Biophysical and Biochemical Investigation on the Structure, Function and Dynamics of Biomolecules Under Physiologically Relevant Environments

THESIS

SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY (SCIENCE)

IN MICROBIOLOGY

BY

PRITAM BISWAS

POST GRADUATE & RESEARCH DEPARTMENT OF

MICROBIOLOGY

ST. XAVIER'S COLLEGE

KOLKATA (AUTONOMOUS)

AFFILIATED TO THE UNIVERSITY OF CALCUTTA

2021

To My Father

Acknowledgments

"I can no other answer make but thanks,

And thanks; and ever thanks..."

William Shakespeare

As I sit down to pen my PhD dissertation, I reminisce of what has been an amazing journey. A journey made memorable by a group of outstanding and wonderful people to whom I owe a great debt of gratitude. Here, I take the opportunity to acknowledge them all for their unwavering support, motivation, and contribution not only to my work but also influencing me and my growth as a human being.

First of all, I would to like to express my deepest gratitude to my supervisor Dr. Sudeshna Shyam Choudhury for accepting me, providing me the opportunity to work and be a part of such an esteemed institution. I am grateful to her for letting me chase my dream of pursuing research. She gave me the freedom and the guidance to work and instilled her confidence on me, which helped me get through the initial days of my journey. Her enthusiasm and warmth touched me. She has been a constant source of motivation; her scientific inputs and constructive criticism have been immensely helpful towards the fruition of my work. She was a constant source of encouragement and inspiration. She was always there to help me, guide me and support me. Her profound knowledge and understanding of biological sciences reinstated my own confidence and steered me through this work. Her undeniable influence on me has left an indelible imprint on this thesis. I cannot thank her enough for bringing me under her wings and providing me with a platform where I could learn and grow. I am forever grateful to her. Thank you, ma'am for believing in me and giving me the chance to pursue my dreams.

I would like to express my heartfelt appreciation to my co-supervisor Prof. (Dr.) Samir Kumar Pal for believing in me and giving me the confidence to push forward in pursuit of my goals. I am ever so grateful to him for accepting me into his lab and providing me with a remarkable intellectual climate. I am grateful to him for his constant motivation, his belief in my abilities and the confidence he instilled. I found his enthusiastic personality infectious. His exemplary ability to constantly push the envelope and view science through every possible lens had a pronounced effect on me. His profound knowledge in inter-disciplinary field of science and technology and his unwavering commitment and uncompromising

attitude towards the highest standards of scientific study is what made me fully commit myself to this thesis. Throughout my journey I have never seen him take a leave of absence, always there for every one of his students. No lackadaisical approach, always pushing me to get better than I was yesterday. It is said that "Success is the natural consequence of consistently applying the basic fundamentals"- well I consider myself fortunate enough to have learnt the fundamentals of research from him. The guidance, the thought-provoking discussions, the constructive criticism, the encouragements and his unshakable belief in me, helped in achieving my goal. This thesis is an embodiment of everything he has taught and engrained in me. To describe his influence on me is beyond the scope of this piece of literature. Thank you, sir for believing in me and realizing a little nobody's dream of becoming somebody.

I am grateful to get the opportunity to actively work with many talented researchers from various national institutes. I would like to thank Prof. Ranjan Das, WBSU, India for his invaluable advices and for his critical review of my works. I would also like to thank Prof. Tanusri Saha-Dasgupta, SNBNCBS, India, Dr. Arpita Chattopadhyay, TINT, India, Dr. Mala Mitra, TINT, India, Dr. Debasish Pal, Uluberia College, India and, Dr. Siddhartha Bhattacharya, Uluberia College, India for fruitful collaboration.

I am especially grateful to Aniruddha da, a brother, mentor and friend. You introduced me to the world of biophysical experiments. You were patient with me, believed in me, gave me all the courage and freedom to work. You were with me through thick and thin, supporting me, keeping me in check and focused. You brushed aside every negativity that came my way. I cannot thank you enough. I am blessed to have shared this journey with you.

It would be unjust and unfair not to express my gratitude to Uttam da. My proper introduction to the world of computational biology came through you. It was a joy working with you. Observing you, how you designed and approached a computational problem was a lesson in itself, many of which I have tried to inculcate into my work. I am blessed to have learned from you. PS: I shall never forget the discussions we had about movies, books, manga, and anime and also holding the rare distinction- to be scolded by you!

"Coming together is a beginning; keeping together is progress; working together is success."

Henry Ford

This thesis is a result of coming together, keeping together and working together of so many brilliant and wonderful people. I'd like to express my sincerest gratitude to all my seniors and juniors for providing me with an enriched lab environment. I want to thank all members past and present of AOSL lab for all the helpful discussion and for creating a reassuring and productive atmosphere. Damayanti di, Priya di, Tanushree di, Gulmi di, Joyita di, Probir da, Prosenjit da, Tuhin da, Animesh da, Soumendra da, Soumendra da (Dr. Soumendra Darbar), Susmita, Ria, Monojit, Dipanjan, Nur, Nivedita, Arka, Arpan, Amrita, Neha, Lopa, Manali, Aman, Mahasweta, Deep Shikha, Abhinaba, Nairit, Sayantika, Oindrila, Sayan, Debanjana. I thank you all for taking care of me, for accepting me as your own.

Special thanks to Aniruddha da, Susmita, Ria and Monojit for the careful proofreading of this thesis. My PhD dissertation would still be an outlandish thought if not for Aniruddha da, Uttam da, Susmita, Ria and Monojit. Working with you all has been a great joy. From the morning tea to the late-night coffee breaks I have cherished every single moment. There was no dull or dreary moment when you all were around. I can't thank you all enough for everything you have done for me. I thank you from the bottom of my heart.

I express my deepest gratitude to St. Xavier's College, Kolkata for giving me the opportunity to pursue my dreams and providing me with all the necessary facilities to complete this work. I would also like to thank all my teachers and scholar friends at the Dept. of Microbiology, St. Xavier's College, Kolkata. They have been so warm and welcoming to me from the very beginning. I must acknowledge Dr. Riddhi Majumder without whom this journey would have been a distant dream.

I also owe a lot to my teachers at the Dept. of Microbiology, University of Kalyani. Especially to Prof. Samir Kumar Mukherjee who introduced me to Microbiology in all its vastness and richness. I must also acknowledge late Prof. Kanak Ranjan Samaddar for all his suggestions and encouragement. Their unwavering support helped me get through my graduate studies.

I am also thankful to all my teachers at my previous college and school for encouraging me to pursue my dreams. I express my deepest gratitude to Mr. Prosenjit Roy, Mr. Mir Hamim Ali and Mr. Sushen Chakraborty for their encouragement which sparked my interest in science.

Finally, I express my deepest gratitude to my family, for all the sacrifices they made to see my dream come to fruition. My father, to whom this thesis is dedicated. There doesn't a day that goes by when I don't miss him. His words, his encouragement, his wisdom, his support and his belief in me shaped my life. The principles he engrained in me, seared in me built the person I am today. I know he is happily looking at me from above. I owe him everything.

In the end, I am touched and humbled by everyone and anyone with whom I had the pleasure of sharing this journey. This little boy from Puruliya with dreams in his eyes has no way repaying the immense debt he is in. All he has is a humble thank you.

Pritam Biswas

Behala

Kolkata-34

December, 2021

"We must find time to stop and thank the people who make a difference in our lives."

John F. Kennedy

CONTENTS

Abstract

Essential dynamics is strongly correlated to biomolecular structure and function. The correlation between structure-dynamics of biomolecules is the key to understanding biomolecular function under different physiological environments. With time there has been an increase in demand for functional biomolecules by the biotechnology-industrial complex both in food, beverages as well as in the health care and pharmaceutical industries. However, a lack of understanding of the structure, function and dynamics of functional biomolecules under different physiologically relevant environments prevents our ability to utilize these biomolecules to their highest efficiency. Thus, exploring an in-depth structural, functional and dynamics study of the biomolecules under the different physiologically relevant environments are the motives of this thesis. Using a combined biophysical, biochemical and aided by various computational methods a correlation between structure-function and dynamics of biomolecules has been studied. In one set of work the influence of temperature on the structure, function and dynamics on a thermostable enzyme β-glucosidase (BGL) has been studied. Spectroscopic and computational studies confirmed the presence of four loop regions (LRs) acting as a gatekeeper to the incoming substrate. While, detailed studies of α-Chymotrypsin (CHT) at different temperatures revealed the essential motions modulating the catalytic activity. Both spectroscopic and theoretical calculations aided in understanding the loop motions and corroborated with the CHT catalytic activity. Finally, a homeothermic enzyme α-chymotrypsin was treated with a common covalent crosslinker formalin which resulted in reduced catalytic activity with increase in formalin concentration. Spectroscopic studies and computational calculations found remarkable change in molecular recognition of CHT and catalytic activity owning to residue modification by formalin. This study, will shed some light on the close relationship between structure-function-dynamics of biomolecules. Thus, opening up new avenues of research for the biotechnology-industrial complex to explore and also the possibility of engineering new enzymes.

Chapter 1

Introduction

1.1. Background:

The traditional view of biomolecules particularly protein is a static arrangement of atoms in space, the structures derived from x-ray crystallographic studies usually report the biomolecules as a single set of atomic coordinates best fitting the diffraction data. The lock and key hypothesis of Emil Fisher considered proteins (lock) having a static or fixed structure to which molecules having complementary shape (key) could bind [1]. The induced fit model introduced by Koshland [2] suggested that ligand binding induces a change at the binding, resulting in complementarity. While, the Monod, Wyman and Changeux model [3, 4] suggests that proteins can be either in a tensed or relaxed conformational state. Even before the first structure of biomolecule was available, work by Linderstrom-Lang [5] proposed that the amino acid units of a protein can form structurally ordered motifs called secondary structure elements, capable of folding into three-dimensional structure [6]. Pioneering works by Anfinsen et al., [7], Chotia et al., [8] and others showed that the three-dimensional structure of the protein depends upon the amino acid sequence of the said protein and, that the structure is conserved during evolution for a particular biomolecule. These studies indicated biomolecular structure as a key determinant to its function. Over the last few years, substantial progress has been made in the different experimental techniques particularly in time resolved fluorescence spectroscopy, X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy revealing the three-dimensional structure and projected a dynamical view of biomolecules [9]. The current view of biomolecules is much more complex than Koshland's or Monod's model, numerous studies concerning protein behavior have demonstrated the dynamical nature of protein or an ensemble of conformations [10]. The different ensemble states are characterized by the thermodynamics and kinetics of the biomolecule that can be defined by a Boltzmann distribution of conformational energy. Biomolecules can transition between different conformation, which depends upon the difference in energy between the conformational states. The lesser the difference, more is the probability of transition [11]. The basic biomolecular structure consists of potential regions of motion- the backbone and the side chain single bonds, the motions around which are limited only by steric hinderance [12-14]. Biomolecular function is affected by both structure and dynamics as well as the environment surrounding the

biomolecule [15]. Understanding the role of conformational dynamics and structure will provide better insights into biomolecular function. An in-depth exploration of structure-dynamics would help in the understanding of biomolecular function, which may have immense application in both industry and health care.

1.2. Structure-Function-Dynamics Modulation of Biomolecules in Physiologically Relevant Environments:

Essential dynamics is strongly correlated with biomolecular structure and function. Any change to the biomolecular structure will have an impact on the dynamics and function of the biomolecule [16, 17]. The relation between the conformational variability and structure allows biomolecules to adapt and function under different environmental stimuli. Change in environmental condition triggers a change in conformation of the biomolecule (Temperature, pH, ligand binding, etc.) [18, 19]. Essential dynamics is characterized by local motions, motions of large side chain, secondary elements and domains within a large time scale. Studies have shown many of such motions occur within the timescale of several biochemical processes. The local motion (10^{-15} to 10^{-10} sec) which correspond to the events occurring during enzyme catalyzed reaction, while studies on the larger side chain and domain motion correspond to proton transport (10^{-9}-10^{-4} sec); water structure reorganization (10^{-8} sec) [20-23]. Such studies indicate that essential dynamics is indeed imperative to biomolecular function. Any change to the inherent essential dynamics by changing the environmental stimuli would trigger a change in the essential dynamics affecting the function of the biomolecule, studies by Vogel and Parvizpour et al., [24, 25] have shown a change in temperature modulates the protein dynamics thereby, affecting the function. Recent study of biomolecules particularly enzymes, have shown collective motions or global motions between different protein residues modulates enzyme function [26]. Investigations involving mesophilic and thermophilic adenylate kinase have established a relationship between the local fluctuations of the hinge region and the catalytic efficiency of the enzyme [14, 27]. While, in separate studies catalytic efficiency of an enzyme was linked with the fluctuations of protein segments functioning as lids [28, 29]. Studies have also shown, the binding of ligand to an enzyme can trigger a conformational selection towards a favorable binding state [20, 30, 31]. Recently, studies have confirmed the enzyme-ligand interaction via conformational selection and induced-fit pathway at different pH [19]. Ligand induced allosteric effects, in which binding of ligand molecule at the allosteric site induce a conformational change at another distant site were found to be influenced by

essential dynamics of enzyme [32, 33]. The essential dynamics and the conformational plasticity of a biomolecule is defined by the structure, often referred to as "structure encoded dynamics" which in turn governs biomolecular function. Recent studies by Manjegowda et al., [34]; Schaefer et al., [35]; Fields et al., [36] and Chao et al., [37] suggest that changes to the biomolecular structure effects a change in the function as well as influences biomolecular adaptation to external environmental stimuli. This investigation on the structure-dynamics of biomolecules will find immense use in understanding biomolecular function e.g., protein (enzyme), DNA, lipid etc. under physiologically relevant environment.

1.3. Scope of the Thesis:

Functional biomolecules are fast emerging as a staple need for the biotechnology-industrial complex. Biomolecules such as enzymes as gaining recognition for their widespread use in food, agriculture, chemicals, pharmaceuticals and energy industries [38]. With time there has been an increase in demand for functional biomolecules and are favored over the more traditional approaches due to reduced process time, low energy input, cost effective, non-toxic and ecofriendly [39, 40]. However, several challenges remain to utilize them with highest efficiency. The development of recombinant DNA technology and protein engineering has opened new windows for manipulation and development of efficient biomolecules [41]. Still, a lack of understanding of the biomolecular dynamics-structure-function relationship hinders the development and utilization of functional biomolecules to their highest efficiency. The major limiting factors are not only the lack of comprehensive studies correlating biomolecular function and dynamics but also that of essential motions regulating the function. The seeming lack of knowledge regarding the local and global motions as well as the amplitude and the direction of motion of biomolecules [27], which is characterized by atoms or group motions within a large time scale hinders their utility to the maximum efficiency. These challenges create the scope of this thesis.

The key focus of this thesis is understanding the structure-function-dynamics correlation of biomolecules, in particular this study describes and correlates the essential dynamics of enzymes with enzyme structure and function using different spectroscopic and biochemical methods. Further, the trajectories of motion were studied and correlated to enzyme function using computational methods. Though several studies involving NMR and computational methods have reported on the dynamical nature of enzymes and other functional biomolecules. Very few studies on the essential motions of biomolecules responsible for

function are reported. In this regard, to explore the correlation between the essential dynamics of a thermotolerant enzyme and catalysis, we performed a spectroscopic and computational study of an industrially important enzyme β-Glucosidase (BGL) [EC 3.2.1.21] at different temperatures. Spectroscopic and computational studies confirm the presence of four loop regions (LRs) acting as a gate to the incoming substrate. The motional studies of the loop regions (LRs) was found to correlate with the catalytic activity of BGL. To further investigate the essential dynamics and its correlation to enzyme structure-function, we studied a homeothermic enzyme α-Chymotrypsin (CHT). While studying, both spectroscopic and computational studies established that the presence of loop regions (LRs) surrounding the catalytic play a pivotal role in catalysis of CHT. Both the studies involving BGL and CHT confirmed that the concerted motion of the LRs are essential for the catalytic activity of the enzymes. They serve as a gate, screening the incoming substrate. Any change to the concertedness of the LR motion causes change in catalytic activity of the enzymes. A full understanding of structure-dynamics-function requires an appreciation of the structural effects on biomolecular function. To correlate structural effects on functional attributes of a biomolecule we treated CHT with a common preservative formalin. Amino acid residue modifications to the primary structure of CHT translated into remarkable change in catalytic as well as molecular recognition attributes of the enzyme.

In brief, this thesis provides a comprehensive and in-depth study on the structure-dynamics-function of biomolecules utilizing both spectroscopic and computational approaches. This study will shed new light on the close relationship between structure-dynamics-function of biomolecules. Thus, opening new avenues for biotechnology-industrial complex to explore and potentially develop new engineered functional biomolecules with highest efficiency.

1.4. Objective:

Dynamics of biomolecules is the integral to the understanding of their function, which is also dependent upon the structure. The modern view of biomolecular structure has changed from a rigid and static to a dynamic one, where the structure exists as a conformational ensemble. Investigating how dynamics modulates function and the correlation between the structure-function-dynamics of biomolecules under physiologically relevant environment is the major aims of the study. Studying the correlation between structure-function-dynamics will open new avenues of knowledge in regards to almost every physiological process including cell signaling,

motility, catalysis, cellular transport. As well as, drug designing, bioengineering new enzymes for industrial and pharmaceutical application.

In this regard, for one set of work, we have studied the dynamical modulation of catalytic activity in a thermostable enzyme β-Glucosidase (BGL) at different temperatures using optical spectroscopy and computational techniques. BGL is an industrially important enzyme. Playing pivotal roles in biotechnological processes such as hydrolysis of isoflavone glucosides and production of bio-ethanol. BGLs are integral part of beverage industry, including wine, tea and juice industry. Catalytic hydrolysis of Rutin into Quercetin followed the enzymatic activity of BGL at different temperatures. Meanwhile, picosecond resolved fluorescence anisotropy using 8-anilino-1-naphthalenesulfonic acid ammonium salt (ANS) followed the dynamics of the enzyme. Structural stability was monitored at different temperature by CD spectroscopy. Molecular dynamics (MD) simulation followed the trajectory of motion. The root mean square deviation (RMSD), root mean square fluctuation (RMSF) and the radius of gyration (R_g) monitored the dynamics, geometry over time at different temperatures. while, normal mode analysis or principal component analysis (PCA) provided the direction of motion of the enzyme indicating the essential motion of BGL. To further correlate BGL dynamics to catalysis the energy required to remove the bound substrate Rutin was calculated using steered molecular dynamics (SMD).

In another study, structure-physiological function-dynamics was correlated in a homeothermic enzyme, α-Chymotrypsin (CHT). Ala-Ala-Phe-7-amido-4 methyl coumarin (Ala-Ala-Phe AMC) hydrolysis followed the enzymatic activity of CHT. Chymotrypsin is a model homeothermic enzyme and plays important roles in digestion, immune response and insect molting. To study the structural changes CD spectroscopy analysis was done at different temperatures. Picosecond resolved fluorescence spectroscopy of ANS-CHT complex followed the CHT dynamics. To further study and corroborate structure and dynamics with CHT function, we performed molecular dynamics (MD) simulation at different temperatures. The inherent dynamics of CHT at different temperatures was analyzed. The changes to the enzyme structure with time were computed in terms of root mean square deviation (RMSD). The root mean square fluctuation (RMSF) and the radius of gyration (Rg) were also calculated using the simulation trajectories. Furthermore, the essential dynamics was analyzed by normal mode analysis or principal component analysis (PCA). We studied trajectory of CHT motion and determined the essential dynamics, geometry and corroborated with CHT function.

Apart from the essential motion, the biomolecular structure also effects the function. In this regard, we chemically modified the homeothermic enzyme α-Chymotrypsin (CHT) using formalin. The effect of such modification on catalysis and dynamics was monitored. The alteration of catalytic activity was followed by hydrolysis of Ala-Ala-Phe-7-amido-4 methyl coumarin (Ala-Ala-Phe AMC). While, the picosecond resolved fluorescence anisotropy of ANS-CHT complex indicated the change in dynamics. CD spectroscopy and dynamic light scattering (DLS) indicated the extent to change to the secondary and globular tertiary structure of CHT. Finally, we studied the molecular recognition of CHT using molecular docking which showed a divergence from the wild type unmodified CHT to formalin modified CHT.

1.5. Summary of the Work Done:

1.5.1. Studies on the Structure, Function, and Dynamics of a Thermostable Enzyme at Physiologically Relevant Temperatures:

1.5.1.1. Flexibility Modulates Catalytic Activity of a Thermostable Enzyme: Key Information from Optical Spectroscopy and Molecular Dynamics Simulation [42]:

Enzymes are dynamical macromolecules and their conformation can be altered via local fluctuations of side chains, large scale loop and even domain motions which are intimately linked to their function. Herein, we have addressed the role of dynamic flexibility in the catalytic activity of a thermostable enzyme almond β-glucosidase (BGL). Optical spectroscopy and classical molecular dynamics (MD) simulation were employed to study the thermal stability, catalytic activity and dynamical flexibility of the enzyme. An enzyme assay reveals high thermal stability and optimum catalytic activity at 333 K. Polarization-gated fluorescence anisotropy measurements employing 8-anilino-1-napthelenesulfonic acid (ANS) have indicated increasing flexibility of the enzyme with an increase in temperature. A study of the atomic 3D structure of the enzyme shows the presence of four loop regions (LRs) strategically placed over the catalytic barrel as a lid. MD simulations have indicated that the flexibility of BGL increases concurrently with temperature through different fluctuating characteristics of the enzyme's LRs. Principal Component Analysis (PCA) and the Steered Molecular Dynamics (SMD) simulation manifest the gatekeeper role of the four LRs through their dynamic fluctuations surrounding the active site which controls the catalytic activity of BGL

1.5.2. Studies on the Correlation Between Structure, Physiological Function, and Essential Dynamics of a Homeothermic Enzyme:

1.5.2.1. Essential Loop Dynamics Modulates Catalytic Activity in α-Chymotrypsin [43]:

Conformational dynamics of macromolecules including enzymes are essential for their function. Herein, we report the role of essential dynamics in α-chymotrypsin (CHT) which correlates with its catalytic activity. Detailed optical spectroscopy and classical molecular dynamics (MD) simulation were used to study thermal stability, catalytic activity and dynamical flexibility of the enzyme. The study of the enzyme kinetics reveals an optimum catalytic efficiency at 308K. Polarization gated fluorescence anisotropy with 8-anilino-1-napthelene sulfonate (ANS) have indicated increasing flexibility of the enzyme with an increase in temperature. Examination of the structure of CHT reveal the presence of five loop regions (LRs) around the catalytic S1 pocket. MD simulations have indicated that flexibility increases concurrently with temperature which decreases beyond optimum temperature. Principal component analysis (PCA) of the eigenvectors manifests essential dynamics and gatekeeping role of the five LRs surrounding the catalytic pocket which controls the enzyme activity.

1.5.3. Studies on the Effect of Chemical Modification on the Structure, Dynamics, and Function of a Model Enzyme:

1.5.3.1. A Combined Spectroscopic and Molecular Modeling Study on Structure-Function-Dynamics under Chemical Modification: α-Chymotrypsin with Formalin Preservative [44]:

Enzyme conformations can be altered via modification of its amino acid residues, side chains and large-scale domain modifications, which are closely linked to its function. Herein, we have addressed the role of residue modification in catalytic activity and molecular recognition of an enzyme alpha-chymotrypsin (CHT) in presence of covalent cross-linker formalin. Optical spectroscopy studies exhibit reduced catalytic activity of the enzyme with increased formalin concentration. Polarization gated anisotropy studies of a fluorophore 8-anilino-1-napthelenesulfonic acid (ANS) in CHT show a dip rise pattern in presence of formalin which is consistent with the generation of multiple ANS binding sites in the enzyme owing to modifications of its local amino acid residues. Molecular docking study on minimal local residue modifications in CHT reveals formation of a stable enzyme-substrate complex even with the serine-histidine cross-linked enzyme which prohibits product formation giving rise to reduced catalytic activity.

1.6. Plan of Thesis:

The plan of the thesis is as follows:

Chapter 1: This chapter provides a brief introduction to the scope and motivation behind the thesis work. A brief summary of the work done is also included in this chapter.

Chapter 2: This chapter provides a brief overview of the steady-state and dynamical tools, computational methods, the structural aspects of biomolecule system (enzymes) and chemical probes used in the experiments.

Chapter 3: This chapter describes the instrumental details, experimental protocols and data analysis procedures.

Chapter 4: In this chapter, the role of flexibility in modulating catalytic activity of a thermostable enzyme β-Glucosidase (BGL) has been investigated.

Chapter 5: In this chapter, the role of essential dynamics in biomolecular function of a homeothermic enzyme α-Chymotrypsin (CHT) has been investigated at physiologically relevant temperatures.

Chapter 6: In this chapter, the role of chemical modification on the structure-dynamics and physiological function of α-Chymotrypsin has been investigated using the common chemical preservative formalin.

References

[1] E. Fischer, Einfluss der Configuration auf die Wirkung der Enzyme, *Chem. Ber.*, 27 (1894) 2985-2993.

[2] D.E. Koshland Jr, Application of a theory of enzyme specificity to protein synthesis, *Proc. Natl. Acad. Sci. U. S. A.*, 44 (1958) 98-104.

[3] J.-P. Changeux, S.J. Edelstein, Allosteric mechanisms of signal transduction, *Science*, 308 (2005) 1424-1428.

[4] J.-P. Changeux, S.J. Edelstein, Conformational selection or induced fit? 50 years of debate resolved, *F1000 Biol. Rep.*, 3 (2011) 1-15.

[5] K.U. Linderstrøm-Lang, Lane medical lectures: proteins and enzymes, *Stanford University Press, USA*, (1952).

[6] S.K. Mishra, Protein sequence-structure-dynamics-function relationships: The close association of dynamics with protein function, *Doctoral dissertation, Iowa State University, USA*, (2018).

[7] C.B. Anfinsen, E. Haber, Studies on the reduction and re-formation of protein disulfide bonds, *J. Biol. Chem.*, 236 (1961) 1361-1363.

[8] C. Chothia, A.M. Lesk, The relation between the divergence of sequence and structure in proteins, *The EMBO journal*, 5 (1986) 823-826.

[9] R.V. Mannige, Dynamic new world: refining our view of protein structure, function and evolution, *Proteomes*, 2 (2014) 128-153.

[10] K.L. Damm, H.A. Carlson, Exploring experimental sources of multiple protein conformations in structure-based drug design, *J. Am. Chem. Soc.*, 129 (2007) 8225-8235.

[11] K.A. Dill, Polymer principles and protein folding, *Protein Sci.*, 8 (1999) 1166-1180.

[12] M. Gerstein, N. Echols, Exploring the range of protein flexibility, from a structural proteomics perspective, *Curr. Opin. Chem. Biol.*, 8 (2004) 14-19.

[13] J. Evenäs, S. Forsén, A. Malmendal, M. Akke, Backbone dynamics and energetics of a calmodulin domain mutant exchanging between closed and open conformations, *J. Mol. Biol.*, 289 (1999) 603-617.

[14] K.A. Henzler-Wildman, M. Lei, V. Thai, S.J. Kerns, M. Karplus, D. Kern, A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, *Nature*, 450 (2007) 913-916.

[15] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, *Proteins: Struct. Funct. and Bioinf.* 17 (1993) 412-425.

[16] Ø. Halskau Jr, R. Perez-Jimenez, B. Ibarra-Molero, J. Underhaug, V. Muñoz, A. Martinez, J.M. Sanchez-Ruiz, Large-scale modulation of thermodynamic protein folding barriers linked to electrostatics, *Proc. Natl. Acad. Sci. U.S.A*, 105 (2008) 8625-8630.

[17] J.A. McCammon, Protein dynamics, *Rep. Prog. Phys.*, 47 (1984) 1-46.

[18] J. Peon, S.K. Pal, A.H. Zewail, Hydration at the surface of the protein Monellin: dynamics with femtosecond resolution, *Proc. Natl. Acad. Sci. U.S.A*, 99 (2002) 10964-10969.

[19] S. Choudhury, S. Batabyal, P.K. Mondal, P. Singh, P. Lemmens, S.K. Pal, Direct observation of kinetic pathways of biomolecular recognition, *Chem. Eur. J.*, 21 (2015) 16172-16177.

[20] M. Akke, NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis, *Curr. Opin. Struct. Biol.*, 12 (2002) 642-647.

[21] M.F. Dunn, Protein–ligand interactions: general description, *e LS*, (2001) 1-13.

[22] K. Teilum, J.G. Olsen, B.B. Kragelund, Functional aspects of protein flexibility, *Cell. Mol. Life Sci.*, 66 (2009) 2231-2247.

[23] D. Zhong, Ultrafast catalytic processes in enzymes, *Curr. Opin. Chem. Biol.*, 11 (2007) 174-181.

[24] M. Vogel, Temperature-dependent mechanisms for the dynamics of protein-hydration waters: a molecular dynamics simulation study, *J. Phys. Chem. B*, 113 (2009) 9386-9392.

[25] S. Parvizpour, J. Razmara, M.S. Shamsir, Temperature adaptation analysis of a psychrophilic mannanase through structural, functional and molecular dynamics simulation, *Mol. Simul.*, 44 (2018) 1270-1277.

[26] T. Haliloglu, I. Bahar, Adaptability of protein structures to enable functional interactions and evolutionary implications, *Curr. Opin. Struct. Biol*., 35 (2015) 17-23.

[27] K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins, *Nature*, 450 (2007) 964-972.

[28] M. Wolf-Watz, V. Thai, K. Henzler-Wildman, G. Hadjipavlou, E.Z. Eisenmesser, D. Kern, Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair, *Nat. Struct. Mol. Biol.*, 11 (2004) 945-949.

[29] R.M. Crean, M. Biler, M.W. van der Kamp, A.C. Hengge, S.C.L. Kamerlin, Loop dynamics and enzyme catalysis in protein tyrosine phosphatases, *J. Am. Chem. Soc.*, 143 (2021) 3830-3845.

[30] K. Lindorff-Larsen, R.B. Best, M.A. DePristo, C.M. Dobson, M. Vendruscolo, Simultaneous determination of protein structure and dynamics, *Nature*, 433 (2005) 128-132.

[31] B.G. Vértessy, F. Orosz, From “fluctuation fit” to “conformational selection”: evolution, rediscovery, and integration of a concept, *Bioessays*, 33 (2011) 30-34.
[32] P.-D. Coureux, H.L. Sweeney, A. Houdusse, Three myosin V structures delineate essential features of chemo-mechanical transduction, *The EMBO journal*, 23 (2004) 4527-4537.
[33] I.J. General, Y. Liu, M.E. Blackburn, W. Mao, L.M. Gierasch, I. Bahar, ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones, *PLoS Comput. Biol.*, 10 (2014) e1003624.
[34] D.S. Manjegowda, P. Karunakar, N.B. Ramachandra, Effect of structural changes in proteins derived from GATA4 nonsynonymous single nucleotide polymorphisms in congenital heart disease, *Indian J. Pharm. Sci.*, 77 (2015) 735-741.
[35] C. Schaefer, B. Rost, Predict impact of single amino acid change upon protein structure, *BMC Genomics*, 13 (2012), 1-10.
[36] P.A. Fields, Y. Dong, X. Meng, G.N. Somero, Adaptations of protein structure and function to temperature: there is more than one way to ‘skin a cat’, *J. Exp. Biol.*, 218 (2015) 1801-1811.
[37] Y.-C. Chao, M. Merritt, D. Schaefferkoetter, T.G. Evans, High-throughput quantification of protein structural change reveals potential mechanisms of temperature adaptation in Mytilus mussels, *BMC Evol. Biol.*, 20 (2020) 1-18.
[38] R. Singh, M. Kumar, A. Mittal, P.K. Mehta, Microbial enzymes: industrial progress in 21st century, *3 Biotech*, 6 (2016) 1-15.
[39] S. Li, X. Yang, S. Yang, M. Zhu, X. Wang, Technology prospecting on enzymes: application, marketing and engineering, *Comput. Struct. Biotechnol. J.*, 2 (2012) e201209017.
[40] J.-M. Choi, S.-S. Han, H.-S. Kim, Industrial applications of enzyme biocatalysis: Current status and future aspects, *Biotechnol. Adv.*, 33 (2015) 1443-1454.
[41] L. Liu, H. Yang, H.-d. Shin, R.R. Chen, J. Li, G. Du, J. Chen, How to achieve high-level expression of microbial enzymes: strategies and perspectives, *Bioeng.*, 4 (2013) 212-223.
[42] P. Biswas, A. Adhikari, U. Pal, P. Singh, M. Das, T. Saha-Dasgupta, S.S. Choudhury, R. Das, S.K. Pal, Flexibility modulates the catalytic activity of a thermostable enzyme: key information from optical spectroscopy and molecular dynamics simulation, *Soft matter*, 16 (2020) 3050-3062.
[43] P. Biswas, U. Pal, A. Adhikari, S. Mondal, R. Ghosh, D. Mukherjee, T. Saha-Dasgupta, S.S. Choudhury, R. Das, S.K. Pal, Essential Loop Dynamics Modulates Catalytic Activity in α-Chymotrypsin, *bioRxiv*, (2021) https://doi.org/10.1101/2021.08.11.455937.

[44] P. Biswas, A. Adhikari, U. Pal, S. Mondal, D. Mukherjee, R. Ghosh, T. Saha-Dasgupta, S.S. Chowdhury, R. Das, S.K. Pal, A Combined Spectroscopic and Molecular Modeling Study on Structure-Function-Dynamics under Chemical Modification: Alpha-Chymotrypsin with Formalin Preservative, *bioRxiv*, (2021) https://doi.org/10.1101/2021.07.24.453635.

Chapter 2

Overview of Experimental, Computational Techniques and Systems

In order to investigate the dynamical processes involved in the course of study on the structure, function and dynamics of biomolecules different steady state, dynamical tools including fluorescence anisotropy, dynamic light scattering have been used. To understand and study the inherent dynamics and molecular recognition several computational methods were employed like classical molecular dynamics (MD) simulation, steered molecular dynamics simulation (SMD), molecular docking. In this chapter, a brief overview of the theoretical aspects of the spectroscopic tools and computational methods are provided. Also, the various systems used in the study are briefly discussed.

2.1. Steady state and Dynamical Tools:

2.1.1. Fluorescence Lifetime:

The fluorescence lifetime of a fluorophore is defined by the average time the molecule spends in the excited state before returning to the ground state via loss of energy through fluorescence and other non-radiative processes (Figure 2.1).

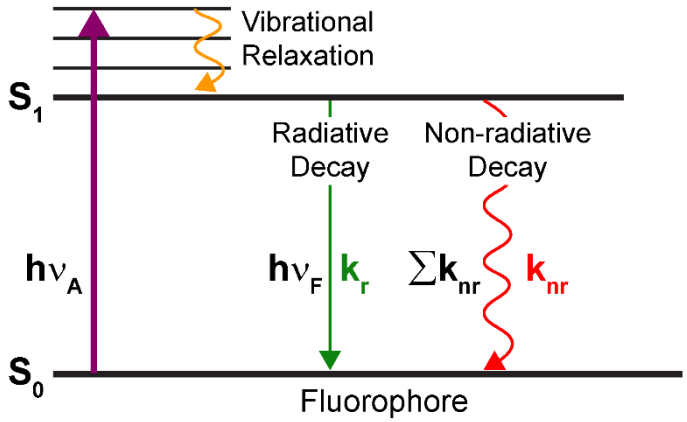

Figure 2.1. *A simplified Jablonski diagram to illustrate the meaning of fluorescence lifetime. k_r and k_{nr} are the radiative and non-radiative rates of decay.*

2.1.1.1. Theory:

The excited state lifetime, τ, is defined by the following equation

$$\tau = \frac{1}{k_r + k_{nr}} \quad (2.1)$$

Where, kr and knr is the radiative and non-radiative rate of decay. The lifetime of the fluorophore in the absence of non-radiative processes is called the intrinsic or natural lifetime, and is given by, Where, k_r and k_{nr} is the radiative and non-radiative rate of decay. The lifetime of the fluorophore in the absence of non-radiative processes is called the intrinsic or natural lifetime, and is given by, $\tau = \frac{1}{k_r}$.

2.1.1.2. Experimental Methods:

The fluorescence transients were measured with commercially available time correlated single photon counting (TCSPC) setup from Edinburgh Instruments Ltd. (Livingston, UK) having an instrument response function, IRF=80 ps.

Curve fitting of the time-resolved fluorescence transients was performed using a nonlinear least square fitting procedure to a function (equation (2.2)) comprised of convolution of the IRF (E(t)) with a sum of exponentials (equation (2.3)) with pre-exponential factors (Bi), characteristic lifetimes (τi), and a background (A).

$$\left(X(t) = \int_0^t E(t')R(t-t')dt' \right) \quad (2.2)$$

$$\left(R(t) = A + \sum_{i=1}^{N} B_i e^{\frac{-t}{\tau_i}} \right) \quad (2.3)$$

Relative concentration in a multi-exponential decay is expressed as (equation 2.4)

$$C_n = \frac{B_n}{\sum_{i=1}^{N} B_i} \times 100 \quad (2.4)$$

The average lifetime (amplitude-weighted) of a multi-exponential decay is finally expressed as,

$$\tau_{avg} = \sum_{i=1}^{N} C_i \tau_i \quad (2.5)$$

The quality of the curve fitting is evaluated by reduced $\chi 2$ and residual data.

2.1.2. Fluorescence Anisotropy:

Anisotropy is defined as the extent of polarization of the emission from a fluorophore. Anisotropy measurements are commonly used in biochemical applications of fluorescence. It provides information about the size and shape of proteins or the rigidity of various molecular environments. Anisotropy measurements have also been used to measure protein-protein associations, fluidity of membranes and for immunoassays of numerous substances. These measurements are based on the principle of photo-selective excitation of those fluorophore molecules whose absorption transition dipoles are parallel to the electric vector of polarized excitation light. In an isotropic solution, fluorophores are oriented randomly. However, upon selective excitation, partially oriented population of fluorophores with polarized fluorescence emission results. The relative angle between the absorption and emission transition dipole moments determines the maximum measured anisotropy (r_0). The fluorescence anisotropy (r) and polarization (P) are defined by,

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad (2.6)$$

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad (2.7)$$

where the fluorescence intensities of vertically and horizontally polarized emission when the fluorophore is excited with vertically polarized light. Polarization and anisotropy are interrelated as,

$$r = \frac{2P}{3 - P} \quad (2.8)$$

$$P = \frac{3r}{2 + r} \quad (2.9)$$

Although polarization and anisotropy provide the same information, anisotropy is preferred since the latter is normalized by total fluorescence intensity $I_T = I_{\parallel} + 2I_{\parallel}$ and in case of multiple emissive species anisotropy is additive while polarization is not. Several phenomena, including rotational diffusion and energy transfer, can decrease the measured anisotropy to values lower than maximum theoretical values. Following a pulsed excitation, the fluorescence anisotropy, r(t) of a sphere is given by,

$$r(t) = r_0 \exp(^{-t}/_{\tau_{rot}}) \tag{2.10}$$

where r_0 is the anisotropy at time t = 0 and τ_{rot} is the rotational correlation time of the sphere.

2.1.2.1. Theory:

For a radiating dipole the intensity of light emitted is proportional to the square of the projection of the electric field of the radiating dipole onto the transmission axis of the polarizer. The intensity of parallel and perpendicular projections are given by,

$$I_{II}(\theta, \psi) = \cos^2 \theta \tag{2.11}$$

$$I_{\perp}(\theta, \psi) = \sin^2 \theta \sin^2 \psi \tag{2.12}$$

where θ and ψ are the orientational angles of a single fluorophore relative to the z and y-axis, respectively (Figure 2.2.a). In solution, fluorophores remain in random distribution and the anisotropy is calculated by excitation photoselection. Upon photoexcitation by polarized light, the molecules having absorption transition moments aligned parallel to the electric vector of the polarized light have the highest probability of absorption. For the excitation polarization along z-axis, all molecules having an angle ψ with respect to the y-axis will be excited. The population will be symmetrically distributed about the z-axis. For experimentally accessible molecules, the value of ψ will be in the range from 0 to 2 ψ with equal probability. Thus, the ψ dependency can be eliminated.

$$\langle \sin^2 \psi \rangle = \frac{\int_0^{2\pi} \sin^2 \psi d\psi}{\int_0^{2\pi} d\psi} = \frac{1}{2} \tag{2.13}$$

$$\text{and } I_{\parallel}(\theta) = cos^2\theta \tag{2.14}$$

$$I_{\parallel}(\theta) = \frac{1}{2} sin^2\theta \tag{2.15}$$

Consider a collection of molecules oriented relative to the z-axis with probability *f(*θ*).* Then, measured fluorescence intensities for this collection after photoexcitation are,

$$I_{II} = \int_0^{\pi/2} f(\theta)\cos^2\theta d\theta = k\langle\cos^2\theta\rangle \tag{2.16}$$

$$\mathrm{I}_{\perp} = \frac{1}{2}\int_0^{\pi/2} f(\theta)\sin^2\theta d\theta = \frac{k}{2}\langle\sin^2\theta\rangle \tag{2.17}$$

where *f(*θ*)d*θ is the probability that a fluorophore is oriented between θ and θ + *d*θ and is given by,

$$f(\theta)d\theta = \cos^2\theta\sin\theta d\theta \tag{2.18}$$

k is the instrumental constant. Thus, the anisotropy (*r*) is defined as,

$$r = \frac{3\langle\cos^2\theta\rangle - 1}{2} \tag{2.19}$$

when θ = 54.70 i.e. when *cos2*θ = 1/3, the complete loss of anisotropy occurs. Thus, the fluorescence taken at θ = 54.70 with respect to the excitation polarization is expected to be free from the effect of anisotropy and is known as magic angle emission. For collinear absorption and emission dipoles, the value of <cos2θ> is given by the following equation,

$$\langle\cos^2\theta\rangle = \frac{\int_0^{\pi/2}\cos^2\theta f(\theta)d\theta}{\int_0^{\pi/2} f(\theta)d\theta} \tag{2.20}$$

Substituting equation (2.18) in equation (2.20) one can get the value of <cos2θ> = 3/5 and anisotropy value to be 0.4 (from equation (2.19)). This is the maximum value of anisotropy obtained when the absorption and emission dipoles are collinear and when no other depolarization process takes place. However, for most fluorophore, the value of anisotropy is less than 0.4 and it is dependent on the excitation wavelength. It is demonstrated that as the displacement of the absorption and emission dipole occurs by an angle γ relative to each other, it causes further loss of anisotropy (reduction by a factor 2/5) [1] from the value obtained from equation (2.19). Thus, the value of fundamental anisotropy, r_0 is given by,

$$r_0 = \frac{2}{5}\left(\frac{3\cos^2\gamma - 1}{2}\right) \tag{2.21}$$

For any fluorophore randomly distributed in solution, with one-photon excitation, the value of r_0 varies from -0.20 to 0.40 for γ values varying from 90° to 0°.

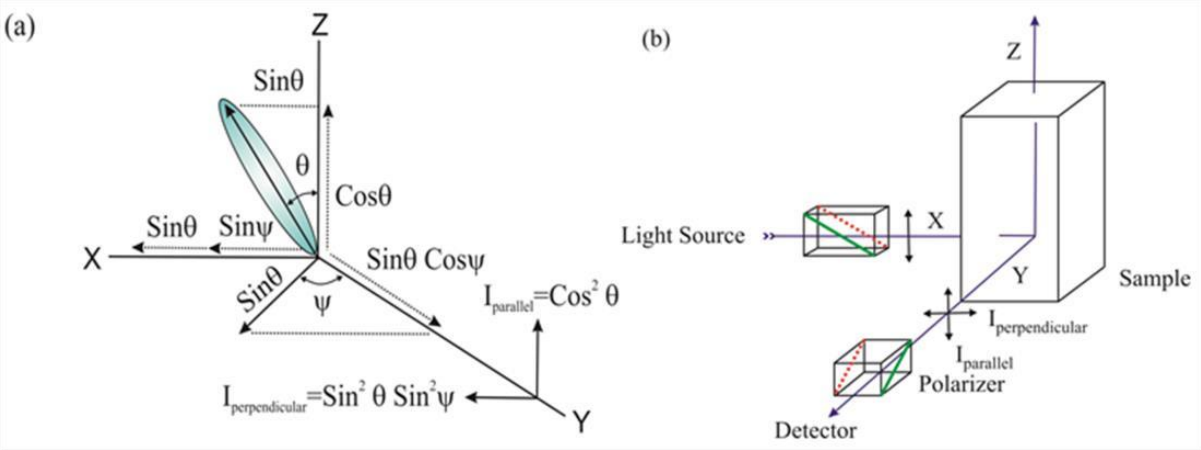

Figure 2.2. *(a) Emission intensity of a single fluorophore (blue ellipsoid) in a coordinate system. (b) Schematic representation of the measurement of fluorescence anisotropy.*

2.1.2.2. Experimental Methods:

For time resolved anisotropy (*r(t)*) measurements (Figure 2.2.b), emission polarization is adjusted to be parallel and perpendicular to that of the excitation polarization. [2] have derived the relevant equations for the time dependence of $I_{II}(t)$ (equation (2.22)) and $I_{\perp}(t)$ (equation (2.23)) for single rotational and fluorescence relaxation times, τ_{rot} and τ_f, respectively,

$$I_{II}(t) = \exp\left(-t/\tau_f\right)\left(1 + 2r_0\exp(-t/\tau_{rot})\right) \tag{2.22}$$

$$I_{\perp}(t) = \exp\left(-t/\tau_f\right)\left(1 - r_0\exp(-t/\tau_{rot})\right) \tag{2.23}$$

The total fluorescence is given by,

$$F(t) = I_{II}(t) + 2I_{\perp}(t) = 3\exp\left(-t/\tau_f\right) = F_0\exp\left(-t/\tau_f\right) \tag{2.24}$$

The time dependent anisotropy, r(t) is given by,

$$r(t) = \frac{I_{II}(t) - I_{\perp}(t)}{I_{II}(t) + 2I_{\perp}(t)} = r_0 \exp(-t/\tau_{rot}) \quad (2.25)$$

F(t) depends upon τf and *r(t)* depends upon τrot so that these two lifetimes can be separated. This separation is not possible in steady-state measurements. It should be noted that the degree of polarization (*P*) is not independent of τf and is therefore not as useful quantity as *r*. For reliable measurement of *r(t)*, three limiting cases can be considered.

(a) If $\tau f < \tau rot$, the fluorescence decays before the anisotropy decays, and hence only r_0 can be measured.

(b) If $\tau_{rot} < \tau_f$, in contrast to steady-state measurements, τ_{rot} can be measured in principle. The equations (2.22) and (2.23) show that the decay of the parallel and perpendicular components depends only upon τ_{rot}. The only experimental disadvantage of this case is that those photons emitted after the period of a few times τrot cannot contribute to the determination of τ_{rot}, but provided the signal-to-noise ratio is favourable, this need not be of great concern.

(c) If $\tau_{rot} \approx \tau_f$, then it becomes the ideal situation since almost all photons are counted within the time (equal to several rotational relaxation times) in which *r(t)* shows measurable changes.

For systems with multiple rotational correlation times, *r(t)* is given by,

$$r(t) = r_0 \sum_i \beta_i e^{-t/\tau_i} \quad (2.26)$$

Where $\sum_i \beta_i = 1$. It should be noted that the instrument monitoring the fluorescence, particularly the spectral dispersion element, responds differently to different polarizations of light, thus emerging the need for a correction factor. For example, the use of diffraction gratings can yield intensities of emission, which depend strongly upon orientation with respect to the plane of the grating. It is inevitably necessary when using such instruments to correct for the anisotropy in response. This instrumental anisotropy is usually termed as G-factor (grating factor) and is defined as the ratio of the transmission efficiency for vertically polarized light to that for horizontally polarized light ($G = I_{II} + I_{\perp}$). Hence, values of fluorescence anisotropy, *r(t)* corrected for instrumental response, would be given by [3],

$$r(t) = \frac{I_{II}(t) - GI_{\perp}(t)}{I_{II}(t) + 2GI_{\perp}(t)} \quad (2.27)$$

The G-factor at a given wavelength can be determined by exciting the sample with horizontally polarized excitation beam and collecting the two polarized fluorescence decays, one parallel and other perpendicular to the horizontally polarized excitation beam. G-factor can also be determined following longtime tail matching technique [4]. If $\tau_{rot} < \tau_f$, it can be seen that the curves for $I_{II}(t)$ and $I_{\perp}(t)$ should become identical. If in any experiment they are not, it can usually be assumed that this is due to a nonunitary G-factor. Hence normalizing the two decay curves on the tail of the decay eliminates the G-factor in the anisotropy measurement.

2.1.3. Circular Dichroism (CD):

CD is a routinely used tool for probing the structures of biomolecules. Particularly, in determining the secondary structures (α-helix and β-sheet) of proteins [5, 6].

2.1.3.1. Theory:

A plane polarised light when passes through an optically active substance, not only do the left (L) and right (R) circularly polarized light rays travel at different speeds, $c_L \neq c_R$, but these two rays are absorbed to a different extent, i.e. $A_L \neq A_R$. The difference in the absorbance of the left and right circularly polarized light, i.e., $\Delta A = A_L - A_R$, is defined as circular dichroism (CD). Circular dichroism spectroscopy follows the Beer-Lambert's law. If I_0 is the intensity of incident light on the cell, and I, that of emergent light, then the absorbance of the sample is given by,

$$A = \log_{10}\left(\frac{I_0}{I}\right) = \varepsilon c l \qquad (2.28)$$

Where, A is proportional to concentration (c) of the optically active sample and optical path length (l). If 'c' is in mol L^{-1} and 'l' is in cm, then ε is called the molar absorptivity or molar extinction coefficient. For an optically active medium having two absorbances A_L and A_R, the intensity of the right and left circularly polarized light are considered the same, i.e. $I_0 = I_L = I_R$ at the time of incidence. Any dicrograph passes periodically changing light through the medium, oscillating between left and right circular polarization, and the difference in absorbances are recorded directly.

$$\Delta A = A_L - A_R = \log_{10}\left(\frac{I_0}{I_L}\right) - \log_{10}\left(\frac{I_0}{I_R}\right) = \log_{10}\left(\frac{I_R}{I_L}\right) \qquad (2.29)$$

$$\Delta A = (\Delta \varepsilon) cl \qquad (2.30)$$

In the above equation (2.29), I_0 does not appear in the final equation. Hence, there is no need for a reference beam. The instruments are, therefore, of single beam type.

Light is changed in two aspects after passing through an optically active substance. The maximal amplitude of intensity is no longer confined to a plane; instead it traces out an ellipse. Ellipticity is defined as the arctangent of the ratio of minor axis to the major axis of the ellipse (Figure 2.3). The orientation of ellipse is another aspect. The major axis of the ellipse no longer remains parallel to the polarization of the incident light. Thus, after passing through an optically active substance, neither do the absorbance nor do the radii of the emergent left and right circularly polarized light remains same. Hence, CD is equivalent to ellipticity.

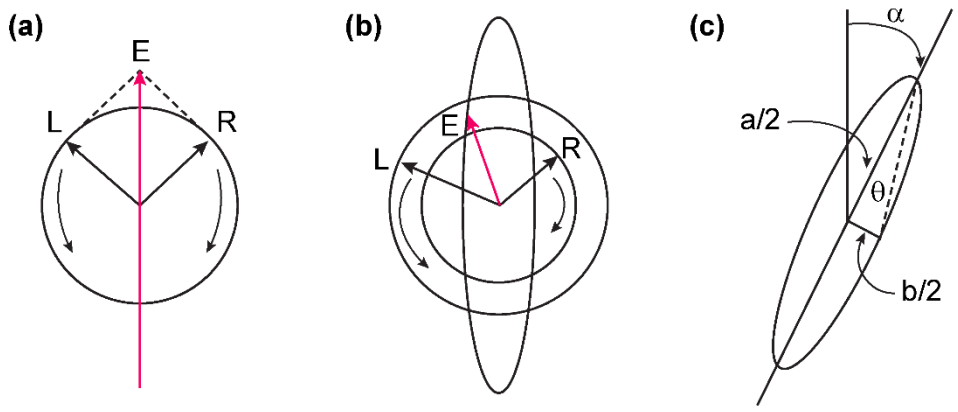

Figure 2.3. *(a) Left (L) and right (R) circularly polarized light component having same intensities and phases lying in one plane and oscillating with same magnitude. (b) R component being less intense (more absorbed) than L component leading to elliptically polarized light. (c) θ, ellipticity is the angle made by semi-major and semi-minor axes of the ellipse. The major axis has rotated through angle α corresponding to optical rotation.*

Most of the CD spectropolarimeters, although they measure differential absorption, produce a CD spectrum in units of ellipticity (θ) expressed in milli-degrees versus λ, rather than ΔA versus λ. The relation between ellipticity and CD is given by,

$$\theta = \frac{2.303 \times 180 \times (A_L - A_R)}{4\pi} \text{ (in degrees)} \quad (2.31)$$

To compare the results from different samples, optical activity is computed on a molar or residue basis. Molar ellipticity, [θ] is defined as,

$$[\theta] = \frac{\theta}{cl} \quad (2.32)$$

Where (θ) is in degrees, *'c'* is in mol L^{-1} and *'l'* is in cm. The unit of molar ellipticity is $deg.M^{-1}.cm^{-1}$. Sometimes, CD is reported as $\Delta\varepsilon = \Delta\varepsilon_L = \Delta\varepsilon_R$. From Beer-Lambert law and molar ellipticity relation it can be shown that,

$$[\theta] = 3300\Delta\varepsilon \quad (2.33)$$

2.1.3.2. Experimental Methods:

In biophysical studies, CD is mostly used to determine the secondary structures of proteins and nucleic acids and the changes in secondary structures upon recognition by small molecules and other biomolecules. Through CD spectropolarimeter, we obtain CD spectrograph having a plot of optical rotation in millidegrees versus wavelength in nm. In order to obtain information about the secondary structures of proteins, the graph is fitted with non-linear least square fitting method using freely available software. The percentages of different secondary structures are calculated by matching the experimental data with that of reference standard. In proteins, the secondary structural content includes α-helix, β-sheet, β-turn and random coil. The CD spectrum of α-helix contains two negative peaks, one at 208 nm (π-π* transition) and 222 nm (n-π* transition). β-sheet has a negative band at 216 nm and a positive band of similar magnitude at 195 nm. β-turn has weak negative peak at 225 nm (n-π* transition), a strong positive peak between 200 nm and 205 nm due to π-π* transition and a strong negative band between 180 nm and 190 nm. Random coil or unordered conformation has a strong negative band below 200 nm; a positive band at 218 nm and in some cases has a very weak negative band at 235 nm.

2.1.4. Dynamic Light Scattering:

According to semi-classical theory of light scattering, when light impinges on matter, the electric field of the light induces an oscillating polarization of electrons in the molecules. Hence, the molecules provide a secondary source of light and subsequently scatter light. The

frequency shift, the angular distribution, the polarization, and the intensity of scattered light are determined by size, shape and molecular interactions in the scattering material. Dynamic light scattering (DLS) also known as photon correlation spectroscopy (PCS) or quasi-elastic light scattering (QELS) is one of the most popular technique used to determine the size of the particle.

2.1.4.1. Theory:

DLS experiments are based on two assumptions:

(a) Particles exhibit Brownian motion (also called 'random walk'). The probability density function is given by the formula,

$$P(r,t|0,0)=(4\pi DT)^{\frac{3}{2}}e^{\frac{-r^2}{4Dt}} \quad (2.34)$$

Where D is the translational diffusion coefficient.

(b) The particles are spherical in shape with particles of molecular dimensions. If it is so, then it is possible to apply the Stoke-Einstein relation and hence have a formula that easily gives the diffusion constant,

$$D=\frac{k_B T}{3\pi\eta d_H} \quad (2.35)$$

Where d_H is the hydrodynamic diameter of the particles, k_B is the Boltzmann constant, T is the temperature in Kelvin and η is the viscosity of the solvent.

2.1.4.2. Experimental Methods:

It has been seen that particles in dispersion are in a constant, random Brownian motion and this causes the intensity of scattered light to fluctuate as a function of time. The correlator used in a DLS instrument constructs the intensity autocorrelation function $G(\tau)$ of the scattered intensity,

$$G(\tau)=\langle I(t)I(t+\tau)\rangle \quad (2.36)$$

Where τ is the time difference (the sample time) of the correlator. For a large number of monodisperse particles in Brownian motion, the correlation function (given the symbol G) is an exponential decaying function of the correlator time delay τ,

$$G(\tau)=A[1+Be^{-2\Gamma\tau}] \quad (2.37)$$

Where A is the baseline of the correlation function, B is the intercept of the correlation function. Γ is the first cumulant and is related to the translational diffusion coefficient as, $\Gamma=Dq^2$, where q is the scattering vector and its magnitude is defined as,

$$q=\left(\frac{4\pi n}{\lambda_0}\right)\sin\left(\frac{\theta}{2}\right) \quad (2.38)$$

Where n is the refractive index of dispersant, λ_0 is the wavelength of the laser and θ, the scattering angle. For polydisperse samples, the equation can be written as,

$$G(\tau)=A\left[1+B\left|g^{(1)}(\tau)\right|^2\right] \quad (2.39)$$

Where the correlation function $g^{(1)}(\tau)$ is no longer a single exponential decay and can be written as the Laplace transform of a continuous distribution G(Γ) of decay times,

$$g^{(1)}(\tau)=\int_0^\infty G(\Gamma)e^{(-\Gamma\tau)d\Gamma} \quad (2.4)$$

The size distribution of the particles is obtained by fitting, using non-linear least square fitting or CONTIN program, the autocorrelation function to multi-exponential function. The size distribution obtained is a plot of the relative intensity of light scattered by particles and is therefore known as an intensity size distribution. However, in the intensity distribution graph, the area of the peak for the larger particle appears at least 10^6 times larger than the peak for the smaller particle. This is because large particles scatter much more light than small particles, as the intensity of light scattering by a particle is proportional to the sixth power of its diameter (Rayleigh's approximation).

2.2. Enzyme Kinetics:

2.2.1. Michaelis-Menten Kinetics:

A classical enzyme catalyzed reaction contains two elementary reactions in which the substrate forms a complex with the enzyme to produce product and enzyme. The equation can be expressed as followed:

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E+P \tag{2.31}$$

Here E, S, ES & P denote the enzyme, substrate, enzyme-substrate complex and product respectively. So, the rate of enzyme-substrate complex production i.e. ES complex is the difference between its rate of formation and the rate of its decomposition to form the product.

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] \tag{2.42}$$

Now two assumptions can be made,

(a) **Assumption of equilibrium:** It can be assumed that $k_{-1} >> k_2$, so that the first step of the reaction achieves equilibrium.

$$k_s = \frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]} \tag{2.43}$$

Here, k_s is the dissociation constant for the first elementary reaction step.

(b) **Assumption of steady state:** In this assumption the concentration of ES complex is thought to be remained at equilibrium, i.e., the rate of synthesis of ES is exactly equal with the rate of decomposition of ES complex. That means [ES] maintains a steady state.

$$\frac{d[ES]}{dt} = 0 \tag{2.44}$$

To derive the kinetic equation the overall reaction must be converted to measurable quantities. The total amount of enzyme, i.e., [E] can be written as

$$[E]_T = [E] + [ES] \tag{2.44}$$

Now the combination of equation (2.42) & equation (2.43) yields equation (2.46).

$$k_1([E]_T - [ES])[S] = (k_{-1} + k_2)[ES] \tag{2.46}$$

Rearranging and dividing both side by k_1 we get

$$[ES] = \frac{[E]_T[E]}{K_M + [S]} \tag{2.47}$$

Where K_M is known as Michaelis constant.

The initial velocity v_0 can be expressed in the terms of measurable quantities $[E]_T$ and [S]:

$$v_0 = \left(\frac{d[P]}{dt}\right) = k_2[ES] = \frac{k_2[E]_T[S]}{K_M + [S]} \tag{2.48}$$

The maximal velocity can be denoted as V_{max} and be expressed as:

$$V_{max} = k_2[E]_T \tag{2.49}$$

Combination of equation (2.48) & equation (2.49) results into equation (2.50).

$$v_0 = \frac{V_{max}[S]}{K_M + [S]} \tag{2.50}$$

This expression is known as Michaelis-Menten equation, which is a basic equation of enzyme kinetics.

To determine the value of V_{max} and K_M more accurately Hans Lineweaver and Dean Burk formulated a reciprocal equation of equation (2.50) known as Lineweaver-Burk Plot or Double-reciprocal Plot:

$$\frac{1}{v_0} = \left(\frac{K_M}{V_{max}}\right)\frac{1}{[S]} + \frac{1}{V_{max}} \tag{2.51}$$

If v_0 and [S] are plotted in this equation then the slope of the line will be equal with K_M/V_{max}, the intercept at Y axis will be $1/V_{max}$ and intercept at X axis will be $-1/K_M$ (Figure 2.4)

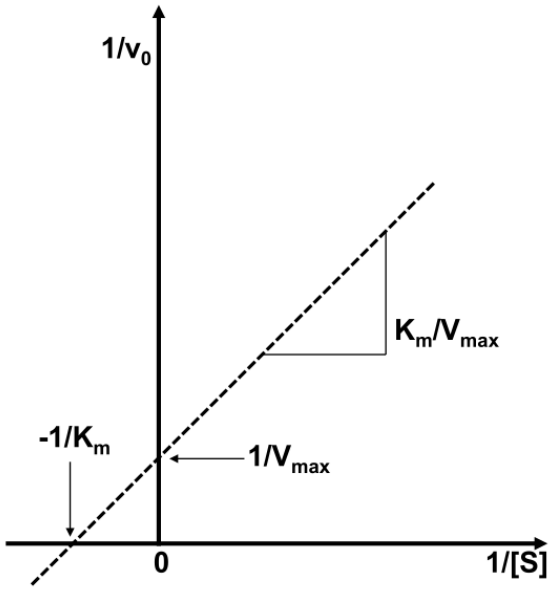

Figure 2.4. *Lineweaver-Burk plot or Double reciprocal plot of enzyme kinetics. $1/v_0$ is plotted against 1/[S]. The slope is K_m/V_{max}. The intercept on the vertical axis is $1/V_{max}$ while, the intercept on the horizontal axis is $-1/K_m$.*

2.2.2. Catalytic Efficiency:

Another kinetic parameter of enzyme to measure is the catalytic efficiency i.e. k_{cat}/K_M. Here k_{cat} is catalytic constant and can be defined as:

$$k_{cat} = \frac{V_{max}}{[E]_T} \tag{2.52}$$

This constant is also known as turnover number as it is the number of turnover that each active site catalyzes per unit time.

When $[S] << K_M$, the formation of ES is very low and subsequently $[E] \approx [E]_T$, so the equation (2.32) becomes a second order rate equation (equation (2.53)):

$$v_0 \approx \left(\frac{k_2}{K_M}\right)[E]_T[S] \approx \left(\frac{k_{cat}}{K_M}\right)[E][S] \tag{2.53}$$

This k_{cat}/K_M is the apparent second-order rate constant. The rate of this reaction directly proportional with the chances of encounter of enzyme and substrate in the solution. So, this quantity k_{cat}/K_M actually determine the catalytic efficiency.

2.3. Computational Techniques:

2.3.1. Homology Modelling:

Homology modelling, also known as comparative modelling is the most accurate method for creating reliable structural models based on its native sequence and template coordinates. The process of homology modelling generally consist of four different steps: identification of the target, sequence alignment, building of the 3D model, and model refinement [7]. Over the past decade, different programs have been developed. Modeller, is one of the most popular homology modelling program packages, Modeller calculates the non-hydrogen atoms to generate the 3D-model of the query sequence [8], PRIMO (Protein Interactive Modelling) is a pipeline for monomer homology modelling. It also provides functionality for modelling protein-ligand complexes [9]. SWISS-MODEL is a server based homology modelling platform, the modelled 3-D structure are also validated by the SWISS-MODEL workspace [10]. The modelled structure is subjected to verification. Usually, Ramachandran Plot also known as a [φ,ψ] plot [11]. Tools such as PRO-CHECK [12] utilizes the Phi/Psi Ramachandran plot to validate the modelled structure also, Verify-3D [13] which estimates the number of residues in favourable environment and ERRAT [14] which evaluates and compare the modelled structure in terms of non-bonded atomic interactions.

Generally, the 3-D model structure of a protein is generated with the following steps: (i) appropriate template search using NCBI BLAST [15], (ii) sequence alignment, (iii) analysis and alignment corrections, functionally conserved residues are aligned, (iv) 3-D modelling, (v) structural optimization by energy minimization, and (vi) structural validation.

2.3.2. Molecular Docking:

Proteins are the favoured target for various bioactive molecules for the treatment of diseases in humans, animals and plants. The first protein crystal structure was solved in 1958, however, its implication in drug discovery was not realized until later [16]. Recognition and binding of such small molecules known as ligands are critical to protein function or biomolecular function in general. In the early 80s, protein structure determination became an accepted tool to facilitate drug discovery [17]. Molecular docking is a computational tool that attempts to screen and predict the binding of such ligands to their respective receptors. The first docking program DOCK was introduced in 1982 [18]. The developed program was relatively simple, and it relied on shape complementarity to explore the binding site of a rigid ligand. Over the past decade,

many docking algorithms have been developed, including Glide [19], Fred [20], AutoDock [21], AutoDock Vina [22], and Gold [23] that not only explores the conformational space of ligand at run time but also allows partial flexibility of the receptor side chains in the binding site. In any docking algorithm, two issues must be considered paramount, (a) the accuracy of binding conformation prediction (search algorithm), and (b) rapid and efficient calculation of binding affinity using reasonable resources (scoring function). The search algorithm is used to generate conformation, position and orientation of the ligand within the active site and to orient the amino acid side chains in the binding cavity as well. Whereas, the scoring function is used to estimate binding affinity in order to identify the most likely pose for an individual ligand and to assign a priority order to a set of diverse ligands docked to the same protein [17]. In short, molecular docking predicts ligand-receptor bound conformations and binding affinity. Molecular docking plays an important role in screening potential drug molecules, drug designing and protein structure-function research.

Several optimization algorithms including genetic algorithm [24], fragment based method [25], simulated annealing [26], and iterated local search global optimizer [27, 28] can combine with the local optimization protocol and speed up the optimization process. Genetic algorithm mimics the process of evolution [29]. In this algorithm, the search space, which is defined as the area either covering the binding site or the whole of the protein, is occupied with ligands carrying mutations (active torsions, rotation and translation in three dimensions) and only the best affinity conformation (i.e. the fittest conformer) is selected. In the fragment based method the ligand is fragmented and reconstructed in the active site/ligand binding site. While the iterated local search global optimizer uses successive steps of mutation and local optimization, with each step accepted according to the Metropolis criterion [22].

In this thesis, we have used AutoDock Vina (Version 1.1.2) as the standard docking program. AutoDock Vina uses the gradient optimization method [22]. While Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [30] is used as the local optimization algorithm, which is an efficient quasi-Newton method. The BFGS method not only uses the scoring function but also its gradient. i.e. the derivatives of the scoring function with respect to its arguments. The arguments include the position and the orientation of the ligand, also the torsion of the active rotatable bonds in the ligand and any flexible residue. AutoDock Vina uses a modified semi-empherical scoring function to calculate the protein-ligand binding free energies [22]. This method combines classical force fields with empherical parameters to estimate and

rank bound conformers rapidly. The general functional form of the conformation-dependent part of the scoring function of AutoDock Vina is given by,

$$c = \sum_{i<j} f_{t_i t_j}(r_{ij}) \tag{2.54}$$

Where the summation is considered over all the pair of atoms that can move relative to each other. Here, each atom of the molecule *i* is assigned a type *ti*, and a symmetric set of interaction functions *ftitj* having interatomic distance *rij* is defined. This value is the sum of the inter and intramolecular contributions given as,

$$c = c_{\text{inter}} + c_{\text{intra}} \tag{2.55}$$

The free energy of binding is predicted and calculated from the intermolecular part of the lowest scoring conformation, and is designated as:

$$s_1 = g(c_1 - c_{\text{intra1}}) = g(c_{\text{inter 1}}) \tag{2.56}$$

Where *g* is an arbitrarily increasing non-linear function. The other low scoring conformations are denoted an *s* value using the c_{intra} of the best binding mode, and is given as,

$$s_i = g(c_i - c_{\text{intra1}}) \tag{2.57}$$

The simplified scoring function of AutoDock Vina estimates the free energy of binding, while the hydrogen bonding and the hydrophobic contribution are estimated as relative weights in the output.

2.3.3. Solvent Accessible Surface Area (SASA):

The concept of solvent accessibility of residues in globular proteins was first introduced by Lee and Richards in 1971 [31]. Solvent accessible surface area (SASA) is a measure of formation of contacts between the atoms on the surface of the protein and the solvent molecules. It represents the locus of the center of the solvent molecule as it rolls along the protein, making the maximum permitted van der Waals contacts without penetrating any other atom [31]. Based on SASA values subunits of a biomolecule such as amino acid residues of a protein can be classified as buried or exposed. A direct estimation of SASA of biomolecules such as proteins experimentally is not possible. But, SASA of the native protein can be calculated computationally from atomic coordinates [32]. Several different algorithms have been

developed over the years for calculating SASA. Z-layer integration [31], Intersection method [33], The Shrake and Rupley Algorithm [34], Linear Combinations of Pairwise Overlaps (LCPO) method [35], and Power diagram model [36]. Several tools and software programs have been developed to estimate SASA. PDBePISA [37, 38], ProtSA [39], GETAREA [40] and, Define Secondary Structure of Protein (DSSP) [41]. During our studies we used the GETAREA integration of PyMOL (Educational Version) for calculating SASA.

2.3.4. Molecular Dynamics (MD) Simulation:

First developed by Fermi in 1955 molecular dynamics (MD) simulation is one of the primary tools for studying biomolecular motion. In an MD simulation particle interact with each other over a specified time period, producing a trajectory, which can provide insights into the dynamic behavior of the system. MD simulation on biological system were first performed by Levitt and Warshel [42], while first MD simulation of macromolecule of biological interest was investigated by Karplus and McCammon [43] on bovine pancreatic trypsin inhibitor. During an MD simulation two types of atomic forces govern the molecular motions: forces resulting from the interactions of chemically bonded atoms and the forces arising due to the non-bonded atoms. The MD potential can describe bonded interactions (bond lengths, bond angles and dihedral angles) and non-bonded interactions (electrostatic and van der Waal interactions). Each MD run typically consists of minimization of the molecule of interest with respect to the underlying potential. MD simulation solves Newton's equation of motion for the interacting molecules (e.g., enzymes). Every particle in the system is assigned an initial velocity and acceleration, then the new position of the particle is calculated after time *t* using Newton's equations. The force on each particle is calculated using the new position of the particles from the slope of the potential, while the acceleration is calculated from the force. This is repeated for a length of time, which provides a trajectory describing the dynamics of all the atoms of the system. GROMACS [44], NAMD [45], TINKER [46, 47], and Desmond [48, 49] are some of the most commonly used programs used for performing MD simulations. Most of the following concepts are interpreted in the framework of Schrodinger Maestro 2018-1 (Academic Release) unless stated otherwise.

2.3.4.1. Integration Algorithm:

MD simulation generates a set of coordinates and velocities as a function of time for a given set of atoms/particles in a system. When considering a simple atomic system, Newton's second law of motion can be written in the following form:

$$\boldsymbol{F} = m\frac{d^2r}{dt^2} \tag{2.58}$$

Here, *m* is the mass of the atom and *r* is the positional coordinate. Integrating the above equation at a particular time step, the position and velocities of the next frame which are separated by a small-time interval δt can be calculated. Generally, the equation of motion is integrated by applying a finite difference method. Various algorithms have been developed and used to integrate the above equation of motion. For any algorithm to be a reliable integrator, it must be, (a) accurate to the trajectory of particle motion, (b) energetically stable and, (c) robust, allowing for longer propagation time of the system in phase space. Verlet algorithm [50], velocity-Verlet algorithm [51], leap-frog algorithm [52], RESPA algorithm [53, 54] are few of the most popular. For our purpose, the RESPA algorithm was used as the integration scheme.

2.3.4.2. Force Field:

The term "force field" denotes to the functional form, the mathematical formula and associated parameters employed to calculate the energy and the forces of the molecular system. The associated parameters include the values for the different degrees of freedom, the partial atomic charges on each atom for different environment. Each of the parameter are a result of extensive quantum mechanical calculation. The basic component of a typical forcefield contains bond, angle, torsion and non-bonded terms. The parameters and the functional form are used to calculate the potential energy of the system.

The basic functional form common for the bonded part of the force field are described by the following equation:

$$V(\boldsymbol{r}^N) = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\in_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}\right) + \sum_{\text{bonds}} k_b(b-b_0)^2 + \sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 + \sum_{\text{dihedrals}}\sum_{n=1}^{N} k_\varphi^{(n)}[1+\cos(n\varphi-\delta)] + \sum_{\text{impropers}} k_\omega(\omega-\omega_0)^2 \tag{2.59}$$

Where, $V(r^N)$ or the total potential defines the function of coordinates of *N* sites of systems. The first term of the equation represents the van der Waals (VDW) interaction between atoms (I and j), having interatomic distance r_{ij}. The VDW force originates due to the repulsive and attractive forces, which yields an energy minimum at the interatomic distance. As the atoms are brought closer from infinite, the negative term ($1/r^6$) in the square prevails initially, as a result the energy gradually becomes negative. As the distance between atoms becomes closer

to the energy minimum, the positive term ($1/r^{12}$) prevails and the energy becomes positive. ϵ_{ij} of the equation represents depth of the potential well, while σ_{ij} defines the interatomic distance at which the potential energy becomes zero. The second term of equation (2.54) shows the non-bonded terms as defined by Coulomb's law owing to the presence of point charges on different interacting sites on the system. The atoms with charge q_i and q_j are separated an interatomic distance r_{ij}. The third part of the equation (2.54) represents the potential energy for bond energy or bond vibration is modelled by harmonic potential; here, b denotes the bond length at any given time and b_0 represents the equilibrium bond length, while k_b is the force constant. Likewise, the fourth term of the above equation (2.54) represents the change in energy during bending motion of the molecule. The term k_θ is the force constant, while θ and θ_0 indicate the angle formed by three consecutive atoms and its equilibrium value. The fifth term of the equation calculates the potential energy change due to torsional or dihedral motion. Herein, φ is the torsion angle and k_φ denotes the height of rotational energy barrier while n indicates the minimum number of points in the function during a 360^o rotation of the bond. Phase factor δ decides the torsion angle passing through the minimum values. The cosine function in the dihedral term represents periodicity of the function. The improper term in the equation (2.54) denotes the improper dihedral potential with improper torsion or dihedral angle ω and equilibrium value ω_0, while k_ω represents the force constant. All of the popular force fields are of pair-additive form, the total potential energy is the sum of the pairwise bonded and the non-bonded terms [55].

AMBER [56], CharmM [57], OPLS [58], GROMOS [59] etc. are among the most popular force fields. The validity of results from a simulation event depends upon the quality of force field, application of the selected force field to the system, and ability to calculate the properties. In our work we adopted the OPLS force field for the biomolecule and SPC models of water for running MD simulation [60, 61].

2.3.4.2.1. OPLS Force Field:

One of the most popular force field is the OPLS (Optimized Potential for Liquid Simulations). The force field is based on the widely used AMBER force field. The functional form of the OPLS force field is defined as under:

$$V(\{\vec{r}_N\}) = \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{torsions} \left\{ \frac{V_1}{2}[1 + \cos(\phi)] + \frac{V_2}{2}[1 - \cos(2\phi)] + \frac{V_3}{2}[1 + \cos(3\phi)] + \frac{V_4}{2}[1 - \cos(4\phi)] \right\} + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_i j} \right)^{12} + \left(\frac{\sigma_{ij}}{r_i j} \right)^{6} \right] + \frac{q_i q_j e^2}{r_{ij}} \right\} f_{ij} \quad (2.60)$$

Here, the angles and bonds are described by harmonic potential. At room temperature, they fluctuate slightly around the equilibrium and are very strong. The cosine function describes the dihedral potential, the value of which depend on the height of the barrier between the low energy conformations and may take any value within 360°. Generally, dihedral potential has a symmetry around 180°. The long-range interactions are calculated for atoms which are three or more bonds apart. The long-range interactions consist of Coulomb and Lennard Jones two-body interaction terms where, the Lennard Jones potential is described as the combination of attractive van der Waals forces due to the dipole-dipole interactions and the empirical repulsive forces due to Pauli's repulsion [56, 62].

2.3.4.3. Boundary Conditions:

Prior to any MD simulation event, an appropriate boundary condition must be set to mimic the system of interest. Periodic boundary condition is one of the most commonly employed boundary conditions. Such a boundary condition not only maintains a fixed number of atoms/particles in the simulation box but also creates a bulk environment by removing the edge effects.

2.3.4.3.1. Periodic Boundary Condition (PBC):

Most common method of minimizing edge effects in a finite system is to use the periodic boundary condition. Atoms or particles to be simulated are put in a cubic, orthorhombic or triclinic box. The box consists of repeating units translated identically in the x, y and z directions. The periodic boundary condition affects not only the computation of the forces, but also the atomic position. PBC is implemented with minimum image convention, only the closest image of the particle is chosen for calculating the short range non-bonded interactions with other particles. A cut-off is normally set for short range non-bonded interactions, while

the long-range van der Waals interaction which are beyond the cut-off are ignored. The long-range electrostatic interactions are usually recovered with Particle-Mesh-Ewald method [63].

During a simulation event, it is necessary to keep the atoms in the central computational box. Periodic boundary condition ensures that under any condition if an atom leaves the central box, it re-enters from the opposite side with an identical velocity at the translated image position of the box.

2.3.4.4. Energy Minimization:

The initial structure used for a simulation in a molecular system may possess poor interactions. Hence, it is required to run an energy minimization of the structure before initiating an MD simulation to remove the poor and bad interactions. The minimization algorithm for a first defines the energy of the system as a function of coordinates. In addition to the energy terms external restraining terms are also included. The conformation is adjusted lowering the value of the target function. Generally, there are three energy minimization method, namely (a) steepest descent (SD), (b) conjugate gradient (CG), and (c) L-BFGS algorithm. SD algorithm [64] updates the position of the particles using the potential energy and forces in an iterative way. Compared to SD algorithm, CG [65] is more efficient when the system used is closer to the energy minimum. The limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [66, 67] which can approximate the original BFGS algorithm using the inverse Hessian matrix but with limited memory requirement. BFGS algorithm converges faster than the conjugate gradient (CG) and is more efficient in free energy calculations. The efficiency of minimization is dictated not only by the time required to evaluate the target function, but also the number of iterations or structural adjustments required to converge to the minimum. During our work a hybrid steepest descent (SD) and limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm was used during energy minimization.

2.3.4.5. Ensemble:

MD simulations at non-zero temperatures require a statistical representative set of configurations, known as ensemble. All of which are replicas of the system of interest, these include macroscopic parameters like number of particles in the system (N), volume of the system (V), pressure (P), temperature (T), etc. The ensembles differ from each other by the assigned coordinates and the particle momenta in the system. The system properties are defined as ensemble averages. For any multi-particle system, the averaging will include several degrees of freedom. Thus, can be carried out over a part of the configuration space. Smaller

configuration space lead to better ensemble average. During a simulation event only, those degrees of freedom are included on which the property to be calculated is dependent.

2.3.4.5.1. Microcanonical (NVE) Ensemble:

A system where the ensemble has the same number of particle (N), same volume (V), and same total energy (E), such an ensemble is referred to as microcanonical ensemble (NVE) and represents a real system. NVE ensemble can either exchange heat nor matter with the surrounding.

2.3.4.5.2. Canonical (NVT) Ensemble:

A collection of systems characterized by same number of particles (N), same volume (V) and same temperature (T). There are several algorithms for temperature coupling including the Berendsen [61], and Nosé-Hoover thermostats [68, 69].

In the Berendsen algorithm, the temperature deviation from the reference T_0 is corrected according to:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \tag{2.61}$$

Where, τ is the coupling parameter.

Even though the Berendsen algorithm is easy to implement. However, kinetic energy fluctuations are suppressed by the thermostat suggesting that a correct canonical ensemble is not generated [61].

The Nosé-Hoover thermostat algorithm, the Hamiltonian of the system can be shown as under [68, 69]

$$\mathcal{H}(P, R, p_s, s) = \sum_i \frac{\mathbf{p}_i^2}{2ms^2} + \frac{1}{2}\sum_{ij, i\neq j} U(\mathbf{r_i} - \mathbf{r_j}) + \frac{p_s^2}{2Q} + gkT\ln(s) \tag{2.62}$$

Here, g indicates the number of independent momentum degrees of freedom of the system. While R and P refers to the coordinates of r_i and P_i. s denotes the degree of freedom for heat bath.

2.3.4.5.3. Isothermal-Isobaric (NPT) Ensemble:

System characterized by this ensemble have same number of particle (N), same pressure (P) and same temperature (T). Similar to a thermostat, a barostat is utilized to control the pressure

of the system. In an NPT ensemble, the Berendsen algorithm scales the atomic coordinates, while the Parrinello-Rahman barostat [70] or the Martyna-Tobias-Klein barostat [71, 72] with an extended ensemble of Nosé-Hoover thermostat is used. For our work, NPT was the favoured ensemble type, using a Nosé-Hoover thermostat with Martyna-Tobias-Klein barostat during the simulation events.

2.3.4.6. Constraints:

During an MD simulation constraint are applied to satisfy the Newtonian motion of a rigid body. Bond vibration or angle vibration is constrained during a simulation to its equilibrium value using a constraint algorithm. A constraint algorithm maintains the distance between the mass points during a simulation event. SHAKE algorithm [73], LINCS algorithm [74], M-SHAKE algorithm [75] are a few of the most popularly used constraint algorithms.

2.3.4.7. MD Trajectory:

MD trajectory is a series of conformation, which are calculated as a function of time under specified conditions (temperature, pressure, volume, etc.) from the desired ensemble as a result of equilibration and production run for a sufficient time.

2.3.4.7.1. MD Trajectory Analysis:

The preliminary trajectory analysis focuses on structural mobility and conformational changes. The analysis is represented in three ways, (a) the root mean square deviation (RMSD), (b) root mean square fluctuation (RMSF) and, (c) radius of gyration (R_g).

Root mean square deviation (RMSD) is a measure of the average distance between a group of atoms. The RMSD calculated from a given MD trajectory indicates the overall change in the geometry of the molecule with respect to its initial structure. RMSD is calculated by fitting the MD trajectory to the initial conformation using the following equation.

$$RMSD = \frac{1}{N}\left(\sum_{i=1}^{N} \|r_i(t) - r_i(0)\|^2\right)^2 \tag{2.63}$$

Where N is the number of atoms. While $r_i(t)$ and $r_i(0)$ are the coordinates of the trajectory and the reference i^{th} atoms.

Root mean square fluctuation (RMSF) is the time average of root mean square deviation (RMSD). It is calculated using the following equation.

$$RMSF = \sqrt{\langle (\mathrm{r}_i - \langle \mathrm{r}_i \rangle)^2 \rangle} \quad (2.64)$$

RMSF measures the average deviation of the atoms/particles over time with respect to the reference position, which is usually the time-averaged position of the atoms/particles. It allows for the detection of the regions of high fluctuation in a molecule (e.g., protein, DNA, etc.).

Radius of gyration (R_g) refers to the average distance of atoms of a molecule (e.g., protein, DNA, etc.) from their common centre of mass. R_g provides insight into the overall shape and compactness of a molecule. The radius of gyration is calculated using the following equation.

$$R_g^2 = \frac{1}{M}\sum_{i=1}^{N} m_i(\mathbf{r}_i - \mathbf{R})^2 \quad (2.65)$$

Where $M = \sum_{i=1}^{N} m_i$ denotes the total mass. $\mathbf{R} = N^{-1}\sum_{i=1}^{N} \mathbf{r}_i$ refers to the centre of mass of the molecule consisting of *N* number of atoms.

2.3.5. Normal Mode Analysis (NMA):

The normal mode analysis (NMA) is used to characterize motions of oscillating systems near equilibrium state and is based on vibrational modes of molecules. Generally, normal mode is defined as a motion where all the particles in a system are in the same phase and move with the same frequency. A molecule (e.g., biomolecule) can vibrate only in certain way, by the virtue of its structure. Such vibrations denote the accessible conformation of any molecule (e.g., protein, DNA, etc.) and strongly correlate with the functional dynamics of the molecule. In short, NMA is used to identify the overall pattern of functional dynamics of a molecule. The NMA is calculated by applying the principal component analysis (PCA). First proposed by Karl Pearson [76], PCA is a statistical method to study multivariate data. PCA takes the MD trajectory and extracts the dominant modes of motion. Normal mode analysis is calculated from covariance matrix of the MD simulation trajectory. Mostly, a large proportion of the overall fluctuation observed from RMSF analysis are accounted for by a few low frequency eigenvector with large eigenvalues. When the motions are similar, the eigenvectors and eigenvalues from an MD trajectory must be similar to each other. Hence, the MD trajectories are filtered along each eigenvector to identify the dominant motion of a molecule (e.g., protein, DNA, etc.) during a simulation event. The calculation of the principal component is calculated by a 3N by 3N dimensional Hessian matrix. It is a matrix of force constant describing the force

experienced by the atoms. The eigenvectors for this matrix are denoted to as normal mode. Each of the normal mode describes the displacement of atoms from their equilibrium position along X, Y and Z directions. The eigenvalues obtained correspond to the square of the frequency of the mode. The low frequency slow modes represent the collective motions (e.g., inter domain motions in protein) while the fast modes describe the local motions (e.g., side chain vibrations in protein). For our work, we used the Normal Mode Wizard of VMD (1.9.3) program [77] using ProDy [78].

2.3.6. Steered Molecular Dynamics (SMD):

Molecular dynamics (MD) simulation is the method of choice for understanding molecular interactions, interaction mechanism and motions. However, due to limited simulation timescale, which ranges from nanoseconds to microseconds in MD simulation, results obtained from ergodic systems are only considered as useful. Steered molecular dynamics (SMD) induces unbinding of ligands and conformational changes in biomolecules on time scales accessible to molecular dynamics simulations [79]. In SMD, a time-dependent external force is applied on a ligand to facilitate its unbinding from a protein. Analysis of the interactions formed by the dissociating ligand, as well as the analysis of the applied force and ligand position can yield important regarding structure-function-dynamics relationship, binding pathways and underlying mechanisms of substrate recognition by enzymes. SMD can also be applied to investigate the molecular mechanisms that determine elastic properties exhibited by proteins subjected to deformations in AFM and optical tweezer experiments such as stretching of titin leading to unfolding of its immunoglobulin domains [80] or, or stretching of tenascin which results in unfolding of its fibronectin-III domains [81]. Besides yielding qualitative information, SMD can also yield quantitative information about the binding potential of the ligand-receptor complex [82]. One way to apply external forces to a protein-ligand complex is to restrain the ligand to a point in space (restraint point) by an external, e.g., harmonic, potential. The restraint point is then shifted in a chosen direction [83-87], forcing the ligand to move from its initial position in the protein and allowing the ligand to explore new contacts along its unbinding path. Let us assume a single reaction coordinate x, and an external potential $U = K(x - x_0)^2/2$, where K is the stiffness of the restraint, and x_0 is the initial position of the restraint point having a constant velocity v, the external force applied on the system is expressed as,

$$F = K(x_0 + vt - x) \quad (2.66)$$

This corresponds to the force, the ligand is being pulled with stiffness *K* having a velocity *v*. Alternatively, Alternatively, a fixed restraint point at a distance much larger than the length of the unbinding pathway may be chosen. In this case, the end of the spring does not move and its stiffness is linearly increased with time [82], i.e., $K = \alpha t$, and the force is,

$$F = \alpha t(x_0 - x) \quad (2.67)$$

External forces such as constant force, or torques can also be used to induce rotational motion to a biomolecule. Also, ligand-biomolecule binding free energy can be estimated by calculating the potential of mean force (PMF). The binding free energy is calculated by the Jarzynski's equation [88] or weighted histogram analysis method (WHAM) [89].

2.4. Systems and Molecular Probes:

2.4.1. Enzyme System:

Two types of enzymes; β-Glucosidase (BGL) from almonds and α-Chymotrypsin (CHT) have been used in our study.

2.4.1.1. β-Glucosidase (BGL):

β-Glucosidase (BGL) are a subgroup of β-D-glucoside glucohydrolases (EC3.2.1.21). BGLs are among the most widely studied enzymes due to its universal distribution and well defined wide variety of substrate and simple nature of enzyme assay [90]. Structurally, BGL is characteristically similar to family 1 glycosyl hydrolase. The catalytic domain consists of alternating α-helices and β-sheets arranged in a TIM barrel fold [(α/β)8] [91] (Figure.2.5) also known as a β-barrel fold. BGLs are dyads and catalyse the transfer of glycosyl groups between oxygen nucleophiles. These transfer reaction results in the hydrolysis of beta-glucosidic linkage present between carbohydrate residues in aryl-amino-, or alkyl-beta-D-glucosides, cyanogenic-glucosides, short chain oligosaccharides and disaccharides [92]. BGL is an industrially important enzyme. Playing pivotal roles in biotechnological processes such as hydrolysis of isoflavone glucosides and production of bio-ethanol. BGLs are integral part of beverage industry, including wine, tea and juice industry. Beside industries, BGLs are pivotal to several physiological process. β-glucosidase is involved in cellulase induction (due to its transglycosylation activities) and cellulose hydrolysis. In plants, the enzyme is involved in

beta-glucan synthesis during cell wall development, pigment metabolism, fruit ripening, and defense mechanisms whereas, in humans and other mammals, BGL is involved in the hydrolysis of glucosyl ceramides. So defects in BGL activity in humans are associated with Gaucher's disease, a non-neuropathic lysosomal storage disorder [92, 93].

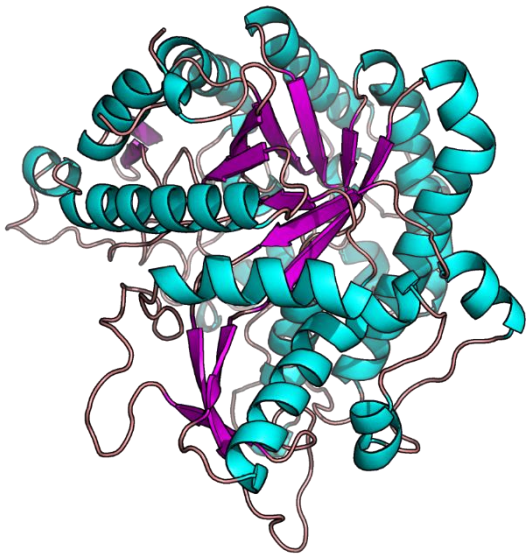

Figure 2.5. *Structure of almond β-Glucosidase. The different colours indicate the secondary structures.*

2.4.1.2. α-Chymotrypsin (CHT):

α-Chymotrypsin (Figure 2.6) isolated from bovine pancreas, is a member of the family serine endopeptidase (molecular weight of 25,191 Da) catalysing the hydrolysis of peptides in the small intestine [94]. The three dimensional structure of CHT was solved by Birktoft and Blow [95]. The molecule is a three-dimensional ellipsoid of dimensions 51 × 40 × 40 Å. And comprises of 245 amino acid residues. CHT contains several antiparallel β-pleated sheet regions and little α-helix. Other than playing vital role in protein digestion, CHT plays pivotal role in immune response and insect molting. CHT is a widely used oral proteolytic enzyme combination to hasten repair of traumatic, surgical, and orthopedic injuries. It shows high bioavailability without losing its biological activities as an anti-inflammatory, anti-edematous, fibrinolytic, antioxidant, and anti-infective agent [96]. Industrially, CHT is mainstay of the

meat industry for meat tenderization [97], brewery for chillproofing [98] and pharmaceutical industry as a digestive supplement [99].

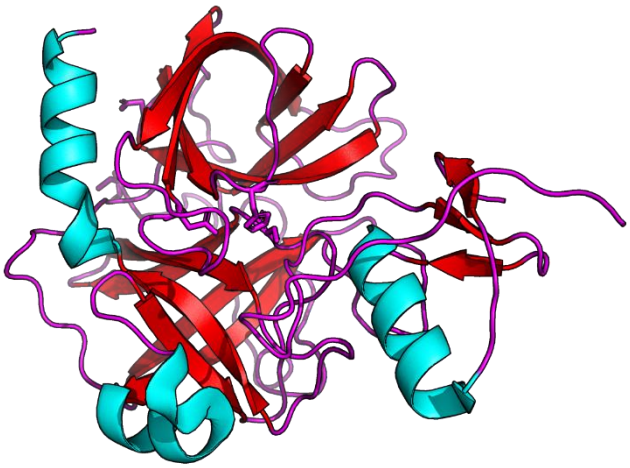

Figure 2.6. *Structure of α-Chymotrypsin. The different colours indicate the secondary structures.*

2.4.2. Molecular Probes:

2.4.2.1. 8-Anilino-1-Naphthalenesulfonic Acid Ammonium Salt (ANS):

ANS is a well-known molecular probe which binds selectively to the native state of certain proteins and enzymes in their hydrophobic as well as polar sites (Figure 2.7). In aqueous solution, the emission quantum yield of ANS is very small (0.004) with emission peak at ~520 nm and a lifetime of ~0.25 ns. The steady-state emission is quenched dramatically in polar solvents. Because of its bichromophoric structure, ANS undergoes charge transfer (CT) from one aromatic moiety to the other ring and solvation. In case of nonpolar solvents, the emission is strong and is mostly from the locally excited state i.e., before charge separation. In polar solvents, the fluorescence decreases and is dominated by emission from the CT state. The solvent polarity and rigidity determine the wavelength and yield of emission and that is why ANS is a useful biological probe [100].

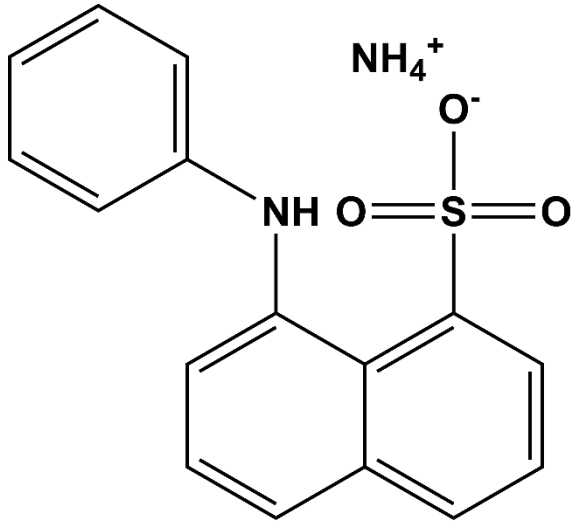

Figure 2.7. *Schematic representation of the molecular probe 8-Anilino-1-Naphthalenesulfonic Acid Ammonium Salt (ANS).*

References

[1] A.S.R. Koti, M.M.G. Krishna, N. Periasamy, Time-resolved area-normalized emission spectroscopy (TRANES): a novel method for confirming emission from two excited states, *J. Phys. Chem. A*, 105 (2001) 1767-1771.

[2] R.D. Spencer, G. Weber, Measurements of subnanosecond fluorescence lifetimes with a cross-correlation phase fluorometer, *Ann. N. Y. Acad. Sci.*, 158 (1969) 361-376.

[3] L.A. Philips, S.P. Webb, J.H. Clark, High-pressure studies of rotational reorientation dynamics: The role of dielectric friction, *J. Chem. Phys.*, 83 (1985) 5810-5821.

[4] B. Kalman, N. Clarke, L.B.A. Johansson, Dynamics of a new fluorescent probe, 2,5,8,11-tetra-tert-butylperylene in solution, cubic lyotropic liquid crystals, and model membranes, *J. Phys. Chem.*, 93 (1989) 4608-4615.

[5] A. Rodger, B. Nordén, Circular dichroism and linear dichroism, *Oxford University Press, USA*, (1997).

[6] N. Berova, K. Nakanishi, R.W. Woody, Circular dichroism: principles and applications, *John Wiley & Sons, USA*, (2000).

[7] A. Meier, J. Söding, Automatic prediction of protein 3D structures by probabilistic multi-template homology modeling, *PLoS Comput. Biol.*, 11 (2015) e1004343.

[8] A. Fiser, A. Šali, Modeller: generation and refinement of homology-based protein structure models, *Methods Enzymol.*, 374 (2003) 461-491.

[9] R. Hatherley, D.K. Brown, M. Glenister, Ö. Tastan Bishop, PRIMO: an interactive homology modeling pipeline, *PLoS One*, 11 (2016) e0166698.

[10] L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey, T. Schwede, Protein structure homology modeling using SWISS-MODEL workspace, *Nat. Protoc.*, 4 (2009) 1-13.

[11] C. Ramakrishnan, G.N. Ramachandran, Stereochemical criteria for polypeptide and protein chain conformations: II. Allowed conformations for a pair of peptide units, *Biophys. J.*, 5 (1965) 909-933.

[12] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 26 (1993) 283-291.

[13] R. Lüthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature*, 356 (1992) 83-85.

[14] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, *Protein Sci.*, 2 (1993) 1511-1519.

[15] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, 25 (1997) 3389-3402.

[16] J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wyckoff, D.C. Phillips, A three-dimensional model of the myoglobin molecule obtained by x-ray analysis, *Nature*, 181 (1958) 662-666.

[17] U. Pal, Interaction of proteins with small molecules and peptides, *Doctoral dissertation, Jadavpur University, India*, (2016).

[18] I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge, T.E. Ferrin, A geometric approach to macromolecule-ligand interactions, *J. Mol. Biol.*, 161 (1982) 269-288.

[19] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, P.C. Sanschagrin, D.T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes, *J. Med. Chem.*, 49 (2006) 6177-6196.

[20] M. McGann, FRED pose prediction and virtual screening accuracy, *J. Chem. Inf. Model.*, 51 (2011) 578-596.

[21] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comput. Chem.*, 30 (2009) 2785-2791.

[22] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, 31 (2010) 455-461.

[23] M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, Improved protein–ligand docking using GOLD, *Proteins: Struct. Funct. and Bioinf.* 52 (2003) 609-623.

[24] D.E. Goldberg, J.H. Holland, Genetic algorithms and machine learning, *Springer, Berlin/Heidelberg, Germany*, (1988).

[25] B. Kramer, M. Rarey, T. Lengauer, Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking, *Proteins: Struct. Funct. and Bioinf.* 37 (1999) 228-241.

[26] S. Kirkpatrick, C.D. Gelatt Jr, M.P. Vecchi, Optimization by simulated annealing, *Science*, 220 (1983) 671-680.

[27] J. Baxter, Local optima avoidance in depot location, *J. Oper. Res. Soc.*, 32 (1981) 815-819.

[28] C. Blum, A. Roli, M. Sampels, Hybrid metaheuristics: an emerging approach to optimization, *Springer, Berlin/Heidelberg, Germany*, (2008).

[29] D.E. Clark, D.R. Westhead, Evolutionary algorithms in computer-aided molecular design, *J. Comput. Aided Mol. Des.*, 10 (1996) 337-358.

[30] S. Wright, J. Nocedal, Numerical optimization, *Springer Sci.*, 35 (1999) 7.

[31] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, *J. Mol. Biol.*, 55 (1971) 379-400.

[32] S. Ausaf Ali, M.I. Hassan, A. Islam, F. Ahmad, A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states, *Curr. Protein Pept. Sci.*, 15 (2014) 456-476.

[33] K.D. Gibson, H.A. Scheraga, Exact calculation of the volume and surface area of fused hard-sphere molecules with unequal atomic radii, *Mol. Phys.*, 62 (1987) 1247-1265.

[34] A. Shrake, J.A. Rupley, Environment and exposure to solvent of protein atoms. Lysozyme and insulin, *J. Mol. Biol.*, 79 (1973) 351-371.

[35] J. Weiser, P.S. Shenkin, W.C. Still, Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities, *Biopolymers*, 50 (1999) 373-380.

[36] R. Bryant, H. Edelsbrunner, P. Koehl, M. Levitt, The area derivative of a space-filling diagram, *Discrete Comput. Geom.*, 32 (2004) 293-308.

[37] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.*, 372 (2007) 774-797.

[38] E. Krissinel, K. Henrick, Detection of protein assemblies in crystals, *International Symposium on Computational Life Science*, *Springer*, (2005) 163-174.

[39] J. Estrada, P. Bernadó, M. Blackledge, J. Sancho, ProtSA: a web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble, *BMC Bioinform.*, 10 (2009) 104.

[40] R. Fraczkiewicz, W. Braun, Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, *J. Comput. Chem.*, 19 (1998) 319-333.

[41] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers*, 22 (1983) 2577-2637.

[42] M. Levitt, A. Warshel, Computer simulation of protein folding, *Nature*, 253 (1975) 694-698.

[43] J.A. McCammon, B.R. Gelin, M. Karplus, Dynamics of folded proteins, *Nature*, 267 (1977) 585-590.

[44] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: Fast, flexible, and free, *J. Comput. Chem.*, 26 (2005) 1701-1718.

[45] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.*, 26 (2005) 1781-1802.

[46] J.W. Ponder, F.M. Richards, An efficient newton-like method for molecular mechanics energy minimization of large molecules, *J. Comput. Chem.*, 8 (1987) 1016-1024.

[47] J.W. Ponder, C. Wu, P. Ren, V.S. Pande, J.D. Chodera, M.J. Schnieders, I. Haque, D.L. Mobley, D.S. Lambrecht, R.A. DiStasio Jr, M. Head-Gordon, G.N.I. Clark, M.E. Johnson, T. Head-Gordon, Current Status of the AMOEBA Polarizable Force Field, *J. Phys. Chem. B*, 114 (2010) 2549-2564.

[48] K.J. Bowers, D.E. Chow, H. Xu, R.O. Dror, M.P. Eastwood, B.A. Gregersen, J.L. Klepeis, I. Kolossvary, M.A. Moraes, F.D. Sacerdoti, J.K. Salmon, Y. Shan, D.E. Shaw, Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters, SC '06: *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, (2006) 43.

[49] R.A. Lippert, K.J. Bowers, R.O. Dror, M.P. Eastwood, B.A. Gregersen, J.L. Klepeis, I. Kolossvary, D.E. Shaw, A common, avoidable source of error in molecular dynamics integrators, *J. Chem. Phys.*, 126 (2007) 046101.

[50] L. Verlet, Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules, *Phys. Rev.*, 159 (1967) 98-103.

[51] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters, *J. Chem. Phys.*, 76 (1982) 637-649.

[52] R.W. Hockney, The potential calculation and some applications, *Methods Comput. Phys.*, 9 (1970) 136.

[53] M. Tuckerman, B.J. Berne, G.J. Martyna, Reversible multiple time scale molecular dynamics, *J. Chem. Phys.*, 97 (1992) 1990-2001.

[54] D.D. Humphreys, R.A. Friesner, B.J. Berne, A Multiple-Time-Step Molecular Dynamics Algorithm for Macromolecules, *J. Phys. Chem.*, 98 (1994) 6885-6892.

[55] J.W. Ponder, D.A. Case, Force Fields for Protein Simulations, Advances in Protein Chemistry, *Academic Press, USA*, 66 (2003) 27-85.

[56] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.*, 117 (1995) 5179-5197.

[57] A.D. MacKerell Jr, J. Wiorkiewicz-Kuczera, M. Karplus, An all-atom empirical energy function for the simulation of nucleic acids, *J. Am. Chem. Soc.*, 117 (1995) 11946-11975.
[58] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc.*, 118 (1996) 11225-11236.
[59] W.R.P. Scott, P.H. Hünenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Krüger, W.F. van Gunsteren, The GROMOS Biomolecular Simulation Program Package, *J. Phys. Chem. A*, 103 (1999) 3596-3607.
[60] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 79 (1983) 926-935.
[61] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 81 (1984) 3684-3690.
[62] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, *J. Comput. Chem.*, 24 (2003) 1999-2012.
[63] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, *J. Chem. Phys.*, 98 (1993) 10089-10092.
[64] K.B. Wiberg, A Scheme for Strain Energy Minimization. Application to the Cycloalkanes1, *J. Am. Chem. Soc.*, 87 (1965) 1070-1078.
[65] D.F. Shanno, Conjugate Gradient Methods with Inexact Searches, *Math. Oper. Res.*, 3 (1978) 244-256.
[66] R.H. Byrd, P. Lu, J. Nocedal, C. Zhu, A Limited Memory Algorithm for Bound Constrained Optimization, *SIAM J. Sci. Comput.*, 16 (1995) 1190-1208.
[67] C. Zhu, R.H. Byrd, P. Lu, J. Nocedal, Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, *ACM Trans. Math. Softw.*, 23 (1997) 550–560.
[68] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.*, 81 (1984) 511-519.
[69] W.G. Hoover, B.L. Holian, Kinetic moments method for the canonical ensemble distribution, *Phys. Lett. A*, 211 (1996) 253-257.
[70] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 52 (1981) 7182-7190.
[71] G.J. Martyna, M.L. Klein, M. Tuckerman, Nosé–Hoover chains: The canonical ensemble via continuous dynamics, *J. Chem. Phys.*, 97 (1992) 2635-2643.

[72] G.J. Martyna, M.E. Tuckerman, D.J. Tobias, M.L. Klein, Explicit reversible integrators for extended systems dynamics, *Mol. Phys.*, 87 (1996) 1117-1157.
[73] J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 23 (1977) 327-341.
[74] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem.*, 18 (1997) 1463-1472.
[75] S.G. Lambrakos, J.P. Boris, E.S. Oran, I. Chandrasekhar, M. Nagumo, A modified shake algorithm for maintaining rigid bonds in molecular dynamics simulations of large molecules, *J. Comput. Phys.*, 85 (1989) 473-486.
[76] K. Pearson, LIII. On lines and planes of closest fit to systems of points in space, *Philos. Mag.*, 2 (1901) 559-572.
[77] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, *J. Mol. Graph.*, 14 (1996) 33-38.
[78] A. Bakan, L.M. Meireles, I. Bahar, ProDy: Protein Dynamics Inferred from Theory and Experiments, *Bioinformatics*, 27 (2011) 1575-1577.
[79] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, W. Wriggers, K. Schulten, Steered Molecular Dynamics, *Springer, Berlin/Heidelberg, Germany*, (1999) 39-65.
[80] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM, *Science*, 276 (1997) 1109-1112.
[81] A.F. Oberhauser, P.E. Marszalek, H.P. Erickson, J.M. Fernandez, The molecular elasticity of the extracellular matrix protein tenascin, *Nature*, 393 (1998) 181-185.
[82] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, Molecular dynamics study of unbinding of the avidin-biotin complex, *Biophys. J.*, 72 (1997) 1568-1581.
[83] H. Grubmüller, B. Heymann, P. Tavan, Ligand Binding: Molecular Mechanics Calculation of the Streptavidin-Biotin Rupture Force, *Science*, 271 (1996) 997-999.
[84] B. Isralewitz, S. Izrailev, K. Schulten, Binding pathway of retinal to bacterio-opsin: a prediction by molecular dynamics simulations, *Biophys. J.*, 73 (1997) 2972-2979.
[85] S. Stepaniants, S. Izrailev, K. Schulten, Extraction of Lipids from Phospholipid Membranes by Steered Molecular Dynamics, *J. Mol. Model.*, 3 (1997) 473-475.
[86] S.-J. Marrink, O. Berger, P. Tieleman, F. Jähnig, Adhesion Forces of Lipids in a Phospholipid Membrane Studied by Molecular Dynamics Simulations, *Biophys. J.*, 74 (1998) 931-943.

[87] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation, *Biophys. J.*, 75 (1998) 662-671.

[88] C. Jarzynski, Nonequilibrium Equality for Free Energy Differences, *Phys. Rev. Lett.*, 78 (1997) 2690-2693.

[89] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, *J. Comput. Chem.*, 13 (1992) 1011-1021.

[90] J.G. Shewale, β-Glucosidase: Its role in cellulase synthesis and hydrolysis of cellulose, *Int. J. Biochem.*, 14 (1982) 435-443.

[91] R. Opassiri, B. Pomthong, T. Onkoksoong, T. Akiyama, A. Esen, J.R. Ketudat Cairns, Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase, *BMC Plant Biol.*, 6 (2006) 33.

[92] R.R. Singhania, A.K. Patel, R.K. Sukumaran, C. Larroche, A. Pandey, Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production, *Bioresour. Technol.*, 127 (2013) 500-507.

[93] R.L. Lieberman, B.A. Wustman, P. Huertas, A.C. Powe Jr, C.W. Pine, R. Khanna, M.G. Schlossmacher, D. Ringe, G.A. Petsko, Structure of acid β-glucosidase with pharmacological chaperone provides insight into Gaucher disease, *Nat. Chem. Biol.*, 3 (2007) 101-107.

[94] J.D. Johnson, M.A. El-Bayoumi, L.D. Weber, A. Tulinsky, Interaction of .alpha.-chymotrypsin with the fluorescent probe 1-anilinonaphthalene-8-sulfonate in solution, *Biochemistry*, 18 (1979) 1292-1296.

[95] J.J. Birktoft, D.M. Blow, Structure of crystalline α-chymotrypsin: V. The atomic structure of tosyl-α-chymotrypsin at 2 Å resolution, *J. Mol. Biol.*, 68 (1972) 187-240.

[96] D. Shah, K. Mital, The Role of Trypsin:Chymotrypsin in Tissue Repair, *Adv. Ther.*, 35 (2018) 31-42.

[97] G.D.M.P. Madhusankha, R.C.N. Thilakarathna, Meat tenderization mechanism and the impact of plant exogenous proteases: A review, *Arab. J. Chem.*, 14 (2021) 102967.

[98] J.F. Kennedy, V.W. Pike, Papain, chymotrypsin and related proteins—a comparative study of their beer chill-proofing abilities and characteristics, *Enzyme Microb. Technol.*, 3 (1981) 59-63.

[99] O.L. Tavano, A. Berenguer-Murcia, F. Secundo, R. Fernandez-Lafuente, Biotechnological Applications of Proteases in Food Technology, *Compr. Rev. Food Sci. Food Saf.*, 17 (2018) 412-436.

[100] P. Singh, S. Choudhury, S. Dutta, A. Adhikari, S. Bhattacharya, D. Pal, S.K. Pal, Ultrafast spectroscopy on DNA-cleavage by endonuclease in molecular crowding, *Int. J. Biol. Macromol.*, 103 (2017) 395-402.

Chapter 3

Instrumentation and Sample Preparation

In this chapter, details regarding the instrumental setup, sample preparation techniques and experimental protocols used in the studies have been described. Furthermore, the details about the computational studies have been included in this chapter.

3.1. Instrument Setup:

3.1.1. Steady-State UV-Vis Absorption and Emission Measurement:

Steady-state UV-Vis absorption and emission spectra of the molecules of interest were measured using Shimadzu Model UV-2600 spectrophotometer (Shimadzu Corporation, Kyoto, Japan) and Jobin Yvon Fluoromax-3® spectrofluorimeter (Horiba, Kyoto, Japan), respectively. Schematic ray diagrams of these two instruments are shown in Figures 3.1 and 3.2.

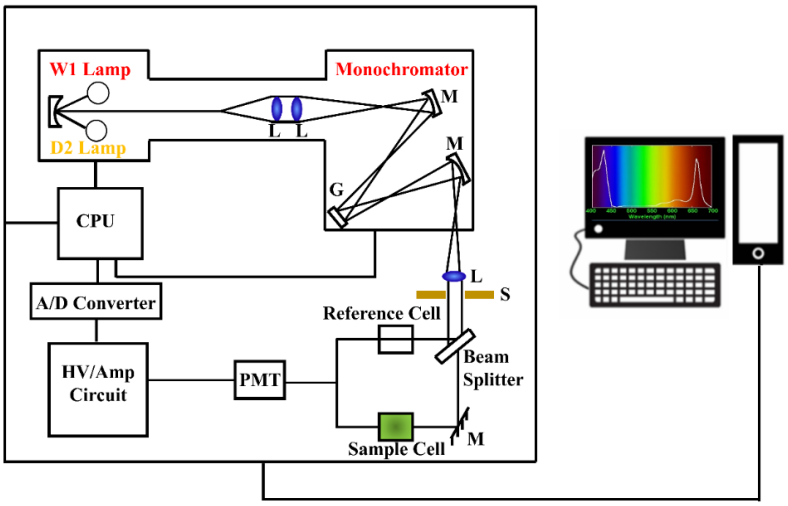

Figure 3.1. *Schematic ray diagram of an absorption spectrophotometer. Tungsten halogen (W1) and deuterium lamps (D2) are used as light sources in the visible and UV regions, respectively. M, G, L, S, PMT designate mirror, grating, lens, shutter and photomultiplier tube, respectively. CPU, A/D converter and HV/amp indicate central processing unit, analog to digital converter and high-voltage/amplifier circuit, respectively.*

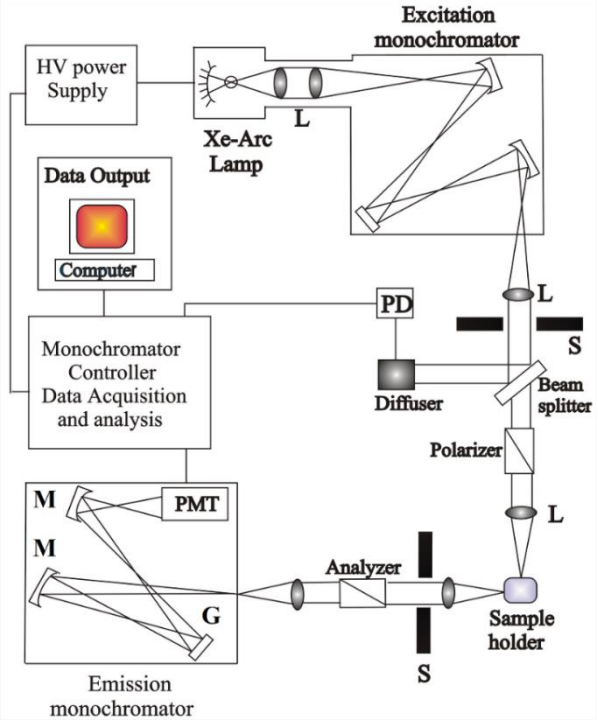

Figure 3.2. *Schematic ray diagram of an emission spectrofluorometer. M, G, L, S, PMT and PD represent mirror, grating, lens, shutter, and photomultiplier tube and reference photodiode, respectively.*

3.1.2. Circular Dichroism (CD) Measurement:

CD spectroscopy is based on the differential absorption of left and right-handed circularly polarized light. CD spectroscopy is used to determine the structure of macromolecules (including the secondary structure of proteins). The measurements were done in a JASCO spectropolarimeter with a temperature controller attachment (Peltier) (Figure 3.3). The CD spectra were acquired using quartz cells of 0.1 and 1.0 cm path length. For proteins and DNA, the typical concentration used for CD measurements were within 10 μM.

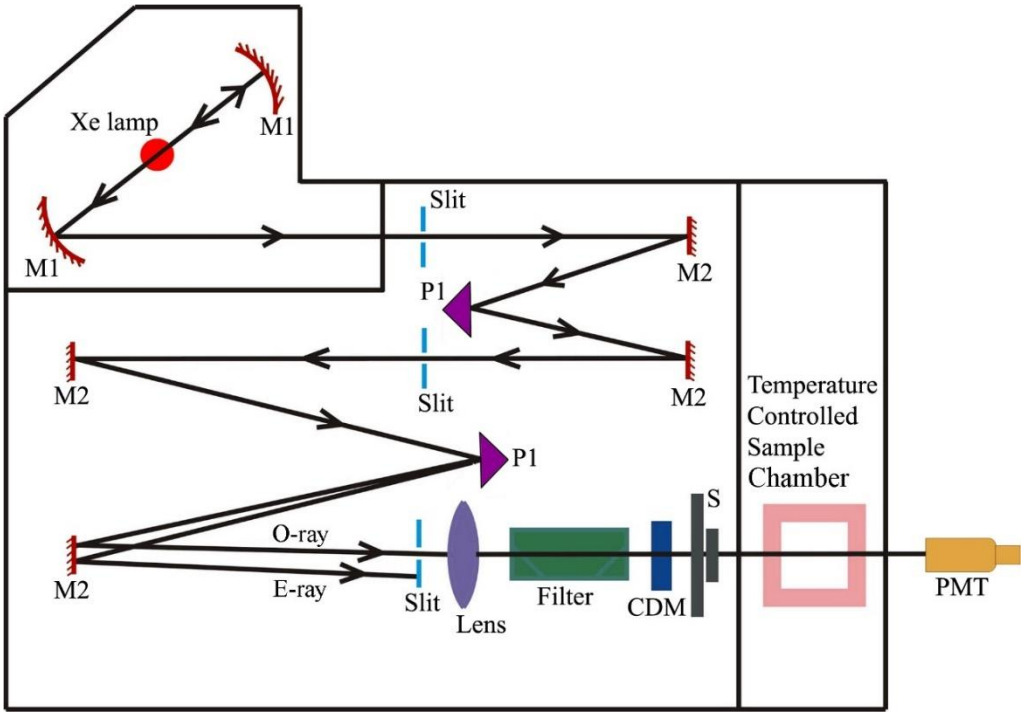

Figure 3.3. *Schematic ray diagram of a circular dichroism (CD) spectropolarimeter. M1, M2, P1, S, PMT, CDM, O-ray and E-ray represent concave mirror, plain mirror, reflecting prism, shutter, photomultiplier tube, CD-modulator, ordinary ray and extraordinary ray, respectively.*

3.1.3. Time Correlated Single Photon Counting (TCSPC) Technique:

All the picosecond-resolved fluorescence transients were measured using TCSPC technique. The schematic block diagram of a TCSPC system is shown in Figure 3.4. TCSPC setup from Edinburgh instruments, U.K., was used during fluorescence decay acquisitions. The instrument response functions (IRFs) of the laser sources at different excitation wavelengths varied between 70 ps to 80 ps. The fluorescence from the sample was detected by a photomultiplier after dispersion through a grating monochromator [1]. For all transients, the polarizer in the emission side was adjusted to be at 54.70 (magic angle) with respect to the polarization axis of excitation beam. In order to measure anisotropy, fluorescent transients were measured with the emission polarizer aligned in parallel and perpendicular directions with respect to vertical polarization of excitation light.

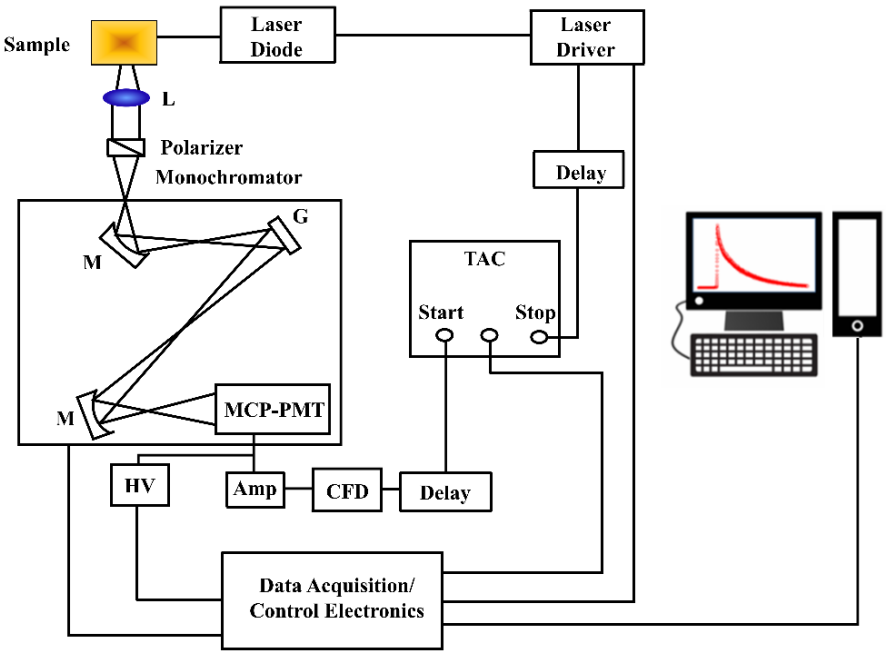

Figure 3.4. *Schematic ray diagram of a time correlated single photon counting (TCSPC) spectrophotometer. A signal from microchannel plate photomultiplier tube (MCP-PMT) is amplified (Amp) and connected to start channel of time to amplitude converter (TAC) via constant fraction discriminator (CFD) and delay. The stop channel of the TAC is connected to the laser driver via a delay line. L, M, G and HV represent lens, mirror, grating and high voltage source, respectively.*

3.1.4. Dynamic Light Scattering (DLS):

Dynamic light scattering (DLS), also known as Photon Correlation Spectroscopy (PCS) or Quasi-Elastic Light Scattering (QELS), is one of the most popular techniques used to determine the hydrodynamic size of the particle. DLS measurements were performed on a Nano S Malvern instruments, U.K. employing a 4 mW He-Ne laser (λ= 632.8 nm) and equipped with a thermostatic sample chamber. The instrument allows DLS measurements in which all the scattered photons are collected at 173˚ scattering angle (Figure 3.5).

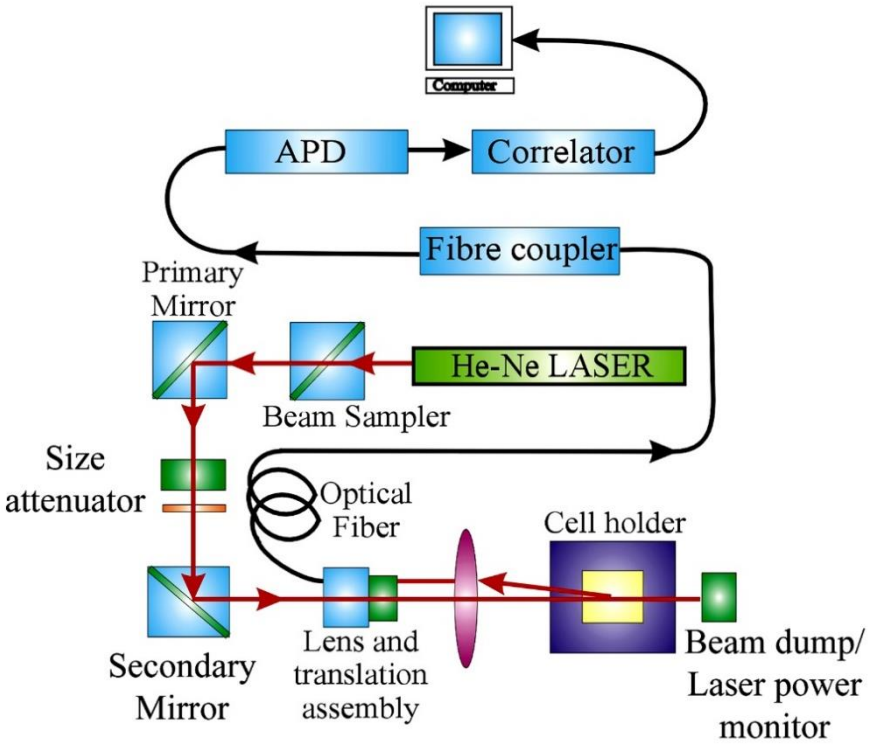

Figure 3.5. *Schematic ray diagram of dynamic light scattering (DLS) instrument. The avalanche photo diode (APD) is connected to preamplifier/amplifier assembly and finally to correlator. It has to be noted that lens and translational assembly, laser power monitor, size attenuator, laser are controlled by the computer.*

3.2. Sample Preparation:

In this section, the different sample preparation methods have been discussed.

3.2.1. Chemicals Used:

Rutin, β-Glucosidase (BGL), α-Chymotrypsin (CHT), Ala-Ala-Phe-7-amido-4-methylcoumarin (AMC), 8-anilino-1-naphthalenesulfonic acid ammonium salt (ANS) were purchased from Sigma (St. Louis, MO, USA) unless stated otherwise. Ethanol and Formalin were purchased from Merck (Darmstadt, Germany) unless stated otherwise. All reagents used were of analytical grade and used without further purification. Nanopure water (resistivity ≥18 MΩ cm) from Milli-Q system (Millipore GmbH, Germany) was used whenever required.

3.2.2. Preparation of ANS-Enzyme Complex:

3.2.2.1. ANS-β-Glucosidase (BGL) Complex:

The ANS-BGL complex was prepared by mixing ANS (0.6 mM) with BGL (6 mM) in phosphate buffer (pH 7.0) with continuous stirring for 5 h at 4^{o}C [2].

3.2.2.2. ANS-α-Chymotrypsin (CHT) Complex:

The ANS-CHT complex was prepared by mixing ANS (0.5 µM) with CHT (5 µM) in phosphate buffer (pH 7.0). The mix was stirred continuously for 5 h at 4^{o}C.

3.2.3. Enzyme Kinetics:

3.2.3.1. Enzyme Kinetics Assay of β-Glucosidase (BGL):

Temperature-dependent (278–353 K) enzymatic activity studies were conducted with 0.5 µM BGL concentration and the concentration of the substrate rutin (λ_{max} ~ 358 nm) was varied in the range of 4–20 mM. The substrate was cleaved to produce quercetin (λ_{max} ~ 374 nm) and the absorbance was recorded. The Michaelis–Menten kinetics parameters were determined from the Lineweaver–Burk plot [See Chapter 4 for Data].

3.2.3.1. Enzyme Kinetics Assay of α-Chymotrypsin (CHT):

Temperature dependent (278K-338K) enzymatic activity studies were conducted using 0.6 µM CHT concentration and the substrate Ala-Ala-Phe-AMC (λ_{max} 325 nm) concentration was varied between 20-200 µM. Substrate was cleaved to produce 7-amido-4-methyl coumarin (λ_{max} 370 nm). The absorbance was monitored at 370 nm. Michaelis-Menten kinetics parameters were calculated using a Lineweaver-Burk plot [See Chapter 5 for Data].

Enzymatic activity of CHT was conducted with 0.5 µM of CHT at different formalin concentrations (1%, 2%, 4%). The substrate Ala-Ala-Phe-AMC (λ_{max} 325 nm) was cleaved to produce 7-amido-4-methyl-coumarin (λ_{max} 370 nm), the absorbance was monitored at 370 nm [See Chapter 6 for Data].

3.3. Computational Study:

3.3.1. Homology Modelling:

Homology modelling was utilized for constructing a 3-D structure of BGL. The amino acid sequence [Accession number: AFH35017.1] was obtained from NCBI GenBank database. Homology modelling was used to construct the 3-D model for almond β-Glucosidase (BGL) using *Trifolium repens* cyanogenic β glucosidase (1CBG) as the template. The SWISS-MODEL server was used to construct the 3-D model of BGL. The structural assessment was

done by VERIFY-3D [3], PRO-CHECK [4], and ERRAT [5] web servers [See Chapter 4 for Data].

3.3.2. Molecular Docking:

3.3.2.1. Molecular Docking Between BGL-ANS and BGL-Rutin:

Molecular docking studies were performed to identify the binding site of ANS and rutin in BGL. The ligand (ANS & rutin) structures were obtained from the PubChem database. Targeted docking was performed at the active site using AutoDock Vina (Version 1.1.2) [6]. The molecular interactions were viewed using Schrodinger Maestro 2018-1 (Academic release) [See Chapter 4 for Data].

3.3.2.2. Molecular Docking Between CHT-AMC and CHT-ANS:

Molecular docking was performed to determine the catalytic (i.e., substrate-AMC) and the ANS binding residues using AutoDock Vina (Version 1.1.2). Solvent exposed residues susceptible to formalin modification were modified followed by minimization of the modified residues using Schrodinger Maestro 2018-1 (Academic release). All covalent modifications were conducted according to Kamps et al., [7]. Molecular docking studies were performed to study the influence of formalin induced covalent modification on substrate (AMC) and molecular recognition of ANS by CHT [See Chapter 6 for Data].

3.3.3. Solvent Accessible Surface Area (SASA):

Solvent accessible surface area (SASA) is a measure of formation of contacts between the atoms on the surface of the protein and the solvent molecules. Thus, solvent accessibility provides a better insight on the amino acid residues of the enzyme (CHT) vulnerable to formalin induced modification. The validation and selection of the appropriate protein data bank (PDB) structure of CHT was made according to Chakraborty et al., [8]. Solvent accessible surface area (SASA) of CHT (PDB ID: 1CGJ) is computed using the GETAREA integration in PyMOL software (Educational version). The catalytic and residues of ANS binding site were identified as the prime targets for the SASA analysis [See Chapter 6 for Data].

3.3.4. Molecular Dynamics (MD) Simulation:

3.3.4.1. MD Simulation of BGL:

The temperature dependent dynamical behaviour of BGL was further studied by molecular dynamics (MD) simulation. The constructed BGL structure was subjected to energy minimization by Schrodinger Maestro 2018-1 (Academic release). Simulations were done at temperatures of 278 K, 313 K, 333 K and 353 K. A protein molecule was placed in the centre of a periodic boundary box at a minimum distance of 10 Å from each side. Then, the protein was solvated with pre-optimized simple point charged (SPC) water molecules. Additionally, the system was charge neutralized by adding the necessary counter ions (seven Na^+). A five step relaxation protocol was used to equilibrate the entire system [9]. For the final runs, a normal pressure and a defined temperature were used and no particles were restrained. The ensemble type was NPT with a constant number of particles, normal pressure (1 bar) and a defined temperature. A Nosé-Hoover thermostat and a Martyna–Tobias–Klein barostat maintained the temperature and pressure of the system, respectively. MD was run in an OPLS 2005 force field [10] with a short-range coulombic interaction cut-off of 9 Å. Long range coulombic interactions were handled by a smooth particle mesh Ewald method (Ewald tolerance of 10^{-9}). The simulations were run for 48 ns at the said temperatures. The overall structural changes in the protein with time were computed in terms of the root mean square deviation (RMSD). The radius of gyration (R_g) and residue-wise fluctuations (RMSF) were also computed from the simulation trajectory [See Chapter 4 for Data].

3.3.4.2. MD Simulation of CHT:

Temperature dependent dynamical behaviour of CHT was further studied by molecular dynamics (MD) simulation. CHT structure (1CGJ) was subjected to energy minimization by Schrodinger Maestro 2018-1 (Academic release). Simulations were carried out at 278K, 288K, 308K, 328K and 338K. CHT was placed in a periodic boundary box at a minimum distance of 10Å from each side. The enzyme was solvated with pre-optimized simple point charged (SPC) water molecules. Then, the charge of the system was neutralized by adding necessary ions (seven Na^+). The system was equilibrated using a five step relaxation protocol [9]. Normal pressure (1 bar) and defined temperature with no particle restraints were used for the final run. A Nosé-Hoover thermostat and a Martyna–Tobias–Klein barostat maintained the temperature and pressure of the system, respectively. MD simulation was run in an OPLS 2005 force field [10]. with a short-range coulombic interaction cut-off of 9 Å. Long range coulombic interactions were handled by a smooth particle mesh Ewald method (Ewald tolerance of 10^{-9}). The simulations were run for 30 ns at the said temperatures. The overall structural changes in the protein with time were computed in terms of the root mean square deviation (RMSD). The

radius of gyration (R_g) and residue-wise fluctuations (RMSF) were also computed from the simulation trajectory [See Chapter 5 for Data].

3.3.5. Normal Mode Analysis (NMA):

3.3.5.1. Normal Mode Analysis (NMA) of BGL:

To identify the regions of dominant motion in BGL a principal component analysis (PCA) was conducted. The principal component analysis (PCA) uses the molecular dynamics (MD) trajectory and extracts the dominant modes of motion of the molecules. These dominant motions correspond to the correlated vibrational modes or collective motions of the groups of atoms in the normal mode analysis [11]. The dominant motions along MD simulation trajectory was computed by the Normal Mode Wizard in VMD 1.9.3 program [12] using ProDy [13] [See Chapter 4 for Data].

3.3.5.2. Normal Mode Analysis of CHT:

The principal component analysis (PCA) takes the trajectory of a molecular dynamics simulation and extracts the dominant modes in the motion of the molecule [14]. These dominant motions correspond to the correlated vibrational modes or collective motions of the groups of atoms in the normal mode analysis. The principal component analysis (PCA) method depends on the covariance matrix to predict the atomic displacement of protein [15]. The direction and amplitude of the dominant motions along a simulation trajectory by principal component analysis (PCA) were computed by Normal Mode Wizard in VMD 1.9.3 program using ProDy [See Chapter 5 for Data].

3.3.6. Steered Molecular Dynamics (SMD) Simulation:

3.3.6.1. SMD Simulation of BGL-Rutin Complex:

A steered molecular dynamics simulation was performed in Schrodinger Maestro 2018-1 (Academic release) using the meta-dynamics module. Rutin bound BGL, as obtained by molecular docking, was subjected to an MD simulation following the previously mentioned protocol (Section 3.3.4.1.) at four different temperatures. Then, starting with the temperature equilibrated substrate (rutin) bound BGL conformations, a steered dynamics simulation was performed to probe the energy required to pull the substrate out of the active site. Therefore, the distance between the centre of mass of BGL and the centre of mass of rutin was gradually

increased. A cut-off value for the subsequent increase in the energy was set to be 20 kcal mol^{-1}. The other parameters were the same as the MD simulation [See Chapter 4 for Data].

References

[1] D. O'Connor, Time-correlated single photon counting, *Academic Press*, *USA*, (2012).

[2] S.K. Pal, J. Peon, A.H. Zewail, Ultrafast surface hydration dynamics and expression of protein functionality: α-Chymotrypsin, *Proc. Natl. Acad. Sci. U.S.A*, 99 (2002) 15297-15302.

[3] R. Lüthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature*, 356 (1992) 83-85.

[4] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 26 (1993) 283-291.

[5] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, *Protein Sci.*, 2 (1993) 1511-1519.

[6] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.*, 31 (2010) 455-461.

[7] J.J.A.G. Kamps, R.J. Hopkinson, C.J. Schofield, T.D.W. Claridge, How formaldehyde reacts with amino acids, *Commun. Chem.*, 2 (2019) 1-14.

[8] B. Chakraborty, C. Sengupta, U. Pal, S. Basu, Probing the hydrogen bond involving acridone trapped in a hydrophobic biological nanocavity: Integrated spectroscopic and docking analyses, *Langmuir*, 36 (2020) 1241-1251.

[9] U. Pal, S.K. Pramanik, B. Bhattacharya, B. Banerji, N.C. Maiti, Binding interaction of a gamma-aminobutyric acid derivative with serum albumin: an insight by fluorescence and molecular modeling analysis, *SpringerPlus*, 5 (2016) 1-17.

[10] A. Amadei, A.B.M. Linssen, B.L. De Groot, D.M.F. Van Aalten, H.J.C. Berendsen, An efficient method for sampling the essential subspace of proteins, *J. Biomol. Struct. Dyn.*, 13 (1996) 615-625.

[11] H. Yamaguchi, D.M.F. Van Aalten, M. Pinak, A. Furukawa, R. Osman, Essential dynamics of DNA containing a cis. syn cyclobutane thymine dimer lesion, *Nucleic Acids Res.*, 26 (1998) 1939-1946.

[12] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.*, 14 (1996) 33-38.

[13] A. Bakan, L.M. Meireles, I. Bahar, ProDy: protein dynamics inferred from theory and experiments, *Bioinformatics*, 27 (2011) 1575-1577.

[14] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, *Proteins: Struct., Funct., and Bioinf.*, 17 (1993) 412-425.
[15] A.M. Martinez, A.C. Kak, Pca versus lda, *IEEE PAMI*, 23 (2001) 228-233.

Chapter 4

Studies on the Structure, Function, and Dynamics of a Thermostable Enzyme at Physiologically Relevant Temperatures

4.1. Introduction:

β-Glucosidases (BGLs) constitute a subgroup of β-D-glucoside glucohydrolases (EC 3.2.1.21) and are among the most widely studied enzymes owing to their wide variety of substrates and broad distribution in the biological strata. BGLs are well characterized biologically important enzymes which catalyse the transfer of glycosyl groups between oxygen nucleophiles and hydrolyze the terminal, non-reducing β-D-glucosyl residues and thereby the terminal sugar moiety [1-3]. BGL have attracted considerable interests due to its importance in various biotechnological processes like hydrolysis of pharmaceutical ingredients, production of bio-ethanol from agricultural wastes [4, 5]. BGL is used extensively in the wine, tea and juice industries [6-8]. In addition to their use in industry, BGLs play pivotal roles in several biological processes [9]. The physiological roles associated with these enzymes are diverse, ranging from facilitators of plant immune defence mechanisms to enzymatic breakdown of glucocerebroside and in treatment of Gaucher's disease in humans [10, 11]. At present, the tertiary structures of several BGLs have been determined. All of these BGLs share a similar structural motif, and similar catalytic reaction mechanism [12, 13].

Almond β-glucosidase activity has been well investigated. However, studies investigating correlation between its flexibility and activity is not as well understood as that between structure and activity. Previous studies have investigated the dynamical nature of BGL upon substrate binding [14, 15], which provided little to no information regarding the essential motions responsible for catalysis. Upon an inspection of molecular structure of the almond BGL four loop regions (LRs) were found to be placed strategically over the catalytic barrel of the enzyme like a lid. We conjecture that these LRs are involved in a gatekeeping mechanism to control enzymatic activity at different temperatures.

The present study examines the essential structural dynamics of almond BGL which was correlated with its catalytic activity at different temperatures using steady state, circular

dichroism (CD) and picosecond resolved fluorescence spectroscopy as well as all atom molecular dynamics (MD) simulation, principal component analysis (PCA) and steered molecular dynamics simulation (SMD).

4.2. Results and Discussion:

4.2.1. Flexibility Modulates Catalytic Activity of a Thermostable Enzyme: Key Information from Optical Spectroscopy and Molecular Dynamics Simulation [16]:

Almond BGL [GenBank accession number AFH35017] exhibited a low similarity with structures within the PDB database based on the results of sequence analysis tools. The sequence search by NCBI BLAST [17] over structures within PDB revealed that BGL matches with cyanogenic β-Glucosidase of *Trifolium repens* (1CBG) having 63% sequence identity. Accordingly, homology modelling method was employed to construct a 3D-model for BGL. The amino acid sequence was analysed using a set of known fold recognition online servers including PSI-BLAST [17], Phyre2 [18], and HHpred [19]. All servers commonly showed the cyanogenic β-Glucosidase from *T. repens* (1CBG) [20] with the highest identity and lowest E-value. The BGL structure was found to belong to the same TIM barrel family of 1CBG: family 1 of glycosyl hydrolase. Consequently, 1CBG was proven to be a suitable template structure for modelling the BGL structure. The 3D-model was built for BGL sequence using the SWISS-MODEL server [21]. The 1CBG protein structure was used as a template.

The constructed 3D-model was submitted to different tools for quality evaluation. The backbone conformation study by the PRO-CHECK web server based on the Psi/Phi Ramachandran plot revealed that 90% of the residues were located in the most favourable region, 9.4% in the additional allowed region, 0.4% in the generously allowed region and 0.2% residues in the disallowed regions (Figure 4.1). The evaluation by VERIFY-3D web server showed that 94.87% of the residues resides in a favourable environment having an average 3D-1D score of above 0.2 [22]. Moreover, the statistics of highly refined structures were compared using the ERRAT web tool to study the overall quality of the model for non-bonded atomic interactions. The computed score by ERRAT (94.51%) indicates that the model is acceptable in the normal range (above 50%). The quality of the constructed model was confirmed based on the scores obtained from the above evaluation tools. The summary of the model analysis is shown in Table 4.1.

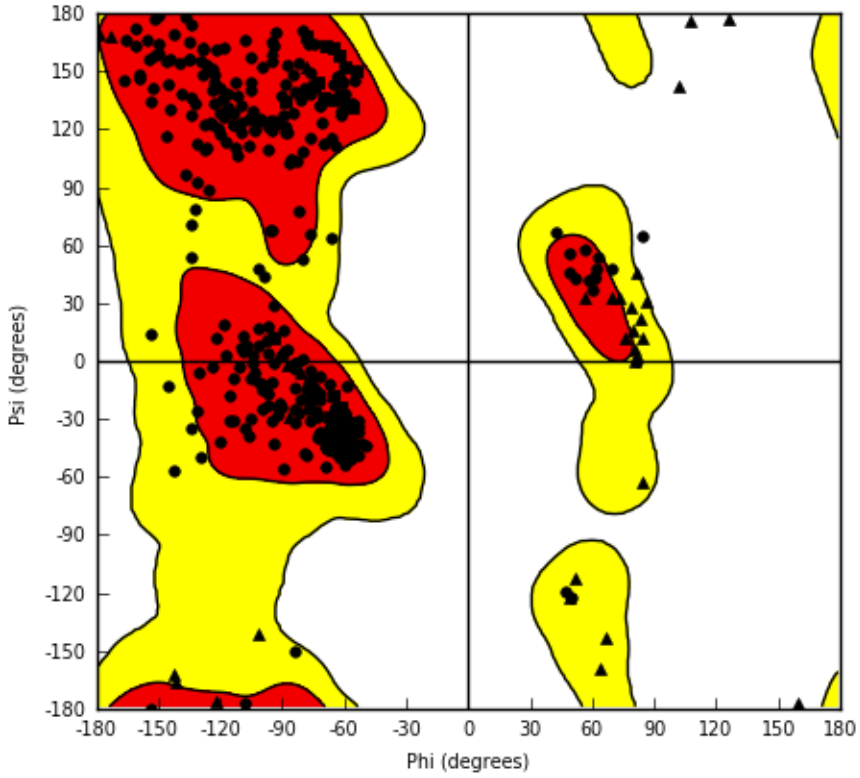

Figure 4.1. *Showing the Ramachandran Plot for the modelled structure of almond BGL.*

Table 4.1. Model evaluation summary.

Model evaluation tool	Assessment system	Obtained score	Normal range of the score
PROCHECK	The number of residues in the allowed region based on Psi/Phi Ramachandran plot	90%	≥90%
VERIFY 3D	The number of residues having an average 3D-1D score above 0.2	94.87%	≥80%
ERRAT	The overall quality for non-bonded atomic interactions	94.51%	≥50%

Figure 4.2.a shows the homology model of almond BGL. The BGL model has similar characteristics of other family members of family 1 glycosyl hydrolases. The catalytic domain of the enzyme (BGL) consists of alternating pattern of β-strands and α-helices that is a TIM barrel fold [(α/β)8] [23]. The BGL structures mostly consist of a TIM barrel as a common

folding motif [24]. Evolutionarily, family 1 glycosyl hydrolases are dyad and figure 4.2.b shows the conserved catalytic sites on the β-barrel (Glu424; Glu211) of BGL. The catalytic site on the β-barrel is covered by four mostly unstructured regions guarding the entry gate of substrate as a lid. The regions are marked as follows: LR1 (residue no: 77-100); LR2 (residue no: 227-242); LR3a (residue no: 359-376); LR3b (residue no: 377-387); LR3c (residue no: 388-402) and LR4 (residue no: 426-449) as shown in the figure 4.2.c and 4.2.d.

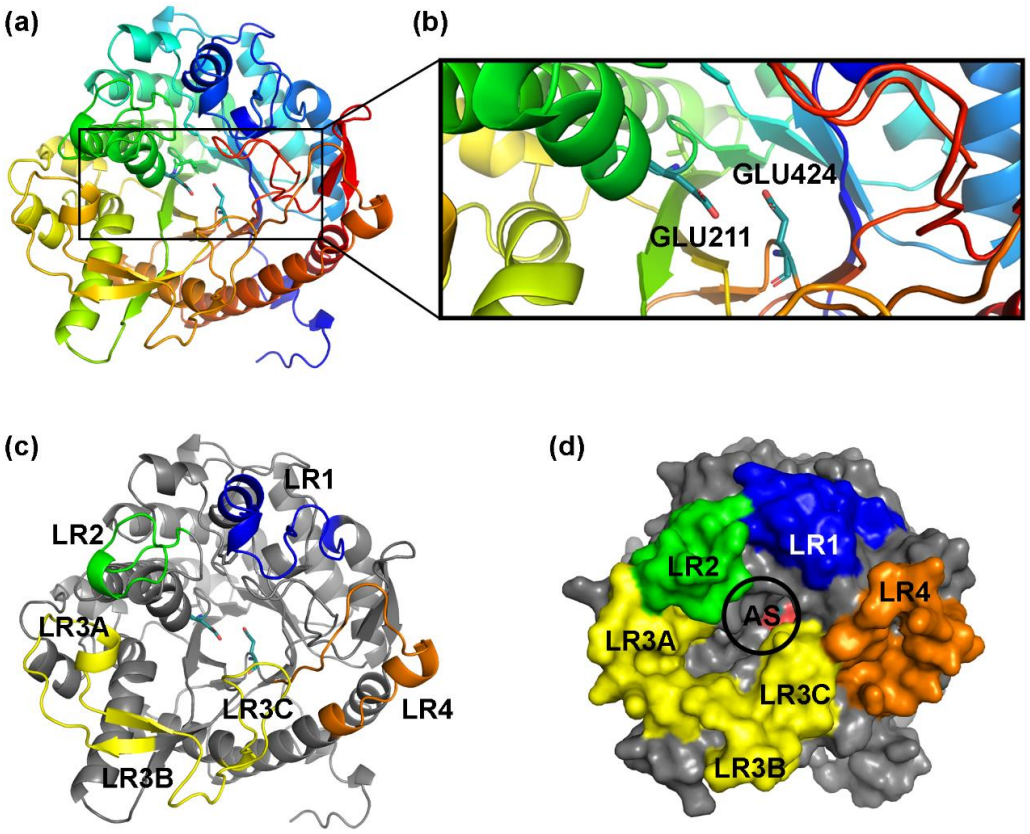

Figure 4.2. *Schematic showing the build structure of almond Beta-Glucosidase (BGL). (a) The modelled 3D structure. (b) The stick representation showing the conserved active site Glutamate residues. (c) Showing the four looped region (LRs). The coloured regions represent each of the LRs, LR1 [Blue]; LR2 [Green]; LR3 [Yellow]; LR4 [Orange]. (d) Space filling model showing the strategic positioning of the LRs along the active site [AS], LRs represented by the same colour scheme.*

Almond β-glucosidase is relatively more thermostable than other β-glucosidases of plant origin [25] including vanilla [26], grape [27], maize [28], and tea leaves β-glucosidase [29]. The thermal stability of BGL was quantified by circular dichroism (CD) spectroscopy in the far UV region. The far UV CD spectra of BGL were characterized by two minima at 208 nm and 220 nm [30] (Figure 4.3). Upon deconvolution of the spectra insignificant changes in

the secondary structure was observed indicating to the absence of thermal denaturation within the temperature range of our study.

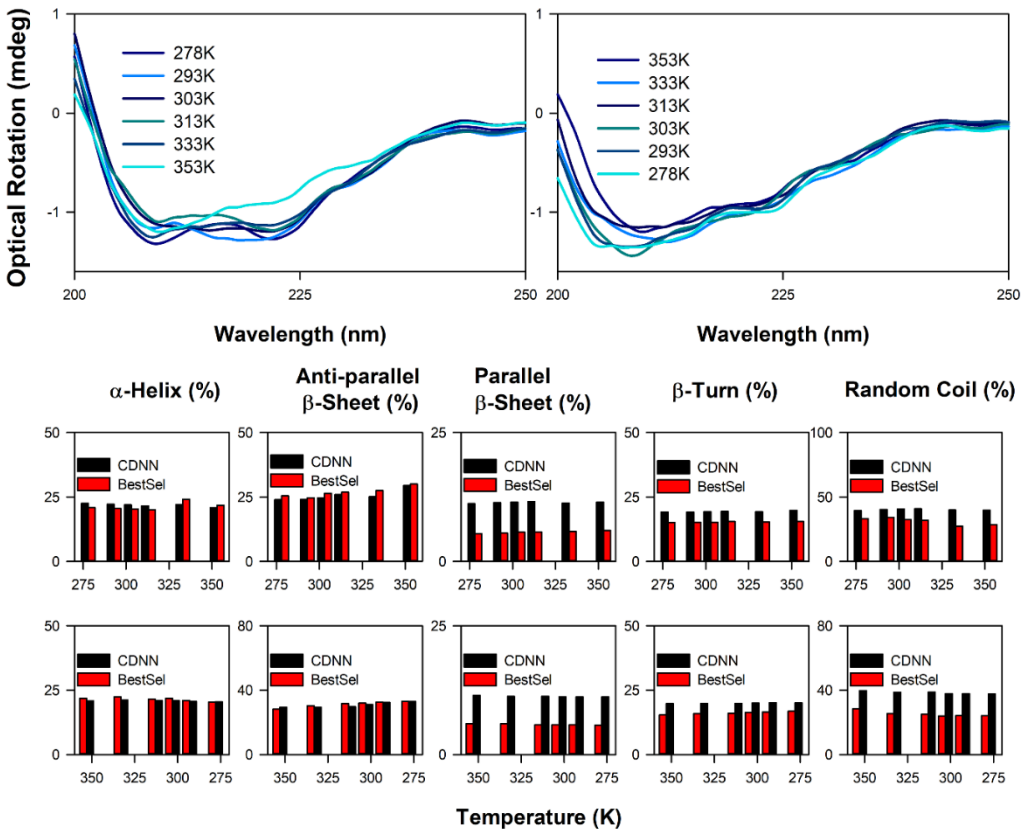

Figure 4.3. *Far UV CD spectroscopy study and thermal stability analysis of almond BGL secondary structure. (a) The far UV CD spectra representing the thermal unfolding of almond BGL. (b) The far UV CD spectra representing the thermal refolding of almond BGL. (c) Percentage composition of secondary structure shown as bar plots. Analysis of secondary structure indicated thermal stability of almond BGL.*

Only 1.7% decrease in the α-helix, 5.5% increase in antiparallel β-sheet, an increase of 0.3% parallel β-sheet, 0.6% increase in β-turn and 0.4% increase in random coil were observed (Table 4.2.a). The refolding of the BGL molecule was studied through gradual cooling of the enzyme to 278K and subsequently, deconvolution of the CD spectra was found to manifest the retention of the enzyme native structure with 1.0% recovery of the α-helix and 2.0% in the antiparallel β-sheet (Table 4.2.b). No significant changes in the random coil, turns and parallel β-sheets were observed (Table 4.2.b).

Table 4.2.a. Forward Temperature Dependent Structure of β-Glucosidase as Deconvoluted by CDNN & BestSel.

Temp (K) (Forward)	Secondary structure				
	β-Glucosidase				
	% α-Helix	% β-Sheet (Anti)	% β-Sheet (Parallel)	% β-Turn	% Random Coil
278	22.50/20.90	23.90/25.40	11.20/5.40	19.20/15.10	39.30/33.20
293	22.20/20.50	24.00/24.70	11.40/5.50	19.20/15.20	40.20/34.10
303	21.90/20.30	24.70/26.40	11.50/5.70	19.30/15.20	40.60/32.40
313	21.50/20.00	25.90/26.90	11.60/5.70	19.40/15.50	40.80/31.90
333	22.00/21.10	25.10/27.50	11.30/5.80	19.30/15.30	39.80/30.30
353	20.80/21.80	29.40/28.20	11.50/6.00	19.80/15.50	39.70.28.50

Table 4.2.b. Reverse Temperature Dependent Structure of Beta-Glucosidase as Deconvoluted by CDNN & BestSel.

Temp (K) (Reverse)	Secondary structure				
	β-Glucosidase				
	% α-Helix	% β-Sheet (Anti)	% β-Sheet (Parallel)	% β-Turn	% Random Coil
353	20.80/21.80	29.40/28.20	11.50/6.00	19.80/15.50	39.70/28.50
333	21.10/22.40	29.40/30.30	11.30/6.00	19.80/15.90	38.70/25.40
313	21.00/21.50	29.70/31.70	11.30/5.80	19.80/16.00	38.80/25.00
303	20.90/21.80	31.10/32.00	11.20/5.80	20.00/16.40	37.90/24.00
293	20.60/21.00	32.40/32.50	11.205.80	20.10/16.50	37.90/24.20
278	20.50/20.40	32.90/33.00	11.20/5.70	20.10/16.80	37.60/24.10

Thus, from the secondary structure analysis no appreciable structural denaturation of the enzyme was noticed upon thermal treatment. Upon re-evaluation of the CD spectra using the BestSel CD deconvolution server similar trends in the results were obtained. No significant change in the secondary structure was observed (Table 4.2.a & 4.2.b), indicating and reaffirming the thermal stability of the enzyme. (Table 4.2.b).

The Enzyme kinetics of BGL glucosidase activity were conducted in details at temperatures ranging from 278K to 353K. The initial velocity of the glycoside hydrolysis by BGL was measured at different rutin (substrate) concentrations (4–20 mM). Lineweaver-Burk plots were used to determine the reaction kinetic parameters (K_m, V_{max}, k_{cat}, k_{cat}/K_m). BGL belongs to the class of glycohydrolases that catalyzes the hydrolysis of its natural substrate rutin into quercetin (Figure 4.4.a and 4.4.b). Figure 4.4.c and 4.4.d illustrate the effect of increasing temperature on the product formation and catalytic efficiency.

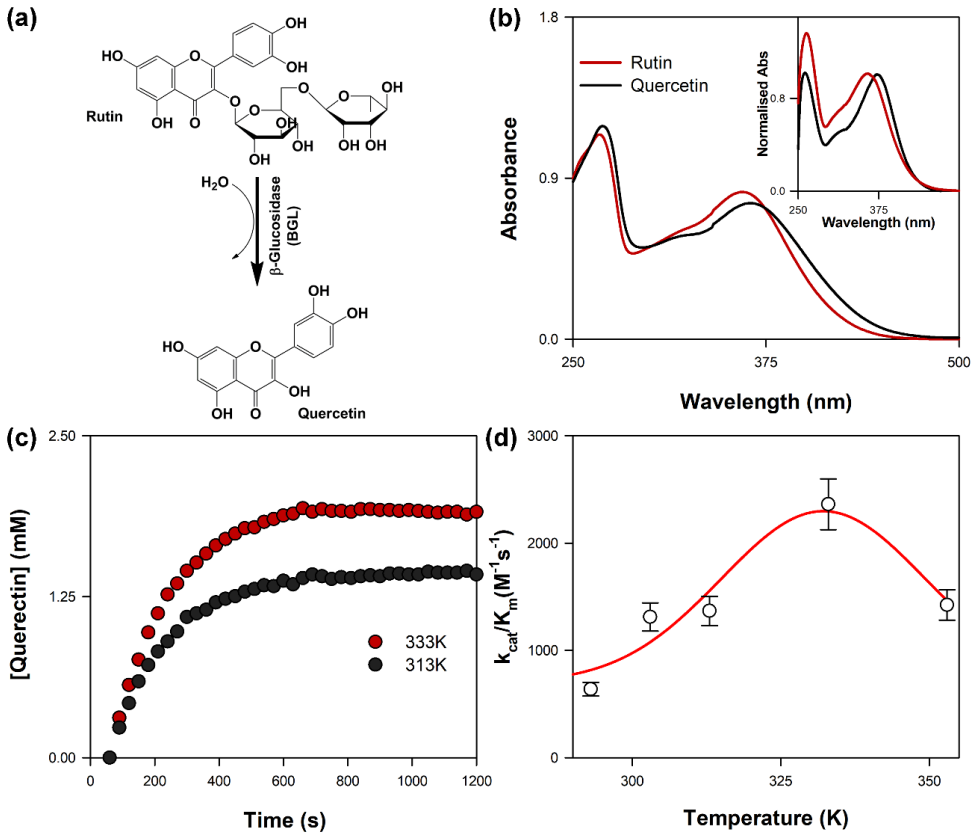

Figure 4.4. *Schematic representation of temperature dependent enzyme catalysis. (a) Simple model outline of the reaction catalysed by BGL, the flavonoid glycoside Rutin hydrolysed to produce Quercetin. (b) Absorbance spectra of the substrate Rutin and product Quercetin during catalysis. The inset representing normalised absorption spectra of the same. (c) Product formation at different representative temperatures showing enhanced product formation at 333K. (d) Showing the effect of temperature on the catalytic efficiency (k_{cat}/K_m) of BGL. The enzyme showed enhancement of catalytic efficiency with increased temperature generating a bell-shaped curve. Temperature optima was observed at 333K. Fall in efficiency was observed beyond 333K.*

Thermal treatment was found to enhance the catalytic efficiency of the enzyme (almond BGL) which becomes maximum at 333K, and a further increase of temperature to 353K resulted in

decreased catalytic efficiency relative to that at 333K. It is also noteworthy that no discernible product formation was observed at temperature as low as 278K, while low product formation and reduced catalytic efficiency were observed at a much higher temperature (353K). This is consistent with the thermal stability of almond BGL over a wide range of temperatures indicating the formation of a stable enzyme-substrate complex even at a temperature as high as 353K. The summary of the kinetics parameters is shown in Table 4.3.

Table 4.3. Temperature-Dependent Kinetics of BGL Catalyzed Hydrolysis of Rutin.

Temperature (K)	K_m (M)	V_{max} (µM/s)	K_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
293K	$(4.7\pm2.35) \times 10^{-5}$	0.15±0.007	0.3±0.015	638±31.90
303K	$(5.1\pm2.55) \times 10^{-4}$	0.4±0.020	0.8±0.040	1311±65.55
313K	$(7.9\pm3.95) \times 10^{-4}$	0.54±0.027	1.08±0.054	1367±68.35
333K	$(1.27\pm6.35) \times 10^{-3}$	1.5±0.075	3.0±0.150	2362±118.10
353K	$(1.04\pm5.2) \times 10^{-3}$	0.74±0.037	1.4±0.070	1423±71.15

The Molecular docking studies of ANS in the presence of the enzyme (BGL) show its' molecular interactions with the amino acid residues of BGL (Figure 4.5). The targeted docking found interactions between the probe and the tryptophan residues (Trp396 and Trp475) at the catalytic barrel of BGL, which was close to the active site. The dynamical flexibility of BGL was revealed from time-resolved anisotropy decay (r(t)) of ANS [31-35] in its complex with the enzyme at different temperatures (Figure 4.6) over the range of 278K to 353K. A three exponential model was required to obtain satisfactory fits to the polarization gated anisotropy decay, r(t). The fast-rotational correlation time (φ_1) of ~100-400 ps corresponds to rapid local motion of the ANS probe bound to Tryptophan (Trp396 and Trp475) residues, whereas the slower one (φ_2) of 0.5-2.36 ns (Table 4.4) is reflective of segmental motion of the neighbouring loop regions close to the active (catalytic) site or the catalytic barrel. The much longer time constant (φ_3 > 20 ns, Table 4.4) in the anisotropy decay (Figure 4.6) corresponds to global tumbling motion of the overall probe-enzyme complex.

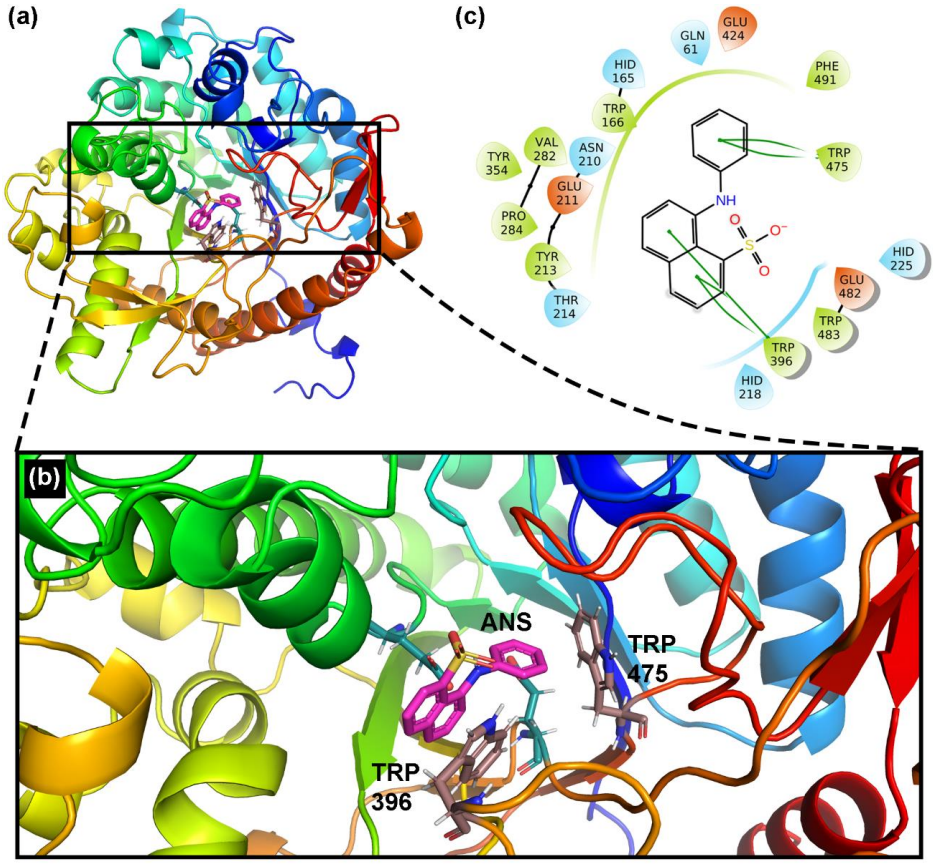

Figure 4.5. *Ribbon representation of docked conformation of ANS-BGL complex. (a) Molecular docking representation of ANS-BGL complex, ANS [shown in pink] was found to be bound at the catalytic barrel of BGL. (b) ANS docked at the catalytic barrel. The fluorescence probe was found to interact with two Tryptophan residues (Trp396; Trp475). (c) Molecular interaction diagram between ANS and Tryptophan residues via π-stacking [shown in green lines] at the catalytic barrel.*

A monotonous decrease in both the correlation times (φ_1 and φ_2) was noted with rise in temperature. In comparison to the lower temperature range (278K to 313K) a significant decrease in φ_1 and φ_2 was noted at 333K or 353K due to increased mobility of the dye local environment that results from significantly increased dynamical flexibility of the amino acid residues (Trp396 and Trp475) and the loop regions close to the catalytic barrel. The significant decrease, especially in φ_2, in going from 313K to 333K indicates significantly increased dynamic flexibility of the loop regions which likely facilitates substrate binding to the catalytic barrel resulting in enhanced catalytic efficiency at the higher temperature (333K). Upon further increase of temperature to 353K, a remarkable decrease in the rotational time constants was observed which is consistent with rapid fluctuations of the amino acid residues and loop regions near the catalytic barrel which may sometimes impede substrate entry into the catalytic barrel or lower the stability of the enzyme-substrate complex.

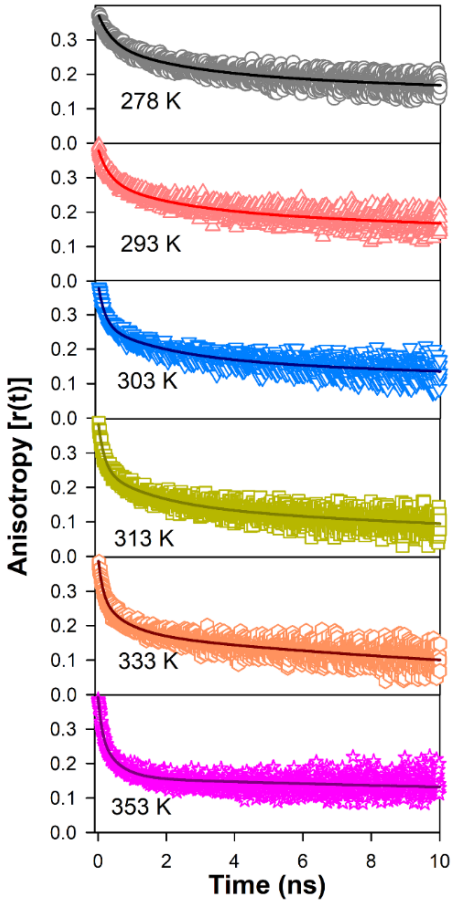

Figure 4.6. *Temperature dependent anisotropy decay of BGL-ANS complex.*

Table 4.4. Time-resolved Anisotropy of ANS in BGL at Different Temperatures

Temperature(K)	r_0	φ_1/ns	φ_2/ns	φ_3
278K	0.37	0.400 ± 0.040 (21%)	2.36 ± 0.20 (25%)	> 20 ns
293K	0.38	0.330 ± 0.030 (22%)	2.14 ± 0.20 (23%)	> 20 ns
303K	0.38	0.206 ± 0.020 (29%)	2.03 ± 0.15 (26%)	> 20 ns
313K	0.38	0.185 ± 0.015 (34%)	1.61 ± 0.15 (27%)	> 20 ns
333K	0.387	0.140 ± 0.010 (28%)	0.81 ± 0.08 (26%)	> 20 ns
353K	0.39	0.110 ± 0.010 (27%)	0.54 ± 0.05 (33%)	> 20 ns

The stability and mobility of the BGL structure were studied by subjecting the structure to 48 ns all atom MD simulation at low (278K), moderate (313K), moderately high (333K) and

high (353K) temperatures. Figure 4.7. shows the RMSD and R_g plot of BGL at different temperatures. RMSD indicates the overall change in the geometry of the enzyme with respect to its initial structure.

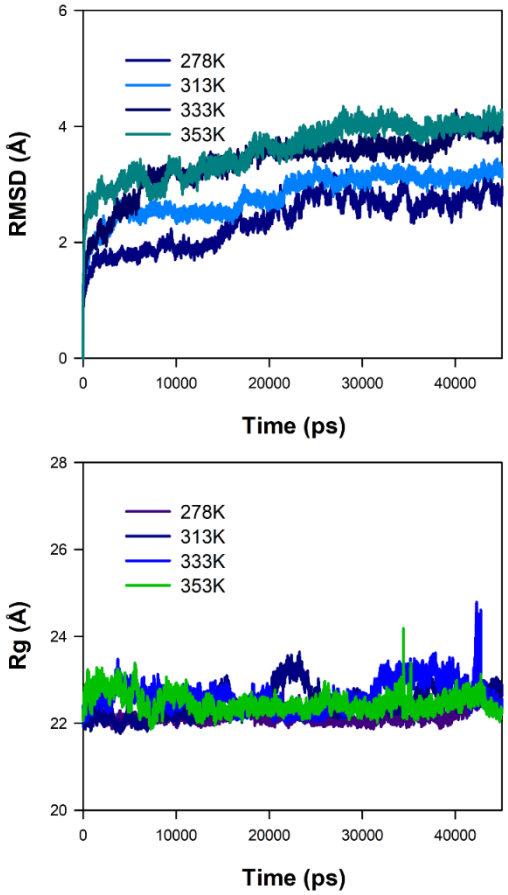

Figure 4.7. *Representation of almond BGL RMSD and Rg values over 48 ns simulation event at different temperatures.*

From figure 4.7 it becomes evident that after ~25 ns, the structure reaches an equilibrium geometry. Furthermore, there is a monotonous increase in the RMSD value with respect to the temperature echoing the effect of temperature on flexibility of the enzyme. In order to have an insight into the effect of temperature on the overall shape and compactness of the enzyme the radius of gyration (R_g) was calculated. R_g is defined as the average distance of the collection of atoms from their common centre of mass. Higher R_g values indicate looser packing of the protein structure whereas lower values correspond to a much compact structure. Figure 4.7 shows further that there is no significant change in the overall shape and compactness of BGL over the course of the simulation event corroborating our data from circular dichroism study.

Furthermore, flexibility in different parts of the enzyme (BGL) at different temperatures was assessed by calculating the Root mean squared fluctuation (RMSF) which is a measure of the displacement of an amino acid residue around its averaged position during a defined period of time and allows detection of the regions of high flexibility in a protein.

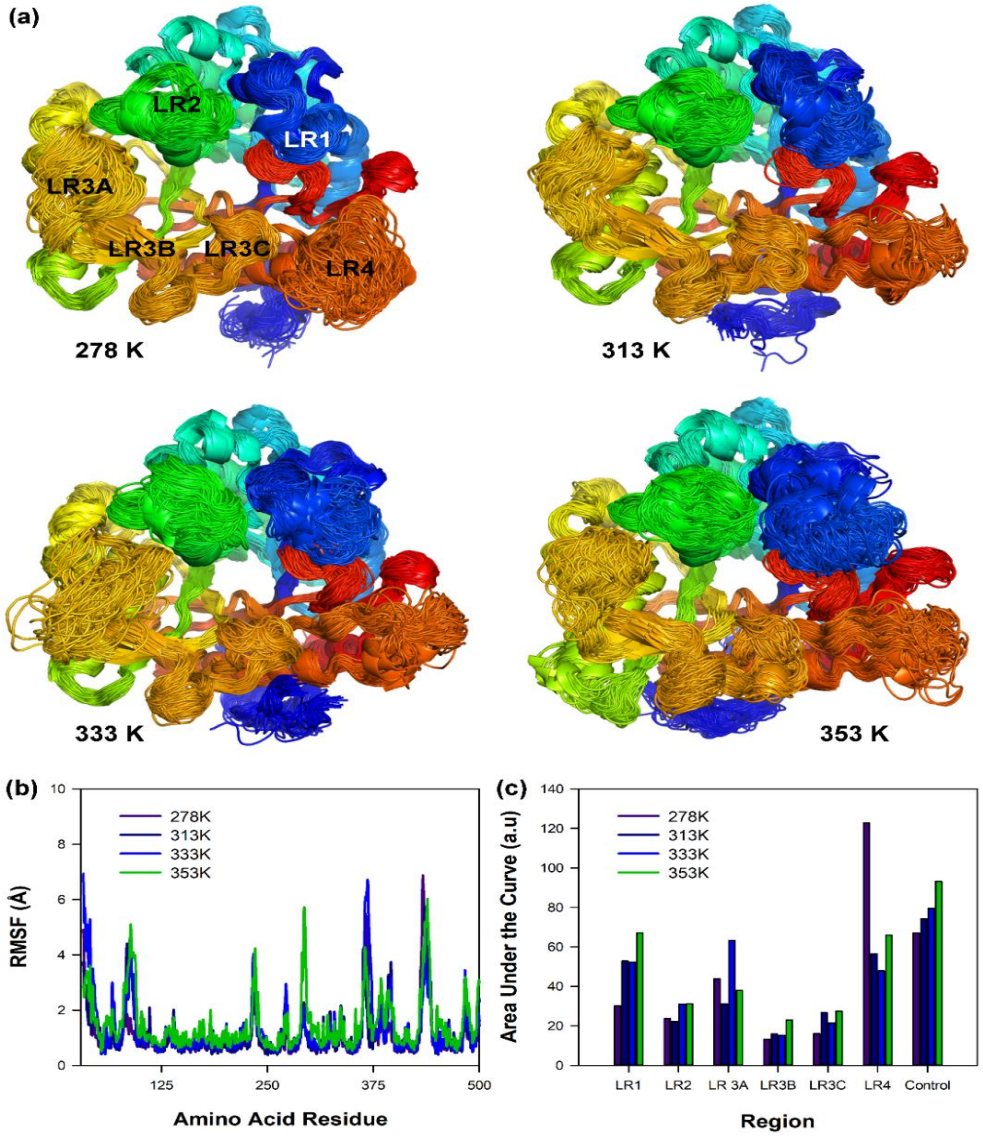

Figure 4.8. *Showing the all atom fluctuations of almond BGL at different temperatures. (a) Ribbon representation of RMS fluctuation at different temperatures. The LRs [colour code maintained as Figure 4.1] are the regions of high fluctuation [RMSF > 4Å]. (b) All atom RMSF of BGL over 48 ns of simulation event at different temperatures. The pattern shows atoms along the LRs as sharp peaks interspersed by low fluctuating atoms. (c) Calculated area under the curve for the LRs and Control at different temperatures. The corresponding amino acid residue for LR1 (77-100); LR2 (227-242); LR3A (359-376); LR3B (377-387); LR3C (388-402); LR4 (426-449); Control (132-222) [Control is a non-LR].*

Figure 4.8 shows RMS fluctuations of BGL at different temperatures. It is observed from figure 4.8.a and 4.8.b that the loop regions experience the most of the residue fluctuations in BGL. Four loop regions in particular form the epicentre of residue fluctuations: LR1, LR2, LR3, and LR4 (Figure 4.2) in the enzyme. These regions comprise of residues 77-100 (LR1), 227-242 (LR2), 359-376 (LR3a), 377-387 (LR3b), 388-402 (LR3c), and 426-449 (LR4). All these regions experience fluctuations greater than 4Å whereas the regions experiencing fluctuations lesser than 4Å were mostly found to be located in the binding site or the catalytic barrel having conserved residues. Figure 4.8.c shows fluctuations experienced by the LRs at different temperatures. It is clearly evident that the LR1 undergoes a monotonous increase in fluctuations as does the LR2 with rise in temperature. However, the LR3 region shows a form of synchronous fluctuations at its sub regions LR3a and LR3c while the LR3b remains rigid acting like a pivot. The LR4 region however displays an unusual fluctuation. The LR4 region shows increased fluctuations at 278K, the movement becomes fairly rigid at 313 and 333K, increasing again at 353K. It is noteworthy that figure 4.8.c does not indicate the magnitude of fluctuations but rather the effect of temperature change on the LRs with respect to control (non-LRs). The changing pattern of fluctuations in the LRs indicate a possible pattern over the course of temperature range. It should be noted that the LRs cover the catalytic barrel like a lid whose fluctuations likely contribute to or control the catalytic activity of the enzyme.

To obtain better insights and identify the overall pattern of motions of the entire protein principal component analysis (PCA) was used. The eigenvectors are calculated from the covariance matrix of a simulation trajectory. The trajectories are filtered along each eigenvector to identify the dominant motions during a simulation event. A large portion of the overall fluctuations of the macromolecules observed from RMSF can often be accounted for by a few low-frequency eigenvectors with large eigenvalues. If the motions are similar, the eigenvectors and eigenvalues coming from individual trajectories should be similar to each other. All atom principal component analysis was applied to study the major motions of BGL at different temperatures. Figure 4.9 and figure 4.10 display the porcupine plot of the first two eigenvectors and the mobility plots of BGL. The mobility plot (Figure 4.10) clearly shows an initial increase in motion in the LR3 with a decrease after 333K, which directly correlates with the catalytic activity of the enzyme. From the porcupine plots (Figure 4.9) we can clearly visualize the major motions of the protein at different temperatures. Figure 4.9 indicates that the LR region is the epicentre of major motions in the protein. At 278K the LR4 region serves as the centre of major

fluctuations while the other LR regions remain rigid. In case of 313K major motions in the LR3 region are observed.

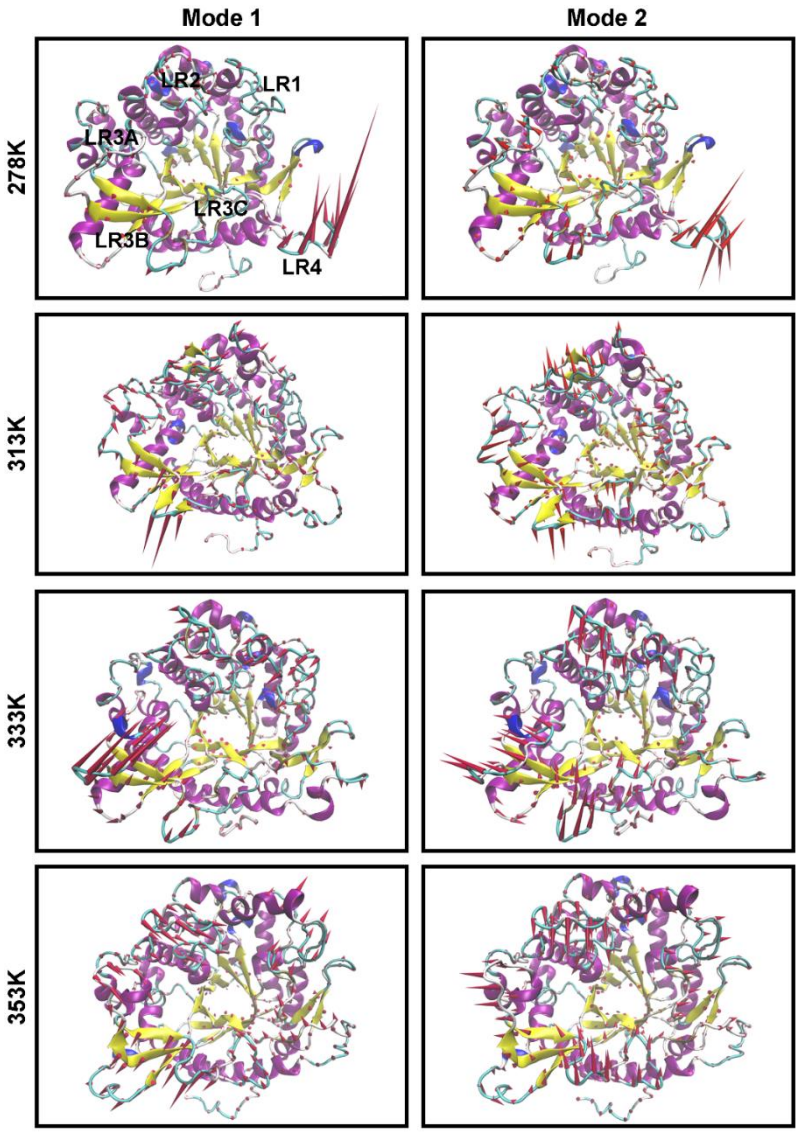

Figure 4.9. *Porcupine plots of first two eigenvector [Mode 1 and Mode 2] of almond BGL at different temperatures. The model shown here represents all atom trace. The arrows attached to each represent the direction of eigenvector and the size of each arrow signifies the magnitude of the corresponding eigenvalue. The motion of eigenvectors shows concerted motion of the LRs across 333K. The motion of eigenvectors become independent of each other and more random as temperature is raised to 353K.*

The porcupine as well as the mobility plots indicate the presence of three different motions at the LR3 region. The LR3a and LR3c regions show a propeller motion; moving in clockwise direction, while the LR3b region remains rigid acting like a pivot. The motion of the entire

protein is clockwise in the first mode. In the second mode, the motion of LR3 is not the primary motion and the protein as a whole move in an anti-clockwise direction relative to that of LR3. This suggests a possible change in the orientation of the loop region covering the catalytic barrel making it more accessible to the incoming substrate.

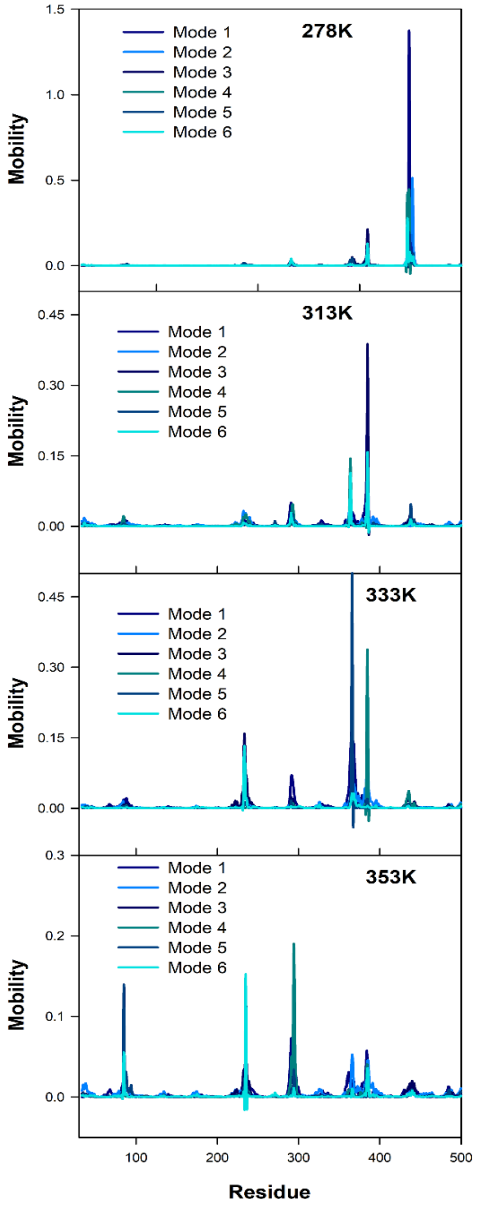

Figure 4.10. *Residue based mobility plot of almond BGL at different temperatures. The plot shows the regions of major motions along the residues across a simulation event. The atoms along residues of LRs are indicated as the epicentres of major fluctuations. Residues along the catalytic barrel are found to be rigid and experience less fluctuations.*

The principal motions for BGL at 333K illustrate the clockwise propeller motion of the LR3 region, while the entire protein experiences a counter-clockwise motion. Both the modes in the porcupine plot and the mobility plot indicate the presence of significant motions in the LR2 region. The presence of new sites of prominent motion could aid in a suitable orientation of the catalytic site for the approaching substrate, thereby, facilitating the enzymatic activity of BGL. The mobility plot at 353K shows the presence of several new sites of motion (LR1 and LR2). From the porcupine plots random and independent motions of different sites are visualized and the fluctuations of the LRs become highly inordinate and asynchronous which may affect the enzymatic activity at high temperatures.

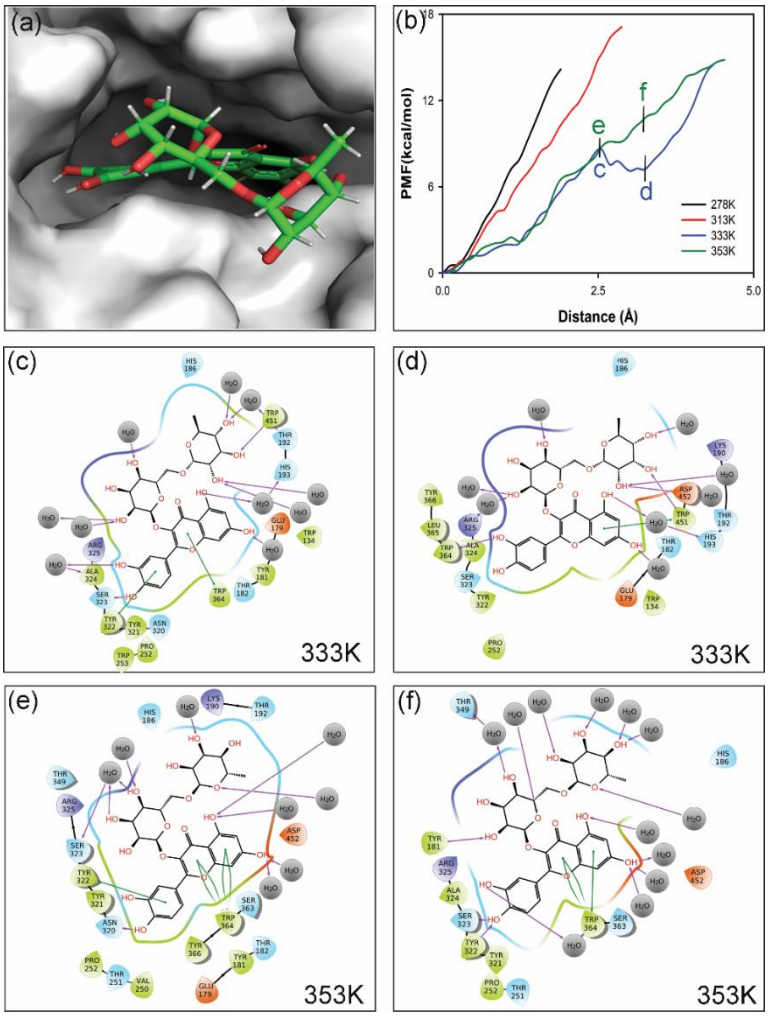

Figure 4.11. *Showing the potential of mean force (PMF) of almond β-glucosidase (BGL) over the distance along the catalytic barrel at different temperatures. (a) Representative figure showing the bound substrate (Rutin). (b) The potential of mean force (PMF) of rutin bound to the almond BGL over the distance along the catalytic barrel at different temperatures. Four different sections over the energy minimum were selected & the molecular interactions observed at 333K & 353K [(c); (d); (e); (f)].*

The Study of the energy required to pull a substrate molecule (rutin) out of the binding site of BGL (Figure 4.11.a), gave further clarity to the notion of ease of substrate binding at higher temperature. Following the methodology of steered molecular dynamics simulation (SMD) described in the earlier work [36], it was observed (Figure 4.11.b) that with increasing temperature less energy was required to break the enzyme substrate interactions and pull the substrate out. 333K was found to be the optimum temperature at which a meta-stable energy minimum was observed after a short pull (Figure 4.11.b). The molecular interaction studies at the minimum indicated compensating bond formation and breakdown with the substrate (Figure 4.11.c & 4.11.d). No compensating interactions were found for 353K, which resulted in the monotonous increase in the potential of mean force (PMF) (Figure 4.11.b, 4.11.e, 4.11.f). Detailed analysis of the interacting residues found the presence of water-bridge connecting Arg325 to the hydroxylic group adjacent to the glycosidic bond at the meta-stable state (Figure 4.11.c & 4.11.d). Our previous work had discussed the role of bulk and bound water molecules in biological recognition and hydrolysis [37]. The bound water molecule in the meta-stable state may be involved in the catalytic activity of the enzyme. Table 4.5 lists the equilibrium distances between the centre of mass of BGL and rutin.

Table 4.5. Distance from the Center of Mass of the Protein (BGL) to the Center of Mass of the Ligand (Rutin).

Temperature (K)	Distance (Å)
278K	12.02
313K	11.05
333K	9.78
353K	10.46

It was evident from the table that rutin (substrate) can go further into the active site of BGL at 333K than at any other temperature. Solvent exposure of the substrate at different temperatures (Table 4.6) was also consistent with the finding. Almond BGL causes the catalytic hydrolysis of the flavonoid glycoside Rutin to release Quercetin. Temperature dependent enzymatic activity was monitored in the temperature range of 278K-353K. We observed no product

formation at temperatures as low as 278K. Increased product formation at temperatures as high as 333K suggested the ease of the enzyme substrate complex formation.

Table 4.6. Solvent Exposure of the Ligand (Rutin) in Terms of the Interaction Energies with Water.

Temperature (K)	Energy (kcal/mol)
278K	-35.47
313K	-31.70
333K	-31.20
353K	-53.35

The kinetics parameter were calculated from the Lineweaver-Burk plot. The catalytic efficiency (kcat/Km) was a typical bell-shaped curve, with the temperature optima at 333K, which was in agreement with previous studies (Figure 4.4) [25]. Loss of catalytic activity was observed at 353K, however the thermal stability observed at this temperature pointed towards dynamical flexibility being the possible architect of the structure-activity relationship.

The results of activity study point to the dynamical flexibility as the possible facilitator of enzymatic activity at different temperatures. Studies by Banerjee and Pal [38], Karplus et al., [39] Min et al., [40] found a direct correlation between picosecond to nanosecond timescale dynamics and enzymatic activity. Picosecond resolved fluorescence anisotropy of the ANS-BGL complex comprehensively shows a monotonous increase in dynamical flexibility of the protein with an increase in temperatures (Figure 4.6 & Table 4.4). The rise in temperature creates a more labile microenvironment of the probe enhancing the orientational relaxation of ANS. The RMSD calculations further support the activity-flexibility correlation being solely responsible for catalysis (Figure 4.7). While the RMSF studies identify four loop regions [LR1; LR2; LR3; LR4] as the possible regions of high fluctuations (Figure 4.8). The loop regions (LRs) are placed strategically as a lid over the catalytic barrel. Other regions of the enzyme experiences fluctuations of lesser magnitude than the loop regions. Time-resolved anisotropy measurements and RMSF analysis as well as the SMD simulations indicate to the possible gatekeeping function of the LRs where the change in fluctuations of the LRs associated with their possible orientation and re-orientation at an optimized position dictates the entry of an

incoming substrate at the active site of the catalytic barrel. Our study also indicates that higher degree of fluctuations beyond the threshold level at 353K (Figure 4.6) likely cause the LRs to an unfavourable orientation for entry of the substrate at the active site, leading to reduced catalytic efficiency (k_{cat}/K_m) (Figure 4.4).

The RMSF studies indicate that the four strategically placed LRs as the possible gatekeepers of enzymatic activity at different temperatures. The overall fluctuations of the macromolecule observed from RMSF cannot provide concrete information regarding major motions along the trajectories of simulation. The trajectories need to be filtered along each eigenvector to identify the dominant motions during a simulation event. The principal component analysis (PCA) of BGL at different temperatures identified the LRs as the epicentre of dominant motions at different temperatures. Our study also found that residues of the catalytic barrel experienced no major fluctuations across all temperature ranges. The dynamic rigidity of the LRs at 278K barring LR4 likely results in a possible closed gate conformation checking the access of an incoming substrate in the catalytic site and as a consequence no enzymatic activity at the said temperature (Figure 4.9 & 4.10) has been noted. The concerted motions of LR3 and rest of the protein, along with the propeller motion of LR3, indicate a change in orientation to an open gate conformation at 313K (Figure 4.9 & 4.10). The open gate conformation would allow access to the substrate resulting in enzymatic activity at 313K. The concerted motion was maintained at 333K and new centres of prominent motions at LR1 and LR2 were observed (Figure 4.9). These new sites of prominent motion could aid in the optimum orientation of the lid LRs aiding in enhanced catalysis and turnover observed at 333K (Table 4.3). the ease of substrate access at this temperature also reverberated in the least pulling force as observed in the SMD simulations (Figure 4.11). At 353K there is a loss of concerted motions of the LRs and fluctuations become independent of each other and random with several new sites of major motions (Figure 4.9). This loss in the concerted motion of the LRs and increased randomness may rationalize the increased flexibility of the ANS-BGL complex observed in polarization gated anisotropy measurements (Figure 4.6) and the fall in the catalytic activity and turnover at 353K (Figure 4.4 & Table 4.3). The random fluctuations of the LRs may lead to the inability or possibly dissociation of the substrate from the active site of the enzyme (BGL) contributing to reduced catalytic efficiency.

4.3. Conclusion:

The 3D-model and secondary structure analysis of an enzyme almond BGL illustrated various remarkable characteristics of glycosyl hydrolase. The structure of this thermostable enzyme consists of four looped regions (LRs) placed strategically over the catalytic barrel as a lid. The concerted motion of these LRs likely acts as a gatekeeper of catalytic activity of the enzyme at different temperatures, and consequently, provides it with the ability to be active at moderately high temperatures. The findings of this study can provide a background for future efforts to enhance enzymatic capabilities through rational design of recombinant enzymes.

References

[1] Y. Bhatia, S. Mishra, V.S. Bisaria, Microbial β-glucosidases: cloning, properties, and applications, *Crit. Rev. Biotechnol.*, 22 (2002) 375-407.

[2] J.R.K. Cairns, A. Esen, β-Glucosidases, *Cell. Mol. Life Sci.*, 67 (2010) 3389-3405.

[3] J.D. McCarter, G.S. Withers, Mechanisms of enzymatic glycoside hydrolysis, *Curr. Opin. Struct. Biol.*, 4 (1994) 885-892.

[4] J. Xiao, T.S. Muzashvili, M.I. Georgiev, Advances in the biotechnological glycosylation of valuable flavonoids, *Biotechnol. Adv.*, 32 (2014) 1145-1156.

[5] R.R. Singhania, A.K. Patel, R.K. Sukumaran, C. Larroche, A. Pandey, Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production, *Bioresour. Technol.*, 127 (2013) 500-507.

[6] O. Shoseyov, B.A. Bravdo, D. Siegel, A. Goldman, S. Cohen, L. Shoseyov, R. Ikan, Immobilized endo-. beta.-glucosidase enriches flavor of wine and passion fruit juice, *J. Agric. Food Chem.*, 38 (1990) 1387-1390.

[7] E. Su, T. Xia, L. Gao, Q. Dai, Z. Zhang, Immobilization of β-glucosidase and its aroma-increasing effect on tea beverage, *Food Bioprod. Process.*, 88 (2010) 83-89.

[8] R. Agrawal, A. Verma, A. Satlewal, Application of nanoparticle-immobilized thermostable β-glucosidase for improving the sugarcane juice properties, *Innov. Food Sci. Emerg. Technol.*, 33 (2016) 472-482.

[9] G. Singh, A.K. Verma, V. Kumar, Catalytic properties, functional attributes and industrial applications of β-glucosidases, *3 Biotech*, 6 (2016) 3.

[10] A.V. Morant, K. Jørgensen, C. Jørgensen, S.M. Paquette, R. Sánchez-Pérez, B.L. Møller, S. Bak, β-Glucosidases as detonators of plant chemical defense, *Phytochemistry*, 69 (2008) 1795-1813.

[11] P.E. Belchetz, J.C.W. Crawley, I.P. Braidman, G. Gregoriadis, Treatment of Gaucher's disease with liposome-entrapped glucocerebroside: β-glucosidase, *The Lancet*, 310 (1977) 116-117.

[12] B. Henrissat, G. Davies, Structural and sequence-based classification of glycoside hydrolases, *Curr. Opin. Struct. Biol.*, 7 (1997) 637-644.

[13] H.M. Jespersen, E.A. MacGregor, B. Henrissat, M.R. Sierks, B. Svensson, Starch-and glycogen-debranching and branching enzymes: prediction of structural features of the catalytic (β/α) 8-barrel domain and evolutionary relationship to other amylolytic enzymes, *J. Protein Chem.*, 12 (1993) 791-805.

[14] Y. Li, X. Hu, J. Sang, Y. Zhang, H. Zhang, F. Lu, F. Liu, An acid-stable β-glucosidase from Aspergillus aculeatus: gene expression, biochemical characterization and molecular dynamics simulation, *Int. J. Biol. Macromol.*, 119 (2018) 462-469.
[15] N.S.F. Mazlan, N.B. Ahmad Khairudin, A molecular dynamics study of Beta-Glucosidase B upon small substrate binding, *J. Biomol. Struct. Dyn.*, 34 (2016) 1486-1494.
[16] P. Biswas, A. Adhikari, U. Pal, P. Singh, M. Das, T. Saha-Dasgupta, S.S. Choudhury, R. Das, S.K. Pal, Flexibility modulates the catalytic activity of a thermostable enzyme: key information from optical spectroscopy and molecular dynamics simulation, *Soft matter*, 16 (2020) 3050-3062.
[17] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, 25 (1997) 3389-3402.
[18] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, *Nat. Protoc.*, 10 (2015) 845-858.
[19] J. Söding, A. Biegert, A.N. Lupas, The HHpred interactive server for protein homology detection and structure prediction, *Nucleic Acids Res.*, 33 (2005) W244-W248.
[20] T. Barrett, C.G. Suresh, S.P. Tolley, E.J. Dodson, M.A. Hughes, The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase, *Structure*, 3 (1995) 951-960.
[21] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.*, 31 (2003) 3381-3385.
[22] J.U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure, *Science*, 253 (1991) 164-170.
[23] R. Opassiri, B. Pomthong, T. Onkoksoong, T. Akiyama, A. Esen, J.R.K. Cairns, Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase, *BMC Plant Biol.*, 6 (2006) 1-19.
[24] G. Davies, B. Henrissat, Structures and mechanisms of glycosyl hydrolases, *Structure*, 3 (1995) 853-859.
[25] N.S. Terefe, P. Sheean, S. Fernando, C. Versteeg, The stability of almond β-glucosidase during combined high pressure–thermal processing: a kinetic study, *Appl. Microbiol. Biotechnol.*, 97 (2013) 2917-2928.
[26] O. Márquez, K.N. Waliszewski, The effect of thermal treatment on β-glucosidase inactivation in vanilla bean (Vanilla planifolia Andrews), *Int. J. Food Sci.*, 43 (2008) 1993-1999.

[27] M. Lecas, Z.Y. Gunata, J.-C. Sapis, C.L. Bayonove, Purification and partial characterization of β-glucosidase from grape, *Phytochemistry*, 30 (1991) 451-454.
[28] A. Esen, Purification and partial characterization of maize (Zea mays L.) β-glucosidase, *Plant Physiol.*, 98 (1992) 174-182.
[29] Y.-Y. Li, C.-J. Jiang, X.-C. Wan, Z.-Z. Zhang, D.-X. Li, Purification and partial characterization of β-glucosidase from fresh leaves of tea plants (Camellia sinensis (L.) O. Kuntze), *Acta Biochim. Biophys. Sin.*, 37 (2005) 363-370.
[30] M.A. Shah, S. Mishra, T.K. Chaudhuri, Structural stability and unfolding transition of β-glucosidases: a comparative investigation on isozymes from a thermo-tolerant yeast, *Eur. Biophys. J.*, 40 (2011) 877-889.
[31] P. Hazra, D. Chakrabarty, A. Chakraborty, N. Sarkar, Probing protein-surfactant interaction by steady state and time-resolved fluorescence spectroscopy, *Biochem. Biophys. Res. Commun.*, 314 (2004) 543-549.
[32] S.K. Pal, J. Peon, A.H. Zewail, Ultrafast surface hydration dynamics and expression of protein functionality: α-Chymotrypsin, *Proc. Natl. Acad. Sci. U.S.A*, 99 (2002) 15297-15302.
[33] K. Sahu, S.K. Mondal, S. Ghosh, D. Roy, K. Bhattacharyya, Temperature dependence of solvation dynamics and anisotropy decay in a protein: ANS in bovine serum albumin, *J. Chem. Phys.*, 124 (2006) 124909.
[34] G. Wang, Y. Gao, M.L. Geng, Analysis of heterogeneous fluorescence decays in proteins. Using fluorescence lifetime of 8-anilino-1-naphthalenesulfonate to probe apomyoglobin unfolding at equilibrium, *Biochim. Biophys. Acta Gen. Subj.*, 1760 (2006) 1125-1137.
[35] P. Sarkar, S. Bharill, I. Gryczynski, Z. Gryczynski, M.P. Nair, A.G. Lacko, Binding of 8-anilino-1-naphthalenesulfonate to lecithin: cholesterol acyltransferase studied by fluorescence techniques, *J. Photochem. Photobiol. B, Biol.*, 92 (2008) 19-23.
[36] S. Mao, J.-W. Wang, F. Liu, Z. Zhu, D. Gao, Q. Guo, P. Xu, Z. Ma, Y. Hou, X. Cheng, Engineering of 3-ketosteroid-Δ 1-dehydrogenase based site-directed saturation mutagenesis for efficient biotransformation of steroidal substrates, *Microb. Cell Factories*, 17 (2018) 1-13.
[37] S.K. Pal, A.H. Zewail, Dynamics of water in biological recognition, *Chem. Rev.*, 104 (2004) 2099-2124.
[38] D. Banerjee, S.K. Pal, Conformational dynamics at the active site of α-chymotrypsin and enzymatic activity, *Langmuir*, 24 (2008) 8163-8168.
[39] M. Karplus, J. Kuriyan, Molecular dynamics and protein function, *Proc. Natl. Acad. Sci. U.S.A*, 102 (2005) 6679-6685.

[40] W. Min, X.S. Xie, B. Bagchi, Two-dimensional reaction free energy surfaces of catalytic reaction: Effects of protein conformational dynamics on enzyme catalysis, *J. Phys. Chem. B*, 112 (2008) 454-466.

Chapter 5

Studies on the Correlation between Structure, Physiological Function, and Essential Dynamics of a Homeothermic Enzyme

5.1. Introduction:

α-Chymotrypsin (CHT) is a proteolytic enzyme of the group S1 serine protease (EC 3.4.21.1) catalyze the hydrolysis of a peptide bond and are among the most widely distributed enzyme in the biological strata. CHT has attracted considerable interest due to its importance in understanding protein folding or unfolding. CHT play crucial role in several important biological processes such as protein digestion [1, 2], immune response [3, 4] and insect molting [5, 6]. CHTs are also known to play a vital role in the intracellular protein turnover [7]. CHTs are used extensively in meat industry [8], brewery [9] and in pharmaceutical industries [10, 11].

Studies investigating the role of essential motions in enzyme activity are poorly understood. Previous studies have investigated the correlation between structure and activity. However, most of these have been confined either at the active site or upon binding of small molecules to CHT [12, 13], and provide very little information on the essential motions of enzyme controlling catalysis. Enzyme activity is dependent on the mobility of the protein scaffold and any change associated to such motions can produce different level of selectivity and activity [14, 15]. Therefore, it is of our interest to investigate and study the essential motions governing catalysis in CHT.

In the present study we observe the presence of five loop regions (LRs) near the catalytic S1 pocket of CHT. We hypothesize that these LRs manifest a gatekeeping role controlling catalysis at different temperatures. In the current study, we correlate the essential motions of CHT with its catalytic activity using steady state, circular dichroism (CD) and picosecond resolved spectroscopy along with all atom molecular dynamics (MD) simulation and principal component analysis (PCA).

5.2. Results and Discussion:

5.2.1. Essential Loop Dynamics Modulates Catalytic Activity in α-Chymotrypsin [16]:

Thermal stability of CHT was quantified by circular dichroism (CD) spectroscopy in the far UV region. The far UV-CD spectra of CHT was characterized by a minimum at ~202 nm with no positive band [17, 18] (Figure 5.1.a).

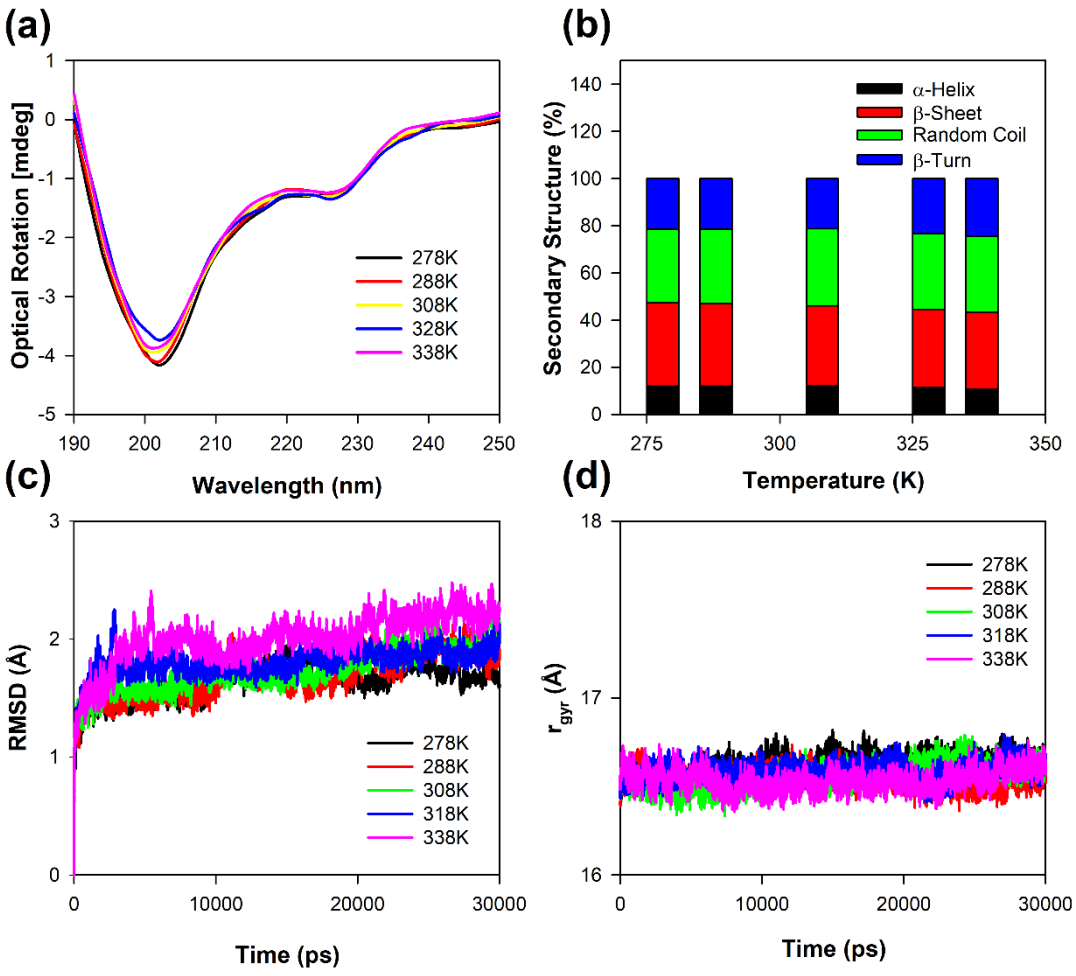

Figure 5.1. *Structural analysis and thermal stability of CHT. (a) Far UV CD spectra of CHT at different temperatures. (b) Secondary structure analysis of CHT. Insignificant changes to the CHT secondary structure was observed. (c) RMSD analysis of CHT over 30 ns MD simulation event at different temperatures. (d)* r_{gyr} *of CHT over 30 ns simulation event at different temperatures. Both the RMSD and* r_{gyr} *analysis indicated no major disruption to the CHT architecture.*

Deconvolution of the spectra displays insignificant changes in the CHT secondary structure indicating to its thermal stability and absence of widespread thermal denaturation of the enzyme within the temperature range of our study. 1.4% decrease in α-helix, 2.6% decrease in β-sheet, 1.1% increase in random coil were observed (Table 5.1) (Figure 5.1.b).

Table 5.1 Temperature dependent secondary structure of CHT as deconvoluted by CDNN

Temperature (K)	% α-Helix	% β-Sheet	% Random Coil	% β-Turn
278K	12	35.30	31.20	21.50
288K	12	35.10	31.60	21.30
308K	12.10	34	32.70	21.20
328K	11.40	33.20	32.10	23.30
338K	10.60	32.70	32.30	24.40

These are reflective of an insignificant change in the secondary structure of CHT upon thermal treatment, suggesting thermal stability of the enzyme within the limitations of our study.

To understand and correlate the effect of temperature on the overall geometry and compactness of the enzyme 30 ns all atom MD simulations were run at different temperatures- 278K, 288K, 308K, 318K, 338K. The MD trajectories were investigated for RMSD and radius of gyration (r_{gyr}). RMSD is a parameter for assessing the stability of MD trajectories. RMSD signifies the deviation from the initial structure which indicates a change in geometry of the enzyme to the initial structure [15]. It was quite evident from figure 5.1.c that the structure reached an equilibrium geometry after ~20 ns. The RMSD was steady over time at different temperatures and had reached a stable conformation. To understand the effect of temperature on the overall compactness of the enzyme the radius of gyration (r_{gyr}) was calculated. r_{gyr} is a measure of the average distribution of atoms from their common center of mass. Lower r_{gyr} values corresponds to a compact structure while higher r_{gyr} describe a loose packing of the structure. As documented from figure 5.1.d we observe no significant change in the overall compactness of the enzyme during the course of the simulation event corroborating with our circular dichroism (CD) study.

A detailed kinetics of the CHT protease activity was studied in temperatures ranging from 278K to 338K. For different concentrations (20-200 µM) of Ala-Ala-Phe-AMC (substrate) the initial velocity corresponding to protease activity were measured. Lineweaver-Burk plot was used to determine the kinetics parameters (K_m, V_{max}, k_{cat}, and k_{cat}/K_m). CHT belongs to the class of serine protease that catalyzes the hydrolysis of Ala-Ala-Phe-AMC (substrate) into 7-amido-4-methycoumarin (Figure 5.2.a & 5.2.b). Figure 5.2.c & 5.2.d illustrate the change in the product formation and catalytic efficiency upon temperature change.

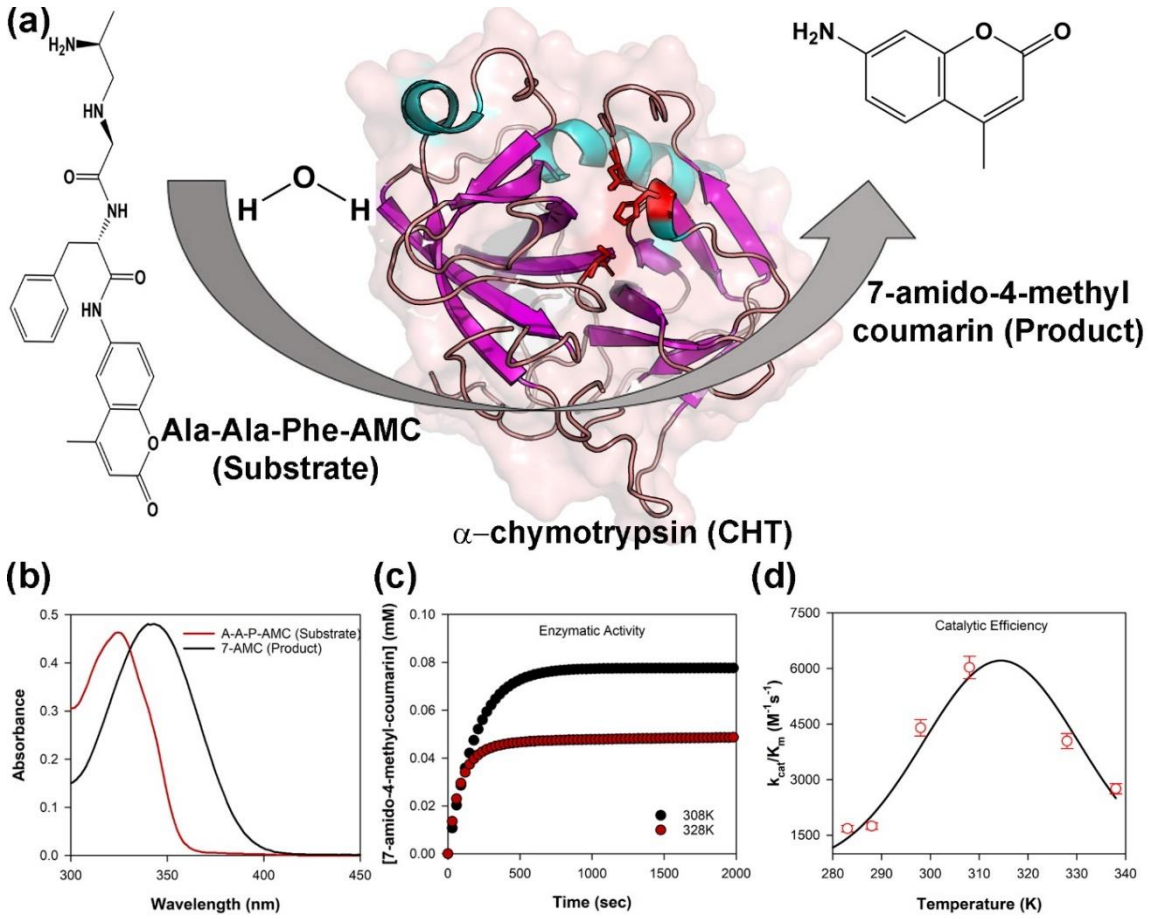

Figure 5.2. *Showing the temperature dependent enzyme catalysis of CHT. (a) Skeletal outline of reaction catalysed by CHT. (b) Absorbance spectra of substrate Ala-Ala-Phe-AMC and the catalysed product 7-amido-4-methyl coumarin. (c) Effect of different representative temperature on product formation. An increase in temperature showed reduced product formation beyond 308K. (d) Catalytic efficiency (k_{cat}/K_m) of CHT at different representative temperature generate a typical bell-shaped curve.*

In the study thermal treatment was found to enhance catalytic efficiency, reaching maximum efficiency at 308K. Further increase in temperature beyond 308K led to fall in catalytic

In the study thermal treatment was found to enhance catalytic efficiency, reaching maximum efficiency at 308K. Further increase in temperature beyond 308K led to fall in catalytic efficiency. It should also be noted that low catalytic efficiency and turnover was observed at temperatures below 308K. The result is consistent with our thermal stability study indicating the favourable formation of a stable enzyme-substrate complex over a wide range of temperatures. The summary of the kinetics parameters are shown in Table 5.2.

Table 5.2. Temperature dependent kinetics of CHT catalysed hydrolysis of Ala-Ala-Phe-AMC

Temperature(K)	K_m (M)	V_{max} (μM s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
278K	$(0.75\pm0.03) \times 10^{-3}$	0.32±0.016	0.64±0.032	853±42.65
283K	$(1.0\pm0.05) \times 10^{-3}$	0.89±0.044	1.78±0.089	1679±83.95
288K	$(1.27\pm0.06) \times 10^{-3}$	1.11±0.055	2.22±0.111	1748±87.40
298K	$(1.75\pm0.08) \times 10^{-3}$	3.85±0.192	7.7±0.385	4400±220
308K	$(2.0\pm0.1) \times 10^{-3}$	5.97±0.298	11.94±0.597	6030±301.50
328K	$(0.89\pm0.04) \times 10^{-3}$	1.88±0.094	3.6±0.180	4044±202.20
338K	$(0.1\pm0.005) \times 10^{-3}$	0.142±0.007	0.28±0.014	2757±137.85

Time resolved fluorescence anisotropy (r(t), Figure 5.3) decays of ANS [19-21] in CHT at different temperatures are characterized by three rotational correlation times (Table 5.3). The time constant (φ_1) of ~54-67 ps corresponds to orientational relaxation of the dye in water due to its close resemblance to the rotational relaxation time of ANS in water (~70 ps) [19].

Table 5.3. Time resolved anisotropy of ANS-CHT at different temperatures (During fitting of anisotropy decay the longer time constant (φ_3) is held fixed at 40 ns)

Temperature (K)	φ_1 (ns)	φ_2 (ns)	φ_3 (ns)
278K	0.063 (33 %)	0.705 (40 %)	40
288K	0.059 (52 %)	0.707 (24 %)	40
298K	0.067 (54 %)	0.841 (22 %)	40
308K	0.054 (50 %)	0.555 (22 %)	40
328K	0.055 (48 %)	0.676 (10 %)	40

On the other hand, the longer sub-nanosecond component (φ_2) may be ascribed to segmental motions of the amino acid residues Cys1-Cys122 of the protein, because, ANS binds rigidly at a single site on the surface of CHT near the Cys1-Cys122 disulfide bond [19] which is localized almost opposite to the catalytic centre of the enzyme.

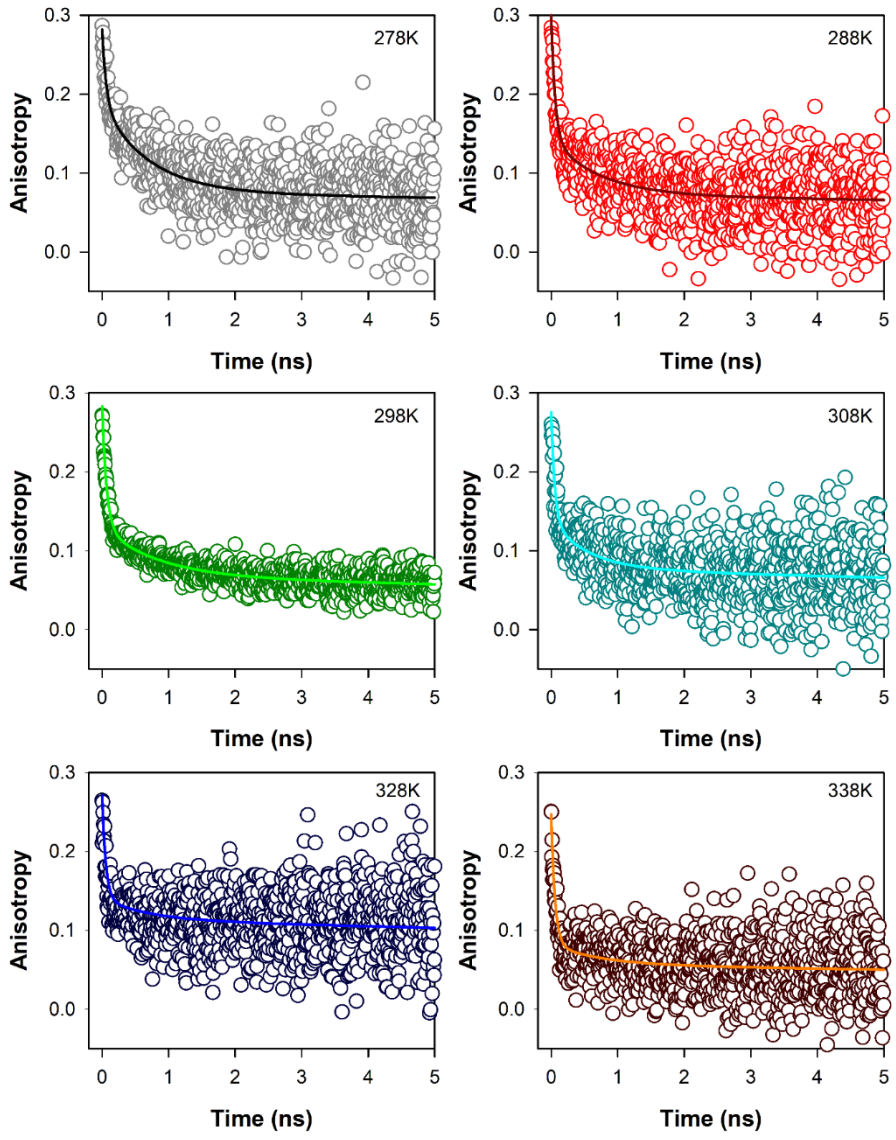

Figure 5.3. *Temperature dependent fluorescence anisotropy of ANS in CHT; solid coloured lines represent fits to raw anisotropy data.*

The longest time constant (φ_3) represents the global tumbling motion of the ANS/CHT complex. Upon an increase of temperature the sub-nanosecond component (φ_2) remains more or less unchanged, until at 308K it decreases significantly to 550 ps (Table 5.3). This is reflective of significantly increased segmental mobility of the amino acid residues opposite to

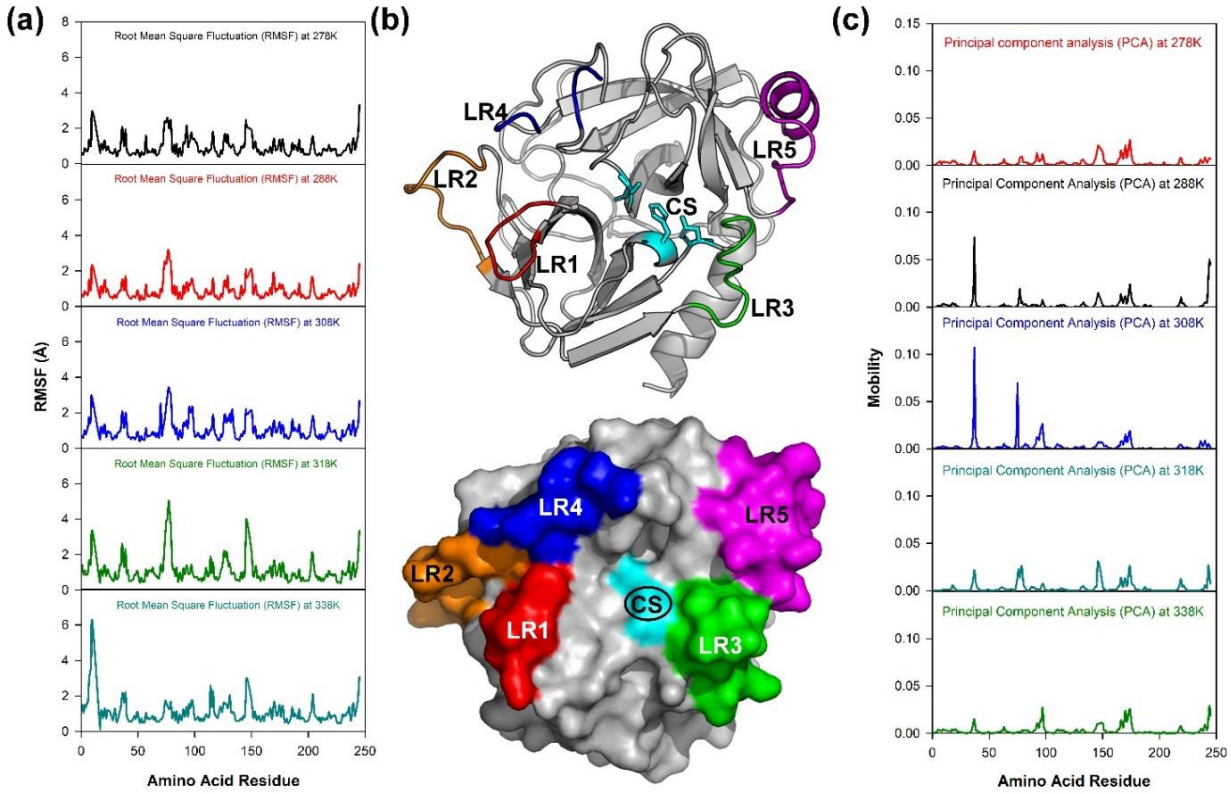

Figure 5.4. *(a) All atom RMSF of CHT over 30 ns of simulation event at different temperature. The RMSF pattern shows regions of high fluctuation along with regions of low fluctuations. (b) Ribbon representation showing the various loop regions (LRs). The coloured region represents each LRs. LR1 [Red]; LR2 [Orange]; LR3 [Green]; LR4 [Blue]; LR5 [Magenta]. The corresponding amino acid residue for LR1 (34-40), LR2 (73-81), LR3 (91-99), LR4 (143-153), LR5 (165-178). Space filling model showing the positioning of the LRs around the catalytic site (CS); the LRs are represented by same colour scheme. (c) Mobility plot of CHT at different temperatures. The plot shows the epicentre of fluctuations along the amino acid residues across a simulation event. The residues along the LRs are the epicentre of major motions. Regions other than the LRs including the catalytic pocket experience no major fluctuations.*

the catalytic centre. As a consequence of such increased flexibility of the amino acid residues opposite to the catalytic centre it likely becomes increasingly accessible to the substrate resulting in optimal enzymatic activity of CHT. On further increase of temperature, the rotational relaxation time (φ_2) increases indicating decreased flexibility of the amino acid residues Cys1-Cys122 with a consequent decrease in the catalytic efficiency of the enzyme.

To assess the mobility and flexibility in the different parts of the enzyme root mean square fluctuations (RMSF) of CHT at different temperatures was calculated. RMSF is a measure of the displacement of an amino acid residue around its averaged position during a defined period of time and allows the detection of the regions of high flexibility in a protein. Figure 5.4.a shows the RMS fluctuations of CHT at different temperatures. As documented from figure 5.4.a that the loop regions (LRs) experience most of the amino acid residue fluctuations in CHT. There are 5 loop regions (LRs) in particular that form the epicentre of the residue fluctuations: LR1, LR2, LR3, LR4, LR5 (Figure 5.4.b.). These LRs comprise of amino acid residues ranging from- LR1 (34-40), LR2 (73-81), LR3 (91-99), LR4 (143-153), LR5 (165-178). All these regions experience fluctuations greater than 2Å, while the other regions experience fluctuations lesser than 2Å with most of them residing at the catalytic S1 pocket. A quick scrutiny of the CHT structure show that the LRs are located around the catalytic pocket the fluctuations of which may control the catalytic activity of CHT. RMS fluctuations however cannot describe the magnitude and direction of the major motions of the protein. Hence normal mode analysis was performed. Protein dynamics is demonstrated as a change in molecular conformation as a function of time. To describe the essential motions of the protein over a broad time and spatial scale, protein conformations are best characterized as vector space that spans a large number of dimensions equal to the number of degrees of freedom (DOF) selected to characterize the motions [22]. To generate better insights into the degree of freedom (DOF) and the interpretation of the trajectories principal component analysis (PCA) was performed. The eigenvectors are calculated from the covariance matrix of a simulation trajectory. Along each eigenvector the trajectories are filtered to identify the dominant motion during a simulation event [23, 24]. Overall fluctuations of the macromolecules observed from RMSF can often be accounted for by a few low-frequency eigenvectors with large eigenvalues. Study of such eigenvectors and eigenvalues can interpret the essential motions of macromolecules. To study the major motions of CHT all atom principal component analysis (PCA) was applied at different temperatures. Figure 5.4.c and figure 5.5 illustrates the mobility plot of CHT and the porcupine plot of the eigenvector respectively. The motion of LR1 and LR2 (Figure 5.4.c)

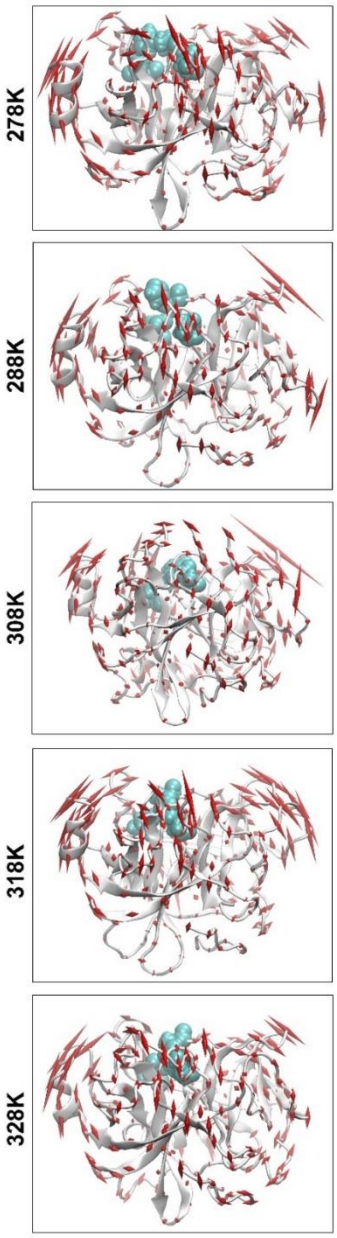

Figure 5.5. *Porcupine plots of CHT at different temperatures. The above model is a backbone trace of CHT. The arrows attached correspond to the direction of each eigenvector while the size of each arrow represents the magnitude of each eigenvalue. The motion of eigenvectors demonstrates synchronous motions of the LRs at 308K. Eigenvector motions become more random and there is loss of concerted eigenvector motion beyond 308K.*

clearly shows increase in synchronous motion with subsequent fall after 308K which correlates with the CHT activity. The porcupine plots (Figure 5.5) clearly shows the major motions of CHT at different temperatures. Figure 5.4.c & figure 5.5 specifies the LR regions as the centre

of major motions in the protein. The mobility and the porcupine plot specify the LR regions as the centre of major motions in the protein. At 278K all the LRs experience and act as the centre of major motions. Magnitude of LR motions increase with increase in temperature. At 308K LR1, LR2 and LR5 experience an anti-clockwise motion while LR3 and LR4 moves in clockwise direction. Such synchronous motion can re-orient the LRs positioned around the catalytic pocket making it more accessible to the incoming substrate. Both the porcupine and the mobility plots as well as the motional analysis indicate the synchronous motion of the LRs with increase in temperature. The presence of sites of synchronous motion could aid in a suitable orientation of the catalytic site for the approaching substrate, thereby, facilitating the enzymatic activity of CHT. However, temperatures above 308K shows the presence of new sites of motion. The mobility plots clearly illustrate new random and independent sites of motion and the loss of synchronous motion in the LRs which may affect catalysis at higher temperatures.

CHT causes the catalytic hydrolysis of substrate Ala-Ala-Phe-AMC into 7-amido-4-methyl coumarin. The catalytic activity was monitored at different temperatures [278K-338K]. Enhanced product formation suggested the ease of product formation at increased temperatures. Kinetic parameters were calculated from Lineweaver-Burk plot. The catalytic efficiency (k_{cat}/K_m) was a typical bell-shaped curve, optimum temperature at 308K (Figure 5.2) which agrees with a previous study [12]. Although a loss in the catalytic efficiency of CHT was noted at 328 and 338K, our thermal stability study at these temperatures indicate conformational flexibility as the major architect of the catalytic activity (Figure 5.1 & Table 5.1). The RMSD and r_{gyr} calculations also corroborate with our thermal stability analysis. Time resolved fluorescence anisotropy of ANS shows significant increase in segmental mobility of the amino acid residues at 308K opposite to the catalytic centre of CHT in comparison to other temperatures below and above 308K. The RMSF calculations show five loop regions (LR1, LR2, LR3, LR4, LR5) as the possible centres of major motions (Figure 5.4.a). A closer inspection of the CHT structure show the LRs are strategically located near the catalytic pocket (Figure 5.4.b). The study further indicates that five LRs may act as possible facilitators of catalytic activity at different temperatures. However, the fluctuations of CHT observed from RMSF cannot provide tangible information regarding the centres of major motions along the simulation trajectory. Principal component analysis (PCA) also indicate the LRs as the epicenter of major motions while the other regions of CHT experience fluctuations of lesser magnitude across all temperatures. The motions of the LRs become less concerted at 278K

resulting in reduced catalytic activity (Figure 5.4.c & Figure 5.5) whereas the concerted motion is increased at 288K reflecting on a change in orientation of the LRs into an open conformation. The concerted synchronous motion of the LRs was maintained at 308K (Figure 5.4.c & Figure.5.5). The concerted synchronous motion of the LRs was maintained at 308K (Figure 5.4.c & Figure 5.5) and their open conformation would facilitate an easy access to the incoming substrate leading to increased catalytic efficiency and enhanced turnover at 308K (Table 5.2). At 338K we observe a loss in concerted motion coupled with the generation of new sites of fluctuations (Figure 5.4.c & Figure 5.5). This loss in concerted motion of the LRs and amplified randomness may justify the reduction in catalytic efficiency and turnover at temperatures above 308K because increased randomness of the macromolecule and loss of synchronous concerted motion of LRs may lead to a possible dissociation from or inability of the substrate to enter the catalytic pocket.

5.3. Conclusion:

Study of the catalytic efficiency of CHT over a range of temperature from 278K to 338K displayed a typical bell shape nature with optimum efficiency being shown at 308K. Secondary structure analysis has indicated maintenance of thermal stability within the studied range of temperature. Polarization-gated fluorescence anisotropy of ANS in CHT display significantly increased segmental mobility of the amino acid residues opposite to the entry of the catalytic centre of the enzyme at 308K in comparison to other temperatures. MD simulations show that synchronous motion of five strategically placed loop regions (LRs) around the catalytic S1 pocket influences and modulates the catalytic activity of CHT at different temperatures. This study manifests the correlation between essential motions in the enzyme to that of its catalytic efficiency which can provide information for future efforts in protein engineering.

References

[1] S. Mazumdar-Leighton, R.M. Broadway, Identification of six chymotrypsin cDNAs from larval midguts of Helicoverpa zea and Agrotis ipsilon feeding on the soybean (Kunitz) trypsin inhibitor, *Insect Biochem. Mol. Biol.*, 31 (2001) 633-644.

[2] G. Broehan, M. Kemper, D. Driemeier, I. Vogelpohl, H. Merzendorfer, Cloning and expression analysis of midgut chymotrypsin-like proteinases in the tobacco hornworm, *J. Insect Physiol.*, 54 (2008) 1243-1252.

[3] H. Jiang, Y. Wang, X.-Q. Yu, M.R. Kanost, Prophenoloxidase-activating Proteinase-2 from Hemolymph of Manduca sexta: A bacteria-inducible serine proteinase containing two clip domains, *J. Biol. Chem.*, 278 (2003) 3552-3561.

[4] H. Jiang, Y. Wang, X.-Q. Yu, Y. Zhu, M. Kanost, Prophenoloxidase-activating proteinase-3 (PAP-3) from Manduca sexta hemolymph: a clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs, *Insect Biochem. Mol. Biol.*, 33 (2003) 1049-1060.

[5] R.I. Samuels, S.E. Reynolds, Molting fluid enzymes of the tobacco hornworm, Manduca sexta: Timing of proteolytic and chitinolytic activity in relation to pre-ecdysial development, *Arch. Insect Biochem. Physiol.*, 24 (1993) 33-44.

[6] W.Y. He, Y.P. Zheng, L. Tang, S.C. Zheng, C. Béliveau, D. Doucet, M. Cusson, Q.L. Feng, Cloning, expression and localization of a trypsin-like serine protease in the spruce budworm, Choristoneura fumiferana, *Insect Sci.*, 16 (2009) 455-464.

[7] G. Broehan, Y. Arakane, R.W. Beeman, K.J. Kramer, S. Muthukrishnan, H. Merzendorfer, Chymotrypsin-like peptidases from Tribolium castaneum: a role in molting revealed by RNA interference, *Insect Biochem. Mol. Biol.*, 40 (2010) 274-283.

[8] A. Sumantha, C. Larroche, A. Pandey, Microbiology and industrial biotechnology of food-grade proteases: a perspective, *Food Technol. Biotechnol.*, 44 (2006) 211-220.

[9] J.F. Kennedy, V.W. Pike, Papain, chymotrypsin and related proteins—a comparative study of their beer chill-proofing abilities and characteristics, *Enzyme Microb. Technol.*, 3 (1981) 59-63.

[10] O.L. Tavano, A. Berenguer-Murcia, F. Secundo, R. Fernandez-Lafuente, Biotechnological applications of proteases in food technology, *Compr. Rev. Food Sci. Food Saf.*, 17 (2018) 412-436.

[11] A. Chandanwale, D. Langade, D. Sonawane, P. Gavai, A randomized, clinical trial to evaluate efficacy and tolerability of trypsin: chymotrypsin as compared to serratiopeptidase and trypsin: bromelain: rutoside in wound management, *Adv. Ther.*, 34 (2017) 180-198.

[12] D. Banerjee, S.K. Pal, Conformational dynamics at the active site of α-chymotrypsin and enzymatic activity, *Langmuir*, 24 (2008) 8163-8168.
[13] S. Farhadian, B. Shareghi, A.A. Saboury, Exploring the thermal stability and activity of α-chymotrypsin in the presence of spermine, *J. Biomol. Struct. Dyn.*, 35 (2017) 435-448.
[14] A. Karshikoff, L. Nilsson, R. Ladenstein, Rigidity versus flexibility: the dilemma of understanding protein thermal stability, *FEBS J.*, 282 (2015) 3899-3917.
[15] P. Biswas, A. Adhikari, U. Pal, P. Singh, M. Das, T. Saha-Dasgupta, S.S. Choudhury, R. Das, S.K. Pal, Flexibility modulates the catalytic activity of a thermostable enzyme: key information from optical spectroscopy and molecular dynamics simulation, *Soft matter*, 16 (2020) 3050-3062.
[16] P. Biswas, U. Pal, A. Adhikari, S. Mondal, R. Ghosh, D. Mukherjee, T. Saha-Dasgupta, S.S. Choudhury, R. Das, S.K. Pal, Essential Loop Dynamics Modulates Catalytic Activity in α-Chymotrypsin, *bioRxiv*, (2021) https://doi.org/10.1101/2021.08.11.455937.
[17] T. De Diego, P. Lozano, S. Gmouh, M. Vaultier, J.L. Iborra, Fluorescence and CD spectroscopic analysis of the α-chymotrypsin stabilization by the ionic liquid, 1-ethyl-3-methylimidazolium bis [(trifluoromethyl) sulfonyl] amide, *Biotechnol. Bioeng.*, 88 (2004) 916-924.
[18] L.M. Simon, M. Kotorman, G. Garab, I. Laczko, Structure and activity of α-chymotrypsin and trypsin in aqueous organic media, *Biochem. Biophys. Res. Commun.*, 280 (2001) 1367-1371.
[19] S.K. Pal, J. Peon, A.H. Zewail, Ultrafast surface hydration dynamics and expression of protein functionality: α-Chymotrypsin, *Proc. Natl. Acad. Sci. U.S.A*, 99 (2002) 15297-15302.
[20] P. Hazra, D. Chakrabarty, A. Chakraborty, N. Sarkar, Probing protein-surfactant interaction by steady state and time-resolved fluorescence spectroscopy, *Biochem. Biophys. Res. Commun.*, 314 (2004) 543-549.
[21] G. Wang, Y. Gao, M.L. Geng, Analysis of heterogeneous fluorescence decays in proteins. Using fluorescence lifetime of 8-anilino-1-naphthalenesulfonate to probe apomyoglobin unfolding at equilibrium, *Biochim. Biophys. Acta Gen. Subj.*, 1760 (2006) 1125-1137.
[22] C.C. David, D.J. Jacobs, Principal component analysis: a method for determining the essential dynamics of proteins, Protein dynamics, *Springer, Berlin/Heidelberg, Germany*, (2014) 193-226.
[23] G.G. Maisuradze, A. Liwo, H.A. Scheraga, Principal component analysis for protein folding dynamics, *J. Mol. Biol.*, 385 (2009) 312-329.

[24] J. Mongan, Interactive essential dynamics, *J. Comput. Aided Mol. Des.*, 18 (2004) 433-436.

Chapter 6

Studies on the Effect of Chemical Modification on the Structure, Dynamics, and Function of a Model Enzyme

6.1. Introduction

Covalent modifications introduce addition or removal of groups from the amino acid residues of a protein altering its structure and function. Formalin is known to react with a plethora of functional groups in small peptides and other substrates to generate covalently modified products [1, 2] which can include both intramolecular and intermolecular cross-linked species. From a chemical standpoint, formalin can react with biological nucleophiles such as protein and DNA and facilitates *in vitro* formation of intra-strand and DNA-protein cross links [2-4]. Formalin related reactions in cells are likely responsible for its toxic/carcinogenic effects, however the precise chemistry of such reactions and their cellular prevalence is still not clearly defined [5]. Pioneering works by McGhee and Von Hippel [6], and a recent study by Kamps et al., [7] on the reactions of nucleosides/nucleotides and amino acids with formalin have envisaged to understand the chemistry of such reactions.

To have a better understanding how formalin influence health and diseases, it is important to investigate the correlation between formalin induced structural modification and function of biological macromolecules. Previous studies on interaction of formalin with biological macromolecules were restricted to model peptides and other proteins [4, 8-10]. These studies, however, revealed very little information on residue modification, and its correlation to protein function and its molecular recognition.

The present study examines and correlates residue modification of a common digestive enzyme α-chymotrypsin (CHT; EC 3.4.21.1) by formalin with its catalytic activity and molecular recognition of a ligand (ANS) using steady state and pico-second resolved fluorescence spectroscopy, circular dichroism (CD) in conjunction with molecular docking studies.

6.2. Results and Discussion:

6.2.1. A Combined Spectroscopic and Molecular Modeling Study on Structure-Function-Dynamics under Chemical Modification: α-Chymotrypsin with Formalin Preservative [11]:

α-chymotrypsin (CHT) is a proteolytic enzyme of the class serine protease (EC 3.4.21.1) which is associated with hydrolysis of peptide bonds in the mammalian digestive system. The catalytic triad of CHT is located in the hydrophobic S1 pocket of CHT [12], molecular docking study of CHT-ANS indicates the presence of two different sites-one located at the S1 pocket (buried internal site) while the other at the Cys1-Cys122 disulfide bond (exposed external site). Figure 6.1 shows the 3D structure of CHT. Figure 6.1.a highlights the catalytic triad comprising the residues His57; Asp102 and Ser195. The site at Cys1-Cys122 disulfide bond (Figure 6.1.b) serves as one of the ANS binding sites and is located almost opposite to the catalytic site.

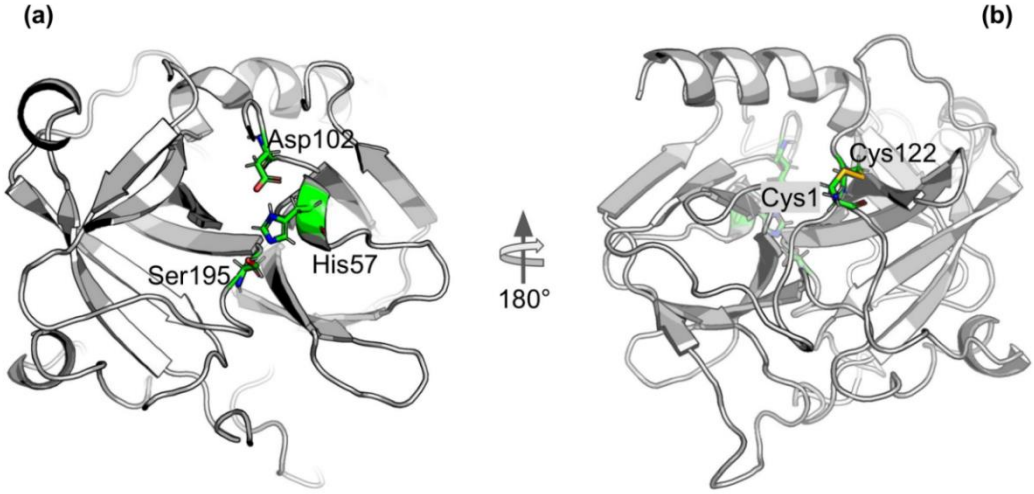

Figure 6.1. *Cartoon representation of alpha-Chymotrypsin (CHT) [PDB ID:1CGJ]. (a) The coloured stick representation of the active site catalytic triad; showing the active site residues [His57; Asp102; Ser195]. (b) At the N-terminal, the coloured stick representation [Cys1; Cys122] shows the molecular recognition site [External ANS binding site]. This site is located almost opposite to the catalytic site.*

The catalytic activity of CHT was measured with different AMC (substrate) concentrations (5-40 μM) at varying formalin concentrations (1% to 4%). CHT belongs to the class of serine protease that catalyzes hydrolysis of AMC to 7-amido-4-methy-coumarin (Figure 6.2.a &

6.2.b). Figure 6.2.c & 6.2.d illustrate the effects of increasing formalin concentration on the product formation and the rate of product formation.

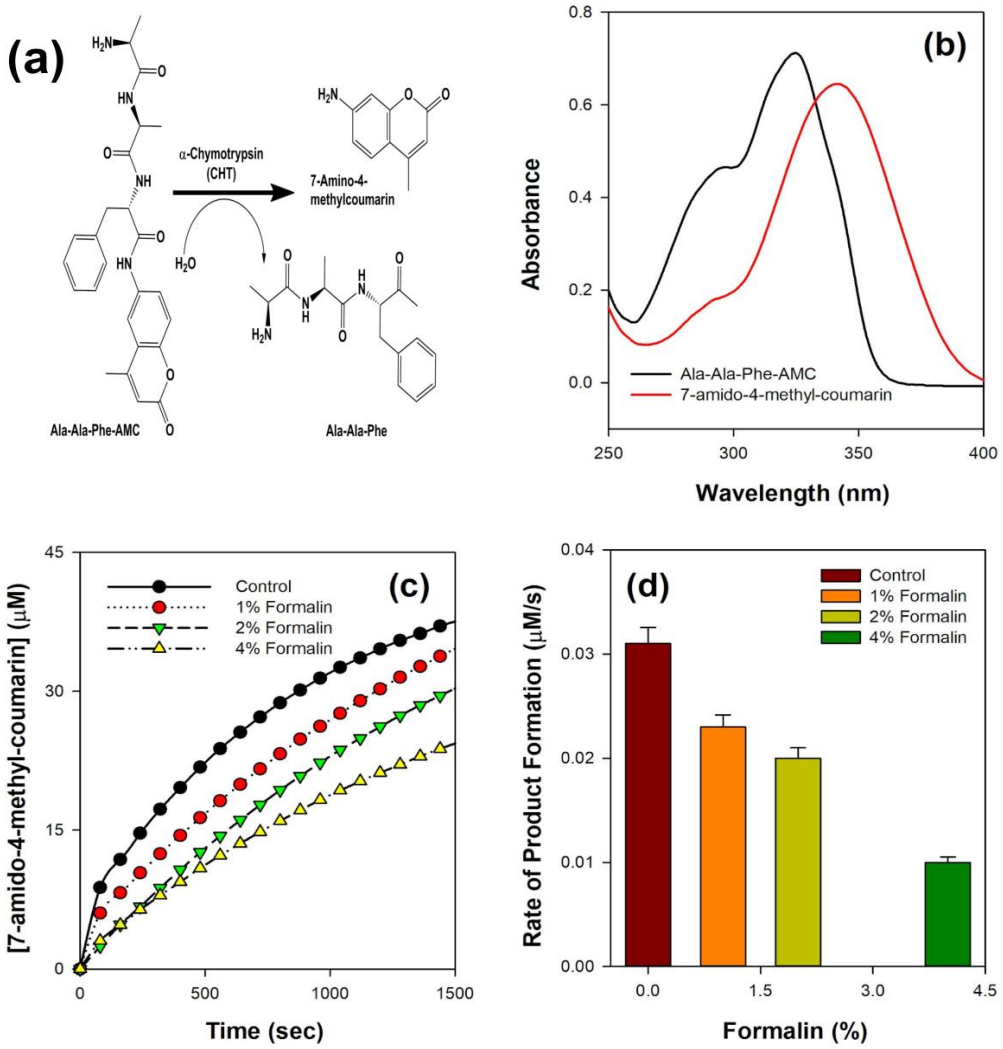

Figure 6.2. *Schematic representation of the formalin treated enzyme catalysis. (a) Simple reaction outline catalysed by CHT; the substrate AMC hydrolysed to produce 7-amido-4-methyl-coumarin. (b) Absorbance spectra of the substrate AMC and product 7-amido-4-methyl-coumarin during catalysis. (c) Product formation at different representative formalin concentration; diminishing product formation with increased formalin concentration. (d) Rate of product formation at different representative formalin concentration; retarding rate of product formation consistent with decreasing product formation with increasing formalin concentration.*

A reduction in the rate of CHT catalyzed product formation was observed (Table 6.1) with increasing formalin concentration reflecting on reduced catalytic activity which likely originates from a possible modification of the substrate binding site.

Close analysis of the 3D structure of CHT (Figure 6.1) shows that the enzyme consists of anti-parallel β-sheets which are highly distorted, forming very short irregular strands.

Table 6.1. Rate of product formation of α-Chymotrypsin at different formalin concentration

Formalin Concentration (%)	Rate of product formation (µM/s)
Control	$(0.031\pm1.55) \times 10^{-3}$
1	$(0.023\pm1.15) \times 10^{-3}$
2	$(0.020\pm1.0) \times 10^{-3}$
4	$(0.010\pm5.0) \times 10^{-4}$

Such conformation can lead to a shift in the negative band from the ideal β-sheet position (210-220 nm) towards 200 nm. The effect of formalin induced modification on CHT was quantified by circular dichroism (CD) spectroscopy in the far UV region.

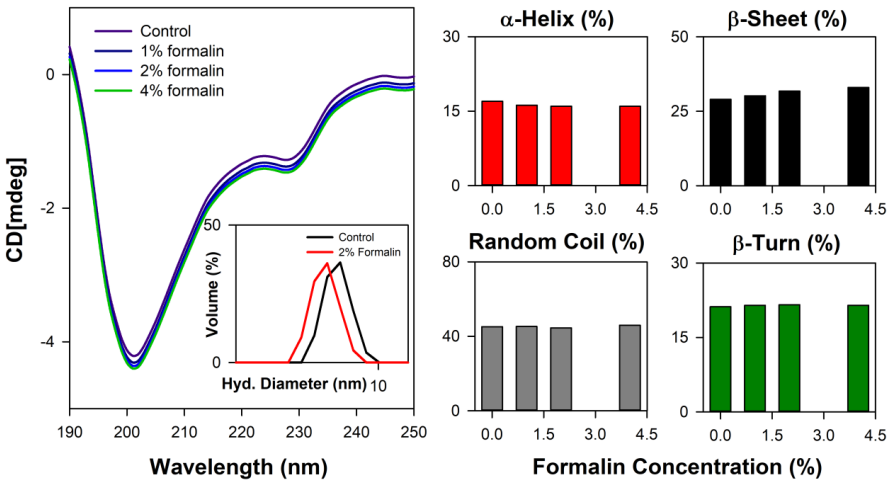

Figure 6.3. *Far UV CD spectroscopy and DLS study of CHT. (a) The far UV CD spectra representing structural integrity of formalin treated CHT. The inset shows the DLS spectra of CHT at representative formalin concentration. (b) Percentage composition of secondary structure shown as bar plots. Analysis of secondary structure indicated structural stability within the experimental setup.*

The far UV spectra of CHT were characterized by a minimum at 202 nm with no positive band (Figure 6.3). The spectral data for all formalin concentration (1%, 2% and 4%) was recorded.

Upon de-convolution, a 4% increase in anti-parallel β-sheet content was observed along with a nominal decrease in α-helix and insignificant changes in parallel β-sheet, turns and random coils (Table 6.2).

Table 6.2. Secondary structure analysis of α-Chymotrypsin in presence of formalin

Formalin conc (%)	% α-helix	% β-sheet	% random coil	% β-turn
Control	17.00	29.00	45.20	21.20
1	16.20	30.20	45.30	21.50
2	16.00	31.80	44.60	21.60
4	16.00	33.00	46.00	21.50

These results indicate an insignificant change in the secondary structure of the enzyme reflecting on its structural stability.

The effect of formalin on the globular tertiary structure of CHT was monitored by dynamic light scattering (DLS). The DLS study (Figure 6.3, inset) reveals the hydrodynamic diameter of CHT to be ~7 nm, which agrees with previous studies [13, 14]. Upon treatment with 4% formalin, a decrease in the hydrodynamic diameter of CHT from ~ 7 nm (control) to ~ 5 nm was noted (Table 6.3) indicating increased compactness and rigidity rather than structural unfolding of the enzyme over the range of formalin concentration used.

Table 6.3. Globular Tertiary Structure of α-Chymotrypsin upon Formalin Treatment

Formalin conc. (%)	Hydrodynamic Diameter (nm)
Control	~ 7.0
1	~6.0
2	~ 6.0
4	~ 5.0

In water fluorescence of ANS decays rapidly with a time constant of 0.25 ns [15]. The fluorescence transients of ANS in CHT at different formalin concentrations (Figure 6.4) are characterized by three-time constants (Table 6.4) indicating its binding to the enzyme. The faster component (τ_1) of few hundreds of picoseconds (~ 0.25 ns) originates due to free ANS in water [15]. ANS bound to CHT reveals two different lifetimes (~ 2.1 ns and ~ 7.1 ns) which

may be ascribed to two different binding sites in the enzyme; the shorter lifetime (τ_2) corresponds to an external binding site partially exposed to water, whereas, the longer one (τ_3)

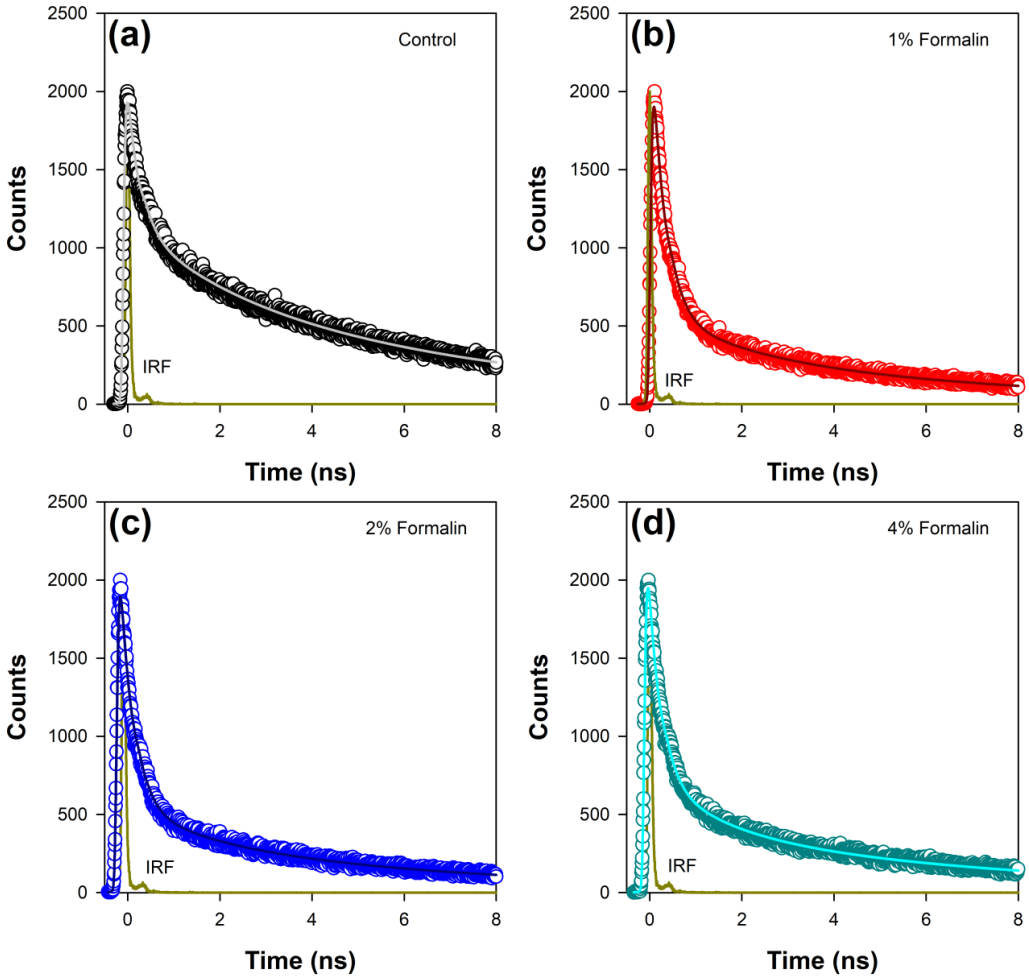

Figure 6.4. *Time resolved fluorescence transients of ANS in (a) pure CHT (control) and at varying formalin concentrations (b-d).*

may be ascribed to an internal binding site located inside the enzyme which is screened from water. Upon increasing formalin content, the longer time constant remains unchanged whereas the shorter fluorescence lifetime (τ_2) decreases to 1.4 ns (Table 6.4). This may be ascribed to structural modification in the external binding site leading to increased exposure of the probe (ANS) to water resulting in a shortening of the shorter fluorescence lifetime (τ_2).

Table 6.4: Fluorescence lifetimes of ANS in α-Chymotrypsin in presence of formalin

Formalin conc. (%)	τ_1 (ps)	τ_2 (ps)	τ_3 (ps)	τ_{avg} (ps)
Control	0.235 (55)	2.11 (14)	7.05 (31)	2.71
1	0.259 (78)	1.94 (10)	7.21 (12)	1.29
2	0.254 (77)	1.76 (10)	7.19 (13)	1.37
4	0.240 (72)	1.41 (13)	7.22 (15)	1.49

Polarization-gated fluorescence anisotropy of ANS in CHT was measured (Figure 6.5) at different formalin concentrations to monitor changes around the ANS binding sites of the enzyme. The anisotropy decay in CHT (Figure 6.5; control) is characterized by three rotational time constants of ~ 56 ps (θ_1), ~ 712 ps (θ_2) and 30 ns (θ_3), respectively, (Table 6.5). The faster

Table. 6.5. Time-resolved Fluorescence anisotropy parameters of ANS-CHT in presence of formalin (The longest rotational correlation time is kept constant at 30 ns)

System	θ_1/ns	θ_2/ns	θ_3/ns
ANS in CHT	0.056	0.712	30
ANS in CHT+1 % formalin	0.098	0.584	30
ANS in CHT+2 % formalin	0.094	0.425	30
ANS in CHT+4 % formalin	0.092	0.244	30

The anisotropy decay of ANS in CHT is fitted by the following equation

$$r(t) = r_0\left(\beta_1 \exp\left(-\frac{t}{\theta_1}\right) + \beta_2 \exp\left(-\frac{t}{\theta_2}\right) + \beta_3 \exp\left(-\frac{t}{\theta_3}\right)\right)$$

where r_0 is initial anisotropy, β_is and θ_is are the amplitudes and corresponding rotational time constants, respectively. The longest rotational correlation time is kept constant at 30 ns. In presence of formalin, the anisotropy decays are fitted with equation (6.1) as described in the text.

(θ_1) and the relatively slower correlation time (θ_2) is attributed to orientational motion of the dye in water [16] and that (ANS) bound to an external binding site in CHT, respectively. On the other hand, the much longer time constant (θ_3) originates from rotational motion of the ANS/CHT complex where the probe is strongly bound to an internal binding site of the enzyme so that it rotates with the enzyme as a protein/dye complex [17]. Upon interaction with formalin the fluorescence anisotropy decay of ANS in CHT is remarkably modified and a dip-rise pattern becomes evident (Figure 6.5) owing to heterogeneous distribution of ANS in different

microenvironments with different fluorescence lifetimes and associated rotational correlation times [18-20]. When a fluorophore exists in different microenvironments, its anisotropy decay, r(t) can be expressed as: [21, 22]

$$r(t) = \sum_{m} f_m(t) r_m(t) \tag{6.1}$$

where the fractional intensity of the m^{th} component at any time t is given by

$$f_m(t) = \frac{\alpha_m \exp\left(-\frac{t}{\tau_m}\right)}{\sum_m \alpha_m \exp\left(-\frac{t}{\tau_m}\right)} \tag{6.2}$$

α_m is the amplitude corresponding to the fluorescence lifetime τ_m at t = 0 and

$$r_m(t) = r_{0m} \exp\left(-\frac{t}{\theta_m}\right) \tag{6.3}$$

where θ_m is the single rotational correlation time of the m^{th} component and r_{0m} is corresponding initial anisotropy.

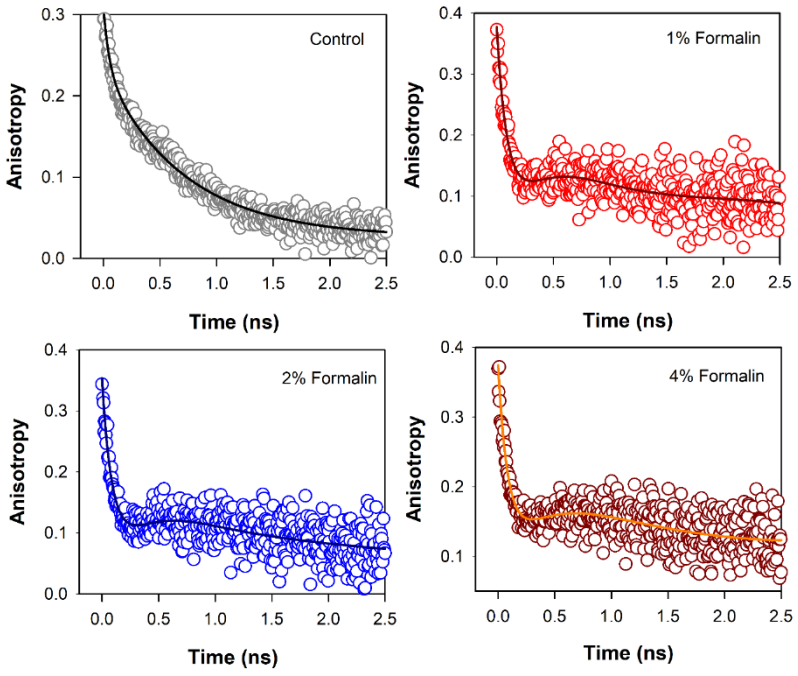

Figure 6.5. *Anisotropy decay (r(t)) of ANS in CHT in absence and presence of formalin.*

The analysis of the dip and rise anisotropy decay (Figure 6.5) using equation (6.1) shows presence of three rotational correlation times (Table 6.5) where the fastest (θ_1) and the longest (θ_3) rotational correlation times remain almost unchanged with a variation in the formalin content. However, a gradual decrease in the intermediate correlation time (θ_2) from 712 ps to 244 ps with increasing formalin content is observed which may be attributed to covalent modification of the amino acid residues in the enzyme by formalin. As a result of such residue modification the external binding site of ANS in the enzyme likely becomes more exposed to water giving rise to faster rotational relaxation of the dye (ANS).

Figure 6.6 shows solvent accessible surface area of CHT (1CGJ). It becomes evident that several amino acid residues of CHT are exposed to solvent and hence, are vulnerable to formalin induced modification. However, we have restricted our study to the substrate and the probe (ANS) binding sites in the enzyme. At the catalytic S1 pocket of CHT which also serves

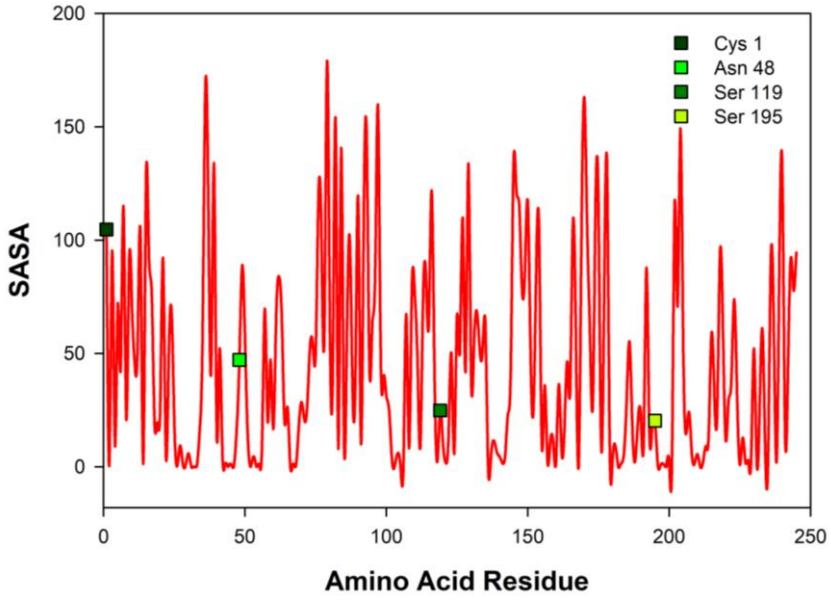

Figure 6.6. *Solvent accessible surface area (SASA) of CHT [PDB ID: 1CGJ]. Four different residues corresponding to the catalytic site and molecular recognition were selected [Cys1; Asn48; Ser119; Ser195] for formalin induced modification.*

as the buried internal site of ANS binding, Ser195 is solvent exposed and contains a free hydroxyl group making it susceptible to formalin induced modification [7]. Furthermore, study of the external ANS binding site indicates Cys1 to be vulnerable to modification by formalin owing to its high extent of solvent exposure and the presence of a free amino ($-NH_2$) group. In

contrary, Cys122 has no free amino (NH_2) group making it almost immune to formalin induced modification. This study also reveals the presence of susceptible residues Asn48 and Ser119 within 3Å of the ANS binding site, making them suitable candidates for our study.

To simulate and study the influence of minimal local residue modification on the function of CHT and its ability to recognize molecular binding partner, the residues selected from the SASA study (Ser195; Cys1; Asn48; Ser119) were subjected to residue modification according to Kamps et al. [7]. Figure 6.7 shows possible modifications due to reaction with

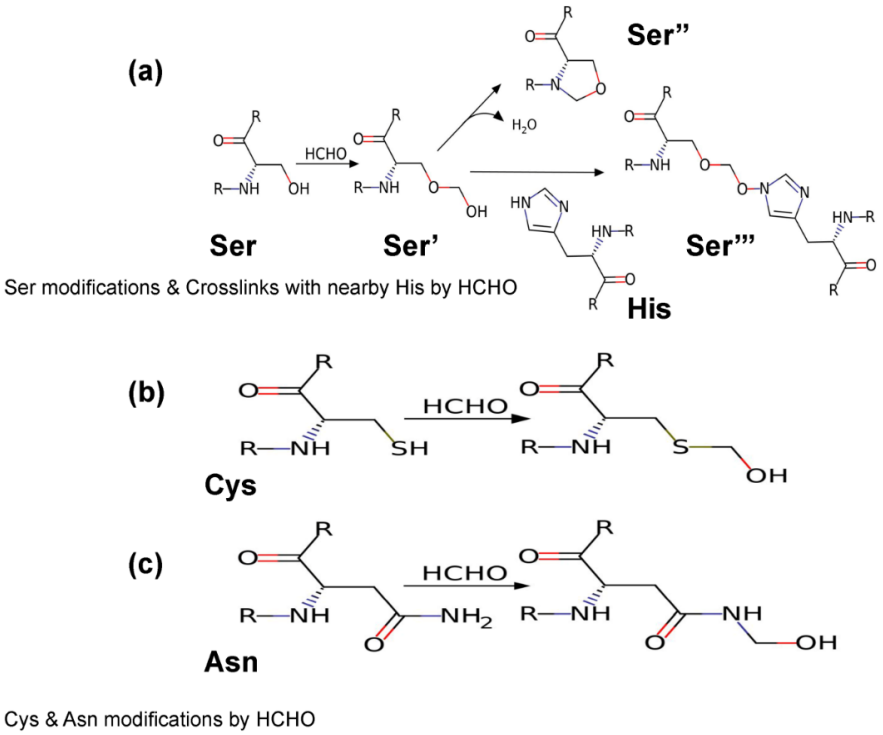

Figure 6.7. *Reaction products of HCHO with different amino acids. (a) Amino acid Serine (Ser)reacting with HCHO form methylol derivative [Ser’]; which undergo cyclisation to form oxazolidine adduct [Ser’’]; the methylol derivative can react with nearby amino acid [case in point amino acid Histidine (His)] to form a crosslink derivative [Ser’’’]. (b& c) Reaction of HCHO with amino acids Cystine (Cys) and Asparagine (Asn) forming their respective methylol derivative.*

formalin. From figure 6.7.a it becomes evident that serine has three possible modified outcomes. The free hydroxyl (OH) group undergoes modification forming a methylol adduct (Ser’), which can undergo internal cyclization, forming oxazolidone adduct (Ser’’) or forming a crosslink with the nearby His57 residue (Ser’’’). Similarly, Cys1 and Asn48, undergo modifications forming a methylol adduct at the free thiol (-SH) or amino (-NH_2) group (Figure

6.7.b & 6.7.c). However, due to the formation of a disulfide bond between Cys1 and Cys122, no thiol group is available for formalin induced modification. Cys1 being the N-terminal amino acid, contains a free amino group (-NH_2) which reacts with the formalin.

Molecular docking was performed on the modified CHT structure. Figure 6.8-6.11 illustrate different ligand bound structures. It becomes evident from the docking studies that residue modification at the catalytic site has no influence on binding of CHT to the substrate (AMC) (Figure 6.8).

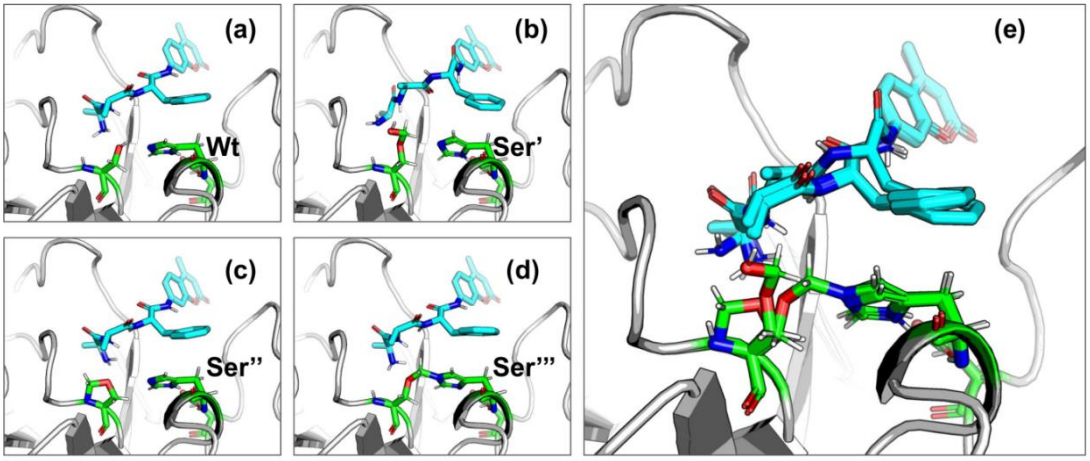

Figure 6.8. *Binding of substrate (AMC) withWt and HCHO modified CHT as obtained by molecular docking simulations. (a-d) Different modes of AMC binding with CHT on modification of the active site Serine residue (Ser195) as mentioned previously [Wt indicates the wild type Ser195 residue]. (e) Overlapping view of the different binding interactions; residue modification causes no major change in binding mode of AMC with CHT.*

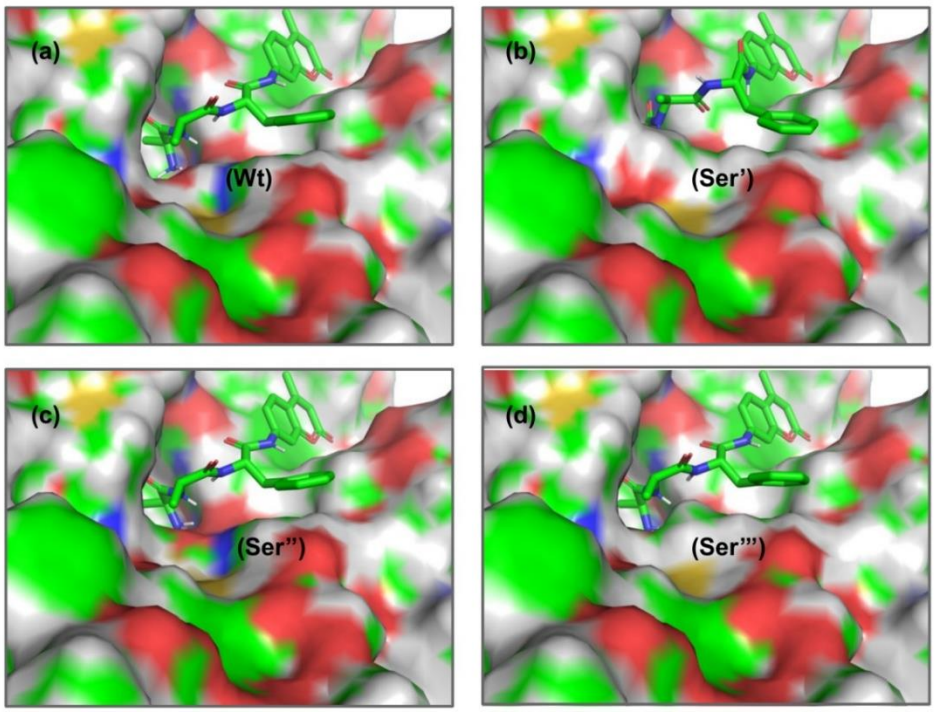

Figure 6.9. *Surface view of the binding cavity at Wt and HCHO modified CHT. (a-d) Insignificant changes to the binding cavity upon residue modification at Ser195.*

Furthermore, a study of the van der Waals surface indicated no major structural distortions of the binding pocket (Figure 6.9). The possible formation of Ser195-His57 crosslink upon formalin treatment may cause a disruption of the serine protease activity at the catalytic triad; thus, impairing the product formation and release, leading to reduced catalytic activity. Molecular docking between ANS and CHT indicates presence of two binding sites, namely, the buried internal site at the S1 pocket and an external binding site at Cys1-Cys122 di-sulfide bond. It becomes evident from (Figure 6.10) that, formalin induced residue modification at the external binding site of ANS may lead to the generation of multiple probe (ANS) binding sites and a heterogeneous probe binding distribution which is indeed reflected by the dip-rise pattern in the anisotropy decay (Figure 6.5). Docking simulations of CHT modified at the external binding site Cys1 and its nearby susceptible residues Ser119, Asn48 generated binding models (Figure 6.10.b-h) which deviate from the unmodified CHT (Figure 6.10.a).

Although molecular docking does not provide pronounced solvent effects in protein/ligand interactions, however, by probing the solvent accessibility of the bound ligand (ANS), solvent effects in protein-ligand interactions can be elucidated from molecular docking

studies. SASA analysis of the bound probe found a diverse range of solvent effects upon formalin modification of the external site. The varied solvent effects ranged from mimicking

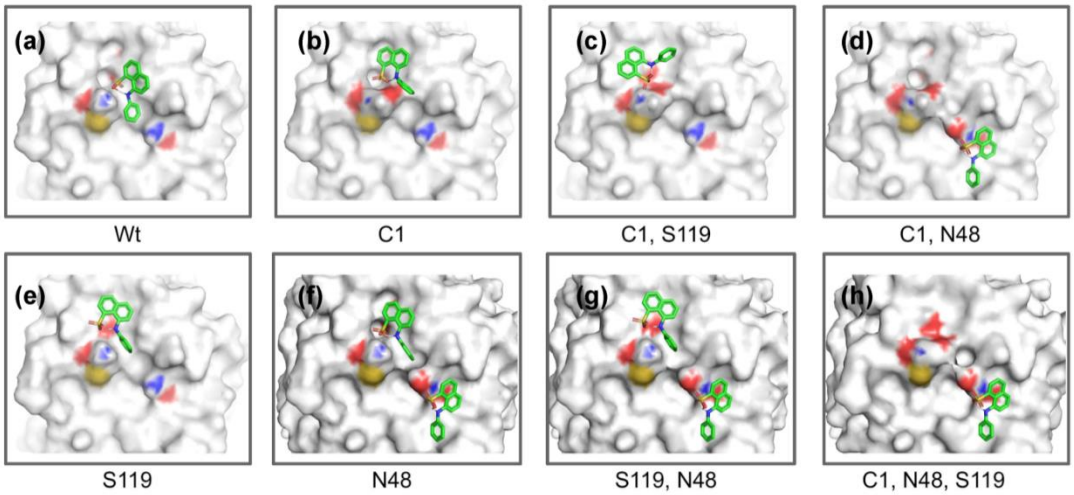

Figure 6.10. *Different binding modes of ANS with CHT (external site) as obtained by molecular docking simulation. (a-h) Best binding modes of ANS with on residue modification [Cys1; Asn48; Ser119] as mentioned previously; extensive change in binding modes resulting in heterogeneity of ANS binding is evident.*

the unmodified ANS binding site [modification at Cys1;1244Å^2 & Ser119;1236Å^2] and lesser degree of solvent accessibility [modifications at Asn48;1150Å^2] to experiencing higher solvent effects [modifications in conjunction at Cys1 Asn48;(1374Å^2), Cys1 Ser119;(1402Å^2), Ser119 Asn48;(1369Å^2), Cys1 Asn48 Ser119;(1360Å^2)] (Table 6.6.a) signifying the bound probe to be more exposed and susceptible to solvent effects.

Table 6.6.a. SASA of ANS upon binding to formalin modified CHT (External Site)

Modification	SASA (Å^2)	Binding Site
Unmodified (Cys1)	1214	Near Cys1
Cys1	1244.40	Toward Ser119
Ser119	1236.36	Toward Ser119
Asn48	1150/1369	Over Cys1 & Asn48
Cys1 Ser119	1402.37	Toward Ser119
Cys1 Asn48	1374.44	Over Asn48
Ser119 Asn48	1212.76/1369	Toward Ser119 & Over Asn48
Cys1 Ser119 Asn48	1360.71	Over Asn48

Similar to the binding interactions of substrate (AMC), residue modification at the internal ANS binding site in CHT (Ser195) has no effect on binding of the probe (Figure 6.11). SASA analysis of ANS at the internal binding site of CHT suggests no major changes in solvent exposure of ANS (Table 6.6.b) suggesting that the bound probe is shielded from the surrounding solvent even after residue modification.

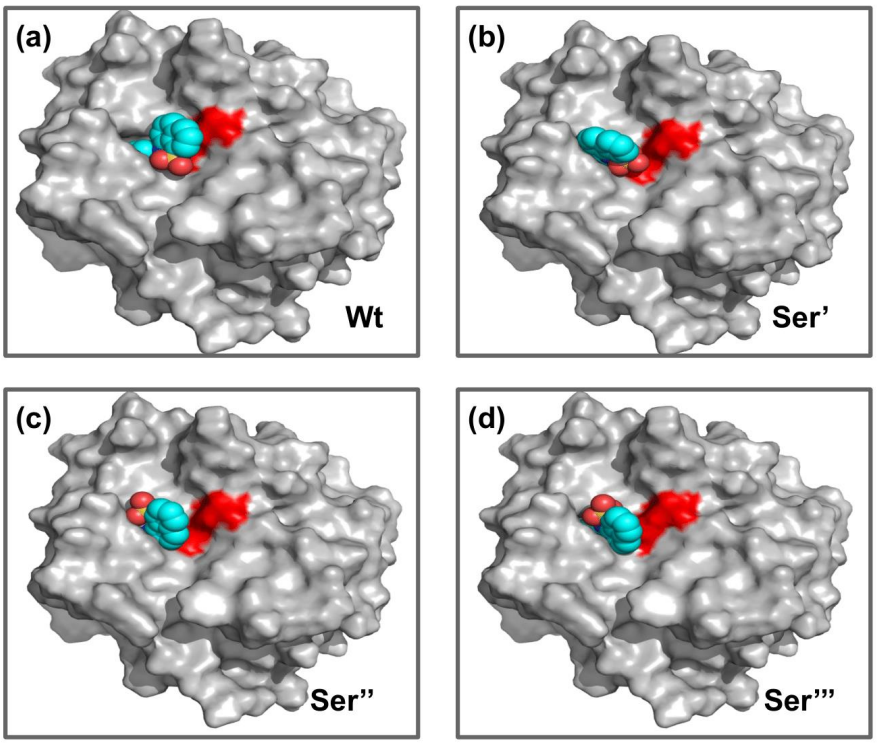

Figure 6.11. *Different binding modes of ANS with CHT (internal site) as obtained by molecular docking simulation. (a-d) ANS binding modes with CHT on active site serine (Ser195) modification as previously mentioned. [Wt indicates the wild type/unmodified CHT].*

Table 6.6.b. SASA of ANS upon binding to formalin modified CHT (Internal Site)

Modifications	SASA (Å^2)
Unmodified Ser195	1027
Ser'	1061
Ser''	1024
Ser'''	1013

CHT causes the catalytic hydrolysis of AMC to 7-amido 4-methyl-coumarin [23]. The effect of formalin induced modification was monitored at different concentrations of formalin

(1-4%). We have observed a gradual decrease in the rate of product formation with an increase in formalin concentration. However, stability of both the secondary and the tertiary structure indicated to local residue modification at the catalytic S1 pocket as the possible architect in reduction of catalytic activity of the enzyme in presence of formalin. Previous studies by Martin et al., [24] and Zhang et al., [25] found direct correlation between modification of the catalytic site residues and activity. Ser195 at the catalytic triad was solvent exposed. Ser195 modification and molecular docking **s**howed no change in substrate (AMC) binding at the catalytic triad. Furthermore, no major structural distortions in the binding pocket were observed. The analysis implied a stable enzyme-substrate formation; however, disruption in the proton transfer cascade from Ser195 to His57 to Asp102 due to serine-histidine cross-linking by formalin impairs product formation and release leading to the reduced catalytic activity. Molecular docking simulations further support that modifications in the local amino acid residues are solely responsible for the impaired catalysis.

Time resolved fluorescence studies of ANS in CHT manifest heterogeneity in binding of ANS upon formalin induced modification. As a result of formalin induced structural modification in the enzyme, ANS bound to an internal binding site becomes increasingly exposed to water thereby leading to a decreased fluorescence lifetime and rotational relaxation time.

Our SASA calculations identified four amino acid residues [Ser195; Cys1; Asn48 & Ser119] as one of the possible centers of local residue modification. Molecular docking analysis of external ANS binding site indicated generation of new ANS binding sites having different solvent accessibility on formalin modification. However, no change in binding was observed for the internal binding site of ANS in CHT. The SASA of bound ANS suggested the bound probe to be buried and shielded from the surrounding solvent. Our solvent effect and docking simulation studies of the bound probe was found to corroborate with the time resolved fluorescence anisotropy studies manifesting heterogeneity in ANS binding to CHT following formalin induced modification.

6.3. Conclusion:

The present work primarily focuses on the influence of residue modification on the catalytic activity and molecular recognition of CHT. The study found remarkable changes in the catalytic activity and molecular recognition of the enzyme (CHT) upon formalin induced chemical modification of its amino acid residues. The structural and molecular docking

analysis provide evidence of no impediment to the formation of a stable enzyme-substrate complex due to formalin induced modifications; rather impairment of the proton transfer cascade in the catalytic triad due to Ser195-His57 cross-linking leads to reduced catalytic activity. Picosecond resolved fluorescence studies reveal generation of multiple binding sites of a fluorophore ANS in CHT, and upon treatment with formalin a reduction in fluorescence lifetime and rotational relaxation time of the probe is observed when it is bound to an external binding site. Subsequent molecular docking studies and SASA analysis provide evidence towards heterogeneity of ligand (ANS) interaction with the enzyme (CHT) upon minimal change in the local amino acid residues. The findings of the present study offer a background towards fine tuning enzymatic capabilities, drug molecule recognition and protein engineering.

References

[1] H. Fraenkel-Conrat, H.S. Olcott, Reaction of formaldehyde with proteins VI. cross-linking of amino groups with phenol, imidazole, or indole groups, *J. Biol. Chem.*, 174 (1948) 827-843.

[2] E.A. Hoffman, B.L. Frey, L.M. Smith, D.T. Auble, Formaldehyde crosslinking: a tool for the study of chromatin complexes, *J. Biol. Chem.*, 290 (2015) 26404-26411.

[3] D. Brutlag, C. Schlehuber, J. Bonner, Properties of formaldehyde-treated nucleohistone, *Biochemistry*, 8 (1969) 3214-3218.

[4] B. Metz, G.F.A. Kersten, P. Hoogerhout, H.F. Brugghe, H.A.M. Timmermans, A. De Jong, H. Meiring, J. ten Hove, W.E. Hennink, D.J.A. Crommelin, W. Jiskoot, Identification of formaldehyde-induced modifications in proteins: reactions with model peptides, *J. Biol. Chem.*, 279 (2004) 6235-6243.

[5] K. Lu, W. Ye, L. Zhou, L.B. Collins, X. Chen, A. Gold, L.M. Ball, J.A. Swenberg, Structural characterization of formaldehyde-induced cross-links between amino acids and deoxynucleosides and their oligomers, *J. Am. Chem. Soc.*, 132 (2010) 3388-3399.

[6] J.D. McGhee, P.H. Von Hippel, Formaldehyde as a probe of DNA structure. 3. Equilibrium denaturation of DNA and synthetic polynucleotides, *Biochemistry*, 16 (1977) 3267-3276.

[7] J.J.A.G. Kamps, R.J. Hopkinson, C.J. Schofield, T.D.W. Claridge, How formaldehyde reacts with amino acids, *Commun. Chem.*, 2 (2019) 1-14.

[8] B. Metz, G.F.A. Kersten, G.J.E. Baart, A. de Jong, H. Meiring, J. ten Hove, M.J. van Steenbergen, W.E. Hennink, D.J.A. Crommelin, W. Jiskoot, Identification of formaldehyde-induced modifications in proteins: reactions with insulin, *Bioconjug. Chem.*, 17 (2006) 815-822.

[9] B. Metz, T. Michiels, J. Uittenbogaard, M. Danial, W. Tilstra, H.D. Meiring, W.E. Hennink, D.J.A. Crommelin, G.F.A. Kersten, W. Jiskoot, Identification of formaldehyde-induced modifications in diphtheria toxin, *J. Pharm. Sci.*, 109 (2020) 543-557.

[10] K. Lu, G. Boysen, L. Gao, L.B. Collins, J.A. Swenberg, Formaldehyde-induced histone modifications in vitro, *Chem. Res. Toxicol.*, 21 (2008) 1586-1593.

[11] P. Biswas, A. Adhikari, U. Pal, S. Mondal, D. Mukherjee, R. Ghosh, T. Saha-Dasgupta, S.S. Chowdhury, R. Das, S.K. Pal, A Combined Spectroscopic and Molecular Modeling Study on Structure-Function-Dynamics under Chemical Modification: Alpha-Chymotrypsin with Formalin Preservative, *bioRxiv*, (2021) https://doi.org/10.1101/2021.07.24.453635.

[12] W. Ma, C. Tang, L. Lai, Specificity of trypsin and chymotrypsin: loop-motion-controlled dynamic correlation as a determinant, *Biophys. J.*, 89 (2005) 1183-1193.

[13] D. Banerjee, S.K. Pal, Conformational dynamics at the active site of α-chymotrypsin and enzymatic activity, *Langmuir*, 24 (2008) 8163-8168.
[14] A.K. Shaw, R. Sarkar, D. Banerjee, S. Hintschich, A. Monkman, S.K. Pal, Direct observation of protein residue solvation dynamics, *J. Photochem. Photobiol. A*, 185 (2007) 76-85.
[15] M. Collini, L. D'Alfonso, H. Molinari, L. Ragona, M. Catalano, G. Baldini, Competitive binding of fatty acids and the fluorescent probe 1-8-anilinonaphthalene sulfonate to bovine β-lactoglobulin, *Protein Sci.*, 12 (2003) 1596-1603.
[16] S.K. Pal, J. Peon, A.H. Zewail, Ultrafast surface hydration dynamics and expression of protein functionality: α-Chymotrypsin, *Proc. Natl. Acad. Sci. U.S.A*, 99 (2002) 15297-15302.
[17] J.D. Johnson, M.A. El-Bayoumi, L.D. Weber, A. Tulinsky, Interaction of. alpha.-chymotrypsin with the fluorescent probe 1-anilinonaphthalene-8-sulfonate in solution, *Biochemistry*, 18 (1979) 1292-1296.
[18] J.R. Lakowicz, Principles of fluorescence spectroscopy, *Springer, Berlin/Heidelberg, Germany*, (2013).
[19] B.K. Paul, N. Guchhait, Modulated photophysics of an ESIPT probe 1-hydroxy-2-naphthaldehyde within motionally restricted environments of liposome membranes having varying surface charges, *J. Phys. Chem. B*, 114 (2010) 12528-12540.
[20] S.S. Sinha, R.K. Mitra, S.K. Pal, Temperature-dependent simultaneous ligand binding in human serum albumin, *J. Phys. Chem. B*, 112 (2008) 4884-4891.
[21] J. Broos, A.J.W.G. Visser, J.F.J. Engbersen, W. Verboom, A. van Hoek, D.N. Reinhoudt, Flexibility of enzymes suspended in organic solvents probed by time-resolved fluorescence anisotropy. Evidence that enzyme activity and enantioselectivity are directly related to enzyme flexibility, *J. Am. Chem. Soc.*, 117 (1995) 12657-12663.
[22] B.K. Paul, N. Ghosh, S. Mukherjee, Interplay of multiple interaction forces: binding of norfloxacin to human serum albumin, *J. Phys. Chem. B*, 119 (2015) 13093-13102.
[23] R. Biswas, S.K. Pal, Caging enzyme function: α-chymotrypsin in reverse micelle, *Chem. Phys. Lett.*, 387 (2004) 221-226.
[24] C.J. Martin, N.B. Oza, M.A. Marini, Formaldehyde as an Active Site Label of α-Chymotrypsin, *Can. J. Biochem.*, 50 (1972) 1114-1121.
[25] W.-N. Zhang, D.-P. Bai, X.-Y. Lin, Q.-X. Chen, X.-H. Huang, Y.-F. Huang, Inactivation kinetics of formaldehyde on N-acetyl-β-d-glucosaminidase from Nile tilapia (Oreochromis niloticus), *Fish Physiol. Biochem.*, 40 (2014) 561-569.

LIST OF PUBLICATIONS

International Peer-Reviewed Journals

[1] **Biswas, P.**, Adhikari, A., Pal, U., Singh, P., Das, M., Saha-Dasgupta, T., Choudhury, S. S., Das, R., Pal, S. K. (2020). Flexibility Modulates Catalytic Activity of a Thermostable Enzyme: Key Information from Optical Spectroscopy and Molecular Dynamics Simulation. *Soft Matter*, 16(12), 3050-302.

https://doi.org/10.1039/C9SM02479D

[2] **Biswas, P.**, Pal, U., Adhikari, A., Mondal, S., Ghosh, R., Mukherjee, D., Saha Dasgupta, T., Choudhury, S. S., Das, R. & Pal, S. K. (2021). Essential Loop Dynamics Modulates Catalytic Activity in α-Chymotrypsin. *bioRxiv*. https://doi.org/10.1101/2021.08.11.455937

[3] **Biswas, P.**, Adhikari, A., Pal, U., Mondal, S., Mukherjee, D., Ghosh, R., Saha-Dasgupta, T., Chowdhury, S. S., Das, R. & Pal, S. K. (2021). A Combined Spectroscopic and Molecular Modeling Study on Structure-Function-Dynamics under Chemical Modification: Alpha-Chymotrypsin with Formalin Preservative. *bioRxiv*. https://doi.org/10.1101/2021.07.24.453635

[4] * Adhikari, A., **Biswas, P.**, Mondal, S., Das, M., Darbar, S., Hameed, A. M., Alharbi, A., Ahmed, S. A., Bhattacharya, S. S., Pal, D., Pal, S. K. (2020). A Smart Nanotherapeutic Agent for in vitro and in vivo Reversal of Heavy Metal Induced Causality: Key Information from Optical Spectroscopy. *ChemMedChem*, 15(5), 420-429. [Front Cover Highlighted]
https://doi.org/10.1002/cmdc.201900543

[5] * Adhikari, A., Mondal, S., Das, M., **Biswas, P.**, Pal, U., Darbar, S., Bhattacharya, S. S., Pal, D., Saha-Dasgupta, T., Das, A. K., Mallick, A. K., Pal, S. K. (2021). Incorporation of a Biocompatible Nanozyme in Cellular Antioxidant Enzyme Cascade Reverses Huntington's Like Disorder in Preclinical Model. *Advanced Healthcare Materials*, 10(7), 2001736.

https://doi.org/10.1002/adhm.202001736

[6] * Adhikari, A., Mondal, S., Das, M., Ghosh, R., **Biswas, P.**, Darbar, S., Singh, S., Das, A. K., Bhattacharya, S. S. & Pal, D. (2021). Redox Buffering Capacity of Nanomaterials as an Index of ROS-based Therapeutics and Toxicity: A Preclinical Animal Study. *ACS Biomaterials Science & Engineering*, 7(6), 2475-2484. https://doi.org/10.1021/acsbiomaterials.1c00402

[7] * Adhikari, A., Mondal, S., Chatterjee, T., Das, M., **Biswas, P.**, Darbar, S., Alessa, H., Al-Thakafy, J. T., Ahmed, S. A., Das, A. K., Bhattacharyya, M., Pal, S. K. (2021). Redox nanomedicine ameliorates chronic kidney disease (CKD) by mitochondrial reconditioning in mice. *Communications biology,* 4, 1-15. https://doi.org/10.1038/s42003-021-02546-8

[8] * Mondal, S., Ghosh, R., Adhikari, A., Pal, U., Mukherjee, D., **Biswas, P.**, Darbar, S., Singh, S., Bose, S. & Saha-Dasgupta, T, Pal, S. K. (2021). In-vitro and microbiological assay of functionalized nano-hybrids to validate their differential efficacy in nano-theranostics: A combined spectroscopic and computational study. *ChemMedChem*, 16, 1-13. https://doi.org/10.1002/cmdc.202100494

Peer-reviewed Conference Publication

[1] * **Biswas, P.**, Adhikari, A., Mondal, S., Das, M., Bhattacharya, S. S., Pal, D., Choudhury, S. S. & Pal, S. K. (2021). Synthesis and spectroscopic characterization of a zinc oxide-polyphenol nanohybrid from natural resources for enhanced antioxidant activity with less cytotoxicity. *Materials Today: Proceedings,* 43**,** 3481-3486. https://doi.org/10.1016/j.matpr.2020.09.567

[2] * Mondal, S., Adhikari, A., Singh, M., Ghosh, R., Goswami, M., **Biswas, P.** & Pal, S. K. (2021). Spectroscopic study on the interaction of Co^{2+} with citrate-Mn_3O_4: Towards the development of nanotherapy against cobalt toxicity. *Materials Today: Proceedings,* 43**,** 3692-3697. https://doi.org/10.1016/j.matpr.2020.10.981

* are not included in the thesis.

List of International/ National Conferences Attended

[1] Participated in a two-day National Level *'Workshop on Advanced Nanotherapeutics'*; Uluberia College, Uluberia, Howrah, West Bengal, India on 10th and 11th March, 2017.

[2] Participated and presented a paper in the *'International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology (5NANO)'* organized by Mangalam College of Engineering, Ettumanoor, Kottayam, Kerela, India on 23rd and 24th July, 2020.