

# **Single Molecular Level Study of Proteins: Insights from Conformational Fluctuation Dynamics and Structural Parameters**

*A Thesis  
Submitted in Partial Fulfillment of the  
Requirements for the Degree of*

**DOCTOR OF PHILOSOPHY**

*by*

**BHASWATI SENGUPTA**



*to the*

**Department of Chemistry  
Indian Institute of Technology Kanpur  
Kanpur, India**

August, 2017



## STATEMENT

I hereby declare that the work presented in the thesis entitled “**Single Molecular Level Study of Proteins: Insights from Conformational Fluctuation Dynamics and Structural Parameters**” is the result of original work carried out by me in the Department of Chemistry, Indian Institute of Technology Kanpur under the supervision of **Dr. Pratik Sen**.

In keeping with general practice of reporting scientific observations, due acknowledgements have been made wherever the work described is based on the findings of other investigators.

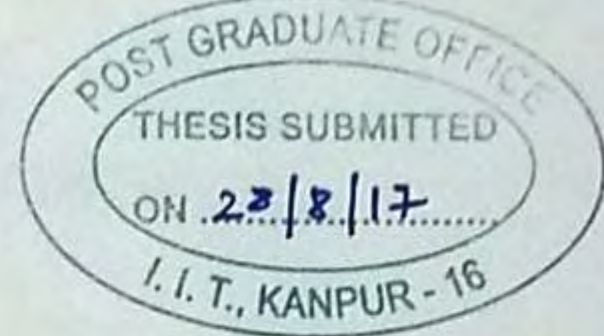
August, 2017

IIT Kanpur

*Bhaswati Sengupta*  
(Bhaswati Sengupta)



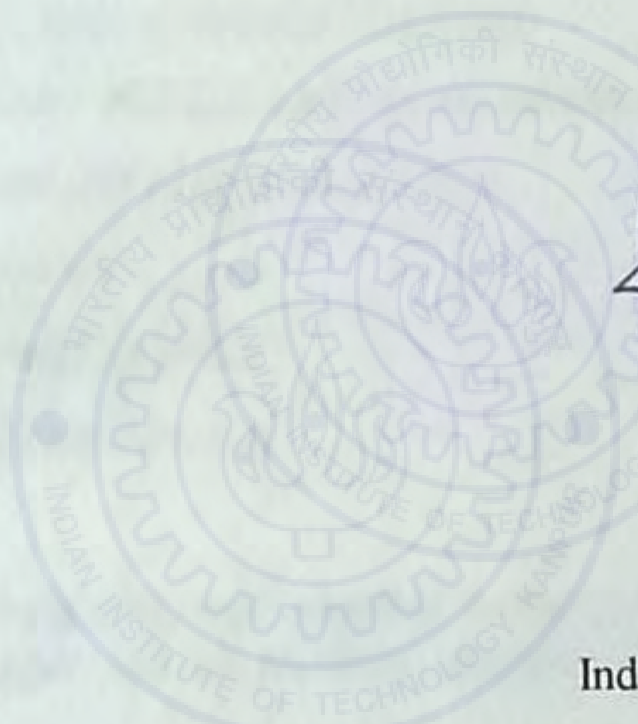




## CERTIFICATE

It is certified that the work reported in the thesis entitled “Single Molecular Level Study of Proteins: Insights from Conformational Fluctuation Dynamics and Structural Parameters” has been carried out by Ms. Bhaswati Sengupta under my supervision and has not been submitted elsewhere for a degree.

August, 2017  
IIT Kanpur



*Pratik Sen*  
27-08-2017

(Dr. Pratik Sen)

Thesis Supervisor

Department of Chemistry

Indian Institute of Technology Kanpur

Kanpur – 208016, India



डा० प्रतीक सेन / Dr. Pratik Sen  
सह-प्राध्यापक, रसायन विभाग  
Associate Professor  
Department of Chemistry  
भारतीय प्रौद्योगिकी संस्थान कानपुर  
Indian Institute of Technology Kanpur  
Kanpur-208016, INDIA





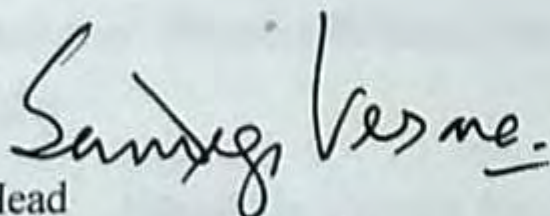
**Department of Chemistry**  
**Indian Institute of Technology Kanpur**

**Certificate of course work**

This is to certify that **Ms. Bhaswati Sengupta** has satisfactorily completed all the courses required for Ph. D. degree. The courses include:

AE698	Intro to Virtual Instrumentation
CHM636	Physical Photochemistry
CHM654	Supramolecular Chemistry
CHM629	Principles of Physical Chemistry
CHM664	Modern Physical Methods in Chemistry
CHM679	Molecular Reaction Dynamics
CHM799	Research
CHM800	General Seminar
CHM801	Graduate Seminar

**Ms. Bhaswati Sengupta** was admitted to the candidacy of the Ph. D. degree in February 2013 after she successfully completed the written and oral qualifying examinations.



Head  
Department of Chemistry  
Indian Institute of Technology Kanpur  
Kanpur – 208016, India

**Sandeep Verma**  
Professor & Head  
Department of Chemistry  
I.I.T. Kanpur-208016 (India)



28.08.2017  
Convenor, DPGC  
Department of Chemistry  
Indian Institute of Technology Kanpur  
Kanpur – 208016, India





## ***Acknowledgements***

*I still remember the day I arrived at IIT Kanpur and got registered. One sentence that the then Head of the Department Prof. R. N. Mukherjee said during the orientation session still echoes in my ears. It goes like “You want to add a Dr. in front of your name, you must achieve it. No one will give you the title, you must fetch it yourself.” But honestly in my case, if I am able to achieve it, an extensive part of credit goes to my thesis supervisor Dr. Pratik Sen. He has been a friend, philosopher and guide throughout this long journey called Ph. D. This thesis literally belongs to him. From the day I joined, he nurtured me in every possible way he could. From fruitful academic discussion to amiable personal suggestions, he was always there for me. I will always remain indebted to him for being such a great mentor and shaping my career for a great future.*

*I want to thank Prof. S. Manogaran, Prof. P. K. Bharadwaj, Prof. A. Chandra, Prof K. Srihari, Dr. P. Sen, Prof. K. Poddar (Dept. of Aerospace Engg) for teaching me courses. I also thank Dr. D. H. Dethe and Dr. R. Ramapanicker for the generous help regarding the organic chemistry part of my thesis work. My sincere thanks to Prof. S. Matheswaran, Prof. D. Katti and Dr. A. K. Thankur from the Department of Biological Science and Biological Engineering for letting me use their lab facility. I wholeheartedly acknowledge Dr. Sachin Dev Verma for teaching me the details of building a single molecular fluorescence measurement set up.*

*Major part of my thesis work has been based upon single molecular fluorescence experiments which I learnt from Prof. Sudipta Maiti from TIFR, Mumbai and Prof. Thorsten Wohland from NUS, Singapore. I convey my sincere gratitude to them.*

*I thank the Current Head of the Department Prof. Sandeep Verma and the former Heads of the Department Prof. R. N. Mukherjee and Prof. P. K. Bharadwaj for their cooperation throughout my Ph. D tenure. I also thank other faculty members who always extended their helping hands while I asked for their valuable suggestions. The*

*chemistry office staffs, especially Sudha madam and Geeta madam were really cooperative and I express my thanks in exchange.*

*I am indebted to Dr. Samragnee Dutta for her constant support. She has been a tremendous source of motivation and always held my hand while I was paving my way through bad times.*

*I wish to extend my gratitude to Mrs. Soma Ghosh Dastidar and Prof. P. S. Ghosh Dastidar for providing me a home far from home in IIT Kanpur. Soma Madam has always bestowed me with her never ending maternal love which I will always be missing once I move out of this place.*

*My lab has been a home and the members have been my family for this five and a half years. I thank all of them from the core of my heart. My lab seniors Rajeevda, Shahnawazda and Shradheyda are the ones from whom I gathered practical knowledge regarding experiments. Me, Vaisakh and Puspall has been together for five long years through all the ups and downs. My heartfelt love to my junior Nilimesh who has been like a brother to me for since he joined. Other lab mates Navin, Shovon, Aritra, Vipin, Vijaykantda, Faizida and Gyanesh made my stay in lab memorable. I also thank the M. Sc. Students Indrani, Arusha and Apala for helping me in building precious memories in lab.*

*As it is said that Ph. D. is a long and tough journey and no one can walk it alone, I am blessed to have a lot of good friends during my stay in campus. The list is long but Rahul and Paramita need a special mention. I also thank Ashabari, Gaurav, Nabanitadi, Sancharidi, Aviru, Hansaraj, Subuhi, Priyank, Sonam, Jyoti, Arvind, Suparno, Santoshda, Virender for making my IITK days beautiful. Special thanks to Pritam without whom this journey would have been incomplete. A warm acknowledgement to my friends Poulomi, Isha, Debolina, Sritama, Sohini for being there in my life. I thank the almighty for creating beautiful human beings like Rick, Nayanika, Anindita, Saheli, Rita and Arghya and gifting them to me as best friends. They will always occupy special place in my heart.*

*Devout thanks to my beloved parents for making me who I am today. They have made all the necessary sacrifice to make sure that I reach here. They inspired me, loved me, guided me and enriched me with moral values to build me as a good human being. Special thanks to my Dadu who always stood by me and believed that I can never fail. My Boropisi who is the dearest to me left me few days ago, I thank her for making me realize that even if I fail, she will hold me, love me, pamper me and prepare me for the next venture. Sincere thanks to my other family members and neighbours who are the sweet inseparable parts of my beautiful world.*

*Finally I thank almighty God for making my life fulfilled with happiness.*

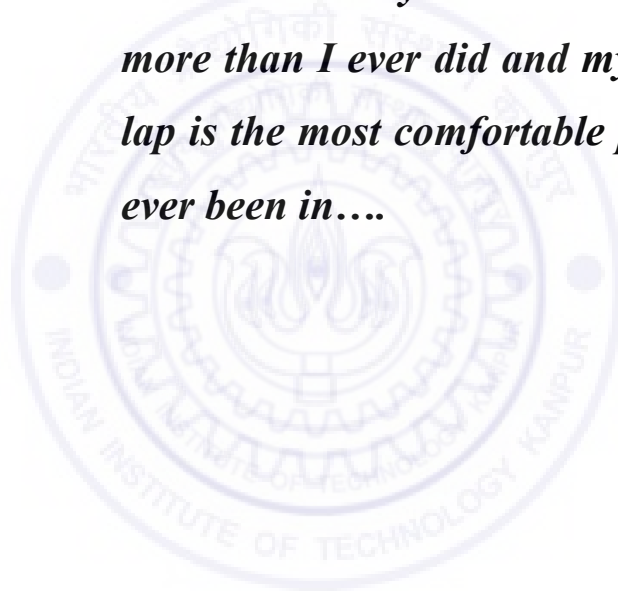
**Bhaswati Sengupta**







*Dedicated to my Dadu who believed in me  
more than I ever did and my Boropisi whose  
lap is the most comfortable place that I have  
ever been in....*





## SYNOPSIS

---

Name of the Student	Bhaswati Sengupta
Roll Number:	11207064
Degree for which Thesis is Submitted:	Ph. D.
Department:	Chemistry
Supervisor:	Dr. Pratik Sen
Thesis Title:	Single Molecular Level Study of Proteins: Insights from Conformational Fluctuation Dynamics and Structural Parameters
Month and Year of Submission:	August, 2017

---

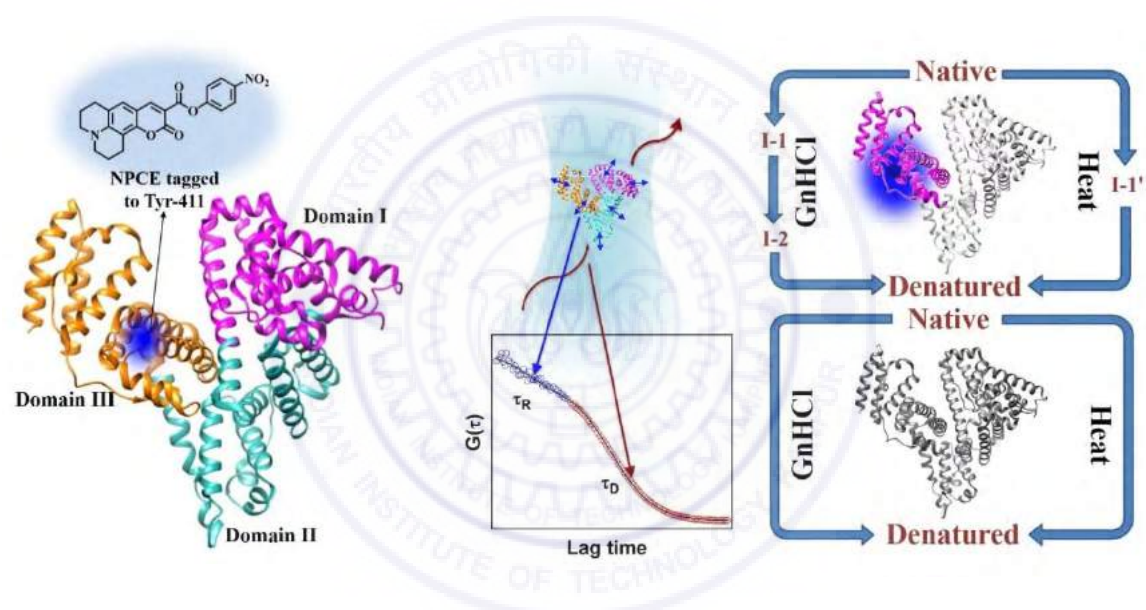
This thesis elucidates my work on the application of fluorescence correlation spectroscopy (FCS), a single molecule fluorescence spectroscopic technique, to simultaneously study protein structure and dynamics simultaneously. The study was performed on human serum albumin (HSA),  $\beta$ -lactoglobulin (BLG) and papain. HSA is a large multidomain protein with a molecular weight of 66.5 kDa whereas BLG and papain are smaller proteins with molecular weights of 18.4 kDa and 23.5 kDa respectively. The proteins were exposed to (i) chemical and thermal denaturation conditions, (ii) a confined environment and (iii) protease digestion to investigate the change in their structure and dynamics under adverse condition. Highly fluorescent thiol specific rhodamine and coumarin dyes were used to tag the free cysteine residues, whereas a hydroxyl specific coumarin dye was used to tag tyrosine residues to obtain the fluorescent labelled proteins; supplementary ensemble experiments were performed to support the conclusions.

### Summary of the work done:

#### (a) Elucidation of $\mu$ s Dynamics of Domain-III of Human Serum Albumin during the Chemical and Thermal Unfolding

The local structural dynamics and denaturation profile of domain-III of HSA against guanidine hydrochloride (GnHCl) and temperature were studied using a

coumarin based solvatochromic fluorescent probe *p*-nitrophenylcoumarin ester (NPCE), covalently attached to Tyr-411 residue. Through steady-state, time-resolved and single molecule fluorescence studies it was ascertained that (i) domain-III of HSA is very sensitive to GnHCl but somewhat resistant to increasing temperature and (ii) domain-III unfolding proceeds in a way altered to the overall unfolding of HSA. While the overall denaturation of HSA is a two-state process for both GnHCl and heat, domain-III adopts two intermediate states upon GnHCl induced denaturation and one intermediate state upon temperature induced denaturation. Fluorescence correlation spectroscopic investigation revealed the conformational dynamics of domain-III of HSA in the native, intermediates and denatured state.

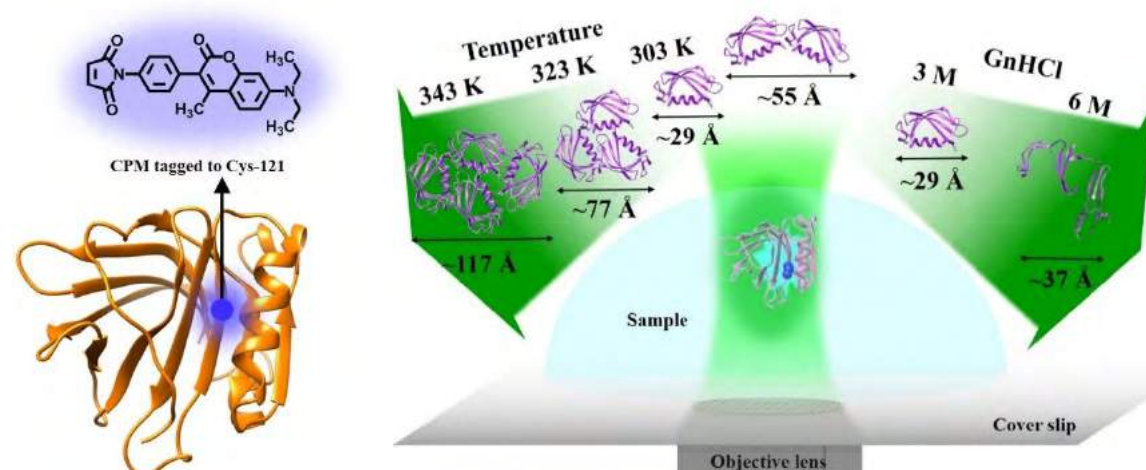


### (b) Monomerization and Aggregation of $\beta$ -lactoglobulin under Adverse Condition

The structural change of  $\beta$ -lactoglobulin at pH 7.4 under the influence of heat and GnHCl has been studied. For this purpose, the only free cysteine residue of  $\beta$ -lactoglobulin was tagged with fluorescent 7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM) and investigated using steady-state spectroscopy, Förster resonance energy transfer (FRET) and single molecule fluorescence techniques. It was found that at  $\sim 3$  M GnHCl the monomer content is almost 100%. Beyond this, the monomers start unfolding. By contrast, upon heating, both reversible and irreversible formation of differently sized aggregates was observed. This aggregation is triggered



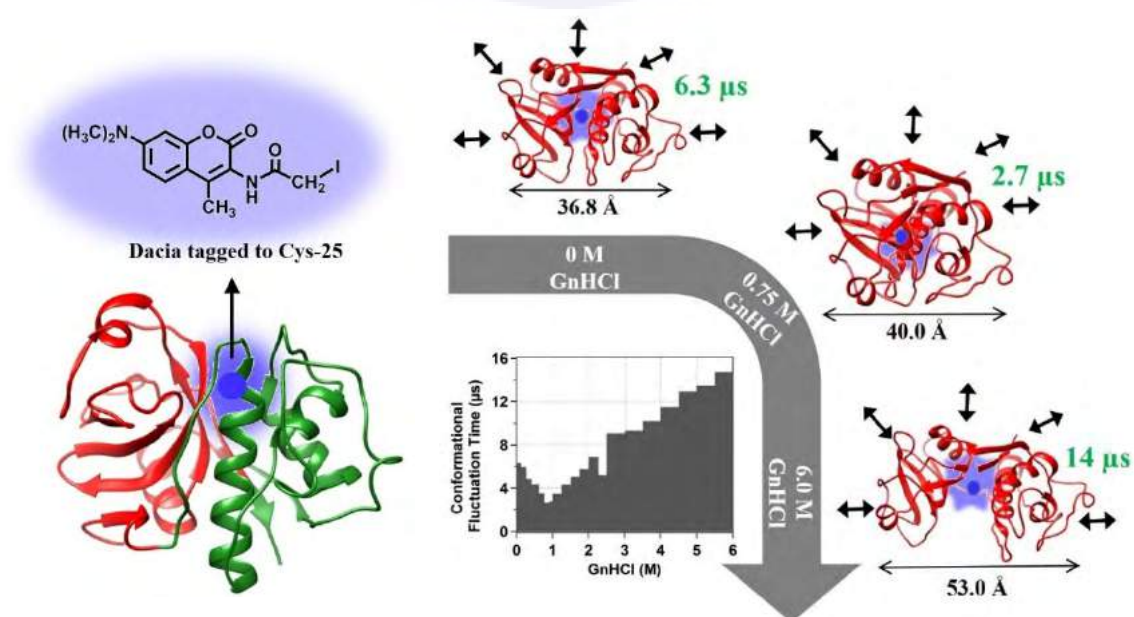
by the initial formation of the monomeric form of  $\beta$ -lactoglobulin. The formation of small aggregates was found out to be reversible whereas that of larger aggregates is



irreversible.

### (c) Investigation of Structure and Dynamics of Papain under Denaturation

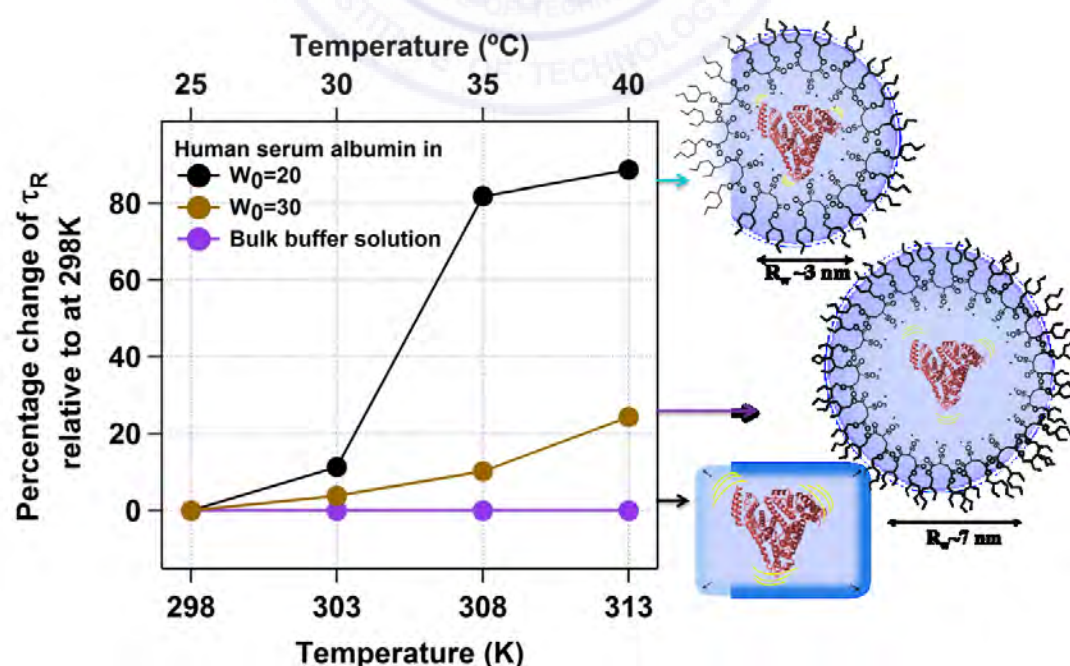
GnHCl induced structural and dynamic variation of papain was studied by site-specific tagging of papain at the Cys-25 residue with N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA). Steady-state, time-resolved, and single molecule fluorescence techniques were used to map the changes in overall size and local environment of the protein by monitoring the fluorescence response of the probe.



It was found that papain undergoes a two-step denaturation in the presence of GnHCl. Fluorescence correlation spectroscopic (FCS) data revealed that, in the presence of small amounts of GnHCl the active site of papain adopts a more compact structure than in the native state (even though the overall size increases), which was designated the intermediate state. Upon further increase in GnHCl concentration, papain was found to denature, as indicated by an increase in hydrodynamic radius. By contrast, the measurement of ellipticity, or gross helical structure, by circular dichroism spectroscopy was found to not detect this structural transformation.

**(d) Startling Temperature Effect on Proteins when Confined: Human Serum Albumin in Reverse Micelle**

The effect of confinement, in combination with temperature, on the conformational fluctuations of domain-I of HSA was studied using FCS. To this end, HSA was tagged with tetramethylrhodamine-5-maleimide (TMR). Sodium bis(2-ethylhexyl) sulfosuccinate (AOT) reverse “water pool” micelles were used for confinement. It was observed that the conformational fluctuation time when confined in a water pool of 3.5 nm radius is 6 fold shorter than in bulk medium. Upon increasing the size of the water-pool the conformational fluctuation time was found to increase monotonically and approach the bulk value. The effect of confinement is consistent



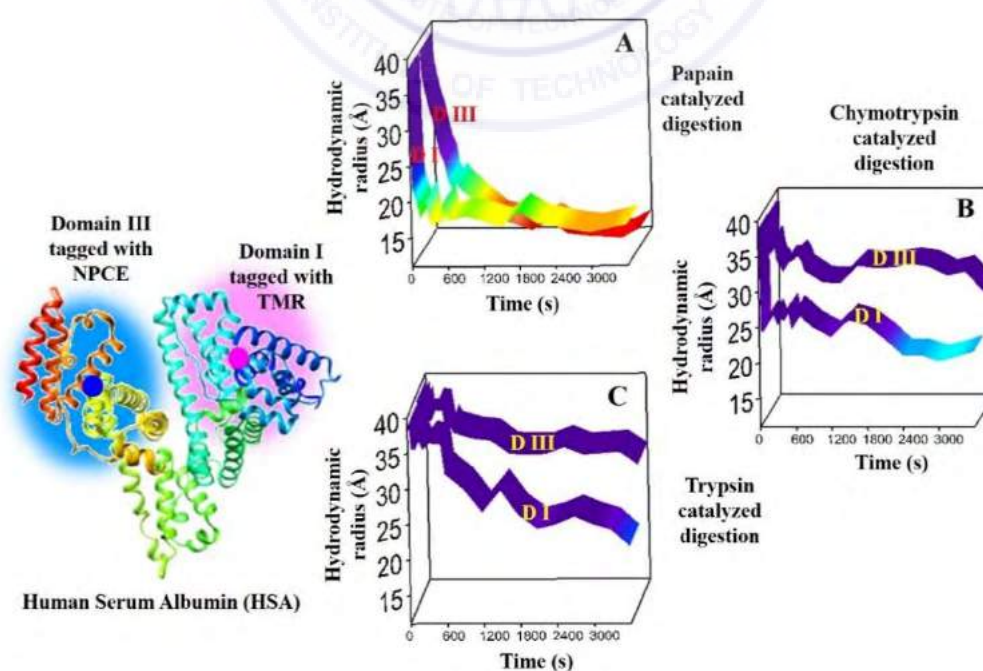
with the general expectation of restricted motion of a macromolecule upon confinement.

By contrast, the effect of temperature was found to be surprising. In bulk medium, no change in conformational fluctuation time was found in the temperature range of 298 K and 313 K, whereas; when HSA is confined in AOT water-pools of radius 3.5 nm; an 88% increase in the fluctuation time was observed over this range. We conclude that the observed prominent thermal effect on the conformational dynamics of domain-I of HSA in the water-pool of AOT reverse micelles compared to bulk medium arises from the confined solvent effect.

#### (e) Directionality of Peptidase Activity on Human Serum Albumin Revealed by Fluorescence Correlation Spectroscopy

Enzymatic proteolysis or protein digestion is the fragmentation of protein into smaller peptide units under the action of peptidase enzymes. In this chapter, the digestion of HSA by papain, chymotrypsin and trypsin was studied using FCS. Domain-I of HSA was tagged with tetramethylrhodamine-5-maleimide (TMR) and Domain-III separately with *p*-nitrophenylcoumarin ester (NPCE) separately and the protein was subjected to the peptidases maintaining a 1:1 molar ratio at 323 K.

Following the change in hydrodynamic radius, as monitored by FCS, it was



confirmed that under similar experimental conditions, the order of digestion efficiency

is papain> chymotrypsin> trypsin. More interestingly, a faster decrease of hydrodynamic radius was observed when the fluorescence from domain-I was monitored in FCS, compared to that of domain-III. This observation suggests that all of these enzymes preferably cleave HSA in domain-I, despite the fact that the possible cleavage sites for the three enzymes are evenly distributed over the peptide chain of HSA. Such specificity of enzymatic digestion was observed here for the very first time and suggests unexpected local accessibility differences.





## Table of Contents

<b>Statement</b>	iii
<b>Certificate-I</b>	v
<b>Certificate-II</b>	vii
<b>Acknowledgements</b>	ix
<b>Synopsis</b>	xv
<b>Chapter 1. Introduction</b>	1-28
<hr/>	
1.1. Proteins: A Brief Overview	
1.1.1. Basic Structural Features	
1.1.2. Domains of proteins	
1.1.3. Classification of proteins	
1.2. Protein Folding	
1.2.1. Thermodynamics of Protein Folding	
1.2.2. Kinetics of Protein Folding	
1.3. Conformational Dynamics of Protein	
1.4. Protein Denaturation	
1.4.1. Chemical denaturation	
1.4.2. Thermal Denaturation	
1.5. Protein Aggregation	
1.6. Effect of Confinement on Proteins	
1.7. Digestion of Proteins	
1.8. Protein spectroscopy	
1.9. Protein Tagging	
1.10. Single Molecular Spectroscopy of Proteins	
1.11. Proteins under Investigation	
1.11.1. Human Serum Albumin	
1.11.2. $\beta$ -Lactoglobulin	
1.11.3. Papain	
<b>Chapter 2. Experimental Methods</b>	29-68
<hr/>	
2.1. Instruments	
2.1.1. Steady State Absorption Measurements using Spectrophotometer	

- 2.1.2. Steady State Fluorescence Measurements using Spectrofluorometer
- 2.1.3. Circular Dichroism Spectroscopy using Circular Dichroism Spectrometer
- 2.1.4. Time Resolved Fluorescence Measurements
- 2.1.5. Single Molecular Fluorescence Measurements
  - 2.1.5.1. History and Invention of FCS
  - 2.1.5.2. Theory of FCS
  - 2.1.5.3. Factors Contributing in the Fluorescence Fluctuation
  - 2.1.5.4. Mathematical Understanding of the Fitting Equation
  - 2.1.5.5. Extracting Structural Information from the Diffusion Time
  - 2.1.5.6. FCS Set up
  - 2.1.5.7. Specification of the Components of FCS Set up
  - 2.1.5.8. Calibration of the Instrument
  - 2.1.5.9. Temperature Dependent FCS Measurements
- 2.2. Calculation of Thermodynamic Parameters Related to Unfolding and Intermediate Formation
- 2.3. Materials
- 2.4. Protein Labelling
  - 2.4.1. Labelling of HSA with *p*-nitrophenylcoumarin ester (NPCE)
    - 2.4.1.1. Synthesis and Characterization of NPCE
    - 2.4.1.2. Labelling of HSA with NPCE
  - 2.4.2. Labelling of  $\beta$ -lactoglobulin with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM)
  - 2.4.3. Labelling of Papain with DACIA
  - 2.4.4. Labelling of HSA with TMR
- 2.5. Sample preparation
- 2.6. Auxiliary instruments

### **Chapter 3. Elucidation of $\mu$ s Dynamics of Domain-III of Human Serum Albumin during the Chemical and Thermal Unfolding**

69-89

- 
- 3.1. Introduction
  - 3.2. Results
    - 3.2.1. Circular Dichroism Spectroscopy
    - 3.2.2. Steady State Absorption and Fluorescence Spectroscopy
    - 3.2.3. Fluorescence Lifetime Measurement

- 3.2.4. Fluorescence Correlation Spectroscopic Measurements
- 3.3. Discussion
- 3.4. Calculation of Thermodynamic Parameters
- 3.5. Conclusion

---

**Chapter 4. Monomerization and Aggregation of  $\beta$ -lactoglobulin under Adverse Condition** 91-113

---

- 4.1. Introduction
- 4.2. Results
  - 4.2.1. Circular Dichroism Spectroscopy
  - 4.2.2. Steady State Absorption and Fluorescence Spectroscopy
  - 4.2.3. Förster Resonance Energy Transfer (FRET) Experiment
  - 4.2.4. Fluorescence Correlation Spectroscopic Measurements
- 4.3. Discussion
- 4.4. Calculation of Thermodynamic Parameters
- 4.5. Conclusion

---

**Chapter 5. Investigation of Structure and Dynamics of Papain under Denaturation** 115-129

---

- 5.1. Introduction
- 5.2. Results
  - 5.2.1. Circular Dichroism Spectroscopy
  - 5.2.2. Steady State Absorption and Fluorescence Spectroscopy
  - 5.2.3. Fluorescence Lifetime Measurement
  - 5.2.4. Fluorescence Correlation Spectroscopic Measurements
- 5.3. Discussion
- 5.4. Calculation of Thermodynamic Parameters
- 5.5. Conclusion

---

**Chapter 6. Startling Temperature Effect on Proteins when Confined: Human Serum Albumin in Reverse Micelle** 131-149

---

- 6.1. Introduction
- 6.2. Results
  - 6.2.1. Circular Dichroism Spectroscopy
  - 6.2.2. Steady State Absorption and Fluorescence Spectroscopy

- 6.2.3. Fluorescence Correlation Spectroscopic Measurements
- 6.3. Discussion
- 6.4. Conclusion

---

**Chapter 7. Directionality of Peptidase Activity on Human Serum Albumin  
Revealed by Fluorescence Correlation Spectroscopy** 151-166

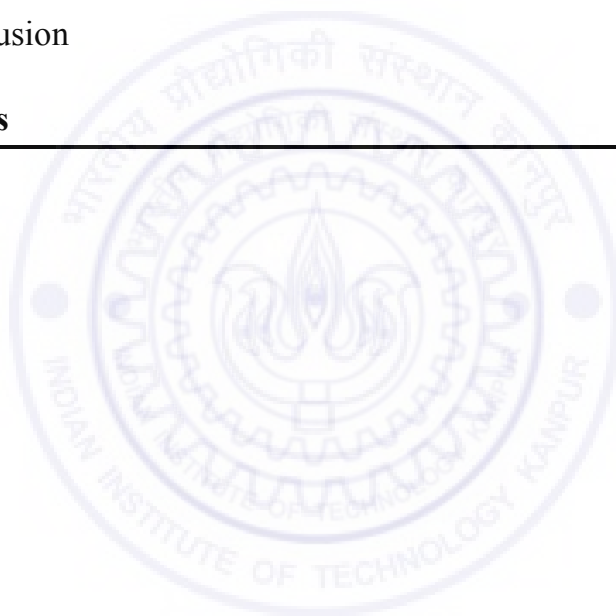
---

- 7.1. Introduction
- 7.2. Results
  - 7.2.1. Circular Dichroism Spectroscopy
  - 7.2.2. Steady State Absorption and Fluorescence Spectroscopy
  - 7.2.3. Fluorescence Correlation Spectroscopic Measurements
- 7.3. Discussion
- 7.4. Conclusion

---

**List of Publications** 167

---





# *Chapter 1*

---

## *Introduction*





## 1.1. Proteins: A Brief Overview

Proteins are kind of factotums for biological systems. They are the most abundant biological macromolecules and perform almost all kinds of biotic activities. The major roles of proteins in living organisms are replication of DNA, catalysing biological reactions, acting as hormones to regulate physiological behaviour, carrying drugs and other molecules to their respective location, taking part in the immunity system *etc.*<sup>1-3</sup> The loss of proteins in cells by degradation or transport is compensated by the biosynthesis of proteins. Biosynthesis of proteins is a complex process, which involves several steps and is controlled by the ribonucleic acids (RNAs). This is the internal source of proteins in human body apart from which the external source is also required for the supply of essential amino acids. These external sources can be meat, milk, fish, eggs, spinach, nuts, beans, grains *etc.*

**1.1.1. Basic Structural Features:** Proteins are made of amino acid units.<sup>3-4</sup> Amino acids are molecules which have carboxyl and amine functional groups with a side chain which is different for each amino acid. They are covalently connected by amide linkage which are better known as peptide bonds. There are a total of twenty different amino acids reported till date. When the number of amino acids connected along chain is less than ~20 it is usually termed peptide whereas the larger collections are called proteins. Amongst these amino acids some are hydrophobic and rest are hydrophilic. Under physiological conditions, proteins stay in folded state which is called the native state of the protein and corresponds to the minima in the potential energy landscape as a function of structural parameters. Proteins are structurally characterized by four structural specifications.

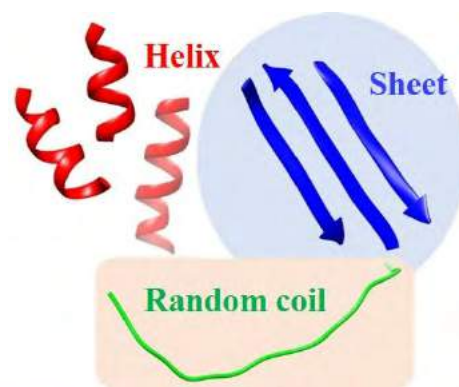
**I. Primary Structure:** The sequence of amino acids in the protein chain is defined as the primary structure of the protein (scheme 1.1). The N-terminal side of a protein

Residue No	
1-50	M T F E A R L V Q G S I L K K V L E A L K D L I N E A C W D I S S S G V N L O S M D S S H V S L V Q Q
51-100	T I R S E G F D T Y R C D R N L A M G V N L T S M S K I L K C A G N E D I T L R A E D N A D T I A
101-150	L V F E A P N Q E K V S D Y E M R L M D L D V E Q L G I P E Q E Y S C V V R M P S G E F A R I C R D
151-200	L S H I G D A V V I S C A R D G V K F S A S G E L G N C N I K L S Q T S N V D K E E A V T I E M N
201-250	E P V Q L T F A L R V L N F F T K A T P L S S T V T L S M S A D V P L V V E Y K I A D M G H L K V Y
251-261	L A P K I E D E E G S

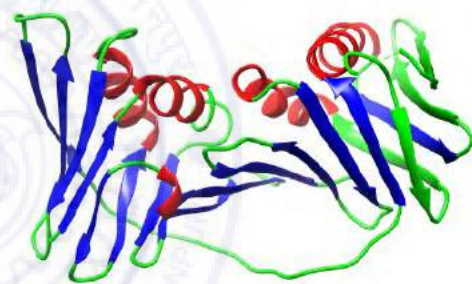
**Scheme 1.1:** Primary structure of human PCNA (proliferating cell nuclear antigen).

is the end where there is a free  $\text{-NH}_2$  group present and the C-terminal side is where there is a free  $\text{-COOH}$  residue.

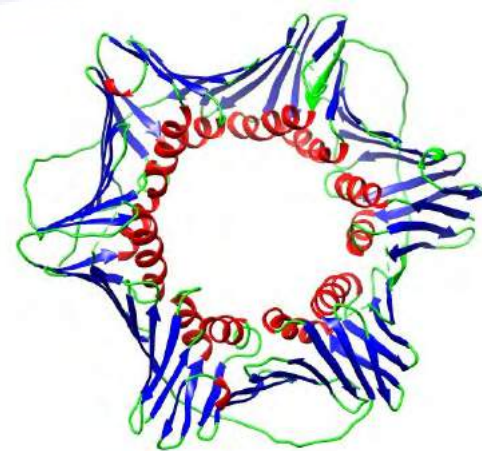
**II. Secondary Structure:** Secondary structure of the protein is a measure of regular repetitive structures in the three dimensional form of the proteins (scheme 1.2). The most common motifs present in proteins are the helices and the sheets. More commonly they are denoted as alpha helices and beta sheets as they were discovered in alpha keratin and beta keratin respectively.<sup>3</sup> The alpha helix is a spiral configuration where each turn is connected to its previous turn *via* hydrogen bonds between the hydrogen and the oxygen of the peptide bond. In most of the cases alpha helix are found to be right handed. The beta sheet is a flat configuration where two consecutive strands connected and stabilized by hydrogen bonds. The unstructured portion of the protein chain is often denoted as the random coil.



**Scheme 1.2:** Secondary structure of human PCNA (proliferating cell nuclear antigen).



**Scheme 1.3:** Tertiary structure of human PCNA (proliferating cell nuclear antigen).



**Scheme 1.4:** Quarternary structure of human PCNA (proliferating cell nuclear antigen).

**III. Tertiary Structure:** The tertiary structure is the three dimensional structure of the protein. As already discussed, proteins adopt specific three dimensional structure to maximize stabilizing hydrophilic interactions. Tertiary structure results from the folding of protein to reach the most stabilized state or the native state.

**IV. Quaternary Structure:** Proteins often stay as a collection of subunits. Many proteins are dimeric, trimeric, tetrameric



or oligomeric in the native state. The orientation of the subunits in the three dimensional space is known as the quaternary structure of the protein (scheme 1.4).

**1.1.2. Domains of Proteins:** Domains are distinct structural regions of protein which behave in a particular fashion as compared to the rest of the framework.<sup>5-7</sup> The number of different domains and their size can vary in different proteins. While most of the mammalian proteins consist of one or two domains, larger proteins can consist more than twenty domains. The average number of amino acid residues forming a domain is 100 with minimum and maximum numbers as ~40 and ~300 respectively. The function, stability and folding properties of different domains can diverge significantly from each other.

**1.1.3. Classification of Proteins:** The legion consisting thousands of proteins can be classified based on their structure or activity. This classification of proteins facilitates the easy identification of a protein's structural and functional properties by looking at the family it belongs to. Proteins can be categorized on the basis of following fields

**I. Composition of Proteins:**

- Simple proteins which are made of amino acids only, e.g. albumin, keratin *etc.*
- Conjugate proteins which contain a non-protein part along with the protein part, e.g. lypoproteins (protein + fat), glycoproteins ( protein + carbohydrate).

**II. Structure of Proteins:**

- Proteins having elongated fibre-like shape are called fibrous proteins, e.g. collagen, elastin.
- Most of the proteins maintain a quite compact structure with a somewhat spherical outline. These proteins are called globular proteins. e.g. lactoglobulin, haemoglobin *etc.*
- The proteins that lack specific native structure are referred to as intrinsically disordered proteins.

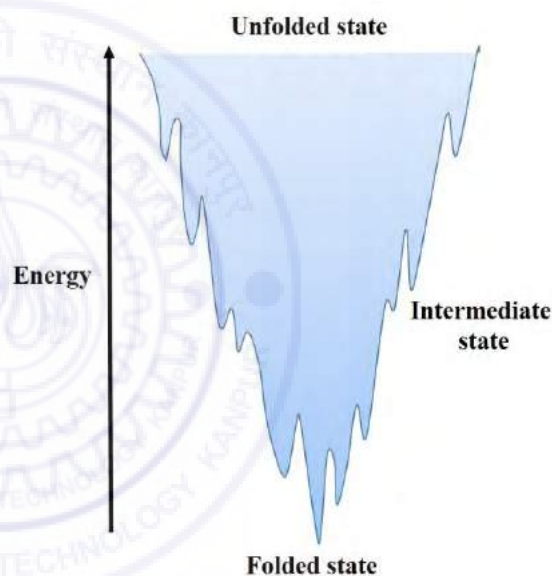
**III. Function of Proteins:**

- Proteins that perform carriage and transport of drugs and other small biologically active molecules are better known as transport proteins, e.g. haemoglobin, transferrin, human serum albumin *etc.*

- Proteins that catalyse biological reactions are known as enzymes, e.g. papain, chymotrypsin *etc.*
- Proteins that act as messenger to control major biological functions are known as hormones, e.g. insulin, glucagon *etc.*
- Proteins that act as warehouse of the biological systems are known as storage proteins, e.g. ferritin, egg yolk phosvitin *etc.*

## 1.2. Protein Folding

The process by which proteins adopt their native structure from being a straight long chain of amino acids is called protein folding.<sup>8-21</sup> This folding process is highly important because, unless the protein is properly folded, it fails to perform the biological role that it is supposed to play. The folding process is driven by the presence of stabilizing and destabilizing forces acting on it. Considering their biological importance, protein folding and misfolding have gained attention throughout several decades. The first ground-breaking work in this regard was performed by Nobel laureate



**Scheme 1.5:** Free energy funnel for protein folding.

Christian Anfinsen who postulated that the amino acid sequence solely administer the native structure of a protein and a protein would return to its well-defined native state upon refolding.<sup>8-9</sup> In his reports he proved this postulate using ribonuclease as a specimen. Since then, it has been accepted that the native state corresponds to the global minima in the potential energy landscape. Recent reports contradict this statement and support the fact that a protein can return to a misfolded state or even can lack a specific native state. These proteins without a well-defined native state are termed intrinsically disordered proteins (IDPs) and are responsible for several diseases like Alzheimer's disease and Creutzfeldt-Jakob disease.<sup>22-24</sup> However, most proteins

are still believed to follow the Anfinsen's dogma and return to the minima of the protein folding funnel upon refolding.<sup>25-26</sup> The protein folding funnel is a potential energy surface as a function of conformational space and is different in each protein. The minima and maxima of the folding funnel (scheme 1.5) correspond to the native and unfolded states of the protein respectively. Additionally, there are many local minima which correspond to the intermediate states. The route that a protein takes to reach the native state from the unfolded state is governed by the slope and shape of the funnel.

**1.2.1. Thermodynamics of Protein Folding:** There are two thermodynamics factors, entropy ( $S$ ) and enthalpy ( $H$ ) that need to be addressed when discussing protein folding.<sup>21</sup> In order for a reaction to proceed the Gibbs free energy ( $G$ ) of the final state must be less than that of the initial state, i.e.  $\Delta G$  must be negative. According to the basic equation of thermodynamics, the Gibbs free energy change is governed by  $\Delta H$  and  $\Delta S$  counterbalancing each other.

$$\Delta G = \Delta H - T\Delta S \quad (\text{Eq. 1.1})$$

In the unfolded state of the protein the system is chaotic due to the probability of thousands of conformations due to unrestricted movement of polypeptide chains and upon folding this messiness decreases which furnishes a decrease in entropy. This decrease in entropy makes  $\Delta G$  more positive leaving the process less feasible. This unfavourable situation can be overcome when  $\Delta H$  acquires a high negative value and compensates for the entropy change. The major contribution of  $\Delta H$  arises from the hydrogen bonding, formation of salt bridges and stabilizing hydrophobic or hydrophilic interactions. Upon protein folding the hydrophobic amino acid residues are buried inside the protein core and the hydrophilic ones are exposed towards the surface where they can experience stabilizing hydrophilic interactions. Besides, the formation of disulphide bonds is also an important factor in the process of folding.<sup>16</sup> The exact rationale behind this remains unsettled till date. People claim that the disulphide bonds are present even in the unfolded states, making them more ordered and curtailing the unfavourable entropic contribution towards folding. Typically, the total energy difference between the folded and unfolded state is usually around 20-40

$\text{kJ mol}^{-1}$  and the equilibrium constant is in the order of  $10^4$ - $10^7$  for a small natural protein.<sup>17</sup>

**1.2.2. Kinetics of Protein Folding:** It is well established that the timescale of protein folding ranges from microseconds to seconds. This is surprising given the fact that the statistically calculated time is nowhere near this. For a 100-residue protein, given each of the residue can have any of the three probable conformational states, the time taken for sampling each conformational state and arrive at the lowest energy conformation would need  $\sim 10^{30}$  years assuming that conversion between conformational states occurs at picosecond timescales. This was predicted by Cyrus Levinthal back in 1969, and the disagreement of experimental and theoretical timescale of protein folding is known as Levinthal paradox.<sup>27</sup> The best solution of this paradox states that proteins are folded gradually guided by local interactions rather than global ones. These locally stabilized structures are designated as intermediate states during the protein folding. In the course of folding, a protein tries out these locally stabilized states to find the global minimum or the native state as shown in the protein folding funnel (scheme 1.5).

### 1.3. Conformational Dynamics of Proteins

Back during 1970s it was believed that proteins possess a rigid structure, which is solely determined by their amino acid sequence. With the advancement of theoretical and experimental research, it was proved that the protein structure is an average of several closely related structure known as different conformational states of a protein. These are owing to the possibility of (a) free rotation around the bonds and (b) the different orientation of side chains in three dimensional space and these states differ only in the local coordinates of the amino acid residues. Protein molecules always undergo transitions between these states rendering the protein structure as a dynamic one rather than a static one. The interconversion between these states is known as the conformational dynamics of the protein.<sup>28-37</sup> The amplitude and timescale of conformational dynamics for various proteins can range from femtoseconds to hours, depending on the environment of the protein. Conformational dynamics is not a futile agitation of proteins, in fact they do serve key roles in protein functions like catalysis, drug binding, hormonal activities *etc.* As the conformational dynamics result from a

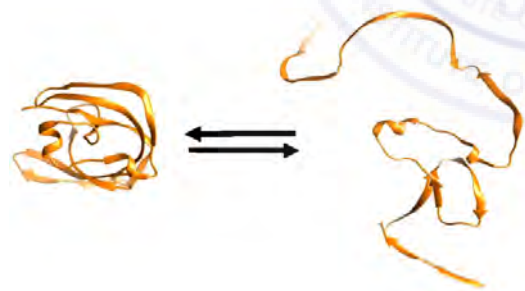


change in local coordinates of the amino acid residues, it is apposite to state that this phenomenon will differ domainwise. This domainwise disparity of conformational dynamics is believed to play a pivotal role in the domain specific activity of a protein.

The conformational fluctuations are not synchronized for an ensemble of protein molecules, due to which the experimental realization of the timescale remains difficult. A variety of theoretical methods based on molecular dynamics simulation and some experimental methods based on nuclear magnetic resonance (NMR), mass spectrometry, fluorescence measurements *etc.* has been practiced to scrutinize conformational dynamics. In recent times, single molecule measurements have emerged as a fine technique for this motive as it successfully takes care of the desynchronicity of the fluctuation.

#### 1.4. Protein Denaturation

Under adverse conditions, the stabilizing interactions which lead to protein folding are broken leading to breakdown of the native state. The opening up of the protein molecule in response to the exposure to these conditions is called protein denaturation or unfolding (scheme 1.6).<sup>38-45</sup> According to a review by Hans Neurath in 1944 the term denaturation has been defined as “any non-proteolytic modification of



**Scheme 1.6:** Protein denaturation.

the unique structure of a native protein, giving rise to definite changes in chemical, physical, or biological properties”.<sup>38</sup> The process of denaturation is as complex as the process of folding. It is believed that the protein unfolding follows the reverse pathway as compared

to folding and hence the thorough understanding of the unfolding pathway may lead us to the understanding of the folding pathway. Some proteins denature following simple two state model while some involve intermediate state(s) in the course of denaturation. Most of the reports on protein denaturation till date are based on induced unfolding where external agents such as chemical denaturant, high temperature and pressure *etc.* are applied to denature the protein. Although both temperature and

chemical agitation trigger unfolding, their action on the protein might be significantly different. Several investigations have been performed to shed light on the similarities and dissimilarities between the pathways that a protein uptakes in the course of chemical and thermal denaturation. It is reported that the thermally and chemically denatured states are different for most proteins.<sup>46,47</sup> It is also examined that these external agents affect aggregation and refolding in notably different ways.<sup>48</sup> The mechanism behind the denaturing capability of these agents must be better understood to grab a clear idea about the two processes.

**1.4.1. Chemical Denaturation:** This group of denaturants include guanidine hydrochloride (GnHCl), urea, surfactants, ionic liquids *etc.*<sup>49-53</sup> For the chemical denaturation cases, at a constant temperature, the Gibbs free energy change ( $\Delta G$ ) follow a linear dependence with the denaturant concentration<sup>45</sup>

$$\Delta G = \Delta G^0 - m[\text{denaturant}] \quad (\text{Eq. 1.2})$$

where  $m$  denotes the slope of change of  $\Delta G$ . Taking  $G^f$  and  $G^u$  to be the energies of the folded and unfolded states of the protein, the Gibbs free energy change for the folding-unfolding equilibrium can be written as,

$$\Delta G = G^u - G^f \quad (\text{Eq. 1.3})$$

The value of  $m$  reflects the denaturation ability of the chemical denaturant and may significantly vary for different proteins. Along with this, the pathway as well as the mechanism of chemical denaturation highly depends on the nature of the denaturant. For example, urea is non polar whereas GnHCl is ionic so their course of action is expected to be different. There are reports which say that they follow same mechanism and reports which state otherwise.<sup>41,45,53</sup> In 2009, Lim *et al.* reported that the formation of hydrogen bonds between urea and the hydrogen and oxygen atom of the amide group is the root of its denaturing ability.<sup>53</sup> In the same report they mentioned that guanidine does not have similar hydrogen bonding capability. As per a report by Meuzelaar *et al.* in 2015, GnHCl breaks down protein's secondary structure by breaking the salt bridges which has significant contribution in the secondary structure of protein.<sup>49</sup> Thus no

generalized model can be derived for the chemical denaturation of protein and a detailed scrutiny is required to arrive at a conclusion about a specific denaturant.

**1.4.2. Thermal Denaturation:** The temperature dependence of the equilibrium existing between folded and unfolded states of proteins has been explained thermodynamically, by considering the effect of temperature on the enthalpy and entropy change of the equilibrium.<sup>54</sup>

At the transition temperature between folded to unfolded state ( $T_m$ ), both the states have equal energy. Hence,

$$\Delta G(T_m) = 0 \quad (\text{Eq. 1.4})$$

$$\Delta H(T_m) - T_m \Delta S(T_m) = 0 \quad (\text{Eq. 1.5})$$

where  $\Delta H(T_m)$  is the enthalpy change and  $\Delta S(T_m)$  is the entropy change at  $T_m$ . Thermodynamically, for a system with heat capacity change at constant pressure  $\Delta C_p$  the enthalpy change can be written as a function of temperature,

$$\Delta H(T) = \Delta H(T_m) + \int_{T_m}^T \Delta C_p(T) dT \quad (\text{Eq. 1.6})$$

Considering  $\Delta C_p$  to be independent of temperature, equation 1.6 can be simplified to

$$\Delta H(T) = \Delta H(T_m) + (T - T_m) \Delta C_p \quad (\text{Eq. 1.7})$$

And the temperature dependent entropy change can be written as,

$$\Delta S(T) = \Delta S(T_m) + \int_{T_m}^T \frac{\partial \Delta S(T)}{\partial T} dT \quad (\text{Eq. 1.8})$$

$$\begin{aligned} \Delta S(T) &= \frac{\Delta H(T_m)}{T_m} + \int_{T_m}^T \Delta C_p(T) d(\ln T) \\ &= \frac{\Delta H(T_m)}{T_m} + \Delta C_p \ln \frac{T}{T_m} \end{aligned} \quad (\text{Eq. 1.9})$$

The temperature dependence of Gibb's free energy of unfolding can now be written by combining equations 1.7 and 1.9,

$$\Delta G(T) = \Delta H - T \Delta S$$

$$= \frac{(T_m - T)}{T_m} \Delta H(T_m) + (T - T_m) \Delta C_p - T \Delta C_p \ln \frac{T}{T_m} \quad (\text{Eq. 1.10})$$

Equation 1.10 shows that  $\Delta G$  depends on  $T_m$ ,  $\Delta H(T_m)$  and  $\Delta C_p$ . The  $\Delta G$  vs  $T$  plot is convex shaped giving two  $\Delta G=0$  points indicating two transition temperatures between folded and unfolded structures. Solving for  $\Delta G=0$  from equation 1.10 using Taylor series expansion till second order for the logarithmic term, we can obtain two solutions:

$$T_1 = T_m \quad (\text{Eq. 1.11})$$

$$T_2^2 \approx \frac{T_m^2 \Delta C_p}{2\Delta H(T_m) + T_m \Delta C_p} \quad (\text{Eq. 1.12})$$

The hot denaturation occurs for  $T = T_m$  and for most proteins the cold denaturation temperature is calculated to be much below the freezing point of water. This temperature increases when the convexity of the curve is increased i.e. the curvature of the plot given by  $\frac{\partial^2 G}{\partial T^2} = -\frac{\Delta C_p}{T}$  becomes more negative. For large globular or long chain proteins with highly positive  $\Delta C_p$  and  $T_m$  values, the cold denaturation temperature increases above zero degrees and can be accessed experimentally. Weakening of the folded structure by other denaturants like pressure, chemical also contribute towards making the curve more convex and hence increasing the cold denaturation temperature.

### 1.5. Protein Aggregation

Most of the proteins fold to their native conformation after their synthesis. As already discussed, this folding process is driven by the hydrophobic, hydrophilic and other non-covalent interactions. In certain cases, these interactions are interfered by change in amino acid sequence or change in environment which leads to misfolding of the proteins.<sup>55-62</sup> Besides this, there are IDPs which are disordered due to intrinsic genetic mutations and have minimum stabilizing interactions present. This misfolded or disordered protein is not stable as native state and tends to form aggregates by stabilizing interactions by individual units. The aggregation of proteins, in general, is detrimental for biological systems and some of the common outcomes of protein



aggregation in human body is Parkinson's disease (aggregation of  $\alpha$ -synuclein), Alzheimer's disease (aggregation of short amyloid- $\beta$  proteins), Huntington's disease (aggregation of Huntington's protein). Protein aggregation depends on the external environment such as pH, ionic strength, temperature *etc* and can be triggered by the formation of unfolding intermediates and unfolded states or direct interaction between different molecules in the native state or chemical interactions.<sup>61</sup> Both reversible or irreversible aggregation processes are reported. Usually, early stages of aggregation is reversible and late stages are irreversible. Protein aggregation is vastly studied by researchers due to its important effect on human health conditions. The protocols which are used for this purpose are well described in a review article by Mahler *et al.*<sup>62</sup>

#### 1.6. Effect of Confinement on Proteins

Upon exposure to environments deviating from physiological conditions, a protein undergoes structural deformations in order to cope with the changed milieu. *In vivo*, the proteins reside inside or outside the cell, depending on which they are classified as intracellular or extracellular proteins. Additionally, there are membrane proteins, which are found in the lipid bilayer of cell membrane. The cellular proteins are always incarcerated inside the cell and in there their behaviour may significantly be altered as compared to when they are free. The major effect of encapsulation of proteins in cells is the limitation of its degrees of freedom and the alteration of the neighbouring solvent atmosphere. Besides, in cellular matrix, there are other chemical species present which effectively interact with the protein. This interaction is known as the background interaction. The behaviour of biological water (the water surrounding the biomolecules, in this case proteins) is unlike inside and outside cells. All these factors together contribute in the changed folding/unfolding behaviour of proteins.<sup>63,67</sup> The dynamics of protein inside confined environment is also found to be significantly different than that of the protein in diluted solution.<sup>63-70</sup> For experimental convenience, most of the *in vitro* studies of proteins has been carried out in bulk buffer solutions. Hence, the outcome of the experiments provides us information about the protein only and not the information about the protein in cellular environment. For this

reason, recently nanoporous and micellar systems are dragging attention as confined environments to better understanding of proteins in cells.<sup>63-67</sup>

### 1.7. Digestion of Proteins

Digestion of proteins is the tearing of proteins into shreds resulting from the cleavage of peptide bonds. The breakdown of protein into peptides is also referred to as proteolysis.<sup>71-74</sup> A healthy digestive system is responsible for the physical wellbeing of an individual and when disturbed, it may cause severe problems including stomach upset, diarrhoea, deficiency of nutrients *etc.* Additionally, the protein digestion ensures the supply of requisite amount of essential amino acids in human body. Apart from *in vivo* health benefits, protein digestion is a necessary step in the protein identification and characterization protocol. In human body this process takes place in small and large intestine and is efficiently catalyzed by enzymes known as peptidases or proteases and in absence of peptidases, the proteolysis is very slow. These enzymes are also proteins structurally and they are secreted from salivary glands, secretory cells in stomach, pancreas and small intestine. The working principle of different peptidases varies widely depending on the mechanism, cleaving site working conditions *etc.* The presence of active site is a common feature of almost all the proteases. Depending on the active site residues the proteases can be classified as serine protease, cysteine protease *etc.* Some enzymes choose their cleavage site based on their locations while some prefer specific amino acid residues over others. The first group is called exopeptidase whereas the later is known as endopeptidase. Exopeptidases sequentially cut small fragments from the terminals containing  $\leq$  three amino acid residues and can be further classified to carboxypeptidase and aminopeptidase depending on their preference of C-terminal and N-terminal side. Due to digestion the protein structure and function is totally lost. The mechanism of digestion varies depending on the enzyme.

### 1.8. Protein Spectroscopy

Typically all proteins show absorption bands in the UV region.<sup>75-78</sup> The near UV peaks at 280 nm are generated from the absorption of tyrosine and tryptophan.<sup>75</sup>

The phenylalanine residues absorb with a peak at 250 nm. The far UV peaks ~190 nm are generated from the absorption of peptide bonds. In water medium the tryptophan, tyrosine and phenylalanine fluoresce with bands centred at 350 nm, 303 nm and 282 nm when excited at respective absorption maxima.<sup>75</sup> In practical cases to measure protein fluorescence, samples are excited at 295 nm to record the fluorescence coming from tryptophan only as at this wavelength, the absorption of tyrosine and phenylalanine is close to zero. Most of the proteins contain ~1 mol% tryptophan residue owing to the high metabolic expense of its synthesis. Presence of less number of tryptophan residue reduces the complexity of the emission spectra and facilitates the spectroscopic investigation of the protein. At 296 K, the quantum yield of tryptophan is 0.14 in neutral aqueous solution.<sup>75</sup> The emission maxima, quantum yield and bandwidth of tryptophan fluorescence of a protein is governed by the location of tryptophan residue in the protein matrix. A tryptophan residue which is buried inside hydrophobic protein core will have a blue shifted peak with a narrow bandwidth as compared to the tryptophan at surface. The environmental dependence of tryptophan fluorescence on the local environment can be utilized to study protein folding, unfolding and binding events.<sup>79-82</sup>

Apart from this general behaviour of proteins, green fluorescent proteins (GFP) which are a special group of proteins, exhibit bright fluorescence when exposed to blue violet light. They are found in the many of the marine organisms but the term GFP usually refers to the fluorescent protein found in jellyfish.

### 1.9. Protein Tagging

Although most of the proteins contain one or more tryptophan residues, the intrinsic tryptophan fluorescence is not sufficient always to explore protein folding. Domains lacking any tryptophan residue on multidomain proteins require attachment of a fluorescent tag for domain specific study of proteins. Besides, tryptophan being a weakly fluorescent one does not suffice for fluorescence imaging and single molecular fluorescence experiments and hence proteins must be stained with highly fluorescent dyes for the purpose. These external tagging agents are known as extrinsic fluorescence labels. When this dye attaches to a specific amino acid residue it is called site specific

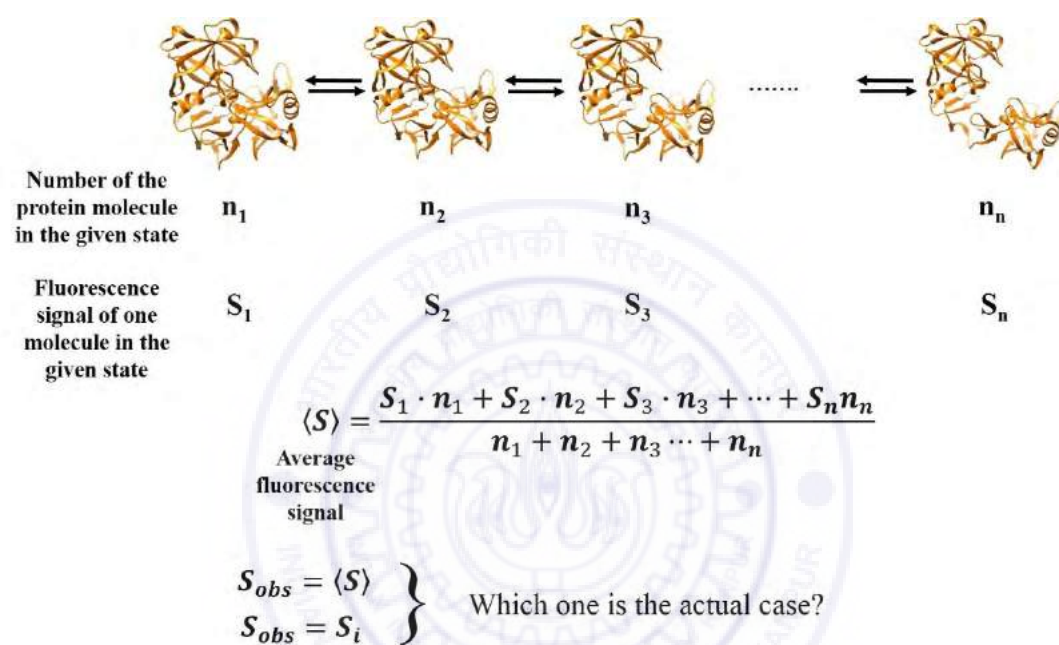
fluorescence labelling of protein. Site specific fluorescence labelling facilitates the study of local environment in proteins because of the dependence of photophysics of the dye on the neighbouring environment. There are two types of extrinsic fluorescent tags based on the mode of attachment to proteins: covalent protein tags which form proper chemical bonds with specific amino acid residue and non-covalent tags which bind *via* weak interactions. Covalent protein tags are preferred over the other one due to the fact that non covalent ones can get detached from the protein upon alteration of solution conditions. Covalent protein tags, based on their functionalization, may attach to cysteine (thiol reactive fluorescent label), lysine or N-terminal amino acid (amine reactive probe), C-terminal residue (carboxyl specific probe) or tyrosines (hydroxyl specific probe).<sup>83-92</sup> Most of the commonly used fluorescent labels contain coumarin or rhodamine as the chromophore. The fluorescent labels which show solvent dependent spectroscopic behaviour are known as solvatochromic labels and are widely used to unveil structural information of proteins.<sup>88-92</sup>

#### **1.10. Single Molecular Spectroscopy of Proteins**

Spectroscopic techniques are the most common and useful tool to understand behaviour of proteins. Fluorescence signal of the tryptophan residues has been studied to understand the folding/unfolding and dynamics of proteins.<sup>79-82</sup> Additionally, fluorescence signal from external fluorescence tags has also been utilized for the same purpose.<sup>83-92</sup> Nuclear magnetic resonance spectroscopy is proven to be utilitarian in studying dynamics of protein.<sup>93-95</sup> The secondary structure of the protein can be efficiently probed by circular dichroism spectroscopy.<sup>97, 98</sup>

Although all these techniques have served the purpose of protein research for a long time, they only produce ensemble-average result. This information is sufficient when all the molecules in a protein solution have the same spectroscopic property. But this condition is not always valid. All the protein molecules in a solution may not have the same structural features at a given point of time. This phenomenon renders the study of proteins difficult with respect to structural features. As already discussed, the fluorescence signal of a protein highly depends on the environment around the

fluorophore. So the signal will be different for different molecules in an inhomogeneous solution. Practically in a one millilitre one micromolar protein solution, there are  $\sim 10^{14}$  number of molecules. These molecules can have several different conformations. Recording the fluorescence of this solution at any specific instance will not deliver any information about the population of each conformational state (scheme 1.7). While seeing each molecule individually helps commenting on the distribution. Besides, the unsynchronized breathing motion or the conformational



**Scheme 1.7:** Limitations of ensemble experiments.

dynamics of the protein molecules make the protein solution heterogeneous, i.e. they continuously undergo transition between closely related conformational states. Single molecular level measurement is also advantageous in fetching information about this dynamics.

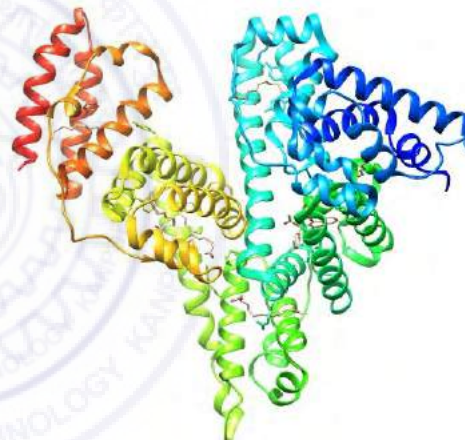
There are several techniques based on single molecule fluorescence measurements, namely single molecular fluorescence imaging, fluorescence correlation spectroscopy, fluorescence lifetime correlation spectroscopy *etc.*<sup>100-104</sup> The technique that has been used for current study is fluorescence correlation spectroscopy (FCS). In FCS, fluorescence fluctuation is monitored and utilized to have knowledge about the size parameters and dynamics of the fluorescent particle. The foundation work on FCS was carried out on DNA on 1972.<sup>105</sup> Only after the invention of confocal



microscopy in late 1990s people have started using FCS extensively. Bioresearch relies on FCS due to its noninvasive property and vast range of information that it may provide. P. Schwille and coworkers have extensively used FCS to study the dynamics of several proteins and other biosystems.<sup>106-108</sup> J. È. Langowski, W. W. Webb, R. Rigler and other groups have also performed FCS experiments on several proteins.<sup>108-114</sup> Apart from protein dynamics, FCS has also been applied to investigate the transport of proteins, oligomerization, protein-protein interaction and interaction of proteins with other molecules.<sup>106-122</sup> The theory and instrumentation of FCS have been discussed in the next chapter in details.

### 1.11. Proteins under Investigation

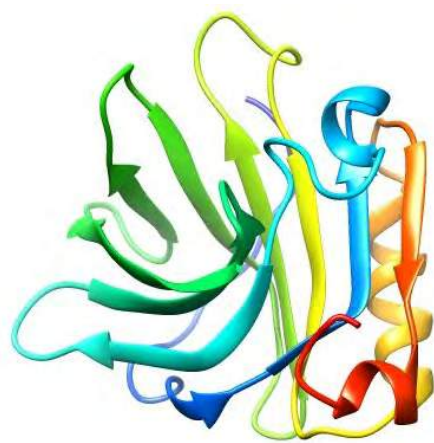
**1.11.1. Human Serum Albumin:** Human serum albumin (HSA) is a multi-domain transport protein (molecular weight = 66.5 kD) present in blood plasma having 585 amino acid residues and 17 disulfide bridges (scheme 1.8).<sup>123-128</sup> The duty of HSA is to carry drug and metabolites to different parts of human body. There is one free cysteine residue at position 34 which can be tagged with thiol specific dyes.<sup>89-91,123</sup> HSA has 18 tyrosine residues out of which the most reactive tyr-411 can be tagged specifically



**Scheme 1.8:** Native state structure of human serum albumin (PDB ID: 1HA2).

with hydroxyl reactive molecules. Crystallographic characterization of HSA disclosed the existence of three structurally similar  $\alpha$ -helical domains, each which are further divided into two sub-domains.<sup>123-128</sup> Each of these domains has different binding properties. The two main binding site namely Sudlow's site I and Sudlow's site II are located at subdomain IIA and subdomain IIIA respectively.<sup>123-125</sup> Sudlow's site I binds bulky heterocyclic anions whereas Sudlow's site II binds aromatic carboxylates.<sup>129, 130</sup> Thus these two subdomains are of immense importance for the biological activity of HSA. Domain I of HSA binds fatty acids and heme and helps in the transport.<sup>131, 132</sup>

**1.11.2.  $\beta$ -lactoglobulin:**  $\beta$ -lactoglobulin is a whey protein belonging to lipocalin family which is found in major quantity in cow and goat milk (3g/l).<sup>133-138</sup> The exact



**Scheme 1.9:** Native state structure of  $\beta$ -lactoglobulin (PDB ID: 5IO6).

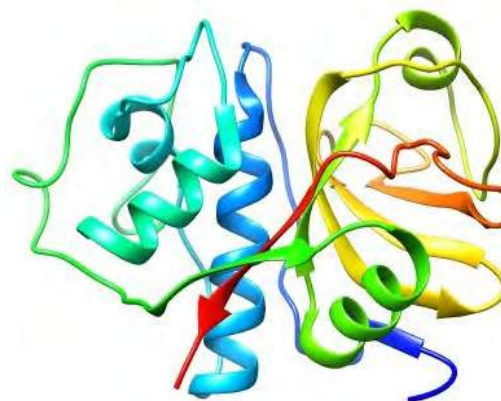
function of  $\beta$ -lactoglobulin is still unknown but is believed to work as a transport protein and binds fatty acid, cholesterol, vitamin D<sub>2</sub> and some other small molecules.<sup>133-136</sup> There are five cysteine residues in  $\beta$ -lactoglobulin located at 66<sup>th</sup>, 106<sup>th</sup>, 119<sup>th</sup>, 121<sup>st</sup>, 160<sup>th</sup> position of the peptide chain, among them Cys-66 forms a disulphide bond with Cys-160 and Cys-106 is bonded to Cys-119. The Cys-121 has a free thiol group, which can be tagged using thiol

reactive fluorophores like N-dansylaziridine, bodipy-derivatives, 2-(4'-Maleimidylanilino) naphthalene-6-sulfonate, *etc.* as reported by several groups.<sup>140-143</sup>  $\beta$ -lactoglobulin also works as a pathogen owing to its ability to bind iron via siderophores.<sup>136</sup> It is normally absent in human breast milk but regular consumption of bovine milk induce the presence in marginal amount.<sup>144, 145</sup> Its ability to initiate allergic reaction to ~2% of infants who consume it makes it the most common food allergen. Thirteen genetic variants of  $\beta$ -lactoglobulin with slightly different amino acid sequence are known among which the most common are  $\beta$ -lactoglobulin A and  $\beta$ -lactoglobulin B. It is a small globular protein containing 162 amino acid residues with a molecular weight of 18.4 kDa (scheme 1.9). Structurally, it contains nine anti parallel  $\beta$  sheets (labelled as A-I) and one  $\alpha$ -helix.<sup>138</sup> The monomeric, dimeric, tetrameric, octameric and multimeric forms of  $\beta$ -lactoglobulin is reported under different atmosphere.<sup>146</sup>

**1.11.3. Papain:** Papain is a cysteine protease enzyme present in papaya and known to help in digesting peptide.<sup>147-154</sup> It breaks the protein fibers into smaller fragments (small peptide or even amino acids) by cleaving the peptide linkage.<sup>147</sup> There are 11 papain like cysteine protease which are part of the human digestive system. Apart from this major enzymatic property, papain is also believed to have significant healing properties and is used to treat burns, irritations, wounds, ulcers and bedsores.<sup>155, 156</sup> The complete

amino acid sequence of papain has been elucidated by Mitchel *et al.* in the year 1970, while the three-dimensional structure of papain has been described at 2.8 Å resolution by Drenth *et al.* in 1968.<sup>148,150</sup>

Structurally, papain contains 21%  $\beta$ -sheet and 25%  $\alpha$ -helix, and the rest is the ordered non-repetitive structure.<sup>147,150</sup> It has a molecular weight of 23.4 kDa and contains a single peptide chain of 212 amino acids structurally divided into two domains namely L and R, which are



**Scheme 1.10:** Native state structure of papain (PDB ID: 9PAP).

The tertiary structure of papain consists of three disulfide bonds, which connect Cys-22 to Cys-63, Cys-56 to Cys-95, and Cys-153 to Cys-200. Domain R consists of residue 1-10 and 112-208, while rest of the residues are present in domain L.<sup>150</sup> Crystallographic data reveals that cysteine-25, which is the only free cysteine residue of papain and histidine-159 are located in the cleft between the two domains and they are mainly responsible for the enzymatic activity of papain.<sup>151</sup>

## References

1. Cox, M. M.; Nelson, D. L. *Lehninger, Principles of Biochemistry*, 5<sup>th</sup> Ed. W. H. Freeman Co. NY, **2008**.
2. Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 5<sup>th</sup> Ed. W.H. Freeman Co. NY, **2002**.
3. Berezovsky, I. N.; Guarnera, E.; Zheng, Z. *Prog. Biophys. Mol. Biol.* **2017**, *128*, 85-99.
4. Orengo, C. A.; Michie, A. D.; Jones, S.; Jones, D. T.; Swindells, M. B.; Thornton, J. M. *Structure* **1997**, *5*, 1093–1109.
5. Chothia, C. *Nature* **1992**, *357*, 543–544.
6. Ponting, C. P.; Russell, R. R. *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 45–71.
7. Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167–339.
8. Anfinsen, C. B. *Science* **1973**, *181*, 223–230.
9. Anfinsen, C. B.; Haber, E.; Sela, M.; White, J. F. H. *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1309–1314.
10. Daggett, V.; Fersht, A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 497–502.
11. Dill, K. A.; MacCallum, J. L. *Science* **2012**, *338*, 1042–1046.
12. Rose, G. D.; Fleming, P. J.; Banavar, J. R.; Maritan, A. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16623–16633.
13. Dobson, C. M.; Šali, A.; Karplus, M. *Angew. Chem. Int. Ed.* **1998**, *37*, 868–893.
14. Englander, S. W.; Mayne, L. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15873–15880.
15. Dobson, C. M. *Nature* **2003**, *426*, 884–890.
16. Betz, S. F. *Protein Science : A Publication of the Protein Society* **1993**, *2*, 1551–1558.
17. Creighton, T. E. *Biochem. J.* **1990**, *270*, 1–16.
18. Shea, J. E.; Brooks Iii, C. L. *Ann. Rev. Phys. Chem.* **2001**, *52*, 499–535.
19. Levitt, M. *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 251–271.
20. Chouard, T. *Nature* **2011**, *471*, 151–153.



21. Privalov, P. L. *J. Chem. Thermodyn.* **1997**, 29, 447–474.
22. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z. *J. Mol. Graph. Mod.* **2001**, 19, 26–59.
23. Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell. Biol.* **2005**, 6, 197–208.
24. Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L. *Curr. Opin. Struct. Biol.* **2008**, 18, 756–764.
25. Clark, P. L. *Trends in Biochem. Sci.* **2004**, 29, 527–534.
26. Schultz, C. P. *Nat. Struct. Biol.* **2000**, 7, 7–10.
27. Levinthal, C. *J. Chem. Phys.* **1968**, 65, 44–45.
28. Kahsai, A. W.; Rajagopal, S.; Sun, J.; Xiao, K. *Nat. Protocols* **2014**, 9, 1301–1319.
29. Guo, J.; Zhou, H.-X. *Chem. Rev.* **2016**, 116, 6503–6515.
30. Gangupomu, Vamshi K.; Wagner, Jeffrey R.; Park, I.-H.; Jain, A.; Vaidehi, N. *Biophys. J.* **2013**, 104, 1999–2008.
31. Forest, E.; Man, P. Mus–Veteau, I. Springer New York: New York, NY, **2016**, 1432, 269–279.
32. Shi, Z.; Castro, C. E.; Arya, G. *ACS Nano* **2017**, 11, 4617–4630.
33. Torres, T.; Levitus, M. *J. Phys. Chem. B* **2007**, 111, 7392–7400.
34. Weiss, S. *Nat. Struct. Mol. Bio.* **2000**, 7, 724–729.
35. Moroni, L.; Gellini, C.; Salvi, P. R. *World J. Chem. Educ.* **2015**, 3, 59–63.
36. Zhang, F.; Chen, N.; Wu, R. *J. Chem. Inf. Model.* **2016**, 56, 877–885.
37. Seo, M.-H.; Park, J.; Kim, E.; Hohng, S.; Kim, H.S. *Nat. Commun.* **2014**, 5, 3724(1)1–3724(7).
38. Neurath, H.; Greenstein, J. P.; Putnam, F. W.; Erickson, J. A. *Chem. Rev.* **1944**, 34, 157–265.
39. Dinner, A. R.; Šali, A.; Smith, L. J.; Dobson, C. M.; Karplus, M. *Trends in Biochemical Sciences* **2000**, 25, 331–339.
40. Loughlin, W. J. *Biochem. J.* **1932**, 26, 1557–1565.



41. Camilloni, C.; Rocco, A. G.; Eberini, I.; Gianazza, E.; Broglia, R. A.; Tiana, G. *Biophys. J.* **2008**, *94*, 4654–4661.
42. Kishore, D.; Kundu, S.; Kayastha, A. M. *PLoS One* **2012**, *7*, e50380(1)–e50380(9).
43. Daggett, V.; Levitt, M. *J. Mol. Biol.* **1993**, *232*, 600–619.
44. Malecki, J.; Wasylewski, Z. *J. Prot. Chem.* **1998**, *17*, 745–755.
45. Wang, Q.; Christiansen, A.; Samiotakis, A.; Wittung–Stafshede, P.; Cheung, M. S., *J. Chem. Phys.* **2011**, *135*, 175102(1)–175102(12).
46. Ramprakash, J.; Dosseva, V.; Galkin, A.; Krajewski, W.; Muthukumar, L.; Pullalarevu, S.; Demirkan, E.; Herzberg, O.; Moulton, J.; Schwarz, F. P. *Anal. Biochem.* **2008**, *374*, 221–230.
47. Anand, U.; Mukherjee, S. *Biochim. Biophys. Acta* **2013**, *1830*, 5394–5404.
48. Kishore, D.; Kundu, S.; Kayastha, A. M. *PLoS One* **2012**, *7*, e53080.
49. Meuzelaar, H.; Panman, M. R.; Woutersen, S. *Angew. Chem. Int. Ed.* **2015**, *54*, 15255–15259.
50. Watlafer, D. B.; Malik, S. K.; Stoller, L.; Coffin, R. L. *J. Am. Chem. Soc.* **1964**, *86*, 508–514.
51. Tanford, C. *J. Am. Chem. Soc.* **1964**, *86*, 2050–2059.
52. Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2462–2470.
53. Lim, W. K.; Rösger, J.; Englander, S. W. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2595–2600.
54. Dias, C. L.; Ala–Nissila, T.; Wong–ekkabut, J.; Vattulainen, I.; Grant, M.; Karttunen, M. *Cryobiology* **2010**, *60*, 91–99.
55. Aguzzi, A.; O'Connor, T. *Nat. Rev. Drug. Discov.* **2010**, *9*, 237–248.
56. Stefani, M.; Dobson, C. M. *J. Mol. Med.* **2003**, *81*, 678–699.
57. De Felice, F. G.; Vieira, M. N.; Meirelles, M. N.; Morozova–Roche, L. A.; Dobson, C. M.; Ferreira, S. T. *FASEB J.* **2004**, *18*, 1099–101.
58. Tanzi, R. E.; Bertram, L. *Cell* **2005**, *120*, 545–555.
59. Moussa, C. E. H.; Mahmoodian, F.; Tomita, Y.; Sidhu, A. *Biochem. Biophys. Res. Comm.* **2008**, *365*, 833–839.
60. LaFerla, F. M.; Green, K. N.; Oddo, S. *Nat. Rev. Neurosci.* **2007**, *8*, 499–509.

61. Wang, W.; Nema, S.; Teagarden, D. *Int. J. Pharm.* **2010**, *390*, 89–99.
62. Mahler, H.-C.; Friess, W.; Grauschopf, U.; Kiese, S. *J. Pharm. Sci.* **2009**, *98*, 2909–2934.
63. Martinez, A. V.; DeSensi, S. C.; Dominguez, L.; Rivera, E.; Straub, J. E. *J. Chem. Phys.* **2011**, *134*, 055107(1)–055107(9).
64. Zhao, X.; Ao, Q.; Chen, F.; Xue, W.; Li, L.; Liu, J. *Food Chemistry* **2009**, *116*, 176–182.
65. Melo, E. P.; Fojan, P.; Cabral, J. M. S.; Petersen, S. B. *Chem. Phys. Lipids* **2000**, *106*, 181–189.
66. Kumbhakar, M.; Goel, T.; Mukherjee, T.; Pal, H. *J. Phys. Chem. B* **2004**, *108*, 19246–19254.
67. Lucent, D.; Vishal, V.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10430–10434.
68. Eggers, D. K.; Valentine, J. S. *Protein Science : A Publication of the Protein Society* **2001**, *10*, 250–261.
69. Zhou, H.-X.; Dill, K. A. *Biochemistry* **2001**, *40*, 11289–11293.
70. Rathore, N.; Knotts, T. A. I. V.; de Pablo, J. J. *Biophys. J.* **2006**, *90*, 1767–1773.
71. Freeman, H. J.; Kim, Y. S. *Annu. Review of Medicine* **1978**, *29*, 99–116.
72. Vahdatpour, S.; Mamaghani, A. P.; Goloujeh, M. S.; Sis, N. M.; Mahmoodpour, H.; Vahdatpour, T. *Electr. J. Biol.* **2016**, *12*, 265–275.
73. Spellman, D.; McEvoy, E.; O’Cuinn, G.; FitzGerald, R. J. *Int. Dairy J.* **2003**, *13*, 447–453.
74. Holmes, J. H. G.; Bayley, H. S.; Leadbeater, P. A.; Horney, F. D. *Br. J. Nutr.* **2007**, *32*, 479–489.
75. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd Ed.; Springer: NY, **2006**.
76. Goldfarb, A. R.; Saidel, L. J.; Mosovich, E. *J. Biol. Chem.* **1951**, *193*, 397–404.
77. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Prot. Sci.* **1995**, *4*, 2411–2423.
78. Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319–326.
79. Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093–2109.

80. Chen, Y.; Barkley, M. D. *Biochemistry* **1998**, *37*, 9976–9982.
81. Ghisaidoobe, A. B. T.; Chung, S. J. *Int. J. Mol. Sci.* **2014**, *15*, 22518–22538.
82. Royer, C. A. *Chem. Rev.* **2006**, *106*, 1769–1784.
83. Strack, R. *Nat. Meth.* **2016**, *13*, 33.
84. Romanini, D. W.; Cornish, V. W. *Nat. Chem.* **2012**, *4*, 248–250.
85. Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786–791.
86. Toseland, C. P. *J. Chem. Biol.* **2013**, *6*, 85–95.
87. Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. *Nat. Chem.* **2016**, *8*, 103–113.
88. Klymchenko, A. S. *Acc. Chem. Res.* **2017**, *50*, 366–375.
89. Yadav, R.; Sengupta, B.; Sen, P. *Biophys. Chem.* **2016**, *211*, 59–69.
90. Flora, K.; Brennan, J. D.; Baker, G. A.; Doody, M. A.; Bright, F. V. *Biophys. J.* **1998**, *75*, 1084–1096.
91. Sasmal, D. K.; Mondal, T.; Mojumdar, S. S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
92. Gerasov, A.; Shandura, M.; Kovtun, Y.; Losytskyy, M.; Negrutska, V.; Dubey, I. *Anal. Biochem.* **2012**, *420*, 115–120.
93. Sapienza, P. J.; Lee, A. L. *Curr. Opin. Pharmacol.* **2010**, *10*, 723–730.
94. Kleckner, I. R.; Foster, M. P. *Biochim. Biophys. Acta* **2011**, *1814*, 942–968.
95. Ishima, R.; Torchia, D. A., *Nat. Struct. Mol. Biol.* **2000**, *7*, 740–743.
96. Kay, L. E. *J. Magn. Res.* **2005**, *173*, 193–207.
97. Greenfield, N. J. *Nat. Prot.* **2006**, *1*, 2876–2890.
98. Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y.-H.; Goto, Y.; Réfrégiers, M.; Kardos, J. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3095–E3103.
99. Sherman, E.; Itkin, A.; Kuttner, Y. Y.; Rhoades, E.; Amir, D.; Haas, E.; Haran, G. *Biophys. J.* **2008**, *94*, 4819–4827.
100. Basak, S.; Chattopadhyay, K. *Phys. Chem. Chem. Phys.* **2014**, *16*, 11139–11149.
101. Lu, H. P. *Protein–Ligand Interactions: Methods and Applications*, Ulrich Nienhaus, G., Ed. Humana Press: Totowa, NJ, **2005**; pp 385–413.
102. Schuler, B. J. *Nanobiotechnol.* **2013**, *11*, S2(1)–S2(17).

103. Godin, Antoine G.; Lounis, B.; Cognet, L. *Biophys. J.* **2014**, *107*, 1777–1784.
104. Petrášek, Z.; Hoege, C.; Mashaghi, A.; Ohrt, T.; Hyman, A. A.; Schwille, P. *Biophys. J.* **2008**, *95*, 5476–5486.
105. Magde, D.; Elson, E.; Webb, W. W. *Phys. Rev. Lett.* **1972**, *29*, 705–708.
106. Chiantia, S.; Ries, J.; Schwille, P. *Biochim. Biophys. Acta (BBA) -Biomembranes* **2009**, *1788*, 225–233.
107. Heinemann, F.; Betaneli, V.; Thomas, F. A.; Schwille, P. *Langmuir* **2012**, *28*, 13395–13404.
108. Betaneli, V.; Schwille, P. *Lipid-Protein Interactions: Methods and Protocols*, Kleinschmidt, J. H., Ed. Humana Press: Totowa, NJ, **2013**; pp 253–278.
109. Kohler, R. H.; Schwille, P.; Webb, W. W.; Hanson, M. R. *J. Cell Sci.* **2000**, *113*, 3921–3930.
110. Haupts, U.; Maiti, S.; Schwille, P.; Webb, W. W. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13573–13578.
111. Renz, M.; Langowski, J. *Chromosome Res.* **2008**, *16*, 427–437.
112. Lagerkvist, A. C.; Földes-Papp, Z.; Persson, M. A. A.; Rigler, R. *Protein Science: A Publication of the Protein Society* **2001**, *10*, 1522–1528.
113. Bahns, J. T.; Liu, C.-M.; Chen, L. *Protein Science: A Publication of the Protein Society* **2004**, *13*, 2578–2587.
114. Zhang, D.; Lans, H.; Vermeulen, W.; Lenferink, A.; Otto, C. *Biophys. J.* **2008**, *95*, 3439–3446.
115. Chattopadhyay, K.; Saffarian, S.; Elson, E. L.; Frieden, C. *Biophys. J.* **2005**, *88*, 1413–1422.
116. Michalet, X.; Weiss, S.; Jäger, M. *Chem. Rev.* **2006**, *106*, 1785–1813.
117. Chowdhury, P. K. *J. Prot. Proteomics* **2013**, *2*, 145–169.
118. Rusu, L.; Gambhir, A.; McLaughlin, S.; Rädler, J. *Biophys. J.* **2004**, *87*, 1044–1053.
119. Hess, S. T.; Huang, S.; Heikal, A. A.; Webb, W. W. *Biochemistry* **2002**, *41*, 697–705.
120. Edman, L. *J. Phys. Chem. A* **2000**, *104*, 6165–6170.
121. Levitus, M. *J. Phys. Chem. Lett.* **2010**, *1*, 1346–1350.

122. Rigler, R.; Grasselli, P.; Ehrenberg, M. *Physica Scripta* **1979**, *19*, 486–490.
123. He, X. M., and Carter, D. C. *Nature* **1992**, *358*, 209–215.
124. Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1975**, *11*, 824–832.
125. Peters, T. Jr. *Adv. Protein Chem.* **1985**, *37*, 161–245.
126. Peters, T. Jr. *All about albumin: Biochemistry, Genetics, and Medical Applications*, Academic press, San Diego, **1996**.
127. Curry, S.; Mandelkow, H.; Brick, P.; Franks, N. *Nat. Struc. Biol.* **1998**, *5*, 827–835.
128. Sugio, S.; Kashima, A.; Mochizuki, S.; Noda, M.; Kobayashi, K. *Protein Eng.* **1999**, *12*, 439–446.
129. Kragh-Hansen, U.; Chuang, V. T. G.; Otagiri, M. *Biol. Pharm. Bull.* **2002**, *25*, 695–704.
130. Kosa, T.; Moruyama, T.; Otagiri, M. *Pharm. Res.* **1997**, *14*, 1607–1612.
131. Curry, S.; Brick, P.; Franks, N. P. *Biochim. Biophys. Acta (BBA) – Mol. Cell Biol. Lipids* **1999**, *1441*, 131–140.
132. Fasano, M.; Curry, S.; Terreno, E.; Galliano, M.; Fanali, G.; Narciso, P.; Notari, S.; Ascenzi, P. *IUBMB Life* **2005**, *57*, 787–796.
133. Kontopidis, G.; Holt, C.; Sawyer, L. *J. Dairy Sci.* **2004**, *87*, 785–796.
134. Spector, A. A.; Fletcher, J. E. *Lipids* **1970**, *5*, 403–411.
135. Chos, Y.; Batt, C. A.; Sawyer, L. *J. Biol. Chem.* **1994**, *269*, 11102–11107.
136. Maux, S. L.; Bouhallab, S.; Giblin, L.; Brodkorb, A.; Croguennec, T. *Bovine Dairy Sci. Technol.* **2014**, *94*, 409–426.
137. Roth-Walter, F.; Pacios, L. F.; Gomez-Casado, C.; Hofstetter, G.; Roth, G. A.; Singer, J.; Diaz-Perales, A.; Jensen-Jarolim, E. *PLoS One* **2014**, *9*, 1–8.
138. Hamada, D.; Segawa, S.; Goto, Y. *Nat. Struct. Biol.* **1996**, *3*, 868–873.
139. Burova, T. V.; Choiset, Y.; Tran, V.; Haertlé, T. *Protein Eng.* **1998**, *11*, 1065–1073.
140. Scouten, W. H.; Lubcher, R.; Baughman, W. *Biochim. Biophys. Acta.* **1974**, *336*, 421–426.
141. Tyagarajan, K.; Pretzer, E.; Wiktorowicz, J. E. *Electrophoresis* **2003**, *24*, 2348–2358.



142. Stapelfeldt, H.; Olsen, C. E.; Skibsted, L. H. *J. Agric. Food. Chem.* **1999**, *47*, 3986–3990.
143. Puljung, M. C.; Zagotta, W. N. *Curr. Protoc. Protein Sci.* **2012**, *14*, 1–10.
144. Fukusima, Y.; Kawata, Y.; Onda, T.; Kitagawa, M. *Am. J. Clin. Nutr.* **1997**, *65*, 30–35.
145. Fukushima, Y.; Kawata, Y.; Onda, T.; Kitagawa, M. *J. Nutr. Sci. Vitaminol. (Tokyo)* **1997**, *43*, 673–678.
146. Gottschalk, M.; Nilsson, H.; Roos, H.; Halle, B. *Protein Science: A Publication of the Protein Society* **2003**, *12*, 2404–2411.
147. Amri, E.; Mamboya, F. *Am. J. Biochem. Biotechnol.* **2012**, *8*, 99–104.
148. Mitchel, R. E. J.; Chaiken, I. M.; Smith, E. L. *J. Biol. Chem.* **1970**, *245*, 3485–3492.
149. Dubey, V. K.; Pande, M.; Singh, B. K.; Jagannadham, M. V. *African J. Biotechnol.* **2007**, *6*, 1077–1086.
150. Drenth, J.; Jansonius, J. N.; Koekoek, R.; Swen, H. M.; Wolthers, B. G. *Nature* **1968**, *218*, 929–932.
151. Wang, J.; Xiang, Y. F.; Lim, C. *Protein Eng.* **1994**, *7*, 75–82.
152. Lown, J. A.; Dale, B. J. *Immunohematology* **1995**, *11*, 140–142.
153. Lowe, G. *Philos. Trans. R Soc. L. B Biol. Sci.* **1970**, *257*, 237–248.
154. Kamphuis, I. G.; Kalk, K. H.; Swarte, M. B. A.; Drenth, J. *J. Mol. Biol.* **1984**, *179*, 233–256.
155. Farnsworth, N. R.; Akerele, O.; Bingel, A. S.; Soejarto, D. D.; Guo, Z. Medicinal plants in therapy. *Bull. World Health Organ.* **1985**, *63*, 965–981.
156. Krishna, K. L.; Paridhavi, M.; Patel, J. A. *Indian J. Nat. Prod. Resour.* **2008**, *7*, 364–373.

## *Chapter 2*

---

### *Experimental Methods*



## Chapter 2

---

*This chapter describes the experimental techniques that I have used to carry out the work presented in this thesis. The basics of working principles and instrumental set up will be elaborated for the following instruments*

- *Spectrophotometer*
- *Spectrofluorometer*
- *Circular dichroism spectrometer*
- *Time correlated single photon counting instrument*
- *Fluorescence correlation spectroscopic system*



## 2.1. Instruments

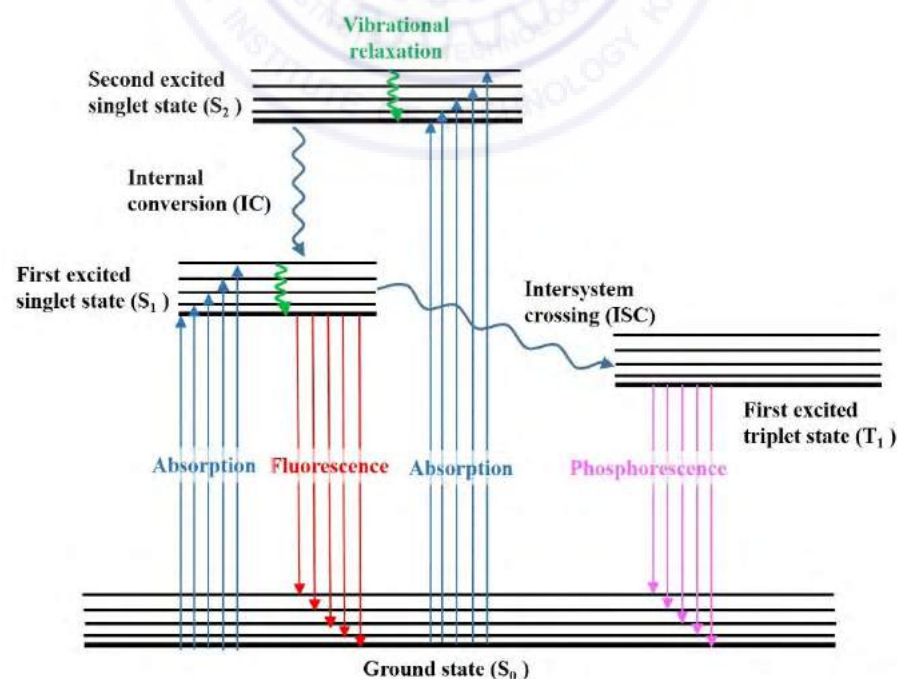
Steady-state and time-resolved experiments were performed with commercial instruments, whereas the single molecule level experiments were performed in a home built set up.

**2.1.1. Steady-State Absorption Measurements Using a Spectrophotometer:** The steady-state absorption of a material is a measure of its ability to absorb a particular light generally expressed as the absorption spectra in the form of an  $XY$ -graph with wavelength in the  $X$ -axis and absorbance in the  $Y$ -axis. The absorbance of a molecule is defined as the capability of a molecule to absorb photons and can be defined as<sup>1,2</sup>

$$A = \log_{10} \frac{I_i}{I_f} \quad (\text{Eq. 2.1})$$

where,  $I_i$  and  $I_f$  denotes the intensity of the light received by the material and the intensity of the light transmitted by the material respectively.

The difference in the initial and final energy states corresponds to the energy of the absorbed photon which in turn is related to the wavelength of the photon. Due to the presence of several vibrational and rotational energy levels in a single electronic energy level (Jablonski diagram, scheme 2.1) the absorption spectra gets a certain

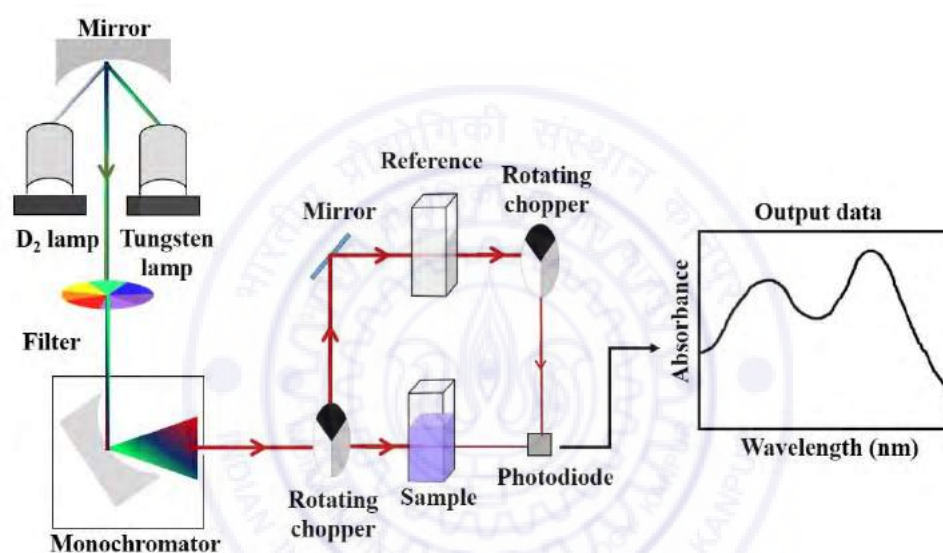


**Scheme 2.1:** Jablonski Diagram.

width.<sup>3,4</sup> The most relevant equation associated with the absorbance of a material is the Lambert-Beer's law which is described as<sup>5</sup>

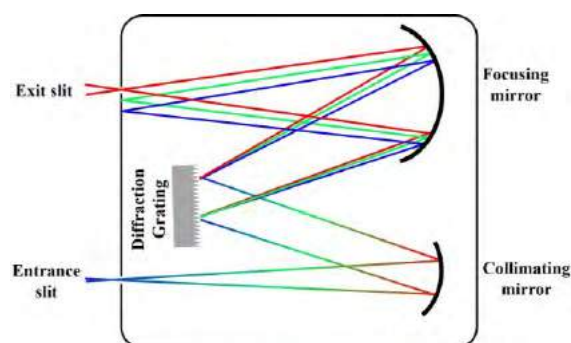
$$A(\lambda) = \varepsilon(\lambda)cl \quad (\text{Eq. 2.2})$$

where  $\varepsilon(\lambda)$  is the extinction coefficient of the material at the associated wavelength ( $\lambda$ ),  $c$  is the concentration and  $l$  is the pathlength which the light travels after incidence and before getting transmitted. The exact knowledge of extinction coefficient of a material and the pathlength enables us to calculate the concentration of a material in a solution from its absorbance.



**Scheme 2.2:** Schematic representation of UV-Vis spectrophotometer.

The basic instrumentation (scheme 2.2) includes a light source, a sample holder, a detector and necessary optical accessories. A commercial spectrophotometer



**Scheme 2.3:** Schematic representation of monochromator.

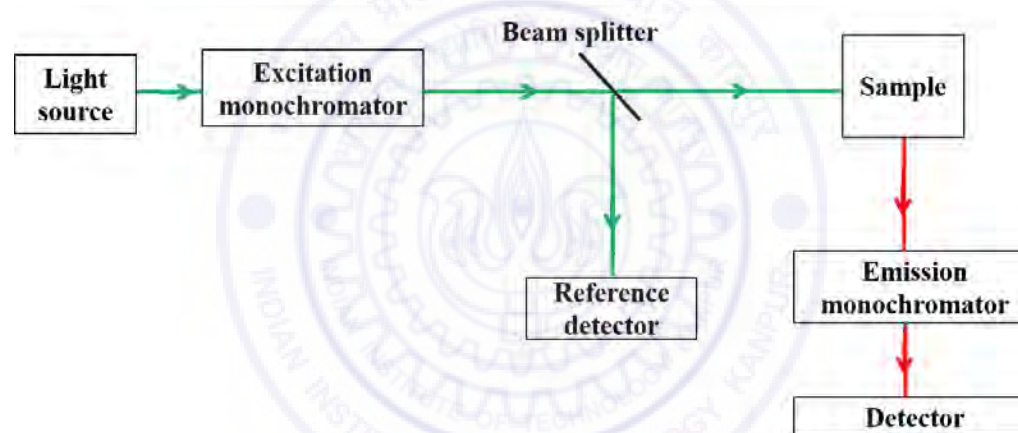
(Shimadzu 2450, Japan) was used for the steady-state absorption measurements. Here, D<sub>2</sub>-lamp is used for UV and near UV wavelength range whereas tungsten lamp is used for visible wavelength range. The light passes through filter, monochromator (scheme 2.3) and a rotating chopper before passing through

the sample kept in quartz cuvette having 10 mm pathlength. The path of the light is



dictated by the rotating chopper. The transmitted photons from the reference and sample are detected by a single photomultiplier tube (PMT), processed and the result is plotted.

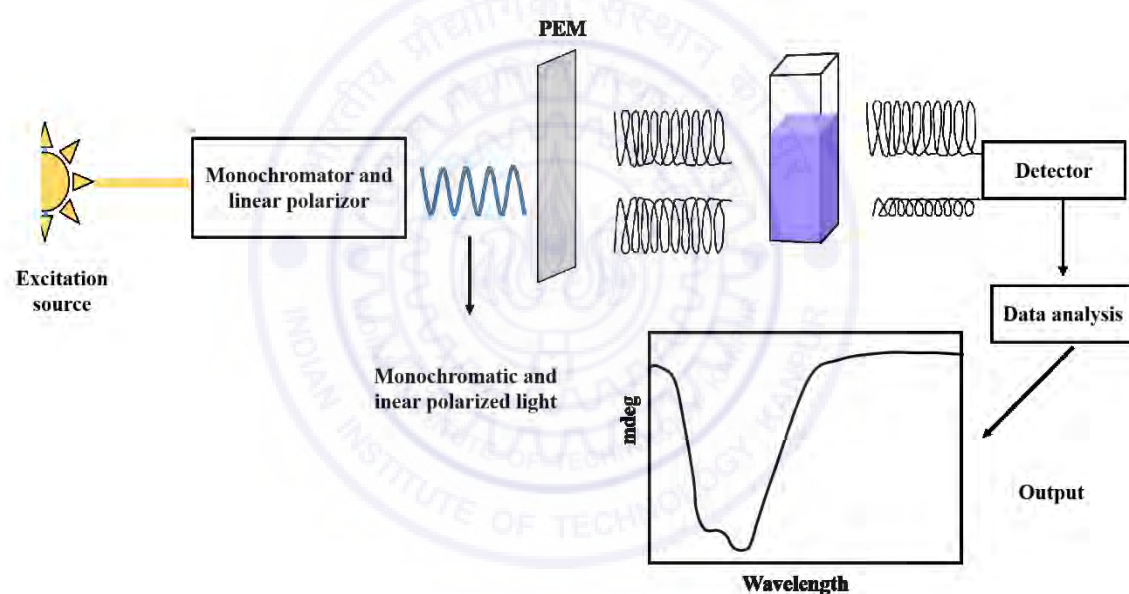
**2.1.2. Steady-State Fluorescence Measurements Using a Spectrofluorometer:** A spectrofluorometer is used to record the fluorescence spectra of a sample. The fluorescence results from the radiative pathway of a molecule while coming back from the excited electronic state to the ground electronic state.<sup>6</sup> During this relaxation the molecule usually emits photons with energy corresponding to the energy gap between the lowest vibrational levels of the initial and final electronic states. The Jablonski diagram (scheme 2.1) of a molecule illustrates the transitions between electronic states of a molecule.



**Scheme 2.4:** Schematic representation of spectrofluorometer.

For carrying out the work described in this thesis, I have used Fluoromax-4 spectrofluorometer from Jobin Yvon, USA. The basic components of this set up consists of a Xenon lamp (excitation source), an excitation monochromator, a sample chamber, an emission monochromator, one photodiode for reference detection and one photomultiplier tube for signal detection and other necessary optical accessories (scheme 2.4). Fluorescence measurements were carried out in quartz cuvette having 10 mm path length. Coumarin 343 in ethanol ( $\phi_f = 0.63$ ), tryptophan in water ( $\phi_f = 0.12$ ) and coumarin 480 in water ( $\phi_f = 0.66$ ) were used as standards for quantum yield measurements.<sup>7,8</sup>

**2.1.3. Circular Dichroism Spectroscopy Using a Circular Dichroism Spectrometer:** Circular dichroism (CD) spectroscopy is useful to collect information about the secondary structure of the proteins. It is basically an absorption spectroscopic technique operating mainly in the near and far ultraviolet (near-UV and far-UV) region.<sup>9</sup> The  $\pi$ - $\pi^*$  and  $n$ - $\pi^*$  transition of amide bonds are highly dependent on the protein backbone and is informative for the extent of regular structural motifs such as helices and sheets. Depending on the structure of the protein, the extent of absorption of the left-handed circularly polarized light (LCPL) and the right-handed circularly polarized light (RCPL) differs. The output signal of CD spectrometer is simply the difference of these two absorptions.<sup>10</sup> The basic instrumentation of CD-spectrometer is shown in scheme 2.5.



**Scheme 2.5:** Schematic representation of circular dichroism spectrometer.

The light emitted from the high power excitation source is dispensed in a monochromator which enables the wavelength selection and then passes through a polarizer to get the plane polarized light. This polarized light then passes through a photo-elastic modulator (PEM). The PEM divides the linearly polarized light into left-handed circularly polarized and right-handed circularly polarized light. This light passes through a protein solution, part of it gets absorbed and the transmitted light hits the detector. Now in absence of any optically active sample, the intensity of LCPL and RCPL reaching the detector will be same but in presence of the protein solution they

will be different. If the total light intensity is  $I_t$  and the difference in the absorbance for two different polarizations is  $I_d$  then the CD signal will be

$$CD = \frac{I_d}{I_t} G \quad (\text{Eq. 2.3})$$

Where  $G$  is the calibration scaling factor. For recording the CD spectroscopic results reported in this thesis a commercial CD spectrometer (J-815, Jasco, Japan) has been used.

The proper analysis of CD spectra of proteins can deliver information about the content of regular motifs such as helix and sheets. Proteins having the  $\alpha$ -helical pattern show negative bands at 208 nm and 222 nm and a positive band at 183 nm. Negative bands at 218 nm and positive band at 195 nm are symbolic of antiparallel  $\beta$ -sheets.<sup>9</sup> Proteins lacking well defined structural motifs or disordered proteins show very less intensity above 210 nm and negative bands at 195 nm. In this thesis, I have mainly dealt with the  $\alpha$ -helical content of the protein which can be calculated from the mean residual ellipticity at 208 nm.<sup>11-13</sup>  $MRE_{208nm}$  is defined as,

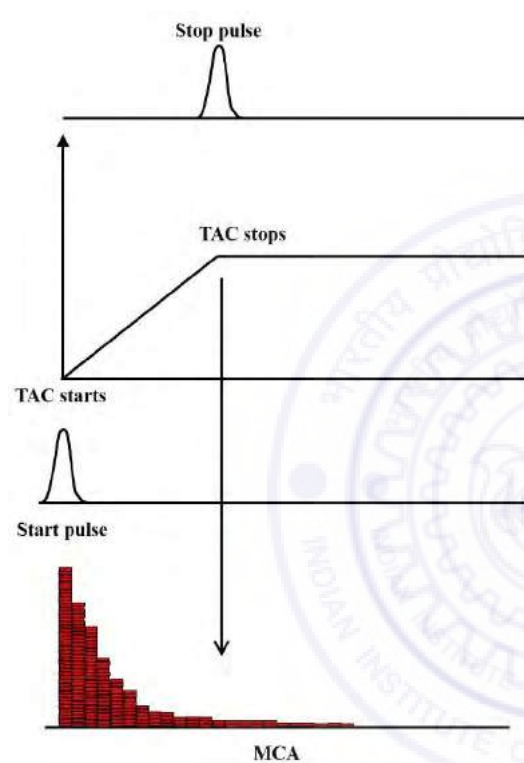
$$MRE_{208nm} = (\alpha_{208nm} \times M) / (n \times c \times l) \quad (\text{Eq. 2.4})$$

where,  $\alpha_{208nm}$  is the CD signal at 208 nm,  $M$  is the molecular weight of the protein,  $n$  is the total number of amino acid residues in the protein,  $l$  is the path length of the cuvette used for CD measurement and  $c$  is the concentration of the protein in  $\text{gm l}^{-1}$ . The mean residual ellipticity of the protein is related to the percentage of  $\alpha$ -helix present in the protein by the following equation,

$$\% \alpha - helicity = (MRE_{208nm} - 4000) / (33000 - 4000) \quad (\text{Eq. 2.5})$$

**2.1.4. Time-Resolved Fluorescence Measurements:** Time-resolved fluorescence measurements are carried out to calculate the average time that the molecule spends in its excited state which is better known as the fluorescence lifetime of the molecule.<sup>6</sup> There are several techniques that can be applied for fluorescence lifetime measurements namely time correlated single photon counting (TCSPC), fluorescence up-conversion and streak camera. For my research work, I have used a TCSPC

instrument with few tens of picosecond resolution. The principle of TCSPC instrument is to plot the time delay of a single emitted photon for each excitation pulse with respect to a reference photon which serves as the start time (scheme 2.6).<sup>14-16</sup> By repeating this process for numerous excitation events, a probability distribution of the time delay of the photons is obtained. The specification of the important components of Lifespec II TCSPC set up from Edinburgh Instruments, UK, which I have used for my thesis work, is described below (scheme 2.7).



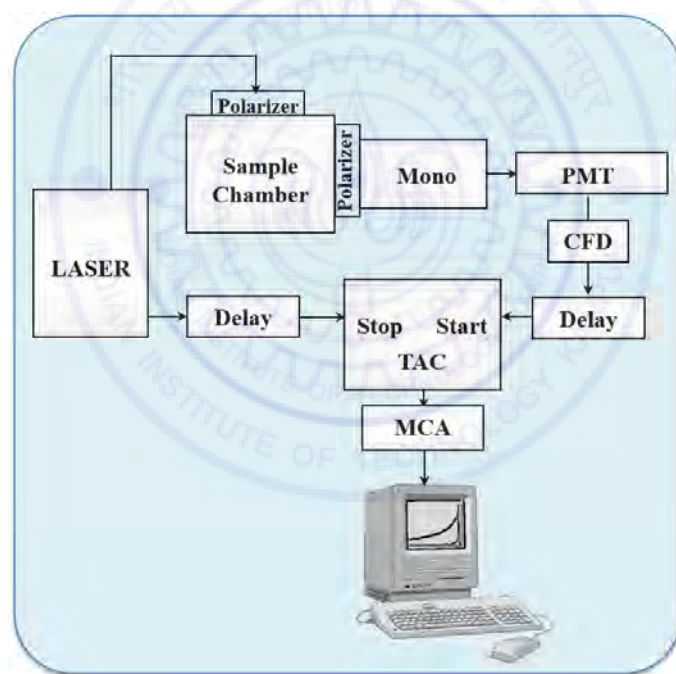
**Scheme 2.6:** Schematic understanding of principle of fluorescence lifetime measurement.

A picosecond diode laser of the required wavelength is used as the excitation source. This laser operates at high repetition rate ( $10^7 \text{ s}^{-1}$ ) and low intensity (average power  $\sim 20 \text{ } \mu\text{W}$ ) and is linearly polarized. For fluorescence lifetime measurements, vertical polarization was used. The sample chamber holds the sample in a cuvette having 10 mm path length. An analyser was used to measure the fluorescence transients at the magic angle ( $54.7^\circ$ ) polarisation to nullify any error which can be induced by the change in reorientation of transition moment of the molecule. A micro-channel plate photomultiplier tube (MCP-PMT) is used to detect the photons.

The difference of MCP-PMT with normal PMT is that in MCP-PMT there are multiple PMTs of smaller radius, which significantly reduce the transit time spread. As already discussed the function of PMT is to generate a definite voltage when it receives a photon. There is a certain fluctuation in this voltage generation, which induces an amplitude jitter and this amplitude jitter may introduce significant error in the time measurement, which is known as the time jitter. As a result, the fluorescence transient may become inaccurate. In order to minimize this, a constant fraction discriminator

(CFD) is used which process the voltage profile generated by the PMT in such a way that the time jitter decreases significantly.

After passing through the CFD, the signal is received by the time to amplitude converter (TAC). This time to amplitude converter is the most important component of a TCSPC instrument. The task of TAC is to generate a voltage ramp, which is proportional to the time lag between the start and the stop pulse. The start pulse is sent from the excitation source directly and can be considered as the zero time. This start pulse triggers the capacitor inside the TAC, which then starts to generate a voltage from 0 to 10 volts over a time window specified by the user. The minimum and maximum time window allowed for Lifespec II is 2.5 ns and one 200  $\mu$ s respectively. The voltage generation stops when a stop pulse, which is actually the first detected photon by the PMT, reaches the TAC. Thus the generated voltage gives a measure of the time



**Scheme 2.7:** Schematic representation of TCSPC.

difference between the excitation event and the emission of the photon. The TAC output is then amplified through a programmable gain amplifier and converted to numerical value using a analog-to-digital converter (ADC) respectively. Finally the detected photon is sorted in a multichannel analyser (MCA) and assigned to different time bins according to their arrival time. For our instrument, the maximum number of



channels is 4096 which gives us a resolution of 0.61 picosecond per channel for a TAC range of 5 ns.

The measured intensity decay profile of a sample can be fitted with proper models to extract information about the fluorescence lifetime of the sample. However, before this, the data must be processed in order to eliminate the interference from the laser excitation source and the error induced by the finite response time of the associated electronics. The principle of intensity decay measurement is based on the assumption that the laser source is a delta pulse, whereas in practice, it has a certain width. To eliminate the error, the response of the instrument for the zero lifetime sample (IRF) is measured using a scatterer solution such as colloidal silica (ludox solution) or milk, *etc.* and then deconvoluted from the measured decay profile of the sample. For deconvolution it is assumed that the finite width excitation source is actually a combination of several delta excitation pulses having different amplitudes. Thus, the total intensity at time  $t$  will be the summation of the responses of the sample to all the delta pulses. Mathematically, it can be expressed as

$$N(t) = \int_0^t L(t')I(t - t') dt' \quad (\text{Eq. 2.6})$$

where,  $I(t - t')$  is the intensity of the sample at time  $t$  as a result of excitation by a delta pulse of amplitude  $L(t')$  at time  $t'$ . The above equation is called the convolution integral. The job now is to find a suitable decay function which gives the best overall fit between  $N(t)$  and  $L(t)$ . This process is easier when the decay is single exponential. With increment in the number of exponential components, the situation becomes more complex. However, one should only try higher exponential fitting model when lower ones fail.

The deconvolution is based on an iterative least square regenerative convolution method.<sup>14-18</sup> After we record the IRF and the fluorescence transient, a new reconvoluted function is generated by mixing the IRF and a projected decay. This reconvoluted function is then taken as the fit function for the recorded decay and  $\chi^2$  value is obtained for the fit. The  $\chi^2$  value is a measure of the goodness of fit and can be written as

$$\chi^2 = \sum_{k=0}^n \frac{1}{\sigma_k^2} [F(t_k) - F_P(t_k)]^2 \quad (\text{Eq. 2.7})$$

where,  $F(t_k)$  and  $F_P(t_k)$  are the measured and projected data at two different times,  $n$  is the number of data points and  $\sigma_k$  is the standard deviation of each data point.  $\sigma_k$  in case of TCSPC is expressed as

$$\sigma_k = [N(t_k)]^{1/2} \quad (\text{Eq. 2.8})$$

based on the fact that TCSPC works on the Poisson statistics.  $N$  is taken as the number of channels used for measurement.

The process of reconvolution and calculation of  $\chi^2$  is repeated until the improvement of  $\chi^2$  value for two successive iterations is negligible. When the process comes to an end, a reduced  $\chi^2$  is calculated following the expression

$$\chi_r^2 = \frac{\chi^2}{n-f} = \frac{\chi^2}{\nu} \quad (\text{Eq. 2.9})$$

where  $\nu$  is the degree of freedom which is difference between the number of data points ( $n$ ) and the floating parameters ( $f$ ).

For data analysis, FAST (fluorescence analysis decay technology) software developed by Edinburgh Instruments, UK, has been used.

**2.1.5. Single Molecular Fluorescence Measurements:** Major part of this thesis work consist of single molecular level fluorescence measurements. Commonly practised spectroscopic methods such as steady-state absorption emission or time-resolved fluorescence spectroscopy deliver information about the average properties of a system. This information suffices for isotropic and homogeneous systems. When the properties of the constituents of a system is not homogeneous, thorough investigations are necessary to understand the inhomogeneity. That is why observing a single molecule is crucial in such cases. Moreover, there are properties associated with a system which can only be described by single molecular measurements. Besides all these, one of the major technical benefits of single molecular experiments are that they can be operated at nanomolar to picomolar concentrations. There are several single

molecular spectroscopic techniques among which I have used fluorescence correlation spectroscopy (FCS) to look into protein structure and dynamics thoroughly.

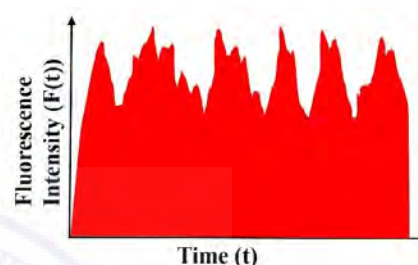
**2.1.5.1. History and Invention of FCS:** FCS deals with the temporal fluctuation of fluorescence intensity collected from a small observation volume, which may arise due to diffusion, chemical reaction or any photophysical process.<sup>6</sup> The idea of FCS was first conceived by Elliot L. Elson, Douglas Magde and Watt. W. Webb in the year 1972. It was illustrated by them using the case of binding of ethidium bromide to DNA.<sup>22</sup> The fluctuation of fluorescence intensity of ET-DNA fluorescent complex was measured to monitor the concentration fluctuation of a reactive system and to determine the reaction rate constants and diffusion coefficients. Encouraged by the success of the previous experiments, they wanted to perform FCS experiments extensively on different biological systems. But, although their first experiment served well to provide the chemical numbers they were interested to calculate, their experimental set up was not enough with respect to resolution for single molecular level measurements. It was important to make a single molecule visible to the measurement for better output. Now, it is difficult to locate someone in a huge overcrowded place as compared to a small room. Hence, a very low concentration and a tiny observation volume was necessary for achieving single molecular resolution. For this reason in 1976, E. Elson and W. W. Webb coupled confocal microscopy with FCS to study the mobility of the molecules on the cell surface.<sup>23</sup> At that time, fluorescence recovery after photobleaching (FRAP) was also considered as an effective way to measure molecular diffusion. But FCS was better in a way that it was non-invasive as compared to FRAP which destroys the fluorophores by photobleaching. From the 1970s onwards they published a lot of works along with R. Rigler, M Ehrenberg, M. Eigen while gradually improving experimental set up for ameliorating the signal to noise ratio.<sup>24-32</sup> With the invention of advanced optical accessories, the confocal system was also improved and detection in the femtoliter observation volume became possible. That is when FCS gained its current status as a highly efficient single molecular fluorescence measurement technique. At present, FCS is widely used to study molecular diffusion, size alteration under environmental effect, molecular interactions, chemical kinetics, conformational dynamics of macromolecule *etc.*<sup>33-44</sup>

In recent times, several sophisticated techniques has been invented based on FCS which have better output, lower artefacts and easier analysis. Some of these are fluorescence cross correlation spectroscopy (FCCS), dual focus FCS, near field optical scanning FCS (NSOM-FCS), fluorescence lifetime correlation spectroscopy (FLCS), two dimensional FCS (2D-FCS), two dimesion FLCS (2D-FLCS) *etc.*

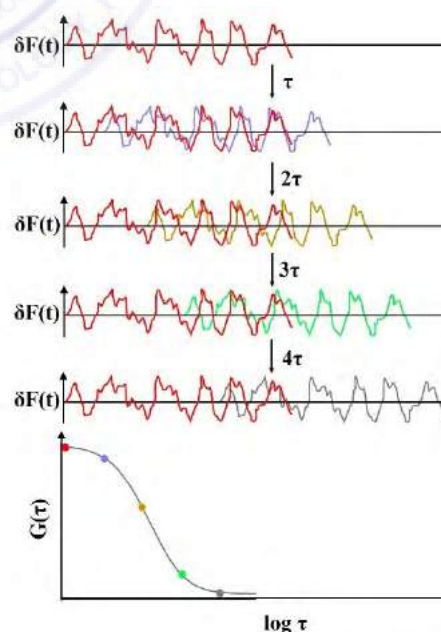
**2.1.5.2. Theory of FCS:** The dictionary meaning of the word autocorrelation is “correlation between the elements of a series and others from the same series separated from them by a given interval”. In case of FCS this “element” is fluorescence intensity and the “series” is the fluorescence intensity as a function of time (scheme 2.8). This fluorescence intensity is practically the number of photons collected from a tiny observation volume. When the fluorescent particle is in this observation volume, the emitted photons are efficiently collected by the detector and when the particle is outside this volume, no photons from this particle reach the detector. If the average fluorescence intensity from this volume is  $\langle f(t) \rangle$  and the deviation of the fluorescence intensity at time  $t$  and  $t+\tau$  from this average intensity is  $\delta F(t)$  and  $\delta F(t+\tau)$  then the extent of self similarity of the fluorescence signal is expressed as<sup>6</sup>

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t+\tau) \rangle}{\langle F(t) \rangle^2} \quad (\text{Eq. 2.10})$$

$G(\tau)$  is called the autocorrelation function. Here ‘ $\tau$ ’ is the time interval over which the autocorrelation is being calculated. A plot of  $G(\tau)$  versus  $\tau$  gives us how the self similarity decays with time. Scheme 2.9 illustrates the decay of autocorrelation function with time. To extract information about diffusion or any other event that may cause fluorescence fluctuation,



**Scheme 2.8:** Schematic representation of fluorescence intensity from the observation volume as a function of time.

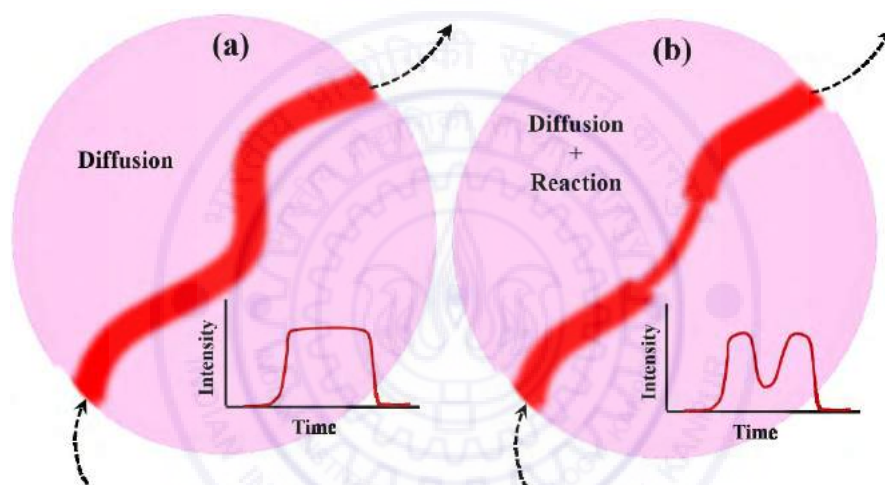


**Scheme 2.9:** Schematic understanding of generation of autocorrelation function.



one must have a clear idea about the focal volume and the system under consideration. The fitting equation which will produce useful parameter stems from the understanding of the origin of fluorescence fluctuation coupled with the mathematical derivation of the exact model that will justify the fluctuation.

**2.1.5.3. Factors Contributing in the Fluorescence Fluctuation:** The main reason of fluorescence fluctuation in fluorescence correlation spectroscopy is diffusion of the fluorescent particle in and out of the observation volume. Besides this, there are several reasons that can play crucial role in the fluctuation. Amongst them, some of the most contributing ones are triplet state blinking, other photophysical intensity fluctuations, chemical reactions having timescale lower than that of diffusion timescale etc. The



**Scheme 2.10:** Intensity fluctuation in case of (a) pure diffusion and (b) diffusion + reaction.

most significant reaction of this kind is the quenching reaction. Scheme 2.10 illustrates two different scenarios of diffusion in absence (scheme 2.10a) or presence (scheme 2.10b) of such reaction. As evident, additional fluctuations can be observed while the fluorescent particle is in the focal volume. In case of such an event, the resultant autocorrelation curve is altered as compared to the pure diffusion case and proper analysis of this autocorrelation curve can provide the timescales of both the processes.

**2.1.5.4. Mathematical Understanding of the Fitting Equation:** The fitting equation that will produce useful parameter stems from the understanding of the origin of fluorescence fluctuation coupled with the mathematical derivation of the exact model that will justify the fluctuation.<sup>19,30,45</sup>



The number of molecules in the focal volume at any point of time is governed by Poissonian distribution. For this distribution, the root mean square fluctuation of the particle number  $N$  is given by

$$\frac{\sqrt{\langle(\delta N)^2\rangle}}{\langle N\rangle} = \frac{\sqrt{\langle(N-\langle N\rangle)^2\rangle}}{\langle N\rangle} \quad (\text{Eq. 2.11})$$

It is necessary to optimize the number of molecules within the focal volume with respect to two opposing factors. As the number of fluorescent particles in the observation volume increases the relative intensity fluctuations become smaller which is detrimental for the measurements. On the other hand, decreasing the number of particles makes the background signal more prominent and the signal to noise ratio decreases. Considering all these, the number of particles in the focal volume must be in the nanomolar range.

The fluctuation of the fluorescence signal ( $\delta F(t)$ ) is defined as the deviations from the temporal average of the signal ( $\langle F(t) \rangle$ ).

$$\begin{aligned} \delta F(t) &= F(t) - \langle F(t) \rangle \\ \langle F(t) \rangle &= \frac{1}{T} \int_0^T F(t) dt \end{aligned} \quad (\text{Eq. 2.12})$$

If the fluorescence fluctuation is solely caused by the changes in the local concentration changes ( $\delta C$ ) within the observation volume ( $V_{eff}$ ), then  $\delta F(t)$  can be written as,

$$\delta F(t) = \kappa \int_V I_{ex}(\underline{r}) \cdot S(\underline{r}) \cdot \delta(\sigma \cdot q \cdot C(\underline{r}, t)) \cdot dV \quad (\text{Eq. 2.13})$$

Here,  $\kappa$  is the overall detection efficiency,  $I_{ex}(r)$  is the intensity at  $r$  with maximum intensity at amplitude  $I_0$ ,  $S(r)$  is a dimensionless parameter which reflects the optical transfer function of for the objective pinhole combination,  $\delta(\sigma \cdot q \cdot C(r, t))$  depicts the dynamics of the fluorophore in the single particle level.  $\delta(\sigma \cdot q \cdot C(r, t))$  is a combination of fluctuations in the molecular absorption cross section ( $\delta\sigma$ ), fluctuations in the quantum yield ( $\delta q$ ) and fluctuations in the local particle concentration at time  $t$  mainly due to Brownian motion ( $\delta C(r, t)$ ). Determining these parameters experimentally is a very difficult job. For simplifying the overall term, a new term  $W(r)$  is introduced which is actually a combination of  $I_{ex}(r)/I_0 * S(r)$ . This term is usually approximated by a three

dimensional Gaussian with an intensity  $I_0/e^2$  at  $r=r_0$  in lateral direction and at  $z=z_0$  in axial direction and is expressed as,

$$W(r) = e^{-2\frac{x^2+y^2}{r_0^2}} e^{-2\frac{z^2}{z_0^2}} \quad (\text{Eq. 2.14})$$

Rest of the parameters  $\kappa$ ,  $\sigma$  and  $q$  is combined with  $I_0$  to generate a new parameter  $\eta_0$  that determines the photon count rate per second and is also a measure of signal to noise ratio.  $\eta_0$  is given by,

$$\eta_0 = I_0 \cdot \kappa \cdot \sigma \cdot q \quad (\text{Eq. 2.15})$$

Introducing these new terms in equation 2.13 becomes,

$$\delta F(t) = \kappa \int_V W(r) \cdot \delta(\eta \cdot C(\underline{r}, t)) \cdot dV \quad (\text{Eq. 2.16})$$

The normalized autocorrelation function is given by equation 2.10. Substitution of eq. 2.16 in eq. 2.10 gives,

$$G(\tau) = \frac{\int \int W(\underline{r}) W(\underline{r}') \langle \delta(\eta \cdot C(\underline{r}, t)) \delta(\eta \cdot C(\underline{r}', t+\tau)) \rangle dV dV'}{\left( \int W(\underline{r}) \langle \delta(\eta \cdot C(\underline{r}, t)) \rangle dV \right)^2} \quad (\text{Eq. 2.17})$$

Now, the fluctuation term can be separated as,

$$\delta(\eta \cdot C(\underline{r}, t)) = C \delta \eta + \eta \delta C$$

To avoid the complication it is assumed that the fluorophore's fluorescence properties are not changing when it is in the observation volume that is within the observation time i.e.  $\delta \eta = 0$ . So equation 2.17 can be rewritten as,

$$G(\tau) = \frac{\int \int W(\underline{r}) W(\underline{r}') \langle \delta(C(\underline{r}, 0)) \delta(C(\underline{r}', \tau)) \rangle dV dV'}{\left( \langle C \rangle \int W(\underline{r}) dV \right)^2} \quad (\text{Eq. 2.18})$$

For particles freely diffusing in three dimensions with a diffusion coefficient 3D, the autocorrelation for number density can be written as,

$$\langle \delta(C(\underline{r}, 0)) \delta(C(\underline{r}', \tau)) \rangle = \langle C \rangle \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} \cdot e^{-\frac{(\underline{r}-\underline{r}')^2}{4D\tau}} \quad (\text{Eq. 2.19})$$

Now  $G(\tau)$  becomes

$$G(\tau) = \frac{\int \int W(\underline{r})W(\underline{r}')\langle C \rangle \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} e^{-\frac{(\underline{r}-\underline{r}')^2}{4D\tau}} dV dV'}{(\langle C \rangle \int W(\underline{r})dV)^2}$$

$$= \frac{1}{\langle C \rangle (4\pi D\tau)^{\frac{3}{2}}} \cdot \frac{\int \int W(\underline{r})W(\underline{r}')\langle C \rangle \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} e^{-\frac{(\underline{r}-\underline{r}')^2}{4D\tau}} dV dV'}{(\langle C \rangle \int W(\underline{r})dV)^2} \quad (\text{Eq. 2.20})$$

The relationship between the lateral diffusion time ( $\tau_D$ ) and the diffusion coefficient (D) is

$$\tau_D = \frac{r_0^2}{4 \cdot D} \quad (\text{Eq. 2.21})$$

The effective focal volume is given by,

$$V_{eff} = \frac{(\int W(\underline{r})dV)^2}{\int W^2(\underline{r})dV} \quad (\text{Eq. 2.22})$$

Using equation 2.14,

$$V_{eff} = \frac{\left( \int e^{-2\frac{x^2+y^2}{r_0^2}} e^{-2\frac{z^2}{z_0^2}} dV \right)^2}{\int e^{-4\frac{x^2+y^2}{r_0^2}} e^{-4\frac{z^2}{z_0^2}} dV} \quad (\text{Eq. 2.23})$$

Integrating over space,

$$V_{eff} = \pi^{\frac{3}{2}} \cdot r_0^2 \cdot z_0 \quad (\text{Eq. 2.24})$$

So, finally the autocorrelation function for a freely diffusing molecule can be written as,

$$G(\tau) = \frac{1}{V_{eff}\langle C \rangle} \cdot \frac{1}{\left(1 + \left(\frac{\tau}{\tau_D}\right)\right)} \cdot \frac{1}{\sqrt{1 + \left(\frac{r_0}{z_0}\right)^2 \cdot \frac{\tau}{\tau_D}}} \quad (\text{Eq. 2.25})$$

The first factor in equation 2.25 is the inverse of the average particle number in the observation volume and is expressed as  $G(0)$ . So,

$$G(0) = \frac{1}{\langle N \rangle} = \frac{1}{V_{eff}\langle C \rangle} \quad (\text{Eq. 2.26})$$

Hence eq. 2.25 can also be written as,

$$G(\tau) = G(0) \cdot \left(1 + \left(\frac{\tau}{\tau_D}\right)\right)^{-1} \cdot \left(1 + (w_{rz})^2 \cdot \frac{\tau}{\tau_D}\right)^{-\frac{1}{2}} \quad (\text{Eq. 2.27})$$

where  $W_{rz}$  denotes the ratio of  $r_0$  and  $z_0$ .

Till now, it is assumed that the fluorophore does not undergo any other fluorescence fluctuation while it is in the observation volume. But in most of the cases, this assumption does not hold true. In presence of any other process, an additional term must be added which takes care of the additional processes. This way, the modified autocorrelation curve for this case looks like

$$G_{total}(\tau) = G_{total}(\tau) \cdot X_{kinetics}(\tau) \quad (\text{Eq. 2.28})$$

In case of reversible fluorescence quenching or triplet state blinking, the additional term can be denoted as

$$X_R(\tau) = 1 - A + A \cdot e^{-\frac{\tau}{\tau_R}} \quad (\text{Eq. 2.29})$$

Where,  $A$  is the fraction of fluorescence fluctuation caused by the process. Upon normalization by  $(1-A)$  the expression is modified as

$$X_R(\tau) = 1 + a \cdot e^{-\frac{\tau}{\tau_R}} \quad (\text{Eq. 2.30})$$

Where,  $a = 1/(1 - A)$ .

So the final form of autocorrelation for diffusion coupled with any other process now becomes,

$$G(\tau) = \frac{1}{V_{eff}\langle C \rangle} \cdot \frac{1}{\left(1 + \left(\frac{\tau}{\tau_D}\right)\right)} \cdot \frac{1}{\sqrt{1 + \left(\frac{r_0}{z_0}\right)^2 \cdot \frac{\tau}{\tau_D}}} \cdot \left(1 + a \cdot e^{-\frac{\tau}{\tau_R}}\right)$$

$$= G(0) \cdot \left(1 + \left(\frac{\tau}{\tau_D}\right)\right)^{-1} \cdot \left(1 + (w_{rz})^2 \cdot \frac{\tau}{\tau_D}\right)^{-\frac{1}{2}} \cdot \left(1 + a \cdot e^{-\frac{\tau}{\tau_R}}\right) \quad (\text{Eq. 2.31})$$

Both equation 2.27 and 2.31 are derived with an assumption that the system is monodisperse, i.e. the all the diffusing particles are having same size. For polydisperse systems, however, these equation containing single diffusion time can not fit the data accurately owing to the fact that particles having different size have different diffusion coefficients. In that case a distribution term needs to be introduced to the diffusion time component. As an example, the micellar system can be considered. Here, the micellar size is not uniform and follows a Gaussian distribution with peak diffusion time  $\tau_{pD}$  with a distribution width  $\sigma_D$ . For this case, the modified equation 2.27 becomes,

$$G(\tau) = \sum_{i=1}^m b_i(\tau_{D_i}) \left(1 + \frac{\tau}{\tau_{D_i}}\right)^{-1} \left(1 + (w_{rz})^2 \frac{\tau}{\tau_{D_i}}\right)^{-1/2} \quad (\text{Eq. 2.32})$$

Where the term  $b_i(\tau_{D_i})$  is expressed as

$$b_i(\tau_{D_i}) = B_i e^{-\left(\frac{\ln(\tau_{D_i}) - \ln(\tau_{pD})}{\sigma_D}\right)^2} \quad (\text{Eq. 2.33})$$

with  $B_i$  as the relative amplitudes of diffusion components.

When this distributed diffusion is assisted by a single relaxation component, the equation to best fit the autocorrelation curve would be,

$$G(\tau) = \sum_{i=1}^m b_i(\tau_{D_i}) \left(1 + \frac{\tau}{\tau_{D_i}}\right)^{-1} \left(1 + (w_{rz})^2 \frac{\tau}{\tau_{D_i}}\right)^{-1/2} \left(1 + a \cdot e^{-\frac{\tau}{\tau_R}}\right) \quad (\text{Eq. 2.34})$$

But when the associated relaxation component also follows distribution, the modified equation 2.34 becomes,

$$G(\tau) = \sum_{i=1}^m b_i(\tau_{D_i}) \left(1 + \frac{\tau}{\tau_{D_i}}\right)^{-1} \left(1 + (w_{rz})^2 \frac{\tau}{\tau_{D_i}}\right)^{-1/2} \sum_{j=1}^n \left(1 + a_j(\tau_{R_j}) e^{-\frac{\tau}{\tau_{R_j}}}\right) \quad (\text{Eq. 2.35})$$

The term  $a_j(\tau_{R_j})$  is expresses as,



$$a_j(\tau_{Rj}) = A_j e^{-\left(\frac{\ln(\tau_{Rj}) - \ln(\tau_{pR})}{\sigma_R}\right)} \quad (\text{Eq. 2.36})$$

where  $A_j$  is the relative amplitudes relaxation components,  $\tau_{pR}$  and  $\sigma_R$  are the peak relaxation time and the distribution width of relation time.

In equation (7) and (8)  $B_i$  and  $A_j$  are the relative amplitudes of components,  $\tau_{pD}$  and  $\tau_{pR}$  are the peak diffusion of relaxation times having a width of  $\sigma_D$  and  $\sigma_R$  respectively. The choice of fitting model in the present study has been elaborated in results and discussion section.

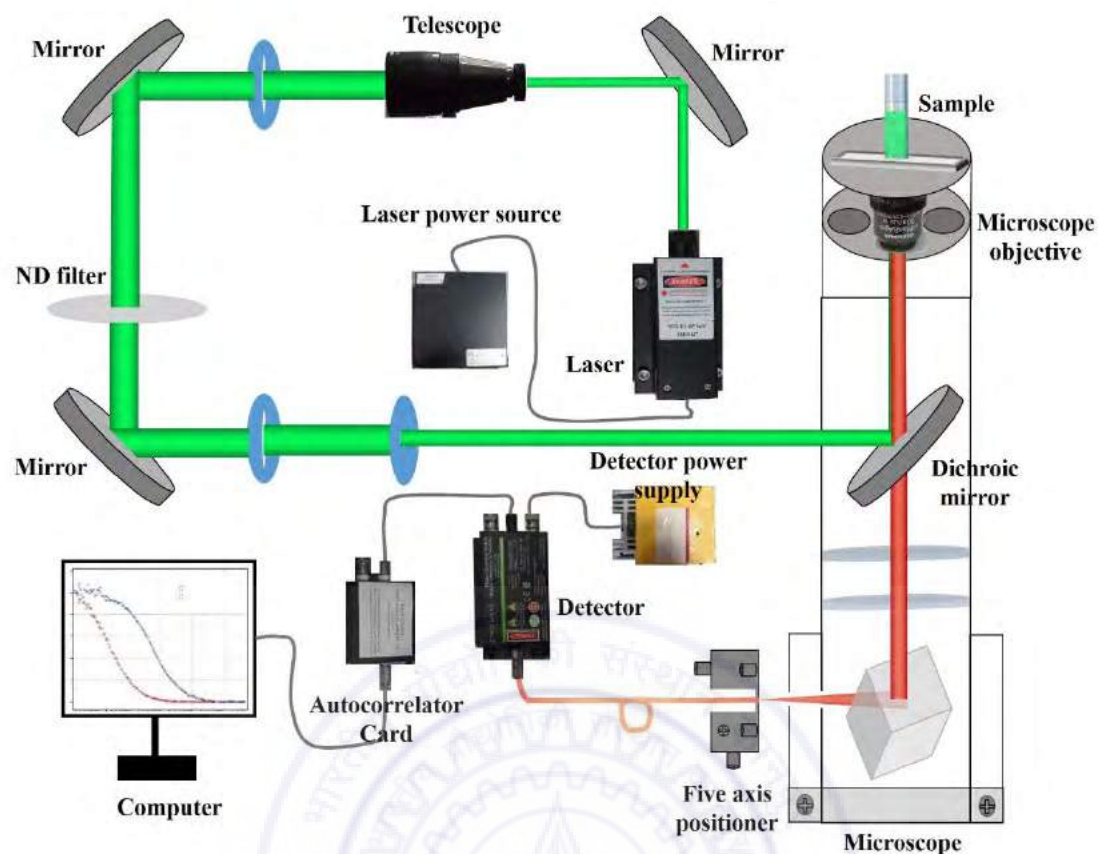
**2.1.5.5. Extracting Structural Information from the Diffusion Time:** Once the diffusion time ( $\tau_D$ ) and diffusion coefficient ( $D$ ) of the particle under investigation is calculated by fitting the autocorrelation curves the hydrodynamic radius ( $r_H$ ) of the particle can be calculated from it using the Stokes-Einstein equation. The Stokes-Einstein equation deals with the Brownian motion of a spherical particle with radius  $r_H$  in a fluid having constant temperature  $T$ . The equation look like,<sup>46</sup>

$$r_H = \frac{k_B T}{6\pi\eta D} \quad (\text{Eq. 2.37})$$

where  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature and  $\eta$  is the viscosity of the system.

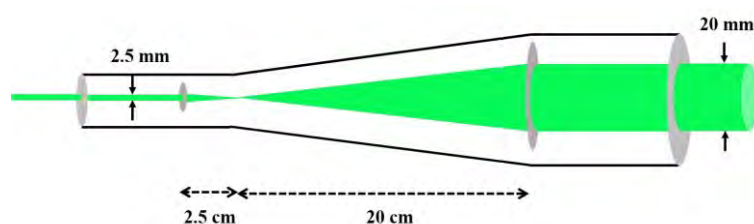
**2.1.5.6. FCS Set up:** A home built FCS set up has been used for performing single molecule fluorescence experiments performed for this thesis work. The schematic diagram has been shown as scheme 2.11 and the photograph of the actual set up has been shown in figure 2.1. For a simple understanding, the discussion of the whole set up has been divided into three basic units as (a) excitation unit, (b) detection unit and (c) data analysis unit and they are discussed in the following paragraphs.

**(a) Excitation Unit:** The excitation unit comprises of a laser source of required wavelength, a telescope, mirrors (M1 to M4) and iris (I1 to I4). The diameter of the laser output is  $\sim 2.5$  mm. The generation of a tiny focal volume is only possible when the beam diameter is larger. So this beam is passed through a telescope. The beam is  $\sim 8$  fold expanded when it passes through the telescope (scheme 2.12). This expanded



**Scheme 2.11:** Schematic diagram of components of our home built FCS set up.

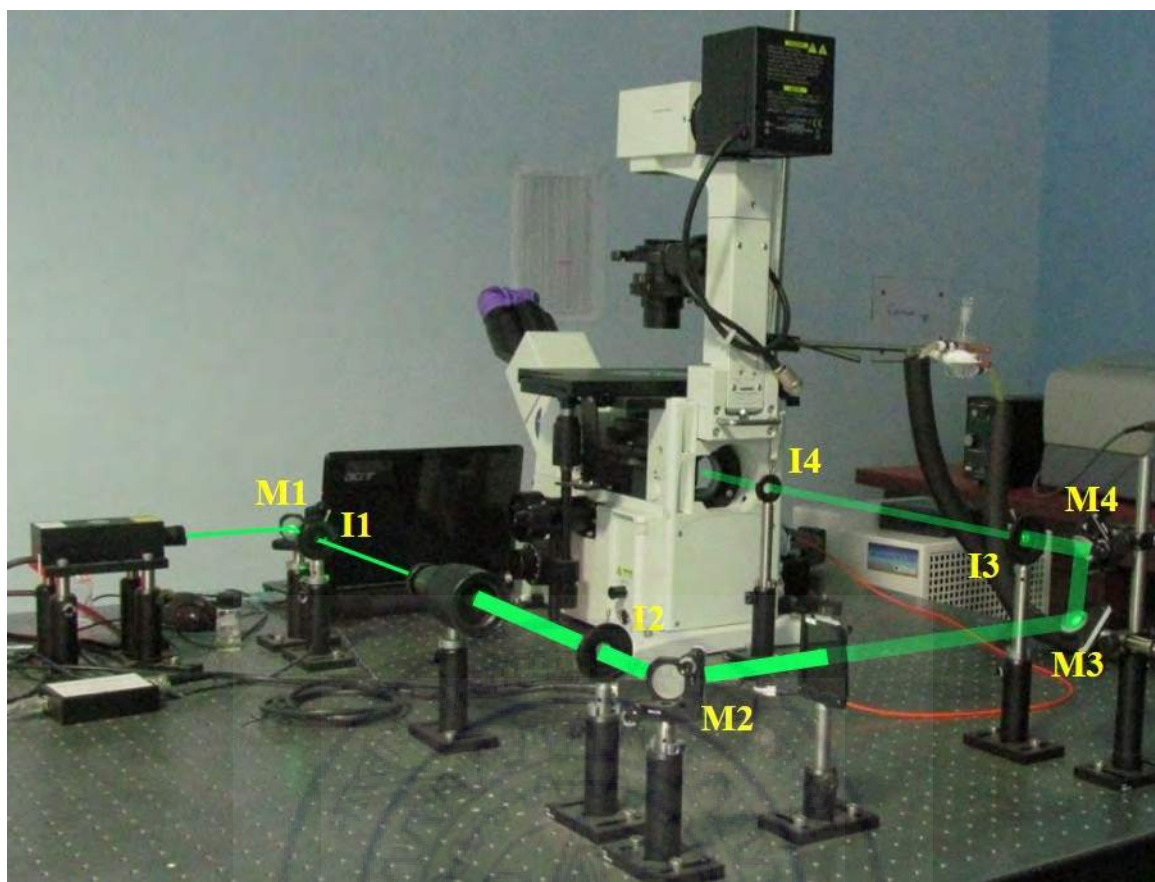
beam is then directed towards the microscope, which is used for the generation of the confocal excitation profile. A set of three mirrors are used for this purpose. Into the path of this beam, a slot for neutral density (ND) filter is installed to control the laser intensity. The final beam diameter can be controlled by the iris placed right before the microscope. Matching with the dimensions of the objective that we use as a part of the confocal set up, the beam diameter is maintained as 10 mm for achieving the lowest



**Scheme 2.12:** Schematic representation of expansion of excitation beam using telescope.

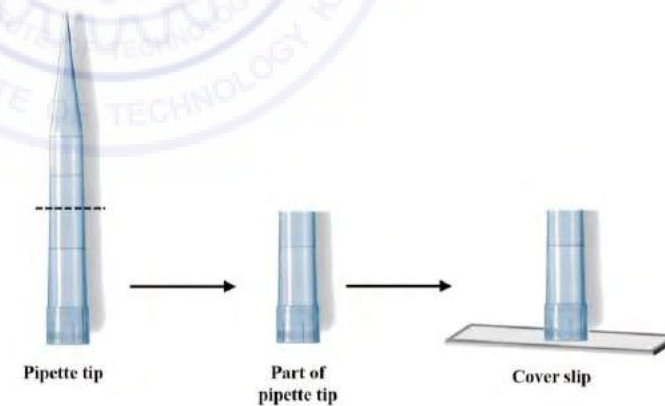
possible excitation volume. This beam then enters the microscope body and gets directed towards the objective after passing through a dichroic mirror.

The objective is actually a collection of lenses, which is used to focus the collimated beam onto the sample.



**Figure 2.1:** Home built FCS set up in our lab

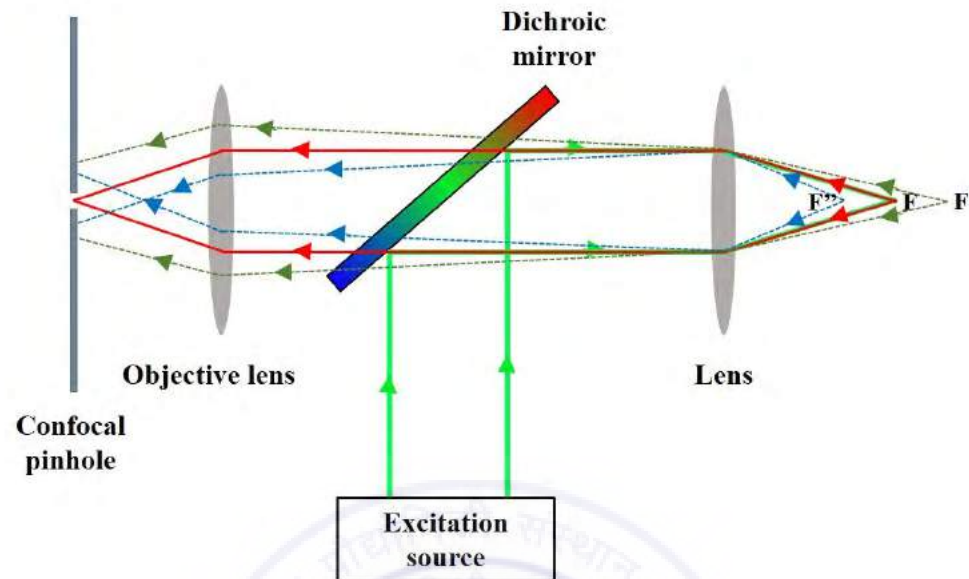
The sample is placed on a clean cover slip kept onto the sample platform. In most of our experiments, we have used specially made sample cells (scheme 2.13). The wider part of a divided pipette tip is attached to a clean cover slip with fevikwik. This cell is useful only when the solvent of the sample is non-interactive with the fevikwik or the pipette tip.



**Scheme 2.13:** Preparation of FCS sample cells.

**(b) Detection Unit:** Upon excitation, the fluorescent molecule starts emitting and the emitted photons are headed towards all the directions. Only the photons ejected towards the objective are detected. These photons are then collimated by the same objective, passes through a dichroic mirror and an emission filter and goes through a

focusing lens which finally focuses it on the tiny opening of the optical fibre. This tiny opening plays a significant role in the confocal detection (scheme 2.14). It makes sure



**Scheme 2.14:** Schematic representation of the confocal detection unit.

that only the light coming from a specific volume centred at the focal point of the excitation beam gets detected and is commonly called as the confocal pinhole. As shown in scheme 2.14, photons originating from  $F$  is only passing through the pinhole whereas those from surrounding areas are efficiently rejected. By varying the size of this pinhole, one can actually control the detection volume. Scheme 2.14 illustrates the role of the pinhole in controlling the focal volume. The optical fibre is firmly held in its position by a five axis positioner while enables the placement of the pinhole at the focal point of the emitted photons. The optical fibre carries the photons to the single photon counting avalanche photodiode that is used as the detector.

**(c) Analysis Unit:** The number of detected photons or the photon count is then sent to the autocorrelator card from the APD. The job of the autocorrelator card is to bin a particular section of this intensity profile and then calculate the extent of autocorrelation of this section with itself as a function of this lag time. For doing this, the autocorrelator card calculates  $\langle f(t) \rangle$  and  $\delta f(t)$  at first. The  $\delta f(t)$  versus  $t$  profile is then given a definite time lag depending on the specifications of the correlator card. Our autocorrelator card has a resolution of 0.2 microsecond which means the minimum time lag it can give to the fluctuation profile is 0.2 microsecond while the maximum



lag time is 3435969984 microsecond or almost an hour. The systems that can be studied with this autocorrelator card is limited by the minimum and maximum lag time of the card. The experimental autocorrelation curve is generated by autocorrelator card and subsequently sent to the displaying unit which is a labview platform in a computer in this case. The data is saved as ASCII format and further analysis is done using Igor Pro software.

**2.1.5.7. Specification of the Components of FCS Set up:**

**Laser:** MDL-III-405-5mW (405 nm), MGL-III-532-5mW (532 nm), CNI, China

**Telescope:** BE10M-A, Thorlabs, USA

**Mirrors:** KM100, Thorlabs, USA

**Microscope:** Inverted confocal microscope, IX-71, Olympus, Japan

**Objective lens:** 60X water immersion objective with numerical aperture 1.2, working distance 0.28 mm, Olympus, Japan

**Dichroic mirror:** ZT532RDC, ZT405RDC, Chroma, USA

**Bandpass filters:** ET605/70m, Chroma, USA

**Neutral density filters:** NE01B, NE02B, ..., NE10B, Thorlabs, USA

**Fibre patch chord:** M43L01, M42L01, M67L01, Thorlabs, USA

**Detector:** SPCM-AQRH-13-FC-24420, Excelitas technologies

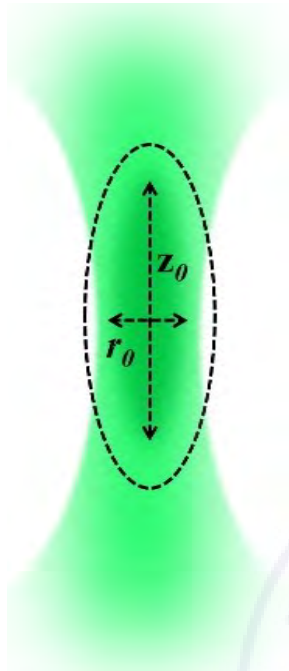
**Correlator card:** Flex 990EM-12E, USA

**Opto-mechanics:** Post, rods, iris and holders from Thorlabs, USA and Holmarc, India.

**2.1.5.8. Calibration of the Instrument:** The experimental autocorrelation curves produces information about the diffusion time through the detection volume upon being fitted. But the diffusion time is an extensive property and hence depends on the detection volume, solution viscosity of the medium *etc.* To acquire information regarding size of a molecule utilizing the diffusion time, one must calculate the focal volume first. Also, the fitting equations are derived based on the assumption that the profile of the excitation beam is Gaussian. For these reasons, calibration of the system before experiment is necessary. This includes several steps as follows.



**(a) Alignment of the Excitation Beam:** The beam height must be constant throughout the path of travelling till M3 (see figure 2.1). This can be checked by using a ruler and if found otherwise, can be adjusted by fine tuning the mirrors M1 and M2.



**Scheme 2.15:** Gaussian shaped observation volume.

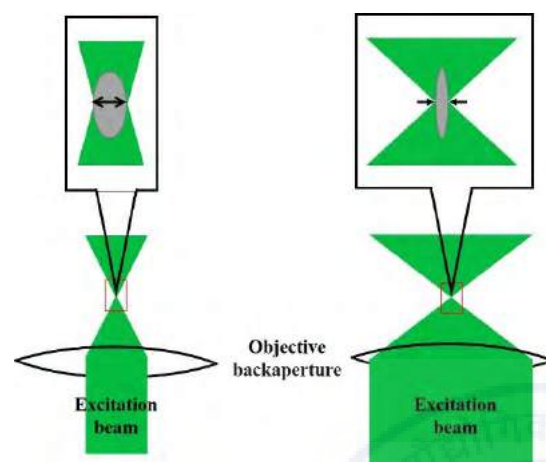
Upon being reflected by M3 the beam must reach M4 exactly vertically. Finally, the excitation light beam must follow the same height throughout the path from M4 to the dichroic mirror inside the microscope. When projected to a plane perpendicular to the direction of travelling, the excitation beam must have highest intensity at the centre and must decay towards every direction in that plane in a similar fashion. All this can be controlled by adjusting the mirrors and the iris. The light must be exactly aligned with the objective when it reaches the objective backaperture to make sure that the shape and size of the focal volume is unperturbed. This can be confirmed by putting a scatterer solution in the sample cell and switching on the excitation unit. The

focal volume can be seen with naked eye and the symmetry of the focal volume is checked in all the directions. If found asymmetric, it can be corrected by fine tuning the position and angle of the mirrors.

**(b) Alignment of the Detector:** The opening of the fibre patch cord, which serves as the confocal pinhole must be properly aligned with the focal point of the emitted photons to assure that it captures the photons coming from a Gaussian volume. As already discussed, the detector is held in position by a five axis positioner. This five axis positioner has a slot where the fibre can be attached tightly with a screw. The five screw gauge (adjustment of x-axis, y-axis, z-axis, angle with the vertical plane and angle with the horizontal plane), which are part of the five axis positioner enables positioning the opening of the fibre at the exact location where the emitted photons are focused. The alignment can be roughly achieved by using a concentrated dye solution as the sample and adjusting the five axis positioner till one observes a glow at the other end of the optical fibre with naked eye. For the perfect alignment, a dilute solution of

the dye ( $\sim 1-5$  nm) is used as a sample and the detected photon counts are maximized by fine tuning the five axis positioner.

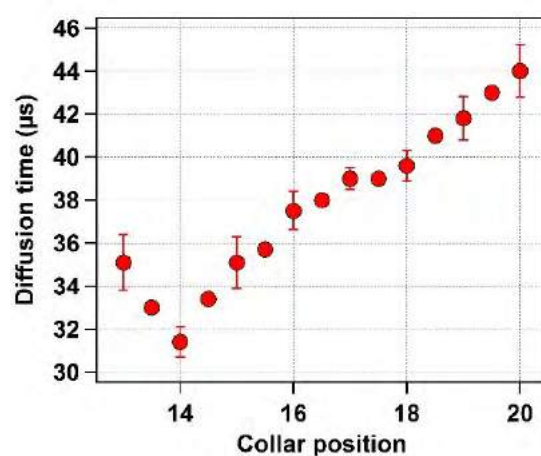
**(c) Achieving the Tiny Detection Volume:** To maximize the efficiency of detection one must achieve the tiniest detection volume possible after the proper



**Scheme 2.16:** Schematic of controlling the focal volume by changing beam diameter.

alignment of the system. The diameter of the excitation beam plays a critical role in the volume of the excitation. More the diameter, less is the excitation volume (scheme 2.16). For achieving the smallest focal volume, all the experiments have been performed with an overfilled objective backaperture. In order to calibrate the instrument for tiny focal volume a dye

solution having concentration  $\sim 5$  nM is kept as sample and focal volume is scanned against the objective collar setting and height of the focal point from the surface of the cover slip. The purpose of the objective collar is to adjust the focal volume according to the refractive index of the sample for the specific excitation wavelength for the solvent under concern. This reading is constant for a constant value of the refractive index. Figure 2.2 shows the dependence of diffusion time of rhodamine 6G in nanomolar aqueous solution on the collar position. As evident, the lowest diffusion

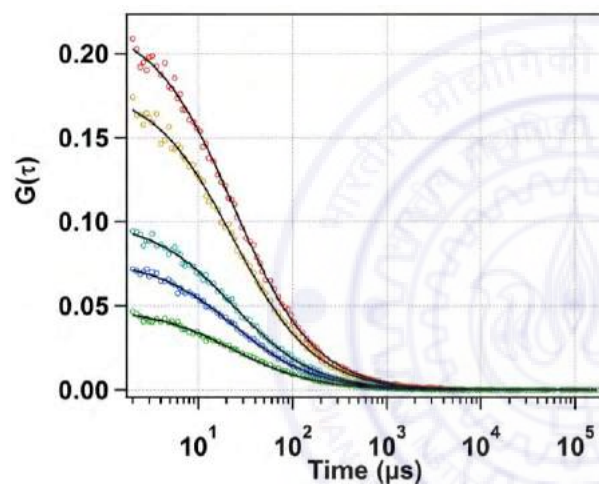


**Figure 2.2:** Diffusion time of rhodamine 6G in water as a function of collar position.

time is attained when the collar position is set at 14. The height of the focal point from the surface of the cover slip can be controlled by the knob provided in the microscope to vary the z-axis of the objective. Both the upper and lower surfaces of the cover slip can be located by finding a sudden jump of the detected photon counts. It is preferable to fix the focal point at a height where no

interference from the slide is experienced. For the current measurements, I have maintained a distance of 40 micrometer, which has been optimized as per abovementioned subject.

**(d) Measurement of Dimensions of the Observation Volume:** To calculate the Gaussian shaped observation volume (scheme 2.15) a dye having known diffusion coefficient is chosen and several autocorrelation curves with varying dye concentrations is recorded. These curves are then globally fitted with the one component diffusion equation. The global fitting yields the diffusion time (which is independent of the concentration and hence same for all the curves) and the ratio of the radial and axial diameter of the observation volume ( $w_{rz}$ ) which is an instrumental



**Figure 2.3:** Global fitting of autocorrelation curves using equation 2.27 in order to measure the dimensions of the observation volume.

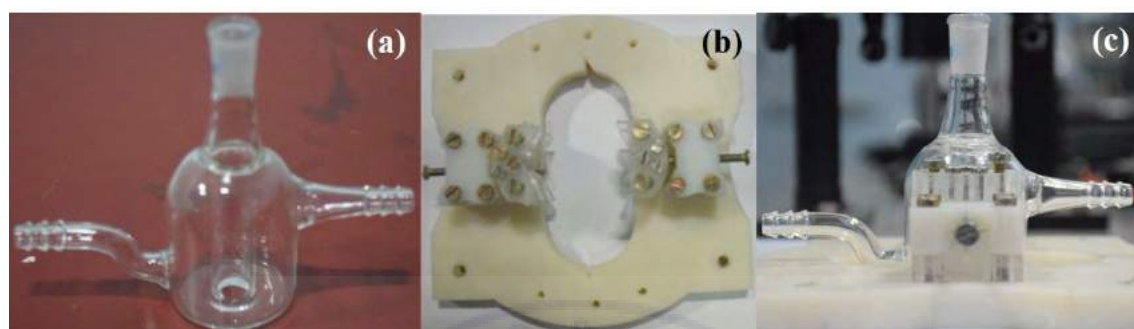
parameter. From the known diffusion coefficient of the standard dye one can calculate the radial diameter of the observation volume using the Stokes-Einstein equation. Following, the axial radius can be calculated using the relationship between  $w$ ,  $r_0$  and  $z_0$ , which are further used to estimate the observation volume using equation 2.24. During my work, I have used

rhodamine 6G (R6G) in water as a standard (diffusion coefficient of  $4.14 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$ ). Figure 2.3 shows the autocorrelation curves of aqueous R6G solution having different concentration and their global fitting lines.

**(e) Checking the Effect of Power:** Additional fluctuation of fluorescence intensity caused by photobleaching and triplet state conversion at high excitation powers can be detrimental for FCS measurements. To avoid these interferences, the laser power must be optimized depending on the fluorophore under concern. For this purpose, prior to running any experiment, the corresponding system is exposed to different laser powers and their autocorrelation curves are recorded. For highly photostable systems such as rhodamine the tolerance level was high whereas for others, above a certain power, the

autocorrelation curves starts getting effected by change in excitation power. So, experiments involving these dyes having low tolerance to high excitation power must be performed at or below the threshold laser power.

**2.1.5.9. Temperature Dependent FCS Measurements:** Temperature dependent FCS measurements have been performed in a specially prepared glass cell (figure 2.4). The



**Figure 2.4:** Temperature dependent FCS cell. (a) sample cell, (b) holder, (c) cell attached to the holder.

cell consists of a sample chamber with a water jacket surrounding it. In the water jacket there are one inlet and one outlet for water circulation. Labocon LLCB-202 temperature controller unit has been used to circulate water of a specific temperature through the water jacket to control the temperature of the sample cell. This cell can be attached firmly to the sample platform by an in home prepared holder.

**2.2. Calculation of Thermodynamic Parameters Related to Unfolding and Intermediate Formation:** As already discussed in chapter 1, protein unfolding is a complex process and may involve several steps. In the next three chapters the chemical and thermal unfolding of three different proteins has been discussed. To reckon the free energy associated with the structural changes the method described by Naidu *et al.* has been adopted.<sup>47</sup>

To start with, a model needs to be used depending on the nature of the structural alteration. For a simple one step or direct unfolding model, with only the native ( $N$ ) and the unfolded ( $U$ ) state (i.e.  $N \leftrightarrow U$ ), the variation of any observable which shows regular change can be written as

$$S = \frac{S_N + (S_U e^{-X})}{1 + e^{-X}} \quad (\text{Eq. 2.38})$$



In the above equation,  $S$  denotes the observable,  $S_N$  is the value of the observable for the native state,  $S_U$  is the same for the unfolded state and  $X$  is denoted as

$$X = \frac{(\Delta G^0 - m[\text{denaturant}])}{RT} \quad (\text{Eq. 2.39})$$

where  $R$  and  $T$  bears the usual meaning as universal gas constant and temperature in K. In equations 2.38 and 2.39, the dependence of free energy change of denaturation ( $\Delta G$ ) with denaturant ( $D$ ) concentration is taken as  $\Delta G = \Delta G^0 - m[D]$ , where,  $\Delta G^0$  is the free energy change in the absence of denaturant and  $m$  is the slope of free energy change against the denaturant concentration.<sup>48</sup>

The midpoint of unfolding can be expressed as

$$[D]_{\frac{1}{2}} = \frac{\Delta G^0}{m} \quad (\text{Eq. 2.40})$$

For complex structural transitions, which involves  $n$  number of intermediate states, the model is reconstructed as,<sup>28, 48</sup>

$$N \leftrightarrow I_1 \leftrightarrow I_2 \leftrightarrow \dots I_n \leftrightarrow U$$

where  $N$ ,  $I$  and  $U$  denote native, intermediate and unfolded states. For this model the variation of signal can be fitted using the following equation

$$S = \frac{S_N + S_{I_1}e^{-X_1} + S_{I_2}e^{-X_2} + \dots S_{I_{n-1}}e^{-X_{n-1}} + S_Ue^{-X_{n+1}}}{1 + e^{-X_1} + e^{-X_2} + \dots e^{-X_n} + e^{-X_{n+1}}} \quad (\text{Eq. 2.41})$$

where  $X_1, X_2, \dots, X_{n+1}$  are denoted as

$$X_1 = \frac{\Delta G_1^0 - m_1[D]}{RT} \quad (\text{Eq. 2.42a})$$

$$X_2 = \frac{\Delta G_1^0 + \Delta G_2^0 - (m_1 + m_2)[D]}{RT}$$

.....

.....

$$X_{n+1} = \frac{\Delta G_1^0 + \Delta G_2^0 + \dots \Delta G_n^0 + \Delta G_{n+1}^0 - (m_1 + m_2 + \dots m_n + m_{n+1})[D]}{RT} \quad (\text{Eq. 2.42b})$$



**2.3. Materials:** Human serum albumin (HSA, essentially fatty acid free), papain, bovine  $\beta$ -lactoglobulin, bovine chymotrypsin, coumarin-343, 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM), tetramethylrhodamine-5-maleimide (TMR), guanidine hydrochloride (GnHCl), 4-dimethylamino pyridine (DMAP), N,N-dicyclohexylcarbodiimide (DCC), cysteine hydrochloride, di-sodium salt of ethylene diamine tetraacetic acid (Na<sub>2</sub>-EDTA) were purchased from Sigma-Aldrich, USA and used as received. N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA) was purchased from Molecular Probes, Inc. and used as received. Analytical grade di-sodium hydrogen phosphate and sodium di-hydrogen phosphate were purchased from Merck, India and used to prepare 50 mM buffer (pH 7.4). Docusate sodium salt (AOT) was purchased from Sigma and used after drying under vacuum for 24 hours each time before the solution preparation. Dialysis membrane tubing (12kDa cut-off) were purchased from Sigma-Aldrich, USA and used after removing the glycerol and sulfur compounds according to the procedure given by Sigma-Aldrich. Centrifugal filter units (Amicon Ultra, 10kDa cut-off) have been purchased from Merck Millipore, Germany. HPLC grade dimethyl sulfoxide (DMSO), dichloromethane (DCM) were purchased from S. D. Fine Chemicals Limited and used after distillation. Deuteriated chloroform (CDCl<sub>3</sub>) for NMR measurements was purchased from Sigma-Aldrich. Analytical grade sodium chloride and EDTA were purchased from Merck and used for the tagging reaction of papain with DACIA. Trypsin solution was generously provided by Dr. A. K. Thakur of biological sciences department, IIT Kanpur.

**2.4. Protein Labelling:** All the proteins under study were labelled following the standard procedure with minor modifications.<sup>48,49</sup> Briefly, labelling dye solution in DMSO was added drop by drop to protein stock solution in phosphate buffer (50 mM, pH 7.4). Upon completion of reaction, the reaction mixture was dialysed against DMSO:buffer mixture (having the same DMSO: buffer ratio as in the reaction mixture) at 4°C to remove unreacted dye. The dialysis solution was replaced 4 times at an interval of 12 h. Then the solution was replaced with pure phosphate buffer (50 mM, pH 7.4) several times during the next 6-8 days. The dialysis solution was checked each

time for the presence of unreacted dye molecules. Dialysis was finally stopped when the dialysis solution was found totally free of any dye. The labelled protein was concentrated using the 10 kDa cutoff centrifugal filtration unit and lyophilized to obtain the solid labelled protein.

Cys-34 of HSA, Cys-25 of papain and Cys-121 of  $\beta$ -lactoglobulin have free thiol groups, which was tagged with thiol reactive fluorophores like N-dansylaziridine, bodipy-derivatives, 2-(4'-Maleimidylanilino)naphthalene-6-sulfonate, *etc.* as reported by several groups.<sup>50-57</sup> In the present study, HSA has been tagged with tetramethylrhodamine-5-maleimide (TMR) at Cys-34. Papain was tagged at Cys-25 with thiol reactive N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA). The Cys-121 of  $\beta$ -lactoglobulin has been tagged with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM). Besides this, for domain specific study the Tyr-411 of the domain III of HSA has been tagged with hydroxyl reactive *p*-nitrophenylcoumarin ester (NPCE). All the three thiol reactive dyes were commercially available whereas NPCE was designed and synthesized in our laboratory.

#### 2.4.1. HSA Labelling with *p*-nitrophenylcoumarin ester (NPCE)

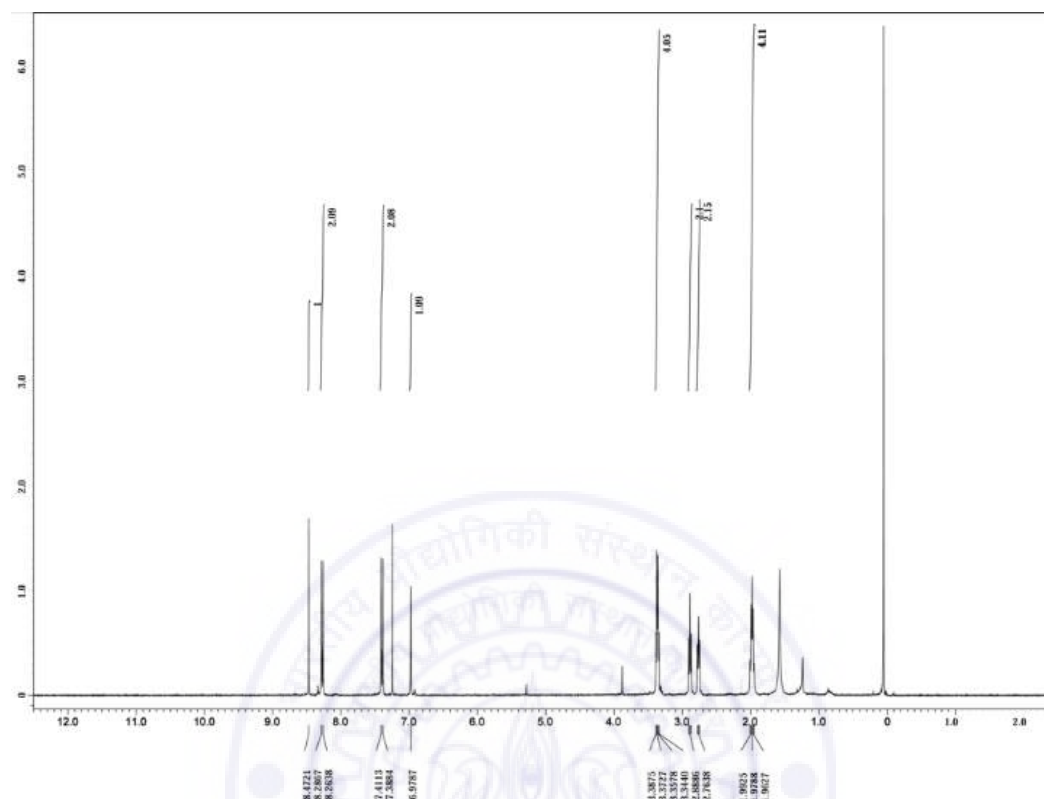
##### 2.4.1.1. Synthesis and characterization of *p*-nitrophenylcoumarin ester (NPCE):

NPCE was synthesized following the common procedure of esterification.<sup>61</sup> The schematic of the reaction has been shown as scheme 2.17. In brief, coumarin-343 (0.38 mmol), *p*-nitrophenol (0.38 mmol) and 4-dimethylamino pyridine (DMAP) (0.38 mmol) were taken in 5 mL DCM and stirred in ice bath for 10 min. 0.38 mmol of N,N-dicyclohexylcarbodiimide (DCC) was then added to the reaction mixture under nitrogen atmosphere with stirring. The reaction mixture was stirred at 0°C for 20 minutes and then at 20°C for 24 hours. The organic layer was washed successively with 35 mL of 1.2 M HCl and 35 mL saturated NaHCO<sub>3</sub> solution and dried over MgSO<sub>4</sub>. The residue was suspended in MeOH and the precipitate was washed with methanol to

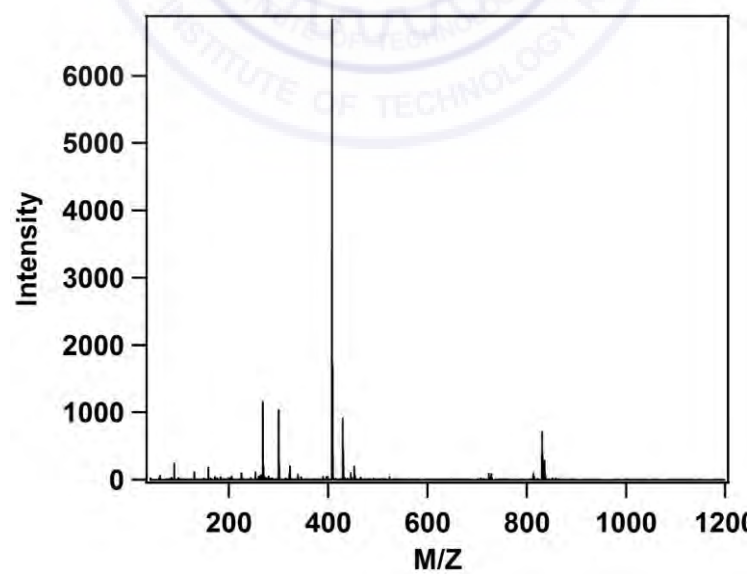


**Scheme 2.17:** Synthetic scheme of *p*-nitrophenylcoumarin ester (NPCE).

remove the unreacted coumarin-343. The precipitate was collected in DCM and characterized by  $^1\text{H}$ -NMR, IR and mass spectrometric studies.

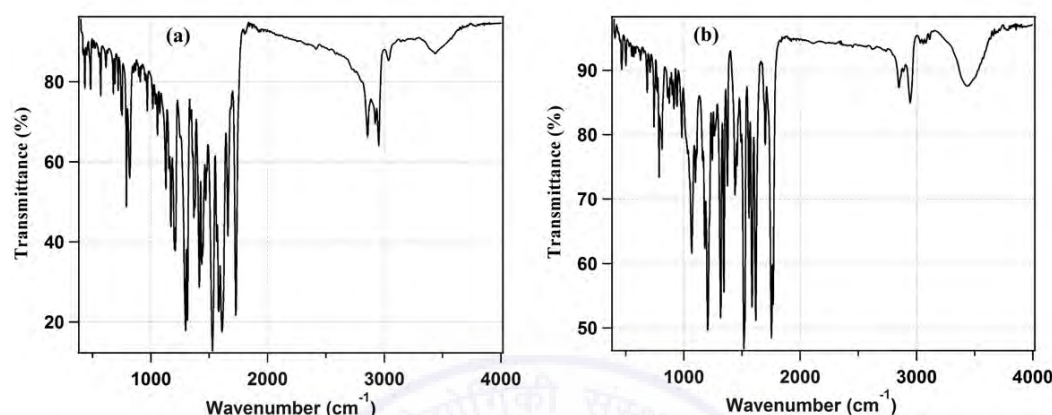


**Figure 2.5:**  $^1\text{H}$  NMR spectra of *p*-nitrophenylcoumarin ester (NPCE), 400MHz, solvent  $\text{CDCl}_3$ .



**Figure 2.6:** Mass analysis of *p*-nitrophenylcoumarin ester (NPCE); inset shows the comparison of experimental and calculated mass.

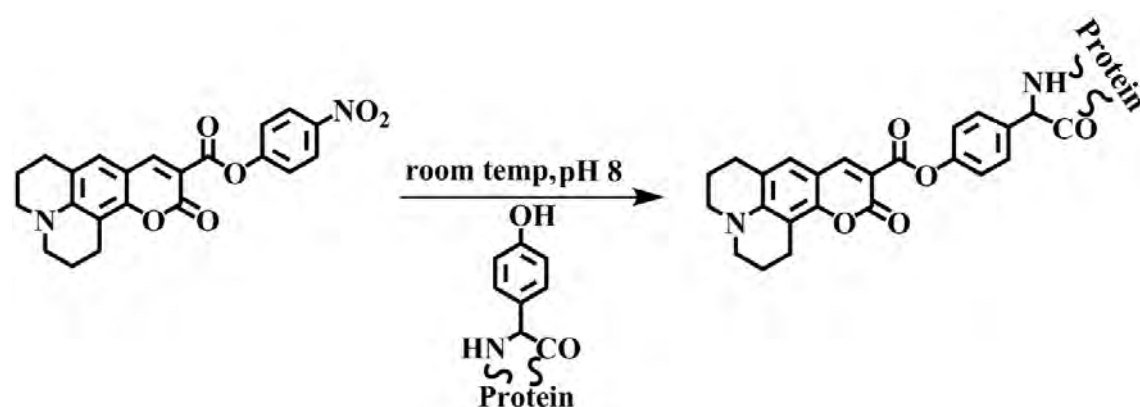
$^1\text{H}$ -NMR spectra (figure 2.5) shows two doublet peaks at ca. 8.28 and 7.48 ppm, respectively, which clearly indicate the formation of NPCE. Other peaks match closely with that of the parent compound to confirm the whole structure of the molecule. ESI mass analysis (figure 2.6) gives a strong peak at 407.1242 and relatively weak peaks at



**Figure 2.7:** IR spectra of (a) Coumarin 343, (b) *p*-nitrophenylcoumarin ester (NPCE).

408.1270 and 409.1300 for  $[\text{M}+\text{H}^+]$ . This is in accordance with the expected result considering the isotopic distribution as predicted by ChemDraw Ultra 11.0 from PerkinElmer, USA and is shown in the inset of figure 2.6. IR spectra of coumarin 343 shows a characteristic peak at  $1728\text{ cm}^{-1}$  for carbonyl group which breaks to two peaks with an increase in wavenumber to  $1753\text{ cm}^{-1}$  and  $1768\text{ cm}^{-1}$  when it forms the ester (see figure 2.7).

**2.4.1.2. Labelling of HSA with NPCE:** 40 mg of HSA was dissolved in 9 mL of 50 mM phosphate buffer (pH=8.0) and 1 mL of corresponding dye solution in DMSO was added slowly into it such that the molar ratio of protein and dye in the reaction mixture

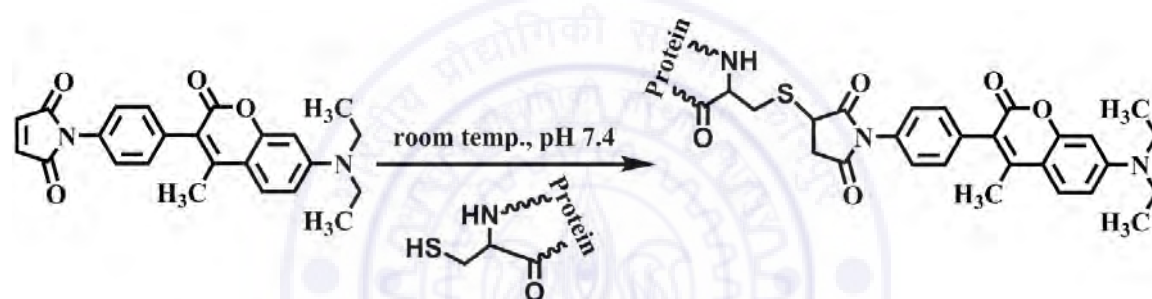


**Scheme 2.18:** Tagging of HSA with *p*-nitrophenylcoumarin ester (NPCE).



is 6:6.5. The reaction mixture was stirred at room temperature for 24 hrs followed by dialysis. The scheme of the reaction is provided below (scheme 2.18).

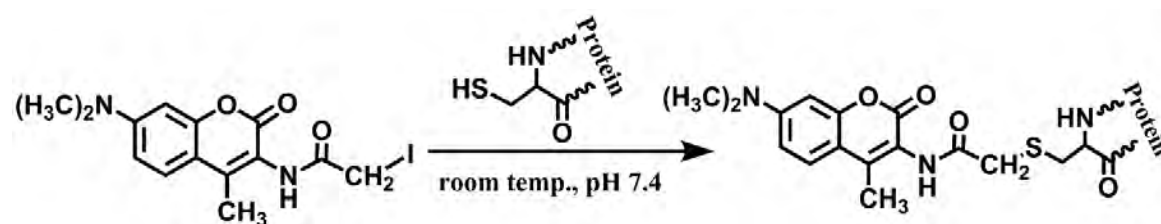
**2.4.2.  $\beta$ -lactoglobulin Labeling with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM):** In the present study, I have tagged the Cys-121 of  $\beta$ -lactoglobulin with CPM using the common procedure of CPM tagging (scheme 2.19).<sup>51</sup> Briefly, 50 mg of  $\beta$ -lactoglobulin was dissolved in 4 ml 50 mM phosphate buffer (pH 7.4) and then 400  $\mu$ l of 5.5 mM CPM in DMSO was added dropwise at room temperature and stirred mildly for 12 hours. The protein:dye ratio was maintained as 1:1. The reaction mixture was then dialyzed against buffer for 10 days at 4°C. The completion of dialysis was checked by recording the fluorescence of the dialysis buffer.



**Scheme 2.19:** Tagging  $\beta$ -lactoglobulin with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM).

The tagging ratio was calculated using the molar extinction coefficient for  $\beta$ -lactoglobulin ( $\epsilon_{280}=17,600 \text{ M}^{-1} \text{ cm}^{-1}$ ) and CPM ( $\epsilon_{390}=33,000 \text{ M}^{-1} \text{ cm}^{-1}$ ) and found to be 1:0.47.<sup>62,63</sup>

**2.4.3. Papain Labelling with DACIA:** Papain was labelled with DACIA following the procedure described by Lindahl *et al.* (scheme 2.20).<sup>49</sup> Papain (400  $\mu$ M) was dissolved in 9 mL of 50 mM tris-HCl buffer (pH 7.4) containing 0.1 M NaCl and 100  $\mu$ M EDTA. DACIA (200  $\mu$ M; 0.58 mg) was dissolved in 1 mL of DMSO and added

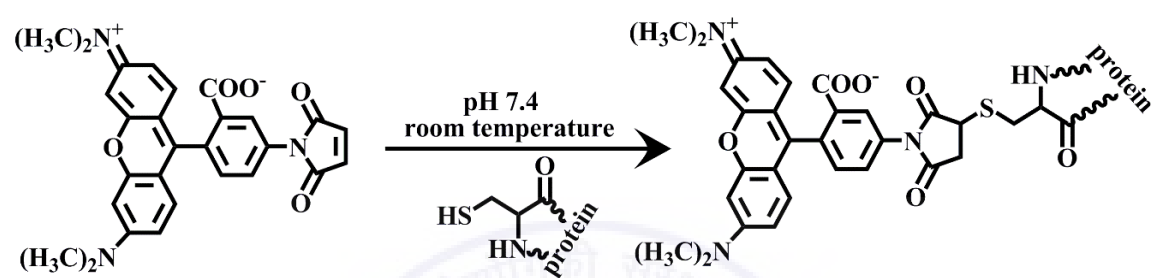


**Scheme 2.20:** Tagging of papain with N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA).



drop wise to the stirred solution of papain in buffer. DACIA is light sensitive, so the reaction was performed under low-light conditions. The reaction mixture was stirred at 20-22°C for two hours followed by dialysis.

**2.4.4. HSA Labelling with TMR:** To a stock solution of HSA (90  $\mu\text{M}$ ) in phosphate buffer (50 mM, pH 7.4), a stock solution of TMR in DMSO was added under stirring condition with a final HSA:dye molar ratio of 1:3.5. The tagging reaction has been



**Scheme 2.21:** Tagging of HSA with tetramethylrhodamine-5-maleimide (TMR).

shown in scheme 2.21. After the completion of labelling, the HSA:TMR ratio was estimated as 1:0.6 from the known value of molar extinction coefficients (for HSA  $\epsilon_{280} = 36500 \text{ M}^{-1} \text{ cm}^{-1}$  and for TMR  $\epsilon_{555} = 75000 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>58-60</sup>

**2.5. Sample preparation:** For ensemble measurements, protein concentration was maintained at micromolar level unless stated otherwise. A concentrated stock solution of the tagged protein was prepared and the concentration was measured from its absorbance at 280 nm. The extinction coefficient of all the proteins under study are well reported at 280 nm. For denaturation experiments, required amount of stock solutions of protein, denaturant and buffer solution was mixed and incubated for 2 hours before performing any experiment. FCS measurements were performed at nanomolar protein concentration. For FCS measurements, the operational concentration is crucial because too high concentration leads to a highly occupied focal volume giving rise to high background signal; on the other hand at too low concentration the number of photons reaching the detector will be less making the output erroneous. So, before preparation of sample for FCS measurement, the ideal concentration of the fluorophore under which the experiment should be performed is determined by recording the FCS curves of the sample at different concentration

ranges. Typically,  $G(0)$  was tried to be maintained at one for ensuring the average number of molecule(s) inside the focal volume is unity. AOT reverse micelles of different water-pool sizes was prepared by injecting required amount of protein sample in 25 mM AOT solution in iso-octane. All the experiments were done at  $298 \pm 1$  K if not stated otherwise.

For digestion reaction, papain was activated using cysteine and EDTA. To 0.9 ml solution of 1 mM EDTA and 10 mM cysteine hydrochloride at pH 7 buffer, 0.1 ml papain solution (10 mg/ml in pH 7 phosphate buffer) was added and the mixture was heated at 310 K for 10 minutes. Chymotrypsin and trypsin was used as it received. The digestion reaction was performed in water bath with 1  $\mu$ M HSA solution in presence appropriate amount of enzyme (stock solution in buffer) to maintain 1:1 ratio of HSA:enzyme. Both the solution was brought to the operational temperature before mixing. Immediately after addition of papain to HSA, the stopwatch was started. At definite time intervals, 20  $\mu$ l aliquot from the solution was taken and diluted to 200  $\mu$ l with chilled phosphate buffer in order to arrest the reaction. This solutions (100 nM with respect to HSA) were subjected to FCS experiments after bringing them to room temperature.

**2.6. Auxiliary instruments:** NMR spectra in  $\text{CDCl}_3$  were recorded on a 400 MHz NMR spectrometer (JEOL, Japan). FT-IR spectra were recorded in KBr pellets on a FTIR spectrometer (Bruker Vector 22). ESI-MS spectra were recorded on a Waters Micromass Quattro Micro triplequadrupole mass spectrometer. Centrifugation was done in eppendorf centrifuge 5810R, at 5000 rpm. Lyophilisation was done in Heto VR-1 setup.

## References

1. Owen, T. *Fundamentals of Modern UV-Visible Spectroscopy*. Agilent Technologies, Germany **2000**.
2. Tkachenko, N. V. *Optical Spectroscopy Methods and Instrumentations*, Elsevier **2006**.
3. Elumalai, P., Atkins, P., de Paula, J. *Atkins' Physical Chemistry*, Oxford University Press, **2002**.
4. Jabłoński, A. *Nature* **1933**, *131*, 839–840.
5. Rodger, A. Beer-Lambert Law Derivation. In *Encyclopedia of Biophysics*, Springer Berlin Heidelberg: Berlin, Heidelberg, **2013**, 184–185.
6. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd Edition, Springer **2006**.
7. Chen, R. F. *J. Res. Natl. Bur. Stand.* **1972**, *76*, 593–606.
8. Jones II, G.; Jackson, W. R.; Choi, C.; Bergmark, W. R. *J. Phys. Chem.* **1985**, *89*, 294–300.
9. Whitmore, L.; Wallace, B. A. *Biopolymers* **2007**, *89*, 392–400.
10. Johnson, W. C. Jr. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 145–166.
11. R. W. Woody, *Theory of circular dichroism of proteins*. In *circular dichroism and conformational analysis of biomolecules*, G. D. Fasman, Ed. Plenum publishing corp., New York, **1996**.
12. Corrêa, D. H. A.; Ramos, C. H. I. *African J. Biochem. Res.* **2009**, *3*, 164–173.
13. Greenfield, N. J. *Nat. Protoc.* **2006**, *1*, 2876–2890.
14. O' Connor, D. V.; Phillips, D. *Time Correlated Single Photon Counting*, Academic Press: New York, **1984**.
15. Becker, W. *Advanced Time-Correlated Single Photon Counting Techniques*, Springer, **2005**.
16. Fleming, G. R. *Chemical Applications of Ultrafast Spectroscopy*, Oxford University Press: New York, **1986**.
17. Bevington, P. R. *Data Reduction and Error Analyses for the Physical Sciences*, McGraw-Hill, **1969**.
18. Eastern, J. H.; DeTomma, R. P.; Brand, L. *Biophys. J.* **1976**, *16*, 571–583.

19. Schwille, P. *In Fluorescence correlation spectroscopy. Theory and applications*, Elson, E. L.; Rigler, R. Eds., Springer, Berlin, **2001**.
20. Magde, D.; Elson, E. L.; Webb, W. W. *Biopolymers* **1974**, *13*, 29–61.
21. Elson, E. L.; Magde, D. *Biopolymers* **1974**, *13*, 1–27.
22. Magde, D.; Elson, E.; Webb, W. W. *Phys. Rev. Lett.* **1972**, *29*, 705–708.
23. Koppel, D. E.; Axelrod, D.; Schlessinger, J.; Elson, E. L.; Webb, W. W. *Biophys. J.* **1976**, *16*, 1315–1329.
24. Magde, D.; Webb, W. W.; Elson, E. L. *Biopolymers* **1978**, *17*, 361–376.
25. Fahey, P. F.; Koppel, D. E.; Barak, L. S.; Wolf, D. E.; Elson, E. L.; Webb, W. W. *Science* **1977**, *195*, 305–306.
26. Ehrenberg, M.; Rigler, R. *Chemical Phys.* **1974**, *4*, 390–401.
27. Rigler, R.; Mets, Ü.; Widengren, J.; Kask, P. *Eur. Biophys. J.* **1993**, *22*, 169–175.
28. Eigen, M.; Rigler, R. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5740–5747.
29. Rigler, R. *J. Biotechnol.* **1995**, *41*, 177–186.
30. Rigler, R.; Widengren, J.; Mets, Ü. *Fluorescence Spectroscopy: New Methods an Applications*, Ed. Wolfbeis, O S., Berlin, Springer **1992**, pp- 13.
31. Widengren, J.; Rigler, R. *Cell. Mol. Biol.* **1998**, *44*, 857–879.
32. Schlessinger, J.; Koppel, D. E.; Axelrod, D.; Jacobson, K.; Webb, W. W.; Elson, E. L. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 2409–2413.
33. Ehrenberg, M.; Rigler, R. *Q. Rev. Biophys.* **2009**, *9*, 69–81.
34. Maiti, S.; Haupts, U.; Webb, W. W. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11753–11757.
35. Wohland, T.; Friedrich, K.; Hovius, R.; Vogel, H. *Biochemistry* **1999**, *38*, 8671–8681.
36. Hess, S. T.; Huang, S.; Heikal, A. A.; Webb, W. W. *Biochemistry* **2002**, *41*, 697–705.
37. Sánchez, S. A.; Gratton, E. *Acc. Chem. Res.* **2005**, *38*, 469–477.
38. Haustein, E.; Schwille, P. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 151–169.
39. Li, X.; Xing, J.; Qiu, Z.; He, Q.; Lin, J. *Molecular Plant* **2016**, *9*, 1229–1239.

40. Schwille, P.; Kummer, S.; Heikal, A. A.; Moerner, W. E.; Webb, W. W. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 151–156.
41. Schuler, B.; Hofmann, H. *Curr. Opin. Struct. Biol.* **2013**, *23*, 36–47.
42. Herrick-Davis, K.; Grinde, E.; Cowan, A.; Mazurkiewicz, J. E. *Mol. Pharmacol.* **2013**, *84*, 630–642.
43. Kaur, G.; Costa, M. W.; Nefzger, C. M.; Silva, J.; Fierro-González, J. C.; Polo, J. M.; Bell, T. D. M.; Plachta, N. *Nat. Commun.* **2013**, *4*, 1637.
44. Nag, M.; Das, D.; Bandyopadhyay, D.; Basak, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19139–19148.
45. Widengren, J.; Mets, Ü. *Conceptual basis of fluorescence correlatrion spectroscopy and related techniques at tools in biosciences. In single molecule detection in solution*, Ed. Ch. Zander, J Enderlein, RA Keller. Wiley-VCD, Darmstadt, Germany, **2002**, pp 69–120.
46. Müller, C. B.; Loman, A.; Pacheco, V.; Koberling, F.; Willbold, D.; Richtering, W.; Enderlein, J. *EPL* **2008**, *83*, 46001(1)–46001(5).
47. K.T. Naidum, N.P. Prabhu, Protein\_Surfactant Interaction: Sodium Dodecyl Sulfate-Induced Unfolding of Ribonuclease A. *J. Phys. Chem. B* **2011**, *115*, 14760–14767.
48. Wang, R.; Sun, S.; Bekos, E. J.; Bright, F. V. *Anal. Chem.* **1995**, *67*, 149–159.
49. Lindahl, P.; Raub-Segall, E.; Olson, S. T.; Björk, I. *Biochem. J.* **1991**, *276*, 387–94.
50. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
51. Sasmal, D. K.; Mondal, T.; Mojumdar, S. S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
52. Flora, K.; Brennan, J. D.; Baker, G. A.; Doody, M. A.; Bright, F. V. *Biophys. J.* **1998**, *75*, 1084–1096.
53. Scouten, W. H.; Lubcher, R.; Baughman, W. *Biochim. Biophys. Acta.* **1974**, *336*, 421–426.
54. Tyagarajan, K.; Pretzer, E.; Wiktorowicz, J. E. *Electrophoresis* **2003**, *24*, 2348–2358.



55. Stapelfeldt, H.; Olsen, C. E.; Skibsted, L. H. *J. Agric. Food. Chem.* **1999**, *47*, 3986–3990.
56. Puljung, M. C.; Zagotta, W. N. *Curr. Protoc. Protein Sci.* **2012**, *14*, 1–10.
57. Wang, J.; Xiang, Y. F.; Lim, C. *Protein Eng.* **1994**, *7*, 75–82.
58. Beaven, G. H.; Chen, S.-H.; D'Albis, A.; Gratzer, W. B. *Euro. J. Biochem.* **1974**, *41*, 539–546.
59. 2.3 Thiol-reactive probes excited with ultraviolet light, Invitrogen.
60. Kwon, G.; Remmers, A. E.; Datta, S.; Neubig, R. R. *Biochemistry* **1993**, *32*, 2401–2408.
61. Aujard, I.; Benbrahim, C.; Gouget, M.; Ruel, O.; Baudin, J. B.; Neveu, P.; Jullien, L. *Chem. Eur. J.* **2006**, *12*, 6865–6879.
62. Albani, J. R.; Vogelaer, J.; Bretesche, L.; Kmiecik, D. *J. Pharm. Biomed. Anal.* **2014**, *91*, 144–150.
63. Kojima, K.; Kitada, S.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **2001**, *276*, 2115–2121.

## Chapter 3

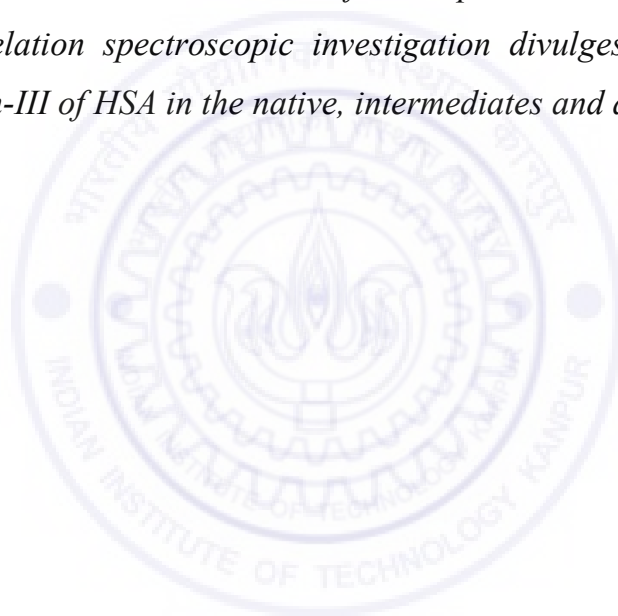
---

### *Elucidation of $\mu$ s Dynamics of Domain-III of Human Serum Albumin during the Chemical and Thermal Unfolding*

Sengupta, B.; Acharyya, A.; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, 18, 28548–28555.

Sengupta, B.; Das, N.; Sen, P. *Biophys. Chem.* **2016**, 221, 17–25.

*The local structural dynamics and denaturation profile of domain-III of HSA against guanidine hydrochloride (GnHCl) and temperature has been studied using a newly synthesized coumarin based fluorescent probe p-nitrophenylcoumarin ester (NPCE), covalently tagged to Tyr-411 residue. By the steady-state, time-resolved and single molecular level fluorescence studies it has been established that the domain-III of HSA is very sensitive to GnHCl but somewhat resistant to temperature and the domain specific unfolding proceeds in an altered way as compared to the overall unfolding of HSA. While the overall denaturation of HSA is a two-state process for both GnHCl and heat, domain-III adopts two intermediate states for GnHCl induced denaturation and one intermediate state for temperature induced denaturation. Fluorescence correlation spectroscopic investigation divulges the conformational dynamics of domain-III of HSA in the native, intermediates and denatured state.*

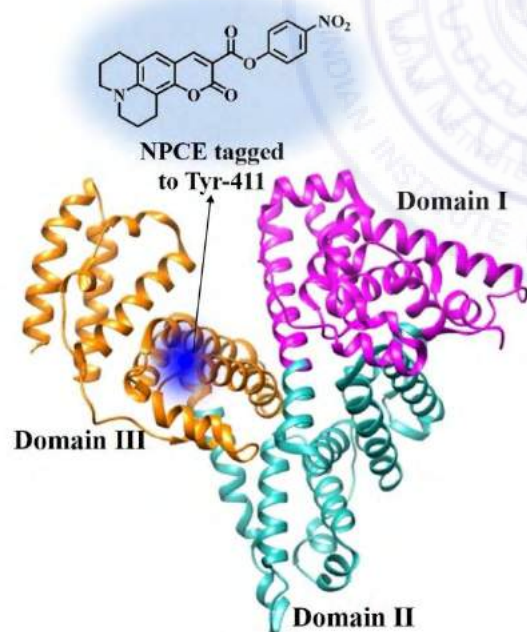


### 3.1. Introduction

Human serum albumin (HSA) is the most abundant transport protein in human body having three distinct domains.<sup>1-13</sup> It transports different types of metabolites including hormones, drugs, fatty acids, *etc.* and has significant esterase activity.<sup>5-7</sup> Fluorescence techniques have been employed to understand the structural and functional details of serum albumins since long time.<sup>14-25</sup> The single tryptophan residue, which is located at the domain-II facilitates spectroscopic investigation of domain-II by intrinsic fluorescence of tryptophan, while scrutiny of other two domains by fluorescence methods require attachment of external fluorophores to the respective domains.<sup>26-27</sup> Non-covalent probes are less preferred as compared to covalent ones because there is always a chance that the non-covalent probe may slip out of the protein core under structural perturbation. Here it is to note that the covalent fluorescent labelling can provide specific information of the tagging site within the protein. The covalent attachment of thiol specific fluorescent tagging agents to the single free cysteine residue in domain-I of HSA makes the study of domain-I feasible.<sup>28-33</sup> Yadav *et al.* have tagged domain-I of HSA with tetramethylrhodamine-5-maleimide (TMR) and N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA) to study the unfolding and conformational dynamics pertaining to domain-I.<sup>28-30</sup> Other thiol specific probes such as 7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM), acrylodan are also reported to tag domain-I of HSA.<sup>31-33</sup> However, domain-III of HSA is not studied thoroughly, in spite of holding important role in ligand binding and esterase activity, probably due to the limited scope of site specific labelling of this domain.

The denaturation of HSA has been studied by various groups worldwide.<sup>34-38</sup> Being a multi-domain protein, HSA shows different denaturation profiles for different domains.<sup>17</sup> In a recent review by Anand *et al.*, the overall unfolding and refolding pathways of serum albumins are summarized.<sup>19</sup> Bright and co-workers have studied the dynamics of acrylodan labeled HSA in the course of thermal and chemical denaturation and reported the presence of intermediate states in the unfolding process.<sup>31</sup> Ahmad *et al.* proposed that chemical denaturation of HSA initiates by a local unfolding

of stable loops in domain-III.<sup>15</sup> In another work, Ahmad *et al.* reported that urea induced unfolding of fast moving isomer (F-isomer) of HSA does not proceed through any intermediate state, whereas N-isomer does.<sup>35</sup> Vlasova *et al.* confirmed the existence of two stages in the SDS induced denaturation of HSA. First one being the disintegration of globules, the second one involves the complete unfolding of the amino acid chain of HSA.<sup>21</sup> Analysing the intrinsic tryptophan fluorescence of HSA, they commented that both stages of denaturation is operational at a pH below the isoelectric point ( $pI = 4.7$ ) of HSA, but when the pH is above  $pI$ , denaturation stops in the first stage only. Picó adopted a model *native state*  $\leftrightarrow$  *unfolded reversible state*  $\rightarrow$  *unfolded irreversible state* to represent the thermal unfolding behaviour of HSA.<sup>39</sup> According to him, the irreversible change happens only above 347 K. Previously Yadav *et al.* have shown that in the presence of guanidine hydrochloride (GnHCl), domain-III unfolds first followed by domain-I and domain-II.<sup>17</sup> They have also elucidated the domain specific unfolding pathway of domain-I in the course of chemical and thermal denaturation using bulk and single molecular level spectroscopic studies.<sup>28</sup>



**Scheme 3.1:** Schematic representation of *p*-nitrophenylcoumarin ester (NPCE) and its adduct to the tyrosine-411 residue located in the domain-III of human serum albumin (PDB ID: 1HA2).

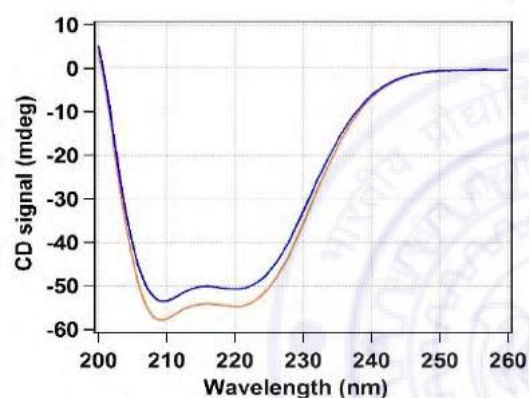
Here I report the unfolding characteristics of domain-III of HSA using NPCE as a site specific covalent tag. I have synthesized a tyrosine specific probe, *p*-nitrophenylcoumarin ester (NPCE) to selectively label tyrosine-411 of HSA residing in the domain-III (Scheme 3.1).<sup>40</sup> The details of synthesis and tagging has been provided in section 2.4.1 of chapter 2. By solvation dynamics, anisotropy and single molecular fluorescence experiments I have characterized the local environment within the domain-III of HSA. Steady-



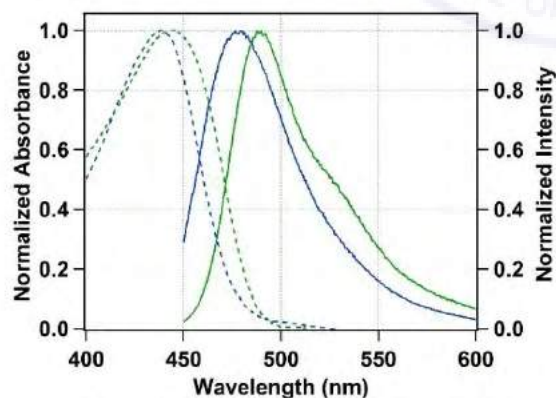
state, time-resolved and single molecular level fluorescence experiments have been performed to elucidate the change in conformational dynamics during the course of denaturation. The change in hydrodynamic radius has been monitored to understand the overall structural change whereas the specific spectroscopic signals from the attached fluorescence probe reflects the change in local environment around the probe which has been related to the local structural change.

### 3.2. Results

**3.2.1. Circular Dichroism Spectroscopy:** In order to confirm that the fluorescent tag



**Figure 3.1:** Circular dichroism spectra of untagged (4.36  $\mu$ M, brown) and NPCE tagged human serum albumin (4.11  $\mu$ M, blue) in phosphate buffer (pH 7.4, 50 mM).



**Figure 3.2:** Normalized absorption (broken lines) and emission (solid lines) spectra of *p*-nitrophenylcoumarin ester when free (green) and when tagged to human serum albumin (blue) in phosphate buffer (pH 7.4, 50 mM), (Excitation wavelength is 446 nm).

does not cause any changes in the secondary structure of HSA, I have recorded the circular dichroism (CD) spectra of HSA in 50 mM phosphate buffer (pH = 7.4) at the room temperature. Figure 3.1 shows the CD spectra of untagged and NPCE tagged HSA. The percentage of  $\alpha$ -helicity of the untagged HSA found to be 63% (using the method described in section 2.1.3 of chapter 2), which is slightly reduced to 61% when NPCE is tagged to HSA.

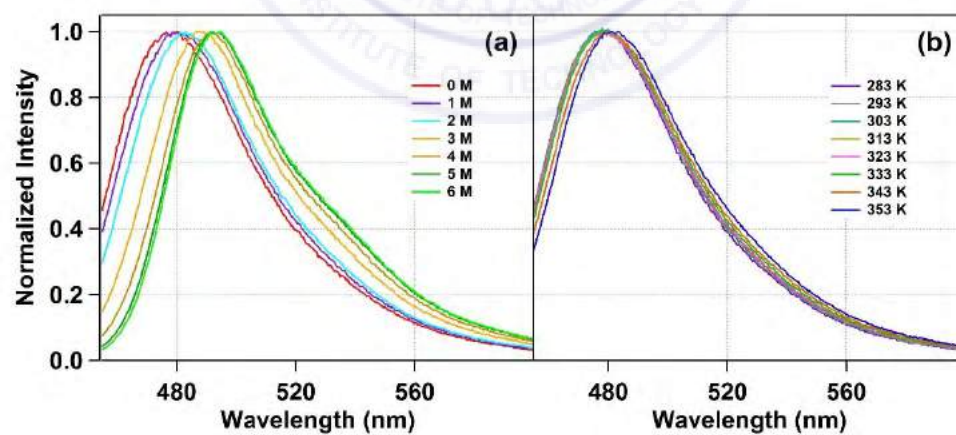
**3.2.2. Steady-State Absorption and Fluorescence Spectroscopy:** The steady-state spectroscopy results are summarized in table 3.1. The absorption and emission spectra of NPCE gives corresponding

maxima at 447 nm and 488 nm respectively in phosphate buffer medium (pH = 7.4, 50 mM) and has been depicted in figure 3.2. When NPCE is tagged in HSA, the

**Table 3.1:** Spectroscopic data for NPCE and NPCE tagged HSA in phosphate buffer (pH= 7.4, 50 mM)

	$(\lambda_{\text{abs}})_{\text{max}}$ (nm)	$(\lambda_{\text{em}})_{\text{max}}$ (nm)	Fluorescence quantum yield	Fluorescence lifetime (ns)
NPCE	447.0 $\pm$ 0.5	489.0 $\pm$ 1	0.45 $\pm$ 0.03	4.1 $\pm$ 0.02
NPCE tagged to HSA	440.0 $\pm$ 0.5	478.5 $\pm$ 1	0.33 $\pm$ 0.04	3.7 $\pm$ 0.02

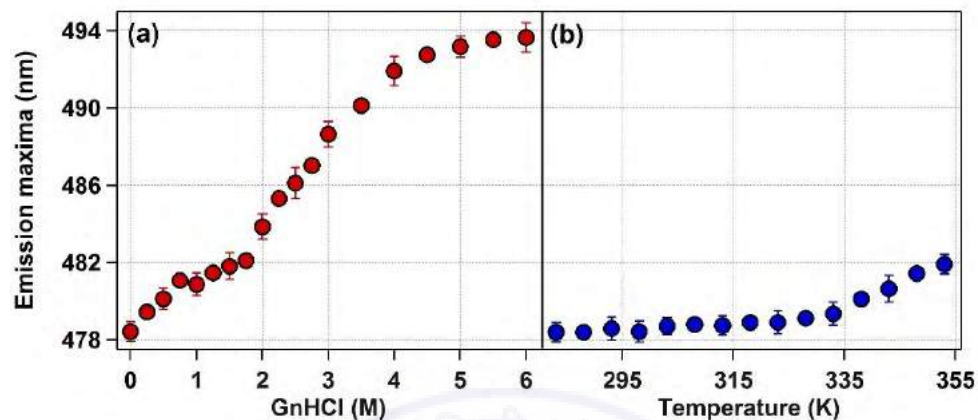
absorption and emission maxima change to 440 nm and 478 nm respectively. The quantum yield of NPCE is 0.45 in phosphate buffer (pH= 7.4, 50 mM), which decreases to 0.33 when tagged with HSA. To study unfolding, I have measured the emission spectra of the NPCE-HSA conjugate in presence of different concentrations of guanidine hydrochloride (GnHCl). Some of the normalized emission spectra has been shown in figure 3.3a. The observed variation of emission maxima of NPCE in HSA as a function of GnHCl concentration, as depicted in figure 3.4a, is in accordance with the expectation that owing to the denaturation the emission maxima should get red shifted due to its exposure to more hydrophilic environment. At 6 M GnHCl the emission maxima reaches a value of 493 nm. The extent of change of emission maxima decreases after 4.5 M GnHCl. It is to be noted that, although the shift is monotonous, there is a

**Figure 3.3:** Normalized emission spectra of NPCE tagged to HSA in phosphate buffer (pH 7.4, 50 mM) at (a) different GnHCl concentrations and (b) different temperatures ( $\lambda_{\text{ex}} = 442$  nm).

small hump around the 1-2 M GnHCl region.

The temperature dependent study of the steady-state emission property of NPCE tagged HSA is also performed in the range 283-353 K. The normalized emission

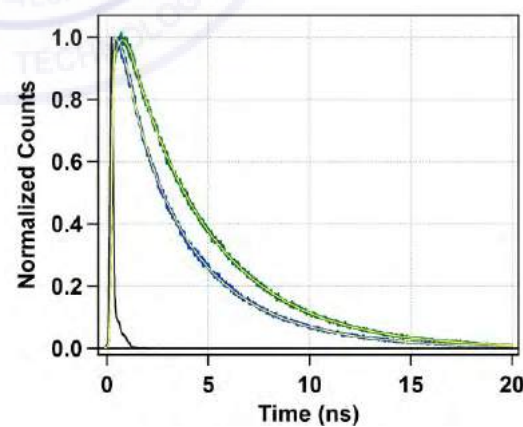
spectra for some of the temperatures are shown in figure 3.3b, whereas the change in emission maxima as the function of temperature is shown in figure 3.4b. I have observed no perceivable change in the emission maxima till 323 K, however, beyond



**Figure 3.4:** Variation of emission maxima of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) with (a) GnHCl and (b) temperature.

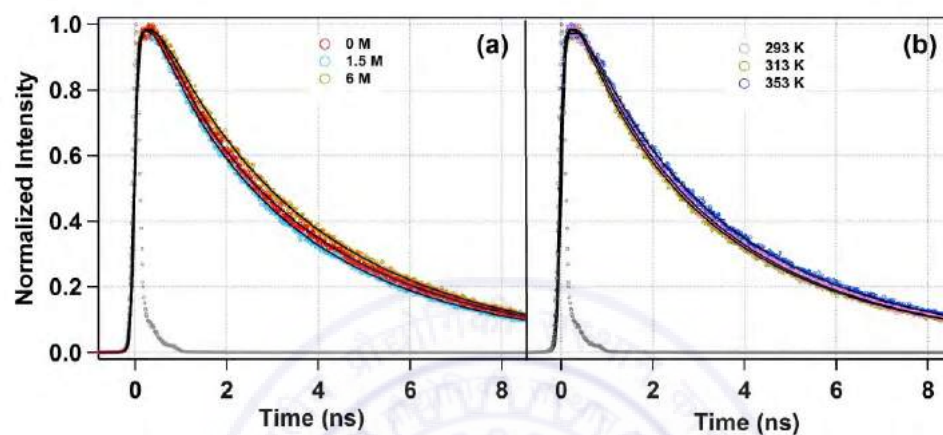
323 K a red shift has been observed and it becomes 482 nm at 353 K.

**3.2.3. Fluorescence Lifetime Measurement:** Fluorescence lifetimes of the NPCE in phosphate buffer can be fitted satisfactorily with single exponential decay function with a time constant of 4.1 ns. When tagged inside HSA, the fluorescence decay becomes multi-exponential in nature and the average lifetime of NPCE decreases to 3.7 ns. The fluorescence decays along with the fits have been shown in figure 3.5. Fluorescence transients were collected at 480 nm upon exciting at 442 nm. All the decays have been best fitted with three exponential decay function. In the native state, the average lifetime of NPCE is found to be  $\sim 3.7$  ns inside HSA with the time components as 0.47 ns (2.5%), 2.3 ns (26.8%), 4.33



**Figure 3.5:** Fluorescence decay profiles of *p*-nitrophenylcoumarin ester when free in buffer (green) and when tagged to human serum albumin (blue). Black solid line represents instrument response function (IRF). Solid yellow lines are the best fits of the fluorescent transient following deconvolution; ( $\lambda_{\text{ex}} = 442$  nm)

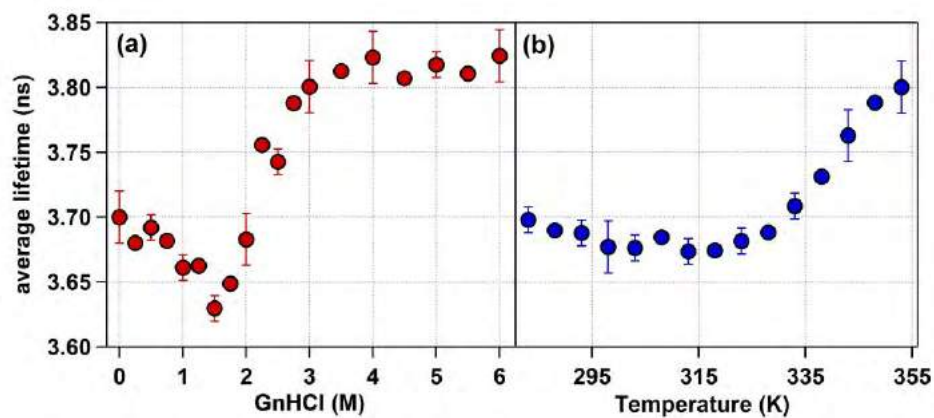
ns (70.7%). On treating with GnHCl, the lifetime decreases till 1.5 M GnHCl ( $\sim 3.6$  ns) and then increases monotonously to  $\sim 3.8$  ns at 6M GnHCl. The normalized fluorescence transients and corresponding fittings for 0 M, 1 M and 6 M GnHCl are shown in figure 3.6a and the variation of lifetime as a function of GnHCl concentration is shown in figure 3.7a. The change in average lifetime in this concentration range is



**Figure 3.6:** Fluorescence transients of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) (a) in presence of 0 M, 1.5 M, 6 M GnHCl and (b) at 293 K, 313 K and 353 K; ( $\lambda_{\text{ex}} = 442$  nm)

very small but a local minima is observed at around  $\sim 1.5$  M GnHCl.

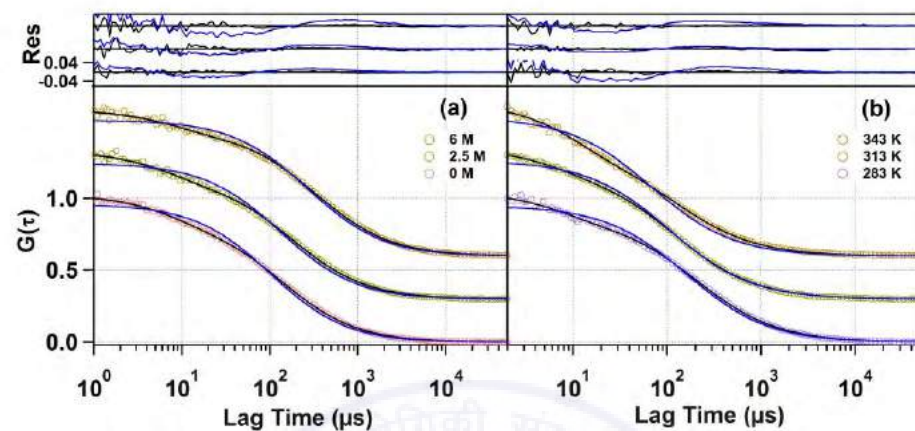
The fluorescence lifetime of NPCE tagged HSA is found to be 3.7 ns at 283 K, which decreases initially till 310 K and then increases to  $\sim 3.8$  ns at 353 K. The normalized data and the corresponding fitted lines for few temperatures are shown in figure 3.6b and the variation of lifetime with temperature is displayed in figure 3.7b. Overall, a shallow dip in the variation of average lifetime is observed around  $\sim 310$  K.



**Figure 3.7:** Variation of fluorescence lifetime of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) with (a) GnHCl and (b) temperature; ( $\lambda_{\text{ex}} = 442$  nm).

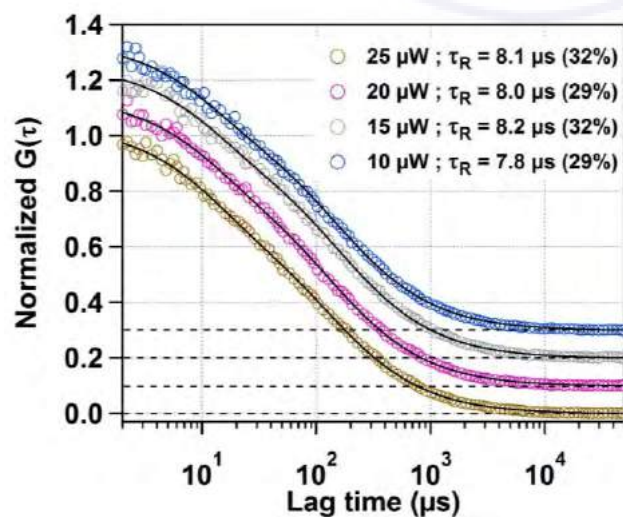


**3.2.4. Fluorescence Correlation Spectroscopic Measurements:** Fluorescence correlation spectroscopic (FCS) technique has been emerged as a powerful technique to monitor single molecular level properties in solution phase.<sup>41-44</sup> I have applied FCS



**Figure 3.8:** Fluorescence autocorrelation curves of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) (a) in presence of 0 M, 2.5 M, 6 M GnHCl and (b) at 283 K, 313 K and 353 K and comparison of fits with corresponding residuals of the autocorrelation curve using single diffusion equation (blue line) and one diffusion + one relaxation component equation (black line).

for further investigation of domain-III of HSA during chemical and thermal denaturation. In FCS, the self similarity of the fluorescence fluctuation collected from a very small sample volume is processed to get the information about its origin.<sup>46,47</sup> This can be diffusion over the small observation volume, quenching interactions, triplet state formation, *etc.* FCS study has been performed for the NPCE tagged HSA as a

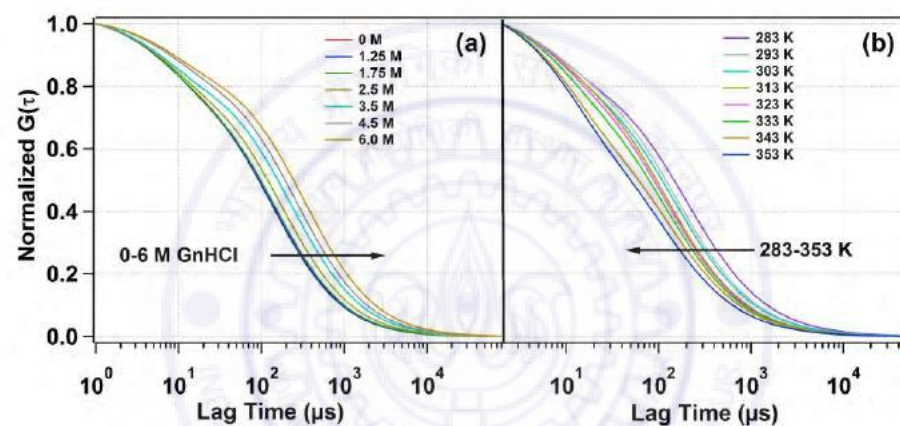


**Figure 3.9:** Effect of power on autocorrelation curves of HSA tagged with NPCE in buffer.

function of GnHCl concentration and temperature. In this case, the autocorrelation curve for the NPCE tagged HSA in the native state can not be fitted satisfactorily by a single component diffusion equation (equation 2.27), however, upon incorporation of a relaxation time component



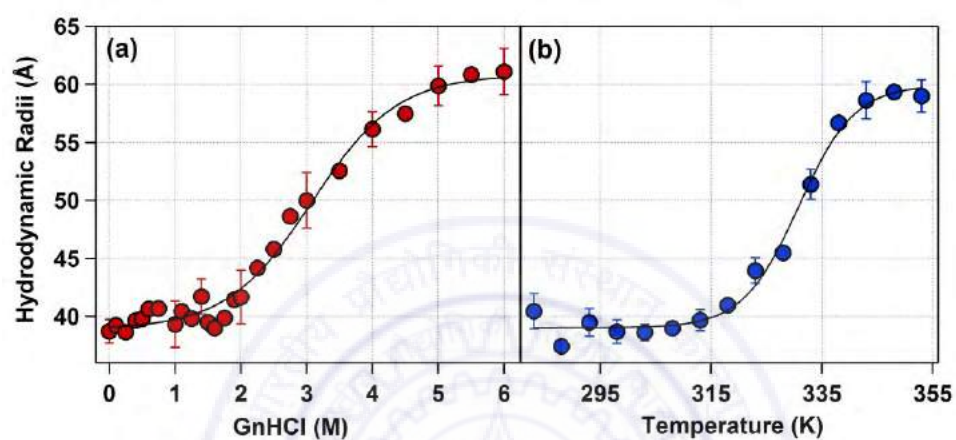
(equation 2.31) the fitting quality improves considerably. A comparison of fitting for a few autocorrelation curves in buffer and with 2.5 M, and 6 M GnHCl are shown in figure 3.8a. The inclusion of the relaxation time component actually takes care of the additional process other than diffusion, which is responsible for the fluctuation in fluorescence intensity. In order to confirm the origin of this time component we performed laser-power dependent FCS measurements on NPCE tagged HSA in buffer which are shown in figure 3.9. For a wide variation of laser power (10–25 mW) the observed relaxation time component and its amplitude remain unchanged indicating that the origin of this 8  $\mu$ s time component is not from the triplet state and has been



**Figure 3.10:** Normalized best fitted lines of fluorescence autocorrelation curves of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) for (a) different GnHCl concentration and (b) different temperatures.

assigned to the conformational relaxation time of domain-III of HSA.<sup>40</sup> Some of the normalized fitted lines for different GnHCl concentrations are shown in figure 3.10a. From the measured diffusion time of NPCE tagged HSA and R6G, the hydrodynamic radius of HSA in its native state is calculated to be 38.7 Å using equations 2.21 and 2.37. This value is in accordance with the already reported value of HSA.<sup>28,40</sup> With increasing GnHCl concentration, the diffusion time found to increase monotonically. There are two contributing factors to this increase of  $\tau_D$ . The first one is the increase in viscosity due to increasing GnHCl concentration and the second one is the increase of size of HSA due to denaturation. The effect of viscosity has been taken care of by performing control experiments with R6G at the corresponding GnHCl concentrations. R6G is a rigid molecule and does not undergo any structural change even at very high

GnHCl concentration. Thus any increase in  $\tau_D$  of R6G will be solely due to the viscosity effect. Thus, the diffusion time of R6G in the experimental range of GnHCl concentration was calculated as a measure of change in viscosity and was used to calculate the correct viscosity of the GnHCl solution having different concentration. The diffusion coefficient of NPCE tagged HSA was calculated for each GnHCl concentration using the observed diffusion time (equation 2.21). Using the diffusion

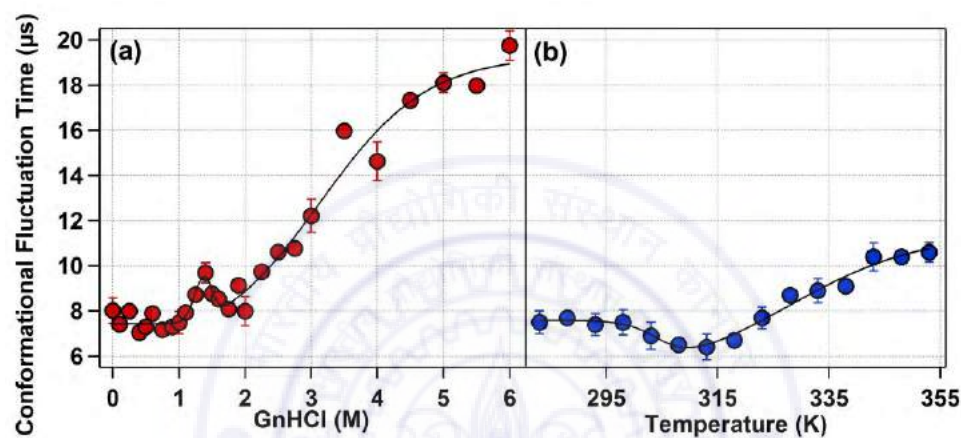


**Figure 3.11:** Variation of hydrodynamic radius of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) with (a) GnHCl and (b) temperature.

coefficient, and the corrected viscosity the hydrodynamic radius was calculated using equation 2.37. It is found that with increasing GnHCl concentration the hydrodynamic radius remains almost same till 1.75 M GnHCl and then starts increasing monotonically till 61.1 Å at 6 M GnHCl. The change in hydrodynamic radius with GnHCl has been shown in figure 3.11a.

The conformational fluctuation time component is found to be 8  $\mu$ s in the native state of HSA, which is in good agreement with previous report.<sup>40</sup> With increasing GnHCl concentration till 1 M, the  $\tau_R$  remains unchanged reflecting that the side chain dynamics within domain-III of HSA remain same till this point. At  $\sim$ 1.4 M GnHCl,  $\tau_R$  increases to  $\sim$ 10  $\mu$ s. Further, at 2 M GnHCl it again drops to  $\sim$ 8  $\mu$ s. After this point it increases monotonically to 19.7  $\mu$ s at 6 M GnHCl (figure 3.12a). Thus the variation of conformational relaxation time shows a rise and dip character at 1.4 M and 2 M GnHCl.

Temperature dependent FCS experiments have been performed for the 283-353 K range. The fluorescence autocorrelation curves at some selected temperatures along with their best fits are shown in figure 3.8b and all the normalized fitted lines are shown in figure 3.10b. As evident from the normalized fits shown in figure 3.10b,  $\tau_D$  decreases at first till 318 K and then remains unchanged upon further increase of temperature till 353 K. Here the two contributing factors, responsible for the change of  $\tau_D$ , are the thermal unfolding of HSA, leading to the increase in the  $\tau_D$  value and the



**Figure 3.12:** Variation of conformational fluctuation time of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) with (a) GnHCl and (b) temperature.

decrease in the viscosity of the medium by increasing the temperature. The temperature dependent viscosity factor has been taken care of by performing control experiments with R6G as discussed before. The calculated value of hydrodynamic radius is found to remain unchanged till 313 K and then increases sharply till 353 K (figure 3.11b). At 353 K, the hydrodynamic radius is calculated to be 59.0 Å, which indicates the almost unfolded structure of HSA.

The conformational fluctuation time at 283 K is found to be 7.5 μs, which remains unchanged till 298 K. Then a small decrease is observed in the range 303-318 K, where the value of  $\tau_R$  reaches a minimum of 6.4 μs at 313 K. Upon further increase of temperature beyond 318 K, an increase of  $\tau_R$  value is observed. At 353 K the value of  $\tau_R$  reaches a value of 10.6 μs. The overall variation of  $\tau_R$  with temperature is shown in figure 3.12b.

### 3.3. Discussion

The calculated  $\alpha$ -helicity of the native and the NPCE tagged HSA has a difference of 2%. This is to mention here that electrostatic interactions are responsible for the secondary structure of the protein. The protein molecule orients itself in such a way that the amino acids containing polar side chains are on the surface exposed to water and the non-polar residues remain buried inside the hydrophobic protein core. This way the energy is minimized. If any external agent induces significant electrostatic forces in the protein matrix, the protein structure gets perturbed, and often, this perturbation causes a change in the helical content of the protein. The similar CD spectra of the NPCE tagged HSA with that of the pure HSA signify minimum structural perturbation by incorporation of NPCE. Thus it is evident that NPCE does not disturb the electrostatic forces acting within the HSA, and retaining the native structure of the protein. NPCE tagged in HSA shows absorption and emission maxima at 440 nm and 478.5 nm respectively. It is widely accepted that the environment inside the protein core is more hydrophobic than outside environment.<sup>45,46</sup> The 15 nm blue shift in emission maxima compared to the free NPCE in buffer is attributed to the covalent labeling of NPCE to the hydrophobic non-polar environment inside the HSA as compared to the bulk.<sup>40</sup> Being present in the core of the protein cavity, the fluorescence of NPCE may be quenched by the nearby electron rich amino acid residues, resulting in the lowering of the fluorescence quantum yield. Similarly, the decrease in lifetime of NPCE upon being tagged has been attributed to the change in environment, especially the different nature of water inside protein cavity around the dye when tagged to HSA.

As the chromophoric part of NPCE is a coumarin moiety with positive solvatochromism, one expects a red shift of the emission maxima of NPCE tagged to HSA, as HSA gets denatured leaving the NPCE in a relatively more polar environment. Moreover, as NPCE is tagged to the domain-III, it reflects the local structural change of domain-III of HSA only. Thus any red shift happening with application of any external agent can be attributed to the unfolding of domain-III of HSA. From the steady-state spectroscopic data it can be inferred that domain-III of HSA is more



sensitive to GnHCl concentration than heat. The emission maxima changes regularly even at small concentration of GnHCl. All along a 15 nm red shift in the emission maxima has been observed from the 0 M to 6 M GnHCl. The minimal change of emission maxima in 4.5-6.0 M GnHCl range confirms that the protein is almost denatured at GnHCl concentration  $> 4.5$  M (see figure 3.4a). A small dip is observed around 1-2 M GnHCl, which can be attributed to any local structural change. In case of thermal denaturation, the unfolding starts only after 328 K. A total 4 nm red shift of the NPCE emission maxima has been observed at 353 K (see figure 3.4b).

The change in fluorescence lifetime in the given GnHCl concentration and temperature range is not very drastic, which may be pertaining to the photophysics of NPCE. However, there is a prominent dip around 1.5 M GnHCl as can be seen in figure 3.7a. This is to be noted that I observed similar dip in the variation of emission maxima. Both these results hint towards the presence of an intermediate state at this concentration range. Fluorescence lifetime change with the change in the temperature is also non-monotonous and shows a faint dip at around 310 K (figure 3.7b). Intermediate state formation seems possible in this case too.

Fluorescence correlation spectroscopic measurement has been undertaken to reveal important information about the overall as well as domain specific denaturation of domain-III of HSA. The change in diffusion time of HSA is associated with the overall unfolding of the protein, whereas the conformational relaxation time is associated with the local side chain dynamics around the attached probe. It has been noticed from the calculated hydrodynamic radius (depicted in figure 3.11a) that till 1 M GnHCl the overall dimension of the protein does not change appreciably. The sharp increase in the hydrodynamic radius in 1-5 M GnHCl indicate a considerable unfolding of HSA at this range. Like steady-state experiments, here also the overall denaturation is found to be complete around 5 M GnHCl. The overall HSA structure is found to be resistant to heat till 313 K, then it shows sharp increase in hydrodynamic radius (Figure 3.11b). Clearly, the overall unfolding of HSA by both GnHCl and temperature is a two state process, where the two states are the native state and the unfolded state. The



current results match quite well with the pattern of overall unfolding of HSA reported by Yadav *et al.*<sup>28</sup>

On the other hand, considering the change in conformational fluctuation time, I have predicted the presence of two intermediate states for GnHCl induced unfolding and one intermediate state for thermal unfolding of HSA. As already mentioned,  $\tau_R$  reflects the local structural change around the fluorescent probe (NPCE) located at the domain-III of HSA. I assign this fluctuation to the movement of electron rich amino acid residues close and away from NPCE resulting in a fluctuation of fluorescence intensity. There are seven amino acid residues that are in close proximity to the monitoring site of domain-III (*i.e.* around the Tyr-411 residue) which are Leu-407, Val-409, Arg-410, Thr-412, Lys-414, Val-415, Ser-427, Phe-488 and Ser-489. Among them lysine, arginine and serine are electron rich and may take part in the quenching of NPCE fluorescence. I propose that, in the first intermediate state at 1.4 M GnHCl the quenching residues move far from the probe resulting in an increase in the conformational relaxation time, whereas the low fluctuation time in the second intermediate state at 2 M GnHCl depicts that the distance between the quencher and the probe is comparatively small (see figure 3.12a). At similar concentration of GnHCl, domain-I also forms intermediate states in the course of GnHCl induced denaturation of HSA.<sup>28</sup> Thus it can be concluded that the domain specific GnHCl induced unfolding pathways of HSA are similar for both domain-I and domain-III. However, the free energy for the corresponding transitions are quite different for the two domains as described in the next section and the values are tabulated in table 3.2. The variation of  $\tau_R$  with temperature established that there is a distinct intermediate state around 310 K, in the course of thermal denaturation of HSA (see figure 3.12b). The lower value of  $\tau_R$  at 353 K as compared to that of the chemically denatured state of HSA (in presence of 6 M GnHCl) confirms that the effect of temperature in domain-III is not same as of GnHCl, although the global unfolding is quite similar. The intermediate state shows a small conformational fluctuation time, probably due to the close proximity of NPCE with the quenching amino acid residues. The local dynamics of this intermediate state

is clearly different from the intermediate state formed during the GnHCl induced denaturation as evident from the difference in conformational fluctuation time.

### 3.4. Calculation of Thermodynamic Parameters

FCS data have been used to estimate the thermodynamic parameters of thermal and chemical unfolding of HSA using the method described by Naidu *et al.* (section 2.2 of chapter 2).<sup>48</sup> Hydrodynamic radius has been taken as the observable for overall denaturation whereas conformational fluctuation time has been considered for unveiling information about structural changes of domain III of HSA.

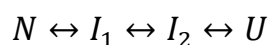
For the GnHCl induced overall unfolding, a two state model was used as  $N \leftrightarrow U$  where  $N$  denotes the native state and  $U$  denotes the unfolded state.<sup>48</sup> The denaturation midpoint ( $[D]_{1/2}$ ) of the overall unfolding is calculated as 3.08 M which is in accordance with the previous reports.<sup>17, 28</sup> The value of the thermodynamic parameters calculated using equation 2.38 also matches well with the already reported data for the overall unfolding of HSA.<sup>28</sup> The fitting line is shown in figure 3.11a.

**Table 3.2.** Thermodynamic parameters for the unfolding of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) against GnHCl concentration.

Conformational fluctuation time	$N \leftrightarrow I_1$			$I_1 \leftrightarrow I_2$			$I_2 \leftrightarrow D$		
	$\Delta G_1^0$	$m_1$	$[D]_{1/2}$	$\Delta G_2^0$	$m_2$	$[D]_{1/2}$	$\Delta G_3^0$	$m_3$	$[D]_{1/2}$
	9.85	7.92	1.24	14.69	9.73	1.51	2.27	0.73	3.13
Hydrodynamic radius	$N \leftrightarrow D$								
	$\Delta G^0$			$m$			$[D]_{1/2}$		
	2.81			0.913			3.08		

( $\Delta G$  in kcal mol<sup>-1</sup>,  $m$  is in kcal mol<sup>-1</sup> M<sup>-1</sup>,  $[D]_{1/2}$  is in M).

The variation of conformational relaxation time with GnHCl was best fitted with a four state unfolding model with two intermediate states (see the fitted line in figure 3.12a). The model can be represented as follows<sup>28, 48</sup>



where  $N$ ,  $I$  and  $U$  denote native, intermediate and unfolded states. The midpoints for the corresponding transitions were calculated using equation 2.40 and they are 1.24 M,

1.51 M and 3.13 M respectively. The values of  $G$ ,  $m$  and the transition midpoints for all the transitions (overall and domainwise) are summarized in table 3.2.

The thermodynamic analysis of the variation of  $r_H$  and  $\tau_R$  with temperature has been done following similar procedure using equation 2.41. A two state and a three state model has been chosen to fit the variation of  $r_H$  and  $\tau_R$  respectively. The midpoint of overall transition was found to be 332 K, whereas the midpoint for the thermal intermediate formation and formation of denatured state from the intermediate state were calculated as 307 K and 328 K respectively. The corresponding fitted lines for hydrodynamic radius and conformational fluctuation time have been shown in figures 3.11b and 3.12b respectively and all the thermodynamic parameters for the thermal transitions are listed in table 3.3.

**Table 3.3.** Thermodynamic parameters for the unfolding of NPCE tagged HSA in phosphate buffer (pH 7.4, 50 mM) against temperature.

	$N \leftrightarrow I_1$			$I_1 \leftrightarrow D$		
	$\Delta G_1^0$	$m_1$	$[D]_{1/2}$	$\Delta G_3^0$	$m_3$	$[D]_{1/2}$
Concerted chain motion time	61.94	0.2	309.7	16.26	0.05	327.8
Hydrodynamic radius	$N \leftrightarrow D$					
	$\Delta G^0$	$m$	$[D]_{1/2}$			
	38.84	0.12	331.9			

( $\Delta G$  in kcal mol<sup>-1</sup>,  $m$  is in kcal mol<sup>-1</sup> M<sup>-1</sup>,  $[D]_{1/2}$  is in M).

### 3.5. Conclusion

Domain-III of HSA has been covalently labelled with NPCE and the domain specific chemical and thermal unfolding has been studied. Steady-state and time-resolved fluorescence experiments hint towards the presence of intermediate states, although the states could not be resolved completely. However, fluorescence correlation spectroscopic experiment clearly enunciate the structural changes in the course of denaturation. It is found that for GnHCl induced denaturation, the overall unfolding is a monotonous two-state process, which is in accordance with previous studies. The free energy of unfolding and the transition midpoint are found to be 2.81 kcal mol<sup>-1</sup> and 3.08 M GnHCl, respectively. On the other hand, the unfolding pertaining

to the domain-III of HSA, as elucidated from the conformational fluctuation dynamics, involves two intermediate states, which has been identified for the first time. At similar concentration, domain-I also forms intermediate states in the course of GnHCl induced denaturation.<sup>28</sup> The overall thermal unfolding of HSA is a two state process, where the free energy of transition is calculated as 38.84 kcal mol<sup>-1</sup> and the transition midpoint is at 332.0 K. However, one intermediate state has been detected in the domain-III of HSA during the course of thermal denaturation. The presence of such intermediate state was not known earlier. The value of  $\tau_R$  at this intermediate state has been found to be 6.5  $\mu$ s. This denotes that the quenching amino acid residues move closer to the attached probe at this intermediate state. On the other hand, domain-I of HSA, has been reported to unfold monotonously with temperature without the involvement of any intermediate state. Thus I conclude that the GnHCl induced denaturation of domain-I and domain-III of HSA follows similar pathways but the thermal unfolding involves different structural changes within the two domains.

## References

1. Dockal, M.; Carter, D. C.; Rüker, F. *J. Biol. Chem.* **1999**, *274*, 29303–29310.
2. Matsushita, S.; Isima, Y.; Chuang, V. T. G.; Watanabe, H.; Tanase, S.; Maruyama, T.; Otagiri, M. *Pharm. Res.* **2004**, *21*, 1924–1932.
3. Dugiaczyk, A.; Law, S.; Dennison, O. E. *Proc. Natl. Acad. Sci. USA.* **1982**, *79*, 71–75.
4. Peters, T. All about albumin: Biochemistry, genetics, and medical applications. Academic Press: San Diego, **1996**.
5. Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1976**, *12*, 1052–1061.
6. Salvi, A.; Carrupt, P. A.; Mayer, J. M.; Testa, B. *Drug Metab. Dispos.* **1997**, *25*, 395–398.
7. Fasano, M.; Curry, S.; Terreno, E.; Galliano, M.; Fanali, G.; Narciso, Notari, S.; Ascenzi, P. *IUBMB Life* **2005**, *57*, 787–796.
8. He, X.M.; Carter, D. C. *Nature* **1992**, *358*, 209–215.
9. Curry, S.; Mandelkow, H.; Brick, P.; Franks, N. *Nat. Struct. Biol.* **1998**, *5*, 827–835.
10. Chamani, J.; Vahedian-Movahed, H.; Saberi, M. R. *J. Pharm. Biomed. Anal.* **2011**, *55*, 114–124.
11. Azimi, O.; Emami, Z.; Salari, H.; Chamani, J. *Molecules* **2011**, *16*, 9792–9818.
12. Movahed, H. V.; Saberi, M. R.; Chamani, J. *J. Biomol. Struct. Dyn.* **2011**, *28*, 483–502.
13. Hosainzadeh, A.; Gharanfoli, M.; Saberi, M. R.; Chamani, J. *J. Biomol. Struct. Dyn.* **2012**, *29*, 1013–1050.
14. Farruggia, B.; Rodriguez, F.; Rigatuso, R.; Fidelio, G.; Picó, G. *J. Protein Chem.* **2001**, 81–89.
15. Ahmad, B.; Ahmed, Md. Z.; Haq, S. K.; Khan, R. H. *Biochim. Biophys. Acta* **2005**, *1750*, 93–102.
16. Nerli, B.; Pico, G. *Pharmacol. Toxicol.* **1993**, *73*, 297–300.
17. Yadav, R.; Sen, P. *Protein Sci.* **2013**, *22*, 1571–1581.
18. Vlasova, I. M.; Saletsky, A.M. *J. Appl. Spectrosc.* **2009**, *76*, 536–541.



19. Anand, U.; Mukherjee, S. *Biochim. Biophys. Acta* **2013**, *1830*, 5394–5404.
20. Shaw, A. K.; Pal, S. K. *J. Photochem. Photobiol., B* **2008**, *90*, 69–77.
21. Vlasova, I. M.; Saletsky, A. M. *Moscow University Physics Bulletin* **2011**, *66*, 59–63.
22. Santra, M. K.; Banerjee, A.; Rahaman, O.; Panda, D. *Int. J. Biol. Macromolec.* **2005**, *37*, 200–204.
23. Nanda, R.K.; Sarkar, N.; Banerjee, R. *J. Photochem. Photobiol. A* **2007**, *192*, 152–158.
24. Kuchlyan, J.; Kundu, N.; Banik, D.; Roy, A.; Sarkar, N. *Langmuir* **2015**, *31*, 13793–13801.
25. Chakraborty, A.; Seth, D.; Setua, P.; Sarkar, N. *J Phys Chem B.* **2006**, *110*, 16607–17.
26. Moriyama, Y.; Ohta, D.; Hachiya, K.; Mitsui, Y.; Takeda, K. *J. Protein Chem.* **1996**, *15*, 265–272.
27. Otsu, T.; Nishimoto, E.; Yamashita, S. *J Biochem.* **2010**, *147*, 191–200.
28. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
29. Sengupta, B.; Yadav, R.; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, *18*, 14350–14358.
30. Yadav, R.; Sengupta, B.; Sen, P. *Biophys. Chem.* **2016**, *211*, 59–69.
31. Flora, K.; Brennan, J. D.; Baker, G. A.; Doody, M. A.; Bright, F. V. *Biophys. J.* **1998**, *75*, 1084–1096.
32. Sasmal, D. K.; Mondal, T.; Mojumdar, S. S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
33. Chowdhury, R.; Chatteraj, S.; Mojumdar, S. S.; Bhattacharyya, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16286–16293.
34. Ahmad, B.; Ankita; Khan, R. H. *Arch. Biochem. Biophys.* **2005**, *437*, 159–167.
35. Ahmad, B.; Muteeb, G.; Alam, P.; Varshney, A.; Zaidi, N.; Ishtikhar, M.; Badr, G.; Mahmoud, M. H.; Khan, R. H. *Int. J. Biol. Macromolec.* **2015**, *75*, 447–452.
36. Heller, W. T. *J. Phys. Chem. B* **2013**, *117*, 2378–2383.

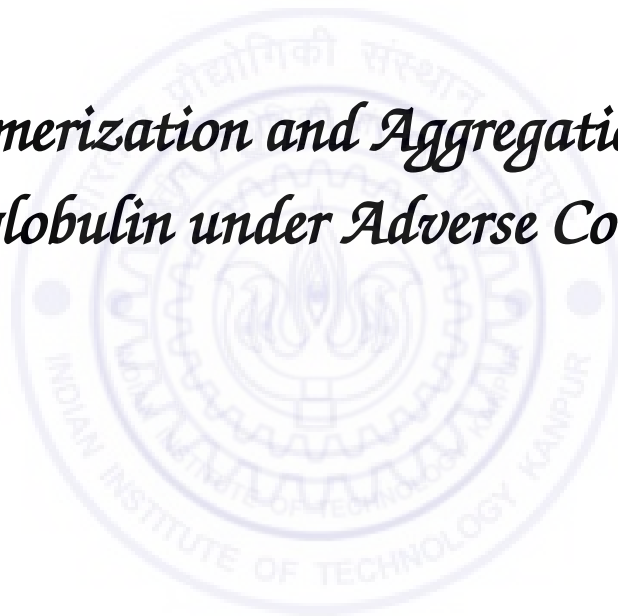
37. Das, D. K.; Das, A. K.; Mandal, A. K.; Mondal, T.; Bhattacharyya, K. *ChemPhysChem* **2012**, *13*, 1949–1955.
38. Wallevik, K. *J. Biol. Chem.* **1973**, *248*, 2650–2655.
39. Picó, G.A. *Int. J. Biol. Macromolec.* **1997**, *20*, 63–73.
40. Sengupta, B.; Acharya, A.; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, *18*, 14350–14358.
41. Ishii, K.; Tahara, T. *Chem. Phys. Lett.* **2012**, *519-520*, 130–133.
42. Ishii, K.; Tahara, T. *J. Phys. Chem. B* **2010**, *114*, 12383–12391.
43. Ishii, K.; Tahara, T. *J. Phys. Chem. B* **2013**, *117*, 11414–11422.
44. Elson, E. L. *Biophys. J.* **2011**, *101*, 2855–2870.
45. Bhattacharyya, K. *Acc. Chem. Res.* **2003**, *36*, 95–101.
46. Lakowicz, J. R. Principles of fluorescence spectroscopy. Springer, third ed., **2006**.
47. Ries, J.; Schwille, P. *Bioassays* **2012**, *34*, 361–368.
48. Naidu, K. T.; Prabhu, N. P. *J. Phys. Chem. B* **2011**, *115*, 14760–14767.



## *Chapter 4*

---

### *Monomerization and Aggregation of $\beta$ -lactoglobulin under Adverse Condition*



**Sengupta, B.;** Das, N.; Sen, P. *BBA- Proteins and Proteomics* **2017**, 1866, 316–326.

*$\beta$ -lactoglobulin is one of the major components of bovine milk and it remains in a dimeric form under physiological conditions. The work described in this chapter elucidates the structural change of  $\beta$ -lactoglobulin at pH 7.4 under the action of guanidine hydrochloride (GnHCl) and heat at the single molecule level. The only free cysteine (Cys-121) of  $\beta$ -lactoglobulin was tagged with 7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM) for this purpose. The dimeric structure of  $\beta$ -lactoglobulin was found to undergo a monomerization prior to the unfolding process upon being subjected to GnHCl. The hydrodynamic diameter of the native dimer, native monomer and the unfolded monomer has been estimated as  $\sim 55$  Å,  $\sim 29$  Å and  $\sim 37$  Å, respectively. The free energy change for the monomerization and denaturation are respectively  $1.57$  kcal mol<sup>-1</sup> and  $8.93$  kcal mol<sup>-1</sup>, respectively<sup>1</sup>. With an increase in temperature, the development of two types of aggregates (small aggregates and large aggregates) was observed, which is triggered by the formation of the monomeric structure of  $\beta$ -lactoglobulin. The hydrodynamic diameters of the smaller and larger aggregates was estimated to be  $\sim 77$  Å and  $\sim 117$  Å, respectively. The formation of small aggregates turns out to be reversible whereas that of larger aggregates is irreversible. The free energy associated with these two steps is  $0.69$  kcal mol<sup>-1</sup> and  $9.09$  kcal mol<sup>-1</sup>, respectively<sup>1</sup>. Based on the size parameters, the smaller and larger aggregates have been proposed to contain twenty and sixty monomeric units. It has also been concluded that the monomeric subunits retain their native state, in particular their secondary structure in these aggregates.*



#### 4.1. Introduction

$\beta$ -lactoglobulin is a whey protein containing mainly  $\beta$ -sheets.<sup>1-7</sup> It is reported to stay in monomer form below pH 3 and above pH 8.<sup>8</sup> In the intermediate pH range, it stays as a dimer with a characteristic hydrophobic cage, which is anticipated to be responsible for ligand binding and transport.<sup>4</sup> The main force responsible for dimerization is the hydrophobic force with additional contributions from peptide-peptide interaction.<sup>9-11</sup> It is also proposed that hydrogen bonding and tight packing of the residues at the interface are also responsible to stabilize the dimeric structure.<sup>11</sup>

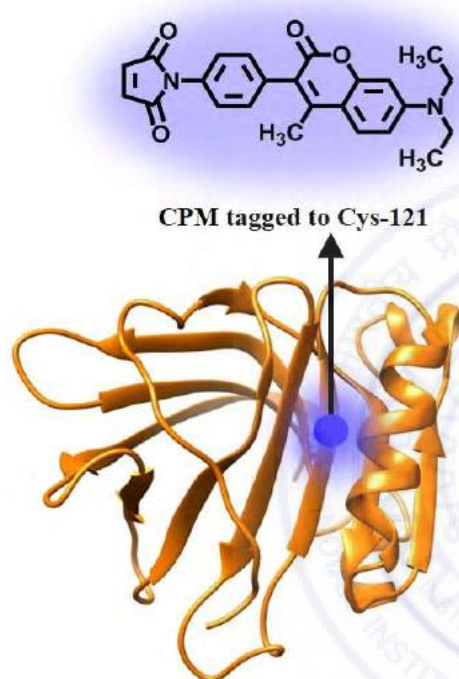
The monomer-dimer equilibrium of  $\beta$ -lactoglobulin has been widely studied under different environmental conditions.<sup>7,10,11,12-18</sup> Apart from pH, the secondary, tertiary and quaternary structure of  $\beta$ -lactoglobulin is known to depend strongly on temperature, ionic strength, chemical denaturants, *etc.*<sup>7,19-22</sup> The effect of GnHCl on the structure of  $\beta$ -lactoglobulin has been studied at a wide pH range and the possibilities of monomerization, formation of molten globule state, *etc.* has been proposed.<sup>20,21,23</sup> In an early report by Greene *et al.*, they characterized the GnHCl ( $\geq 3$  M GnHCl) induced denaturated state at pH 3.2 of  $\beta$ -lactoglobulin as mostly random coils.<sup>20</sup> On the contrary, Sakurai *et al.* studied the effect of GnHCl on the structure of  $\beta$ -lactoglobulin and reported that low concentration of GnHCl stabilizes the dimeric species at pH 3.0.<sup>10</sup> Dar *et al.* proposed that at pH 2.0, where the native state is the monomeric one, the GnHCl induced structural transition involves two distinct steps being the formation of additional secondary structures at 2.0 M GnHCl and melting of all secondary structures at a higher concentration of GnHCl.<sup>23</sup> In these low pH region the native state is monomeric and hence the transition pathway must be straightforward than at the neutral pH, where the native state is dimeric. In this case the possibility of monomerization is coupled with other structural alterations, making the overall pathway complex. Using size exclusion fast protein liquid chromatographic techniques Valente-Mesquita *et al.* concluded that at pH 7, high GnHCl concentration ( $>3$  M) dissociates two subunits of  $\beta$ -lactoglobulin dimer giving rise to the monomers.<sup>21</sup> D'Alfonso *et al.* observed an increment in the size of the protein as compared to the native dimeric state at pH 6.2 in the presence of GnHCl at low concentration, which they attributed to the

higher degree of aggregation under this condition. Additionally, they stated that at high GnHCl concentration  $\beta$ -lactoglobulin retains a compact structure although the structure expands a little bit.<sup>24</sup> Thus, several statements have been proposed regarding the possible action of GnHCl on the native state of  $\beta$ -lactoglobulin.

The thermal unfolding of  $\beta$ -lactoglobulin is also widely explored and the most prominent effect of heating was concluded to be the formation of aggregates at high

temperature.<sup>16,25-31</sup> Iametti *et al.*

observed that the proportion of covalently bound and non-covalently bound dimers changes with increasing temperatures and at temperatures around 358 K, higher molecular weight species were formed. These high molecular weight species were recognized as aggregates and their relative abundance was found to increase with increasing protein concentration.<sup>26</sup> Immediately after that, Elofsson *et al.* proposed that the monomer content increases significantly while the temperature is raised to 336 K and the aggregation rate at higher temperatures depends on the



**Scheme 4.1:** Tagging of  $\beta$ -lactoglobulin with CPM. The structure file of  $\beta$ -lactoglobulin is downloaded from Protein Data Bank. (PDB ID: 1BEB)

concentration of monomers.<sup>27</sup> Two facts related to the thermal transition of  $\beta$ -lactoglobulin are widely accepted now. First, with increasing temperature the equilibrium gets shifted towards the monomeric side and the structural changes induced by heating upto 333 K are reversible in nature. Secondly, an exposure to higher temperatures lead to aggregation of  $\beta$ -lactoglobulin, which is irreversible in nature.<sup>16,25</sup> On the contrary, Tolkach *et al.* described the overall transition with four distinct transitions as i) monomerization, ii) formation of another state, which differs with the

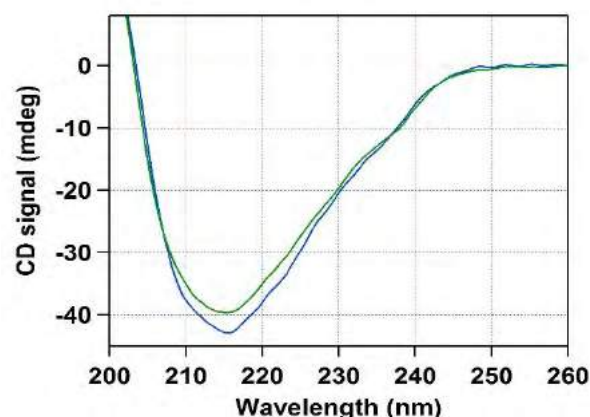
monomer only by minor conformational changes, iii) formation of thermally induced molten globule state and iv) formation of unfolded monomeric state.<sup>28</sup>

In spite of all these reports, the exact structural transition pathway and the size factors therein with increasing GnHCl concentration and temperature are still undisclosed. Beside this, all the structural transitions of  $\beta$ -lactoglobulin are reported to be highly concentration dependent and most of the studies have been carried out in the concentration range of 0.1-10 g l<sup>-1</sup> (~5-500  $\mu$ M) depending upon the technique that is being used, since this is the concentration of  $\beta$ -lactoglobulin in cow milk.<sup>28</sup> However, the regular consumption of cow milk by lactating mothers induce the presence of  $\beta$ -lactoglobulin in the low  $\mu$ g l<sup>-1</sup> level.<sup>32,33</sup> Moreover, the concentration of  $\beta$ -lactoglobulin in human serum is found to be in the nanomolar concentration regime after consumption of cow milk.<sup>34</sup> Thus it is important to examine the properties and transitions of  $\beta$ -lactoglobulin in the lower concentration range.

In this work, I have performed ensemble and single molecule spectroscopic techniques to unveil the behaviour of  $\beta$ -lactoglobulin under the action of GnHCl and heat. For this purpose,  $\beta$ -lactoglobulin has been tagged with 7-diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin (CPM) at Cys-121 (section 2.4.2 of chapter 2).

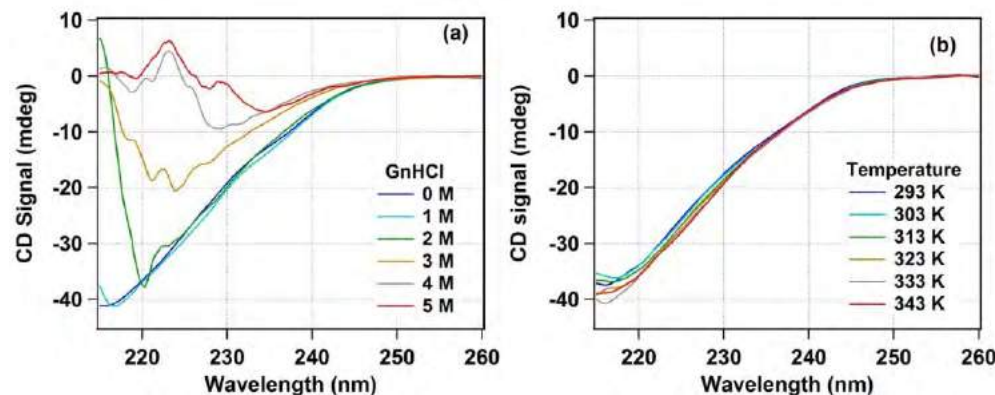
## 4.2. Results

**4.2.1. Circular Dichroism Spectroscopy:** The circular dichroism spectra for the



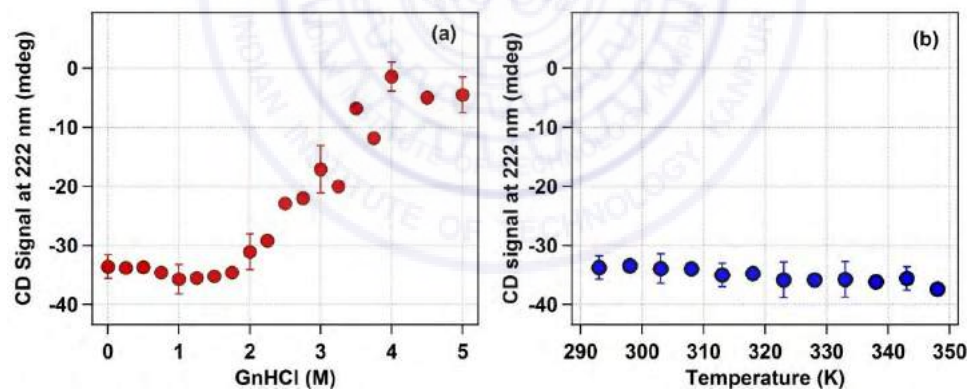
**Figure 4.1:** CD spectra of untagged (green, 10.7  $\mu$ M) and CPM tagged (blue, 10.9  $\mu$ M)  $\beta$ -lactoglobulin in phosphate buffer (pH 7.4, 50 mM).

tagged and untagged  $\beta$ -lactoglobulin were recorded under the same experimental condition (see figure 4.1) and the  $\alpha$ -helicity was found to be 16.3% and 17.0% respectively, which matches quite well with the reported value of %-helicity of  $\beta$ -lactoglobulin in its native state.<sup>13</sup> Next we have studied the effect of GnHCl and temperature on the  $\alpha$ -helicity of  $\beta$ -



**Figure 4.2:** CD spectra of untagged  $\beta$ -lactoglobulin ( $\sim 10 \mu\text{M}$ ) at varying (a) GnHCl concentration (0 M- 5 M) and (b) temperatures (293 K- 348 K) in phosphate buffer (pH 7.4, 50 mM).

lactoglobulin. Figure 4.2a and 4.2b depicts some of the measured CD spectrum. Unfortunately, at high GnHCl concentration the signal below 215 nm contains a lot of interference from GnHCl and in these cases the magnitude of the negative peak at 222 nm is taken as a measure of  $\alpha$ -helicity<sup>35</sup> followed by the measure of the denaturation. Figure 4.3a and 4.3b show the variation of the magnitude of CD signal at 222 nm as a function of GnHCl concentration and temperature, respectively. Clearly, with increase

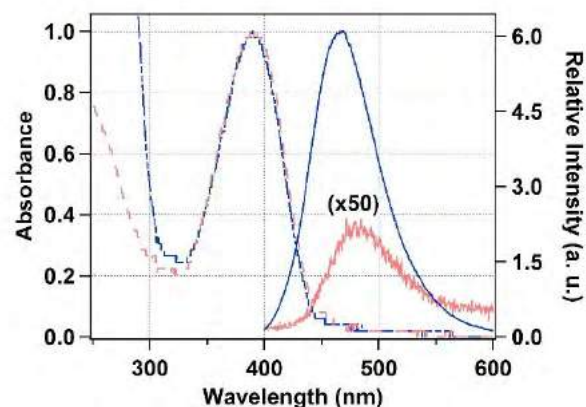


**Figure 4.3:** Change of circular dichroism (CD) signal of  $\beta$ -lactoglobulin ( $10 \mu\text{M}$ ) at 222 nm as function of (a) GnHCl concentration and (b) temperatures in phosphate buffer (pH 7.4, 50 mM).

in GnHCl concentration a significant decrease of negative signal at 222 nm has been observed between 2.0 to 4.0 M GnHCl, whereas with increase in temperature no significant change has been observed.

**4.2.2. Steady-State Absorption and Fluorescence Measurements:** Absorption spectra of CPM shows one major peak at 390 nm in buffer. Upon being excited at 390 nm, CPM shows a very weak emission ( $\phi_f = 2.8 \times 10^{-3}$ ) with an emission maxima at



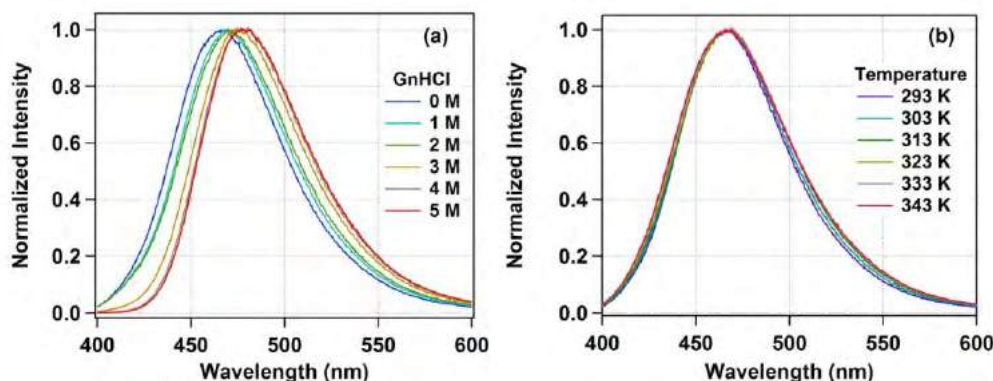


**Figure 4.4:** Normalized absorption spectra (dashed lines) and emission spectra (solid lines) of untagged (red) and tagged (blue) CPM in phosphate buffer (pH 7.4, 50 mM). The emission spectra of untagged CPM has been multiplied by a factor of 50 to remove its flatness and make it visible.

481 nm. The negligible emission intensity of CPM in buffer describes it as a non-fluorescent compound. Upon being tagged in  $\beta$ -lactoglobulin the absorption peak remains unaltered whereas the emission peak shifts to 467 nm with a 160 fold increase in the fluorescence quantum yield ( $\phi_f = 0.45$ ).

The absorption spectra (normalized to unity) and emission spectra of untagged and tagged CPM are shown in figure 4.4.

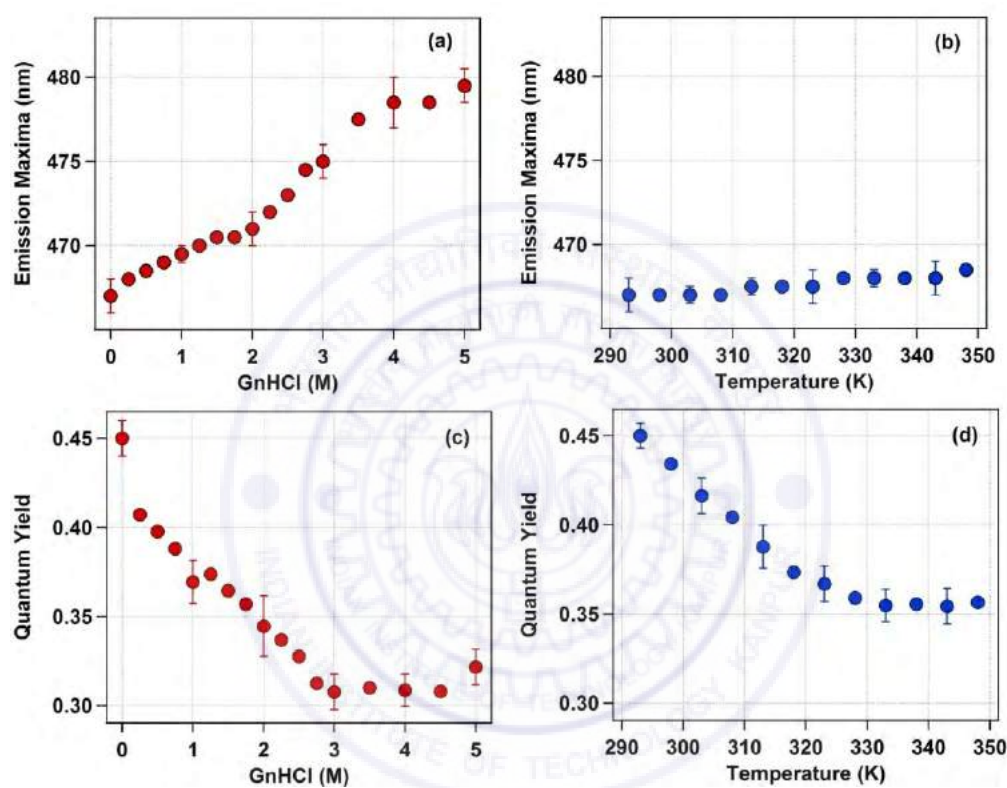
To study the structural changes in  $\beta$ -lactoglobulin with GnHCl and heat, emission spectra of both the tryptophan and the tagged-CPM was recorded with increasing concentration of GnHCl and with increasing temperature. We have noticed a gradual red shift in the emission maxima of CPM with increasing GnHCl concentration and at 5 M GnHCl, the emission maxima is found to be 12.5 nm red shifted from the native state of  $\beta$ -lactoglobulin (figure 4.5a). The quantum yield decreases regularly till 3.0 M GnHCl and then becomes constant. However, the effect of temperature on emission maxima is found to be small. At 353 K the maxima is found to be at 468.5 nm (see figure 4.5b). However, the fluorescence quantum yield of CPM



**Figure 4.5:** Normalized emission spectra of CPM tagged  $\beta$ -lactoglobulin at (a) different GnHCl concentration (0 M-5 M) and (b) different temperatures (293 K-343 K) in phosphate buffer (pH 7.4, 50 mM); ( $\lambda_{ex} = 390$  nm).

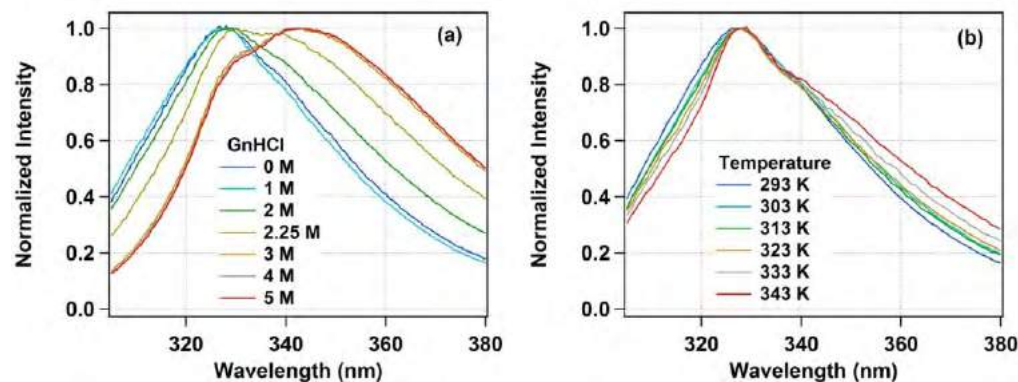


decreases gradually till 333 K ( $\Phi_f = 0.35$ ) and then remains unchanged. The variation of emission maxima of CPM with GnHCl and temperature are depicted in figure 4.6a and 4.6b respectively and the variation of quantum yield of CPM with GnHCl and temperature are shown in figure 4.6c and 4.6d respectively. The intrinsic tryptophan emission maxima of native  $\beta$ -lactoglobulin is found to be 327.5 nm, which starts shifting towards higher wavelength with increase in the GnHCl concentration and



**Figure 4.6:** The variation of emission maxima of CPM tagged to  $\beta$ -lactoglobulin with (a) GnHCl and (b) temperature. The variation of emission quantum yield of tryptophan in  $\beta$ -lactoglobulin with (c) GnHCl and (d) temperature.

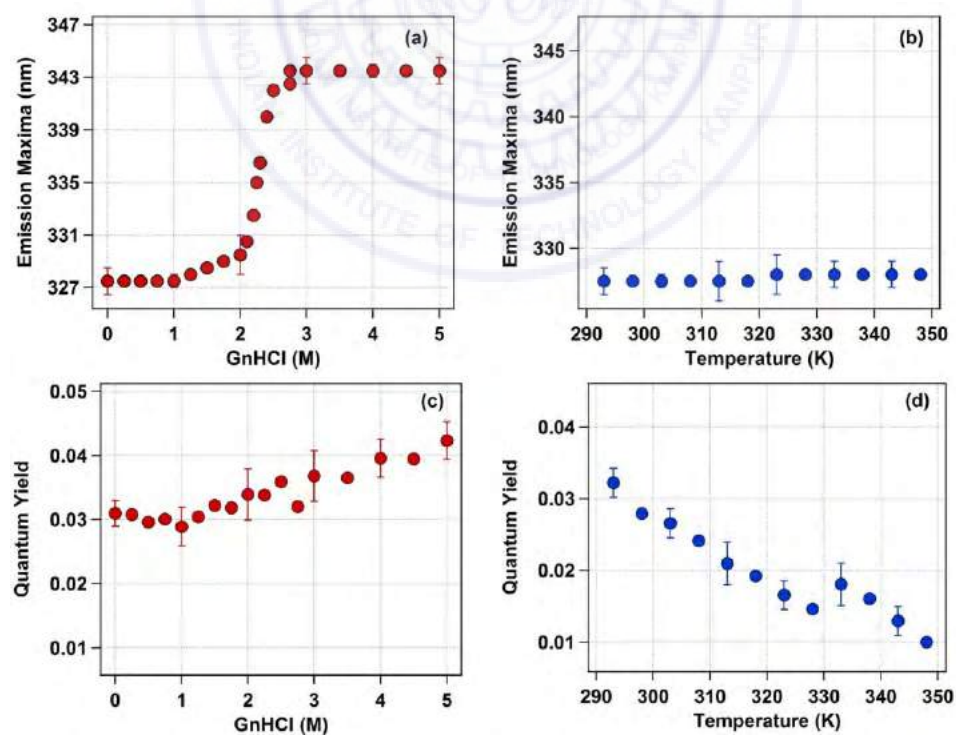
reaches to 343.5 nm at 4 M GnHCl. Some of the normalized emission spectra are shown in figure 4.7a and a plot of the emission maxima as a function of GnHCl is depicted in figure 4.8a. The full width at half maxima (FWHM) is found to be  $\sim 46.0$  nm in absence of GnHCl and  $\sim 60.0$  nm in presence of 4 M GnHCl. On the other hand, at higher temperatures one may notice the formation of shoulder at the higher wavelength side although with increase in temperature the position of the tryptophan emission maxima remains unchanged. Figure 4.7b shows emission spectra and figure 4.8b shows change in emission maxima with temperature. In the native state of  $\beta$ -lactoglobulin, the



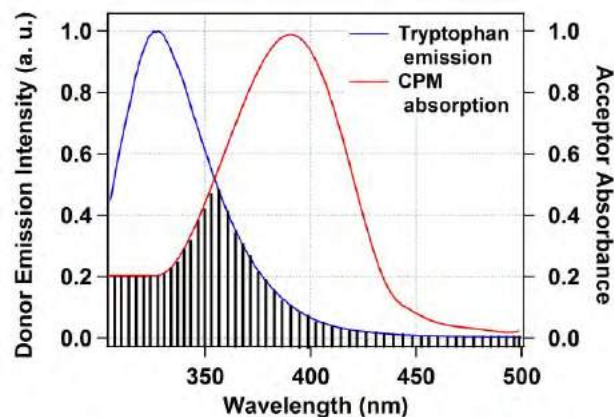
**Figure 4.7:** Normalized tryptophan emission spectra of  $\beta$ -lactoglobulin while excited at 295 nm at (a) different GnHCl concentration (0 M-5 M) and (b) different temperatures (293 K-343 K) in phosphate buffer (pH 7.4, 50 mM); ( $\lambda_{\text{ex}} = 295$  nm).

fluorescence quantum yield of tryptophan emission is found to be 0.03. An increase of quantum yield beyond 1 M GnHCl concentration and a regular decrease of quantum yield with temperature has been observed (figure 4.8c and 4.8d).

**4.2.3. Förster Resonance Energy Transfer (FRET) Experiment:** The significant spectral overlap (see figure 4.9) between tryptophan emission and CPM absorption



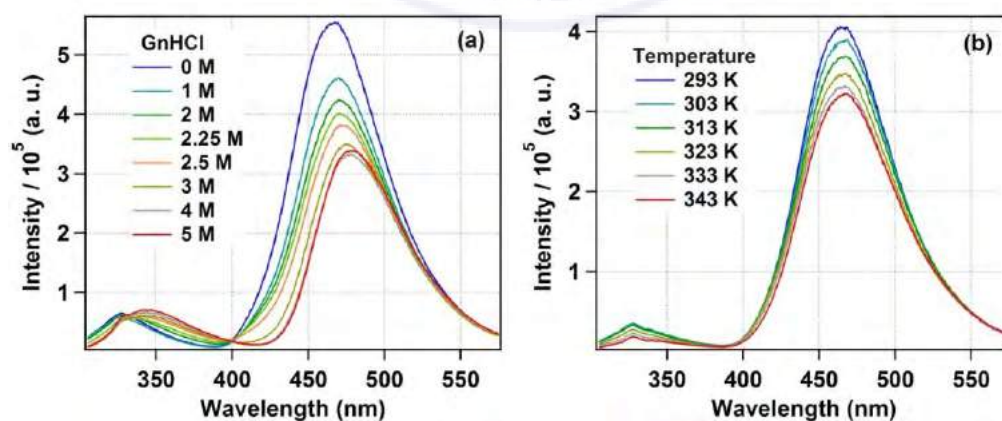
**Figure 4.8:** The variation of emission maxima of tryptophan in  $\beta$ -lactoglobulin with (a) GnHCl and (b) temperature. The variation of emission quantum yield of tryptophan in  $\beta$ -lactoglobulin with (c) GnHCl and (d) temperature.



**Figure 4.9:** Spectral overlap between tryptophan emission and CPM absorption.

allows us to perform FRET experiment for CPM tagged  $\beta$ -lactoglobulin. The Förster distance for this donor acceptor pair was calculated as 21.9 Å considering the emission spectra of tryptophan and the absorption spectra of CPM in the native state of the protein. The efficiency of energy transfer in the native state of  $\beta$ -lactoglobulin is

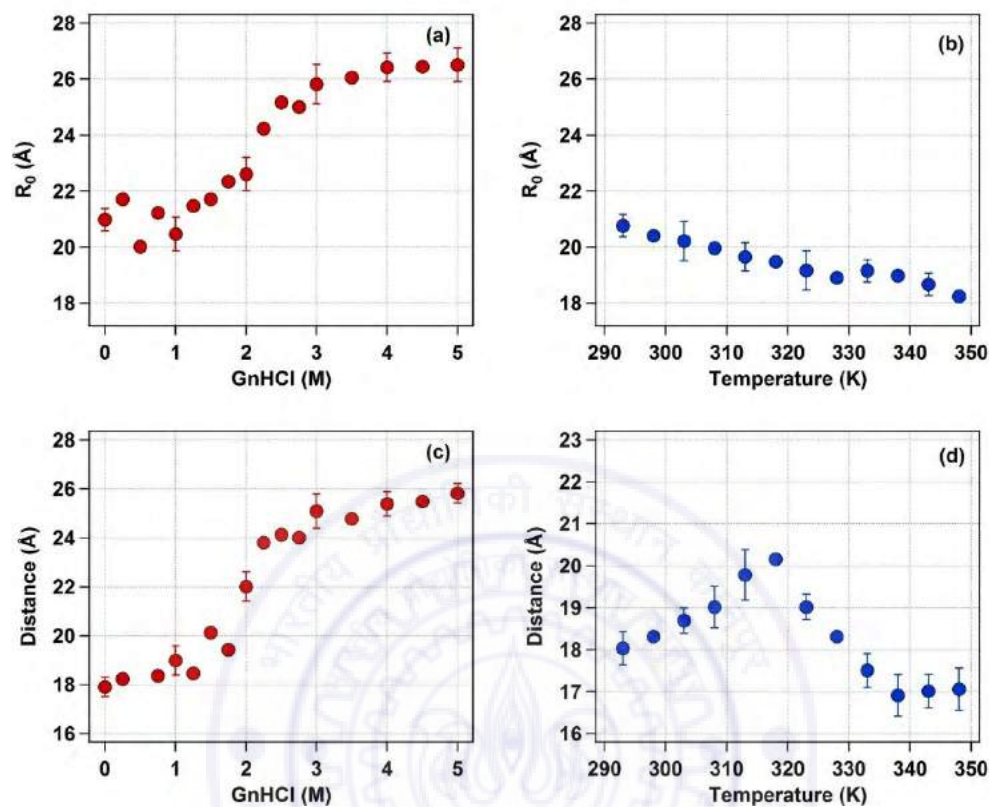
calculated to be 0.70. Using these values, the distance between donor and acceptor was calculated to be 17.9 Å. A few emission spectra corresponding to FRET measurements for different GnHCl and temperature are shown in figures 4.10a and 4.10b, respectively. With varying GnHCl concentration both the spectral overlap, thus the Förster distance (which has been taken care accordingly, figure 4.11a and 4.11b) and the extent of energy transfer changes. Calculation reveals that the average distance between tryptophan and CPM remains constant as in the native state till 1.25 M GnHCl and then increases sharply to reach a maximum value of ~25.4 Å at 4 M GnHCl (figure 4.11c). On application of heat till 318 K, the distance keeps increasing and reaches 20.2



**Figure 4.10:** Fluorescence emission spectra for FRET measurements of CPM labelled  $\beta$ -lactoglobulin while excited at 295 nm at (a) different GnHCl concentration (0 M-5 M) and (b) different temperatures (293 K-343 K) in phosphate buffer (pH 7.4, 50 mM); ( $\lambda_{\text{ex}} = 295$  nm).



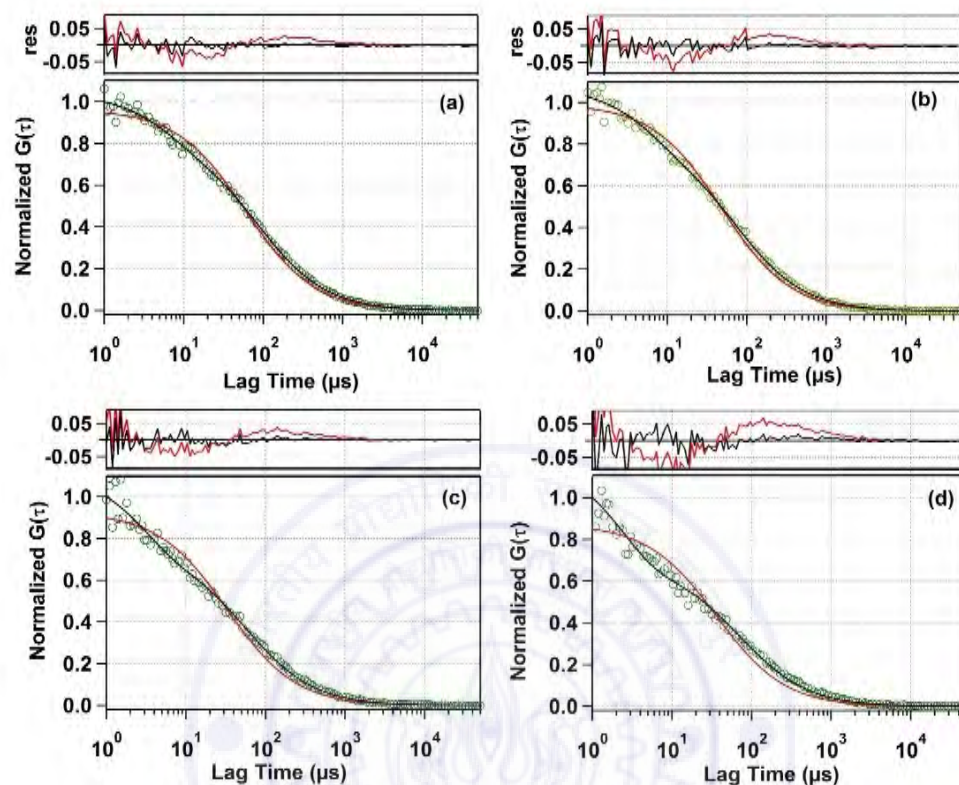
Å and then decreases to 16.9 Å at 338 K. Further increase of temperature does not affect the distance (figure 4.11d).



**Figure 4.11:** Change of Förster radius between tryptophan (donor) and CPM (acceptor) as a function of (a) GnHCl concentration and (b) temperature. Change of distance between tryptophan (donor) and CPM (acceptor) as a function of (c) GnHCl concentration and (d) temperature.

**4.2.4. Fluorescence Correlation Spectroscopic Measurements:** Autocorrelation curves of CPM labelled  $\beta$ -lactoglobulin were recorded at different GnHCl concentrations and temperatures. The curves were all fitted using equation 2.31 (one diffusion component along with one relaxation component) as equation 2.27 (only diffusion component) fails to fit them successfully. A comparison of fit for autocorrelation curve of native state of  $\beta$ -lactoglobulin at 293 K and  $\beta$ -lactoglobulin at 318 K are shown in figures 4.12a and 4.12b, respectively. Figure 4.12c and 4.12d show the comparison of fitting of autocorrelation curves of CPM tagged  $\beta$ -lactoglobulin in the presence of 3 M and 5 M GnHCl, respectively. Some of the normalized fitted lines for the fluorescence autocorrelation function for different GnHCl concentrations and different temperatures are depicted in figures 4.13a and 4.13b, respectively. In the

analysis, the diffusion time ( $\tau_D$ ) is a measure of the diffusion of the whole protein and hence provides information about the size of the protein. The hydrodynamic diameter



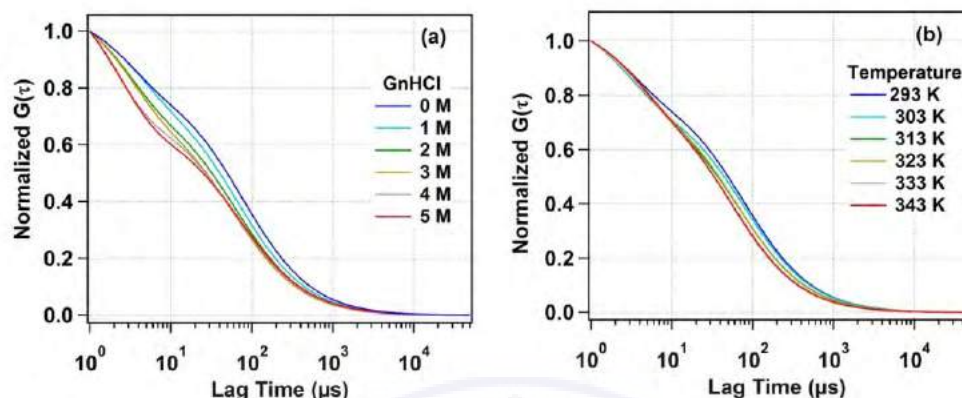
**Figure 4.12:** Experimentally obtained fluorescence autocorrelation curve (green open circles) of CPM-tagged  $\beta$ -lactoglobulin and the comparison of fitting between one diffusion (red solid line) and one diffusion + one exponential (black solid line) model at (a) 293 K and (b) 318 K (c) 293 K, 3 M GnHCl and (d) 293 K, 5 M GnHCl.

of  $\beta$ -lactoglobulin was calculated to be 55.0 Å in its native state using equations 2.21 and 2.37 after proper correction of viscosity as described in section 3.2.4 of chapter 3.

Inspired from similar power dependent study as described in section 3.2.4, this additional exponential time component ( $\tau_R$ ) is ascribed to the quenching of the attached fluorophore by the nearby amino acid residue (discussed in detail in the discussion section) and referred to as the conformational fluctuation time. The timescale of this fluctuation is found out to be 6.0  $\mu$ s in the native state. With increasing GnHCl concentration the hydrodynamic diameter decreases to ~29.0 Å at 3.0 M GnHCl (see figure 4.14a). Upon further increase in GnHCl concentration the diameter starts increasing and reaches 35 Å at 5 M GnHCl. In case of temperature dependent experiments, we have observed a complex variation of the hydrodynamic diameter of



$\beta$ -lactoglobulin. As evident from figure 4.14b the diameter remains constant till 303 K followed by three distinct regions; first one from 303 K to 308 K, second one from 308 K to 338 K and the third one further till 348 K. The hydrodynamic diameter of  $\beta$ -

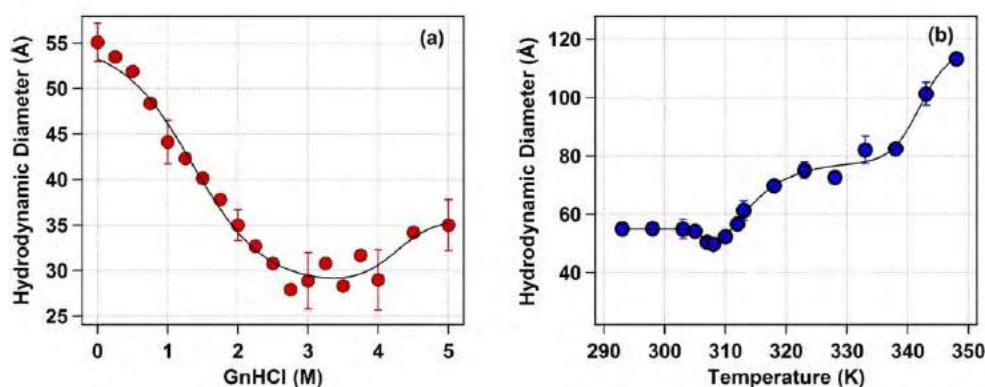


**Figure 4.13:** Normalized fitting lines of autocorrelation curves for CPM-tagged  $\beta$ -lactoglobulin for (a) different GnHCl concentrations and (b) for different temperatures. All data are recorded in 50 mM, pH 7.4 phosphate buffer.

lactoglobulin is found to be 55.0 Å, 49.7 Å, 82.5 Å and 113.3 Å at 293 K, 308 K, 338 K and 348 K, respectively (figure 4.14b). The conformational fluctuation time, however, shows no definite trend with increase in GnHCl or temperature.

### 4.3. Discussion

The similar value of  $\alpha$ -helicity of native and CPM tagged  $\beta$ -lactoglobulin is indicative of negligible or no perturbation of secondary structure upon tagging. The blue shift of CPM emission maxima upon being tagged in  $\beta$ -lactoglobulin is ascribed to the burial of the probe to the hydrophobic core of the protein.<sup>48</sup> The increment in



**Figure 4.14:** Variation of hydrodynamic diameter of  $\beta$ -lactoglobulin as function of (a) GnHCl concentration (red solid circles) and (b) temperature (blue solid circles).

fluorescence quantum yield of CPM is due to the restricted rotation of the coumarin side chain inside the protein matrix. Under the action of GnHCl the reverse effect is observed *i.e.* the emission maxima shifts towards red side. This denotes some structural change, which exposes the CPM molecule to the outside environment. In the native state the 327.5 nm emission maxima and 46 nm bandwidth of the tryptophan emission denote effective emission from a single tryptophan residue located at a hydrophobic region.<sup>41</sup> With increasing GnHCl concentration, this emission maxima is red shifted. However, as shown in figure 4.7a, the shift of emission spectra reflects the generation of a new spectra rather than the alteration of the current spectra. Following an important report by Burstein *et al.* we have assigned the emission of the native state to emission of Trp-19 and the new emission peak to Trp-61, which is located at the surface of  $\beta$ -lactoglobulin.<sup>41</sup> The transition occurs between 1 M to 2.5 M GnHCl. So it is reasonable to comment that under the action of GnHCl, the tryptophan residue (Trp-61), which was silent in the native state, starts fluorescing. This observation is in agreement with the results reported earlier.<sup>41</sup>

The calculated hydrodynamic diameter of CPM tagged  $\beta$ -lactoglobulin at pH 7.4, from the FCS study, matches well with the reported value of hydrodynamic diameter associated with the dimer.<sup>25</sup> Thus, it is established that  $\beta$ -lactoglobulin remains in dimeric state at this pH. As already presented, the hydrodynamic diameter decreases gradually till 3.0 M GnHCl. This is to note here that a change in the hydrodynamic diameter may arise due to two reasons; i) monomerisation and ii) denaturation of the individual units present in the dimer. Had it been the denaturation of the protein, the diameter would have increased. On the other hand, the decrease of hydrodynamic diameter is indicative of the monomerization process. It was expected that in case of co-existing monomer and dimer, the fitting of fluorescence autocorrelation curves should require two diffusion time. However, the data were well fitted with one diffusion component (with one relaxation component), which affirms that the monomer dimer equilibrium is dynamic and the interconversion time is so fast that it can not be captured in our measurement. Such cases have been discussed by Novo *et al.*, Al-Soufi *et al.* and others.<sup>36-39</sup> All of them adopted a model for data fitting

with the concept of average diffusion time. In the present case too, the instrument detects only one average diffusion time depending on the ratio of monomer and dimer. So, it can be affirmed that the monomerization continues till 3.0 M GnHCl and the change in emission maxima of CPM and tryptophan till 3.0 M GnHCl is associated with the local structural change assisting the dimer to monomer conversion. The decrease of  $G(0)$  with increasing GnHCl concentration indicates an increase in the number of particles which is in line with the prediction of monomerization. After 3.0 M GnHCl concentration, the diameter starts increasing again and continues to increase till 5 M GnHCl, which we have assigned to the unfolding of the monomers. Both these processes are found to be reversible in nature.

FRET experiment reveals that the distance between the tagged CPM and the tryptophan residue in the native state to be 17.9 Å. From the crystallographic data the distance of Cys-121 to Trp-19 has been calculated as 9.4 Å (pdb file: 1BEB). This mismatch in the distance may be due to the assumption of the orientation factor ( $\kappa$ ) to be  $2/3$  while calculating the FRET distance, as practically it can deviate significantly from this value. Besides, there is also a chance of energy transfer from Trp-19 of one subunit (subunit A) to CPM of other subunit (subunit B) in the dimeric form. The distance between Cys-121 of monomer A to Trp-19 of monomer B is calculated to be 34.5 Å. Thus, if Cys-121 of sub-unit A accepts energy from Trp-19 of both the sub-units A and B then the calculated distance between Cys and Trp must be somewhere between 9.4 Å and 34.5 Å. Considering all these possibilities, we have ascribed the measured distance as the average distance between Trp and CPM in  $\beta$ -lactoglobulin. However, we would focus more on the variation of this particular distance in the course of chemical and thermal agitation rather than concentrating on its absolute value. With the addition of GnHCl, the distance increases sharply after 1.25 M and reaches 25.4 Å at 4 M. This increment is a result of both the structural change of the protein and the evolution of Trp-61 from being silent to fluorescent. The distance between Trp-61 and Cys-121 of the same subunit is calculated (from the structure file, PDB ID: 1BEB) to be 20 Å in the native state, whereas it is 30.1 Å if the Cys-121 and Trp-61 belong to different sub-units of the dimer. It is evident from FCS results that till 3.0 M GnHCl

monomerization is the prime process, but the increase in distance can not be simply attributed to the monomerization because the system becomes more complicated owing to the presence of a new donor (Trp-61) and the detachment of other subunit, which renders zero energy transfer from Trp-19 of the other subunit. Moreover, in the intermediate range, where a mixture of population is present, the calculation of the associated FRET parameters becomes much complex due to the fractional contribution of all the species. Hence, we restrain from drawing any conclusion from distance dependent FRET.

An overall 2 nm bathochromic shift of tagged-CPM emission maxima on increasing the temperature by 75 K from an initial value of 293 K tells us that no significant change in the local environment around the probe is caused due to heating. The sharp decrease of fluorescence quantum yield with increasing temperature is ascribed to the temperature dependent photophysics of the CPM molecule itself. With increase in the temperature the tryptophan emission intensity decreases but the peak position remains same. This implies the less profound effect of temperature than GnHCl on the local environment of tryptophan residues. At high temperatures the appearance of the shoulder at lower energy side is assigned to the emission of Trp-61 as already discussed. The increasing average distance between Cys and Trp (calculated from FRET measurements) till 318 K indicates the structural change similar to chemical denaturation case. Above this temperature, the average distance decreases indicating a structural change which is different than in the lower temperature range. But, as already discussed, due to involvement of several interferences it will be a poor idea to conclude any structural change from this measurements. FCS data indicates a three step transition as evident from the change of hydrodynamic diameter (figure 4.14b). As reported earlier, monomerization occurs around 313 K followed by reversible and irreversible aggregation till 338 K and above 338 K, respectively.<sup>16,25</sup> There are several reports stating that with application of heat the dimeric structure breaks down to monomers and then these monomeric units form aggregates at higher temperatures.<sup>16,25-31</sup> We ascribe the first transition (293 K to 308 K) to the monomerization, the second one (308 K to 338 K) to the formation of small aggregates (SM) and the last one beyond 338 K to the formation of larger aggregates (LA). Upon incubating the sample at normal temperature (293 K) after measurement, the small aggregates dissociate and the native dimeric state is re-established, which is



not observed for large aggregates. Thus we can conclude that the second step is reversible whereas the third step is irreversible.

To understand the difference of action of GnHCl and heat on the structure of  $\beta$ -lactoglobulin, one must comprehend the nature of forces that control the monomer dimer equilibrium. As already mentioned, the hydrophobic force is one of the main reasons of dimerization. Besides, the interfacial interactions also stabilize the dimeric structure. These interactions include the packing of the ninth  $\beta$ -strands of the two sub-units by the formation of antiparallel two stranded  $\beta$  sheet and the development of hydrogen bonds between the amino acid residues of each sub-unit.<sup>40</sup> O'Brien *et al.* reported that the ability of GnHCl to directly bind on the protein surface and to destroy charge-charge and hydrogen bonding interactions is the key reason of its denaturing effect. In line with all these reports, we propose that, in presence of GnHCl the effective hydrogen bonding network between two monomers breaks down and effective hydrogen bonding between GnHCl and the amino acid residues of the monomeric unit is established. The hydrophobic interactions inside the protein cavity remain affected and it successfully resists the denaturation of the dimeric or the monomeric unit till 3.0 M GnHCl. After the monomerization is complete, the monomeric units start unfolding.

The effect of temperature on the secondary structure of protein has been justified by Dias *et al.*, who introduced a microscopic model of “Clathrate Cages” made up of particularly ordered water molecules around non polar residues to describe thermal denaturation.<sup>41</sup> In the aqueous solution, the hydrophobic protein molecules have a tendency to break the intermolecular hydrogen bond network between the water molecules. Water molecules form cage like ordered structures around the hydrophobic parts to retain the network. The formation of cage like structures costs high entropic loss. In order to compensate the effect of unfavorable entropy change, the protein molecules tend to fold so as to reduce the hydrophobic surface area of contact with the water molecules. The entropy change associated with hydrophobic effect depends on the temperature, and with increase in temperature the ordered cage like structures break down gradually leading to decrease in the negative entropy effect. Thus the



denaturation of the protein is facilitated at higher temperature. In the present case, the entropic effect may lead to the breakage of dimeric structure to monomeric structure and then the hydrophobic surface area is exposed as already discussed. The high temperature aggregation is initiated by the hydrophobic interaction between the monomers.<sup>25-31</sup>

#### 4.4. Calculation of Thermodynamic Parameters

Next, we attend to fit the variation of hydrodynamic diameter with GnHCl and temperature as described in section 2.2 of chapter 2. in order to elucidate the thermodynamic parameters associated with the transitions. The equilibrium constant for the monomerization and denaturation process has been calculated by fitting the variation of hydrodynamic diameter with a three state model<sup>42</sup>  $N_D \rightleftharpoons N_M \rightleftharpoons U_M$  where  $N_D$ ,  $N_M$  and  $U_M$  denotes the dimeric, monomeric and the unfolded monomeric state. Fitting of GnHCl dependent variation of hydrodynamics diameter using equation 2.41 yields the values of  $\Delta G_M^0$  and  $\Delta G_U^0$  as 1.57 kcal mol<sup>-1</sup> and 8.93 kcal mol<sup>-1</sup>, respectively (see figure 4.14a and table 4.1). The midpoints of the two GnHCl induced structural transition,  $N_D \rightarrow N_M$  and  $N_M \rightarrow U_M$ , are found to be 1.35 M and 4.25 M GnHCl, respectively. From the fitting parameters, the size of the monomer and unfolded  $\beta$ -lactoglobulin are estimated as ~29 Å and ~37 Å, respectively (see scheme 2).

For the temperature dependent change of measured hydrodynamics diameter of  $\beta$ -lactoglobulin, a four state model has been adopted as  $N_D \rightleftharpoons N_M \rightleftharpoons SA \rightarrow LA$ . Equation 2.41 is used to fit the temperature dependent variation of the hydrodynamic diameter of  $\beta$ -lactoglobulin considering the dimeric form exists till 305 K (see figure 4.14b) and the fitted parameters are tabulated in table 4.1. In this case, we observed that  $\Delta G_M^0 = 1.76$  kcal mol<sup>-1</sup>,  $\Delta G_{SA}^0 = 0.69$  kcal mol<sup>-1</sup>,  $\Delta G_{LA}^0 = 9.09$  kcal mol<sup>-1</sup> and the hydrodynamic diameters of the monomer, small aggregate (SA) and large aggregate (LA) are estimated to be ~29 Å, ~77 Å and ~117 Å, respectively (scheme 4.2). The size of the large aggregate also matches well with the reported value.<sup>25</sup>

**Table 4.1.** Thermodynamic parameters for structural changes of  $\beta$ -lactoglobulin induced by GnHCl and temperature in phosphate buffer (pH 7.4, 50 mM).

GnHCl	$N_D \rightleftharpoons N_M$			$N_M \rightleftharpoons U_M$		
	$\Delta G_M^0$ (kcal mol <sup>-1</sup> )	$m_M$	$[D]_{1/2}$ (M)	$\Delta G_U^0$ (kcal mol <sup>-1</sup> )	$m_U$	$[D]_{1/2}$ (M)
	1.57	1.16	1.35	8.93	2.10	4.25

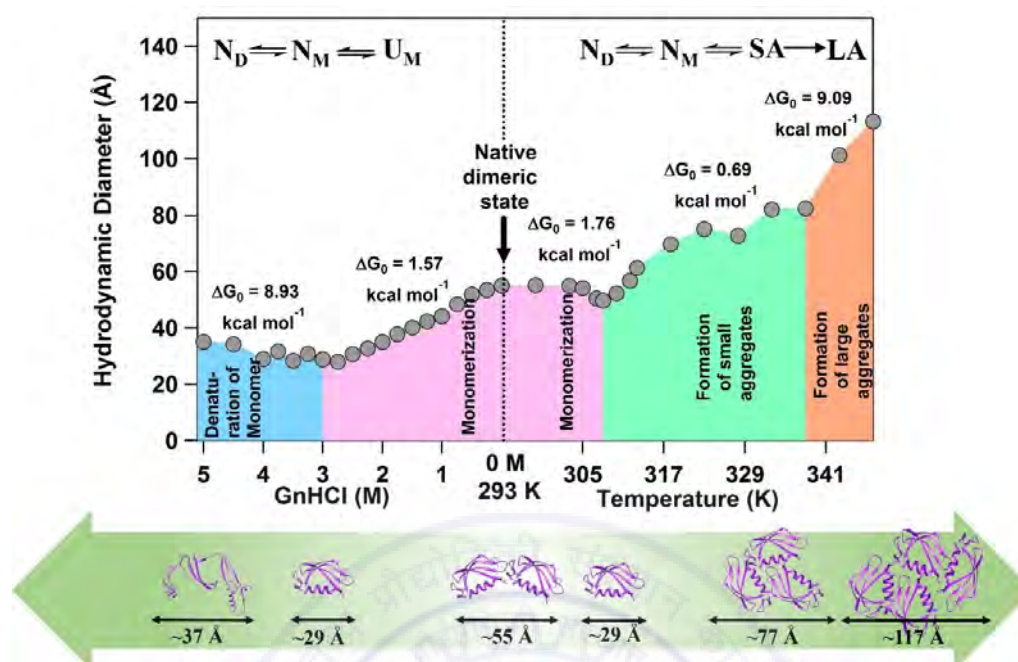
  

Tempe -rature	$N_D \rightleftharpoons N_M$			$N_M \rightleftharpoons SA$			$SA \rightarrow LA$		
	$\Delta G_M^0$ (kcal mol <sup>-1</sup> )	$m_M$	$[D]_{1/2}$ (K)	$\Delta G_{SA}^0$ (kcal mol <sup>-1</sup> )	$m_{SA}$	$[D]_{1/2}$ (K)	$\Delta G_{LA}^0$ (kcal mol <sup>-1</sup> )	$m_{LA}$	$[D]_{1/2}$ (K)
	1.76	0.68	307.6	0.69	0.13	310.3	9.09	0.24	342.9

( $\Delta G$  in kcal mol<sup>-1</sup>,  $m$  is in kcal mol<sup>-1</sup> M<sup>-1</sup>,  $[D]_{1/2}$  is in M).

#### 4.5. Conclusion

A careful observation of the effect of GnHCl and temperature on the structural transformation of  $\beta$ -lactoglobulin reveals that under both the conditions, monomerization is first initiated. In case of GnHCl, the monomerization is followed by the denaturation of the monomeric species, whereas under the effect of temperature, aggregate formation has been noticed. The radius of the small and large aggregates are about  $\sim 2.6$  times and  $\sim 4$  times larger than the monomeric unit. Considering a spherical morphology of the monomeric unit and aggregates with no change in the density upon aggregation, we have estimated that as many as  $\sim 20$  and  $\sim 60$  monomeric units are present in the small and large aggregates, respectively. Distance dependent FRET experiment reveals that upon monomerization, the distance between CPM and tryptophan decreases citing a probability of a loosened structure as compared to the dimer. Upon aggregation, however, the CPM – tryptophan distance is found to be similar to the dimer. This is seconded by CD experiments where we observed no significant change in  $\alpha$ -helicity upon increasing the temperature as compared to the huge change upon addition of GnHCl. The overall GnHCl and temperature dependent structural behaviour of  $\beta$ -lactoglobulin is summarized in scheme 2.



**Scheme 4.2:** Structural evolution of β-lactoglobulin under chemical and thermal perturbation and the associated free energy changes ( $\Delta G_0$ ).

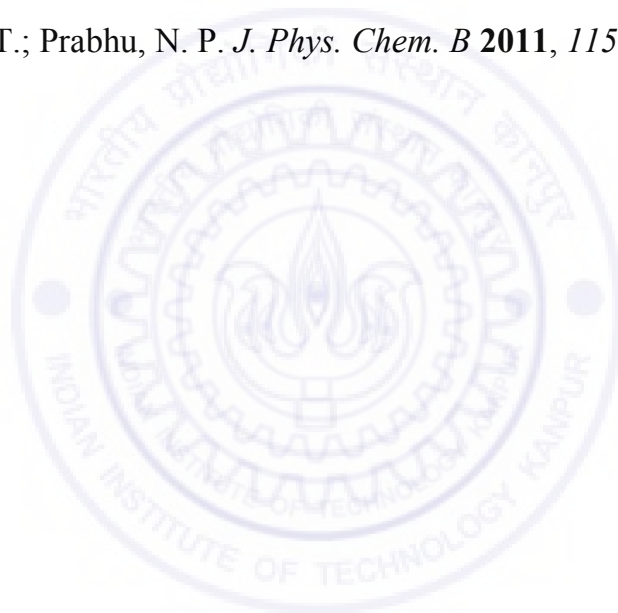
## References

1. Kontopidis, G.; Holt, C.; Sawyer, L. *J. Dairy Sci.* **2004**, *87*, 785–796.
2. Spector, A. A.; Fletcher, J. E. *Lipids* **1970**, *5*, 403–411.
3. Chos, Y.; Batt, C. A.; Sawyer, L. *J. Biol. Chem.* **1994**, *269*, 11102–11107.
4. Maux, S. L.; Bouhallab, S.; Giblin, L.; Brodkorb, A.; Croguennec, T. *Dairy Sci. Technol.* **2014**, *94*, 409–426.
5. Roth-Walter, F.; Pacios, L. F.; Gomez-Casado, C.; Hofstetter, G.; Roth, G. A.; Singer, J.; Diaz-Perales, A.; Jensen-Jarolim, E. *PLoS One* **2014**, *9*, 1–8.
6. Hamada, D.; Segawa, S.; Goto, Y. *Nat. Struct. Biol.* **1996**, *3*, 868–873.
7. Uhrínová, D.; Smith, M. H.; Jameson, G. B.; Uhrín, D.; Sawyer, L.; Barlow, P. N. *Biochemistry* **2000**, *2*, 3565–3574.
8. Yan, Y.; Seeman, D.; Zheng, B.; Kizilay, E.; Xu, Y.; Dubin, P. L. *Langmuir* **2013**, *29*, 4584–4593.
9. Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. *Faseb J.* **2016**, *10*, 75–83.
10. Sakurai, K.; Goto, Y. *J. Biol. Chem.* **2002**, *277*, 25735–25740.
11. Sakurai, K.; Oobatake, M.; Goto, Y. *Protein Sci.* **2001**, *10*, 2325–2335.
12. Burova, T. V.; Choiset, Y.; Tran, V.; Haertlé, T. *Protein Eng.* **1998**, *11*, 1065–1073.
13. Mercadante, D.; Melton, L. D.; Norris, G. E.; Loo, T. S.; Williams, M. A. K.; Dobson, R. C. J.; Jameson, G. B. *Biophys. J.* **2012**, *103*, 303–312.
14. Taulier, N.; Chalikian, T. V. *J. Mol. Biol.* **2001**, *314*, 873–889.
15. Kelly, M. J.; Reithel, F. J. *Biochemistry* **1964**, *10*, 2639–2644.
16. Aymard, P.; Durand, D.; Nicolai, T. *Int. J. Biol. Macromol.* **1996**, *19*, 213–221.
17. Renard, D.; Lefebvre, J.; Griffin, M. C. A.; Griffin, W. G. *Int. J. Biol. Macromol.* **1998**, *22*, 41–49.
18. Verheul, M.; Pedersen, J. S.; Roefss, S. P.; Kruif, K. G. D. *Biopolymers* **1999**, *49*, 11–20.
19. Hattori, M.; Hiramatsu, K.; Kurata, T.; Nishiura, M.; Takahashi, K.; Ametani, A.; Kaminogawa, S. *Biochim. Biophys. Acta. Proteins Proteomics* **2005**, *1752*, 154–165.

20. Greene, R. F. Jr.; Pace, C. N. *J. Biol. Chem.* **1974**, *17*, 5388–5393.
21. Valente-mesquita, V. L.; Botelho, M. M.; Ferreira, T. *Biophys. J.* **1998**, *75*, 471–476.
22. Qi, X. L.; Holt, C.; McNulty, D.; Clarke, D. T.; Brownlow, S.; Jones, G. R. *Biochem. J.* **1997**, *324*, 341–346.
23. Dar, T. A.; Singh, L. R.; Islam, A.; Anjum, F.; Moosavi-movahedi, A. A.; Ahmad, F. *Biophys. Chem.* **2007**, *127*, 140–148.
24. Alfonso, L. D.; Collini, M.; Baldini, G. *Biochemistry* **2002**, *41*, 326–333.
25. Griffin, W. G.; Griffin, M. C. A.; Martin, S. R.; Price, J. *J. Chem. Soc. Faraday Trans.* **1993**, *89*, 3395–3406.
26. Iametti, S.; Gregori, B. D.; Vecchio, G.; Bonomi, F. *Eur. J. Biochem.* **1996**, *112*, 106–112.
27. Elofsson, U. M.; Dejmek, P.; Paulsson, M. A. *Int. Dairy J.* **1996**, *6*, 343–357.
28. Tolkach, A.; Kulozik, U. *Lait* **2007**, *87*, 301–315.
29. Seo, J. A.; Hédoux, A.; Guinet, Y.; Paccou, L.; Affouard, F.; Lerbret, A.; Descamps, M. *J. Phys. Chem. B* **2010**, *114*, 6675–6684.
30. Iametti, S.; Cairoli, S.; Gregori, B. D.; Bonomi, F. *J. Agric. Food. Chem.* **1995**, *43*, 53–58.
31. Qi, X. L.; Brownlow, S.; Holt, C.; Sellers, P. *Biochim. Biophys. Acta* **1995**, *1248*, 43–49.
32. Fukusima, Y.; Kawata, Y.; Onda, T.; Kitagawa, M. *Am. J. Clin. Nutr.* **1997**, *65*, 30–35.
33. Fukushima, Y.; Kawata, Y.; Onda, T.; Kitagawa, M. *J. Nutr. Sci. Vitaminol. (Tokyo)* **1997**, *43*, 673–678.
34. Lovegrove, J. A.; Osman, D. L.; Morgan, J. B.; Hampton, S. M. *Gut* **1993**, *34*, 203–207.
35. Kelly, S. M.; Price, N. C. *Curr. Protein Pept. Sci.* **2000**, *1*, 349–384.
36. Al-soufi, W.; Novo, M.; Felekyan, S.; Kühnemuth, R.; Seidel, C. A. M. *J. Am. Chem. Soc.* **2005**, *109*, 8775–8784.



37. Suthari, P.; P, H. K.; Doddi, S.; Bangal, P. R. *J. Photochem. Photobiol. A: Chem.* **2014**, *284*, 27–35.
38. Al-soufi, W.; Reija, B.; Felekyan, S.; Seidel, C. A. M.; Novo, M. *ChemPhysChem* **2008**, *9*, 1819–1827.
39. Novo, M.; Felekyan, S.; Seidel, C. A. M.; Al-soufi, W. *J. Phys. Chem. B* **2007**, *111*, 3614–3624.
40. Brownlow, S.; Cabral, J. H. M.; Cooper, R.; Flower, D. R.; Yewdall, S. J.; Polikarpov, I.; North, A. C. T.; Sawyer, L. *Structure* **1997**, *5*, 481–495.
41. Dias, C. L.; Ala-Nissila, T.; Wong-ekkabut, J.; Vattulainen, I.; Grant, M.; Karttunen, M. *Cryobiology* **2010**, *60*, 91–99.
42. Naidum, K. T.; Prabhu, N. P. *J. Phys. Chem. B* **2011**, *115*, 14760–14767.





## *Chapter 5*

---

### *Investigation of Structure and Dynamics of Papain under Denaturation*

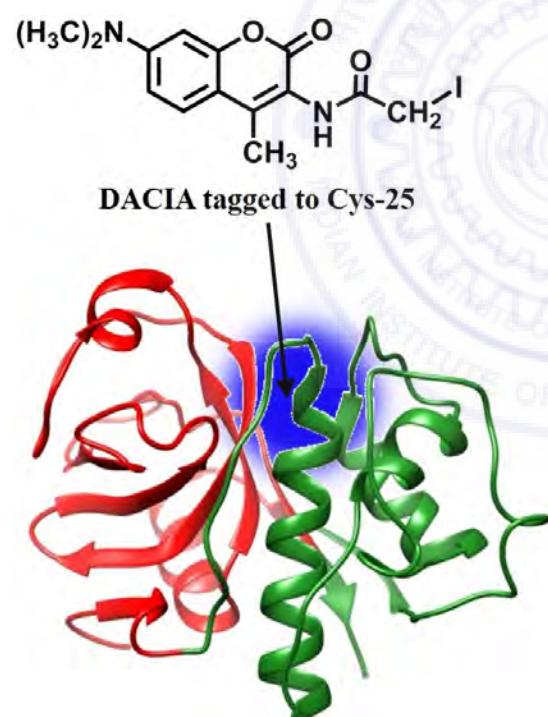
**Sengupta, B.;** Chaudhury, A.; Das, N.; Sen, P. *Protein Pept. Lett.* **2017**  
(DOI: 10.2174/0929866524666170811145838)

*Guanidine hydrochloride (GnHCl) induced structural and dynamical variation of papain is studied by site-specific tagging of papain at Cys-25 residue with N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA). Steady-state, time-resolved, and single molecule level fluorescence techniques have been applied to map the change in the local environment by monitoring the fluorescence response of the probe. It was found that papain undergoes a two-step denaturation in the presence of GnHCl. Fluorescence correlation spectroscopic (FCS) data indicates that the size (hydrodynamic diameter) of native papain is  $\sim 36.8$  Å, which steadily increases to  $\sim 53$  Å in the presence of 6 M GnHCl. FCS study also revealed that the conformational fluctuation time of papain is 6.3  $\mu$ s in its native state, which decreases to 2.7  $\mu$ s in the presence of 0.75 M GnHCl. It is believed that in the presence of small amount of GnHCl the active site of papain takes up a more compact structure (although the overall size increases) than in the native state, which has been designated as the intermediate state. Upon further increase in GnHCl concentration, the conformational fluctuation time increase monotonically till 6 M GnHCl, where the time constant is measured as 14  $\mu$ s. On the other hand, the measurement of ellipticity, hence the helical structure, by circular dichroism spectroscopy is found to be incapable to capture such structural transformation.*

### 5.1. Introduction

Enzymes are vital components to keep the physiological system properly functional. Almost all the cellular reactions are catalyzed and controlled by enzymes failing which, serious biological disorder may arise. Enzymes are present in all types of cells such as procaryotes and eucaryotes and most of them are proteins. Papain is a plant enzyme which helps to digest proteins.<sup>1-10</sup> It has got two domains connected by a cleft and the active site of papain is centered around this cleft. The single cysteine residue which is a prime member of the active site triad, can be site specifically labelled using thiol specific dyes facilitating the spectroscopic study of the active site of papain.

The structural transformation of papain under environmental effects has been studied by various researchers till date.<sup>11-22</sup> Edwin *et al.* found that papain stays in native state till pH 2.5 and below this pH it loses most of its tertiary interactions.<sup>16</sup>



**Scheme 5.1:** Schematic structural representation of papain (input file downloaded from protein data bank; PDB ID: 9PAP).

Rašković *et al.* reported that papain undergoes cold denaturation and aggregation under preservation in cold storage.<sup>21</sup> There have been several reports on the formation of an intermediate state of papain under influence of several denaturing agents. Arana *et al.* concluded that under thermal denaturation the two domains of papain unfold independently through an intermediate state.<sup>15</sup> The intermediate state here corresponds to a state where domain L is unfolded but domain R is intact.<sup>15</sup> Briggs *et al.* found that the change in emission maxima of the intrinsic fluorescence increases stepwise

on gradual increase of guanidine hydrochloride (GnHCl) concentration, which hints towards the formation of an intermediate state.<sup>14</sup> On the other hand when Zhuang *et al.*



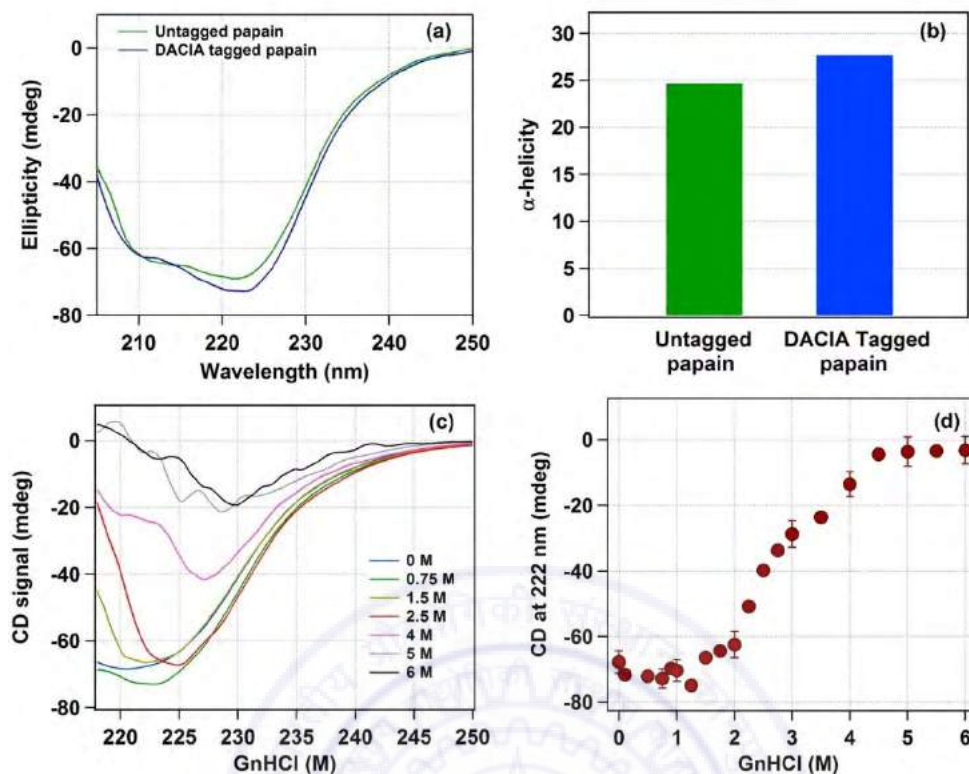
tagged domain L with acrylodan and studied the denaturation by monitoring acrylodan fluorescence, they failed to detect any intermediate state.<sup>12</sup> So it is still not clear whether the domains unfold monotonously or it involves one or more intermediate states.

Here steady-state, time-resolved and single-molecule fluorescence techniques were used to study the behaviour of papain under chemical denaturant GnHCl. The single free cysteine residue at position 25 in domain L is covalently labeled with N-(7-dimethylamino-4-methylcoumarin-3-yl) iodoacetamide (DACIA) for this purpose. DACIA is a coumarin based solvatochromic dye (scheme 5.1). It is reported that DACIA successfully tags free cysteine residue of proteins.<sup>23</sup> It is also reported that DACIA shows significant change in fluorescence property upon a change in its local environment, which makes it suitable probe for studying structural transformation of proteins.<sup>23</sup> Single molecular spectroscopic investigations are employed to measure the hydrodynamic radius of papain as a function of GnHCl and also the conformational dynamics of the active site. The spectroscopic parameters also deliver the information about the local structural change of papain around DACIA.

## 5.2. Results

**5.2.1. Circular Dichroism Spectroscopy:** In order to confirm that the fluorescent labelling does not cause any significant change in the secondary structure of papain, the circular dichroism (CD) spectra of untagged and tagged papain in 50 mM phosphate buffer (pH = 7.4) has been recorded (figure 5.1a). The  $\alpha$ -helicity of native unlabelled papain is found to be 24.7% and it matches with the reported value for the native papain.<sup>24-26</sup> A bar diagram of comparison of  $\alpha$ -helicity is shown as figure 5.1b.

The CD spectra of DACIA tagged papain were also recorded for varying GnHCl concentration (figure 5.1c), where a monotonous change in the CD signal has been observed with increase in GnHCl concentration. This is to note that at high GnHCl concentrations, data could not be recorded below 215 nm. Thus the  $\alpha$ -helicity could not be calculated in this case. On the other hand, the negative signal at 222 nm is also known to represent the folded 3D structure of papain and in the present case the loss of

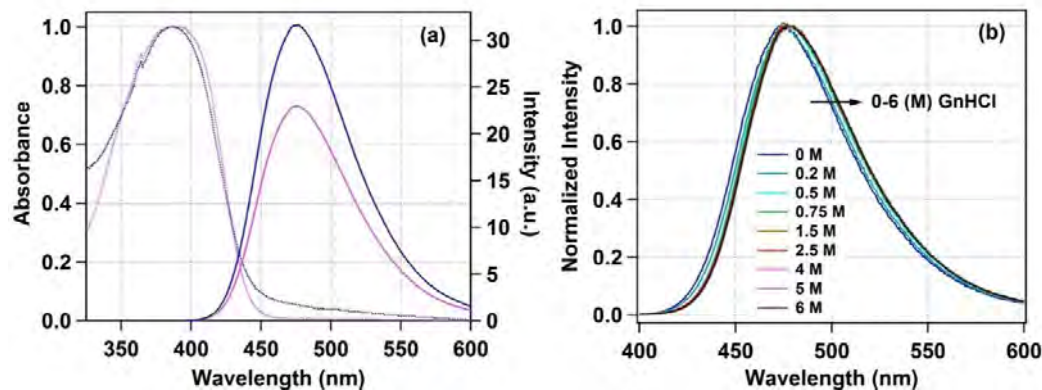


**Figure 5.1:** (a) CD spectra of untagged papain (11.4  $\mu$ M) and DACIA tagged papain (11.2  $\mu$ M). (b) Comparison of  $\alpha$ -helicity of untagged and tagged papain as bar diagram. (c) CD spectra of DACIA tagged papain (11.3  $\mu$ M) with different concentrations of GnHCl. (d) Variation of CD signal at 222 nm as a function of GnHCl concentration. All data are recorded in phosphate buffer (pH 7.4, 50 mM).

222 nm signal is monitored as a function of GnHCl concentration, which has been depicted in figure 5.1d. The variation CD signal of papain at 222 nm as the function of GnHCl concentration clearly shows three regions. From 0 M to 1.25 M GnHCl, no significant change in the CD signal at 222 nm (-68 mdeg) was observed, while from 1.25 – 4.5 M a sharp decrease in the magnitude of CD signal and remains constant till 6 M GnHCl (-3 mdeg) was noticed.

**5.2.2. Steady-State UV-Vis Absorption and Fluorescence Spectroscopy:** The absorption spectra of DACIA in 50 mM phosphate buffer (pH 7.4) shows a peak at 389 nm. The corresponding emission maxima is found to be at 478.5 nm upon exciting at 389 nm. When DACIA is tagged in papain, the absorption maximum has been observed at 387 nm and an excitation at this wavelength results in a fluorescence band with an emission maximum of 473.5 nm, which is 5 nm blue shifted than that in buffer (figure

5.2a). The spectroscopic parameters for untagged and tagged DACIA are listed in table 5.1.

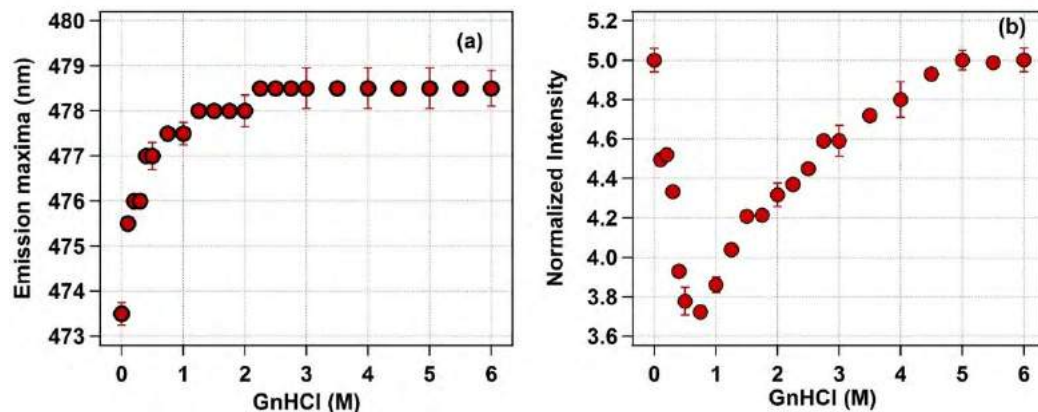


**Figure 5.2:** (a) Absorption (dotted lines) and emission spectra (solid lines) of free DACIA (pink) and DACIA tagged to papain (blue). Absorption spectra has been normalized to 1 and emission spectra has been normalized with respect to absorbance at the excitation wavelength in order to show the relative difference in the fluorescence intensity. (b) Fluorescence spectra of DACIA tagged papain at different GnHCl concentration; ( $\lambda_{\text{ex}} = 387 \text{ nm}$ )

Addition of GnHCl also induces change in the absorption and emission properties of DACIA tagged in papain. Some of the normalized emission spectra has been shown in figure 5.2b. The emission maxima shifts steeply from 473.5 nm to 477.5 nm upon changing the GnHCl concentration from 0 M to 1 M and become 478.5 nm beyond 2.25 M GnHCl (figure 5.3a). On the other hand, about 33% decrease in emission intensity has been noticed with 0.75 M GnHCl in comparison with the native state of papain, which then fully recovered monotonically till 6 M GnHCl. The variation of relative emission intensity with GnHCl concentration is shown as figure 5.3b. At 6 M

**Table 5.1.** Spectroscopic parameters for untagged DACIA and DACIA tagged in papain in phosphate buffer (pH 7.4, 50 mM).

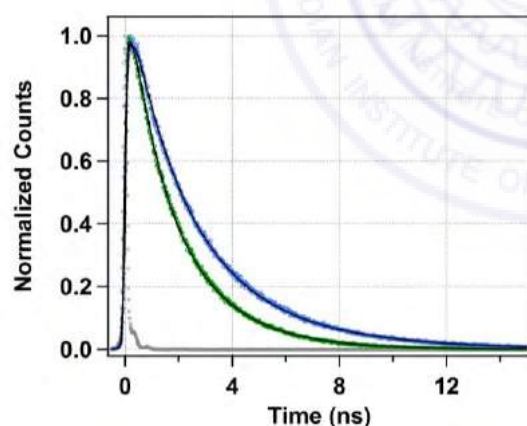
	Absorption Maximum (nm)	Emission Maximum (nm)	Fluorescence Lifetime (ns)
DACIA in buffer	389.0	478.5	1.97
DACIA tagged in papain in buffer	387.0	473.5	2.98



**Figure 5.3:** (a) Variation of emission maxima of DACIA tagged to papain with GnHCl concentration. (b) Variation of emission intensity of DACIA tagged to papain with GnHCl concentration.

GnHCl the fluorescence intensity is found to be comparable with the native state of papain.

**5.2.3. Fluorescence Lifetime Measurements:** Fluorescence transient of free DACIA in buffer measured at 478 nm shows a bi-exponential decay with time components of 0.95 ns (15%), 2.16 ns (85%). The average lifetime is calculated to be 1.97 ns. When DACIA is tagged to papain, its average fluorescence lifetime recorded at the emission

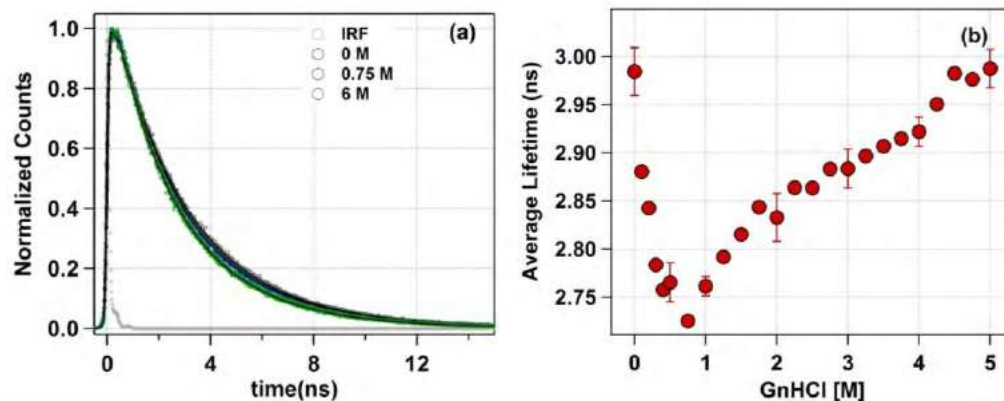


**Figure 5.4:** Normalised fluorescence decay curves of free DACIA and DACIA tagged to papain upon excitation at 375 nm. All data are recorded in phosphate buffer (pH 7.4, 50 mM); ( $\lambda_{\text{ex}} = 375$  nm).

maxima increases to 2.98 ns characterized by two decay components of 1.85 ns (36%) and 3.61 ns (64%) (figure 5.4). On addition of GnHCl, the bi-exponential decay nature of DACIA in papain remains similar, which are shown in figure 5.5a. In figure 5.5b, the variation of average lifetime of DACIA in papain with GnHCl concentration is depicted. Clearly the average lifetime decreases

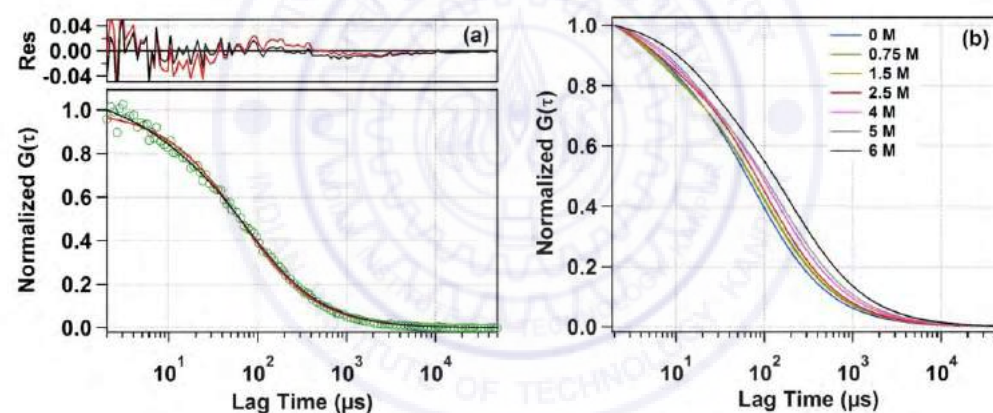
from 2.98 ns at 0 M GnHCl to 2.72 ns at 0.75 M GnHCl and then starts rising again till 6 M GnHCl to reach a value of ~3.00 ns.





**Figure 5.5:** (a) Fluorescence decay for DACIA tagged papain for different GnHCl concentrations. (b) Variation of average lifetime of DACIA tagged papain with change in GnHCl concentration. All data are recorded in phosphate buffer (pH 7.4, 50 mM); ( $\lambda_{\text{ex}} = 375$  nm)

**5.2.4. Fluorescence Correlation Spectroscopic Studies:** To understand the effect of GnHCl on the overall dimension and conformational fluctuation time of papain,

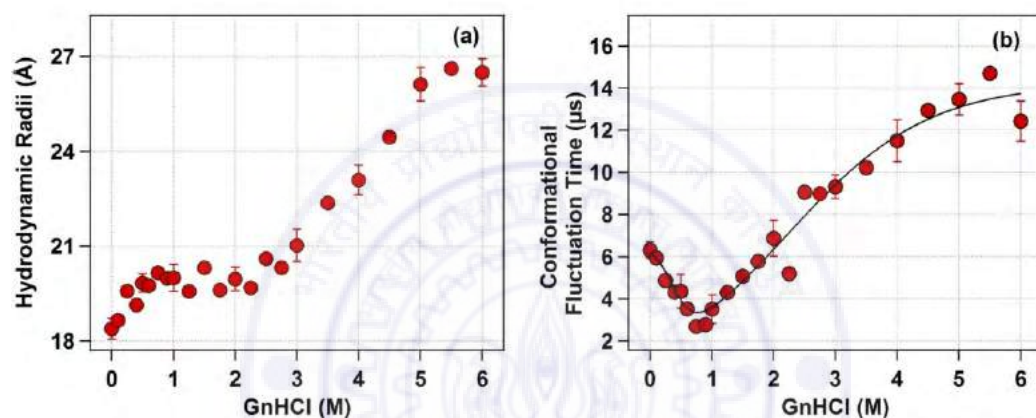


**Figure 5.6:** (a) Comparison of fitting of the fluorescence auto-correlation curves with single diffusion equation (in red) and after incorporation of a relaxation time component (in black). (b) Normalized autocorrelation curves of DACIA tagged papain for different GnHCl concentration.

fluorescence correlation spectroscopic measurements of DACIA tagged papain as a function of GnHCl concentration has been performed. The autocorrelation curves were attempted to fit with a single diffusion model (equation 2.27) at first but all of them returned a bad fitting residual. Upon incorporation of an additional relaxation term (equation 2.31) the fitting residual improved significantly. A comparison of fitting for the autocorrelation curve in buffer is shown in figure 5.6a. The detail description of all the fitting models has been elaborated in section 2.1.5.4 of chapter 2. Few normalized fits of the autocorrelation curves are shown in figure 5.6b. Hydrodynamic radius ( $r_H$ )



of papain at different GnHCl concentration were calculated from their diffusion time using equations 2.21 and 2.37, which are depicted in figure 5.7a. The hydrodynamic radius of papain in the native state is found to be 18.4 Å, which upon addition of 0.75 M GnHCl increases to ~20.0 Å. After that it remains constant till 2 M and then again starts increasing till 5.5 M GnHCl and reaches a value of ~26.6 Å. The relaxation time  $\tau_i$  is found to be 6.3  $\mu$ s in the native state. With increasing GnHCl concentration  $\tau_i$  first decreases to a minimum 2.7  $\mu$ s at 0.75 M GnHCl and then increases to ~14  $\mu$ s for 6 M GnHCl (figure 5.7b).



**Figure 5.7:** (a) Variation of hydrodynamic radii of papain with GnHCl concentration. (b) Variation of conformational fluctuation time of papain with GnHCl concentration. Data are recorded in phosphate buffer (pH 7.4, 50 mM).

### 5.3. Discussion

$\alpha$ -helicity of a protein denotes its helical content, which in turn is a parameter associated with the secondary structure of the protein. From CD experiment it is seen that the  $\alpha$ -helicity of the untagged and tagged papain are similar, which indicates that the secondary structure of papain is retained after tagging with DACIA. From the variation of CD signal of papain at 222 nm with varying GnHCl concentrations (Figure 5.1d), it can be seen that the secondary structure starts rupturing at ~1.25 M GnHCl and it continues till 4.5 M GnHCl. From this study it has been concluded that the maximum loss of the secondary structure of papain took place at 4.5 M GnHCl and further increase of GnHCl concentration does not affect papain structure anymore. The monotonous change in the CD signal indicates that the overall denaturation does not involve any intermediate state.

The emission maxima of DACIA shows ~5 nm blue shift upon being tagged with papain. This has been attributed to the burial of DACIA in the more hydrophobic protein core of papain upon tagging as compared to the bulk buffer. Thus it is expected that as the protein unfolds, DACIA will be exposed to more hydrophilic environment and will show a red shift in the emission maximum. Indeed, a sharp increase of emission maximum till 1 M GnHCl is observed, which is indicative of denaturation of papain with low GnHCl concentration. After 2.25 M GnHCl no further change in emission maxima has been observed and it has been concluded that denaturation of papain completes at 2.25 M GnHCl. It is to be noted here that this change reflects the change of local environment of papain around DACIA only.

As already discussed, both the fluorescence intensity and the lifetime reaches a minima at ~0.75 M GnHCl concentration and then again increases till 6 M GnHCl concentration. If the protein underwent monotonous denaturation under application of chemical denaturant, it is expected to get a regular trend of observable with increasing GnHCl concentration. The minimum fluorescence intensity and lifetime at ~0.75 M GnHCl hints towards the formation of an intermediate state at this point. Thus the overall effect of GnHCl can be summed up as a three state process involving two transitions. The first transition gives rise to an intermediate state between 0 to 0.75 M GnHCl. The midpoint of this transition lies at 0.2 M GnHCl. The second transition takes place from 0.75 to 6 M GnHCl with a midpoint at 2.25 M GnHCl. Here, it is surprising to observe that the emission intensity and lifetime of DACIA in the native state and the unfolded state are almost similar. The emission intensity and lifetime is mainly guided by the large amplitude torsional motion of the 7-dimethyl amino group of DACIA. The similarity of the said parameters in the native and the unfolded state is probably due to a minimum alteration of the local micro-environment (e.g. steric hindrance or viscosity) of DACIA associated with the structural change of papain in native and unfolded state, which consequently renders the fluorescence intensity and lifetime same in those two states. However, in the proposed intermediate state at around ~0.75 M GnHCl, the microenvironment is different leading to a decrease in the fluorescence intensity and lifetime. Thus, no inference regarding the unfolding of

papain can be drawn from the intensity and lifetime change of DACIA, which actually motivated us to perform single molecular level investigation.

Fluorescence autocorrelation curves of DACIA tagged papain at different GnHCl concentrations were best fitted with one diffusion and one relaxation component model. The relaxation component is assigned to the breathing motion or the conformational motion of papain owing to its independence on laser power (as described in section 3.2.4 of chapter 3). Due to this kind of dynamics the fluorescence intensity of DACIA fluctuates based on the distance of the nearby electron rich amino acid residues. The time component associated with this motion is termed as conformational fluctuation time ( $\tau$ ). With the increment of GnHCl concentration, the hydrodynamic radii increases as expected. There is a flat region from 0.75 M to 2 M GnHCl where the extent of change of hydrodynamic radii is minimum. Thus at this concentration range the overall size of the protein remains constant. After 2 M GnHCl the hydrodynamic radii increases sharply till 6 M GnHCl. Thus it can be inferred that the overall denaturation is most prominent in 2-6 M GnHCl concentration range. Here it is to be mentioned that even the size increases from 0-0.75 M GnHCl, the  $\alpha$ -helicity remains unchanged and even when the size remains constant from 0.75 to 2 M GnHCl concentration, the  $\alpha$ -helicity changes significantly. The  $\alpha$ -helicity is indicative of the secondary structure, whereas the hydrodynamic radius reflects the tertiary structure. So, it can be stated on a broad sense that the secondary structure of papain is less perturbed at lower GnHCl concentration than the tertiary structure. This is accounted by the fact that the driving force behind the formation of helices and the formation of tertiary structure are different and hence it is not unnatural that the breakdown of these two structures will need different extent of perturbation.

The change of conformational fluctuation time with GnHCl concentration closely matches with that of the fluorescence lifetime. The two distinct regions denote the formation of intermediate and the unfolding process respectively. The conformational fluctuation time of the intermediate state is significantly less than the native and the unfolded state of papain indicating that the local structure around the probe becomes more compact in this intermediate state. In this compact structure the

amino acid residues (which are responsible for fluorescence quenching) can not move too far from the probe due to the size constraint. This explanation also justifies the lower fluorescence intensity and lower fluorescence lifetime at this state.

#### 5.4. Calculation of Thermodynamic Parameters

The transition of papain with GnHCl concentration can be summarized as a two step model,  $N \leftrightarrow I \leftrightarrow U$ , where N denotes native state, I denotes intermediate state and U denotes the unfolded state, respectively.<sup>27</sup> The thermodynamic parameters associated with the two steps i.e. intermediate state formation and denaturation has been calculated using the method described in section 2.2 of chapter 2 using one intermediate state. The conformational fluctuation time has been taken as the observable and the variation of conformational fluctuation time with GnHCl has been fitted with equation 2.41 with  $n=1$ , i.e. one intermediate state, and the corresponding thermodynamic parameters are listed in table 5.2. The denaturation midpoint for intermediate formation has been calculated as 0.43 M and that for denaturation has been calculated as 2.20 M GnHCl.

**Table 5.2:** Thermodynamic parameters for the GnHCl induced unfolding of DACIA tagged papain in phosphate buffer (pH 7.4, 50 mM) as calculated by fitting the variation of conformational fluctuation time with GnHCl concentration. ( $\Delta G$  in kcal mol<sup>-1</sup>,  $m$  is in kcal mol<sup>-1</sup> M<sup>-1</sup>,  $[D]_{1/2}$  is in M).

$N \leftrightarrow I$			$I \leftrightarrow D$		
$\Delta G_1^0$	$m_1$	$[D]_{1/2}$	$\Delta G_2^0$	$m_2$	$[D]_{1/2}$
1.46	3.36	0.43	1.13	0.51	2.20

( $\Delta G$  in kcal mol<sup>-1</sup>,  $m$  is in kcal mol<sup>-1</sup> M<sup>-1</sup>,  $[D]_{1/2}$  is in M).

#### 5.5. Conclusion

The existence of intermediate state during the GnHCl induced unfolding of papain has been confirmed in the present investigation. The time scale of conformational fluctuation dynamics of papain in its native state has also been reported along with how it varies during the unfolding process. The conformational fluctuation time of papain in the intermediate state is found to be 2.3 and 5.2 times smaller than in

native and denatured state, whereas the overall size of the intermediate state lies between the two extremes. The possible structural features of this intermediate state are (i) it retains the  $\alpha$ -helicity as the native state, (ii) The overall size of the protein is a little bigger than the native state, and (iii) active site cysteine comes in close proximity with the nearby quenching amino acid residues indicating a more compact structure of the active site.





## References

1. Amri, E.; Mamboya, F. *Am. J. Biochem. Biotechnol.* **2012**, *8*, 99–104.
2. Mitchel, R. E. J.; Chaiken, I. M.; Smith, E. L. *J. Biol. Chem.* **1970**, *245*, 3485–3492.
3. Dubey, V. K.; Pande, M.; Singh, B. K.; Jagannadham, M. V. *African J. Biotechnol.* **2007**, *6*, 1077–1086.
4. Drenth, J.; Jansonius, J. N.; Koekoek, R.; Swen, H. M.; Wolthers, B. G. *Nature* **1968**, *218*, 929–932.
5. Wang, J.; Xiang, Y. F.; Lim, C. *Protein Eng.* **1994**, *7*, 75–82.
6. Lown, J. A.; Dale, B. J. *Immunohematology* **1995**, *11*, 140–142.
7. Lowe, G. *Philos. Trans. R Soc. L. B Biol. Sci.* **1970**, *257*, 237–248.
8. Kamphuis, I. G.; Kalk, K. H.; Swarte, M. B. A.; Drenth, J. *J. Mol. Biol.* **1984**, *179*, 233–256.
9. Farnsworth, N. R.; Akerele, O.; Bingel, A. S.; Soejarto, D. D.; Guo, Z. *Bull. World Health Organ.* **1985**, *63*, 965–981.
10. Krishna, K. L.; Paridhavi, M.; Patel, J. A. *Indian J. Nat. Prod. Resour.* **2008**, *7*, 364–373.
11. Sharma, Y. V.; Jagannadham, M. V. *Protein Pept. Lett.* **2003**, *10*, 83–90.
12. Zhuang, P.; Butterfield, D. A. *Biophys. J.* **1991**, *60*, 623–628.
13. Rialdi, G.; Battistel, E. *J. Therm. Anal.* **1996**, *47*, 17–25.
14. Briggs, G. S.; Freedman, R. B.; Goodenough, P. W.; Sumner, I. G. *Biochem. Soc. Trans.* **1992**, *20*, 257S.
15. Hernández-Arana, A.; Soriano-García, M. *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.* **1988**, *954*, 170–175.
16. Edwin, F.; Jagannadham, M. V. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **2000**, *1479*, 69–82.
17. Naeem, A.; Khan, K. A.; Khan, R. H. *Arch. Biochem. Biophys.* **2004**, *432*, 79–87.
18. Chamani, J.; Heshmati, M.; Rajabi, O.; Parivar, K. *Open Surf. Sci. J.* **2009**, *1*, 20–29.

19. Tiktopulo, E. I.; Privalov, P. L. *FEBS Lett.* **1978**, *1*, 57–58.
20. Edwin, F.; Jagannadham, M. V. *Biochem. Biophys. Res. Commun.* **1998**, *252*, 654–660.
21. Rašković, B.; Popović, M.; Ostojić, S.; Anđelković, B.; Tešević, V.; Polović, N. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2015**, *150*, 238–246.
22. Ghosh, S. *Coll. Surf. B Biointerfaces* **2005**, *41*, 209–216.
23. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
24. Woody, R. W. *Theory of Circular Dichroism of Proteins. In Circular Dichroism and Conformational Analysis of Biomolecules*, G. D. Fasman, Ed. Plenum publishing corp., New York, **1996**.
25. Corrêa, D.; Ramos, C. *African J. Biochem. Res.* **2009**, *3*, 164–173.
26. Qadeer, A.; Zaman, M.; Khan, R. H. *Biochemistry (Moscow)* **2014**, *79*, 785–796.
27. Naidu, K. T.; Prabhu, N. P. *J. Phys. Chem. B* **2011**, *115*, 14760–14767.



## *Chapter 6*

---

### *Startling Temperature Effect on Proteins when Confined: Human Serum Albumin in Reverse Micelle*

**Sengupta, B.;** Yadav, R; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, 18, 14350–14358.

*The present work reports the effect of confinement, and temperature therein, on the conformational fluctuation dynamics of domain-I of human serum albumin (HSA) by fluorescence correlation spectroscopy (FCS). The water-pool of sodium bis(2-ethylhexyl) sulfosuccinate (AOT) reverse micelle was used as the confined environment. It was observed that the conformational fluctuation time of domain I of HSA is about 6 times smaller compared to bulk medium when confined in a water-pool of 3.5 nm radius. On increasing the size of the water-pool the conformational fluctuation time was found to increase monotonically and approaches to the bulk value. The effect of confinement is at par with the general belief of restricted motion of a macromolecule upon confinement. However, the effect of temperature was found to be surprising. An increase in the temperature from 298 K to 313 K induces larger change in conformational fluctuation time in HSA, when confined. In the bulk medium, apparently there is no change in conformational fluctuation time in the aforementioned temperature range, whereas, when HSA is present in AOT water-pool of radius 3.5 nm, about 88% increase in the fluctuation time was observed. The observed prominent thermal effect on the conformational dynamics of domain-I of HSA in the water-pool of AOT reverse micelle as compared to bulk medium was concluded to arise from the confined solvent effect.*



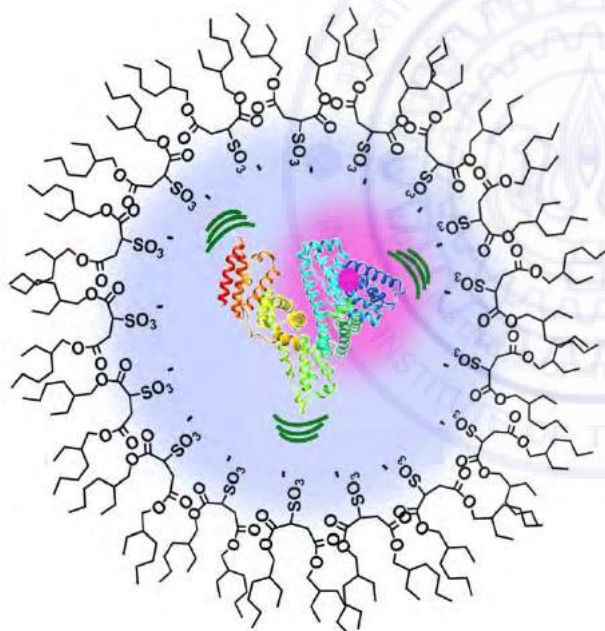
### 6.1. Introduction

The physiological importance of proteins makes them one of the most significant classes of biological molecules and have always catch hold of attention among researchers since early times.<sup>1,2</sup> Although proteins are characterised by their unique structures, recent studies indicate that the unique structures of proteins are not static in nature and is associated with low-frequency breathing motion, which is related to the conformational fluctuation dynamics of the protein.<sup>3,4</sup> Such a conformational dynamics is believed to be responsible for several biological activities of a protein.<sup>5</sup> This is expected to be sensitive to the overall structure of the protein and found to be mostly appreciable during the course of its denaturation.<sup>6-13</sup> In a recent work, the effect of denaturation on the conformational dynamics of human serum albumin (HSA) has been reported.<sup>8</sup> On denaturation by guanidine hydrochloride, the conformational fluctuation time of domain-I of HSA was found to increase from the native state and has been assigned due to the increased size of the protein. Moreover, an intermediate state was identified through this study, which was not otherwise observed. However, no signature of such an intermediate state was observed during the thermal denaturation of HSA. Recently Bhattacharyya and co-workers have studied the effect of room temperature ionic liquid (RTIL) on the conformational fluctuation time of HSA and cytochrome-c in both native and unfolded state. They have reported that the addition of RTIL slows down the dynamics of the protein in the native state, whereas in the denatured state the dynamics become faster in presence of RTIL.<sup>9,10</sup> Apart from experimental methods, theoretical approach has also been utilised to visualize protein conformational dynamics.<sup>11,14-17</sup> Gollapalli *et al.* performed molecular dynamics simulation of a heat shock protein derived from mycobacterium tuberculosis to map the conformational change during the protein folding process. They could trace the dynamics related to the two rigid body movements across the hinge region, which are assigned as the movement of the apical domain away from the equatorial domain and inward movement of the intermediate domain, respectively.<sup>15</sup>

On the other hand, behaviour of proteins in reverse micellar system is also a topic of contemporary research as it closely mimics the biological cell.<sup>18-24</sup> Malik *et al.*

studied the interaction of 1-anilinonaphthalenesulfonate (ANS) with HSA under confinement and reported that binding of ANS to HSA in AOT reverse micelle is significantly retarded as compared to buffer and suggested the non-specificity of domain specific binding sites in the reverse micellar system as the probable reason.<sup>25</sup> It has been generally confirmed that proteins behave differently in confined environment as compared to bulk.<sup>21,22,24,25</sup> Considering the idea of confinement, it is expected that proteins will undergo faster conformational dynamics under confined environment due to restriction in the free movement of the polypeptide chains and the effect of temperature should also be less prominent under confinement compared that in bulk.

This study has been performed with an intension to elucidate the conformational dynamics of HSA in a nano cavity using fluorescence correlation spectroscopy (FCS). HSA has been



**Scheme 6.1:** Schematic representation of tetramethylrhodamine-5-maleimide (TMR) and its adduct to the Cys-34 residue located in the domain-I of human serum albumin.

tagged with tetramethylrhodamine-5-maleimide (TMR) at Cys-34 residue located in domain I (section 2.4.4 of chapter 2). As a model for the confined environment the well defined water-pool of sodium bis(2-ethylhexyl) sulfosuccinate (AOT) reverse micelle, which is known from early days as a mimic of biological confined environment (Scheme 6.1) was chosen.<sup>26, 27</sup> The size of the water-pool can be controlled easily and efficiently and is characterized by the ratio of the

concentration of polar phase to concentration of surfactant ( $W_0$ ), and expressed as<sup>28</sup>

$$W_0 = \frac{[water]}{[AOT]} \quad (\text{Eq. 6.1})$$

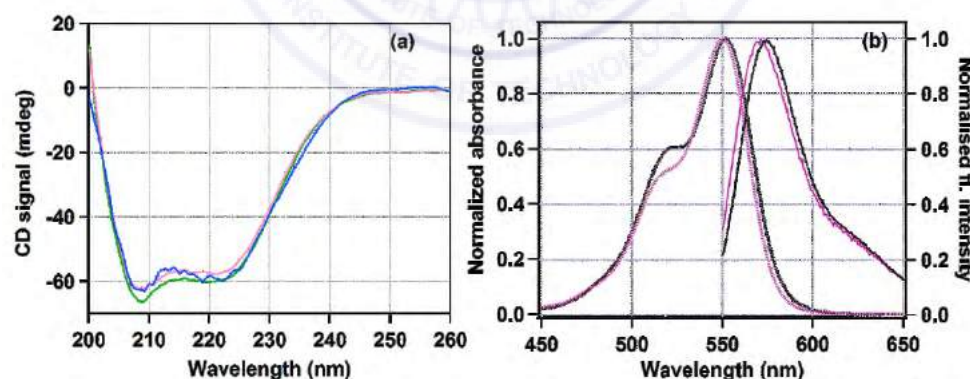
With an increase in  $W_0$ , the radius of the AOT water-pool ( $r_w$ ) increases. The radius of the water-pool ( $r_w$ ) inside the AOT reverse micelle can be written as a function of  $W_0$  as

$$r_w = \left( \frac{90W_0n_{ag}}{4\pi} \right)^{1/3} \quad (\text{Eq. 6.2})$$

where,  $n_{ag}$  is the average aggregation number of surfactants per reverse micelle.<sup>29</sup> The value of  $W_0$  were varied from 20 to 46 and the conformational dynamics of domain-I of HSA encapsulated within it was monitored. I have also checked the variation of conformational dynamics as a function of temperature from 298 K to 313K.

## 6.2. Results

**6.2.1. Circular Dichroism Spectroscopy:** The far-UV circular dichroism spectra of untagged HSA, TMR tagged HSA and HSA inside reverse micelle having  $W_0=20$  were recorded in order to confirm the structural perturbation due to labelling or encapsulation. The CD spectra have been shown in figure 6.1a. The  $\alpha$ -helicity has been found to be 64% and 60% for untagged and tagged protein. The  $\alpha$ -helicity of tagged HSA in buffer and in the water-pool of AOT reverse micelle of  $W_0=20$  has calculated to be same.

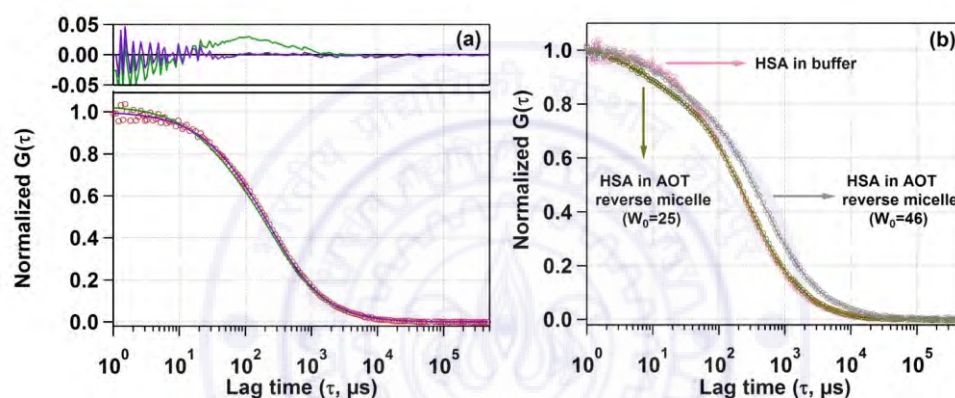


**Figure 6.1:** (a) Circular dichroism spectra of untagged human serum albumin in buffer (green), TMR-tagged human serum albumin in buffer (pink) and tagged human serum albumin in water-pool of AOT reverse micelle  $W_0=20$  (blue); overall concentration of HSA is maintained at  $5 \mu\text{M}$  in each case. (b) Absorption (dotted lines) and emission (solid lines) spectra of untagged TMR (Black) and TMR tagged to Cys-34 of HSA (Pink).

**6.2.2. Steady-State Absorption and Fluorescence Spectroscopy:** Figure 6.1b shows the absorption and emission spectra of free TMR and TMR tagged HSA in 50 mM

phosphate buffer, pH 7.4. The absorption maxima of pure TMR in phosphate buffer is at 550 nm. Upon excitation at 540 nm, TMR gives an emission maxima at 571 nm in phosphate buffer. When TMR is tagged with HSA, both the absorption and the emission maxima remains almost unchanged with final values of 552 nm and 574 nm, respectively.

**6.2.3. Fluorescence Correlation Spectroscopic Study:** The fluorescence autocorrelation curve of TMR-HSA in phosphate buffer is shown in figure 6.2a, which was best fitted with equation 2.31 having a conformational fluctuation term along with

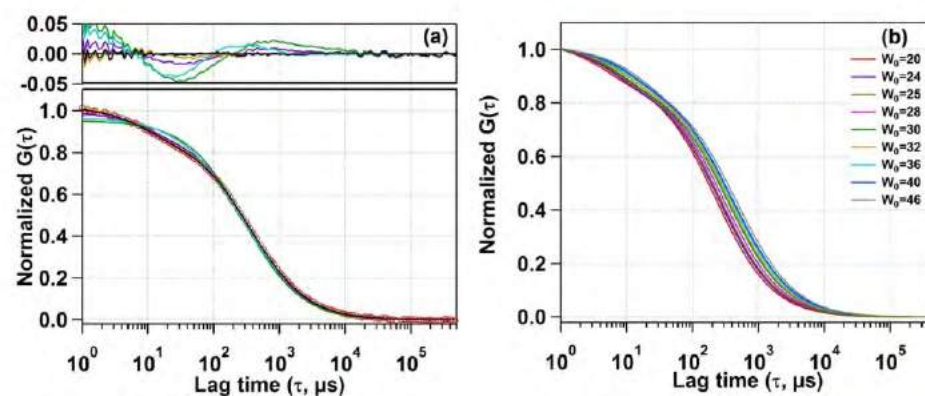


**Figure 6.2:** Fluorescence autocorrelation curves of (a) TMR tagged human serum albumin (HSA) in buffer and the comparison of fitting using only diffusion time (green) and diffusion time + relaxation time component (violet) models. (b) TMR tagged human serum albumin (HSA) in buffer and in AOT reverse micelle ( $W_0=25$  and  $W_0=46$ ). Solid black lines indicate best fits of the corresponding data.

one diffusion component. The value of  $\tau_D$  and  $\tau_R$  obtained from the fitting are 226  $\mu s$  and 26.9  $\mu s$ , respectively. Next, the autocorrelation curves of the TMR tagged HSA inside the water-pool of AOT reverse micelle with varying sizes were recorded. The autocorrelation curves for AOT encapsulated HSA look significantly different from the one in pure buffer solution (autocorrelation curves for  $W_0=25$  and 46 and bulk buffer has been shown in figure 6.2b). The experimental autocorrelation curves of the reverse micellar systems were first attempted to fit by a single component diffusion model (equation 2.27). But it returned bad fitting residual and the quality of fit was also considerably bad for all  $W_0$  values. Consequently, equation 2.31 was tried, which incorporates a conformational fluctuation term along with one diffusion component, to



fit the data. Although this modified equation fits the autocorrelation curve for HSA in pure buffer solution quite well, but the fitting in case of the HSA inside the water-pool

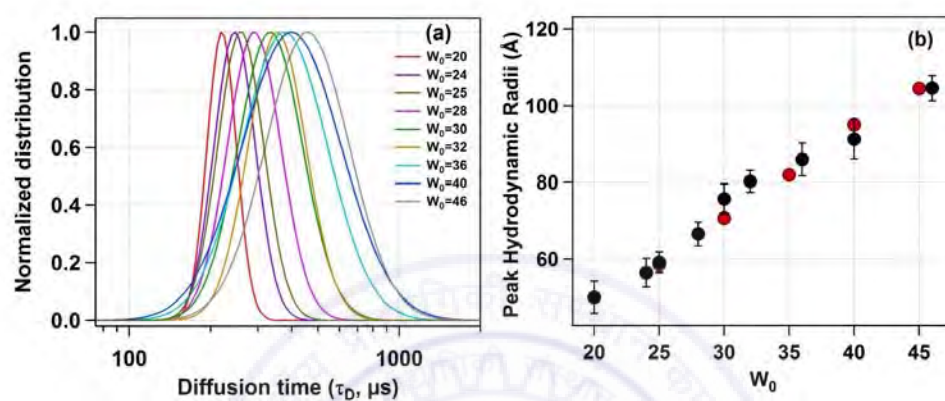


**Figure 6.3:** (a) Experimental fluorescence autocorrelation curve of TMR tagged HSA in the water-pool of AOT reverse micelle ( $W_0=25$ ) and its fitting with (i) 1 diffusion model (red line), (ii) 1 diffusion + 1 conformational relaxation component model (brown line), (iii) Gaussian distribution on diffusion model (blue line), (iv) Gaussian distribution on diffusion + 1 conformational relaxation component model (green line), and (v) Gaussian distribution on diffusion + Gaussian distribution on conformational relaxation component model (black line). Upper panel shows the residuals of the corresponding fits. In the inset, a comparison of the fluorescence autocorrelation is shown between TMR tagged HSA in the water-pool of AOT reverse micelle ( $W_0=25$ ) and untagged HSA along with Rhodamine 6G in the water-pool of AOT reverse micelle ( $W_0=25$ ). (b) Normalized fits of fluorescence autocorrelation curves of TMR tagged HSA in AOT reverse micelle with increasing  $W_0$ .

of AOT reverse micelle are still found to be unsatisfactory. Hence a Gaussian distribution model for diffusion time as reported by Pal *et al.* was tried.<sup>28</sup> However, this model (equation 2.32) also fails to deliver a satisfactory fitting. Finally the data was best fitted when I introduced a Gaussian distribution in the conformational fluctuation time along with a distribution in the diffusion time (as in equation 2.35). A comparison of all these fitting results for  $W_0=25$  are shown in figure 6.3a. Evidently, the residual and the fit are best in the case of one relaxation component and one diffusion component both having a Gaussian type of distribution. This model was found best in case of all other  $W_0$  values. To validate the observation of protein fluctuation time control experiment was performed, where, untagged HSA along with Rhodamine 6G was incorporated in the water-pool of AOT reverse micelle of  $W_0=25$ . In this control experiment the data could be fitted satisfactorily with distributed diffusion model. In case of TMR tagged HSA inside reverse micelle a regular change in the fluorescence



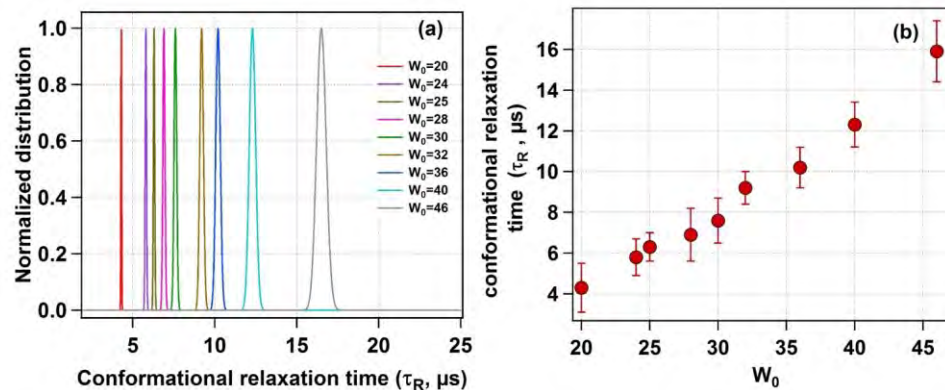
autocorrelation curves with increase in  $W_0$  value of the AOT reverse micelle was observed. The normalized best fits has been shown in figure 6.3b. There is a monotonous increase in the peak diffusion time with increase in  $W_0$  value which is shown in figure 6.4a. At  $W_0 = 20$  ( $r_w = 3.5$  nm) the peak  $\tau_D$  value is found to be 218  $\mu$ s, which increases to 456  $\mu$ s for a larger reverse micelle having  $W_0 = 46$  ( $r_w = 7.0$  nm). The



**Figure 6.4:** (a) Distribution of diffusion time ( $\tau_D$ ) of TMR tagged HSA for different water pool sizes. (b) Variation of hydrodynamic radii of AOT reverse micelle with increasing water pool size (black) as estimated from the peak  $\tau_D$  values along with reported value of hydrodynamic radii (red) in reference 28.

**Table 6.1.** Fitting result of fluorescence autocorrelation curves of TMR tagged HSA in the water-pool of AOT reverse micelle of varying sizes at 298 K.

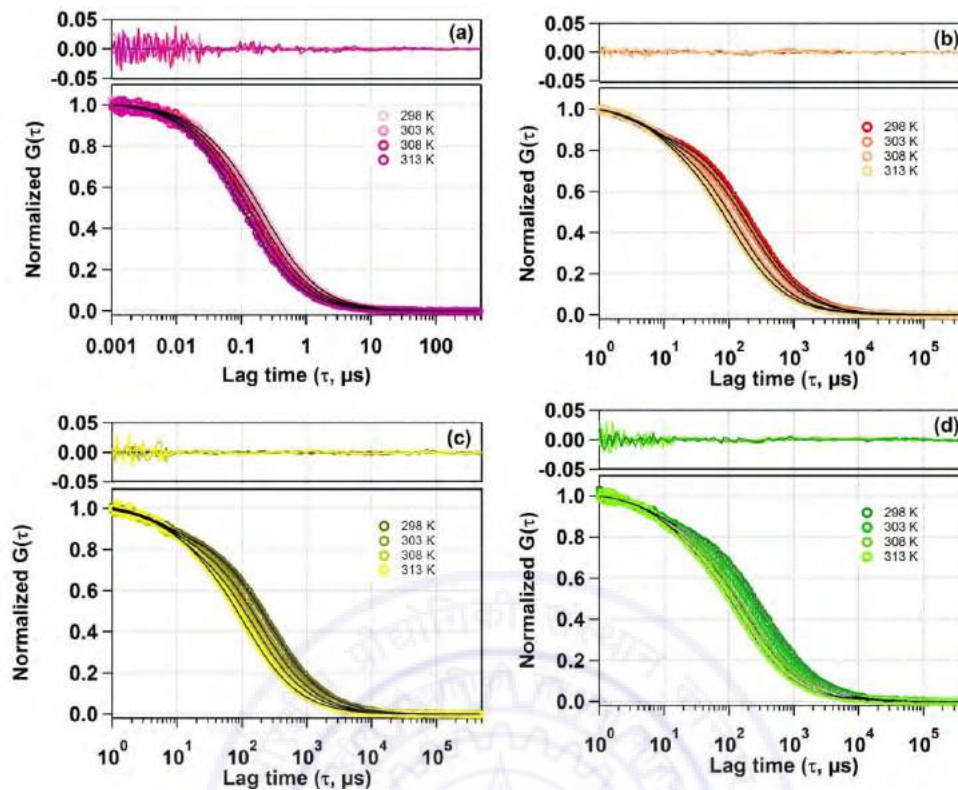
$W_0$	Peak value of diffusion time, $\tau_D$ ( $\mu$ s)	Peak value of conformational fluctuation time, $\tau_R$ ( $\mu$ s)	Hydrodynamic radii (nm) and FWHM (nm)	Reported hydrodynamic radii (nm) and FWHM (nm) <sup>38</sup>
20	240	4.4	5.08 (5.64)	-
24	274	5.5	5.77 (6.71)	-
25	-	-	-	5.85 (6.25)
28	355	7.9	7.48 (5.84)	-
30	371	8.2	7.82 (5.35)	7.07 (7.27)
32	390	8.9	8.22 (4.20)	-
35	-	-	-	8.20 (8.37)
36	415	11.0	8.75 (12.38)	-
40	447	12.2	9.40 (8.68)	9.51 (8.85)
45	-	-	-	10.45 (10.61)
46	515	16.7	10.85 (8.07)	-



**Figure 6.5:** (a) Distribution of conformational relaxation time ( $\tau_R$ ) of TMR tagged HSA for different water-pool sizes. (b) Change in conformational relaxation time ( $\tau_R$ ) with increasing water-pool size of AOT reverse micelle.

diffusion coefficient and the hydrodynamic radii of the reverse micelle in each case is calculated from the peak value of the Gaussian distribution for the diffusion component ( $\tau_D$ ) using equation 2.21 and 2.37 respectively. The variation of hydrodynamic radii of reverse micelle has been shown in figure 6.4b and is tabulated in table 6.1. Figure 6.5a shows the distribution of the conformational fluctuation time ( $\tau_R$ ) for domain I of HSA confined in different sized water-pool of AOT reverse micelles. The peak value of the Gaussian distribution of  $\tau_R$  in case of confined systems are found to be much lower as compared to the value in bulk buffer. In reverse micelle having  $W_0 = 20$  ( $r_w = 3.5$  nm) this value is 4.3  $\mu s$ . Further, with increasing  $W_0$  value,  $\tau_R$  also increases uniformly and become 15.9  $\mu s$  for  $W_0 = 46$  ( $r_w = 7.0$  nm), which is shown in figure 6.5b.

Next, the effect of temperature on HSA inside AOT reverse micelle in the range of 298 K-313 K was studied. As it is known that the AOT reverse micellar system remain unchanged in this temperature region the study was restricted in the above temperature range.<sup>30</sup> In this temperature range the conformational fluctuation time of domain-I of HSA in bulk buffer medium was found to be unaltered (see figure 6.6a for the autocorrelation curves and table 6.2 for the data). However, when HSA is present in the water-pool of AOT reverse micelle the fluctuation time was found to increase as the temperature was increased. For  $W_0 = 20$  ( $r_w = 3.5$  nm), the peak  $\tau_R$  value was increased from 4.3  $\mu s$  at 298 K to 7.8  $\mu s$  at 313 K (see figure 6.6b for the autocorrelation curves and table 6.2 for the data). An 81% increase in the fluctuation time has been



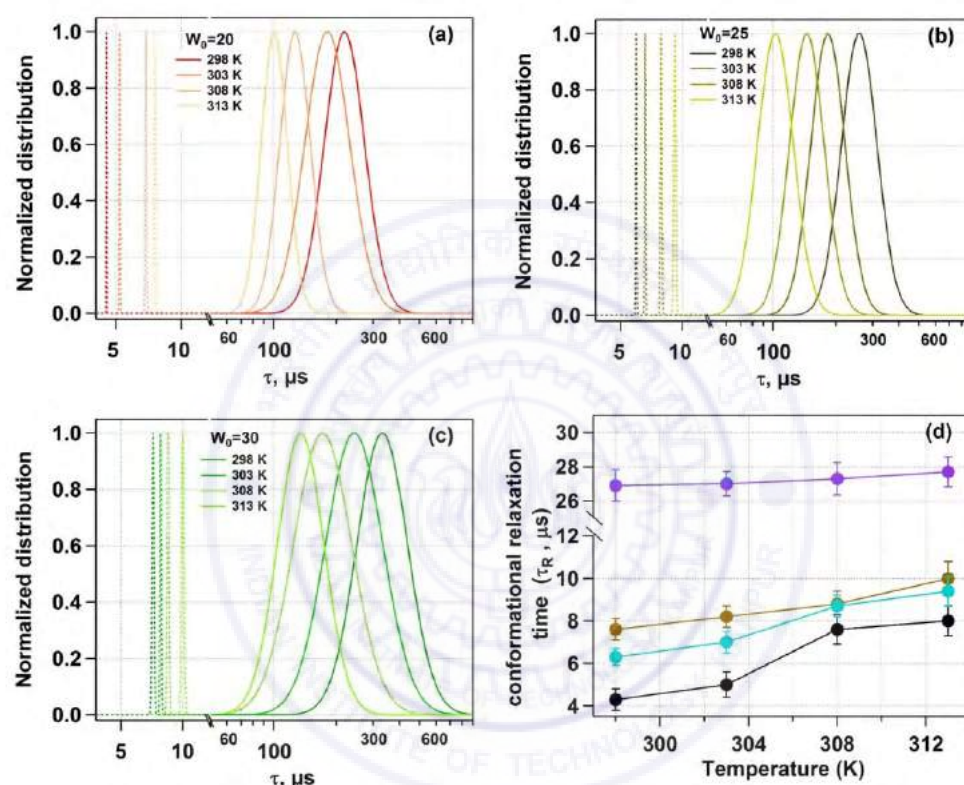
**Figure 6.6:** Normalized fits of autocorrelation curves of (a) TMR tagged HSA in buffer at different temperatures; (b) Normalized fits of autocorrelation curves of TMR tagged HSA in  $W_0=20$  AOT reverse micelle at different temperatures; (c) Normalized fits of autocorrelation curves of TMR tagged HSA in  $W_0=25$  AOT reverse micelle at different temperatures; (d) Normalized fits of autocorrelation curves of TMR tagged HSA in  $W_0=30$  AOT reverse micelle at different temperatures.

observed in this case. On increasing the size of the water-pool to  $W_0=25$  ( $r_w=4.3$  nm)<sup>29</sup> and  $W_0=30$  ( $r_w=5.1$  nm)<sup>29</sup>, the effect of temperature on the conformational fluctuation dynamics decreases gradually compared to  $W_0=20$ . At  $W_0=25$  the fluctuation time increases from  $6.3$   $\mu$ s at  $298$  K to  $9.4$   $\mu$ s at  $313$  K with an increase of  $\sim 49\%$ . At  $W_0=$

**Table 6. 2.** Data of temperature dependence of conformational relaxation time of domain-I of HSA inside the water-pool of AOT reverse micelle of  $W_0 = 20, 25, 30$  and in bulk buffer medium.

Temperature (K)	$\tau_R$ of domain-I of HSA inside			
	AOT reverse micelle $W_0=20$ ( $\mu$ s)	AOT reverse micelle $W_0=25$ ( $\mu$ s)	AOT reverse micelle $W_0=30$ ( $\mu$ s)	Buffer ( $\mu$ s)
298	4.4	6.0	7.8	27.8
303	4.9	6.5	8.1	27.6
308	8.0	8.3	8.6	27.9
313	8.3	8.9	9.7	28.2

30 there is only ~31% increase in fluctuation time on increasing the temperature from 298 K (7.8  $\mu$ s) to 313 K (9.7  $\mu$ s). The normalized autocorrelation curves for  $W_0=25$  and  $W_0=30$  are shown in figure 6.6c and 6.6d. The distribution of  $\tau_D$  and  $\tau_R$  for  $W_0=20$ ,  $W_0=25$  and  $W_0=30$  at different temperatures has been shown in figure 6.7a, 6.7b and 6.7c respectively and the change in the peak  $\tau_R$  for the corresponding  $W_0$  values has been shown in figure 6.7d.



**Figure 6.7:** Normalized distribution of diffusion time ( $\tau_D$ ) and conformational fluctuation time ( $\tau_R$ ) of TMR tagged HSA in (a)  $W_0=20$  AOT reverse micelle at different temperatures; (b)  $W_0=25$  AOT reverse micelle at different temperatures; (c)  $W_0=30$  AOT reverse micelle at different temperatures; (d) Change in conformational fluctuation time ( $\tau_R$ ) of domain-I of HSA with temperature for bulk buffer (violet), inside water-pool of AOT reverse micelle of  $W_0=20$  (black)  $W_0=25$  (skyblue) and  $W_0=30$  (brown).

### 6.3. Discussion

The goal of the present work is to study the conformational fluctuation dynamics of domain-I of HSA in a confined environment and to see the effect of temperature in it. For this purpose TMR labelled HSA was confined inside AOT reverse micelles having variable water content and size. To make sure that the observed



effect of confinement is not coupled with the deformation of structure of HSA upon encapsulation I confirmed that the enfolding of HSA inside reverse micelle does not induce significant alteration in the protein structure through CD spectra. As shown in figure 6.1a, the signature of CD spectra of HSA in buffer and in reverse micelle is similar. The presence of excess noise in the second case might be due to the light scattering by the reverse micelles. The similar percentage of  $\alpha$ -helical content allows us to assume that the confinement effect will be similar or even less for larger water-pools. Thus we confirm that any change in the conformational dynamics of HSA inside the water-pool would be solely due to the confinement effect.

The calculated value of hydrodynamic radius of HSA matches well with the previous reports. The 26.9  $\mu$ s conformational fluctuation time component in the native state of HSA is similar to our previously reported data. This conformational fluctuation component arises from the time dependent quenching of TMR fluorescence by some amino acid residues due to protein fluctuation. In case of HSA inside reverse micelle the most efficient fitting model consists of a Gaussian distribution in the conformational fluctuation time along with a Gaussian distribution in the diffusion time. As already reported by Pal *et al.*, AOT reverse micelles are not monodisperse in nature.<sup>28</sup> They follow a definite size distribution which varies as  $W_0$  changes. The diffusion time, which is actually the diffusion time of the whole reverse micelle here, will definitely not be single valued and likely to follow a distribution. Also when the proteins are incorporated within AOT reverse micelle, the conformational dynamics may depend on the reverse miceller size. As the size follows a distribution, the conformational fluctuation time is also likely to follow a distribution. Thus, goodness of fitting with the combination of distributed diffusion and distributed conformational fluctuation is justified.

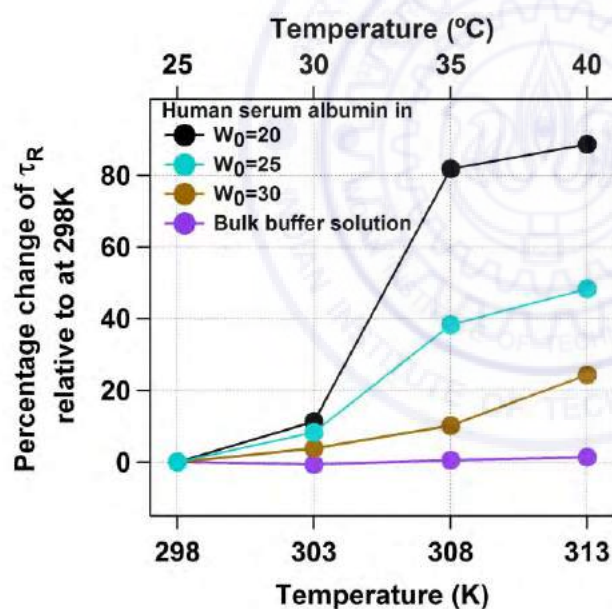
In the control experiment with rhodamine the fluctuation time component is found to be almost negligible and the data could be fitted satisfactorily with distributed diffusion model only. This vividly confirmed the observation of protein fluctuation time in the present experiment. The regular increase in the peak value of  $\tau_D$  indicates the increase in hydrodynamic radii, thus the size of water pool, as this  $\tau_D$  corresponds



to the reverse micelle as a whole, not the individual HSA molecule in it. The value of hydrodynamic radius of the reverse micelles having different  $W_0$  values was calculated from the peak diffusion time. As evident from the figure 6.4b, the present values of hydrodynamic radii are in quite good agreement with the reported values.<sup>28</sup> This in turn suggest that the well defined water-pool of AOT reverse micelle does not get affected by the incorporation of HSA.

Upon being confined in reverse micelle having  $W_0=20$ , the conformational fluctuation time of domain I of HSA decreases drastically from 26.9  $\mu\text{s}$  to 4.3  $\mu\text{s}$ . The reason of this difference is attributed to the change in environment upon confinement of the protein. The protein inside the water-pool is under constrained condition and hence the movement of the side chains are restricted as compared to that in bulk, where there is no restriction on the breathing motion for the protein. Because of this

confinement, the quencher molecules could not diffuse far away from the fluorescent dye due to the lack of available space and hence travels less distance for the breathing motion. This results in a low value of  $\tau_R$ . With increasing water pool size a regular increment of peak  $\tau_R$  (figure 6.5a and 6.5b) is observed. This is at per with our expectation and is ascribed to the increment in water pool size of the reverse micelle, which gives the protein side chain more free



**Figure 6.8:** Percentage change in conformational fluctuation time ( $\tau_R$ ) of domain-I of HSA with temperature for bulk buffer (violet), inside water-pool of AOT reverse micelle of  $W_0=20$  (black)  $W_0=25$  (skyblue) and  $W_0=30$  (brown).

space for movement. The value of  $\tau_R$  approaches towards the bulk value of 26.9  $\mu\text{s}$  but due to experimental constraints we could not reach to that point. This effect of temperature is more prominent at smaller reverse micelle as compared to bigger ones.

However, as mentioned, even for larger water-pools the effect of temperature is more compared to bulk water, where less than a percent increase in the fluctuation time was observed for the present temperature range (figure 6.8). The higher temperature dependence of the conformational fluctuation time of HSA inside the water-pool compared to bulk medium is surprising because inside the water-pool the motion of protein is restricted, and it is expected that the temperature induced fluctuation will be much smaller inside the water-pool compared to the bulk medium. Thus we expected a less change in conformational fluctuation time with increase in temperature in the water-pool than in bulk is expected. The observed opposite trend in the temperature dependence broach dubiousness on the general belief of imposed rigidity in macromolecules upon confinement. This startling behaviour either can stem from the altered nature of the confined protein itself or can arise due to the significantly different behaviour of confined water as compared to bulk water.<sup>31,32</sup> For protein in solution, two different types of water have been proposed. Water molecules that make a stronger hydrogen bond with protein known as “bound water” and the other type water, which make hydrogen bonding among themselves known as “free water”.<sup>33</sup> It is now well established that this “bound water” is responsible for the slow solvation dynamics in protein and helps the protein to retain its native functional structure under small perturbation.<sup>33-35</sup> Once the “bound water” detaches from the protein it becomes “free water” and *vice-versa*. It has been recognized that this bound-free exchange has an intrinsic activation energy barrier.<sup>36</sup> For bovine serum albumin (BSA) in bulk water it is ~20 kcal/mole.<sup>37</sup> This energy is considerably high and at room temperature the same cannot be achieved without any external stimuli. We have already shown that the fluctuation time of domain-I of human serum albumin (HSA), having a very similar structure that of BSA, in buffer responds to thermal agitation only at >323 K.<sup>8</sup> This can be attributed to the availability of the activation energy required for the bound to free exchange. On the other hand, the large amplitude motion or fluctuation in protein may also associated with the bound-free exchange as in this case the micro-environment of the protein changes with time. As already discussed, the nature of water inside the water-pool of a reverse micelle is different from that of bulk water.<sup>31,32</sup> Thus, it is very

likely that the transition of protein bound water to the water-pool confined water will not be the same as in bulk solution. Considering the experimental results, we suggest that the corresponding activation energy inside AOT reverse micelle is less than the bulk water, at least in the experimental range. With increase in temperature, the bound to free water transition is facilitated and the hydrogen bond network is broken; hence the protein can breathe more freely, showing a larger fluctuation time. With the increase in reverse micellar size, the property of the confined water in the water-pool starts approaching to the bulk water and thus the effect of temperature decreases monotonically. The possibility of protein hydration in similar line has also been concluded in the literature.<sup>38-40</sup> Pal and co-workers reported that the protein is more hydrated at elevated temperatures due to the transition of the surface bound water to free water.<sup>39</sup> Lucent *et al.* delineated that when a protein is under confinement, it does not reach the exact denatured structure that of in the bulk, rather it lands up to a different unfolded state.<sup>40</sup> They attributed this irregularity mainly to the confined solvent effect. In my case the observed prominent thermal effect on the conformational dynamics of domain-I of HSA in the water-pool of AOT reverse micelle as compared to bulk medium may arise due to the confined solvent effect, although I can not rule out the effect of geometrical restriction to the protein towards my observation. At this point thus this startling behaviour can not be addressed in a confirmed way. We look forward to further investigation to validate the actual reason for this mystifying behaviour.

#### **6.4. Conclusion**

The present study elucidates the nature of conformational fluctuation dynamics of domain-I of human serum albumin (HSA) when confined inside the water-pool of AOT reverse micelle. Fluorescence correlation spectroscopy was used and the domain-I of HSA was labelled with fluorescent probe tetramethylrhodamine-5-maleimide for this purpose. It is shown that the dynamics varies significantly under confinement compared to in bulk medium. As the size of the water-pool increases, the fluctuation time is found to increase monotonically. The effect of temperature on conformational fluctuation of domain-I of HSA also differs inside the water-pool as

compared to in bulk medium. In bulk, the measured fluctuation time is found to remain same till 313 K, however, inside the water-pool of AOT reverse micelle we have observed an increase in the fluctuation time beyond 303 K. This unexpected behaviour raise question on the general belief of imposed rigidity in macromolecules upon confinement.



## References

1. Karplus, M.; Kuriyan, J. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6679–6685.
2. Petsko, G. A.; Ringe, D. *Protein Structure and Function: Primers in Biology*. New Science Press. Ltd, **2004**.
3. Yang, L. Q.; Sang, P.; Tao, Y.; Fu, Y. X.; Zhang, K. Q.; Xie, X. H.; Liu, S. Q. *J. Biomol. Struct. Dyn.* **2014**, *32*, 372–393.
4. Orozco, M. A. *Chem. Soc. Rev.* **2014**, *43*, 5051–5066.
5. Krishnamoorthy, G. *Curr. Sci.* **2012**, *2*, 266–276.
6. Neuweiler, H.; Banachewicz, W.; Fersht, A. R. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22106–22110.
7. Chattopadhyay, K.; Elson, E. L.; Frieden, C. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2385–2389.
8. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
9. Sasmal, D. K.; Mondal, T.; Mojumdar, S. S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
10. Mojumdar, S. S.; Chowdhury, R.; Chattoraj, S.; Bhattacharyya, K. *J. Phys. Chem. B* **2012**, *116*, 12189–12198.
11. Gruebele, M. *J. Am. Chem. Soc.* **2014**, *136*, 16695–16697.
12. Good, D. B.; Wang, S.; Ward, M. E.; Struppe, J.; Brown, L. S.; Lewandowski, J. R.; Ladizhansky, V. *J. Am. Chem. Soc.* **2014**, *136*, 2833–2842.
13. Yang, H.; Luo, G.; Karnchanaphanurach, P.; Louie, T. M.; Rech, I.; Cova, S.; Xun, L.; Xie, X. S. *Science* **2003**, *302*, 262–266.
14. Langosch, D.; Arkin, I. T. *Science* **2009**, *18*, 1343–1358.
15. Gollapalli, P.; Hanumanthappa, M. *Int. J. Comput. Biol.* **2015**, *4*, 31–45.
16. Li, D.; Liu, M. S.; Ji, B. *Biophys. J.* **2015**, *109*, 647–660.
17. Li, D.; Ji, B.; Hwang, K.; Huang, Y. *J. Phys. Chem. B* **2010**, *114*, 3060–3069.
18. Martinez, A. V.; DeSensi, S. C.; Dominguez, L.; Rivera, E.; Straub, J. E. *J. Chem. Phys.* **2011**, *134*, 055107(1)–055107(9).
19. Nicot, C.; Waks, M. *Biotechnol. Genet. Eng. Rev.* **1995**, *13*, 267–314.



20. Roche, C. J.; Dantsker, D.; Heller, E. R.; Sabat, J. E.; Friedman, J. M. *Chem. Phys.* **2013**, *422*, 88–97.
21. Ulmer, T. S.; Bax, A. *J. Biol. Chem.* **2005**, *280*, 43179–43187.
22. Yeung, P. S. W.; Gözde Eskici, G.; Paul H. Axelsen P. H. *Biochim. Biophys. Acta* **2013**, *1828*, 2314–2318.
23. Murakami, H.; Toyota, Y.; Nishi, T.; Nashima, S. *Chem. Phys. Lett.* **2012**, *519*–*520*, 105–109.
24. Melo, E. P.; Fojan, P.; Cabral, J. M. S.; Petersen, S. B. *Chem. Phys. Lipids* **2000**, *106*, 181–189.
25. Malik, A.; Kundu, J.; Karmakar, S.; Lai, S.; Chowdhury, P. K. *J. Lumin.* **2015**, *167*, 316–326.
26. Martinez, A. V.; Dominguez, L.; Małolepsza, E.; Moser, A.; Ziegler, Z.; Straub, J. E. *J. Phys. Chem. B* **2013**, *117*, 7345–7351.
27. Dutt, G. B. *J. Phys. Chem. B* **2008**, *112*, 7220–7226.
28. Pal, N.; Verma, S. D.; Singh, M. K.; Sen, S. *Anal. Chem.* **2011**, *83*, 7736–7744.
29. Maitra, A. *J. Phys. Chem.* **1984**, *88*, 5122–5125.
30. Mitra, R. K.; Sinha, S. S.; Pal, S. K. *Langmuir* **2008**, *24*, 49–56.
31. Levinger, N. E. *Science* **2002**, *298*, 1722–1723.
32. Spehr, T. L.; Frick, B.; Zamponi, M.; Stühn, B. *Soft Matter* **2011**, *7*, 5745–5755.
33. Nandi, N.; Bhattacharyya, K.; Bagchi, B. *Chem. Rev.* **2000**, *100*, 2013–2045.
34. Pal, S. K.; Peon, J.; Zewail, A. H. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1763–1768.
35. Pal, S. K.; Zewail, A. H. *Chem. Rev.* **2004**, *104*, 2099–2123.
36. Han, K. L.; Zhao, G. J. *Hydrogen Bonding and Transfer in the Excited State* John Wiley & Sons. **2011**.
37. Bokor, M.; Csizmók, V.; Kovács, D.; Bánki, P.; Friedrich, P.; Tompa, P.; Tompa, K. *Biophys. J.* **2005**, *88*, 2030–2037.
38. Mukherjee, S.; Chowdhury, P.; Gai, F. *J. Phys. Chem. B* **2006**, *110*, 11615–11619.
39. Rakshit, S.; Saha, R.; Pal, S. K. *J. Phys. Chem. B* **2013**, *117*, 11565–11574.

40. Lucent, D.; Vishal, V.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10430–10434.





## *Chapter 7*

---

### *Directionality of Peptidase Activity on Human Serum Albumin Revealed by Fluorescence Correlation Spectroscopy*

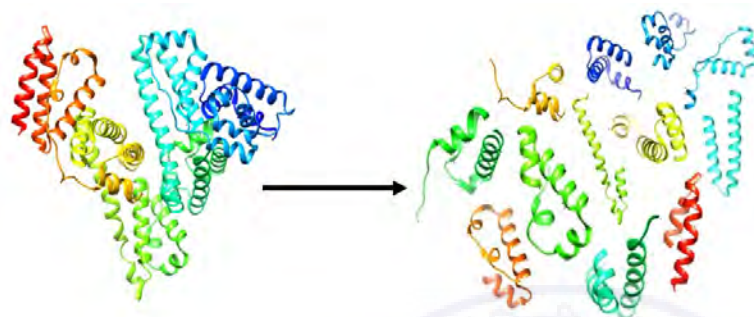
Sengupta, B.; Das, N.; Sen, P. *Manuscript under preparation*

*Enzymatic proteolysis or protein digestion is the fragmentation of protein into smaller peptide units under the action of peptidase enzymes. In this work, the enzymatic digestion of human serum albumin (HSA) was studied using fluorescence correlation spectroscopy. Papain, chymotrypsin and trypsin were used as the enzymes for this purpose. HSA was tagged with TMR to Cys-34 located in domain I and p-nitrophenylcoumarin ester to Tyr-411 located in domain III separately and subjected to the enzymes maintaining 1:1 molar ratio at 323 K. The change in hydrodynamic radius as a function of digestion time was monitored. Following the change in hydrodynamic radius, as monitored by FCS, it has been confirmed that under similar experimental conditions, the order of efficiency of digestion is papain > chymotrypsin > trypsin. More interestingly, a faster decrease of hydrodynamic radius was observed when the fluorescence from domain-I was monitored in FCS, compared to that of domain-III. This observation clearly indicates that all these enzymes prefer to cleave HSA from the N-terminal side (domain-I), while the possible cleavage sites for all these three enzymes are distributed over the entire span of amino acid chain in HSA. Such directionality of enzymatic digestion is observed for the very first time and it is quite interesting.*



### 7.1. Introduction

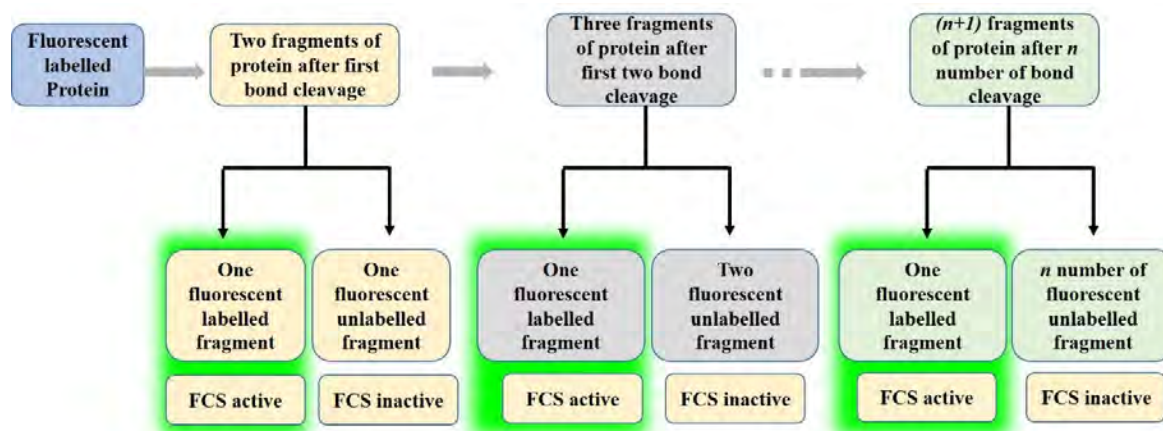
Protein digestion or proteolysis is the breakdown of protein to small peptides (scheme 7.1) or to its constituent amino acids and when this process is catalysed by enzymes it is known as enzymatic digestion of proteins.<sup>1-5</sup> Apart from *in vivo* health benefits related to physical fitness, protein digestion is a necessary step in the protein



**Scheme 7.1:** Schematic representation of digestion of protein.

identification and characterization protocol. The enzymes that are responsible for proteolysis are generally termed as peptidase and in human body they are mostly secreted from

salivary glands, secretory cells in stomach, pancreas and small intestine. It is well known that in the folded state or native state of the protein, the hydrophilic amino acid residues are located at the surface of the protein maximizing the stabilizing hydrophilic interaction with the water molecules, whereas the hydrophobic ones are buried deep inside the core minimizing destabilizing hydrophobic interactions with the water molecules.<sup>1,6-9</sup> Hence, in a general sense, hydrophilic residues are always more prone to proteolysis with respect to accessibility. But among several peptide linkages present in a peptide bond, which one will be cleaved first is decided not only by the accessibility of the peptide bond but also by the selectivity of the peptidase. Peptidases are broadly distributed into two classes: exopeptidase and endopeptidase.<sup>10-12</sup> Exopeptidases are more specific to the proximity of locations of the scissile peptide linkage to the terminal residue whereas the endopeptidases choose a target peptide bond depending on the constituent amino acids participating in the peptide bond. Till date, the most commonly practised methods to study the enzymatic digestion of proteins are mass spectroscopy, gel electrophoresis, size exclusion chromatography (SEC), high performance liquid chromatography (HPLC) and reverse phase high performance liquid chromatography (reverse phase-HPLC).<sup>2,13-16</sup> However, a moderate



**Scheme 7.2:** Protocol of monitoring digestion by FCS.

concentration of protein is necessary to perform all these experimental techniques which practically is a major drawback. In the current study, we have used fluorescence correlation spectroscopy (FCS), a single molecular fluorescence technique to study digestion of human serum albumin (HSA). Being a single molecular experimental technique, FCS operates at a very low concentration such as nanomolar or picomolar.<sup>17</sup> Way back in 1967, the initial fragmentation in the tryptic digestion of human serum albumin was studied by Markus *et al.* using gel filtration and electrophoresis techniques.<sup>18</sup> Ten years later, similar study on the peptic digestion was carried out by Kieras *et al.*<sup>19</sup> Griffiths *et al.* studied the digestion of human serum proteins by molecular weight dependent acrylamide gel electrophoresis.<sup>20</sup> Recently, tryptic digestion of HSA was utilized to create a mass spectroscopic library for peptides.<sup>21</sup> As being a large protein, the proper understanding of fragmentation for HSA is difficult. Herein, a protocol is proposed to monitor the fragmentation of HSA using FCS and consequently the applicability of the protocol is shown. Here it is to be noted that FCS relies on fluorescence fluctuation originating mainly from the diffusive motion of the particle in and out of a small observation volume.<sup>22</sup> As only observable in FCS is fluorescence, in case of site specific fluorescence labelled proteins, the fragment that contains the fluorescent tag will be monitored. Prior to digestion, the diffusion of the whole protein will be responsible for fluorescence fluctuation and as the reaction proceeds, the size of the protein decreases which can be efficiently detected by FCS using the fluorescence fluctuation arising from the diffusion of the fragment containing

the fluorescence tag through the measured diffusion coefficient (scheme 7.2). This can be used to decipher the kinetics of digestion. In addition, the directionality of the digestion can also be investigated by introducing fluorescence labels at the two ends of the protein chain. In this work, the domain I of 585 residue long amino acid sequence of HSA has been tagged at Cys-34 residue by thiol specific rhodamine based dye tetramethylrhodamine-5-maleimide (TMR) for single molecular experiments whereas Domain III of HSA has been tagged at Tyr-411 with a hydroxyl specific coumarin based dye *p*-nitrophenylcoumarin ester (NPCE) respectively.<sup>23,24</sup> The details of tagging process has been described in section 2.4.1 and 2.4.4 of chapter 2 for NPCE and TMR respectively. Papain, chymotrypsin and trypsin has been used as enzymes in the current study. Papain is a cysteine plant protease with 212 amino acid residues found in papaya latex.<sup>25-29</sup> The active site of papain consists of Cys-25, His-159 and Asn-175.<sup>28</sup> Rice *et al.* studied the digestion of native, denatured and stabilized HSA in 1945 and concluded that upon denaturation of HSA the rate of digestion increases significantly which is more prominent in case of urea induced denaturation as compared to thermal denaturation.<sup>29</sup> Due to the enzymatic capability of papain, papaya latex is exclusively used in household cooking of high protein substances. In human body, 11 papain like cysteine proteases are found. Chymotrypsin and trypsin both are found in pancreatic juice.<sup>30</sup> The catalytic mechanisms of chymotrypsin and trypsin are also similar although the substrate specificity is different. Chymotrypsin contains 245 amino acid residues among which Ser-195, His-57 and Asp-102 forms the active catalytic triad. Chymotrypsin favors to cleave at peptide bonds containing aromatic amino acid residues such as tryptophan, phenylalanine and tyrosine.<sup>31-36</sup> Trypsin is a globular protein which specifically cleaves at the C-terminal side of arginine and lysine residues. The active site of trypsin contains Asp-102, His-57, and Ser-195.<sup>37-39</sup>

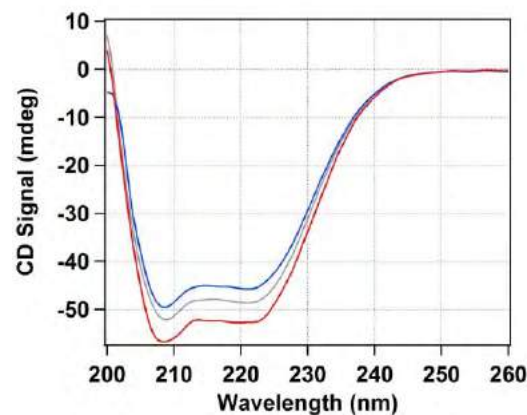
## 7.2. Results

### 7.2.1. Circular Dichroism Spectroscopy:

That the secondary structure of HSA does not get perturbed upon tagging with TMR and NPCE, has been checked with CD spectroscopy. As can be seen in figure 7.1 the CD spectra of untagged native HSA, TMR tagged HSA and NPCE tagged HSA are very similar.

From the intensity at 208 nm the  $\alpha$ -

helicity are calculated to be 63%, 60% and 65% for the native, TMR tagged and NPCE tagged HSA respectively.

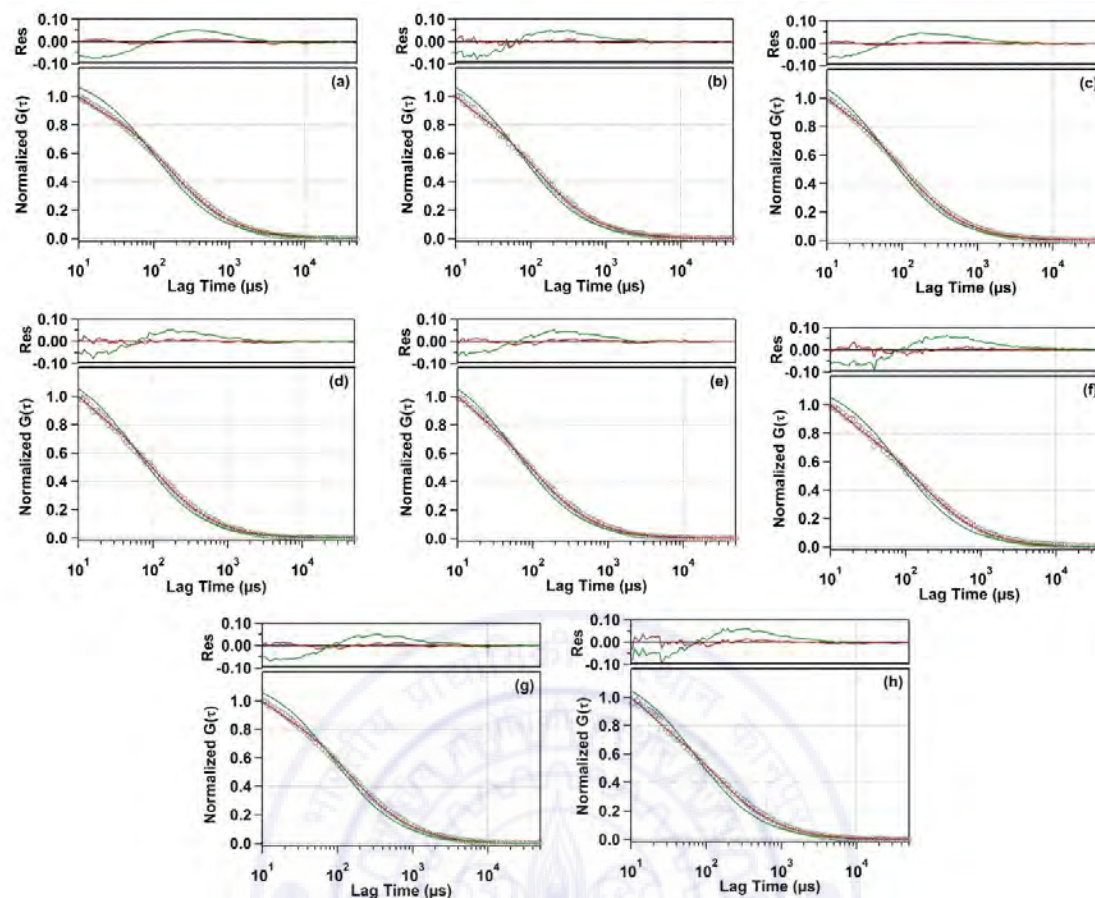


**Figure 7.1:** CD spectra of native HSA (ash, 1.98  $\mu$ M), TMR tagged HSA (red, 2.18  $\mu$ M), NPCE tagged HSA (blue, 1.84  $\mu$ M) in phosphate buffer (pH 7.4, 50 mM).

**7.2.2. Fluorescence Correlation Spectroscopy:** The fluorescence autocorrelation curves were recorded separately for TMR tagged and NPCE tagged HSA as a function of digestion times for all the three enzymes chymotrypsin, trypsin and papain. None of the autocorrelation curves were satisfactorily fitted with the pure diffusion equation (equation 2.27) and an additional time component ( $\tau_R$ ) with an amplitude  $a$  was introduced in the fitting equation (equation 2.31) to fit the fluorescence autocorrelation data. A comparison of fitting for the native state HSA tagged in domain I and domain III has been shown in figure 7.2a and 7.2b respectively (domain I is tagged with TMR and domain III is tagged with NPCE). The hydrodynamic radius of the native state of HSA is calculated to be  $38.7 \pm 0.2$  and  $38.8 \pm 0.3$  Å, for the above two cases respectively. The very similar value obtained in both the cases validates our measurements. After addition of enzymes, the diffusion time decreases for all the cases. Figure 7.2c-h shows the fitting of autocorrelation curves at 30 seconds after the starting the enzymatic reaction with all three enzymes for two different domains. In all these cases the fitting residuals show significant improvement of fitting quality while the additional fluctuation term is incorporated. From now onwards, for ease of writing, we will refer to the different cases as follows

Domain I labelled with TMR and digested by papain: HPD1





**Figure 7.2:** Comparison of fit using only diffusion (green line) and diffusion + relaxation (red line) for (a) TMR tagged HSA before starting the digestion reaction, (b) NPCE tagged HSA before starting the digestion reaction, (c) HPD1 after 30 seconds from the starting of digestion reaction, (d) HPD3 after 30 seconds from the starting of digestion reaction, (e) HCD1 after 30 seconds from the starting of digestion reaction, (f) HCD3 after 30 seconds from the starting of digestion reaction, (g) HTD1 after 30 seconds from the starting of digestion reaction, (h) HTD3 after 30 seconds from the starting of digestion reaction.

Domain I labelled with TMR and digested by chymotrypsin: HCD1

Domain I labelled with TMR and digested by trypsin: HTD1

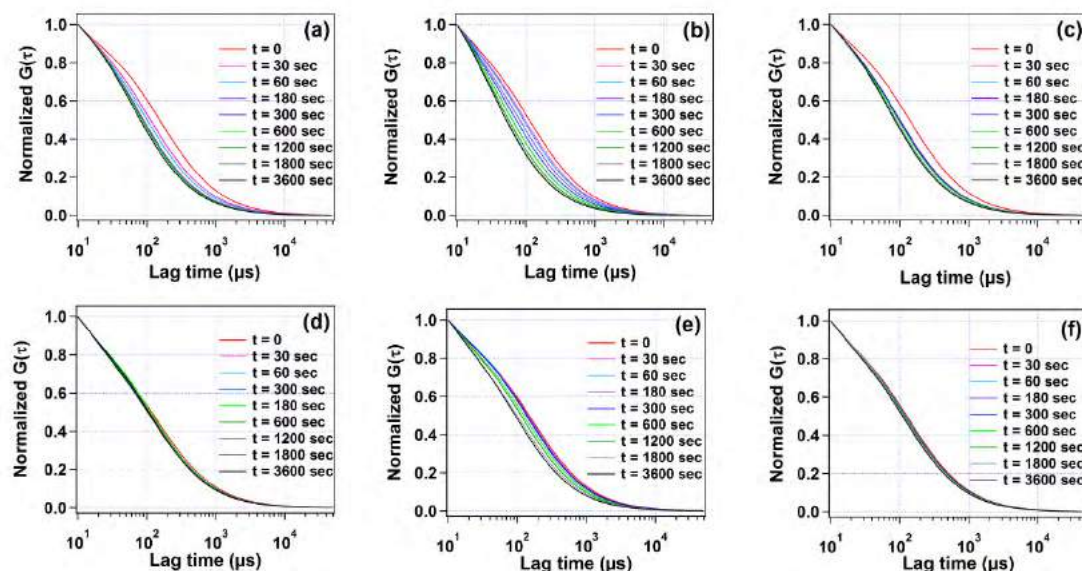
Domain III labelled with NPCE and digested by papain: HPD3

Domain III labelled with NPCE and digested by chymotrypsin: HCD3

Domain III labelled with NPCE and digested by trypsin: HTD3

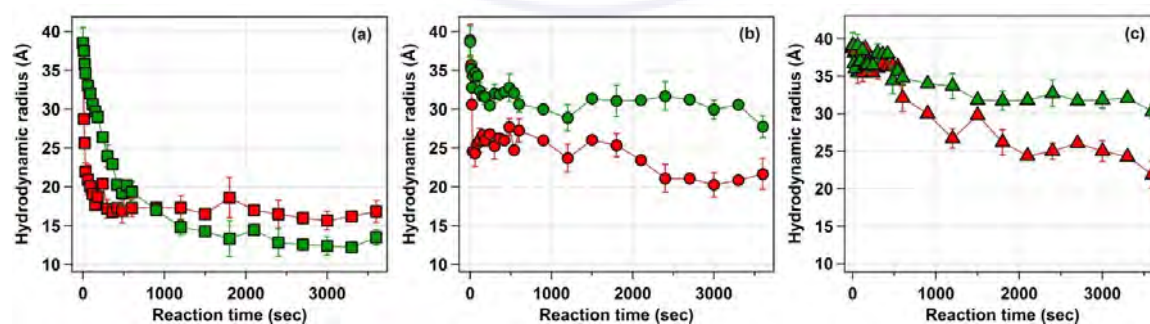
A few normalized best fitted autocorrelation curves at different times from the start of digestion for HPD1, HPD3, HCD1, HCD3, HTD1, HCD3 are shown in figure 7.3a-f respectively. The measurements clearly indicates that decay of the autocorrelation becomes faster with the progress of digestion for all the cases. This decrease was then readily connected with the decrease of the hydrodynamic radius of the protein





**Figure 7.3:** Normalized autocorrelation curves at different digestion times for (a) HPD1, (b) HPD3, (c) HCD1, (d) HCD3, (e) HTD1, (f) HTD3.

containing fluorescence tag. The decrease of hydrodynamic radius with digestion time for both the domain in presence of papain, chymotrypsin and trypsin are plotted in figure 7.4a-c respectively. Evidently, the decrease of the hydrodynamic radius for two different domains are different for a given enzyme. This difference is prominent in case of papain and chymotrypsin from early times whereas for trypsin this difference is observed only at longer times. For HPD3 the value of hydrodynamic radius is found to be least (13.5 Å) after one hour, whereas for HTD3 it is the largest (32.1 Å). The initial rate of digestion is also found to depend upon the peptidase used in the study. The



**Figure 7.4:** Decrease of hydrodynamic radius for domain I (red) and domain III (green) as a function against digestion time for (a) papain (square box), (b) chymotrypsin (solid circles) and (c) trypsin (triangles) induced denaturation.

hydrodynamic radius of the TMR tagged fragment after 10 minutes of the digestion becomes 17.3 Å, 27.3 Å and 32.1 Å for papain, chymotrypsin and trypsin respectively.

On the other hand for NPCE tagged fragment the radii is found to be 18.1 Å, 30.6 Å and 34.7 Å after 10 minutes of digestion for papain, chymotrypsin and trypsin, respectively. For HPD1, the measured hydrodynamic radius for the fragment having the fluorescence tag TMR is found to reach a minimum of 16.8 Å after 6 minutes of digestion and remains unchanged further, whereas for HPD3, the minimum value of 13.3 is reached at 30 minutes and then remains unchanged. Likewise, for HCD1 and HCD3, the minimum values of hydrodynamic radius are 21.1 Å and 30.6 Å which are reached at 40 minutes and 10 minutes respectively. For HTD1 and HTD3 the values of hydrodynamic radii become minimum (24.4 Å and 31.8 Å respectively) at 35 minutes and 25 minutes respectively.

### 7.3. Discussion

The similar  $\alpha$ -helicity of native, TMR tagged and NPCE tagged HSA indicates that the tagging procedure does not perturb the secondary structure of HSA and hence it can be assumed that in all the FCS experiments the starting point is the native HSA. The failure of the single diffusion fitting equation to fit the autocorrelation curves implies that there is an additional fluctuation component apart from that caused by the diffusion. In most of the cases, this kind of fluctuation generates due to the conformational fluctuation of the protein.<sup>20-43</sup> As already described in chapter 3 and chapter 6, HSA shows conformational fluctuation with a timescale of 27.0 microseconds and 8.0 microseconds for domain I and domain III respectively. In the present case we have also got similar timescales for native HSA but in presence of the enzymes this timescale changes significantly. This alteration can be attributed to the interaction of the enzymes with HSA. To avoid complication, the discussion will be focused only on the diffusion timescale as that will provide us the crucial parameters that we are interested in. To begin with, it is necessary to illustrate the principle behind this idea of detection of directionality of the digestion. A long chain protein (in this case HSA) is site specifically tagged somewhere near the two terminals separately and the diffusion profile is monitored against digestion time. As already mentioned in section 7.1, all the untagged fragments will be silent in FCS experiments and only the tagged fragment will be visible. So, initially, the diffusion of the whole protein will be

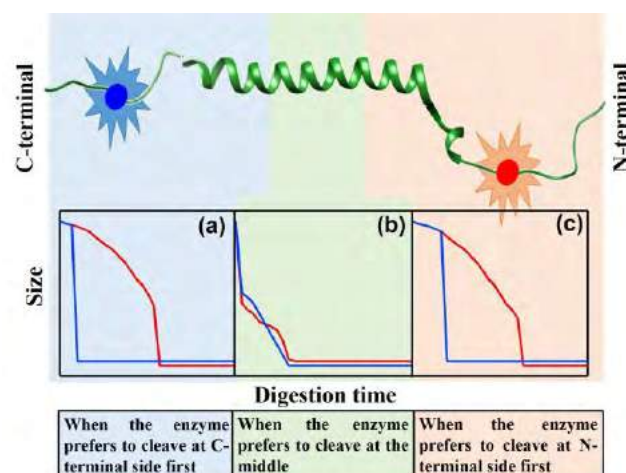
responsible for the fluorescence fluctuation and with progress of digestion, once the fluorescence tag containing fragment is separated from the protein via peptide bond cleavage, we will only observe the diffusion of the fluorescent fragment. Hence, the gradual decrease of the diffusion time will be a measure of how the tagged fragment size changes with the progress of the digestion. The relative change in diffusion time for this two different fluorescence tags at the two ends of the protein will be governed by the site specificity of the enzyme. There could be a maximum of three possibilities viz. (I) the enzyme prefers to cleave at the domain I, (II) the enzyme prefers operating from domain III and (III) the enzyme does not have any specificity regarding the domains. Considering the change in hydrodynamic radii of the protein (tagged at domain I and domain III) as the progress of the digestion, the abovementioned cases can be summarized as follows,

**Case I:** The enzyme prefers to cleave the peptide bonds located in domain I. In this case, the protein tagged in domain I should demonstrate a faster decrease of hydrodynamic radius as compared to the protein tagged at domain III.

**Case II:** This scenario is the exact opposite of case I where the enzyme prefers peptide bonds of domain III and one expects to observe a faster decay of hydrodynamic radius when the fluorescence tag is attached in domain III as compared to that of domain I.

**Case III:** The cleavage site is non-preferential. Expectedly, in this case one may observe no significant difference in the variation of hydrodynamic radius for the domain I and domain III tagged protein.



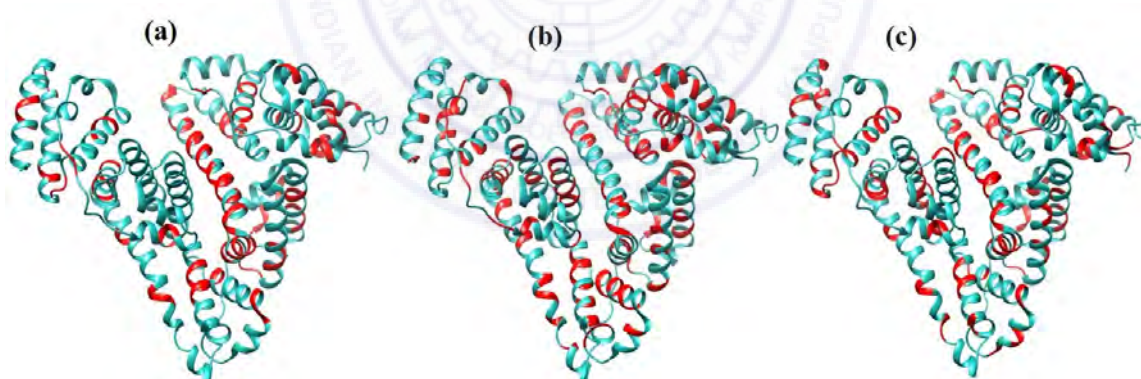


**Scheme 7.3:** Protocol of monitoring directionality with the help of FCS.

Scheme 7.3 depicts the probable pattern of variation of hydrodynamic radius with time for these three different cases.

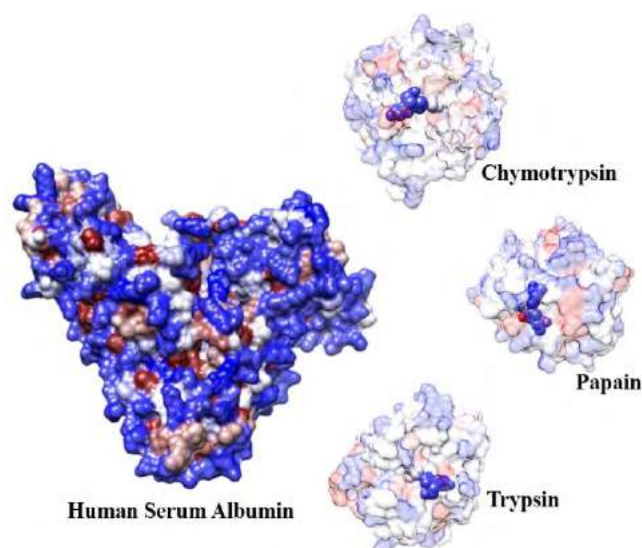
In the present work, the faster decrease of hydrodynamic radius for HPD1, HCD1 and HTD1 as compared to HPD3, HCD3 and HTD3 respectively indicates that papain, chymotrypsin and trypsin prefer to cleave at domain I side as

compared to domain III. However, the preference of cleaving sites of papain, chymotrypsin and trypsin are present throughout the protein chain and they are highlighted in red in scheme 7.4 (PDB ID: 1HA2). For papain there are 21 possible targets in both domain I and domain II whereas domain III contains 15 cleaving site which is defined as [21, 21, 15].<sup>25-29</sup> For chymotrypsin and trypsin the numbers are [39,



**Scheme 7.4:** Peptide linkages more prone to cleavage (marked in red) under enzymatic action of (a) papain, (b) chymotrypsin and (c) trypsin. (PDB ID of HSA: 1HA2)

38, 32] and [24, 24, 31] respectively for the three domains.<sup>30-39</sup> The fact that the preferred peptide bonds are spread over the whole region of a protein makes our observation even more interesting. The directionality that we observed for all three enzymes clearly indicate that even the number of preferred peptide bonds are similar in domain I and domain III, there is some other factor which makes the enzymes to cleave at domain I preferentially. One common reason can be the presence of a single



**Scheme 7.5:** Hydrophilic nature of HSA and the active sites of papain, chymotrypsin and trypsin. (PDB ID of HSA: 1HA2; PDB ID of papain: 9PAP; PDB ID of chymotrypsin: 1YPH; PDB ID of trypsin: 1F0T)

free cysteine residue in domain I which may initiate the HSA-enzyme interaction, but in the present case, Cys-34 is tagged with TMR and the free thiol group is unavailable. So any involvement of Cys-34 in this preferential behaviour is eliminated. It is well known that the ligand binding property of these two domains of HSA are different.<sup>44-46</sup> Domain III has a hydrophobic pocket which makes

it the most suitable binding location in the whole protein structure. Domain I, on the other hand, lacks such a binding pocket due to the effective packing of a couple of helix in this region. Domain I mainly binds fatty acids and metal containing complexes. Domain III of HSA has a positive charge of +10 whereas domain I has negative charge of -10.<sup>47</sup> The first step of the general mechanism of proteolysis by cysteine or serine proteases is the deprotonation of thiol group followed by attack on the carbonyl carbon of the target peptide bond.<sup>48</sup> The electrophilicity of the carbonyl carbon may be a determining factor in selection of the target peptide bond but that factor has already been incorporated while assigning the number of susceptible bonds in each domain. Next, I consider the initial interaction of domain I and domain III of HSA with the enzymes, which can dictate the attachment of the enzyme with either domain of HSA. This can be hydrophilic/ hydrophobic interaction which allows the enzyme to interact preferentially at any part of the protein depending on the nature of the enzyme active site. To visualize that, I have plotted the surface of HSA and mark the hydrophilic regions as blue, hydrophobic regions as red and the neutral regions as white (scheme 7.5). Scheme 7.5 clearly shows that the hydrophilicity of domain I surface is a little more as compared to domain III. Likewise, when the nature of active site residues for the three enzymes under investigation are considered, it is seen that the involved



residues are mostly hydrophilic. This makes the interaction of the enzyme active site with domain I of HSA more favourable than that of domain III and has been assigned as the possible reason of enzymatic attack on domain I preferentially over domain III.

#### 7.4. Conclusion

The digestion of HSA has been investigated as a function of time in the presence of three enzymes papain, chymotrypsin and trypsin. To the best of my knowledge this is the first time FCS has been utilized to monitor digestion of HSA. It has been noticed that papain digests HSA most efficiently followed by chymotrypsin and trypsin. The initial rate of digestion is visibly more for domain I case as compared to domain III for all the three enzymes. The hydrodynamic radius of HSA decreases by 56.9% in just 6 minutes for HPD1. Although the maximum decrease is 65.9% for HPD3, which is greater than HPD1 but this value is reached only after 30 minutes from the start of the reaction i.e. the fragmentation of this part is slower as compared to that of domain I. Similarly for chymotrypsin, 45.9% and 21.5% decrease of hydrodynamic radius is observed for domain I and domain III respectively. For trypsin the decrease is 37.5% and 18.5% for domain I and domain III respectively. The reason for faster digestion of domain I over domain III by all these three enzymes can be accounted for the favourable hydrophilic interaction of domain I with the active site of the enzymes. However at this stage, with the current set of results, the exact reason of the enzymatic behaviour can not be specified with confidence and supplementary experiments are necessary before drawing final conclusion. Besides, digestion of denatured substrate and temperature dependence of the digestion reaction may prove to be useful in further understanding.

## References

1. Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry*, 5th Ed. W.H. Freeman Co. NY, **2002**.
2. Vahdatpour, S.; Mamaghani, A. P.; Goloujeh, M. S.; Sis, N. M.; Mahmoodpour, H.; Vahdatpour, T. *Electr. J. Biol.* **2016**, *12*, 265–275.
3. Freeman, H. J.; Kim, Y. S. *Annu. Rev. Med.* **1978**, *29*, 99–116.
4. Spellman, D.; McEvoy, E.; O’Cuinn, G.; FitzGerald, R. J. *Int. Dairy J.* **2003**, *13*, 447–453.
5. Holmes, J. H. G.; Bayley, H. S.; Leadbeater, P. A.; Horney, F. D. *Br. J. Nutr.* **2007**, *32*, 479–489.
6. Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105–132.
7. Berger, B.; Leighton, T. *J. Comput. Biol.* **1998**, *5*, 27–40.
8. Shea, J. E.; Brooks Iii, C. L. *Ann. Rev. Phys. Chem.* **2001**, *52*, 499–535.
9. Joshi, R. R. *J. Biomol. Struct. Dyn.* **2013**, *31*, 965–966.
10. Rawlings, N. D.; Barrett, A. J. *Biochemical J.* **1993**, *290*, 205–218.
11. Rawlings, N. D.; Barrett, A. J.; Bateman, A. *J. Biol. Chem.* **2011**, *286* (44), 38321–38328.
12. Rawlings, N. D.; Barrett, A. J. *Peptidases in eLS*. John Wiley & Sons Ltd, Chichester, **2014**.
13. Switzar, L.; Giera, M.; Niessen, W. M. A. *J. Proteome Res.* **2013**, *12*, 1067–1077.
14. Williams, K. R.; Stone, K. L. *Mol. Biotech.* **1997**, *8*, 155–167.
15. Lapolla, A.; Fedele, D.; Reitano, R.; Aricò, N. C.; Seraglia, R.; Traldi, P.; Marotta, E.; Tonani, R. *J. Am. Soc. Mass Spect.* **2004**, *15*, 496–509.
16. Peters, T. Jr.; Feldhoff, R. C. *Biochemistry* **1975**, *14*, 3384–3391.
17. Schwille, P. *In Fluorescence correlation spectroscopy. Theory and applications*, Elson, E.L.; Rigler, R. Eds., Springer, Berlin, **2001**.
18. Markus, G.; McClintock, D. K.; Castellani, B. A. *J. Biol. Chem.* **1967**, *242*, 4395–4401.
19. Kieras, J. H.; King, T. P. *J. Biol. Chem.* **1977**, *252*, 4326–4329.

20. Griffiths, W. C.; Bogaars, H. A.; Sheehan, C. F.; Diamond, B. S. *Ann. Clin. Lab. Sci.* **1976**, *6*, 177–183.
21. Dong, Q.; Yan, X.; Kilpatrick, L. E.; Liang, Y.; Mirokhin, Y. A.; Roth, J. S.; Rudnick, P. A.; Stein, S. E. *Mol. Cell. Proteomics* **2014**, *13*, 2435–2449.
22. Lakowicz, J. R. Principles of fluorescence spectroscopy. Springer, third ed., **2006**.
23. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
24. Sengupta, B.; Yadav, R.; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, *18*, 14350–14358.
25. Amri, E.; Mamboya, F. *Am. J. Biochem. Biotechnol.* **2012**, *8*, 99–104.
26. Mitchel, R. E. J.; Chaiken, I. M.; Smith, E. L. *J. Biol. Chem.* **1970**, *245*, 3485–3492.
27. Berger, A.; Schechter, I. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* **1970**, *257*, 249–264.
28. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1968**, *32*, 898–902.
29. Rice, R. G.; Ballou, G. A.; Boyer, P. D.; Luck, J. M.; Lum, F. G. *J. Biol. Chem.* **1945**, *158*, 609–617.
30. Hedstrom, L.; Perona, J. J.; Rutter, W. J. *Biochemistry* **1994**, *33*, 8757–8763.
31. Schoellmann, G.; Shaw, E. *Biochemistry* **1963**, *2*, 252–255.
32. Steitz, T. A.; Hendekson, R.; Blow, D. M. *J. Mol. Biol.* **1969**, *46*, 337–348.
33. Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337–340.
34. Perona, J. J.; Craik, C. S. *J. Biol. Chem.* **1997**, *272*, 29987–29990.
35. Perona, J. J.; Craik, C. S. *Protein Science* **1995**, *4*, 337–360.
36. Zimmerman, M.; Ashe, B. M. *Biochim. Biophys. Acta – Enzymol.* **1977**, *480*, 241–245.
37. Graf, L.; Craik, C. S.; Patthy, A.; Rocznik, S.; Fletterick, R. J.; Rutter, W. J. *Biochemistry* **1987**, *26*, 2616–2623.
38. Evnin, L. B.; Vásquez, J. R.; Craik, C. S. *Proc. Nat. Acad. Sci. USA* **1990**, *87* (17), 6659–6663.
39. Hedstrom, L. *Biol Chem* **1996**, *377*, 465–470.
40. Yadav, R.; Sengupta, B.; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.

41. Sasmal, D. K.; Mondal, T.; Mojumdar, S. S.; Choudhury, A.; Banerjee, R.; Bhattacharyya, K. *J. Phys. Chem. B* **2011**, *115*, 13075–13083.
42. Chattopadhyay, K.; Saffarian, S.; Elson, E. L.; Frieden, C. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14171–14176.
43. Chattopadhyay, K.; Elson, E. L.; Frieden, C. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2385–2389.
44. Kamal, J. K.; Behere, D. V. *Indian J. Biochem. Biophys.* **2005**, *42*, 7–12.
45. Dockal, M.; Carter, D. C.; Rüker, F. *J. Biol. Chem.* **1999**, *274*, 29303–29310.
46. Kragh-Hansen, U. *Pharmacolog. Rev.* **1981**, *33*, 17–53.
47. Ramezani, F.; Rafii-Tabar, H. *Mol. BioSys.* **2015**, *11*, 454–462.
48. Erez, E.; Fass, D.; Bibi, E. *Nature* **2009**, *459*, 371–378.



## List of Publications

---

1. Conformational Fluctuation Dynamics of Domain I of Human Serum Albumin in the Course of Chemically and Thermally Induced Unfolding Using Fluorescence Correlation Spectroscopy. Yadav, R.; **Sengupta, B.**; Sen, P. *J. Phys. Chem. B* **2014**, *118*, 5428–5438.
2. Femtosecond Dynamics of Photoinduced *cis-trans* Isomerization of Ethyl-3-(1H-indole-3-yl) acrylate. **Sengupta, B.**; Mukherjee, P.; Das, S.; Rafiq, S.; Gupta, S.; Dethe, D. H.; Sen, P. *Chem. Phys. Lett.* **2015**, *638*, 31–37.
- \*3. Startling Temperature Effect on Proteins when Confined: Single Molecular Level Behaviour of Human Serum Albumin in Reverse Micelle. **Sengupta, B.**; Yadav, R.; Sen, P. *Phys. Chem. Chem. Phys.*, **2016**, *18*, 14350–14358.
4. Effect of Sucrose on Chemically and Thermally Induced Unfolding Of Domain-I of Human Serum Albumin: Solvation Dynamics and Fluorescence Anisotropy Decay Kinetics Study. Yadav, R.; **Sengupta, B.**; Sen, P. *Biophys. Chem.* **2016**, *211*, 59–69.
- \*5. Elucidation of Local Dynamics of Domain-III of Human Serum Albumin over ps- $\mu$ s Time Regime using a New Fluorescent Label. **Sengupta, B.**; Acharyya, A.; Sen, P. *Phys. Chem. Chem. Phys.* **2016**, *18*, 28548–28555.
- \*6. Elucidation of  $\mu$ s Dynamics of Domain-III of Human Serum Albumin during the Chemical and Thermal Unfolding: A Fluorescence Correlation Spectroscopic Investigation. **Sengupta, B.**; Das, N.; Sen, P. *Biophys. Chem.* **2016**, *221*, 17–25.
- \*7. Single Molecular Level Investigation of Structure and Dynamics of Papain under Denaturation. **Sengupta, B.**; Chaudhury, A.; Das, N.; Sen, P. *Protein Pept. Lett.* **2017** DOI: 10.2174/0929866524666170811145838.
8. Detail Modes of Binding Assessed by Bulk and Single Molecular Level Fluorescence, MD Simulation and its Temperature Dependence: Coumarin 152 with



Human Serum Albumin Revisited. Yadav, R.; **Sengupta, B.**; Sen, P. *Directions*, **2017** in Press.

9. Dynamical Response in Methanol-Chloroform Binary Solvent Mixture over fs- $\mu$ s Time Regime. Gupta, S.; Mukherjee, P.; **Sengupta, B.**; Sen, P. *Phys. Chem. Liq.* **2017**, DOI: 10.1080/00319104.2017.1346649.

\*10. Monomerization and Aggregation of  $\beta$ -Lactoglobulin under Adverse Condition: A Fluorescence Correlation Spectroscopic Investigation. **Sengupta, B.**; Das, N.; Sen, P. *BBA- Proteins Proteomics* **2018**, 1866, 316–326

\*11. Directionality in Digestion of Human Serum Albumin by Peptidases Revealed by Fluorescence Correlation Spectroscopy. **Sengupta, B.**; Das, N.; Sen, P. (Manuscript under preparation).

12. Mapping the Complete Pathway of Domain Specific Double Denaturation of Human Serum Albumin: A Spectroscopic Study. **Sengupta, B.**; Mohan, V.; Acharyya, A.; Yadav, R.; Sen, P. (Manuscript under preparation).

13. Monitoring the Effect of Encapsulation within Cationic and Anionic Reverse Micelle on the Active Site Dynamics of Papain. Mohan, V.; **Sengupta, B.**; Sen, P. (Manuscript under preparation).

14. Domain Specific Stabilization of Human Serum Albumin by Sucrose. Mohan, V.; **Sengupta, B.**; Sen, P. (Manuscript under preparation).

15. Albumin Based Assay and Formulation for Targeting Huntingtin Aggregation. Singh, V.; **Sengupta, B.**; Sharma, R.; Deepak, K.; Sinha, N.; Sen, P.; Thakur, A. K. (Manuscript under preparation).

16. Direct Observation of Intermediate State(s) in the Mechanistic Investigation of Domain Specific Protein-Surfactant Interaction. Yadav, R.; Das, S.; **Sengupta, B.**; Sen, P. (Manuscript under preparation).

17. Tracking Wormlike Micelle Formation in Solution: A Fluorescence Correlation Spectroscopic Study. Subba, N.; **Sengupta, B.**; Sen, P. (Manuscript under preparation).

18. Spectroscopic Investigation of ZnO Nanoparticle Size and Shape Controlled by using EtOH-BzOH Binary Solvent Mixture. Gupta, S.; Mukherjee, P.; **Sengupta, B.**; Sen, P. (Manuscript under preparation).

\* Included in thesis

