and reverse, (5′-ACTAGTGTCGACGCTCGGGTTTCTGATCATGG-3). *M. commoda* genomic DNA was used as a template for PCR amplification of the gene encoding *M. commoda* NAAA using primers forward, (5`- ACTAGTGAATTCATGAGGGTTTGGTCGCGTGT -3) and reverse, (5′-ACTAGTTCTAGAGATGGGAGAGCAACCCTCAC -3). The underlined sequences correspond to the restriction sites. These sites were selected to insert the gene into the pPICZα-A vector, in order to ensure that the recombinant DNA fragment had an alpha-factor signal peptide preceding its N-terminal, followed by a C-terminus 6× His extension under the AOX promoter. The gene was amplified using Ex Taq DNA polymerase (TaKaRa) under the following conditions: 1 cycle of pre-denaturation at 94°C (1 min) followed by 28 cycles of denaturation at 94°C (30 s), primer annealing at 58°C (30 s) and extension at 72°C (1:30 min), a final extension at 72°C (10 min) and then cooled to 4°C. The PCR amplified products were analyzed on 1% agarose gel. The amplicon of the desired size was gel extracted using the QIAQuick Gel Purification kit (Qiagen, Cat#28704).

3.2.1.3 Restriction digestion and ligation

The mouse gene was digested with EcoRI and SalI and *M. commoda* with EcoRI and XbaI at 37°C with suitable buffers for 4 hours and ligated into pPICZα-A vector. Approximately 1:3 ratio of vector to insert was incubated with T4 DNA ligase (Invitrogen) with a suitable buffer at 16°C overnight.

3.2.1.4 Transformation of cloned plasmid

Both constructs were transformed into competent *E. coli* DH5α. Positive colonies were selected on low salt LB zeocin plates and gene insertion was confirmed by colony PCR using 5′AOX1 vector specific forward primer and insert specific reverse primer. The clone was further confirmed by RE double digestion and sequencing.

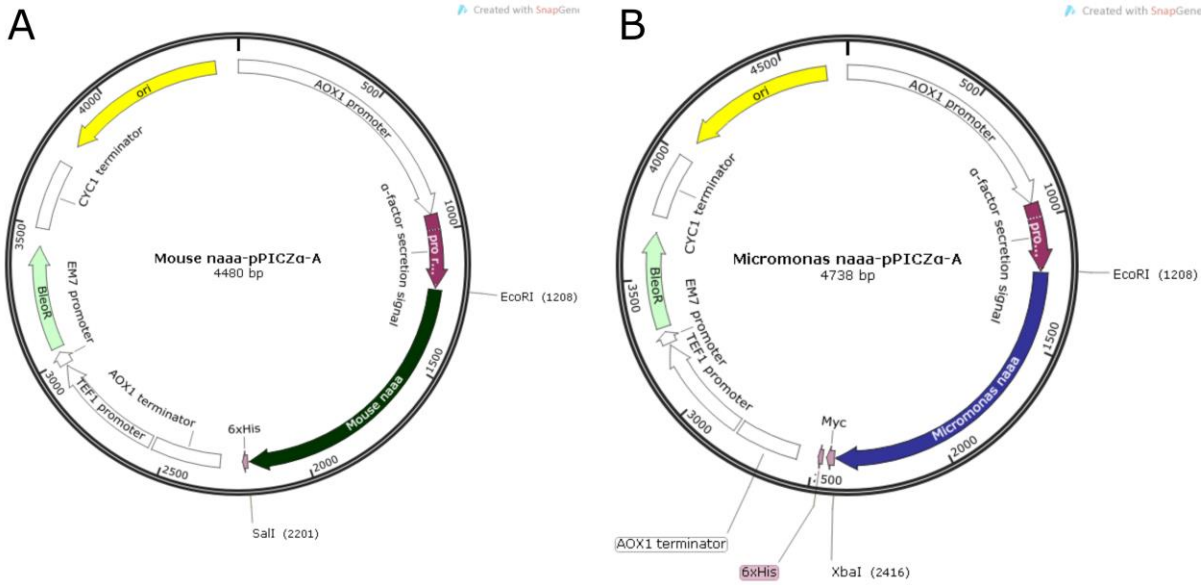


Figure 3.1: Map of the vector. A: Mouse naaa-pPICZα-A. B: *M. commoda* naaa-pPICZα-A.

3.2.2 Expression asah1b in *Pichia pastoris*

Both constructs were linearized using PmeI restriction enzyme and transformed into competent *Pichia pastoris* (GS115 strain) by electroporation. The transformants were selected on YPDS-Zeocin plates and colony PCR using the 5′α-Factor and 3′AOX1 primers. The selected clone was cultured into 25 ml of BMGY medium and grown at 28 °*C* till OD_{600} reaches to 2–6. These cells were re-suspended in 300 ml of BMMY medium and incubated at 28°*C*. This secondary culture was induced by the addition of 0.5% (v/v) methanol at 24 h intervals. The cell-free supernatant was concentrated by lyophilization and then was dialyzed through Millipore 12-kDa cut-off membrane to exclude ions and salts.

3.2.3 Confirmation of expression

The purity, molecular weight and autoproteolytic processing of the protein were assessed on SDS-PAGE by coomassie brilliant blue (CBB) R-250 and silver staining. The molecular weight was estimated by using a prestained dual color protein ladder (10-250 kDa, Bio-Rad) as standard. Western blotting was achieved using anti-His mouse monoclonal antibodies (Qiagen), HRP coupled secondary antibodies and SIGMAFAST^{TM} 3, 3`-Diaminobenzidine developer tablet. The glycosylation staining was done by glycoprotein staining kit (Pierce). Deglycosylation was done with protein deglycosylation Mix II (New England Biolabs) according to the manufacturer's protocol.

3.2.4 Purification

3.2.4.1 Ni-NTA affinity chromatography

The C-terminal His-tagged NAAA protein was purified using Ni-NTA affinity chromatography. The supernatant with Ni^{2+}- agarose beads were incubated for overnight for binding at 4°C with shaking. After incubation supernatant was loaded on to a column. The matrix was then washed with equilibration buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM DTT, pH 8) followed by removal of non-specific and weakly bound proteins by washing with wash buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM DTT and 10 mM imidazole, pH 8). The matrix-bound NAAA enzyme was eluted by passing the elution buffer (50 mM Tris-HCl, 300 mM NaCl, 10% v/v glycerol, 2 mM DTT and 250mM imidazole, pH 8). The high concentration of imidazole in elution fraction was removed by buffer exchange using a 10-kDa cut-off ultrafiltration membrane (Millipore).

3.2.4.2 Size exclusion chromatography

The mouse NAAA protein solution obtained was finally subjected to size exclusion chromatography using sephacryl S200(16/60) column, pre-equilibrated with Tris buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 2 mM DTT and 10% v/v glycerol). Fractions containing NAAA activity were pooled and analyzed by running 12% SDS-PAGE and western blot. Gel filtration standard (Bio-Rad) was used as a standard for estimation of the native molecular mass of protein.

3.2.4.2 Concanavalin A–sepharose chromatography

1gm of lyophilized powder CNBr-activated sepharose^{TM} (Amersham Pharmacia Biotech AB) 4B in 200 ml of 1mM HCl washed by 1mM HCl until final volume 3.5 ml

volume for 15 min. CNBr-activated sepharose^{TM} with 35 mg concanavalin A (Con A) (Sigma) in 5 ml coupling buffer (200 mM sodium bicarbonate, 500 mM NaCl, 200 mM and competitive sugar (Methyl α-D-glucopyranoside) (Sigma), pH8.3) incubated for 24hrs at 4°C with gentle stirring. Excess ligand washed with 5 column volume (CV) by coupling buffer (without competitive monosaccharide). Final matrix with 1M glycine solution incubated for 3 hrs at room temperature. Column washed by coupling buffer with 0.02% sodium azide followed by ice cold milli q water. The packed column of Con A-sepharose column was equilibrated by equilibration buffer (5 mM sodium acetate, 100 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$,1 mM $MgCl_2$ and 0.02% sodium azide, pH 5.2). After equilibration 1ml sample was loaded on to a column. The matrix was then washed with equilibration buffer followed by removal of non-specific and weakly bound proteins by washing with elution buffer I (equilibration buffer with 10 mM methyl α-D-glucopyranoside). The matrix-bound NAAA enzyme was eluted by passing the elution buffer II (equilibration buffer with 300 mM methyl α-D-glucopyranoside).

3.2.5 Quantitative enzyme assay

NAAA enzyme activity is measured by estimating ethanolamine released by hydrolysis activity on incubating the NAAA enzyme with N-Palmitoylethanolamine (PEA) (Avanti polar lipid) at 37°C in Assay buffer III (100 mM acetate buffer pH 4, 0.1 % v/v NP-40 substitute). 500 µl of the reaction mixture was withdrawn and an equal volume of OPA was added and kept at RT for 2 min. The product formed by the reaction was recorded at 340 nm. One unit of NAAA activity is defined as the amount of enzyme that liberates 1 µmol of Ethanolamine from the substrate per minute. Specific activity was defined as the number of units of activity per milligram of the pure protein. Protein concentration was measured by the Bradford method. Ethanolamine (Sigma) standard graph was prepared to quantify the product released on substrate hydrolysis (Figure. 3.3).

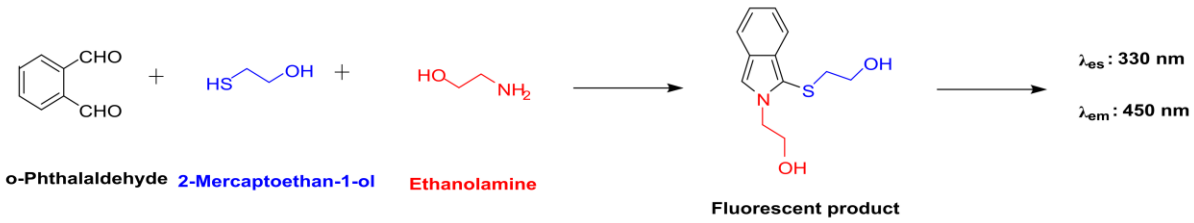


Figure 3.2: Detection of primary amines of ethanolamine with OPA. Hydrolyzed ethanolamine reacts with OPA in the presence of a nucleophile (2-mercaptoethanol) and under alkaline conditions yields highly fluorescent product that is quantified photometrically at 340 nm.

3.2.6 Effect of time, pH, temperature and DTT on NAAA activity

The NAAA activity was assayed at different time value in the range 1-36 hours, 0-10 DTT mM, pH values in the range 3.0-10.0 and temperatures 20-70°C to determine the optimum conditions for enzyme activity.

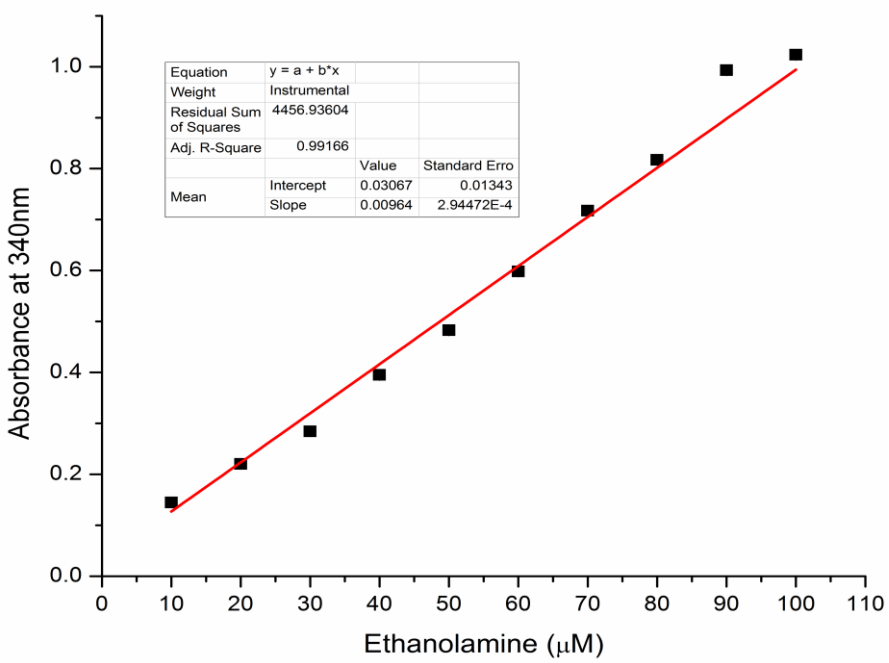


Figure 3.3: Standard graph of Ethanolamine (µM) estimation.

3.2.7 Steady-state kinetic parameters for mouse NAAA

The kinetic constants $K_{0.5}$ and V_{max} of Mouse NAAA were estimated by assaying the enzyme activity with increasing concentrations of PEA as substrates in the range 10-500 µM following the standard assay protocol. The assay was carried out in triplicates and the kinetic constants ($K_{0.5}$ V_{max}, hill coefficient 'h') were determined by non-linear regression by fitting the kinetic data in Origin8. The *k*cat value for PEA calculator by non-linear regression by fitting the kinetic data (real velocity in nM/s) in GraphPad Prism version 5.0 with constraints enzyme concentration (44.4 nM) and K_m (67866.9 nM).

$[V_{max} .Sh] / [K_{0.5} + Sh + (S2h/Ki)]$ ……..... (3.1)

3.3.8 Steady-state fluorescence measurement

Fluorescence of protein was measured using a Perkin Elmer Life Sciences LS50 fluorescence. Fluorescence emission spectra of the enzyme were measured with a slit width of 7 nm for both the monochromators. Spectra for either buffer or buffer containing substrate were recorded in order to eliminate the background emission noise by correcting with observed fluorescence of the sample. Suitable amount of enzyme in 20 mM sodium citrate-phosphate buffer, pH 5.0 buffer were filtered with the 0.22 um fitters and protein was centrifuged at high-speed 13000 rpm in order to remove the particulate and aggregates in the solution. 2 ml samples were used throughout the experiment at constant temperature (25°C ± 0.1°C) in a quartz cuvette to record the fluorescence spectra with the help of a Julabo F 25 circulating cryobath. Samples were excited at 295 nm and the emission spectra were recorded from 300 to 400 nm.

3.2.9 Far UV-circular dichroism spectroscopic studies

Purified mouse NAAA enzyme at a concentration of 0.348 mg ml^{-1} was used for the entire sample. Far UV-CD spectroscopic method described in chapter 2 (section 2.2.11).

3.2.10 Differential Scanning Fluorimetry

Measurements thermal unfolding was performed with Prometheus NT.48 nano DSF (NanoTemper technology, Germany). DSF experiment was based on the change in intrinsic fluorescence of the protein where the shift in fluorescence emission is plotted as the ratio between 350 and 330 nm. The capillaries were filled with a mixture of 20 µl of protein placed onto the capillary tray of the Prometheus NT.48. Start to end temperatures were set as 20°C to 95°C and the heating rate was defined as 1°C/min. The inflection point of the resulting sigmoidal curve provided the melting temperature (T_m) of the protein. The T_m of the unprocessed *M. commoda* NAAA (inactive) was studied in native condition and in the presences of substrates.

3.2.11 Ceramidase hydrolase activity

Ceramidase hydrolase qualitative assay and quantitative assay method described in chapter 2 (section 3.2.7 and 3.2.9).

3.3 Results

3.3.1 Cloning of naaa gene from a mouse and *M. commoda*

Mouse NAAA consists of 362 amino acid residues with a 31 amino acid at N-terminal that, functions as a signal peptide. Active site residues consist of Cys 131, Arg 147 and Asp 150; these were predicted via sequences conservation analysis of acid ceramidase family. The mouse naaa gene of size 987 bp without the signal peptide sequence was fished out mouse cDNA (Figure 3.4 A: lane 4). We also amplify the asah1 gene to check the presences of acid ceramidase gene in mouse. Mouse naaa was cloned into the vector pPICZα-A between N-terminal yeast α-factor secretion signal and c-terminal 6XHis tag for secretion and purification, respectively (Figure 3.4 C: lane 5). Mouse naaa translates into a 341 AA protein, theoretical molecular weight of 38 kDa having pI at pH of 5.89.

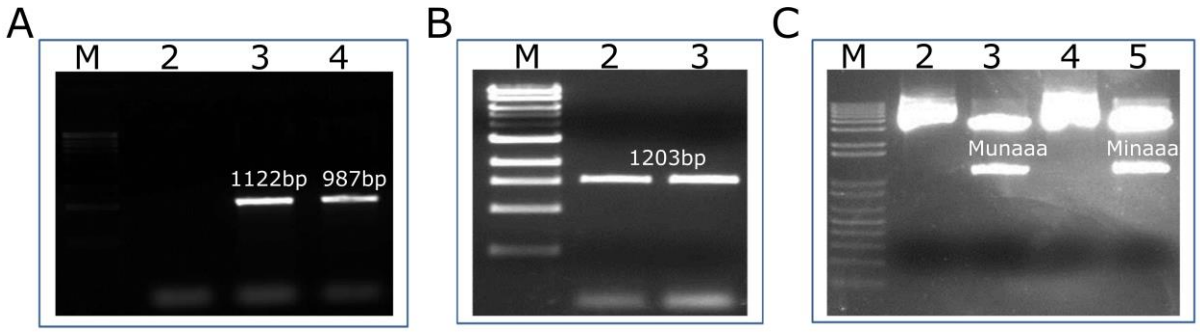


Figure 3.4: Cloning of naaa of Mouse and *M. commoda*. A: 1% agarose gel showing PCR product of mouse naaa gene. Lane 1, NEB 1Kb DNA ladder; Lane 2, PCR blank; Lane 3, 1122 bp PCR product of asah 1a; Lane 4, 987 bp PCR product of naaa. B: 1% agarose gel showing the PCR product of *M. commoda* naaa. Lane 1, NEB 1Kb DNA ladder; Lane 2 and 3, 1203 bp PCR product of naaa. C: Restriction digestion. Lane 1, NEB 1Kb plus DNA ladder; Lane 2, Double digestion (*EcoRI/SalI*) of cloned plasmid (Mouse naaa-pPICZα-A); Lane 3, Undigested cloned plasmid (Mouse naaa-pPICZα-A); Lane 4, Double digestion (*EcoRI/XbaI*) of cloned plasmid (*M. commoda* naaa-pPICZα-A); Lane 5, Undigested cloned plasmid (*M. commoda* naaa-pPICZα-A).

M. commoda NAAA consists of 400 amino acid residues. Active site residues consist of Cys 146, Arg 164 and Asp 167; these were predicted from via sequences conservation analysis of acid ceramidase family family. M. commoda naaa gene size 1203 bp was cloned into the vector pPICZα-A between N-terminal α-factor secretion signal and c-terminal 6XHis and myc tag (Figure 3.4 B: lane 2 and 3, C: lane 5). *M. commoda* naaa translates into a 425 aa protein, theoretical molecular weight of 47 kDa having pI at pH of 5.56

3.3.2 Heterologous expression of mouse NAAA

The recombinant P. pastoris GS115 strain having mouse naaa gene was inoculated in a 500 ml shake flask with 25 ml BMGY and grown for 48 h at 28 °C, 200 rpm. Cells pellets were harvested by centrifugation and resuspended in a 2 L shake flask with 300 ml BMMY. Methanol induction was given every 24 h for naaa gene expression. The culture was grown in BMMY medium for 144 h at 28 °C, 200 rpm. Since the inserted gene is present downstream to the secretory signal sequences, it was expected that the expressed gene product would be secreted out in the medium. The culture medium was duly collected and cell debris was removed. The cell-free supernatant was concentrated and desalted by dialysis.

3.3.3 Purification and confirmation of mouse NAAA

Mouse NAAA was purified from approximately 300 ml of cell-free medium, 72 h after induction. A two-step purification protocol, using Ni-NTA affinity chromatography followed by size exclusion chromatography was employed to purify mouse NAAA. As shown in the chromatogram, four peaks were obtained (Figure 3.5 C). These were further analyzed by SDS-PAGE and checked for activity. Only two peaks have activity and three peaks have bands of interest on SDS-PAGE (Figure 3.5 A: lane 3-5). Majority of the purified enzyme was obtained in the unprocessed form; however, small amounts of α- and β-subunits were also obtained (Figure 3.5 A). The broad peak in size exclusion chromatography and SDS-PAGE suggests that the purified enzyme fraction is heterogeneous due to glycosylation. Western blotting using anti-His antibodies detected unprocessed NAAA band along with the β-subunit. (Figure 3.5 B). The β-subunit had two bands in western blot, suggesting that there were two different poly peptide, as a result of carboxy-terminal processing or which may reflect different glycosylation status, similar to hAC (He *et al.*, 2003). The purified mouse NAAA fraction showed activity towards the substrate PEA as a substrate. Protein yield obtained was ~5 mg/l of culture. This study illustrates that the P. pastoris expression system is efficient for the heterologous expression of proteins undergoing autoproteolytic processing such as mouse NAAA.

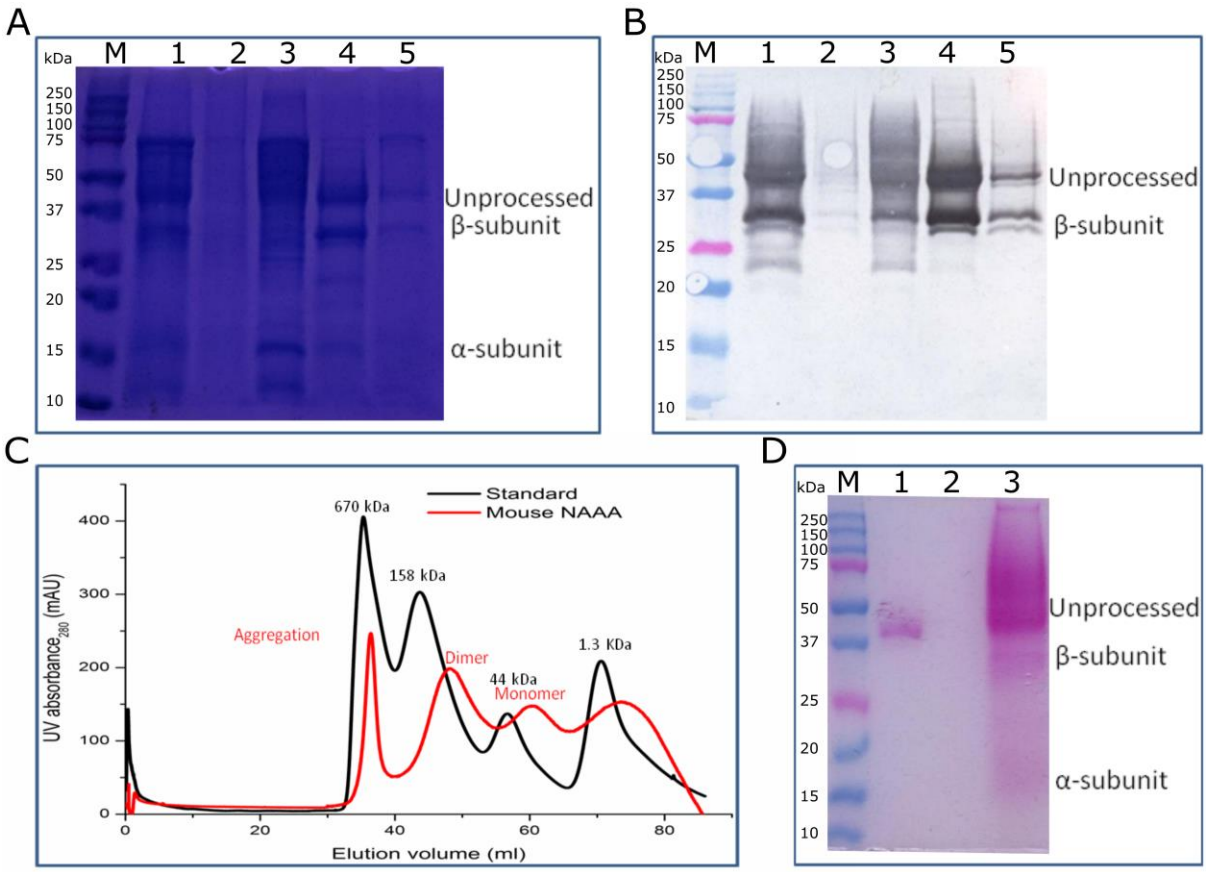


Figure 3.5: Purification of mouse NAAA. A: 12% SDS-PAGE analysis shows purification of mouse NAAA protein by sequential chromatography. M, Prestained protein ladder; Lane 1, Mouse NAAA protein was purified using Ni-NTA affinity chromatography; Lane 2-5, Mouse NAAA protein was purified by AKTA purifier with size exclusion chromatography (2: aggregation, 3: dimer, 4: monomer 5: late eluted).B: Western blotting analysis confirms mouse NAAA protein expression with monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 1, Mouse NAAA protein was purified using Ni-NTA affinity chromatography; Lane 2-5, Mouse NAAA protein was purified by AKTA purifier with size exclusion chromatography (2: aggregation, 3: dimer, 4: monomer 5: late eluted). C: Chromatogram of size exclusion chromatography using sephacryl S200 (16/60) column for purification of mouse NAAA protein. D: Glycosylation analysis of mouse NAAA protein by glycosylation staining kit. M, Prestained protein ladder; Lane 1, Horseradish peroxidase as a positive control; Lane 2, Soybean trypsin inhibitor as negative control; Lane 3, Mouse NAAA protein was purified by AKTA purifier with size exclusion chromatography.

3.3.4 Molecular mass determination mouse NAAA

The native molecular mass of mouse NAAA was found to be ~100 kDa, as estimated by size-exclusion chromatography using a sephacryl-S200 column in FPLC system when compared with molecular weight standards to determine the mass (Figure 3.5 C). The molecular mass corresponding to ~45-55, ~34 and ~11-15 kDa was α, β and unprocessed mouse NAAA respectively estimated on SDS-PAGE (Figure 3.5 A). The difference between

masses on size-exclusion chromatography, SDS-PAGE suggested that mouse NAAA exists as a homodimer of the heterodimer (α + β) having two α-subunit and two β-subunits. Our findings resemble previous findings of the human AC as homotrimers of the heterodimer in urine (Bernardo *et al.*, 1995).

3.3.5 Autoproteolytic activation of mouse NAAA

A mouse NAAA like human NAAA is also subjected to proteolytic cleavage during maturation. With respect to the similarity between primary structure of mouse NAAA to that of hNAAA, mouse NAAA Cys 131 corresponded to Cys 126 of hNAAA, the N-terminal residue of the β-subunit. Western blotting confirmed mouse NAAA cleavage; the anti-his antibody binds to the only the C-terminal 6X His-tagged mouse NAAA. Based on subunit molecular masses obtained from SDS-PAGE Cys 131 was predicated to be the cleavage site. The difference in electrophoretic mobility between mature and immature mouse NAAA on SDS-PAGE in the presences and absences of β-mercaptoethanol (Figure 3.7A), suggest that like hAC, mouse NAAA does form an αβ heterodimer but not connected through a disulfide bond (Bernardo *et al.*, 1995). The ratio of immature protein to mature α and β-subunit in the purified varied to a moderate extent from batch to batch. This may be due to, the mouse NAAA is secreted as an unprocessed non-numeric precursor (immature form) and then slowly gets converted into matures heterodimer during expression and purification. The unprocessed form did not completely processed into the mature form may be due to the oxidation of the active site cysteine.

3.3.6 Glycoprotein analysis of mouse NAAA

On the basis of gel electrophoretic results, the molecular masses corresponding to the bands ~45-50, ~34 and ~11-15 kDa were observed with slight variation in the molecular weight than the predicted in silico amino acid sequence 38, 27 and 11 kDa for unprocessed NAAA, β-and α- subunit, respectively. Five putative N-glycosylation sites (Asn-Xaa-Thr/ser) and one potential O-glycosylation site were predicted in the mouse NAAA when the amino acid sequence was analyzed using NetNGlyc 1.0 Server (Gupta, Jung and Brunak, 2016) and NetOGlyc 4.0 Server, respectively (Steentoft *et al.*, 2013). Five N glycosylation sites at amino acid positions 42, 112, 314, 338 and 360 and a single O-glycosylation site at positions 307 present on the β-subunit (Figure 3.6).

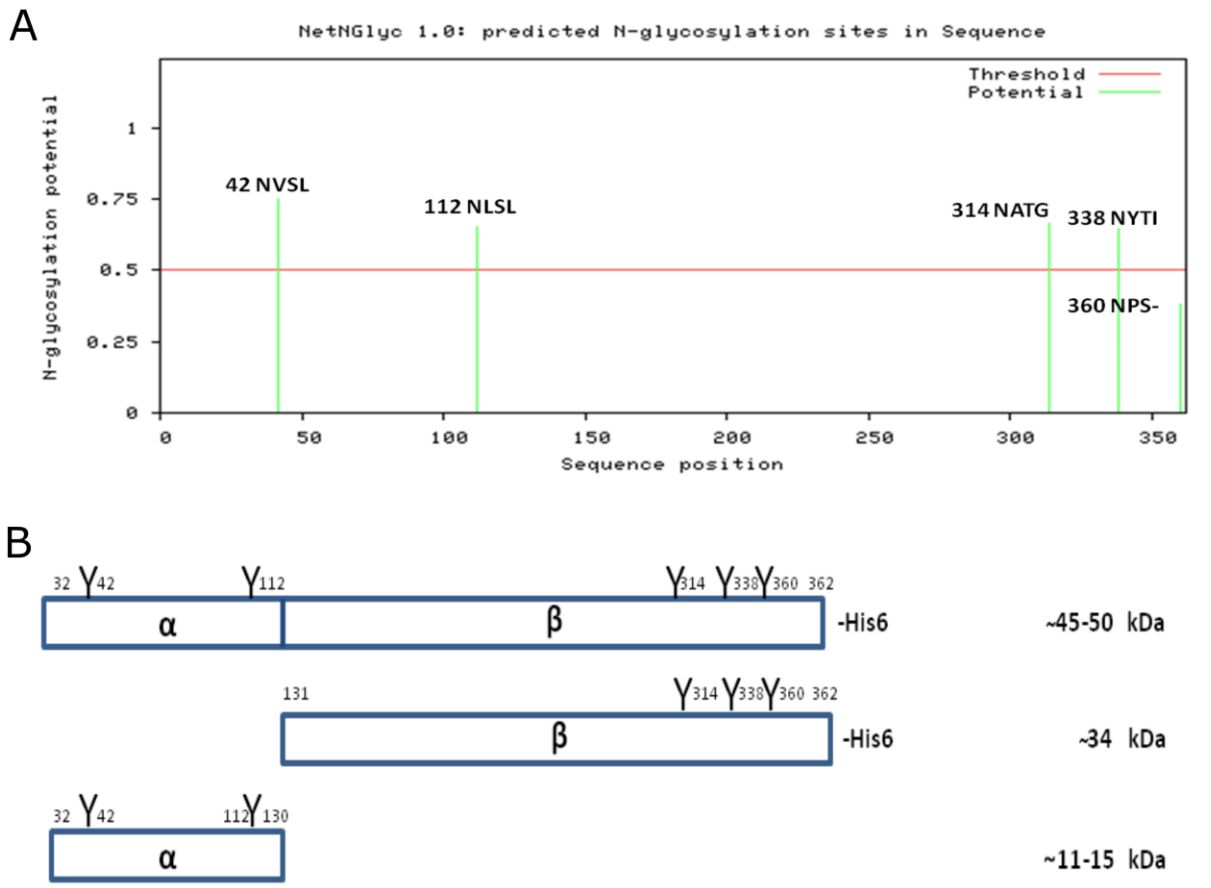


Figure 3.6: The predicted N-glycosylation site and observed molecular weight of mouse NAAA. A: Predicted N-glycosylation site in mouse NAAA using NetNGlyc 1.0 server B: The schematic diagram of recombinant mouse NAAA showed that different subunit with sequence number, glycosylation number and molecular weight.

The purified mouse NAAA is known to have heterogeneity due to post-translational processes such as glycosylation and autoproteolytic clevage. Kit based glycosylation staining on mouse NAAA confirmed glycosylation. The ~45-50, ~34 and ~11-15 kDa bands were stained magenta, implicating glycosylation that the ~45-50, ~34 and ~11-15 kDa bands were glycosylated which were the unprocessed, α-subunit and β-subunit of mouse NAAA respectively (Figure 3.5 D). The enzymatic deglycosylation using PDM II appeared to decrease the protein size to about 38, 27 and 11 kDa (Figure 3.7 B and C). The glycosylation content of the recombinant enzyme clearly confirmed by deglycosylation using PDM II and glycoprotein staining kit but the structure of oligosaccharides has not been investigated.

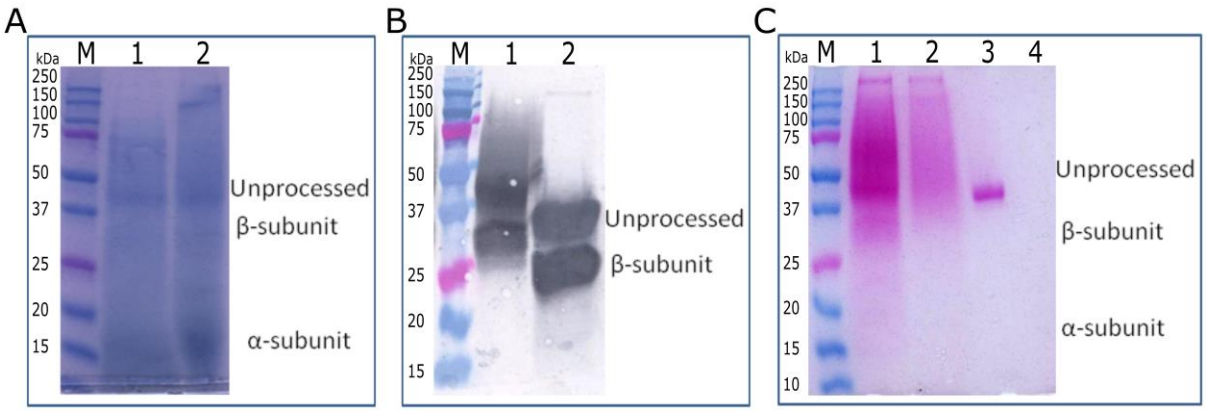


Figure 3.7: N-glycosylation analysis of mouse NAAA. A: 8-16% SDS-PAGE analysis showing that the electrophoretic mobility of the cleaved form of mouse NAAA changes in reducing and non-reducing conditions. M, Prestained protein ladder; Lane 1, Mouse NAAA protein with 5% BME and heat; Lane 2, Mouse NAAA protein without β mercaptoethanol (BME) and heat. B: 100 µg of mouse NAAA protein were treated with deglycosylation mix for 16 h, separated by SDS–PAGE and analyzed by western blotting using with monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 2, Mouse NAAA protein; Lane 1, Mouse NAAA protein was treated with deglycosylation mix. C: 100 µg of mouse NAAA protein were treated with deglycosylation mix for 16 h, separated by SDS–PAGE and analyzed by glycosylation staining kit. M, Prestained protein ladder; Lane 1, Mouse NAAA protein; Lane 2, Mouse NAAA protein was treated with deglycosylation mix; Lane 3, Horseradish peroxidase as a positive control; Lane 4, Soybean trypsin inhibitor as a negative control.

3.3.7 Enzyme assay and steady-state kinetics of mouse NAAA

The O-Pthaladehyd assay has used the estimation of the hydrolyzed by-product ethanolamine by recording absorbance at 340 nm. The purified enzyme mouse NAAA was assayed for NAE hydrolase activity with PEA as substrate and the reaction was carried out at pH 4 and temperature 37°C. Optimal enzyme activity was at pH 4 (sodium acetate buffer) in the assay condition (Figure 3.8 C). The optimum temperature for enzyme activity was found to be 35°C (Figure 3.8 D). The optimum time and DTT concentration also observe 24 hrs and 8 mM, respectively (Figure 3.8 A and B). Optimum pH of hNAAA is 4.5 for the hydrolysis of N-palmitoylethanolamine as substrate (Tsuboi *et al.*, 2005; Wang *et al.*, 2008), activity stimulated by DTT and Nonidet P-40 (Tsuboi *et al.*, 2005).

Mouse enzyme followed the Michaelis-Menten (MM) kinetics (Figure 3.9). The values of K0.5, h and turnover number (kcat) for substrate PEA was determined by incubating the mouse NAAA with a range of concentrations of PEA (10-500 µM) under the standard assay conditions.

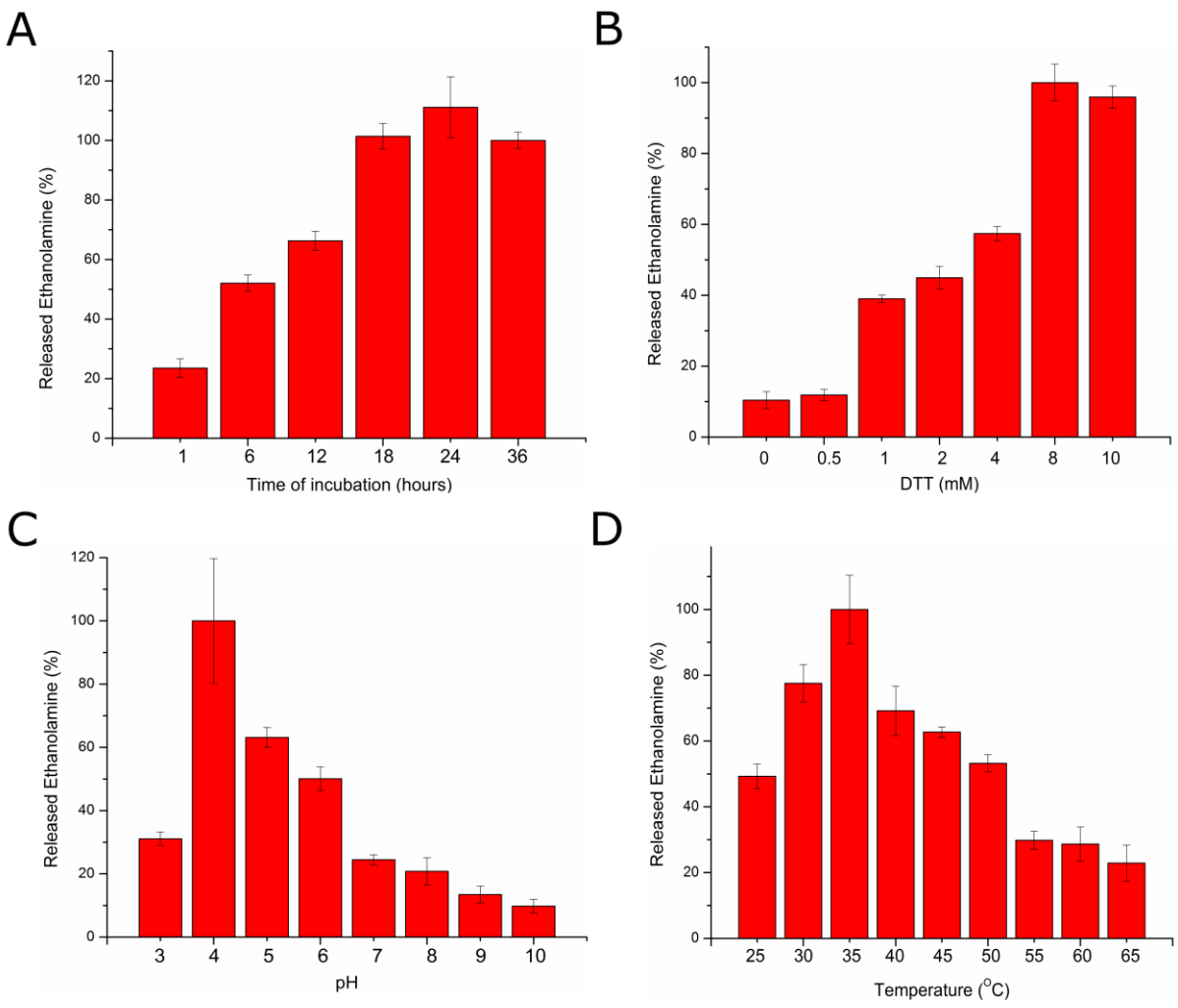


Figure 3.8: Optimum assay condition of mouse NAAA. A: Optimum time. B: DTT concentration. C: Optimum pH. D: Optimum temperature determined for the recombinant mouse NAAA.

The concentration of mouse NAAA used was 1µg/reaction for all kinetic studies (The final concentration of enzyme in 500 µl assay reaction is 44.4 nM). The kinatics parameters K0.5, h and Vmax were calculated using Hill's equation. The kinetic parameters of mouse NAAA calculated for substrate PEA are summarised in Table 3.1.

Table 3.1: Steady-state kinetic parameters of mouse NAAA.

Substrate	**$K_{0.5}$**	**h**	**V_{max}**	**Kcat (S^{-1})**	**Kcat/$K_{0.5}$ (M^{-1} S^{-1})**
PEA	67.8669±2.12769	2.2298±0.2897	214.9658±6.6592	0.2537±0.01866	0.0037

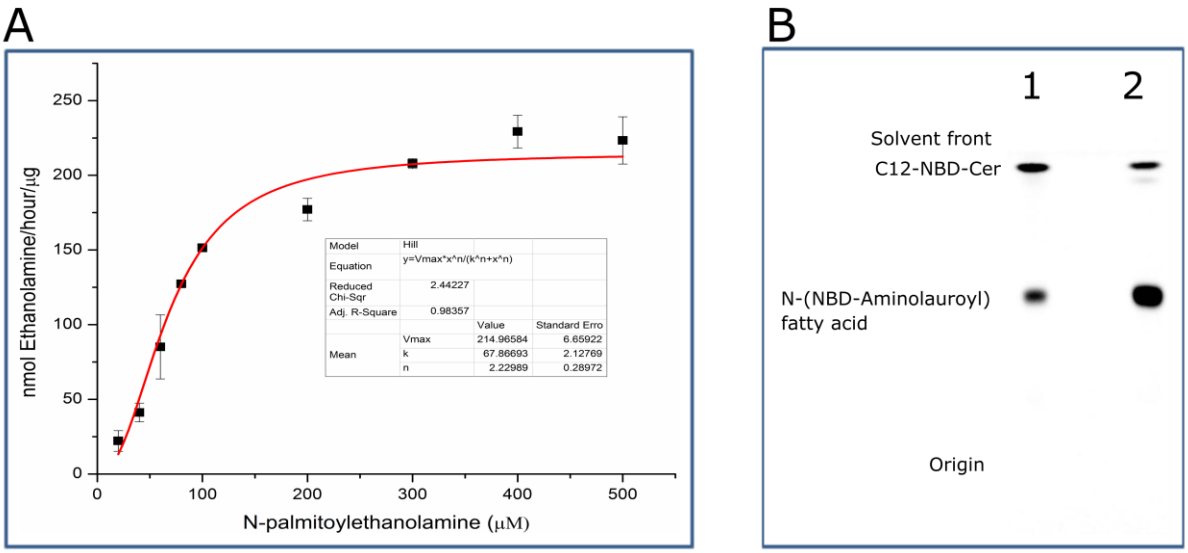


Figure 3.9: Steady-state kinetics of recombinant mouse NAAA. A: Steady-state kinetic graphs of recombinant mouse NAAA with PEA. B: Detections of fluorescent fatty acid released from labeled ceramides by the action of an NAAA Lanes 1, C12-NBD-Ceramide extracted after treated with mouse NAAA; Lane 2, C12-NBD-Ceramide extracted after treated with *M. commoda* NAAA.

Table 3.2: Comparison of mouse NAAA activity with other reported hNAAA.

Sr. No.	Substrate	$K_{0.5}$	V_{max}	pH	Temp	Reference No.
1	PEA	97 µM		4.5	37	(Tsuboi *et al.*, 2005)
2	PEA	21 µM	5.4 nmol/min/mg	4.5	37	(West *et al.*, 2012)
3	PEA	67 µM	214 nmol/h/ µg	4	37	This Study

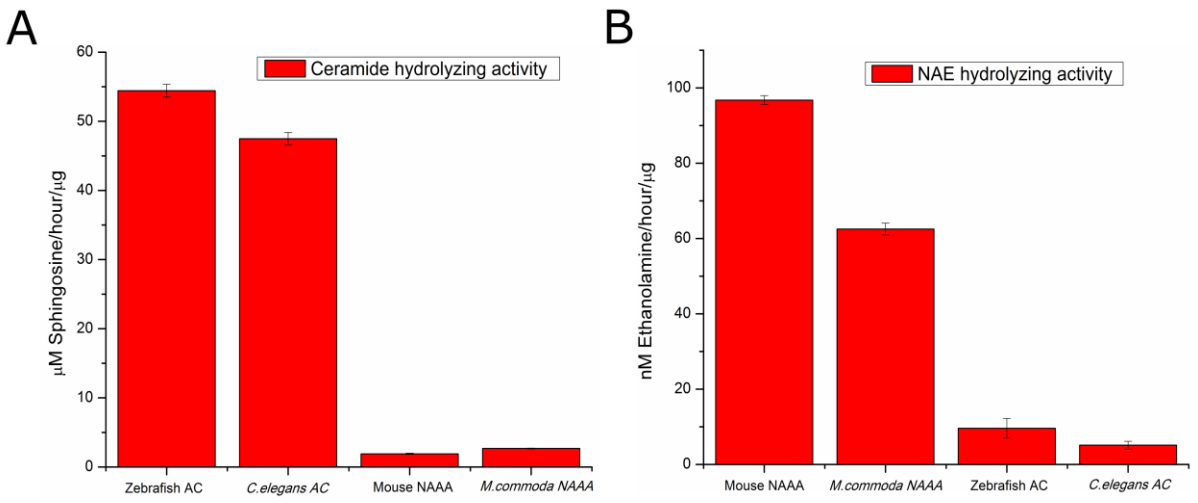


Figure 3.10: Ceramide and NAE hydrolyzing activity of NAAA and AC. A: N-Lauroyl-D-sphingosine (200 µm) hydrolyzed by AC and NAAA B: N-palmitoylethanolamine (50 µm) hydrolyzed by NAAA and AC.

Mouse NAAA show cross-reactivity against C12-NBD-ceramide (Figure 3.9 B: lane 1). Mouse NAAA and *M. commoda* NAAA also hydrolyzed ceramide N-Lauroyl-D-sphingosine (Figure 3.10 A). It is observed that *M. commoda* NAAA showing more activity against N-Lauroyl-D-sphingosine but less activity against PEA compared to mouse NAAA (Figure 3.10 A and B). Zebrafish AC and C. elegans show very less activity but detectable against PEA (Figure 3.10 B).

3.3.8 Circular dichroism and Fluorescence spectroscopy of mouse NAAA

Intrinsic fluorescence spectra of mouse NAAA recorded using fluorescence spectroscopy, where protein show λ max at 353.5 nm and spectra also indicate that most of the tryptophan exposed to surface (Figure 3.11 A).

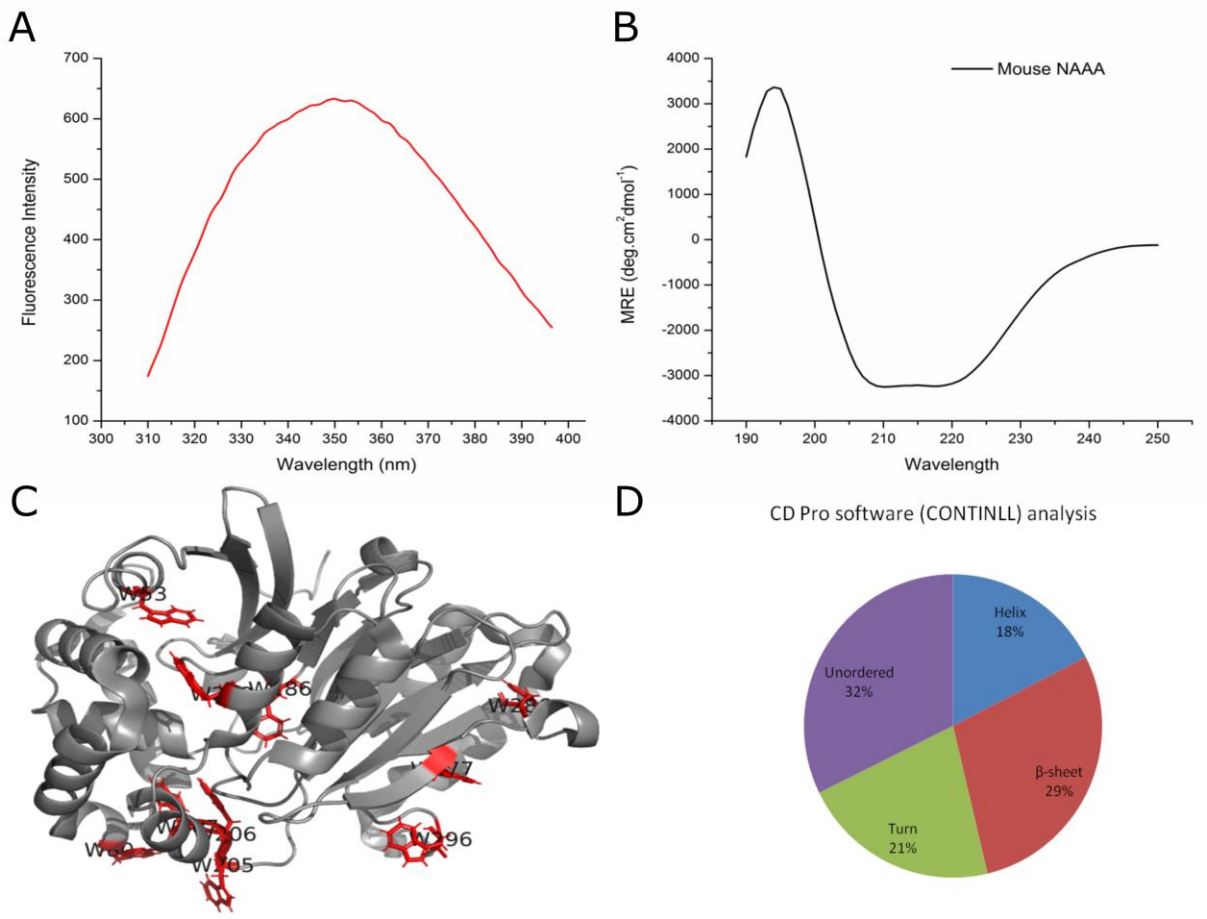


Figure 3.11: Fluorescence and CD spectra of mouse NAAA. A: Fluorescence scan of mouse NAAA at pH 7.0 (0.1 mg ml^{-1}). B: Far-UV CD scan of mouse NAAA at pH 7.0 (0.348 mg ml^{-1}). C: Ribbon diagram of mouse NAAA showing the location of tryptophan residues (red) in protein (PDB: 6dxy). The diagram is prepared in PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrodinger, LLC.). D: CD Pro software analysis of mouse NAAA.

The far UV CD spectrum of mouse NAAA had positive ellipticity maxima at 194 nm and negative minima at 211 nm and 217 nm (Figure 3.11 B). The secondary structure content of mouse NAAA was estimated as 18% α-helix and 29% β-sheet from spectrum using CDpro, this data suggested that mouse NAAA is rich in β-sheet with slight α-helix. The ordered CD profile also indicates that the recombinant protein was in a properly folded form. Thermal stability of the protein provides information about the non-covalent and covalent forces involved in stabilizing the protein structure. Secondary structure monitoring during thermal denaturation by CD spectroscopy revealed almost unperturbed CD profile up to 75°C. CONTINLL analysis indicates that the β-sheet content of the protein decreased significantly above 75°C with a simultaneous increase in the α-helix.

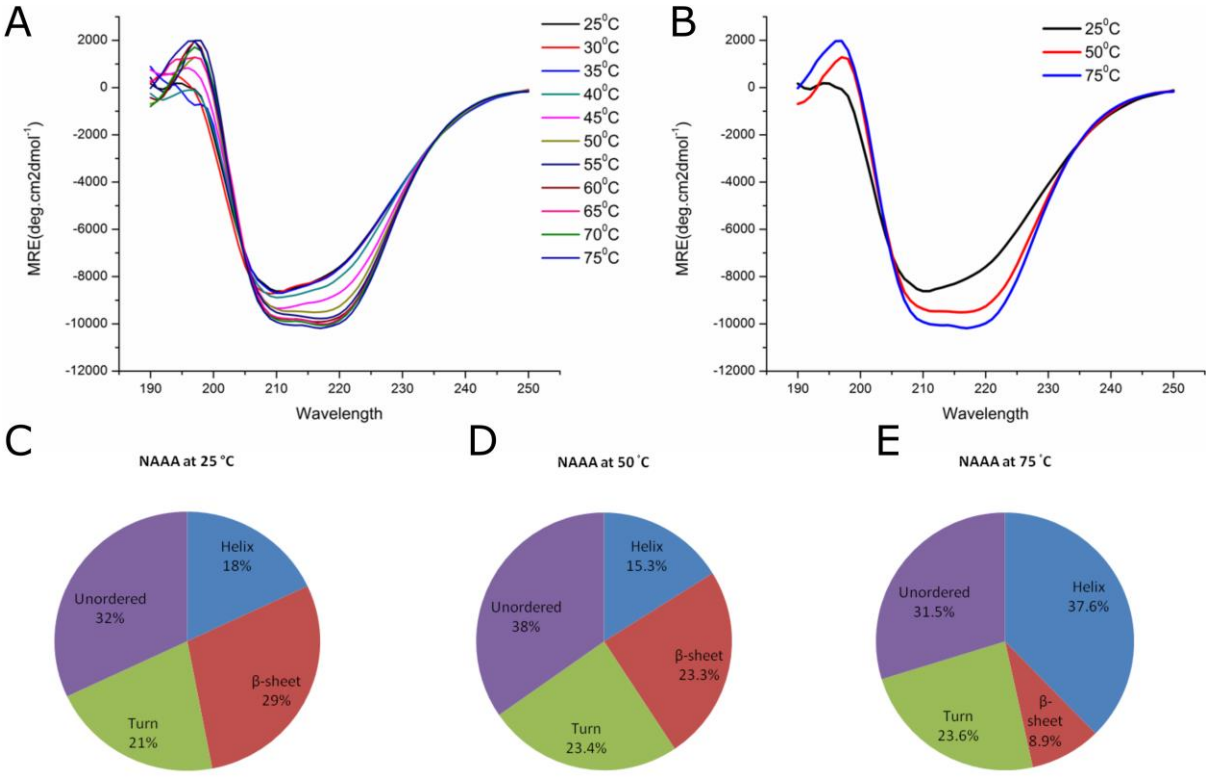


Figure 3.12: Thermal denaturation of mouse NAAA using the CD. B: Far-UV CD spectra of mouse NAAA incubated at different temperatures for 5 min, 0.348 mg ml^{-1} of protein was used, each sample was incubated for 5 min at the respective temperature and then the scans were recorded.

3.3.9 Heterologous expression of *M. commoda* NAAA

The recombinant P. pastoris GS115 strain having *M. commoda* NAAA gene was inoculated in a 500 ml shake flask with 25 ml BMGY and grown for 48 h at 28 °C, 200 rpm. Cells pellets were harvested by centrifugation and resuspended in a 2 L shake flask with 300

ml BMMY. The culture was grown in BMMY medium for 144 h at 28 °C, 200 rpm and methanol induction was given every 24 h for expression naaa gene. The culture medium was duly collected and cell debris was removed. The cell-free supernatant was concentrated and desalted by dialysis. The supernatant was subsequently analyzed on 10% SDS-PAGE. Protein bands of unprocessed NAAA approximately ~50-45 kDa were observed along with a significant amount of processed form of NAAA approximately 30-40 and14 kDa, α and β-subunits respectively (Figure 3.14 A). Expression of *M. commoda* naaa gradually increased over time until the 4th day (94 hrs) but after 72 hrs unprocessed protein intensity decrease, as well as processed bands, increased. After72 hrs grown culture supernatant did not show any band in western blot but show activity against PEA. After 72 hrs grown culture supernatant purified Ni-NTA affinity chromatography does not have any band in SDS-PAGE and also did not show activity also.

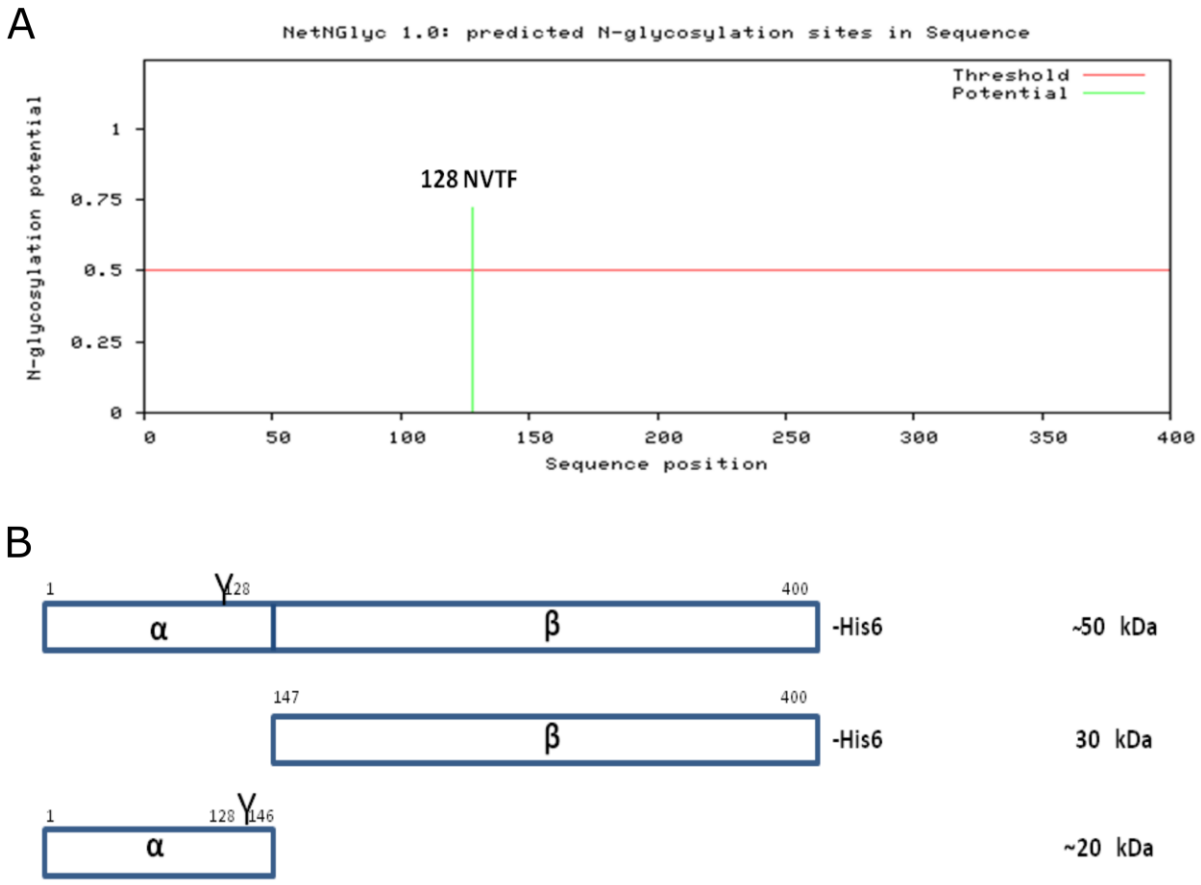


Figure 3.13: The predicted N-glycosylation site and observed molecular weight of *M. commoda* NAAA. A: Predicted N-glycosylation site in *M. commoda* NAAA using NetNGlyc 1.0 server B: The schematic diagram of recombinant *M. commoda* NAAA showed that different subunit with sequence number, glycosylation number and molecular weight.

This observation clearly suggests that after 72 hrs M commoda NAAA processed into α-subunit and β-subunit in pH 8. The processed NAAA difficult to catch by Ni-NTA affinity chromatography, which was probably the result of carboxy-terminal processing similar to human AC (He *et al.*, 2003).

3.3.10 Purification of *M. commoda* NAAA

M. commoda NAAA was purified using Ni-NTA affinity chromatography after 48 hrs. Majority of the purified enzyme was obtained in the unprocessed form (Figure 3.14 B).

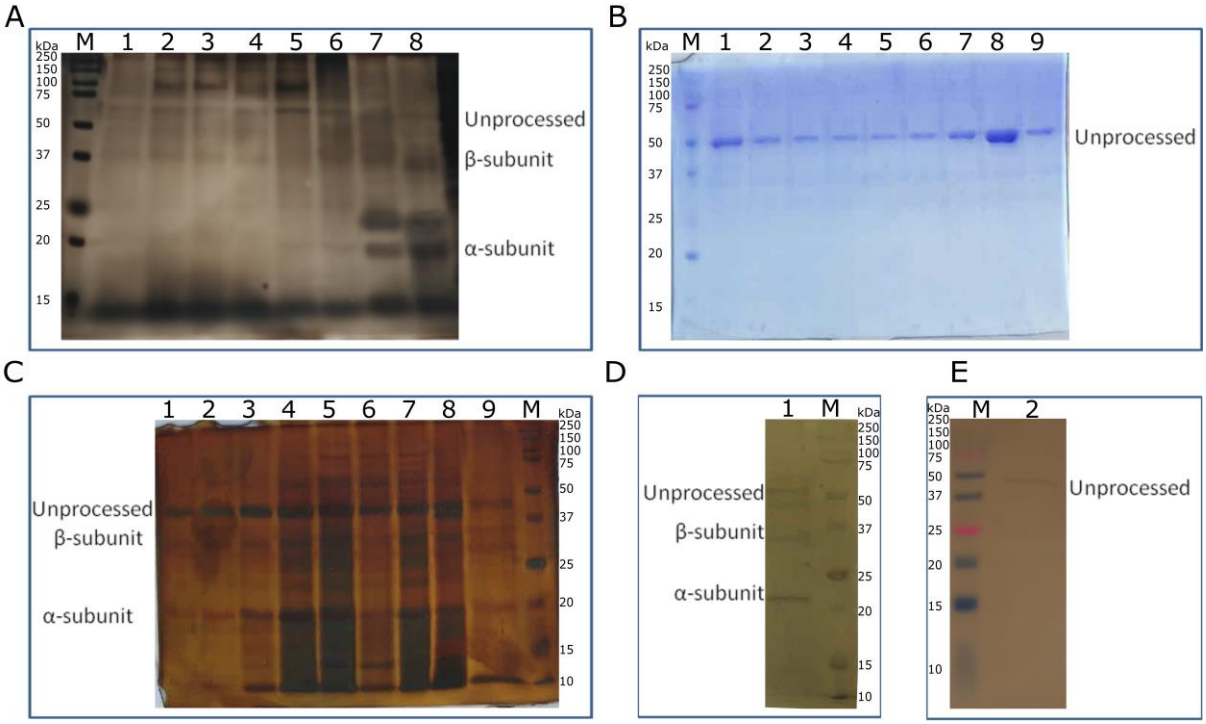


Figure 3.14: Purification and characterization of *M. commoda* NAAA. A: 12 % SDS-PAGE analysis shows expression of *M. commoda* NAAA protein in Pichia pastoris. M, Prestained protein ladder; Lanes 1-8, 20 µl of the concentrated supernatant after being induced by methanol for 12, 24, 36 48, 60, 72, 84 and 96, h, respectively. B: 12% SDS-PAGE analysis shows purification of *M. commoda* NAAA protein using Ni-NTA affinity chromatography. M, Prestained protein ladder; Lane 1, Unbound; Lane 2-5, Wash fraction; Lane 6-9, Eluted fraction. C: 10% SDS-PAGE analysis shows purification of *M. commoda* NAAA protein using Concanavalin A–sepharose affinity chromatography. Lane 1-3, Eluted fraction (300 mM methyl α-D-glucopyranoside); Lane 4-6, Eluted fraction (10 mM methyl α-D-glucopyranoside); Lane 7-8, Wash fraction; Lane 9, Unbound, M, Prestained protein ladder. D: Self-processing of purified recombinant *M. commoda* NAAA. Lanes 1-2, *M. commoda* NAAA of incubated 24 hrs at 4 °C and analyzed on 10% SDS-PAGE; M, Prestained protein ladder. E: Western blotting analysis confirms *M. commoda* protein expression with monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 1, *M. commoda* protein was purified using Ni-NTA affinity chromatography.

Protein detected by western blot after purification with monoclonal anti-His antibodies (Figure 3.14E). The catching his tag in western blot more difficult compared to catching his tag in Ni-NTA affinity chromatography because of anti-penta his-tag antibody required at least intact 5 his tag. The purified protein after His-Tag affinity chromatography incubated for 4 °C for unprocessed precursor (immature form) convert into matures heterodimer (Figure 3.14 D). The purified M commoda NAAA containing the mixture of glycosylated and nonglycosylated forms. By Concanavalin A-sepharose chromatography, the nonglycosylated NAAA was successfully separated from the glycosylated one (Fig. 3.14 C: lane 1, 2 and 3).

On the basis of gel electrophoretic results, the molecular masses corresponding to the bands ~50, 30 and ~20 kDa were predicted with slight variation in the molecular weight than the predicted in silico amino acid sequence 47, 30 and 17 kDa for unprocessed *M. commoda* NAAA, β- and α-subunit, respectively (Figure 3.14). One putative N-glycosylation sites at amino acid positions 128 (Figure 3.13) and two potential O-glycosylation site at 49 and 50 were predicted in the *M. commoda* NAAA when the amino acid sequence was analyzed using NetNGlyc 1.0 Server (Gupta, Jung and Brunak, 2016) and NetOGlyc4.0 Server, respectively (Steentoft *et al.*, 2013). The β-subunit does not have any putative glycosylation site. *M. commoda* NAAA show cross-reactivity against C12-NBD-ceramide (Figure 3.10 B). *M. commoda* shows more activity compare to mouse NAAA.

3.3.11 DSF studies of *M. commoda* NAAA

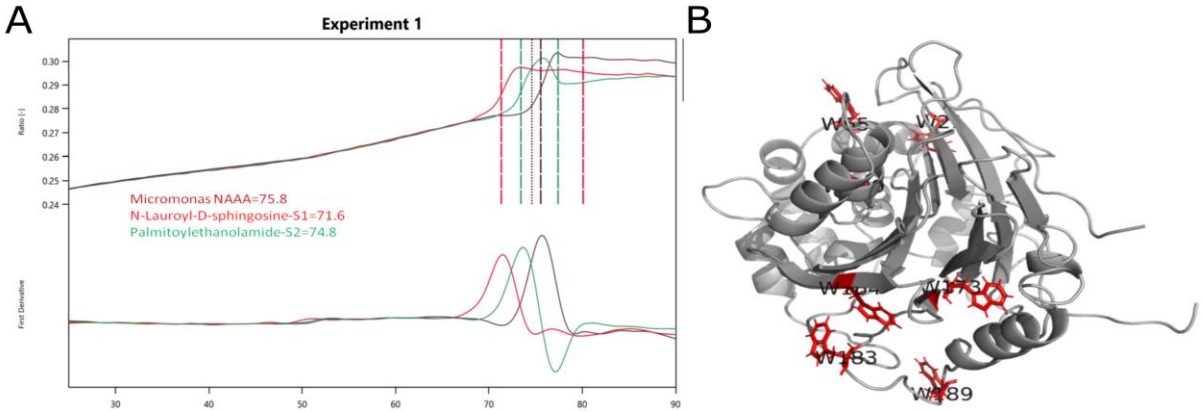
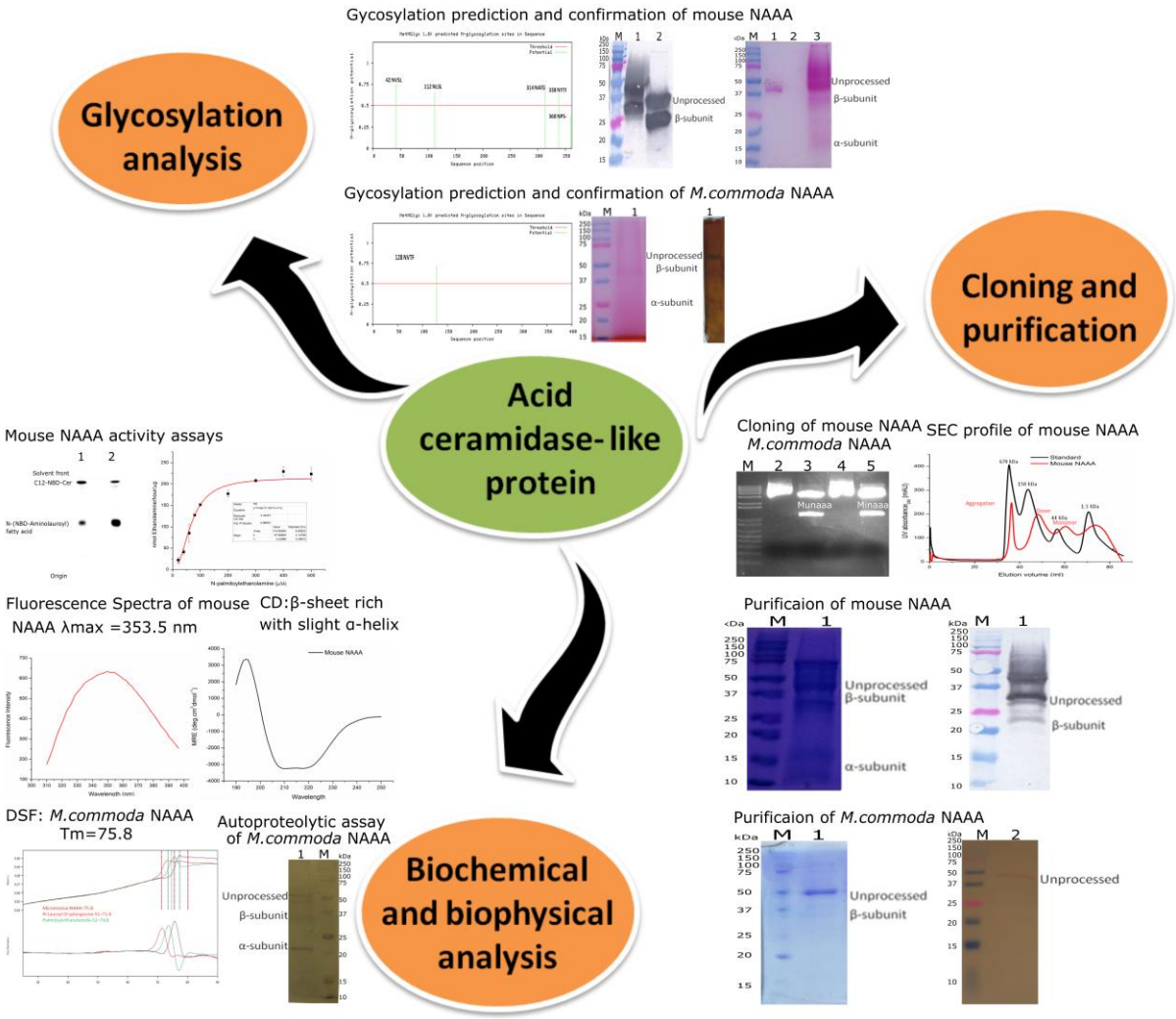


Figure 3.15: Thermal unfolding curves of *M. commoda* NAAA using DSF. A: Thermal unfolding curves, indicating the denaturation of *M. commoda* NAAA using DSF with or without substrates. B: Ribbon diagram of *M. commoda* NAAA showing the location of tryptophan residues (red) in protein.

Temperature-induced protein denaturation was studied using DSF. In DSF of an enzyme, intrinsic fluorescence emitted by tryptophan residue quantifies to study the directed movement of protein molecules along microscale temperature gradients (Jerabek-Willemsen *et al.*, 2011). From the thermophoretogram, The Tm of *M. commoda* NAAA and with substrates N-lauroyl-D-sphingosine and N-palmitoylethanolamine was determined as 75.8, 71.6 and 74.8, respectively (Figure 3.15 A: lane 2).

3.4 Graphical summary

Schematic illustration of NAAA characterization



Chapter 4: In silico study of acid ceramidase (AC) and acid ceramidase-like protein (NAAA)

4.1 Introduction

Acid ceramidase (AC) is encoded by asah1 gene and catalyzes the hydrolysis of the amide bond in ceramide. Mutations in this gene have been associated with FD (Koch *et al.*, 1996) and SMA-PME (Zhou *et al.*, 2012). It is synthesized as inactive premature precursor polypeptide (signal, peptide, α-chain, β-chain), which is subsequently processed into mature active heterodimer consisting of α- and β-chains in lysosomes.

Farber lipogranulomatosis shows autosomal recessive inheritances as described and characterized by Sidney Farber (FARBER, COHEN and UZMAN, 1957; Alves *et al.*, 2013). Characteristic features of this disease are the formation of subcutaneous nodules, painful and progressively deformed joints and hoarseness (Alves *et al.*, 2013). This disease further classified into five types. Accumulation of ceramide in subcutaneous nodule and absence of ceramidase activity in kidney and cerebellum from a patient with FD suggests deficiency of ceramidase in the patient (Sugita *et al.*, 1975). Lysosomal acid ceramidase deficiency in this disease has also been confirmed by analysis of ceramide and acid ceramidase activity in Farber fibroblast cell and normal fibroblast cell (Jameson, Holt and Keen, 1987; Bär *et al.*, 2001). First, homozygous point mutation (T222K) in FD was identified by cloning and full-length cDNA characterization (Koch *et al.*, 1996). There are many more novel mutations identified in the asah1 gene causing FD (Amirhakimi *et al.*, 1976; Fensom *et al.*, 1979; Cartigny *et al.*, 1985; Ben-Yoseph Y, Gagné R, Parvathy MR, Mitchell DA, 1989; Kattner, Schäfer and Harzer, 1997; Li *et al.*, 1999; Zhang *et al.*, 2000; Bär *et al.*, 2001; Muramatsu *et al.*, 2002; Devi *et al.*, 2006; Bedia *et al.*, 2010; Cvitanovic-Sojat *et al.*, 2011; Lucki *et al.*, 2012; Chedrawi *et al.*, 2012; Alves *et al.*, 2013; Knorr, Rudolf and Nuernberger, 2013; Dyment *et al.*, 2014; Bashyam *et al.*, 2014; Kim *et al.*, 2016).

A homozygous missense mutation (C125T) (Thr42Met) in asah1 was identified by genomic-wide SNP genotyping and exome sequencing of families affected by childhood SMA-PME. An acid ceramidase activity below 32% causes SMA-PME but not FD. Knockdown of asah1 in zebrafish results in defective motor neurons (Zhou *et al.*, 2012). ASAH1 gene of hAC (N-acylsphingosine acyl hydrolase) is located on chromosome 8 (8p21.3-22) & contain 14 exons and 13 introns. In 2018, Ahmad et al. resolved the crystal structure of both immature and mature forms of AC and was described in the immature form.

AC consists of two subunits, namely an α-subunit (22-142 aa) and a β-subunit (143-395 aa) and a disulfide bridge connects these two subunits. The first asah1 mutation was reported in 1996 (Koch *et al.*, 1996). Till date, 47 mutations associated with FD and SMA-PME have been deposited and documented in the **H**uman **G**ene **M**utation **D**atabase (HGMD) (Stenson *et al.*, 2017). The recent crystal structure of hAC has endorsed investigators to analyze the effect of disease-associated mutations on the structure and function of the enzyme

In this context, understanding the pathogenic or neutral role of *asah1* mutations in the development of FD and SMA-PME has significant implications for early diagnosis and treatment. Understanding a structural and functional context for all mutations is expensive and time-consuming. Therefore, we performed an *in silico* analysis as an efficient alternative.

4.2 Materials and Methods

4.2.1 Comparative modeling

The sequence of zebrafish AC and *M. commoda* NAAA was taken from the UniProtKB (http://uniprot.org) protein databank (Entry number: C1E7W0 and Q6PH71). A similarity search carried out using the BLAST server (Altschul *et al.*, 1990), identified in Protein Databank (PDB) (http://www.rcsb.org), the acid ceramidase (PDB ID: 5U7Z) of mammals, which had 57% sequence identity with the corresponding zebrafish AC and the acid ceramidase like-protein (PDB ID: 6DXW) of mammals, which had 34% sequence identity with the corresponding *M. commoda* NAAA. The inactive zebrafish AC was modeled using Robetta server (http://robetta.bakerlab.org) (Kim, Chivian and Baker, 2004), whereas active NAAA was modeled using an acid ceramidase structure of human (resolution: 2.5 A). The inactive *M. commoda* NAAA was modeled using an acid ceramidase like-protein structure of mouse (resolution: 2.3 A), whereas active NAAA was modeled using an acid ceramidase structure of mice (resolution: 1.8 A). The homology model of AC and NAAA was built with Prime 3.1 in Schrödinger Suite (Schrödinger, LLC, New York, NY). The secondary structure of NAAA was predicted using the SSpro program bundled with Prime. The target (AC and NAAA) and template (AC and NAAA) sequences were aligned using the Clustal W method employed in Prime, followed by manual adjustment to avoid big gaps in the secondary structure domain.

4.2.2 Validations of homology models and refinement

The protein model was imported to the maestro window (Schrödinger, Inc.). The protein structure was energy minimized using the protein preparation wizard by applying the OPLS_2005 force field. Progressively weaker restraints were applied to non-hydrogen atoms only. The molecular models were evaluated using PROCHECK (Laskowski *et al.*, 1993), ERRAT 2 (Colovos and Yeates, 1993), ProSA-web Protein structure analysis (Wiederstein and Sippl, 2007). The ERRAT program verifies the quality of the model. This program plot error values as a function of position in the sequence by sliding a nine-residue window along the sequence. The error function is based on the statistics of non-bonded atom-atom interaction in the template structure. ProSA is a tool widely used to check errors in 3D protein models. An overall quality score or Z score is estimated and shown in a plot where scores estimated from experimentally determined structures in PDB are plotted. The Z score is an indication of the overall quality of the model and measures the deviation of the total energy of the modeled structure from an energy distribution derived from random conformations.

4.2.3 Docking of ceramide in zebrafish AC

The homology model of mature zebrafish AC used to dock N-hexanoyl-D-*erythro*-sphingosine using Glide (Schrodinger) (Friesner *et al.*, 2006). The model structure was imported and centroid receptor grid was generated around the residue Cys 143 known to interact with ceramide. The ligand molecule N-hexanoyl-D-*erythro*-sphingosine (ceramide) (CID 16219484) was downloaded from the PubChem database (Bolton *et al.*, 2008) and prepared in the LigPrep v2.5 of the Schrodinger suite. This was then docked in mature zebrafish AC structure.

4.2.4 Retrieving Disease-causing nsSNPs

Information regarding the nsSNPs in the coding region of the *asah1* gene was retrieved from the dbSNP (Bhagwat, 2010), UniProt, HGMDand OMIM databases. 26 nsSNPs caused FD, and four caused FD nsSNPs SMA-PME, according to published reports was chosen for *in silico* analysis. Structure of the human AC protein was obtained from PDB protein structure database (PDB ID: 5U7Z).

4.2.5 Analysis of the most damaging nsSNPs

4.2.5.1. SIFT

Sorting **I**ntolerant **F**rom **T**olerant (SIFT 2) (http://sift-dna.org/sift4g) predicts the tolerated and deleterious SNPs and identifies the impact of amino acid substitution on protein function and phenotype alterations. According to the previously reported data, SIFT discriminates between functionally neutral and deleterious polymorphisms during mutagenesis studies in humans (Vaser *et al.*, 2016). SIFT performs analysis on the basis of different algorithms and interprets the homologous sequences using the Swiss-Prot (version 51.3) and TrEMBL (version 34.3). The threshold intolerance score for SNPs is 0.05 or less.

4.2.5.2. PolyPhen

Polymorphism **Phen**otyping (PolyPhen2) (http://genetics.bwh.harvard.edu/pph2/) predicts the influence of amino acid substitution on the structure and function of proteins by using the specific empirical rules. Protein sequence, database ID/accession number, amino acid positionand amino acid variant details are the input options for PolyPhen (Adzhubei, Jordan and Sunyaev, 2013). The tool estimates the **P**osition-**S**pecific **I**ndependent **C**ount (PSIC) score for every variant and calculates the score difference between variants.

4.2.5.3 PROVEAN

PROtein **V**ariation **E**ffect **AN**alyzer (PROVEAN) (http://provean.jcvi.org) predicts the functional effect of amino acid substitutions. The threshold of prediction is -2.5, above this score prediction, is supposed to be neutral and below or equal -2.5 prediction is deleterious (Choi and Chan, 2015).

4.2.5.4 PhD-SNP

Predictor of **h**uman **D**eleterious **S**ingle **N**ucleotide **P**olymorphisms (PhD-SNP) (http://snps.biofold.org/phd-snpg) is a single-sequence SVM method (SVM-Sequence) that discriminates disease-related mutations based on the local sequence environment of the mutation (Capriotti and Fariselli, 2017).

4.2.4.5. SNPs and GO

Single **N**ucleotide **P**olymorphism Database (SNPs) and **G**ene **O**ntology (GO) (http://snps.biofold.org/snps-and-go/snps-and-go.html) are support vector machine (SVM) based accurate methods used to predict the disease-related mutations from protein sequences

with a scoring accuracy of 82% and Matthews correlation coefficient of 0.63. For SNPs and GO, the FASTA sequence of the whole protein is considered to be an input option and output will be the prediction results based on the discrimination among disease-related and neutral variations of the protein sequence. The probability score higher than 0.5 reveals the disease-related effect of the mutation on the parent protein function (Thomas *et al.*, 2003; Calabrese *et al.*, 2009).

4.2.6. Prediction of change instability due to mutation

4.2.6.1. I-Mutant

I-Mutant 2 (http://folding.uib.es/i-mutant/i-mutant2.0.html) is a support vector machine (SVM) based server capable of doing automatic prediction of protein stability changes arising from a single point mutation We used sequence and structure-based version of I-mutant 2 and the protein stability change was predicted as an increase or decrease for AC protein sequences retrieved from NCBI (Capriotti, Fariselli and Casadio, 2005).

4.2.6.2 FOLD-X

FOLD-X is a computer algorithm for quantitative estimation of interactions facilitating the stability of proteins. The FOLD-X tool provides a comparison between wild type and mutant models in the form of van der Waals clashes, which greatly influence the energy decomposition. Sometimes the mutations cause strain in the original native structure and sometimes reduce (Van Durme *et al.*, 2011).

4.2.6.3 PoPMuSiC

PoPMuSiC is a web server that can predict the changes in protein stability caused by point mutations through using a linear combination of statistical potentials whose coefficients depend on the solvent accessibility of the mutated residue (Dehouck *et al.*, 2011). Being different from I-Mutant, the ΔΔG value predicted by FoldX and PoPMuSiC is positive when the mutation is destabilizing and negative when it is stabilizing.

4.2.7 Biophysical characterization

4.2.7.1 SNPeffect 4.0

SNPeffect 4.0 (De Baets *et al.*, 2012) integrates aggregation prediction (TANGO), amyloid prediction (WALTZ), chaperone-binding prediction (LIMBO)and protein stability analysis (FoldX) for structural phenotyping. Mutations are classified as mutations that

increase (dTANGO>50), decrease (dTANGO< -50), or do not affect (dTANGO between -50 and 50) the propensity of the protein to aggregate and as mutations that increase (dWALTZ>50), decrease (dWALTZ< -50), or do not affect (dWALTZ between -50 and 50) the amyloid propensity of the protein.

4.2.8 Evolutionary conservation analysis

4.2.8.1 ConSurf

ConSurf is a bioinformatics tool that we used to estimate the evolutionary conservation of amino acid positions using protein sequence (Ashkenazy *et al.*, 2016). The analysis is based on phylogenetic relations between homologous sequences. Degree of conservation of amino acid residues was estimated using 50 homologous sequences. We selected those highly conserved residues for further analysis which were located at the sites of high-risk nsSNPs.

4.2.8.2 NetSurfP

NetSurfP2 server (http://www.cbs.dtu.dk/services/NetSurfP/) predicts the surface accessibility, secondary structure, disorderand backbone dihedral angles of each amino acids when they are in a sequence. The reliability of this prediction method is in the form of Z-score. The Z-score highlights the surface prediction reliability, but not associated with a secondary structure (Klausen *et al.*, 2019). The Z-score is related to the surface predictionand not the secondary structure.

4.2.9 Modeling of mutant structures

4.2.9.1 Fold X

Mutations were created using the protein design tool of FoldX (version 3 beta) using the PDB 5U7Z as the wild-type (WT) structure.

4.2.9.2 Project HOPE

Project **H**ave y**O**ur **P**rotein **E**xplained (HOPE) (http://www.cmbi.ru.nl/hope/home) is an automatic mutant analysis server to study the insight structural features of the native protein and the variant models. HOPE provides the 3D structural visualization of mutated proteins and gives the results by using UniProt and DAS prediction servers. The input method of Project HOPE carries the protein sequence and selection of Mutant variants.

HOPE server predicts the output in the form of structural variation between mutant and wild type residues (Venselaar *et al.*, 2010).

4.2.10 Identifying structural and functional properties

4.2.10.1 MutPred

MutPred 1.2 was consulted which is a web-based application tool that effectively screens amino acid substitutions (Pejaver *et al.*, 2017). It also helps into predict the molecular cause of the disease. MutPred is based upon gain/loss of 14 different functional and structural properties like loss of a phosphorylation site or gain of helical propensity. Protein sequence (FASTA format) of AC and its amino acid substitutions were submitted. The output provides p-value where; p <0.05 and p <0.01 were considered as confident and very confident hypotheses, respectively.

4.2.10.2 Predicting Post Translational Modification (PTM) sites

The putative methylation sites in the AC protein sequence were predicted by **P**rediction **S**pecies-**S**pecific of **Me**thylation (PSSMe) and **P**rediction of **UB**iquitination sites with **B**ayesian **D**iscriminate **M**ethod (BPB-PPMS). The former tool identifies methylation sites based on information gain feature optimization method. The higher **S**upport **V**ector **M**achines (SVMs) probability indicates a higher probability of lysine (or arginine) to get methylated. False-positive predictions were controlled by focusing on sites with stringency setting higher than 50% (Wen *et al.*, 2016). Likely phosphorylation sites in AC protein at serine, threonine and tyrosine residues were predicted using NetPhos 3.1. NetPhos 3.1 uses ensembles of neural networks to complete this task and residues having scores >0.5 thresholds are considered phosphorylated (Blom, Gammeltoft and Brunak, 1999). Likewise in GPS 3.0, higher value depicts the higher potential of the residue to get phosphorylated (Xue *et al.*, 2008). Putative protein ubiquitylation sites were predicted by BDM-PUB. In BDM-PUB, the balanced cut-off option was selected.

4.2.11 MD simulations

MD simulations of wild type and mutant models were performed using the GROMACS 4.6.3 software package (Hess *et al.*, 2008) with the AMBER99SB-ILDN protein, nucleic AMBER94 force field implemented on a parallel architecture (Lindorff-Larsen *et al.*, 2010). In each case, the treated structure was initially soaked in a periodic dodecahedron box filled with the single point charge (SPC216)water molecules (Berendsen *et al.*, 1981), with

the minimum distance of solute-box edges of 1.0 nm. Periodic boundary conditions were applied in all directions. To compensate the net charge of the system counter ions (Na+ and Cl^-) were added to a final concentration of 0.15 M. Subsequently, the system was subjected to energy minimization to remove spacial clashes (1000 steps of steepest descent) until no significant energy change could be detected.

The molecular mechanic's optimization was followed by four successive 100-ps position-restrained dynamics runs at 300 K with a gradual decrease in harmonic positional restraint force constant of1000 kJ mol-1 nm-2 on the solute, which enables the water to pack around the protein. During equilibration, the system was coupled to an external bath by the Berendsen(Berendsen *et al.,* 1984) temperature and pressure method respectively. Complete mixing and equilibration of the contents were confirmed by the convergence of the potential energy and temperature of the system. Finally, the equilibrated system was subjected to 65 ns production MD run without position restraints on the protein. Long-range electrostatic interactions beyond the cutoff were treated with the **P**artical-**M**esh **E**wald (PME) algorithm (Darden *et al.,* 1999) with an interpolation order of 4, Fourier grid spacing of 0.16 nmand Coulomb radius of 1.0 nm; and van der Waals was applied with a cut-off distance of 1.0 nm. The linear constraint (LINCS) (Hess *et al.,* 1997) algorithm with an order of 4 was used to constrain all bond lengths to their equilibrium positions, allowing an integration time step of 1 fs. V-rescale and Parrinello-rahman algorithms were utilized to control the temperature at 300 K with a coupling constant of 0.1 ps and pressure at 1 atm with a coupling constant of 0.5 ps, respectively (Berendsen *et al.*, 1984). The MD motion equations were integrated through the leap-frog algorithm using 2 fs time steps and the structural frames were saved every 2 ps.

4.2.12 Analysis of trajectory files

The trajectory files generated by MD simulations were analyzed using the GROMACS basic utilities g_rmsd, g_rmsf, g_hbond, g_sasand g_gyrate to obtain the root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), hydrogen bonding (HB), solvent-accessible surface area (SASA) and radius of gyration (Rg), respectively. The RMSD, RMSF, hydrogen bonding, solvent-accessible surface area (SASA)and radius of gyration (Rg) analysis were plotted for all 8 simulations using the **G**raphing, **A**dvanced **C**omputationand **E**xploration (GRACE) program.

4.2.13 Sequence comparison and phylogenetic analysis

Protein sequences of assigned and unassigned AC and NAAA were retrieved from Merops database (Rawlings *et al.*, 2014). A dataset of 715 protein sequences of AC and NAAA was obtained and checked in UniProtKB/Swiss-Prot (http://www.uniprot.org/) and gene bank. The partial sequences, unassigned, low-quality prediction, isoform and without active site residue sequence removed from the analysis. Protein sequences were initially aligned using **Mu**ltiple **S**equence **C**omparison by **L**og-**E**xpectation (MUSCLE) (Edgar, 2004) employing the **U**nweighted **P**air **G**roup **M**ethod with **A**rithmetic Mean (UPGMA) based clustering method with gap open and gap extension penalties set to 22.9 and 0. This alignment was used to build the Phylogenetic tree using the **N**eighbor-**J**oining (NJ) algorithm of **M**olecular **E**volutionary **G**enetics **A**nalysis 7 (MEGA) (Kumar, Stecher and Tamura, 2016) with the Dayhoff substitution matrix (PAM 250) and bootstrap value set to 1000. A well-defined cluster for AC and NAAA was observed in the dendrogram comprising thirty protein sequences. Multiple sequence alignment (MSA) of these thirty AC and NAAA protein sequences was done to analyze the extent of conservation among the sequences using the Clustal X program. A percent identity matrix of AC and NAAA was prepared with the help of the Clustal X program. The Phylogenetic trees generated were visualized in iTOL (Letunic and Bork, 2016).

4.3 Results

4.3.1 Homology modeling and model validation of mature and immature AC

Due to the absence of the three-dimensional structure of zebrafish AC, mature 3D homology models were build using hAC structure as a template, and immature 3D homology models were built using Robetta server (Figure 4.1 and 4.2).

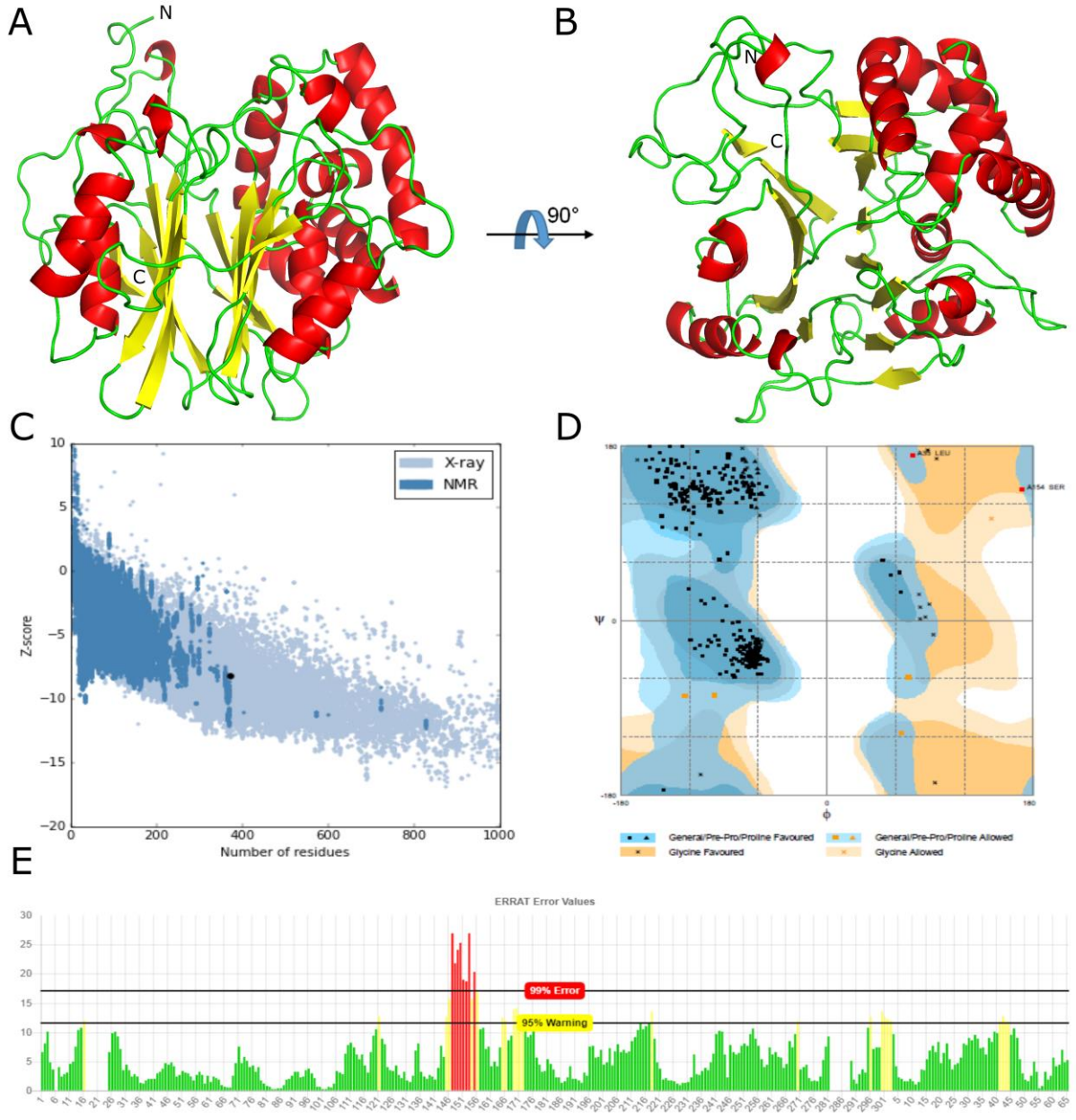


Figure 4.1: Homology model of immature zebrafish AC. A: Molecular model of immature zebrafish AC prepared using Robetta software. B: Ramachandran plot. C: ProSA analysis. D: ERRAT analysis.

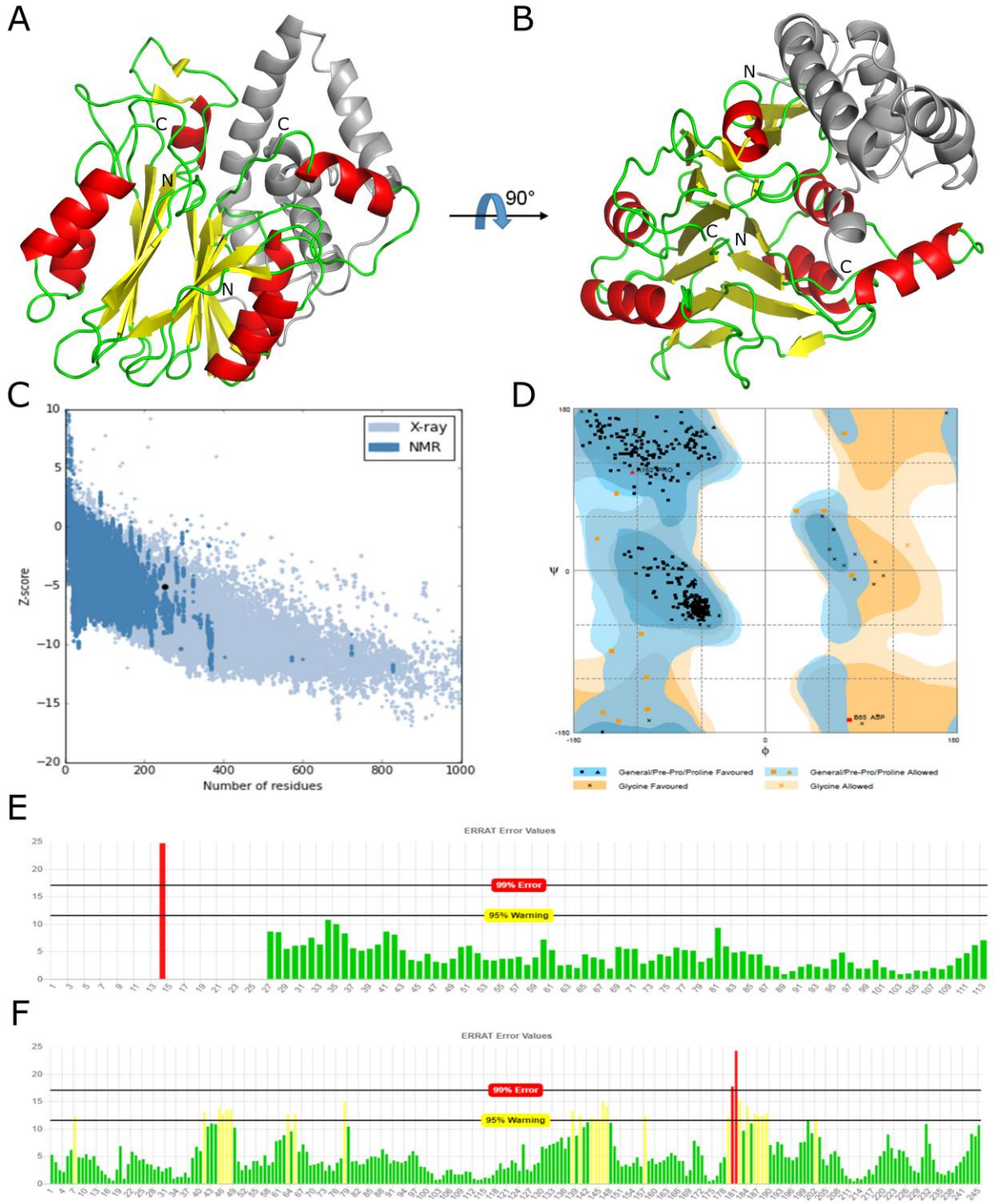


Figure 4.2: Homology model of mature zebrafish AC. A: Molecular model of mature zebrafish AC prepared using pdb structure 5u7z (C and D-chain) as a template (Grey color: α-subunit). B: Ramachandran plot. C: ProSA analysis. D: ERRAT analysis for chain A. E: ERRAT analysis for chain B.

Five molecular models were generated for immature zebrafish AC, best among which were chosen depending on Ramachandran plot score, ERRAT score for protein structure analysis (Table 4.1) (Figure 4.1). Model 1 with good Ramachandran plot and ERRAT

statistic score was further refined using Galaxy Web server (Ko *et al.*, 2012).

Table 4.1: Estimated values of evaluation parameters to assess the quality of immature zebrafish AC model.

Immaturezebrafish AC model	ERRAT Score	Ramachandran Plot number of residues		
		Core region %	Allowed region %	Disallowed region %
1	82.521	95.2	3.5	1.3
2	86.389	93.3	5.9	0.8
3	86.592	94.4	4.0	1.6
4	90.830	94.1	4.6	1.3
5	89.589	92.7	5.1	2.2

The signal sequence was not modeled. The α-subunit is primarily helical in nature, while β-subunit and immature AC comprise of loops and β-sheets, along with few helices. Sequence comparison has shown that α-subunit of template sequences were more than 46 % identical to hAC, where β-subunit was more than 67 % identical. (Table 4.2).

Table 4. 2: Templates used for modeling the subunits of the zebrafish AC

Zebrafish AC model	Template						
	PDB ID	Chain	Sequence identity (%)	Resolution (A°)	Positives (%)	Gaps (%)	Coverage (%)
α-subunit	5U7Z	C	46	2.5	71	6	28-140
β-subunit	5U7Z	D	67	2.5	85	0	143-395

The catalytic key residue of an enzyme, Cys 143, Asp 162 and Arg 333 are well conserved among all AC (Figure 4.3 A and B). The tertiary structure models built for zebrafish AC were observed to be of high quality validated by model validation programs (Table 4.3).

Table 4.3: Estimated values of evaluation parameters to assess the quality of zebrafish AC model.

Zebrafish AC model	ERRAT Score		Ramachandran Plot number of residues		Z Score
	α-subunit	β-subunit	Allowed region %	Disallowed region %	
Immature	90.8571		1.3	0.5	-8.2
Mature	88.9796	98.8636	3.5	0.5	-5.13

Ramachandran plot calculated using PROCHECK estimated that more than 99 % residues of each model of zebrafish AC present in the favoured and allowed regions (Figure 4.1 B and 4.2 B). ERRAT program evaluates the quality of a model by analyzing the statistics of non-bonded atom-atom interaction. All models were having an overall quality factor >88 %, close to the value for a high-quality structure (Figure 4.1 D and 4.2 D). The Z-score of each model is inside the satisfactory range calculated using ProSA and it is matching to similar length proteins Z-score (Figure 4.1 C and 4.2 C). Thus, the validation analysis confirmed the acceptable quality of all models. Also, each of the model structures superposed very well with the template structure hAC.

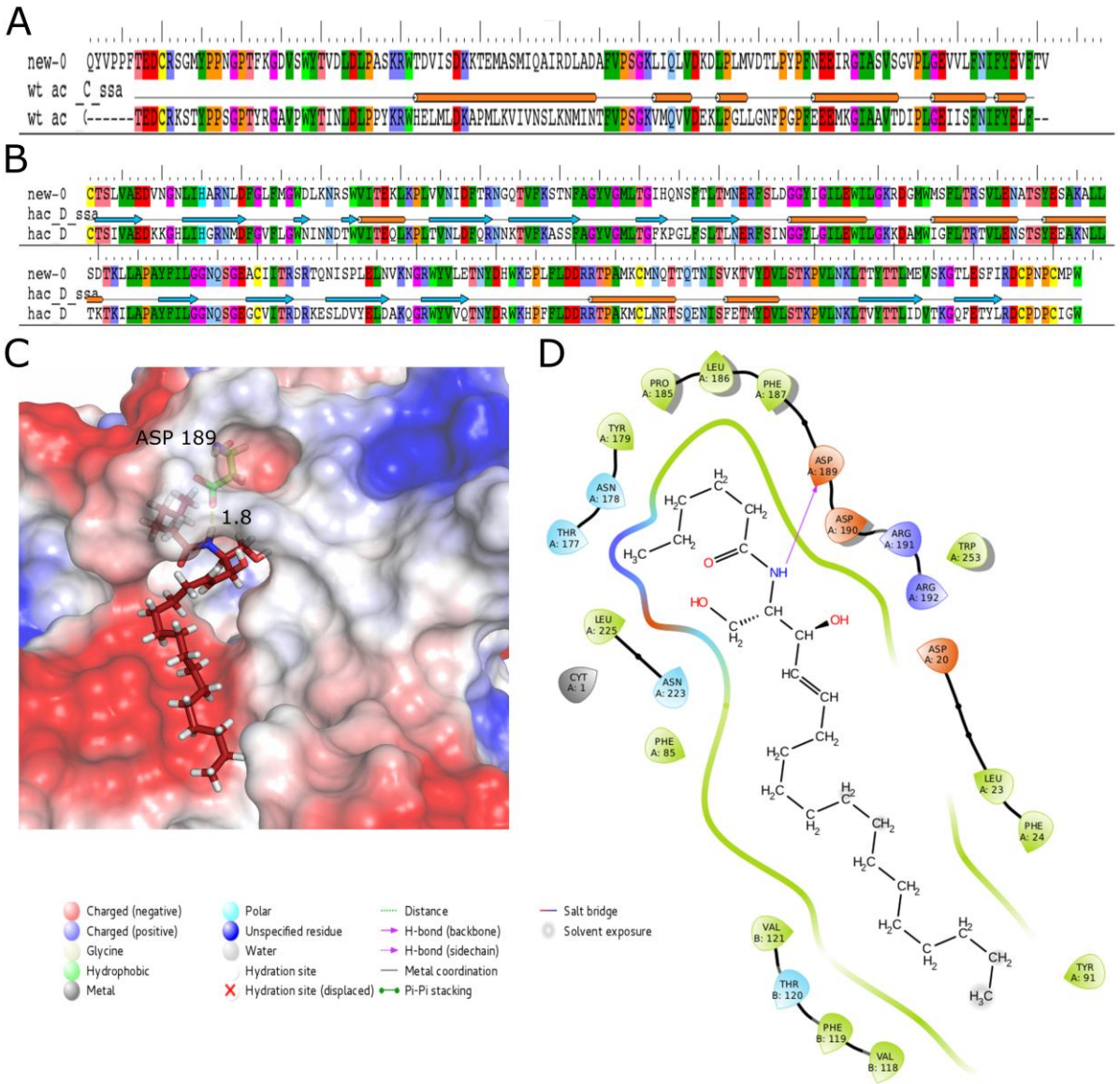


Figure 4.3: Docked ceramide displaying the interaction with zebrafish AC. A: Sequence comparison analysis for chain A. B: Sequence comparison analysis for chain B. C: Surface

view of the active site of zebrafish AC. D: Ligand interaction diagram showing the interaction of ceramide with the amino acid residues of zebrafish AC.

4.3.2 Modeled structures of mature and immature NAAA

Ten molecular models were generated for mature NAAA and immature NAAA, using corresponding templates (Table 4.4). The signal sequences were not modeled. The best among them were chosen depending on the score of PROCHECK, ERRATand ProSA-web protein structure analysis (Figure 4.4 and 4.5).

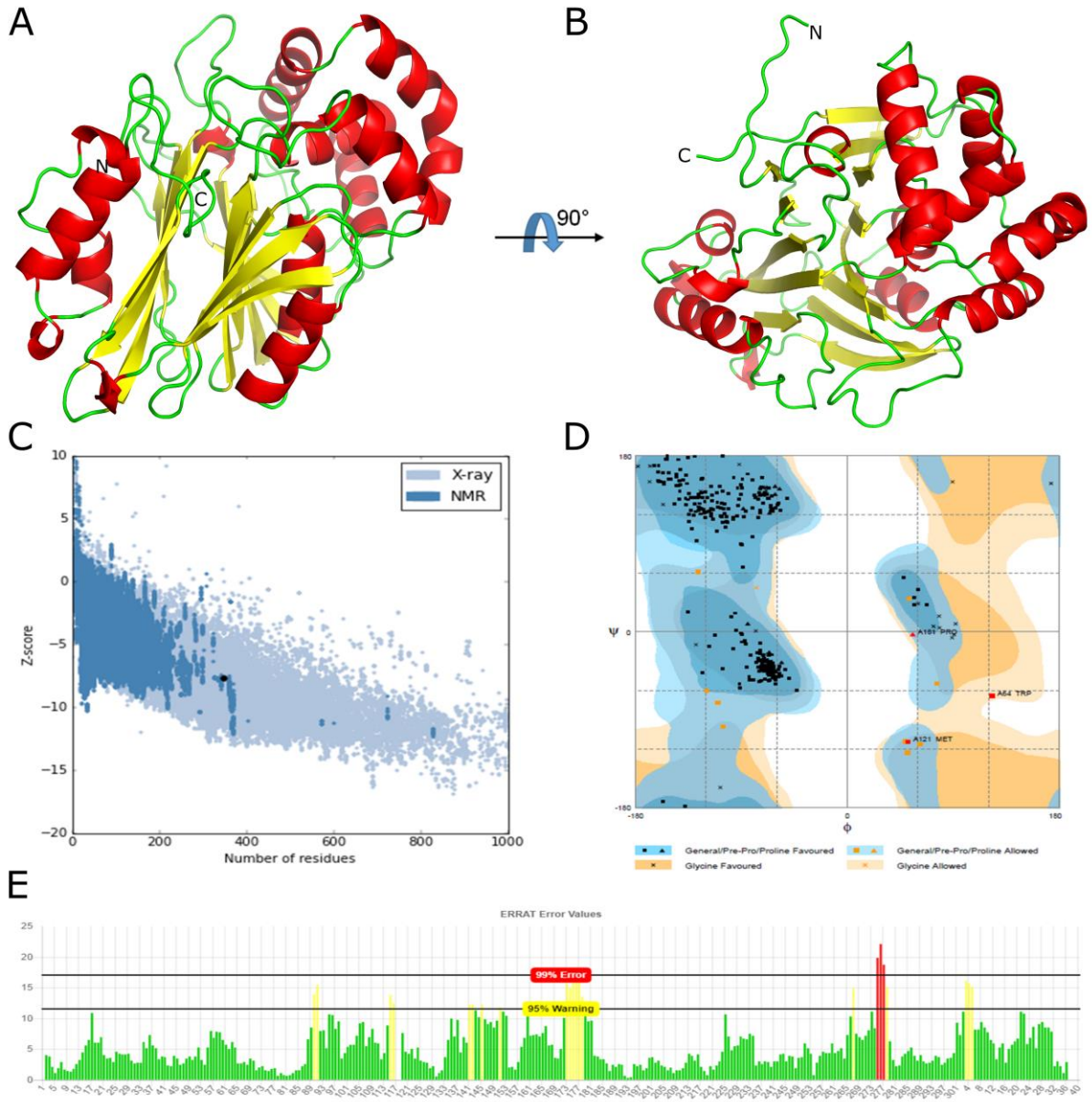


Figure 4.4: Homology model of immature *M. commoda* NAAA. A: Molecular model of immature *M. commoda* NAAA prepared using pdb structure 6dxw (A-chain) as a template. B: Ramachandran plot. C: ProSA analysis. D: ERRAT analysis

The models having the best validation parameters were chosen for further study (Table 4.5). The α-subunit is primarily helical structures, while β subunit and inactive NAAA comprise of loops and β-sheets, along with few helices.

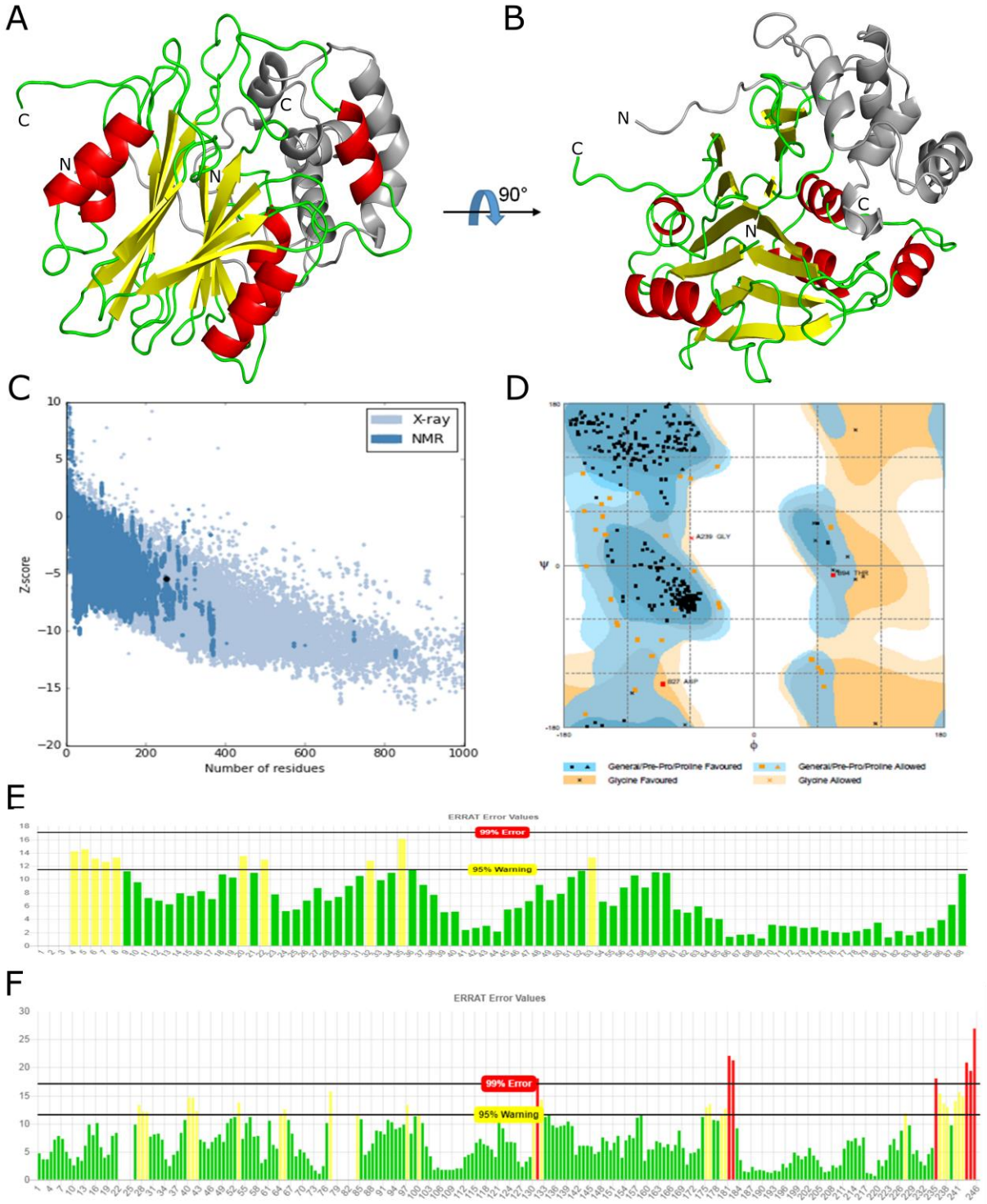


Figure 4.5: Homology model of mature *M. commoda* NAAA. A: Molecular model of mature *M. commoda* NAAA AC prepared using pdb structure 6DXY (A and B-chain) as a template (Grey color: α-subunit).B: Ramachandran plot. C: ProSA analysis. D: ERRAT analysis for chain A. E: ERRAT analysis for chain B.

Table 4.4: Templates used for modeling the subunits of the *M. commoda* NAAA.

Model *M. commoda* NAAA	Template						
	PDB ID	Chain	Sequence identity (%)	Resolution (A°)	Positives (%)	Gaps (%)	Coverage (%)
Immature	6DXW	A	34	2.3	50	7	22-395
α-subunit	6DXY	A	27	1.8	47	3	22-140
β-subunit	6DXY	B	37	1.8	53	9	143-395

Table 4.5: Estimated values of evaluation parameters to assess the quality of *M. commoda* NAAA model.

Model *M. commoda*NAAA	ERRAT Score		Ramachandran Plot number of residues		Z Score
	α-subunit	β-subunit	Allowed region %	Disallowed region %	
Immature	93.413		2.9	0.9	-7.65
Mature	88.2353	85.9574	9.0	0.9	-5.45

4.3.3 Orientation of ceramide in Zebrafish AC

C6 Ceramide (6:0) [N-hexanoyl-D-*erythro*-sphingosine] was docked into the homology modeled structure of zebrafish AC. The ligand interaction diagram for ceramide showing the interaction of the ligand bound to protein (Figure 4.3 D) depicts. The residue Cys 1 of β-subunits (equivalent to Cys 143 of immature AC) closes proximate with ligand and Asp 189 (equivalent to Asp331 of immature AC) form H bond with the amine group of ceramide. Another important active site residue Asp 20 (equivalent to Asp 162 of immature AC) also close proximate toward the substrate. The residues of Val l118, Phe 119, Thr 120and Val 121 and of α-subunit proximity with sphingosine. The electrostatic surface view of zebrafish AC shows the active site to be primarily made up of hydrophobic residues which suit the binding of its hydrophobic substrate (ceramide) (Figure 4.3 C).

4.3.4 Selection of SNP data set

The ASAH1 gene has three isoforms due to alternative splicing and to has a total of 9664 SNPs in the SNP database to date. The Exome Aggregation Consortium (ExAC) database reported a total of 740 SNPs for the asah1 gene (Transcript: ENST00000262097), out of which 190 coding nsSNPs, 56 coding synonymous, 5 frameshift, 10 stop gained, 3 start lost, 2 inframe deletion, 290 in the intronic region, 41 in the mRNA 5'UTR region, 21 in the mRNA 3' UTR region, 76 in upstream gene, 4 in splice donor region, 4 in splice acceptor region and 38 in splice region (Figure 4.6). The most of SNPs of ASAH1 located in the

coding region and also illustrious that the number of nsSNPs were greater than compared to other SNPs.

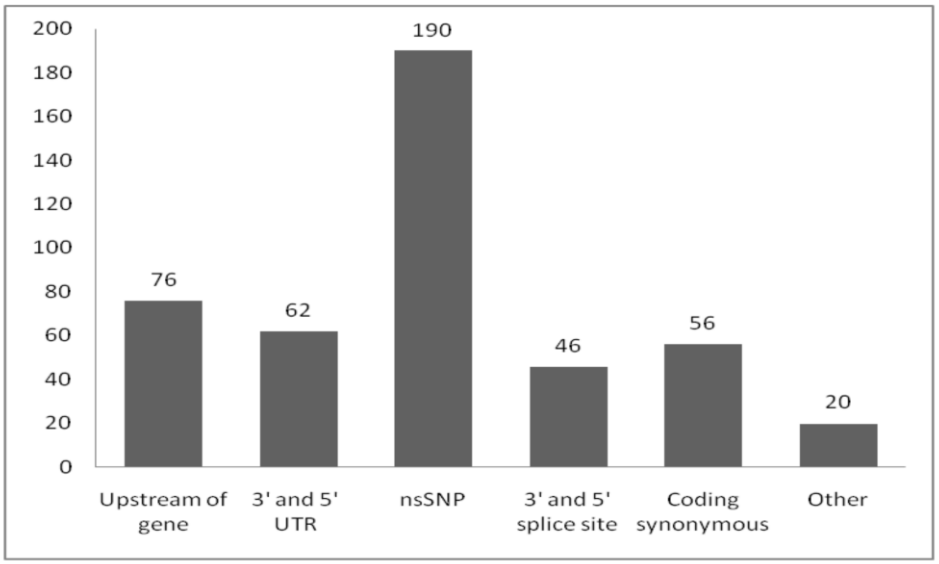


Figure 4.6: SNP distribution of asah1 gene.

Table 4.6: A list of selected Farber disease mutations reported in the α-subunit of hAC.

Farber disease mutation in α-subunit			
Sr. No.	**Mutation**	**Number of cases**	**Reference**
1	Q22H	1	(Zhang *et al.*, 2000)
2	H23D	1	(Zhang *et al.*, 2000)
3	Y36C	4	(Bär *et al.*, 2001)
4	V97E	1	(Muramatsu *et al.*, 2002)
5	V97G	2	(Chedrawi *et al.*, 2012)
6	F136L	1	(Bashyam *et al.*, 2014)
7	E138V	5	(Li *et al.*, 1999)

Table 4.7: A list of selected SMA-PME disease mutations reported in the α-subunit of hAC

SMA-PME disease mutation in α-subunit			
Sr. No.	**Mutation**	**Number of cases**	**Reference**
8	T42A	4	(Filosto *et al.*, 2016)
9	T42M	12	(Zhou *et al.*, 2012)

Table 4.8: A list of selected Farber disease mutations reported in the β-subunit of hAC.

Farber disease mutation in β-subunit			
Sr. no.	**Mutation**	**Number of cases**	**Reference**
10	G168W	1	(Cvitanovic-Sojat *et al.*, 2011)
11	W169R	7	(Bonafé *et al.*, 2016)
12	E180K	1	(Bashyam *et al.*, 2014)
13	L182V	4	(Devi *et al.*, 2006)
14	T222K	1	(Koch *et al.*, 1996)
15	R226P	1	(Bashyam *et al.*, 2014)
16	G235R	3	(Muramatsu *et al.*, 2002)
17	G235D	1	(Knorr, Rudolf and Nuernberger, 2013)
18	R254G	4	(Li *et al.*, 1999)
19	N320D	1	(Bär *et al.*, 2001)
20	N320S	1	(Bashyam *et al.*, 2014)
21	D331N	1	(Bär *et al.*, 2001)
22	R333C	3	(Kim *et al.*, 2016)
23	R333G	4	(Bashyam *et al.*, 2014)
24	R333H	1	(Bashyam *et al.*, 2014)
25	P362R	2	(Li *et al.*, 1999)
26	P362T	1	(Bashyam *et al.*, 2014)

Table 4.9: A list of selected SMA-PME disease mutations reported in the β-subunit of hAC.

SMA-PME disease mutation in β-subunit			
Sr. no.	**Mutation**	**Number of cases**	**Reference**
27	K152N	5	(Dyment *et al.*, 2014)
28	T179I	3	(Sathe and Pearson, 2014)

Table 4.10: A list of selected Farber disease mutations reported in the β-subunit of isoform 2.

Farber disease mutation in β-subunit of isoform 2		
Sr. no.	**Mutation**	**Reference**
29	W185R	(Al Jasmi, 2012)
30	K382Q	(Al Jasmi, 2012)

Since nsSNPs could alter function of hAC and are likely to be disease-causing. The asah1 gene investigated in this work has a total of 30 disease-causing nsSNPs reported, from which 26 caused FD (Table 4.6 and 4.7) and 4 caused SMA-PME disease (Table 4.8 and 4.9). Only disease-causing nsSNPs were chosen for further analysis. Two nsSNPs (W185R and K382Q) associated with FD are present in isoform 2 of AC protein (Table 4.10).

4.3.5 Prediction more deleterious SNPs

Five different tools (SIFT, PROVEAN, PolyPhen2, PhD-SNP and SNPs and GO) were used to predict and compare the consequence of an amino acid substitution on protein. The Y36C, V97E, E138V, G169R, L182V, T222K, R226P, G235R, G235D, R254G, N320S, D331N, R333G, R333H, P362Rand W185R are more damaging mutant compared to other mutants (Table 4.12). This 16 nsSNPs were identified as highly deleterious using five different tool scores (Figure 4.7 A). From this analysis, it is observed that FD mutants are more damaging compared to SMA-PME disease mutants.

4.3.6 Prediction of protein structural stability

The SNPs from published reports analyzed throw pathogenic prediction tools were also investigated to enzyme stability using tools such as, PoPMuSiC, FoldX and I-mutant.

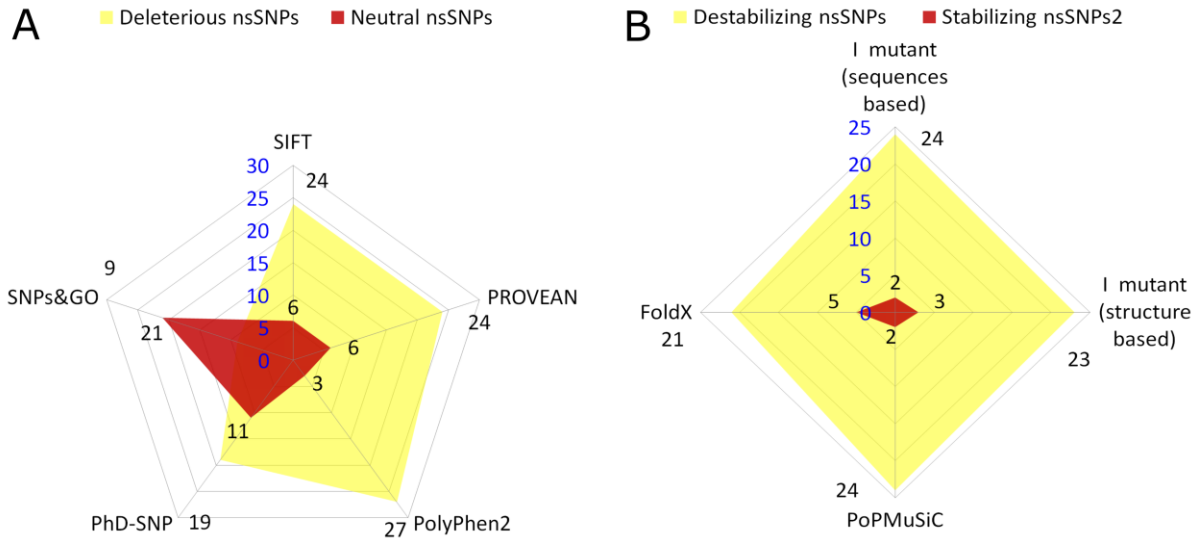


Figure 4.7: Prediction of the most damaging nsSNPs. A: Distribution of predicted deleterious nsSNPs. B: Distribution of predicted destabilizing nsSNPs.

The V97E, V97G, F136L, G168W, G169R, L182V, T222K, R226P, G235R, G235D, R254G, N320D, R333C, R333G, R333H, P362Rand P362T, are more unstable compared to other mutants (Table 4.13). These 17 mutations were identified as highly unstabilizing based on the prediction scores obtained using these tools (Figure 4.7 A).

4.3.7 Molecular phenotype analysis

SNPeffect4.0 was used to predict changes in structural stability (FoldX), chaperone binding (LIMBO) amyloidogenicity (WALTZ) and aggregation (TANGO). Ten nsSNPs were found to be associated with one or more of these changes (Figure 4.8) (Table 4.11). Notably, one nsSNPs, T179I, was found by WALTZ and TANGO to be associated with hAC enzyme to aggregate.

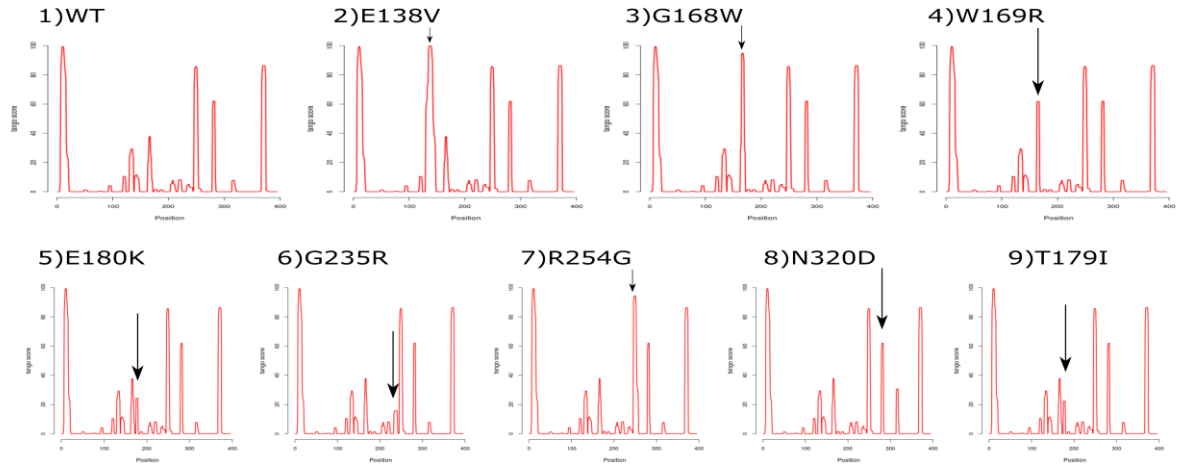


Figure 4.8: Per-residue TANGO scores of WT and mutants.

Table 4. 11: Molecular phenotyping of nsSNPs in AC by SNPeffector 4.0.

Sr. No.	mutant	SNP Effect		
		TANGO	WALTZ	LIMBO
1	F136L	Decreases	Decreases	Not Affect
2	E138V	Increases	Decreases	Not Affect
3	G168W	Increases	Decreases	Not Affect
4	W169R	Increases	Decreases	Not Affect
5	E180K	Increases	Not Affect	Not Affect
6	G235R	Increases	Not Affect	Increases
7	G235D	Not Affect	Not Affect	Increases
8	R254G	Increases	Not Affect	Not Affect
9	N320D	Increases	Not Affect	Not Affect
10	T179I	Increases	Increases	Not Affect

Table 4. 12: The results for prediction pathological effect of all 30nsSNPs by five *in silico* tools.

Sr. No.	Mutation	SIFT		PROVEAN		PolyPhen2				PhD-SNP		SNPs and GO	
						(Hum Div)		(Hum Var)					
		Prediction	Tolerance index	Score	Cutoff=-2.5	Effect	Score	Effect	Score	Prediction	Reliable Index	Prediction	Reliable Index
1	Q22H	Damaging	0	Neutral	-1.35	Pro-damg	0.99	Pos-damg	0.87	Neutral	6	Neutral	9
2	H23D	Tolerated	0.26	Neutral	-0.06	Benign	0	Benign	0	Neutral	5	Neutral	8
3	Y36C	Damaging	0	Deleterious	-7.16	Pro-damg	1	Pro-damg	0.99	Disease	6	Neutral	6
4	V97E	Damaging	0	Deleterious	-4.92	Pro-damg	0.99	Pro-damg	0.94	Disease	3	Neutral	7
5	V97G	Damaging	0	Deleterious	-5.74	Pro-damg	0.97	Pro-damg	0.89	Neutral	0	Neutral	8
6	F136L	Damaging	0	Deleterious	-4.97	Pro-damg	0.99	Pro-damg	0.95	Neutral	1	Neutral	7
7	E138V	Damaging	0	Deleterious	-6.16	Pro-damg	1	Pro-damg	0.99	Disease	6	Disease	0
8	G168W	Damaging	0	Deleterious	-7.23	Pro-damg	0.96	Pos-damg	0.86	Disease	5	Disease	1
9	W169R	Damaging	0	Deleterious	-12.93	Pro-damg	0.99	Pro-damg	0.98	Disease	9	Disease	4
10	E180K	Damaging	0.03	Deleterious	-3.17	Pos-damg	0.88	Pos-damg	0.76	Neutral	1	Neutral	7
11	L182V	Damaging	0	Deleterious	-3.00	Pro-damg	1	Pro-damg	1	Disease	3	Neutral	6
12	T222K	Damaging	0	Deleterious	-4.67	Pro-damg	0.99	Pro-damg	0.96	Disease	1	Neutral	7
13	R226P	Damaging	0	Deleterious	-6.58	Pro-damg	1	Pro-damg	0.99	Disease	8	Disease	3
14	G235R	Damaging	0	Deleterious	-7.31	Pro-damg	1	Pro-damg	0.98	Disease	6	Neutral	4
15	G235D	Damaging	0	Deleterious	-6.36	Pro-damg	1	Pro-damg	0.99	Disease	3	Neutral	4
16	R254G	Damaging	0	Deleterious	-6.63	Pro-damg	1	Pro-damg	0.99	Disease	8	Disease	1
17	N320D	Damaging	0.01	Deleterious	-4.48	Pos-damg	0.74	Pos-damg	0.65	Disease	6	Disease	3
18	N320S	Damaging	0	Deleterious	-4.54	Pro-damg	0.99	Pro-damg	0.97	Disease	6	Disease	1
19	D331N	Damaging	0	Deleterious	-4.55	Pro-damg	1	Pro-damg	0.99	Disease	5	Neutral	5
20	R333G	Damaging	0	Deleterious	-6.24	Pro-damg	1	Pro-damg	1	Disease	4	Neutral	1
21	R333C	Damaging	0	Deleterious	-7.263	Pro-damg	1	Pro-damg	1	Neutral	1	Disease	7
22	R333H	Damaging	0	Deleterious	-4.55	Pro-damg	1	Pro-damg	1	Disease	5	Neutral	4
23	P362R	Damaging	0	Deleterious	-7.14	Pro-damg	1	Pro-damg	0.99	Disease	6	Neutral	3
24	P362T	Tolerated	0.15	Deleterious	-6.08	Pro-damg	0.99	Pro-damg	0.97	Disease	2	Neutral	7
25	T42A	Tolerated	0.33	Neutral	-2.14	Benign	0.19	Benign	0.93	Neutral	5	Neutral	8
26	T42M	Tolerated	0.06	Neutral	-2.37	Pro-damg	0.99	Pos-damg	0.83	Neutral	3	Neutral	7
27	K152N	Tolerated	0.42	Neutral	-0.21	Benign	0	Benign	0	Neutral	8	Neutral	9
28	T179I	Damaging	0	Deleterious	-5.75	Pro-damg	0.99	Pro-damg	0.99	Neutral	2	Neutral	7
29	W185R	Damaging	0	Deleterious	-13.07	Pro-damg	0.99	Pro-damg	0.97	Disease	9	Disease	4
30	K382Q	Tolerated	0.07	Neutral	-2.31	Pro-damg	0.98	Pos-damg	0.82	Neutral	4	Neutral	8

Table 4.13: The results for prediction stability of all 26 nsSNPs by four *in silico* tools.

Sr. No.	Mutant	I mutant (sequences based)			I mutant (structure based)				PoPMuSiC (structure based)				FoldX	
		Stability	DDG	Reliable Index	Stability	DDG	Reliable Index	Relative Surface Accessibility	Stability	Secondary structure	Solvent accessibility (%)	ΔΔG	Stability	ddG kcal/mol
1	Y36C	Increase	0.16	3	Decrease	-0.13	1	21.1	Destabilizing	T	19.95	2.21	Reduces	1.21
2	V97E	Decrease	-1.70	7	Decrease	-1.65	9	19.5	Destabilizing	H	17.87	1.79	Reduces	1.29
3	V97G	Decrease	-4.19	10	Decrease	-3.57	10	19.5	Destabilizing	H	17.87	2.64	Reduces	2.21
4	F136L	Decrease	-2.62	7	Decrease	-2.95	4	22.9	Destabilizing	H	21.87	1.46	Slightly reduces	0.96
5	E138V	Increase	0.45	1	Increase	-0.40	4	0.6	Stabilizing	H	0.66	-0.28	Slightly reduces	1.00
6	G168W	Decrease	-1.73	8	Decrease	0.30	2	30.5	Destabilizing	C	27.32	1.77	Reduces	3.89
7	W169R	Decrease	-1.83	8	Decrease	-1.02	6	37	Destabilizing	E	32.20	1.25	Reduces	2.30
8	E180K	Decrease	-0.78	7	Decrease	-0.97	7	18.9	Destabilizing	H	17.21	0.62	Slightly enhances	-0.62
9	L182V	Decrease	-1.21	8	Decrease	-0.18	7	2.7	Destabilizing	H	2.45	2.05	Reduces	3.45
10	T222K	Decrease	-1.26	8	Decrease	-3.01	6	0.0	Destabilizing	E	0.58	2.90	Severely reduces	5.70
11	R226P	Decrease	-0.80	1	Decrease	-3.16	8	7.4	Destabilizing	C	7.22	1.97	Severely reduces	9.59
12	G235R	Decrease	-1.60	7	Decrease	-2.13	4	1.3	Destabilizing	H	0.90	0.17	Severely reduces	10.62
13	G235D	Decrease	-1.03	6	Decrease	-3.21	8	1.3	Destabilizing	H	0.90	1.32	Severely reduces	11.93
14	R254G	Decrease	-1.77	7	Decrease	-2.74	8	0.4	Destabilizing	H	0.18	2.19	Slightly reduces	0.87
15	N320D	Decrease	-1.36	0	Decrease	-2.06	8	13.0	Destabilizing	S	13.69	1.01	Reduces	1.19
16	N320S	Decrease	-1.59	4	Decrease	-1.97	8	13	Destabilizing	S	13.69	1.21	Reduces	1.64
17	D331N	Decrease	-1.40	5	Decrease	-0.96	7	11.1	Destabilizing	C	9.68	0.66	Slightly enhances	-0.89
18	R333G	Decrease	-0.73	8	Decrease	-2.77	9	6.6	Destabilizing	S	6.27	2.57	Reduces	1.89
19	R333C	Decrease	-1.77	6	Decrease	-1.64	9	6.6	Destabilizing	S	6.27	1.30	Reduces	2.22
20	R333H	Decrease	-1.42	9	Decrease	-1.52	9	6.6	Destabilizing	S	6.27	0.53	Reduces	4.29
21	P362R	Decrease	-1.16	7	Decrease	-2.61	9	3.5	Destabilizing	T	3.42	2.25	Reduces	3.23
22	P362T	Decrease	-1.13	6	Decrease	-1.94	8	3.5	Destabilizing	T	3.42	1.73	Reduces	4.44
23	T42A	Decrease	-0.62	6	Decrease	-0.65	8	13.7	Destabilizing	T	18.20	1.47	No effect	-0.13
24	T42M	Decrease	-0.81	5	Decrease	-0.11	3	13.7	Destabilizing	T	18.20	0.85	Slightly enhances	-0.51
25	K152N	Decrease	-0.93	5	Increase	-0.14	7	93.3	Stabilizing	T	85.68	-0.68	Enhances	-1.02
26	T179I	Decrease	-2.07	9	Increase	1.48	3	0	Destabilizing	H	0	0.41	Severely reduces	5.57

4.3.8 AC protein sequence conservation analysis

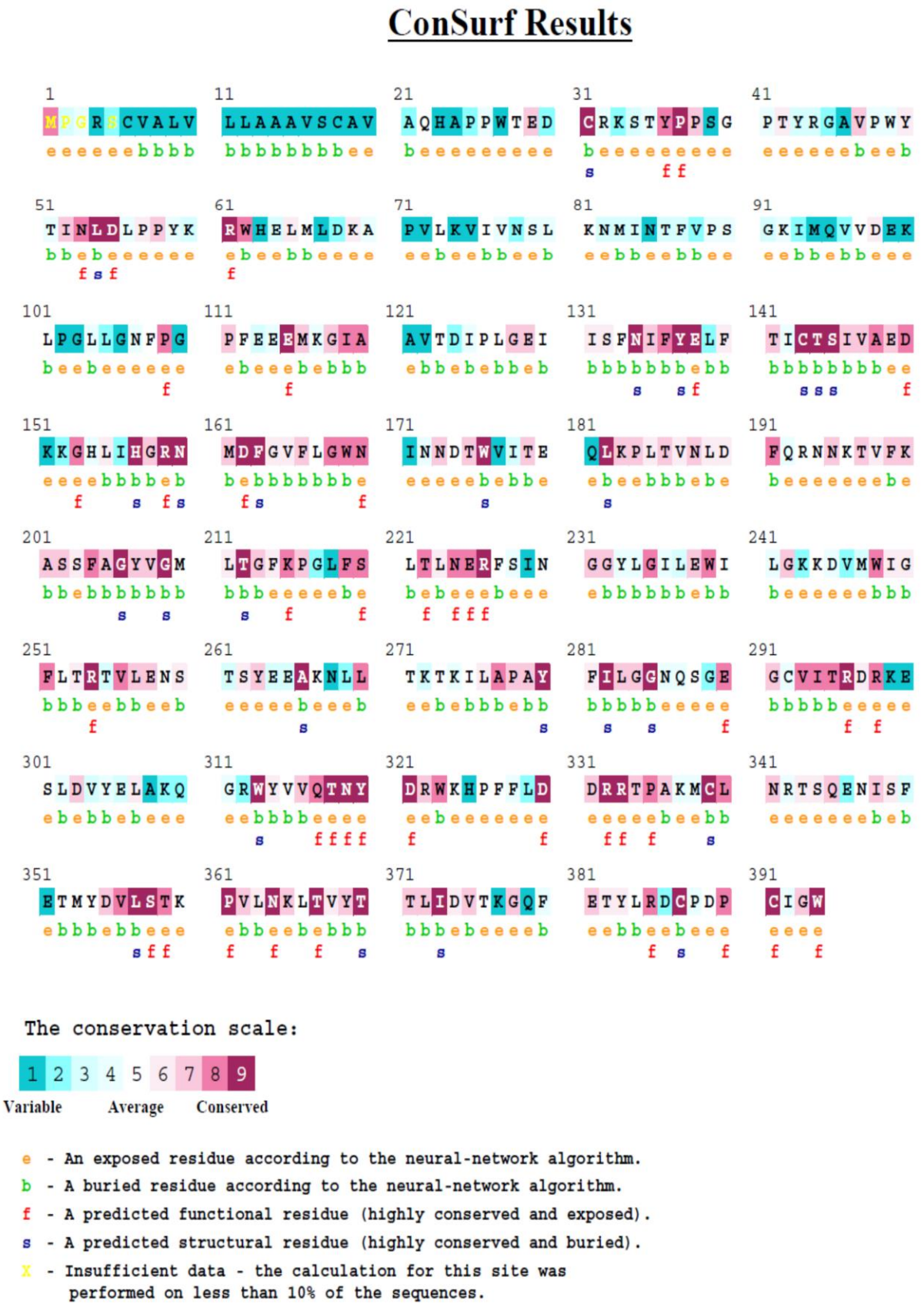


Figure 4.9: Conservation analysis of the AC protein sequence using ConSurf.

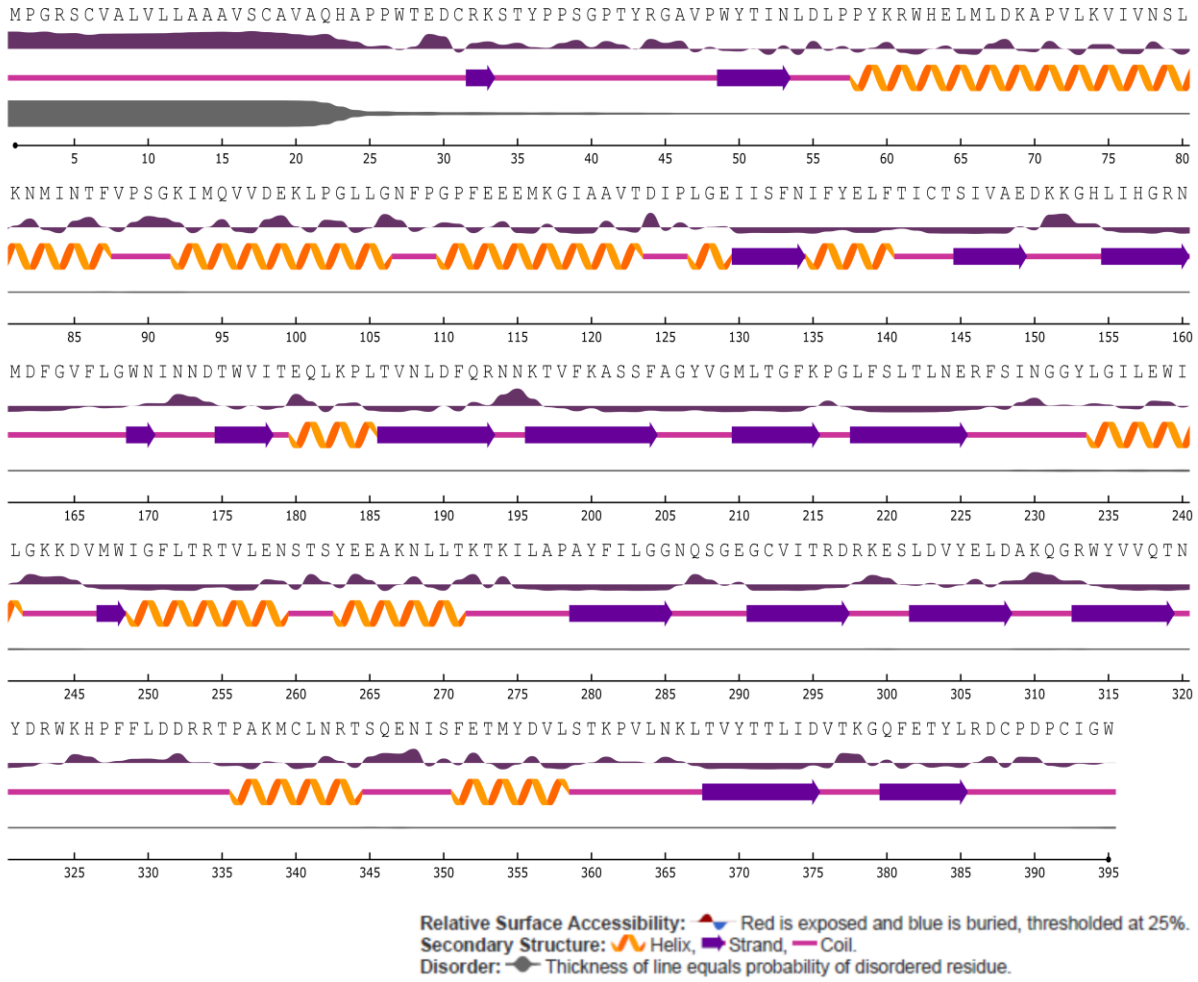


Figure 4.10 : Conservation analysis of the AC protein sequence using NetSurfP 2.

A comparative analysis of amino acid conservation between species based on protein sequence alignment provides an understanding of the importance of specific amino acid residues within a protein and reveals localized evolution. Importantly, the information from the ConSurf using the Bayesian analyzer indicates that most of the amino acids in the AC protein are highly conserved (Figure 4.9). In general, the substitution of conserved residues is deleterious. Consistent with this generalization, the majority of the substituted amino acids in AC were predicted to be harmful in nature by all of the computational prediction methods.

The Y36C, F136L, F138V, L182V, R226P, R254G, N320D, D331N, R333C, R333G and R333H are highly deleterious based on their ConSurf conservation score. NetSurf P was used for predictions about solvent accessibility (exposed, buried and partially buried) of the amino acids (Figure 4.10).

4.3.9 Predicted functional and structural modifications and post-translational modifications

The functional and structural alterations of disease-associated mutant investigated using MutPred server to predict a gain of disulfide linkage, altered transmembrane and secondary structure, and loss of disulfide, glycosylation, catalytic site (Table 4.15). PTMs are essential in regulating functions and structures of enzymes; therefore play a role in many biological processes; for example, methylation regulates the gene expression, phosphorylation regulates cell signaling and ubiquitylation control degradation mechanism for proteins. The PSSMe tool predicted a total of 3 residues (Lye 152, Arg 226 and Arg 333) that can get methylated. NetPhos 3.1 predicted a total of 3 residues (Tyr 36, Thr 42 and Thr 222) having the potential of getting phosphorylated. BDM-PUB predicted that K152 lysine residues undergo ubiquitylation (Table 4.14).

Table 4. 14: Putative methylation, phosphorylation and ubiquitylation sites predicted by PSSMe, NetPhos 3.1 and BDM-PUB, respectively.

Sr. No.	Position	Flanking residues	Score	Kinase	PTM Prediction
1	Y36	RKST-Y-PPSG	0.748		Putative phosphorylation sites predicted by NetPhos 3.1
2	T42	CRKS-T-YPPS	0.749	PKC	
			0.628	PKG	
3	T222	LFSL-T-LNER	0.523	CKII	
4	K152	CTSIVAEDK-K-GHLIHGRNM	0.83474		Putative methylation sites predicted by PSSMe
5	R226	GLFSLTLNE-R-FSINGGYLG	0.61		
6	R333	WKHPFFLDD-R-RTPAKMCLN	0.52378		
7	K152	SIVAEDKKGHLIHGR	0.41		Putative ubiquitylation sites predicted by BDM-PUB

Table 4.15: Effects of nsSNPs on structural and functional properties of *AC* by MutPred server.

Sr. No.	Mutation	Probability of deleterious mutation	Features
1	Y36C	0.822	Gain of Disulfide linkage at C31 (P = 5.1e-03); Altered Transmembrane protein (P = 4.0e-03); Loss of O-linked glycosylation at S39 (P = 0.02)
2	T42A	0.604	Altered Ordered interface (P = 0.02); Altered Transmembrane protein (P = 7.2e-04); Gain of ADP-ribosylation at R44 (P = 0.03); Loss of O-linked glycosylation at S39 (P = 0.03)
3	T42M	0.634	Altered Transmembrane protein (P = 3.5e-04); Gain of ADP-ribosylation at R44 (P = 0.03); Loss of O-linked glycosylation at S39 (P = 0.02)
4	V97E	0.795	Altered Transmembrane protein (P = 1.5e-05); Loss of Helix (P = 0.02)
5	V97G	0.695	Altered Transmembrane protein (P = 1.2e-04); Loss of Helix (P = 0.05); Altered Stability (P = 0.02)
6	E138V	0.887	Altered Metal binding (P = 7.1e-03); Loss of Helix (P = 0.05); Loss of Disulfide linkage at C143 (P = 0.03); Loss of Catalytic site at C143 (P = 0.04)
7	G168W	0.796	Loss of Loop (P = 0.04); Altered Transmembrane protein (P = 1.6e-03); Loss of N-linked glycosylation at N173 (P = 0.01)
8	W169R	0.796	Loss of Loop (P = 0.05); Altered Transmembrane protein (P = 3.0e-03); Loss of N-linked glycosylation at N173 (P = 0.01)
9	T179I	0.54	Altered Transmembrane protein (P = 1.9e-04)
10	T222K	0.706	Altered Transmembrane protein (P = 3.5e-04); Altered Stability (P = 0.03)
11	R226P	0.92	Altered Transmembrane protein (P = 4.5e-04); Loss of ADP-ribosylation at R226 (P = 0.04); Altered Stability (P = 0.01); Gain of GPI-anchor amidation at N230 (P = 0.03)
12	G235D	0.79	Altered Transmembrane protein (P = 1.9e-04); Loss of Helix (P = 0.03); Loss of GPI-anchor amidation at N230 (P = 0.03)
13	G235R	0.8	Altered Transmembrane protein (P = 1.6e-03); Loss of GPI-anchor amidation at N230 (P = 0.03)
14	R254G	0.895	Loss of Helix (P = 0.03); Gain of Strand (P = 0.05); Altered Transmembrane protein (P = 2.3e-03); Gain of N-linked glycosylation at N259 (P = 0.01); Gain of GPI-anchor amidation at N259 (P = 0.01)
15	N320D	0.878	Altered Transmembrane protein (P = 0.02); Gain of Sulfation at Y321 (P = 4.9e-03)
16	N320S	0.801	Altered Transmembrane protein (P = 0.02); Loss of Sulfation at Y321 (P = 0.03)
17	D331N	0.558	Altered Transmembrane protein (P = 0.01); Loss of Proteolytic cleavage at D331 (P = 0.02)
18	R333C	0.768	Altered Metal binding (P = 0.03); Altered Transmembrane protein (P = 4.7e-03); Loss of Proteolytic cleavage at D331 (P = 0.02)
19	R333G	0.82	Altered Transmembrane protein (P = 4.8e-03); Gain of Proteolytic cleavage at D331 (P = 0.02); Altered Stability (P = 0.05)
20	R333H	0.652	Altered Metal binding (P = 4.6e-03); Altered Transmembrane protein (P = 5.1e-03); Loss of Proteolytic cleavage at D331 (P = 0.02)
21	P362R	0.855	Altered Transmembrane protein (P = 2.4e-05); Altered Ordered interface (P = 4.8e-03); Gain of Helix (P = 0.03); Loss of Loop (P = 0.05); Altered Metal binding (P = 0.02)
22	P362T	0.77	Altered Transmembrane protein (P = 1.1e-04); Altered Ordered interface (P = 0.02); Altered Metal binding (P = 6.4e-03)

4.3.10 Structure of AC mutants

In silico 26 mutant structures were modeled by Fold X using 5U7Z as a template. The other four disease-causing nsSNPs structure was not modeled due to unavailability of suitable template structure in the PDB database. Q22H and H23D mutants were not modeled because the residues 22-27 are missing in 5U7Z. Other mutants W185R and K382Q were also not modeled because the structure of isoform 2 of hAC is not available. Thus nsSNPs Q22H, H23D, W185R and K382Q were excluded from further structure-based analysis study.

The comparative investigation of disease mutant model structures with wild type hAC crystal structure was achieved throw using Fold X, project HOPE and Pymol viewer. This mutant model and hAC structures differed with respect to size, polarity, rigidity, hydrophilicity /hydrophobicity, solvent accessibility. These different properties might disrupt the structural and functional of the hAC enzyme.

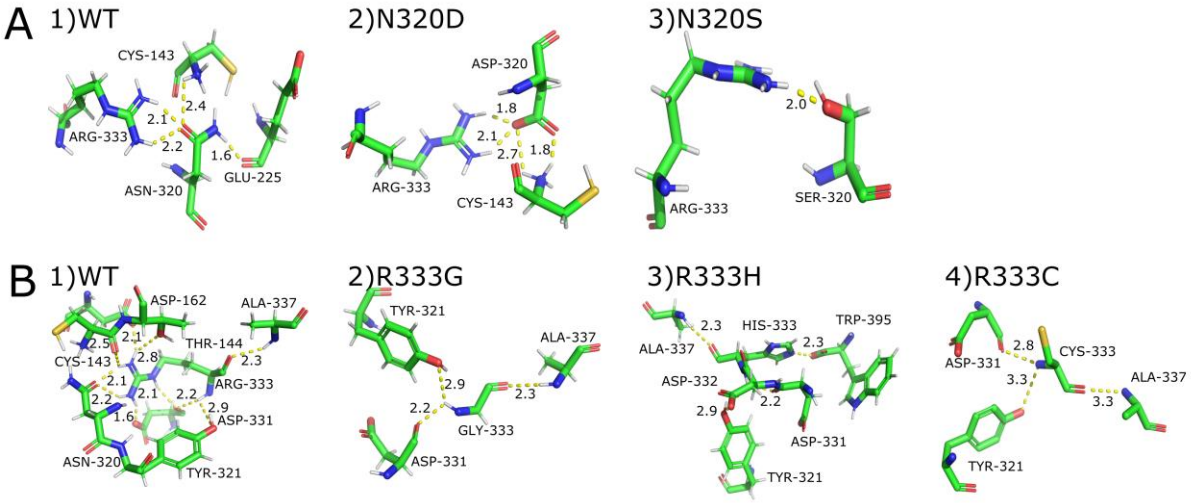


Figure 4.11: Mutations predicted to affect the activation of the enzyme.

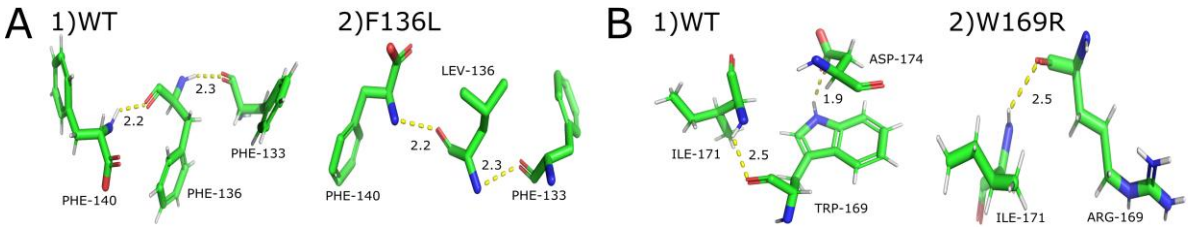


Figure 4.12: Mutations predicted to affecting the hydrophobic surface of the protein.

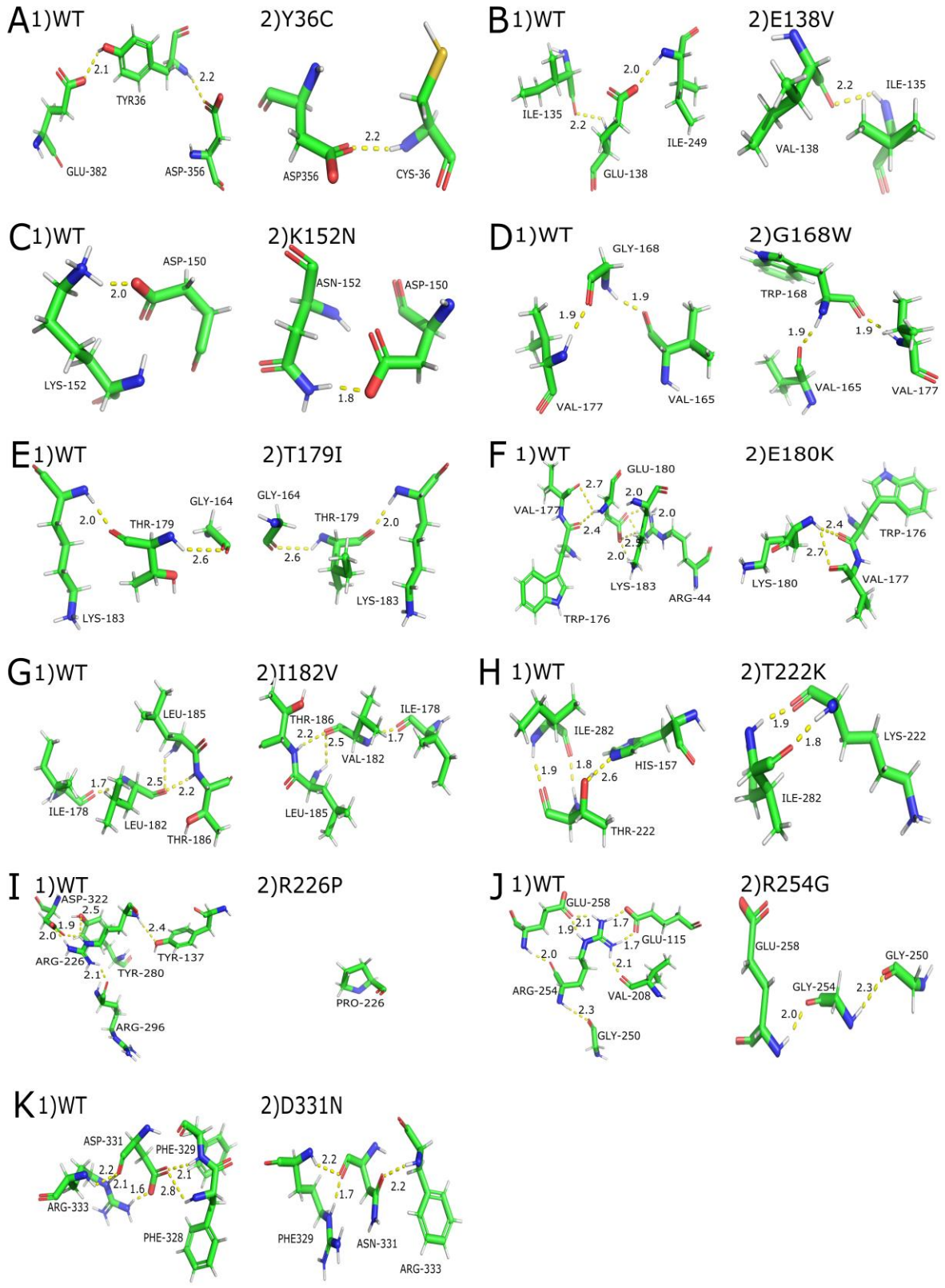


Figure 4.13: Mutations predicted to affecting the folding or stability of the protein.

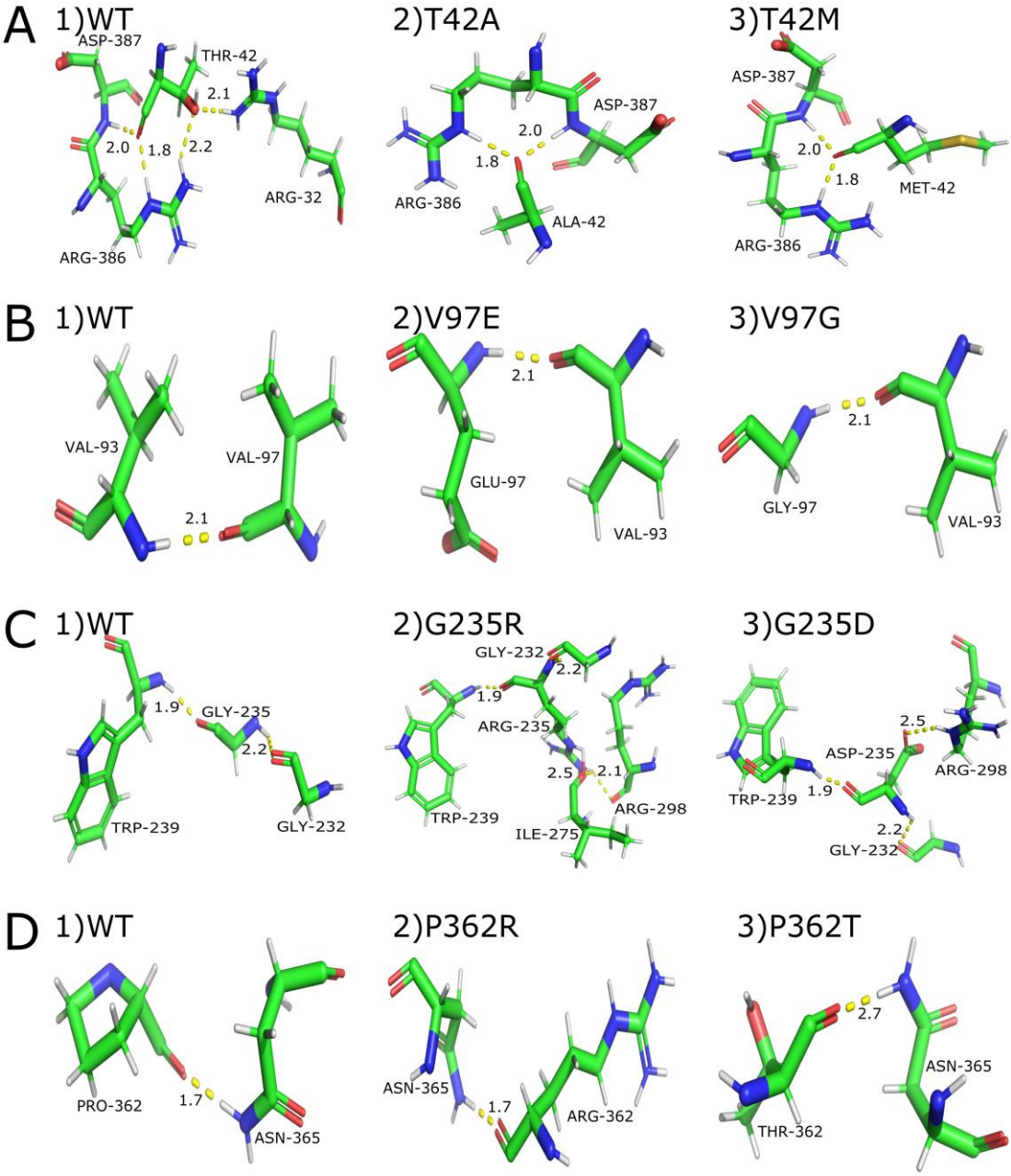


Figure 4.14: Mutations predicted to affecting the folding or stability of the protein.

4.3.10.1 Y36C

The mutant residue being smaller than Tyr, might lead to a loss of interactions. Change in hydrophobicity due to mutation result in the loss of hydrogen bond with Glu 382 of β-subunit, and disturb correct folding (Figure 4.13 A).

4.3.10.2 V97E

Val has neutral charge while Glu charge is negative. The mutation is located in the α-helix region of α-subunit of hAC. The mutant residue Glu 97 is preferable in the α-helix

region compared to Val. Mutation can lead repulsion of other residues and ligands due to change in size. The mutant residue being bigger, might lead to bumps. The hydrophobic interactions in the core of the protein will be lost (Figure 4.10).

Table 4.16: Details of polarity and hydrophobicity of reported desease mutant.

Sr. No.	Mutation	Change in size due to mutation	Change in polarity due to mutation	Change in Hydrophobicity due to mutation
1	Y36C	Big to Small	Neutral to Nonpolar	More hydrophobic to Less hydrophobic
2	V97E	Small to Big	Neutral Non polar to Acidic Polar	More hydrophobic to Less hydrophobic
3	V97G	Big to Small	Neutral non polar to Neutral Nonpolar	More hydrophobic to Less hydrophobic
4	F136L	Big to Small	Neutral to Nonpolar	Less hydrophobic to More hydrophobic
5	E138V	Big to Small	Acidic Polar to Neutral non polar	Less hydrophobic to More hydrophobic
6	G168W	Small to Big	Neutral Nonpolar to Nonpolar	Less hydrophobic to More hydrophobic
7	W169R	Big to Small	Non Polar to Basic Polar	More hydrophobic to Less hydrophobic
8	E180K	Small to Big	Acidic Polar to Basic Polar	More hydrophobic to Less hydrophobic
9	L182V	Big to Small	Nonpolar to Neutral non polar	Less hydrophobic to More hydrophobic
10	T222K	Small to Big	Polar to Basic Polar	More hydrophobic to Less hydrophobic
11	R226P	Big to Small	Basic Polar to Neutral Nonpolar	Less hydrophobic to More hydrophobic
12	G235R	Small to Big	Neutral Nonpolar to Basic Polar	More hydrophobic to Less hydrophobic
13	G235D	Small to Big	Neutral Nonpolar to Acidic Polar	More hydrophobic to Less hydrophobic
14	R254G	Big to Small	Basic Polar to Neutral Nonpolar	Less hydrophobic to More hydrophobic
15	N320D	Big to Small	Neutral to Acidic Polar	Less hydrophobic to More hydrophobic
16	N320S	Big to Small	Neutral to Neutral Polar	Less hydrophobic to More hydrophobic
17	D331N	Small to Big	Acidic Polar to Neutral	More hydrophobic to Less hydrophobic
18	R333G	Big to Small	Basic Polar to Nonpolar	Less hydrophobic to More hydrophobic
19	R333C	Big to Small	Basic Polar to Nonpolar	Less hydrophobic to More hydrophobic
20	R333H	Big to Small	Basic Polar to Basic Polar	Less hydrophobic to More hydrophobic
21	P362R	Small to Big	Nonpolar to Polar	More hydrophobic to Less hydrophobic
22	P362T	Small to Big	Polar to Polar	More hydrophobic to Less hydrophobic
23	T42A	Big to Small	Polar to Nonpolar	Less hydrophobic to More hydrophobic
24	T42M	Small to Big	Polar to Non polar	Less hydrophobic to More hydrophobic
25	K152N	Big to Small	Basic polar to Neutral	Less hydrophobic to More hydrophobic
26	T179I	Big to Small	Polar to Neutral non-polar	Less hydrophobic to More hydrophobic

4.3.10.3 V97G

The mutant residue Gly is smaller; hence there might be a loss of intra-protein interactions. The flexibility of Glycine might disturb the required rigidity of the enzyme at 97 position (Figure 4.14 B).

4.3.10.4 F136L

Phe 136 residue interacts with Phe 140 and Phe 133, and mutation did not affect this interaction (Figure 4.12 A).

4.3.10.5 E138V

The mutant residue Val 138 is smaller and more hydrophobic than Glu. The mutant residue being smaller, might lead to loss of interactions. The charge of Glu 138 will be lost due to mutation and this might loss of interactions with residues and other molecules. The

change in hydrophobicity due to mutation results in loss of hydrogen bond with Ile 249 of β-subunit and might disturb correct folding and hetrodimeric interaction (Table 4.16) (Figure 4.13 B). The Glu is to be located in an α-helix while the mutated residue does not prefer α-helices as secondary structure, these change might affect the conformation of α-helix.

4.3.10.6 G168W

The mutant residue Trp 168 is bigger than Gly. The flexibility of Glycine might be essential for the enzyme function. The residue Gly 168 occurs on the surface of the hAC and mutation might disturb interactions with other parts of the enzyme and other molecules (Figure 4.13 D). The flexibility of glycine is plenty to have unique torsion angles. The local structure will disturb due to the incorrect conformation of local backbone by mutation.

4.3.10.7 W169R

The Trp charge was neutral, while the mutant residue charge is positive. In the hAC 3D-structure, Trp 169 is sited in β-sheet of β-subunit. The β-sheet will be somewhat disrupted due to Arg prefers another secondary structure. The interaction of this Arg 169 residue with Ile 171 and Asp 174 is possible and interaction with Asp 174 affected by this mutation due to change in hydrophobicity and polarity (Table 4.16) (Figure 4.12 B). The change in polarity might cause repulsion between the neighboring residues and mutant residue. Hence protein surface will loss hydrophobic interactions.

4.3.10.8 E180K

Lys 180 residue do not be in the right location to make the same hydrogen bond as the Glu180 did. The Lys 180 residue forms a hydrogen bond with Lys 183. The Glu 180 residue forms a salt bridge with Arg 44 and Lys 183, respectively of α- and β-subunit. This ionic & salt bridge interaction disturb due to the change in polarity & hydrophobicity (Table 4.16), respectively (Figure 4.13 F).

4.3.10.9 L182V

In the crystal structure of hAC, Leu residue is located in an α-helix and buried in the core of a protein. The mutation converts the Leu residue into Val and Val does not favor α-helices. Val 182 will cause an empty space in the core of hAC and disturb the core due to small size. The wild-type and mutant residue interact with Ile 178, Leu 185 and Thr 186 of β-subunit of hAC (Figure 4.13 G).

4.3.10.10 T222K

The mutant residue Lys 222 is bulikier and less hydrophobic than Thr. Thr 222 forms a hydrogen bond with Ser145. Lys 222 do not make the same H bond as the Thr 222 because of the change of position due to size differences. The hydrophobic interactions of the core of the enzyme will be lost. The H bond formation might affect due to the difference in hydrophobicity (Figure 4.13 H). In the 3D-structure of hAC, Thr is located in its preferred secondary structure, a residue of β-subunit. However, Lys prefers to be in another secondary structure thereby slightly destabilizing the residue conformation. This mutation is probably not damaging to the enzyme because Thr residue is not conserved at 222 position in hAC but might disturb the core structure because Thr 222 buried in the core of Thr 222. This mutation creates a protein folding problem due to changing the charge of buried residue (Figure 4.9).

4.3.10.11 R226P

The mutant residue Pro 226 is smaller and more hydrophobic compared to Arg. The wild-type residue Arg 226 forms a salt bridge with Asp 297 and Asp 322. Pro disturb the ionic interaction made by Arg due to charge differences (Figure 4.13 I). The Arg residue is not conserved at 226 position in AC (Figure 4.10). Hence mutation at this position is probably not damaging enough for the protein. This mutation might disturb the core structure of hAC because Arg 226 is buried in the core of an hAC (Figure 4.9). The mutation will result in the formation of empty space in the core of of hAC and as a result, disturb correct folding.

4.3.10.12 G235R

The mutant residue Arg 235 is bulkier and less hydrophobic than Gly. The flexibility of Gly might be necessary for the protein's function. The Arg will add a charge in the core of hAC and create folding problems and hence might destabilize the protein structure (Figure 4.14 C).

4.3.10.13 G235D

The mutant residue Asp 235 is bulkier and more hydrophilic Gly 235 was buried in the hAC and mutation can reduce the flexibility of the core structure of the protein. The Asp introduces a charge in the core of the enzyme and responsible for misfolding.

4.3.10.14 R254G

The mutant residue Gly 254 is smaller and more hydrophobic than Arg. The Arg 254 forms a hydrogen bond with Val 208, Glu 258 and Glu 115 & this interaction loss due to change in size as well as in polarity (Table 4.16). The mutated residue Gly 254 is located on the surface of protein due to the charge difference (Figure 4.10). Arg 254 also forms a salt bridge with Glu115 and Glu 258 & mutation loss this interaction due to change in polarity (Figure 4.13 J). Arg 254 is present in the β-subunit and interacts with Glu 115 of α-subunit. The hydrogen bonds made by Arg to the α-subunit residues will be lost as a result of the mutation and affect the heterodimeric interactions. Gly is very flexible and will disturb the rigidity provided by Arg to the protein structure.

4.3.10.15 N320D

The Asn 320 residue forms a hydrogen bond with Cys 143, Glu 225 and Arg 333. Asp might disturb the core of hAC by introducing a charge in a buried residue and lead to misfolding. This will also affect processing & catalysis of the enzyme due to the gain of new interaction with active site residue Cys 143 (Figure 4.11 A) (Figure 4.10).

4.3.10.16 N320S

Asn 320 forms a H bond with Cys 143 and Glu 225 (Figure 4.11 A). Ser 320 will cause loss of hydrogen bonds in the core of the hAC due to the difference in hydrophobicity and as a result, disturb correct folding. This will also affect processing & catalysis of the enzyme due to loss of interaction with active site residue Cys 143 as well as with Arg 333 of β-subunit. The residue Asn 320 is buried in the core of the hAC and the mutation will lead to the formation of empty space in the core of hAC due to size differences (Figure 4.10).

4.3.10.17 D331N

Asn is disturbing the ionic interaction made by the Asp residue with Arg 333 due to the difference in charge and could possibly affect the catalysis and processing of AC. The Asp residue is not conserved at 331 position in AC. Asn is observed more often at 331 position in other homologous sequences of AC (Figure 4.9). The residue Asp is buried in the core of the hAC and the core structure of this enzyme might be disturbed due to loss of the charge.

4.3.10.18 R333G

The mutant residue Gly 333 is smaller and more hydrophobic than Arg. Arg 333 forms a hydrogen bond with Asp 162. The difference in hydrophobicity and size will affect hydrogen bond formation (Table 4.16). Arg 333 residue forms a salt bridge with Asp 162, Asp 331, Asp 332 and Asp 395 (Figure 4.11 B). The mutation disturbs this interaction due to the change in polarity. The flexibility of mutant residue can disturb the rigidity of Arg. Arg 333 is buried in the enzyme and mutation might disturb the structure of hAC by forming an empty space in the core (Figure 4.10). This mutation also disturbs the correct folding of hAC due to loss of interaction in the core of hAC.

4.3.10.19 R333C

Arg 333 forms a hydrogen bond with Asp 162. It also forms a salt bridge with Asp 162, Asp 331, Asp 332 and Trp 395. The ionic interaction and H bond formation will be disturbed by differences in polarity & size as well as hydrophobicity, respectively. The Arg 333 is buried in the protein, and mutation might disturb structure by generating a hollow space in the core of hAC. This mutation also disturbs the correct folding of hAC due to loss of interaction in the core of hAC. This mutation will disturbs the correct folding of hAC due to loss of interaction in the core of the protein.

4.3.10.20 R333H

Arg 333 residue forms a salt bridge with Asp 162, Asp 331, Asp 332 and Asp 395 and mutation will disturb it due to the difference in charge (Figure 4.11 B).

4.3.10.21 P362R

The rigidity of Pro induces a special backbone conformation and mutation will disturb it. Pro 362 is buried in the core of the protein and the Arg add a charge in a core of hAC residue which affects the folding of hAC. The Arg is bigger and probably will not fit properly in this position (Figure 4.10 D).

4.3.10.22 P362T

The core of the hAC might disturb due to loss of hydrophobic interaction as well as the difference in the size of residues (Figure 4.14 D).

4.3.10.23 T42A

The change in hydrophobicity of residue might disturb the correct folding of hAC. Thr 42 interact with Arg 32 and this interaction lost due to mutation. It also interacts with Asp 387 and Arg 386 of β-subunit of Hac and mutation did not affect this interaction (Figure 4.14 A).

4.3.10.24 T42M

This mutation disturbsthe correct folding of hAC due to change of hydrophobicity (Table 4.16). It also affects the interaction with Arg 32 but not with Asp 387 and Arg 386 of β-subunit like T42A mutation.

4.3.10.25 K152N

The Lys 152 residue forms a salt bridge with Asp 150 & mutation did not affect this interaction (Figure 4.13 C).

4.3.10.26 T179I

The mutant residue Ile 179 is bulkier than Ile. The Thr and Ile form an H bond with Gly 164 & Lys 183. The wild-type residue and mutant residue is not conserved at this position. Thr 179 is buried in the core of the enzyme and might disturb structure (Figure 4.13 E).

4.3.11 Molecular dynamics

Based on the prediction scores obtained using different methods, 7 mutations (L182V, R226P, G235R, G235D, R254G, R333H and P362R) were identified as highly deleterious from 30 disease causing nsSNPs. Molecular dynamics simulation approaches cover insight analysis of the effects of disease mutations on hAC structure. 7 highly deleterious nsSNPs (L182V, R226P, G235R, G235D, R254G, R333H and P362R) were investigated using 65-ns MD simulation trajectories.

The wild type enzyme and selected mutations models were comparatively analyzed throughout the simulation, particularly solvent-accessible surface (SASA), radius of gyration (ROG), total number of intramolecular hydrogen bonds, root mean square fluctuations (RMSF) and root mean square deviation (RMSD) of the enzyme with the time-dependent

function of MD to understand the effect of mutation on conformation of protein. Statical data of MD analysis shown in (Table 4.16).

Table 4.17: Statistical analysis for the MD simulations trajectory of WT and mutated proteins.

Sr. No.	Protein	RMSD		Intramolecular H-Bonds		Radius of Gyration		Solvent-accessible surface area	
		Range	Mean	Range	Mean	Range	Mean	Range	Mean
1	WT	0-1.2	0.7	40-71	53	2.2-2.2	2.2	69.2-81.7	76.0
2	L182V	0-1.5	0.8	44-71	56	2.3-2.3	2.2	69.2-83.3	77.1
3	R226P	0-1.3	0.7	37-66	50	2.2-2.3	2.2	71.0-83.4	77.8
4	G235D	0-2.1	1.0	38-70	55	2.3-2.5	2.4	69.3-83.6	77.3
5	G235R	0-1.9	1.2	46-75	61	2.2-2.4	2.3	66.2-81.3	73.7
6	R254G	0-2.0	1.0	36-65	51	2.3-2.5	2.4	69.5-84.2	78.0
7	R333H	0-2.0	1.1	48-75	60	2.3-2.5	2.4	65.8-80.4	74.5
8	P362R	0-2.3	1.5	34-64	50	2.4-2.5	2.5	70.5-83.2	77.8

4.3.11.1 Protein structure conformational flexibility and stability analysis

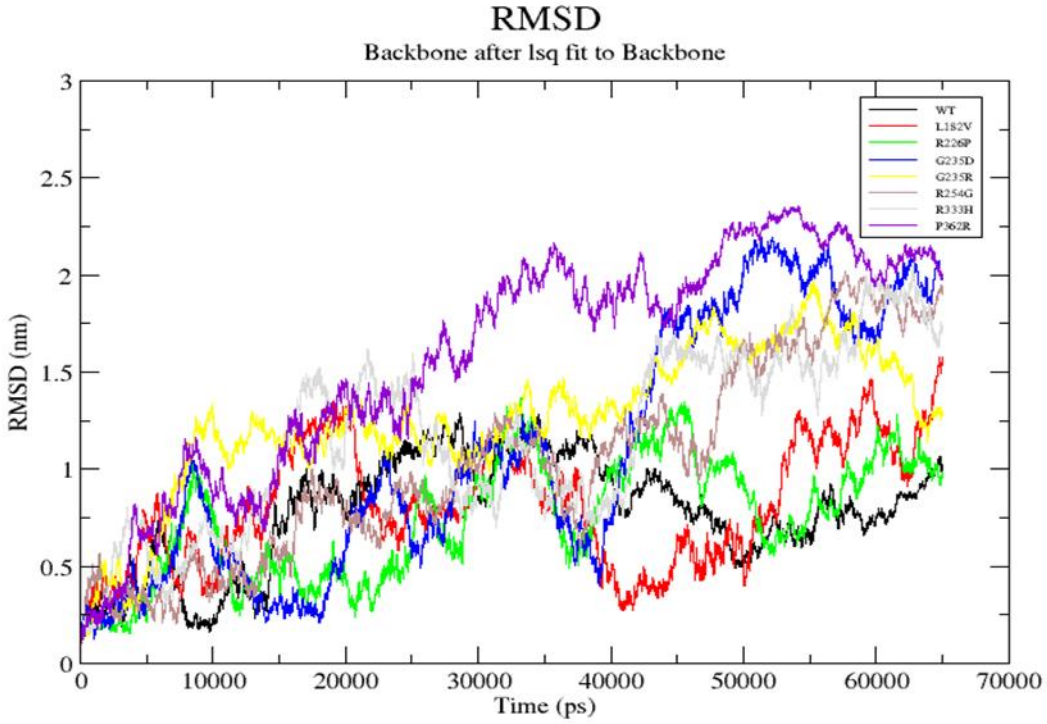


Figure 4.15: Backbone RMSD values of the WT protein along with those of the associated disease mutant proteins.

The RMSD analysis was done for eight disease mutant as well as wild type to interpret the mutation effect on hAC stability. All 8 systems standard deviation range of RMSD was between ~0.7 to ~1.5 nm and reached stable states after 50 ns of simulation. After 50 ns all mutant show higher value compared to wild type hAC. Both the wild type and mutant structures of hAC reached more stable states during ~50 to ~65 ns simulation period of MD. Hence, all further analyses were done at ~55 to ~65 ns. P362R showing high RMSD value indicates that this mutant is highly unstable (Figure 4. 15).

RMSF fluctuations of each residue in the enzyme analyzed to understand the dynamic behavior of each residue. From RMSF plot and statistic data, concluded that residue level fluctuations for G235D, R333H and G235R were quite high when compared with wild type and another mutant. Mutant L182V show fewer fluctuations compared with the wild type and another mutant. From RMSF plot the, it was concluded that, R333H mutant protein was shown the greatest degree of flexibility (Figure 17).

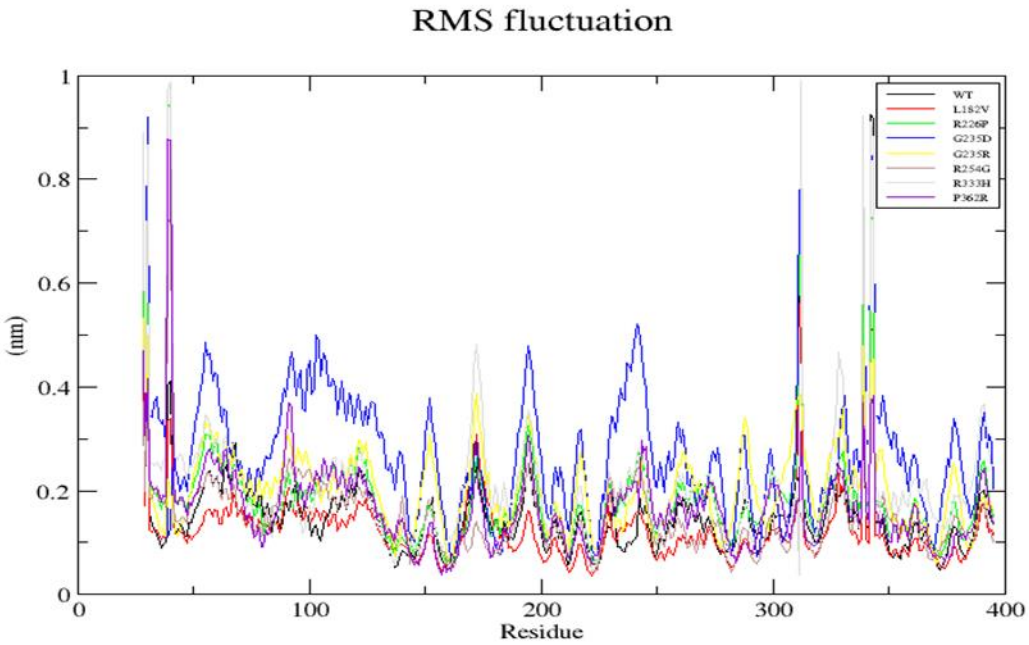


Figure 4.16: Central alpha-carbon RMSF values of the WT protein along with those of the associated disease mutant proteins.

We have analyzed the ROG to understand the effect of the mutation on the compactness of an enzyme. It can be concluded that P362R mutant has higher ROG value 2.5 nm hence its structure is less compact and unstable, while hAC has less ROG value 2.2 nm hence its structure is highly compact and stable. From ROG plot and statistic analysis of

ROG, it was concluded that, all mutation has caused structural destabilizing effects to result in loss of protein compactness (Figure 17).

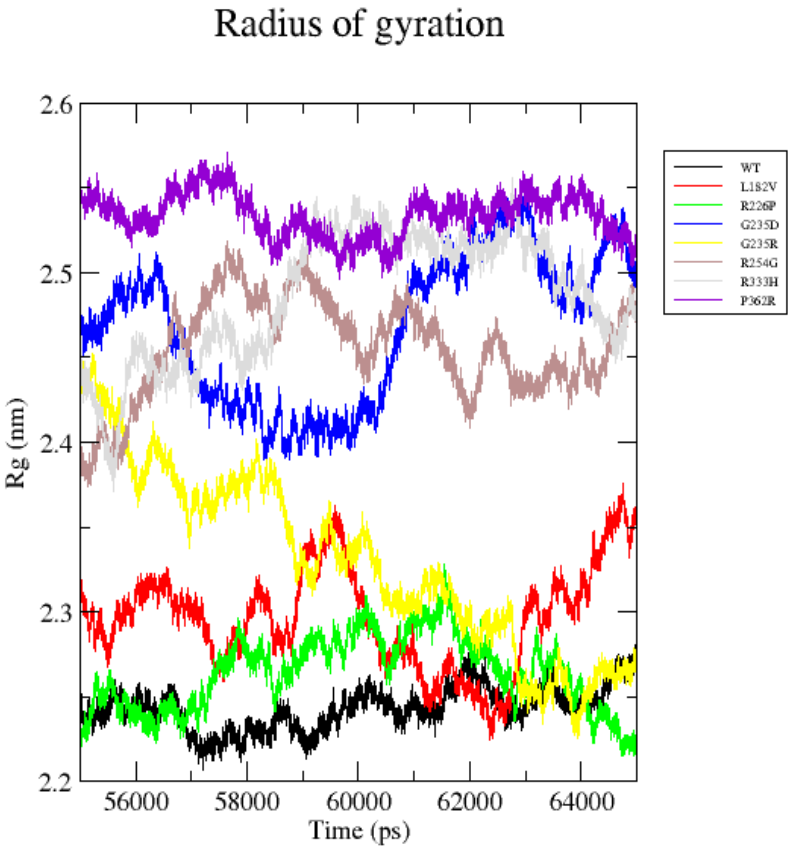


Figure 4.17: Radius of Gyration of for WT protein along with those of the associated disease mutant proteins.

4.3.11.2 Hydrogen bonding in wild and mutant AC proteins

The total number of hydrogen bonds was evaluated for seven mutant structures with comparative wild type AC for the understanding effect of the mutation on the stability of enzyme structure through MD simulation. The quantifying the number of H bonds offers important information about the stability of enzyme as well as molecular recognition. The maximum number of H- bond was observed for G235R (61) followed by R333H (~60), L182V (~56), G235D (~55) and wild type (~53). The R254G, R226P and P362R mutations have less number of H-bond with respect to wild type hAC. Remarkably, the number of H bonds in the seven mutants AC diverse with respect to wild type enzymes, this indicate that

due to a point mutation in hAC responsible decrease or increase number of H bond. Decrease of H bond might decrees the stability of the enzyme (Figure 4.18).

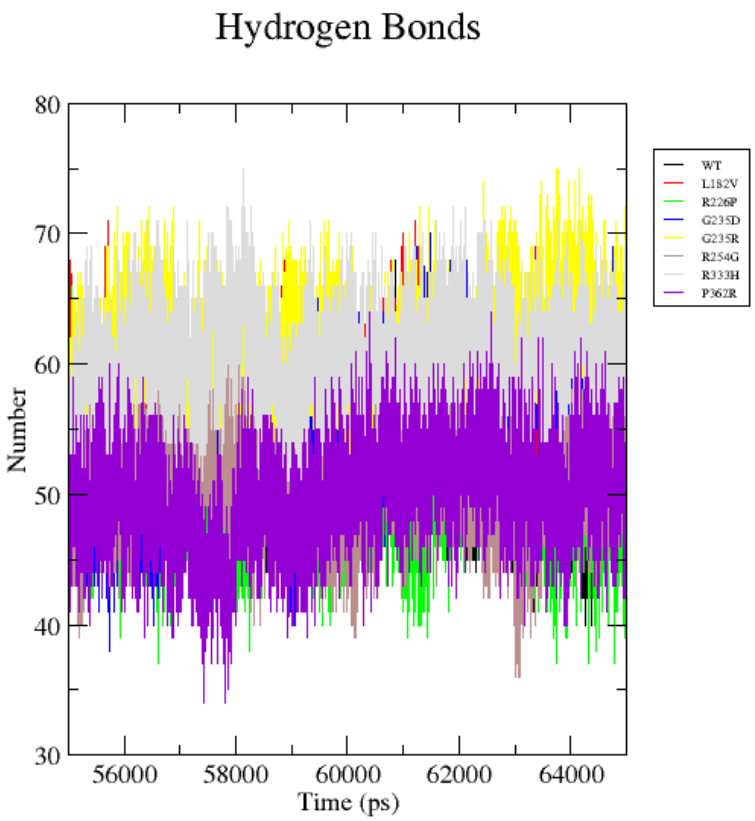


Figure 4.18: Total number of intramolecular hydrogen bond for WT protein along with those of the associated disease mutant proteins.

4.3.11.3 Effect of hAC disease mutant on SASA of hAC

The solvent-accessible surface area (SASA) of a protein is the protein surface accessible to the solvent molecule. The explicit models used to calculate solvation effects (Becker et al., 2001). The solvation effect maintains protein folding and stability. It also participates in the protein binding process and the rearrangement of protein structure. The SASA models have regularly acknowledged within available implied solvent treatments and have been used efficaciously in MD (Ferrara, Apostolakis and Caflisch, 2002) and structure prediction of protein.

The wild type AC protein has a SASA of ~76 nm2, whereas the G235R and R333H have less SASA (~73.7 nm2 and ~74.5 nm2, respectively) with few fluctuations than the wild

type enzyme in the last ten ns of the simulation period. The changes in the SASA of the disease mutant structures of hAC point toward that moving of mutant residues from exposed region to core regions, or vice versa. The changes in SASA of AC structure due mutation affect the folding and stability of the enzyme. Reduced SASA value for G235R and R333H indicatetes that the enzyme assumes a compact conformation upon mutation (Figure 4.19).

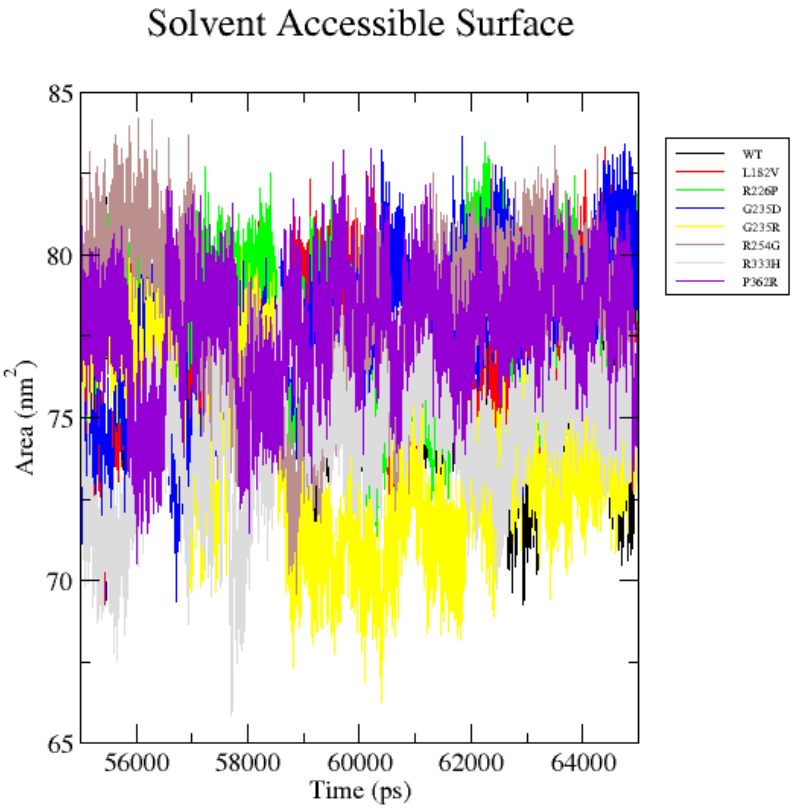


Figure 4.19: **Solvent-accessible surface area of the wild type and mutant structures of the AC.**

4.3.12 The four substrate binding pocket loops of AC and NAAA and comparison with BSH

A structural comparison of AC and NAA with Ntn hydrolase member BSH structure revealed similar positional preference of their catalytic residues in the active site region (Figure 4.20 A). The substrate binding sites consist of four loops, i.e. loop1-loop4, with residues 163-176, 205-211, 273-277 and 360-368, respectively, in hAC (Table 4.18). Out of four binding site loop in this enzyme, loop1 show a significant difference in terms of their length. Loop1 in AC is larger than the NAAA . Loop1 of hAC is important for protein

stability and activation. The Trp 169, a residue mutated in FD mapped on loop1, which indicate that loop 1 play an important role in the function of an enzyme. (Gebai et al., 2018).

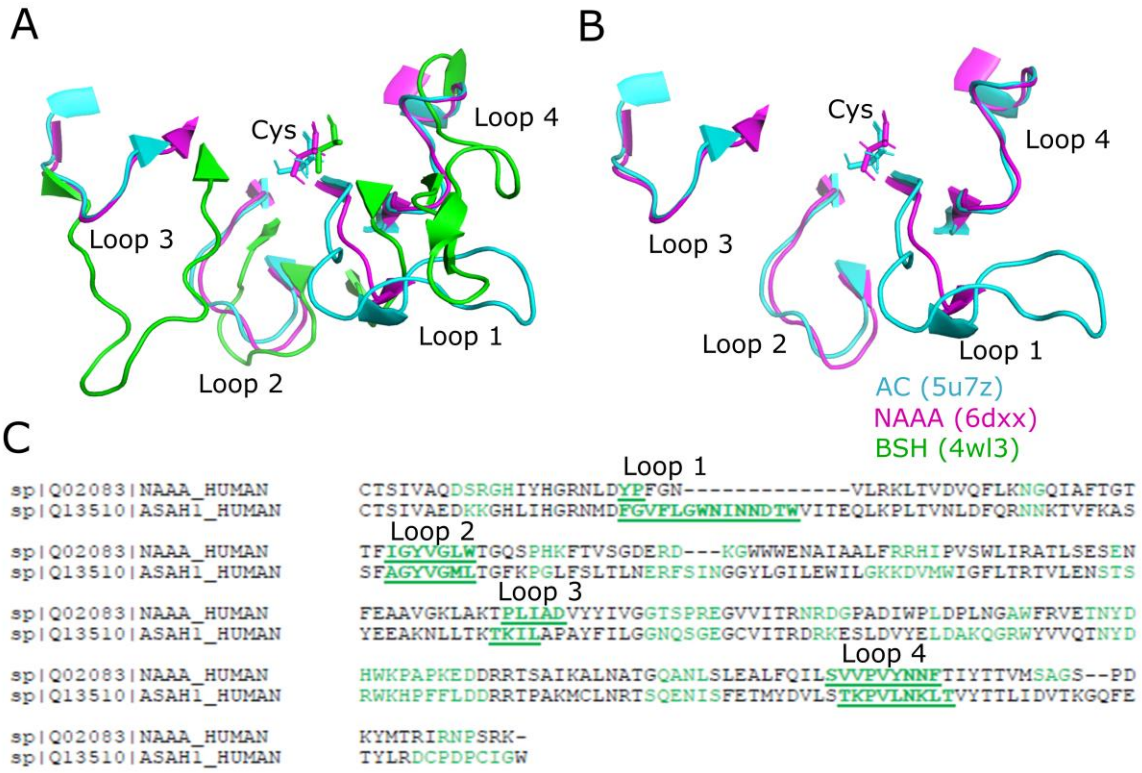


Figure 4.20: Substrate binding pocket loops mapped on the sequence and structure of AC and NAAA. A: The four substrate binding pocket loops of AC and comparison of AC (Cyan color) with EfBSH (Green) and NAAA (Magenta). B: Comparison of the AC loops with NAAA. C: Substrate binding pocket loops mapping on primary sequences of both AC and NAAA

Table 4.18: Substrate binding pocket loops of the hAC and hNAAA comparison with EfBSH.

Loop No.	AC(5u7z)	NAAA(6dxx)	EfBSH(4wl3)
Loop 1	163-176 aa	146-147 aa	21-25 aa
Loop 2	205-211 aa	175-181 aa	57-64 aa
Loop 3	273-277 aa	241-245 aa	125-137 aa
Loop 4	360-368 aa	325-333 aa	255-269 aa

4.3.13 Amino acid-based phylogenetic analysis of AC and NAAA

Initially only assigned 231 AC and 186 NAAA sequences obtained from Merops C89 family from Merops database. This sequence belongs to the animal. These sequences were again cross-checked in Uniprot and GenBank databases to remove partial, unassigned, low-quality prediction sequences, isoforms and sequences without active site residue from the analysis. From animal sources, 253 sequences were studied and most of the animal having only AC or NAAA sequences is removed from the dataset for studying the evolution of AC and NAAA. The final dataset obtained from an animal source has 91 AC and 90 NAAA sequences and from this dataset, a dendrogram was built using MEGA. The AC and NAAA separate in two clusters.

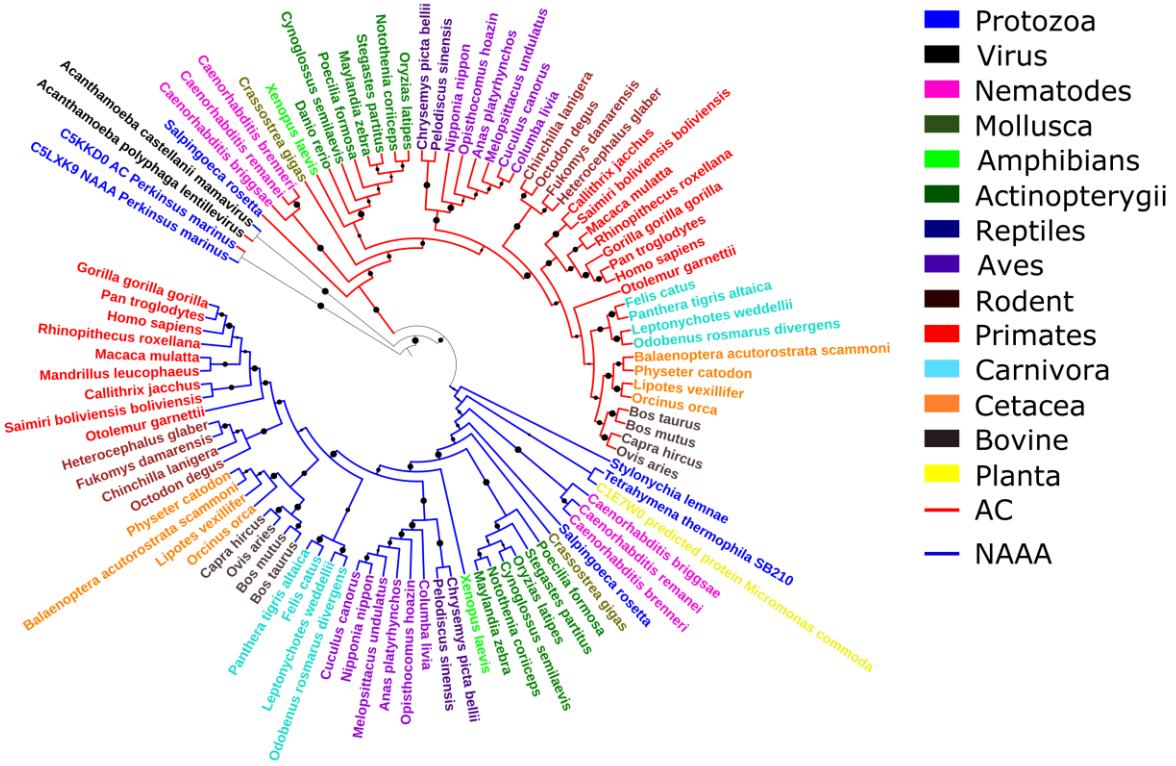


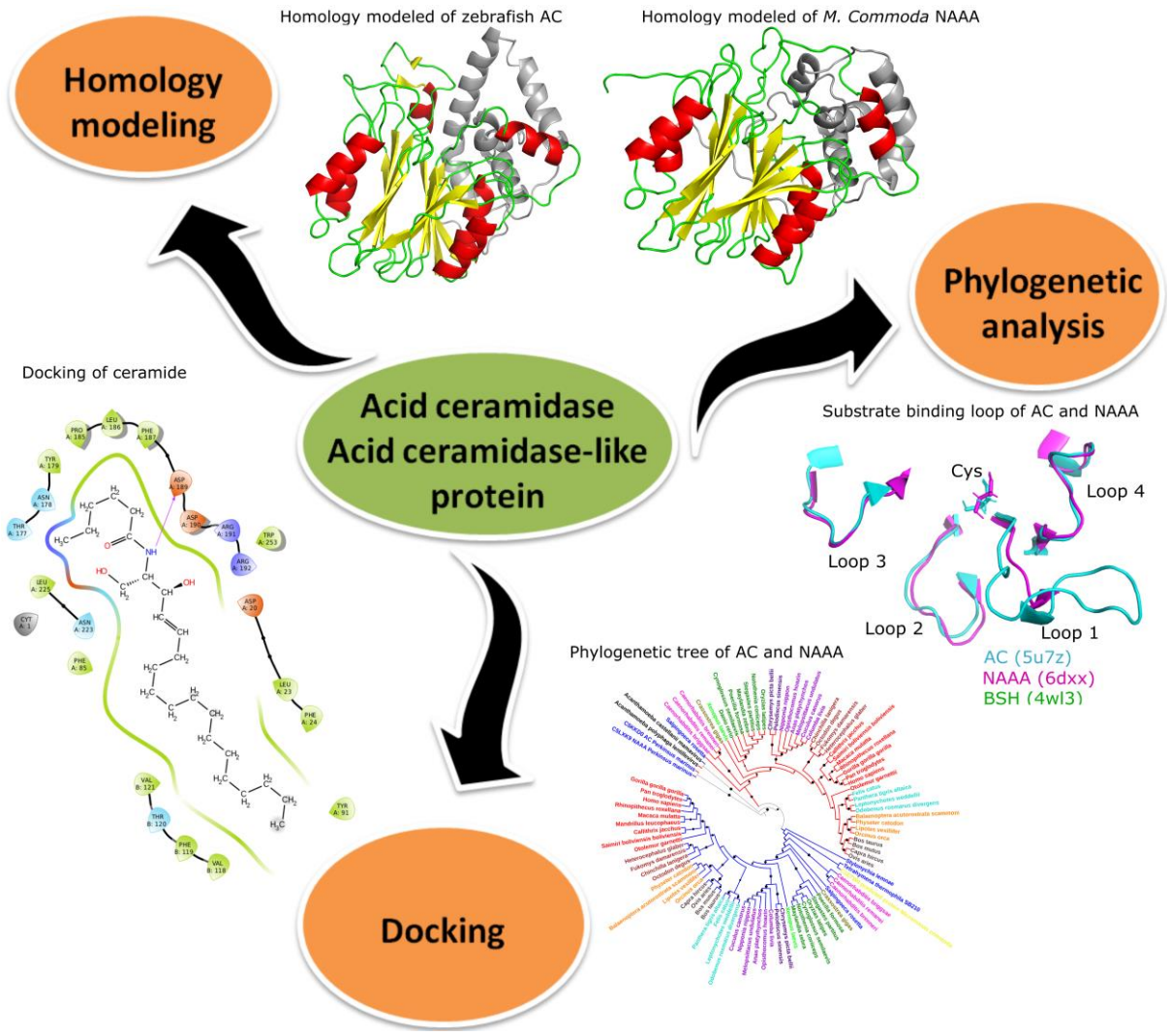
Figure 4.21: The dendrogram prepared based on the phylogenetic analysis of the sequences of AC and NAAA

Nine sequences of AC and NAAA (micromonas sp, protozoa and virus) selected out of 298 unassigned sequences of C89 family in Merops database and cross-checked in Unipart and GenBank for assigned it as AC and NAAA. This nine sequences (One from micromonas, two from viruses and six from protozoa) of C89 famly are added in the final animal dataset (assigned AC and NAAA). From this data set again tree is constructed. The AC and NAAA again separated into two clusters but virus and some protozoa AC and NAAA are not separated. The Virus AC and NAAA made one separate cluster as well as parkinsus protozoa

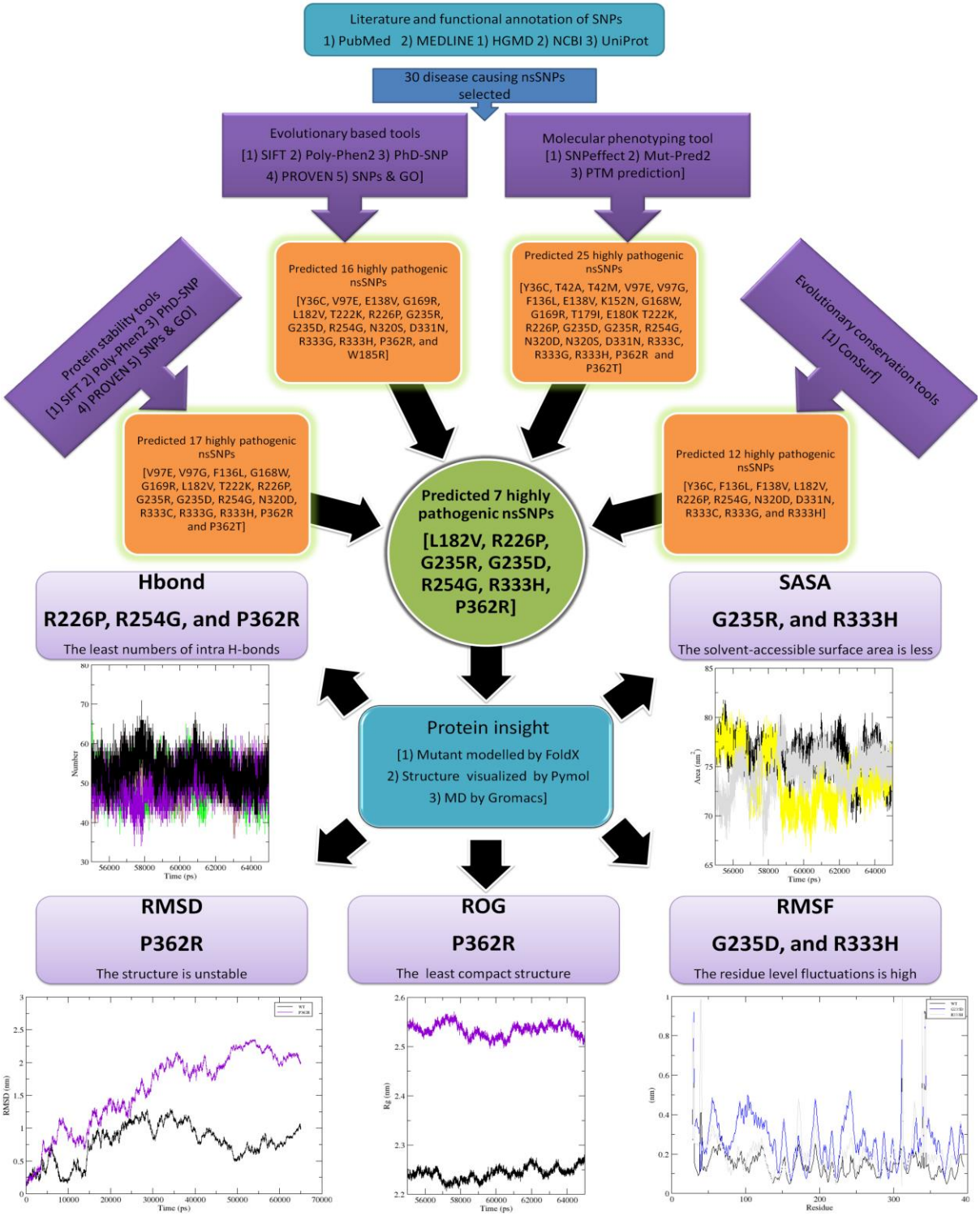
AC and NAAA separated in a separate cluster. AC and NAAA sequences (47 and 46 respectively) were manually removed from the dataset since they belonged to the same group (order, phylum, et cetera) since the original had more sequences so that the tree looked simple without altering the phylogeny. After creating a dataset of 97 (47 AC + 50 NAAA) protein sequences of AC and NAAA, a dendrogram was constructed using MEGA to study AC & NAAA evolutionary relationships. AC and NAAA make a separated cluster in animal and some protozoa but in some protozoa, AC and NAAA are not separated. In planta only micromonas have NAAA but AC is absent and in higher plants both AC and NAAA is absent. Micromonas sp. AC is separated from animal AC (Figure 4.21). In zebrafish NAAA is absent and in drosophila both AC and NAAA is absent.

4.4 Graphical Summary

A: Schematic illustration of AC and NAAA structural analysis



B: Schematic illustration of the effect of FD and SMA-PME mutation on AC structure



Chapter 5: Summary and conclusions

Acid ceramidase (AC) and acid ceramidase-like protein (NAAA) that are clinically and pharmaceutically significant enzymes have been investigated in this thesis work. The enzymes that have been characterized in this study are zebrafish AC, *C. elegans* AC, mice NAAAand *M. commoda* NAAA. These belong to the Ntn-hydrolase superfamily. The main objective of our analyses was to understand the comparative structure-function relation of these enzymes that can enhance their prospective applications. The study involved biochemical, biophysical and computational evaluation of these proteins.

AC and NAAA are enzymes that catalyze the hydrolysis of the amide bond in ceramide and in NAE, respectively. The gene coding for AC from zebrafish and *C.elegans* and NAAA gene coding from mice and *M. commoda* were fished out using gene-specific primers, cloned and purified by heterologous expression using *Pichia pastoris* eukaryotic expression cells. The protein was secreted into the medium with the help of an alpha-factor signal peptide and N-glycosylated in endoplasmic reticulum and O- glycosylated in Golgi bodies.

All recombinant enzymes were purified by Ni-NTA affinity chromatography. The size exclusion chromatography was employed to further purify zebrafish AC and mice NAAA. After initial the affinity chromatography, Concanavalin A-sepharose chromatography was followed to separate glycosylated form of M. commoda NAAA from a nonglycosylated form of M. commoda NAAA. Purified recombinant proteins are intrinsically heterogeneous in nature due to post-translational processing and glycosylation. Purified proteins were qualitatively and quantitatively estimated by western blotting, enzyme activity and autoproteolytic processing assay. The polyclonal anti-human AC antibody showed cross-reactivity against zebrafish AC.

The molecular weight of α- and β-subunits of both the proteins (AC and NAAA) predicted from the highly conserved nature of this autoproteolytic processing site in sequences of AC and NAAA. Predicted molecular weight and autoproteolytic cleavage sites were confirmed by the molecular weight estimation of unprocessed, α-and β-subunit in SDS-PAGE. The autoproteolytic cleavage site was precisely established by western blotting using anti-his tag antibody, which selectively catches only unprocessed protein and β-subunit.

Autoproteolytic cleavage assay again helped to corroborate and characterize unprocessed, α- and β-subunit of AC and NAAA.

Table 5. 1: Glycosylation summary of recombinant AC & NAAA.

Recombinant protein	Subunits	Predicted molecular weight (kDa) *in silico*	Observed molecular weight (kDa) in SDS-PAGE	Predicted glycosylation *in silico*		Observed glycosylation	Comments
				N-glycosylation sites (NetNGlyc 1.0 Server)	O-glycosylation sites (NetOGlyc4.0 Server)		
Zebrafish AC	α-subunit	14	14	No potential sites	No potential sites	No glycosylation observed	Non-glycosylated
	β-subunit	30	~30-40	173, 259, 286, 301, 342 and 348	344	Visualized as pink color bands in glycosylation staining. Observed 4 bands in western blotting. Deglycosylation by PNGase F reduced 4 bands into 2 bands. Deglycosylation by PDM II reduced the apparent molecular size to 30 into 1 band	Glycosylated
***C. elegans* AC**	α-subunit	14	17	118	No potential sites	Visualized as pink color bands in glycosylation staining. Deglycosylation by PDM II reduced the apparent molecular size of bands	Glycosylated
	β-subunit	30	30	No potential sites	No potential sites	No glycosylation observed	Non- glycosylated
Mice NAAA	α-subunit	11	~11-15	42 and 112,	No potential sites	Visualized as pink color bands in glycosylation staining. Deglycosylation by PDM II reduced the apparent molecular size of bands	Non- glycosylated
	β-subunit	27	~34	314, 338and 360	307	Visualized as pink color bands in glycosylation staining. Observed 4 bands in western blotting. Deglycosylation by PNGase F reduced 4 bands into 2 bands. Deglycosylation by PDM II reduced apparent molecular size bands	Glycosylated
***M. commoda* NAAA**	α-subunit	17	~20	128	49 and 50	Glycosylated protein purified by Concanavalin A-sepharose chromatography	Glycosylated
	β-subunit	30	30	No potential sites	No potential sites	No glycosylation observed	Non-glycosylated

Observed molecular weight was higher than the predicted molecular weight (based on primary amino acid sequences) indicating that the protein is glycosylated. Glycosylation of protein was confirmed by the glycosylation staining kit, which positively stained glycosylated bands, appearing as a magenta band with the light pink background in SDS-PAGE.

We also predicted N-glycosylation and O-glycosylation sites by sequence analysis using NetNGlyc 1.0 Server (Gupta, Jung and Brunak, 2016) and NetOGlyc 4.0 Server (Steentoft *et al.*, 2013) respectively. Between both approaches *in silico* glycosylation prediction and glycosylation staining facilitated to determine subunits (α- and β-subunit), which is glycosylated. We further established N-glycosylated subunits (α- and β-subunit) with deglycosylation of protein by PNGase F with the supportive advantage of glycosylation staining and western blotting. The comparative account of three experiments viz, deglycosylation, glycosylation stainingand His-tag based western blotting revealed more precise information about glycosylation of subunits. Whenever PNGase F failed to complete deglycosylation then we have also used PDM II treatment to confirm O-glycosylation. A comparative account of the glycosylation status of purified recombinant proteins is summarised in Table 5.1.

Secreted zebrafish AC and mice NAAA were purified as a mixture of unprocessed and mature form but *C. elegans* AC and *M. commoda* NAAA were secreted only as unprocessed form. This unprocessed form matures into α- and β-subunit. The unprocessed form slowly converts into mature form at pH 4 and 37 °C. This autoproteolytic clevage process reaction also occurs at pH 7 and 4 °C but at a slower rate. The processed form and unprocessed form discriminated in SDS-PAGE on the basis of molecular weight of subunits (α- and β-subunit) and western blotting using anti-His antibodies.

The native molecular mass of zebrafish AC and mice NAAA was found to be ~100 kDa, as estimated by size-exclusion chromatography suggested that these enzymes exist as a homodimer of the heterodimer (α + β) having two α-subunit and two β-subunits.

The electrophoretic mobility (under reducing conditions and further under non-reducing conditions) of the cleaved form of zebrafish AC changes but not in case of mice NAAA. This observation suggests that zebrafish AC does form an αβ heterodimer disulfide-linked but not in case of mice NAAA. The sequences of α-subunit of NAAA did not contain cysteine residues but α-subunit of AC contains one conserved cysteine residue. This in silico and experimental analysis indicates that an αβ heterodimer is connected through a disulfide bond in AC but not in NAAA.

Zebrafish AC and *C. elegans* AC are able to carry out both ceramide degradation and synthesis *in vitro* using fluorescent substrate (NBD-C12-ceramide/ N-(NBD-Aminolauroyl) fatty acid). They also hydrolyze PEA into ethanolamine and palmitic acid. Mice NAAA, *M. commoda* NAAA can carry out both ceramide hydrolysis and NAE hydrolysis *in vitro*. Kinetic analysis of the zebrafish AC and mice NAAA was carried out using ceramide substrate (N-Lauroyl-D Sphingosine) and NAE substrate (PEA). The optimal pH, temperature and time for both this reaction were near pH4, 35 °C and 24 hours, respectively. The optimum DTT concentration for zebrafish AC for ceramide hydrolysis and mice NAAA for NAE hydrolysis is 0.5 mM and 8 mM respectively. The dynamic properties of zebrafish AC and mice NAAA were investigated using CD; in case of mice NAAA and *M. commoda* NAAA intrinsic fluorescence measurement using fluorescence spectroscopy and DSF, respectively were used to study conformational transition.

Structures of zebrafish AC and NAAA were generated through homology modeling. Ceramide molecule was docked into the zebrafish AC homology model structure for characterizing the active site of the enzyme.

30 disease causing nsSNPs are analyzed for the screening of the most damaging nsSNPs. 11nsSNPs (V97E, G169R, L182V, T222K, R226P, G235R, G235D, R254G, R333G, R333HandP362R) were identified as highly deleterious using 10 different tools (SIFT, PROVEAN, PolyPhen2, PhD-SNP and SNPs and GO, I mutant (sequences based), I mutant (structure based), PoPMuSiC, FoldX). 14nsSNPs (Y36C, L182V, R226P, G235D, G235R, R254G, N320D, N320S, D331N, R333C, R333G, R333H, P362Rand P362T) were identified as highly damaging based on the result obtained from 5 different servers (ConSurf, Mut-Pred 2, SNPeffect and PTM prediction server). Finally based on the prediction result obtained from all in silico tools, 7 nsSNPs (L182V, R226P, G235R, G235D, R254G, R333Hand P362R) were identified as highly deleterious from 30 disease causing nsSNPs. Further, in silico mutation were generated for clinically reported disease mutants of hAC using FoldX program and visualized through Pymol to understand the molecular level interaction of mutant residues compared to wild-type residue. The 7 highly deleterious nsSNPs were investigated using 65-ns MD simulation to understand the effect of FD and SMA mutation on the hAC enzyme structure. The RMSD values of the mutant structure suggest that all 7 mutant structures are unstable compared to the wild-type protein. The residue level fluctuations for G235D, R333Hand G235R were quite high when compared with wild-type and other mutants. Based on the radius of gyration (ROG) it was inferred that P362R mutant has the least compactness of its structure with 2.5 nm, whereas the wild-type protein has

shown to be highly compacted with 2.2 nm. The less number of H-bonds were observed for R254G, R226P and P362R mutations when compared to the wild-type protein. Lower SASA value observed for G235R and R333H than the wild-type hAC enzyme.

Following are the major highlights of the observations recorded and conclusion of the thesis

- Sequences analysis of AC and NAAA show identical sequence, conserved active site residues and lower conservation in α-subunit compared to β-subunit.
- The catalytic triad of zebrafish AC consists of Cys 143, Arg 159 and Asp 162; mice NAAA consist of Cys 131, Arg 147 and Asp 150; *M. commoda* NAAA consist of Cys 146, Arg 164 and Asp 167; *C. elegans* AC consist of Cys 140, Arg 156 and Asp 159 amino acid residues.
- AC and NAAA expression failed in a prokaryotic expression system (*E. coli*). This result might indicate that expressuin failed due to the absence of post-translational processing machinery. Eukaryotic expression system (*P. Pastoris*) is more efficient for expression of lysosomal glycosylated and autoprocessing protein like AC and NAAA.
- The yield of zebrafish AC is ~10 mg/l as compared to mice NAAA is ~5mg/l
- *M. Commoda* NAAA glycosylated form was purified from the non-glycosylated form.
- Disulfide linkage is present between α and β-subunit of zebrafish AC but absent in mice NAAA
- Oligomeric state of mice NAAA is a homodimer of the heterodimer (α + β) like zebrafish AC but intermediate monomer form was also observed.
- Glycosylation prediction analysis revealed six N-glycosylation sites along with a single O-glycosylation site on the β-subunit of zebrafish AC. Out of six, only five N-glycosylation sites are conserved in human and potential. *C. elegans* AC and *M. Commoda* NAAA have only one N-glycosylation site on the α-subunit. The α-subunit of human and mice AC does not have any N-glycosylation site on the alpha subunit. *M. Commoda* NAAA α-subunit N-glycosylation site is conserved in human and mice NAAA. Mice NAAA has 5 N-glycosylation sites out of which two are on the α-subunit conserved and potential in hNAAA and three on β-subunit only two of which are conserved and potential in hNAAA. Mice NAAA has one O-glycosylation site on β-subunit and *M. Commoda* NAAA has two O-glycosylation sites on the α-subunit.
- Glycosylation analysis: The β-subunit of zebrafish AC is both N- and O- glycosylated. At amino acid positions 344 O- glycosylation is present. The α-subunit of *C. elegans* AC and *M. Commoda* NAAA are glycosylated. In the case of mice NAAA, both subunits are

glycosylated. It is concluded that at least one N-glycosylation required for proper targeting and functioning of both NAAA.

- Autoproteolytic cleavage: All AC and NAAA studied in this thesis are heterodimer consisting of α- and β-subunit. These subunits are generated by autocleavage of precursor protein called as unprocessed. This autoproteolytic process is pH, time, detergent and salt-dependent vary from organism to organism. In case *M. Commoda* NAAA carboxy-terminal processing is also observed similar to β-subunit of hAC.
- Mice NAAA activity was enhanced many folds by DTT but not in case of zebrafish AC.
- Kinetics constant: K_m and V_{max} of zebrafish AC for N-Lauroyl-D Sphingosine are 273 μM and 170 μmol/h/μg, respectively at pH 4 and temperature 37 °C. K_m and V_{max} of mice NAAA for PEA is 67 μM and 214 nmol/h/ μg, respectively at pH 4 and temperature 37 °C.
- Reverse activity: Zebrafish AC and *C. elegans* AC synthesizes ceramide at pH 6.
- Cross-reactivity: Zebrafish AC and *C. elegans* AC are able to hydrolyze PEA at low rates. Mice NAAA and *M. Commoda* NAAA are able to hydrolyze N-Lauroyl-D Sphingosine at low rates. *M. Commoda* has good activity against ceramide compared to Mice NAAA.
- CD spectrum: zebrafish AC is rich in α-helix with slight β-sheet and mice NAAA is β-sheet rich with less α-helix content.
- Intrinsic fluorescence spectra of mouse NAAA shows λ max at 353.5 nm which implies that most of the tryptophan exposed to the surface.
- DSF spectrum of *M. Commoda* NAAA revealed that melting temperature (T_m) is 75.8 °C.
- Modeling of zebrafish AC and *M. Commoda NAAA* provides insights into the structural architecture of the enzyme
- SMA-PME mutants Y36C, T42M are more prone to aggregates and less active when compared to mutant N152K and wild type zebrafish AC.
- Molecular docking throws lights on enzyme-substrate interaction
- *In silico* disease mutants modeling helped to compare the stability of mutants with respect to the wild-type protein.
- Out of 30 Disease causing nsSNPs, 7 nsSNPs were identified (L182V, R226P, G235R, G235D, R254G, R333Hand P362R) as highly deleterious.
- Molecular dynamics simulation established the effects of mutations on the local and global stability of the protein.

- Four substrate binding loops have been mapped on the primary sequences as well as the structure of both AC and NAAA. Loop1 shows a significant difference in terms of their length and useful to separate AC and NAAA from each other.
- Phylogenetic analysis separated the NAAA and AC on the basis of primary structure and loop 1. Phylogenetic analysis also separated ceramidases into different isoforms.

In conclusion, our research reported here provides insight towards understanding the comparative structural and functional relationship between AC and NAAA with using cloning, expressing and characterizing the enzyme using various biochemical and biophysical techniques. In silico studies understand the molecular effect of a pathological mutation in AC leading genetic disease. We report for the first time, an efficient expression system for the heterologous overexpression of biologically active AC and NAAA in this study..

In direction to understand more about the relationship between the structure and function of lysosomal amidase (AC and NAAA), further research is crucial. This thesis study is centric to the characterization of AC and NAAA *in vitro* and *in silico* therefore *in vivo study* need to be investigated. Although these methods are useful to explore and improve analysis of genetic disease, understanding the diagnosis and treatment of FD and SMA-PME requires more work. Recently mammalian AC structure has been solved but the structure of AC with ceramide remains to be solved. Model of NAAA interacting with membrane lipids via the helix has been proposed but AC interaction with membrane via saposin D is still needs to be investigated.

Appendix A

A.1 Introduction

Penicillin G acylase (PGA) convert penicillin G into phenylacetic acid and 6-aminopenicillanic acid (6APA). The 6-APA used as precursor for produce semi-synthetic penicillin, which is more effective than natural penicillin (Arroyo *et al.*, 2003). This enzyme also ussefull for peptide synthesis and resolving racemic mixture of chiral compound. ***Escherichia coli*** (***Ec***PGA) enzyme unstable beyond 30°C (SIO and QUAX, 2004) but thermostability moderately improves by immobilization (Erarslan and Koçer, 1992) and site directed mutagenesis (Polizzi *et al.*, 2006). The Achromobacter xylosoxidans (AxPGA) (Verhaert *et al.*, 1997) and Alcaligenes faecalis (AfPGA) (Cai *et al.*, 2004) are more thermostable enzyme than ***Ec***PGA having half life at 55°C is 55 min and 15 min respectively. The reason of thermostability of ***Af***PGA enzyme may be due to present of disulfides bond in structure(Varshney *et al.*, 2012) . The ***Ax***PGA more thermostable than ***Af***PGA may be due to high ratio of Arg to Lys, high number of Pro residues and stable ion pair and absent of thermo labile amino acid residues and bond (Cai *et al.*, 2004). All PGA from MEROPS database compares with known thermostabile PGA & identified novel probable thermostable ***Acinetobacter oliverances*** PGA (***Ao***PGA) using computational approaches. Also identified site directed mutation site in ***Ao***PGA, which mutation could improve thermostability (Panigrahi *et al.*, 2015).

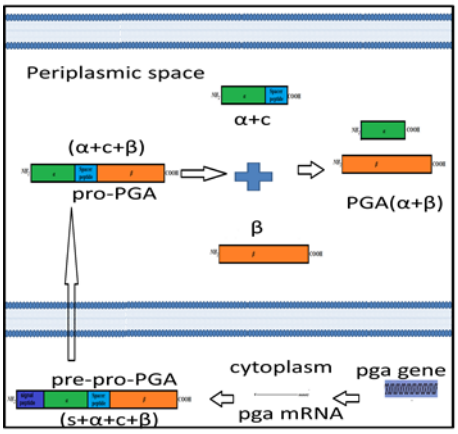


Figure A.1 Schematic representation of the synthesis and maturation by autocatalysis processing of the penicillin G acylases. The enzyme is synthesized in the cytosol with a signal peptide which transports the inactive pre-pro-peptide to the periplasm where autocatalysis processing takes place.

A.2 Materials & Methods

A.2.1 Cloning of *AoPGA* with *AoPGA* signal peptide

The genomic DNA was isolated from *Acinetobacter oleivorans* (NCMC) using DNA isolation Kit (Quagen). The *AoPGA* gene sequence was retrieved from GenBank (Accession No. NC_014259). Gene was amplified by gene specific primers, forward primer *AoPGA_NdeI* 5'-ACTAGTCATATGCTGAATTCTCGGCCT-3' and reverse primer *AoPGA_XhoI* 5'-ACTAGTCTCGAGGAGTTTAAAGTGATGCTTTGTTG-3'. The restriction sites are highlighted. ***Ao***PGA gene was amplified from gDNA. PCR condition; denaturation at 94^{o} C/3min, (94°C/30 sec, 55°C/30 sec, 72/2:30 min) x 25 cycles and final extension at 72°C for 10 min. The resulting PCR fragment was digested with *NdeI* and *XhoI* , cloned between the *NdeI* and *XhoI* sites of the expression vector pET22b(+). The ligation mixture was transformed in the DH5α competent cells and incubated overnight at 37oC. Colonies were screened for positive clones by using the colony PCR. In addition, to confirm the gene insertion in the cloned plasmid, double digestion was performed (*NdeI/XhoI*) in which the inserted gene was released.

A.2.2 Cloning of *AoPGA* with pel B signal peptide

In order to reach the AoPGA to periplasmic space for proper processing into mature form new construct was created using a new forward primer, *AoPGA_BamHI* 5'-ACTAGTGATCCGTCAACACCAACACAAAAGC-3' and reverse primer *AoPGA_XhoI*. These sites were selected to insert the gene into the pET22b(+) vector, in order to ensure that the recombinant DNA fragment had an pel B signal peptide preceding its N-terminal, followed by a C-terminus 6-His extension under the T7 promoter.

A.2.3 Expression in *E. coli*

The constructs transformed into BL21 (DE3) competent ***E. coli*** cell & selected by ampicilin selection marker. Expression cell having ***AoPGA*** gene grown in LB broth at 37°C for obtaining cell mass & induced 0.1M IPTG overnight grown at 16°C for protein expression. Cell harvested by centrifugation, disrupted by lysis solution with lysozyme & ultrasonication. Cell debris removal by centrifugation & purified by Ni-NTA affinity chromatography & size exclusion chromatography. The expression confirmed by western blotting.

A.2.4 PGA assay

The enzyme activity of *Ao*PGA was determined by the purified enzyme preparation reacted with its substrate penicillin G (20 mg/ml) in 50 mM phosphate buffer (pH 7.5), at 45˚C for 10 min. The reaction was quenched by adding 1 ml 300 mM citrate phosphate buffer (pH 2.5). The product 6-APA released was estimated spectrophotometrically at 415 nm after reacting it with 2 ml of 0.6% (w/v) p- dimethylaminobenzaldehyde in methanol.

A.3 Results

A.3.1 Cloning, expression & purification of *Acinetobacter oliverances* (*Ao*PGA)

Two construct designed one with AoPGA signal peptide another with Pel B signal for target the enzyme to periplasmic space of *E. coli*. ***Ao***PGA expression with ***Ao***PGA signal peptide was confirmed by western blotting (Figure A.1 A: lane 2) as well as penicillin G acylase enzyme activity (Figure A.I 2), therefore no expression (Figure A.1 G: lane 4 and 5) and activity (Figure A.I 1) was observed for construct with pel B signal. From this observation, it concluded that, *Acinetobacter oliverances* signal sequences work very well in *E. coli* expression system. The purified enzyme using Ni-NTA affinity chromatography followed by gel filtration (SEC) showed five bands in SDS-PAGE with respect to α-subunit (~23 kDa), α-subunit with spacer peptide (~30 kDa), β-subunit(~61 kDa), truncated or processed β-subunit (~33 kDa) and immature ***Ao***PGA (α-subunit, spacer peptide and β-subunit) (~92 kDa). In western blotting immature ***Ao***PGA detected along with β-subunit of enzyme. One extra band also observed, this may be further processing of β-subunit or it might be a truncation of β-subunit of ***Ao***PGA (Figure A. G: lane 2 and 3, and A.H: lane 2). Processing of β-subunit is not reported in case of PGA but another member of Ntn hydrolase reported such as in mouse lysosomal 66.3 kDa proteinand zebrafish acid ceramidase. Unprocessed forms slowly convert into processed form at 4°C (Figure A. H: lane 2). In SEC chromatogram PGA observed as monomeric as well as in the oligomeric state. According to the molecular weight of PGA observed in SEC might form oligomer state as pentamar. This enzyme was degraded during purification as well as storage at 4°C.

Our aim was to screen thermostable PGA for industrial applicaion, unfortunately, the result form experimental characterization of AoPGA indicating that this enzyme was unstable. Therefore it is not suitable for caring out further research on it.

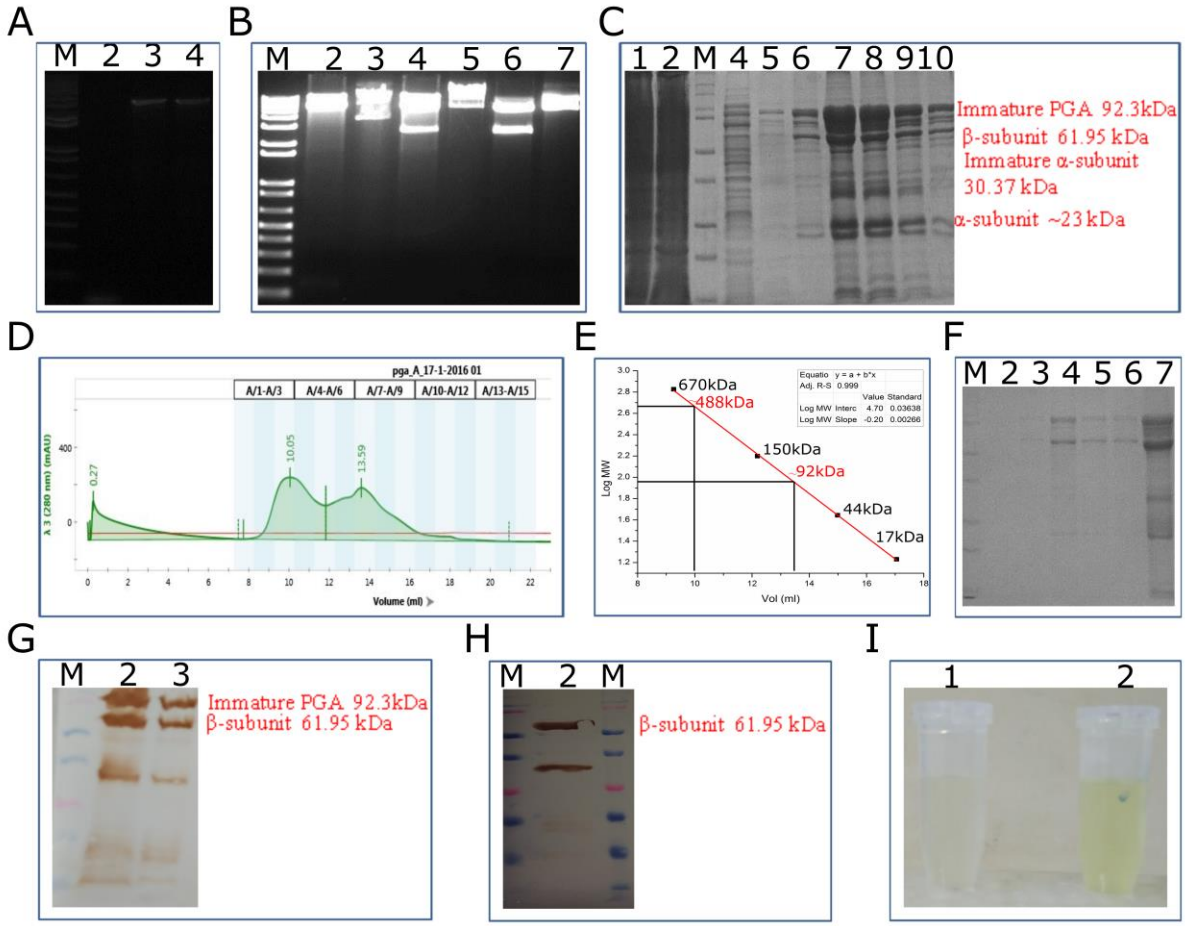


Figure A.1: Purification and characterization of *AoPGA*. B: 1% agarose gel showing PCR product of *AoPGA* gene. Lane 1, NEB 1Kb DNA ladder; Lane 2, PCR blank as negative control; Lane 3, 2.4 kb PCR product of *AoPGA* with *AoPGA* signal peptide; Lane 3, 2.3 kb PCR product of **AoPGA** with pel B signal peptide. C: Restriction digestion. Lane 1, NEB 1Kb plus DNA ladder; Lane 2, Undigested plasmid (pET22b); Lane 3, Double digestion (NdeI/XhoI) of plasmid (pET22b); Lane 4, Double digestion (NdeI/XhoI) of cloned plasmid (*AoPGA*-pET22b); Lane 5, Undigested cloned plasmid (*AoPGA*-pET22b); Lane 6, Double digestion (NdeI/ BamHI) of cloned plasmid (pel B-*AoPGA*-pET22b); Lane 7, Undigested cloned Plasmid (pel B-*AoPGA*-pET22b). C: AoPGA protein was purified using Ni-NTA affinity chromatography. Lane 1, Supernatant and Lane 2, Flowthrough; M, Prestained protein ladder; Lane 4-5, Wash; Lane 6-10, Eluted fractions. D: Chromatogram of size exclusion chromatography using sephacryl S200 (6/15) column for purification of AoPGA protein. E: Plot of logarithm of molecular weight against volume eluted. The molecular weight of recombinant AoPGA was found to be about 92 kDa. F: AoPGA protein was purified by AKTA purifier with size exclusion chromatography. M, Prestained protein ladder; SEC fraction A4-A8 loaded into Lane 2-6, respectively; Lane 7, Ni-NTA AoPGA. G: Western blotting analysis confirms AoPGA protein expression with monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 1, Pellet fraction of after lysis of AoPGA protein with intrinsic signal peptide; Lane 2, Supernatant after lysis of AoPGA protein with intrinsic signal peptide. Lane 3, Pellet fraction of after lysis of AoPGA protein with pel B signal peptide; Lane 4, Supernatant after lysis of AoPGA protein with pel B signal peptide. H: Western blotting analysis autoprocessing of AoPGA using monoclonal anti-His antibodies. M, Prestained protein ladder; Lane 2, Zebrafish AC protein was purified using Ni NTA affinity chromatography. I: PGA assay. 1, With pel B signal give no activity; 2, With AoPGA signal give activity.

Appendix B

B.1 Introduction

Micromonas pusilla BBS3 is ortholog of human ARL6/ BBS3 (ADP-ribosylation factor 6/ Bardet-Biedl syndrome 3) small GTP binding protein. Homozygous mutation in this gene associated with a rare human disorder Bardet-Biedl syndrome (BBS 3). BBS 3 is a genetically heterogeneous disorder(Young *et al.*, 1998) characterized by retinal degeneration, polydactyly, obesity, and mental retardation. ARL6/ BBS3 also play an important role in Wnt signaling and assembly and disassembly cilia of vertebrates (Wiens *et al.*, 2010).

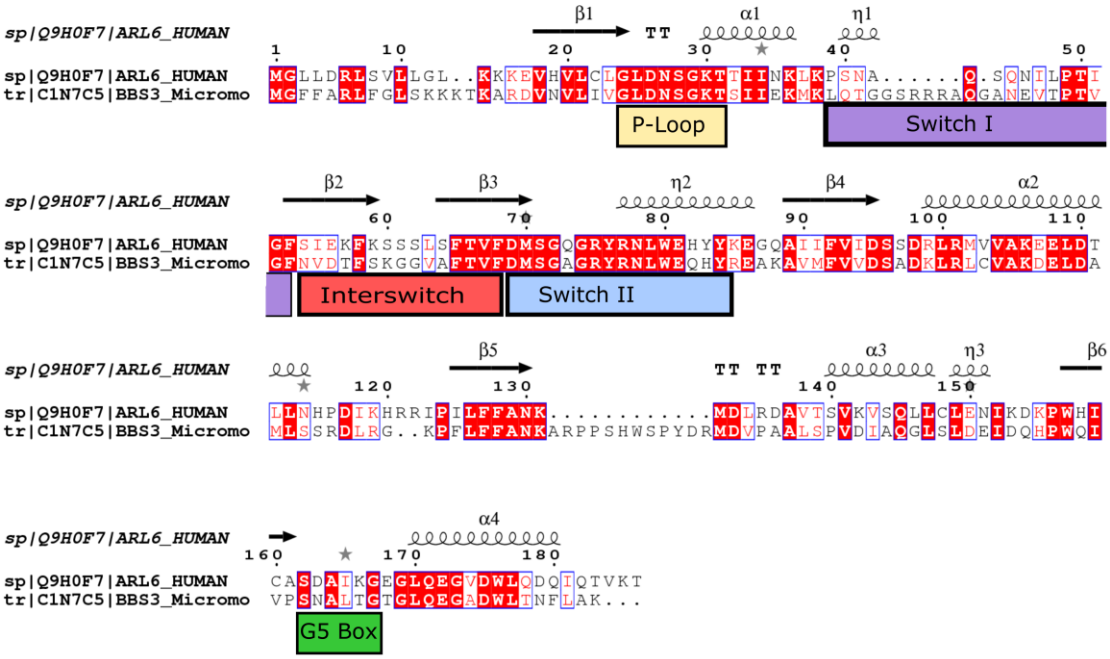


Figure A.1 Sequences alignment of *M. pusilla* BBS3 with human BBS3 (important domain marked).

B.2 Materials and Methods

B.2.1 Expression of Bardet-biedl syndrome 3 protein (BBS3)

The clone of Bardet-biedl syndrome 3 (Gene ID: 9689316 and Protein ID: XP_003063712) of *Micromonas pusilla (strain CCMP1545) (Picoplanktonic green alga)* was gifted by Dr. Sureshkumar Ramasamy (Dr. Clemons Lab, Caltech). The positive clone was expressed in the BL21 Rosetta (DE3) cells. The cells were grown in the LB media (with chloramphenicol and kanamycin) and induced with IPTG for the overexpression of the protein and incubated at 16°C overnight. The cells were harvested and lysed with the lysis

buffer (Tris-Cl pH 7.5, 500 mM NaCl, and 1mM DTT). Protein was purified by the Ni-NTA affinity chromatography. Expression was confirmed by western blotting.

B.2.2 Refolding of Bardet-biedl syndrome 3 protein by on-column methods

B.2.3 Isolation of inclusion bodies

The bacterial cells were resuspended in lysis buffer (50 mM Tris-HCl, 10 mM $MgCl_2$ 0.1 % Tritron-X 100, 300 mM NaCl, 2 mM β-mercaptoethanol, pH 7.5) and sonicated twice for 5 minutes with 5 sec on/off pulse and 45% amplitude. After the cell disruption, inclusion bodies were isolated by centrifugation at 12000 rpm for 15 min at 4°C. Inclusion bodies that were sedimented as pellet washed twice with buffer I (50 mM Tris-HCl, 1 M urea, 0.1 % Tritron-X 100, 300 mM NaCl, 2 mM β-mercaptoethanol, pH 7.5) and repeated ultrasonication and centrifugation steps. A final was given with was buffer II (50 mM Tris-HCl, pH 7.5) and repeated ultrasonication and centrifugation steps. Isolated inclusion bodies resuspended in 1 ml wash buffer II.

B.2.4 Solubilisation of inclusion bodies

The isolated inclusion bodies were then solubilised in 30 ml of denaturation (solubilizing) buffer. After solubilisation inclusion bodies, the supernatant was obtained by centrifugation at 12000 rpm for 15 min at 4°C.

B.2.5 On column refolding

The supernatant (solubilised inclusion bodies) with Ni-NTA (pre-equilibrated in denaturation buffer) beads were incubated for overnight for binding at 4°C with shaking. The resin was packed into a column. To remove contaminant protein column was washed with 10 times column volume by denaturation was buffer III (20 mM Tris-HCl, 8 M urea, 100 mM NaCl, 20 mM imidazole, 10 mM B-mercaptoethanol, pH 8). In the next step column was washed with 10 times column volume NaCl wash buffer IV (20 mM Tris-HCl, 0.1 % Tritron-X 100, 500 mM NaCl, 10 mM β-mercaptoethanol, pH 7.5). To remove the detergent and allow the protein to refold column was washed with 10 times column volume β-cyclodextrin wash buffer AII (20 mM Tris-HCl, 5 mM β-cyclodextrin, 100 mM NaCl, 10 mM β-mercaptoethanol, pH 7.5). To remove β-cyclodextrin column was washed with 10 times column volume wash buffer AIII (20 mM Tris-HCl, 100 mM NaCl, 10 mM β-mercaptoethanol, pH 7.5). In the last step, the Refolded protein was eluted with buffer A (20 mM Tris-HCl, 100 mM NaCl, 300 mM imidazole, 10 mM β-mercaptoethanol, pH 7.5). Eluted protein was dialyzed against the dialysis buffer (20 mM Tris-HCl, 100 mM NaCl, 2

mM β-mercaptoethanol, pH 7.5) in remove imidazole. Protein refolding was confirmed by CD.

B.3 Results

B.3.1 Expression and purification of Bardet-biedl syndrome protein

The clone of *BBS3* was expressed in Bl21 rosseta (DE3) expression host & protein was expressing in the inclusion bodies. To produce CutC protein in the soluble form various parameters like, IPTG concentration, temperature and incubation time were tried (Figure 3.14 A: lane 2).

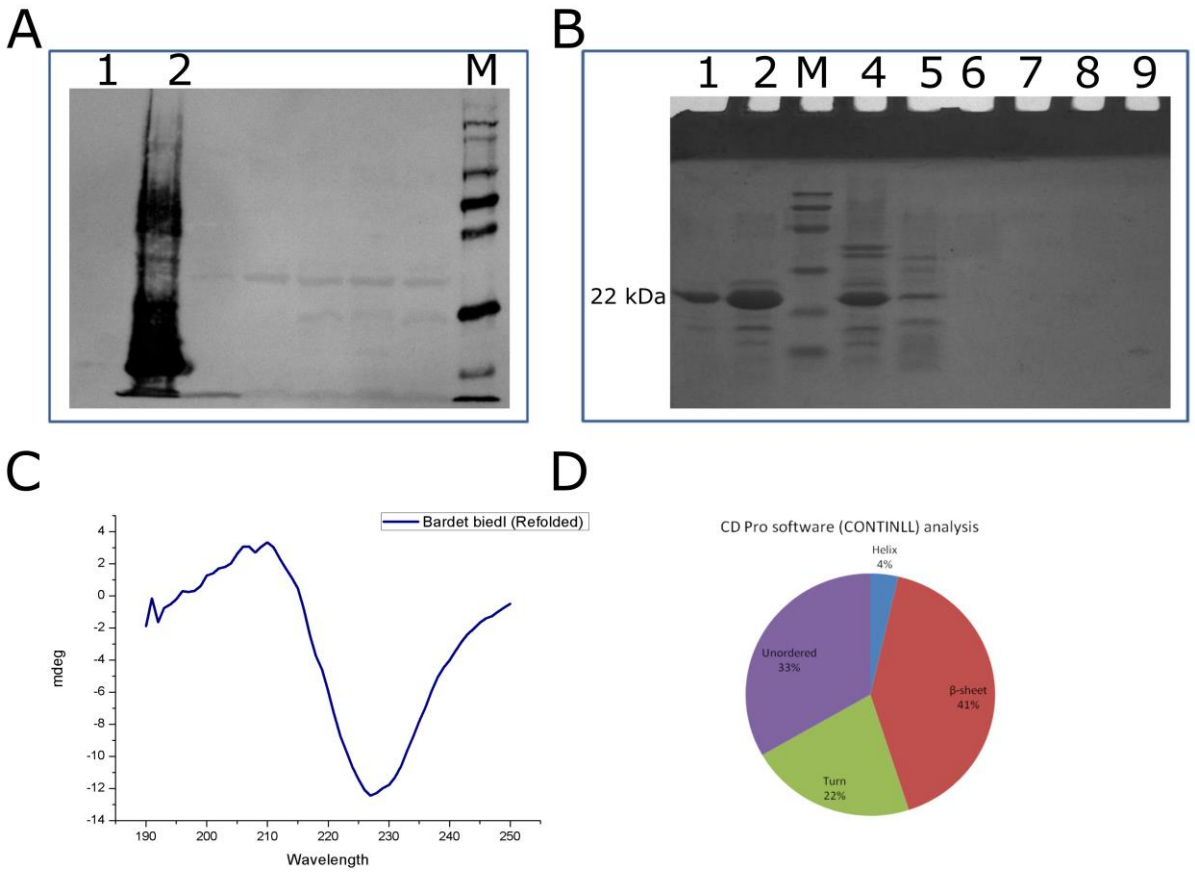


Figure A.1: Purification of BBS3 protein. A: Western blotting analysis confirms BBS3 protein expression with monoclonal anti-His antibodies. Lane 1, Supernatant after lysis of BBS3; Lane 2, Pellet fraction of after lysis of BBS3; M, Protein ladder. B: Modified column refolding methods. Lane 1, Eluted BBS3 (soluble); Lane 2, Pellet fraction of of BBS3 (insoluble) ; M, Protein ladder, Lane 4, Flowthrough; Lane 5, Wash (denaturation was buffer III); Lane 6, Wash (NaCl wash buffer IV); Lane 7, Wash (β-cyclodextrin wash buffer AII); Lane 8, Wash (buffer A). C: Far-UV CD scan of BBS3 at pH 7.0 (0.4435 mg ml^{-1}). D: CD Pro software analysis of BBS3.

Finally protein was successfully solubilized using refolding of Bardet-biedl syndrome 3 by on-column methods. Refolding of Bardet-biedl syndrome 3 analysed by CD experiment. From the spectra of CD and CDpro analysis concluded that protein refolded properly.

Bibliography

Ahern, G. P. (2003) 'Activation of TRPV1 by the satiety factor oleoylethanolamide.', *The Journal of biological chemistry*, 278(33), pp. 30429–34.

Aktories, K. (1997) 'Identification of the catalytic site of clostridial ADP-ribosyltransferases.', *Advances in experimental medicine and biology*, 419, pp. 53–60.

Al, B. J. M. *et al.* (1989) 'Properties of acid ceramidase from human spleen', *Biochimica et Biophysica Acta (BBA)/Lipids and Lipid Metabolism*, 1004(2), pp. 245–251.

Alhouayek, M. *et al.* (2017) 'N-acylethanolamine-hydrolyzing acid amidase and fatty acid amide hydrolase inhibition differentially affect N-acylethanolamine levels and macrophage activation.', *Biochimica et biophysica acta. Molecular and cell biology of lipids*, 1862(5), pp. 474–484.

Alhouayek, M. and Muccioli, G. G. (2014) 'Harnessing the anti-inflammatory potential of palmitoylethanolamide', *Drug Discovery Today*, 19(10), pp. 1632–1639.

Alves, M. Q. *et al.* (2013) 'Molecular basis of acid ceramidase deficiency in a neonatal form of Farber disease: identification of the first large deletion in ASAH1 gene.', *Molecular genetics and metabolism*, 109(3), pp. 276–81.

Arroyo, M. *et al.* (2003) 'Biotechnological applications of penicillin acylases: state-of-the-art', *Applied Microbiology and Biotechnology*, 60(5), pp. 507–514.

Augé, N. *et al.* (1999) 'Role of Sphingosine 1-Phosphate in the Mitogenesis Induced by Oxidized Low Density Lipoprotein in Smooth Muscle Cells via Activation of Sphingomyelinase, Ceramidase, and Sphingosine Kinase', *Journal of Biological Chemistry*, 274(31), pp. 21533–21538.

Avery, K. *et al.* (2008) 'Sphingosine-1-phosphate mediates transcriptional regulation of key

targets associated with survival, proliferation, and pluripotency in human embryonic stem cells.', *Stem cells and development*, 17(6), pp. 1195–205.

Azuma, N. *et al.* (1994) 'Stimulation of acid ceramidase activity by saposin D.', *Archives of biochemistry and biophysics*, 311(2), pp. 354–7.

Bach, A. *et al.* (2015) 'Benzoxazolone Carboxamides as Potent Acid Ceramidase Inhibitors: Synthesis and Structure-Activity Relationship (SAR) Studies.', *Journal of medicinal chemistry*, 58(23), pp. 9258–72.

Beckham, T. H. *et al.* (2012) 'Acid ceramidase-mediated production of sphingosine 1-phosphate promotes prostate cancer invasion through upregulation of cathepsin B.', *International journal of cancer*, 131(9), pp. 2034–43.

Beckham, T. H. *et al.* (2013) 'Acid ceramidase promotes nuclear export of PTEN through sphingosine 1-phosphate mediated Akt signaling.', *PloS one*. Edited by X. Zi, 8(10), p. e76593.

Adzhubei, I., Jordan, D. M. and Sunyaev, S. R. (2013) 'Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2', *Current Protocols in Human Genetics*, 76(1), pp. 7.20.1-7.20.41.

Altschul, S. F. *et al.* (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410.

Alves, M. Q. *et al.* (2013) 'Molecular basis of acid ceramidase deficiency in a neonatal form of Farber disease: identification of the first large deletion in ASAH1 gene.', *Molecular genetics and metabolism*, 109(3), pp. 276–81.

Amirhakimi, G. H. *et al.* (1976) 'Familial lipogranulomatosis (Farber's disease).', *Clinical genetics*, 9(6), pp. 625–30.

Ashkenazy, H. *et al.* (2016) 'ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules', *Nucleic Acids Research*, 44(W1), pp. W344–W350.

De Baets, G. *et al.* (2012) 'SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants', *Nucleic Acids Research*, 40(D1), pp. D935–D939.

Bär, J. *et al.* (2001) 'Molecular analysis of acid ceramidase deficiency in patients with Farber disease', *Human Mutation*, 17(3), pp. 199–209.

Bashyam, M. D. *et al.* (2014) 'Molecular analyses of novel *ASAH1* mutations causing Farber lipogranulomatosis: analyses of exonic splicing enhancer inactivating mutation', *Clinical Genetics*, 86(6), pp. 530–538.

Becker, O. M. *et al.* (2001) *Computational Biochemistry and Biophysics*. Edited by O. Becker, A. MacKerell Jr, and Benoît Roux. CRC Press.

Bedia, C. *et al.* (2008) 'Cytotoxicity and acid ceramidase inhibitory activity of 2-substituted aminoethanol amides.', *Chemistry and physics of lipids*, 156(1–2), pp. 33–40.

Bedia, C. *et al.* (2010) 'A simple fluorogenic method for determination of acid ceramidase activity and diagnosis of Farber disease.', *Journal of lipid research*. American Society for Biochemistry and Molecular Biology, 51(12), pp. 3542–7.

Bektas, M. *et al.* (2005) 'Sphingosine kinase activity counteracts ceramide-mediated cell death in human melanoma cells: role of Bcl-2 expression', *Oncogene*, 24(1), pp. 178–187.

Berdyshev, E. V *et al.* (1997) 'Influence of fatty acid ethanolamides and delta9-tetrahydrocannabinol on cytokine and arachidonate release by mononuclear cells.', *European journal of pharmacology*, 330(2–3), pp. 231–40.

Ben-Yoseph Y, Gagné R, Parvathy MR, Mitchell DA, M. T. (1989) 'Leukocyte and plasma

N-laurylsphingosine deacylase (ceramidase) in Farber disease. - PubMed - NCBI', *clin genet*, 36(1), pp. 38–42.

Berendsen, H. J. C. *et al.* (1981) 'Interaction Models for Water in Relation to Protein Hydration', in. Springer, Dordrecht, pp. 331–342.

Berendsen, H. J. C. *et al.* (1984) 'Molecular dynamics with coupling to an external bath', *The Journal of Chemical Physics*. American Institute of Physics, 81(8), pp. 3684–3690.

Bernardo, K. *et al.* (1995) 'Purification, characterization, and biosynthesis of human acid ceramidase.', *The Journal of biological chemistry*, 270(19), pp. 11098–102.

Berndt, N. *et al.* (2013) 'Akt2 and acid ceramidase cooperate to induce cell invasion and resistance to apoptosis.', *Cell cycle (Georgetown, Tex.)*, 12(13), pp. 2024–32.

Bhagwat, M. (2010) 'Searching NCBI's dbSNP Database', in *Current Protocols in Bioinformatics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Bielawska AE, Shapiro JP, Jiang L, Melkonyan HS, Piot C, Wolfe CL, Tomei LD, Hannun YA, U. S. (1997) 'Ceramide is involved in triggering of cardiomyocyte apoptosis induced by ischemia and reperfusion. - PubMed - NCBI', *Am J Pathol*, 151(5), pp. 1257–63.

Blom, N., Gammeltoft, S. and Brunak, S. (1999) 'Sequence and structure-based prediction of eukaryotic protein phosphorylation sites', *Journal of Molecular Biology*, 294(5), pp. 1351–1362.

Bolton, E. E. *et al.* (2008) 'PubChem: Integrated Platform of Small Molecules and Biological Activities', *Annual Reports in Computational Chemistry*. Elsevier, 4, pp. 217–241.

Bonafé, L. *et al.* (2016) 'Brief Report: Peripheral Osteolysis in Adults Linked to *ASAH1* (Acid Ceramidase) Mutations: A New Presentation of Farber's Disease', *Arthritis & Rheumatology*, 68(9), pp. 2323–2327.

Bonezzi, F. T. *et al.* (2016) 'An Important Role for N-Acylethanolamine Acid Amidase in the Complete Freund's Adjuvant Rat Model of Arthritis.', *The Journal of pharmacology and experimental therapeutics*, 356(3), pp. 656–63.

Bourbon, N. A., Sandirasegarane, L. and Kester, M. (2002) 'Ceramide-induced inhibition of Akt is mediated through protein kinase Czeta: implications for growth arrest.', *The Journal of biological chemistry*, 277(5), pp. 3286–92.

Bracey, M. H. (2002) 'Structural Adaptations in a Membrane Enzyme That Terminates Endocannabinoid Signaling', *Science*, 298(5599), pp. 1793–1796.

Brann, A. B. *et al.* (2002) 'Nerve growth factor-induced p75-mediated death of cultured hippocampal neurons is age-dependent and transduced through ceramide generated by neutral sphingomyelinase.', *The Journal of biological chemistry*, 277(12), pp. 9812–8.

Brannigan, J. A. *et al.* (1995) 'A protein catalytic framework with an N-terminal nucleophile is capable of self-activation', *Nature*, 378(6555), pp. 416–419.

Burek, C. *et al.* (2001) 'The role of ceramide in receptor- and stress-induced apoptosis studied in acidic ceramidase-deficient Farber disease cells.', *Oncogene*, 20(45), pp. 6493–502.

Burley, S. K. *et al.* (1999) 'Structural genomics: beyond the Human Genome Project', *Nature Genetics*, 23(2), pp. 151–157.

Butler, A. *et al.* (2002) 'Reproductive Pathology and Sperm Physiology in Acid Sphingomyelinase-Deficient Mice', *The American Journal of Pathology*, 161(3), pp. 1061–1075.

Calabrese, R. *et al.* (2009) 'Functional annotations improve the predictive score of human disease-related mutations in proteins', *Human Mutation*, 30(8), pp. 1237–1244.

Cai, G. *et al.* (2004) 'Cloning, Overexpression, and Characterization of a Novel Thermostable Penicillin G Acylase from Achromobacter xylosoxidans: Probing the Molecular Basis for Its High Thermostability', *Applied and Environmental Microbiology*, 70(5), pp. 2764–2770.

Calignano, A., La Rana, G. and Piomelli, D. (2001) 'Antinociceptive activity of the endogenous fatty acid amide, palmitylethanolamide.', *European journal of pharmacology*, 419(2–3), pp. 191–8.

Capriotti, E. and Fariselli, P. (2017) 'PhD-SNPg: a webserver and lightweight tool for scoring single nucleotide variants.', *Nucleic acids research*. Oxford University Press, 45(W1), pp. W247–W252.

Capriotti, E., Fariselli, P. and Casadio, R. (2005) 'I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure', *Nucleic Acids Research*, 33(Web Server), pp. W306–W310.

Cartigny, B. *et al.* (1985) 'Clinical diagnosis of a new case of ceramidase deficiency (Farber's disease).', *Journal of inherited metabolic disease*, 8(1), p. 8.

Chatelut, M. *et al.* (1997) 'Model SV40-transformed fibroblast lines for metabolic studies of human prosaposin and acid ceramidase deficiencies.', *Clinica chimica acta; international journal of clinical chemistry*, 262(1–2), pp. 61–76.

Chavez, J. A. *et al.* (2005) 'Acid ceramidase overexpression prevents the inhibitory effects of saturated fatty acids on insulin signaling.', *The Journal of biological chemistry*, 280(20), pp. 20148–53.

Chen, W. W., Moser, A. B. and Moser, H. W. (1981) 'Role of lysosomal acid ceramidase in the metabolism of ceramide in human skin fibroblasts.', *Archives of biochemistry and biophysics*, 208(2), pp. 444–55.

Chien, C. T. (2009) *Overexpression of human acid ceramidase precursor and variants of the catalytic center in Sf9 cells : Analysis of ceramidase maturation, autocatalytic processing and interaction with Sap-D*.

Chedrawi, A. K. *et al.* (2012) 'Novel V97G ASAH1 mutation found in Farber disease patients: Unique appearance of the disease with an intermediate severity, and marked early involvement of central and peripheral nervous system', *Brain and Development*, 34(5), pp. 400–404.

Choi, Y. and Chan, A. P. (2015) 'PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels.', *Bioinformatics (Oxford, England)*. Oxford University Press, 31(16), pp. 2745–7.

Colovos, C. and Yeates, T. O. (1993) 'Verification of protein structures: Patterns of nonbonded atomic interactions', *Protein Science*, 2(9), pp. 1511–1519.

Cvitanovic-Sojat, L. *et al.* (2011) 'Farber lipogranulomatosis type 1 – Late presentation and early death in a Croatian boy with a novel homozygous ASAH1 mutation', *European Journal of Paediatric Neurology*, 15(2), pp. 171–173.

Darden, T. *et al.* (1999) 'New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations.', *Structure (London, England : 1993)*, 7(3), pp. R55-60.

Dehouck, Y. *et al.* (2011) 'PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality', *BMC Bioinformatics*, 12(1), p. 151.

Dementiev, A. *et al.* (2019) 'Molecular Mechanism of Inhibition of Acid Ceramidase by Carmofur', *Journal of Medicinal Chemistry*, 62(2), pp. 987–992.

Devi, A. R. R. *et al.* (2006) 'Farber lipogranulomatosis: clinical and molecular genetic analysis reveals a novel mutation in an Indian family', *Journal of Human Genetics*, 51(9), pp.

811–814.

Deng, X., Gao, F. and May, W. S. (2009) ‘Protein phosphatase 2A inactivates Bcl2’s antiapoptotic function by dephosphorylation and up-regulation of Bcl2-p53 binding.’, *Blood*, 113(2), pp. 422–8.

Duggleby, H. J. *et al.* (1995) ‘Penicillin acylase has a single-amino-acid catalytic centre’, *Nature*, 373(6511), pp. 264–268.

Duranti, A. *et al.* (2012) ‘N-(2-oxo-3-oxetanyl)carbamic acid esters as N-acylethanolamine acid amidase inhibitors: synthesis and structure-activity and structure-property relationships.’, *Journal of medicinal chemistry*, 55(10), pp. 4824–36.

Dyment, D. A. *et al.* (2014) ‘Evidence for clinical, genetic and biochemical variability in spinal muscular atrophy with progressive myoclonic epilepsy’, *Clinical Genetics*, 86(6), pp. 558–563. Van Durme, J. *et al.* (2011) ‘A graphical interface for the FoldX forcefield’, *Bioinformatics*, 27(12), pp. 1711–1712.

Edgar, R. C. (2004) ‘MUSCLE: multiple sequence alignment with high accuracy and high throughput’, *Nucleic Acids Research*, 32(5), pp. 1792–1797.

Ekici, O. D., Paetzel, M. and Dalbey, R. E. (2008) ‘Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration.’, *Protein science : a publication of the Protein Society*, 17(12), pp. 2023–37.

Eliyahu, E. *et al.* (2007) ‘Acid ceramidase is a novel factor required for early embryo survival.’, *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 21(7), pp. 1403–9.

Eliyahu, E. *et al.* (2011) ‘Identification of Cystatin SA as a Novel Inhibitor of Acid Ceramidase’, *Journal of Biological Chemistry*, 286(41), pp. 35624–35633.

Elojeimy, S. *et al.* (2006) 'New insights on the use of desipramine as an inhibitor for acid ceramidase.', *FEBS letters*, 580(19), pp. 4751–6.

Erarslan, A. and Koçer, H. (1992) 'Thermal inactivation kinetics of penicillin G acylase obtained from a mutant derivative of Escherichia coli ATCC 11105.', *Journal of chemical technology and biotechnology (Oxford, Oxfordshire : 1986)*, 55(1), pp. 79–84.

Facci, L. *et al.* (1995) 'Mast cells express a peripheral cannabinoid receptor with differential sensitivity to anandamide and palmitoylethanolamide.', *Proceedings of the National Academy of Sciences*, 92(8), pp. 3376–3380.

FARBER, S., COHEN, J. and UZMAN, L. L. (1957) 'Lipogranulomatosis; a new lipo-glycoprotein storage disease.', *Journal of the Mount Sinai Hospital, New York*, 24(6), pp. 816–37.

Fensom, A. H. *et al.* (1979) 'Prenatal diagnosis of Farber's disease.', *Lancet (London, England)*, 2(8150), pp. 990–2.

Ferlinz, K. *et al.* (2001) 'Human acid ceramidase: processing, glycosylation, and lysosomal targeting.', *The Journal of biological chemistry*, 276(38), pp. 35352–60.

Ferrara, P., Apostolakis, J. and Caflisch, A. (2002) 'Evaluation of a fast implicit solvent model for molecular dynamics simulations.', *Proteins*, 46(1), pp. 24–33.

Filosto, M. *et al.* (2016) 'ASAH1 variant causing a mild SMA phenotype with no myoclonic epilepsy: a clinical, biochemical and molecular study', *European Journal of Human Genetics*, 24(11), pp. 1578–1583.

Friesner, R. A. *et al.* (2006) 'Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes.', *Journal of medicinal chemistry*, 49(21), pp. 6177–96.

Gabaldón, T. and Koonin, E. V. (2013) 'Functional and evolutionary implications of gene orthology', *Nature Reviews Genetics*, 14(5), pp. 360–366.

Gangoiti, P. *et al.* (2008) 'Ceramide 1-phosphate stimulates macrophage proliferation through activation of the PI3-kinase/PKB, JNK and ERK1/2 pathways.', *Cellular signalling*, 20(4), pp. 726–36.

Gangoiti, P. *et al.* (2010) 'Control of metabolism and signaling of simple bioactive sphingolipids: Implications in disease', *Progress in Lipid Research*, 49(4), pp. 316–334.

Gangoiti, P. *et al.* (2012) 'Ceramide 1-phosphate stimulates proliferation of C2C12 myoblasts.', *Biochimie*, 94(3), pp. 597–607.

GATT, S. (1963) 'ENZYMIC HYDROLYSIS AND SYNTHESIS OF CERAMIDES.', *The Journal of biological chemistry*, 238, pp. 3131–3.

Gebai, A. *et al.* (2018) 'Structural basis for the activation of acid ceramidase', *Nature Communications*, 9(1), p. 1621.

Gómez-Muñoz, A. *et al.* (2004) 'Ceramide-1-phosphate blocks apoptosis through inhibition of acid sphingomyelinase in macrophages.', *Journal of lipid research*, 45(1), pp. 99–105.

Gorelik, A. *et al.* (2018) 'Molecular mechanism of activation of the immunoregulatory amidase NAAA', *Proceedings of the National Academy of Sciences*, 115(43), pp. E10032–E10040.

Gouazé-Andersson, V. *et al.* (2011) 'Inhibition of acid ceramidase by a 2-substituted aminoethanol amide synergistically sensitizes prostate cancer cells to N-(4-hydroxyphenyl) retinamide.', *The Prostate*, 71(10), pp. 1064–73.

Granado, M. H. *et al.* (2009) 'Ceramide 1-phosphate inhibits serine palmitoyltransferase and blocks apoptosis in alveolar macrophages', *Biochimica et Biophysica Acta (BBA) - Molecular*

and Cell Biology of Lipids, 1791(4), pp. 263–272. doi: 10.1016/j.bbalip.2009.01.023.

Gulaya, N. M. *et al.* (1998) 'Long-chain N-acylethanolamines inhibit lipid peroxidation in rat liver mitochondria under acute hypoxic hypoxia.', *Chemistry and physics of lipids*, 97(1), pp. 49–54.

Gupta, R., Jung, E. and Brunak, S. (2016) 'Prediction of N-glycosylation sites in human proteins. 2004', *Ref Type: Unpublished Work*.

Gutteridge, A. and Thornton, J. M. (2005) 'Understanding nature's catalytic toolkit.', *Trends in biochemical sciences*, 30(11), pp. 622–9.

von Haefen, C. *et al.* (2002) 'Ceramide induces mitochondrial activation and apoptosis via a Bax-dependent pathway in human carcinoma cells.', *Oncogene*, 21(25), pp. 4009–19.

Haliloglu, G. *et al.* (2002) 'Spinal muscular atrophy with progressive myoclonic epilepsy: report of new cases and review of the literature.', *Neuropediatrics*, 33(6), pp. 314–9.

Hara, S. *et al.* (2004) 'p53-Independent ceramide formation in human glioma cells during gamma-radiation-induced apoptosis.', *Cell death and differentiation*, 11(8), pp. 853–61.

He, X. *et al.* (1999) 'A Fluorescence-Based High-Performance Liquid Chromatographic Assay to Determine Acid Ceramidase Activity', *Analytical Biochemistry*, 274(2), pp. 264–269.

He, X. *et al.* (2003) 'Purification and characterization of recombinant, human acid ceramidase. Catalytic reactions and interactions with acid sphingomyelinase.', *The Journal of biological chemistry*, 278(35), pp. 32978–86.

Heinrich, M. *et al.* (2004) 'Cathepsin D links TNF-induced acid sphingomyelinase to Bid-mediated caspase-9 and -3 activation.', *Cell death and differentiation*, 11(5), pp. 550–63.

Hess, B. *et al.* (1997) 'LINCS: A linear constraint solver for molecular simulations', *Journal of Computational Chemistry*. John Wiley & Sons, Ltd, 18(12), pp. 1463–1472.

Hess, B. *et al.* (2008) 'GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation', *Journal of Chemical Theory and Computation*, 4(3), pp. 435–447.

HILL, A. V., HILL, A. and Hill, A. V. (1910) 'The possible effects of the aggregation of the molecules of haemoglobin on its oxygen dissociation curve'.

Holliday, G. L., Mitchell, J. B. O. and Thornton, J. M. (2009) 'Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis', *Journal of Molecular Biology*, 390(3), pp. 560–577.

Hong, S.-B. *et al.* (1999) 'Molecular Cloning and Characterization of a Human cDNA and Gene Encoding a Novel Acid Ceramidase-like Protein', *Genomics*, 62(2), pp. 232–241.

Huang, Y. *et al.* (2004) 'Elevation of the level and activity of acid ceramidase in Alzheimer's disease brain.', *The European journal of neuroscience*, 20(12), pp. 3489–97.

Hussain, Z. *et al.* (2017) 'Mammalian enzymes responsible for the biosynthesis of N - acylethanolamines', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1862(12), pp. 1546–1561.

Jameson, R. A., Holt, P. J. and Keen, J. H. (1987) 'Farber's disease (lysosomal acid ceramidase deficiency).', *Annals of the rheumatic diseases*. BMJ Publishing Group, 46(7), pp. 559–61.

Jankovic, J. and Rivera, V. M. (1979) 'Hereditary myoclonus and progressive distal muscular atrophy.', *Annals of neurology*, 6(3), pp. 227–31.

Jan, J.-T., Chatterjee, S. and Griffin, D. E. (2000) 'Sindbis Virus Entry into Cells Triggers

Apoptosis by Activating Sphingomyelinase, Leading to the Release of Ceramide', *Journal of Virology*, 74(14), pp. 6425–6432.

Jankovic, J. and Rivera, V. M. (1979) 'Hereditary myoclonus and progressive distal muscular atrophy.', *Annals of neurology*, 6(3), pp. 227–31.

Jerabek-Willemsen, M. *et al.* (2011) 'Molecular Interaction Studies Using Microscale Thermophoresis', *ASSAY and Drug Development Technologies*, 9(4), pp. 342–353.

Karp, P. D. (1998) 'What we do not know about sequence analysis and sequence databases', *Bioinformatics*, 14(9), pp. 753–754.

Al Jasmi, F. (2012) 'A novel mutation in an atypical presentation of the rare infantile Farber disease', *Brain and Development*, 34(6), pp. 533–535.

Kattner, E., Schäfer, A. and Harzer, K. (1997) 'Hydrops fetalis: manifestation in lysosomal storage diseases including Farber disease.', *European journal of pediatrics*, 156(4), pp. 292–5.

Kim, D. E., Chivian, D. and Baker, D. (2004) 'Protein structure prediction and analysis using the Robetta server', *Nucleic Acids Research*. Narnia, 32(Web Server), pp. W526–W531.

Kim, S. Y. *et al.* (2016) 'Atypical presentation of infantile-onset farber disease with novel *ASAH1* mutations', *American Journal of Medical Genetics Part A*, 170(11), pp. 3023–3027.

Kishimoto, Y., Hiraiwa, M. and O'Brien, J. S. (1992) 'Saposins: structure, function, distribution, and molecular genetics.', *Journal of lipid research*, 33(9), pp. 1255–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1402395 (Accessed: 10 June 2019).

Klausen, M. S. *et al.* (2019) 'NetSurfP- 2.0: Improved prediction of protein structural features by integrated deep learning', *Proteins: Structure, Function, and Bioinformatics*, 87(6), pp. 520–527.

Klein, A. *et al.* (1994) 'Sphingolipid activator protein D (sap-D) stimulates the lysosomal degradation of ceramide in vivo.', *Biochemical and biophysical research communications*, 200(3), pp. 1440–8.

Knorr, J., Rudolf, P. and Nuernberger, P. (2013) 'A comparative study on chirped-pulse upconversion and direct multichannel MCT detection', *Optics Express*, 21(25), p. 30693.

Ko, J. *et al.* (2012) 'GalaxyWEB server for protein structure prediction and refinement', *Nucleic Acids Research*. Narnia, 40(W1), pp. W294–W297.

Koch, J. *et al.* (1996) 'Molecular cloning and characterization of a full-length complementary DNA encoding human acid ceramidase. Identification Of the first molecular lesion causing Farber disease.', *The Journal of biological chemistry*, 271(51), pp. 33110–5.

Koehl, P. and Delarue, M. (1994) 'Polar and nonpolar atomic environments in the protein core: Implications for folding and binding', *Proteins: Structure, Function, and Genetics*, 20(3), pp. 264–278.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular Biology and Evolution*, 33(7), pp. 1870–1874.

Laskowski, R. A. *et al.* (1993) 'PROCHECK: a program to check the stereochemical quality of protein structures', *Journal of Applied Crystallography*. International Union of Crystallography (IUCr), 26(2), pp. 283–291.

Lauffer, L. M., Iakoubov, R. and Brubaker, P. L. (2009) 'GPR119 is essential for oleoylethanolamide-induced glucagon-like peptide-1 secretion from the intestinal enteroendocrine L-cell.', *Diabetes*, 58(5), pp. 1058–66.

Laurier-Laurin, M.-E. *et al.* (2014) 'Blockade of lysosomal acid ceramidase induces GluN2B-dependent Tau phosphorylation in rat hippocampal slices.', *Neural plasticity*, 2014,

p. 196812.

Lefebvre, S. *et al.* (1995) 'Identification and characterization of a spinal muscular atrophy-determining gene.', *Cell*, 80(1), pp. 155–65.

Letunic, I. and Bork, P. (2016) 'Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees', *Nucleic Acids Research*, 44(W1), pp. W242–W245.

Li, C.-M. *et al.* (1999) 'The Human Acid Ceramidase Gene (ASAH): Structure, Chromosomal Location, Mutation Analysis, and Expression', *Genomics*, 62(2), pp. 223–231.

Li, C.-M. *et al.* (2002) 'Insertional mutagenesis of the mouse acid ceramidase gene leads to early embryonic lethality in homozygotes and progressive lipid storage disease in heterozygotes.', *Genomics*, 79(2), pp. 218–24.

Li, C. M. *et al.* (1998) 'Cloning and characterization of the full-length cDNA and genomic sequences encoding murine acid ceramidase.', *Genomics*, 50(2), pp. 267–74.

Li, Y. *et al.* (2012) 'Design and synthesis of potent N-acylethanolamine-hydrolyzing acid amidase (NAAA) inhibitor as anti-inflammatory compounds.', *PloS one*. Edited by R. Seifert, 7(8), p. e43023.

Li, Y. *et al.* (2018) 'Inflammation-restricted anti-inflammatory activities of a N -acylethanolamine acid amidase (NAAA) inhibitor F215', *Pharmacological Research*, 132, pp. 7–14.

Li, Yuhang *et al.* (2017) 'Identification of highly potent N-acylethanolamine acid amidase (NAAA) inhibitors: Optimization of the terminal phenyl moiety of oxazolidone derivatives.', *European journal of medicinal chemistry*, 139, pp. 214–221.

Lindorff-Larsen, K. *et al.* (2010) 'Improved side-chain torsion potentials for the Amber

ff99SB protein force field', *Proteins: Structure, Function, and Bioinformatics*, 78(8), p. NA-NA.

Linke, T. *et al.* (2001) 'Interfacial regulation of acid ceramidase activity. Stimulation of ceramide degradation by lysosomal lipids and sphingolipid activator proteins.', *The Journal of biological chemistry*, 276(8), pp. 5760–8.

Liu, J., Beckman, B. S. and Foroozesh, M. (2013) 'A review of ceramide analogs as potential anticancer agents.', *Future medicinal chemistry*, 5(12), pp. 1405–21.

Lucki, N. C. *et al.* (2012) 'Acid Ceramidase (ASAH1) Represses Steroidogenic Factor 1-Dependent Gene Transcription in H295R Human Adrenocortical Cells by Binding to the Receptor', *Molecular and Cellular Biology*, 32(21), pp. 4419–4431.

Maceyka, M. *et al.* (2002) 'Sphingosine kinase, sphingosine-1-phosphate, and apoptosis.', *Biochimica et biophysica acta*, 1585(2–3), pp. 193–201.

Mallat, Z. and Tedgui, A. (2001) 'Current perspective on the role of apoptosis in atherothrombotic disease.', *Circulation research*, 88(10), pp. 998–1003.

Mazzari, S. *et al.* (1996) 'N-(2-hydroxyethyl)hexadecanamide is orally active in reducing edema formation and inflammatory hyperalgesia by down-modulating mast cell activation.', *European journal of pharmacology*, 300(3), pp. 227–36.

Migliore, M. *et al.* (2016) 'Second-Generation Non-Covalent NAAA Inhibitors are Protective in a Model of Multiple Sclerosis', *Angewandte Chemie International Edition*, 55(37), pp. 11193–11197.

Mitra, P. *et al.* (2006) 'Role of ABCC1 in export of sphingosine-1-phosphate from mast cells.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(44), pp. 16394–9.

Mizutani, N. *et al.* (2015) 'Increased acid ceramidase expression depends on upregulation of androgen-dependent deubiquitinases, USP2, in a human prostate cancer cell line, LNCaP.', *Journal of biochemistry*, 158(4), pp. 309–19.

Monick, M. M. *et al.* (2004) 'Cooperative prosurvival activity by ERK and Akt in human alveolar macrophages is dependent on high levels of acid ceramidase activity.', *Journal of immunology (Baltimore, Md. : 1950)*, 173(1), pp. 123–35.

Morad, S. A. F. *et al.* (2013) 'Novel off-target effect of tamoxifen — Inhibition of acid ceramidase activity in cancer cells', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*. Elsevier, 1831(12), pp. 1657–1664.

Morita, Y. *et al.* (2000) 'Oocyte apoptosis is suppressed by disruption of the acid sphingomyelinase gene or by sphingosine -1-phosphate therapy', *Nature Medicine*, 6(10), pp. 1109–1114.

Morrison, K. L. and Weiss, G. A. (2001) 'Combinatorial alanine-scanning.', *Current opinion in chemical biology*, 5(3), pp. 302–7.

Muramatsu, T. *et al.* (2002) 'Mutation analysis of the acid ceramidase gene in Japanese patients with Farber disease.', *Journal of inherited metabolic disease*, 25(7), pp. 585–92.

Nuzzi, A. *et al.* (2016) 'Potent α-amino-β-lactam carbamic acid ester as NAAA inhibitors. Synthesis and structure-activity relationship (SAR) studies.', *European journal of medicinal chemistry*, 111, pp. 138–59.

O'Sullivan, S. E. (2016) 'An update on PPAR activation by cannabinoids.', *British journal of pharmacology*, 173(12), pp. 1899–910.

Ogura, Y. *et al.* (2016) 'A calcium-dependent acyltransferase that produces N-acyl phosphatidylethanolamines', *Nature Chemical Biology*, 12(9), pp. 669–671.

Oinonen, C. and Rouvinen, J. (2000) 'Structural comparison of Ntn-hydrolases.', *Protein science : a publication of the Protein Society*, 9(12), pp. 2329–37.

Panigrahi, P. *et al.* (2014) 'An improved method for specificity annotation shows a distinct evolutionary divergence among the microbial enzymes of the cholylglycine hydrolase family.', *Microbiology (Reading, England)*, 160(Pt 6), pp. 1162–74.

Panigrahi, P. *et al.* (2015) 'Sequence and structure-based comparative analysis to assess, identify and improve the thermostability of penicillin G acylases.', *Journal of industrial microbiology & biotechnology*, 42(11), pp. 1493–506.

Park, J.-H. and Schuchman, E. H. (2006) 'Acid ceramidase and human disease', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1758(12), pp. 2133–2138.

Pauling, L. (1946) 'Molecular Architecture and Biological Reactions', *Chem. Eng. News*, 24(10), pp. 1375–1377.

Pavlopoulos, S. *et al.* (2018) 'Secretion, isotopic labeling and deglycosylation of N -acylethanolamine acid amidase for biophysical studies', *Protein Expression and Purification*, 145, pp. 108–117.

Payen, A. & Persoz, J. (1833) 'Memoire sur la Diastase, les principaux Produits de ses Reactions, et leurs applications aux arts industriels', *Ann. chim. phys*, 53, pp. 73–92.

Payne, S. G., Brindley, D. N. and Guilbert, L. J. (1999) 'Epidermal growth factor inhibits ceramide-induced apoptosis and lowers ceramide levels in primary placental trophoblasts.', *Journal of cellular physiology*, 180(2), pp. 263–70.

Pejaver, V. *et al.* (2017) 'MutPred2: inferring the molecular and phenotypic impact of amino acid variants', *bioRxiv*. Cold Spring Harbor Laboratory, p. 134981.

Petracca, R. *et al.* (2017) 'Novel activity-based probes for N-acylethanolamine acid amidase',

Chem. Commun., 53(86), pp. 11810–11813.

Petrosino, S. *et al.* (2015) 'Diacerein is a potent and selective inhibitor of palmitoylethanolamide inactivation with analgesic activity in a rat model of acute inflammatory pain.', *Pharmacological research*, 91, pp. 9–14.

Pizzirani, D. *et al.* (2013) 'Discovery of a new class of highly potent inhibitors of acid ceramidase: synthesis and structure-activity relationship (SAR).', *Journal of medicinal chemistry*, 56(9), pp. 3518–30.

Pizzirani, D. *et al.* (2015) 'Benzoxazolone carboxamides: potent and systemically active inhibitors of intracellular acid ceramidase.', *Angewandte Chemie (International ed. in English)*, 54(2), pp. 485–9.

Ponzano, S. *et al.* (2013) 'Synthesis and structure-activity relationship (SAR) of 2-methyl-4-oxo-3-oxetanylcarbamic acid esters, a class of potent N-acylethanolamine acid amidase (NAAA) inhibitors.', *Journal of medicinal chemistry*, 56(17), pp. 6917–34.

Proksch, D., Klein, J. J. and Arenz, C. (2011) 'Potent Inhibition of Acid Ceramidase by Novel B-13 Analogues', *Journal of Lipids*, 2011, pp. 1–8.

Pyne, N. J. and Pyne, S. (2010) 'Sphingosine 1-phosphate and cancer.', *Nature reviews. Cancer*, 10(7), pp. 489–503.

Raisova, M. *et al.* (2002) 'Bcl-2 overexpression prevents apoptosis induced by ceramidase inhibitors in malignant melanoma and HaCaT keratinocytes.', *FEBS letters*, 516(1–3), pp. 47–52.

Ramsubir, S. *et al.* (2008) 'In vivo delivery of human acid ceramidase via cord blood transplantation and direct injection of lentivirus as novel treatment approaches for Farber disease.', *Molecular genetics and metabolism*, 95(3), pp. 133–41.

Rawlings, N. D. *et al.* (2014) 'MEROPS: the database of proteolytic enzymes, their substrates and inhibitors.', *Nucleic acids research*. Oxford University Press, 42(Database issue), pp. D503-9.

Realini, N. *et al.* (2013) 'Discovery of highly potent acid ceramidase inhibitors with in vitro tumor chemosensitizing activity.', *Scientific reports*, 3(1), p. 1035.

Ren, J. *et al.* (2017) 'Design, synthesis, and biological evaluation of oxazolidone derivatives as highly potent N-acylethanolamine acid amidase (NAAA) inhibitors', *RSC Advances*. The Royal Society of Chemistry, 7(21), pp. 12455–12463.

Ribeiro, A. *et al.* (2015) 'A Potent Systemically Active *N* -Acylethanolamine Acid Amidase Inhibitor that Suppresses Inflammation and Human Macrophage Activation', *ACS Chemical Biology*, 10(8), pp. 1838–1846.

Rossocha, M. *et al.* (2005) 'Conjugated bile acid hydrolase is a tetrameric N-terminal thiol hydrolase with specific recognition of its cholyl but not of its tauryl product.', *Biochemistry*, 44(15), pp. 5739–48.

Samsel, L. *et al.* (2004) 'The ceramide analog, B13, induces apoptosis in prostate cancer cell lines and inhibits tumor growth in prostate cancer xenografts.', *The Prostate*, 58(4), pp. 382–93.

Sasso, O. *et al.* (2013) 'Antinociceptive effects of the N-acylethanolamine acid amidase inhibitor ARN077 in rodent pain models.', *Pain*, 154(3), pp. 350–60.

Sasso, O. *et al.* (2018) 'The N-Acylethanolamine Acid Amidase Inhibitor ARN077 Suppresses Inflammation and Pruritus in a Mouse Model of Allergic Dermatitis', *Journal of Investigative Dermatology*, 138(3), pp. 562–569.

Sathe, S. and Pearson, T. (2014) 'Phenotypic characterization of the spinal muscular atrophy with progressive myoclonus epilepsy syndrome caused by ASAH1 mutations', *Molecular*

Genetics and Metabolism, 111(2), p. S93.

Saturnino, C. *et al.* (2010) 'Synthesis and biological evaluation of new potential inhibitors of N-acylethanolamine hydrolyzing acid amidase', *Bioorganic & Medicinal Chemistry Letters*, 20(3), pp. 1210–1213.

Sauer, B. *et al.* (2005) 'Sphingosine 1-phosphate is involved in cytoprotective actions of calcitriol in human fibroblasts and enhances the intracellular Bcl-2/Bax rheostat.', *Die Pharmazie*, 60(4), pp. 298–304.

Schnabel, D. *et al.* (1992) 'Simultaneous deficiency of sphingolipid activator proteins 1 and 2 is caused by a mutation in the initiation codon of their common gene.', *The Journal of biological chemistry*, 267(5), pp. 3312–5.

Schulze, H., Schepers, U. and Sandhoff, K. (2007) 'Overexpression and mass spectrometry analysis of mature human acid ceramidase.', *Biological chemistry*, 388(12), pp. 1333–43.

Seelan, R. S. *et al.* (2000) 'Human acid ceramidase is overexpressed but not mutated in prostate cancer.', *Genes, chromosomes & cancer*, 29(2), pp. 137–46.

Ségui, B. *et al.* (2000) 'Stress-induced apoptosis is not mediated by endolysosomal ceramide.', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 14(1), pp. 36–47.

Selzner, M. *et al.* (2001) 'Induction of apoptotic cell death and prevention of tumor growth by ceramide analogues in metastatic human colon cancer.', *Cancer research*, 61(3), pp. 1233–40.

Shapiro, L. and Harris, T. (2000) 'Finding function through structural genomics.', *Current opinion in biotechnology*, 11(1), pp. 31–5.

Shtraizent, N. *et al.* (2008) 'Autoproteolytic Cleavage and Activation of Human Acid

Ceramidase', *Journal of Biological Chemistry*, 283(17), pp. 11253–11259.

Skaper, S. D. *et al.* (1996) 'The ALIAmide palmitoylethanolamide and cannabinoids, but not anandamide, are protective in a delayed postglutamate paradigm of excitotoxic death in cerebellar granule neurons.', *Proceedings of the National Academy of Sciences*, 93(9), pp. 3984–3989.

Solorzano, C. *et al.* (2009) 'Selective N-acylethanolamine-hydrolyzing acid amidase inhibition reveals a key role for endogenous palmitoylethanolamide in inflammation.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(49), pp. 20966–71.

Sreerama, N. and Woody, R. W. (2000) 'Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set', *Analytical Biochemistry*, 287(2), pp. 252–260. Steentoft, C. *et al.* (2013) 'Precision mapping of the human O- GalNAc glycoproteome through SimpleCell technology', *The EMBO Journal*. EMBO Press, 32(10), pp. 1478–1488.

Stenson, P. D. *et al.* (2017) 'The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies.', *Human genetics*. Springer, 136(6), pp. 665–677.

Stratford, S. *et al.* (2004) 'Regulation of Insulin Action by Ceramide', *Journal of Biological Chemistry*, 279(35), pp. 36608–36615.

Sugita, M. *et al.* (1975) 'Ceramidase and ceramide synthesis in human kidney and cerebellum. Description of a new alkaline ceramidase.', *Biochimica et biophysica acta*, 398(1), pp. 125–31.

Sumner, J. (1926) 'The isolation and crystallization of the enzyme urease', *J. Biol. Chem*, 69, pp. 435–441.

Thomas, P. D. *et al.* (2003) 'PANTHER: a library of protein families and subfamilies indexed by function.', *Genome research*. Cold Spring Harbor Laboratory Press, 13(9), pp. 2129–41.

Tsuboi, K. *et al.* (2005) 'Molecular Characterization of N-Acylethanolamine-hydrolyzing Acid Amidase, a Novel Member of the Choloylglycine Hydrolase Family with Structural and Functional Similarity to Acid Ceramidase', *Journal of Biological Chemistry*, 280(12), pp. 11082–11092.

Tsuboi, K., Takezaki, N. and Ueda, N. (2007) 'The N-acylethanolamine-hydrolyzing acid amidase (NAAA).', *Chemistry & biodiversity*, 4(8), pp. 1914–25.

Ueda, N. *et al.* (1999) 'An acid amidase hydrolyzing anandamide as an endogenous ligand for cannabinoid receptors.', *FEBS letters*, 454(3), pp. 267–70.

Ueda, N. *et al.* (2000) 'The fatty acid amide hydrolase (FAAH).', *Chemistry and physics of lipids*, 108(1–2), pp. 107–21.

Ueda, N., Yamanaka, K. and Yamamoto, S. (2001) 'Purification and characterization of an acid amidase selective for N-palmitoylethanolamine, a putative endogenous anti-inflammatory substance.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 276(38), pp. 35552–7.

Urquhart, P., Nicolaou, A. and Woodward, D. F. (2015) 'Endocannabinoids and their oxygenation by cyclo-oxygenases, lipoxygenases and other oxygenases.', *Biochimica et biophysica acta*, 1851(4), pp. 366–76.

Vaser, R. *et al.* (2016) 'SIFT missense predictions for genomes', *Nature Protocols*, 11(1), pp. 1–9.

Venselaar, H. *et al.* (2010) 'Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces', *BMC Bioinformatics*, 11(1), p. 548.

Varshney, N. K. *et al.* (2012) 'Crystallization and X-ray structure analysis of a thermostable penicillin G acylase from Alcaligenes faecalis.', *Acta crystallographica. Section F, Structural biology and crystallization communications*. International Union of Crystallography, 68(Pt 3), pp. 273–7.

Verhaert, R. M. *et al.* (1997) 'Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from Alcaligenes faecalis.', *Applied and environmental microbiology*, 63(9), pp. 3412–8.

Wang, G. *et al.* (2005) 'Direct binding to ceramide activates protein kinase Czeta before the formation of a pro-apoptotic complex with PAR-4 in differentiating stem cells.', *The Journal of biological chemistry*, 280(28), pp. 26415–24.

Wang, J. *et al.* (2008) 'Amino acid residues crucial in pH regulation and proteolytic activation of N-acylethanolamine-hydrolyzing acid amidase.', *Biochimica et biophysica acta*, 1781(11–12), pp. 710–7.

Wen, P.-P. *et al.* (2016) 'Accurate *in silico* prediction of species-specific methylation sites based on information gain feature optimization', *Bioinformatics*, 32(20), pp. 3107–3115.

West, J. M. *et al.* (2012) 'Mass spectrometric characterization of human N-acylethanolamine-hydrolyzing acid amidase.', *Journal of proteome research*, 11(2), pp. 972–81.

Wiederstein, M. and Sippl, M. J. (2007) 'ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.', *Nucleic acids research*. Oxford University Press, 35(Web Server issue), pp. W407-10.

Wiens, C. J. *et al.* (2010) 'Bardet-Biedl Syndrome-associated Small GTPase ARL6 (BBS3) Functions at or near the Ciliary Gate and Modulates Wnt Signaling', *Journal of Biological Chemistry*, 285(21), pp. 16218–16230.

Yang, L. *et al.* (2015) 'Potential analgesic effects of a novel N-acylethanolamine acid

amidase inhibitor F96 through PPAR-α.', *Scientific reports*, 5(1), p. 13565.

Yoshimura, Y. *et al.* (2002) 'Molecular Cloning and Characterization of a Secretory Neutral Ceramidase of Drosophila melanogaster', *Journal of Biochemistry*, 132(2), pp. 229–236.

Young, T. L. *et al.* (1998) 'Canadian Bardet-Biedl syndrome family reduces the critical region of BBS3 (3p) and presents with a variable phenotype.', *American journal of medical genetics*, 78(5), pp. 461–7.

Yu, F. P. S. *et al.* (2018) 'Acid ceramidase deficiency: Farber disease and SMA-PME', *Orphanet Journal of Rare Diseases*, 13(1), p. 121.

Xue, Y. *et al.* (2008) 'GPS 2.0, a Tool to Predict Kinase-specific Phosphorylation Sites in Hierarchy', *Molecular & Cellular Proteomics*, 7(9), pp. 1598–1608.

Zeidan, Y. H. *et al.* (2006) 'Acid Ceramidase but Not Acid Sphingomyelinase Is Required for Tumor Necrosis Factor-α-induced PGE $_2$ Production', *Journal of Biological Chemistry*, 281(34), pp. 24695–24703.

Zeidan, Y. H. *et al.* (2008) 'Molecular targeting of acid ceramidase: implications to cancer therapy.', *Current drug targets*, 9(8), pp. 653–61.

Zeidan, Y. H. and Hannun, Y. A. (2007) 'Translational aspects of sphingolipid metabolism.', *Trends in molecular medicine*, 13(8), pp. 327–36.

Zhang, Z. *et al.* (2000) 'Human Acid Ceramidase Gene: Novel Mutations in Farber Disease', *Molecular Genetics and Metabolism*, 70(4), pp. 301–309.

Zhao, L.-Y. *et al.* (2007) 'Proteolytic activation and glycosylation of N-acylethanolamine-hydrolyzing acid amidase, a lysosomal enzyme involved in the endocannabinoid metabolism', *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1771(11), pp. 1397–1405.

Zhou, J. *et al.* (2012) 'Spinal Muscular Atrophy Associated with Progressive Myoclonic Epilepsy Is Caused by Mutations in ASAH1', *The American Journal of Human Genetics*, 91(1), pp. 5–14.

List of publications

1. **Rajput, V. B**., Karthikeyan, M. and Ramasamy, S. (2019) 'Zebrafish acid ceramidase: Expression in Pichia pastoris GS115 and biochemical characterization', *International Journal of Biological Macromolecules*, 122, pp. 587–593. doi: 10.1016/j.ijbiomac.2018.10.222.

2. **Rajput** *et al.*, 'Mouse NAAA: expression in *pichia pastoris* GS115 & biophysical characterization', (communicated).

3. **Rajput** *et al.*, '*In silico* analysis & modeling of pathogenic point mutations in ASAH1 causing Farber disease & spinal muscular atrophy with progressive myoclonic epilepsy (SMA-PME) ', Manuscript under communication.

4. **Rajput** *et al.*, 'Acid ceramidase and Acid ceramidase-like protein: structure function distribution & molecular genetics', Review in preparation.