Some Studies On The Conformational Stability, Hydration Dynamics And Activity Of Biomolecules In Presence Of Co-Solutes

A thesis submitted for the award of the degree of

Doctor of Philosophy (Science)

In

Chemistry (Physical)

By

Ria Saha

Department of Chemistry (Physical)

University of Calcutta

2024

Dedicated To

my Nation,

my Family

& the respected Supervisor

Declaration

I hereby declare that the research works manifested in the thesis *"Some Studies On The Conformational Stability, Hydration Dynamics And Activity Of Biomolecules In Presence Of Co-Solutes"* are original. All the experiments were carried out by me with the help of my lab-mates, collaborators and my supervisor, Prof. Rajib Kumar Mitra. I further declare that these results have not been used elsewhere for the award of any degree/ diploma from any university or institute.

Date: 26th July, 2024

Ria Saha

Ria Saha

(email: riasaha94@gmail.com)

S N Bose National Centre for Basic Sciences

Block JD, Sec-III, Salt Lake, Kolkata-700106, India

The page is left blank intentionally

Abstract

Molecular insight of proteins is fundamental to comprehending the biophysical, chemical and physiological processes involved with life. Co-solutes (like ions) can modulate the biomolecular function, conformational stability and dynamical arrangements and also lead to pathologically imbalanced states or could induce diseases. In this thesis work, we focus on the intricate role of ions in shaping the structural stability of proteins and how they regulate the biophysical activities of these biomolecules. Additionally, we delve into the associated hydration dynamics, which often lead to malfunctions and various diseases. The thesis contains five articles, namely (i) Trivalent cation-induced phase separation in proteins: ion specific contribution in hydration also counts, (ii) Thermo-Resistive Phase Behavior of Trivalent Ion-Induced Microscopic Protein-Rich Phases: Correlating with Ion-Specific Protein Hydration, (iii) Lanthanides driven microphase behaviours of proteins: ions hydration matters, (iv) Ion-pairing Propensity in Guanidinium Salts Dictates Their Protein (De)Stabilization Behaviour and (v) Thiocyanate ion (SCN^-) offers a major impact in rapid protein amyloidosis: a salient role played by the protein solvation. Lanthanide (III) ions are found to induce microscopic phases (liquid–liquid phase separation (LLPS), re-entrant condensation) in negatively charged globular bovine serum albumin (BSA) proteins at ambient temperature which is contrasting as compared to mono- and di-valent cations. These ions also offer distinct signature at temperature above the melting temperature of BSA and retain structural integrity that is revealed by optical microscopy and turbidity measurements. Terahertz (THz) hydration studies clearly demonstrate that consideration of mere charge (+III in this case) on the metal ions is not sufficient to explain their behaviour towards proteins, ion-specific hydration should also play a key role in such multi-component system. Additionally, the thesis highlights the unique influence of guanidinium (Gdm) ions on protein stability, which varies depending on the counter ions. Through THz absorption spectroscopy, THz time domain spectroscopy, and simulation, the molecular mechanisms underlying this contrasting behavior are elucidated. Furthermore, the impact of SCN^- ions on naturally occurring proteins (ovalbumin, bovine serum albumin, and lysozyme) is investigated, revealing four distinct fibrillar aggregates and demonstrating that the counter-cations (Na^+ and Gdm^+) play a pivotal role in shaping the final arrangements of mature fibrils.

List of Publications

Inside Thesis:

1. R. Saha and R. K. Mitra, "Trivalent cation-induced phase separation in proteins: ion specific contribution in hydration also counts", ***Phys. Chem. Chem. Phys***., vol. 24, pp. 23661-23668, 2022.
2. R. Saha and R. K. Mitra, "Thermo-Resistive Phase Behavior of Trivalent Ion-Induced Microscopic Protein-Rich Phases: Correlating with Ion-Specific Protein Hydration", ***Langmuir***, vol. 39, pp. 4601-4610, 2023.
3. R. Saha, A. Marick and R. K. Mitra, "Lanthanides driven microphase behaviours of proteins: ions hydration matters.", ***Manuscript to be submitted***.
4. R. Saha, S. Chakrabarty, K. Sinha, P. Pyne, S. Pal, A. Barman, S. Chakrabarty and R. K. Mitra, "Ion-pairing Propensity in Guanidinium Salts Dictates Their Protein (De)Stabilization Behaviour", ***Under Review***.
5. R. Saha, I. Bhattacharya, S. Pyne and R. K. Mitra, "Thiocyanate ion (SCN^-) offers a major impact in rapid protein amyloidosis: a salient role played by the protein solvation", ***Under Review.***

Outside Thesis:

1. I. Bhattacharya, R. Saha**,** S. Pyne, A. Bera and R. K. Mitra, "Excipient induced unusual phase separation in bovine serum albumin solution: an explicit role played by ion-hydration", ***Under Review***.
2. S. Pyne, S. I. Haque, R. Saha**,** and R. K. Mitra, "Structural alteration in lipid hydrocarbon part alters the surrounding hydration: THz investigation", ***To be submitted.***
3. V. Rupreo, R. Saha, R. K. Mitra, and J. Bhattacharyya, "Effect of Lobeline alkaloid on lysozyme fibrillation", ***Manuscript to be submitted***.
4. A. Bera, R. Saha, R. K. Mitra, "Spherulite like aggregation observed in amino acids: solvation elucidates the matter", ***Manuscript under preparation***.
5. V. Rupreo, R. Saha, R. K. Mitra, and J. Bhattacharyya, "Exploring the anti-fibrillation potential of Yohimbine and Lobeline alkaloids", ***Manuscript under preparation***.

Acknowledgement

"PhD" is not just an academic journey of an individual, it's a bunch of experiences with some good companies. It has not proved wrong for me as well. So, I should acknowledge those moments and the people who are the part of my journey.

Before writing about the journey, I have to convey my respect to my country for giving me the rights to pursue my dreams. First and foremost, I would like to thank my parents, who have always been my backbone, supporting me at every step. My mother is my first teacher and her ideology, love and care brought me here today. And my father is the one who continuously motivated me to join the PhD program.

It will be a brief space to express my gratitude towards my supervisor, Prof. Rajib Kumar Mitra. According to him, '*He just follows the role of supervisor*', however, his role is uncountable. His continuous support, motivation and fruitful discussions help me to deal with different scientific and non-scientific problems. He taught me how to observe the real world and accept a thing by proper judgement. Beyond his knowledge of basic science, he is really a kind human being. I feel fortunate to work with someone who embodies goodness, calmness, and a welcoming nature. It has been a privilege to have him as my supervisor.

Beside my supervisor I would like to convey my gratitude to my collaborators Prof. Anjan Barman and Dr. Suman Chakrabarty for giving me the opportunity to access their laboratories and research facilities and also for their valuable suggestions.

I express my deepest gratitude to all my lab mates; without their company the journey would have been very difficult. At first, I would like to thank my lab seniors Imadul da, Saikat da, Partha da, Subhadip da and Sumana di. I have spent the maximum part of my journey under their guidance. Saikat da trained me to make protein samples and run a measurement with deep focus. Imadul da is the one who helped me to learn the plotting and fitting using a software like 'Sigma' and 'Origin'. His extra curriculum activities besides research help me to engage with SNB family. Now, coming to Partha da and Sumana di, we shared a lot of experiences together. The COVID restrictions brought us closer to each other which not only nourishes my scientific thinking ability also taught me the subject of life by their experiences. The challenges given by Partha da encouraged me to learn coding in 'Mathematica' to minutely analyze a dataset and how to apply physics on bio-chemical systems. He taught me how to carefully handle an instrument and lab equipment and run the whole lab with gentle care like as one's own home. The love and care by Sumana di are blessings for me in the work place. I would like to specially thank Subhadip da with whom I have shared a great time by discussing so many cultural topics beyond science. Valuable suggestions by him and Subash da encourage me to build a good human nature. I express my love to the recent members of 'Bio-photonics lab'. I am blessed to have two seniors like Indrani di and Tanushree di. I have learnt many things from them. I have '*Three Musketeers*' in lab; Aritra, Rahul and Asesh, the three naughty brothers. Thanks to them for always making me smile and providing a joyful

working place. Thanks to Bidisha, the little and emotional sister, I enjoy her company in every space. I would like to thank Nirnay da, Animesh da, Debashish da, Bivab da, Didhiti di, Sudip da, Anulekha di, Chaitrali di, Anusree, Parama, Sudhanshu, Anushka, Debojyoti, Swati for cheerful discussions and for the friendly atmosphere in the lab. I am lucky to have Jez as a very good friend. Her magical suggestions helped me a lot.

I would like to thank Sir and madam (Dr. Sukanya Chakrabarty) for such delicious lunch at their place and sharing some delightful and memorable times with us. I have had a lot of sweet interactions with madam. A special acknowledgement goes to Sir for serving my favorite *fish dishes*, which were wonderfully prepared by him.

Besides the lab members, I would like to acknowledge some wonderful people I have in my journey. Thanks to Subhamita di, Priyanka di, Anupam da, Swarnali di, Dheeraj da, Debayan da, Anwesha di, Abhik da, Susmita di, Ria (Ghosh) di, Dipanjan, Amrita, Ardhendu, Sreya, Soma, Ria (Barick), Rajdeep, Pritam, *SNB Dance group*, *SNB social wings*, *SNB Students*, *SNB mess committee*, *SNB natok* dol, *Muktangan cultural group.* With these groups I had great fun and refreshment. I am happy to have been a part of the groups.

I express my love to my three friends here. We (Manjari, Nibedita and me) share a tremendous hostel life together. We fight, laugh, make gossips at dinner time. Besides friendship, Krishnendu and me share some good academic discussions. He helped me to learn basic simulation, Python and so many softwares. I am very much delighted to have them.

Special thanks to our *SNB mess* cooking teams and *Bhagirathi* canteen staffs for arranging the delicious foods and giving a homely behaviour. Thanks to Hori da for making my day with cups of tea and coffee with some crispy news. Thanks to the security team, Kamakhya da, cleaning staffs and gardeners to provide a pleasant and beautiful campus with their cares. Thanks to the technical staff (Urmi di, Amit da, Jay da, Debarghya da) and all academic and non-academic staffs for their help in different aspects.

I express my heartily thanks to my family members (my two elder brothers (Bom and Mem) and two sister-in laws (boro boudi and choto boudi)) to always be with me in my ups and downs. A deep love to my nephew (bablush) and niece (Chulku) for making me happy with their sweet care. I would like to thank my husband, my childhood friend, my best friend Koushik for his kind support, suggestions, motivations and be a good listener. Thanks to his family also. Thanks to all my uncles, aunts and cousins to have faith on me.

I am grateful to all my teachers who taught and inspired me at different academic phases.

I am thankful to all the reviewers of different journals for their judicious evaluation and thanks to the thesis committee. Finally, I would like to acknowledge S.N. Bose centre and DST for the financial supports and providing the opportunity to carry out my research work. Last but not the least, thanks to all my doctors for whom I have been able to do my work in good health.

Acronyms

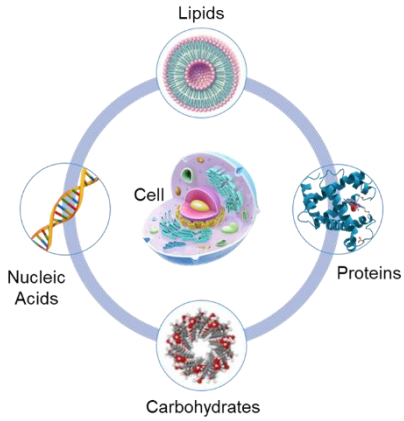
AFM	Atomic Force Microscopy
ANS	Anilinonaphthalene sulfonic acid
ATR	Attenuated Total Reflection
BBB	Blood-Brain-Barrier
BSA	Bovine Serum Albumin
CD	Circular Dichroism
CIP	Contact Ion Pair
CSF	Cerebrospinal Fluid
DHO	Damped Harmonic Oscillator
DIC	Differential Interference Contrast
DLS	Dynamic Light Scattering
EIS	Electrochemical Impedance Spectroscopy
FESEM	Field Emission Scanning Electron Microscopy
FIR	Far-Infra Red
FTIR	Fourier Transform Infrared
Gdm	Guanidinium
HB	Hydrogen Bond
Lib	Librational
LLPS	Liquid Liquid Phase Separation
NMA	N-methyl acetamide
PBS	Phosphate Buffer Saline
REE	Rare Earth Elements
rtl	Rattling
SIP	Solvent separated Ion Pair
ThT	Thioflavin-T
THz	Tera Hertz
UV	Ultra-Violet

Table of Contents

Chapter 1

1. Introduction

Biological systems are hierarchically organized with some emergent properties. Eukaryotic unicellular organisms represent simpler systems that sustain life at the level of individual cells. However, multicellular organisms show hierarchically organized complexity in each level.[1] Biomolecules (proteins, nucleic acids, lipids, carbohydrates, metabolites etc.; scheme 1.1) are the fundamental organic compounds crucial for the maintenance and metabolic functions within these living organisms.[2] These biomolecules are not only the building blocks of the cellular organelles, they are also essential for comprehending the biophysical, chemical and physiological processes involved in life.[3, 4] Among the biomolecules, proteins have the most complexed three dimensional structure with diverse functionality. In a brief line, twenty natural amino acids are sequentially linked together by peptide bonds and form polypeptide chain which further thermodynamically stabilized to folded native conformation by different covalent (di-sulphide bonds) and non-covalent (H-bonds, electrostatics, van der Waals forces, etc.) interactions. In a cell, proteins are associated with cellular membranes, including the plasma membrane and internal organelle membranes (membrane proteins), and are also located within the nucleus (nuclear proteins), in cytoplasm (cytoplasmic proteins) and inside the both membrane-bound and membrane less organelles (organelle specific proteins). Cellular proteins contribute to structural support, enzymatic activities and cell signaling. Some proteins are also present outside the cell (extracellular proteins), play important role in intercellular communication and maintaining the integrity of tissues and organs.[5] However, proteins are biologically active mostly in its native form and misfolding could often lead to malfunction and diseases.[6] Therefore, a deep understanding of their conformational stability and dynamical arrangements helps to grasp how they maintain cellular equilibrium under healthy and pathological conditions.



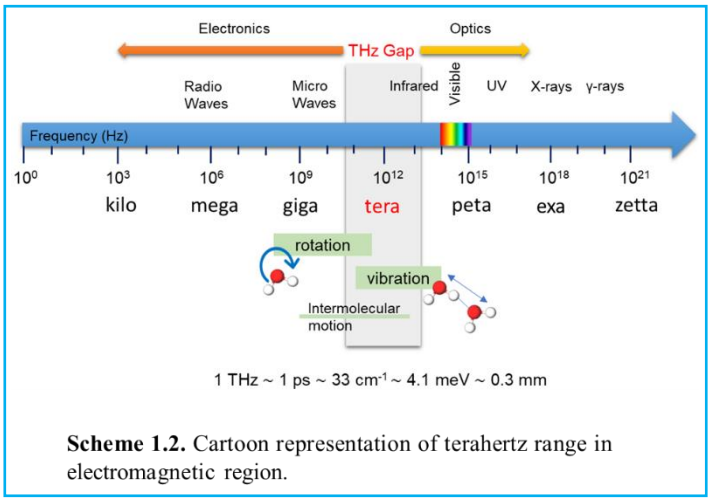
Scheme 1.1. Cartoon representation of biomolecules

The cellular organisms and body fluids (both intra- and extra-cellular fluids) consist of some substances that are known as "*co-solutes*" which have significant effects on the biomolecular structure, stability and function.[7-9] Researchers found that the co-solutes like ions, sugars, polyols, osmolytes, crowders etc. can modulate the biomolecular activities and lead to pathologically imbalanced states or induce diseases.[10-15] Among the co-solutes, ions have major contribution in interactions between biomolecules to support life. Ions belong to the fundamental charged particles in colloidal domain. Beyond non-specific coulombic interactions ions also offer specific effects on bio-macromolecular stability.[16, 17] They, as co-solutes, are integral to numerous physiological processes, ranging from the fundamental cellular functions, nerve signaling, and muscle contraction to the maintenance of blood pressure, pH balance, and mineral homeostasis.[18-20] Some metal (trace metal) ions are naturally abundant in biological systems and are essential nutrients which contribute to fundamental biological processes such as enzyme catalysis, oxygen transport, antioxidant defense, immune function, neurological activity, and growth and development.[21-23] Maintaining appropriate levels of these metals is essential for optimal health and well-being. In this age, technological progress and the surge in industrialization expand routes of exposure to toxic metals. Environmental pollution and occupational exposure (such as mining, smelting, welding, and manufacturing) lead to a range of adverse health effects.[24, 25] Therefore, researchers gain interests in investigating the effects of ions at the molecular level. Accordingly, ions have been observed to play a key role in many biophysical processes involving proteins.[26-28] Inorganic salts with metal cations offer intriguing protein phase behaviors depending on ions valency.[29] These phase behaviors encompass various microscopic phenomena, such as re-entrant condensation of globular proteins,[30] liquid-liquid phase separation (LLPS),[31] oligomerization[32] as well as macroscopic phases like nucleation,[33] fibrillation[34] and cross-linked gelation.[35] Additionally, some bulky polyatomic positively charged ions (cations) act as chemical denaturants and mild stabilizer grouping with negatively charged ions (anions)[36] and commonly used in biochemistry and molecular biology. Ion induced unfolding and/or misfolding of proteins can have significant impact on human health and could lead to the formation of toxic protein aggregates, contributing to neurodegenerative diseases.

Water is the base medium of these types of protein aggregation as it is one of the major components of body fluids.[37, 38] An unusual large dielectric constant is crucial for its solvent activity and a high melting and boiling points offers a wide temperature range over which it remains liquid.[39] Life, as

we understand it, relies on water, or more specifically salty water[40, 41] and that water molecules affect structure and dynamics of biomolecules within it.[37, 41] The interactions between water and biomolecules are mostly through H-bond[42] and long range Coulombic forces with polar or ionic surface residues and weak hydrophobic forces.[43] In liquid phase, each water molecule generally possesses less than or equal to three H-bonds with the possibility one/two via hydrogen donating OH group and two/one via hydrogen acceptor a O atom.[44] Such water structure is dynamic and from breaking to reforming of one H-bond between two neighboring water molecules takes place in picosecond (ps) timescale.[45] This dynamic nature involves distinct intra- and inter-molecular vibrational modes of water; intramolecular OH bond stretching and bending occurs in 10 and 20 femtoseconds (fs)[46] whereas intermolecular H-bond stretching and bending vibrations have time period of 200 fs and 600-800 fs, respectively.[46, 47] Again, intermolecular restricted rotation (libration) appears within tens to hundreds of fs time-scale in H-bonded environment.[46] Water, at the immediate vicinity of a biomoleculer surface, has spatial arrangement than that in the bulk[48, 49] and such structural heterogeneity in the hydration layer is caused by H-bond and electrostatic interaction between the water molecules and the polar and/or charge moieties of the biomolecules.[50] Such bound waters are termed as “biological water” or “hydration shell” and possess slower rotational and vibrational degrees of freedom with a concomitant increase in the H bond lifetime.[51] Researchers have developed number of experimental methods to study the water structure and dynamics around biomolecules. Those experimental techniques give information from different view point like, nuclear magnetic resonance spectroscopy (NMR), neutron scattering, and infrared spectroscopy provide the dynamics of individual water molecules averaging over the collective response. Nuclear Overhauser effect (NOE) coupled NMR spectroscopy offer the hydration around biomolecules[52, 53] in response of the long range dipolar coupling with the water molecules.[54] Again, femtosecond pump probe spectroscopy is one of the powerful techniques that has been used to deploy the reorientation water dynamic around biomolecular surface.[55] Some techniques including optical Kerr-effect (OKE)[56] probe different collective dynamics of water molecules associated with the biomolecules; small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) can determine the radial distribution of counterion around biomolecules and can directly probe the hydration shell structure.[57, 58] The displacement of individual hydrogen atom of water molecules and the translational and rotational dynamical motion of those molecules can be probed by incoherent elastic and quasi-elastic neutron

scattering (EINS and QENS) techniques.[59] Another widely explored tool is time resolved fluorescence spectroscopy (both time correlated single photon counting and fluorescence up-conversion spectroscopy)[51, 60] which indirectly measures the water dynamics at the biomolecular local environment by externally employed fluorophore or by the intrinsic fluorophore present in the biomolecules (e.g. Trp (Tryptophan), Phe (Phenylalanine), Tyr (Tyrosine) residues present in the protein.[51] All these techniques have their own advantages and restrictions. Recently, terahertz (THz) spectroscopy has emerged as a valuable tool for exploring the unique characteristics of water molecules within the THz region (0.3 THz to 22 THz), which lies between the microwave and infrared regions (scheme 1.2) of the electromagnetic spectrum (1THz = 10^{12} Hz = $1ps^{-1}$). THz spectroscopy captures ultrafast fluctuations in the collective dipole moment of water molecules. This technique enables direct investigation of the dynamical and structural responses, as well as the amplitudes of various collective water network motions, including translational, rotational diffusion, and librational motion around biomolecular surfaces.[61-63] One of its main advantages is its label-free, non-invasive nature, with the low power of THz radiation (~4 meV) being insufficient to harm biomolecules. Despite our comprehensive understanding of the structure and dynamics of bulk water molecules, there remains a notable gap in studies concerning the behavior of water molecules buried within biomolecules or within the hydration shell surrounding them.



Scheme 1.2. Cartoon representation of terahertz range in electromagnetic region.

The primary objective of this thesis is to explore various conformational stability of protein molecules and how their dynamical hydration shell picture is perturbed under different charged ionic environments using THz spectroscopy. Water molecules surround individual proteins, forming a structured hydration shell[61] which influences the stability of the protein structure and the interactions between proteins.[64] De-solvation of protein surface is an important factor in protein aggregation both in absence and in presence of co-solutes. Therefore, hydrophobic and hydrophilic interactions between water molecules on protein surface are crucial for determining protein conformation and aggregation propensity.[65] As co-solute, different charged ions or salts have been

chosen to investigate the complex interactions between ions and proteins through the following studies: I. Trivalent cation-induced phase separation in proteins: ion specific contribution in hydration also counts; II. Thermo-Resistive Phase Behavior of Trivalent Ion-Induced Microscopic Protein-Rich Phases: Correlating with Ion-Specific Protein Hydration; III. Lanthanides driven microphase behaviours of proteins: ions hydration matters; IV. Ion-pairing Propensity in Guanidinium Salts Dictates Their Protein (De)Stabilization Behaviour; V. Thiocyanate ion (SCN^-) offers a major impact in rapid protein amyloidosis: a salient role played by the protein solvation. Throughout the thesis works, we characterize each protein-ion system using a range of spectroscopic techniques and detect the morphologies using some microscopic tools. UV-Visible spectroscopy is used to measure the turbidity of protein solutions. Dynamic light scattering spectroscopy (DLS) is utilized for the determination of size specifically hydrodynamic diameter of bare protein and protein in presence of different salts. Zetapotential measurement using the same tool provides the surface charge information of proteins. Steady-state and time resolved fluorescence spectroscopy help to probe the local environment of dye staffs in the different protein environments. The structural conformation of proteins has been studied using circular dichroism (CD) and mid-infrared FTIR techniques. Differential interference contrast (DIC) microscope, atomic force microscope (AFM) and field emission scanning electron microscope (FESEM) used to monitor morphology of different protein aggregates and to probe structural modifications i.e., analysis of height and length distributions. As hydration has an important role in protein conformational stability and functions, the thesis explores THz spectroscopy (both time domain THz spectroscopy and far infrared Fourier transform spectroscopy in the THz region) as a potential tool to probe the changes in collective (sub)ps hydrogen bond structure and dynamics of liquid water explicitly associated with the protein surface and of the ions water cage.

This thesis consists of following chapters as described briefly:

- ***Chapter 1*** deals with the general introduction of biomolecules and how their structure and conformational stability are relevant to maintaining healthy life in presence of co-solutes, and the necessity of studying their physicochemical response using different spectroscopic techniques, especially the state of art THz spectroscopy to probe their hydration behavior.
- ***Chapter 2*** comprises of fundamental theories, mathematical models, and a description of the chemicals utilized in the thesis. Additionally, it delves into the operational principles of

the instrument and the analysis techniques employed to measure or derive the parameters outlined in this study.

- ***Chapter 3*** contains the details investigation of multivalent (specifically trivalent) metal ions induced microscopic liquid–liquid phase separation (LLPS) in negatively charged globular proteins. We experimentally investigate hydration alteration by FIR-THz absorption spectroscopy during the ion-mediated LLPS process in a model protein bovine serum albumin (BSA) at ambient temperature. We have used chloride salts of: Na^{+}, K^{+}, Ca^{2+}, Mg^{2+}, La^{3+}, Y^{3+}, Ho^{3+} and Al^{3+}. We visually detect sample turbidity and measure size of the protein in absence and in presence of these salts using DLS technique. CD measurements confirm compact globular structure of BSA in LLPS state. We observe drastic inversion in the protein surface charge in presence of the trivalent salts. We clearly demonstrate that consideration of mere charge (III in this case) on the metal ions is not sufficient to explain their behaviour towards macromolecules, ion-specific hydrophilic hydration needs to count.
- ***Chapter 4*** explores whether trivalent lanthanide ion induced LLPS offer distinct signature at temperature above the melting temperature of the protein ($T>T_M$). Optical microscopy and turbidity measurements unambiguously reveal that trivalent salt induced LLPS does retain their structural integrity while bare and mono- as well as bi-valent ions containing protein solutions do undergo conformational rupture followed by phase transition upon heating. We finally explore using THz spectroscopy that hydration differs as thermo-sensitive systems undergo phase transition while M^{3+} containing systems retain it.
- ***Chapter 5*** reveals the effects of lanthanide ions on the phase behaviour of proteins beyond LLPS. We choose La^{3+}, Nd^{3+}, Gd^{3+}, Ho^{3+} and Lu^{3+} ions and visualize the change of phases with addition of salt in the protein solution and illustrate in phase diagram. Phase diagrams are classified in three regimes; regime I (R_1) and regime III (R_3) are visually transparent whereas regime II signifies a turbid phase. The starting (C*) and the end (C**) points of regime II are the two critical concentrations and the middle point of that particular zone is depicted by C_m. Further, we characterize lanthanide-induced protein phases (protein homogeneous phase → turbid phase (metastable) → re-entrant to the homogeneous phase) by turbidity (absorption at 400 nm) measurements, zeta-potential measurements and DIC morphologies. Our findings reveal that these microscopic phase behaviors of proteins are

driven not only by changes in the protein surface charge potential but also by ion-protein hydration.

- ***Chapter 6*** covers the experimental validation of guanidinium (Gdm) salts induce contrasting effect on proteins depending on the counter ion(s). Through a combination of electrochemical impedance spectroscopy (EIS) and THz spectroscopy, we highlight the impact of GdmCl and Gdm_2SO_4 on a model amide molecule N-methyl acetamide (NMA). MD simulation studies predict that Gdm_2SO_4 forms hetero-ion pairing in water which inhibits Gdm^+ ions to approach NMA molecules however in GdmCl, the Gdm^+ ions directly interact with NMA. Our experimental results on ion hydration excellently corroborate this idea.
- ***Chapter 7*** describes the rapid protein amyloidosis of three naturally occurring proteins: ovalbumin (Ova), bovine serum albumin (BSA), and lysozyme (Lys) in presence of SCN^- ions and investigate the alteration of water network during the fibrillation pathway. Amyloidosis is important in pathology as it is associated with many neurodegenerative diseases. The study observes that an elevated level of SCN- ions (~0.5 M) induces rapid fibrillization of these proteins at pH 2 and ambient temperature, a phenomenon not typically observed. Field emission scanning electron microscopy (FESEM) and atomic force microscopy (AFM) reveal two distinct initial amyloid-aggregated states: nuclei and protofibrils. After 24 hours of incubation, two mature fibril states emerge: a crosslinked network or matrix and bundle-like structures. Formation of these morphologies depends on the counterion: Na^+ and guanidinium (Gdm^+). We determine the protein and the salt hydration at different phases using THz-FTIR spectroscopy and the findings can be a potential foundation for future handling of amyloidosis.
- ***Chapter 8*** summarizes the key finding of chapter 3, 4, 5, 6 and 7. It also consists of some ideas of future research in this direction.

1.I. References.

1. Y. I. Wolf, M. I. Katsnelson and E. V. Koonin, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E8678-e8687.
2. A. J. Bruce Alberts, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter., *Molecular Biology of the Cell*, Garland Science, New York, 4 edn., 2002.
3. A. L. L. a. M. M. C. D. L. Nelson, *Lehninger Principles of Biochemistry*, W. H. Freeman, 2008.
4. D. E. G. a. R. F. Goldberger, *Molecular Insights Into the Living Process*, Academic Press, 1967.

5. K. Dzobo and C. Dandara, *Biomimetics.*, 2023, **8**, 146.6. F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333-366.
7. H. Cinar and R. Winter, *Sci. Rep.*, 2020, **10**, 17245.
8. J. R. de Xammar Oro, *J. Biol. Phys.*, 2001, **27**, 73-79.
9. B. Mateos, G. Bernardo-Seisdedos, V. Dietrich, N. Zalba, G. Ortega, F. Peccati, G. Jiménez-Osés, R. Konrat, M. Tollinger and O. Millet, *Biophys. J.*, 2021, **120**, 2067-2077.
10. T. Fujii, P. Podbevšek, J. Plavec and N. Sugimoto, *J. Inorg. Biochem.*, 2017, **166**, 190-198.
11. J. Li, J. Chen, L. An, X. Yuan and L. Yao, *Commun. Biol.*, 2020, **3**, 528.
12. W. G. Lilyestrom, S. J. Shire and T. M. Scherer, *J. Phys. Chem. B*, 2012, **116**, 9611-9618.
13. P. Pyne and R. K. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
14. M. Senske, D. Constantinescu-Aruxandei, M. Havenith, C. Herrmann, H. Weingärtner and S. Ebbinghaus, *Phys. Chem. Chem. Phys.*, 2016, **18**, 29698-29708.
15. A. Tsoga, R. K. Richardson and E. R. Morris, *Food Hydrocolloids*, 2004, **18**, 907-919.
16. K. P. Gregory, G. R. Elliott, H. Robertson, A. Kumar, E. J. Wanless, G. B. Webber, V. S. J. Craig, G. G. Andersson and A. J. Page, *Phys. Chem. Chem. Phys.*, 2022, **24**, 12682-12718.
17. P. Lo Nostro and B. W. Ninham, *Chem. Rev.*, 2012, **112**, 2286-2322.
18. W. Aoi and Y. Marunaka, *Biomed. Res. Int.*, 2014, **2014**, 598986.
19. K. Jomova, M. Makova, S. Y. Alomar, S. H. Alwasel, E. Nepovimova, K. Kuca, C. J. Rhodes and M. Valko, *Chem. Biol. Interact.*, 2022, **367**, 110173.
20. U. K. Udensi and P. B. Tchounwou, *Int. J. Clin. Exp. Physiol.*, 2017, **4**, 111-122.
21. N. R. C. U. C. o. D. a. Health., *Diet and Health: Implications for Reducing Chronic Disease Risk.*, Washington (DC): National Academies Press (US), 1989.
22. A. Mehri, *Int. J. Prev. Med.*, 2020, **11**, 2.
23. E. A. B. Pajarillo, E. Lee and D. K. Kang, *Anim. Nutr.*, 2021, **7**, 750-761.
24. J. Briffa, E. Sinagra and R. Blundell, *Heliyon*, 2020, **6**, e04691.
25. C. Nishida and K. Yatera, *Int. J. Environ. Res. Public. Health.*, 2022, **19**.
26. R. L. Baldwin, *Biophys. J.*, 1996, **71**, 2056-2063.
27. P. Jungwirth and P. S. Cremer, *Nat. Chem.*, 2014, **6**, 261-263.
28. H. I. Okur, J. Hladílková, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer and P. Jungwirth, *J. Phys. Chem. B*, 2017, **121**, 1997-2014.
29. O. Matsarskaia, F. Roosen-Runge and F. Schreiber, *ChemPhysChem*, 2020, **21**, 1742-1767.
30. F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, S. Zorn, R. A. Martin, C. M. Martin, G. F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *Phys. Rev. Lett.*, 2008, **101**, 148101.
31. F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. A. Skoda, R. M. J. Jacobs, M. Stzucki and F. Schreiber, *Soft Matter*, 2012, **8**, 1313-1316.
32. L. Xin, X. Duan and N. Liu, *Nat. Comm.*, 2021, **12**, 3207.
33. W. F. Xue, *Biophys. J.*, 2015, **109**, 1999-2000.
34. P. C. Ke, R. Zhou, L. C. Serpell, R. Riek, T. P. J. Knowles, H. A. Lashuel, E. Gazit, I. W. Hamley, T. P. Davis, M. Fändrich, D. E. Otzen, M. R. Chapman, C. M. Dobson, D. S. Eisenberg and R. Mezzenga, *Chem. Soc. Rev.*, 2020, **49**, 5473-5509.
35. B. Li, K. Ren, Y. Wang, Y. Qi, X. Chen and Y. Huang, *ACS Appl. Mater. Interfaces*, 2016, **8**, 30788-30796.
36. S. K. Jha and S. Marqusee, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 4856-4861.
37. Y. Levy and J. N. Onuchic, *Annu. Rev. Biophys. Biomol. Struct.*, 2006, **35**, 389-415.
38. G. M. Cooper, *The Cell: A molecular approach*, Oxford University press, U. S. A.: New York, 8 edn., 2001.
39. N. E. Levinger, *Science*, 2002, **298**, 1722-1723.
40. L. J. Rothschild and R. L. Mancinelli, *Nature*, 2001, **409**, 1092-1101.

41. P. Ball, *Chem. Rev.*, 2008, **108**, 74-108.
42. E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg and P. Hobza, *Pure Appl. Chem.*, 2011, **83**, 1619-1636.
43. D. Chandler, *Nature*, 2005, **437**, 640-647.
44. D. Laage and J. T. Hynes, *Science*, 2006, **311**, 832-835.
45. T. Steinel, J. B. Asbury, J. Zheng and M. Fayer, *J. Phys. Chem. A*, 2004, **108**, 10957-10964.
46. D. Laage, T. Elsaesser and J. T. Hynes, *Chem. Rev.*, 2017, **117**, 10694-10725.
47. C. Fecko, J. Eaves, J. Loparo, A. Tokmakoff and P. Geissler, *Science*, 2003, **301**, 1698-1702.
48. B. Bagchi, *Chem. Rev.*, 2005, **105**, 3197-3219.
49. L. Fumagalli, A. Esfandiar, R. Fabregas, S. Hu, P. Ares, A. Janardanan, Q. Yang, B. Radha, T. Taniguchi and K. Watanabe, *Science*, 2018, **360**, 1339-1342.
50. B. Schneider, K. Patel and H. M. Berman, *Biophys. J.*, 1998, **75**, 2422-2434.
51. S. K. Pal and A. H. Zewail, *Chem. Rev.*, 2004, **104**, 2099-2124.
52. G. Otting, E. Liepinsh and K. Wüthrich, *Science*, 1991, **254**, 974-980.
53. E. Liepinsh, G. Otting and K. Wüthrich, *Nucleic Acids Res.*, 1992, **20**, 6549-6553.
54. B. Halle, *J. Chem. Phys.*, 2003, **119**, 12372-12385.
55. Z. Brotzakis, C. Groot, W. H. Brandeburgo, H. Bakker and P. Bolhuis, *J. Phys. Chem. B*, 2016, **120**, 4756-4766.
56. Q. Zhong and J. T. Fourkas, *J. Phys. Chem. B*, 2008, **112**, 15529-15539.
57. S. A. Pabit, S. P. Meisburger, L. Li, J. M. Blose, C. D. Jones and L. Pollack, *J. Am. Chem. Soc.*, 2010, **132**, 16334-16336.
58. D. Svergun, S. Richard, M. Koch, Z. Sayers, S. Kuprin and G. Zaccai, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 2267-2272.
59. S. HsináChen, *Faraday Discuss.*, 1996, **103**, 281-294.
60. S. K. Pal, J. Peon and A. H. Zewail, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 1763-1768.
61. S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, 2007, **104**, 20749-20752.
62. M. Heyden, S. Ebbinghaus and M. Havenith, in *Encycl. Anal. Chem.*, 2010.
63. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 12068-12073.
64. H. Vondracek, J. Dielmann-Gessner, W. Lubitz, M. Knipp and M. Havenith, *J. Chem. Phys.*, 2014, **141**, 22d534.
65. B. Qiao, F. Jiménez-Ángeles, T. D. Nguyen and M. Olvera de la Cruz, *Proc. Natl. Acad. Sci. U.S.A.*, 2019, **116**, 19274-19281.

Chapter 2

2. Theories, Chemicals, Experimental instrumentations and Analyses

2.I. Basic Theories.

2.I.A. *Protein surface charge density from Zeta potential.*

Zeta-potential (ζ) measurements by *Nano S Malvern* instrument are mainly based on the basic mean-field approach of the Poisson-Boltzmann (PB) theory. The ion distribution around the dispersed charged object creates an electrostatic double layer (EDL), which is a reason for charge screening in electrolyte solutions, over the Debye screening length.

The Debye Huckel parameter (κ) in m^{-1} is first calculated by,[1]

$$\kappa = \sqrt{4\pi\lambda_B N_A \sum_i n_i Z_i^2} \qquad (2.I.A1)$$

with the Bjerrum length[2] in m,

$$\lambda_B = \frac{e^2}{4\pi\varepsilon_0\varepsilon_{H_2O}k_BT} \qquad (2.I.A2)$$

Here, i-th number of ionic species are accounted with Z_i , valences and n_i, number concentration. e is the elementary charge, k_BT denotes the product of Boltzmann constant and absolute temperature. ε_0 and ε_{H_2O} are the vacuum dielectric permittivity and the total dielectric permittivity of water at the respective temperature.[3] The value of $\varepsilon_0$and ε_{H_2O} are taken $8.85 \times 10^{-12}C^2N^{-1}m^{-2}$ and 80, respectively.

Electrophoretic mobility μ of the charged colloidal particle with spherical in shape towards the oppositely charged electrodes under an external electric field relates to the measured ζ by Henry's equation,[4]

$$\mu = \frac{2}{3}\frac{\varepsilon_r\varepsilon_0}{\eta}\zeta f(\kappa a) \qquad (2.I.A3)$$

where η is the viscosity of the medium and a is the radius of the spherical particle.

Here we have measured the ζ of BSA with a value of 3.3 nm. The default Henry function $f(\kappa a) = 1.5$ used in Zetasizer settings, which does not show the distinct effect on the investigated systems. Therefore, $f(\kappa a)$can be calculated as given by Ohshima,[5]

$$f(\kappa a) = 1 + \frac{1}{2}\left[1 + \left(\frac{2.5}{\kappa a[1+2exp(-\kappa a)]}\right)\right]^{-3} \quad (2.I.A4)$$

To get the corrected zeta-potential (ζ_{corr}), ζ for each system was multiplied by its corresponding calculated $f(\kappa a)$ value. Then ζ_{corr} was multiplying with $\frac{e}{k_B T}$ to rescale it as to ζ^*.[6]

The surface charge density of BSA protein, σ (C/m^2), was obtained by,[7]

$$\sigma = \frac{\varepsilon_0 \varepsilon_{H_2O} \kappa k_B T}{e}\left(2sinh\left(\frac{\zeta^*}{2}\right) + \left(\frac{4}{(\kappa a)}tanh\left(\frac{\zeta^*}{4}\right)\right)\right) \quad (2.I.A5)$$

2.I.B. *Dielectric polarization: theory behind the use of Debye model.*

Dielectric relaxation is a phenomenon essential to understand the polarization of a dielectric material changes over time in response to an applied electric field.[8] In the thesis work, we use the terahertz (THz) time domain technique which helps to measure the dielectric relaxation of collective water dipoles under THz electric field.[9] Here the time or frequency domain response of the water dipoles or free charges with electric field variation measures polarization that is the orientation of the dipoles or the charges in the system and provides a dimensionless quantity, the complex dielectric constant. This phenomenon is characterized by a relaxation time, which represents the time it takes for the polarization to reach equilibrium after the removal of the electric field. A brief discussion (detail is found elsewhere[8]) about the basic principle of dielectric polarization is discussed in the following section.

Dielectric polarization (linear) in static electric field: Polarization ($\vec{P}$) is well described as the dipole moment $< \vec{M} >$ (ensemble average) per unit volume of the sample.

$$\vec{P} = \frac{<\vec{M}>}{V} \quad (2.I.B1)$$

The *macroscopic dipole moment* arises due to the charge displacement within the dielectric materials in an external electric field ($\vec{E}$). If the external electric field applied to an isotropic and

uniform dielectric material of dielectric susceptibility χ is static, then the collective polarization can be explained as,

$$\vec{P} = \varepsilon_0 \chi \vec{E} \qquad (2.I.B2)$$

where ε_0 is the dielectric permittivity of the free space ($8.85 \times 10^{-12} C^2 N^{-1} m^{-2}$).

The electric displacement vector ($\vec{D}$) can be written in terms of polarization by the following relation,

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} = \varepsilon_0 (1 + \chi) \vec{E} = \varepsilon \vec{E} \qquad (2.I.B3)$$

ε is the permittivity of the material as the relative dielectric permittivity or *dielectric constant* (a dimensionless quantity) is defined as, $\varepsilon_r = \frac{\varepsilon}{\varepsilon_0} = 1 + \chi$.

Dielectric polarization also depends on the frequency of the applied electric field. Total polarization is additive of the polarization arises from three different types of mechanisms in different frequency regions; termed as: i) *Ionic polarization* ($\vec{P}_i$) in , ii) *Orientation or dipolar rotation polarization* ($\vec{P}_o$), iii) *Deformation (Atomic* ($\vec{P}_a$) *and Electronic* ($\vec{P}_e$)*) polarization*.

Now, total polarization is $\vec{P} = \vec{P}_o + \vec{P}_a + \vec{P}_e + \vec{P}_i$. $\vec{P}$ also, can be written as the sum of *induced dipole polarization* ($\vec{P}_\alpha$) and *permanent dipole polarization* ($\vec{P}_\mu$).

Therefore, $$\vec{P}_\alpha + \vec{P}_\mu = \varepsilon_0 (\varepsilon_r - 1) \vec{E} \qquad (2.I.B4)$$

In *polar dielectrics*, individual molecules inherently possess a permanent dipole moment, even without an external electric field. However, *nonpolar dielectrics* lack a dipole moment unless subjected to an electric field. The long range of dipolar forces makes precise calculation of interactions among dipoles (within a sample) highly complex. A reasonable approximation can be achieved by treating dipoles beyond a certain distance as a continuum with macroscopic dielectric properties. Considering a continuum with a dielectric constant ε_∞, within which point dipoles with a moment μ_d are embedded then we can express the induced polarization as,

$$\vec{P}_\alpha = \varepsilon_0 (\varepsilon_\infty - 1) \vec{E} \qquad (2.I.B5)$$

And dipole polarization ($\vec{P}_\mu$) can be written as follows,

$$\vec{P}_{\mu} = \frac{<\vec{M}_d>}{V} = \frac{\sum_i (\vec{\mu}_d)_i}{V} \qquad (2.I.B6)$$

Dielectric polarization (linear) in time dependent electric field: In dynamic scenarios, when a time-dependent electric field is introduced, the molecules within the material require some time to reach a stable polarization state. This differs from the static case, where polarization instantly aligns with the applied field. The time-dependent displacement vector, proportional to the time-dependent applied electric field, can be expressed as follows:

$$\vec{E}^*(t) = E_0 e^{i\omega t} \qquad (2.I.B7)$$

$$\vec{D}^*(t) = D_0 e^{i(\omega t - \delta(\omega))} \qquad (2.I.B8)$$

where $\delta(\omega)$ is the phase difference between the $\vec{E}$ and $\vec{D}$. In line with the relationship ($\vec{D} = \varepsilon\vec{E}$) observed in the static case, in dynamic scenarios, we can introduce a complex frequency-dependent dielectric permittivity to express this relationship as,

$$\varepsilon^*(\omega) = \frac{D_0}{E_0} e^{-i\delta(\omega)} \qquad (2.I.B9)$$

By using Euler's relations, $\varepsilon^*(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega)$ (2.I.B10)

where $\varepsilon'(\omega) = \frac{D_0(\omega)}{E_0(\omega)}\cos[\delta(\omega)]$, $\varepsilon''(\omega) = \frac{D_0(\omega)}{E_0(\omega)}\sin[\delta(\omega)]$, $D_0(\omega) = E_0(\omega)\sqrt{\varepsilon'(\omega)^2 + \varepsilon''(\omega)^2}$

and $\tan[\delta(\omega)] = \frac{\varepsilon''(\omega)}{\varepsilon'(\omega)}$.

Origin of Debye relaxation model: When the time dependent field is suddenly removed the relaxation function or decay function of the dielectric polarization is given by,

$$\phi(t) = \frac{\vec{P}(t)}{\vec{P}(0)} \qquad (2.I.B11)$$

where $\vec{P}(t)$ is the time dependent polarization decay vector.

Therefore, time dependent displacement vector can be written like, $\vec{D}(t) = \varepsilon_0\vec{E}(t) + \vec{P}(t)$. The polarization of a material cannot instantaneously respond to an external electric field, its response depends on the characteristics of the material. A more general formulation of $\vec{P}(t)$ is provided as follows:

$$\vec{P}(t) = \varepsilon_0 \int_{-\infty}^{t} \Phi(t')\,\vec{E}(t-t')dt' \qquad (2.I.B12)$$

$$\vec{D}(t) = \varepsilon_0[\varepsilon_\infty \vec{E}(t) + \int_{-\infty}^{t} \Phi(t')\,\vec{E}(t-t')dt'] \qquad (2.I.B13)$$

where $\Phi(t)$ is the dielectric response function of the material, $\Phi(t) = (\varepsilon_s - \varepsilon_\infty)/[1-\varphi(t)]$. ε_s and ε_∞ are the dielectric permittivity at very low and high frequency, respectively. The relationship between the frequency-dependent complex permittivity ($\varepsilon^*(\omega)$) and the time derivative of the relaxation function is established via Laplace transformation,[10]

$$\frac{\varepsilon^*(\omega)-\varepsilon_\infty}{\varepsilon_s-\varepsilon_\infty} = \hat{L}[-\frac{d}{dt}\varphi(t)] \qquad (2.I.B14)$$

where $\hat{L}$ is the Laplace operator, which is defined for the arbitrary time dependent function *f(t)* as,

$$\hat{L}[f(t)] \equiv F(\omega) = \int_0^\infty e^{-pt} f(t)dt \qquad (2.I.B15)$$

where $p = x + i\omega$ and $x \to 0$.

Equation 2.I.B14 gives equivalent information on dielectric relaxation properties of the sample both in frequency domain and time domain measurements. Consequently, dielectric response can be experimentally determined as a function of either frequency or time. It provides data in the form of a dielectric spectrum $\varepsilon^*(\omega)$ or the macroscopic relaxation function $\varphi(t)$. For instance, neglecting inertia effects we assume that the change of the polarization is directly proportional to its actual value (only in first order),

$$\frac{d\vec{P}(t)}{dt} = -\frac{1}{\tau_D}\vec{P}(t) \qquad (2.I.B16)$$

where τ_D indicates the relaxation time. The macroscopic relaxation function follows a simple exponential law,

$$\varphi(t) = \exp\,(-\frac{t}{\tau_D}) \qquad (2.I.B17)$$

Equation 2.I.B15 and 2.I.B17 give rise to the following relation,

$$\frac{\varepsilon^*(\omega)-\varepsilon_\infty}{\varepsilon_s-\varepsilon_\infty} = \frac{1}{1+i\omega\tau_D} \qquad (2.I.B18)$$

$$\varepsilon^*(\omega) = \varepsilon_\infty + \frac{\varepsilon_s-\varepsilon_\infty}{1+i\omega\tau_D} \qquad (2.I.B19)$$

This is referred to as the *Debye relaxation model* for frequency-dependent dielectric permittivity. However, it's rare for this relaxation behavior to precisely match experimentally measured dielectric data. While the Debye conjecture is simple and elegant, in many instances, a superposition of multiple Debye functions with varying relaxation times must be employed. This approach considers both the dispersion (real part) and the loss factors (imaginary contribution) in terms of,

$$\varepsilon'(\omega) = \varepsilon_\infty + \frac{\varepsilon - \varepsilon_\infty}{1+\omega^2{\tau_D}^2} \quad (2.I.B20)$$

$$\varepsilon''(\omega) = \omega\tau \frac{\varepsilon - \varepsilon_\infty}{1+\omega^2{\tau_D}^2} \quad (2.I.B21)$$

The real part exhibits a monotonically decreasing trend, whereas the imaginary part reaches a maximum at $\omega = \frac{1}{\tau_D}$.

2.I.C. *Quantum mechanical approach to determine the isotropic absorption coefficient in the far-IR (THz) window.*

As mentioned earlier (*Introduction section*), THz absorption spectroscopy has emerged as a potential tool to probe the changes in the ultrafast fluctuations of water collective dipole moments[11] and any change of water network structure at biomolecular surface offers contrasting features compared to bulk water.[12] Lineshape computation of the THz spectra depends on the assumption that the Thz field (external field) employed to probe the system is only weakly coupled to the system itself.[13]

Here, we employ a two-state model within the framework of quantum mechanics to assess how the absorption coefficient of molecules depends on their dipole moments.[13] Let H_0 denotes the Hamiltonian of the unperturbed time-independent system, initially in the state ψ_i with energy E_i. The molecules can transit from the initial state to the final state ψ_f (with energy E_f) upon interacting weekly with an external electric field of frequency ω ($\approx \frac{E_f - E_i}{h}$). A monochromatic field can be described as $\vec{E}(t) = \epsilon E_0 cos\omega t$. where ϵ is the unit vector that defines the direction of the incident radiation.

If the wavelength of the light is much larger than the molecular dimension, the interaction between the molecule and the light can be written as:

$$\hat{V}(t) = -\hat{\mu} \,.\, \vec{E}(t) \qquad \text{(2.I.C1)}$$

where $\hat{\mu}$ is the dipole moment operator relates with the collective dipole moment of the system.

Now, we can write the total Hamiltonian of the system as $H = H_0 + \hat{V}(t)$, where the explicit time dependence is due to the oscillatory external field, and not to the operator itself (*Schrödinger picture*)

Therefore, the transition probability from ψ_i to ψ_f (according to the time dependent perturbation theory) is

$$P_{if}(\omega) = \frac{1}{\hbar^2} \left| \int_0^\infty < \psi_f | \hat{V}(t) | \psi_i > e^{i\omega_{fi}t} dt \right|^2 \qquad \text{(2.I.C2)}$$

Here we can apply the Fermi Golden rule as IR light source excites vibrational states and these states are almost continuum

$$P_{if}(\omega) = \frac{\pi {E_0}^2}{2\hbar^2} | < \psi_f | \epsilon\hat{\mu} | \psi_i > |^2 [\delta(\omega_{fi} + \omega) + \delta(\omega_{fi} - \omega)] \qquad \text{(2.I.C3)}$$

As the δ term corresponds to energy conservation, to prevent energy loss, $\delta(\omega_{fi} + \omega)$ will be invalid, and equation (2.I.C3) leads to

$$P_{if}(\omega) = \frac{\pi {E_0}^2}{2\hbar^2} | < \psi_f | \epsilon\hat{\mu} | \psi_i > |^2 [\delta(\omega_{fi} - \omega)] \qquad \text{(2.I.C4)}$$

The transition rate for all molecules is given by the average of $P_{if}(\omega)$ over the initial state. Therefore, for a canonical ensemble, the average transition rate is

$$P(\omega) = \frac{\pi {E_0}^2}{2\hbar^2} \sum_i \rho_i | < \psi_f | \epsilon\hat{\mu} | \psi_i > |^2 [\delta(\omega_{fi} - \omega)] \qquad \text{(2.I.C5)}$$

where $\rho_i (= \frac{e^{-\beta E_i}}{Tr(e^{-\beta H_0})})$ is the density matrix eigenvalue.

Now, we can switch to interaction picture from the Schrödinger picture as

$$P(\omega) = \frac{{E_0}^2}{4\hbar^2} \int dt \sum_i \rho_i < \psi_i | \epsilon\hat{\mu} | \psi_f >< \psi_f | \epsilon\hat{\mu} | \psi_i > e^{i(\omega_{fi} - \omega)t}$$

$$\Longrightarrow P(\omega) = \frac{{E_0}^2}{4\hbar^2}\int dt\, e^{-i\omega t}\sum_i \rho_i < \psi_i|\epsilon\hat{\mu}|\psi_f >< \psi_f|\epsilon\hat{\mu}|\psi_i > \qquad (2.I.C6)$$

Now, they form a complete basis of Hilbert space as ψ_f is the eigenstate of unperturbed Hamiltonian. Therefore, the closure relationship $|\psi_f >< \psi_f| = \mathbb{I}$ ($\mathbb{I}$ is the identity operator) holds good and the equation (2.I.C6) reduces to,

$$P(\omega) = \frac{{E_0}^2}{4\hbar^2}\int dt\, e^{-i\omega t}\sum_i \rho_i < \psi_i|\epsilon\hat{\mu}(0)\epsilon\hat{\mu}(t)|\psi_i > \qquad (2.I.C7)$$

Finally, putting the value of ρ_i and summing over we get an equilibrium ensemble average

$$P(\omega) = \frac{1}{3}\frac{{E_0}^2}{4\hbar^2}\int dt\, e^{-i\omega t}\sum_i < \hat{\mu}(0)\hat{\mu}(t) > \qquad (2.I.C8)$$

The "1/3" term arises from considering an isotropic system, where the response is independent of the direction of the radiation.

Another case that may arise is for $\omega < 0$; in this scenario, $\delta(\omega_{fi} - \omega)$ will be invalid. Physically, this case implies an emission process, and the transition rate for such a process can be recalculated by replacing ψ_i to ψ_f and interchanging the indices i to f in equation (2.I.C5) (when the system is in thermal equilibrium),

$$P(-\omega) = \frac{\pi{E_0}^2}{2\hbar^2}e^{-\beta\hbar\omega}\sum_{i,f}\rho_i|< \psi_i|\epsilon\hat{\mu}|\psi_f >|^2[\delta(\omega_{fi} - \omega)] = e^{-\beta\hbar\omega}P(\omega) \quad (2.I.C9)$$

$$[\text{As } \rho_f = \frac{e^{-\beta E_f}}{Tr(e^{-\beta H_0})} = \frac{e^{-\beta(E_i+\hbar\omega)}}{Tr(e^{-\beta H_0})} = \rho_i e^{-\beta\hbar\omega}]$$

Therefore, net energy absorbed per unit time, $\Pi(\omega) = \hbar\omega[P(\omega) - P(-\omega)]$

$$\Rightarrow \Pi(\omega) = \frac{{E_0}^2}{12\hbar}\int dt\, e^{-i\omega t} < \hat{\mu}(0)\hat{\mu}(t) > (1 - e^{-\beta\hbar\omega}) \qquad (2.I.C10)$$

Now the absorption coefficient can be estimated as $\alpha(\omega) = \frac{\Pi(\omega)}{V|S|}$, where V is the volume of the system, and |S| the magnitude of the incident flux of the radiation, the Poynting vector.

Putting the value of Poynting vector, |S| and assuming $(1 - e^{-\beta\hbar\omega}) \approx \beta\hbar\omega$, we finally get,

$$\alpha(\omega) = \frac{1}{4\pi\varepsilon_0}\frac{2\beta\omega^2\pi}{3Vcn(\omega)}\int dt e^{-i\omega t} < \hat{\mu}(0)\hat{\mu}(t) > \qquad (2.I.C11)$$

Hence, equation (2.I.C11) establishes a direct correlation between the absorption coefficient (measured directly from THz spectroscopy) and the collective dipole moment of the sample.

2.I.D. *How THz absorption co-efficient related to Damped Harmonic Oscillator model?*

Due to the significant difference in mass between the nucleus and the electron, the bond between two atoms in a molecule behaves analogously to a spring. This bond holds an intrinsic frequency, and as a result, the molecule exhibits simple harmonic oscillator motion with a natural frequency.[14] The natural frequency of oscillation can be derived from Hooke's law and Newton's second law. When molecules are in motion, frictional forces or damping become relevant, which is proportional to the velocity of molecular motion. If the THz radiation is applied on the sample, THz field driven one dimensional equation of motion of atoms can be written as,[15]

$$\frac{d^2x}{dt^2} + 2\gamma\frac{dx}{dt} + \omega_0^2 x = f cos\omega t \qquad (2.I.D1)$$

with $\gamma = \frac{b}{2m}$; $\omega_0 = \sqrt{\frac{k}{m}}$; $f = -\frac{eE}{m}$.

Now, the solution of equation (2.I.D1) provides: $x(t) = \frac{f cos(\omega t - \psi)}{\sqrt{(\omega_0^2-\omega^2)^2+4\gamma^2\omega^2}}$ with $\psi = tan^{-1}(\frac{2\gamma\omega}{\omega_0^2-\omega^2})$

Again, the absorption of radiation by the ions per unit time is

$$P = -N < \vec{F}.\vec{v} > \qquad (2.I.D2)$$

where $\vec{F} = -e\vec{E}\cos(\omega t)$ and $\vec{v} = \frac{d\vec{x}}{dt}$

We can calculate the power by putting the value of $x(t)$ in equation 2.I.D2 and get,

$$P(\omega) = \frac{N(eE)^2\omega^2\gamma}{m[(\omega_0^2-\omega^2)^2+4\gamma^2\omega^2]} \qquad (2.I.D3)$$

Again, the absorption coefficient, $\alpha(\omega)$ is related with the dissipation as

$$P(\omega) = \alpha(\omega)\frac{c}{n}\cdot\frac{1}{2}\epsilon E^2; \qquad (2.I.D4)$$

where “n” is the refractive index of the medium, “c” is the velocity of the light; "ϵ" is the dielectric constant of the medium.

From simplifying equations 2.I.D3 and 2.I.D4, we get,

$$\alpha(\omega) = A\frac{\omega^2\gamma}{[(\omega_0^2-\omega^2)^2+4\gamma^2\omega^2]} \qquad (2.I.D5)$$

where $A = 2\frac{n}{c}\frac{Ne^2}{m\epsilon}$.

If the molecule undergoes multiple underdamped motions, then the absorption coefficient can be expressed as the sum of these damping motions as

$$\alpha(\omega) = \sum_i A_i\frac{\omega^2\gamma_i}{[(\omega_0^2-\omega^2)^2+4\gamma^2\omega^2]} \qquad (2.I.D6)$$

2.I.E. *Significance of THz absorption spectroscopy to probe the solute hydration: Three components model.*

Water, as the primary solvent and a major constituent of body fluids, exerts a significant influence on shaping the structure and stability of bio-molecules. Solutes (here, proteins and co-solutes) immersed in an aqueous environment feature a unique hydration sheath (nearest water layer surrounding the solute) that significantly differs from the properties of bulk water. The hydration of protein surfaces plays a central role in regulating key processes involved in protein aggregation. THz spectroscopy is a label free tool which can directly probe the hydration sheath and provides the response of water dipole fluctuations in terms of absorption co-efficient.[11, 16] Therefore, any alteration in collective dipole moment of the solution will reflect on the change in absorption coefficient (already explained in the section 2.I.C). For any solute, absorption coefficient can be written as the average weight of absorption coefficient of each component; for solute itself (α_{solute}), for dynamical hydration shell around the solute (α_{shell}) and for bulk water (α_{bulk}).

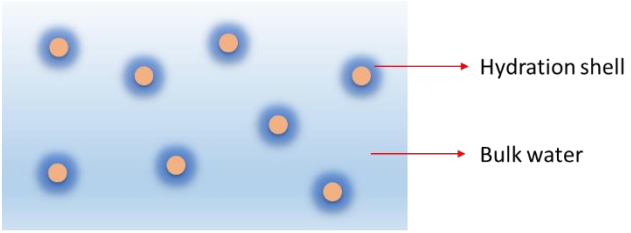
$$\alpha_{total}(\nu) = \alpha_{solute}(\nu)\frac{V_{solute}}{V_{total}} + \alpha_{shell}(\nu)\frac{V_{shell}}{V_{total}} + \alpha_{bulk}(\nu)\frac{V_{bulk}}{V_{total}} \qquad (2.I.E1)$$

where V_{total} is total volume of the solution and V_{solute}, V_{shell} and V_{bulk} are the volume occupied by the solute molecule, dynamical hydration shell and bulk water, respectively.

The three components model (equation 2.I.E1) can also rewrite as,

$$\alpha_{total}(\nu) = \alpha_{solute}(\nu)\phi_{solute} + \alpha_{shell}(\nu)\phi_{shell} + \alpha_{bulk}(\nu)\phi_{bulk} \quad (2.I.E2)$$

where ϕ_i is the volume fraction of the i-th component.



Scheme 2.I.E1. Schematic representation of solute (defined by orange circle) solvation with dynamical hydration shell (dark blue) in bulk water environment (light blue).

2.I.F. *Molecular Dynamics simulations: Background and technique.*

The first computer simulation study in 1953 computed the thermodynamic properties of liquids modeled as hard spheres, laying the foundation for the Metropolis Monte Carlo simulation.[17] Alder and Wainright conducted the initial molecular dynamics study,[18] using hard spheres to obtain dynamic properties from Newton's equations of motion. In 1964, Rahman's molecular dynamics simulation of liquid Argon with a soft sphere potential marked a significant advancement.[19] Following these early studies, the first protein molecular dynamics simulation in 1977 led to rapid progress in the field,[20] now enabling simulations of systems with over a million atoms and long production runs (ns-μs).

Molecular dynamics simulation calculates the positions and velocities of atoms in an N-particle system using Newton's equations of motion:

$$F_i(t) = m_i\ddot{r}_i(t) \quad (2.I.F1)$$

Here, $F_i$is the force on the ith atom at time t, $r_i(t)$ is its position, and m_i is its mass. The force is derived from the potential energy function based on particle interactions:

$$F_i(t) = -\frac{\delta V_{(r_i)}}{\delta r_i} \quad (2.I.F2)$$

where $V_{(r_i)}$ is the potential of the system.

By numerically solving Newton's equations, the positions, velocities, and accelerations of atoms are updated over discrete time steps, generating an ensemble of successive conformations to explore the system's energy landscape.

Steps in Running Molecular Dynamics Simulation

1) System Setup
   - Generate the initial conformation of the system from structural data or by randomly placing molecules in a simulation box.
   - Select an appropriate force field to assign non-bonded and bonded parameters.
   - Solvate the system by adding water and ions to replicate biological conditions.
2) Energy Minimization and Equilibration
   - Perform energy minimization using steepest descent or conjugate gradient methods to remove atom overlaps and steric stress.
   - Assign initial velocities randomly from a Maxwellian distribution and scale to the desired temperature.
   - Conduct NVT and NPT equilibration steps by coupling the system to a thermostat and barostat, respectively, to reach equilibrium.
   - Optionally, restrain the solute to equilibrate the solvent.
3) Production Run
   - Allow the equilibrated system to evolve through conformational space.
   - Use a time step (~2 fs) adequate for the fastest events in the simulation.
   - Ensure the production run is long enough for proper sampling of phase space.
4) Data Analysis
   - Analyze the data on atomic positions and velocities to calculate physical properties.

2.II. Chemicals used.

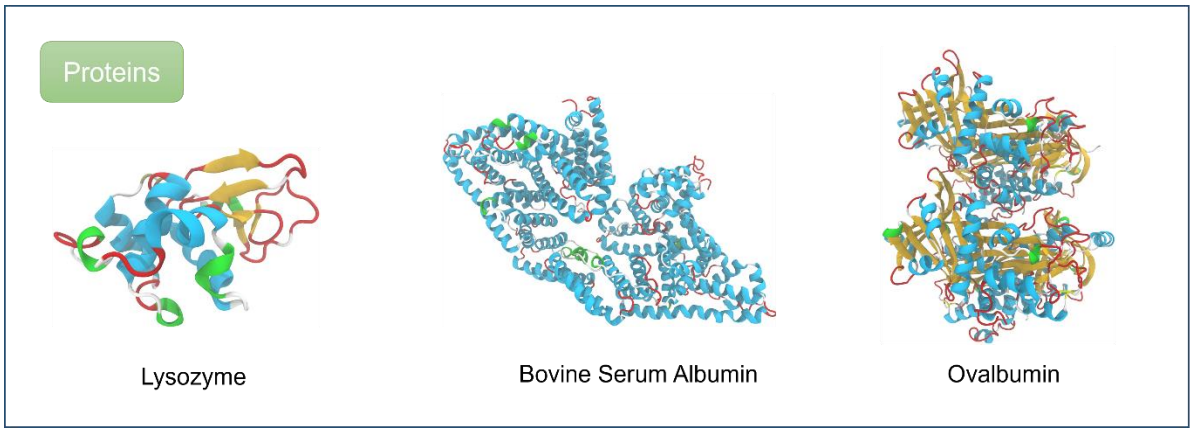
2.II.A. *Proteins.*

a. Bovine Serum Albumin (BSA). Serum albumin is the most abundant protein in the mammal bloodstream.[21] The heart-shaped globular protein (~65% α-helix) Bovine serum albumin (extracted from cow) molecules are analogous to human serum albumin (~70%

similar) with three homologous domains (I, II, and III; each domain consists of two subdomains IIA and IIIA)[22] and has two fluorescence active tryptophan moieties (Trp 134 and Trp 213). ~66 KDa average molecular weight contained BSA molecule has 583 amino acid residues with 17 disulphide bridges. Isoelectric point of BSA is ~4.6.[23]

b. ***Hen Egg White Lysozyme (Lys)*** is a hydrolytic glycosidase and 129 amino acid residues containing single chain polypeptide with ~ 14.3 KDa average molecular weight. Isoelectric point of Lys is 11.[24]

c. ***Ovalbumin (OVA)*** is the main protein of egg white and 55% of the total protein content.[25] It is also a globular protein with average molecular weight 45 kDa and 385 amino acid residues. Ovalbumin has an isoelectric point of 4.5.[26] It belongs to the serpin superfamily.[27]



Scheme 2.II.1. Crystallin structures of Lysozyme (Lys; 1LYZ), Bovine Serum Albumin (BSA; 4F5S) and Ovalbumin (OVA; 1OVA) proteins from protein data bank (PDB). [All the PDB files are visualized and modified in VMD software]

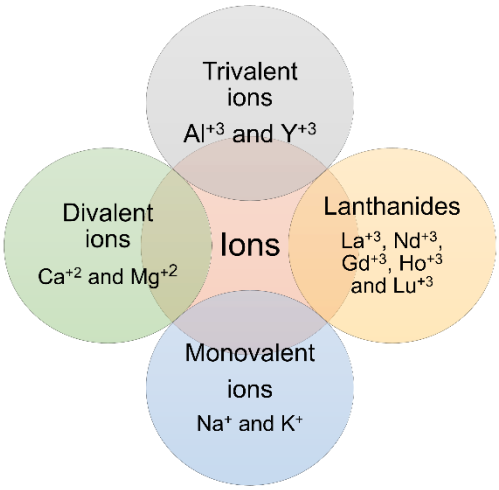
2.II.B. *Main group cation chlorides.*

a. ***Monovalent chlorides.*** Sodium chloride (NaCl; molecular weight 58.44g) and Potassium chloride (KCl; molecular weight 74.55g).

b. ***Divalent chlorides.*** Magnesium chloride ($MgCl_2$; molecular weight 95.21g) and Calcium chloride ($CaCl_2$; molecular weight 110.98g).

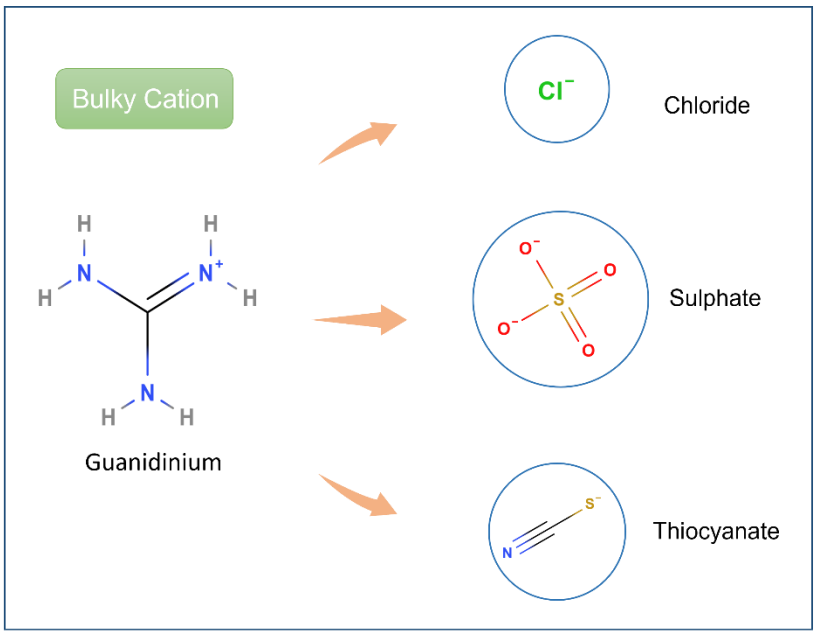
c. ***Trivalent chloride.*** Aluminium chloride ($AlCl_3$; molecular weight 133.34g) and Yttrium chloride (YCl_3; molecular weight 195.26g).

2.II.C. ***Lanthanide chlorides.*** Lanthanum chloride ($LaCl_3$; molecular weight 245.26g), Neodymium chloride ($NdCl_3$; molecular weight 250.6g), Gadolinium chloride ($GdCl_3$; molecular weight 263.61g), Holmium chloride ($HoCl_3$; molecular weight 271.29g) and Lutecium chloride ($LuCl_3$; molecular weight 281.33g).



Scheme 2.II.2. Schematic representation of ions.

2.II.D. ***Inorganic salts with bulky ions.*** Guanidinium chloride (GdmCl; molecular weight 95.53g), Guanidinium sulphate (Gdm_2SO_4; molecular weight 108.11g), Guanidinium thiocyanate (GdmSCN; molecular weight 118.16g) and Sodium thiocyanate (NaSCN; molecular weight 81.07g).



Scheme 2.II.3. Chemical structures of guanidinium (Gdm^+) salts. [The structures have been collected from *www.molview.org*]

2.II.E. *Model amide.*

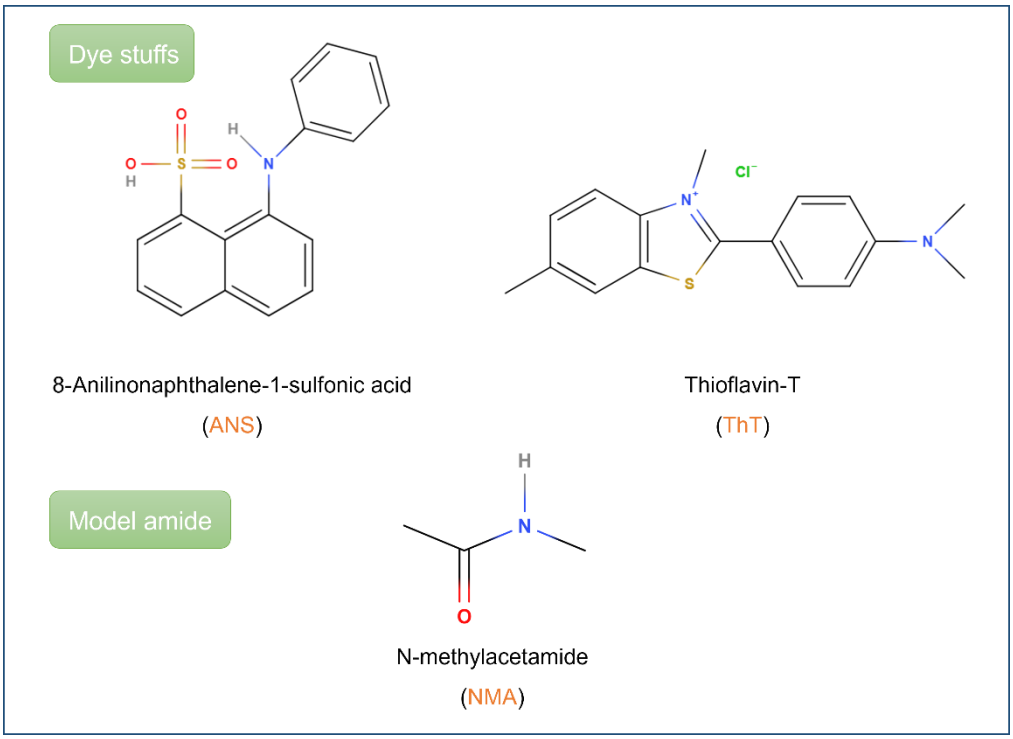
N-methyl-acetamide (NMA) is the simplest molecule containing the peptide bond and a structural unit of the polypeptide chain.[28] NMA has average molecular weight 73g and crystalline in form at room temperature with 0.957 gml^{-1} density. It can be both donor and acceptor of H-bonds.[29]

2.II.F. *Dyes.*

a. ***8-Anilinonaphthalene-1-sulfonic acid (ANS)*** is a charged hydrophobic fluorescent dye and frequently used to probe the hydrophobic environment in protein system.[30] Minimally soluble in polar solvent (quantum yield about 0.004) and highly soluble in nonpolar solvent (quantum yield about 0.98).[31, 32] Molar extinction co-efficient at 350 nm excitation wavelength is 5000 M^{-1} cm^{-1}.[33]

b. ***Thioflavin-T (ThT)*** is a benzothiazole-based fluorophore and the widely used amyloid marker. ThT dye minimally fluoresces in water however, the restricted rotation of bound ThT molecule in cross-β structure of fibril originates intense emission peak at ~480 nm by

excitation at ~440 nm. Molar extinction co-efficient at 412 nm excitation wavelength is 31600 M^{-1} cm^{-1}.[33]



Scheme 2.II.4. Chemical structures of ANS, ThT dyes and NMA model amide. [The structures are made by using *www.molview.org*]

2.III. Experimental Techniques and Instrumentations.

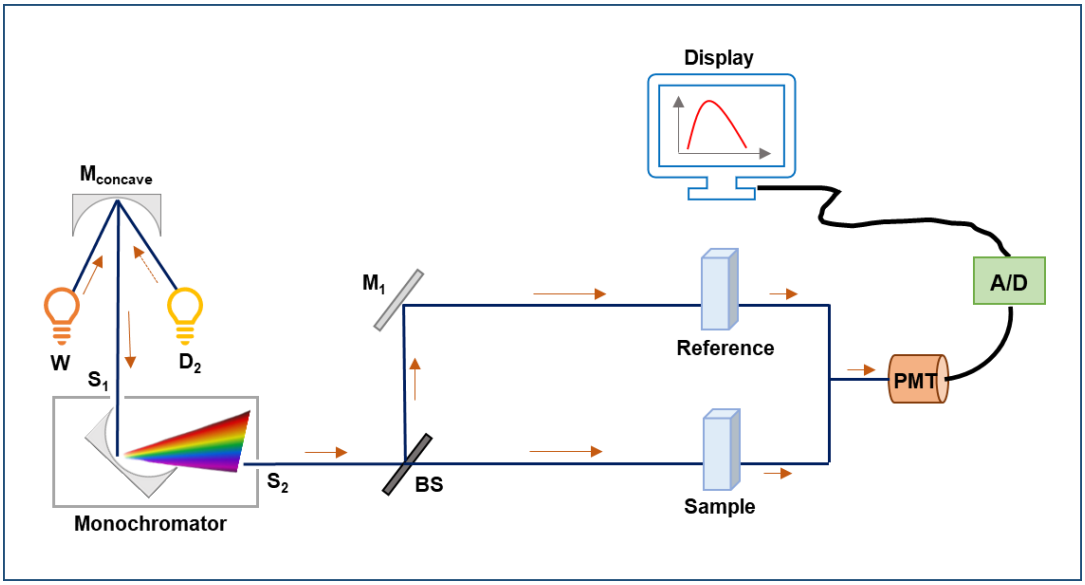
2.III.A. *UV-Visible absorption spectrophotometer.*

UV-Visible spectrophotometer is one of the basic spectroscopic tools which is used to record absorption spectra. This spectroscopic technique insights into the transitions of molecules from their ground state (S_0) to higher electronic states (e.g., S_1, S_2…S_n etc.). In the thesis, *Shimadzu-2600* spectrophotometer (Scheme 2.III.A) has been utilized with a tungsten (W) lamp as visible range source and deuterium (D_2) lamp as Ultra-Violet (UV) range source. 282-293 nm wavelength is the lamp interchange region. The light beam from the source splits into different component wavelengths by the monochromator. High performance blazed holographic grating presents inside

the monochromator is responsible for tuning the white light into a single component. When the sample is continuously irradiated with the radiation, some part of the radiation gets absorbed and the residual part transmitted through the sample. The transmitted light is then collected and detected by a photo-multiplayer tube (PMT). In mathematical point of view the absorption of a molecule can be defined as $A=\log_{10}(\frac{I_0}{I})$ and according to the Beer-Lambert's law the absorption is, $A_\lambda = \varepsilon_\lambda cl$. Here I_0 and I is the incident and transmitted light intensity, respectively. ε_λ is the molar extinction coefficient of the sample at wavelength λ, c is the concentration of sample solution and l is the sample path length.

Therefore; $A_\lambda = \log_{10}(\frac{I_0}{I}) = \varepsilon_\lambda cl$ (2.III.A1)

Turbidity measurement of the sample solution also has been done by monitoring the absorbance at the higher wavelength (for protein sample at 400 nm) where the sample itself does not absorb.[34] A temperature Peltier attached with the instrument was used during temperature dependent turbidity measurements.



Scheme 2.III.A. Schematic diagram of UV-Visible absorption spectrophotometer. $M_{concave}$ is a concave mirror and M_1 is a plane mirror. W and D_2 is the tungsten and deuterium lamp, respectively. S_1 and S_2 are the slits. BS is the beam splitter.

2.III.B. *Dynamic light scattering spectroscopy.*

Dynamic light scattering (DLS) is a spectroscopic technique (also known as Quasi-Elastic light Scattering) which provides the hydrodynamic diameter (d_H) of the sample particles.[35, 36] We use *Nano S Malvern* instrument (Scheme 2.III.B), engaging a He-Ne laser (λ=632.8 nm) with lasing power 4mW. The instrument has a thermostatic sample chamber and it is enriched with zeta-potential measurement method. Photon collides with the particles in solution and scatters in all direction. The detector only collects the scattered photons angled at 173º. The intensity of collected scatter light fluctuates over time due to the Brownian motion of the particles. The velocity of this Brownian motion correlates directly with the diffusion coefficient of the particles in the solution. The definition of the scattered intensity correlation function is,

$$G(t) =< I(t)I(t+\delta t) > \quad \text{(2.III.B1)}$$

For nondispersive particle, correlation function is fitted by the exponential decay function,

$$G(t) = A[1 + Be^{-2\Gamma\tau}] \quad \text{(2.III.B2)}$$

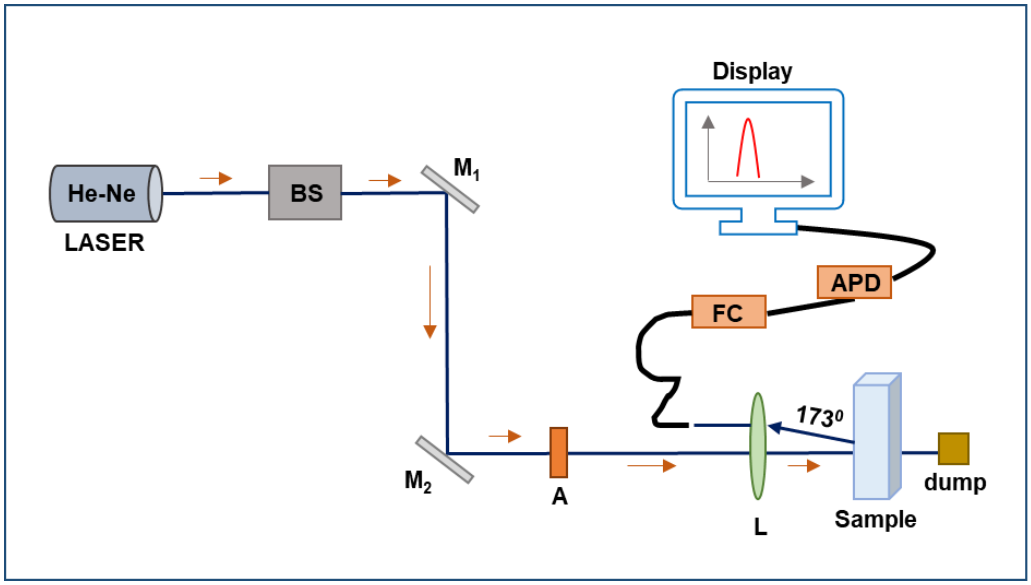
Where A and B are the baseline and intercept of the correlation function and $\Gamma = Dq^2$ and $q = \frac{4\pi n}{\lambda} Sin(\frac{\theta}{2})$; D is the diffusion coefficient, n is the refractive index of the dispersant, λ is the wavelength of the laser and θ is the scattering angle.

For polydisperse particle, the correlation function can be fitted by,

$$G(t) = A[1 + Bg(t)^2] \quad \text{(2.III.B3)}$$

where g(t) is the sum of all exponential decay functions.

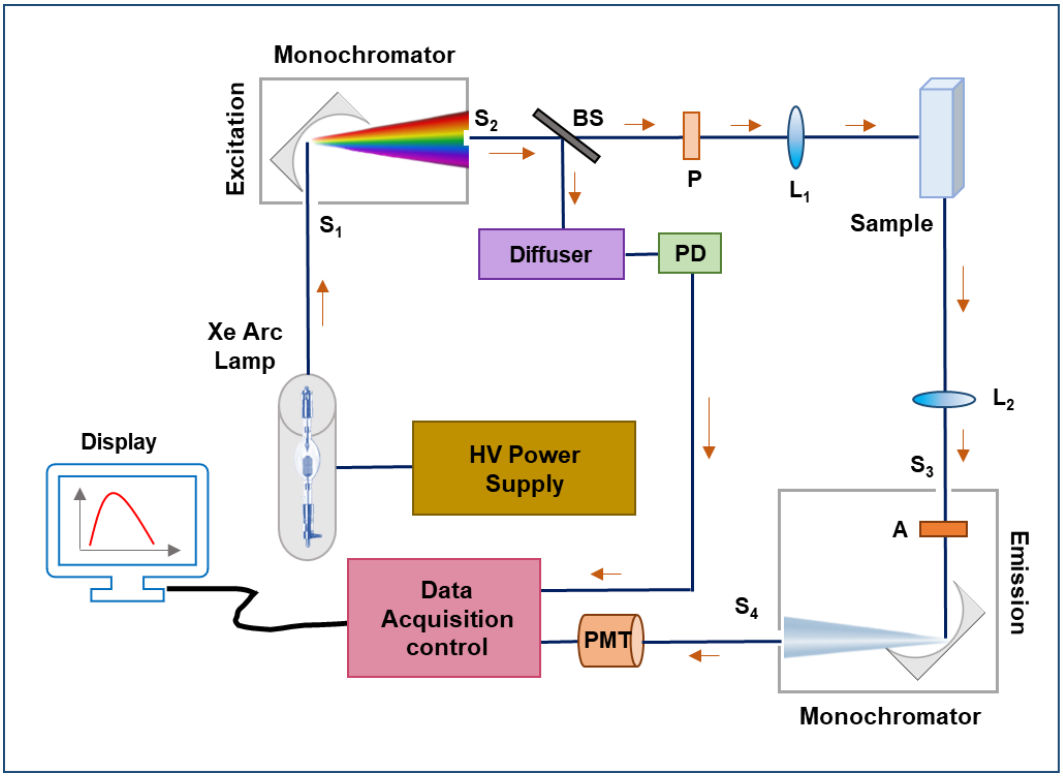
Now, the d_H of the particle is directly related to the diffusion coefficient of the solution by the Stokes Einstein model: $d_H = \frac{k_B T}{3\pi\eta D}$ where k_B is the Boltzmann constant η is the viscosity of the solution.



Scheme 2.III.B. Schematic diagram of dynamic light scattering spectrometer. M_1 and M_2 are the plane mirrors. L is the lens and A is the attenuator. BS, FC, APD and dump stands for the beam splitter, fibre coupler, avalanche photodiode and beam dump, respectively.

2.III.C. *Steady state fluorescence spectroscopy.*

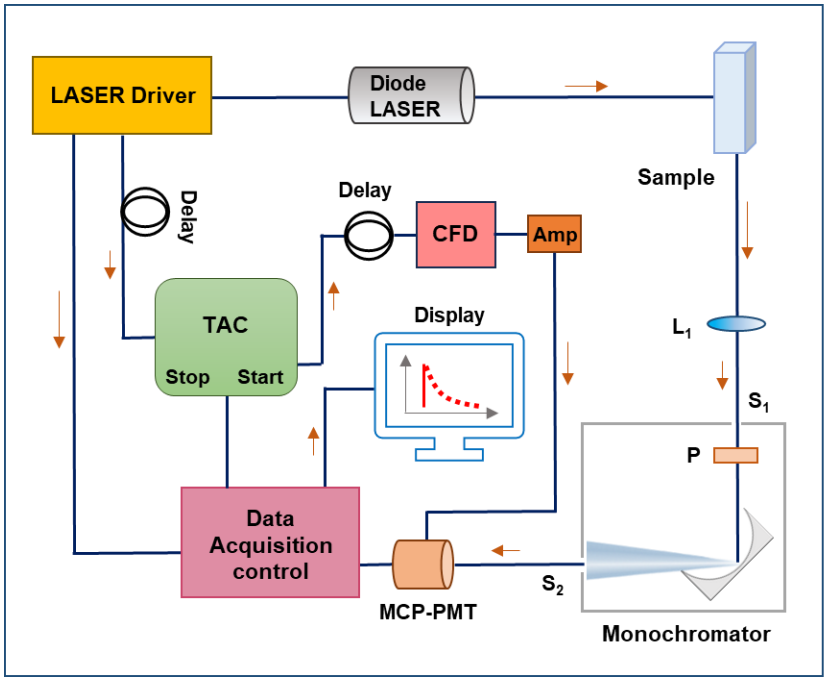
We employed the *Fluorolog* instrument (*Horiba Jobin Yvon*) for the steady state fluorescence measurements. Detailed information of the working principal and the instrumentation (Scheme 2.III.C) can be found elsewhere.[37] Ozone free Xenon lamp is the continuous light source. Light emitted from the source splits into tunable monochromatic component by excitation monochromator which excite the sample. According to Kasha's rule, the sample fluoresces from its lowest excited electronic state (S_1), with the exception of malachite green, which fluoresces from the S_2 state. Emitted light from the sample is collected in perpendicular direction to minimize the stray light (scattering). Then the emission monochromator is used to collect the fluorescence intensity at a specific wavelength. Finally, the signal is captured by a PMT and data acquisition is performed using a computer.



Scheme 2.III.C. Schematic of steady state fluorescence technique. S_i is the split and L_i is the lens. BS, P and A stands for the beam splitter, polarizer and attenuator, respectively.

2.III.D. *Time resolved fluorescence spectroscopy.*

We conducted picosecond (ps) resolved fluorescence transients using the time-correlated single photon counting (TCSPC) technique (*Life Spec II Edinburgh Instruments, UK*). A schematic diagram of the instrument is shown in Scheme 2.III.D. Instrument response function (IRF) is laser specific and ranges from 70-250 ps. Emission polarizer angle is kept at the magic angle 54.7º corresponds to the polarization axis of the excitation beam to minimize the anisotropic effect.



Scheme 2.III.D. Schematic of TCSPC instrument. S_i is the split and L_1 is the lens. TAC and CFD is the time to amplitude converter and constant fraction discriminator, respectively.

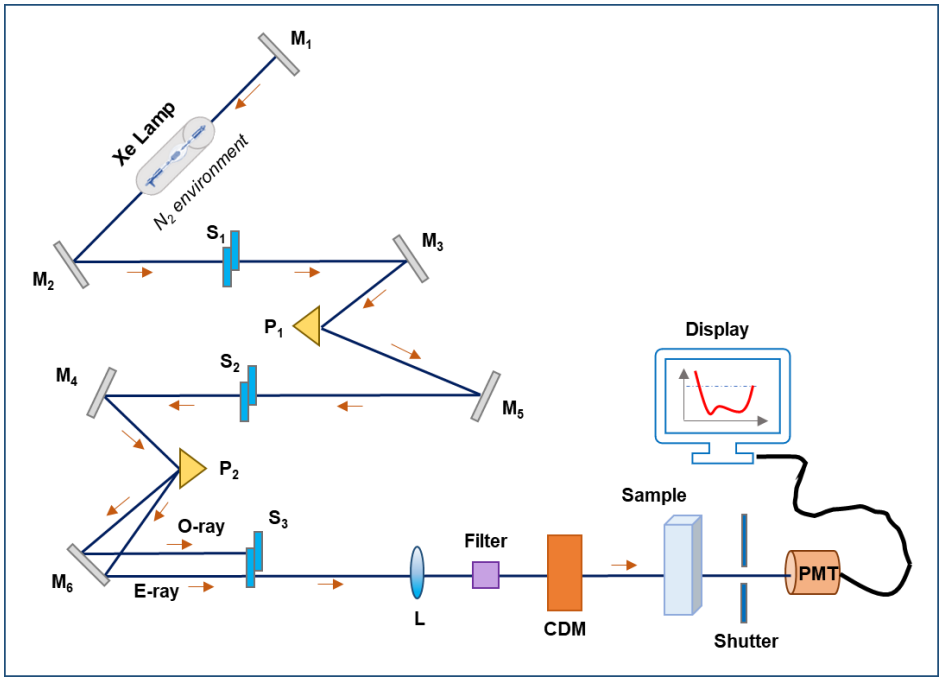
2.III.E. *Circular Dichroism spectroscopy.*

Circular dichroism spectroscopy (*Jasco J-815*) is one of the renowned methods to analyse the protein secondary structures. A schematic representation of CD instrument has been provided in Scheme 2.III.E. A Xenon lamp is used as a source and N_2 gas is continuously purged to remove ozone gas from the source chamber. Chiral molecule absorbs the left and right circularly polarized light differently ($A_\lambda(LCP) - A_\lambda(RCP)$) which results the ultimate CD signal (θ_λ in mdeg).[38]

Molar ellipticity, ε is defined as[39]

$$\varepsilon\ (\mathrm{deg}\, cm^2 mol^{-1}) = \frac{\theta_\lambda(mdeg)}{C_M \times l \times 10} \qquad \text{(2.III.E1)}$$

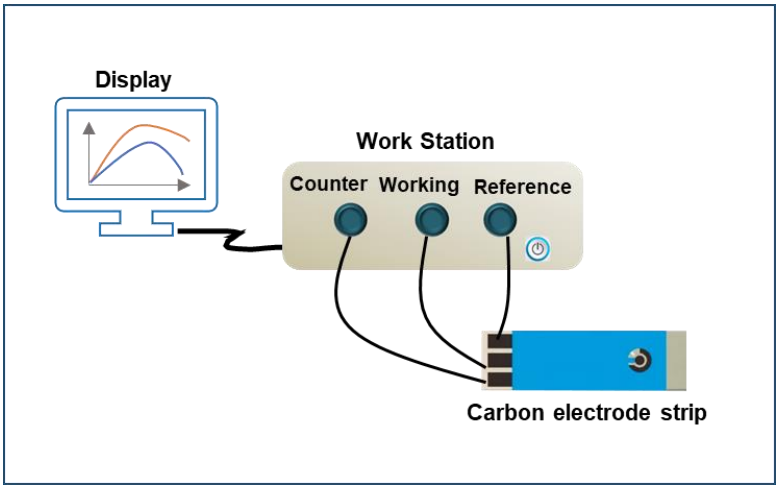
where C_M is the molar concentration of sample and l is path length of the sample in cm.



Scheme 2.III.E. Schematic representation of CD instrument. S_i is the splits; M_i is the plane mirrors and P_i is the reflecting prisms. O-ray and E-ray stands for ordinary and extra-ordinary ray, respectively. CDM and PMT specifies the circular dichroism modulator and photomultiplier tube, respectively.

2.III.F. *Electrochemical Impedance spectroscopy.*

Electrochemical workstation (*CH Instruments Model 600E Series*) has been used to perform counter-ion dynamics measurements. It measures the impedance a complex quantity of an electrochemical system as a function of frequency. The work station consists a working electrode (the material or surface of interest), a reference electrode, and a counter electrode. A drop of ~100 µL volume placed on the top of *Zensor* three-screen printed electrode strip in case of each measurement. We have changed the screen electrode prior to any measurement. All the simulation analysis were performed with equivalent circuit modeling using inbuilt software of the instrument.

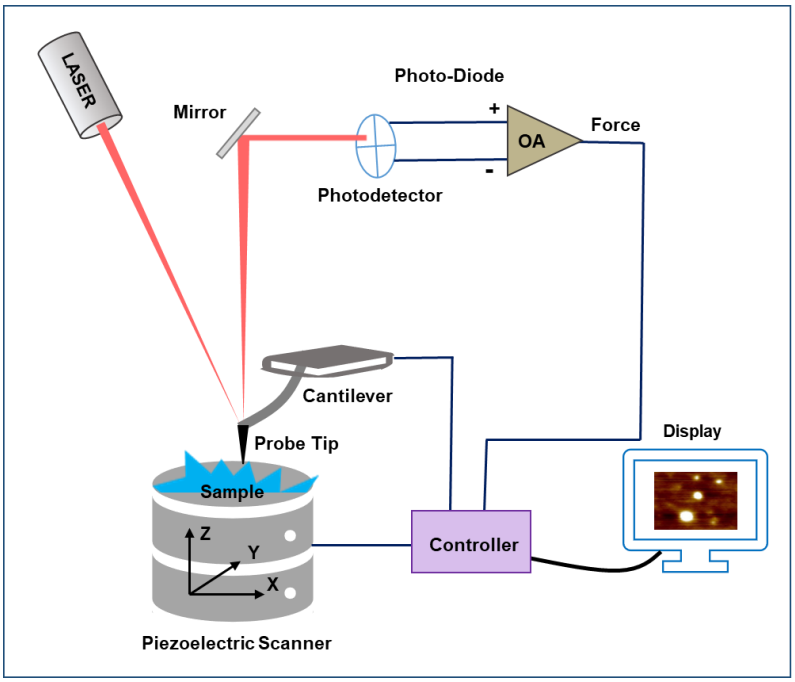


Scheme 2.III.F. View of EIS instrument. The strip of carbon electrode is connected with the work station through the cable.

2.III.G. *Atomic Force Microscopy.*

Morphologies of different aggregation states of proteins were collected using atomic force microscopy (AFM; model number *di INNOVA*). AFM is a type of scanning probe microscope that enables the acquisition of topographical images of sample surfaces.[40, 41] The technique basically helps in acquiring the information by measuring the force between a probe and the sample.[42, 43] The typical "AFM probe" consists of a cantilever with a sharp tip (made of silicon or silicon nitride) at its end (schematically shown in Scheme 2.III.G). The tip experiences vertical and lateral deflections coming into proximity of the sample surface. The deflection is mapped by using a laser (of wavelength 600-700 nm) spot reflected from the top surface of the cantilever. The reflected laser beam hits a position-sensitive four quadrant photodetector. The signals arise from the different segments of the photo detector are different and also give the estimate of the angular deflections of the cantilever. In absence of any force between tip and sample (i.e. both are far away), laser beam strikes the middle of the photo detector and the signal comes from the upper and lower panel of the photo detector remains almost unchanged. Therefore, the ultimate signals provide by the operational amplifier (OP-AMP) is zero. However, when the tip interacts with the sample surface, even with a small force, both the cantilever and the reflected beam deflect. This deflection leads to a non-zero output force from the OP-AMP. It is noteworthy that in general the

tip is kept in fixed position, and a small area of the sample is raster scanned in the x, y, z direction using piezoelectric material and construct 3D-topography of images. Depending on the sample specification and the acting force between the probe and the sample, AFM can be operated in one of the three modes, namely contact mode or non-contact mode (for hard sample) or tapping mode (for soft bio-samples). In this thesis, we have utilized tapping mode to acquire the topography image of the samples.

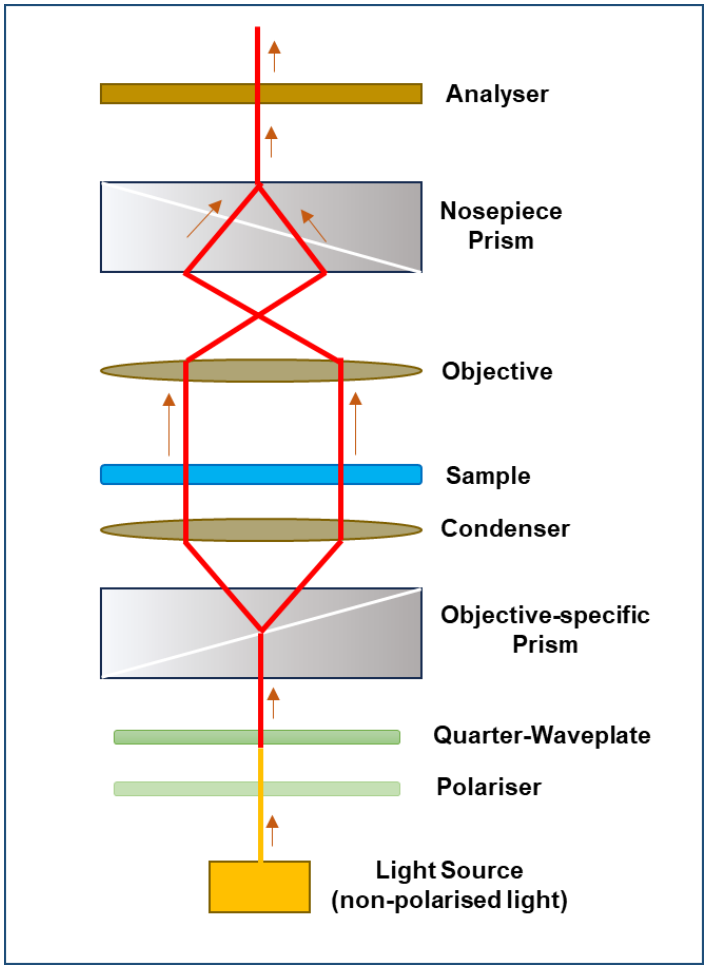


Scheme 2.III.G. Block diagram of AFM instrument.

2.III.H. *Differential Interference Contrast Microscopy.*

We have performed differential interference contrast (DIC) microscopy measurements in Leica DM6M optical microscope. We prefer differential interference contrast (DIC) method over simple bright-field microscopy technique,[44] as undetectable phase differences to the human eye produce high contrast images with a pseudo three-dimensional relief shading. In the DIC microscope,[45] a linearly polarized light source illuminates the sample. Initially, the polarized light is split into two distinct rays with perpendicular polarizations. Although these rays travel very close to each other, they do not interfere due to their orthogonal polarizations. As these rays pass through the specimen, the different thickness and refractive index of the sample cause an alteration in the effective path

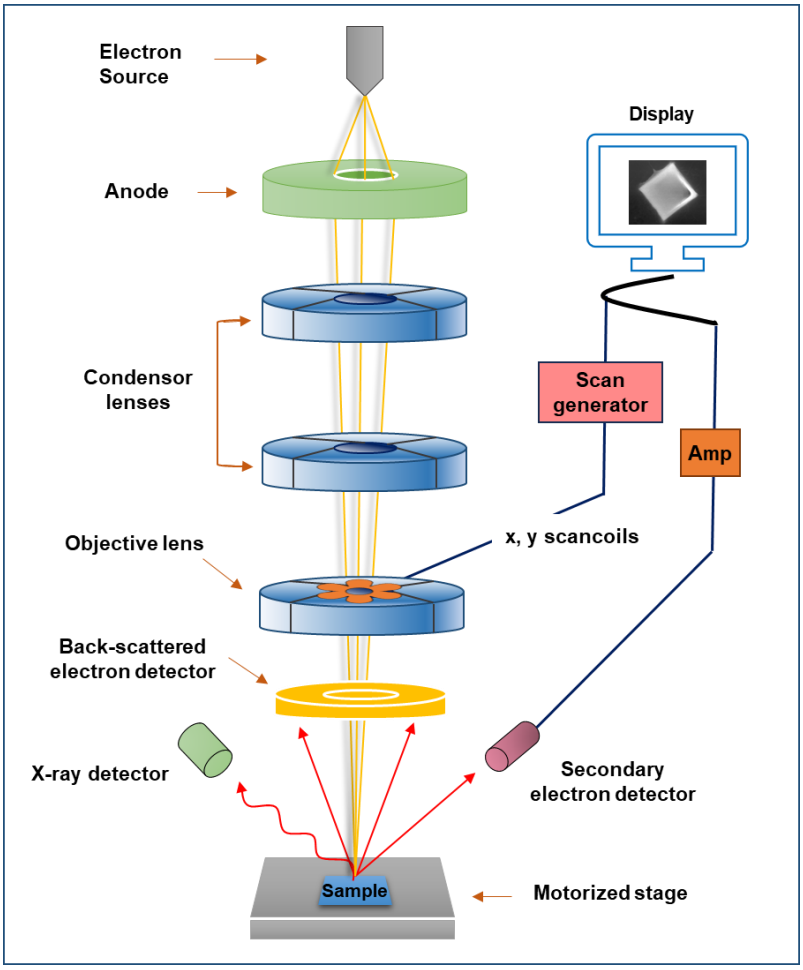
length (means the path travel by light multiplied by the refractive index of the medium) of both rays, resulting in a phase shift (equivalent to $2\pi\lambda$ path difference; λ is the wavelength). Then they pass through the objective lens, where they are focused above their rear focal plane. Upon recombination, the rays interfere with each other, and the resultant light becomes elliptically polarized. This polarization can be converted into an amplitude shift using an analyzer. Finally, the light passes through the objective or camera, where the image is formed based on differences in intensity and color. A schematic diagram of this DIC microscope setup is illustrated in Scheme 2.III.H.



Scheme 2.III.H. Illustrative diagram of DIC microscope.

2.III.I. *Field Emission Scanning Electron Microscopy.*

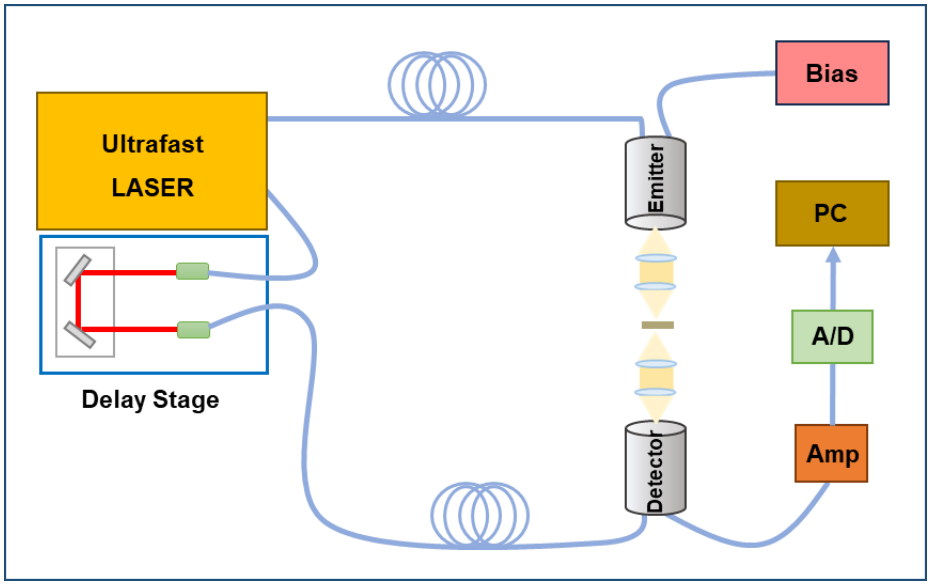
Scanning Electron Microscopy (SEM) is an electron microscopy technique which is utilized to analyze surface chemistry and morphology with submicron resolution.[46, 47]s In this instrument, an electron gun serves as the source, accelerating electrons through accelerating voltages ranging from 1 to 30 kV. Electromagnetic condenser lenses focus these electrons into a beam, which is then scanned across the sample surface to generate images. To prevent electron-air interactions, the SEM chamber is maintained under vacuum conditions (typically with a pressure of 0.1-10^{-4} Pa). As the beam hit the sample surface, electrons lose energy due to scattering and absorption within the interaction volume. The generated signals can be detected by specialized detectors. These signals include secondary electrons (produce SEM images), backscattered electrons (due to elastic scattering), auger electrons, transmitted electrons, photons (characteristic X-rays that are used for elemental analysis and continuum X-rays), visible light (cathodoluminescence–CL), and heat. Secondary electrons originate from a few nanometers beneath the sample surface and provide detailed topographic information. To prevent any scanning faults and other image artifacts the specimens should be electrically conductive at the surface and electrically grounded, because a nonconductive surface accumulates electric charge when scanned by the electron beam. We use field-emission scanning electron microscope (*FEI QUANTA FEG 250*) in the thesis (Scheme 2.III.I).



Scheme 2.III.I. Detail scheme of FESEM microscope according to the working principal.

2.III.J. *Terahertz Time Domain Spectroscopy.*

THz measurements were carried out in a commercial compact spectrophotometer, *TeraSmart of Menlo System GmbH, Munich, Germany* (Scheme 2.III.J). A 1550 nm, integrated femtosecond laser source having a pulse width of <100 fs and a repetition rate of 100 MHz was used to excite a Fe:InGaAs/InAsAs THz emitter antenna which produce THz radiation with a bandwidth of up to 6 THz. The THz radiation was then focused on the sample, and the transmitted THz radiation was further focused on a LT InGaAs/InAlAs THz detector antenna. Before each measurement, the optical chamber was purged with dry nitrogen to avoid water vapour absorption and to control humidity of <10% at 20^0 C. For each sample at least three repeated spectra were taken using a liquid cell with z-cut quartz windows and a Teflon spacer of 100 μm thickness.



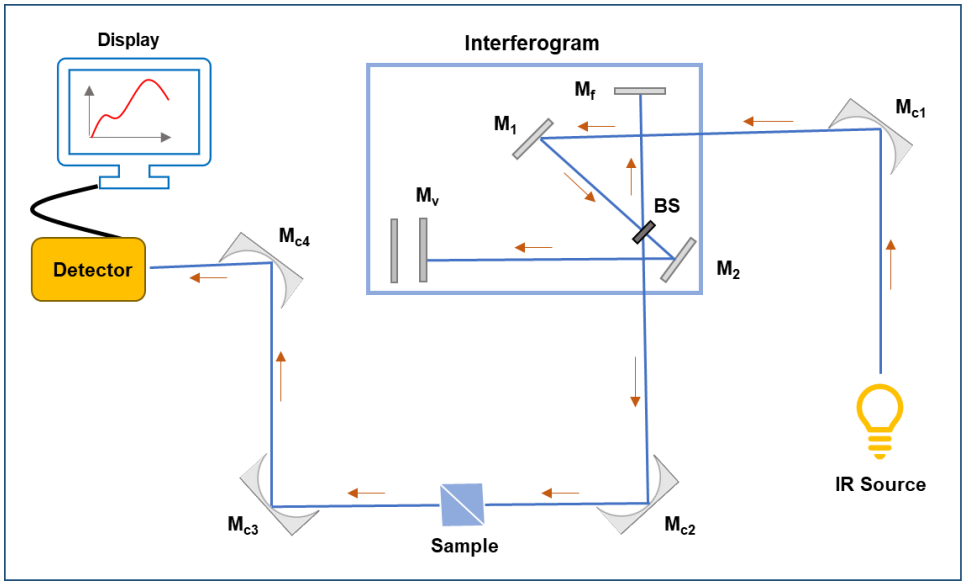
Scheme 2.III.J. Schematic representation of TTDS instrument.

2.III.K. *Terahertz Far-Infra Red Absorption Spectroscopy.*

In this thesis work, we perform terahertz far infra-red (THz-FIR) absorption measurements by using Fourier transform infrared (FTIR) instrument (model number: *Vertex 70V, Bruker, Germany*). Here Globar is the IR source and a complex material DLaTGS is used as a detector. The FIR spectra were collected in the range of 50-750 cm^{-1}(1.5-22.5 THz). The instrument has an ATR (attenuated total reflection) unit, attached with a single reflection diamond crystal (refractive index 2.38). Vacuum level (<1 hPa) of the sample compartment was maintained by a dry vacuum pump (prior to each measurement). A polyvinyl chloride (PVC) disk was utilized to safeguard the sample from being displaced from the diamond crystal after evacuating the instrument. The working principle of FTIR aligns with the Michelson interferometer principle (schematic diagram has been shown in Scheme 2.III.K).[48] This technique is employed to investigate the hydration behaviour of solute molecules and biomolecules.[49-52] It measures the average change in the ultrafast fluctuations of collective water dipole moment as it directly probes the permanent and induced dipole moment of the systems in terms of absorption coefficient.[50]

$$\alpha(\nu) = \left(\frac{1}{4\pi\epsilon_0}\right)\left(\frac{2\pi\nu^2}{3k_BTVcn(\nu)}\right)\int_{-\infty}^{\infty} dt e^{-ivt} < \vec{M}(0).\vec{M}(t) > \qquad (2.III.K1)$$

where $\vec{M}(t)$ is the total dipole moment of the system of volume V, T is the temperature, $n(v)$ is the refractive index and c is the speed of light.



Scheme 2.III.K. Schematic profile of FIR-THz instrument. M_i, M_f, M_v, M_{ci} and BS is the plane mirror, plane fixed mirror, movable mirror (related to the delay stage), concave mirror and beam-splitter respectively.

2.IV. Experimental Analyses.

2.IV.A. *Turbidity measurement.*

UV-Visible absorption spectroscopy is a pivotal tool to verify the turbidity of protein samples.[53] All samples, were individually scanned under the spectrophotometer from 500 nm to 400 nm wavelength range with baseline correction. For temperature dependent turbidity measurement, we have used the temperature controller Peltier attachment. Sample cell temperature tuned by increment of each 5 ˚C per 10 mins. Turbidity calculated from the observed absorbance at 400 nm by,

$$Turbidity = \frac{2.303 \times Absorbance}{Cell\ Pathlength} \qquad (2.IV.A1)$$

2.IV.B. *Steady state fluorescence anisotropy measurement.*

Anisotropy is a parameter which helps to measure the local rigidity of a fluorescence active molecule in various solvent environment. The basic principle behind the measurement is the photo-selective excitation of fluorophores by polarized light. The fluorophores that exposed to polarized light absorb mostly that component of light which is parallel to the absorption transition moment of them. In steady-state anisotropy measurements, emission polarization is adjusted to be parallel or perpendicular to the excitation polarization. Anisotropy is then determined using the following equation,

$$r(t) = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \qquad (2.IV.B1)$$

Where $I_{\parallel}$ and $I_{\perp}$ are intensities measured in parallel or perpendicular orientation of a linear emission polarizer with respect to the excitation polarization.

2.IV.C. *Time resolved fluorescence anisotropy measurement.*

Picosecond time resolved fluorescence spectroscopy (TCSPC) allows to measure the time that a fluorophore takes to relax from the excited state by a radiative pathway which commonly known as fluorophore lifetime. This measurement is fluorophore concentration independent and provides all distinct lifetimes related to the different local environments of the fluorophore. Steady state measurements unable to resolve the emission originated from two different environment. The fluorophore excited with short pulse width laser and repetition rate of the laser pulses is adjusted to be at least four times greater than the lifetime of the fluorophores. The fluorescence decay transients are globally fitted with a single or multi exponential function after convoluting with laser pulse. The decay can be mathematically defined in terms of intensity as,

$$I(t) = \sum_i a_i e^{-\frac{t}{\tau_i}} \qquad (2.IV.C1)$$

where τ_i is the lifetime of the i^{th} species with contribution a_i. The total contribution should be $\sum_i a_i = 1$.

For anisotropy measurements, emissions were collected at 0˚ and 90˚ angles of emission polariser with respect to excitation polariser. The experimental transients were then analyzed and anisotropy values were calculated using the equation,

$$r(t) = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \qquad (2.IV.C2)$$

where G is the grating factor, calculated by longtime tail matching technique. The anisotropy decay profiles were fitted well with the equation,

$$r(t) = r(0)(a_1 \exp\left(-\frac{t}{\tau_1}\right) + a_2 \exp\left(-\frac{t}{\tau_2}\right)) \qquad (2.IV.C3)$$

Here r(0) is the initial anisotropy. τ_1 and τ_2 are the two rotational relaxation times; fast component and slow component, respectively. The a_1 and a_2 are the corresponding amplitudes. The quality of the fitting was evaluated by reduced χ^2 and residual data.

2.IV.D. *Protein secondary structure analysis.*

Secondary structure of protein can be analysed by two renowned techniques; CD spectroscopy and mid-IR spectroscopy. In this thesis work, we have used both methods to determine the change in protein secondary structure. Here mid-IR spectroscopy has been employed to those protein samples which have some limitation to be analysed by CD spectrometer (model number: *JASCO J-815*).

Protein structure analysis by CD spectroscopy: CD absorption spectroscopy consists two wavelength range specification: 190-250 nm as far UV CD and 250-310 nm as near UV CD which can sense the secondary and tertiary structure, respectively. The far UV CD spectra shows two negative bands at 222 nm (due to $n \rightarrow \pi^*$ transition) and 208 nm (due to $\pi \rightarrow \pi^*$ transition) and a positive band for α-helix component; a negative band at 218 nm and a positive band at 196 nm for β-sheets component and random coil display two bands at 212 nm (positive band) and 195 nm (negative band).[54] The near UV CD spectrum defines the presence of aromatic residues (phenylalanine, tyrosine, tryptophan) and disulphide bonds. For quantitative analysis of all structural components % contribution, we use the CDNN software.

Secondary structure analysis by mid-IR spectroscopy: A typical protein mid-IR spectrum often contains nine amide bands, representing vibrational contributions from both the protein backbone and amino acid side chains.[55] Among these bands, amide I (~1600-1700 cm^{-1}) and amide II (~1470-1570 cm^{-1}) are particularly relevant to protein secondary structure. The absorptions associated with C=O stretching and N–H bending are labeled as amide I and amide II, respectively.

Since both C=O and N–H bonds participate in the hydrogen bonding between different components of secondary structure, the positions of both amide I and amide II bands are sensitive to the secondary structure composition of a protein. However, amide I band is considered more effective for accurately quantifying protein secondary structure compared to amide II. Here we get a single broad peak (amide I) which is further deconvoluted by mathematical techniques like Fourier self-deconvolution and second derivatives for the quantitative analysis of protein secondary structure.[56] Deconvolution of the amide I band yields contributions of five major structural components: β-sheet at 1620-1625 cm^{-1}, random coil at 1635-1643 cm^{-1}, α-helix at 1653-1659 cm^{-1}, β-tern at 1667-1676 cm^{-1} and anti-parallel β-sheet at 1682-1690 cm^{-1}. It is noteworthy that all these assignments are depicted as ranges, as exact peak position can vary among proteins due to differences in hydrogen bonding interactions and environmental factors.

2.IV.E. *Microscopy sample specimen preparation and image analysis.*

Atomic force microscopy: A thin layer of mica substrate was cleaved gently. Then a drop of sample was casted on the freshly cleaved mica substrate and kept in a desiccator for 30 mins to ensure complete adsorption on the substrate. Then flow of milli-Q water was passed over the substrate to discard excess sample thickness. Finally, the remaining water was evaporated by a gentle nitrogen gas flow prior to each measurement. The collected morphologies of protein aggregations were then analysed using *Image J* and *Mathematica 13.2* softwares.

Optical (DIC) microscopy: To prepare specimens, first new transparent glass slides and cover slips were taken in separate petri dish with help of a tweezers and dipped into a freshly prepared approx. 90% aqueous ethanol solution for 1 hour (at R.T). Then each one was thoroughly washed with distilled water and dried in warm air. Next, 10-15 µl sample was pipetted out and drop casted on the centre of a glass slide and then a glass cover slip was slowly placed on the drop (at 45˚ angle) using tweezers. The specimen slide then placed on the stage of *Leica DM6M* microscope with a 20x / 0.5 dry objective. All the microscopic images were analysed by *Mathematica 13.2* software.

Field emission scanning electron microscopy: A sample specimen was prepared using properly cleaned silicon wafer substrate. A squared chip like silicon wafer was dipped in milli-Q water and 90% ethanol mixture and sonicated for 10 mins. Then sample was drop casted on that substrate and complete de-wetting was done by keeping it in desiccator overnight. A 1nm thick conducting

gold layer was deposited on the sample to get the information. Then the collected morphologies were analysed by *Image J* and *Mathematica 13.2* softwares.

2.IV.F. *TTDS data analysis.*

By TTDS technique, we measure the THz electric field amplitude ($E_{THz}(t)$) over time. The time delay between the probe and pump beams were adjusted using the inbuilt programming. The measured $E_{THz}(t)$ were then converted to the frequency dependent power by Fourier analysis. We have also calculated the phase shift of the transmitted pulse from the reference signal. The frequency domain THz signal in air is fairly smooth with a high signal to noise ratio and minimal water absorption lines up to ~2.5 THz. This characteristic allows for the extraction of some optical parameters with high precision up to ~2 THz. Those optical parameters are then derived from the frequency-domain data.

First, the absorption coefficient was measured by,

$$\alpha(\nu) = \frac{2}{d} \ln \left(\frac{reference\ signal}{sample\ signal}\right) \qquad (2.IV.F1)$$

where d indicates the sample thickness.

The real part of the complex refractive index then was calculated by measuring the phase of the transmitted pulse (frequency dependent),

$$n(\nu) = \frac{\phi_{sample}(\nu) - \phi_{reference}(\nu)}{2\pi\nu d} \qquad (2.IV.F2)$$

The imaginary part of the complex refractive index,

$$k(\nu) = \frac{c.\alpha(\nu)}{4\pi\nu} \qquad (2.IV.F3)$$

Where the complex refractive index is, $\tilde{n}(\nu) = n(\nu) - i\,k(\nu)$.

From the $n(\nu)$ and $k(\nu)$, the frequency dependent real (ε') and imaginary (ε'') dielectric constants of the samples were extracted for each sample using the following expressions,

$$\varepsilon'(\nu) = n^2(\nu) - k^2(\nu) \qquad (2.IV.F4)$$

$$\varepsilon''(\nu) = 2n(\nu)k(\nu) \qquad (2.IV.F5)$$

We have extracted all these optical parameters using *Mathematica 13.2* software and then the ε' and ε'' profiles were fitted simultaneously with multiple Debye relaxation model,[57, 58]

$$\tilde{\varepsilon}(\nu) = \varepsilon_\infty + \sum_{j=1}^{3} \frac{S_j}{1+i\omega\tau_j} + \frac{\sigma}{i\omega\varepsilon_0} \qquad (2.IV.F6)$$

where τ_j and S_j $(= \varepsilon_j - \varepsilon_{j+1})$ are the relaxation time and strength of the j-th relaxation mode, respectively. ε_∞ is the extrapolated dielectric constant at very high frequency, ε_0 is the permittivity in free space, ω is the angular frequency and σ is the DC conductivity of the solution.

2.IV.G. *Data analysis as measured from FIR-THz absorption spectroscopy.*

All the experiments were done in ATR mode as the *Vertex 70V* instrument has an ATR attachment with a diamond crystal cell surface (refractive index 2.38). The sample chamber was evacuated prior to each measurement using a pump. Each sample was scanned with a typical spectral resolution of 4 cm^{-1} over 128 scans average. OPUS software provides us frequency-dependent ATR intensity profile.

ATR intensity was converted to the FIR-absorbance using the following relation:

$$Abs(\nu) = \frac{Intensity\ in\ ATR\ unit\ (\nu) \times 1000}{wavenumber} \qquad (2.IV.G1)$$

From the absorbance, the frequency-dependent absorption co-efficient α was estimated using Beer Lambert law:

$$\alpha\ (\nu) = \frac{Abs_{sample} - Abs_{reference}}{d_p} \qquad (2.IV.G2)$$

where reference (air) represents the evacuated bare diamond crystal surface and d_p is the penetration depth,[52] defined as,

$$d_p = \frac{\lambda}{2\pi * \sqrt{n_{diamond}^2 * (\sin\theta)^2 - n_{sample}^2}} \qquad (2.IV.G3)$$

where λ is the radiation wavelength, θ is the incident angle (fixed at 45°), $n_{diamond}$ and n_{sample} are the refractive indices of the diamond crystal and the sample, respectively. We have measured n_{sample} for each sample using a Digital refractometer (*RUDOLPH J357*).

The difference absorption coefficient (Δα) was obtained as,

$$\Delta\alpha(\nu) = \alpha_{sample}(\nu) - \alpha_{bulk}(\nu) \quad \text{(2.IV.G4)}$$

The $\Delta\alpha(\nu)$ profiles were fitted using a sum of damped harmonic oscillator (DHO) functions as derived in equation (2.I.D6),

$$\Delta\alpha(\nu) = \sum_{i=1}^{N} \frac{a_i \omega_{0,i} \nu^2}{4\pi^3 \left[\frac{\nu^2 \omega_{0,i}^2}{\pi^2} + \left(\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2} - \nu^2\right)^2\right]} \quad \text{(2.IV.G5)}$$

where a_i, $\omega_{0,i,}$ and $\nu_{d,i}$ elucidate the amplitude, the damping width, and the center frequency of the i^{th}damped harmonic oscillator modes, respectively. The unperturbed center frequency, $\nu_{c,i}$, can be calculated from the relation $\nu_{c,i} = \sqrt{\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2}}$.

2.V. References.

1. A. V. Delgado, F. González-Caballero, R. J. Hunter, L. K. Koopal and J. Lyklema, *J. Colloid Interface Sci.*, 2007, **309**, 194-224.
2. J. N. Israelachvili, *Intermolecular and surface forces*, Academic press, 2011.
3. B. B. Owen, R. C. Miller, C. E. Milner and H. L. Cogan, *J. Phys. Chem.*, 1961, **65**, 2065-2070.
4. S. Bhattacharjee, *Journal of controlled release : official journal of the Controlled Release Society*, 2016, **235**, 337-351.
5. H. Ohshima, *J. Colloid Interface Sci.*, 1994, **168**, 269-271.
6. O. Matsarskaia, F. Roosen-Runge, G. Lotze, J. Möller, A. Mariani, F. Zhang and F. Schreiber, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27214-27225.
7. H. Ohshima, T. W. Healy and L. R. White, *J. Colloid Interface Sci.*, 1982, **90**, 17-26.
8. P. B. I. Yuri Feldman, Alexander Puzenko, Valerică Raicu, Oxford University Press, Oxford, 2015.
9. Y.-S. Lee, *Principles of Terahertz Science and Technology*, Springer, USA, 2009.
10. H. Fröhlich, *Theory of dielectrics: dielectric constant and dielectric loss*, Clarendon Press, 1958.
11. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 12068-12073.
12. B. Born and M. Havenith, 2009, **30**, 1245-1254.
13. V. Vitale, J. Dziedzic, S. M.-M. Dubois, H. Fangohr and C.-K. Skylaris, *J. Chem. Theory Comput.*, 2015, **11**, 3321-3332.
14. C. N. M. Banwell, E. M., *Fundamentals of Molecular Spectroscopy*, McGraw-Hill Publishing Company, 4th edn.
15. T. Dodo, M. Sugawa, E. Nonaka, H. Honda and S. i. Ikawa, *J. Chem. Phys.*, 1995, **102**, 6208-6211.
16. D. Sharma and A. Chandra, *J. Phys. Chem. B*, 2023, **127**, 9323-9335.
17. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087-1092.
18. B. J. Alder and T. E. Wainwright, *J. Chem. Phys.*, 1957, **27**, 1208-1209.
19. A. Rahman, *Phys. Rev.*, 1964, **136**, A405-A411.
20. J. A. McCammon, B. R. Gelin and M. Karplus, *Nature*, 1977, **267**, 585-590.
21. S. Curry, H. Mandelkow, P. Brick and N. Franks, *Nat. Struct. Biol.*, 1998, **5**, 827-835.
22. A. Bujacz, *Acta Crystallogr., Sect. D: Biol. Crystallog.*, 2012, **68**, 1278-1289.

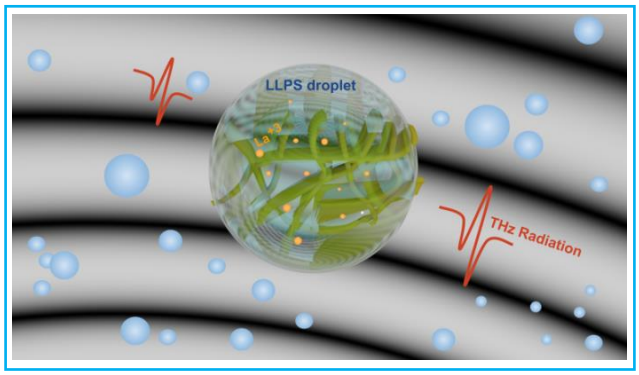
23. L. Medda, M. Monduzzi and A. Salis, *Chem. Commun.*, 2015, **51**, 6663-6666.
24. V. A. Proctor, F. Cunningham and D. Y. Fung, *C R C Critical Reviews in Food Science and Nutrition*, 1988, **26**, 359-395.
25. H. Sugino, T. Nitoda and L. Juneja, in *Hen eggs*, CRC press, 2018, pp. 13-24.
26. P. E. Stein, A. G. W. Leslie, J. T. Finch, W. G. Turnell, P. J. McLaughlin and R. W. Carrell, *Nature*, 1990, **347**, 99-102.
27. R. H. Law, Q. Zhang, S. McGowan, A. M. Buckle, G. A. Silverman, W. Wong, C. J. Rosado, C. G. Langendorf, R. N. Pike, P. I. Bird and J. C. Whisstock, *Genome Biol.*, 2006, **7**, 216.
28. S.-i. Mizushima, T. Simanouti, S. Nagakura, K. Kuratani, M. Tsuboi, H. Baba and O. Fujioka, *J. Am. Chem. Soc.*, 1950, **72**, 3490-3494.
29. E. Salamatova, A. V. Cunha, R. Bloem, S. J. Roeters, S. Woutersen, T. L. Jansen and M. S. Pshenichnikov, *J. Phys. Chem. A*, 2018, **122**, 2468-2478.
30. V. Ali, K. Prakash, S. Kulkarni, A. Ahmad, K. Madhusudan and V. Bhakuni, *Biochemistry*, 1999, **38**, 13635-13642.
31. L. Stryer, *J. Mol. Biol.*, 1965, **13**, 482-495.
32. D. C. Turner and L. Brand, *Biochemistry*, 1968, **7**, 3381-3390.
33. E. Bismuto, E. Gratton and D. C. Lamb, *Biophys. J.*, 2001, **81**, 3510-3521.
34. H. Cinar and R. Winter, *Sci. Rep.*, 2020, **10**, 17245.
35. R. K. Mitra, S. S. Sinha and S. K. Pal, *Langmuir*, 2007, **23**, 10224-10229.
36. C. M. Hoo, N. Starostin, P. West and M. L. Mecartney, *J. Nanopart. Res.*, 2008, **10**, 89-96.
37. J. R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer, 2006.
38. G. D. Fasman, *Circular Dichroism and the Conformational Analysis of Biomolecules*, Plenum Press, NewYork, 1996.
39. N. J. Greenfield, *Nature Protocols*, 2006, **1**, 2876–2890.
40. Y. F. Dufrêne, T. Ando, R. Garcia, D. Alsteens, D. Martinez-Martin, A. Engel, C. Gerber and D. J. Müller, *Nat. Nanotechnol*, 2017, **12**, 295-307.
41. G. Lamour, C. K. Yip, H. Li and J. Gsponer, *ACS Nano*, 2014, **8**, 3851-3861.
42. R. W. Carpick and M. Salmeron, *Chem. Rev.*, 1997, **97**, 1163-1194.
43. N. Gavara, *Microsc. Res. Tech.*, 2017, **80**, 75-84.
44. S. Ray, N. Singh, R. Kumar, K. Patel, S. Pandey, D. Datta, J. Mahato, R. Panigrahi, A. Navalkar and S. Mehra, *Nat. Chem.*, 2020, **12**, 705-716.
45. D. B. Murphy, *Fundamentals of Light Microscopy and Electronic Imaging*, A John Willey & Sons, NewYork, 2001.
46. G. Hübschen, I. Altpeter, R. Tschuncky and H.-G. Herrmann, *Materials characterization using nondestructive evaluation (NDE) methods*, Woodhead publishing, 2016.
47. A. C. Reimschuessel, *J. Chem. Educ.*, 1972, **49**, A413.
48. V. Saptari, *Fourier transform spectroscopy instrumentation engineering*, SPIE Optical Engineering Press Bellingham Washington, DC, 2003.
49. F. Sebastiani, C. Y. Ma, S. Funke, A. Bäumer, D. Decka, C. Hoberg, A. Esser, H. Forbert, G. Schwaab and D. Marx, *Angew. Chem. Int. Ed.*, 2021, **60**, 3768-3772.
50. V. Conti Nibali and M. Havenith, *J. Am. Chem. Soc.*, 2014, **136**, 12800-12807.
51. G. Schwaab, F. Sebastiani and M. Havenith, *Angew. Chem. Intl. Ed.*, 2019, **58**, 3000-3013.
52. J. Ahlers, E. M. Adams, V. Bader, S. Pezzotti, K. F. Winklhofer, J. Tatzelt and M. Havenith, *Biophys. J.*, 2021, **120**, 1266-1275.
53. E. Layne, in *Methods in Enzymology*, Academic Press, 1957, vol. 3, pp. 447-454.
54. N. J. Greenfield, *Nature Protocols*, 2006, **1**, 2876-2890.
55. A. Elliott and E. J. Ambrose, *Nature*, 1950, **165**, 921-922.
56. A. Barth, *Biochim. Biophys. Acta.*, 2007, **1767**, 1073-1101.

57. N. Nandi and B. Bagchi, *J. Phys. Chem. B*, 1997, **101**, 10954-10961.
58. J. T. Kindt and C. A. Schmuttenmaer, *J. Phys. Chem.*, 1996, **100**, 10373-10379.

Chapter 3

3. Trivalent cation-induced phase separation in proteins: ion specific contribution in hydration also counts

Multivalent (specifically trivalent) metal ions are known to induce microscopic phase separation (commonly termed as liquid–liquid phase separation (LLPS)) in negatively charged globular proteins even at ambient temperatures, the process being mostly driven by protein charge neutralization followed by aggregation. Recent simulation studies have revealed that such self-aggregation of proteins is entropy driven; however, it is associated with a solvation effect, which could as well be different from the usual notion of hydrophobic hydration. In this contribution we have experimentally probed the explicit change in hydration associated with ion-induced LLPS formation of a globular protein bovine serum albumin (BSA) at ambient temperature using FIR-THz FTIR spectroscopy (50–750 cm^{-1}; 1.5–22.5 THz). We have used ions of different charges: Na^{+}, K^{+}, Ca^{2+}, Mg^{2+}, La^{3+}, Y^{3+}, Ho^{3+} and Al^{3+}. We found that all the trivalent ions induce LLPS; the formation of large aggregates has been evidenced from dynamic light scattering (DLS) measurements, but without perturbing the protein structure as confirmed from circular dichroism (CD) measurements. From the frequency dependent absorption coefficient ($\alpha(\nu)$) measurements in the THz frequency domain we estimate the various stretching/vibrational modes of water and we found that ions, forming LLPS, produce definite perturbation in the overall hydration, the extent of which is ion specific, invoking the definite role of hydrophilic (electrostatic) hydration of ions in the observed LLPS process.



3.I. Introduction.

Charged colloidal particles, when placed in solutions containing multi-valent ions, could offer unique features including reversal of their electrophoretic mobility.[1] Accordingly, ions have been observed to play a key role in many biophysical processes involving proteins.[2-4] The effects include reentrant condensation of proteins in the presence of ions,[5] ion concentration dependent liquid–

liquid phase separation (LLPS),[6] lower critical solution temperature,[7] etc. Among such interactions, LLPS is of special interest, in which liquid–liquid micro-phases do coexist, analogues to liquid–gas transition in van der Waals gases, stabilized by short range inter-particle attractions.[8] Since the pioneering report by Ishimoto and Tanaka,[9] there have been several attempts to describe protein phase behaviour in the presence of mono- and multi-valent salts.[7, 10-12] In these studies, salt concentration dependency on protein phase separation has been represented by phase diagrams, which revealed that at a particular protein concentration and between two critical salt concentrations, metastable LLPS states are formed at ambient as well as at high temperatures. The thermodynamic aspect of such phase separation follows Flory–Huggin's theory[13, 14] according to which (de)mixing of two macromolecular ensembles could be tuned by the free energy of mixing, which in turn depends upon both enthalpic and entropic contributions. It has now been established that entropy offers the key driving force during this phenomenon[7] and the total entropy cost involves contributions from the solvation processes of both the protein and the cation.[15]

It has mostly been observed that cations of heavy metals, specifically lanthanides, show very rich phase behaviour with negatively charged globular proteins.[10, 16, 17] The heavy metal cations induce short range interactions between proteins forming cation-bridges, which eventually results in phenomena like protein crystallization, cluster formation, LLPS and reentrant condensation, etc.[5, 6, 8, 18-20] It has also been identified that for salts like $FeCl_3$ and $AlCl_3$, salt hydration and consequently the pH of the solution influence the protein's phase behaviour.[12] A recent simulation result[15] has shown that a fraction of tightly bound water molecules is released from the protein surface during a cation-mediated LLPS process, which provides an entropic impetus towards the process. While such a claim provides a fundamental understanding of the molecular mechanism responsible for the LLPS process, an experimental validation of the same is equally important. In the present contribution we aim to experimentally investigate hydration alteration during the ion-mediated LLPS process in a model protein bovine serum albumin (BSA) using the THz-FTIR (Tera hertz-Fourier transform infrared) spectroscopy technique (50–750 cm^{-1}; 1.5–22.5 THz).

THz absorption spectroscopy has emerged as a potential tool to probe the changes in collective (sub)ps hydrogen bond dynamics and to detect low-frequency vibrational modes in liquids.[21] Water acts as an exceptional THz absorber as the ultrafast fluctuations of water collective

dipole moments show their imprint in this frequency window.[22] Any change of water network structure at the biomolecular surface offers contrasting features compared to the bulk water absorption profile and such measurements provide a label-free estimation of biomolecular-solvation.[23] While biomolecular hydration is a fascinating field of research, hydration of heavy (transition and lanthanide) metal ions also offers distinct features compared to main group elements owing to ligand-field stabilization and higher atomic mass.[24-26] Such results are reflected in the absorption coefficient profiles with characteristic rattling modes of strongly hydrated metal ions within their hydration cage.[27] With the notion that metal ions substantially perturb the hydrogen bonded network structure of bulk water, it is deemed interesting to understand how such a change modulates the overall hydration of negatively charged globular proteins during salt-mediated LLPS processes.

We investigate the effect of multivalent salts ($AlCl_3$, $LaCl_3$, YCl_3, and $HoCl_3$) along with monovalent (NaCl, KCl) and divalent ($MgCl_2$, $CaCl_2$) salts on the stability and phase separation process of bovine serum albumin (BSA). The salt (C_S) and the protein (C_P) concentration were fixed at 14 mM and 40 mg ml^{-1} (=0.6 mM) to maintain the overall C_S : C_P ratio of 23.33. The choice of this ratio is based on the fact that earlier studies[17] have suggested that at this particular C_S : C_P ratio, multivalent ions do induce BSA to undergo LLPS. We visually detect a less transparent solution of BSA in the presence of the tri-valent salts compared to the monovalent and divalent ones. We use dynamic light scattering (DLS) technique to determine the size of the protein in the presence of these salts. The surface charge density of BSA is determined using zeta potential measurements. The secondary structure of the protein is estimated using circular dichroism measurements. Finally, protein solvation in the absence and in the presence of these salts is probed using THz-FTIR measurements.

3.II. Materials and Methods.

Sodium chloride (NaCl), potassium chloride (KCl), magnesium chloride ($MgCl_2$), calcium chloride ($CaCl_2$), aluminium chloride ($AlCl_3$), lanthanum chloride ($LaCl_3$), yttrium chloride (YCl_3), holmium chloride ($HoCl_3$) and bovine serum albumin (BSA) were purchased from Sigma Aldrich with ≥99% purity and were used without further refinement. The pH of the samples was maintained at 7.4 using freshly prepared phosphate buffer (PBS) as the solvent throughout the experiments. PBS buffer was prepared by dissolving acid NaH_2PO_4 and salt Na_2HPO_4 in water of

10 mM concentration. All the measurements were carried out at 25° C. The details of all the used instruments are described in the section 2.III and the experimental analysis details are also mentioned prior in the section 2.IV.

3.III. Results and Discussions.

Visual analysis of the protein solutions: We prepared a series of BSA solutions in the presence of monovalent alkali metal chlorides: NaCl and KCl, divalent alkaline earth metal chlorides: $MgCl_2$ and $CaCl_2$, a trivalent transition metal chloride: YCl_3, trivalent Lanthanide chlorides: $LaCl_3$ and $HoCl_3$, and a trivalent metal chloride: $AlCl_3$ maintaining the ratio C_S : C_P 23.33. Photographs of these solutions (after ensuring complete equilibrium) are depicted in figure 3.III.1. Visual inspection of these solutions concludes that protein solutions containing trivalent salts are less transparent and bluish-translucent[17] in appearance compared to the mono- and di-valent salt solutions.

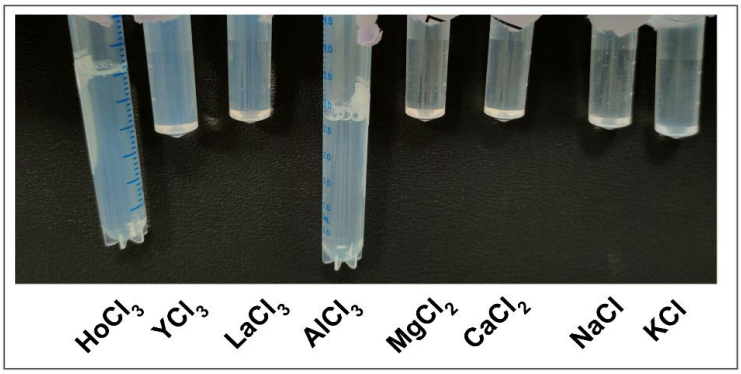


Figure 3.III.1. Images of different phase of samples, containing 40 mg ml^{-1} (0.6 mM) of BSA and Chloride salts with different valences. In every sample, the protein concentration (C_p) is 0.6 mM and the salt concentration (C_s) is 14 mM to maintain the C_s : C_p ratio of 23.33. Images show that the protein solutions get turbid in presence of trivalent salts compared to mono and divalent salts at room temperature.

DLS measurements: Such bluish appearance of these protein solutions might indicate the formation of larger particles producing higher scattering. To ensure this we measure the size of the protein in the presence of these salts using the DLS technique. Native BSA (0.6 mM) shows a hydrodynamic diameter of 4.9 nm (figure 3.III.2a).[28] This confirms that at this concentration BSA does not self-aggregate in the buffer solution. The size remains almost the same in the presence of mono- and di-valent salts (figure 3.III.2b). In the presence of YCl_3, however, we observe the appearance of

larger sized particles with hydrodynamic diameter of ~60 nm (figure 3.III.2a). We also note the formation of such larger particles (30–90 nm diameters) in the case of all the trivalent salts (figure 3.III.2b). Formation of these larger particles is responsible for the observed bluish appearance of the protein solutions (figure 3.III.1). DLS measurements thus reveal that the size of the protein increases many-fold in the presence of trivalent metal ions.

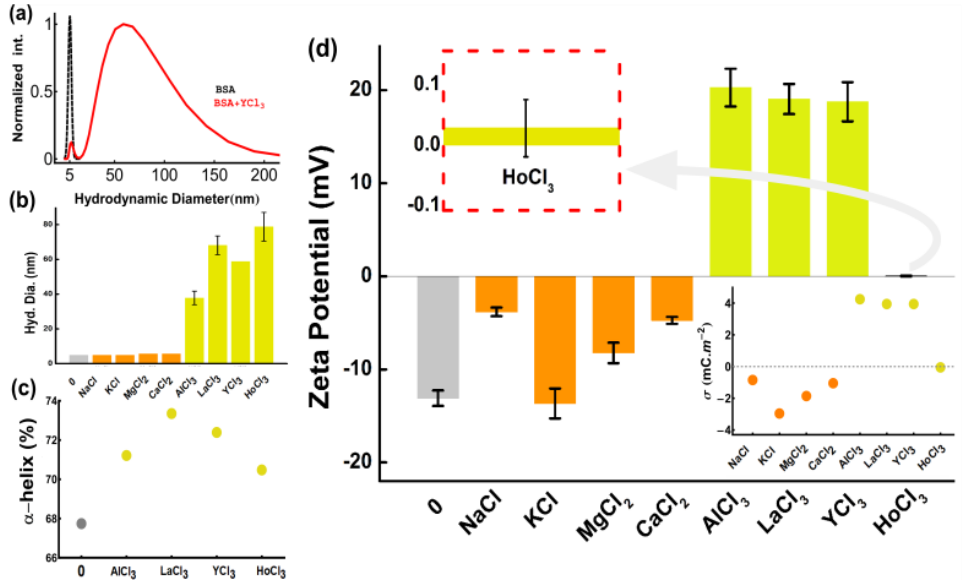


Figure 3.III.2. (a) Representative dynamic light scattering profiles of BSA (black dashed line) and BSA in presence of 14 mM YCl_3 (Red solid line). (b) Hydrodynamic diameter of different BSA-Salts samples. BSA and salts concentrations are fixed at 40 mg ml^{-1} (0.6 mM) and 14 mM respectively. The grey bar with 0 label indicates bare BSA solution, the orange bars specify BSA in presence of mono- and divalent salts, and the yellowish-green bars represent trivalent salts-protein samples. (c) α-helix abundance (in percentage) of BSA in presence of trivalent salts. (d) Zeta Potential of salt containing BSA solution. Zeta potential value of $HoCl_3$-BSA solution is shown in the red box. The corresponding surface charge density value, as estimated from zeta potential, is depicted in the inset. The error bars signify standard deviations.

Secondary structural analysis of BSA in LLPS: Next, it is important to check whether the protein itself undergoes structural perturbation in the presence of these metal ions. We determine the secondary structures of BSA in the presence of these ions using CD spectroscopy (figure 3.III.3.). For the CD measurements C_P is kept fixed at 10 mM and C_S at 23.33 mM to retain the C_S : C_P ratio

of 23.33. Earlier studies suggest that the protein does not undergo significant structural perturbation in the presence of monovalent ions.[29] The present measurements conclude that the secondary structure of BSA does not get noticeably perturbed in the presence of trivalent salts as well. BSA is an α-helix rich (~68%) globular protein. It is observed that the abundance of α -helix increases further in the presence of these salts (figure 3.III.2c), indicating a further stabilization of the compact globular structure of BSA, the effect being found to be maximum in La^{3+}. CD measurements thus confirm that the salts do not bring about structural perturbation in the protein, which could have been anticipated as a possible reason for the increased size of the droplets.

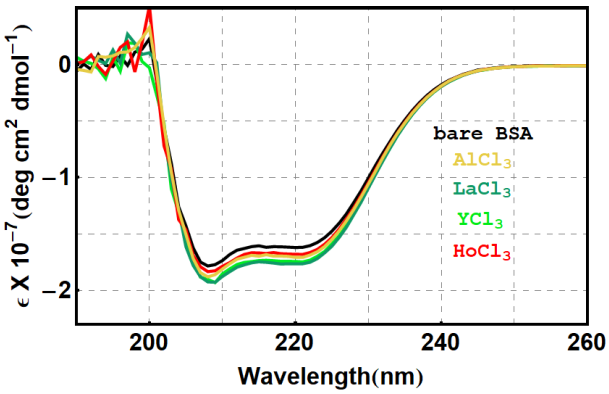


Figure 3.III.3. Representative CD signals (analysed in molar ellipticity, ε) of bare BSA protein, and in presence of the mentioned trivalent salts.

Role of protein surface potential: The protein surface is generally charged due to the presence of charged amino acids at the surface and therefore, electrostatic interaction plays a crucial role securing protein stability and its functionality.[30, 31] Anisotropic charge interaction on the protein surface could affect the phase behaviour of the protein solution.[32] BSA (isoelectric point, pI ~ 4.6)[33] is negatively charged in PBS buffer (pH 7.4). Ion distribution around the negatively charged globular protein produces an electrical double layer (EDL) which dissects into two layers: the Stern layer and the diffuse layer. Metal ions are distributed around the charged protein surface, such and distribution could theoretically be described following Poisson-Boltzmann (PB) and/or Derjaguin-Landau-Verwey-Overbeek (DLVO theory).[34] The presence of multivalent ions could as well induce charge inversion or overcompensation of surface charge of proteins by the binding of significant numbers of counter-ions at the surface.[35] A distinct phase behaviour, called reentrant

condensation (RC),[36] results from such charge inversion of globular, negatively charged proteins in the presence of multivalent cations at specific C_S : C_P ratios. We measure zeta-potential (ZP) and consequently the surface charge density (refer to the section 2.I.A) of BSA in the presence of the metal ions (figure 3.III.2d). It is observed that in the presence of mono- and di-valent salts BSA retains its intrinsic negative surface charge. However, in the presence of Al^{3+}, La^{3+} and Y^{3+} we observe a drastic inversion in the protein surface charge as it becomes positively charged. The effect is found to be equivalent in all these three ions and corroborates the previous studies[12] at this particular C_S : C_P ratio. In Ho^{3+}, however, the change is only marginal and evidently at that point the protein-ion electrostatic contributions to the binding seem to be minimized.[7] DLS and zeta potential measurements thus conclude that trivalent ions bring about remarkable changes both in the phase behaviour and the electrostatics in the protein solution, the effect being comparable in La^{3+} and Y^{3+}, while in Al^{3+} and Ho^{3+}, it is contrasting. It therefore appears interesting to investigate to what extent such changes are correlated with the corresponding change in the protein as well as in salt hydration.

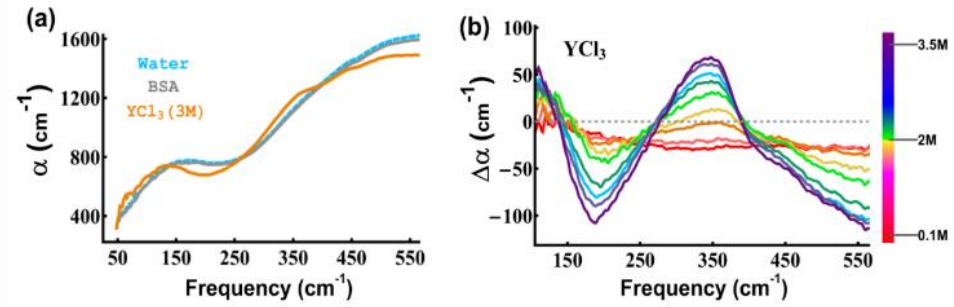


Figure 3.III.4. (a) Representative frequency-dependent absorption coefficient (α(ν)) of water, BSA (0.6 mM), and 3M YCl_3. (b) The difference of absorption coefficients Δα (= α_{sample} - α_{bulk}) at different YCl_3 concentration (0.1 to 3.5M) as a function of frequency.

Hydration study: We extensively examine the hydration structure of BSA in the absence and in the presence of the salts in the FIR-THz region (50–750 cm^{-1}; 1.5–22.5 THz). We measure the frequency dependent absorption coefficient, $\alpha(\nu)$, in this frequency window and representative profiles for bulk water, BSA in buffer and Y^{3+} in water are depicted in figure 3.III.4a. As observed from the figure, water strongly absorbs in this frequency region as different intermolecular hydrogen bond vibrating modes show their imprints in this frequency window.[22] During ultrafast

(sub-ps) breaking and making of hydrogen bonds in bulk water, its collective dipole moment fluctuates and a Fourier transform of such autocorrelation yields the absorption coefficient $\alpha(\nu)$ profile. Thus, experimental determination of $\alpha(\nu)$ is a label-free explicit description of the collective vibrational motion of water. BSA in buffer solution produces a diminished $\alpha(\nu)$ profile and alters these water modes as poor THz absorbing protein molecules replace high absorbing water, this key concept follows the well-known phenomenon of a ''Terahertz defect''.[37] A lowering of the $\alpha(\nu)$ profile also manifests the fact that the protein hydration water is more structured than the bulk water.[38] It is evident from figure 3.III.4a that Y^{3+} also induces noticeable perturbation in the $\alpha(\nu)$ profile of water. To understand such change quantitatively, we calculate the difference absorption profile $\Delta\alpha(\nu) = \alpha_{salt} - \alpha_{bulk}$ (figure 3.III.4b). Unlike in proteins, we observe the appearance of a distinct positive band, which could be explained by the rattling motions of the ions (i.e. ion oscillations in its solvation cage).[39] The $\Delta\alpha(\nu)$ profile of ion rattling motion in low frequency (<200 cm^{-1}) can be overlapped with water absorption and above this range, the increase in absorption frequency is extended to the librational motion of water. Extension of the rattling motion to >200 cm^{-1} in the presence of several cations is related to a specific cationic resonance (M^{n+}-O) stretching mode, mostly evidenced in lanthanides (III).[40] The ion hydration cage is considered as a 3D harmonic potential and the oscillation of ions in this potential well depends on the ion mass and force constant (curvature of the potential well), which eventually manifests H-bond strength within the H-bond water network cage.[41]

To reach into a more quantitative insight, we deconvolute all spectra into a multiple damped harmonic oscillator model,[25, 41]

$$\Delta\alpha(\nu) = \sum_{i=1}^{N} \frac{a_i \omega_{0,i} \nu^2}{4\pi^3 \left[\frac{\nu^2 \omega_{0,i}^2}{\pi^2} + \left(\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2} - \nu^2 \right)^2 \right]}$$

where a_i, $\omega_{0,i,}$ and $\nu_{d,i}$ elucidate the amplitude, the damping width, and the center frequency of the i^{th}damped harmonic oscillator modes, respectively. Subsequent modes can be extracted from the unperturbed center frequency, $\nu_{c,i}$, obtained by, $\nu_{c,i} = \sqrt{\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2}}$.

From the deconvolution of the bulk water spectrum (figure 3.III.5), two intense bands are observed: at ~129 cm^{-1} associated with the intermolecular H-bond stretching and at ~569 cm^{-1} assigned to the intermolecular librational motion of water molecules.[42] We then deconvolute the difference spectra of the protein solution (figure 3.III.6a) in which we observe negative contributions of two FIR active modes.

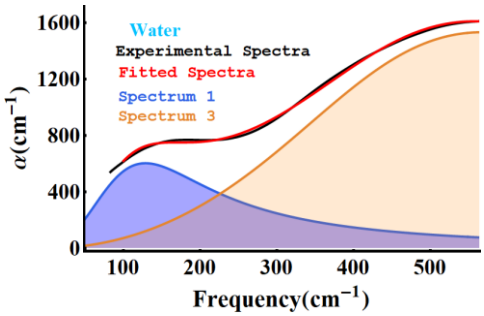


Figure 3.III.5. Deconvoluted frequency dependent absorption coefficient, α(ν) spectra of bulk water. Two damped modes are observed: HB stretching mode (~129 cm^{-1}), and librational mode (~569 cm^{-1}).

The intermolecular hydrogen bond stretching frequency is centred at ~73 cm^{-1} and that of the librational mode is located at ~474 cm^{-1}. The observed red shift in the peak frequency ascertains weakening of H-bonds and restriction in the librational mode. Since we are interested in extracting the change in protein hydration, we obtain $\Delta\alpha(\nu) = \alpha_{protein+salt+water} - \alpha_{salt+water}$ for the protein solutions in the presence of the salts. As we deconvolute the $\Delta\alpha(\nu)$ profiles we observe three bands with the corresponding centre frequencies appearing at ~190-204 cm^{-1}, 333-375 cm^{-1} and 568-615 cm^{-1} (Table 3.III.1). The first and the third bands are the intermolecular H-bond stretching vibrations and the constrained librational motion of water molecules, as discussed earlier.[43] The band with a positive contribution appearing in the middle emanates from the rattling modes of ions along with their hydration cages and is found to be ion specific. It is important to recollect here that the trivalent salts induce the formation of larger aggregates (LLPS) with proteins and the observed changes in the protein hydration are expected to correlate that. We plot the central frequencies of the three fitted modes in figure 3.III.6(d-f) for different metal ions.

For mono- and di-valent salts, the ion rattling modes in the presence of BSA are centered in the 300–400 cm^{-1} region, which is mildly red shifted (~30–50 cm^{-1}) compared to that observed for these ions in aqueous solutions. It should also be remembered here that the salts used in protein solutions are low in concentration (14 mM) compared to the concentration range (~1 M) studied in aqueous salt solutions. It is intriguing to note that despite this low salt concentration, the rattling peak, specific to ion hydration,[39] is prominent. In bare La^{3+} and Y^{3+}, this band appears at 315 and 340 cm^{-1}, and interestingly, in the presence of proteins, they suffer a distinct blue shift to appear at

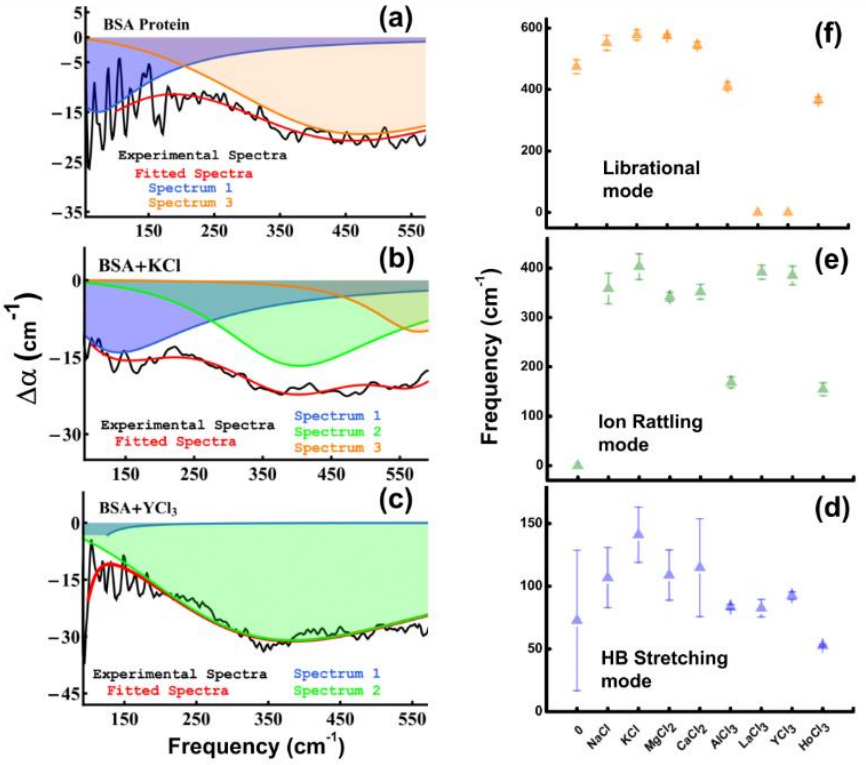


Figure 3.III.6. Dissection of experimental Δα(ν) spectra fitted using a damped harmonic oscillator model for (a) pure BSA, (b) BSA-KCl and (c) BSA-YCl_3 solutions. The black solid lines represent experimental data and the red lines stand for the overall fit. Three damped modes are observed: HB stretching mode (70-140 cm^{-1}), rattling motion of ions (380-410 cm^{-1}), and librational mode (470-580 cm^{-1}). The corresponding peak frequencies for the different salts are presented in (d-f). '0' stands for the bare protein solution. The error bar signifies the standard deviations for the respective system.

~390 cm^{-1}. This signifies that the cationic resonance (M^{n+}–O) stretching is the major contributor to this mode. The observed blue shift thus correlates with the strengthening of the La^{+3}–O and Y^{+3}–O stretching modes within the hydration cage of these ions. Protein–ion binding during LLPS formation leads to extended confinement of these ions (La^{3+} and Y^{3+}) in their hydration cage. On the contrary, the other two trivalent ions Ho^{3+} (327 cm^{-1}) and Al^{3+} (258 cm^{-1}) exhibit drastic red shift to appear at ~154 and 168 cm^{-1} in the presence of the protein. Noticeably, these ions also induce LLPS in the studied protein–salt composition. It can be noticed here that this specific vibrational mode is ion specific and any change in it is due to the modified ion-solvation as the ions interact with proteins. It is evident that the effect of the same protein on the salts is different and it does not depend solely on the charge of the ions as contrasting solvation features are observed between La^{3+}, Y^{3+} and Ho^{3+} and Al^{3+}.

We now focus on the solvation of the protein. As discussed earlier, the HB stretching mode of bulk water suffers a pronounced red shift in the presence of BSA to appear at ~73 cm^{-1}. As mono- and di-valent salts are added, a comparative blue shift in the mode is observed to appear at ~120 cm^{-1}, the peak frequency being similar to that observed in bulk water. In the presence of trivalent salts, however, the blue shift is comparatively less (~10–15 cm^{-1}), and most intriguingly, in Ho^{3+}, it is even more red shifted to appear at ~53 cm^{-1}. Focusing on the librational band, in the presence of BSA, the bond is significantly red shifted (~90 cm^{-1}). In the presence of mono- and di-valent salts, it more or less regains a value comparable to that of bulk water. A strikingly different behaviour is yet again observed in trivalent ions. In Al^{3+} and Ho^{3+}, it is even more red shifted than in the protein solution itself. The most captivating result is obtained in La^{3+} and Y^{3+} in which we do not recover any contribution of the librational mode. These ions do induce LLPS and at ambient conditions, the microscopically separated protein-rich phase contains more tetrahedrally coordinated water molecules, which perhaps inhibits its librational motion.[44] Unlike these two salts, during LLPS formation, $AlCl_3$ and $HoCl_3$, however, could not dissipate the librational band and the band eventually shifts towards a lower frequency.

Measurement of the absorption co-efficient (α) in the THz frequency region can provide useful and explicit information on solute solvation.[45] As observed, water offers a huge absorption in this frequency window (figure 3.III.4a), and the addition of solutes (here, protein and salts) modifies the pristine profile. The experimentally observed $\alpha(\nu)$ profile is expected to be explained

as a summation of contributions from each component: $\alpha(\nu) = \sum_i \phi_i \alpha_i(\nu)$ where ϕ_i is the volume fraction of the i[th] component. However, it has been concluded that mere consideration of the bulk components does not reproduce the experimental $\alpha(\nu)$ profile, and one needs to consider the contribution from the solute hydration layer also.[45] In the present case, while describing the total absorption profile of BSA solution, one needs to consider the absorption contribution from the protein hydration shell also. The difference absorbance $\Delta\alpha_p(\nu)$, therefore takes the form:

$$\Delta\alpha_p(\nu) = \phi_p\alpha_p(\nu) + \phi_{ph}\alpha_{ph}(\nu) + (\phi_{bulk} - 1)\,\alpha_{bulk}(\nu) \qquad (3.1)$$

where ϕ_{ph} is the volume fraction of the protein hydration sheath and α_{ph} is the corresponding absorption coefficient. Keeping in mind that the absorption co-efficient and volume fraction are always positive quantities; it is evident from this equation that an observed negative profile of $\Delta\alpha_p(\nu)$ (figure 3.III.6a) emanates only if the negative value of the third term overwhelms the first two positive contributions. Note that in this equation, we have two unknown terms ϕ_{ph} and $\alpha_{ph}(\nu)$, and previous studies suggest that $\alpha_{ph}(\nu)$ is a function of the solute type and also distance from the surface.[46] So, it is not possible to explicitly identify the individual contribution of either ϕ_{ph} or α_{ph}, however, it can unambiguously be stated that the product of these two terms has a lower value than the third term. As salt(s) is added to the protein solution, an additional term due to salt hydration (α_{sh}) also needs to be accommodated, and $\Delta\alpha_{p+s}(\nu)$ thus takes the form:

$$\Delta\alpha_{p+s}(\nu) = \phi'_p\alpha_p(\nu) + \phi'_{ph}\alpha_{ph}(\nu) + \phi'_s\alpha_s(\nu) + \phi'_{sh}\alpha_{sh}(\nu) + (\phi'_{bulk} - 1)\alpha_{bulk}(\nu) \qquad (3.2)$$

To quantify the extent of the ion-induced perturbation in protein hydration we estimate the parameter $\Delta\Delta\alpha(\nu)(= \Delta\alpha_{p+s}(\nu) - \Delta\alpha_p(\nu))$ for different salt-BSA pairs (figure 3.III.7). We observe that $\Delta\Delta\alpha(\nu)$ is sensitive towards the salt type. Using eqn (3.1) and (3.2) the analytical form of $\Delta\Delta\alpha(\nu)$ could approximately be expressed as (keeping in mind that both BSA and the salts absorbs much less in this frequency window):

$$\Delta\Delta\alpha(\nu) = (\phi'_{ph} - \phi_{ph})\alpha_{ph}(\nu) + \phi'_{sh}\alpha_{sh}(\nu) + (\phi'_{bulk} - \phi_{bulk})\alpha_{bulk}(\nu) \qquad (3.3)$$

We now try to explain the experimentally obtained results (figure 3.III.7) in the light of this equation remembering that the second term always contributes a positive share and thus the sign of $\Delta\Delta\alpha(\nu)$ is determined by the two terminal difference terms.

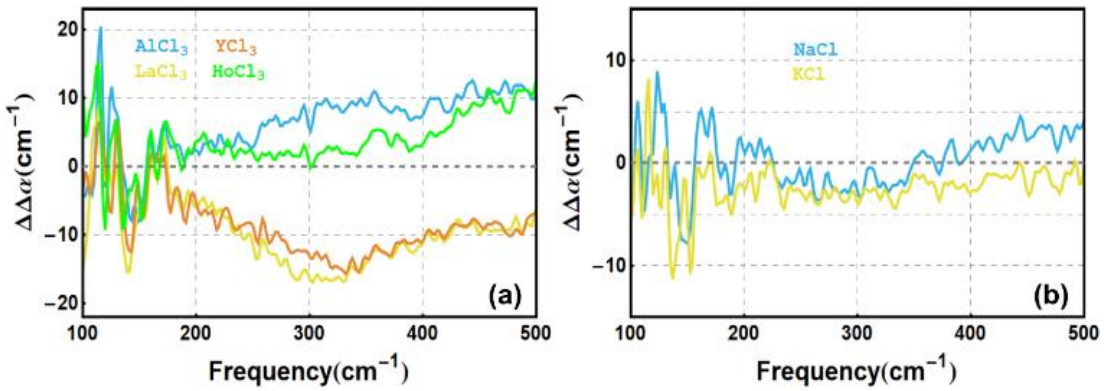


Figure 3.III.7. Frequency dependent change in difference absorption coefficients ΔΔα(ν) (= $\Delta\alpha_{(p+s)}(\nu)$- $\Delta\alpha_{p}(\nu)$) (a) of trivalent salts, and (b) of monovalent salts.

Extracting thermodynamic parameters, specifically solvation free energy ($\Delta G_{hyd} = \Delta H_{hyd} - T\Delta S_{hyd}$), provides an understanding of protein–ion binding processes. In a recent study Ahlers et al.[44] have proposed that minimization of protein solvation plays a decisive force during LLPS. Studies on elastin like polypeptides and polymers have also revealed that LLPS is entropically driven by the loss of hydration structure around the macromolecules during LLPS.[47] Our group has also recently described how LLPS is stabilized or inhibited depending on the change of protein hydration and consequently on the free energy.[42] Thermodynamic parameters related to protein-multivalent cation binding have mostly been quantitatively understood following a 2PT (Two phase thermodynamic) model.[15] According to this theory, the ‘‘hydration effect’’ associates partial de-solvation of both the protein and the ion(s) during an ion-mediated LLPS process. In the 2PT method, protein–ion binding entropy can be estimated for the collective vibrational modes of the non-interacting harmonic oscillators in terms of density of states (DOS) of fluids.

While the analytically calculated DOS relates to the velocity autocorrelation function, the experimentally determined THz absorption coefficient $\alpha(\nu)$ is an outcome of the total dipole moment autocorrelation function.[22] The analytical relation between the velocity autocorrelation function and total dipole moment autocorrelation function is still an uncovered field, which restricts us from quantitatively disclosing the change in protein–ion binding entropy in terms of the experimentally obtained $\alpha(\nu)$ values. However, we could appear at a qualitative interpretation

of the change in the protein hydration entropy ($\Delta S_{hyd} = \Delta S_{p,s} - \Delta S_{p,b}$), in terms of the observed $\Delta\Delta\alpha(\nu)$ profiles.

We first consider that the monovalent salts do not produce any perturbation and/or modification on protein solubility, and we observe that for these ions $\Delta\Delta\alpha(\nu) \to 0$ (figure 3.III.7b) and correspondingly the hydration entropy gain $\Delta S_{hyd} \to 0$ forming no LLPS. A previous simulation study[15] suggests that these ions do accommodate at the protein surface, which compromises their hydration structure (figure 3.III.6b of the ref. 15) during binding. It also has been evidenced from the red shift of the monovalent ions rattling motion from the bulk (~385 cm^{-1}) to the protein solution (~358 cm^{-1}), which supports the weakening of the ion hydration. This effect should be felt in the term $\phi'_{sh}\alpha_{sh}(\nu)$ in eqn (3.3). However, since the proteins are not aggregating, the bulk water contribution does not change ($\phi'_{bulk} \approx \phi_{bulk}$) and so any change in the protein solvation is compensated by the $\phi'_{sh}\alpha_{sh}(\nu)$ term to produce $\Delta\Delta\alpha(\nu) \to 0$. As expected, a similar observation ($\Delta\Delta\alpha(\nu) \to 0$) is also obtained in the case of the bivalent salts.

We now focus on the trivalent salts, all of them inducing LLPS. As mentioned earlier, these salts also offer distinct behaviour depending on their position in the periodic table. For $LaCl_3$ and YCl_3 we observe $\Delta\Delta\alpha(\nu) < 0$ (figure 3.III.7a). These ions do induce LLPS following BSA charge reversal and this thereby confirms their strong interaction with the protein surface. Note that we could not recover any librational model for these ions in the presence of BSA, and perhaps an otherwise wide ion rattling mode overlaps with the librational mode; a possible rationale for the same could be strong confinement of the ions within the LLPS hydration realm. The formation of LLPS defines the release of protein solvation in the form of bulk water implying a positive $\Delta\phi_{bulk}$contribution. It therefore appears that the decrease in protein hydration $\Delta\phi_{ph}$ more than compensates for the two positive contributions (eqn (3.3)) to produce an overall negative $\Delta\Delta\alpha(\nu)$. Noticeably the salt hydration contribution is not sufficient to compensate that.

Finally, we discuss the unusually different behaviour of the two ions: Ho^{3+} and Al^{3+}. We have summarized the key results in Table 3.III.2. For Ho we interestingly observe a near-zero profile of $\Delta\Delta\alpha(\nu)$. For Al^{3+} we even notice that $\Delta\Delta\alpha(\nu)$ is positive. Since these ions also induce LLPS, we therefore can safely conclude a release of bulk water into the solution, which implies

$\Delta\phi_{bulk} > 0$ and $\Delta\phi_{ph} < 0$, entropically favouring the process. A positive value of $\Delta\Delta\alpha(\nu)$ thus can be achieved either by $\Delta\phi_{bulk} > \Delta\phi_{ph}$ or with a high positive contribution of $\phi'_{sh}\alpha_{sh}(\nu)$ or both. It is interesting to note here that both Ho^{3+} and Al^{3+} produce varied solvation behaviour. A previous EXAFS (extended X-ray absorption fine structure spectroscopy) study provides varying hydration structures of different ions of the lanthanide subgroup (Ln(III)) depending on their hydration geometry.[48] The square antiprismatic (8 coordinated) distribution of hydration water molecules around Ho^{3+} (heavy Ln(III)) ions have shorter mean Ho–O bond length than lighter Ln(III) ions (La^{3+})) leads to weakening of the water exchange rate. This argument also corroborates with the marginal change in the protein surface potential in the presence of Ho^{3+} (figure 3.III.2d) and also the noticed red shift in its rattling mode in the protein solution, consequently implying a high positive contribution of ϕ'_{sh} which eventually provides a near-positive $\Delta\Delta\alpha(\nu)$ profile. Finally, the distinctly positive $\Delta\Delta\alpha(\nu)$ profile of Al^{3+} can be explained from a different perspective: the substantial red shift of the rattling mode of the ion in the presence of BSA signifies a major de-solvation of the octahedrally coordinated Al^{3+} ions[49] resulting in the further enhancement of bulk water, thereby making the third term in eqn (3.3) predominantly positive.

3.IV. Conclusions.

Our study has been aimed to experimentally reveal the associated hydration change as trivalent cations impose LLPS in BSA using THz-FTIR spectroscopy. Monovalent and divalent ions do not produce any such phase separation, as expected. All the trivalent cations do induce LLPS. They also induce charge reversal of the protein, the effect being although very subtle in Ho^{3+}. From the deconvolution of the experimental $\alpha(\nu)$ profiles, we obtain three modes, one of which is specific to ions (rattling mode), which manifests the solvation states of the ions, and it is found to strongly depend on the ion type. A detailed analysis of the overall hydration change (as defined by $\Delta\Delta\alpha(\nu)$) reveals very interesting ion-specific protein–ion hydration, specifically for the LLPS forming trivalent ions. We observe three distinct types of solvation nature: La^{3+} and Y^{3+}, Ho^{3+} and Al^{3+} (figure 3.III.7a). Since all these ions induce LLPS, the overall process must be entropically driven; however, it might not solely be driven by the hydrophobic hydration (protein aggregation followed by the release of hydration water), rather one needs to as well consider the corresponding contribution from ion (hydrophilic)–hydration, which mostly is electrostatic in nature and is very much ion specific, corroborating the conclusion drawn from a recent simulation study.[15] Our

experiments thus clearly demonstrate that consideration of mere charge (III in this case) on the metal ions is not sufficient to explain their behaviour towards macromolecules (formation of LLPS in the present case). It is noteworthy here that the ions have different ionic radius which follows the order: Al^{3+} (71 pm) < Y^{3+} (88 pm) < La^{3+} (108 pm) < Ho^{3+}(175 pm), and the change in hydration does not follow any correlation with the ionic size (figure 3.III.7a). So, a mere consideration of the surface charge density of the ions is not sufficient to explain the hydration behaviour, specifically in the presence of proteins. A possibility could be a salient role of the electronic configuration of these ions.

We believe that our study will provide impetus towards a combined experimental and simulation investigation to explain the ion specificity from a molecular point of view.

Table 3.III.1. The listed values of the fitted unperturbed center frequency (ν_c): Respective values for hydrogen bond stretching mode, for rattling motions of ions, and for water librational motion of each protein-salt system and only aqueous solution of BSA protein.

Salt	HB stretch ν_c (cm^{-1})	Ion rattling ν_c (cm^{-1})	Water libration ν_c (cm^{-1})
0	72.7(56)	-	474.1(23)
NaCl	106.8(24)	358.6(31)	551.7(25)
KCl	141(22)	403(26)	577.7(17)
$MgCl_2$	108.9(20)	341.9(9)	574.1(5)
$CaCl_2$	114.8(39)	351.8(15)	544.3(11)
$AlCl_3$	83.5(2)	168.5(11)	410.9(13)
$LaCl_3$	82.5(7)	391.8(14)	-
YCl_3	92.4(3)	385 (19)	-
$HoCl_3$	52.7(0.8)	154.5(13)	365.8(10)

Table 3.III.2.

System	$\Delta\Delta\alpha(\nu)$	ΔS_{hyd}	Hydration geometry	Remarks
Ho^{3+}	almost zero or nearly positive	positive	square antiprismatic (8 co-ordinated)	Red shift of the ion rattling mode Weak hydration water exchange
Al^{3+}	Distinctly positive	positive	octahedrally co-ordinated	Red shift of the ion rattling mode. Major desolvation of ocatahedrally co-ordinated ion

3.V. References.

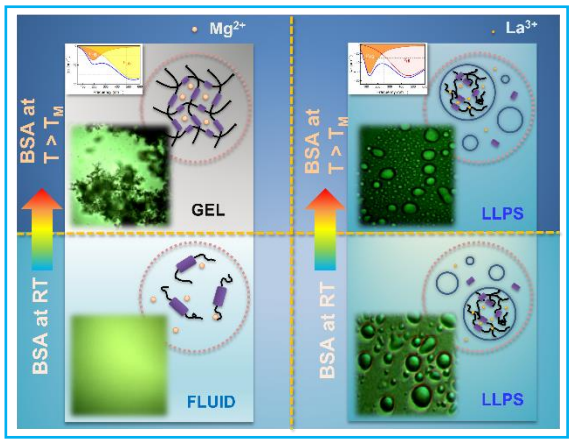
1. S. Pianegonda, M. C. Barbosa and Y. Levin, *Europhys. Lett.*, 2005, **71**, 831.
2. P. Jungwirth and P. S. Cremer, *Nat. Chem.*, 2014, **6**, 261-263.
3. O. Matsarskaia, F. Roosen-Runge and F. Schreiber, *Chemphyschem : a European journal of chemical physics and physical chemistry*, 2020, **21**, 1742-1767.
4. H. I. Okur, J. Hladílková, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer and P. Jungwirth, *J. Phys. Chem. B*, 2017, **121**, 1997-2014.
5. F. Zhang, M. Skoda, R. Jacobs, S. Zorn, R. A. Martin, C. Martin, G. Clark, S. Weggler, A. Hildebrandt and O. Kohlbacher, *Phys. Rev. Lett.*, 2008, **101**, 148101.
6. F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. Skoda, R. M. Jacobs, M. Stzucki and F. Schreiber, *Soft Matter*, 2012, **8**, 1313-1316.
7. O. Matsarskaia, M. K. Braun, F. Roosen-Runge, M. Wolf, F. Zhang, R. Roth and F. Schreiber, *J. Phys. Chem. B*, 2016, **120**, 7731-7736.
8. V. J. Anderson and H. N. Lekkerkerker, *Nature*, 2002, **416**, 811-815.
9. C. Ishimoto and T. Tanaka, *Phys. Rev. Lett.*, 1977, **39**, 474.
10. M. K. Braun, A. Sauter, O. Matsarskaia, M. Wolf, F. Roosen-Runge, M. Sztucki, R. Roth, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2018, **122**, 11978-11985.
11. E. Jordan, F. Roosen-Runge, S. Leibfarth, F. Zhang, M. Sztucki, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *J. Phys. Chem. B*, 2014, **118**, 11365-11374.
12. F. Roosen-Runge, B. S. Heck, F. Zhang, O. Kohlbacher and F. Schreiber, *J. Phys. Chem. B*, 2013, **117**, 5777-5787.
13. P. J. Flory, *J. Chem. Phys.*, 1942, **10**, 51-61.
14. M. L. Huggins, *J. Chem. Phys.*, 1941, **9**, 440-440.
15. A. K. Sahoo, F. Schreiber, R. R. Netz and P. K. Maiti, *Soft Matter*, 2022, **18**, 592-601.
16. O. Matsarskaia, S. Da Vela, A. Mariani, Z. Fu, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2019, **123**, 1913-1919.
17. O. Matsarskaia, F. Roosen-Runge, G. Lotze, J. Möller, A. Mariani, F. Zhang and F. Schreiber, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27214-27225.
18. T. Gibaud and P. Schurtenberger, *J. Phys.: Condens. Matter*, 2009, **21**, 322201.
19. F. Zhang, F. Roosen-Runge, A. Sauter, M. Wolf, R. M. J. Jacobs and F. Schreiber, *Pure Appl. Chem.*, 2014, **86**, 191-202.
20. F. Zhang, G. Zocher, A. Sauter, T. Stehle and F. Schreiber, *J. Appl. Crystallogr.*, 2011, **44**, 755-762.

21. M. Heyden, S. Ebbinghaus and M. Havenith, *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation*, 2006.
22. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 12068-12073.
23. B. Born and M. Havenith, *J. Infrared, Millimeter, Terahertz Waves*, 2009, **30**, 1245-1254.
24. R. Aakesson, L. G. M. Pettersson, M. Sandstroem and U. Wahlgren, *J. Am. Chem. Soc.*, 1994, **116**, 8691-8704.
25. G. Schwaab, F. Sebastiani and M. Havenith, *Angew. Chem. Int. Ed. Engl.*, 2019, **58**, 3000-3013.
26. V. Sharma, F. Böhm, G. Schwaab and M. Havenith, *Phys. Chem. Chem. Phys.*, 2014, **16**, 25101-25110.
27. S. Funkner, G. Niehues, D. A. Schmidt, M. Heyden, G. Schwaab, K. M. Callahan, D. J. Tobias and M. Havenith, *J. Am. Chem. Soc.*, 2012, **134**, 1030-1035.
28. A. Valstar, M. Almgren, W. Brown and M. Vasilescu, *Langmuir*, 2000, **16**, 922-927.
29. D. D. Mahanta, N. Samanta and R. K. Mitra, *J. Mol. Liq.*, 2016, **215**, 197-203.
30. I. Hamachi, R. Eboshi, J.-i. Watanabe and S. Shinkai, *J. Am. Chem. Soc.*, 2000, **122**, 4530-4531.
31. M. Lund and B. Jönsson, *Biochemistry*, 2005, **44**, 5722-5727.
32. E. Bianchi, R. Blaak and C. N. Likos, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6397-6410.
33. L. Medda, M. Monduzzi and A. Salis, *Chem. Commun.*, 2015, **51**, 6663-6666.
34. J. J. López-García, J. Horno and C. Grosse, *Langmuir*, 2011, **27**, 13970-13974.
35. W. M. de Vos and S. Lindhoud, *Adv. Colloid Interface Sci.*, 2019, **274**, 102040.
36. D. Melekaslan and O. Okay, *Macromol. Chem. Phys.*, 2001, **202**, 304-312.
37. B. Born, S. J. Kim, S. Ebbinghaus, M. Gruebele and M. Havenith, *Faraday Disc.*, 2009, **141**, 161-173.
38. V. Bianco and G. Franzese, *Phys. Rev. Lett.*, 2015, **115**, 108101.
39. D. A. Schmidt, O. z. r. Birer, S. Funkner, B. P. Born, R. Gnanasekaran, G. W. Schwaab, D. M. Leitner and M. Havenith, *J. Am. Chem. Soc.*, 2009, **131**, 18512-18517.
40. C. Klinkhammer, F. Böhm, V. Sharma, G. Schwaab, M. Seitz and M. Havenith, *J. Chem. Phys.*, 2018, **148**, 222802.
41. T. Dodo, M. Sugawa, E. Nonaka, H. Honda and S. i. Ikawa, *J. Chem. Phys.*, 1995, **102**, 6208-6211.
42. P. Pyne and R. K. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
43. P. Schienbein, G. Schwaab, H. Forbert, M. Havenith and D. Marx, *J. Phys. Chem. Lett.*, 2017, **8**, 2373-2380.
44. J. Ahlers, E. M. Adams, V. Bader, S. Pezzotti, K. F. Winklhofer, J. Tatzelt and M. Havenith, *Biophys. J.*, 2021, **120**, 1266-1275.
45. U. Heugen, G. Schwaab, E. Bründermann, M. Heyden, X. Yu, D. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 12301-12306.
46. M. Heyden, E. Bründermann, U. Heugen, G. Niehues, D. Leitner and M. Havenith, *J. Am. Chem. Soc.*, 2008, **130**, 5773-5779.
47. S. Park, R. Barnes, Y. Lin, B.-j. Jeon, S. Najafi, K. T. Delaney, G. H. Fredrickson, J.-E. Shea, D. S. Hwang and S. Han, *Commun. Chem.*, 2020, **3**, 83.
48. I. Persson, P. D'Angelo, S. De Panfilis, M. Sandström and L. Eriksson, *Chemistry*, 2008, **14**, 3056-3066.
49. W. W. Rudolph, R. Mason and C. C. Pye, *Phys. Chem. Chem. Phys.*, 2000, **2**, 5030-5040.

Chapter 4

4. Thermo-resistive phase behaviour of trivalent ion induced microscopic protein rich phases: Correlating with the ion-specific protein hydration

Proteins, in presence of tri-valent cations, exhibit intriguing phase behaviour which is contrasting as compared to mono- and di-valent cations. At room temperature (RT) tri-valent cations induce microscopic liquid-liquid phase separation (LLPS) in which a protein rich phase coexists with a dilute phase. The critical solution temperature (CST) related phenomena in these complex fluids are well studied, however, such studies have mostly been restricted below the denaturation temperature (T_M) of the protein(s) involved. Here we probe the phase behaviour of bovine serum albumin (BSA) incubated at 70 ^{0}C (>T_M) in presence of Na^+, Mg^{2+}, La^{3+}, Y^{3+} and Ho^{3+} ions. BSA in presence of mono- and bi-valent ions forms intense gel phase at 70 ^{0}C; however, the tri-valent salts offer a remarkable thermal resistivity and retain the fluid LLPS phase. We determine the microscopic phase behaviour using differential interference contrast (DIC) optical microscopy, which shows that the LLPS droplet structures in M^{3+} ion containing protein solutions prevail upon heating whereas Mg^{2+} forms composed crosslinking gelation upon thermal incubation. We probe the interior environment of the protein aggregates by ps-resolved fluorescence anisotropy measurements using 8-Anilino-1-naphthalenesulfonic acid (ANS) as an extrinsic fluorophore. It reveals that while LLPS phase retains the rotational time-constants upon heating, in case of gelation, the immediate environment of ANS gets significantly perturbed. We investigate the explicit protein hydration at RT and as well as at T>T_M using ATR THz-FTIR (1.5-22.5 THz) spectroscopy technique and found that the hydration shows strong ion-specificity and correlates the phase transition behaviour.



4.I. Introduction.

Multivalent (specifically trivalent) cations offer intriguing effects on biomacromolecules (proteins) as they show diverse phase behaviours which is governed by their distinct electrostatic interactions.[1-3] Such ions could lead to charge inversion of protein surface at room temperature (RT) and subsequently to re-entrant phase behaviour unlike in mono- and di-valent ions.[4] Generic re-entrant ion-condensation implies characteristic isothermal liquid-liquid phase separation (LLPS) of globular protein molecules.[5] However, such isothermal LLPS formation is often fails to explain the lower critical solution temperature (LCST) behaviours observed in many globular proteins.[6] Microscopic phase behaviour allows long range inter-protein cation binding and cation bridging between individual protein strands in LCST-LLPS condition at high temperature.[7] LCST-LLPS is an entropically driven phenomenon at high temperature with ease in phase separation.[8, 9] Most of these earlier studies on the temperature dependency of cation-induced protein LLPS have been limited to a particular range of temperature, which is below the denaturing temperature of the corresponding protein ($T<T_M$).[10] Heat induced protein denaturation leads to unfolding of protein's structural domains followed by aggregation/gelation depending on the protein concentration.[11-14] According to the classical colloidal theories, ionic strength of mono- and di-valent cationic salts offers prominent effect on protein gelation by reducing the Coulomb interactions[15] and retaining a certain degree of order in the gel matrix and predicts[16] an entropy gain upon protein aggregation as hydrated protein releases water into bulk as gelation sets in. This theory also involves the role of ions in the gelation process; however, the exact nature of their contribution has not been explicitly been expressed to be ion-specific. Thus, this theory predicts comparable phase behaviour in presence of ion of any charge. This idea, while holds true for many mono- and bi-valent cations, in not trivial in presence of tri-valent cations. We recently have observed that trivalent cations impose the formation of LLPS microphase(s) in bovine serum albumin (BSA) at RT.[17] Our experimental study reveals an intriguing ion-specific protein and ion hydration, specifically for the LLPS forming trivalent ions. In the present study we take a step forward from the phase isotherm at RT and explore whether they offer distinct signature at $T>T_M$. Moreover, it is expected that hydration should also play a key role in determining the phase behaviour in such multi-component system. Such a notion has very rarely been addressed in the past. The present experimental investigation is a step forward to verify those.

In this work we characterize in vitro phase behaviour of a model globular protein bovine serum albumin (BSA) in presence of monovalent (NaCl), divalent ($MgCl_2$) and trivalent

($LaCl_3$, YCl_3 and $HoCl_3$) salts at RT (25 ^{0}C) and at 70 ^{0}C (>T_M). We carry out all the measurements maintaining a fixed salt concentration (C_s):protein concentration (C_p) ratio of 23.33; the specific choice of this ratio gets its justification from the fact that previous reports show that trivalent salts bring about microscopic phase separation in BSA at this particular ratio at RT.[18] We observe that trivalent ions induce distinct phase behaviour at T>T_M. We found a remarkable thermal stability of the LLPS phase at T>T_M, which is contrasting as compared to bare protein and Mg^{2+} containing protein. We use UV-Visible spectroscopy, dynamic light scattering (DLS) and optical microscopy techniques to characterize these different phases. We also employ steady state and picosecond-resolved fluorescence anisotropy measurements using ANS as the fluorophore to underline the distinct information on the micro-heterogeneous viscous environment around the protein as different phases are formed. We finally investigate the change in protein hydration during such phase transitions using ATR THz-FTIR spectroscopy in the frequency window of 1.5-22.5 THz (50-750 cm^{-1}). THz absorption spectroscopy is a potential to label-free tool to trace the global dynamical at biomolecular interfaces as it directly probes the fluctuation of the collective dipole moment of water and hence can explicitly probe the vibrational modes of water network.[19, 20] Our group has put forward substantial effort to understand protein hydration under various bio-physical conditions using THz spectroscopy measurements.[21-25] We observe that the thermos-responsive behavior of mono- and di-valent ions as well as the thermos-resistive behavior of the tri-valent ions is strongly associated with the corresponding change in the protein hydration.

4.II. Materials and Methods.

All the salts: sodium chloride (NaCl), magnesium chloride ($MgCl_2$), lanthanum chloride ($LaCl_3$), yttrium chloride (YCl_3) and holmium chloride ($HoCl_3$) were purchased from Sigma Aldrich (Merck) with ≥99% purity of analytical grade. Bovine serum albumin (BSA; Sigma-A4503) with ≥96% purity was used without further refinement. The details of all the used instruments are described in the section 2.III and the experimental analysis details are also mentioned prior in the section 2.IV.

4.III. Results and Discussions.

Visual appearance and turbidity measurements: We first investigate the effect of ions on the physical appearance of BSA at RT. Figure 4.III.1a describes the visual appearance of freshly prepared BSA solutions at 25 ^{0}C in different environments. Pure BSA and NaCl, $MgCl_2$ containing solutions look transparent and colourless while the trivalent ($LaCl_3$, YCl_3 and $HoCl_3$) salt(s) containing samples appear slightly bluish and translucent. Overnight incubation at 25 ^{0}C does not alter the visual appearance of the samples. For a more quantitative insight we perform turbidity measurements to determine the inhomogeneity in the samples (figure 4.III.1c). The trivalent salt containing solutions produce ~2 to 3 times higher turbidity than the pure BSA solution; our previous study[17] has confirmed the formation of LLPS droplets in these samples, which results in the increased turbidity. We then incubate all the samples at 70 ^{0}C for

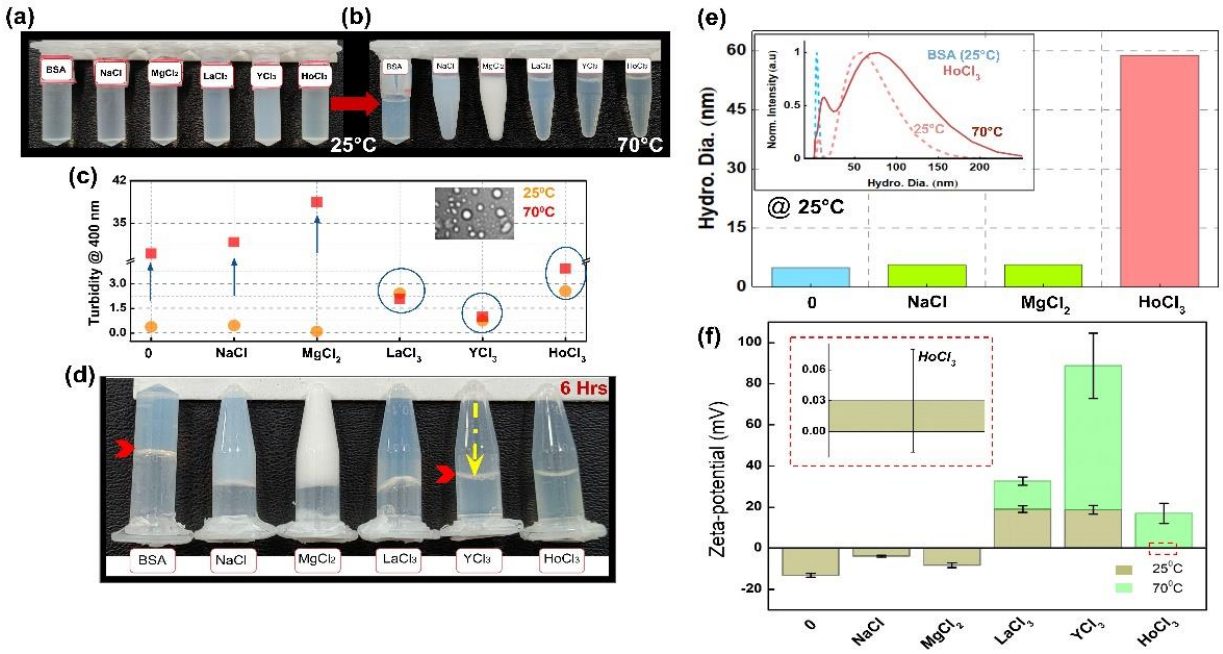


Figure 4.III.1. Thermal behaviour of only BSA protein, and protein-salt solutions w.r.t turbidity, structural analysis and ζ potential measurements. Photographs of the respective samples after incubation (a) at 25 ^{0}C and (b) at 70 ^{0}C for 6 h. (c) Turbidity measurement at 400 nm wavelength using a UV spectrophotometer. Blue arrows and rounds used to guide the visual. (d) Photograph of all incubated samples to isolate their flowing nature at each state. Arrows are used to point out ultimate surface. (e) Hydrodynamic diameter(s) of BSA in the absence and presence of salts at 25 ^{0}C. The inset shows normalised intensity distribution against hydrodynamic diameter of bare BSA and BSA in the presence of $HoCl_3$ at two different temperatures 25 and 70 ^{0}C. (f) ζ Potential of BSA solutions at two different temperatures (25 and 70 ^{0}C). ζ potential value of $HoCl_3$-BSA solution (at 25 ^{0}C) is shown in the red box. The error bars signify standard deviations.

6 hours and visualize them as they equilibrate at RT (figure 4.III.1b). Pure BSA solution is transformed into a less transparent highly viscous gel-like state. It does not flow through the surface of the tube when kept upside down for hours (figure 4.III.1d). Its turbidity also increases by ~30-40 fold (figure 4.III.1c). In presence of $MgCl_2$ it forms an even intense white opaque highly viscous gel like phase. In presence of the trivalent salts, the turbidity of the solutions remains mostly unaltered (figure 4.III.1c), however, we observe ion-dependent phase behaviour: in presence of La^{3+}, it forms a slightly viscous phase while in Y^{3+} and Ho^{3+}, the solutions remain fluid as it was at RT (figure 4.III.1d).

Size measurements by DLS: We then measure the size of the droplets in all the samples using DLS technique (figure 4.III.1e). Compared to the pure protein (diameter ~5 nm) the LLPS droplets are found to be larger in size at 25 ^{0}C; a representative scattering profile for Ho^{3+} is shown in the figure. Noticeably both Na^+ and Mg^{2+} do not alter the size of the protein molecules. As the bare protein is incubated at 70 ^{0}C, it denatures and thereby increasing the size drastically with a wide poly dispersity index (50 nm to 6 μm). Interestingly, as the LLPS solutions were incubated at 70 ^{0}C, the size changes only marginally (figure 4.III.1e; inset). The turbidity and DLS measurements thus indicate that temperature increase above T_m has less pronounced effect on the LLPS droplets.

Zeta-potential study: To understand the charge modification on the protein surface we measure the zeta-potential of BSA in presence of these metal ions (figure 4.III.1f). At RT the BSA surface potential is negative in buffer and remains negative in presence of both Na^+ and Mg^{2+}. We observe a drastic inversion of surface potential at RT in presence of all the trivalent ions as the surface charge becomes positive. The effect is subtle in Ho^{3+}, while for the other trivalent ions it is considerable. Interestingly, as the solutions were incubated at 70 ^{0}C, it appears that the charge inversion becomes even more prominent in presence of the trivalent salts, which supports the thermos-resistivity of the protein surface in presence of the tri-valent ions.

Morphological analysis: In order to obtain insight into the microstructures of these phases we perform differential interference contrast (DIC) optical microscopy measurements (figure 4.III.2). We show three representative systems: pure protein, Mg^{2+} and La^{3+} containing protein solution(s). After incubation at 25 ^{0}C 0.6 mM BSA solution, in absence and in presence of $MgCl_2$, appears clear under the microscope indicating homogeneously dispersed monomeric protein molecules in the solution. In presence of $LaCl_3$ we register the appearance of large LLPS droplets.[17] Similar phenomenon is also observed for the other trivalent ions. When these

samples are incubated at 70 ^{0}C for 6 hours we observe various phases to appear (the middle panel, figure 4.III.2): in pure BSA we found indefinite structures while in $MgCl_2$ it forms a

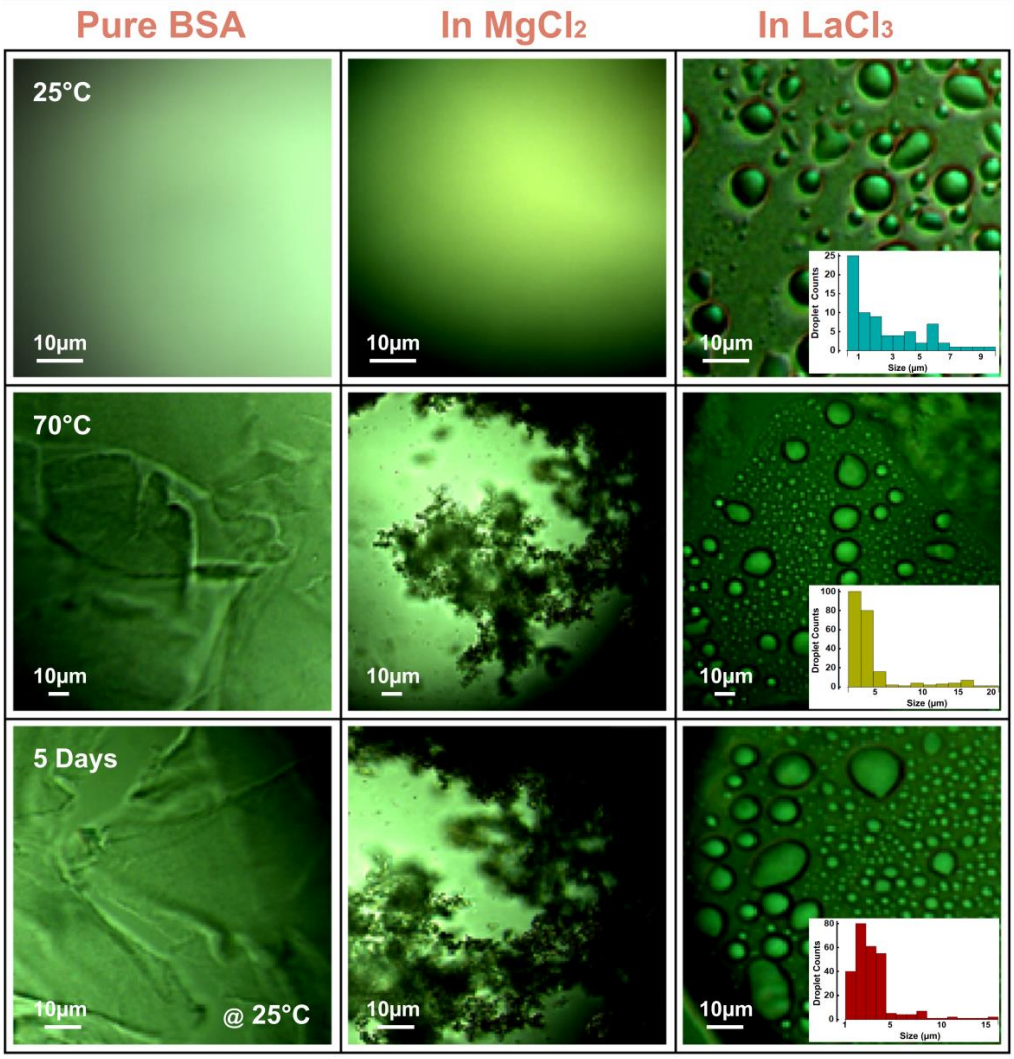


Figure 4.III.2. Optical microscopy (DIC) images of pure BSA protein in the presence of divalent $MgCl_2$ salt and trivalent $LaCl_3$ salt, respectively (in the column) using RGB color scheme at 10 μm scale. Images using the DIC method after (upper panel) overnight incubation at 25˚C; (middle panel) 6 h of incubation at 70 ˚C, and followed by (low panel) incubation at 25˚C for 5 days. Representative size distribution plots (insets) of LLPS liquid droplets corresponds to each incubated state (protein in $LaCl_3$).

huge network of cross-linking aggregates. Note that the curdy opaque appearance in Mg^{2+} (figure 4.III.1b) implies the formation of such huge network structures as evidenced in earlier studies.[26, 27] These micro-structures do not change much upon keeping them at room temperature for 5 days. On the contrary, for La^{3+} we observe that upon incubation at 70 ^{0}C the basic shape of the microstructures as observed at RT (LLPS droplets) remains the same,

however, their number density and size distribution gets subtly modified. The distribution of the droplet size is shown in the corresponding inset. The droplet phase does not change upon keeping the samples for 5 days at room temperature, their size distribution, however, suffers certain modification upon standing. This result unambiguously reveals that trivalent salt induced LLPS does retain their structural integrity while bare and mono- and bi-valent ions containing protein solutions do undergo conformational rupture followed by phase transition upon heating.

Study of micro-environment by fluorescence anisotropy: Before analysing the hydration change during the various phase transitions, it is also important to understand the internal environment of the proteins aggregations as they are formed. An elegant way to monitor this is to perform fluorescence spectroscopy measurements. We use ANS, a well fluorescence probe used in biomolecules,[28] as an extrinsic fluorophore to probe the micro-heterogeneity within the protein aggregates. Measurement of angle dependent fluorescence (polarization angle of 0^0 and 90^0) enables one to calculate the rotational anisotropy (see SI section for details), which in turn manifests the local viscosity around the probe as it binds to the protein. We calculate both steady state and time-resolved rotational anisotropy of ANS in different protein environments. The steady state anisotropy of ANS in buffer is presumably very low (0.06) and decreases mildly at higher temperature (table 4.III.1). In presence of BSA (0.6 mM) at RT the value increases to 0.18 as ANS binds to the hydrophobic sites of the protein and thereby its rotation gets restricted.[29, 30] In protein rich LLPS micro-phase (in presence of salts) anisotropy value also remains high. When BSA forms gel at 70 ^{0}C, we observe a sharp drop in the ANS anisotropy value (table 4.III.1), a value similar to that in bulk water, which manifests increase in the rotational flexibility of ANS in the BSA gel. Similar drop in anisotropy is also observed in presence of Mg^{2+} ions also. Interesting results are observed for trivalent ion containing LLPS phases, as the anisotropy value does not change upon heating (table 4.III.1). To get into a deeper insight we also measure the ps-resolved rotational anisotropy of ANS in all these systems. ANS in buffer offers a single exponential rotational decay with a time constant of ~0.1 ns (figure 4.III.3a, table 4.III.2). In presence of BSA, the rotational decay transient gets noticeably slower, and could only be fitted considering a bi-exponential model (figure 4.III.3b). We obtain two time constants: 0.23 ns (τ_{fast}) and a relatively slower time constant ~19 ns (τ_{slow}). It is to be noted here that while τ_{fast} manifests the restriction imposed on the ANS present in the protein solvation shell, τ_{slow} mostly associates its translational movement of the probe at the protein surface coupled with the tumbling motion of the protein molecule. Since our study is focused

on the modification of hydration during protein aggregation, we make our discussion mostly focusing more on the fast rotational time constant. As the protein forms LLPS at RT in presence of La^{3+} and Ho^{3+} ions, the (τ_{fast}) shows a relatively slower value (0.35-0.4 ns), while in presence of Mg^{2+}, it does not change much

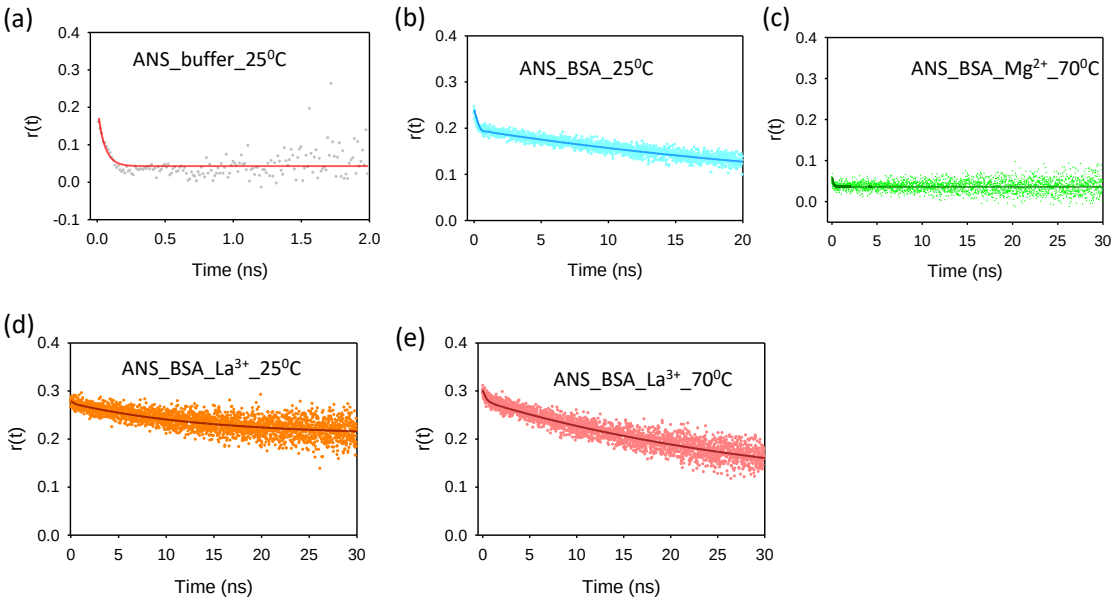


Figure 4.III.3. Time resolved rotational anisotropy (r(t) of ANS in different environments (a) in buffer, (b) in 6 mM BSA, (c) in BSA incubated at 70^0C in presence of Mg^{2+} (forming a gel phase), (d) BSA incubated at 25^0C in presence of La^{3+} (forming an LLPS phase) and (e) BSA incubated at 75^0C in presence of La^{3+} (forming an LLPS phase). The solid lines are exponential fits.

(table 4.III.2). As the BSA solution is incubated at 70 ^{0}C and gel is formed, we observe that interestingly τ_{fast} takes a value comparable to that in bulk water (~0.1 ns) with an offset in the *r(t)* profile (figure 4.III.3c). The presence of an offset indicates that a fraction of the ANS molecules is bound to a phase of protein which is unable to complete its rotation within the experimental time window. On the other hand, the τ_{fast}~0.1 ns value indicates that a fraction of ANS molecules are also present in bulk water, which could result from the de-solvation during the gel formation (see later). A similar observation is found in case of BSA-Mg^{2+} system also (table 4.III.2). Interestingly, BSA, in presence of La^{3+} and Ho^{3+}, when incubated at 70^0C, produces decay transients which are marginally faster, however, keep the bi-exponential feature as has been obtained at RT (figure 4.III.3, table 4.III.2). These observations also corroborate the steady state results and strongly establish the thermos-resistive nature of the LLPS phases, which is in stark contrast with the monomeric one.

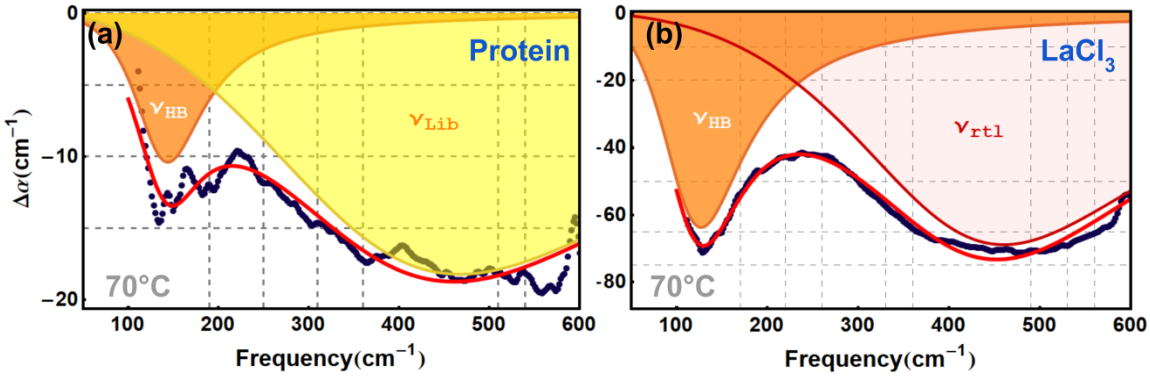


Figure 4.III.4. Dissection of the experimental frequency dependent difference absorption coefficients Δα(ν) fitted profile using the 3D damped harmonic oscillator model: (a) for only BSA protein and (b) in the presence of $LaCl_3$. Deconvoluted spectrums with orange, yellow and light pink colors represent the HB stretching mode, water librational motion and ion rattling mode, respectively. Red line stands for the fitted curve over deep blue experimental data points.

Hydration study: We finally explore that while the LLPS phases retain their structural morphology, what happens to its hydration; or to put the question in other words: how the hydration differs as thermos-sensitive systems undergo phase transition while M^{3+} containing systems retain it. To address this, we determine the protein hydration in these different phases using ATR-THz (50-750 cm^{-1}; 1.5-22.5 THz) FTIR measurements. As the temperature is raised to 70 ^{0}C, the protein solution shows further decrease in the absorption profile. A more explicit insight of protein hydration is obtained from the difference spectra, $\Delta\alpha(\nu) = \alpha(\nu)_{protein+salt+water} - \alpha(\nu)_{salt+water}$ (figure 4.III.4). Two representative fitting profiles for pure BSA and La^{3+} containing BSA (both incubated at 70 ^{0}C) are shown in figure 4.III.4. All the fitting parameters are summarized in table 4.III.3 and 4.III.4. The analysis yields three types of bands with corresponding centre frequencies appearing at ~127-190, 402-460 and 467-548 cm^{-1}. Noteworthy that all these three bands do not necessarily appear in each system manifesting the very distinct solvation nature of the different phases. It is important to remember here that since the fitted parameters are extracted from the difference spectra (Δα(ν)), it correlates the protein surface hydration explicitly, and does not include any contribution from the bulk hydration. Thus, any change in the fitting parameter(s) emanates from the finite change in the protein surface hydration. Dissection of pure water spectrum (α(ν)) produces two bands peaking at ν_{HB} ~130 cm^{-1}, associated with the intermolecular H-bond stretching vibrations of pure water, and ν_{Lib} ~570 cm^{-1}, associated the constrained

rotational (librational) motion of water, respectively.[31] Note that in presence of protein (as we dissect the difference profile, $\Delta\alpha(\nu)$), these peak frequencies change (table 4.III.3 and 4.III.4), which is in accordance to the fact that water at the protein surface experiences altered environments. It is also important to note here that the intermediate band ν_{rtl}~ 402-460 cm^{-1} appears only in presence of ions as the ion(s) rattle along their hydration cages.

We plot the peak frequencies of the three modes in all the systems at two different incubation temperatures (25 ^{0}C and 70 ^{0}C) in figure 4.III.5. Protein solution incubated at 25 ^{0}C shows two signature bands: HB stretching mode (ν_{HB}) at ~73 cm^{-1} and libration band (ν_{Lib}) at ~474 cm^{-1}, both of which are red shifted compared to those in bulk water. When incubated at 70 ^{0}C ν_{HB} appears at ~144 cm^{-1} while ν_{Lib} appears at ~467 cm^{-1}. A blue shift in ν_{HB} with respect to bulk water ascertains the formation of stronger intermolecular HBs at protein surface. Protein in presence of mono- and bi-valent ions shows a blue shift (~30 cm^{-1}) of ν_{HB} upon incubation at 25 ^{0}C, while ν_{Lib} suffers a drastic blue shift. Upon 70 ^{0}C incubation we observe $\nu_{HB}^{Na} \sim \nu_{HB}^{bulk}$, while in Mg^{2+} we observe ν_{HB}^{Mg} ~190 cm^{-1}. Also ν_{lib}^{Mg} is obtained at ~25 cm^{-1} smaller than that at RT. This red shift in the libration motion imparts a comparatively free rotation of water molecules as protein aggregates. In case of trivalent ions: incubation at 25 ^{0}C, the blue shift in ν_{HB} is comparatively less (~10-15 cm^{-1}) and, interestingly ν_{HB}^{Ho} appears at 53 cm^{-1}. Incubation at 70 ^{0}C makes ν_{HB} to regain its value as that in bulk water up to a certain extent. Intriguingly in presence of Y^{3+} and La^{3+} the water librational mode could not be recovered in both the incubated states. In case of Ho^{3+} the otherwise highly red shifted libration contribution at 25 ^{0}C could not be recovered at 70 ^{0}C.

Apart from these two modes, the recovered rattling mode in metal ion(s) (ν_{rtl}^{M}) imprints the individual contribution of ion(s) in protein's collective hydration. This mode is generally observed at low frequency (<200 cm^{-1}) and can as well overlap with the water ν_{HB} band. In lanthanides (III), an extension of the rattling motion to >200 cm^{-1} is observed which is related to a specific cationic resonance (M^{n+}-O) stretching mode.[32] The oscillation of ions in their hydration cage eventually manifests H-bond strength.[33] Incubation at 70 ^{0}C makes trivalent ions to rattle at higher frequencies than at RT and assigns a strong specific cationic resonance M^{n+}-O stretching mode within their hydration cages. It infers that these ions get restricted in the rigid hydration layer following heat treatment. We observe another interesting fact that there is no contribution of the ion-rattling mode in Mg^{2+} containing protein after 70 ^{0}C incubation.

Perhaps a complete engulfment of the Mg^{2+} ions at the denatured protein surface annihilates this mode.

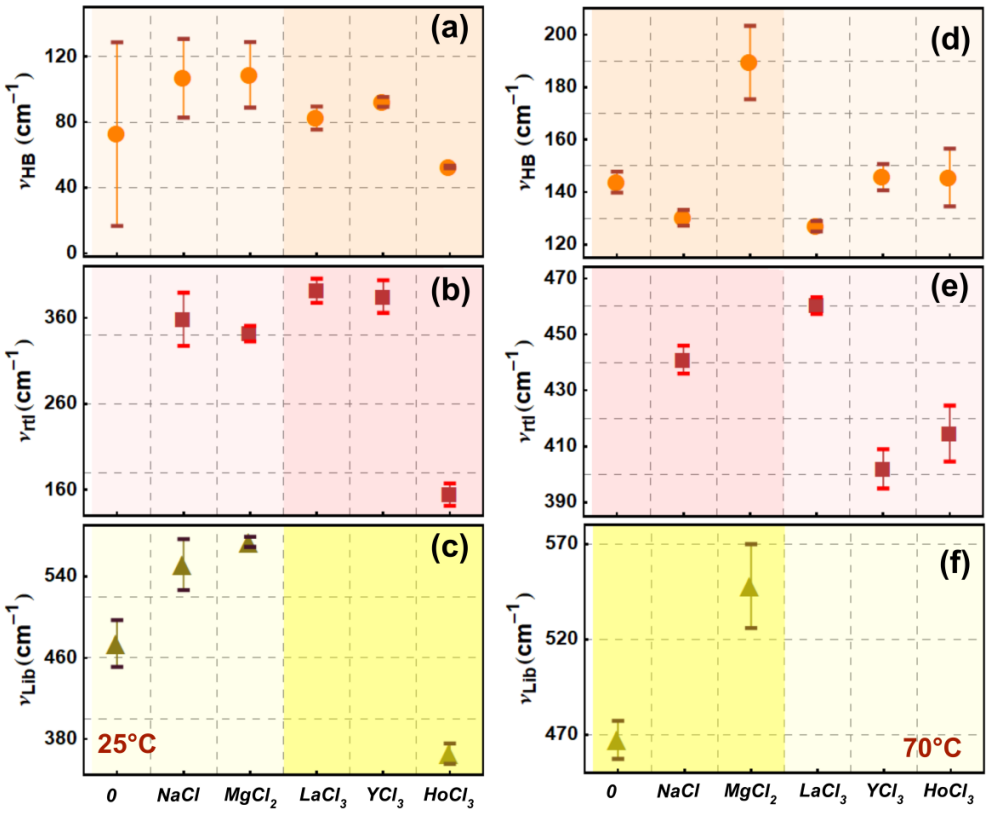


Figure 4.III.5. Comparative visualization of the unperturbed peak positions (a-f) at two different temperatures, for only BSA protein (indicated by 0 here) and all examined protein-salt solutions against three different modes due to HB stretching motion, ions rattling motions and, water librational motion, respectively. The error bar signifies the standard deviations for the respective system and colored background from lighter to darker indicates turbidity increments.

Discussion: Our study has been aimed to understand BSA conformational changes in presence of differently charged cations as the temperature is increased to 70 ^{0}C (>T_M). BSA is known to undergo irreversible thermal denaturation when heated T >65 ^{0}C.[34] Also, it has been known that BSA forms protein rich globular micro-phases in presence of trivalent salt(s) at particular concentration window(s), whereas, mono- and di-valent salts do not usually offer any such structural modifications.[18, 35] Our earlier study has established that tri-valent cations induces unusual hydration effect in BSA, specially as it forms LLPS at RT. Here we extend the idea that how far this unusual hydration prevails at T>T_M. THz measurements enable us to obtain explicit information on protein hydration as it forms various phases at T>T_M. We found that BSA forms a hydrogel when incubated at 70 ^{0}C and in presence of Mg^{2+} it forms even intense

gel (figure 4.III.1a). It is known that bi-valent ions facilitate BSA denaturation followed by the formation of hydrogel;[36, 37] however, a molecular level understanding of the process involved, specially if hydration plays a role, is not very clearly understood yet. In case of trivalent salts, it appears that the LLPS phase is thermo-resistive to 70 ^{0}C incubation. It is also observed that turbidity of the solutions changes only mildly with temperature (figure 4.III.1c). In order to achieve at a microscopic insight of the phase behaviour we construct the DIC micrographs (figure 4.III.2). It appears that pure BSA at elevated temperature forms an indefinite layered structure. The microscopic opaque gel phase in presence of Mg^{2+} actually consists of large network of cross-linked aggregates (figure 4.III.2). The LLPS droplets (formed in presence of trivalent salts) remain more or less intact, however, their size distributions change at elevated temperatures. This conclusion is also well supported from the time-resolved anisotropy study which reveals that the interior environment of LLPS phase does not suffer much modification when heated at $T>T_M$.

Hydration change during the phase transition was determined using ATR-FTIR spectroscopy in the THz frequency region. We observe that the peak frequency(s) (arising mostly due to the various vibrational modes of water) changes as different phases are formed (figure 4.III.5, table 4.III.1). An elegant way to probe protein solvation is to dissect the experimentally obtained absorption co-efficient (α_{THz}), a parameter which manifests explicit solute solvation, into contributions from each component: $\alpha(\nu) = \sum_i \phi_i \alpha_i(\nu)$ where ϕ_i is the volume fraction of the *i-th* component. Also to be noted here that the sole consideration of the pure components only does not reconstruct the experimental $\alpha(\nu)$ profiles; one rather needs to consider the contribution(s) from the solute hydration layer(s), too.[19] We can identify four distinct phases of BSA in presence of the salts: native fluid state (N) at 25 ^{0}C, hydrogel state (G) at 70 ^{0}C, LLPS state (l) at 25 ^{0}C and LLPS state (L) at 70 ^{0}C. We experimentally measure the $\Delta\alpha(\nu)$ profiles for each of these states individually. They can analytically be expressed as:

$$\Delta\alpha_p^N(\nu) = \phi_p^N \alpha_p(\nu) + \phi_{ph}^N \alpha_{ph}^N(\nu) + (\phi_b^N - 1)\alpha_{bulk}(\nu) \qquad (4.1)$$

$$\Delta\alpha_p^G(\nu) = \phi_p^G \alpha_p(\nu) + \phi_{ph}^G \alpha_{ph}^G(\nu) + (\phi_b^G - 1)\alpha_{bulk}(\nu) \qquad (4.2)$$

Note that both protein (α_p) and its surface hydration (α_{ph}) along with bulk water (α_{bulk}) contribute towards the overall $\Delta\alpha_p(\nu)$ profiles in both the protein's native (N) and gel (G) state(s). Also note that since the protein in the gel state possesses a modified, possibly more hydrophobic, interface, $\alpha_{ph}^N(\nu) \neq \alpha_{ph}^G(\nu)$ seems to be a reasonable argument. It is evident that

the negative value of the third term overwhelms the first two positive contributions (in equation 4.2) resulting in the overall negative $\Delta\alpha(\nu)$ profile(s) (figure 4.III.4). To quantify the extent of heat-induced perturbation in protein hydration we further calculate the difference parameter:

$$\Delta\Delta\alpha_p(\nu) = \Delta\alpha_p^G(\nu) - \Delta\alpha_p^N(\nu)$$

$$= \left(\phi_p^G - \phi_p^N\right)\alpha_p(\nu) + \phi_{ph}^G\alpha_{ph}^G(\nu) - \phi_{ph}^N\alpha_{ph}^N(\nu) + (\phi_b^G - \phi_b^N)\alpha_{bulk}(\nu) \quad (4.3)$$

In presence of salt the hydration profile(s) accommodates additional absorption terms emanating from bare salt(s) (α_s) and their hydration (α_{sh}). Thus, the general form of $\Delta\alpha(\nu)$ for protein-salt solutions takes up the form:

$$\Delta\alpha_{p+s}(\nu) = \phi_p\alpha_p(\nu) + \phi_{ph}\alpha_{ph}(\nu) + \phi_s\alpha_s(\nu)$$

$$+ \phi_{sh}\alpha_{sh}(\nu) + (\phi_b' - 1)\alpha_{bulk}(\nu) \quad (4.4)$$

Taking into consideration that both the protein and the salt(s) offer negligible absorbance in the THz frequency window the following sets of equations can be derived:

$$\Delta\alpha_{p+s}^N(\nu) = \phi_{ph}^{N'}\alpha_{ph}(\nu) + \phi_{sh}^{N'}\alpha_{sh}(\nu) + (\phi_b^{N'} - 1)\alpha_{bulk}(\nu) \quad (4.5)$$

$$\Delta\alpha_{p+s}^G(\nu) = \phi_{ph}^{G'}\alpha_{ph}(\nu) + \phi_{sh}^{G'}\alpha_{sh}(\nu) + (\phi_b^{G'} - 1)\alpha_{bulk}(\nu) \quad (4.6)$$

$$\Delta\alpha_{p+s}^l(\nu) = \phi_{ph}^l\alpha_{ph}(\nu) + \phi_{sh}^l\alpha_{sh}(\nu) + (\phi_b^l - 1)\alpha_{bulk}(\nu) \quad (4.7)$$

$$\Delta\alpha_{p+s}^L(\nu) = \phi_{ph}^L\alpha_{ph}(\nu) + \phi_{sh}^L\alpha_{sh}(\nu) + (\phi_b^L - 1)\alpha_{bulk}(\nu) \quad (4.8)$$

The suffix N' and G' manifest contribution of each component corresponding to the native (N) and the gel (G) state(s) of the protein. 'l' implies LLPS state at RT and 'L' corresponds to that at 70 ^{0}C incubation. Consequently, the change in the hydration during thermal incubation in presence of the salts could be summarized as:

$$\Delta\Delta\alpha_{Gel}(\nu) = \phi_{ph}^{G'}\alpha_{ph}^{G'}(\nu) - \phi_{ph}^{N'}\alpha_{ph}^{N'}(\nu) + \left(\phi_{sh}^{G'} - \phi_{sh}^{N'}\right)\alpha_{sh}(\nu)$$

$$+\left(\phi_b^{G'} - \phi_b^{N'}\right)\alpha_{bulk}(\nu) \quad (4.9)$$

$$\Delta\Delta\alpha_{LLPS}(\nu) = \left(\phi_{ph}^L - \phi_{ph}^l\right)\alpha_{ph}(\nu) + \left(\phi_{sh}^L - \phi_{sh}^l\right)\alpha_{sh}(\nu)$$

$$+\left(\phi_b^L - \phi_b^l\right)\alpha_{bulk}(\nu) \quad (4.10)$$

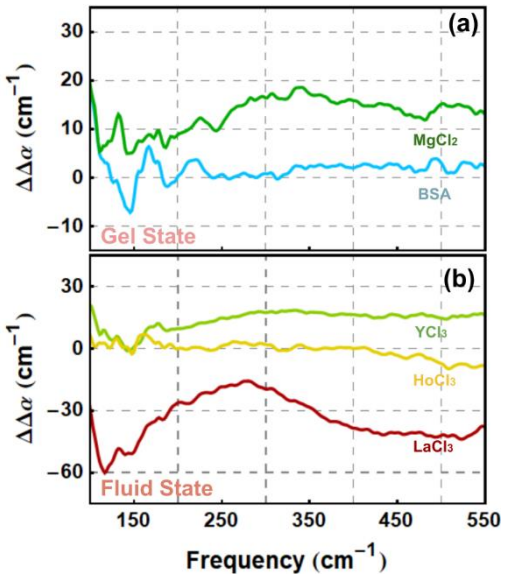


Figure 4.III.6. Representative profile of frequency dependent $\Delta\Delta\alpha(\nu)$ at two different physical state; (a) at gel state, and (b) at liquid state.

In these set of equations $\Delta\Delta\alpha(\nu)(= \Delta\alpha_{70^{\circ}C}(\nu) - \Delta\alpha_{25^{\circ}C}(\nu))$ refer to the explicit hydration of protein at macroscopic gel phase (Eq. 4.9) and microscopic LLPS phase (Eq. 4.10), respectively. The $\Delta\Delta\alpha(\nu)$ profiles in presence of different salts are depicted in figure 4.III.6. These experimentally obtained $\Delta\Delta\alpha(\nu)$ profiles could be explained in terms of the analytical expressions (equations 4.3-4.9) derived above in the following manner:

Let us first consider the bare protein itself. The $\Delta\Delta\alpha_{Gel}(\nu)$ profile is featureless and the value is very close to zero (figure 4.III.6a, blue curve). Consider the different terms in equation 4.3 to understand the contribution from each component. Native protein has a more-or-less hydrophobic interior with complex combinations of multiple non-covalent interactions with only traces of water.[38, 39] Raising temperature breaks these non-covalent bonds and makes the hydrophobic interior getting exposed to the polar solvent. This makes $\phi_b^G > \phi_b^N$ implying that the last term in equation 4.3 is always positive. Also, since the protein dimension also increases with temperature (see the DLS results), it can be assumed that $\phi_{ph}^G > \phi_{ph}^N$. So, the only condition that satisfies $\Delta\Delta\alpha_{Gel}(\nu) \approx 0$ is that $\alpha_{ph}^G < \alpha_{ph}^N$, which corroborates with the fact that in the hydrogel phase, the protein interior gets exposed to water.

We now discuss Mg^{2+}, which forms an intense gel phase and we found that $\Delta\Delta\alpha_{Gel}(\nu) > 0$ for the whole frequency profile. The change in hydration could be extracted considering the different terms in equation (4.9). Mg^{2+} ions compensates the colloidal London

interactions between the protein molecules and induce cross-linking between the thermally unfolded protein molecules having solvent exposed hydrophobic surface milieu, as has also been evident from the microscopy results.[15] Since Mg^{2+} directly binds to the protein at elevated temperatures and forms a gel phase, its individual hydration is highly compromised [$\Delta\phi_{sh} < 0$], a fact strongly corroborated from the disappearance of its rattling mode at 70 ^{0}C incubation (figure 4.III.5b, table 4.III.1). This makes the third term in equation (4.9) negative. Also, to note here that due to the aggregation, the protein's hydration is compromised [$\Delta\phi_{ph} < 0$] and from the fact that $\alpha_{ph}^{G} < \alpha_{ph}^{N}$, it seems evident that the first two terms contribute only negative shares towards $\Delta\Delta\alpha_{Gel}(\nu)$. The only condition that can make $\Delta\Delta\alpha_{Gel}(\nu) > 0$, is that $\Delta\phi_b$ is profoundly positive. This indeed happens as supported from the visual observation of eventual separation of water from the gel phase upon prolonged standing. It indicates to the fact that as protein hydrogel converts to compact gel in presence of Mg^{2+}, it expels the excess hydration present in the hydrogel. This conclusion is also supported by the fluorescence anisotropy measurements (figure 4.III.3, table 4.III.1): a large fraction of the anisotropy decay is not completed within the experimental time window corroborating the formation of a large protein network in which a fraction of the fluorophore ANS is embodied, and the faster component takes a value comparable to that of bulk water signifying the excess water to phase out from the protein aggregate.

The ion specific and apparently contrasting $\Delta\Delta\alpha_{LLPS}(\nu)$ profiles in case of trivalent ions can be explained considering the different terms appearing in equation (4.10). Formation of LLPS involves release of protein solvation in the form of bulk water[40] which implies $\Delta\phi_{bulk} > 0$ in all the trivalent systems.[17] For Y^{3+} we observe $\Delta\Delta\alpha_{LLPS}(\nu) > 0$ over the full frequency range. It can be remembered here Y^{3+} has also induced a drastic change in the protein surface potential after incubation at 70 ^{0}C (figure 4.III.1f) by making it highly positive, clearly indicating more Y^{3+} ions to bind the protein surface. Moreover, Y^{3+} ions are also prone towards bridging the individual protein monomers.[35] It therefore results in a loss of ion hydration and a subsequent release of that into the bulk. As a result, the overall $\Delta\phi_b$ becomes highly positive to compensate the negative contribution of $\Delta\phi_{Ph}$ and eventually making $\Delta\Delta\alpha_{LLPS}(\nu) > 0$ (equation 4.10). Ho^{3+} shows negligible change in protein solvation from RT to 70 ^{0}C incubation (figure 4.III.6b). It can be noted here that Ho^{3+} offers a blue shift in the ion rattling band upon incubation (figure 4.III.5, table 4.III.3) which specifies ion confinement within its hydration cage. In an earlier X-ray absorption fine structure spectroscopy study it has been revealed that hydration structures of lanthanide subgroup (Ln (III)) ion(s) depend strongly on

their hydration geometry.[41] Ho^{3+} belongs to the heavy Ln(III) ions and offers square antiprismatic (8 co-ordinated) distribution of hydration water molecules. The mean value of a Ho-O bond length is shorter than the lighter Ln(III) ions and shows slower hydration layer to bulk water exchange rate. This infers a minimal change in $\Delta\phi_{sh}$. This fact is also supported from a marginal change in the zeta-potential value also (figure 4.III.1f). The two terminal terms in equation (4.10) are balanced in a manner which eventually results in $\Delta\Delta\alpha(\nu) \rightarrow 0$. We find a captivating negative $\Delta\Delta\alpha(\nu)$ profile for La^{3+} (figure 4.III.6b). A blue shift in the La^{3+} rattling motion upon thermal incubation (figure 4.III.5) invokes strongly confined hydration layer. It inhibits the release of salt hydration into bulk and therefore offers a minimal change in $\Delta\phi_{sh}$. This fact is also supported from the observation that La^{3+} containing solutions are less fluid than the other two M^{3+} ions (figure 4.III.1d). We also observe that in the heat induced LLPS phase, larger number of micro-droplets are formed in presence of La^{3+} (figure 4.III.2) and thereby a substantial negative contribution in $\Delta\phi_{Ph}$ is anticipated. It appears that the marginal (positive) changes in both $\Delta\phi_{sh}$ and $\Delta\phi_{bulk}$ is not enough to overwhelm the negative protein hydration $\Delta\phi_{ph}$ contribution making $\Delta\Delta\alpha(\nu)$ negative throughout (equation 4.10).

Finally, we discuss on the thermo-responsive hydrogel formation of protein, specifically in presence of the mono- and di-valent salts and on the rather unique thermo-resistive behaviour of the tri-valent salts. Thermo-responsive gels are macroscopic phase of cross-linked protein network structures. Specifically, hydrogels contain high intra-trapped solvent content which helps to maintain the ordered crosslinking of charged protein molecules which is a prerequisite for hydrogel formation.[42] Classical theory of gelation[43] provides the structural information of the hydrogels in terms of shear modulus. According to this theory any external deformation makes crosslinking of protein networks inducing a shear modulus. BSA forms such hydrogel at $T>T_M$ as evidenced from the DIC micrographs (figure 4.III.2). When mono- and di-valent ions are added, they deform such hydrogel phase, protein-ion binding is one of the major reasons behind the deformation. Such deformation could as well force to squeeze out the inner solvent and to form composed gel.[44] However, the counterions remain inside the gel phase and maintain the overall electro neutrality of the gel.[45] Interestingly, in our explicit hydration study ($\Delta\Delta\alpha_{Gel}(\nu)$ profile) of BSA in Mg^{2+} strongly corroborates such a de-solvation notion. This conclusion gets confirmation from the fluorescence anisotropy

measurements also. The disappearance of its rattling mode at 70 ^{0}C incubation signifies Mg^{2+} ions to compromise its individual hydration [$\Delta\phi_{sh} < 0$] as it directly binds with the protein gel network (figure 4.III.5b, table 4.III.1). The dense network structure observed under DIC microscope (figure 4.III.2 at 70 ^{0}C) and the visual detection of water separation from the condensed gel phase eventually represents Mg^{2+} ion induced gel squeezing. Upon deformation the net negative charged protein backbone and Mg^{2+} ions arrange themselves in a way that the total electrostatic energy gets minimized. It could be noted here that apparently ions in the bulk would experience entropic advantage over those in entrapped in the gel phase, they do relocate in the gel interior in order to maintain electroneutrality. LLPS, on the other hand, is a metastable equilibrium in which proteins with inversed surface potential (as evidenced from potential measurements at $T>T_M$, f figure 4.III.1f) remain in equilibrium with the remaining oppositely charged protein monomers.[35] Close proximate binding of trivalent ions to the negatively charged protein surface leads to LLPS formation at RT. At elevated temperature ($T>T_M$) also this equilibrium sustains and stabilises protein from thermal deformation.

4.IV. Conclusions.

BSA, in presence of trivalent cations, forms microscopic phase separated LLPS droplets at RT; on the other hand, mono- and di-valent cations does not change protein conformation appreciably. When incubated at a temperature $T>T_M$ native BSA forms transparent hydrogel; in presence of Mg^{2+} white opaque composed gel is formed, which on microscopy reveals to containing cross-linked polymer network. Interestingly, in presence of trivalent salts, we observe a rather unusual thermo-resistive behavior as the RT formed LLPS retains their transparent and fluid like appearance. We determine the explicit hydration at the protein interface using THz-FTIR spectroscopy and found an ion-specific effect on the phase behavior. Gelation in biopolymers is a rather complex phenomenon as it involves several interactions: hydrophobic, ionic, H-bond and van Der Waals, the relative magnitude of these interactions being highly system specific.[43] When incubated at $T>T_M$, the denatured protein molecules start aggregating followed by a hydrophobic collapse and eventually gelation. We identify that the ν_{HB} of the protein suffers a substantial blue shift in presence of Mg^{2+}, along with a disappearance of its rattling mode and the fact that $\Delta\Delta\alpha_{Gel}(\nu) > 0$ clearly identifies the role of the ion to squeezing of hydrogel to form composed gel as it directly binds to the protein surface. Fluorescence anisotropy results also corroborate the de-solvation of hydrogel. However, in LLPS phases no such collapse occurs. Time resolved anisotropy results conclude that the internal environment of the LLPS microphases does not show significant alteration at $T>T_M$.

The entropy gained during the RT LLPS formation perhaps overwhelms the energy profit of gelation.

Table 4.III.1. Steady state fluorescence anisotropy of ANS in different environments after incubation at two different temperatures: 25˚C and 70˚C (λ_{ex} =365 nm and λ_{em} =520 nm).

Sample	25˚C	Anisotropy 70˚C
buffer	0.06	0.04
BSA	0.18	0.03
BSA_$MgCl_2$	0.16	0.03
BSA_$LaCl_3$	0.18	0.17
BSA_$HoCl_3$	0.17	0.22

Table 4.III.2. Time-resolved fluorescence anisotropy decay parameters of ANS different environments and incubation temperatures (λ_{ex} =365 nm; λ_{em} 520 nm).

Sample	25˚C		70˚C	
	τ_{fast} (ns)	τ_{slow} (ns)	τ_{fast} (ns)	τ_{slow} (ns)
buffer	0.08	-	0.07	-
BSA	0.21	18.8	0.09	-
BSA_$MgCl_2$	0.25	20.5	0.11	-
BSA_$LaCl_3$	0.35	33.2	0.25	19.3
BSA_$HoCl_3$	0.42	35.1	0.20	22.0

Table 4.III.3. Fitted parameter unperturbed center frequency (ν_c) of various modes (hydrogen bond stretching mode (ν_{HB}), ion rattling motion (ν_{rtl}), and water librational motion (ν_{Lib})) obtained by fitting $\Delta\alpha(\nu)$ profiles using damped harmonic oscillator model for BSA protein in various environments (incubated at 25^0C).

Salt	ν_{HB} (cm^{-1})	ν_{rtl} (cm^{-1})	ν_{Lib} (cm^{-1})
0	72.7(56)	-	474.1(23)
NaCl	106.8(24)	358.6(31)	551.7(25)
$MgCl_2$	108.9(20)	341.9(9)	574.1(5)
$LaCl_3$	82.5(7)	391.8(14)	-
YCl_3	92.4(3)	385 (19)	-

$HoCl_3$	52.7(0.8)	154.5(13)	365.8(10)

Table 4.III.4. Fitted parameter unperturbed center frequency (ν_c) of various modes (hydrogen bond stretching mode (ν_{HB}), ion rattling motion (ν_{rtl}), and water librational motion (ν_{Lib})) obtained by fitting $\Delta\alpha(\nu)$ profiles using damped harmonic oscillator model for BSA protein in various environments (incubated at 70^0C).

Salt	ν_{HB} (cm^{-1})	ν_{rtl} (cm^{-1})	ν_{Lib} (cm^{-1})
0	143.8(4)	-	467.3(10)
NaCl	130.2(3)	441(5)	-
$MgCl_2$	189.4(14)	-	548(22)
$LaCl_3$	127(2)	460.3(3)	-
YCl_3	145.7(5)	402(7)	-
$HoCl_3$	145.6(11)	414.6(10)	-

4.V. References.

1. C. Ishimoto and T. Tanaka, *Phys Rev Lett.*, 1977, **39**, 474.
2. E. Jordan, F. Roosen-Runge, S. Leibfarth, F. Zhang, M. Sztucki, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *J. Phys. Chem. B*, 2014, **118**, 11365-11374.
3. F. Roosen-Runge, B. S. Heck, F. Zhang, O. Kohlbacher and F. Schreiber, *J. Phys. Chem. B*, 2013, **117**, 5777-5787.
4. F. Zhang, M. Skoda, R. Jacobs, S. Zorn, R. A. Martin, C. Martin, G. Clark, S. Weggler, A. Hildebrandt and O. Kohlbacher, *Phys Rev Lett.*, 2008, **101**, 148101.
5. M. K. Braun, A. Sauter, O. Matsarskaia, M. Wolf, F. Roosen-Runge, M. Sztucki, R. Roth, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2018, **122**, 11978-11985.
6. N. Begam, O. Matsarskaia, M. Sztucki, F. Zhang and F. Schreiber, *Soft Matter*, 2020, **16**, 2128-2134.
7. F. Zhang, G. Zocher, A. Sauter, T. Stehle and F. Schreiber, *J. Appl. Crystallogr.*, 2011, **44**, 755-762.
8. J. Li, R. Rajagopalan and J. Jiang, *J. Chem. Phys.*, 2008, **128**, 235104.
9. H. G. Schild and D. A. Tirrell, *J. Phys. Chem.*, 1990, **94**, 4352-4356.
10. O. Matsarskaia, M. K. Braun, F. Roosen-Runge, M. Wolf, F. Zhang, R. Roth and F. Schreiber, *J. Phys. Chem. B*, 2016, **120**, 7731-7736.
11. G. A. Ballou, P. D. Boyer, J. M. Luck and F. G. Lum, *J. Biol. Chem.*, 1944, **153**, 589-605.
12. C. Huggins and E. V. Jensen, *J. Biol. Chem.*, 1949, **179**, 645-654.
13. E. V. Jensen, V. D. Hospelhorn, D. F. Tapley and C. Huggins, *J. biol. Chem.*, 1950, **185**, 411.
14. K. Shimada and S. Matsushita, *Agric. Biol. Chem.,*, 1981, **45**, 1945-1952.
15. C. Nick Pace, R. W. Alston and K. L. Shaw, *Protein Sci.*, 2000, **9**, 1395-1398.
16. F. Tanaka, in *Polymer Physics: Applications to Molecular Association and Thermoreversible Gelation*, Cambridge University Press, Cambridge, 2011, pp. 97-127.
17. R. Saha and R. K. Mitra, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23661-23668.

18. O. Matsarskaia, F. Roosen-Runge, G. Lotze, J. Möller, A. Mariani, F. Zhang and F. Schreiber, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27214-27225.
19. S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 20749-20752.
20. K. Meister, S. Ebbinghaus, Y. Xu, J. G. Duman, A. DeVries, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 1617-1622.
21. D. Das Mahanta, N. Samanta and R. K. Mitra, *J. Phys. Chem. B*, 2017, **121**, 7777-7785.
22. T. Q. Luong, P. K. Verma, R. K. Mitra and M. Havenith, *Biophys. J.*, 2011, **101**, 925-933.
23. S. Pal, N. Samanta, D. Das Mahanta, R. K. Mitra and A. Chattopadhyay, *J. Phys. Chem. B*, 2018, **122**, 5066-5074.
24. P. Pyne and R. K. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
25. P. Pyne, N. Samanta, H. Gohil, S. Prabhu and R. K. Mitra, *Chem. Commun.*, 2021, **57**, 998-1001.
26. A.-M. Hermansson, ACS Publications, 1979, vol. 92, ch. 5.
27. Q. Tang, O. J. McCarthy and P. A. Munro, *J. Agric. Food. Chem.*, 1994, **42**, 2126-2130.
28. M. Bhattacharya, N. Jain and S. Mukhopadhyay, *J. Phys. Chem. B*, 2011, **115**, 4195-4205.
29. M. Cardamone and N. K. Puri, *Biochem J.*, 1992, **282 (Pt 2)**, 589-593.
30. E. Schönbrunn, S. Eschenburg, K. Luger, W. Kabsch and N. Amrhein, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 6345-6349.
31. P. Schienbein, G. Schwaab, H. Forbert, M. Havenith and D. Marx, *J. Phys. Chem. Lett.*, 2017, **8**, 2373-2380.
32. C. Klinkhammer, F. Böhm, V. Sharma, G. Schwaab, M. Seitz and M. Havenith, *J. Chem. Phys.*, 2018, **148**, 222802.
33. T. Dodo, M. Sugawa, E. Nonaka, H. Honda and S. i. Ikawa, *J. Chem. Phys.*, 1995, **102**, 6208-6211.
34. K. Takeda, A. Wada, K. Yamamoto, Y. Moriyama and K. Aoki, *J. Protein Chem.*, 1989, **8**, 653-659.
35. O. Matsarskaia, F. Roosen-Runge and F. Schreiber, *ChemPhysChem*, 2020, **21**, 1742-1767.
36. L. Donato, C. Garnier, J.-L. Doublier and T. Nicolai, *Biomacromolecules*, 2005, **6**, 2157-2163.
37. K. Ako, T. Nicolai and D. Durand, *Biomacromolecules*, 2010, **11**, 864-871.
38. Y. Levy and J. N. Onuchic, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 3325-3326.
39. L. R. Olano and S. W. Rick, *J. Am. Chem. Soc.*, 2004, **126**, 7991-8000.
40. S. Park, R. Barnes, Y. Lin, B.-j. Jeon, S. Najafi, K. T. Delaney, G. H. Fredrickson, J.-E. Shea, D. S. Hwang and S. Han, *Commun. Chem.*, 2020, **3**, 1-12.
41. I. Persson, P. D'Angelo, S. De Panfilis, M. Sandström and L. Eriksson, *Chemistry*, 2008, **14**, 3056-3066.
42. Y. Guo, J. Bae, Z. Fang, P. Li, F. Zhao and G. Yu, *Chem. Rev.*, 2020, **120**, 7642-7707.
43. M. T. Shibayama, T., *Adv. Polym. Sci.*, 1993, **109**, 1–62.
44. K. Matsumoto, N. Sakikawa and T. Miyata, *Nat. Commun.*, 2018, **9**, 1-7.
45. A. Erbas and M. Olvera de la Cruz, *ACS Macro Lett.*, 2015, **4**, 857-861.

Chapter 5

5. Lanthanides driven microphase behaviours of protein: ions hydration matter

Lanthanides, well-known rare earth elements (REE), acquire a major position in the modern world. Their growing application have prompted researchers to study their effects on human life. In presence of Lanthanide (III) cations, proteins exhibit intriguing phase behaviours such as liquid liquid phase separation (LLPS), reentrant condensation, etc which are highly dependent on concentration. In our study, we choose La^{3+}, Nd^{3+}, Gd^{3+}, Ho^{3+} and Lu^{3+} ions and BSA protein to observe the phase changes upon the addition of salt to the protein solution. We illustrate the phase diagram which is classified in three regimes; regime I (R_1) and regime III (R_3) are visually transparent whereas regime II signifies the turbid phase with two critical concentrations C* and C** and the middle point depicts as C_m. The phases are characterized by turbidity, zeta-potential and optical microscopy measurements. Finally, we explore the role of ion hydration in these processes using THz-FTIR (0.2-22.5 THz) measurements. Our experimental results clearly demonstrate that considering only the charge (III) is insufficient to explain the observed phases; ion hydration must also be taken into account.

5.I. Introduction.

Lanthanides (La to Lu) are the group 15 elements in the modern periodic table, also referred to as the "4f-block elements".[1] Chemically identical these elements belong to the rare earth elements (REE) and are claimed as "vitamins of industry".[2-4] Since the 19th century these elements hold huge applications in sectors such as clean-energy, military industry, medicine, agronomy and others.[5-9] They have unique magnetic and optical properties which make them essential in diverse fields such as catalyst,[10, 11] magnets,[12, 13] in photoluminescent devices,[14] as polishing powders for lenses, optical glasses for lasers,[15] and light-converting phosphors in fluorescent lamps,[16] in solar cells,[3] in displays,[17] etc. These numerous applications have led to their presence in various environments, ranging from microbes (such as bacteria and fungi) and animals to the human body.[18] Water soluble compounds of these REEs can migrate through the soil and contaminate ground and surface water resources.[19, 20] Therefore, in recent, researchers prompt to learn their

effect on our ecosystem. Moreover, Lanthanide ions are found to regulate the transport and release of synaptic transmitters and block membrane receptors through their interactions with various enzymes.[21] Consequently, molecular level understanding of their behaviour with biomolecules needs to be explored.

F. Schreiber's group proposes in their recent studies that lanthanides (La^{3+} and Ho^{3+}) exhibit intricate phase behaviors when interacting with negatively charged globular proteins.[22-24] The phases include reentrant condensation of proteins in the presence of ions,[25] ion concentration dependent LLPS,[26] protein crystal growth etc. These phase behaviours have been found to be concentration dependent and are rarely observed in mono- or divalent ions. They first report in their study[25] that addition of multivalent salt in protein aqueous solution leads reentrant condensation which is different than DNA. Through small-angle x-ray scattering measurements, they show that re-entrant condensation of proteins is caused by contrasting short-ranged electrostatic interactions between multivalent ions and proteins than DNA.[27] This re-entrant condensation phase behaviour was explained in terms of protein surface charge inversion.[28] Computational studies have also confirmed the re-entrant behaviour in terms of the protein-protein potential of mean force[29] and seems to be caused from the binding of multivalent ions.[30] LLPS was found as a metastable phase under appropriate experimental conditions, representing a hydration entropy-driven transition.[31] All these studies proposed multivalent charge effect to be the key factor for such phase behaviours of globular protein where the scope of hydrophilic hydration has not been considered.

In the previous two chapters, we limit our study to the specific salt concentration (C_s) and protein concentration (C_p) ratio (23.33) which belongs to the particular microscopic phase separation (LLPS) of the globular protein BSA. This phase separation is notably induced by certain trivalent ions, particularly lanthanides (La^{3+} and Ho^{3+}) showcasing intriguing outcomes in protein hydration regardless of their charge. In this chapter, we investigate the effects of other lanthanides on the phase behaviour of BSA. We choose chloride salts of La^{3+}, Nd^{3+}, Gd^{3+}, Ho^{3+} and Lu^{3+} ions and BSA protein and illustrate the phase diagrams for each salt separately. For all these salts, we observe re-entrant condensation of BSA. Turbidity and zeta-potential measurements are done to characterise each phase and optical microscopy is used to detect the micro-environments of a particular phase. It is noteworthy that all the lanthanide salts are water soluble, however, their

hydration features are less focused. We finally measure the salt hydration and determine the ion-protein solvation in these different phases using THz (50-750 cm^{-1}; 1.5-22.5 THz) FTIR measurements. THz absorption spectroscopy is a state-of-the-art as well as label-free technique. The method provides explicit information of the changes in water network by probing fluctuations in collective water dipole moments.[32-34] The results strongly conclude that these microscopic phase behaviours of proteins are driven not only by the changes in the protein surface charge potential but also by the involved ion-protein hydration.

5.II. Materials and Methods.

Salts, lanthanum chloride ($LaCl_3$), neodymium chloride ($NdCl_3$), gadolinium chloride ($GdCl_3$), holmium chloride ($HoCl_3$) and lutecium chloride ($LuCl_3$) were purchased in anhydrous form from Sigma Aldrich with ≥99% trace metal basis purity of analytical grade. Protein bovine serum albumin (A4503) was also purchased from Sigma Aldrich without further purification. Milli-Q water was filtrated with a 0.2 um filter prior to prepare each sample. All the measurements were carried out in aqueous medium at 25 ^{0}C. The details of all the used instruments are described in the section 2.III and the experimental analysis details are also mentioned prior in the section 2.IV.

5.III. Results and Discussions.

Visual turbidity detection: We first visually detect the change of phases with the addition of salts in the protein solution and illustrate in phase diagram (figure 5.III.1). At a time, we take a fixed protein concentration (C_p) and gradually increases the salt concentration (C_s). After a critical C_s value, the initial clear and transparent protein solution gets turbid. Turbidity increases at certain range of C_s and then gets faint again till a critical salt concentration, C_s and the protein solution re-entrant into a clear and transparent phase. These phases are dependent on both C_p and C_s and we observe the salt critical concentration values are also ion-specific. The phase diagrams are classified in three regimes; regime I (R_1) and regime III (R_3) are visually transparent whereas regime II signifies the turbid phase. The starting (C*) and the end (C**) points of the regime II are the two critical concentrations and the middle point of that particular zone is designated as C_m. The thickness of regime II is found to be different for each salt. However, for $LuCl_3$ the upper critical salt concentration C** is almost similar throughout the entire protein concentration range. We specify a part of the regime II as *metastable* phase where we detect bluish translucent

appearance of the solution. In figure 5.III.2A, we represent each regime of the phase diagram for $HoCl_3$ salt. Further we focus on these phases of protein in concentration 40 mg ml^{-1} throughout the study.

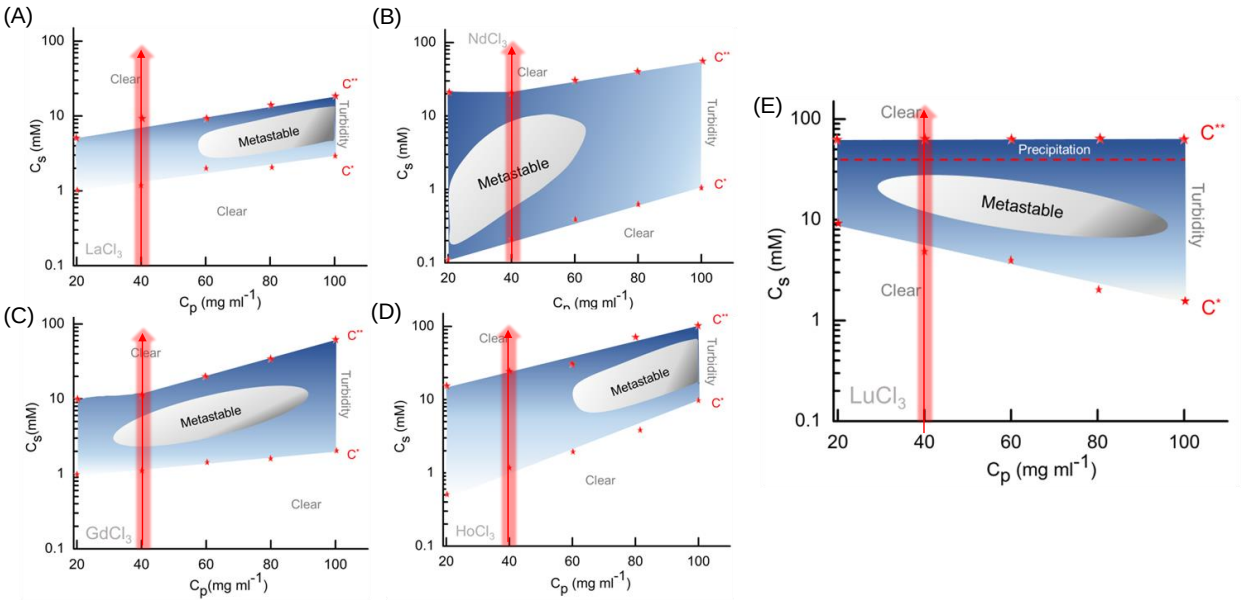


Figure 5.III.1. Phase diagrams (A-E) from the visual detection of different phases of the BSA protein at a particular concentration (C_p) with increasing salt concentration (C_s). Here the blue region signifies turbid appearance of the protein solution where specifically mentioned gray area stands for the bluish translucent appearance of the solution. C* and C** are the two critical concentrations after and before that particular concentration turbidity exists, respectively. Red arrows are given to guide viewer's eye.

Turbidity and zeta-potential measurement: UV-visible spectrophotometer is used to experimentally measure the turbidity of bare protein solution and in all five different phases (R_1, C*, C_m, C** and R_3). Absorption profile at 400 nm wavelength (figure 5.III.2B) exhibits the spectroscopic results of turbidity which follows almost identical trend of the phase diagram. To determine how the phases are related to the alteration of protein surface potential, we measure the zeta-potential (ζ) of each protein sample at room temperature (RT). For all the lanthanides, we observe charge neutralization after the C* region followed by charge reversal of the negatively charged protein surface at the C_m region. After the is C_m reached, ζ shows positive value at high salt concentrations and the overall ζ profile follows a sigmoidal feature (figure 5C).

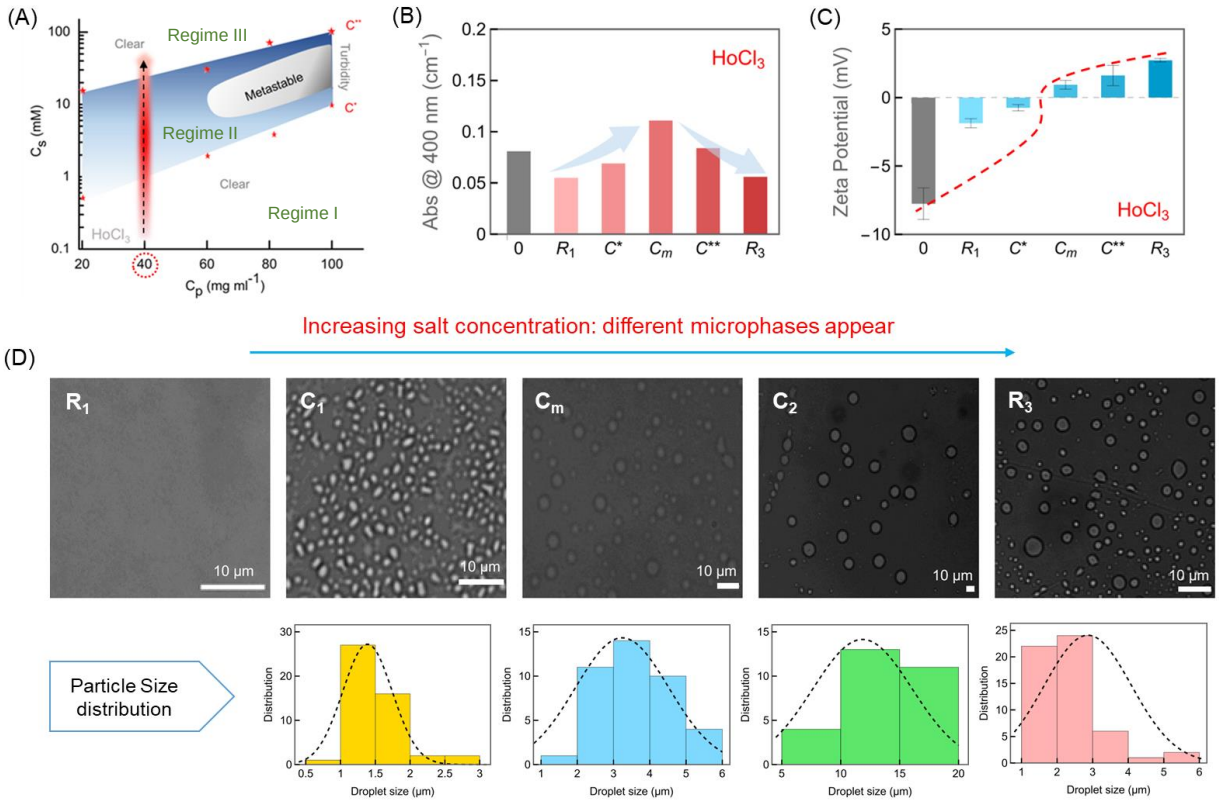


Figure 5.III.2. (A) Representative phase diagram of BSA protein against $HoCl_3$ salt concentration where three regions are mentioned as regime I, II and III depending on appearance and disappearance of turbidity. (B) Turbidity (absorption @ 400 nm) profile of 40 mg ml^{-1} bare protein (represents by '0') and of 5 different consecutive microphases of the protein depending on the $HoCl_3$ salt concentration. (C) ζ potential of the BSA solutions at pure and microphases states. The error bars signify standard deviations. (D) DIC images of protein in $LaCl_3$ using gray color scheme at 10 µm scale. Representative size distribution plots (second row panel) of liquid droplets correspond to each state.

Morphological observation of the phases: In order to obtain insight into the microstructures of the phases appeared, we observe them under a differential interference contrast (DIC) optical microscope. After 24 hours of incubation at RT, the R_1 phase for all the lanthanides appears clear like bare BSA solution under the microscope. It indicates a homogeneously dispersed phase of monomeric protein molecules. The micrographs (figure 5.III.2D) depict the emergence of droplets from the C*(C_1) phase, confirming the formation of protein-rich liquid droplets embedded in a dilute protein phase. Interestingly, droplets are still registered in the visually clear R_3 region. The statistical size distribution profiles of each state (lower panel of figure 5.III.2D) ascribe the gradual increase of the droplet size with the increment of salt concentration up to C**(C_2). However, in R_3

region the number of smaller size droplets increases indicating microscopic regime III at higher salt concentrations.

Far-IR-THz hydration study of Lanthanides: As these ions are beyond the scope of Hofmeister series,[35] a fundamental question arises how they behave towards proteins. To address this question, it is of utmost importance to understand how Ln^{3+} ion solvates. Here, we have used THz spectroscopy to study the hydration behavior of Ln^{3+} salts in 50-700 cm^{-1} (1.5-21 THz) frequency window. The experimentally measured absorption co-efficient $\alpha(\nu)$ is the linear response of the collective dipole fluctuations in the sample (*details in the section 2.I.C*). Typically, water exhibits significant absorption in this fingerprint region.[32] The presence of any solute perturbs the water structure, leading to either "THz defect" or "THz excess" phenomenon, depending on how the solute interacts with the water structure. Here we measure absorption co-efficient $\alpha(\nu)$ of all the

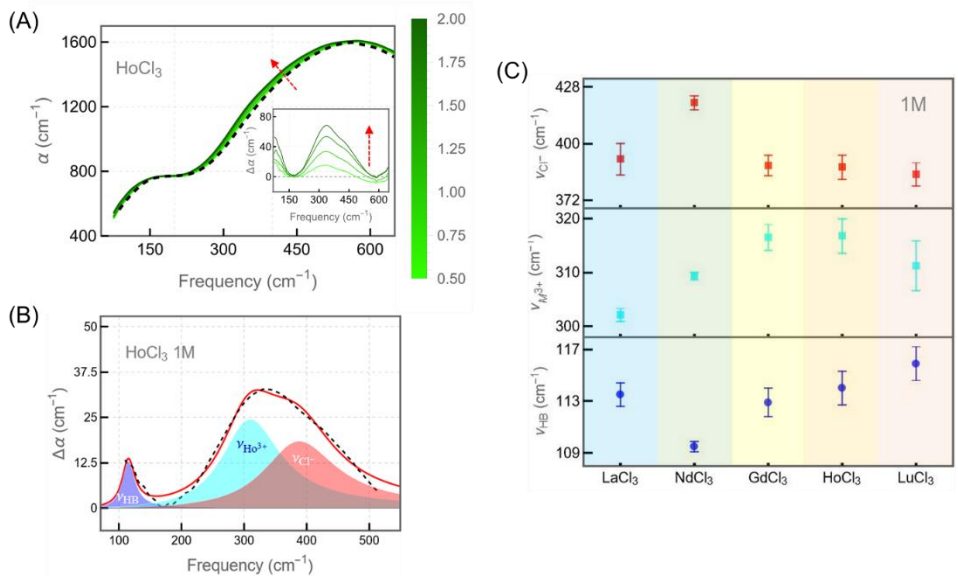


Figure 5.III.3. (A) Representative concentration dependent α(ν) profile of $HoCl_3$ salt aqueous solutions with the corresponding Δα(ν) profile as inset. (B) Δα(ν) profile of 1M $HoCl_3$ salt solution into three vibrational modes; intermolecular H-bond stretching mode (ν_{HB}) and two rattling modes ($\nu_{Ho^{3+}}$ and ν_{Cl^-}). (C) Comparative representation of all three vibrational mode frequencies of all the lanthanide salts. Error bar stands for standard deviations.

aqueous solutions of Ln^{3+} and the $\alpha(\nu)$ profiles (figure 5.III.3A) suggest certain rise at particular region with increasing salt concentration. To get more quantitative insight, we plot $\Delta\alpha(\nu)$ ($=\alpha_{salt+water}(\nu) - \alpha_{water}(\nu)$) which clearly demonstrate (figure 5.III.3A inset) the increment in

absorption at this frequency window. Deconvolution of water $\alpha(\nu)$ profile by damped harmonic oscillator model (DHO),[36, 37]

$$\alpha(\nu) = \sum_{i=1}^{N} \frac{a_i \omega_{0,i} \nu^2}{4\pi^3 \left[\frac{\nu^2 \omega_{0,i}^2}{\pi^2} + \left(\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2} - \nu^2 \right)^2 \right]} \tag{5.1}$$

where a_i, $\omega_{0,i}$, and $\nu_{d,i}$ stand for the amplitude, the damping width, and the centre frequency of the i^{th} damped harmonic oscillator mode, respectively (*see the section 2.I.D*). We find the H-bonds stretching mode for bulk water (ν_{HB}) to appear at ~129 cm^{-1} and a restricted rotational mode of water molecules (libration or ν_{Lib}) to appear at ~565 cm^{-1}. We use the same DHO model to fit the $\Delta\alpha(\nu)$ profile of each salt solution. Dissection of $\Delta\alpha_{salt}(\nu)$ profile depicts four different vibrational modes suggesting two extra positive modes to appear in the region between ~300-430 cm^{-1} window signifying ion-induced contributions on water structure and is commonly defined as the *ion-rattling modes* (ν_{rtl}).[38, 39] It is needed to mention that for further analysis of $\Delta\alpha_{salt}(\nu)$, we fix the frequency window to 50-550 cm^{-1} as we are only concerned about the ion mediated alteration in water structure (table 5.III.1). In figure 5.III.3B, the dissection of $\Delta\alpha_{HoCl_3}(\nu)$ profile of 1M concentration depicts three different vibrational modes; two rattling modes for cation ($\nu_{M^{3+}}$) and anion (ν_{Cl^-}) respectively and another one is intermolecular H-bond stretching mode (ν_{HB})

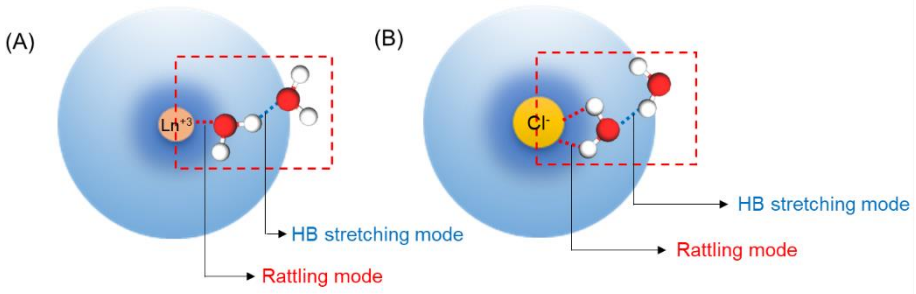


Figure 5.III.4. Cartoon representation of the H-bond stretching mode and rattling modes of Ln^{+3} (A) and Cl^- (B) ions.

where effect of both ions can be encountered (figure 5.III.4). The rattling components are attributed to the motions of both the corresponding anions and cations along with their solvation shells. In figure 5.III.3C, we show the peak frequencies of the different vibrational modes of 1M aqueous solution of each salt. For all of the salts the ν_{HB} mode belongs to the ~112-116 cm^{-1} which is much

red shifted than bulk water. It implies that the lanthanide salts make the intermolecular H-bonds between the neighbouring water molecules weaker than that in bulk water. Our observations indicate that the solvation behavior of the salts depends on the specific type of Ln^{3+} ion present.

Hydration study of protein phases: Our next concern is to find out how protein and ion hydration are responsible for the observed phase behaviours of the protein. We deduce the $\Delta\alpha(\nu)$ ($=\alpha_{Protein+salt+water}(\nu) - \alpha_{salt+water}(\nu)$) profile of each phase for each lanthanide salts. Figure 5.III.4A represents the difference absorption profiles for the bare protein and for all the phases appeared in presence of $HoCl_3$. The $\Delta\alpha(\nu)$ profile (figure 5.III.4A) for $HoCl_3$ clearly shows distinct hydration behaviour for the different phases in compare to the bare protein. In order to quantify the ion induced perturbation in protein hydration we further dissect the $\Delta\alpha(\nu)$ profile by the DHO model (eq. 5.1) fit and get the contributions arising from the different vibrational modes of water dipole(s) and ion charge dipole(s). We observe three different contributions in the form of: translational H- bond motion (ν_{HB}), restricted rotational motion of water molecules (ν_{Lib}) and

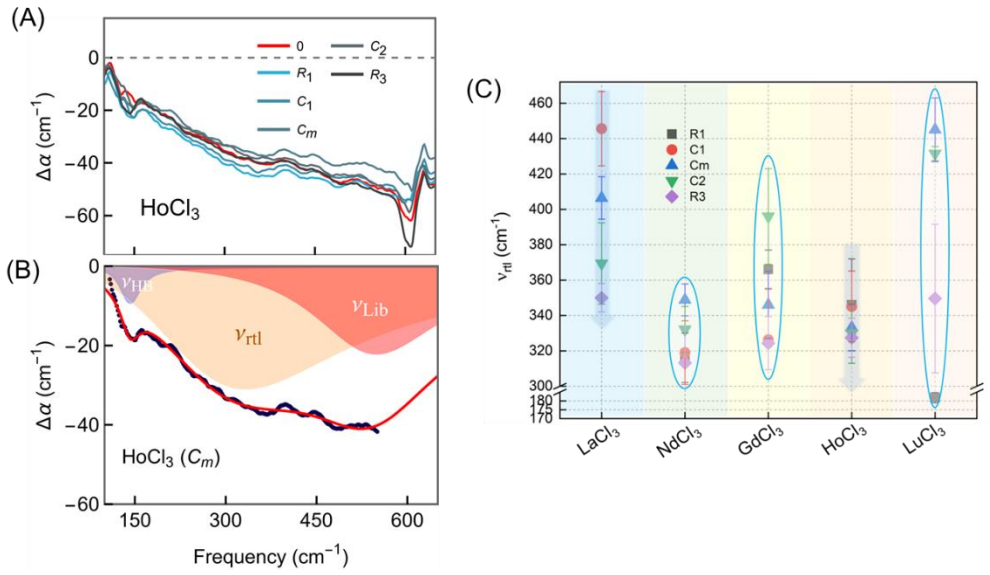


Figure 5.III.4. (A) Frequency dependent difference absorption, Δα(ν) profile of the different phases of protein in absence and presence of $HoCl_3$ salt. (B) Representative dissection of the Δα(ν) profile of protein-$HoCl_3$ system at C_m phase into the HB stretching mode (ν_{HB}), water librational motion (ν_{Lib}), and ion rattling mode (ν_{rtl}), respectively. Red line stands for the fitted curve over deep blue experimental data points. (C) Comparative representation of the different phases for all examined protein-salt solutions against their ion rattling mode. The error bar signifies the standard deviations for the respective system, and the downward arrows and oval shaped markers stand for linear trend and non-linearity, respectively.

ion induced contribution on protein hydration (ν_{rtl}) (figure 5.III.4B; table 5.III.2). It is to be noted here that we do not get the individual ion rattling contribution, rather a broad spectrum is dissected from the total $\Delta\alpha(\nu)$ profile. Considering that $\alpha_{THz}(\nu)$ is composed of contributions from various components, $\Delta\alpha(\nu)$ can be described as: $\alpha(\nu) = \sum \phi_i \alpha_i(\nu)$, where ϕ_i denotes the volume fraction and $\alpha_i(\nu)$ is the frequency dependent absorption co-efficient of the *i*-th component. Each hydration contribution exhibits distinct, non-overlapping features in the THz fingerprint regions. This characteristic allows for the individual analysis of these contributions, facilitating the determination of the solvation status of each constituent(s).

Therefore, we can write the $\alpha(\nu)$ as:

$$\alpha_{p+w}(\nu) == \phi_p \alpha_p(\nu) + \phi_{ph}\alpha_{ph}(\nu) + \phi_b^p \alpha_{bulk}(\nu) \tag{5.2}$$

$$\alpha_{s+w}(\nu) = \phi_s \alpha_s(\nu) + \phi_{sh}\alpha_{sh}(\nu) + \phi_b^s \alpha_{bulk}(\nu) \tag{5.3}$$

$$\alpha_{p+s+w}(\nu) = \phi_p' \alpha_p(\nu) + \phi_{ph}'\alpha_{ph}(\nu) + \phi_s'\alpha_s(\nu) + \phi_{sh}'\alpha_{sh}(\nu) + \phi_b'\alpha_{bulk}(\nu) \tag{5.4}$$

where the subscripts '*p*', '*s*' and '*w*' stands for protein, salt and water, respectively; '*ph*' and '*sh*' denote protein and salt hydration, respectively. Notably, all the ions and proteins individually exhibit very low absorbance in this frequency window, allowing for the neglect of the α_s and α_p terms. Thus, the expression for the explicit protein hydration, $\Delta\alpha_p(\nu)$ takes the form:

$$\Delta\alpha_p(\nu) = \alpha_{p+w}(\nu) - \alpha_w(\nu) = \phi_{ph}\alpha_{ph}(\nu) + (\phi_b^p - 1)\alpha_{bulk}(\nu) \tag{5.5}$$

$$\Delta\alpha_p(\nu) = \alpha_{p+s+w}(\nu) - \alpha_{s+w}(\nu)$$

$$= \phi_{ph}^{'}\alpha_{ph}(\nu) + (\phi_{sh} - \phi_{sh}')\alpha_{sh}(\nu) + (\phi_b^{'} - \phi_b^s)\,\alpha_{bulk}(\nu) \tag{5.6}$$

The first and third terms in equation 5.6 represent ion induced change in protein solvation related to the particular phase. The second term, which basically is the extra contribution added in equation 5.5, signifies alteration in collective ion hydration during phase change of protein. Rattling motion of ions in its solvation cage has direct impact in ion-water interaction contributing to ion solvation. So, the rattling mode (ν_{rtl}) indicates the ion contribution in protein solvation ($\Delta\alpha_p(\nu)$).

We summarize the ν_{rtl} frequencies (figure 5.III.4C) of the lanthanide salts to illustrate the contribution of ion solvation in all the observed phases of BSA. Absence of any particular mode for a phase indicates ion solvation remains unaltered from bare ion hydration during the phase transition; for example, R_1 phase of $LaCl_3$ and C_1 phase of $LuCl_3$. Interestingly, only for $LaCl_3$ and $HoCl_3$, we observe a linear red shift of ν_{rtl} subsequently from R_1 to R_3 phase (figure 5.III.4). The red shift implies the ion-water bonding strength to decrease as the protein-ion interaction increases; it strongly signifies that ions lose their hydration. This phenomenon corroborates very well with the protein surface potential inversion. However, for $NdCl_3$, $GdCl_3$ and $LuCl_3$, a blue shift of the ν_{rtl} is observed at C_m phase from C_1 and further redshift is detected at the R_3 phase. Interestingly, such contrasting ion hydration features demonstrate that the C_m phase for these three ions is altogether different than that for La^{3+} and Ho^{3+}. As observed earlier in the phase diagrams (figure 5.III.1), the C_m phase for Nd^{3+}, Gd^{3+} and Lu^{3+} (figure 5.III.1 (B, C, E)) belong to the metastable phase (bluish translucent) at protein concentration 40 mg ml^{-1}. However, at that specific protein concentration, C_m phase for La^{3+} and Ho^{3+} is not in their metastable phase, which occurs at higher protein concentrations for these ions. We can conclude that in the metastable phase, Ln^{3+} ions are getting caged in their hydration shell.

5.IV. Conclusions

In summary, all the Lanthanides exhibit microscopic self-aggregation of BSA at a certain concentration range. Protein re-entre into a homogeneous phase after reaching an ion-specific critical salt concentration. Our experimental findings demonstrate that these microscopic phase behaviours of proteins are driven not only by changes in the protein surface charge potential but also by ion-protein hydration. Additionally, contrasting ion hydration separates the metastable phase from others at a particular protein concentration.

Table 5.III.1. The listed values of the fitted unperturbed center frequency (ν_c): Respective values for hydrogen bond stretching mode, for rattling motions of cations, and anions of each aqueous salt system.

Sample	Conc. (M)	ν_{HB} (cm^{-1})	$\nu_{M^{3+}}$ (cm^{-1})	ν_{Cl^-} (cm^{-1})
$LaCl_3$	1	115.9(2.0)	313.2(4.7)	392.3(7.8)
	1.5	115(1.1)	303.2(2.0)	399.8(2.5)
	2	113.5(0.9)	302.1(1.2)	406.3(1.8)
$NdCl_3$	1	112.3(0.8)	306.5(2.1)	420.2(3.4)
	1.5	112.1(0.4)	305.1(1.1)	417.7(2.0)
	2	109.5(0.4)	309.3(0.7)	423.4(1.5)
	2.5	111.9(0.6)	308.1(1.1)	418.5(2.2)
$GdCl_3$	1	116.7(2.2)	319.5(4.1)	389.2(5.1)
	1.5	115.6(1.2)	316.5(2.8)	396.6(4.5)
	2	112.9(1.1)	316.5(2.4)	402.3(4.7)
	2.5	113.6(0.8)	319.7(1.7)	413.5(4.3)
	3	112.4(0.7)	319.8(1.5)	417.1(4.4)
	3.5	112.4(0.6)	322.9(1.2)	430.5(4)
	4	111.9(0.6)	324.2(0.9)	440.9(3.8)
$HoCl_3$	1	115.1(1.6)	310.0(3.7)	388.3(6.0)
	1.5	113.6(1.4)	316.5(3.2)	403.6(6.4)
	2	114.0(1.3)	316.7(3.2)	399.8(6.4)
$LuCl_3$	1	117.9(2.2)	318.8(5.7)	384.8(5.7)
	1.5	116.2(1.7)	310.8(4.9)	386.1(5.8)
	2	115.9(1.3)	311.2(4.6)	390.7(6.3)
	2.5	115.5(1.4)	312.6(4.7)	393.9(7.5)
	3	114.5(1.3)	312.7(4.4)	393.3(7.5)
	3.5	114.7(1.3)	319.2(4.7)	409.3(11.2)
	4	114.3(1.2)	323.8(4.1)	425.0(13.1)

Table 5.III.2. The listed values of the fitted unperturbed center frequency (ν_c): Respective values for hydrogen bond stretching mode, for rattling motions of ions, and for water librational motion of each protein-salt system at different microscopic phases of BSA protein.

HB Stretching (cm^{-1})

Sample	R_1	C_1	C_m	C_2	R_3
$LaCl_3$	143.3(2)	147.8(2)	145.7(6)	144.1(3)	257.2(27)
$NdCl_3$	152.4(3)	156.8(4)	144.2(1)	142.6(1)	143.9(3)
$GdCl_3$	138.6(2)	159.7(5)	146(2)	146(5)	143(1)
$HoCl_3$	141.2(2)	143.1(3)	142.6(2)	149.4(2)	138.5(1)
$LuCl_3$	107.7(0.4)	117.2(4)	95.7(3)	104.6(2)	85(5)

Ion rattling (cm^{-1})

Sample	R_1	C_1	C_m	C_2	R_3
$LaCl_3$	-	445.6(21)	406.4(12)	369.4(23)	350(8)
$NdCl_3$	316.1(14)	319(18)	348.8(9)	332.1(13)	313.2(21)
$GdCl_3$	366(11)	326.3(29)	346(19)	396(27)	324.3(15)
$HoCl_3$	346(26)	345.1(20)	333.2(13)	330(17)	327.4(11)
$LuCl_3$	182(3)	-	445(18)	431.5(4)	349.6(42)

Libration (cm^{-1})

Sample	R_1	C_1	C_m	C_2	R_3
$LaCl_3$	617.5(35)	-	-	583.6(16))	541(24)
$NdCl_3$	568.3(34)	584.1(68)	544.8(9)	563.5(19)	572.3(38)
$GdCl_3$	588.8(10)	566.8(65)	608.6(22)	593(9)	556.3(23)
$HoCl_3$	578.9(54)	591(52)	548.8(18)	547.8(32)	561(16)
$LuCl_3$	491.8(15)	508(37)	-	-	531.8(33)

5.V. References.

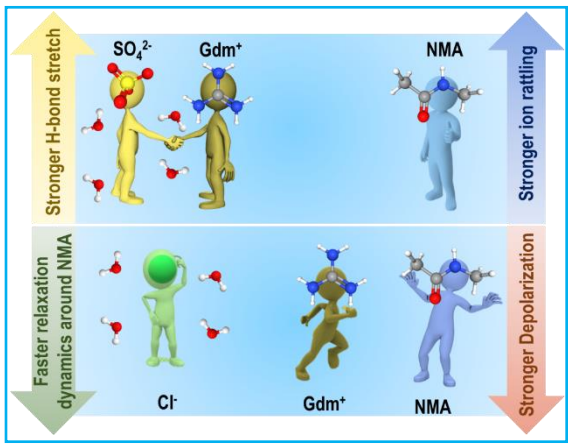
1. S. a. Atkins, *Inorganic Chemistry*, Oxford University Press, New York, 5th edn., 2010.
2. S. V. Eliseeva and J.-C. G. Bünzli, *New J. Chem.*, 2011, **35**, 1165-1176.
3. H.-Q. Wang, M. Batentschuk, A. Osvet, L. Pinna and C. J. Brabec, *Adv. Mater.*, 2011, **23**, 2675-2680.
4. Z. Zeng, Y. Xu, Z. Zhang, Z. Gao, M. Luo, Z. Yin, C. Zhang, J. Xu, B. Huang, F. Luo, Y. Du and C. Yan, *Chem. Soc. Rev.*, 2020, **49**, 1109-1143.
5. A. J. Hutchison, M. Speake and F. Al-Baaj, *Nephrol. Dial. Transplant.* , 2004, **19**, 1902-1906.
6. B. Lacour, A. Lucas, D. Auchère, N. Ruellan, N. M. de Serre Patey and T. B. Drüeke, *Kidney Int.*, 2005, **67**, 1062-1069.
7. W. W. Monafo, S. N. Tandon, V. H. Ayvazian, J. Tuchschmidt, A. M. Skinner and F. Deitz, *Surgery*, 1976, **80**, 465-473.

8. X. Pang, D. Li and A. Peng, *Environ. Sci. Pollut. Res. Int.*, 2002, **9**, 143-148.
9. R. Rajeswari, N. Islavath, M. Raghavender and L. Giribabu, *Chem. Rec.*, 2020, **20**, 65-88.
10. J.-C. G. Bünzli and I. Mcgill, in *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2018, pp. 1-53.
11. Y. Qiao and E. J. Schelter, *Acc. Chem. Res.*, 2018, **51**, 2926-2936.
12. H. Kohlmann, T. C. Hansen and V. Nassif, *Inorg. Chem.*, 2018, **57**, 1702-1704.
13. B. Shen, C. Yu, D. Su, Z. Yin, J. Li, Z. Xi and S. Sun, *Nanoscale*, 2018, **10**, 8735-8740.
14. N. Thejo Kalyani and S. J. Dhoble, *Renewable Sustainable Energy Rev.*, 2012, **16**, 2696-2723.
15. J. Lucas, P. Lucas, T. Le Mercier, A. Rollat and W. Davenport, in *Rare Earths*, Elsevier, Amsterdam, 2015, pp. 319-332.
16. H. Li, C. Chen, J. Jin, W. Bi, B. Zhang, X. Chen, L. Xu, D. Liu, Q. Dai and H. Song, *Nano Energy*, 2018, **50**, 699-709.
17. Y. Lu and B. Yan, *Chem. Commun.*, 2014, **50**, 15443-15446.
18. C. H. Evans, *Biochemistry of the Lanthanides*, Springer Science & Business Media, 2013.
19. M. Bau, A. Knappe and P. Dulski, *Geochemistry*, 2006, **66**, 143-152.
20. M. G. Lawrence, *Mar. Pollut. Bull.*, 2010, **60**, 1113-1116.
21. P. Caravan, *Chem. Soc. Rev.*, 2006, **35**, 512-523.
22. M. K. Braun, A. Sauter, O. Matsarskaia, M. Wolf, F. Roosen-Runge, M. Sztucki, R. Roth, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2018, **122**, 11978-11985.
23. O. Matsarskaia, S. Da Vela, A. Mariani, Z. Fu, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2019, **123**, 1913-1919.
24. O. Matsarskaia, F. Roosen-Runge, G. Lotze, J. Möller, A. Mariani, F. Zhang and F. Schreiber, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27214-27225.
25. F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, S. Zorn, R. A. Martin, C. M. Martin, G. F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *Phys. Rev. Lett.*, 2008, **101**, 148101.
26. F. Zhang, R. Roth, M. Wolf, F. Roosen-Runge, M. W. A. Skoda, R. M. J. Jacobs, M. Stzucki and F. Schreiber, *Soft Matter*, 2012, **8**, 1313-1316.
27. B. I. Shklovskii, *Phys. Rev. Lett.*, 1999, **82**, 3268-3271.28. F. Zhang, S. Weggler, M. J. Ziller, L. Ianeselli, B. S. Heck, A. Hildebrandt, O. Kohlbacher, M. W. A. Skoda, R. M. J. Jacobs and F. Schreiber, *Proteins*, 2010, **78**, 3450-3457.
29. C. Pasquier, M. Vazdar, J. Forsman, P. Jungwirth and M. Lund, *J. Phys. Chem. B*, 2017, **121**, 3000-3006.
30. A. I. Abrikosov, B. Stenqvist and M. Lund, *Soft Matter*, 2017, **13**, 4591-4597.
31. O. Matsarskaia, M. K. Braun, F. Roosen-Runge, M. Wolf, F. Zhang, R. Roth and F. Schreiber, *J. Phys. Chem. B*, 2016, **120**, 7731-7736.
32. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 12068-12073.
33. B. König, S. Pezzotti, S. Ramos, G. Schwaab and M. Havenith, *Biophys. J.*, 2023, **S0006-3495**, 00474-00475.
34. P. Pyne and R. K. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
35. F. Hofmeister, *Arch. Exp. Pathol. Pharmakol.*, 1888, **24**, 247-260.
36. T. Dodo, M. Sugawa, E. Nonaka, H. Honda and S. i. Ikawa, *J. Chem. Phys.*, 1995, **102**, 6208-6211.
37. G. Schwaab, F. Sebastiani and M. Havenith, *Angew. Chem. Int. Ed.*, 2019, **58**, 3000-3013.
38. D. Sharma, B. Das and A. Chandra, *J. Phys. Chem. B*, 2023, **127**, 6714-6725.
39. C. Klinkhammer, F. Böhm, V. Sharma, G. Schwaab, M. Seitz and M. Havenith, *J Chem Phys.*, 2018, **148**, 222802.

Chapter 6

6. Ion-pairing propensity in guanidinium salts dictates their protein (de)stabilization behaviour

Since the proposition of the Hofmeister series, guanidinium (Gdm) salts hold a special mention in protein chemistry owing to its contrasting effect on proteins depending on the counter anion(s). For example, while GdmCl is known to act as a potential protein denaturant, Gdm_2SO_4 offers minimal effect on protein structure. Despite the fact that theoretical studies reckon the formation of ion-paring to be responsible for such behaviour, experimental validation of this hypothesis is still in sparse. In this study we combine electrochemical impedance spectroscopy (EIS) and THz spectroscopy to underline the effect of GdmCl and Gdm_2SO_4 on a model amide molecule N-methyl acetamide (NMA). THz spectroscopy signatures water dipole fluctuations, and therefore could provide explicit picture on amide and salt interfacial hydration individually. Molecular dynamics (MD) simulation studies predict that Gdm_2SO_4 forms hetero-ion pairing in water which inhibits Gdm^+ ions to approach NMA molecules while in case of GdmCl the Gdm^+ ions directly interact with NMA. The experimental findings on ion hydration unambiguously authenticate this hypothesis. Our study establishes the fact that the propensity of ion-pairing in Gdm salts dictates their (de)stabilization effect on proteins.



6.I. Introduction.

Ions, the fundamental charged particles in colloidal domain, beyond the usual non-specific coulombic interactions can also offer specific effects to induce biomolecular stability.[1, 2] Around 130 years back Franz Hofmeister classified a number of cations and anions into a series depending on their specific effect on protein solubility.[3, 4] In general, a particular ion receives a Hofmeister ranking based on its impact on liquid water structure by virtue of its "making" and "breaking" property around proteins.[5-7] The rankings have been modified over ages as it is found that the effect

being not solely dependent on the ionic radii and surface charge density.[8] In a recent study, measurement of site-specific water structure and binding affinity has concluded a strong correlation between ion-specific effect (driven by Hofmeister series) and hydrophobic hydration.[9] In this context, the bulkiest cation guanidinium (Gdm^+) of this series holds a special mention[8, 10-12] due to its structural analogy with the arginine (Arg) residue[13] and its widespread applications in biotechnology and in chemical and material sciences. While GdmCl is one of the most commonly used protein denaturants, intriguingly several other salts of Gdm (e.g. Gdm_2SO_4) is mostly indifferent towards protein stability.[14] So, it remains a fundamental question to address that how guanidinium salts interact with protein; and more importantly, how the counterion(s) dictates its propensity towards protein denaturation.

Structural integrity and stability of biomolecules are intricately linked to the presence of ions in body fluids. In general, two perspectives can be considered when analysing the interactions between ions and biomolecules: (i) direct molecular interaction between ion and biomolecular active site(s) and (ii) hydration mediated indirect mechanism of ion-macromolecular binding. Several theoretical (computational) and experimental measurements have been performed to extract the actual mechanism of guanidinium (Gdm^+) ion driven protein denaturation, however, a molecular level understanding still remains elusive. *Gregory et al.* [15]concluded that polyatomic or molecular ions are exceptions in Hofmeister's original series in terms of specific ion effects (SIE) as Coulombic interactions between such ions and protein segments could not be modeled in terms of point charge, rather one needs to consider the active site of such ions with the highest charge density. Gdm^+ is one of such polyatomic ions consisting of three amino groups bound to the central carbon atom, thereby belonging to the class of planar Y-conjugated quasi-aromatic structures.[10] The most remarkable molecular properties of Gdm^+ is its capacity to form contact-ion pairs in water, even in spite of the likelihood of the presence of electrostatic repulsion.[16] Extensive ab-initio molecular dynamics (AIMD) simulations by *Vazdar et al.* have identified that nonuniform charge distribution in Gdm^+ allows the formation of that homo-ion pair through rigid facial plane in stacking manner.[17] Such homo-ion pairing affinity of Gdm^+ implicates destabilization and aggregation of Arg-rich proteins for the analogous pairing with positively charged Arg residue.[18] Experimental and simulation studies by Meson et al. concluded that Gdm^+ actively participates in preferential interaction with planer aromatic side chains over aliphatic side chains of proteins.[19, 20] However, 2D-IR results by Ding et al.[21] contrastingly found that Gdm^+ does not interact with

aromatic side chains in a peptide environment. Several other simulation and experimental studies suggest that Gdm mediated protein denaturation occurs as a result of altering electrostatic interactions by solvating the charged residues or by H-bonds interaction with the protein back bone.[22-24] Additionally, Gdm^+ ions has also been reported to interact with hydrophobic residues through direct Van der Waals preferential interactions.[25, 26]

Interestingly, counterion(s) has been found to play decisive roles on the protein denaturing activity of Gdm^+.[14, 27-29] *Graziano et al.*[14] has made statistical thermodynamic treatments to show that high charge density of SO_4^{2-} ions makes them to interact strongly with water and thereby making a solvent cavity creation process, a prerequisite to host unfolded protein segments, thermodynamically costly, rending the reported indifference of Gdm_2SO_4 towards protein denaturation. Previous simulation studies by *Mason et al.*[29] also concluded that in Gdm_2SO_4 solution the ions aggregated into mesoscopic (nanometer-scale) clusters which allow sulfate to reverse the denaturant power of Gdm^+.

All these previous reports suggest that one needs to consider a detailed ion-protein-water tri-partied interaction picture to divulge into a conclusive end to this ever-arguing phenomenon. In this aspect, it is interesting to note that Gdm^+ could offer unusual solvation scenarios depending on its immediate environment.[30, 31] Neutron diffraction with isotopic substitution (NDIS) experiments and MD simulations by *Mason et al.*[32, 33] indicate that Gdm^+ does not compete effectively for H-bonding interactions with water molecules and water density at inner hydration shell was found to be similar like bulk water. It is therefore of utmost importance to understand the nature of hydration in such tri-partied systems.

In this contribution we make a combined spectroscopic and MD simulation investigation to explore the contrasting effect GdmCl and Gdm_2SO_4 on N-methyl acetamide (NMA), a model single amide molecular unit that replicates the building unit of protein peptide back bone with small hydrophobic methyl segment, specifically focusing on the perturbation of both salt and amide hydration. We use time-domain THz spectroscopy (in the frequency window 0.2-1.5 THz) to underline the relaxation dynamics and THz-FTIR spectroscopy (in the frequency window:1.5-22.5 THz, 50 to 750 cm^{-1}) to determine the hydration structure in these systems. This non-ionizing, low energy radiation probes the inherent collective fluctuations of water dipoles and therefore can explicitly be used as label-free probe to extract small solutes as well as bimolecular hydration.[34,

[35] Moreover, ionic hydration, which is reflected in the form of rattling mode in the 350-450 cm^{-1} region, does not overlap with any other mode of water, and therefore provides an excellent avenue to detect explicit change in the ion-hydration.[36] We also use electrochemical impedance spectroscopy (EIS) to decipher the effect of counterions on the activity of Gdm^+, specifically at the electrode interface (in the frequency range of 1 Hz-1 MHz). We have used classical MD simulations to further elucidate the interactions between NMA and GdmCl and Gdm_2SO_4 salts. Specifically, we simulated the behaviour of NMA in bulk water, as well as in 2M Gdm_2SO_4 and GdmCl solutions. Through these simulations, we aimed to characterize the spatial distribution of molecules and provide additional insights into the hydration and solvation dynamics of NMA in the presence of these salts. Our MD simulations revealed significant differences in the interactions of NMA with the two guanidinium salts. Particularly, we observed much stronger interactions between Gdm^+ ions and the counterions (X^-) in the case of Gdm_2SO_4 as compared to GdmCl. Our combined experimental and simulation study firmly concludes that hetero-ion association in Gdm_2SO_4 inhibits its direct interaction with protein segments, which makes it an ineffective protein denaturant.

6.II. Materials and Methods.

The salts: Guanidinium chloride (GdmCl) and Guanidinium sulphate (Gdm_2SO_4) were purchased from Sigma Aldrich (Merck) with ≥99% purity of analytical grade. N-methylacetamide (NMA; Sigma-M26305) with ≥99% purity was used without further refinement. NMA is in crystal form at room temperature and it has melting point around 30-40 ^{0}C. We, therefore, first keep the NMA crystal in incubator at 40^{0}C for 30 minutes and prepare NMA aqueous solution maintaining the concentration 2M. We retain the salt concentration 4M in each salt solution. The details of all the used instruments are described in the section 2.III and the experimental analysis details are also mentioned prior in the section 2.IV.

We carried out classical molecular dynamics (MD) simulations using the GROMACS 2019.6 software suite[37] and the Charmm36 force field. The parameters for Gdm^+, Cl^-, NMA, and SO_4^{2-} were generated using the CGenFF server.[38] We used the TIP3P water model to solvate the systems. A 5nm simulation box was used to perform the MD simulations with periodic boundary conditions along all directions. 2M salt concentration (Gdm_2SO_4 & GdmCl) and 2M NMA concentration were used to perform the simulations. Table 6.III.1 provides comprehensive information about the

composition of each system. The steepest descent algorithm was used for energy minimization. The simulations were carried out in the NPT ensemble with a V-rescale thermostat[39] at 300 K and the Parrinello-Rahman[40] barostat at 1 bar. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) summation method[41] with a grid spacing of 0.16 nm and fourth-order cubic interpolation. Short-range electrostatic and van der Waals interactions were truncated at a cut-off distance of 1 nm. The LINCS algorithm was used to constrain all covalent bonds,[42] and the integration time step was set to 2 fs.

6.III. Results and Discussions.

Electrochemical impedance spectroscopy (EIS) measurements: We perform electrochemical impedance measurements to obtain the counter-ion (chloride and sulphate) distribution in the aqueous solutions of 4M GdmCl and Gdm_2SO_4. Figure 6.III.1A shows a schematic of a typical 3-electrode measurement scheme used in this study along with an equivalent circuit to extract various dielectric parameters.[43] Frequency dependent impedance (Bode plots) corresponding to the samples are shown in figure 6.III.1C. We focus on the linear response in the lower frequency range (up to 10^4 Hz) since this frequency window fingerprints the interfacial region of the electrode surface and the sample solution. When the electrode is biased with a positive charge the anions accumulate in its vicinity forming a static electrical layer at the electrode interface. Thickness of such interfacial layer and its corresponding electrical capacitance (referred to as the *interfacial capacitance*) is a potent marker of counter-ion behaviour in any ionic solution. We calculate the interfacial thickness using a parallel plate capacitor model; for 4 M GdmCl and Gdm_2SO_4 solutions the interfacial capacitance is estimated to be 1.5 and 0.7 μF, respectively, while the corresponding thicknesses are 16.6 and 49.5 Å (figure 6.III.1D). The interfacial region between an electrode and an electrolytic solution is highly inhomogeneous in terms of the spatial distributions of ions and solvent molecules. Following Gouy-Chapman-Stern (GCS) Model the total interfacial capacitance can be considered as a series combination of Stern capacitance (static charges) and diffuse capacitance. For highly concentrated ionic solutions Stern capacitance dominates and contributes explicitly towards the overall double layer capacitive response.[44] The observed fact that Gdm_2SO_4 offers a thicker interfacial layer than GdmCl can be explained in terms of accumulation of charges (ions) near the former interface, which could be due to the extent of hetero-ion paring in GdmCl and Gdm_2SO_4.

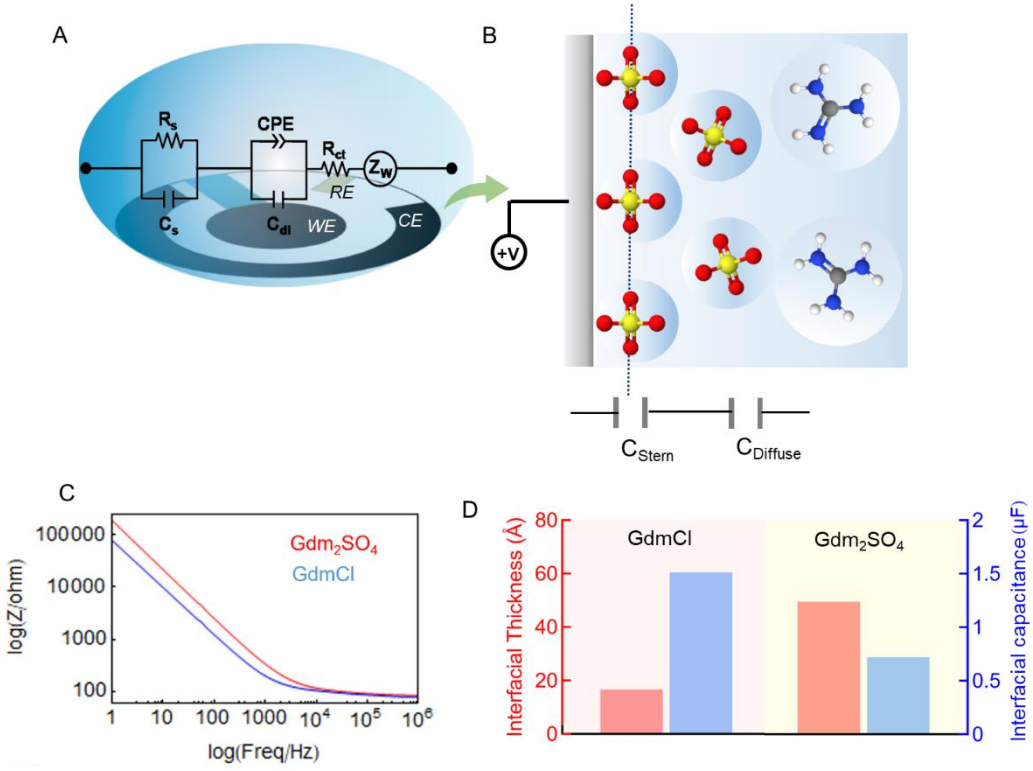


Figure 6.III.1. (A) Schematic representation of the 3-electrode EIS measurement system and the relevant equivalent circuit; (we can specify the circuit components here, R_s and C_s stand for solution source resistance and capacitance, respectively. C_{dl} is the interfacial capacitance in parallel with a constant phase element (CPE); R_{ct} in the charge transfer resistance in series with Warburg impedance (Z_w). (B) GCS model representation of static and mobile charge distribution. (C) Representative Bode plot of 4M aqueous solution of GdmCl and Gdm_2SO_4 salt. (D) Interfacial capacitance and corresponding thickness for GdmCl and Gdm_2SO_4.

MD simulation study: To obtain molecular level insight into the proposed ion-paring phenomenon we perform MD simulations in Gdm salts-water system to underline the various bipartite interactions. Radial distribution function (RDF) provides a quantitative measure of the local structural ordering of other molecules around a central reference molecule. We have calculated the RDF between the C atoms of both the Gdm^+ ions (figure 6.III.2A) to investigate the possibility of the formation of homo-ion pairs,[45] i.e., self-aggregation propensity of the Gdm^+ ions. Our analysis revealed distinct differences in the RDF peaks between GdmCl and Gdm_2SO_4 systems. The RDF peak for GdmCl appeared at 0.38 nm, while for Gdm_2SO_4, it appeared at 0.56 nm. This suggests

that the distance between Gdm^+ ions is shorter in GdmCl as compared to Gdm_2SO_4, indicating a stronger interaction between Gdm^+ ions in the former system. The shorter Gdm-Gdm distance in GdmCl suggests a higher propensity for Gdm^+ ions to form homo-ion pairs.

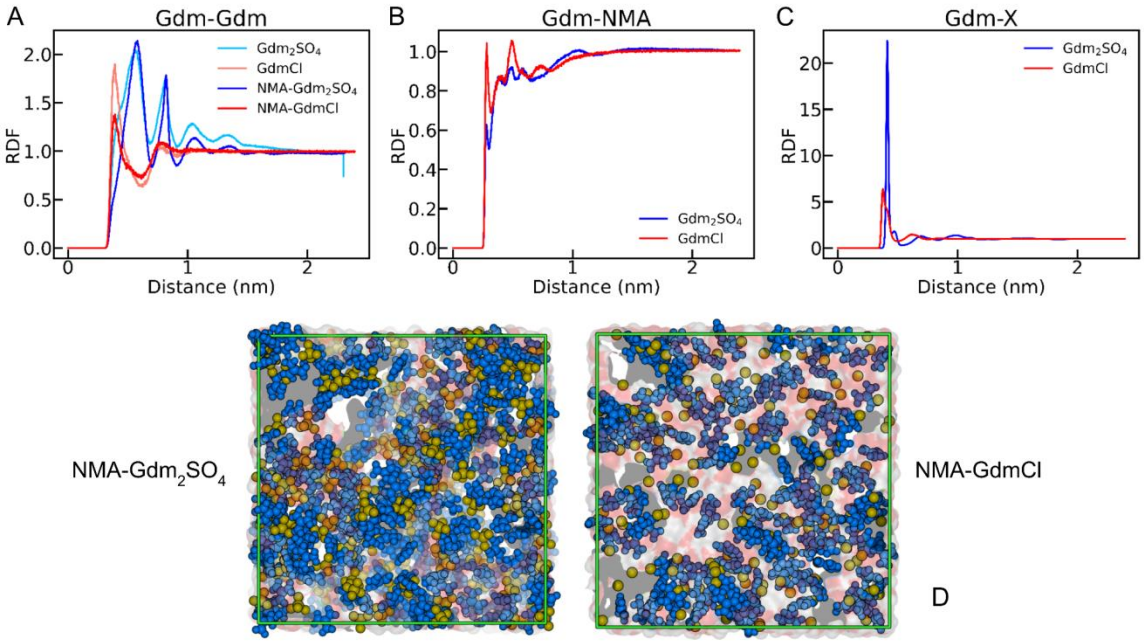


Figure. 6.III 2. (A) Carbon (Gdm)-carbon (Gdm) radial distribution function ($g_{C\text{-}C}$), calculated from MD simulations for Gdm_2SO_4, GdmCl, NMA-Gdm_2SO_4, NMA-GdmCl (B) Oxygen (NMA)-carbon (Gdm) radial distribution function ($g_{O\text{-}C}$), calculated from MD simulations for NMA-Gdm_2SO_4, NMA-GdmCl. (C) Carbon (Gdm)-X radial distribution function ($g_{C\text{-}X}$), calculated from MD simulations for NMA-Gdm_2SO_4, NMA-GdmCl. (D) Representative snapshot of NMA-Gdm_2SO_4 and NMA-GdmCl systems. Gdm^+ ions are shown in blue color and X (Cl^- & SO_4^{2-}) shown in olive color.

To further understand how the presence of NMA affects the ion-paring we introduce NMA into 2M Gdm-water solutions and the various tripartite interactions are estimated. We find that introduction of NMA does not affect the nature of the RDF for Gdm_2SO_4-water solution, however, in case of GdmCl, the peak height gets reduced compared to the GdmCl-water solution. This observation suggests that in GdmCl-NMA solution the Gdm^+ ions interact with the NMA molecules, whereas in Gdm_2SO_4-NMA solution they seemingly do not. We further calculate the RDF between the O atom of NMA and the N atom of Gdm^+ (figure 6.III.2B); the results clearly show that the interaction between NMA and Gdm^+ is prominent in the case of GdmCl, while it is poor in Gdm_2SO_4. This strong hereto-ion interaction between Gdm^+ and SO_4^{2-} does not allow

Gdm^+ ions to approach NMA; this explains the indifference of Gdm_2SO_4 towards proteins. To further support this claim, we calculate the RDF of the C atom of Gdm^+ and X (where X is the S atom for SO_4^{2-} and the Cl atom for Gdm_2SO_4 and GdmCl, respectively) (figure 6.III.2C). The results unambiguously confirm a much stronger interaction between Gdm^+ and X^- in the case of Gdm_2SO_4 compared to GdmCl.

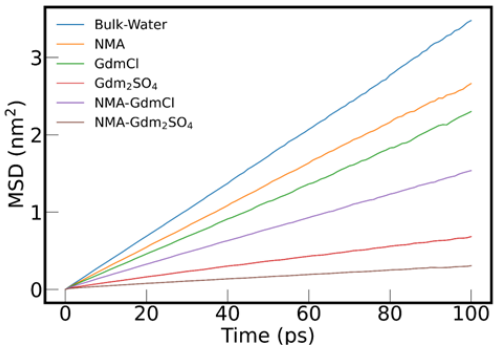


Figure 6.III.3. Mean square displacement of water for all the systems.

Representative snapshots of NMA in presence of Gdm_2SO_4 and in GdmCl (figure 6.III. 2D) reveals the microscopic insight of the system. We observe an extended cluster formation in case of Gdm_2SO_4 which results from the strong interaction between Gdm^+ and SO_4^2. Such cluster formation is expected to a significant dynamical slowdown of water in the system. To verify that we calculate the diffusion coefficient of water in all these systems; the corresponding mean square displacement plots are shown in figure 6.III.3. We observe that the diffusion coefficient indeed is decreased in Gdm_2SO_4 containing systems as compared to the others (Table 6.III.2). The dynamics of water in these systems is calculated using hydrogen bond autocorrelation function C(t) defined as[46, 47]

$$C(t) = \frac{<h(0)h(t)>}{<h>} \qquad (6.1)$$

The function C(t) is a correlation representing the likelihood of a hydrogen bond remaining intact at time t, given that it was unbroken at time zero, without considering any possible breakage during the intervening time. The behavior of C(t) provides insight into how hydrogen bonds undergo structural relaxation over time. In figure 6.III.4A, we have shown the variation of hydrogen bond autocorrelation function with time for all the systems. We fit the C(t) profiles using a double exponential model:

$$C(t) = \alpha_1 e^{-\frac{\tau_1}{t}} + \alpha_2 e^{-\frac{\tau_2}{t}} + c \tag{6.2}$$

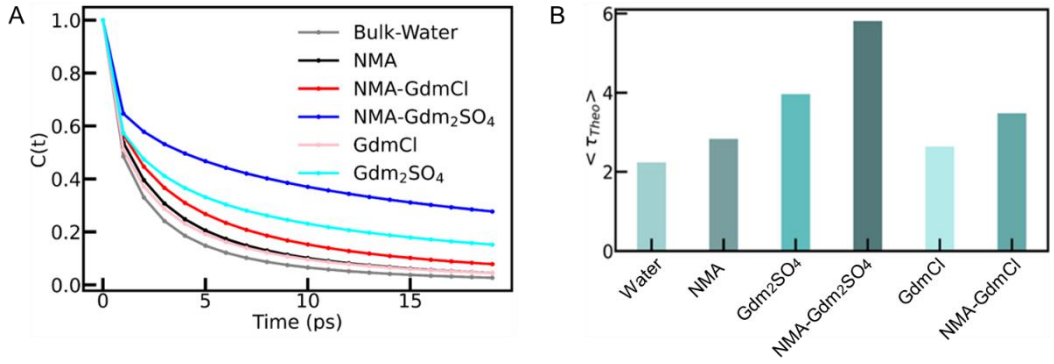


Figure 6.III.4. (A) Water-water hydrogen bond autocorrelation function for all the systems. (B) average timescales obtained from double exponential fitting of the autocorrelation function data. The timescale is significantly slower for Gdm_2SO_4 than GdmCl, which is expected as GdmCl acts as a water "structure breaker".

Some representative fittings are shown in figure 6.III.4A and the fitted parameters are depicted in Table 6.III.3. The average time constant $< \tau >= \frac{(\alpha_1\tau_1+\alpha_2\tau_2)}{(\alpha_1+\alpha_2)}$ is found that in the presence of Gdm_2SO_4, the hydrogen bond dynamics slows down significantly, moreover, the change is significant in presence of Gdm_2SO_4 compared to the less moderate change in GdmCl (figure 6.III.4B).

TTDS measurements: To experimentally determine the water dynamics around these ions, we measure THz absorption coefficient using TTDS (in the frequency window (0.2-1.5 THz) of these salts and NMA, both individually and in combination (figure 6.III.5A). Pure water shows a high absorption coefficient in the THz frequency window ($\alpha \sim 240\ cm^{-1}$ at 1 THz).[48, 49] As water is replaced by low absorbing solute molecules, the $\alpha(\nu)$ profiles decrease. Being a coherent measurement, TTDS allows to extract frequency dependent complex dielectric constant, $\tilde{\varepsilon}(\nu) = \varepsilon_{real}(\nu) - i\varepsilon_{img}(\nu)$ (see details in the 2.IV.F). To obtain the water relaxation dynamics in these solutions we fit the $\tilde{\varepsilon}(\nu)$ profiles using a multiple Debye relaxation model:[50, 51]

$$\tilde{\varepsilon}(\nu) = \varepsilon_\infty + \sum_{j=1}^{3} \frac{S_j}{1+i\omega\tau_j} + \frac{\sigma}{i\omega\varepsilon_0} \tag{6.3}$$

where τ_j and S_j $(= \varepsilon_j - \varepsilon_{j+1})$ are the relaxation time and strength of the j-th relaxation mode, respectively. ε_∞ is the extrapolated dielectric constant at very high frequency, ε_0 is the permittivity in free space, ω is the angular frequency and σ is the DC conductivity of the solution. Figure 6.III.5B represents the Debye model fit of all the $\tilde{\varepsilon}(\nu)$ profiles; the fitting parameters are provided in Table 6.III.4. Pure water shows three distinct relaxation time constants: 8.6 ps (τ_1), 200 fs (τ_2) and 80 fs (τ_3).[49, 52, 53] We primarily focus on the alteration of the cooperative hydrogen bond dynamics in presence of solutes, which is defined by the slowest timescale (τ_1) of water relaxation. The other two ultrafast time constants originate from either quick jump of under coordinated water[54] or a small angular rotation foregoing a large angle jump.[55] Since the reorientation dynamics at the solute interface is mostly reflected through the time

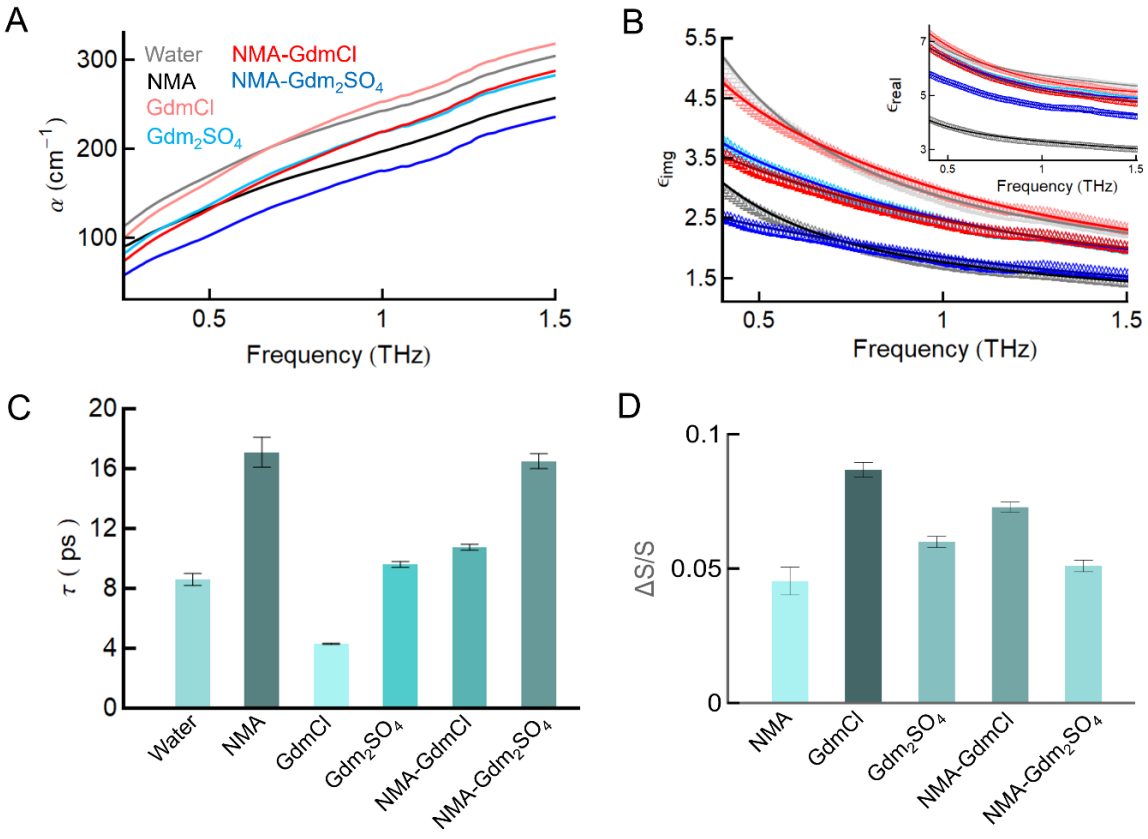


Figure 6.III.5. (A) Frequency dependent absorption co-efficient, α(ν) profile of only water, Gdm^+ salts (at concentration 4M) solutions, and NMA (at concentration 2M) solutions in absence and presence of Gdm^+ salts from TTDS measurements. (B) Representative fittings of ε_{real} (inset) and ε_{img} for the respective systems using multiple Debye relaxation model. (C) Relaxation timescale

($\tau_1 = \tau$) as a function of the respective systems; Explicitly associates with the spontaneous restructuring of the hydrogen bond network. (D) Relative change in the static permittivity ($\Delta S/S$) of the respective samples.

---

constant τ_1 we will make all the further discussions based on the alteration in the time constant τ_1. We observe that 4M GdmCl accelerates the relaxation dynamics (4.5 ps) compared to that of bulk water (figure 6.III.5C).[48] On the contrary, in 4M Gdm_2SO_4 solution the time constant does not change appreciably (9.6 ps). In presence of NMA the dynamics gets noticeably slower (~17 ps) indicating formation of structured water layer around the NMA molecules. As GdmCl is added to NMA, τ_1 reduces to 10.8 ps, whereas addition of Gdm_2SO_4 retains the time constant slower (figure 6.III.5C), which is in comparable agreement with that obtained from simulation results (Table 6.III.3).

We further calculate the relative depolarization of the dielectric strength ($\frac{\Delta S}{S} = \frac{(\varepsilon_S - \varepsilon_\infty) - (\varepsilon_S - \varepsilon_1)}{\varepsilon_S - \varepsilon_\infty} = \frac{\varepsilon_1 - \varepsilon_\infty}{\varepsilon_S - \varepsilon_\infty}$) (figure 6.III.5D). In general depolarization could emanate through a number of mechanisms[56] including dilution effect, kinetic depolarization and static depolarization. *Dilution effect* is caused as solute molecules displace solvent molecules decreasing the effective concentration of the solvent dipoles. *Static depolarization* arises due to strong binding of solvent dipoles with the solute molecules. Consequently, water dipoles fail to follow the external electric field and do not participate in the dielectric relaxation processes. In case solute molecules contain charge, they move towards the external electric field, and consequently the solvation water molecules are forced to reorient through a local electric field in a direction opposite to the external field; this eventually gives rise to *kinetic depolarizatio*n. We observe that the relative depolarization is the smallest in NMA aqueous solution. Chargeless NMA is not expected to offer any kinetic depolarization share towards the total depolarization, and it is only the amide group in NMA that makes H-bond with water dipoles leading to a static depolarization only and consequently we observe smaller depolarization. For the salt solutions on the other hand, depolarization is noticeably higher than NMA as the ions contribute in the form of kinetic depolarization. Interestingly, GdmCl shows a stronger depolarization compared to Gdm_2SO_4. This fact indicates that the contribution of kinetic depolarization and hence the apparent availability of free ions is higher in GdmCl compared

to that in Gdm_2SO_4. This observation supports the hetero-ion paring phenomenon in Gdm_2SO_4 as predicted by simulation results. In the binary mixture of NMA and Gdm salts, the depolarization values are found to be higher than NMA, but lower than the respective salt solutions.

FIR-THz hydration study: Finally, we evaluate the collective hydration structure of all the systems using of THz-FTIR measurements (50-750 cm^{-1}; 1.5-23 THz). Oscillation of O-H bond leads to ultrafast (sub-ps) fluctuations of collective dipole moment of water which makes its imprint in the THz absorption coefficient $\alpha_{THz}(\nu)$ according to the following relation[34]

$$\alpha(\nu) = \frac{1}{4\pi\varepsilon_0}\frac{2\beta\nu^2\pi}{3Vcn(\nu)}\int dt e^{-i\nu t} < M(0)M(t) > \qquad (6.4)$$

in which $M(t)$ $(= \sum_{I=1}^{N}\mu_I(t))$ is the total molecular dipole moments of V volume of sample, $n(\nu)$ is the frequency dependent refractive index at $T = \frac{1}{k_B\beta}$. Detailed experimental analysis of α(ν) are provided in the 2.IV.G section. The $\alpha_{THz}(\nu)$ profile of water (figure 6.III.6) shows characteristic

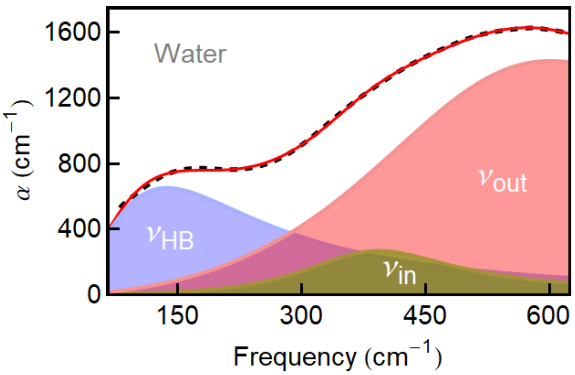


Figure 6.III.6. Frequency dependent absorption co-efficient (α(ν)) profiles of bulk water.

features and is fitted using a damped harmonic oscillator model (DHO) (in the section 2.I.D),

$$\Delta\alpha(\nu) = \sum_{i=1}^{N}\frac{a_i\omega_{0,i}\nu^2}{4\pi^3\left[\frac{\nu^2\omega_{0,i}^2}{\pi^2}+\left(\nu_{d,i}^2+\frac{\omega_{0,i}^2}{4\pi^2}-\nu^2\right)^2\right]} \qquad (6.5)$$

where a_i, $\omega_{0,i,}$ and $\nu_{d,i}$ elucidate the amplitude, the damping width, and the center frequency of the i^{th}damped harmonic oscillator modes, respectively. Subsequent modes can be extracted from

the unperturbed center frequency, $\nu_{c,i}$, obtained by, $\nu_{c,i} = \sqrt{\nu_{d,i}^2 + \frac{\omega_{0,i}^2}{4\pi^2}}$. The model discerns three vibrational modes: ν_{HB} associates intermolecular H-bond stretching mode centred at ~ 136 cm^{-1} and two restricted rotational modes; *in plane libration* (ν_{in} at ~390 cm^{-1}) and *out of plane libration* (ν_{out} at ~600 cm^{-1}). Replacement of water molecules by the less absorbent solutes create "THz defect" and offers distinct hydration features compared to bulk water.[57] For a quantitative insight we calculate the difference absorption co-efficient, $\Delta\alpha(\nu)$ ($=\alpha_{sample}(\nu) - \alpha_{water}(\nu)$) of these solutions (figure 6.III.7A, B, and figure 6.III.8). It is important to note here that being a difference spectrum, $\Delta\alpha(\nu)$ eliminates any bulk contributions and therefore offers explicit information on the solute surface hydration. We use the same DHO model (equation 6.5) to fit all the experimentally obtained $\Delta\alpha(\nu)$ profiles. The fitted parameters are provided in Table 6.III.5. Interestingly, in case of Gdm containing solutions, we observe one additional positive band appearing in the $\Delta\alpha_{Salt}(\nu)$

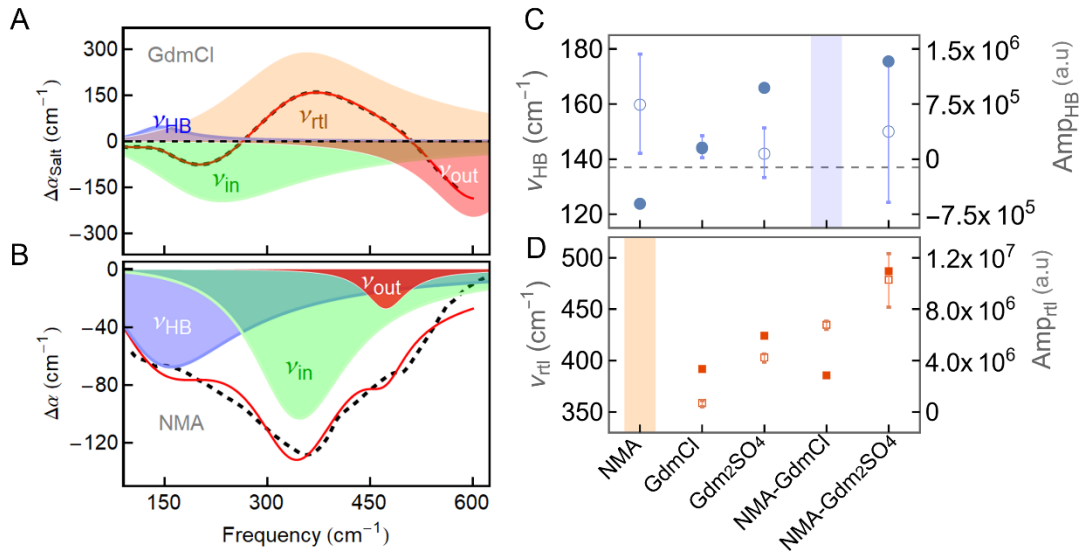


Figure 6.III.7. Dissection of THz (FIR) experimental Δα(ν) spectra of GdmCl aqueous solution (4M) (A), and of NMA aqueous solution (B), fitted using a damped harmonic oscillator model. The black dashed lines represent experimental data and the red solid lines stand for the overall fit. Four damped modes are observed: HB stretching mode (~130-150 cm^{-1}), in-plane librational mode (~230-350 cm^{-1}), rattling motion of ions (~350-450 cm^{-1}), and out of plane librational mode (~500-640 cm^{-1}). The combined representation of peak frequencies (empty symbols) and corresponding amplitudes (filled symbols) of the HB stretching mode and the ion rattling mode for the different

solutions are presented in (C) and (D), respectively. Colored background stands for the absence of that particular mode of the respective system. The error bar signifies the standard deviations.

($=\alpha_{salt+water}(\nu) - \alpha_{water}(\nu)$) profiles in the 350-450 cm^{-1} frequency window (figure 6.III.7A and figure 6.III.8, table 6.III.5). This particular band signifies *ion-rattling mode* (ν_{rtl}) [58-60] which originates from the dipole-dipole auto and cross correlation of ions and water.[36] Thus, this band provides explicit information of salt hydration.[61] We summarise the ν_{HB} and ν_{rtl} mode frequencies and the corresponding amplitudes for all the systems in figure 6.III.7 (C and D). It is important to note that while the mode frequencies offer structural insights, the corresponding amplitude values associates the population of water molecules responsible for that particular mode. For example, if we consider the collective change in H-bond network, it appears that we could not recover the ν_{HB} mode for NMA-GdmCl, while for NMA-Gdm_2SO_4 solution the ν_{HB} mode gets blue shifted compared to the Gdm_2SO_4.aqueous solution. Intriguingly, a significant decrease in the rattling amplitude suggests a loss of collective ion hydration involving Gdm^+ and Cl^- ions as it interacts with NMA.

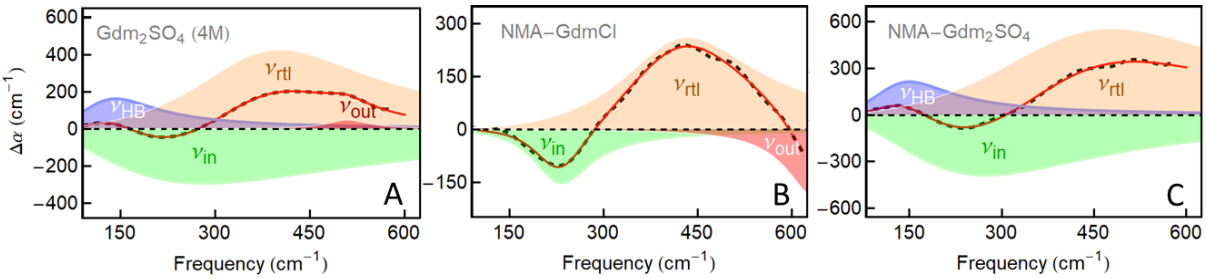


Figure 6.III.8. Dissection of $\Delta\alpha(\nu)$ profiles of Gdm_2SO_4 aqueous solution (4M) and NMA-Salts systems.

Discussion

Our study is aimed to understand the contrasting behavior of GdmCl and Gdm_2SO_4 towards protein stability and we hypothesize that tripartite (salt-protein-water) interactions play the key role in determining the process. To mimic the interaction of these salts with proteins, we use a model peptide NMA and perform various spectroscopic and simulation investigations to determine how these two salts interact with NMA in aqueous medium. Interfacial thickness and interfacial capacitance measurements (figure 6.III.1) of the aqueous salt solutions indicate strong hetero-ion interactions between Gdm^+ and SO_4^{2-}, which, on the other hand, is rather negligible in GdmCl.

Formation of such hetero-ion association is also supported by simulation results (figure 6.III.2). For the trinary systems, simulation predicts the formation of extended cluster-like structures in case of Gdm_2SO_4-NMA (figure 6.III.2D), in which water is entrapped, eventually retarding its dynamics, as also has been evidenced from the MSD calculations (figure 6.III.3). This prediction is supported experimentally by the TTDS measurement; the thus obtained H-bond relaxation time constant confirms a relatively slower dynamic around NMA aqueous solution (figure 6.III.5C, Table 6.III.4). Addition of GdmCl accelerates water relaxation dynamic, whereas addition of Gdm_2SO_4 slows it down. This result unambiguously confirms the formation of hetero-ion paring in Gdm_2SO_4, which makes the Gdm^+ ion less accessible to the NMA molecules and thus leads the NMA dynamics to remain unaltered; in GdmCl, however, the Gdm^+ ions bind to the NMA molecules, making the relaxation dynamics slower compared to that in GdmCl-water solution (Table 6.III.3 and 6.III.4).

$$\Delta\alpha_{Salt} = \alpha_{NMA+Salt+water} - \alpha_{NMA+water}$$

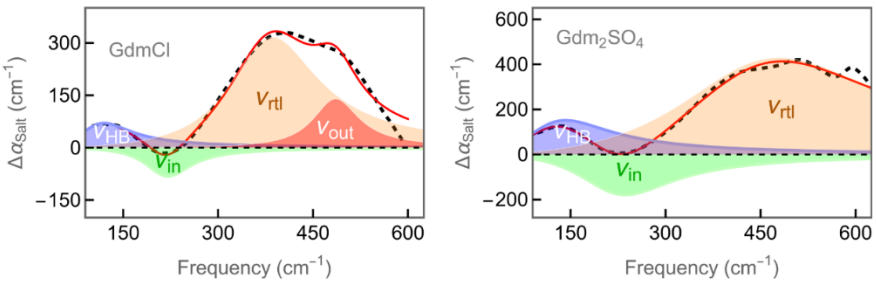


Figure 6.III.9. Dissection of $\Delta\alpha(\nu)$ profiles of guanidinium salts hydration with NMA.

Finally, to delve into the contrasting behaviour of the Gdm salts, we explore the hydration structure of the trinary system using THz-FTIR measurements (figure 6.III.8B,C, Table 6.III.5). The experimentally measured $\alpha_{THz}(\nu)$ is a collective parameter with contributions from different components: $\alpha(\nu) = \sum \phi_i \alpha_i(\nu)$, where ϕ_i denotes the volume fraction and $\alpha_i(\nu)$ is the frequency dependent absorption co-efficient of the *i*-th component. Interestingly, each hydration contribution offers non-overlapping and distinct features in the THz fingerprint regions. This unique feature enables one to analyse these contributions individually to discern the solvation status of each constituent(s). To obtain the explicit ion (both Gdm^+ and the counterions) solvation change in presence of NMA, we calculate: $\Delta\alpha_{salt}(\nu) = \alpha_{salt+NMA+water}(\nu) - \alpha_{NMA+water}(\nu)$ profiles and

fitted them as before. The results are summarized in figure 6.III.9 and Table 6.III.6. We can further write the $\alpha(\nu)$ values in analytical terms:

$$\alpha_{s+N+w}(\nu) = \phi_s\alpha_s(\nu) + \phi_{sh}\alpha_{sh}(\nu) + \phi_N\alpha_N(\nu) + \phi_{Nh}\alpha_{Nh}(\nu) + \phi_b\alpha_{bulk}(\nu) \quad (6.6)$$

$$\alpha_{N+w}(\nu) = \phi'_N\alpha_N(\nu) + \phi'_{Nh}\alpha_{Nh}(\nu) + \phi'_b\alpha_{bulk}(\nu) \quad (6.7)$$

$$\alpha_{s+w}(\nu) = \phi''_s\alpha_s(\nu) + \phi''_{sh}\alpha_{sh}(\nu) + \phi''_b\alpha_{bulk}(\nu) \quad (6.8)$$

where the subscripts '*N*', '*s*' and '*w*' stands for NMA, salt and water, respectively; '*Nh*' and '*sh*' denote NMA and salt hydration, respectively. Noteworthy that all the ions and NMA individually offer very low absorbance in this frequency window and one can neglect the contributions of the α_s and α_N terms. Thus, the expression for the explicit salt hydration, $\Delta\alpha_s(\nu)$ takes the form:

$$\Delta\alpha_s(\nu) = \alpha_{s+N+w}(\nu) - \alpha_{N+w}(\nu)$$

$$= \phi_{sh}\alpha_{sh}(\nu) + (\phi_{Nh} - \phi'_{Nh})\alpha_{Nh}(\nu) + (\phi_b - \phi'_b)\,\alpha_{bulk}(\nu) \quad (6.9)$$

The last two terms in equation (6.9) represent NMA induced change in ion solvation. From the DHO fitting parameters, we estimate the amplitude (A_i) of each damped oscillatory mode measured at the central frequency ($\nu_{d,i}$) of the respective mode (Table 6.III.6). The recovered amplitude for the intermolecular H-bond stretching mode (A_{HB}) is directly correlated to the NMA induced changes in the ion hydration shell. A_{rtl} defines the NMA induced change in the residual rattling population around the ion solvation sheath. Notably, the magnitude of α_{Nh} and α_{bulk} terms are very small compared to that of α_{sh} in the rattling frequency window (350-450 cm^{-1}). Therefore, any modification in A_{rtl} explicitly defines alteration in ion hydration. It is noteworthy that the water dipole correlation of HB and rattling modes do not intersect, and therefore allows one to extract independent hydration information of each mode. We schematically show the HB and rattling mode of vibration for the cation and anion hydration in figure 6.III.10A and B. As we are interested to understand the NMA mediated changes in ion hydration, we define a parameter depicting the change in amplitude: $\Delta A_i = A_{salt\ in\ NMA-water} - A_{salt\ in\ bulk\ water}$, and plot it for the two modes (blue: HB and red: rattling) for the two salts (figure 6.III.10C). We find that ΔA_{HB}^{GdmCl} is negative indicating a lowering of H-bonded water population around GdmCl in presence of NMA. It implies that the presence of NMA perturbs the outer layer of the ion-hydration

sheath and thereby making Gdm^+-NMA interaction feasible. This, in turn, also explains the protein denaturing propensity of GdmCl.

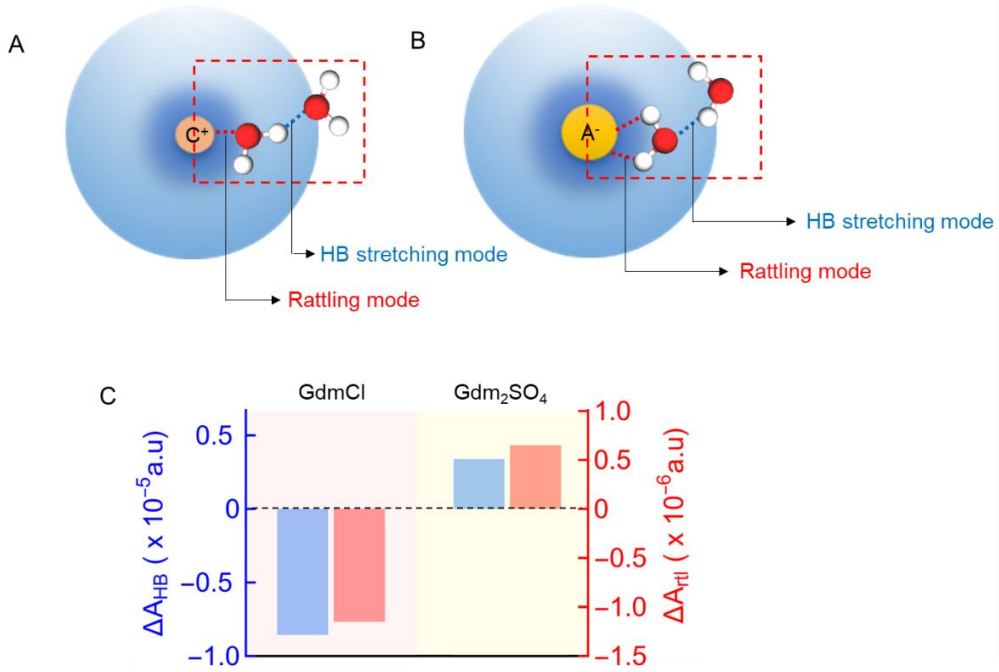


Figure 6.III.10. Schematic images of cation (A) and anion (B) hydration with defined vibrational modes. Changes in the amplitude of (C) intermolecular H-bond stretching mode (ΔA_{HB}), and ion-water rattling (ΔA_{rtl}) mode for the salt(s) solution(s) from the absence to the presence of NMA.

Contrastingly, $\Delta A_{HB}^{Gdm_2SO_4}$ is found to be positive, which unambiguously confirms that Gdm_2SO_4 does not lose its hydration structure in presence of NMA indicating that it does not directly interact with NMA. ΔA_{rtl}, on the other hand, describes interfacial ion-water structure for the Gdm salts in presence of NMA. We found that ΔA_{rtl}^{GdmCl} is negative while $\Delta A_{rtl}^{Gdm_2SO_4}$ is positive. This apparently contrasting trends could be understood in the light of the linewidth (ω_o) of the damping of the vibrational modes within the ion-water complexes. Broadening of solvated ion modes, specifically *out of plane* libration, is governed by the librational motions of the surrounding water molecules.[56] For GdmCl, we observed that the ω_0^{out} decreases from ~ 517 cm^{-1} to ~ 277 cm^{-1} in presence of NMA (table 6.III.6). This reduction in ω_0^{out} corresponds to the Gdm ion association with NMA, which also gets reflected in the reduction of ω_0^{rtl} value from 899 cm^{-1} to ~ 538 cm^{-1} (table 6.III.6). Interestingly, the damping width for Gdm_2SO_4 rattling mode (ω_0^{rtl}) increases in presence of NMA suggesting more restricted water environment around the ion surface. This corroborates excellently

with the co-operative ion pairing picture as depicted from the MD simulation results revealing the existence of hetero-ion pairing in the Gdm_2SO_4 solution.

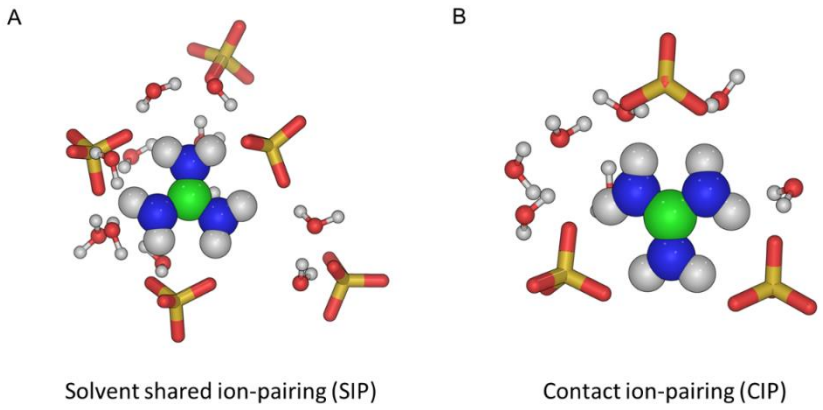


Figure 6.III.11. Snapshots of the two different types of ion pairs; (A) solvent shared ion pairs (SIPs) and (B) contact ion pairs (CIPs) observed in the simulated Gdm_2SO_4 systems at 2M concentration.

Simulation snapshots (figure 6.III.11) for the possible hetero-ion pairing interactions confirm both solvent shared ion pairs (SIPs) and contact ion pairs (CIPs) formation in the Gdm_2SO_4 solutions. The observed positive values of both $\Delta A_{HB}^{Gdm_2SO_4}$ and $\Delta A_{rtl}^{Gdm_2SO_4}$ aligns with the fact that the population of shared water molecules increases, while the population of unshared water molecules decreases correspondingly.[57] The shared water molecules mostly donate H-bonds to SO_4^{2-}.[58] Additionally, the increased ω_0^{rtl} also infers a reduced likelihood of forming CIPs in a crowded NMA environment in Gdm_2SO_4.

6.IV. Conclusions.

Our combined spectroscopic and MD-simulation investigation using a model amide NMA unambiguously concludes that it is the contrasting ion-pairing propensity of the Gdm^+ ion with the counterion (Cl^- or SO_4^-) that decides its protein denaturation ability. While ion-paring in Gdm_2SO_4 does not lead free Gdm^+ ions to interact with NMA, it can abundantly do so in GdmCl due to the lack of any such ion-pairing.

Table 6.III.1. Molecular composition of various simulated systems

Species	Gdm_2SO_4	GdmCl	NMA	NMA-Gdm_2SO_4	NMA-GdmCl
Gdm^+	76	38	0	76	38
X^-	38	38	0	38	38
NMA	0	0	152	152	152

Table 6.III.2. Self-diffusion coefficient of water in different systems

System	Diffusion coefficient ($D \times 10^{-5}$ cm^2s^{-1})
Bulk Water	5.78 ± 0.09
NMA	4.49 ± 0.11
GdmCl	3.78 ± 0.10
NMA-GdmCl	2.51 ± 0.03
Gdm_2SO_4	1.09 ± 0.09
NMA-Gdm_2SO_4	0.48 ± 0.01

Table 6.III.3. Double exponential fitting parameters of hydrogen bond autocorrelation function.

System	α_1	α_2	c	τ_1 (ps)	τ_2 (ps)	$<\tau>$ (ps)	R^2 score
Bulk-water	0.54	0.42	0.02	0.69	4.22	2.23	0.999
GdmCl	0.51	0.45	0.03	0.65	4.88	2.63	0.998
Gdm_2SO_4	0.45	0.43	0.11	0.6	7.49	3.96	0.999
NMA	0.47	0.48	0.03	0.68	4.96	2.84	0.998
NMA-Gdm_2SO_4	0.36	0.42	0.20	0.52	10.31	5.79	0.999
NMA-GdmCl	0.44	0.50	0.05	0.64	6.01	3.49	0.999

Table 6.III.4. Debye relaxation fitting parameters of each NMA-salt system, water and only aqueous solution of NMA and salts.

Sample	ε_s	ε_1	ε_2	ε_∞	τ_1 (ps)	τ_2 (fs)	σ (S)
water	78.3	9.4(0.5)	5.7(0.02)	4.2(0.02)	8.6(0.4)	545(50)	5E-6
NMA	81.6	5.8(0.4)	3.4(0.01)	2.2(0.01)	17.1(1)	604(100)	5E-6
GdmCl	41.8	8.1(0.06)	4.4(0.05)	4.9(0.04)	4.3(0.05)	259(10)	5E-3
Gdm_2SO_4	58	7.8(0.07)	4.4(0.05)	4.6(0.04)	9.6(0.2)	278(20)	4.5E-3
NMA-GdmCl	55	7.9(0.06)	4.4(0.04)	4.2(0.04)	10.75(0.2)	290(10)	3.8E-3
NMA- Gdm_2SO_4	55	6.6(0.06)	3.8(0.06)	4.0(0.05)	16.5(0.5)	263(10)	4.1E-3

Table 6.III.5. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c): Respective values for hydrogen bond stretching mode, for water in plane librational motion, for rattling motions of ions, and for water out of plane librational motion of each NMA-salt system, water and only aqueous solution of NMA and salts.

Sample	a_0 (a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
HB stretch				
water	6627376	809.4(11)	45.5(4)	136.6(4)
NMA	-593282.7	700(94)	115(10)	160(18)
GdmCl	162614.8	275.2(26)	137.8(1)	144.5(4)
Gdm_2SO_4	984745.4	483.2(51)	119.7(4)	142.3(9)
NMA-GdmCl	-	-	-	-
NMA- Gdm_2SO_4	1344510	499.6(159)	127.5(7)	150.3(26)
In-plane libration				
water	2398116	714(49)	375.6(1)	392.4(8)
NMA	-677944	524(42)	336.8(2)	347(7)
GdmCl	-2089400	858(33)	190.2(5)	234.1(7)
Gdm_2SO_4	-6340508	1720.4(61)	68.3(54)	282.2(54)
NMA-GdmCl	-672875.7	353.6(8)	226.2(0.6)	233.1(1)
NMA- Gdm_2SO_4	-6606865	1354.2(534)	168.7(58)	273.7(72)
Ion rattling				
water	-	-	-	-
NMA	-	-	-	-
GdmCl	3291027	899(14)	328.3(2)	358.1(3)
Gdm_2SO_4	5877086	1100.8(25)	362.3(2)	402.5(4)
NMA-GdmCl	2808708	845(22)	413(1)	434.3(4)
NMA- Gdm_2SO_4	10902733	1557.8(118)	408.7(18)	478(26)
Out of plane libration				
water	32996001	1840.3(43)	525(7)	601(10)

NMA	-80051.2	233(49)	471.5(4)	473(9)
GdmCl	-1604304	517.4(1)	595.5(1)	601.2(2)
Gdm_2SO_4	150606	253.9(8)	507.2(0.4)	508.8(1)
NMA-GdmCl	-950996.7	380.6(47)	639.6(14)	642.5(16)
NMA- Gdm_2SO_4	-	-	-	-

Table 6.III.6. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c): Respective values for hydrogen bond stretching mode, for water both librational motions, and for rattling motions of ions are enlisted of each salt system.

Sample	a_0 (a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
HB stretch				
water	6627376	809.4(11)	45.5(4)	136.6(4)
Only GdmCl	162614.8	275.2(26)	137.8(1)	144.5(4)
Only Gdm_2SO_4	984745.4	483.2(51)	119.7(4)	142.3(9)
Explicit GdmCl	77028.3	324.1(88)	108.2(7)	120(16)
Explicit Gdm_2SO_4	1018618.6	534.8(207))	113.7(8)	142(34)
In-plane libration				
water	2398116	714(49)	375.6(1)	392.4(8)
GdmCl	-2089400	858(33)	190.2(5)	234.1(7)
Gdm_2SO_4	-6340508	1720.4(61)	68.3(54)	282.2(54)
Explicit GdmCl	-283371.2	274.2(43)	216.1(3)	220.4(8)
Explicit Gdm_2SO_4	-1297160	574.3(120)	217.8(17)	236.2(26)
Ion rattling				
water	-	-	-	-
GdmCl	3291027	899(14)	328.3(2)	358.1(3)
Gdm_2SO_4	5877086	1100.8(25)	362.3(2)	402.5(4)
Explicit GdmCl	2140369	538.5(25)	376.5(2)	386.1(4)
Explicit Gdm_2SO_4	6744771	1250.4(37)	435.5(6)	478.8(8)
Out of plane libration				
water	32996001	1840.3(43)	525(7)	601(10)
GdmCl	-1604304	517.4(1)	595.5(1)	601.2(2)
Gdm_2SO_4	150606	253.9(8)	507.2(0.4)	508.8(1)
Explicit GdmCl	486167	277.5(26)	484(2)	485.7(5)
Explicit Gdm_2SO_4	-	-	-	-

6.V. References.

1. P. Lo Nostro and B. W. Ninham, *Chem. Rev.*, 2012, **112**, 2286-2322.
2. K. P. Gregory, G. R. Elliott, H. Robertson, A. Kumar, E. J. Wanless, G. B. Webber, V. S. J. Craig, G. G. Andersson and A. J. Page, *Phys. Chem. Chem. Phys.*, 2022, **24**, 12682-12718.
3. F. Hofmeister, *Arch. Exp. Pathol. Pharmakol.*, 1888, **24**, 247-260.
4. W. Kunz, J. Henle and B. W. Ninham, *Curr. Opin. Colloid Interface Sci.*, 2004, **9**, 19-37.

5. K. D. Collins, *Biophys. J.*, 1997, **72**, 65-76.
6. K. D. Collins, *Methods*, 2004, **34**, 300-311.
7. H. Ohtaki and T. Radnai, *Chem. Rev.*, 1993, **93**, 1157-1204.
8. R. Breslow and T. Guo, *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 167-169.
9. B. A. Rogers, H. I. Okur, C. Yan, T. Yang, J. Heyda and P. S. Cremer, *Nat. Chem.*, 2022, **14**, 40-45.
10. P. Gund, *J. Chem. Educ.*, 1972, **49**, 100.
11. S. K. Jha and S. Marqusee, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 4856-4861.
12. P. Mason, G. Neilson, A. Barnes, J. Enderby, J. Brady and M.-L. Saboungi, *J. Chem. Phys.*, 2003, **119**, 3347-3353.
13. M. Vazdar, J. Heyda, P. E. Mason, G. Tesei, C. Allolio, M. Lund and P. Jungwirth, *Acc. Chem. Res.*, 2018, **51**, 1455-1464.
14. G. Graziano, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12008-12014.
15. K. P. Gregory, E. J. Wanless, G. B. Webber, V. S. J. Craig and A. J. Page, *Chem. Sci.*, 2021, **12**, 15007-15015.
16. J.-C. Soetens, C. Millot, C. Chipot, G. Jansen, J. G. Ángyán and B. Maigret, *J. Phys. Chem. B*, 1997, **101**, 10910-10917.
17. M. Vazdar, F. Uhlig and P. Jungwirth, *J. Phys. Chem. Lett.*, 2012, **3**, 2021-2024.
18. A. Kubíčková, T. Křížek, P. Coufal, E. Wernersson, J. Heyda and P. Jungwirth, *J. Phys. Chem. Lett.*, 2011, **2**, 1387-1389.
19. C. E. Dempsey, T. J. Piggot and P. E. Mason, *Biochemistry*, 2005, **44**, 775-781.
20. P. E. Mason, C. E. Dempsey, G. W. Neilson, S. R. Kline and J. W. Brady, *J. Am. Chem. Soc.*, 2009, **131**, 16689-16696.
21. B. Ding, D. Mukherjee, J. Chen and F. Gai, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 1003-1008.
22. R. Perez-Jimenez, R. Godoy-Ruiz, B. Ibarra-Molero and J. Sanchez-Ruiz, *Biopys. J.*, 2004, **86**, 2414-2429.
23. I. Negi, R. Jangra, A. Gharu, J. F. Trant and P. Sharma, *Phys. Chem. Chem. Phys.*, 2023, **25**, 857-869.
24. E. P. O'Brien, R. I. Dima, B. Brooks and D. Thirumalai, *J. Am. Chem. Soc.*, 2007, **129**, 7346-7353.
25. P. Cohen, K. A. Dill and S. S. Jaswal, *J. Phys. Chem. B*, 2014, **118**, 10618-10623.
26. R. Godawat, S. N. Jamadagni and S. Garde, *J. Phys. Chem. B*, 2010, **114**, 2246-2254.
27. V. Balos, M. Bonn and J. Hunger, *Phys. Chem. Chem. Phys.*, 2017, **19**, 9724-9728.
28. Q. Shao, Y. Fan, L. Yang and Y. Q. Gao, *J. Chem. Theory Comput.*, 2012, **8**, 4364-4373.
29. P. Mason, C. Dempsey, G. Neilson and J. Brady, *J. Phys. Chem. B*, 2005, **109**, 24185-24196.
30. S. Heiles, R. J. Cooper, M. J. DiTucci and E. R. Williams, *Chem. Sci.*, 2015, **6**, 3420-3429.
31. R. J. Cooper, S. Heiles, M. J. DiTucci and E. R. Williams, *J. Phys. Chem. A*, 2014, **118**, 5657-5666.
32. P. Mason, G. Neilson, C. Dempsey, A. Barnes and J. Cruickshank, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 4557-4561.
33. P. E. Mason, G. W. Neilson, J. E. Enderby, M.-L. Saboungi, C. E. Dempsey, A. D. MacKerell, Jr. and J. W. Brady, *J. Am. Chem. Soc.*, 2004, **126**, 11462-11470.
34. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 12068-12073.
35. S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 20749-20752.
36. D. Sharma and A. Chandra, *J. Phys. Chem. B*, 2023, **127**, 9323-9335.
37. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19-25.
38. K. Vanommeslaeghe and A. D. MacKerell, Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3144-3154.
39. G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**.
40. M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182-7190.

41. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
42. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463-1472.
43. L. F. Lima, A. L. Vieira, H. Mukai, C. M. Andrade and P. R. Fernandes, *J. Mol. Liq.*, 2017, **241**, 530-539.
44. J. Huang, B. G. Sumpter and V. Meunier, *Angew. Chem. Int. Ed.* , 2008, **47**, 520-524.
45. M. J. Hebert and D. H. Russell, *J. Phys. Chem. Lett.*, 2019, **10**, 1349-1354.
46. A. Striolo, *Nano Lett.*, 2006, **6**, 633-639.
47. A. Luzar, *J. Chem. Phys.*, 2000, **113**, 10663-10675.
48. N. Samanta, D. D. Mahanta and R. K. Mitra, *Phys. Chem. Chem. Phys.*, 2014, **16**, 23308-23315.
49. S. Chakraborty, P. Pyne, R. K. Mitra and D. Das Mahanta, *J. Mol. Liq.*, 2023, **382**, 121998.
50. N. Nandi and B. Bagchi, *J. Phys. Chem. B*, 1997, **101**, 10954-10961.
51. J. T. Kindt and C. A. Schmuttenmaer, *J. Phys. Chem.*, 1996, **100**, 10373-10379.
52. N. Samanta, D. D. Mahanta and R. K. Mitra, *Chem. Asian J.* , 2014, **9**, 3457-3463.
53. D. Das Mahanta, A. Patra, N. Samanta, T. Q. Luong, B. Mukherjee and R. K. Mitra, *J. Chem. Phys.*, 2016, **145**, 164501.
54. D. Laage and J. T. Hynes, *Science*, 2006, **311**, 832-835.
55. A. Y. Zasetsky, *Phys. Rev. Lett.*, 2011, **107**, 117601.
56. K. Tielrooij, S. Van Der Post, J. Hunger, M. Bonn and H. Bakker, *J. Phys. Chem. B*, 2011, **115**, 12638-12647.
57. B. Born, S. J. Kim, S. Ebbinghaus, M. Gruebele and M. Havenith, *Faraday Discuss.*, 2009, **141**, 161-173.
58. C. Klinkhammer, F. Böhm, V. Sharma, G. Schwaab, M. Seitz and M. Havenith, *J. Chem. Phys.*, 2018, **148**, 222802.
59. R. Saha and R. K. Mitra, *Langmuir*, 2023, **39**, 4601-4610.
60. R. Saha and R. K. Mitra, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23661-23668.
61. D. A. Schmidt, Ö. Birer, S. Funkner, B. P. Born, R. Gnanasekaran, G. W. Schwaab, D. M. Leitner and M. Havenith, *J. Am. Chem. Soc.*, 2009, **131**, 18512-18517.

Chapter 7

7. Thiocyanate ion (SCN^-) offers a major impact in rapid protein amyloidosis: a salient role played by the protein solvation

Thiocyanate (SCN^-) is known to be a naive ion abundant in biological fluids, blood, and urine. It is also used as a biomarker as it can penetrate to the brain by crossing the blood brain barrier (BBB) and also gets into the Cerebrospinal Fluid (CSF) through the blood-CSF barrier. Considering its importance in human physiology, we examine the effect of SCN^- ions on three model proteins: ovalbumin (Ova), bovine serum albumin (BSA) and lysozyme (Lys). We observe that an elevated level of SCN^- (~0.5 M) leads to an otherwise unusual instant fibrilization of all these proteins at pH=2 at ambient temperature. Field emission scanning electron microscopy (FESEM) and atomic force microscopy (AFM) reveal two distinct initial amyloid-aggregated states: nucleus, protofibril and two mature fibril states (upon 24 hours of incubation): crosslinked network or matrix and bundle-like structures. Despite the structural variation of the three proteins, formation of these morphologies depends on the counterion: Na^+ and guanidinium (Gdm^+). Since these processes are assisted by the associated alteration in the protein hydration, we determine the protein and the salt hydration at different phases using THz-FTIR spectroscopy in the 1.5-22.5 THz (50-750 cm^{-1}) frequency window. We found that depending on the counterion, interfacial hydration could act either as a *"lubricant"* or as a *"de-wetting"* agent and the findings can be a potential foundation for future handling of amyloidosis.

7.I. Introduction.

Thiocyanate (SCN^-) is an otherwise innocent ion abundantly found in human body[1] in varying concentrations (0.01-2 mM) in biological fluids, blood, and urine.[2, 3] SCN^- in human body is mostly acquired from the dietary sources;[4, 5] tobacco consumption including inhalation of smoke significantly contributes to elevated SCN^- levels.[6] A portion of SCN^- is also generated through a detoxification process involving the reaction between cyanide (CN^-) and thiosulphate ($S_2O_3^{2-}$) in the liver.[7-9] SCN^- is also a widely used biomarker in cerebrospinal fluid (CSF) as it readily penetrates into extracellular spaces coupled with its slow excretion.[10] An elevated level of SCN^- ion acts as marker of a smoker against a non-smoker.[11] The first documentation regarding biological role of SCN^- ions dates back to 1814 when SCN^- was identified as a

substance in human saliva that exhibits a blood-red coloration upon the addition of ferric ions.[12] In the early to the mid- 20th century, research on SCN^- garnered notable interest in pharmacology as an oral antihypertensive agent.[13] A significant discovery raised concerns about the potential role of SCN^- in the gastric cancer development.[14] Moreover, it was found that high serum concentration of SCN^- could result into lung cancer.[15, 16] SCN^- from blood plasma can penetrate to the brain by crossing the blood brain barrier (BBB)[17] and also gets into the CSF through the blood-CSF barrier.[18] The brain can have 5 times higher accumulation of SCN^- ions than in the CSF,[17] arising concerns about its potential effects on brain tissues and/or biomolecules. These findings conclude that SCN^- is a very important component of human physiology.

Ions play a pivotal role in several biophysical processes. It has been found that the presence of several cations at concentrations beyond a threshold limit could bring about substantial changes in biomolecular conformations leading to protein aggregation and consequently phase transitions. While much attention has been paid on the cation selectivity of such biomolecular aggregation, less has been explored about the anionic counterpart. Ion induced biomolecular aggregation has mostly been governed by the well-known Hofmeister series, the ions present at the edge of the series being more prone to bring about protein precipitation.[19] In this respect SCN^- holds a special mention as it occupies a terminal position among the anions. Interestingly, compared to the other anions, SCN^- is weakly solvated and could bring about mild to considerable perturbation in biomolecular conformation depending on the cation it is associated with; grouping with potent protein denaturant guanidinium (Gdm^+) cation its denaturing capacity is enhanced.[19] This unique nature coupled with substantial bioavailability makes the otherwise naive SCN^- a rather interesting species to investigate, specifically its effect on proteins at varied concentration and pH range. In this study, we investigate the impact of SCN^- ions on three naturally occurring proteins: ovalbumin (Ova), the most abundant protein source in chicken; bovine serum albumin (BSA), a dairy source protein, also a serum protein; and lysozyme (Lys), an enzyme presents in human organs. The choice of the proteins lies on the fact that these proteins are essential and human consumes these proteins through dietary sources. Both of Ova and BSA are globular proteins; Ova has the α-helical and β-sheets contents in the 1:1 ratio.[20] BSA is a model serum albumin; analogues to human serum albumin (HSA) with approximate 76% structural similarity.[21] Lys is a small protein with an important role in the innate immunity.[22] We use two counter-cations: Na^+ (the most common cation in biofluids, which is one of the major components of blood plasma and highly permeable to cross

BBB) and guanidinium, Gdm^+ (a weakly hydrating cation that can induce protein denaturation and has structural similarity to one of the amino acid residues, arginine). We observe that NaSCN does not produce significant effect on protein structures at neutral pH (7.4), however, GdmSCN induces precipitation of Ova at this pH condition, the effect on the other two proteins seems mild. As SCN^- is mostly consumed dietarily, we check the effect of this anion on these three proteins at pH =2 to reproduce the digestion track environment. We observe that all the proteins precipitate in presence of both these salts. Field emission scanning electron microscopy (FESEM) and atomic force microscopy (AFM) images reveal that the types of precipitation differ for the two salts. We identify four distinct states of fibrillar aggregates: a nucleus state for NaSCN, a protofibril state for GdmSCN, and mature fibril states characterized by a crosslinked network or matrix for NaSCN and a bundle-like structure for GdmSCN.

Water, the base solvent and the major component of body fluids, exhibits crucial impact in shaping the structure and the stability of protein aggregates. Proteins in aqueous medium possess a distinct hydration environment that differs markedly from the bulk water.[23] Protein surface hydration plays a pivotal role in governing the key processes in protein aggregation: starting from nucleation to the growth and maturation of the nanostructures. Thus, understanding the nature of protein solvation is a key to underline the molecular mechanisms of salt induced protein aggregate formation. Here we study the change in protein hydration during such SCN^- ion induced amyloidosis using ATR THz-FTIR spectroscopy in the 1.5-22.5 THz (50-750 cm^{-1}) frequency window. THz absorption spectroscopy is an efficient label-free technique for analytically probing fluctuations in collective water dipole moments. This method offers explicit insights into changes within the water network by detecting intermolecular vibrational modes of water molecules.[24-29] Our results strongly answered that protein hydration during SCN^- ion induced amyloidosis follows distinct features when it grouped with different cations and plays a pivotal role in understanding the ultimate fibrillar arrangements.

7.II. Materials and Methods.

Salts, guanidinium thiocyanate (G9277; ~99%) and sodium thiocyanate (251410; ~98%) were purchased from Sigma Aldrich. All proteins bovine serum albumin (A4503), Lysozyme (L6876) and Ovalbumin (Albumin from chicken egg white; A5503) were also purchased from Sigma Aldrich without further purification. Thioflavin T (ThT) dye also we brought from

Sigma Aldrich. The details of all the used instruments are described in the section 2.III and the experimental analysis details are also mentioned prior in the section 2.IV.

Buffer preparation: All pH buffers, used in this study were freshly prepared in laboratory.

- **pH 2** We prepare 50mM pH 2 buffer from KCl salt and HCl acid using the Henderson-Hasselbalch equation.
- **pH 5** The buffer (pH 5) was prepared with citric acid and tri-sodium citrate (.$2H_2O$) of 50 mM concentration also using the Henderson-Hasselbalch equation.
- **pH 7.4** Sodium dihydrogen phosphate (NaH_2PO_4) acid and disodium hydrogen phosphate ($Na_2HPO_4.2H_2O$) salt were used to prepare the phosphate buffer of pH 7.4 (physiological pH). The concentration of the buffer kept 50mM using the mentioned equation.
- **pH 12.6** We also prepare 50mM pH 12.6 buffer using KCl and NaOH beads accordingly.

Sample preparation and aggregation assays: Each protein and salt samples were prepared separately by dissolving directly into a specific pH buffer. The buffer solutions were filtered with a 0.2 um filter prior to prepare each sample. Details of preparation of all the used pH buffers have mentioned in supporting information section. Initially, to monitor the aggregation assay different sets of protein samples (at fixed concentration) were prepared by adding different salt concentration. Later for all the measurements, we have prepared the aggregations just adding the pH 2 solution of the thiocyanate salt into the protein pH 2 solution where for each case salt and protein concentration maintained to 0.5M and 200 µM, respectively.

ThT assay: We first prepare a stock ThT solution in fresh mili-Q water and measure its concentration using UV-Visible spectrophotometer (*Shimadzu UV 2600*). That stock solution was added which maintain overall ThT concentration 10 µM in the sample. Here we run all the steady state measurements using a fluorimeter *Fluorolog (HORIBA, JOBIN YVON)*. The ThT solutions were excited at wavelength $\lambda_{ex} \sim$ 440 nm the slits (both excitation and emission) were kept 5 for all measurements. In every batch of protein salt samples, we consistently maintain a ThT concentration of 10 µM. Specifically, for the protein salt samples at pH 2, we assess the ThT kinetics by adding the mother aggregates into the pH 7.4 ThT solution.

7.III. Results and Discussions.

Visual analysis: Figure 7.III.1A depicts the phase behaviour of BSA, Ova and Lys (200 μM) in presence of NaSCN and GdmSCN (0.5 M) at different pH's. The blue ticks represent the appearance of precipitation and/or turbidity while the cross symbols represent clear solutions.

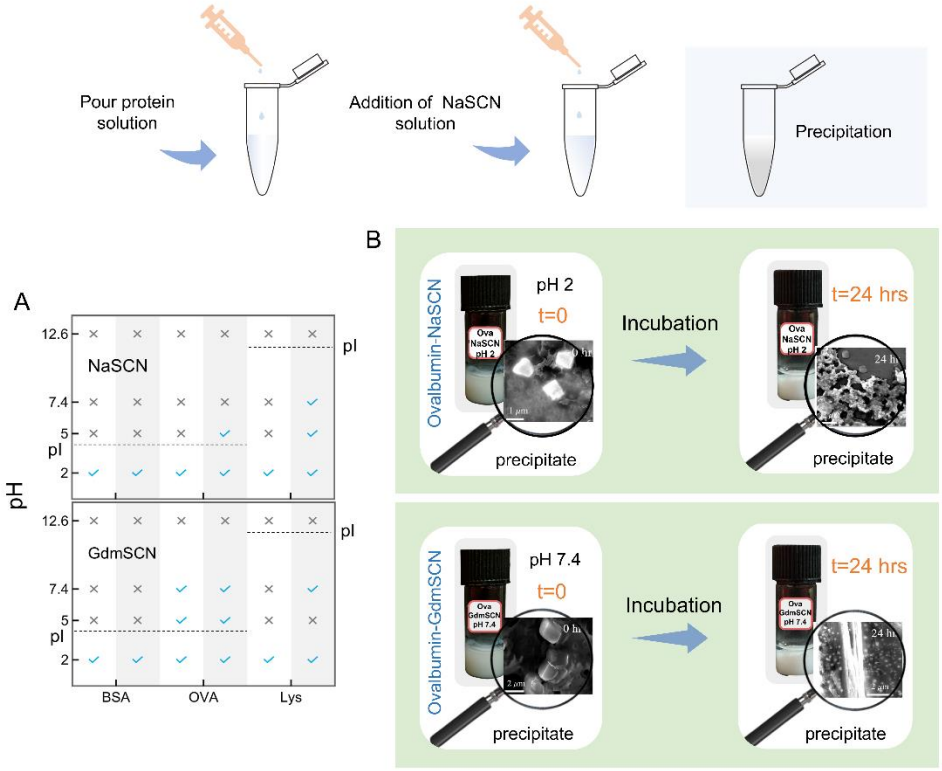


Figure 7.III.1. (A) pH dependent solubilization of BSA, Ova and Lys solutions (200 μM) in presence of GdmSCN and NaSCN (0.5 M). The white and grey panels signify t=0 and t=24 hours, respectively. The tick marks represent the appearance of protein aggregates, the cross marks represent a clear solution. Isoelectric point (pI) of each protein has been marked by dotted lines. (B) Optical images of Ova in presence of NaSCN at pH=2 (upper panel) and GdmSCN at pH=7.4 (lower panel).

It is to be noted that in the chosen experimental condition ([protein]~200 μM and [salt]~ 0.5 M] both the protein and the salt individually remain as monomer(s) (not exemplified). Figure 7.III.1B shows the FESEM images of the precipitate formed upon the addition of the salts in Ova in two different experimental conditions (pH) and on incubation. It is interesting to observe that the microscopic morphology of the aggregates differs with the type on the cation as well as incubation time. For example, Ova-NaSCN (pH=2) forms small cubic crystals which on incubation transforms into mature fibrils (upper panel, for exact phase characterization, see the FESEM studies Figure 7.III.3). Likewise, upon the addition of GdmSCN in Ova (pH=7.4),

micron-sized cubic crystal-like aggregates are formed at t=0, which upon overnight standing produces rod-like structures. Figure 7.III.1A clearly demonstrates that the precipitation phenomenon is protein specific, ion specific and pH specific. Interestingly, all the protein-salt duo produce precipitate at pH=2. We therefore perform most of the further investigations at this pH only.

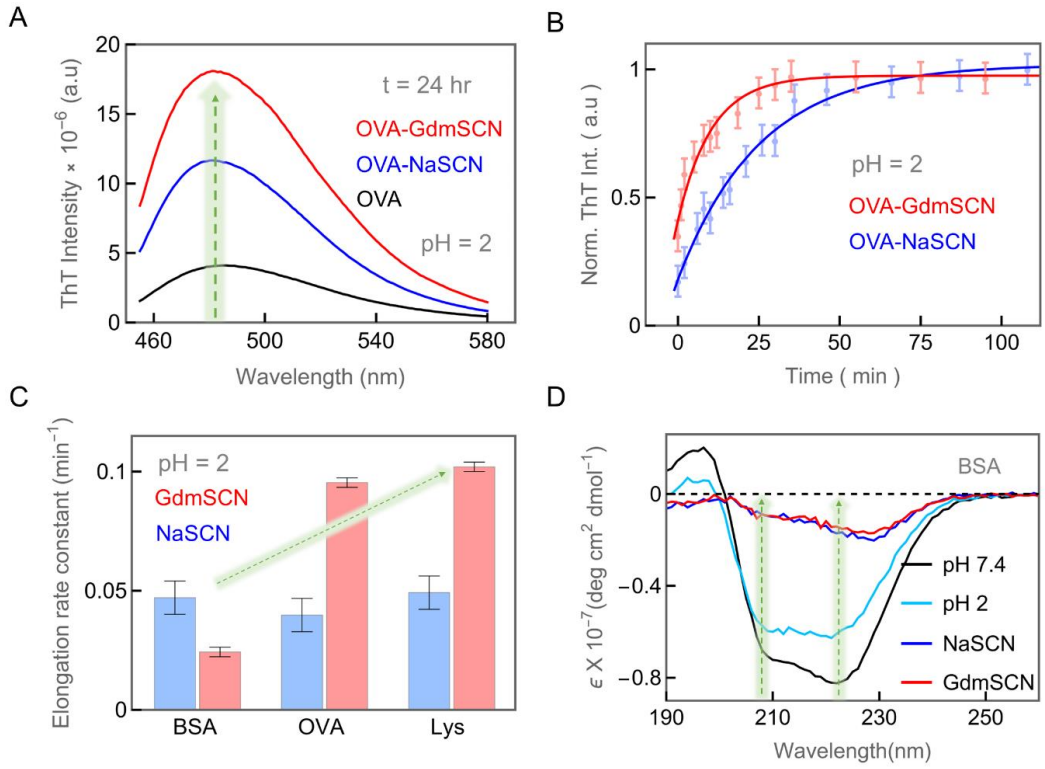


Figure 7.III.2. (A) Fluorescence emission spectra of ThT bound to ovalbumin (at pH 2) with and without GdmSCN and NaSCN salts. (B) Normalised ThT emission kinetic profile in ovalbumin (at pH 2) in presence of GdmSCN or NaSCN. Error bars signify the standard deviations. (C) Elongation rate constants of the three proteins (at pH 2) in presence of GdmSCN or NaSCN. (D) Representative CD spectra of BSA protein at pH 7.4 and at pH 2 in absence and in presence of GdmSCN or NaSCN. The arrows are used to follow the observed trend.

ThT kinetics and secondary structural measurements: To confirm the formation of the aggregates we perform ThT binding and its emission characteristics.[30] Figure 7.III.2A shows the emission profile of ThT in bare Ova solutions at pH 2 in absence and in presence of the salts (at t=24 hrs). We observe noticeable increase in the ThT fluorescence intensity upon the addition of both the salts, strongly suggesting the proteins to form aggregates. To further explore the time dependency of aggregate formation we monitor ThT intensity in the aggregates (at the peak frequency) as a function of time; a representative profile for Ova is

shown in figure 7.III.2B. We observe a near-sigmoidal intensity pattern for all the three protein-salt solutions at pH 2 (representative profile figure 7.III.2B). Such a sigmoidal growth kinetics confirms fibrillar aggregation.[31] We detect two distinct regions: *elongation or growth phase* (steep increase in intensity) and *saturation phase or plateau phase* (in which emission does not change appreciably). It could be noted here that we do not recover any lag phase; perhaps the rapid kinetics of protein aggregation is beyond our instrumental time resolution. It is important to note here that fibrillization mostly initiate through a lag phase, and the growth process could be tuned in this phase only. The SCN^- induced aggregation is rather instantaneous and therefore one cannot control the lag phase and consequently cannot tune the fibril formation. This rather unusual instantaneous nature of aggregation makes the SCN^- ions non-trivial. We fit each kinetic profile using an empirical asymptotic exponential function $y=a-b.c^x$; a is asymptote, b is response range and c is the growth rate. We further calculate the *elongation rate constant* (k) from the growth rate c as $k=-\ln c$. The measured rate constants for the different systems are shown in figure 7.III.2C. We observe that while the elongation rate is protein specific in GdmSCN, it is almost constant in case of NaSCN. We also perform similar kinetic measurements at pH 7.4 solutions in presence of GdmSCN. While in Ova we found an asymptotic exponential growth in the ThT fluorescence intensity, the changes do not follow any particular pattern in Lys.

We perform circular dichroism (CD) and mid-IR measurements to underline the secondary structural changes in the proteins during aggregation. From far-UV (190-260 nm) CD measurements (figure 7.III.2D), we observe a less negative molar ellipticity value at pH 2 compared to pH 7.4. This indicates a decrease in α-helical content, as the native protein exhibits two characteristic negative peaks at 208 nm and 222 nm, indicative of its abundant α-helix structure. Interestingly, the decrease in α-helical contents is more significant in presence of the both thiocyanate salts. Then we quantitatively determine the relative abundance of all secondary structural components (α-helix, parallel and antiparallel β-sheet, β-turn and random coil) for each sample. A significant increase in % of the total β-sheet content is clearly observed in presence of the salts with a concomitant decrease in the α-helix contribution (figure 7.III.3A-B). Notably, the effect of acidic pH on the native state is minimal for Ova and Lys but prominently visible in BSA. On the other hand, deconvolution of the amide I band of the mid-

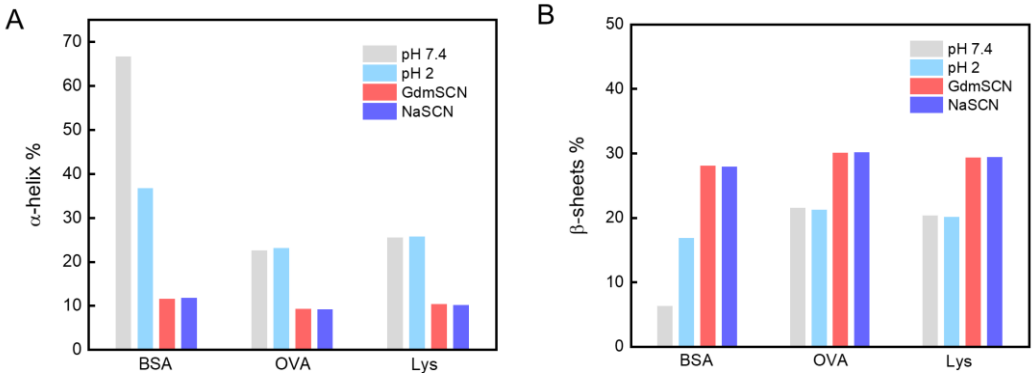


Figure 7.III.3. Percentage change in α-helix (A) and β-sheets (B) structure of all the three proteins at pH 7.4 and at pH 2 in absence or presence of GdmSCN or NaSCN salt.

IR spectra yield contributions of five major structural components: β-sheet at 1620-1625 cm^{-1}, random coil at 1635-1643 cm^{-1}, α-helix at 1653-1659 cm^{-1}, β-turn at 1667-1676 cm^{-1} and anti-parallel β-sheet at 1682-1690 cm^{-1}. All the analysed peak intensity and area % values are summarised in table 7.III.1. The area under each fitted curve corresponds to the relative

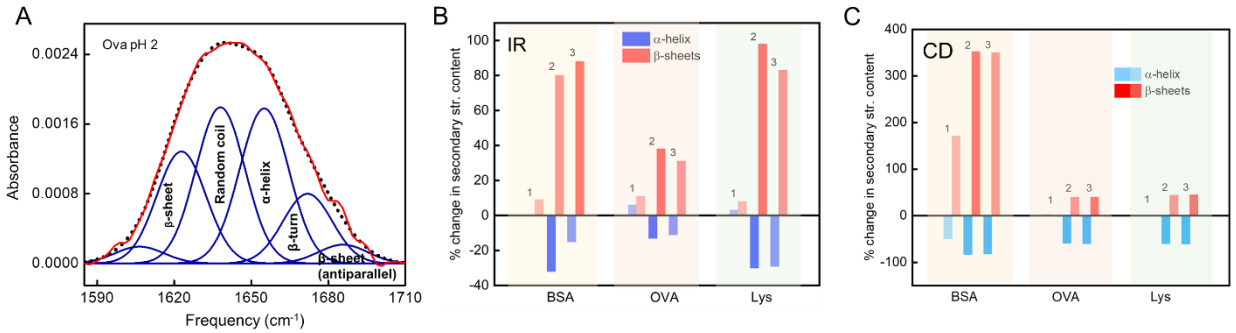


Figure 7.III.4. (A) Representative deconvolute mid-IR spectra of ovalbumin at pH 2. The black dashed line represents raw data and the red line is the overall fitted line. Percentage change in secondary structure content of all the three proteins analysed by mid-IR (B) and CD (C) spectroscopy.

abundance of the different secondary structural components (as marked in the figure 7.III.4A). The % change in secondary structure content (α-helix and total β-sheet content containing anti-parallel β-sheet as well) for all the three proteins are shown in figure 7.III.4B-C. Here we present how the secondary structure of the respective protein gets alter in absence and in presence of salt (GdmSCN or NaSCN) at pH 2 compared to its native state (pH 7.4). Both CD and FTIR results infer a significant alteration in the protein secondary structure in presence of the salts.

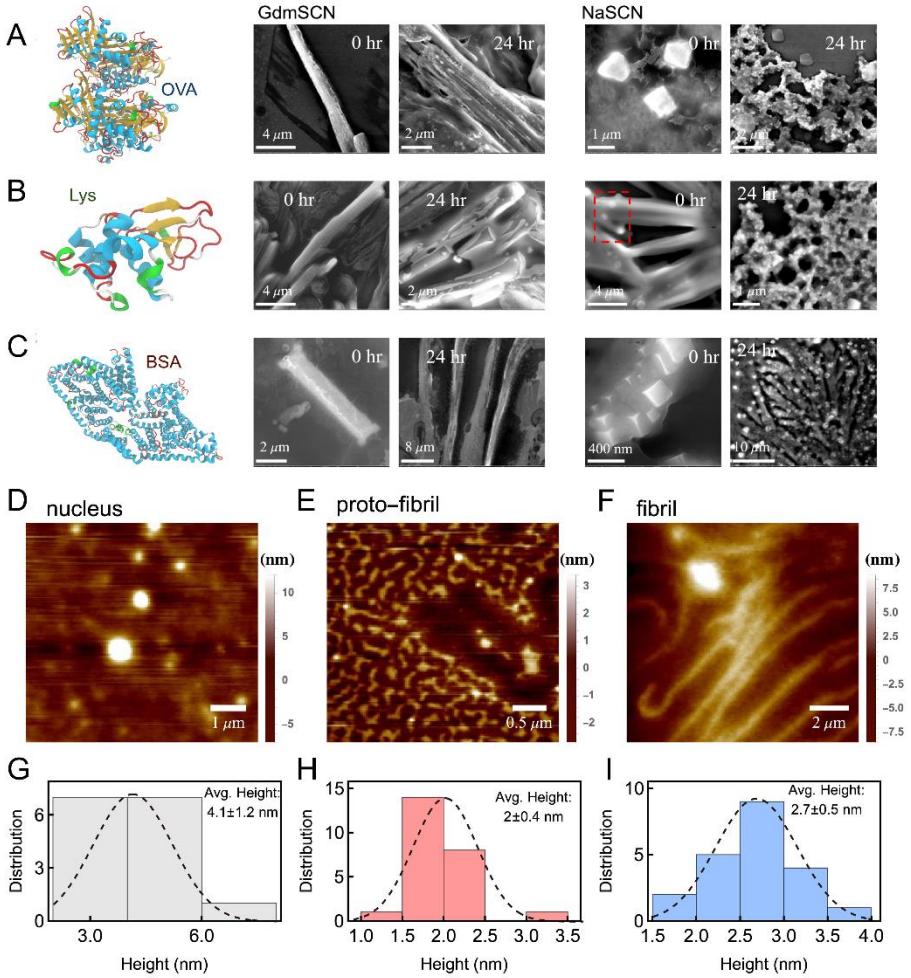


Figure 7.III.5. FESEM images of (A) ovalbumin, (B) lysozyme, and (C) BSA upon the addition of GdmSCN or NaSCN: instantly and after 24 hours standing. (D-F) Representative AFM images of different states of ovalbumin protein fibrillization (e.g., nucleus, proto-fibril and fibril). (G-I) Height distribution profiles as obtained from AFM analysis of these states.

Morphological analysis: We use electron microscopy (FESEM) and atomic force microscopy (AFM) to image the various aggregated states. The morphologies obtained at pH=2 for all the protein-salt pair(s) are depicted in figure 7.III.5. As clearly evident from the images, addition of GdmSCN to proteins forms protofibrils (of dimension ~10-40 μm) at t=0 which then undergo transition to bundle-like mature fibrils upon overnight incubation. In case of NaSCN we observe cubic crystal-like fibrillar nucleus state at t=0 (of dimension ~0.5-2.5 μm) which elongates into fibrillar network like structures upon standing. Interestingly, in case of Ova we notice cross-linking of amyloid protofibrils which co-exists with the cubic nucleuses (figure 7.III.5A-C). The FESEM images thus identify the appearance of three distinct phases during

protein fibrillation process: nucleus, protofibril, and mature fibril. All the three phases are distinctly visualised through AFM micrographs also and the structures exhibit their height in the range of a few *nm* with a *μm* order length. The histograms (figure 7.III.5(G, H, and I) present *nm*-scale height distributions for these fibrillar states. Notably, the observed higher average height of the nucleus phase signifies the initial random accumulation of the monomers, which then gradually organizes and elongates to form the ordered protofibril phases followed by the formation of mature fibrils.

Hydration study: Protein (un)folding or misfolding process is strongly driven by hydrophobic hydration;[33] we employ THz (50-750 cm^{-1}, 1.5-22.5 THz)-FTIR measurements to explore that. The details discussion on the roles of THz frequency window is already provided in the chapter 1. Thus, the measured frequency dependent absorption co-efficient, $\alpha(\nu)$ depicts the linear response of the collective dipole fluctuations in the aqueous system. α(ν) profile of water (buffer pH 7.4; figure 7.III.6A; upper panel) shows two distinct modes which are deconvoluted using a damped harmonic oscillator model (depicted in the section 2.I.D) and we find the H-bonds stretching mode (ν_{HB}) to appear at ~129 cm^{-1} and a restricted rotational mode (libration or ν_{Lib}) to appear at ~565 cm^{-1}. Notably, the $\alpha(\nu)$ profile of buffer at pH 2 (figure 7.III.6A; lower panel) appears slightly different than that observed at pH 7.4; the profile could only be fitted by further deconvoluting the librational band into two sub-Guassian bands with peak positions at ~395 cm^{-1} and ~604 cm^{-1}. The former band is attributed to the intermolecular *"in-plane"* libration (ν_{in}) and the next one for *"out-of-plane"* librational motion (ν_{out}) of water.[34]

α(ν) profile of protein solutions are distinct from that of bulk water as highly absorbing water molecules are replaced by proteins. For a quantitative insight we determine the difference absorption co-efficient, $\Delta\alpha(\nu)$ ($=\alpha_{protein+buffer}(\nu) - \alpha_{buffer}(\nu)$) of these solutions (figure 7.III.6(B, C and D)). It is important to note here that since $\Delta\alpha(\nu)$ is a difference profile, it excludes all the bulk contributions and therefore it offers explicit information on the protein surface hydration. We fit all the experimentally obtained $\Delta\alpha(\nu)$ profiles and the fitted parameters are provided in table 7.III.2. The fitted parameters exemplify the collective properties of protein solvation and they markedly differ from that of bulk. The observed shifts in the vibrational modes imply that all these proteins at pH conditions significantly alter water structure at their surroundings. Interestingly, the bare salt solutions offer an additional vibrational mode in the $\Delta\alpha(\nu)$ ($=\alpha_{salt+buffer}(\nu) - \alpha_{buffer}(\nu)$) profile with positive contributions to appear in the 300-450 cm^{-1} window (figure 7.III.7A-C; table 7.III.2). This

particular mode signifies ion-induced contribution on water structure and is commonly defined as *ion-rattling mode*.[27, 35-37]. It is important to note here that this mode emanates from the dipole-dipole correlation of ions and water, and thus provides with explicit information on the salt hydration.[36, 38, 39]

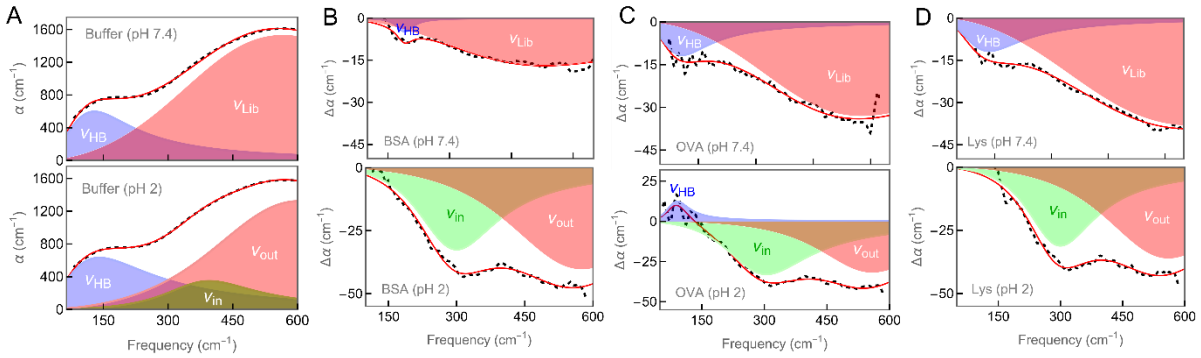


Figure 7.III.6. Frequency dependent absorption co-efficient $\alpha(\nu)$profiles of (A) pH 7.4 (upper) and pH 2 (lower) buffers. Difference absorption co-efficient $\Delta\alpha(\nu)$profiles of protein (B) bovine serum albumin, (C) ovalbumin, and (D) lysozyme at pH 7.4 (above) and at pH 2 (below) respectively.

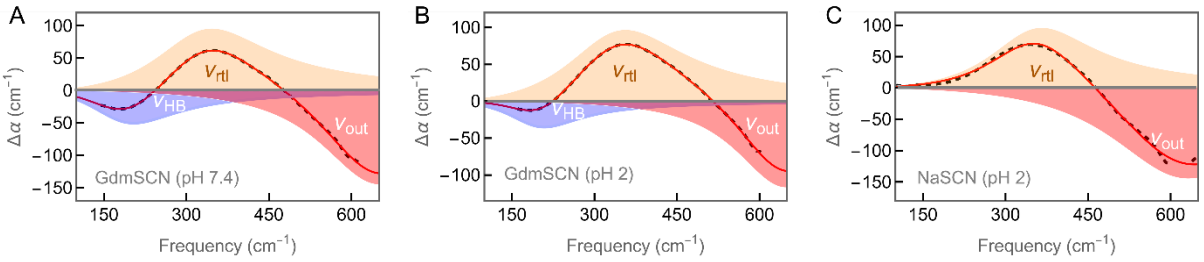


Figure 7.III.7. Frequency dependent difference absorption co-efficient $\Delta\alpha(\nu)$ profile of (A) GdmSCN at pH 7.4, (B) GdmSCN at pH 2 and (C) NaSCN at pH 2.

To obtain solvation change during aggregation we obtain the $\Delta\alpha_{agg}(\nu)$ profiles ($= \alpha_{protein+salt+water}(\nu) - \alpha_{salt+water}(\nu)$) and fitted them as before. The results are summarized in figure 7.III.8, and table 7.III.3-7.III.5. We observe that solvation change is dependent both on the type of the salt as well as on the incubation time. In figure 7.III.8, we present the peak frequencies of the different vibrational modes at different aggregated stages. It is interesting to find that protein-salt pair(s) offer varying solvation features irrespective of their comparable final fibrillar state(s). For example, in NaSCN-Ova system the ν_{HB} mode appears at $\sim$ 120 cm^{-1} which no longer is detectable as mature fibrils are formed upon

incubation. Similar trend is observed in Lys-NaSCN pair. However, in the case of BSA, while the ν_{HB} mode remains undetectable in the nucleus state, it reappears in the matrix-like mature fibrillar form (at ~84 cm^{-1}). In case of GdmSCN-Lys pair the ν_{HB} appears at ~202 cm^{-1} in protofibrils, which experiences a minor red-shift to ~197 cm^{-1} in mature fibrils, while no ν_{HB} mode is detectable in case of the other two proteins. Such changes are also evident in the ion-solvation rattling mode as well (figure 7.III.8, table 7.III.3-7.III.5).

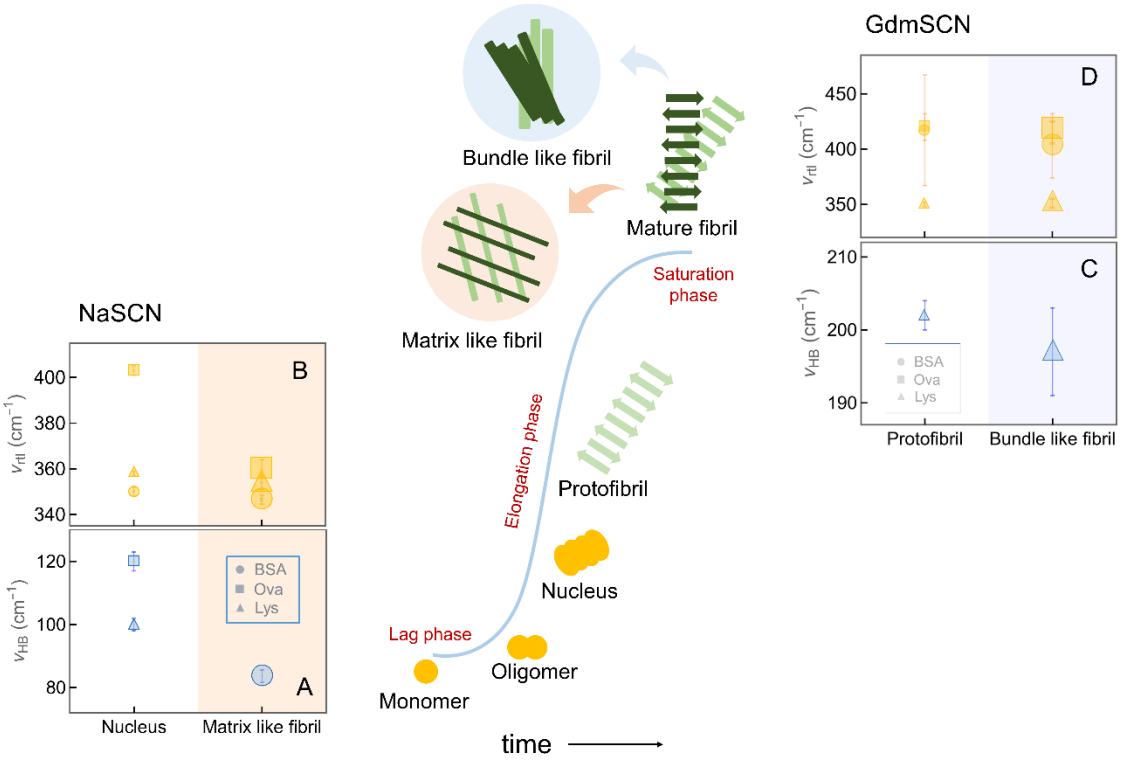


Figure 7.III.8. HB stretching frequency (ν_{HB}) and ion rattling frequency (ν_{rtl}) of water at protein (BSA (filled circle), Ova (filled square) and Lys (filled triangle)) interface in presence of NaSCN (A-B) and GdmSCN (C-D). Note that in NaSCN matrix like fibril are formed while in GdmSCN bundle like fibril are formed.

Discussions: SCN^- instigates unusually instantaneous fibrillar aggregation in proteins at pH=2 at room temperature. Considering the otherwise naive characteristic of NaSCN the observed rapid aggregation of proteins in its presence at the ambient temperature is intriguing. The fact is worth noticing considering the possibility of transient accumulation of the ions in body parts, specifically in the digestive tract, where such a low pH if often encountered. Acidic pH environment, which leads to partial unfolding of the protein monomers,[40] can initiate polymorphisms through nucleation process. It is worthy to note here that acidic pH

environment itself does not induce aggregate formation in the studied proteins as evidenced from the ThT fluorescence measurements (figure 7.III.2A). On the other hand, NaSCN, at neutral pH, does not initiate aggregation. This apparent incongruity strongly concludes the decisive role of the SCN^- ions to induce the partially unfolded proteins to aggregate. We also observe intriguing cation-specific differences on the morphologies of the aggregated states; while GdmSCN forms a collection of rod-like (bundle-like) mature fibrils, NaSCN induces nanofibrillar crosslinking (matrix or network-like aggregates) (figure 7.III.5A-C).

In general, formation of amyloid fibrils results as a consequence of protein misfolding.[41] Stepwise association of un/mis-folded proteins[42] leads to the formation of unordered soluble oligomers;[32] such oligomers serve as an intermediate during monomer to crystal-like nuclei formation. Melting of α-helix and extension of β-sheets in proteins gives rise to the formation of critical nuclei which is a key step in the fibrillization process as it accelerates its further growth.[43, 44] Zipper-like intertwining (along fibre axis) of parallel and/or antiparallel β-sheets of protofilament units results in the formation of flexible protofibrils of *nm* dimensions. These protofibrils further intertwine to form more rigid *μm* length mature fibrils having diameters in the *nm* dimension.[42] A theoretical study by *krone et. al.* revealed intertwining of protofilament is associated with distinct changes in the interfacial solvation: either some water molecules get trapped as a "*lubricant*" and/or "*de-wet*" the protofilament surface;[45] either way, the change in the protein solvation dictates the fibrilization process.[46-50] We explore the solvation status of these fibrillar-aggregates using THz-FTIR hydration (figure 7.III.8, table 7.III.3-7.III.5). In these multicomponent systems, deconvolution of the $\alpha_{THz}(\nu)$ profiles into different contributions ($\alpha(\nu) = \sum \phi_i \alpha_i(\nu)$) where ϕ_i denotes the volume fraction and $\alpha_i(\nu)$ is the frequency dependent absorption co-efficient of the *i*-th component) leads us to obtain the solvation status of each component under different phases[29] considering the fact that protein and salt hydration has non-overlapping and distinct fingerprint regions in the THz window.[23, 37] According to the additive contributions of all the hydration components to the absorption profile, we can further demonstrate the $\alpha(\nu)$ in terms of,

$$\alpha_{p+s+w}(\nu) = \phi_p \alpha_p(\nu) + \phi_{ph} \alpha_{ph}(\nu) + \phi_s \alpha_s(\nu) + \phi_{sh} \alpha_{sh}(\nu) + \phi_b \alpha_{bulk}(\nu) \quad (7.1)$$

$$\alpha_{s+w}(\nu) = \phi'_s \alpha_s(\nu) + \phi'_{sh} \alpha_{sh}(\nu) + \phi'_b \alpha_{bulk}(\nu) \quad (7.2)$$

where subscript '*p*', '*s*', 'w' and '*b*' stands for protein, salt, water content and bulk component respectively whereas '*ph*' and '*sh*' denotes protein and salt hydration shell respectively.

Now, we can write the explicit protein hydration as, $\Delta\alpha_p(\nu) = \alpha_{p+s+w}(\nu) - \alpha_{s+w}(\nu)$;

$$\Delta\alpha_p(\nu) = \phi_{ph}\alpha_{ph}(\nu) + (\phi_{sh} - \phi'_{sh})\alpha_{sh}(\nu) + (\phi_b - \phi'_b)\,\alpha_{bulk}(\nu) \quad (7.3)$$

Since we cannot ascertain the oligomer state during rapid nucleation, the initial perturbation in protein surface hydration from monomeric bare protein surface remains untouched. Rather we choose to monitor evolution of the protein hydration throughout the 24-hour transition period.

We can write $\Delta\alpha_p$ as per the protein hydration at four different aggregation states; Nucleus (n), protofibril (pf), bundle like mature fibril (bf) and matrix like mature fibril (mf) as,

$$\Delta\alpha_p^n(\nu) = \phi_{ph}^n\alpha_{ph}(\nu) + \Delta\phi_{sh}^n\alpha_{sh}(\nu) + \Delta\phi_b^n\alpha_{bulk}(\nu) \quad (7.4)$$

$$\Delta\alpha_p^{pf}(\nu) = \phi_{ph}^{pf}\alpha_{ph}(\nu) + \Delta\phi_{sh}^{pf}\alpha_{sh}(\nu) + \Delta\phi_b^{pf}\alpha_{bulk}(\nu) \quad (7.5)$$

$$\Delta\alpha_p^{bf}(\nu) = \phi_{ph}^{bf}\alpha_{ph}(\nu) + \Delta\phi_{sh}^{bf}\alpha_{sh}(\nu) + \Delta\phi_b^{bf}\alpha_{bulk}(\nu) \quad (7.6)$$

$$\Delta\alpha_p^{mf}(\nu) = \phi_{ph}^{mf}\alpha_{ph}(\nu) + \Delta\phi_{sh}^{mf}\alpha_{sh}(\nu) + \Delta\phi_b^{mf}\alpha_{bulk}(\nu) \quad (7.7)$$

These analytical expressions help in detail understanding of the hydration effect in the observed fibrillar transitions.

We make our further discussion based on the observed changes in the amplitudes (ΔA, measured at the central frequency ($\nu_{d,i}$) of the corresponding mode) which defines the oscillator strength of different vibrational modes of water dipoles (figure 7.III.9). For a proto-fibril *(pf)* to bundle fibril *(bf)* transition (in case of GdmSCN) the change in the frequency dependent amplitude (ΔA) for any particular mode is given by,

$$\Delta A^{pf\to bf} = (\phi_{ph}^{bf} - \phi_{ph}^{pf})\alpha_{ph}(\nu) + (\Delta\phi_{sh}^{bf} - \Delta\phi_{sh}^{pf})\alpha_{sh}(\nu) + (\Delta\phi_b^{bf} - \Delta\phi_b^{pf})\alpha_{bulk}(\nu) \quad (7.8)$$

Correspondingly, for the nucleus *(n)* to matrix fibril *(mf)* transition (in case of NaSCN),

$$\Delta A^{n\to mf} = (\phi_{ph}^{mf} - \phi_{ph}^{n})\alpha_{ph}(\nu) + (\Delta\phi_{sh}^{mf} - \Delta\phi_{sh}^{n})\alpha_{sh}(\nu) + (\Delta\phi_b^{mf} - \Delta\phi_b^{n})\alpha_{bulk}(\nu) \quad (7.9)$$

Any change in the protein as well as salt hydration can be apprehended by these set of equations (7.8-7.9). Note that the sign of ΔA^i depends on the difference terms in volume fractions (ϕ_i) as the α_i terms are always positive. The magnitude, however, is determined by all the terms.

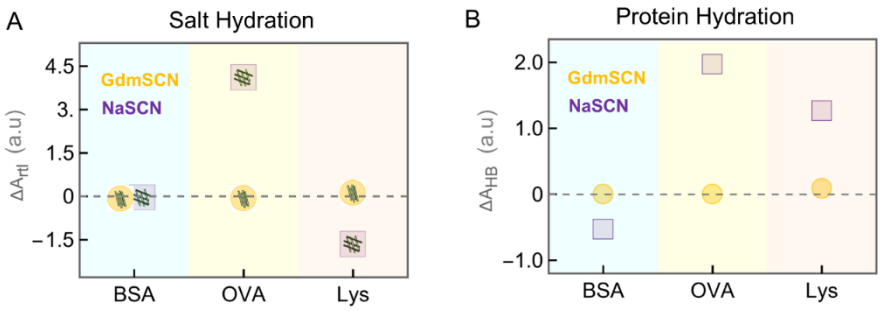


Figure 7.III.9. Changes in the amplitude of (A) ion-water rattling (ΔA_{rtl}) mode and (B) intermolecular H-bond stretching mode (ΔA_{HB}) of BSA, Ova and Lys in presence of salt(s) as bundle like fibril transitions (($\Delta A_{pf\rightarrow bf}$) in presence of GdmSCN, circles) and nucleus to matrix like fibril transition (($\Delta A_{n\rightarrow mf}$) in presence of NaSCN, square symbols) takes place.

We will focus our discussion on two major aspects: the change in hydration of the salts and that of the protein as the phase transitions occur. The two major parameters to be considered are: ΔA_{rtl} and ΔA_{HB}; while the first one is exclusive for the ions, the second one defines protein hydration. It is important to note here that the water dipole correlation of these two modes do not interfere,[35, 36] and therefore enables one to extract independent hydration information. We plot these two parameters for the protein-salt pair(s) in figure 7.III.9. We first discuss ion-hydration (figure 7.III.9A): the parameter ΔA_{rtl} is governed by the second term in the right-hand side of equation 7.8 and 7.9 as the magnitude of the α_{ph} and α_{bulk} terms are very small compared to that of α_{sh} in this frequency. So, ΔA_{rtl} directly correlates the extent of the collective rattling motion between Gdm^+ and SCN^- ions or Na^+ and SCN^- and the surrounding water molecules within their hydration shell. It is not possible to separate out the anion-water and cation-water contributions individually. However, the following rationale could be made: the rattling frequencies and amplitudes of GdmSCN and NaSCN are comparable (table 7.III.2), this suggests that either all the ions contribute equally or the anion dominates. The later can further be argued on that fact that rattling frequency of Na_2SO_4 is ~20 cm^{-1} blue shifted with a higher amplitude than NaCl. As evidenced from figure 7.III.9A, ΔA_{rtl} ~0 for protofibril to fibril transitions (GdmSCN-protein pairs) while for the NaSCN-protein pairs it shows non-zero values. This unambiguously confirms that the hydration state of GdmSCN does not change appreciably during the phase transition process, however, for NaSCN it does. This observation corroborates very well with a theoretical study in which it was concluded that during aggregation Na^+ is partitioned away from the collapsed state of a model peptide and SCN^- ions are excluded in the bulk.[51] Thus, keeping in mind that Na^+ remains in the bulk throughout, the

solvation environment of SCN^- changes on phase transition and the term $(\Delta\phi_{sh}^{mf} - \Delta\phi_{sh}^{n})$ in equation 7.9 offers a non-zero value. The magnitude of the difference term is dependent on the SCN^- immediate environment, which might differ with its specific interaction with protein(s). On the other hand, Gdm^+ always remains in the protein vicinity, both in protofibril and in bundle fibril states, and therefore the term in equation 7.8 $(\Delta\phi_{sh}^{bf} - \Delta\phi_{sh}^{pf}) \approx 0$.

Regarding protein hydration it could be noted that in equations 7.8 and 7.9 both the first and the third terms in the right-hand side contributes to ΔA_{HB}. Any de- or re-hydration in the protein surface, as manifested by the terms $(\phi_{ph}^{f} - \phi_{ph}^{i})$ (here 'i' and 'f' stands for initial and final states) is adjusted in the subsequent change in bulk solvation $(\phi_{b}^{f} - \phi_{b}^{i})$. We found that ΔA_{HB} for all the three proteins are nearly zero in GdmSCN. This suggests that there occurs no further alteration in the intermolecular H-bond structure during protofibril to bundle like mature fibril formation at the protein surface; consequently, both the first and the third term in equation 7.8 becomes zero. In case of NaSCN, however, we obtain nonzero ΔA_{HB} values (figure 7.III.9B) indicating that protein solvation undergoes dehydration when water is released from aggregate interior to the bulk during fibrillar crosslinking. The negative ΔA_{HB}^{BSA} value affirms strong interactions between ions and the protein surface during the elongation process rupturing both of their solvation shells and thereby sum of first two terms in equation 7.8 are negative enough to overweigh the positive contribution from third term. On the other hand, the positive values of ΔA_{HB}^{OVA} and ΔA_{HB}^{Lys} suggest a high positive $\Delta\phi_b$ (equation 7.8) which overwhelms the other two negative terms. The interactions thus are found to be extremely ion-protein specific in nature.

Our study provides a comprehensive understanding of how ion-specific hydration induces distinct arrangements in mature fibrils. Weakly hydrated SCN^- ions initially bind to the protein surface initiating the self-aggregation of protein molecules. Depending upon its counter-cation the growth follows different routes. The bulky cation Gdm^+, which does not associate a finite hydration shell,[52] experiences strong interactions with the aromatic surface residues of the partially unfolded proteins through $\pi - \pi$ stacking.[53] Both Gdm^+ and SCN^- ions participate in binding interactions with the protein surface residues. Gdm^+ ions further form bridges between the neighbouring fibril strands and arrange themselves into a bundle-like structure. The active role of Gdm^+ in the formation of the bundle like structure is substantiated by its solvation characteristics (figure 7.III.9A). On the other hand, strongly hydrated Na^+ ions do not experience such binding interactions with protein residues, rather it remains soluble in

the bulk as supported by the negligible change in the salt solvation (figure 7.III.9A). This results in a faster fibrillation rate in GdmSCN compared to the crosslinking of amyloid-fibrils in NaSCN (see figure 7.III.2B, C). Faster bridging of mature fibrils by GdmSCN maintains the collective hydration unchanged from the protofibril phase, however, the elongation rate for Gdm^+ interactions is highly protein specific (figure 7.III.2C). Additionally, SCN^- ions bound fibrillar nuclei form crosslinked network with equivalent growth rate for all three proteins whereas the collective hydration exhibits a distinct change in the water structure. This leads us to infer strongly that water acts as a *"lubricant"* in the bundle interface, while the formation of crosslinked matrix arrangements is driven by *"de-wetting"*.

7.IV. Conclusions.

Our experimental finding offers a primary evidence that SCN^- instigates instantaneous protein aggregation and amyloid fibril formation at pH 2 environment. We further conclude that it is the counter-cation that leads in shaping the final arrangements of the mature fibrils. It is noteworthy that water behaves differently in the interface of the mature fibrils. Our study confirms that hydration pattern exerts a specific effect on each protein as they undergo fibrillar transition. Our study affirms the fact that accumulation of SCN^- ions and an abrupt elevation of their abundance in blood plasma or extracellular fluid can lead to severe damage of human body. Considering that SCN^- ion is used as a biomarker in the diagnostic of different medical conditions, diffusion of it into the brain and cerebrospinal fluid (CSF) through vascular infusion thus becomes a notable consideration. Also, that both Na^+ and SCN^- ions are permeable through the blood-brain barrier (BBB), increase in plasma concentration could lead to heightened risk of neurodegenerative disorders. Keeping these facts in mind our result finds its novelty in basic understanding of the fibrillar growth process by SCN^-- protein interaction and indulge further studies on this aspect.

Table 7.III.1. The listed values of the fitted parameters after deconvolution of mid-IR spectra. All peak positions, area % and intensity values of each secondary structural components of protein systems.

Components	Protein	Conditions	Peak position	Area %
β sheet (including β-antiparallel)	BSA	pH 7.4	1620 1697	11.7 1.5
		pH 2	1628 1688	11.8 2.6
		GdmSCN (pH 2)	1625 1685	15.6 8.2
		NaSCN (pH 2)	1622 1681	19.6 5.2

	Ova	pH 7.4	1623 1686	16.5 5.7
		GdmSCN (pH 7.4)	1624 1687	18.7 9.4
		pH 2	1623 1686	21.2 3.5
		GdmSCN (pH 2)	1625 1682	20.9 9.7
		NaSCN (pH 2)	1624 1683	25.3 3.8
	Lys	pH 7.4	1622 1690	11.4 4.5
		GdmSCN (pH 7.4)	1624 1693	12.2 5.9
		pH 2	1622 1689	12.6 4.6
		GdmSCN (pH 2)	1621 1682	18.8 12.7
		NaSCN (pH 2)	1624 1683	25.3 3.8
Random Coil	BSA	pH 7.4	1636	26.7
		pH 2	1637	17.7
		GdmSCN (pH 2)	1643	21.4
		NaSCN (pH 2)	1639	26.3
	Ova	pH 7.4	1638	30.0
		GdmSCN (pH 7.4)	1641	23.1
		pH 2	1638	29.5
		GdmSCN (pH 2)	1642	28.3
		NaSCN (pH 2)	1640	31.1
	Lys	pH 7.4	1636	22.7
		GdmSCN (pH 7.4)	1640	19.0
		pH 2	1636	19.3
		GdmSCN pH 2	1641	25.1
		NaSCN pH 2	1640	31.1
α-helix	BSA	pH 7.4	1655	43.0
		pH 2	1653	43.0
		GdmSCN (pH 2)	1660	29.2
		NaSCN (pH 2)	1656	36.7
		pH 7.4	1656	27.7
		GdmSCN (pH 7.4)	1657	25.9

	Ova	pH 2	1655	29.3
		GdmSCN (pH 2)	1660	24.2
		NaSCN (pH 2)	1659	24.7
	Lys	pH 7.4	1654	34.7
		GdmSCN pH 7.4	1657	32.2
		Lys (pH 2)	1653	35.7
		GdmSCN (pH 2)	1660	26.6
		NaSCN (pH 2)	1659	24.7
β-turn	BSA	pH 7.4	1676	15.0
		pH 2	1672	16.7
		GdmSCN (pH 2)	1675	19.9
		NaSCN (pH 2)	1672	8.3
	Ova	pH 7.4	1672	13.9
		GdmSCN pH 7.4	1672	18.5
		pH 2	1672	13.2
		GdmSCN (pH 2)	1669	8.9
		NaSCN (pH 2)	1672	7.2
	Lys	pH 7.4	1672	21.0
		GdmSCN (pH 7.4)	1676	26.4
		pH 2	1671	23.3
		GdmSCN (pH 2)	1671	8.5
		NaSCN (pH 2)	1672	7.2

Table 7.III.2. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c) of respective modes (i.e. for hydrogen bond stretching mode, for ion rattling motion, for water in-plane and out of plane librational motion): Respective values are fitted results of pH 7.4 solution, pH 2 solution, each protein and salt solution at both pH.

Protein	pH	Mode	a_0 ($\times 10^{-6}$ a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
0	pH 7.4	ν_{HB}	4.81	641.2	79.6(2)	129.4(3)
		ν_{Lib}	4.26	2210	443.2(2)	565.8(3)
		ν_{HB}	6.71	845	41.2(4)	140.6(4)
		ν_{in}	3.61	827.1	373(1)	395.5(5)

	pH 2	ν_{out}	29.19	1753.6	536(4)	604.2(5)
Bovine Serum Albumin	pH 7.4	ν_{HB}	-0.009	155.8	189(2)	190.6(3)
		ν_{Lib}	-0.48	2270	389.7(8)	531.5(13)
	pH 2	ν_{in}	-0.28	685.5	279.6(2)	300(4)
		ν_{out}	-0.68	1338	534.8(4)	575.6(23)
Ovalbumin	pH 7.4	ν_{HB}	-0.08.	520.6	70(12)	108.2(16)
		ν_{Lib}	-1.02	2446.7	424.8(14)	576.3(30)
	pH 2	ν_{HB}	0.03	215	84.8(2)	91.4(4)
		ν_{in}	-0.32	779.4	278.6(2)	305(8)
		ν_{out}	-0.41	1031.5	533.5(5)	558.1(21)
Lysozyme	pH 7.4	ν_{HB}	-0.09	643.7	88.3(9)	135.3(11)
		ν_{Lib}	-1.51	3136.1	453.1(12)	674.1(22)
	pH 2	ν_{in}	-0.25	639.8	282(2)	300(6)
		ν_{out}	-0.58	1267.5	522.8(6)	560.4(28)
GdmSCN	pH 7.4	ν_{HB}	-0.39	614.5	179.2(1)	204.2(4)
		ν_{rtl}	0.97	803.7	321.6(2)	346(3)
		ν_{Lib}	-1.54	847.8	637.1(3)	651.3(5)
	pH 2	ν_{HB}	-0.23	510.3	192.4(0.7)	208.8(3)
		ν_{rtl}	1.02	836	330.7(1)	356.5(2)
		ν_{Lib}	-1.06	723	638.1(4)	648.4(5)
NaSCN	pH 2	ν_{rtl}	0.90	746.3	346.1(1)	366(3)
		ν_{Lib}	-2	1099.2	607.3(2)	632(4)

Table 7.III.3. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c) of respective modes (i.e. for hydrogen bond stretching mode, for ion rattling motion, for water in-plane and out of plane librational motion): Respective values are fitted results of each protein-GdmSCN solution at pH 7.4.

Protein	Time (hr)	Mode	a_0 ($\times 10^{-6}$ a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
Ovalbumin	0	ν_{rtl}	-0.60	701.7	351.4(0.7)	368.7(2)
		ν_{Lib}	0.16	272	607.5(1)	609(2)
	24	ν_{HB}	-0.27	369.8	94(1)	111(2)
		ν_{rtl}	-0.71	624.6	336.5(1)	351(3)
		ν_{Lib}	-0.35	575.4	454(3)	463.1(4)

Lysozyme	0	ν_{HB}	0.21	581.2	190(3)	211.3(8)
		ν_{rtl}	-0.61	786.7	325.3(4)	348.6(6)
		ν_{Lib}	0.72	695.7	622.7(3)	632.5(6)
	24	ν_{HB}	0.12	463.6	193.3(3)	207(7)
		ν_{rtl}	-0.44	678.6	346.1(3)	362.6(5)
		ν_{Lib}	0.37	470.7	617.3(3)	621.9(6)

Table 7.III.4. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c) of respective modes (i.e. for hydrogen bond stretching mode, for ion rattling motion, for water in-plane and out of plane librational motion) : Respective values are fitted results of each protein-GdmSCN solution at pH 2.

Protein	Time (hr)	Mode	a_0 ($\times 10^{-6}$ a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
Bovine Serum Albumin	0	ν_{in}	-0.27	438.2	318.6(8)	326(18)
		ν_{rtl}	-0.37	714.3	401(41)	417(50)
		ν_{out}	0.12	279.5	592.3(3)	594(10)
	24	ν_{in}	-0.13	339.3	314(7)	318.6(22)
		ν_{rtl}	-0.45	687.8	387.7(25)	402.8(29)
		ν_{out}	0.11	262	598.4(4)	599.8(10)
Ovalbumin	0	ν_{in}	-0.20	427.3	320(3)	327.2(6)
		ν_{rtl}	-0.24	605.2	409(9)	420.1(13)
		ν_{out}	0.11	276.3	608(1)	610(3)
	24	ν_{in}	-0.17	352.1	323.2(3)	328(6)
		ν_{rtl}	-0.31	560.6	405.7(8)	415.4(11)
		ν_{out}	0.20	265.3	624.3(9)	625.7(10)
Lysozyme	0	ν_{HB}	0.11	405	191.4(1)	202(2)
		ν_{rtl}	-0.63	754.1	329(1)	350.2(2)
		ν_{out}	0.36	549	605.1(1)	611.4(3)
	24	ν_{HB}	0.19	577.8	174.6(2)	197.3(6)
		ν_{rtl}	-0.52	734.7	331.1(2)	351.2(4)
		ν_{out}	0.83	683.3	652.1(9)	661.1(11)

Table 7.III.5. The listed values of the fitted parameters; amplitude (a_0), damping width (ω_0), damped frequency (ν_d), and unperturbed center frequency (ν_c) of respective modes (i.e. for hydrogen bond stretching mode, for ion rattling motion, for water in-plane and out of plane librational motion): Respective values are fitted results of each protein-NaSCN solution at pH 2.

Protein	Time (hr)	Mode	a_0 ($\times 10^{-6}$ a.u)	ω_0 (cm^{-1})	ν_d (cm^{-1})	ν_c (cm^{-1})
Bovine Serum Albumin	0	ν_{rtl}	-0.64	722.7	330.7(0.8)	350.1(2)
		ν_{out}	0.85	805.6	613.2(2)	626.5(5)
	24	ν_{HB}	-0.53	388.8	56(2)	83.6(2)
		ν_{rtl}	-0.69	698.2	328.2	346.5(2)
		ν_{out}	0.98	855	634.5	649(10)
		ν_{HB}	-1.97	522	86.6(3)	120(3)

Ovalbumin	0	ν_{rtl}	-4.45	1061.8	366.3(1)	403.4(2)
	24	ν_{rtl}	-0.35	563.3	348.7(2)	360(4)
		ν_{out}	0.91	784.9	614.7(5)	627.3(10)
Lysozyme	0	ν_{HB}	-1.26	533.6	52.4(1)	100(2)
		ν_{rtl}	1.25	759.1	336.7(0.7)	357.7(1)
		ν_{out}	0.57	666.8	620(1)	629(3)
	24	ν_{rtl}	-0.40	583	342(1)	354.4(3)
		ν_{out}	1.07	803.2	614.1(3)	627.2(7)

7.V. References.

1. Y. Xu, S. Szép and Z. Lu, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 20515-20519.
2. M. A. Fragoso, V. Fernandez, R. Forteza, S. H. Randell, M. Salathe and G. E. Conner, *J. Physiol.*, 2004, **561**, 183-194.
3. P. T. San Gabriel, Y. Liu, A. L. Schroder, H. Zoellner and B. Chami, *Int. J. Mol. Sci.*, 2020, **21**, 6450.
4. A. M. Leung, A. LaMar, X. He, L. E. Braverman and E. N. Pearce, *J. Clin. Endocrinol. Metab.*, 2011, **96**, E1303.
5. P. Felker, R. Bunch and A. M. Leung, *Nutr Rev.*, 2016, **74**, 248-258.
6. C. V. Kalburgi, K. L. Naik, M. V. Kokatnur and S. Warad, *Contemp Clin Dent.*, 2014, **5**, 182-186.
7. R. K. Bhandari, R. P. Oda, I. Petrikovics, D. E. Thompson, M. Brenner, S. B. Mahon, V. S. Bebarta, G. A. Rockwood and B. A. Logue, *J. Anal. Toxicol.*, 2014, **38**, 218-225.
8. S. Buonvino, I. Arciero and S. Melino, *Int. J. Mol. Sci.*, 2022, **23**, 8452.
9. M. Wróbel, H. Jurkowska, L. Śliwa and Z. Srebro, *Toxicol Mech Methods.*, 2004, **14**, 331-337.
10. J. Bitman, H. C. Cecil, M. R. Connolly, R. W. Miller, M. Okamoto, J. W. Thomas and T. R. Wrenn, *J. Dairy Sci.*, 1962, **45**, 872-878.
11. S. Narkowicz, E. Jaszczak, Ż. Polkowska, B. Kiełbratowska, A. Kotłowska and J. Namieśnik, *Biomed. Chromatogr.*, 2018, **32**, e4111.
12. M. T. Ashby, in *Advances in inorganic chemistry*, Elsevier, 2012, vol. 64, pp. 263-303.
13. E. D. Freis, in *Hypertension (Second Edition)*, eds. S. Oparil and M. A. Weber, W.B. Saunders, Philadelphia, 2005, pp. 1-6.
14. W. S. Ruddell, E. S. Bone, M. J. Hill, L. M. Blendis and C. L. Walters, *Lancet.*, 1976, **2**, 1037-1039.
15. M. Heliövaara, M. J. Karvonen, S. Punsar, Y. Rautanen and J. Haapakoski, *J Chronic Dis.*, 1981, **34**, 305-311.
16. W. Lijinsky and R. M. Kovatch, *Toxicol Ind Health.*, 1989, **5**, 25-29.
17. M. Pollay and R. J. Kaplan, *Am J Physiol.*, 1970, **219**, 802-808.
18. D. B. Cheek and A. B. Holt, *Pediatr Res.*, 1978, **12**, 635-645.
19. R. L. Baldwin, *Biophys. J.*, 1996, **71**, 2056-2063.
20. P. E. Stein, A. G. W. Leslie, J. T. Finch and R. W. Carrell, *J. Mol. Biol.*, 1991, **221**, 941-959.
21. T. Topală, A. Bodoki, L. Oprean and R. Oprean, *Clujul Med.*, 2014, **87**, 215-219.
22. P. Ferraboschi, S. Ciceri and P. Grisenti, *Antibiotics (Basel, Switzerland)*, 2021, **10**.
23. S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 20749-20752.
24. M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 12068-12073.
25. B. König, S. Pezzotti, S. Ramos, G. Schwaab and M. Havenith, *Biophys. J.*, 2023, **S0006-3495**, 00474-00475.
26. P. Pyne and R. K. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
27. R. Saha and R. K. Mitra, *Langmuir*, 2023, **39**, 4601-4610.
28. J. Sun, G. Niehues, H. Forbert, D. Decka, G. Schwaab, D. Marx and M. Havenith, *J. Am. Chem. Soc.*, 2014, **136**, 5031-5038.

29. H. Vondracek, J. Dielmann-Gessner, W. Lubitz, M. Knipp and M. Havenith, *J. Chem. Phys*, 2014, **141**, 22d534.
30. C. Xue, T. Lin, D. Chang and Z. Guo, *R. Soc. Open Sci.*, 2017, **4**, 160696.
31. R. Crespo, E. Villar-Alvarez, P. Taboada, F. A. Rocha, A. M. Damas and P. M. Martins, *J. Biol. Chem.*, 2016, **291**.
32. P. Arosio, T. P. J. Knowles and S. Linse, *Phys. Chem. Chem. Phys.*, 2015, **17**, 7606-7618.
33. A. Kessel and N. Ben-Tal, *Introduction to Proteins: Structure, Function, and Motion, Second Edition*, CRC Press, 2018.
34. P. Pyne and R. K. J. T. J. o. P. C. L. Mitra, *J. Phys. Chem. Lett.*, 2022, **13**, 931-938.
35. C. Klinkhammer, F. Böhm, V. S. Sharma, G. W. Schwaab, M. Seitz and M. Havenith, *J Chem Phys.*, 2018, **148 22**, 222802.
36. R. Saha and R. K. Mitra, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23661-23668.
37. D. Sharma and A. Chandra, *J. Phys. Chem. B*, 2023, **127**, 9323-9335.
38. D. A. Schmidt, Ö. Birer, S. Funkner, B. P. Born, R. Gnanasekaran, G. W. Schwaab, D. M. Leitner and M. Havenith, *J. Am. Chem. Soc.*, 2009, **131**, 18512-18517.
39. G. Schwaab, F. Sebastiani and M. Havenith, *Angew. Chem. Int. Ed.*, 2019, **58**, 3000-3013.
40. M. Bhattacharya and S. Mukhopadhyay, *J. Phys. Chem. B*, 2012, **116**, 520-531.
41. F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2017, **86**, 27-68.
42. L. R. Volpatti, M. Vendruscolo, C. M. Dobson and T. P. J. Knowles, *ACS Nano*, 2013, **7**, 10443-10448.
43. T. C. T. Michaels, D. Qian, A. Šarić, M. Vendruscolo, S. Linse and T. P. J. Knowles, *Nat. Rev. Phys.*, 2023, **5**, 379-397.
44. T. R. Jahn, O. S. Makin, K. L. Morris, K. E. Marshall, P. Tian, P. Sikorski and L. C. Serpell, *J. Mol. Biol.*, 2010, **395**, 717-727.
45. M. G. Krone, L. Hua, P. Soto, R. Zhou, B. J. Berne and J.-E. Shea, *J. Am. Chem. Soc.*, 2008, **130**, 11066-11072.
46. J. D. Camino, P. Gracia and N. Cremades, *Biophys. Chem.*, 2021, **269**, 106520.
47. Y. Fichou, G. Schirò, F.-X. Gallat, C. Laguri, M. Moulin, J. Combet, M. Zamponi, M. Härtlein, C. Picart, E. Mossou, H. Lortat-Jacob, J.-P. Colletier, D. J. Tobias and M. Weik, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 6365-6370.
48. L. A. Munishkina, J. Henriques, V. N. Uversky and A. L. Fink, *Biochemistry*, 2004, **43**, 3289-3300.
49. D. Thirumalai, G. Reddy and J. E. Straub, *Acc. Chem. Res.*, 2012, **45**, 83-92.
50. T. Wang, H. Jo, W. F. DeGrado and M. Hong, *J. Am. Chem. Soc.*, 2017, **139**, 6242-6252.
51. J. Heyda, H. I. Okur, J. Hladílková, K. B. Rembert, W. Hunn, T. Yang, J. Dzubiella, P. Jungwirth and P. S. Cremer, *J. Am. Chem. Soc.*, 2017, **139**, 863-870.
52. P. E. Mason, G. W. Neilson, J. E. Enderby, M.-L. Saboungi, C. E. Dempsey, A. D. MacKerell, Jr. and J. W. Brady, *J. Am. Chem. Soc.*, 2004, **126**, 11462-11470.
53. P. E. Mason, C. E. Dempsey, G. W. Neilson, S. R. Kline and J. W. Brady, *J. Am. Chem. Soc.*, 2009, **131**, 16689-16696.

Chapter 8

8. Summary and Future Perspective

8.I. Summary.

The thesis is based on some spectroscopic studies to determine the conformational and functional insights of biomolecules in presence of co-solutes. For that we choose protein as a biomolecule and different types of ions as co-solutes. In the initial three works, we deal with the trivalent Lanthanide ions which induce microscopic phase behaviours (liquid–liquid phase separation (LLPS), re-entrant condensation) to the negatively charged globular protein BSA. As this phase behaviours are strongly dependent on salt (C_s) and protein (C_p) concentrations, we maintain C_s:C_p ratio 23.33 to minutely study the LLPS state of BSA in the prior two chapters. In the last one, we draw the overall phase diagram of protein for the Ln^{3+} ions by varying concentrations and analyze each phase separately. Our findings indicate that mono- and divalent salts do not induce LLPS at this specific ratio under ambient conditions. DLS studies reveal larger particles in the trivalent ion-protein solution. Zeta-potential measurements show an inversion in protein surface charge, while CD measurements confirm that the protein's structural integrity is retained in the LLPS state. Our THz hydration experiments claim that mere charge (III in this case) is insufficient to describe the formation of LLPS by these ions, instead, it is essential to also consider the contribution from ion (hydrophilic)–hydration interactions. Lanthanides also exhibit distinct signature at temperature above the melting temperature (T_M) of BSA. At $T>T_M$ and at C_s:C_p 23.33, where denatured BSA typically undergoes gelation with mono- and divalent ions, Ln^{3+} ions retain the LLPS state. UV-turbidity, DLS and surface potential measurements corroborate well with our visual detection. DIC microscopy images and fluorescence anisotropy also provide clear picture of the micro-environments. Finally, we determine the explicit hydration at the protein interface using THz spectroscopy and found an ion specific effect on both the gel and LLPS phase behavior. As an extension of the previous works, we deal with five lanthanide ions (La^{3+}, Nd^{3+}, Gd^{3+}, Ho^{3+} and Lu^{3+}). The detailed study of lanthanide-induced protein phases (protein homogeneous phase → turbid phase (metastable) → re-entrant to the homogeneous phase) are both concentration-dependent and ion-specific. Our findings reveal that these microscopic phase behaviors of proteins are driven not only by changes in the protein surface charge potential but also by ion-protein

hydration. Additionally, the contrasting ion hydration distinguishes the metastable phase from others at a particular protein concentration.

On the other part of the thesis, we concern over the unique nature of guanidinium (Gdm) ions on protein (de)stability, which varies depending on the counter ions. First, we address the exact reason of the contradict behaviours of GdmCl and Gdm_2SO_4 by combined THz spectroscopic and MD-simulation investigation using a model amide NMA. The simulation results confirm a much stronger interaction between Gdm^+ and X^- in the case of Gdm_2SO_4 compared to that in GdmCl in both aqueous and NMA solution. Snapshots from the simulations suggest the possibility of contact and solvent-shared ion pairing between Gdm^+ and SO_4^{2-}. Our THz results conclude that the hetero ion-pairing propensity in Gdm_2SO_4 prevents free Gdm^+ ions from interacting with NMA, whereas in GdmCl, the absence of such ion-pairing allows abundant interaction with NMA. Second, we monitor the thiocyanate-assisted fibrilization of some model proteins (BSA, Ova and Lys). The fibrillar morphology depends on the counterion(s) (Na^+ and guanidinium (Gdm^+)), specifically in varied concentration and pH range as evidenced by field emission scanning electron microscopy (FESEM) and atomic force microscopy (AFM). We underline the changes in protein hydration during the fibrillar processes using THz spectroscopy technique. We emphasize the mechanism of fibrillar arrangements and conclude that water behaves differently (either as a *"lubricant"* or as a *"de-wetting agent"*) in the interface of the mature fibrils depending on the ion-protein binding interactions.

In brief the key findings of this thesis are:

- Trivalent ions, which form LLPS, cause significant changes in overall protein hydration. The extent of these changes is specific to each ion, highlighting the important role of hydrophilic (electrostatic) hydration in the observed LLPS process.
- Trivalent Lanthanide salts show unusual thermo-resistive behavior of protein as the room temperature formed LLPS retains their transparent and fluid like appearance at $T>T_M$.
- Lanthanide-induced protein phases (protein homogeneous phase → turbid phase (metastable) → re-entrant to the homogeneous phase) are both concentration-dependent and ion-specific. Ion specificity are not charge dependent rather the phase are driven by protein-ion hydration.

- Contrasting ion-pairing propensity of the Gdm^+ ion with the counterion (Cl^- or SO_4^-) decides its protein denaturation ability.
- SCN^- instigates instantaneous protein aggregation and amyloid fibril formation at pH 2 environment. Counter-cation (Na^+ and Gdm^+) that leads in shaping the final arrangements of the mature fibrils. Our results strongly suggest that *elevated level of SCN^- ion in human body could lead to severe medical conditions.*

8.II. Future Perspective.

The thesis deals with both microscopic (LLPS) and macroscopic (gel, amyloid fibril) protein aggregations which are the cause of different neurodegenerative diseases.[1, 2] As the studies confirm, lanthanides are responsible for LLPS of model protein and such phase separation strongly depends on protein and ion hydration, one could consider the facts to deal with lanthanide toxicity. We believe our solvation studies of LLPS will help to attempt the study how toxic the lanthanides are with intrinsically disordered proteins (e.g. α-synuclein, tau, prion, amyloid β protein etc.) which are directly related to the chronic neurodegenerative diseases like Parkinson disease, Alzheimer disease, Creutzfeldt-Jakob's disease etc.[3-5] One should also learn a lesson that the excessive use of these rare earth elements in modern technologies will be an upcoming threat. Detail spectroscopic analysis suggests SCN^- ion triggers amyloidosis. The amyloids are directly associated with various health conditions, including obesity, type 2 diabetes, neurological disorders, cancer, depression etc.[6] This makes one of different scientific fields show grave concern in investigating each minute factor responsible for amyloid fibrillation. As SCN^- ion is a naïve ion and we uptake SCN^- from food resources, our claim that elevated level of the ion in human body could lead to these severe medical conditions.

Our future plan will be to separate the lanthanides from metal mixtures with the help of the THz spectroscopy which will help in modern industries. We also aim to study the oligomeric state and LLPS state to know which one more prone to the neurodegenerative diseases.

8.III. References.

1. S. Ray, N. Singh, R. Kumar, K. Patel, S. Pandey, D. Datta, J. Mahato, R. Panigrahi, A. Navalkar, S. Mehra, L. Gadhe, D. Chatterjee, A. S. Sawner, S. Maiti, S. Bhatia, J. A. Gerez, A. Chowdhury, A. Kumar, R. Padinhateeri, R. Riek, G. Krishnamoorthy and S. K. Maji, *Nat. Chem.*, 2020, **12**, 705-716.

2. A. Siegert, M. Rankovic, F. Favretto, T. Ukmar-Godec, T. Strohäker, S. Becker and M. Zweckstetter, *Protein Sci.*, 2021, **30**, 1326-1336.
3. D. S. Knopman, H. Amieva, R. C. Petersen, G. Chételat, D. M. Holtzman, B. T. Hyman, R. A. Nixon and D. T. Jones, *Nat. Rev. Dis. Primers*, 2021, **7**, 33.
4. W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brundin, J. Volkmann, A.-E. Schrag and A. E. Lang, *Nat. Rev. Dis. Primers*, 2017, **3**, 17013.
5. I. Zerr, A. Ladogana, S. Mead, P. Hermann, G. Forloni and B. S. Appleby, *Nat. Rev. Dis. Primers*, 2024, **10**, 14.
6. R. Al Hamed, A. H. Bazarbachi, A. Bazarbachi, F. Malard, J.-L. Harousseau and M. Mohty, *Blood Cancer J.*, 2021, **11**, 97.