Conformational Fingerprinting of Phase Separation and Amyloid Formation Through the Lens of Vibrational Raman Spectroscopy

Anamika Avni

A thesis submitted for the partial fulfillment of the degree of

Doctor of Philosophy

Department of Chemical Sciences

Indian Institute of Science Education and Research Mohali

Knowledge City, Sector 81, SAS Nagar, Manauli PO, Mohali

140306, Punjab, India

February 2023

Dedicated to my

parents and teachers

“Research is to see what everybody else has seen, and to think what nobody else has thought”

Albert Szent-Gyorgyi

Declaration

The work presented in this thesis has been carried out by me under the guidance of Prof. Samrat Mukhopadhyay at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgment of collaborative research and discussions. This thesis is a bona fide record of original work done by me and all sources listed within have been detailed in the bibliography.

Anamika Avni

Date:

Place:

In my capacity as the supervisor of the candidate's thesis work, I certify that the above statements by the candidate are true to the best of my knowledge.

Prof. Samrat Mukhopadhyay

Date:

Place:

Acknowledgments

This Ph.D. thesis was only possible with the support and contributions of several people. I want to thank every person directly and indirectly involved in this challenging yet beautiful journey.

I owe immense gratitude to my Ph.D. advisor, Prof. Samrat Mukhopadhyay, for allowing me to be a part of this wonderful research group and work with him on some exciting research projects. His enthusiasm, inquisitiveness, and quest for knowledge motivated me each day to perform and extend my boundaries. He made me believe that with hard work and a ton of dedication, much is possible. One of his favorite quotes, *"Success is no accident,"* makes complete sense to me now after seeing him work passionately and tirelessly in this field. The sincerity and dedication with which he handles and executes his duties as a researcher and a faculty are just excellent. I sincerely appreciate the enormous time and effort he invested in me, which enabled me to push my limits, achieve, and learn all I could during my Ph.D. I am incredibly thankful to him for guiding and supporting me over these years. I am also very grateful to him for his scientific advice and many insightful discussions and suggestions. The stimulating conversations over coffee and otherwise also were enough to energize and motivate me on days that seemed dull and gloomy. His passion for learning and contributing to this field of science and various other areas is incredibly unique. I hope I can be as lively, enthusiastic, motivated, and energetic as him in all that I plan to do in my future.

I want to thank my thesis committee members, Dr. Raj Kumar Roy and Dr. Sabyasachi Rakshit, for their guidance and yearly evaluation of my progress. Their discussions and feedback have been invaluable.

I want to thank all my colleagues for their great collaborations. I owe special thanks to Ashish Joshi, Anuja Walimbe, Swastik G. Pattanashetty, Aishwarya Agarwal, and Sandeep K. Rai for their contributions to my thesis work. I thank Prof. Dorothee Dormann (Ludwig-Maximilians University & Institute für Molekulare Biology, Mainz), Prof. Witold Surewicz (Case Western Reserve University, Cleveland), and Prof. Vinod Subramaniam (University of Twente, Netherlands) for their kind gifts of DNA plasmids. I acknowledge IISER Mohali for providing financial support and the Department of Science and Technology (DST), Science and Engineering Research Board (SERB), for the international travel award that enabled me to attend an international scientific meeting in Switzerland.

I would especially like to thank Dr. Mily Bhattacharya for her enthusiasm and constant encouragement. Although I never had an opportunity to work with her in the lab, the random conversations I had with her always left me super energized. I will always look up to her cheerfulness, kindness, and never say die attitude. I am extremely grateful to her for making our manuscripts better and for all her valuable suggestions that immensely helped my research. Her presence has always kept me smiling and inspired.

I gratefully acknowledge the research facilities at IISER Mohali that helped me greatly in my research. I would like to thank Vidya bhaiya for taking care of the lab's cleanliness and making sure we performed our experiments without any problems. I would also like to thank the workers in the hostels, academic blocks, and the administrative building, as well as the library staff, for making this journey better.

I was more than blessed to be surrounded by some fantastic souls in the lab. There was so much to learn from the former and current lab members. Dr. Priyanka Madhu introduced me to the field of recombinant proteins and taught me the basics of protein expression and purification. Thereafter she has always been there for me and closely checked my progress in the lab. I am incredibly thankful to Dr. Swapnil Singh and Dr. Priyanka Dogra, who taught me the principles of protein vibrational Raman spectroscopy and explained in greater detail how to use the Raman instrument during the initial years of my Ph. D. I will always be grateful to them as Raman forms the backbone and foundation of this thesis. The sheer hard work, dedication, and passion for science shown by Dr. Priyanka Dogra have always inspired me, and I thoroughly enjoyed our conversations over coffee and food. I am thankful to her for being there and always helping me in all she could. I am more than grateful to Dr. Aishwarya Agarwal, Dr. Debapriya Das, and Dr. Anupa Majumdar for guiding and helping me in planning some of my experiments and being critical of my work when required. They were always excited to help me out in whichever way possible. It was great to have met Swastik in this lab; a great soul and such a dear friend. He never failed to motivate me and has been a great cheerleader throughout this journey. I will always look up to the positive, energetic, and happy vibrations that he radiates, which always energizes and refreshes me.

This lab has always been like a home away from home, and the people working here are more like family members. We have all shared our happiness, joy, sorrows, and laughter, which I will always cherish. I want to thank Ashish Joshi, Sandeep K. Rai, Lisha Arora, Anuja Walimbe, Anusha Sarbahi, Sayanta Mahapatra, Roopali Khanna, Snehasis Sarkar, Dr.

Dipankar Bhowmik, Harshita Sawdhekar, and Jyoti Swami from the bottom of my heart for making this place so comfortable and stress-free to work. I will never forget our gossip and laughter over our morning/evening tea sessions which always kept me entertained and joyful in the lab. All our lab outings, birthday photography sessions, and parties will always be close to my heart and treasured for life. It was a collaborative effort of all these wonderful people to keep the lab environment so lively, happy, and productive.

I cannot thank Ashish enough for always being my critique, a great collaborator, and a fantastic friend. His scientific aptitude and skill in simplifying complex-looking problems will always be looked up to. He taught me the power of a hot cup of tea in handling stressful situations in science and life in general. Anuja (my little floobie) has been a great work companion and an outstanding junior. I thoroughly enjoyed the scientific conversations and random chats with her, which kept me happy, driven, and motivated. I owe my heartfelt gratitude to Lisha for always being there for me in my moments of happiness and sorrow. Her chirpy nature kept the mood light, and the random deep-life talk sessions were extremely delightful. She has always been there for help in times of need. I am incredibly grateful to Sandeep for always motivating and pushing me to meet my deadlines. My girl gang, Anusha and Roopali, have kept the atmosphere so positive, lively, and joyous with their contagious laughter and friendly nature. I deeply admire the never say die attitude of Sayanta and the pinch of witty humor he added to this lab. I sincerely acknowledge the help Dr. Dipankar extended during this thesis's writing process and in the lab in general. I thank Snehasis, Harshita, and Jyoti for being such amazing juniors in this lab. I greatly acknowledge the help and the role played by Harshita during her stay in this lab as my first summer student during her internship.

This journey was even more enjoyable with some amazing friends around. I am incredibly grateful to have Dr. Labhini, Lona, Dr. Atanu, and Dr. Suman as my closest friends. I will always cherish the moments we spent together exploring fancy hotels and restaurants for some amazing food. I owe a special acknowledgment to my girls, Labhini and Lona, for being there whenever I needed anything. They were always there for me and gave me the best advice whenever I was in doubt. I will always treasure our beautiful friendship, and the moments we spent together. I would like to especially thank Atanu for being a great philosopher friend in my life. He has always been there to help me out, gave me valuable suggestions and advice, and often helped me to make the right decision.

I want to thank my family for their unconditional love and understanding. My parents, who are my first teachers, have always been my role model and encouraged me to pursue research. I am highly thankful to my mother for investing ample time and energy to give me the best of everything. The values, lessons, and training she provided can never be forgotten and will always be of great help in all my future endeavors. My father has always been my most significant cheerleader and taught me the power of knowledge. His sincerity, dedication, and passion for learning have always motivated me to improve. My elder brother has always been my guiding light, who was just a phone call away and eager to help me when my mind was full of doubts and insecurities. His carefree and positive outlook has always eased me and filled me with great energy to perform better.

Finally, I would like to thank all my teachers who have taught me all my life and have been instrumental in shaping my research career. It is rightly said that a teacher takes a hand, opens a mind, and touches a heart. I am grateful to have been mentored by some great teachers who profoundly impacted my life. In the words of a wise man, “true teachers use themselves as bridges over which they invite their students to cross, and having facilitated their crossing joyfully collapse, encouraging them to create bridges of their own.” I will forever be grateful to my parents and all my teachers and dedicate this thesis to them.

Anamika Avni

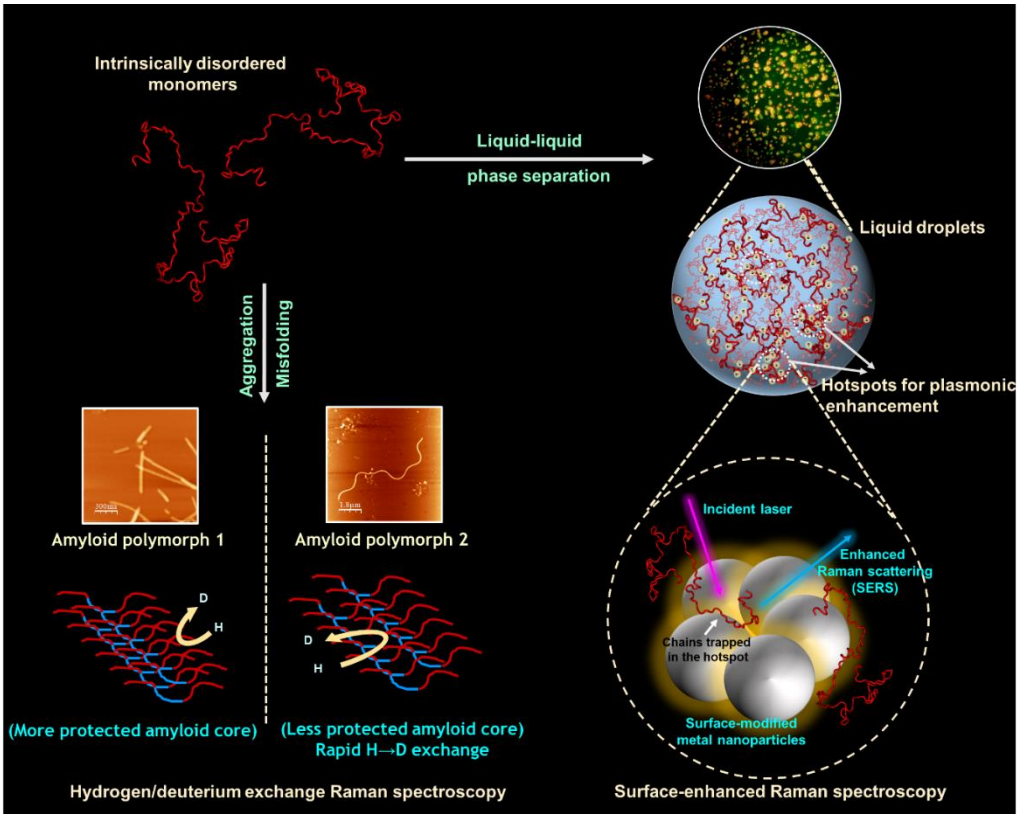
Thesis Synopsis

Conformational Fingerprinting of Phase Separation and Amyloid Formation Through the Lens of Vibrational Raman Spectroscopy

Chapter 1. Introduction

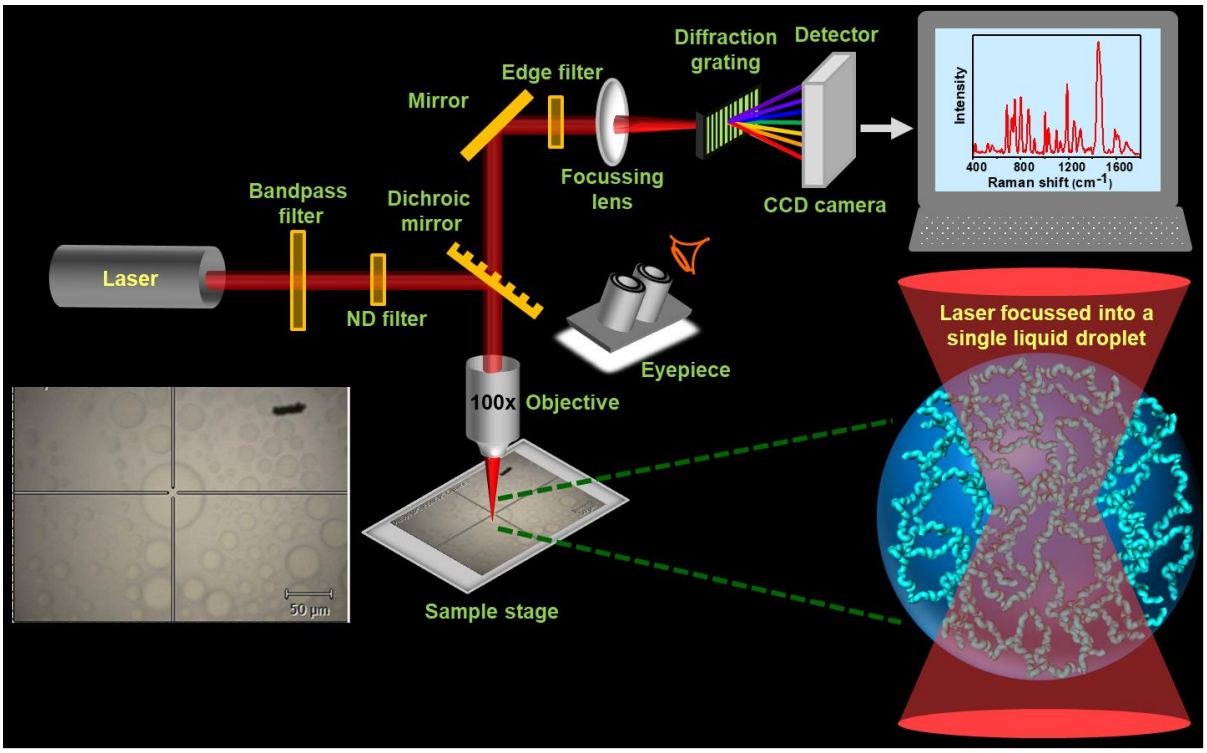
Proteins are complex, versatile macromolecules in the living cell that can perform critical roles in various biological processes, including catalyzing biochemical reactions, immune protection, trafficking of molecules to specific cellular locations, regulation of cellular growth and differentiation, and so on. According to the conventional sequence-structure-function paradigm, the critical physiological function of a protein is governed by its unique, well-defined, 3-dimensional structure encoded by the amino acid sequence. A particular class of proteins confronts this traditional one structure-one function paradigm and fails to fold into an ordered conformational state under physiological conditions autonomously. This emerging class of proteins is known as natively unfolded or intrinsically disordered proteins (IDPs). IDPs have been shown to play essential roles in signaling, cell division, intracellular transport, cell cycle control, transcriptional and translational regulations, and so forth. They either undergo disorder-to-order transitions or form "fuzzy complexes" upon interaction with suitable binding partners to perform critical physiological functions. Misfolding of IDPs allows the polypeptide chains to form partially folded states/folding intermediates, thereby leading to extra/intracellular accumulation and deposition of misfolded proteins. This condition leads to the formation of either amorphous aggregates or amyloids and has been linked to various biological dysfunctions such as Alzheimer's, Parkinson's, Huntington's disease, and so forth. The highly ordered amyloid assemblies share a common core architecture exhibiting a structural diversity in their supramolecular packing arrangement within the amyloid backbone. Such an altered packing results in amyloid polymorphism that is often responsible for distinct amyloid strains. We showed that vibrational Raman spectroscopy coupled with H/D exchange could discern the conformational differences within highly ordered cross-β-core of amyloid polymorphs exhibiting noticeable differences in the hydrogen bonding strength and supramolecular packing within the β-rich backbone and side-chain orientation. Furthermore, increasing evidence suggests that cells can perform various critical functions employing (non-canonical) membrane-less organelles formed via intracellular phase separation of IDPs and/or nucleic acids. These mesoscopic liquid droplets are known to concentrate biopolymers and are involved in critical cellular functions and pathology. In this thesis, efforts were directed toward

understanding key molecular interactions within the condensed phase that is crucial to discern the mechanism of condensate assembly and regulation of biomolecular condensates. Their maturation via liquid-to-solid phase transitions predisposes the protein towards forming aggregates, thereby causing debilitating neurodegenerative diseases. We used single-droplet vibrational Raman spectroscopy, a potent tool that can provide enormous information about intramolecular dynamics and intermolecular interactions within the protein-rich dense phase, thereby aiding in regulating droplet properties. Our single-droplet surface-enhanced Raman scattering (SERS) methodology utilizes surface-engineered, plasmonic, metal nanoparticles to illuminate the inner workings of mesoscopic liquid droplets and captures the complex conformational characteristics governing liquid-liquid phase separation of IDPs. We believe such measurements will pave the way for ultra-sensitive detection, characterization, and quantification of a wide range of biomolecular condensates and amyloids involved in physiology and disease.



Chapter 2. Probing the conformational heterogeneity and structural distribution within biomolecular condensates using single-droplet vibrational Raman spectroscopy

IDPs lack the ability to undergo autonomous folding under native conditions and exist as a dynamic ensemble of rapidly fluctuating interconverting structures. This special class of proteins has an intrinsic preference for conformational disorder and is often characterized by low complexity (LC) domains. A growing body of current research reveals that IDPs/IDRs containing LC regions and prion-like domains that offer conformational heterogeneity and multivalency are excellent candidates for biomolecular condensation to form highly dynamic, liquid-like, mesoscopic liquid droplets that are involved in a range of critical cellular functions and debilitating human pathologies. Understanding the role of intrinsic disorder and



Agarwal, A.; Rai, S. K.; Avni, A.; Mukhopadhyay, S.
Proc. Natl. Acad. Sci. U. S. A. (2021).

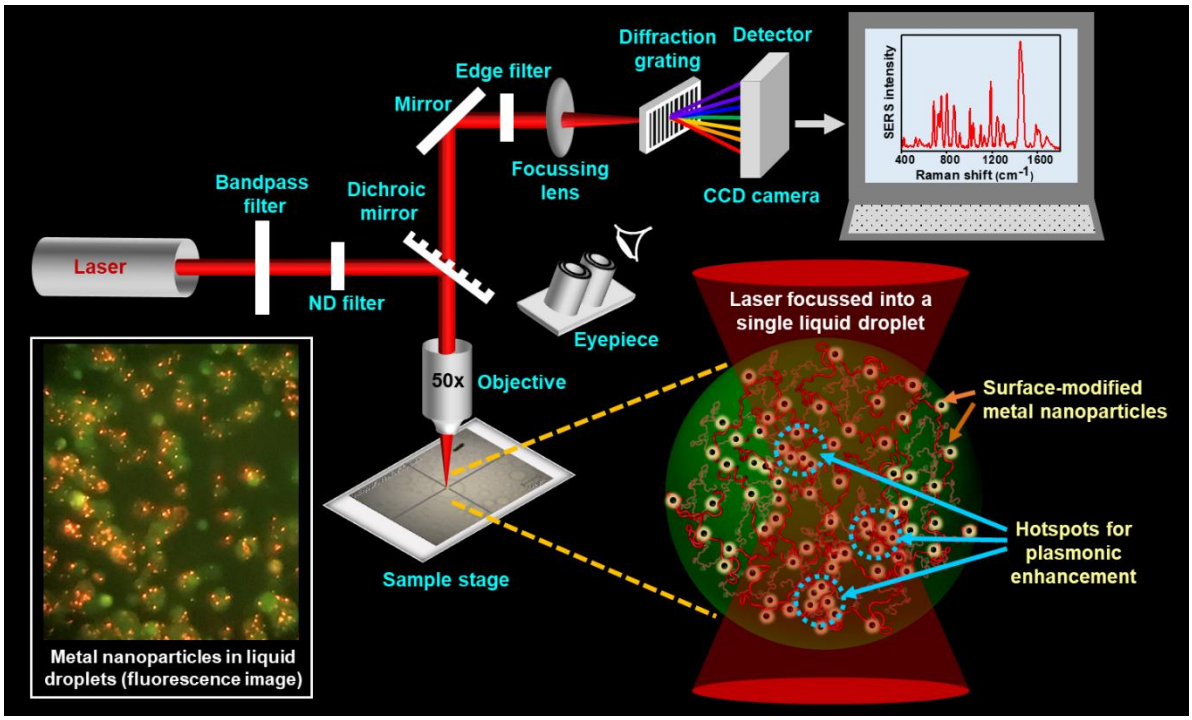
Avni. A.; Joshi. A.; Walimbe. A.; Pattanashetty. S. G.; Mukhopadhyay. S.
Nat. Commun. (2022).

conformational heterogeneity of IDPs/IDRs in the formation of liquid-like droplets is crucial to gain insights into the molecular mechanisms of phase separation. Further elucidating the key molecular drivers of phase separation is of immense importance in order to regulate and modulate the condensate properties. Even though there exists a host of microscopic and

spectroscopic tools to study intracellular phase separation, most of these techniques fail to provide information regarding the conformational heterogeneity and distribution within individual droplets. In this direction, we adapted single-droplet vibrational Raman spectroscopy, which is a powerful tool to capture the complex conformational characteristics governing phase transitions of IDPs. Focusing the laser beam within each protein droplet provides a wealth of molecular information within a single droplet by capturing Raman scattering bands due to different vibrational modes from the mesoscopic protein-rich droplets. Our single-droplet vibrational Raman experiments captured conformational landscape and intrinsic disorder within liquid droplets of two well-known phase-separating systems, PrP-Y145Stop and full-length FUS (Fused in Sarcoma). Further, our results revealed the involvement of several aromatic amino acid residues in the chain collapse and condensation of polypeptide chains upon liquid phase transitions. We envision that such single-droplet measurements can be utilized to characterize the protein-rich droplet's internal environment, further allowing us to modulate the physiological and pathological roles of the biomolecular condensates.

Chapter 3. Single-droplet surface-enhanced Raman scattering captures the conformational landscape within phase-separated liquid condensates of an RNA-binding protein

Biomolecular condensates formed via liquid-liquid phase separation (LLPS) are involved in a myriad of critical cellular functions and debilitating neurodegenerative diseases. We have shown in the previous chapter the utility of single-droplet vibrational Raman spectroscopy in illuminating some key structural details within phase-separated condensates. However, recording vibrational signatures from proteins in aqueous solutions at physiological concentrations is highly challenging due to the low Raman scattering cross-section of proteins. High concentrations and laser power used for such measurements can be detrimental to soft biological samples and lead to laser-induced damages and other artifacts. In this direction, we developed and adapted a novel single-droplet surface-enhanced Raman scattering (SERS) technique using plasmonic silver nanoparticles to unveil the inner workings of the protein droplets and capture the conformational heterogeneity and distribution in a droplet-by-droplet manner. Such measurements allowed us to zoom into the mesoscopic condensed phase to unmask the molecular determinants governing the intriguing condensate biophysics. These



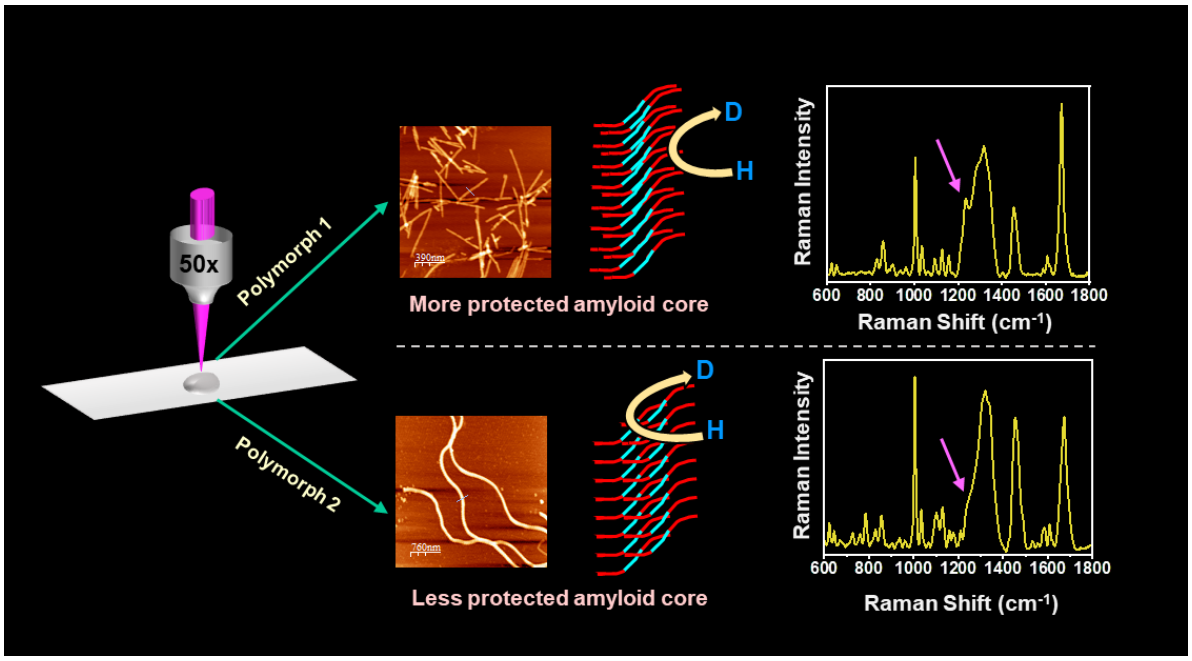
Avni. A.; Joshi. A.; Walimbe. A.; Pattanashetty. S. G.; Mukhopadhyay. S.
Nat. Commun. (2022)

highly sensitive measurements offer unprecedented sensitivity to capture the crucial interactions, conformational heterogeneity, and structural distributions within the condensed phase of a well-known RNA-binding protein, FUS, in a single droplet fashion. Our findings showed that the ordered C-terminal RNA-binding domain of FUS undergoes partial unwinding in the presence of RNA, which increases the polypeptide chain disorder that can promote both homotypic (FUS-FUS) and heterotypic (FUS-RNA) interactions within the condensed phase. Such an ultra-sensitive single-droplet vibrational methodology can serve as a potent tool to decipher the key molecular drivers of biological phase transitions of a wide range of biomolecular condensates involved in physiology and disease by using different surface functionalities and other metals as SERS substrates for enhancements.

Chapter 4. Hydrogen-deuterium exchange vibrational Raman spectroscopy distinguishes distinct amyloid polymorphs comprising altered core architecture

Amyloids are cross-β-sheet rich proteinaceous fibrillar aggregates and are associated with a wide array of deadly human diseases as well as with a range of physiological functions from bacteria to humans. These assemblies share a common amyloid core architecture yet vary in

their supramolecular packing arrangements within the core, exhibiting molecular-level polymorphism that is often responsible for strain-specific disease phenotypes. Elucidating the structures of cross-β-sheet rich amyloids is of immense importance to discern the mechanism of amyloid polymorphism and the structure-pathology relationship. Here, we utilize vibrational Raman spectroscopy coupled with hydrogen/deuterium exchange to gain structural insights from diverse amyloid polymorphs of a presynaptic neuronal protein, α-synuclein, aggregation of which is associated with numerous human pathologies termed synucleinopathies. By directly monitoring and analyzing the key vibrational Raman bands for both protein backbone and sidechains, we were able to capture the conformational heterogeneity and structural distribution within the three distinct amyloid polymorphs. Hydrogen-deuterium exchange Raman



Avni. A.; Joshi. A.; Mukhopadhyay. S.
(Manuscript submitted)

spectroscopy allowed us to structurally distinguish these distinct amyloid polymorphs displaying altered hydrogen bonding ability and supramolecular packing within the cross-β structural motif. Incorporating a heavy atom such as deuterium causes the replacement of the easily accessible protons in the disordered fuzzy coat and loosely packed ordered segments of the protein aggregates and help us to fathom the hydrogen-bonding patterns and solvent accessibility within the extensively H-bonded amyloid core. Our results delineate the key molecular factors governing the structural diversity in amyloids and will pave the way for studying amyloid remodeling by small molecules. Hydrogen-deuterium exchange Raman

spectroscopy will serve as a potent, noninvasive, and label-free tool to explore a wide range of *in vitro* and *ex vivo* generated amyloid polymorphs displaying strain diversity associated with physiology and disease.

Chapter 5. Conclusions and future directions

The work described in this thesis elucidates the utility of vibrational Raman spectroscopy as a potent tool to capture the fundamental interactions, conformational heterogeneity, and structural distribution in a range of protein assemblies from disordered monomers to mesoscopic liquid droplets to amyloid fibrils. This chapter aims to provide a broad overview and discuss some exciting future questions from the results described in the previous chapters. The principal findings of the present work include: (i) probing the conformational heterogeneity within biomolecular condensates using single-droplet vibrational Raman spectroscopy (Chapter 2), (ii) developing single-droplet surface-enhanced Raman spectroscopy method to decipher the key molecular drivers of biological phase transitions (Chapter 3) and (iii) utilizing hydrogen/deuterium exchange vibrational Raman spectroscopy to distinguish distinct amyloid polymorphs (Chapter 4). The work described in this thesis can broaden our understanding of vibrational Raman spectroscopy as a method to study a diverse range of protein assemblies associated with physiology and diseases.

******

List of Publications

- **A. Avni,** A. Joshi, and S. Mukhopadhyay "Hydrogen-Deuterium Exchange Vibrational Raman Spectroscopy Distinguishes Distinct Amyloid Polymorphs Comprising Altered Core Architecture" *(manuscript submitted)*.
- S. Mahapatra, A. Sarbahi, N. Punia, A. Joshi, **A. Avni**, A. Walimbe, and S. Mukhopadhyay (2022) "ATP modulates self-perpetuating conformational conversion generating structurally distinct yeast prion amyloids that limit autocatalytic amplification" *J. Biol. Chem.* 299(5), 104654.
- S. K. Rai, R. Khanna, **A. Avni,** and S. Mukhopadhyay (2022) "Heterotypic Electrostatic Interactions Control Complex Phase Separation of Tau and Prion into Multiphasic Condensates and Co-Aggregates" *Proc. Natl. Acad. Sci. U. S. A.* 118(45), e2100968118.
- **A. Avni**, A. Joshi, A. Walimbe, S. G. Pattanashetty, and S. Mukhopadhyay (2022) "Single-Droplet Surface-Enhanced Raman Scattering Decodes the Molecular Determinants of Liquid-Liquid Phase Separation" *Nat. Commun.* 13, 4378.
- A. Agarwal, L. Arora, S. K. Rai, **A. Avni,** and S. Mukhopadhyay (2022) "Spatiotemporal Modulations in Heterotypic Condensates of Prion and α-Synuclein Control Phase Transitions and Amyloid Conversion" *Nat. Commun.* 13, 1154.
- A. Agarwal, S. K. Rai, **A. Avni,** and S. Mukhopadhyay (2021) "An Intrinsically Disordered Pathological Prion Variant Y145Stop Converts into Self-Seeding Amyloids via Liquid-Liquid Phase Separation" *Proc. Natl. Acad. Sci. U. S. A.* 118(45), e2100968118.
- S. Singh, A. Agarwal, **A. Avni,** and S. Mukhopadhyay (2021) "Ultrasensitive Characterization of the Prion Protein by Surface-Enhanced Raman Scattering: Selective Enhancement via Electrostatic Tethering of the Intrinsically Disordered Domain with Functionalized Silver Nanoparticles" *J. Phys. Chem. Lett.* 12, 3187-3194.
- A. Majumdar, D. Das, P. Madhu, **A. Avni,** and S. Mukhopadhyay (2020) "Excitation Energy Migration Unveils Fuzzy Interfaces within the Amyloid Architecture" *Biophys. J.* 118, 2621-2626.
- **A. Avni**, H. M. Swasthi, A. Majumdar, and S. Mukhopadhyay, "Intrinsically Disordered Proteins in the Formation of Functional Amyloids from Bacteria to Humans" in "***Progress in Molecular Biology and Translational Science***" **2019**, Volume 166, Pages 109-143. (Publisher: Elsevier; Editor: Vladimir Uversky, University of South Florida).

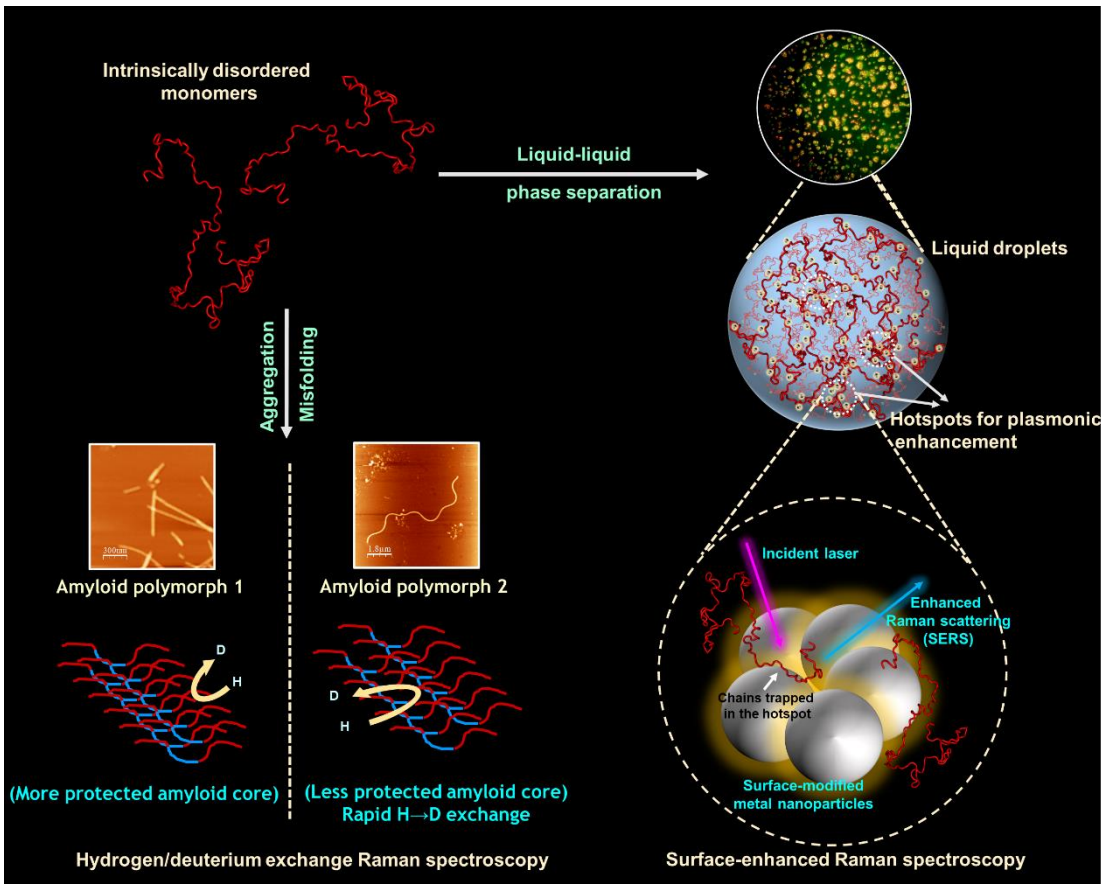
Presentations at Conferences

- **Anamika Avni,** Aishwarya Agarwal, Sandeep K. Rai, Ashish Joshi, Anuja Walimbe, Swastik G. Pattanashetty and Samrat Mukhopadhyay "Single-Droplet Vibrational Raman Spectroscopy Illuminates the Inner Workings of Phase-Separated Biomolecular Condensates". Presented poster at Gordon Research Conference on Intrinsically Disordered Proteins held in Les Diablerets, Switzerland (June 2022).

- **Anamika Avni,** Aishwarya Agarwal, Sandeep K. Rai, Ashish Joshi, Anuja Walimbe, Swastik G. Pattanashetty and Samrat Mukhopadhyay "Single-Droplet Vibrational Raman Spectroscopy Illuminates the Inner Workings of Phase-Separated Biomolecular Condensates". Presented poster at the 29^{th} CRSI-National Symposium in Chemistry and CRSI-ACS Symposium Series in Chemistry at IISER Mohali, India (July 2022).

Table of Contents

Introduction



1.1 Intrinsically disordered proteins

1.1.1 Introduction to intrinsic disorder in proteins

Proteins are complex, versatile macromolecules in the living cell that can perform critical roles in various biological processes, including catalyzing biochemical reactions, immune protection, trafficking of molecules to specific cellular locations, regulation of cellular growth and differentiation, and so on. According to the conventional structure-function paradigm, the critical physiological function of a protein is governed by its unique, well-defined, 3-dimensional structure encoded by the amino acid sequence. Intrinsically disordered proteins or natively unfolded proteins, popularly abbreviated as IDPs, are a class of proteins that confronts the classical *sequence-structure-function paradigm* according to which a protein folds into a stable 3-dimensional structure determined by the amino acid sequence, which is responsible for the unique biological functions of a given protein.[1-7] There also exists folded proteins containing disordered regions/domains referred to as intrinsically disordered protein regions (IDPRs). The existence of this class of proteins exhibiting numerous crucial biological functions despite having a stable folded structure dates back to the 1970s and late 1980s but was primarily unnoticed.[8-14] They are known to be abundantly present in the entire eukaryotic proteome, where nearly 25-30% of the eukaryotic genome codes for IDPs or IDRs.[15-16] IDPs are highly dynamic and are known to exist as heterogeneous ensembles of rapidly fluctuating conformations of approximately equal energies.[15] The folding free energy landscape of IDPs has a relatively flat free-energy surface with multiple local energy minima separated by small energy barriers in contrast to the "funnel-shaped" global energy minimum for folded/globular proteins, which corresponds to the thermodynamically favorable native state (Figure 1.1).[16,17,18] This corresponds to the on-pathway folding of polypeptide chains. IDPs undergo disorder-to-order transitions by interacting with specific binding partners to form folding intermediates/partially folded states funneling down to the native state via intramolecular contacts, which offers functional capabilities to the proteins.[19] As the protein folds into the native lowest energy state, several kinetically trapped off-pathway intermediate (metastable states) get populated along the sides of the funnel, which triggers the formation of amorphous aggregates or pathological amyloid fibrils via intermolecular interactions.[20] Often, there exists a competition between the on-pathway and off-pathway folding processes.[21,22] The cellular systems comprise molecular chaperones that assist in the folding of the unstructured polypeptide chains into the native state by lowering the free energy barrier. During the off-pathway folding process, chaperones prevent the aberrant intermolecular interactions between

the misfolded forms and promote folding into the native conformations.[22,23] IDPs are known to perform numerous biological functions under physiological conditions such as cellular signaling, cell division, cell cycle control, regulation of transcription and translation, protein phosphorylation, allosteric interactions, and so forth.[6,24]

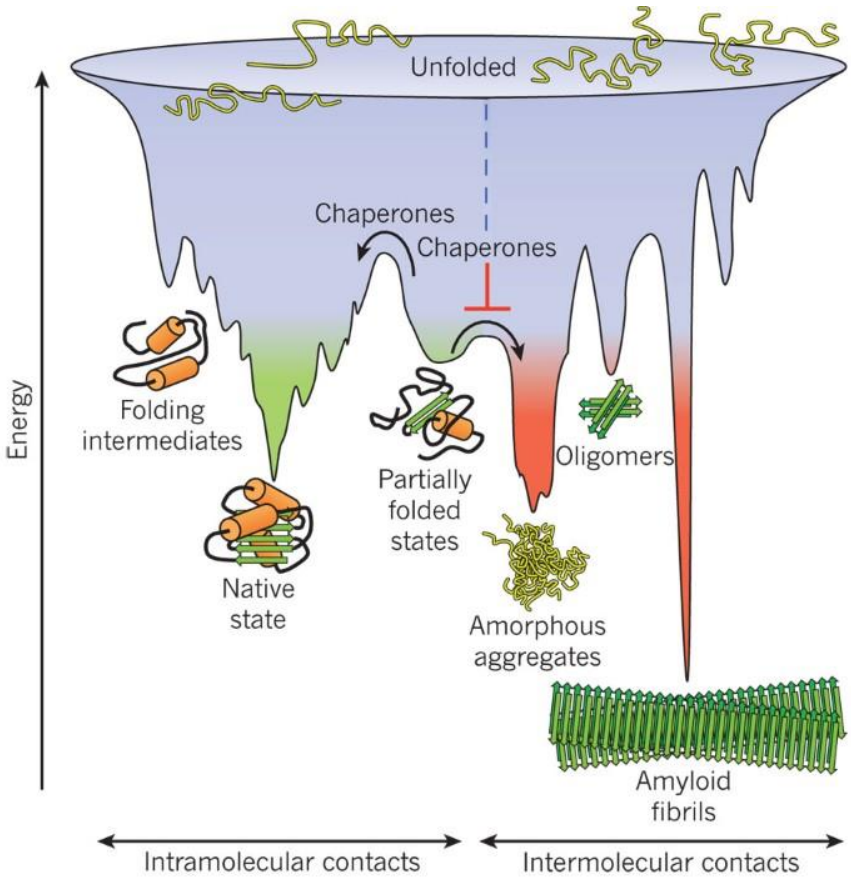


Figure 1.1: A schematic representation of the energy landscape of protein folding (purple) and aggregation (pink) illustrating the formation of native globular and intrinsically disordered proteins, oligomers, amyloid fibrils, and amorphous aggregates. The unfolded polypeptide chains funnel down to the thermodynamically favorable native state via intramolecular interactions by crossing substantial kinetic energy barriers. In the case of off-pathway folding, amyloid fibrils represent the thermodynamic ground state. Off-pathway intermediates are prone to form various forms of aggregates via intermolecular interactions. Reproduced with permission from reference (23).

The folding ability of the protein is encoded in its amino acid sequence.[26] According to the protein quartet model, IDPs can adopt four different conformational states, ordered, molten globule, pre-molten globule, and random-coil, each of which is involved in performing crucial functions.[25,27,28] Comparison of the mean net charge and mean hydrophobicity for a set of folded and natively unstructured proteins revealed that IDPs are typically characterized by low

overall hydrophobicity and high net charge, unlike the structured proteins (Figure 1.2).[29] IDPs are enriched in repeats of disorder-promoting residues such as Arg, Pro, Gln, Gly, Glu, Ser, Ala, Lys and are devoid of order-promoting residues like Trp, Cys, Tyr, Ile, Phe, Val, Asn, and Leu.[29,30] The conformational preferences of IDPs depend on the amino acid compositions and can be modulated by the fraction of charged and polar residues.[25,31] Two parameters, the fraction of charged residues (FCR) and the net charge per residue (NCPR), defined as the sum and difference of a fraction of positively charged (f+) and negatively charged residues (f-), respectively, play crucial roles in predicting the conformational properties of IDPs.[31,32,33] Depending on this, IDPs can be classified as polyelectrolyte and polyampholyte adopting expanded coil-like conformations in water because of high content of charged residues. On the contrary, IDPs rich in polar but uncharged residues adopt collapsed globule structures in the aqueous medium.[34] Therefore, deciphering the sequence-conformation relationship of IDPs can be extremely useful in modulating the properties of IDPs/IDRs and evaluating their influence on the functions.

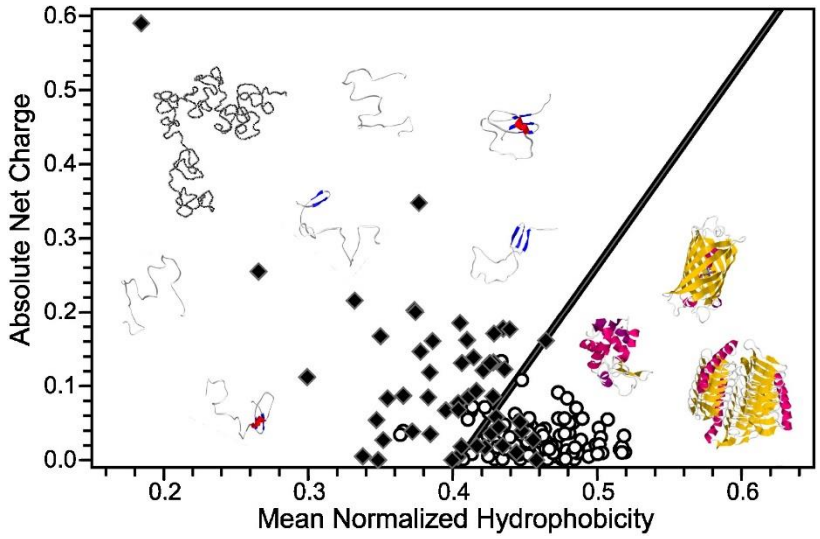


Figure 1.2: Comparison of the mean net charge and the mean hydrophobicity for a set of 275 folded (open circles) and 91 natively unfolded proteins (black diamonds). The solid line represents the border and shows that the extended IDPs (native coils or native pre-molten globules) occupy a unique space well separated from compact globular proteins. Data are shown from some highly disordered proteins possessing little or no tertiary and secondary structures and ordered globular proteins. Reproduced with permission from reference (29).

1.1.2 Physiological and pathological roles of intrinsically disordered proteins

IDPs are known to occupy almost half of the entire human proteome. This special class of proteins has been an area of great interest among researchers despite defying the decades-old paradigm that 3D structure implies function. IDPs are known to interact with themselves or with other biomolecules, forming interaction networks that allow them to perform multiple functions.[5,6,25] The IDP kingdom is often called the "chaotic" system, falling in the region between order and complete randomness. The exception conformational plasticity exhibited by IDPs due to environmental perturbations allows them to adopt multiple structures.[25] This allows them to perform a multitude of specialized physiological functions and expand the functional repertoire of the proteins in an organism to complement the structured (ordered) proteins. This class of proteins can either undergo a disorder-to-order transition upon interacting with specific binding partners or remain disordered in the bound state to form so-called "fuzzy" complexes.[35,36] The fuzziness allows the proteins to participate in one-to-many and many-to-one interactions and are associated with key cellular processes such as allosteric regulation, posttranslational modifications, and alternative splicing. They are also involved in cell signaling, protein-protein interaction networks, and gene regulation networks.[37-39]

In addition to performing critical roles in various cellular processes, IDPs/IDRs are also linked to debilitating human diseases. In some cases, mutations within IDPs/IDRs cause the polypeptide chains to adopt abnormal, aggregation-prone conformations leading to the formation of insoluble structures termed amyloids exposing the hydrophobic domains in the solvent.[40,41] These filamentous amyloid aggregates accumulate in the brain and a variety of other tissues and are known to be the characteristic hallmark of several protein-misfolding diseases in humans, such as Alzheimer’s disease, Parkinson’s disease, Amyotrophic lateral sclerosis, and so on. However, not all amyloid deposits are pathological.[42,43] There exist "Functional amyloids" that are known to play beneficial roles in a diverse range of organisms from bacteria to humans, such as bacterial biofilm formation in *Escherichia coli*, long-term memory in *Drosophila*, regulation of oogenesis in *Xenopus laevis*, sperm selection, clearance, and antimicrobial activity in *Homo sapiens*.[44-47] Table 1.1 tabulates some of the pathological and functional amyloids. Over the last decade, a growing body of research has revealed that several IDPs with low-complexity sequences/domains (LCDs) undergo intercellular phase separation to form membrane-less organelles.[48-50] These supramolecular assemblies are known to perform critical functions in cells and undergo liquid-to-solid phase transitions to form

solid/gel-like aggregates and amyloid fibrils.[50-52] The phenomena of biomolecular condensation will be discussed in detail in the subsequent sections.

Table 1.1: Select examples of some pathological and functional amyloids

IDPs	**Amyloid related pathology/function**	**Affected tissue/organism**
Pathological amyloids		
α-synuclein	Parkinson's disease	Brain
Amyloid-β peptide	Alzheimer's disease	Brain
Prion protein or its fragments	Spongiform encephalopathies	Brain
Tau	Alzheimer's disease	Brain
Amylin	Type II diabetes	Pancreas
Calcitonin	Medullary carcinoma of the thyroid	Thyroid
Apolipoprotein A1 fragments ApoAI	Apolipoprotein A1 amyloidosis	Eyes
Functional amyloids		
Curli	Biofilm formation and host cell adhesion and invasion	*Escherichia coli* (Bacterium)
Sup35 (prion)	Regulation of stop-codon readthrough	*Saccharomyces cerevisiae* (yeast)
Cytoplasmic polyadenylation element-binding protein (CPEB)	Long-term memory storage	*Aplysia and Drosophila*
Spidroin	Involved in the formation of silk fiber of the spider web	Araneus diadematus (Garden spider)
Xvelo	Regulation in oogenesis	*Xenopus laevis*

Semenogelin 1 and 2	Sperm selection, clearance, and antimicrobial activity	*Homo sapiens*
Pmel 17	Melanin biogenesis	*Homo sapiens*

1.2 Liquid-liquid phase separation (LLPS)

1.2.1 Liquid-liquid phase separation as an organizing principle in biology

Liquid-liquid phase separation is widely recognized as an emerging paradigm in the field of cell biology to address significant biological problems such as the mechanisms of gene regulation, organization of the cytoplasm and nucleoplasm, intracellular compartmentalization, and so forth.[48-50,53-56] Phase separation is defined as a process by which a homogeneous solution of macromolecules such as proteins and/or nucleic acids condense into a dense phase that is enriched with the macromolecular components and coexist with a dilute phase devoid of the same macromolecules.[57] The supramolecular assemblies of proteins and/or nucleic acids formed have been termed membrane-less organelles (MLOs) or biomolecular condensates, and examples include nucleolus, nuclear pores, paraspeckles, Cajal bodies, nuclear speckles, chromatin in the nucleus while stress granules, P-bodies, centrosome, neuronal RNA bodies in the cytoplasm (Figure 1.3).[58] These MLOs lack an enclosing membrane in contrast to the canonical vesicle-like organelles enclosed by membranes and are known to play essential roles in several key intracellular functions such as ribosome biogenesis, RNA transcription and processing, regulation of protein translation, and so forth.[50,59] Elucidating the molecular mechanisms for condensate formation *in vitro* and cells and their roles in physiology and pathology is currently under study. The concept of scaffolds and clients has been instrumental in understanding the compositional characteristics of biomolecular condensates.[60,61] Scaffolds can be linear multivalent proteins with distinct binding or folding domains, patchy colloids, or completely disordered multivalent proteins with distinct stickers interspersed by different spacer sequences. A sticker is a short segment or motif that enables specific homotypic or heterotypic non-covalent attractive interactions between protein molecules. In contrast, spacer molecules furnish the additional heterogeneous conformational flexibility and determine the material properties of the condensates.[62] While scaffolds are crucial for condensate formation having multiple valences and are usually present in higher concentrations inside the cell, clients are present in much lower concentrations but can modulate the phase separation propensity of the scaffolds and partition into the assemblies.[60,61]

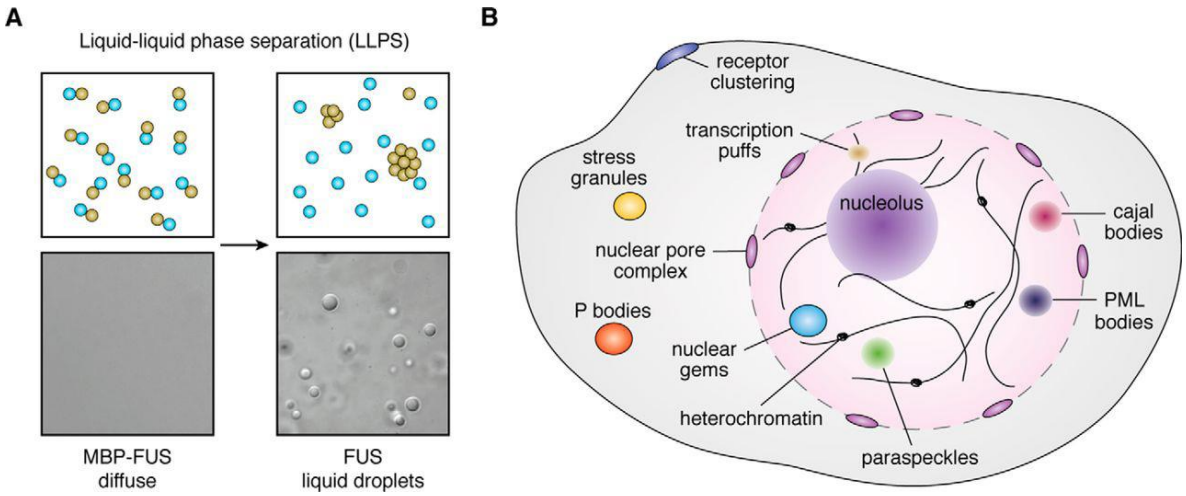


Figure 1.3: (A) In vitro recapitulation of liquid-liquid phase separation where a mixture of macromolecules like proteins and nucleic acids condenses into a dense phase that resembles liquid droplets. This phenomenon is similar to the droplets of oil in water. Green circles represent the RNA-binding protein FUS, and blue-colored circles represent the maltose-binding protein (MBP) tag. Liquid droplets enriched in FUS are formed upon cleavage of the MBP tag. (B) Various membrane-less organelles are found in the cells. The nucleolus, paraspeckles, Cajal bodies, nuclear speckles, and chromatin are present in the nucleus, while stress granules, P-bodies, centrosomes, and neuronal RNA bodies are in the cytoplasm. Reproduced with permission from reference (66).

1.2.2 Molecular drivers of intracellular phase separation

Intrinsically disordered proteins or regions (IDPs/IDRs) are well-known phase-separating systems undergoing biomolecular condensation due to the presence of fundamental units called stickers and spacers that imparts multivalency to the system and makes them the best candidates in the formation and functionalities of biomolecular condensates.[63] Unlike the multidomain proteins comprising several folded regions, IDPs/IDRs lack a well-defined 3-dimensional, folded structure and are described as ensembles of rapidly fluctuating structures with slightly different energies.[15] This protein class is characterized by prion-like low complexity domains (PrLCDs), backbone flexibility, high net charge, and lesser hydrophobic residues. The amino acid composition and the patterning of charged residues often determine the phase separation propensity of IDPs.[64,65] A symphony of weak and noncovalent sequence-specific interactions is known to drive the phase separation of different proteins, such as cation-π interactions, π-π interactions, electrostatic interactions, transient cross-β contacts, and intermolecular charge-transfer interactions (Figure 1.4).[66,67] Cation- π and π-π interactions are

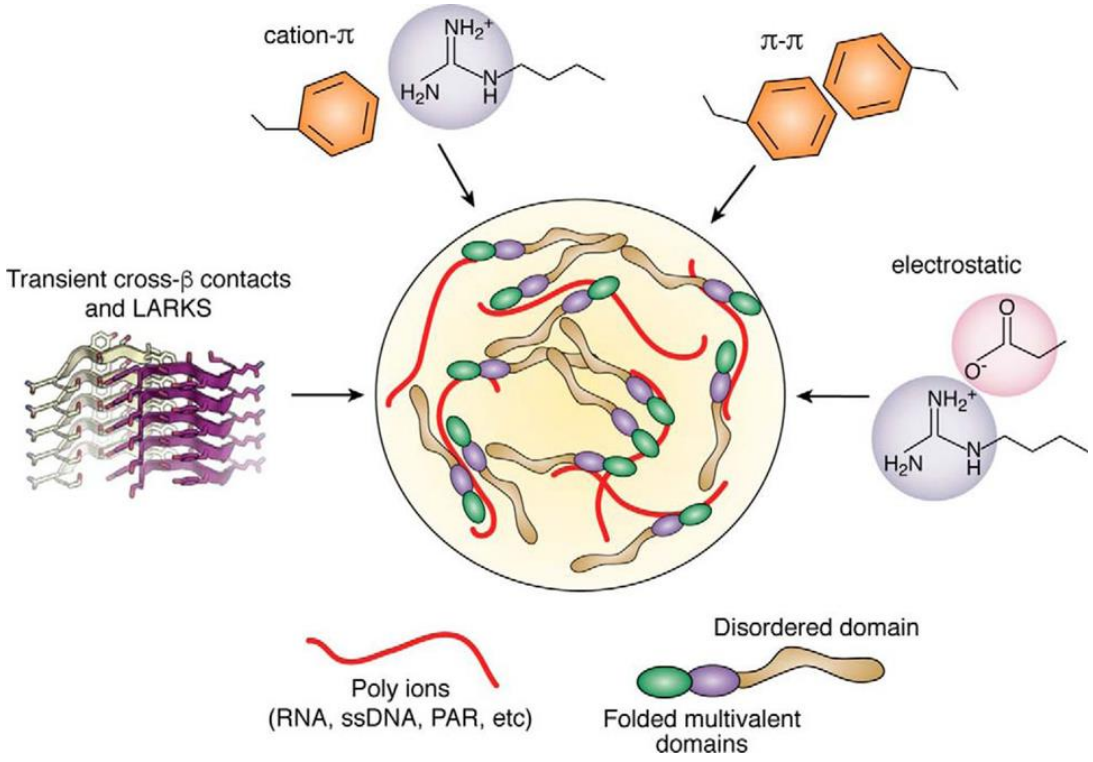


Figure 1.4: Critical interactions that drive LLPS of IDPs/IDRs, including cation-π interactions, π-π interactions, electrostatic interactions, hydrophobic interactions, transient cross-β contacts, and intermolecular charge-transfer interactions. Poly ions such as RNA can act as scaffolds or seeds for LLPS. Reproduced with permission from reference (66).

observed in IDPs possessing electron-rich aromatic amino acid motifs such as tyrosine, tryptophan, phenylalanine, and positively-charged amino acid residues like arginine and lysine. This interaction helps stabilize the condensates formed by proteins comprising low- complexity IDRs like FUS, TDP-43, hnRNPA1.[68,69] Electrostatic interaction or charge-charge interactions are primarily observed due to the presence of high net charge on the surface of IDPs that are known to undergo simple coacervation due the presence of positively and negatively charged residues within the same or different proteins.[70] Additionally, IDPs/IDRs can undergo complex coacervation in the presence of nucleic acids, DNA, and RNA, via charge neutralization.[66,70-72] Furthermore, IDPs have a reduced fraction of hydrophobic residues, contributing to their lack of a well-defined structure and folding ability. These residues, when present, help in the assembly and stabilization of biomolecular condensates. Hydrogen bonding is also considered a major driver of phase separation due to the presence of H-bond donors and acceptors in the proteins undergoing condensate formation.[73-75] Moreover, the liquid-like nature of the condensates formed by short fragments of PrLCDs of RNA-binding proteins is maintained by

the weak and transient cross-β contacts that mature to form cytoplasmic fibrillar aggregates having kinked β-sheets known as low-complexity aromatic-rich kinked segments (LARKS).[66,76] These aggregates are structurally different from typical amyloid fibrillar aggregates possessing strong and persistent cross-β sheet steric zippers.

1.2.3 Phase separation in physiology and disease

The mesoscopic liquid-like supramolecular assemblies can mature into viscoelastic gel-like and pathological solid-like aggregates. Aberrant liquid-to-solid and liquid-to-gel phase transitions are known to be involved in neurodegeneration, cancer, and other infectious diseases (Figure 1.5).[56,77-80] Genetic mutations, abnormal posttranslational modifications (PTMs), altered subcellular localization, and impaired protein quality control can alter the physical properties of condensates from liquid-like dynamic assemblies to a less dynamic state, causing condensate aging and dissolution.[53,81] Their liquid-to-solid phase transitions lead to the formation of protein-trapping cross-linked hydrogels and neurotoxic protein aggregates that disrupt RNP granule functions and impair new proteins synthesis in neuron terminals causing

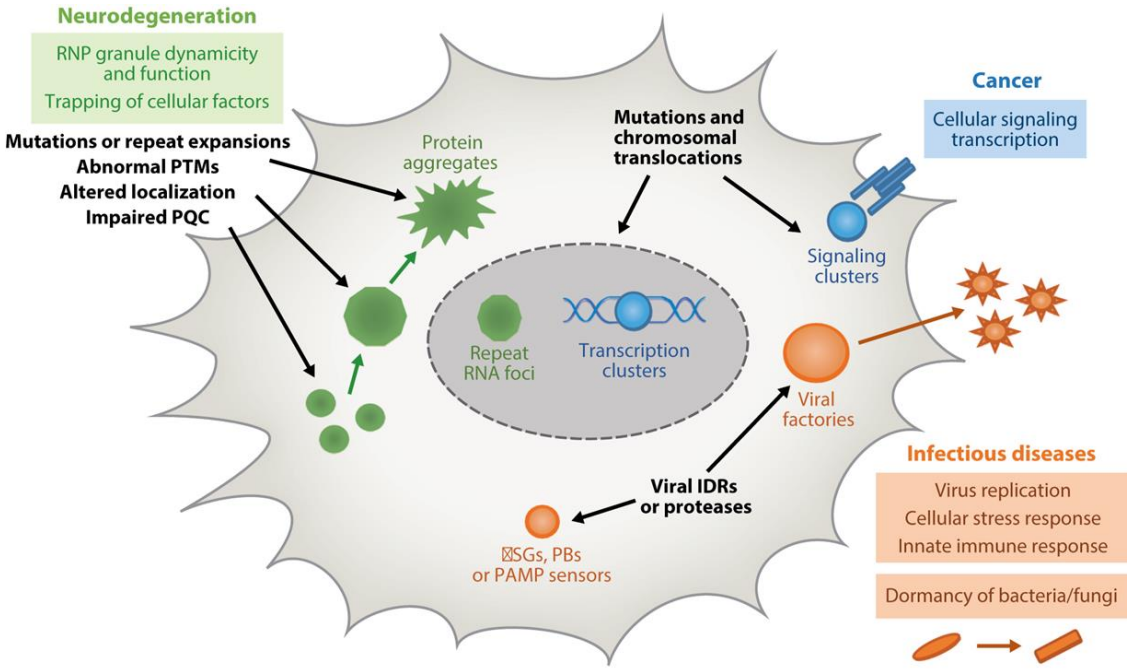


Figure 1.5: Phase separation under stressed conditions leads to aberrant condensate formation leading to neurodegenerative disorders, cancer, and infectious diseases. Assembly and dissolution of these condensates have been linked to aging and age-related disorders characterized by genomic instability, epigenetic alterations, mitochondrial dysfunction, loss of protein quality control, disrupted cell signaling, and DNA damage. Reproduced with permission from reference (81).

neurodegenerative diseases characterized by progressive loss of cognitive or motor functions as in Alzheimer's disease, frontotemporal dementia (FTD), amyotrophic lateral sclerosis (ALS), and so on.[82] In viral infections, viral factories, viroplasm, or viral replication centers formed by the condensation of viral IDPs can cause viral genome replication and gene expression besides altering the antiviral immune response.[83,84] Mutations in the protein quality control (PQC) systems or loss of proteostasis due to biological aging can lead to the accumulation of large amounts of misfolded and aggregation-prone proteins known to form aberrant condensates at physiological conditions.[53,85,86] Assembly and dissolution of these condensates have been linked to aging and age-related disorders characterized by genomic instability, epigenetic alterations, mitochondrial dysfunction, loss of protein quality control, disrupted cell signaling, and DNA damage.[53]

1.2.4 Methods to study phase-separated biomolecular condensates

Various techniques and methods have been developed to examine the composition, viscoelastic properties, structure, dynamics, and fundamental molecular interactions of biomolecular condensates formed via liquid-liquid phase separation (Figure 1.6).[87,88] Several microscopy tools have been routinely used to visualize the number, size, and morphology of the condensed phase. Bright-field contrast-based imaging and fluorescence microscopy can probe the effect of various conditions and factors on droplet formation. Multi-color labeling permits visualization of differently labeled species and provides quantitative information on the dynamics, concentration, and localization of fluorescently labeled macromolecules and morphological changes within membrane-less organelles.[87] Microrheology experiments by particle tracking probes such as a fluorescently labeled species or an inert commercially available fluorescently labeled nano/micro bead can directly quantify the viscosity and density within the phase-separated assemblies. Such measurements are useful in providing insights into the physicochemical properties determining the material state as well as the strength and nature of intermolecular interactions within these droplets.[89] However, these tools fail to provide molecular-level insights into the condensed phase. The incorporation of fluorescent probes can potentially alter the phase separation mechanisms in vitro and in live cells. Structural tools for characterizing the molecular networks within membrane-less organelles, such as nuclear magnetic resonance (NMR), X-ray crystallography, and cryo-EM, can furnish atomic-level 3D structure and protein dynamics in solutions.[68,90,91] These tools pose significant challenges due to the lack of ability to form stable secondary and tertiary structures by the proteins undergoing

phase separation owing to the presence of low sequence complexity. The large molecular size and extensive conformational dynamics of the LC proteins promote intra- and inter-molecular multivalent interactions and cause spectral overlap and signal broadening, making data processing and interpretation tedious in NMR studies.[68,90] Moreover, these measurements require isotopically labeled proteins in high concentrations. Inherent conformational and structural heterogeneity of phase-separated systems can make atomic resolution structural determination quite challenging using X-ray diffraction and cryo-EM, which have been extensively used for the structural characterization of amyloid fibrils and soluble oligomers associated with neurodegenerative diseases.[119,120] Single-molecule fluorescence spectroscopic tools, including single-molecule Förster-resonance energy transfer (FRET) and fluorescence correlation spectroscopy (FCS), are powerful tools to study the structural and dynamical aspects of macromolecules over a broad range of timescales within the dense phase and the way several modulations in the cellular environment influence these properties.[92,93] smFRET investigates inter- and intra-molecular interactions and conformational rearrangements for the phase separating systems by labeling the macromolecule with a suitable donor and acceptor fluorophores and then recording the fluorescence fluctuations within femtoliters of the solution containing a very low concentration of fluorescently labeled proteins enabling single molecule detections.[94] FCS can quantitatively monitor molecular diffusion and conformational fluctuations and is ideal for probing the thermodynamics and kinetics of protein-protein interactions, which dictate the material properties of the organelles.[72,95] Often, such measurements are associated with high fluorescence background and require significant expertise, optimization, and knowledge of the system.

Raman spectroscopy is a powerful technique that can shed enormous light on the 3-dimensional structure of proteins, intramolecular dynamics, and intermolecular interactions. [96-98] In this thesis, we aimed to develop a methodology that combines the capabilities of vibrational spectroscopy and optical microscopy that can illuminate the unique molecular details of the polypeptide chains with unprecedented sensitivity within the mesoscopic liquid condensed phase at the single-droplet resolution. The details of this study are described in Chapter 2 and Chapter 3 of this thesis.

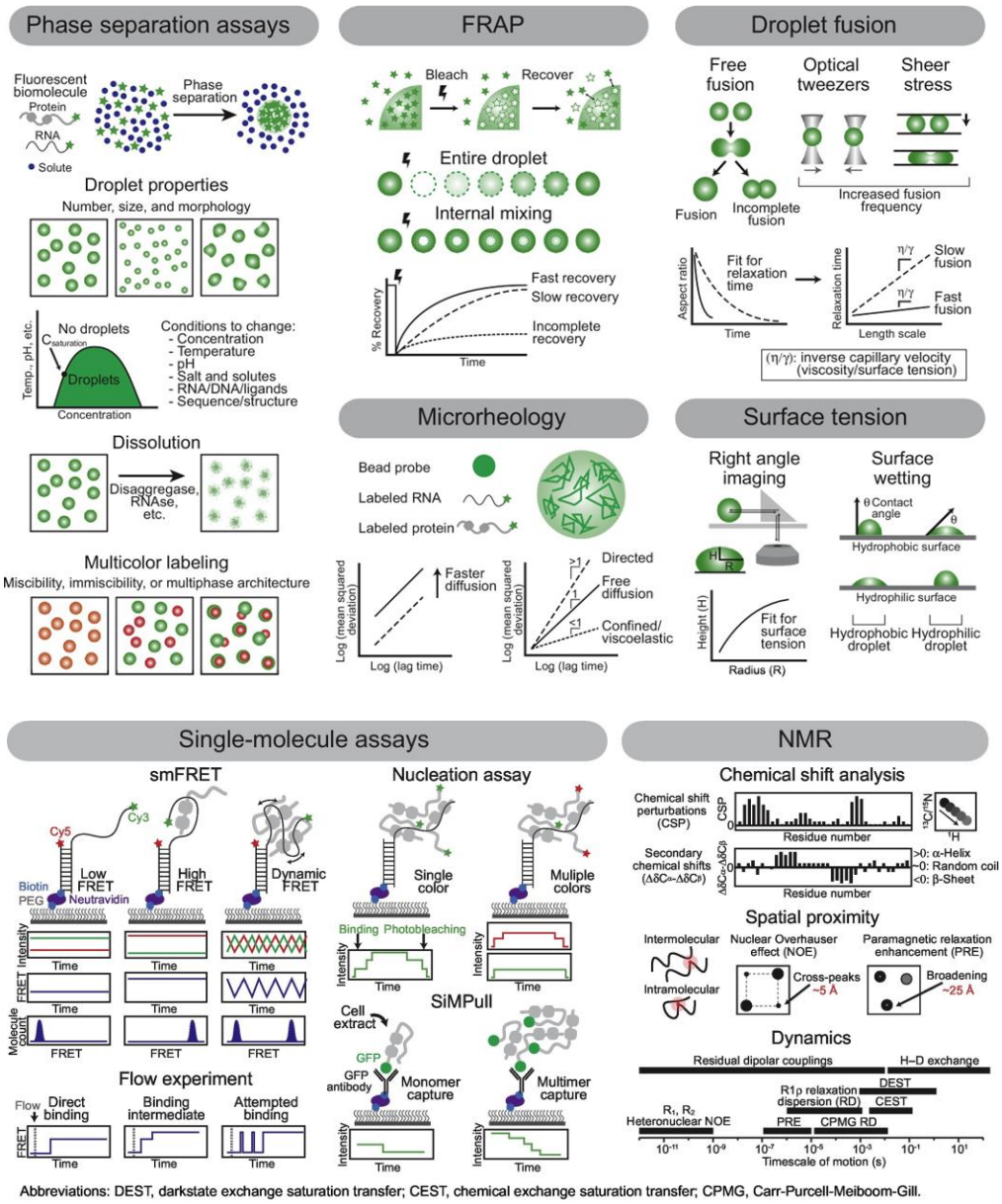


Figure 1.6: Many methods and tools have been developed to unmask the critical molecular interactions within the condensed phase, which is crucial to discern the mechanism and regulation of condensate assembly and dissolution. Reproduced with permission from reference (88).

1.3 Conformational diversity in amyloid structure and consequences

1.3.1 Supramolecular architecture of amyloids

Amyloids are typically hundreds of nanometers to micrometers in length and about 8-12 nm wide. The X-ray diffraction pattern of amyloids indicated that they share a common cross-β sheet structural motif in which the β-strands are oriented perpendicular to the growing fibril axis. The diffraction pattern showed two major reflections, meridional reflection at ~4.7 Å

corresponding to the inter-strand distance, and an equatorial reflection at 6-11 Å corresponding to inter-sheet spacing (Figure 1.7 A).[99,100] Amyloids are known to display unusual stability and tensile strength, and capacity to self-replicate their conformation by self-templating and seeding.[99,101] Each amyloid fibril consists of a close lateral association of typically two or more protofilaments held together by a set of inter-protofilament packing interactions and often wrap around one another, imparting an overall twisted ribbon or helical morphology.[102] Each protofilament consists of a stack of monomeric subunits containing one or more β-strands. Depending on the orientation of the β-strands present in the adjacent β-sheets, amyloid fibrils can have different possible structural arrangements such as parallel in-register packing, antiparallel β-sheet structure, β-solenoid and so forth (Figure 1.7 B).[103] In parallel in-register structures, each subunit contributes a single β-strand per β-sheet, and the adjacent strands are oriented in-register at ~4.7 Å. This is the most common architecture observed in pathological amyloids such as α-synuclein and β-microglobulin.[104] In an anti-parallel arrangement, the adjacent β-strands within the β-sheet run in the opposite direction. This arrangement is rare in amyloids and is typically observed in the fibrils formed by globular proteins.[105] In solenoid cross-β structure, the subunits form layered intramolecular parallel β-sheets by wrapping around an imaginary axis in a multi-layered solenoid fashion. Such an arrangement is adopted by the C-terminal domain of HET-s prion from the fungus *Podospora anserina,* which consists of three parallel β-sheets in which each subunit contributes a pair of stacked β-strand folded in a left-handed β-helical manner to form a two-layered structure.[103]

1.3.2 Mechanism of protein aggregation and amyloid formation

Amyloid deposition is involved in numerous debilitating neurodegenerative disorders. Therefore, elucidating the critical biophysical principles and the molecular mechanism involved in converting soluble proteins, partially folded/collapsed to form β-rich amyloid fibrils, is crucial in deciphering and modulating the onset and progression of human pathologies. Two models have been proposed to study the amyloid formation mechanism, nucleation-dependent polymerization (NDP), and isodesmic polymerization.[106-108] NDP mechanism follows sigmoidal aggregation kinetics and consists of three phases: the nucleation/lag phase, the growth/elongation phase, and the saturation phase (Figure 1.8). The nucleation phase involves the conversion of native proteins in equilibrium with their partially folded/unfolded forms to dynamic, transient, heterogeneous oligomeric species known as a critical nucleus or seed.[109] These highly toxic oligomeric species can further assemble to form higher-order oligomers, rapidly polymerizing into amyloid fibrils by recruiting other

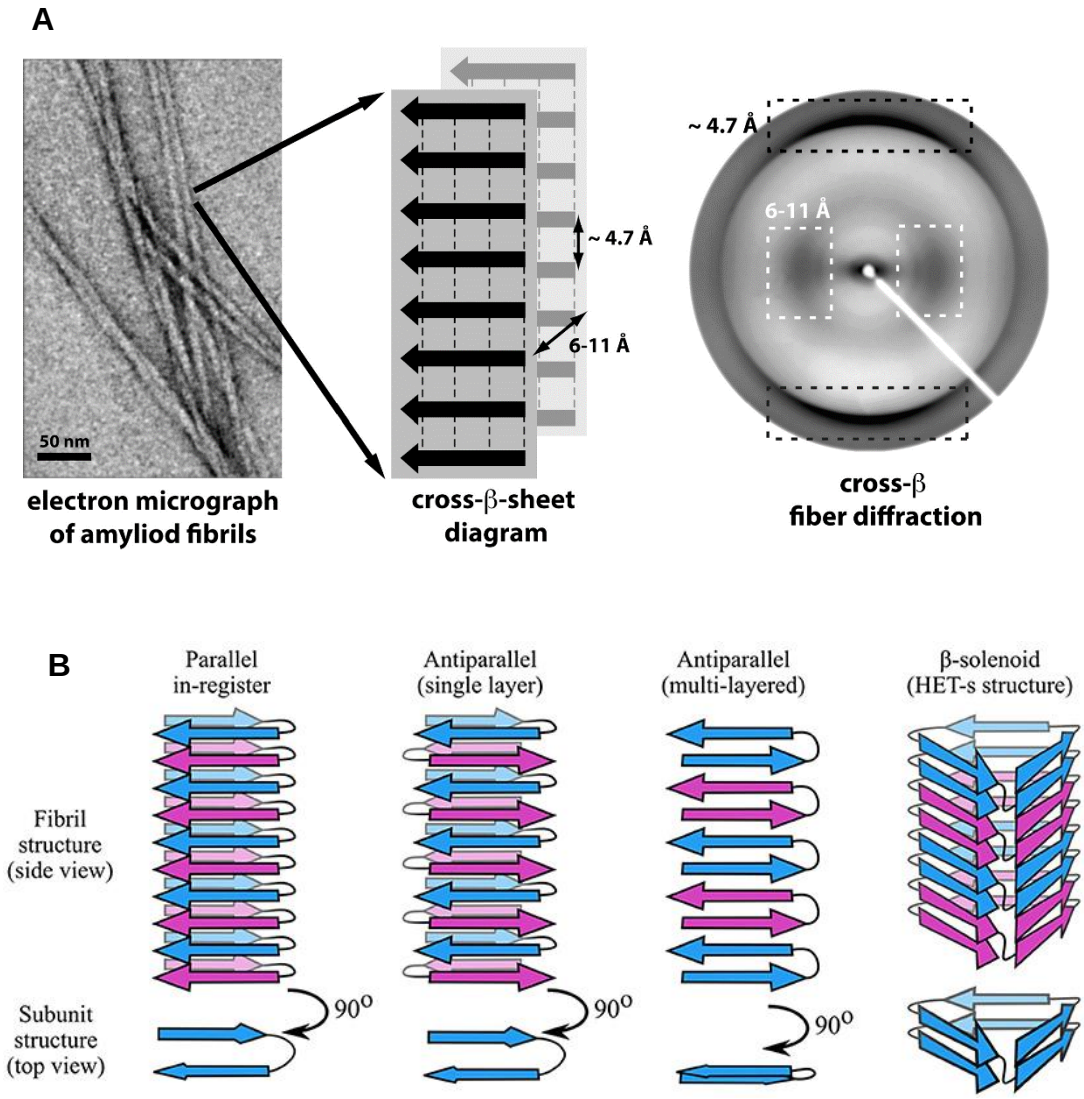


Figure 1.7: (A) Negatively stained transmission electron micrographs showing amyloid fibrils are composed of long filaments. The schematic diagram of the cross-β sheets in a fibril, with the dashed lines denoting the backbone hydrogen bonds. A typical fiber diffraction pattern is shown with a meridional reflection at 4.7 Å (black dashed box) and an equatorial reflection at 6–11 Å (white dashed box). Reproduced with permission from reference (113). (B) Classification of amyloid folds into the parallel arrangement, antiparallel arrangement, and β-solenoid. Alternate monomeric subunits are colored blue and purple in these schematics. In parallel in-register packing, β-strands in adjacent β-sheets are in the same direction, while in antiparallel arrangement, β-strands run in alternate directions. In β-solenoid subunits occupy more than one layer by coiling in a solenoidal fashion. Reproduced with permission from reference (103).

monomeric proteins. These toxic oligomers can also assemble to form off-pathway assemblies that do not form fibrils but are still toxic and may be associated with neurodegenerative diseases. As the fibril mass increases in the elongation phase, it can undergo fragmentation,

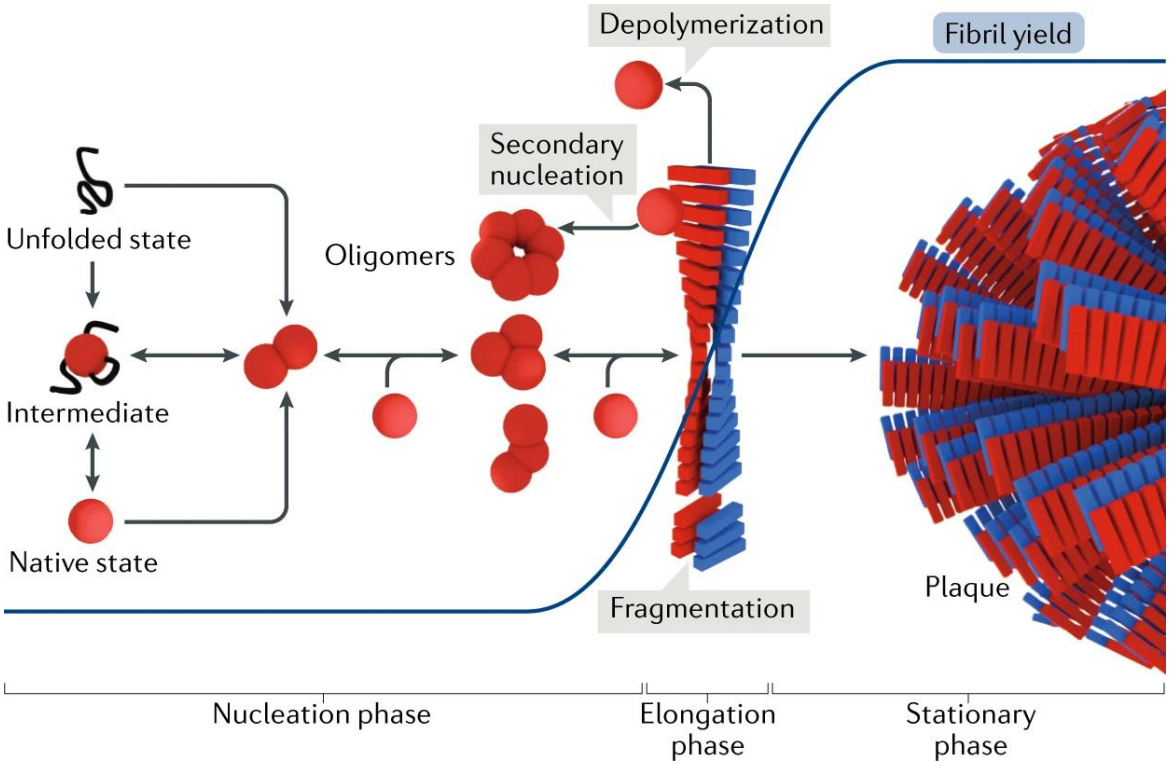


Figure 1.8: Schematic of nucleation-dependent polymerization model of protein aggregation and amyloid formation. Native, partially folded, or unfolded precursors form oligomeric species, which assemble further to form higher-order oligomers. This phase is the nucleation phase of the assembly process. Fibril nucleus rapidly recruits monomers in the elongation phase and enters the stationary phase, yielding amyloid fibrils. Secondary nucleation and fragmentation may occur in the elongation phase. The saturation phase represents mature fibrils. Reproduced with permission from reference (109).

secondary nucleation, or depolymerization to release oligomers from the fibrillar species that may or may not be toxic.[109] The rate of fibril assembly can be drastically increased by fibril fragmentation. New fibril ends generated can recruit more aggregation-prone species or catalyze aggregate formation on the surface of pre-existing fibrils via autocatalytic amplification.[110,111] Finally, mature fibrils possessing β-sheet-rich architecture are formed in the saturation phase, after which there is no further addition to the fibril ends. Adding pre-formed fibril seeds can often abolish the lag phase by bypassing the primary nucleation phase. Such an aggregation mechanism follows a single exponential function known as isodesmic or linear polymerization. For elucidating the seeding mechanisms, pre-formed fibrillar seeds rich in the cross-β structure are introduced into a solution containing monomeric protein. These pre-formed seeds act as templates for monomers and cause monomer polymerization to form amyloids. This process has been proposed as a possible basis for the propagation of infectious prion diseases and other amyloid-related disorders. Protein aggregation is typically studied *in*

vitro using an amyloid marker, Thioflavin T (ThT), which exhibits enhanced fluorescence upon binding to β-sheet rich amyloid fibrils.[112] Furthermore, light scattering experiments can probe the formation of various aggregation-prone species along the assemble pathway, circular dichroism spectroscopy can monitor the secondary structural changes during the amyloid formation while the nanoscale morphology of the aggregates can be visualized using atomic force microscopy (AFM).

1.3.3 Polymorphism in amyloid fibrils

Amyloid polymorphism arises due to the structural variations among different amyloid fibrils formed by a particular polypeptide chain resulting in distinct modes of alternative packing of polypeptide chains.[113-115] A spectrum of fibril morphologies thus generated has been linked to the pathological and clinical heterogeneity of several amyloid-associated diseases.[116] The final morphology of amyloids is determined by the number of protofilaments, variation in the orientation and conformation, and 3-dimensional packing of the protofilaments forming the fibrils. The existence of polymorphism in amyloids can lead to the generation of prion-like strains. A growing body of research involving cell lines, animal models, and human brain extracts has revealed that these amyloid fibrils exhibit molecular-level polymorphisms that are often responsible for distinct amyloid strains.[117,118] The main polymorphic structures formed by amyloids include twisted ribbons, helical ribbons, nanotubes, and crystals.[22] Out of these, the twisted ribbon is the most observed morphology of amyloids, while amyloid crystals occupy the ground state in the energy landscape (Figure 1.9). Twisted ribbon polymorph has been observed for amyloid fibrils formed by globular proteins such as β-lactoglobulin. It can undergo untwisting through thermal fluctuations and enters the deepest point of the funnel corresponding to crystal polymorph possessing low specific surface area and interfacial energy.[22] In another transition, the twisted ribbon can convert into a helical ribbon via the lateral addition of protofilaments, causing an increase in lateral width. This twisted→helical ribbon transition is typically observed when a critical number of protofilaments is reached upon lateral aggregation. The helical ribbon morphology can further reduce energy and convert to nanotubes through the closure of the edges at constant mean curvature and releasing the line tension associated with the external protofilaments.[22]

A plethora of high-resolution structural models such as solid-state NMR, cryo-electron microscopy (cryo-EM), X-ray, microcrystal electron diffraction, hydrogen/deuterium exchange NMR (HDX-NMR), electron paramagnetic resonance (EPR) and so on have been

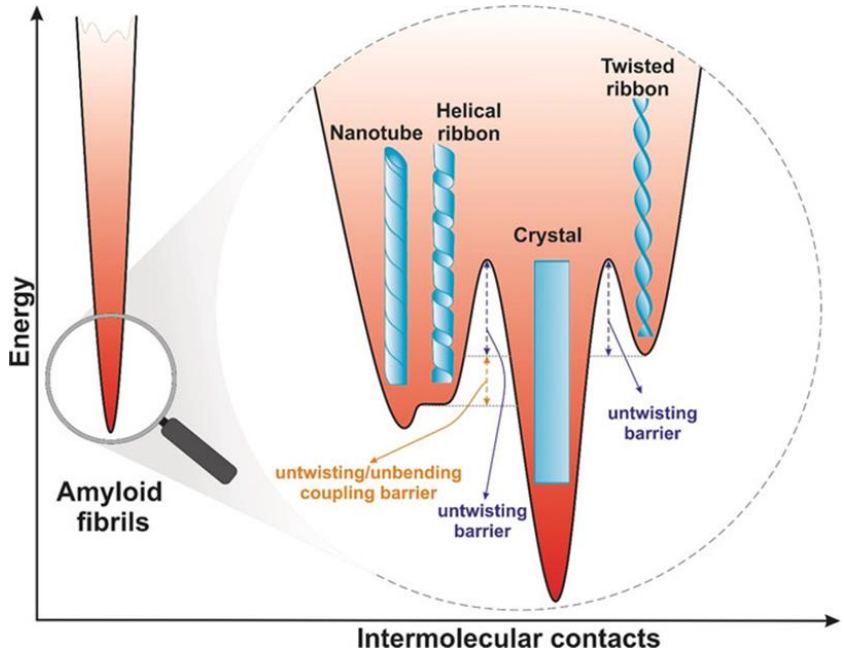


Figure 1.9: Energy landscape for the main amyloid polymorphs: twisted ribbons, helical ribbons, nanotubes, and crystals. Amyloid crystals occupy the lowest-energy ground state in the landscape, whereas all other polymorphs are metastable and occupy relative minima. Reproduced with permission from reference (22).

widely used to study molecular structures of *in vitro* or *ex vivo* generated amyloid polymorphs associated with neurodegenerative and systemic diseases as well as functional amyloids.[119-121] These structural methods can differentiate fibril polymorphs in terms of the fibril diameter, number and packing of protofilaments, presence of twists, side-chain interactions, and so forth, to provide atomic-resolution details of the supramolecular packing arrangement of the amyloid core.[122,123] Recently, cryo-EM has widely been used to delineate the distinct amyloid architectures and amino acid side-chains arrangements of several amyloid polymorphs formed by α-synuclein, tau, amyloid-β, and so on. α-synuclein forms two polymorphs, 'rod' and 'twister,' that share a common protofilament structure but differ in the inter-protofilament packing.[119] Furthermore, two different morphologies were observed and characterized in different tauopathies. The paired helical filaments (PHF) and straight filaments (SF) were observed in Alzheimer's disease patients' brains. In contrast, in the case of Pick's disease patients' brains, narrow Pick filaments (NPF) and wide Pick filaments (WPF) were characterized by cryo-EM.[124,125] However, obtaining high-resolution atomic structures can be quite challenging since these tools either require site-specific ^{13}C and/or ^{15}N labeling of proteins or involve extensive sample preparation. Additionally, techniques such as cryo-EM

demand expensive analysis software and are highly dependent on expert skills and knowledge.[126]

In this thesis, we utilize protein vibrational Raman spectroscopy coupled with hydrogen/deuterium exchange to capture the conformational heterogeneity and structural distribution within three distinct α-synuclein polymorphs exhibiting variable hydrogen bonding patterns. The details of this study are described in Chapter 4 of this thesis.

1.4 Vibrational Raman spectroscopy in proteins

1.4.1 Physical principles of Raman effect

Raman effect is an inelastic scattering phenomenon that involves the inelastic scattering of photons (Raman scattering). Raman scattering was first discovered by the Indian physicist C.V. Raman and his research partner K. S. Krishnan in 1928 and is well-described in the literature to study the vibrational, rotational, and low-frequency modes of the molecules.[127]

Upon the illumination of a molecule with electromagnetic radiation, the oscillating electromagnetic field of the photon induces a polarization of the molecular electron cloud, and an exchange of quantum vibrational energy takes place between the incident photon and the molecule. This leads to the formation of a very short-lived complex between the photon and the molecule, leaving the molecule in an unstable virtual higher energy state. This virtual state is the excited vibrational state of the ground electronic state from where the photon is immediately reemitted as scattered light.[128] Often, the energy of the molecule remains unchanged post its interaction with the incident photon, and the scattered photon is emitted with the same energy/wavelength as that of the incident photon. This process is the elastic scattering of photons and is termed Rayleigh scattering (Figure 1.10).

However, one out of 10^{10} incident photons can undergo inelastic scattering, where the molecule can either gain or lose energy. This rare event is termed Raman scattering. If the molecule gains energy from the photon during the scattering (excited to a higher vibrational level) then the scattered photon loses energy, and its wavelength increases called Stokes Raman scattering. Inversely, if the molecule loses energy by relaxing to a lower vibrational level, the scattered photon gains the corresponding energy, and its wavelength decreases, called Anti-Stokes Raman scattering. Quantum mechanically, Stokes and Anti-Stokes are equally likely processes. However, with an ensemble of molecules, most molecules will be in the ground vibrational level (Boltzmann distribution), and Stokes-Raman scattering is

statistically more probable. As a result, the Stokes scattering is always more intense than the anti-Stokes (Figure 1.11).

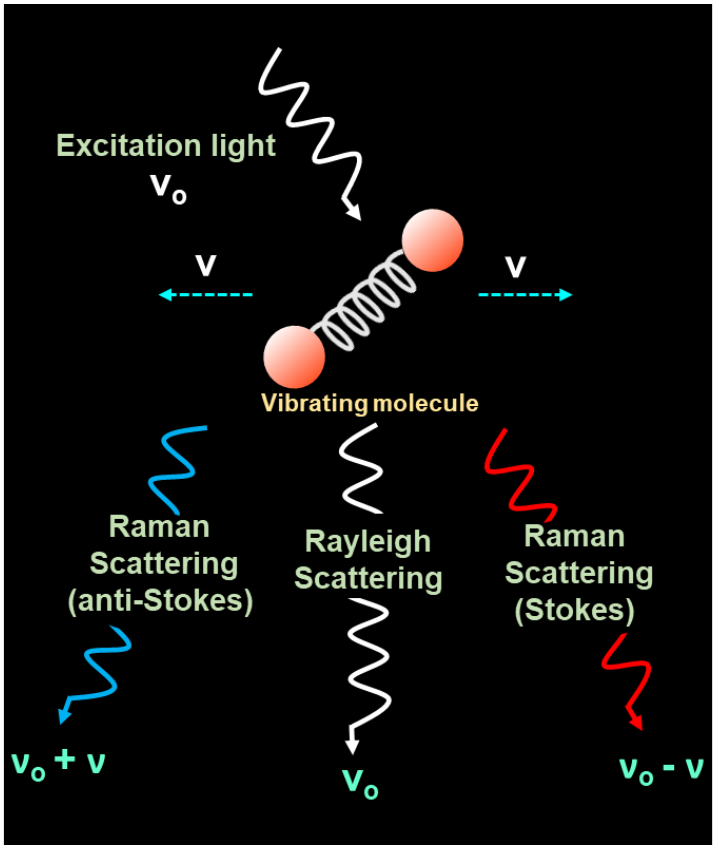


Figure 1.10: Schematic representation depicting scattering of incident photons (ν_o) by vibrating molecules. Stokes Raman scattering occurs when the scattered photons have energy lower than the incident photons ($\nu_{stokes} < \nu_o$). In contrast, anti-Stokes Raman scattering occurs when the scattered photons have energy higher than incident photons ($\nu_{anti\text{-}Stokes} > \nu_o$). Rayleigh scattering occurs when the photons are scattered with identical energy to the incident photons ($\nu_{Rayleigh} = \nu_o$).

1.4.2 Vibrational Raman bands in proteins

Polypeptide structure, conformations, intermolecular interactions, and dynamics can be analyzed using several vibrational modes in proteins, as discussed below.[129-132] Table 1.2 summarizes the most common vibrational bands observed in a protein Raman spectrum (Figure 1.12).[133-136]

Amide bands

Amide I (1630-1700 cm^{-1}): This intense band is observed in protein Raman spectra primarily due to C=O stretching vibrations along with a minimal contribution from the out-of-plane C-N stretching vibrations and is therefore sensitive to the backbone conformation of the proteins,

deuteration, and hydrogen bonding strength between the C=O and N-H groups. This band is one of the secondary structural marker bands and is commonly used to interpret changes in the protein's secondary structure. The Amide I band frequency often depends on the conformation of the polypeptide backbone and the protein sample's intra- and intermolecular hydrogen bonds. For instance, the Amide I band located in the range 1665-1680 cm^{-1} corresponds to the β-sheet structure, while the α-helical secondary structure is observed in the range 1640-1655 cm^{-1}. The position of the Amide I band for β-sheet depends on the number of β strands and the hydrogen-bonding strength that can be used to distinguish parallel and anti-parallel β-sheets. Amide I band located in the range 1655-1665 cm^{-1} and 1680-1700 cm^{-1} is typically assigned to disordered, turns or random coil protein secondary structures.[137]

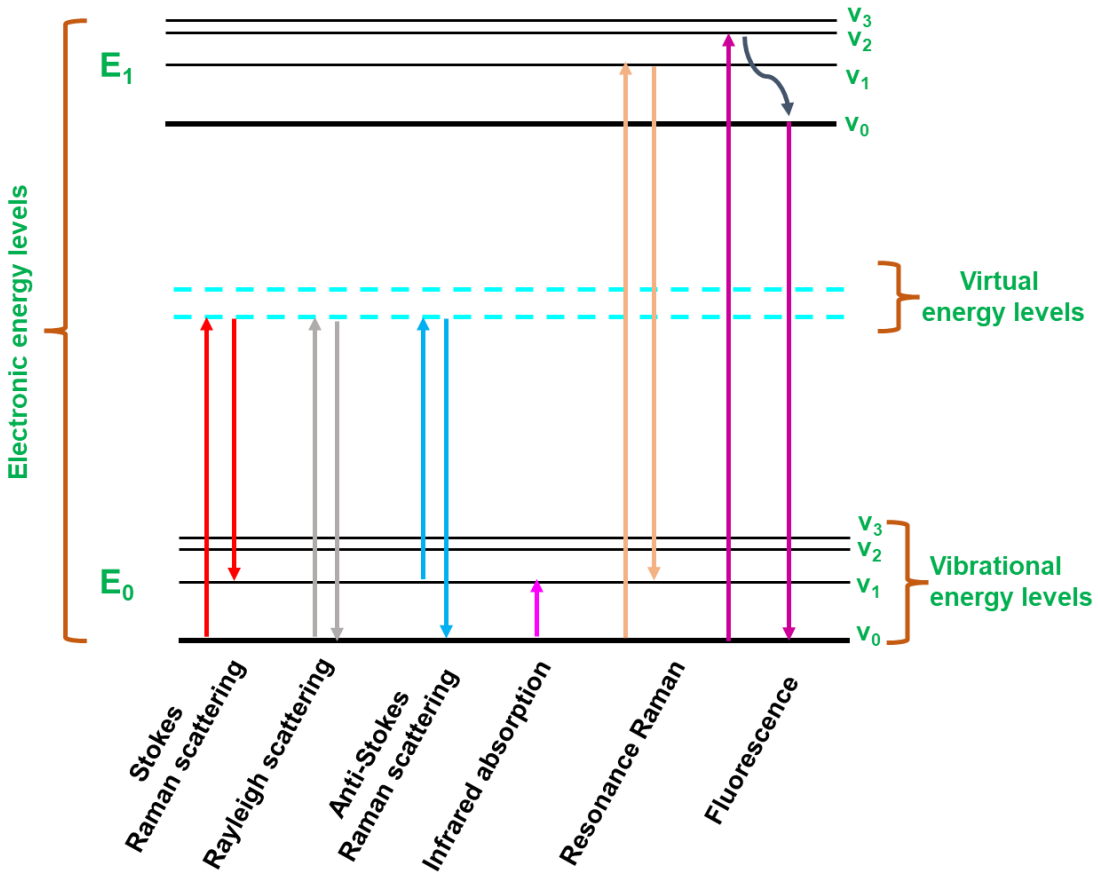


Figure 1.11: Energy level diagram illustrating various transitions occurring for Rayleigh scattering, Raman scattering (Stokes/anti-Stokes), Infrared absorption, Resonance Raman, and Fluorescence transitions. Subscript 0 denotes the ground state, while 1,2,3, and so on show the first, second, and third excited energy levels.

Amide II (1480-1580 cm^{-1}): This band originates from the combination of out-of-plane N-H bending and the C-N stretching vibrations with minor contributions from the C-O bending and the C-C stretching vibrations. This band is typically invisible in the normal Raman spectra of proteins due to the small Raman scattering cross-section of this band at near-IR laser excitation. In contrast, a strong Amide II band is visible in deep UV Resonance Raman (DUVRR) spectra of proteins. Amide II band is susceptible to the H/D exchange in DUVRR, where it undergoes a red-shift to ~1440 cm^{-1} upon deuteriation and is often used for quantitative estimation of the amount of hydrophobic fibril core unavailable for H/D exchange.[138]

Amide III (1230-1320 cm^{-1}): This band originates from the combination of in-phase N-H bending and the C-N stretching vibrations with minor contributions from the C-O in-plane bending and the C-C stretching vibrations. Akin to Amide I band, Amide III band frequencies are also dependent on the secondary structural elements present in the proteins. For instance, Amide III band for α-helices appear ~1300 cm^{-1}, 2.5_1 helix at ~1272 cm^{-1}, β-sheets at ~1240 cm-1, polyproline II (PPII) at ~1252 cm^{-1}, and random coils at ~1240-1250 cm^{-1}. Since the Amide III band frequencies due to various secondary structures have minimal overlap, estimating secondary structural elements is less complicated. Moreover, the Amide III band position is susceptible to the conformations of the polypeptide backbone. Asher and coworkers have shown that the dihedral angle varies with the Amide III band frequencies, and the Ramachandran dihedral angle (ψ) of the backbone can be easily calculated using the Amide III band position using the empirical relation given by Asher and coworkers.[139,140] Amide III is highly sensitive to deuteration, which disappears upon H→D exchange due to the loss of coupling between the N-H bending and C-N stretching vibrations.[138]

Amide IV: This band appears in the range 625-770 cm^{-1} and originates from OCN bending vibrations.

Amide V: This band appears in the range 640-800 cm^{-1} and originates from out-of-plane N-H bending vibrations.

Amide VI: This band appears in the range 540-600 cm^{-1} and originates from out-of-plane C=O bending vibrations.

Amide VII: This band appears around ~200 cm^{-1} and originates primarily from skeletal vibrations.

Amide I and Amide III bands are primarily used for protein folding, unfolding, and interaction studies.

Bands due to aromatic amino acids

Phenylalanine (Phe): Phenylalanine shows many vibrational modes at 620, 1000, 1030, 1205, 1590, and 1610 cm^{-1}. The band at 1000 cm^{-1} due to ring breathing vibrations of the benzene ring is the most critical mode. It is often used for the intensity normalization of the Raman spectra of proteins since this band is insensitive to protein conformations. However, the band intensity varies with the surrounding solvent conditions. Moreover, the band at 620 $cm^{-1,}$ due to in-plane ring deformation, is used as a hydrophobicity marker since the intensity of this band is sensitive to the solvent conditions in the vicinity of this residue.

Tyrosine (Tyr): Seven common Raman markers are typically observed for Tyr at 643, 830, 850, 1178, 1210, 1606, and 1616 cm^{-1}. The intensity ratio at 850 cm^{-1} and 830 cm^{-1} (I_{850}/I_{830}) of the tyrosine Fermi doublet that is observed due to Fermi resonance between the ring-breathing vibration and overtone of an out-of-plane ring-bending vibration (~413 cm^{-1}) of the phenolic ring of tyrosine is often used to decipher the sidechain environment. Therefore, this ratio indicates the solvent-mediated hydrogen bonding propensity of the phenolic (-OH) group and measures the solvent accessibility of tyrosine residues.[141] The I_{850}/I_{830} ratio is typically ≥ 2 for a well-solvated tyrosine and is used as a marker for the hydrophobic/hydrophilic environment around the residue.[141,142] Additionally, bands at 1178 cm^{-1} and 1606 cm^{-1} are often used to probe the hydration environment in the vicinity of Tyr residues.

Tryptophan (Trp): Trp shows multiple vibrational bands at 760, 880, 1010, 1340, 1360, and 1550 cm^{-1}. The intensity ratio at 1360 cm^{-1} and 1340 cm^{-1} (I_{1360}/I_{1340}) is used as a hydrophobicity marker in the vicinity of tryptophan residues. The value for this ratio increases (usually >1) with an increase in the hydrophobic environment around the tryptophan residues. Further, the Raman band at 880 cm^{-1} measures the hydrogen bonding strength between the indole ring and the surrounding solvent molecules. This band undergoes a red shift in the case of strong hydrogen bonding. Additionally, a tryptophan band at 767 cm^{-1} corresponds to the indole ring breathing and is used as a marker for cation-π/CH-π interactions.[143,144]

Bands due to non-aromatic amino acids

Many vibrational modes appear in the range 400-1800 cm^{-1} due to side chain vibrations of non-aromatic amino acids. Most of these bands overlap due to the presence of many residues in

proteins.

Disulfide bonds show a Raman marker at 500-550 cm^{-1} depending on the conformation of the S-S bond in proteins. For instance, *gauche-gauche-gauche* (ggg) conformation shows a band ~505-515 cm^{-1}, *gauche-gauche-trans/gauche-trans-gauche* (ggt/gtg) conformation shows a band at 520-530 cm^{-1}, and *trans-gauche-trans* (tgt) conformation appears at 540-545 cm^{-1}. In addition to these bands, the S-H group appears at ~ 2500 $cm^{-1,}$ and the band position depends on the hydrogen bonding strength of the thiol group. Free cysteine appears at ~ 2585 cm^{-1} while hydrogen bonding shifts this band to a lower wavenumber. Vibrational bands at ~ 2575, 2565, and < 2565 cm^{-1} indicate weak, moderate, and strong H-bonding, respectively.[136]

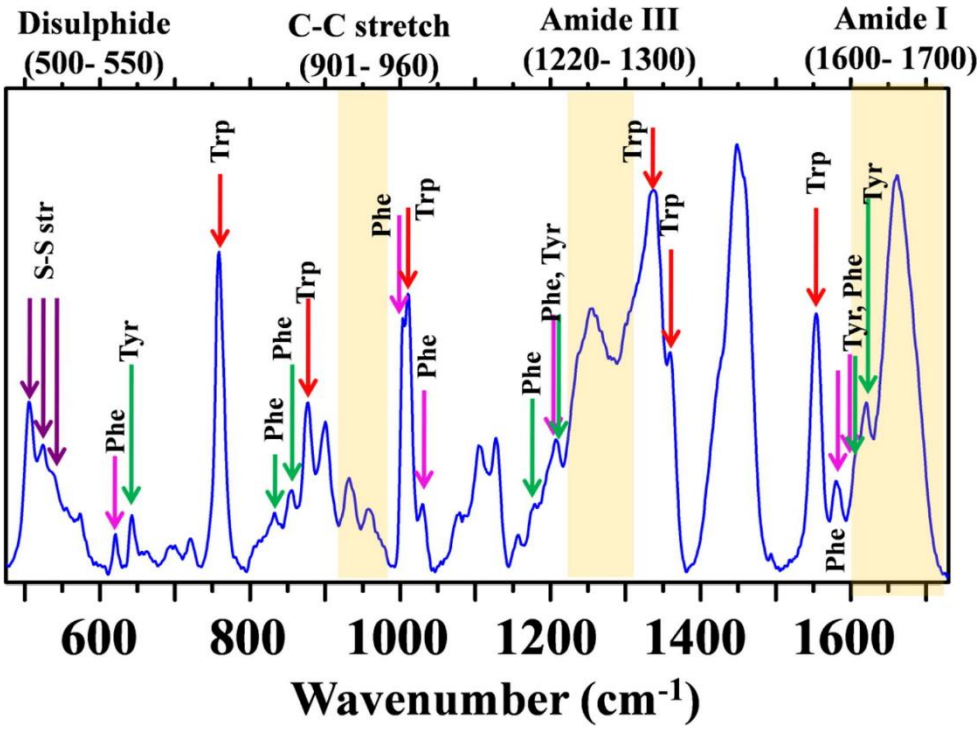


Figure 1.12: Raman spectrum of hen egg white lysozyme in water depicting various Raman scattering bands arising from different vibrational modes of the aromatic amino acids (tyrosine, tryptophan, and phenylalanine) in addition to several other aliphatic side-chain vibrations. Reproduced with permission from (136).

Table 1.2 Common vibrational bands of protein Raman spectra.[133-136]

Amide Bands Amide bands represent the motion of atoms of the peptide backbone and determine the secondary structure of the proteins. **Amide I** (primarily a carbonyl stretching mode) **Amide III** (combines in-plane N-H bending and C-N stretching motions)	 ○ α-helix: ~1630-1655 cm^{-1} ○ β-sheet: ~1669-1675 cm^{-1} ○ Non-regular/disordered structures: ~1655-1669 cm^{-1} and 1675-1690 cm^{-1} ○ α-helix: ~1280-1320 cm^{-1} ○ β-sheet: ~1220-1240 cm^{-1} ○ Non-regular/disordered structures: ~1250-1270 cm^{-1}
Disulfide stretch (S-S) and carbon-sulfur (C-S) bands Raman bands due to disulfide bonds appear in 500-550 cm^{-1} region. Different conformations of atoms around S-S bonds appear at different wavenumbers. C-S bands of cysteine and methionine residues appear depending on sidechains and surrounding environment.	 ○ Gauche-gauche-gauche (g-g-g): ~515 cm^{-1} ○ Gauche-gauche-trans (g-g-t): ~525 cm^{-1} ○ Trans-gauche-trans (t-g-t): ~540 cm^{-1} ○ 630-760 cm^{-1}
Aromatic amino acids **Phenylalanine** • Intense band due to ring breathing vibrations and is often used for the normalization of the Raman spectra of proteins as this band is independent of conformational changes of the proteins. **Tryptophan** • Indole ring-breathing vibrations	 ○ 1000 cm^{-1} ○ ~ 760 cm^{-1} ○ Marker band for cation-π/CH-π interactions ○ Band intensity increases with decreasing hydrophobicity.

• Indole N-H bending vibrations and is sensitive to indole N-H hydrogen bond donation	○ ~ 880 cm^{-1} ○ Blue shift indicates indolyl moiety is located in highly hydrophobic environment. ○ Red shift indicates strong H-bonding representing strength of N-H bond of the indole ring with surrounding solvent molecules.
• Fermi doublet at 1360 cm^{-1} and 1340 cm^{-1}	○ I_{1360}/I_{1340} increases with increasing hydrophobicity of indole ring environment.
Tyrosine	
• Fermi doublet at 850 cm^{-1} and 830 cm^{-1} observed due to Fermi resonance between the ring-breathing vibrations and overtone of an out-of-plane ring-bending vibration of the phenolic ring of tyrosine.	○ I_{850}/I_{830} is an indicator of solvent-mediated hydrogen bonding propensity of the phenolic (-OH) group
• Ring stretching mode	○ 1617 cm^{-1} & 1600 cm^{-1}
• C-O stretching mode	○ 1263 cm^{-1}
• C-C stretching mode	○ 1210 cm^{-1}
• O-H bend + C-H bend	○ 1180 cm^{-1}
Backbone CH_2/CH_3 deformations	~ 1440-1470 cm^{-1}
Backbone $^{\alpha}$C-H bending vibrations	~ 1390 cm^{-1}

1.4.3 Limitations of normal Raman spectroscopy

A multitude of structural methods such as solid-state NMR, cryo-electron microscopy (cryo-EM), X-ray, microcrystal electron diffraction, hydrogen/deuterium exchange NMR (HDX-NMR), electron paramagnetic resonance (EPR), and so on are widely used in structural biology to decipher the behavior and function of biomolecules at the molecular level.[87,88,120,121] However, most of these tools require extensive sample preparation, such as protein crystallization or site-specific isotope labeling of proteins, and can be quite cumbersome, requiring expert skills and knowledge.[126] In this direction, vibrational Raman spectroscopy proves to be a promising technique that can illuminate the 3-dimensional structure of proteins,

intramolecular dynamics, and intermolecular interactions. It is a non-contact, label-free, non-destructive tool to assess protein conformations, dynamics, and specific interactions in physiological conditions with minimum sample preparation.[133,134]

Although Raman spectroscopy has been widely used for the characterization, identification, and detection of a range of proteins and biomolecules as well as pathogenic organisms such as bacteria and viruses, there are several limitations associated with this technique.[145,146] One of the significant limitations arises from the low Raman scattering cross-section of proteins since only one in a million photons is scattered inelastically, reducing its efficiency.[133] This makes normal Raman spectroscopy a highly insensitive tool for biomolecule characterization in aqueous solutions under physiological conditions. Moreover, high sample concentration, laser power, and magnifications required for Raman spectroscopic detection can be destructive for soft biological samples and lead to laser-induced damage and other artifacts.

1.4.4 Surface-enhanced Raman spectroscopy (SERS)

The limitations associated with normal Raman spectroscopy can be elegantly overcome by utilizing an ultrasensitive analytical technique known as surface-enhanced Raman spectroscopy (SERS).[147-150] SERS causes a dramatic increase (by a factor of $\geq 10^8$) in the inelastic scattering of light by molecules when they are bound to or placed in the vicinity of plasmonic nanostructures. SERS primarily results from the enhanced local fields that arise from plasmon excitation in the silver or gold nanoparticles.[148]

Fleischmann, Hendra, and McQuillan first observed the significant increase in Raman scattering signals at the University of Southampton in 1974 during measurements of the Raman scattering of pyridine adsorbed on a roughened silver electrode.[151] This unexpectedly huge Raman signal intensity was believed to be due to the increased surface area of the silver electrodes. Later in 1977, Richard Van Duyne and his graduate student David Jeanmaire came up with the electromagnetic field mechanism of SERS and proposed that the significant amplification of Raman signals is due to "some property of the electrode surface or electrode/solution interface acting to enhance the effective Raman scattering cross-section" (Figure 1.13).[152,153] During the same time, Albrecht and Creighton suggested the chemical (or charge transfer) mechanism for enhancement, hypothesizing that the enhancement occurs due to the formation of a metal-analyte complex.[154] Since its discovery almost five decades ago, SERS has been widely used in protein detection, biophysical characterization of biomolecules, including lipids and cells, and protein-ligand interactions.[148,149,155] Recent advances and

developments allow us to carry out SERS measurements *in vivo* by injecting the SERS active substrates inside the cells and monitoring the processes occurring on cellular membranes. Therefore, this potent analytical tool is gaining immense popularity in the field of medicine, drug delivery, biomedical engineering, biosensing, catalysis, electrochemistry, material science, pharmacology, and so on.

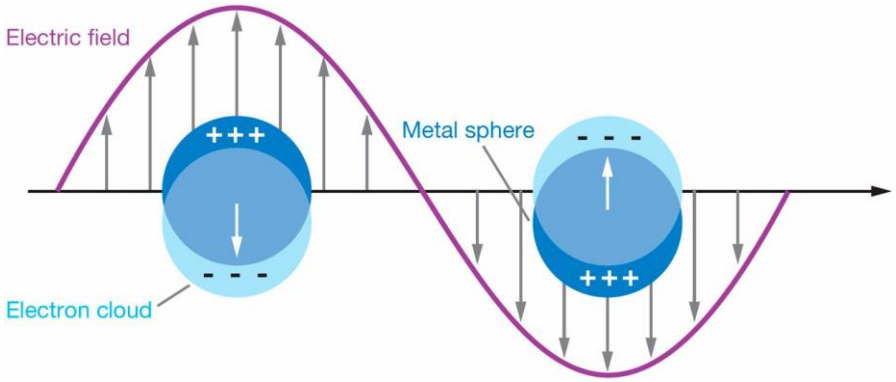


Figure 1.13: The iconic figure that depicts the generation of a localized surface plasmon resonance. Reproduced with permission from (156).

1.4.5 Mechanisms of SERS signal enhancement

Two enhancement mechanisms are widely accepted to explain the increase in Raman scattering intensities; electromagnetic enhancement and chemical enhancement.[157,158] Although electromagnetic enhancement is considered the major contributor to SERS signals, the total SERS enhancement factor is usually a product of electromagnetic and chemical enhancement mechanisms (Figure 1.14). Typical SERS substrates used are coinage metals (Ag, Au, and Cu) due to their strong optical absorptions in the visible region. Ag and Au are highly preferred among these metals for SERS studies due to their strong plasmonic effect, relative chemical inertness, and well-controlled nanostructures. Lately, dielectrics and, in particular, semiconductors have emerged as an alternative to conventional metal-based SERS substrates due to their low invasiveness, reproducibility, versatility, and recyclability.

1.4.5.1 Electromagnetic enhancement mechanism

A positively charged metallic nanosphere is surrounded by a sea of oscillating electrons referred to as plasmon. When an incident electromagnetic field $E_o(\lambda)$ in resonance with the plasmonic frequency is incident on the nanostructure, the surface plasmons get excited (known as localized surface plasmon resonance, LSPR) and generate an oscillating dipole with an electromagnetic field E_s that varies inversely with the twelfth power of the distance of the target

analyte from the nanoparticle surface.[148] Therefore, the analyte experiences an enhanced field $E_i=E_o+E_s$ and behaves as an antenna to transfer the Raman signal from the near field to the far field, and the Raman signal is proportional to the enhanced local electric field at the Raman emission wavelength. This emits three kinds of frequencies, Rayleigh, Stokes, and anti-Stokes. The SERS electromagnetic enhancement is usually in the order of 10^{10}. It is often approximated to the fourth power of the local field (having contribution from the incident and Raman scattered field) in the vicinity of the metal.[148,150]

1.4.5.2 Chemical enhancement mechanism

This enhancement mechanism requires the analytes to be in direct contact with the nanoparticle surface and is, therefore, a short-range effect contributing 10^2-10^4 of the total enhancement factor.[157] The enhancement typically arises due to charge transfer between the metal and analyte, causing a modification in the polarizability of a molecule. Thus, the Raman cross-section of the vibrational mode causes an increase in Raman scattering intensity. Charge transfer transitions take place from the highest occupied molecular orbital (HOMO) of the analyte to the Fermi level of the metal substrate, or such transitions can also take place between the HOMO of the analyte to the conduction band (CB) of the nanostructure or from valence band (VB) to the lowest unoccupied molecular orbital (LUMO) of the analyte in the case of dielectric substances.[158] Typically, the charge transfer can have three types of contributions. First is the ground state charge transfer (GSCT), which is a non-resonant phenomenon involving chemisorption between the molecule and metal in the ground state without excitations. Second is the photo-induced charge transfer (PICT) between the metal-to-molecule or molecule-to-metal in resonance with excitation photons. The charge transfer direction depends on the metal's Fermi level (EF) location and HOMO/LUMO in the molecule. Resonance Raman scattering can also be involved where the excitation frequency matches the electronic transition frequency of the molecule (μ_{mol}). Additional charge transfer transitions can also occur due to the transference of electrons in semiconductors from the valence band to the conduction band to generate electron-hole pairs known as excitons (μ_{ex}).[159]

1.5 Thesis motivation and perspective

It is well established that the physiological functions performed by any given protein are governed by its well-defined 3-dimensional structure, which is often controlled by the highly orchestrated chaperone machinery of the cell. IDPs/IDRs are a class of proteins that defy this

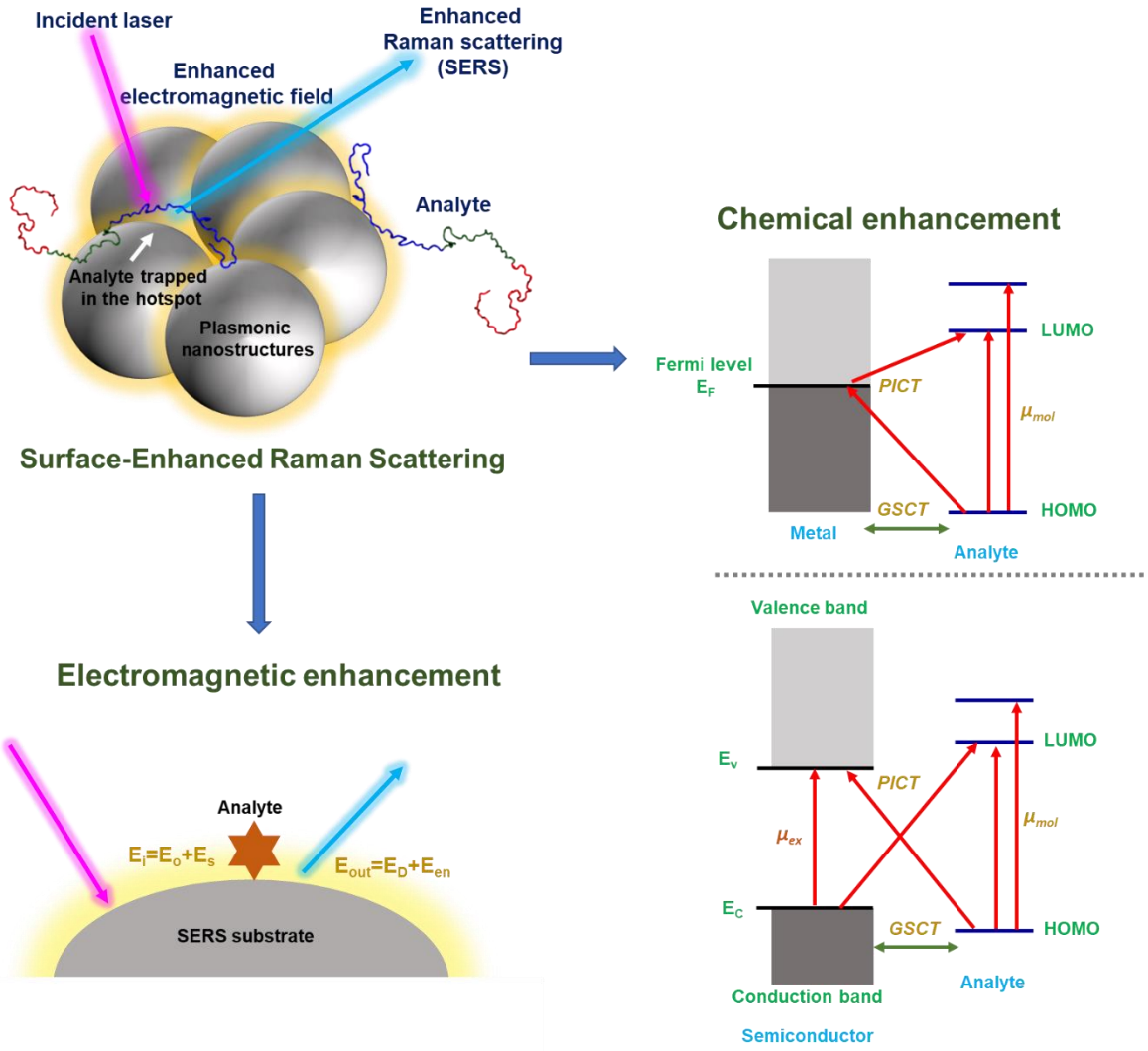


Figure 1.14: Schematic representation of surface-enhanced Raman scattering. Massive enhancements are observed when the analyte gets trapped in the regions between the nanoparticles known as "plasmonic hotspots." Electromagnetic enhancements and chemical enhancement mechanisms drive SERS. E_i: resultant field experienced by the analyte; E_o: incident field; E_s: due to surface plasmon resonance; E_{out}: outgoing field; E_c: conduction band energy; E_v: valence band energy; E_F: metal's Fermi level.

conventional structure-function paradigm and fail to fold autonomously into ordered globular structures rather than exist as an ensemble of rapidly interconverting structures. These dynamic proteins can perform various physiological functions by undergoing disorder-to-order transitions or maintaining their disordered characteristics in "fuzzy complexes." However, under certain pathological conditions, IDPs undergo misfolding to form partially folded states/unfolded intermediates, thereby leading to extra/intracellular accumulation and deposition of misfolded proteins termed amyloidosis. This condition leads to the formation of

either amorphous aggregates or amyloids and has been linked to various fatal neurodegenerative disorders. Polymorphism is one of the remarkable features of amyloid fibril structures and gives rise to morphological and biological strain diversities. Further, a growing body of current research has revealed that cells can perform various critical functions employing (non-canonical) membrane-less organelles formed via intracellular phase separation of IDPs and/or nucleic acids. These liquid-like compartments can transform into viscoelastic gel-like or solid-like aggregates under stressed conditions that can have lethal implications for the cell. Thus, unmasking the key molecular interactions within the condensed phase is crucial to discern the mechanism and regulation of condensate assembly and dissolution. Several methods and techniques have been routinely used to structurally characterize the phase-separated and the amyloid states of IDPs, such as NMR, X-ray diffraction, cryo-EM, single-molecule studies, and so forth.

This thesis describes the conformational and structural characterization of the phase-separated protein condensates and conformational fingerprinting of distinct amyloid polymorphs using vibrational Raman spectroscopy that serves as a highly potent, noninvasive, and label-free tool. In chapter 2 of this thesis, we utilize this tool to study the conformational heterogeneity and structural distributions within several phase-separated liquid droplets in a single droplet-by-droplet fashion for two interesting phase-separating systems, PrP-Y145Stop, and full-length FUS (Fused In Sarcoma). Our single-droplet Raman experiments indicate conformational heterogeneity and intrinsic disorder within the protein droplets. Additionally, this tool allows us to capture the environment in the vicinity of specific amino acid residues as well as chain collapse by recording the Raman scattering bands due to various vibrational modes of the polypeptide chains. Normal Raman spectroscopy is a relatively weak tool due to the proteins' low Raman scattering cross-section. This makes recording vibrational Raman spectra challenging for aqueous solutions under physiological conditions. In chapter 3, we introduce a unique single-droplet surface-enhanced Raman scattering that uses surface-engineered, plasmonic, metal nanoparticles to unveil the inner workings of mesoscopic liquid droplets of a well-known RNA binding protein, full-length Fused in Sarcoma (FUS) in the absence and presence of RNA. These highly sensitive measurements offer unprecedented sensitivity to capture the crucial interactions, conformational heterogeneity, and structural distributions within the condensed phase in a droplet-by-droplet manner. Finally, in chapter 4, we utilize vibrational Raman spectroscopy coupled with hydrogen/deuterium exchange to gain structural insights from diverse amyloid polymorphs of a presynaptic neuronal protein, α-

synuclein, aggregation of which is associated with numerous human pathologies termed as synucleinopathies. Hydrogen-deuterium exchange Raman spectroscopy allowed us to structurally differentiate the higher-order assemblies displaying a common cross-β-structural motif with varying hydrogen bonding strengths within the β-spine and side-chain packing. I believe the work presented in this thesis can broaden our current understanding of vibrational Raman spectroscopy as a potent tool to decipher the key molecular drivers of phase transitions and critical interactions in various biomolecular condensates. Further, hydrogen-deuterium exchange vibrational Raman spectroscopy can help study a wide range of *in vitro* and *ex vivo* generated amyloid polymorphs displaying strain diversity associated with physiology and pathology.

1.6 References

1. Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999,** 293(2), 321-331.

2. Dunker, A. K.; Obradovic, Z.; Romero, P.; Garner, E. C.; Brown, C. J. Intrinsic protein disorder in complete genomes. *Genome Inform Ser. Workshop Genome Inform.* **2000,** 11, 161-171.

3. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z. Intrinsically disordered protein. *J. Mol. Graph Model.* **2001,** 19(1), 26-59.

4. Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradović, Z. Intrinsic disorder and protein function. *Biochemistry* **2002,** 41(21), 6573-6582.

5. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002,** 27(10), 527-533.

6. Dyson, H. J.; Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell. Biol.* **2005,** 6(3), 197-208.

7. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; Han, K. H.; Jones, D. T.; Longhi, S.; Metallo, S. J.; Nishikawa, K.; Nussinov, R.; Obradovic, Z.; Pappu, R. V.; Rost, B.; Selenko, P.; Subramaniam, V.; Sussman, J. L.; Tompa, P.; Uversky, V. N. What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. *Intrinsically Disord. Proteins.* **2013,** 1(1), e24157.

8. Dill, K. A.; Chan, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **1997,** 4(1), 10-9.

9. Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. Navigating the folding routes. *Science* **1995,** 267(5204), 1619-1620.

10. Turoverov, K. K.; Kuznetsova, I. M.; Uversky, V. N. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog. Biophys. Mol. Biol.* **2010,** 102(2-3), 73-84.

11. Jahn, T. R.; Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **2005,** 272(23), 5962-5970.

12. Betts, M. J.; Sternberg, M. J. An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Eng.* **1999,** 12(4), 271-283.

13. Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **1991,** 254(5038), 1598-1603.

14. Uversky, V. N.; Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta.* **2010,** 1804(6), 1231-1264.

15. Uversky, V. N. Intrinsically disordered proteins and their “Mysterious” (Meta)Physics. *Fron. Phys.* **2019,** 7, 1-18

16. Romero, P.; Obradovic, Z.; Kissinger, C. R.; Villafranca, J. E.; Garner, E.; Guilliot, S.; Dunker, A. K. Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput.* **1998,** 437-448.
17. Burger, V. M.; Gurry, T.; Stultz, C. M. Intrinsically Disordered Proteins: Where Computation Meets Experiment. *Polymers* **2014,** 6(10), 2684-2719.
18. Ke, P. C.; Zhou, R.; Serpell, L. C.; Riek, R.; Knowles, T. P. J.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fändrich, M.; Otzen, D. E.; Chapman, M. R.; Dobson, C. M.; Eisenberg, D. S.; Mezzenga, R. Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **2020,** 49(15), 5473-5509.
19. Dobson, C. M. Protein folding and misfolding. *Nature* **2003,** 426(6968), 884-890.
20. Hartl, F. U.; Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **2009,** 16(6), 574-581.
21. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **2008,** 37, 215-246.
22. Adamcik, J.; Mezzenga, R. Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape. *Angew. Chem. Int. Ed. Engl.* **2018,** 57(28), 8370-8382.
23. Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **2011,** 475(7356), 324-32.
24. Chaari, A.; Hoarau-Véchot, J.; Ladjimi, M. Applying chaperones to protein-misfolding disorders: molecular chaperones against α-synuclein in Parkinson's disease. *Int. J. Biol. Macromol.* **2013,** 60, 196-205.
25. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014,** 114(13), 6589-6631.

26. Uversky, V. N. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta.* **2013,** 1834(5), 932-951.

27. Uversky, V. N. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **2002,** 11(4), 739-756.

28. Dunker, A. K.; Oldfield, C. J.; Meng, J.; Romero, P.; Yang, J. Y.; Chen, J. W.; Vacic, V.; Obradovic, Z.; Uversky, V. N. The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* **2008,** 9, S1.

29. Uversky, V. N. Intrinsically disordered proteins from A to Z. Int. *J. Biochem. Cell. Biol.* **2011,** 43(8), 1090-1103.

30. Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing protein intrinsic disorder. *Chem. Rev.* **2014,** 114(13), 6561-6588.

31. Das, R. K.; Ruff, K. M.; Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2015,** 32, 102-112.

32. Forman-Kay, J. D.; Mittag, T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* **2013,** 21(9), 1492-1499.

33. Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** 107(18), 8183-8188.

34. Das, R. K.; Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** 110(33), 13392-13397.

35. Tompa, P.; Fuxreiter, M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **2008,** 33(1), 2-8.

36. Hazy, E.; Tompa, P. Limitations of induced folding in molecular recognition by intrinsically disordered proteins. *Chem. Phys. Chem.* **2009,** 10(9-10), 1415-1419.

37. Galea, C. A.; Wang, Y.; Sivakolundu, S. G.; Kriwacki, R. W. Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits. *Biochemistry* **2008,** 47(29), 7598-7609.
38. Liu, J.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K. Intrinsic disorder in transcription factors. *Biochemistry* **2006,** 45(22), 6873-6888.

39. Wright, P. E.; Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015,** 16(1), 18-29.

40. Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **2011,** 21(3), 432-40.

41. Dobson, C. M. The structural basis of protein folding and its links with human disease. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2001,** 356(1406), 133-145.

42. Babu, M. M. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem. Soc. Trans.* **2016,** 44(5), 1185-1200.

43. Uversky, V. N. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front. Aging Neurosci.* **2015,** 7, 18.

44. Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006,** 75, 333-366.

45. Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W. Functional amyloid--from bacteria to humans. *Trends Biochem. Sci.* **2007,** May, 32(5), 217-224.

46. Otzen, D. Functional amyloid: turning swords into plowshares. *Prion* **2010,** 4(4), 256-264.

47. Barnhart, M. M.; Chapman, M. R. Curli biogenesis and function. *Annu. Rev. Microbiol.* **2006,** 60, 131-147.

48. Brangwynne, C. P. Phase transitions and size scaling of membrane-less organelles. *J. Cell Biol.* **2013,** 203(6), 875-881.

49. Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **2014,** 30, 39-58.

50. Mitrea, D. M.; Kriwacki, R. W. Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **2016,** 14, 1.

51. Patel, A.; Lee, H. O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M. Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T. M.; Pozniakovski, A.; Poser, I.; Maghelli, N.; Royer, L. A.; Weigert, M.; Myers, E. W.; Grill, S.; Drechsel, D.; Hyman, A. A.; Alberti, S. A. Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015,** 162(5), 1066-1077.

52. Wegmann, S.; Eftekharzadeh, B.; Tepper, K.; Zoltowska, K. M.; Bennett, R. E.; Dujardin, S.; Laskowski, P. R.; MacKenzie, D.; Kamath, T.; Commins, C.; Vanderburg, C.; Roe, A. D.; Fan, Z.; Molliex, A. M.; Hernandez-Vega, A.; Muller, D.; Hyman, A. A.; Mandelkow, E.; Taylor, J. P.; Hyman, B. T. Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J.* **2018,** 37(7), e98049.

53. Alberti, S.; Hyman, A. A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **2021,** 22(3), 196-213.

54. Fuxreiter, M.; Vendruscolo, M. Generic nature of the condensed states of proteins. *Nat. Cell Biol.* **2021,** 23(6), 587-594.

55. Sabari, B. R.; Dall'Agnese, A.; Young, R. A. Biomolecular Condensates in the Nucleus. *Trends Biochem. Sci.* **2020,** 45(11), 961-977.

56. Boeynaems, S.; Alberti, S.; Fawzi, N. L.; Mittag, T.; Polymenidou, M.; Rousseau, F.; Schymkowitz, J.; Shorter, J.; Wolozin, B.; Van Den Bosch, L.; Tompa, P.; Fuxreiter, M. Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol.* **2018,** 28(6), 420-435.

57. Forman-Kay, J. D.; Kriwacki, R. W.; Seydoux, G. Phase Separation in Biology and Disease. *J. Mol. Biol.* **2018,** 430(23), 4603-4606.

58. Darling, A. L.; Liu, Y.; Oldfield, C. J.; Uversky, V. N. Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **2018,** 18(5-6), e1700193.

59. Shin, Y.; Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **2017,** 357(6357), eaaf4382.

60. Banani, S. F.; Rice, A. M.; Peeples, W. B.; Lin, Y.; Jain, S.; Parker, R.; Rosen, M. K. Compositional Control of Phase-Separated Cellular Bodies. *Cell* **2016,** 166(3), 651-663.

61. Nakashima, K. K.; Vibhute, M. A.; Spruijt, E. Biomolecular Chemistry in Liquid Phase Separated Compartments. *Front. Mol. Biosci.* **2019,** 6, 21.

62. Posey, A. E.; Holehouse, A. S.; Pappu, R. V. Phase Separation of Intrinsically Disordered Proteins. *Methods Enzymol.* **2018,** 611, 1-30.

63. Wang, J.; Choi, J. M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; Poser, I.; Pappu, R. V.; Alberti, S.; Hyman, A. A. A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell* **2018,** 174(3), 688-699.

64. Martin, E. W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **2018,** 57(17), 2478-2487.

65. Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. *Annu. Rev. Phys. Chem.* **2020,** 71, 53-75.

66. Gomes, E.; Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem.* **2019,** 294(18), 7115-7127.

67. Pak, C. W.; Kosno, M.; Holehouse, A. S.; Padrick, S. B.; Mittal, A.; Ali, R.; Yunus, A. A.; Liu, D. R.; Pappu, R. V.; Rosen, M. K. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell.* **2016,** 63(1), 72-85.

68. Martin, E. W.; Holehouse, A. S.; Peran, I.; Farag, M.; Incicco, J. J.; Bremer, A.; Grace, C. R.; Soranno, A.; Pappu, R. V.; Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020,** 367(6478), 694-699.

69. Qamar, S.; Wang, G.; Randle, S. J.; Ruggeri, F. S.; Varela, J. A, Lin, J. Q.; Phillips, E. C.; Miyashita, A.; Williams, D.; Ströhl, F.; Meadows, W.; Ferry, R.; Dardov, V. J.; Tartaglia, G. G.; Farrer, L. A.; Kaminski, Schierle, G. S.; Kaminski, C. F.; Holt, C. E.; Fraser, P. E.; Schmitt-Ulms, G.; Klenerman, D.; Knowles, T.; Vendruscolo, M.; St George-Hyslop, P. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. *Cell* **2018,** 173(3), 720-734.

70. Zhou, H. X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018,** 118(4), 1691-1741.

71. Banerjee, P. R.; Milin, A. N.; Moosa, M. M.; Onuchic, P. L.; Deniz, A. A. Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angew. Chem. Int. Ed. Engl.* **2017,** 56(38), 11354-11359.

72. Maharana, S.; Wang, J.; Papadopoulos, D. K.; Richter, D.; Pozniakovsky, A.; Poser, I.; Bickle, M.; Rizk, S.; Guillén-Boixet, J.; Franzmann, T. M.; Jahnel, M.; Marrone, L.; Chang, Y. T.; Sterneckert, J.; Tomancak, P.; Hyman, A. A.; Alberti, S. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **2018,** 360(6391), 918-921.

73. Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nature Phys.* **2015,** 11, 899–904.

74. Dignon, G. L.; Zheng, W.; Kim, Y.C.; Best, R.B.; Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS. Computational Biology* **2018,** 14, 1005941.

75. Murthy, A. C.; Dignon, G. L.; Kan, Y.; Zerze, G. H.; Parekh, S. H.; Mittal, J.; Fawzi, N. L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019,** 26(7), 637-648.

76. Luo, F.; Gui, X.; Zhou, H.; Gu, J.; Li, Y.; Liu, X.; Zhao, M.; Li, D.; Li, X.; Liu, C. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **2018,** 25(4), 341-346.

77. Rai, S. K.; Savastano, A.; Singh, P.; Mukhopadhyay, S.; Zweckstetter, M. Liquid–Liquid Phase Separation of Tau: From Molecular Biophysics to Physiology and Disease. *Protein Sci.* **2021,** 30, 1294–1314.

78. Portz, B.; Lee, B. L.; Shorter, J. FUS and TDP-43 phases in health and disease. *Trends Biochem. Sci.* **2021,** 46, 550–563.

79. Aguzzi, A.; Altmeyer, M. Phase separation: Linking cellular compartmentalization to disease. *Trends Cell Biol.* **2016,** 26, 547–558.

80. Agarwal, A.; Rai, S. K.; Avni, A.; Mukhopadhyay, S. An intrinsically disordered pathological prion variant Y145Stop converts into self-seeding amyloids via liquid-liquid phase separation. *Proc. Natl. Acad. Sci. U. S. A.* **2021,** 118(45), e2100968118.

81. Alberti, S.; Dormann, D. Liquid-Liquid Phase Separation in Disease. *Annu. Rev. Genet.* **2019,** 53, 171-194.

82. Elbaum-Garfinkle, S. Matter over mind: Liquid phase separation and neurodegeneration. *J. Biol. Chem.* **2019,** 294(18), 7160-7168.

83. Netherton, C. L.; Wileman, T. Virus factories, double membrane vesicles and viroplasm generated in animal cells. *Curr. Opin. Virol.* **2011,** 1(5), 381-387.

84. Novoa, R. R.; Calderita, G.; Arranz, R.; Fontana, J.; Granzow, H.; Risco, C. Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol. Cell* **2005,** 97(2), 147-172.

85. Labbadia, J.; Morimoto, R. I. The biology of proteostasis in aging and disease. *Annu. Rev. Biochem.* **2015,** 84, 435-464.

86. Hartl, F. U. Cellular Homeostasis and Aging. *Annu. Rev. Biochem.* **2016,** 85, 1-4.

87. Mitrea, D. M.; Chandra, B.; Ferrolino, M. C.; Gibbs, E. B.; Tolbert, M.; White, M. R.; Kriwacki, R. W. Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles. *J. Mol. Biol.* **2018,** 430(23), 4773-4805.

88. Ganser, L. R.; Myong, S. Methods to Study Phase-Separated Condensates and the Underlying Molecular Interactions. *Trends Biochem. Sci.* **2020,** 45(11), 1004-1005.

89. Elbaum-Garfinkle, S.; Kim, Y.; Szczepaniak, K.; Chen, C. C.; Eckmann, C. R.; Myong, S.; Brangwynne, C. P. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** 112(23), 7189-7194.

90. Murthy, A. C.; Fawzi, N. L. The (un)structural biology of biomolecular liquid-liquid phase separation using NMR spectroscopy. *J. Biol. Chem.* **2020,** 295(8), 2375-2384.

91. Brady, J. P.; Farber, P. J.; Sekhar, A.; Lin, Y. H.; Huang, R.; Bah, A.; Nott, T. J.; Chan, H. S.; Baldwin, A. J.; Forman-Kay, J. D.; Kay, L. E. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** 114(39), E8194-E8203.

92. Niaki, A. G.; Sarkar, J.; Cai, X.; Rhine, K.; Vidaurre, V.; Guy, B.; Hurst, M.; Lee, J. C.; Koh, H. R.; Guo, L.; Fare, C. M.; Shorter, J.; Myong, S. Loss of Dynamic RNA Interaction and Aberrant Phase Separation Induced by Two Distinct Types of ALS/FTD-Linked FUS Mutations. *Mol. Cell.* **2020,** 77(1), 82-94.

93. Jain, A.; Liu, R.; Xiang, Y. K.; Ha, T. Single-molecule pull-down for studying protein interactions. *Nat. Protoc.* **2012,** 7(3), 445-452.

94. Mitrea, D. M.; Cika, J. A.; Guy, C. S.; Ban, D.; Banerjee, P. R.; Stanley, C. B.; Nourse, A.; Deniz, A. A.; Kriwacki, R. W. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* **2016,** 5, e13571.

95. Wei, M. T.; Elbaum-Garfinkle, S.; Holehouse, A. S.; Chen, C. C.; Feric, M.; Arnold, C. B.; Priestley, R. D.; Pappu, R. V.; Brangwynne, C. P. Phase behaviour of disordered proteins

underlying low density and high permeability of liquid organelles. *Nat. Chem.* **2017,** 9(11), 1118-1125.

96. Murakami, K.; Kajimoto, S.; Shibata, D.; Kuroi, K.; Fujii, F.; Nakabayashi, T. Observation of liquid-liquid phase separation of ataxin-3 and quantitative evaluation of its concentration in a single droplet using Raman microscopy. *Chem. Sci.* **2021,** 12(21), 7411-7418.

97. Yokosawa, K.; Kajimoto, S.; Shibata, D.; Kuroi, K.; Konno, T.; Nakabayashi, T. Concentration Quantification of the Low-Complexity Domain of Fused in Sarcoma inside a Single Droplet and Effects of Solution Parameters. *J. Phys. Chem. Lett.* **2022,** 13(24), 5692-5697.

98. Shuster, S. O.; Lee, J. C. Watching liquid droplets of TDP-43_{CTD} age by Raman spectroscopy. *J. Biol. Chem.* **2022,** 298(2), 101528.

99. Eisenberg, D. S.; Sawaya, M. R. Structural Studies of Amyloid Proteins at the Molecular Level. *Annu. Rev. Biochem.* **2017,** 86, 69-95.

100. Sunde, M.; Blake, C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Protein Chem.* **1997,** 50, 123-159.

101. Knowles, T. P.; Fitzpatrick, A. W.; Meehan, S.; Mott, H. R.; Vendruscolo, M.; Dobson, C. M.; Welland, M. E. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **2007,** 318(5858), 1900-1903.

102. Toyama, B. H.; Weissman, J. S. Amyloid structure: conformational diversity and consequences. *Annu. Rev. Biochem.* **2011,** 80, 557-585.

103. Taylor, A. I. P, Staniforth, R. A. General Principles Underpinning Amyloid Structure. *Front. Neurosci.* **2022,** 16, 878869.

104. Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **1997,** 273(3), 729-739.

105. Shewmaker, F.; McGlinchey, R. P.; Wickner, R. B. Structural insights into functional and pathological amyloid. *J. Biol. Chem.* **2011,** 286(19), 16533-16540.

106. Arosio, P.; Knowles, T. P.; Linse, S. On the lag phase in amyloid fibril formation. *Phys. Chem. Chem. Phys.* **2015,** 17(12), 7606-7618.

107. Frieden, C. Protein aggregation processes: In search of the mechanism. *Protein Sci.* **2007,** 16(11), 2334-2344.

108. Kumar, S.; Udgaonkar, J. B. Mechanisms of amyloid fibril formation by proteins. *Current Science* **2010,** 98(5), 639–656.

109. Iadanza, M. G.; Jackson, M. P.; Hewitt, E. W.; Ranson, N. A.; Radford, S. E. A new era for understanding amyloid structures and disease. *Nat. Rev. Mol. Cell Biol.* **2018,** 19(12), 755-773.

110. Linse, S. Monomer-dependent secondary nucleation in amyloid formation. *Biophys. Rev.* **2017,** 9(4), 329-338.

111. Törnquist, M.; Michaels, T. C. T.; Sanagavarapu, K.; Yang, X.; Meisl, G.; Cohen, S. I. A.; Knowles, T. P. J.; Linse, S. Secondary nucleation in amyloid formation. *Chem. Commun. (Camb).* **2018,** 54(63), 8667-8684.

112. LeVine, H. Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol.* **1999,** 309, 274-284.

113. Greenwald, J.; Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **2010,** 18(10), 1244-1260.

114. Fändrich, M.; Meinhardt, J.; Grigorieff, N. Structural polymorphism of Alzheimer Abeta and other amyloid fibrils. *Prion* **2009,** 3(2), 89-93.

115. Hu, K. N.; McGlinchey, R. P.; Wickner, R. B.; Tycko, R. Segmental polymorphism in a functional amyloid. *Biophys. J.* **2011,** 101(9), 2242-2250.

116. Cendrowska, U.; Silva, P. J.; Ait-Bouziad, N.; Müller, M.; Guven, Z. P.; Vieweg, S.; Chiki, A.; Radamaker, L.; Kumar, S. T.; Fändrich, M.; Tavanti, F.; Menziani, M. C.; Alexander-Katz, A.; Stellacci, F.; Lashuel, H. A. Unraveling the complexity of amyloid polymorphism using gold nanoparticles and cryo-EM. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** 117(12), 6866-6874.

117. Tycko, R. Physical and structural basis for polymorphism in amyloid fibrils. *Protein Sci.* **2014,** 23(11), 1528-1539.

118. Seuring, C.; Verasdonck, J.; Ringler, P.; Cadalbert, R.; Stahlberg, H.; Böckmann, A.; Meier, B. H.; Riek, R. Amyloid Fibril Polymorphism: Almost Identical on the Atomic Level, Mesoscopically Very Different. *J. Phys. Chem. B.* **2017,** 121(8), 1783-1792.

119. Li, B.; Ge, P.; Murray, K. A.; Sheth, P.; Zhang, M.; Nair, G.; Sawaya, M. R.; Shin, W. S.; Boyer, D. R.; Ye, S.; Eisenberg, D. S.; Zhou, Z. H.; Jiang, L. Cryo-EM of Full-Length α-Synuclein Reveals Fibril Polymorphs with a Common Structural Kernel. *Nat. Commun.* **2018,** 9(1), 3609.

120. Gath, J.; Bousset, L.; Habenstein, B.; Melki, R.; Böckmann, A.; Meier, B. H. Unlike Twins: An NMR Comparison of Two α-Synuclein Polymorphs Featuring Different Toxicity. *PLoS One.* **2014,** 9(3), e90659.

121. Nath, A.; Sammalkorpi, M.; DeWitt, D. C.; Trexler, A. J.; Elbaum-Garfinkle, S.; O'Hern, C. S.; Rhoades, E. The Conformational Ensembles of α-Synuclein and Tau: Combining Single-Molecule FRET and Simulations. *Biophys. J.* **2012,** 103(9), 1940-1949.

122. Li, Y.; Zhao, C.; Luo, F.; Liu, Z.; Gui, X.; Luo, Z.; Zhang, X.; Li, D.; Liu, C.; Li, X. Amyloid Fibril Structure of α-Synuclein Determined by Cryo-Electron Microscopy. *Cell Res.* **2018,** 28(9), 897-903.

123. Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; Willbold, D.; Schröder, G. F. Fibril Structure of Amyloid-β(1-42) by Cryo-Electron Microscopy. *Science* **2017,** 358(6359), 116-119.

124. Falcon, B.; Zhang, W.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Vidal, R.; Crowther, R. A.; Ghetti, B.; Scheres, S. H. W.; Goedert, M. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature* **2018,** 561(7721), 137-140.

125. Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017,** 547(7662), 185-190.

126. Zielinski, M.; Röder, C.; Schröder, G. F. Challenges in sample preparation and structure determination of amyloids by cryo-EM. *J. Biol. Chem.* **2021,** 297(2), 100938.

127. Raman, V.; Krishnan, K. S. A New Type of Secondary Radiation *Nature* **1928,** 121, 501-502.

128. Smith.; Dent, G. Modern Raman Spectroscopy: A Practical Approach 1st ed., Wiley (**2005**).

129. Devitt, G.; Howard, K.; Mudher, A.; Mahajan, S. Raman Spectroscopy: An Emerging Tool in Neurodegenerative Disease Research and Diagnosis. *ACS Chem. Neurosci.* **2018,** 9(3), 404-420.

130. Thomas, G. J. Jr. Raman spectroscopy of protein and nucleic acid assemblies. *Annu. Rev. Biophys. Biomol. Struct.* **1999,** 28, 1-27.

131. Carey, P. R. Raman spectroscopy, the sleeping giant in structural biology, awakes. *J. Biol. Chem.* **1999,** 274(38), 26625-26628.

132. Matousek, P.; Stone, N. Emerging concepts in deep Raman spectroscopy of biological tissue. *Analyst* **2009,** 134(6), 1058-1066.

133. Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. Raman Spectroscopy of Proteins: A Review. *J. Raman Spectrosc.* **2013,** 44, 1061-1076.

134. Tuma, R. Raman Spectroscopy of Proteins: From Peptides to Large Assemblies. *J. Raman Spectrosc.* **2005,** 36, 307-319.

135. Singh, S.; Agarwal, A.; Avni, A.; Mukhopadhyay, S. Ultrasensitive Characterization of the Prion Protein by Surface-Enhanced Raman Scattering: Selective Enhancement via Electrostatic Tethering of the Intrinsically Disordered Domain with Functionalized Silver Nanoparticles. *J. Phys. Chem. Lett.* **2021,** 12, 3187-3194.

136. Kuhar, N.; Sil, S.; Umapathy, S. Potential of Raman spectroscopic techniques to study proteins. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2021,** 258, 119712.

137. Lippert, J. L.; Tyminski, D.; Desmeules, P. J. Determination of the secondary structure of proteins by laser Raman spectroscopy. *J. Am. Chem. Soc.* **1976,** 98, 7075–7080.

138. Kuhar, N.; Sil, S.; Verma, T.; Umapathy, S. Challenges in application of Raman spectroscopy to biology and materials. *RSC Adv.* **2018,** 8(46), 25888-25908.

139. Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. Dihedral psi angle dependence of the amide III vibration: a uniquely sensitive UV resonance Raman secondary structural probe. *J. Am. Chem. Soc.* **2001,** 123(47), 11775-11781.

140. Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. Peptide secondary structure folding reaction coordinate: correlation between UV Raman amide III frequency, Psi Ramachandran angle, and hydrogen bonding. *J. Phys. Chem. B.* **2006,** 110(4), 1928-1943.

141. Hernandez, B.; Coic, Y. M.; Pfluger, F.; Kruglik, S. G.; Ghomi, M. All characteristic Raman markers of tryrosine and tyrosinate originate from phenol ring fundamental vibrations. *J. Raman Spectrosc.* **2015,** 47, 212–220.

142. Bhattacharya, M.; Jain, N.; Dogra, P.; Samai, S.; Mukhopadhyay, S. Nanoscopic amyloid pores formed via stepwise protein assembly. *J. Phys. Chem. Lett.* **2013,** 4, 480–485.

143. Schlamadinger, D. E.; Daschbach, M. M.; Gokel, G. W.; Kim, J. E. UV resonance Raman study of cation-π interactions in an indole crown ether. *J. Raman Spectrosc.* **2011,** 42(4), 633-638.

144. Milan-Garces, E. A.; Mondal, S.; Udgaonkar, J. B.; Puranik, M. Intricate packing in the hydrophobic core of barstar through a CH- π interaction. *J. Raman Spectrosc.* **2014,** 45, 814–821.

145. Benevides, J. M.; Overman, S. A.; Thomas, G. J. Raman spectroscopy of proteins. *Curr. Protoc. Protein Sci.* **2004,** Chapter 17, Unit 17.8.

146. Chatterjee, S.; Kan, Y.; Brzezinski, M.; Koynov, K.; Regy, R. M.; Murthy, A. C.; Burke, K. A.; Michels, J. J.; Mittal, J.; Fawzi, N. L.; Parekh, S. H. Reversible Kinetic Trapping of FUS Biomolecular Condensates. *Adv. Sci. (Weinh).* **2022,** 9(4), e2104247.

147. Feliu, N.; Hassan, M.; Garcia Rico, E.; Cui, D.; Parak, W.; Alvarez-Puebla, R. SERS Quantification and Characterization of Proteins and Other Biomolecules. *Langmuir* **2017,** 33(38), 9711-9730.

148. Langer, J.; Jimenez de Aberasturi, D.; Aizpurua, J.; Alvarez-Puebla, R. A. *et al.* Present and Future of Surface-Enhanced Raman Scattering. *ACS Nano.* **2020,** 14(1), 28-117.

149. Zong, C.; Xu, M.; Xu, L. J.; Wei, T.; Ma, X.; Zheng, X. S.; Hu, R.; Ren, B. Surface-Enhanced Raman Spectroscopy for Bioanalysis: Reliability and Challenges. *Chem. Rev.* **2018,** 118(10), 4946-4980.

150. Bruzas, I.; Lum, W.; Gorunmez, Z.; Sagle, L. Advances in surface-enhanced Raman spectroscopy (SERS) substrates for lipid and protein characterization: sensing and beyond. *Analyst* **2018,** 143(17), 3990-4008.

151. Fleischmann, M.; Hendra, P. J.; McQuillan, A. J. Raman spectra of pyridine adsorbed at a silver electrode. *Chem. Phys. Lett.* **1974,** 26, 163–166.

152. Jeanmaire, D. L.; Van Duyne, R. P. Surface Raman spectroelectrochemistry: Part I. Heterocyclic, aromatic, and aliphatic amines adsorbed on the anodized silver electrode. *J. Electroanal. Chem. Interfacial Electrochem.* **1977,** 84, 1–20.

153. King, F. W.; Van Duyne, R. P.; Schatz, G. C. Theory of Raman scattering by molecules adsorbed on electrode surfaces. *J. Chem. Phys.* **1978,** 69, 4472–4481.

154. Albrecht, M. G.; Creighton, J. A. Anomalously intense Raman spectra of pyridine at a silver electrode. *J. Am. Chem. Soc.* **1977,** 99, 5215-5217.

155. Aggarwal, S.; Mondal, S.; Siddhanta, S.; Bharat, E.; Nagamalleswari, E.; Nagaraja, V.; Narayana, C. Divalent Ion-Induced Switch in DNA Cleavage of KpnI Endonuclease Probed through Surface-Enhanced Raman Spectroscopy. *J. Phys. Chem. B.* **2021,** 125(9), 2241-2250.

156. Willets, K.A.; Van Duyne, R. P. Localized surface plasmon resonance spectroscopy and sensing. *Annu. Rev. Phys. Chem.* **2007,** 58, 267–297.

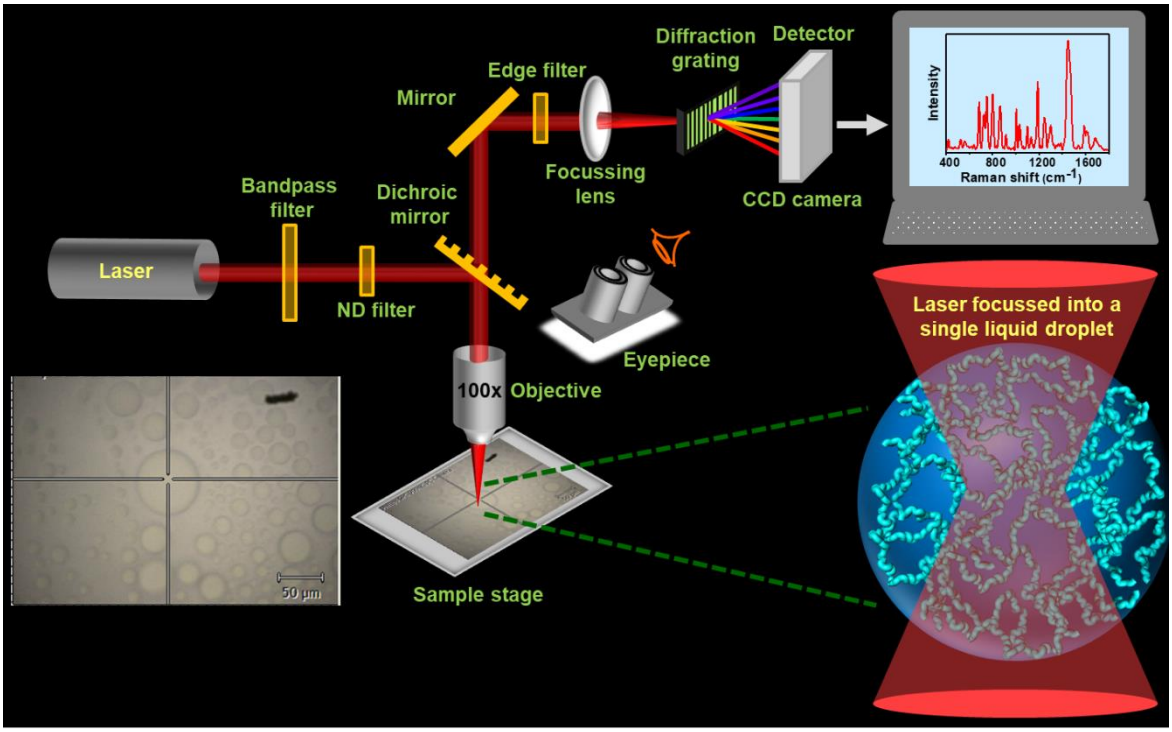
157. Campion, A.; Kambhampati, P. Surface-enhanced Raman scattering. *Chem. Soc. Rev.* **1998,** 27, 241-250.

158. Siddhanta, S.; Narayana, C. Surface Enhanced Raman Spectroscopy of Proteins: Implications for Drug Designing. *Nanomaterials and Nanotechnology* **2012,** 2.

159. Cong, S.; Liu, X.; Jiang, Y.; Zhang, W.; Zhao, Z. Surface Enhanced Raman Scattering Revealed by Interfacial Charge-Transfer Transitions. *Innovation (Camb).* **2020,** 1(3), 100051.

******

Probing the Conformational Heterogeneity and Structural Distribution Within Biomolecular Condensates Using Single-Droplet Vibrational Raman Spectroscopy



The work described in this chapter has been published in the *Proceedings of the National Academy of Sciences of the United States of America* and *Nature Communications*

Agarwal, A.; Rai, S. K.; Avni, A.; Mukhopadhyay, S. An Intrinsically Disordered Pathological Variant of the Prion Protein Y145Stop Transforms into Self-Templating Amyloids via Liquid-Liquid Phase Separation. *Proc. Natl. Acad. Sci. U. S. A.* **2021,** 118(45), e2100968118.

Avni. A.; Joshi. A.; Walimbe. A.; Pattanashetty. S. G.; Mukhopadhyay. S. Single-Droplet Surface-Enhanced Raman Scattering Decodes the Molecular Determinants of Liquid-Liquid Phase separation. *Nat. Commun.* **2022,** 13, 4378.

2.1 Introduction

Eukaryotic cells perform thousands of complex biochemical reactions within their distinct membrane-bound organelles like the nucleus, mitochondria, lysosomes, and so on. In addition to these canonical membrane-bound organelles, there are dozens of highly concentrated, non-stoichiometric, mesoscopic, and dynamic proteins and/or nucleic acid-rich membrane-less organelles such as the nucleolus, P granules, nuclear speckles, and so forth. These compact liquid-like supramolecular species or biomolecular condensates are proposed to be formed via liquid-liquid phase separation (LLPS).[1-13] A current flurry of research suggests that intrinsically disordered proteins/regions (IDPs/IDRs) along with a few repetitive, low-complexity (LC) and prion-like regions are excellent candidates for the process of intracellular phase separation. These features in the polypeptide chain offer an array of multivalent, heterogeneous inter and intrachain, noncovalent interactions such as electrostatic, hydrophobic, hydrogen bonding, dipole-dipole, π-π, and cation-π, thereby imparting conformational flexibility and liquid-like properties such as fusion, dripping, and surface wetting to these biomolecular condensates.[14-23] Many of these varied interactions provide the transient sticker-spacer contacts or "Fuzzy" interactions, which thermodynamically favor these condensate formations. However, the molecular origin and the sequence of events that regulates intracellular phase transition into these liquid-like compartments remain elusive. There are a multitude of methods developed recently to study the key biophysical principles governing the condensate formation and their bulk properties.[24-32] Techniques like bright-field imaging or fluorescence microscopy, high-speed atomic force microscopy (AFM), and fluorescence recovery after photobleaching (FRAP) can directly probe the droplet properties like their number, size, morphology, and diffusion within the condensed phase *in vitro* and *in vivo*.[33,34] While it is possible to examine the effect of various factors altering LLPS and determine the multiphase architecture by multicolor labeling techniques, however, these experiments fail to provide information at the molecular level, and also use of certain fluorescent probes can derange the phase-separating systems. On the other hand, NMR coupled with small-angle X-ray scattering and small-angle neutron scattering can also provide a wealth of information at the atomic level.[35-39] However, all the above-mentioned experiments are ensemble methods giving information from a collection of molecules in the light and dense phase and often require highly concentrated samples with high purity.

In this direction, Raman spectroscopy proves to be a powerful technique that can shed enormous light on the 3-dimensional structure of proteins, intramolecular dynamics, and

intermolecular interactions. It is a non-contact, label-free, non-destructive tool to study protein conformations in physiological conditions with minimum sample preparation.[40-44] A typical protein Raman spectrum is dominated by amide bands, amide I and amide III, as well as bands due to different vibrational modes of aromatic amino acids like Phenylalanine, Tyrosine, Tryptophan in addition to several other sidechain vibrations. Amide I (1620-1700 cm^{-1}) and amide III (1220-1300 cm^{-1}) bands are the secondary structural marker bands and are observed due to C=O stretching vibrations and C-N stretching, N-H bending vibrations of the polypeptide backbone, respectively.[45-48] We utilized this potent tool to study the conformational heterogeneity and structural distributions within several phase-separated liquid droplets in a single droplet-by-droplet fashion for two interesting phase-separating systems, PrP-Y145Stop and FUS (Fused In Sarcoma).[49,50]

Human prion protein (PrP) undergoes a conformational conversion to a β-rich, self-replicating, misfolded amyloid-like form that is associated with a range of lethal transmissible neurodegenerative diseases referred to as transmissible spongiform encephalopathies.[51,52] Cellular PrP is a 253-residue long cell-surface binding (Glycosylphosphatidylinositol) GPI-anchored protein which consists of an N-terminal signal peptide (residues 1-23), an unstructured N-terminal (residues 23-120) domain, a structured C-terminal domain (residues 121-231), and a GPI-anchor signal (residues 231-253).[51-54] The unstructured N-terminal tail is positively charged and comprises two lysine clusters (residues 23-30 and 100-110), five glycine-rich octapeptide repeats (PHGGGWGQ), and a hydrophobic segment (residues 113-135) and can be classified as an IDR. The highly structured C-terminal domain consists of three α-helices (residues 144-154, 175-193, and 200-219) and two short antiparallel β-strands (residues 128-131 and 161-164). An unstructured disease-associated stop-codon mutation of PrP at tyrosine 145 (Y145Stop) results in a C-terminal truncated, N-terminal IDR, which under normal cellular conditions is unstable and is rapidly degraded in the cell via the proteasomal machinery. This intrinsically disordered pathological stop codon variant of the prion protein, namely, Y145Stop, is associated with Gerstmann-Sträussler-Scheinker syndrome and familial cerebral amyloid angiopathy.[55-57] Under stress conditions, this fascinating mutant of the prion protein accumulates in the endoplasmic reticulum (ER), Golgi, and nucleus.[58,59] Recently, we showed in our lab that Y145Stop spontaneously phase separates into dynamic liquid-like droplets, which gradually undergo liquid-to-solid transitions to form more ordered, β-rich, amyloid-like aggregates.[49]

Full-length FUS is amongst the most intensively studied prion-like RNA-binding protein and hence becomes the best prototype for these structural studies. The human genome encodes approximately 30 FUS family proteins, including the FET family of proteins comprising of FUS, EWSR1 (Ewing sarcoma breakpoint 1), and TAF15 (TATA-box binding protein associated factor 15), which are known to be involved in both function and dysfunction such as mRNA splicing, DNA damage repair, formation of stress granules as well as neurodegeneration in Amyotrophic lateral sclerosis (ALS) and Frontotemporal dementia (FTD).[60-67]

2.2 Experimental details

2.2.1 Materials

Sodium phosphate monobasic dihydrate, potassium chloride, sodium phosphate dibasic dihydrate, sodium chloride, sodium hydroxide, 2-mercaptoethanol, 1,4-dithiothreitol (DTT), tris base, L-glutathione reduced, triton X-100, Thrombin from bovine plasma, Poly-U sodium salt, phenylmethylsulfonyl fluoride (PMSF), magnesium chloride hexahydrate, Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), ethylenediaminetetraacetic acid (EDTA), were of MB grade purity, procured from Sigma (St. Louis, MO, USA). Luria Bertani Broth, Miller (LB), zinc chloride, nickel chloride, Lysozyme, HEPES, and maltose were procured from HiMedia laboratories (MB grade). Ampicillin, chloramphenicol, and isopropyl-β-thiogalactopyranoside (IPTG) were ordered from Gold Biocom (USA). Ni-NTA resin and Amylose resin were purchased from Qiagen. PD-10 column was obtained from GE Healthcare Life Sciencs (USA). Amicon membrane filters for concentrating protein were obtained from Merck Millipore. High-purity Milli-Q water was used to prepare all the buffers in this study. A Metrohm 827 lab pH meter was used to adjust pH (± 0.01) of all the buffer solutions prepared at 25 °C and all the buffer solutions were filtered before use.

2.2.2 Recombinant protein expression, purification

Recombinant PrP-Y145Stop was purified by Ni-NTA chromatography as described previously.[49]

Expression conditions and purification of full-length FUS is described in detail in Chapter 3.

2.2.3 Phase separation assay

Phase separation of PrP-Y145Stop was induced by adding 350 mM NaCl to 100 μM of Y145Stop at 37 ºC in 20mM sodium phosphate buffer, pH 7.5. Prolonged incubation of liquid droplets at 37 ºC, transformed these droplets irreversibly to gel-like or solid-like fibrous morphologies that were used for Raman experiments (Phase separation and aggregation experiments were performed in detail by Aishwarya Agarwal and are a part of her thesis work).

Phase separation of FUS was initiated by TEV cleavage in a 1:10 molar ratio (TEV: protein) at room temperature in 20 mM sodium phosphate buffer, pH 7.4. The turbidity of phase-separated samples (protein concentration 20 μM) was then estimated using 96-well NUNC optical bottom plates (Thermo Scientific) on a Multiskan Go (Thermo Scientific) plate reader by recording the absorbance at 350 nm. The total sample volume used was 100 μL for all the measurements, and then background subtracted turbidity was plotted using Origin.

2.2.4 Raman spectroscopy

All the spectra were recorded on an inVia laser Raman microscope (Renishaw, UK). Raman spectra were recorded for Y145Stop monomers, droplets, and aggregates. The sample volume of 2-3 μL was placed onto a glass slide covered with an aluminum sheet. For single-droplet normal Raman measurements droplet reaction (2 μL) of full-length FUS (20 μM) with or without RNA was placed on a glass slide covered with an aluminum foil and single droplets were focused. The sample was focused using a 100x objective lens (Nikon, Japan), and a 785-nm NIR laser was used for excitation, with an exposure time of 10 s and an appropriate laser power depending on the sample. Single-droplet Raman measurements were performed by focusing the laser on individual droplets. The Rayleigh scattering was filtered by using an edge filter of 785 nm. The Raman scattering was collected and dispersed using a 1200 lines/mm diffraction grating and detected using an air-cooled CCD detector. Inbuilt Wire 3.4 software was used for data acquisition. All the data were averaged over 10 scans. Baseline correction and smoothening of the acquired spectra were performed using Wire 3.4 and the spectra were plotted using Origin.

2.3 Results

2.3.1 Experimental design for single-droplet vibrational Raman spectroscopy

Our laser micro-Raman system consists of several components, namely, an excitation source comprising of a near-infrared (NIR) laser, an integrated microscope spectrometer consisting of

a combined system of lenses, mirrors, filters, and a diffraction grating, and a charge-coupled device (CCD) detector (Figure 2.1). This integrated optical setup allows us to irradiate the sample and filter out the (elastic) Rayleigh scattered light and further collimate the (inelastic) Raman scattered light onto the detector to obtain a Raman spectrum. Such a design permits us to focus the laser beam of a suitable power using an objective lens into a small sub-micron spot-size within a single protein-rich droplet and acquire (regular) normal single-droplet Raman spectra.

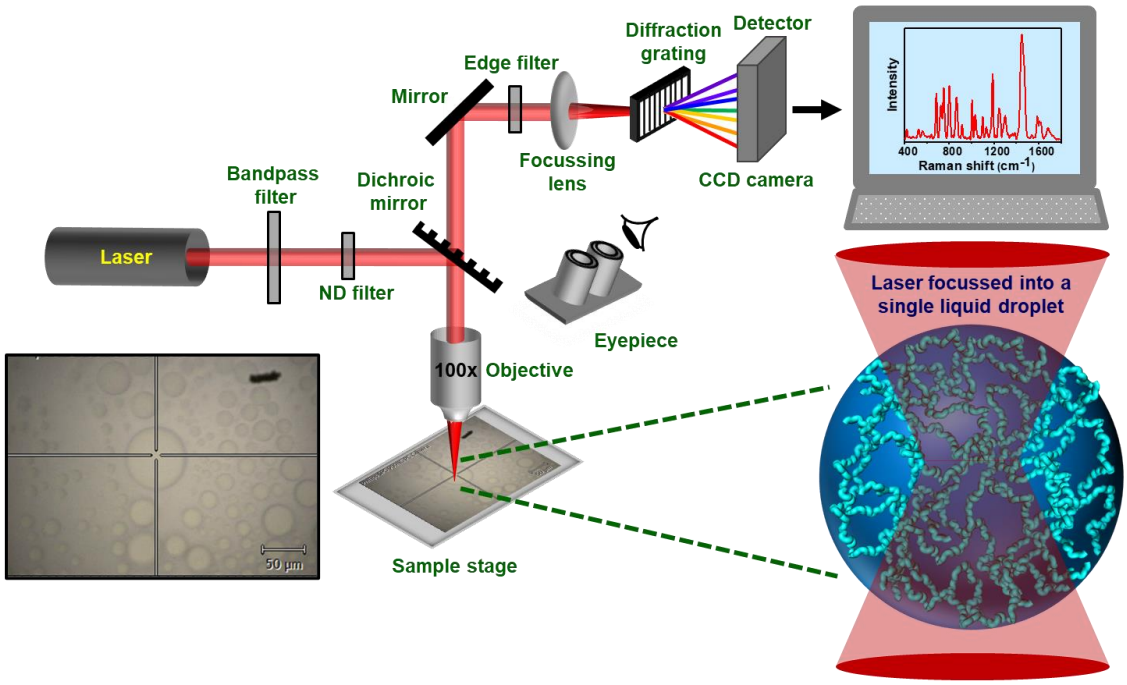


Figure 2.1: Single-droplet vibrational Raman spectroscopy: laser micro-Raman setup. A near-infrared (NIR) laser gets focused within each protein-rich droplet through an integrated system of lenses, mirrors, and filters. Vibrational Raman signals are detected by a CCD detector. The diffraction grating shown here is for representative purpose only.

2.3.2 Conformational heterogeneity and disorder within Y145Stop droplets and aggregates captured by vibrational Raman spectroscopy

In order to elucidate the protein conformational states, we employed vibrational Raman spectroscopy that provides a wealth of molecular information about the polypeptide backbone and sidechains (Figure 2.2). The Raman spectra of dispersed and phase-separated states of Y145Stop clearly showed characteristic bands corresponding to backbone amide I, amide III, Trp, Tyr, Phe, and other vibrational modes. Amide I (1630-1700 cm^{-1}) originate primarily due to the C=O stretching of the polypeptide backbone, whereas amide III (1230-1300 cm^{-1})

corresponds to a combination of N-H bending and C-N stretching modes.[46,47] The amide vibrational bands are often used to identify the secondary structural elements in proteins and are thus referred to as secondary structural marker bands.[68] Broad amide I band for both dispersed and demixed phases exhibited highly disordered conformations (Figure 2.2 A and B). A closer inspection of amide I revealed an increase in full-width at half maximum (FWHM) from the dispersed phase (~ 55 cm^{-1}) to the droplet phase (~ 63 cm^{-1}), indicating a higher conformational heterogeneity in the condensed phase (Figure 2.2 C). Next, we determined the intensity ratio of the tyrosine Fermi doublet (I_{850}/I_{830}) that serves as an indicator of the hydrogen bonding strength between the phenolic hydroxyl group of Tyr and the neighbouring water

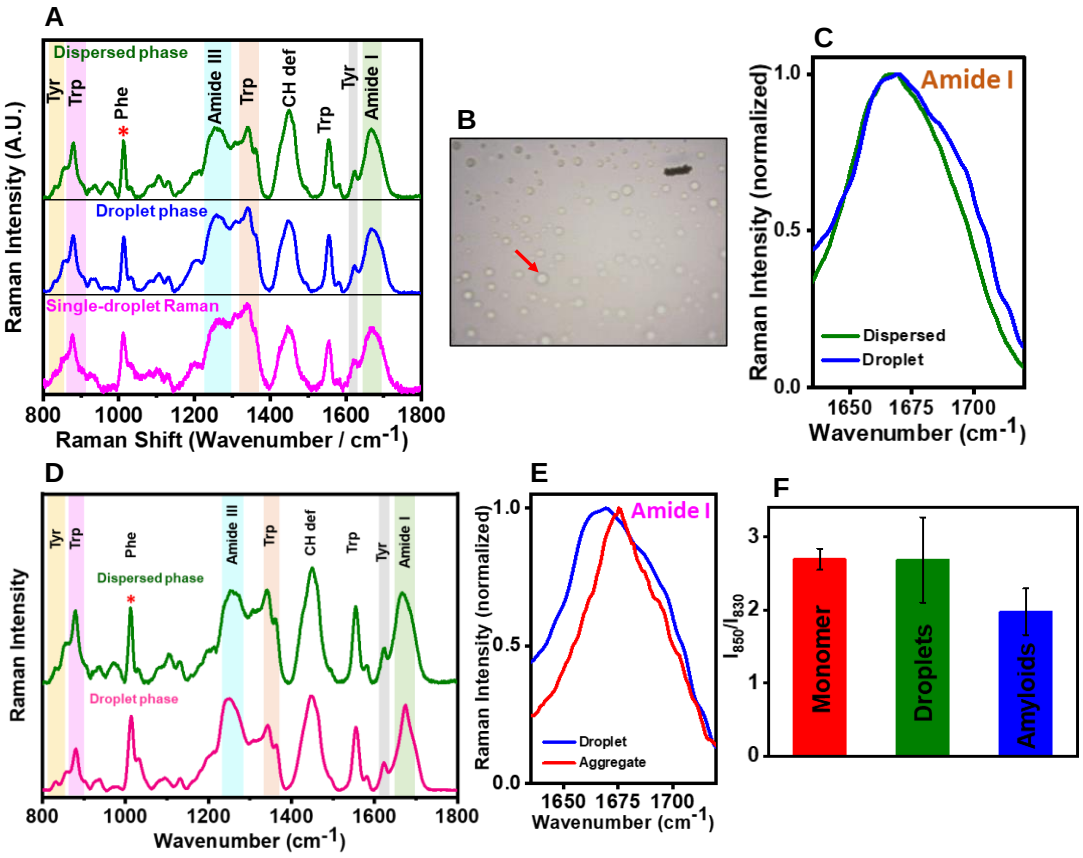


Figure 2.2: (A) Bulk Raman spectra of dispersed (olive) and droplet phases (blue) and single-droplet Raman (magenta). Prominent vibrational signatures of amide backbones and sidechains are shaded and labeled. (B) Single droplets of PrP-Y145Stop were focussed through the Raman microscope. Arrowhead shows the focal point of the NIR laser within the droplet. (C) The amide I regions are shown for dispersed and droplet phases. (D) Vibrational Raman spectra of dispersed monomeric and amyloid aggregates normalized using a sharp band at ~ 1000 cm^{-1} (Phe). (E) Amide I vibrational Raman band shows a narrow peak at 1675 cm^{-1}, indicating an amyloid-like cross-β architecture. (F) The ratio of Tyrosine Fermi doublet (I_{850}/I_{830}) estimated from the peak intensities. The mean ratio and standard error are estimated from at least three independent measurements.

molecules in the vicinity.[69] This ratio was ~ 2.7 for both dispersed and droplets indicating little or no changes in the solvent accessibility of polypeptide chains in the liquid-like condensed phase. The intensity ratio for the tryptophan Fermi doublet (I_{1360}/I_{1340}) (hydrophobicity) and a band at 880 cm^{-1} (hydrogen bonding strength between water and N-H of the indole ring) did not exhibit any measurable changes upon the liquid phase condensation. We were able to recapitulate all these vibrational signatures in our single-droplet Raman studies. Taken together, these results demonstrate that Y145Stop retains its intrinsic disorder with a slight increase in the structural heterogeneity within the demixed liquid droplets.

These droplets of Y145Stop irreversibly transformed into gel-like or solid-like fibrous species after ~5 h. The fibrous species formed bind to thioflavin-T (ThT), a well-known amyloid reporter indicating their amyloid-like nature. "Sea urchin"-like morphologies were observed as reported previously for FUS aggregates formed via LLPS.[67] These observations indicate that the liquid droplets can perhaps act as nucleation centers for the growth of amyloid-like fibers. Our vibrational Raman spectroscopic investigation showed a sharp amide I at 1675 cm^{-1}, which is a hallmark of hydrogen-bonded cross-β architecture in amyloid fibrils (Figure 2.2 D). The FWHM of amide I considerably narrowed down from liquid droplets (~ 63 cm^{-1}) to aggregates (~ 44 cm^{-1}), indicating a more ordered and less heterogeneous conformational state (Figure 2.2 E). The ratio of the tyrosine Fermi doublet (I_{850}/I_{830}) decreased considerably upon liquid-to-solid transition from droplets (~ 2.7) to aggregates (~ 2.0), suggesting a buried environment of Tyr residues with a lower propensity to form hydrogen bonds with water (Figure 2.2 F). Additionally, the polarity around Trp also exhibited a small decrease, as evidenced by a slight blue shift of the band around 880 cm^{-1}, indicating a lower hydrogen bonding ability of N-H of the indole ring with water within the aggregates.

2.3.3 Phase-separation of full-length FUS

FUS consists of an intrinsically disordered low-complexity (LC) prion-like domain (residues 1-163) and a C-terminal RNA-binding domain (RBD) containing both disordered and α-helical secondary structural elements (residues 267-526). The RBD contains two RGG-rich stretches, an RNA-recognition motif (RRM), and a zinc finger domain and carries a net-positive charge (+10.8) at physiological pH (Figure 2.3). The N-terminal disordered LC-domain is the primary driver of phase separation, and the presence of the RBD greatly enhances the propensity for

phase transitions under physiological conditions. We induced LLPS of FUS by cleaving the N-terminal maltose binding protein (MBP) tag using TEV protease, that resulted in the condensation of a homogeneously mixed phase into a demixed phase, as reported previously.[65] The condensed droplet phase was devoid of cleaved MBP, as observed before (Figure 2.4).

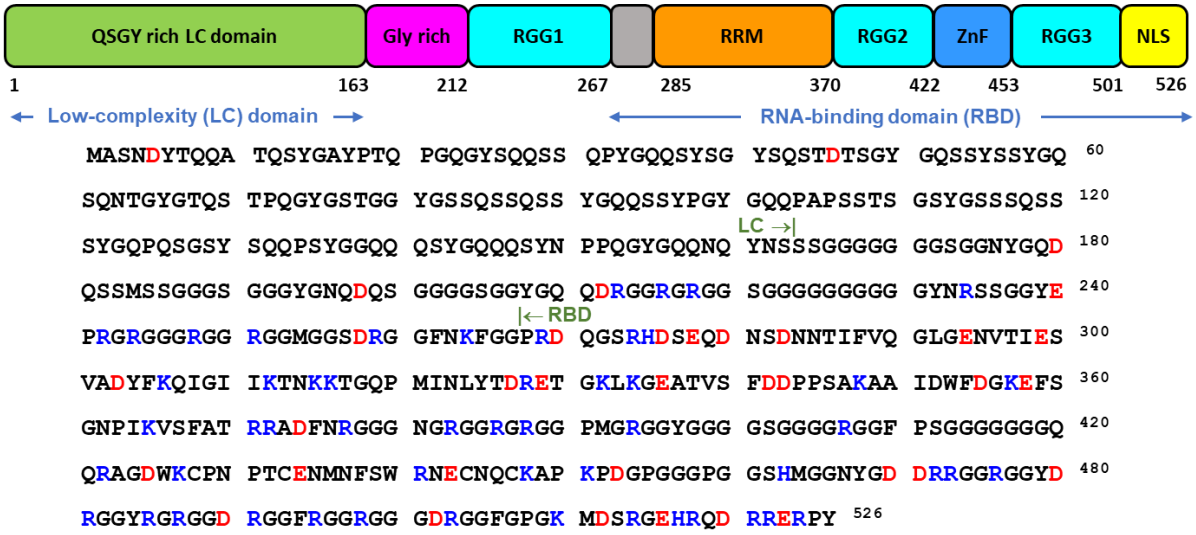


Figure 2.3: Schematic representation of full-length FUS showing all the segments and domains. Amino acid sequence of full-length FUS is shown below. The positively and negatively charged residues are highlighted in blue and red, respectively.

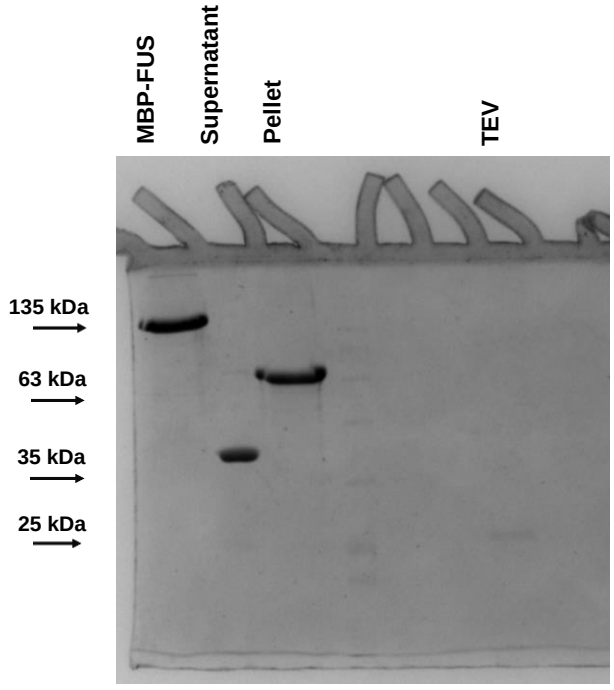


Figure 2.4: SDS-PAGE (12%) depicting that the condensed phase is devoid of the maltose-binding protein (MBP) tag.

2.3.4 Single-droplet normal Raman spectroscopy of FUS condensates

We carried out (regular) normal single-droplet Raman spectroscopy by focusing a high-power (500 mW) 785 nm laser beam using a 100x objective. These studies allowed us to characterize the condensed phase by recording the Raman scattering bands for different vibrational modes of the polypeptide chains in a droplet-by-droplet manner (Figure 2.5). The condensed phase showed a broad amide I band centered at ~ 1671 cm^{-1} with a full width at half maximum (FWHM) ~ 59 cm^{-1} representing a considerable conformational distribution. This was further supported by the amide III band at ~ 1262 cm^{-1} representing highly disordered conformers within the condensed phase. Further, to decode the sidechain environment, we inspected the intensity ratio at 850 cm^{-1} and 830 cm^{-1} (I_{850}/I_{830}) of the tyrosine Fermi doublet that is observed due to Fermi resonance between the ring-breathing vibration and overtone of an out-of-plane ring-bending vibration of the phenolic ring of tyrosine.[69] Therefore, this ratio is an indicator of the solvent-mediated hydrogen bonding propensity of the phenolic (-OH) group and is a measure of the water accessibility of tyrosine residues. The I_{850}/I_{830} ratio is typically ≥ 2 for a well-solvated tyrosine, and this ratio was found to be ~ 0.5 for FUS droplets indicating considerable solvent protection, possibly due to the participation of tyrosine residues in π-π stacking and/or cation-π interactions in the dense phase. Another important sidechain band is the tryptophan Raman band, typically observed at ~ 880 cm^{-1} that arises due to the indole N-H bending and is often used to probe the environment and is a measure of the hydrogen bonding strength between the N-H of the indole ring with the surrounding solvent molecules. This band is highly blue-shifted to 891 cm^{-1} in droplets indicating a reduced hydrogen bonding propensity of the N-H group with the surrounding water molecules implying an apolar microenvironment in the vicinity of tryptophan residues. Additionally, we observed a tryptophan band at 767 cm^{-1} that corresponds to the indole ring breathing and is used as a marker for cation-π/CH-π interactions.[70,71] Therefore, these observations might potentially indicate the presence of π-π/cation-π interactions within FUS condensates. Additionally, we observe considerable variation in the intensities of conventional Raman signatures in our single-droplet measurements. Since each protein droplet has a variable size and is in different stages of formation having different conformations of the polypeptide chains as well as the varying extent of protein-protein interactions, we do expect to observe this intensity variation that suggests a high degree of conformational heterogeneity and structural variation within the protein-rich droplets. We also observed a broad band at around 540 cm^{-1} corresponding to characteristic backbone deformations due to the presence of a large number of highly flexible

glycine residues in FUS.[72,73] Taken together, this set of normal single-droplet Raman experiments indicates conformational heterogeneity and intrinsic disorder of FUS within the droplets. These results also revealed the involvement of aromatic sidechains of tyrosine and tryptophan residues in the chain collapse and condensation of FUS corroborating previous findings.[39]

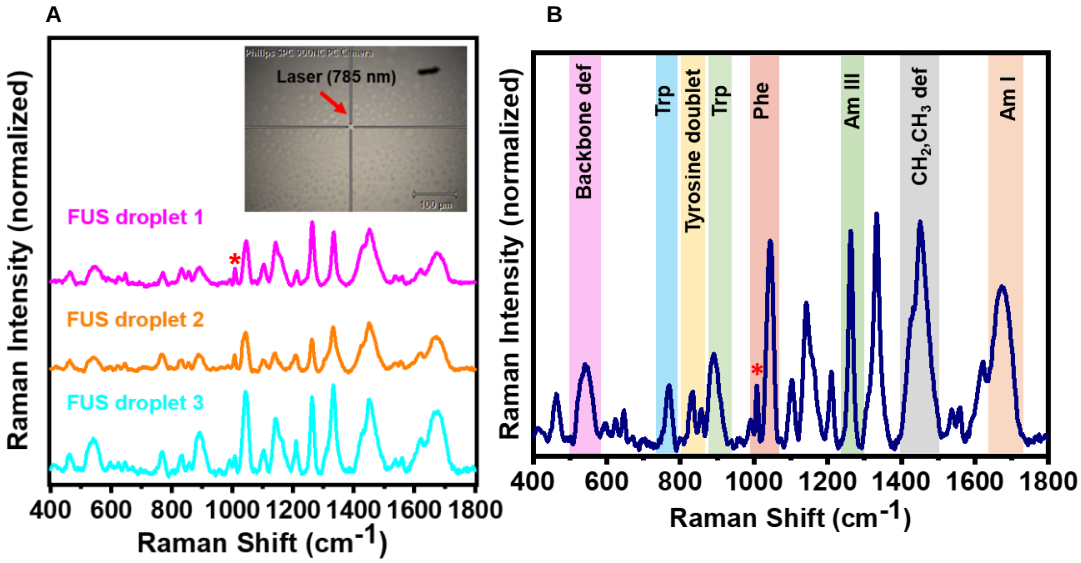


Figure 2.5: (A) Representative single droplet normal Raman spectra of a few individual FUS droplets (spectra recorded at 500mW laser power, ×100 objective; number of droplets, n = 3). Inset shows single droplets of FUS focused through the Raman microscope. Arrowhead shows the focal point of the NIR laser within the droplet. (B) Average Raman spectrum of the FUS condensed phase. All spectra are normalized with respect to the phenylalanine ring breathing band at 1007 cm^{-1} marked by an asterisk.

2.3.5 Illuminating FUS-RNA heterotypic condensates using single-droplet Raman spectroscopy

RNA is known to modulate the phase behavior and biophysical properties of liquid condensates formed by several RNA-binding proteins including FUS[63,74]. With the objective of elucidating the effect of polyU RNA on the chain conformations within the droplets, we performed normal Raman different stoichiometries of RNA and protein (Figure 2.6). A careful inspection of normal single-droplet Raman spectra showed two RNA marker bands, a shoulder band at 782 cm^{-1} and a band at 1230 cm^{-1} corresponding to uracil breathing and ring stretching modes, respectively (Figure 2.7 A and B).[75] This is corroborated by the Raman difference spectrum as well (Figure 2.7 C). An increase in the RNA concentration leads to its greater recruitment

within the phase-separated droplets which is confirmed by the linear plot of peak intensity at 1230 cm^{-1} as a function of RNA concentration (Figure 2.7 D). Additionally, these single-

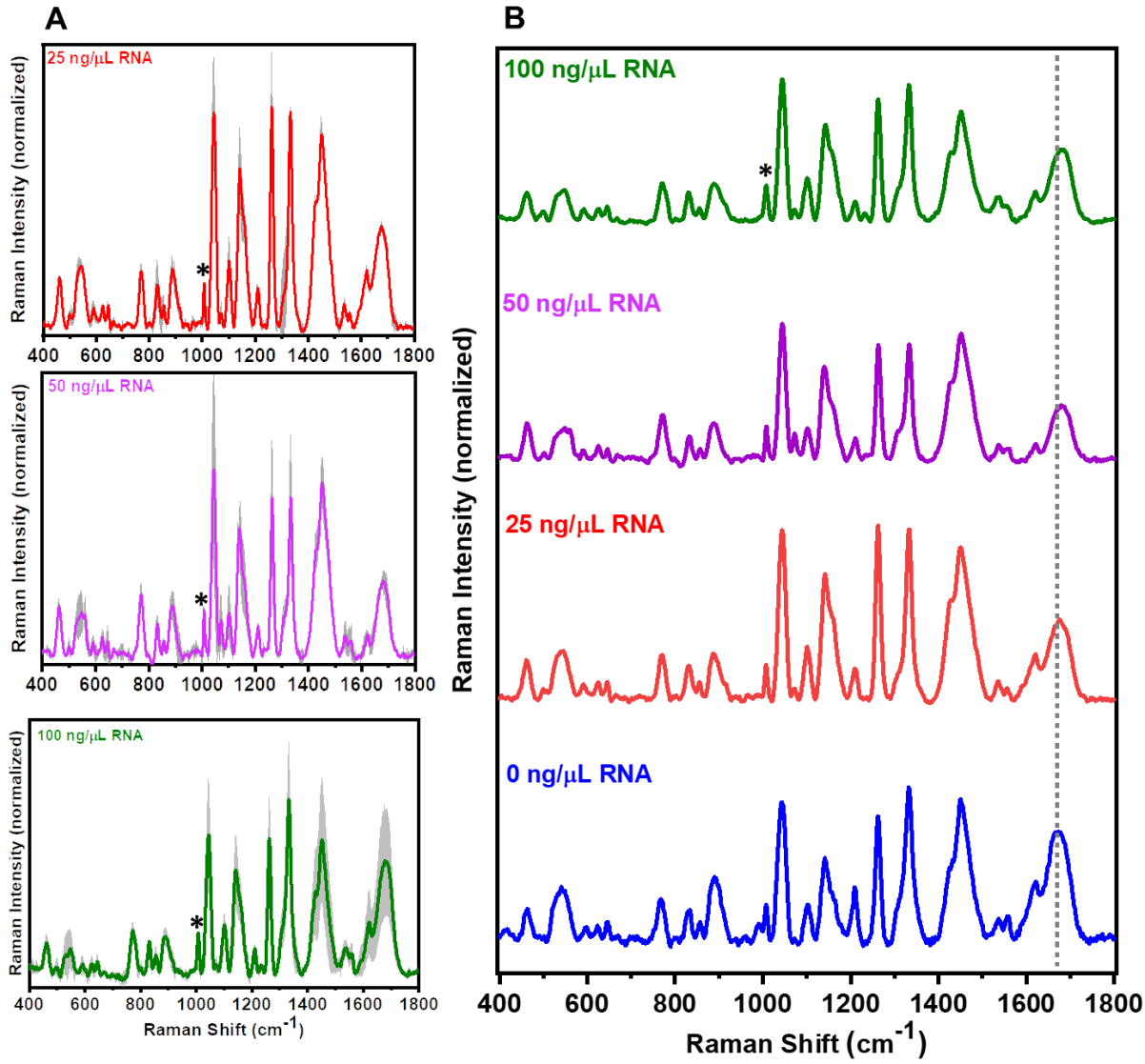


Figure 2.6: (A) Average single-droplet normal Raman spectra in the presence of 25 ng/µL, 50 ng/µL, and 100 ng/µL poly-U RNA (spectra recorded at 500mW laser power with a ×100 objective; number of droplets, n = 3). Solid lines represent the mean, whereas shaded region represents the standard deviation (n = 3). All spectra are normalized with respect to the phenylalanine ring breathing band at 1003 cm^{-1} marked by an asterisk. (B) Stacked average single-droplet normal Raman spectra for different concentrations of RNA are shown in blue (0 ng/µL RNA), red (25 ng/µL RNA), purple (50 ng/µL RNA), and olive (100 ng/µL RNA) for comparison (the dotted line shows the shift in amide I in the presence of RNA).

droplet Raman measurements also allowed to obtain the stoichiometry of RNA and protein within condensates by following the ratio of intensities at 1230 cm^{-1} (uracil ring stretching of

RNA) to 1450 cm^{-1} (CH_2/CH_3 deformation modes of protein). A linear relationship obtained can be used as a calibration line to evaluate the stoichiometry of RNA and protein within the condensed phase (Figure 2.7 E). Such quantitative and ratiometric estimates can be valuable in determining the concentration and composition of complex multi-component and multi-phasic condensates. Upon a closer inspection of the amide I region, we observed a considerable blue shift (Figure 2.7 F) from 1671 cm^{-1} to 1682 cm^{-1} indicating a β→disorder conversion with an increase in RNA concentration. This could potentially be due to the formation of more liquid-like condensates having more disorder and less β-content at higher RNA concentrations. Since the uracil carbonyl (C(4)=O) stretching mode can appear around this region,[76] we next zoomed into the amide III region to independently confirm this unraveling of FUS in the presence of RNA resulting in a more liquid-like behavior of FUS-RNA heterotypic condensates (Figure 2.7 G). In this amide III region as well, in the absence of RNA, we observed primarily random coils (~ 1262 cm^{-1}) with a small contribution of β-structure (broad shoulder at ~ 1248 cm^{-1}) that disappeared at higher RNA to protein stoichiometry. This observation also supports that the observed blue shift in the amide I might be due to an increase in the disordered content rather than the C(4)=O stretching mode of the uracil ring. As in the case of single-droplet normal Raman experiments of FUS droplets in the absence of RNA, we see a variation in the intensities of several Raman signatures within droplets in the presence of varying RNA concentrations. This too, suggests conformational heterogeneity within the condensed phase due to varying degrees of protein-RNA interactions. Additionally, we observed lower α-helical contents within the droplets (amide III: 1325 cm^{-1}; amide I: 1660 cm^{-1}) in the presence of RNA, as shown in the Raman difference spectrum (Figure 2.7 C). This observation also hinted at a possible unwinding of the helical region of RBD.

2.4 Discussion

Various on-demand membrane-less organelles formed in cells via LLPS are involved in a myriad of critical cellular functions and neurodegenerative diseases. Unmasking the role of intrinsic disorder and conformational heterogeneity of IDPs/IDRs in promoting promiscuous and ephemeral interactions resulting in the liquid-like behavior of these condensates is crucial to understand the molecular drivers of LLPS. While a host of existing microscopic and spectroscopic tools are immensely useful for studying LLPS, most of these methodologies are inadequate in illuminating the conformational heterogeneity and distribution within individual droplets. In order to overcome these limitations, we developed and adapted a novel, highly

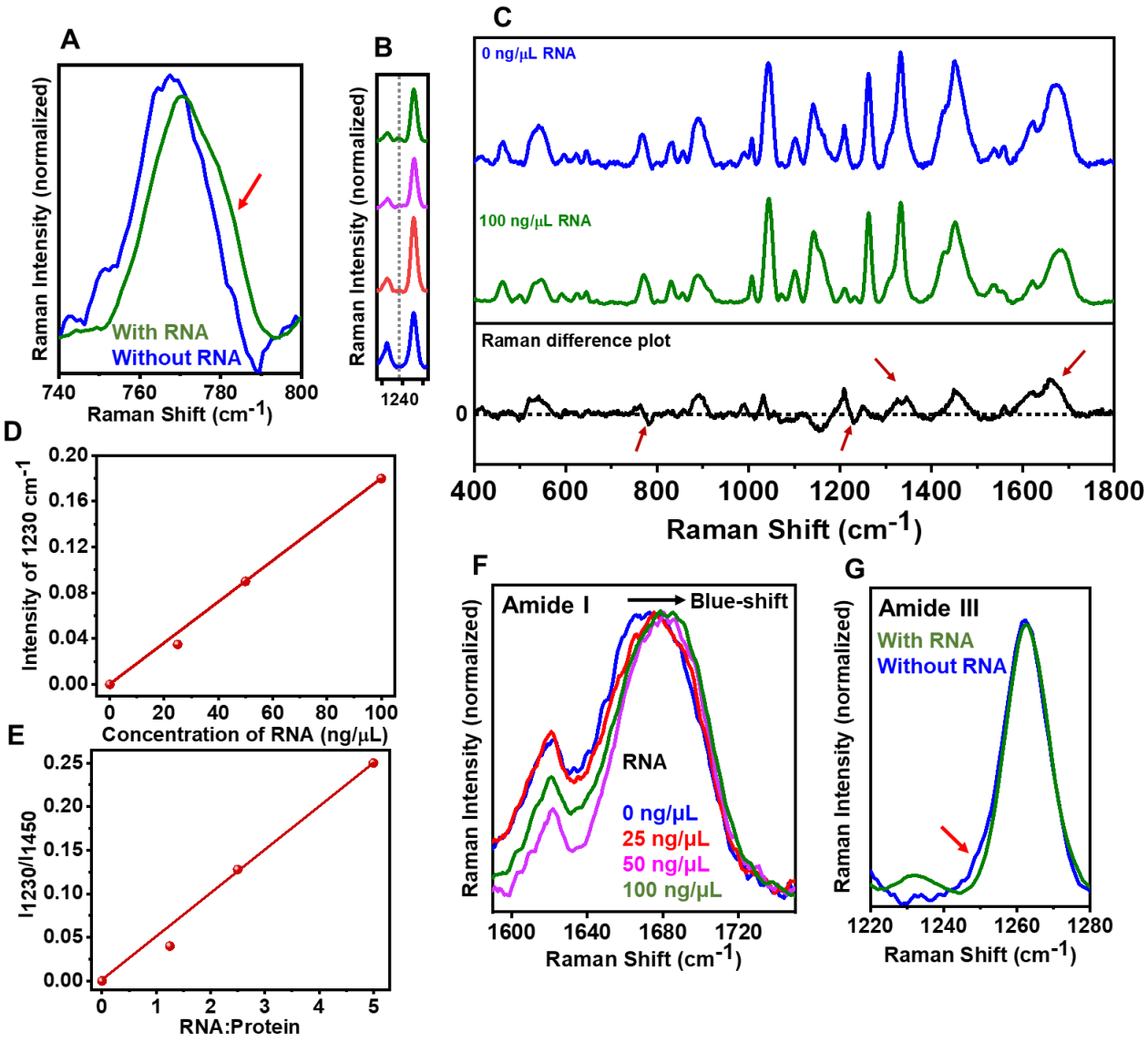


Figure 2.7: (A) Shoulder band at ~782 cm^{-1} indicated by a red arrow corresponding to the uracil ring breathing mode for FUS droplets in the presence of 100 ng/μL RNA. (B) Another RNA marker band at ~1230 cm^{-1} corresponding to the uracil ring stretching mode. (C) Raman difference plot (between 0 ng/ μL RNA and 100 ng/ μL RNA) of single-droplet normal Raman spectra of droplets in the absence and presence of 100 ng/μL RNA (arrows indicate the differences of interest). Arrows at 780 cm^{-1} and 1230 cm^{-1} represent the RNA marker bands. Arrows at 1325 cm^{-1} and 1660 cm^{-1} denote greater α-helical content within FUS droplets in the absence of RNA. (D) Linear plot of RNA marker band at 1230 cm^{-1} versus concentration of RNA (ng/μL) used. (E) Plot of ratio of intensity at 1230 cm^{-1} and 1450 cm^{-1} versus RNA: protein. (F) Zoomed in amide I region for FUS droplets at different concentrations of RNA indicating blue-shift of band maxima. (G) Zoomed in amide III region band for FUS droplets in the absence and presence of RNA (100 ng/μL). Arrow indicates a shoulder at ~1248 cm^{-1} denoting a small contribution of β-sheets in addition to random-coil structures within droplets in the absence of RNA.

sensitive, single-droplet structural tool involving dispersive laser Raman spectroscopy in a microscopy format that offers a wealth of fundamental molecular information within the mesoscopic liquid condensed phase. This design allows us to focus the laser beam into the protein-rich dense phase of individual droplets to capture Raman scattering bands associated with a range of molecular vibrational modes from protein-rich single droplets. Such single-droplet measurements offer a unique capability to capture the key molecular drivers and unveil the fundamental molecular interactions involved in biological phase transitions. This information is crucial to discern the mechanism and regulation of condensate assembly, dissolution, and modulation of aberrant phase transitions and liquid-to-solid maturations involved in debilitating human pathologies.

In summary, our single-droplet experiments allowed us to capture unique conformational characteristics that govern LLPS of an intrinsically disordered pathological stop codon variant of the prion protein, namely, Y145Stop that is associated with Gerstmann-Sträussler-Scheinker syndrome and familial cerebral amyloid angiopathy. The single-droplet vibrations studies illuminate the conformational disorder, heterogeneity, and distribution in Y145Stop liquid droplets and their maturation into ordered, β-rich, autocatalytic, amyloid-like aggregates formed via liquid-to-solid phase transitions. Our single-droplet Raman results showed an increase in the structural heterogeneity within liquid droplets of FUS. Several aromatic amino acid residues, such as tyrosine and tryptophan residues, display characteristics of the conformationally restricted environment in the condensed phase hinting at intermolecular π-π and/or cation-π interactions within liquid phase condensates. Our methodology allowed us to capture the unique spectral markers for droplets formed in the presence of varying RNA-protein ratios and showed that FUS binds stoichiometrically to RNA as shown previously[77,] and this can be used to estimate the stoichiometry of other complex biomolecular condensates of proteins and nucleic acids.

2.5 References

1. Alberti, S.; Hyman, A. A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **2021**, 22, 196-213.

2. Lyon, A. S.; Peeples, W. B.; Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **2021**, 22, 215-235.

3. Fuxreiter, M.; Vendruscolo, M. Generic nature of the condensed states of proteins. *Nat. Cell Biol*. **2021**, 23, 587-594.

4. Roden, C.; Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol*. **2021**, 22, 183-195.

5. Portz, B.; Lee, B. L.; Shorter, J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem. Sci.* **2021,** 46, 550-563.

6. Choi, J.-M.; Holehouse, A. S.; Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys*. **2020**, 49, 107-133.

7. Shapiro, D. M.; Ney, M.; Eghtesadi, S. A.; Chilkoti, A. Protein Phase Separation Arising from Intrinsic disorder: First-Principles to Bespoke Applications. *J. Phys. Chem. B*. **2021**, 125, 6740-6759.

8. Boeynaems, S.; Alberti, S.; Fawzi, N. L.; Mittag, T.; Polymenidou, M.; Rousseau, F.; Schymkowitz, J.; Shorter, J.; Wolozin, B.; Van Den Bosch, L.; Tompa, P.; Fuxreiter M. Protein phase separation: A new phase in cell biology. *Trends Cell Biol*. **2018**, 28, 420-435.

9. Lafontaine, D. L.; J.; Riback, J. A.; Bascetin, R.; Brangwynne, C. P. The nucleolus as a multiphase liquid condensate. *Nat. Rev. Mol. Cell Biol*. **2021,** 22, 165-182.

10. Sabari, B. R.; Dall'Agnese, A.; Young, R. A. Biomolecular condensates in the nucleus. *Trends Biochem. Sci*. **2020,** 45, 961-977.

11. Nesterov, S. V.; Ilyinsky, N. S.; Uversky, V. N. Liquid-liquid phase separation as a common organizing principle of intracellular space and biomembranes providing dynamic adaptive responses. *Biochim. Biophys. Acta (BBA)-Molecular Cell Res*. **2021,** 1868, 1-11.

12. Banani, S. F.; Lee, H. O.; Hyman, A. A.; Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol*. **2017,** 18, 285-298.

13. Fare, C. M.; Villani, A.; Drake, L. E.; Shorter, J. Higher-order organization of biomolecular condensates. *Open Biol*. **2021,** 11, 1-19.

14. Riback, J. A.; Zhu, L.; Ferrolino, M. C.; Tolbert, M.; Mitrea, D. M.; Sanders, D. W.; Wei, M. T.; Kriwacki, R. W.; Brangwynne, C. P. Composition-dependent thermodynamics of intracellular phase separation. *Nature* **2020**, 581, 209-214.

15. Dignon, G. L.; Best, R. B.; Mittal, J. Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem*. **2020**, 71, 53-75.

16. Vernon, R. M.; Chong, P. A.; Tsang, B.; Kim, T. H.; Bah, A.; Farber, P.; Lin, H.; Forman-Kay, J. D. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **2018**, 7, e31486.

17. Qamar, S.; Wang, G.; Randle, S. J.; Ruggeri, F. S.; Varela, J. A.; Lin, J. Q.; Phillips, E. C.; Miyashita, A.; Williams, D.; Ströhl, F.; Meadows, W.; Ferry, R.; Dardov, V. J.; Tartaglia, G. G.; Farrer, L. A.; Kaminski Schierle, G. S.; Kaminski, C. F.; Holt. C. E.; Fraser. P. E.; Schmitt-Ulms, G.; Klenerman, D.; Knowles, T.; Vendruscolo, M.; St George-Hyslop, P. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation- π interactions. *Cell* **2018**, 173, 720-734.

18. Krainer, G.; Welsh, T. J.; Joseph, J. A.; Espinosa, J. R.; Wittmann, S.; de Csilléry, E.; Sridhar, A.; Toprakcioglu, Z.; Gudiškytė, G.; Czekalska. M. A.; Arter, W. E.; Guillén-Boixet, J.; Franzmann, T. M.; Qamar, S.; George-Hyslop, P. S.; Hyman, A.A.; Collepardo-Guevara, R.; Alberti, S.; Knowles, T. P. J. Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* 2021, 12, 1085-1099.

19. Martin, E. W.; Holehouse, A. S.; Peran, I.; Farag, M.; Incicco, J. J.; Bremer, A.; Grace, C. R.; Soranno, A.; Pappu, R. V.; Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020**, 367, 694-699.

20. Wang, J.; Choi, J. M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; Poser, I.; Pappu, R. V.; Alberti, S.; Hyman,

A. A. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **2018**, 174, 688-699.

21. Uversky, V. N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Curr. Opin. Struct. Biol*. **2017**, 44, 18-30.

22. Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nature Phys.* **2015,** 11, 899-904.

23. Mukhopadhyay, S. The Dynamism of Intrinsically Disordered Proteins: Binding-Induced Folding, Amyloid Formation, and Phase Separation. *J. Phys. Chem. B* **2020**, 124, 11541-11560.

24. Mitrea, D. M.; Chandra, B.; Ferrolino, M. C.; Gibbs, E. B.; Tolbert, M.; White, M. R.; Kriwacki, R. W. Methods for Physical Characterization of Phase-Separated Bodies and Membrane-less Organelles. *J. Mol. Biol.* **2018,** 430(23), 4773-4805.

25. Ganser, L. R.; Myong, S. Methods to Study Phase-Separated Condensates and the Underlying Molecular Interactions. *Trends Biochem. Sci.* **2020,** 45(11), 1004-1005.

26. Elbaum-Garfinkle, S.; Kim, Y.; Szczepaniak, K.; Chen, C. C.; Eckmann, C. R.; Myong, S.; Brangwynne, C. P. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** 112(23), 7189-7194.

27. Murthy, A. C.; Fawzi, N. L. The (un)structural biology of biomolecular liquid-liquid phase separation using NMR spectroscopy. *J. Biol. Chem.* **2020,** 295(8), 2375-2384.

28. Brady, J. P.; Farber, P. J.; Sekhar, A.; Lin, Y. H.; Huang, R.; Bah, A.; Nott, T. J.; Chan, H. S.; Baldwin, A. J.; Forman-Kay, J. D.; Kay, L. E. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** 114(39), E8194-E8203.

29. Niaki, A. G.; Sarkar, J.; Cai, X.; Rhine, K.; Vidaurre, V.; Guy, B.; Hurst, M.; Lee, J. C.; Koh, H. R.; Guo, L.; Fare, C. M.; Shorter, J.; Myong, S. Loss of Dynamic RNA Interaction and Aberrant Phase Separation Induced by Two Distinct Types of ALS/FTD-Linked FUS Mutations. *Mol. Cell.* **2020,** 77(1), 82-94.

30. Jain, A.; Liu, R.; Xiang, Y. K.; Ha, T. Single-molecule pull-down for studying protein interactions. *Nat. Protoc.* **2012,** 7(3), 445-452.

31. Mitrea, D. M.; Cika, J. A.; Guy, C. S.; Ban, D.; Banerjee, P. R.; Stanley, C. B.; Nourse, A.; Deniz, A. A.; Kriwacki, R. W. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* **2016,** 5, e13571.

32. Wei, M. T.; Elbaum-Garfinkle, S.; Holehouse, A. S.; Chen, C. C.; Feric, M.; Arnold, C. B.; Priestley, R. D.; Pappu, R. V.; Brangwynne, C. P. Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.* **2017,** 9(11), 1118-1125.

33. Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, 176(3), 419-434.

34. Alberti, S.; Saha, S.; Woodruff, J. B.; Franzmann, T. M.; Wang, J.; Hyman, A. A. A User's Guide for Phase Separation Assays with Purified Proteins. *J. Mol. Biol*. **2018**, 430(23), 4806-4820.

35. Conicella, A. E.; Dignon, G. L.; Zerze, G. H.; Schmidt, H. B.; D'Ordine, A. M.; Kim, Y. C.; Rohatgi, R.; Ayala, Y. M.; Mittal, J.; Fawzi, N. L. TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, 117(11), 5883-5894.

36. Rai, S. K.; Savastano, A.; Singh, P.; Mukhopadhyay, S.; Zweckstetter, M. Liquid-liquid phase separation of tau: From molecular biophysics to physiology and disease. *Protein Sci.* **2021**, 30(7), 1294-1314.

37. Paloni, M.; Bailly, R.; Ciandrini, L.; Barducci, A. Unraveling Molecular Interactions in Liquid-Liquid Phase Separation of Disordered Proteins by Atomistic Simulations. *J. Phys. Chem. B.* **2020**, 124(41), 9009-9016.

38. Kanelis, V.; Forman-Kay, J. D.; Kay, L. E. Multidimensional NMR methods for protein structure determination. *IUBMB Life*. **2001**, 52(6), 291-302.

39. Emmanouilidis, L.; Esteban-Hofer, L.; Damberger, F. F.; de Vries, T.; Nguyen, C. K. X.; Ibáñez, L. F.; Mergenthal, S.; Klotzsch, E.; Yulikov, M.; Jeschke, G.; Allain, F. H. NMR and EPR reveal a compaction of the RNA-binding protein FUS upon droplet formation. *Nat. Chem. Biol*. **2021**, 17(5), 608-614.

40. Benevides, J. M.; Overman, S. A.; Thomas, G. J. Raman spectroscopy of proteins. *Curr. Protoc. Protein Sci*. **2004**.

41. Devitt, G.; Howard, K.; Mudher, A.; Mahajan, S. Raman Spectroscopy: An Emerging Tool in Neurodegenerative Disease Research and Diagnosis. *ACS Chem. Neurosci*. **2018**, 9(3), 404-420.

42. Jayaraman, V.; Rodgers, K. R.; Mukerji, I.; Spiro, T. G. Hemoglobin allostery: resonance Raman spectroscopy of kinetic intermediates. *Science* **1995**, 269(5232), 1843-1848.

43. Carey, P. R. Raman spectroscopy, the sleeping giant in structural biology, awakes. *J. Biol. Chem*. **1999**, 274(38), 26625-26628.

44. Matousek, P.; Stone, N. Emerging concepts in deep Raman spectroscopy of biological tissue. *Analyst* **2009**, 134(6), 1058-1066.

45. Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. Raman spectroscopy of proteins: a review. *J. Raman Spectrosc*. **2013**, 44, 1061-1076.

46. Tuma, R. Raman Spectroscopy of Proteins: from Peptides to Large Assemblies. *J. Raman Spectrosc.* **2005**, 36, 307-319.

47. Kuhar, N.; Sil, S.; Verma, T.; Umapathy, S. Challenges in Application of Raman Spectroscopy to biology and Materials. *RSC Adv*. **2018**, 8, 25888-25908.

48. Maiti, N. C.; Apetri, M. M.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. Raman spectroscopic characterization of secondary structure in natively unfolded proteins: alpha-synuclein. *J. Am. Chem. Soc*. **2004**, 126(8), 2399-2408.

49. Agarwal, A.; Rai, S. K.; Avni, A.; Mukhopadhyay, S. An intrinsically disordered pathological variant of the prion protein Y145Stop transforms into self-templating amyloids via liquid-liquid phase separation. *Proc. Natl. Acad. Sci. U. S. A.* **2021,** 118(45), e2100968118.

50. Avni. A.; Joshi. A.; Walimbe. A.; Pattanashetty. S. G.; Mukhopadhyay. S. Single-Droplet Surface-Enhanced Raman Scattering Captures the Conformational Landscape within Phase Separated Liquid Condensates of an RNA-Binding Protein. *Nat. Commun.* **2022,** 13, 4378.

51. Prusiner, S. B. Prions. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** 95(23), 13363-13383.

52. Prusiner, S. B. Prion Biology. *Cold Spring Harbor Laboratory Press,* **2017.**

53. Scheckel, C.; Aguzzi, A. Prions, prionoids and protein misfolding disorders. *Nat. Rev. Genet.* **2018,** 19(7), 405-418.

54. Zahn. R.; Liu, A.; Lührs, T.; Riek, R.; von Schroetter, C.; López García, F.; Billeter, M.; Calzolai, L.; Wider, G.; Wüthrich, K. NMR solution structure of the human prion protein. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** 97(1), 145-150.

55. Kundu, B.; Maiti, N. R.; Jones, E. M.; Surewicz, K. A.; Vanik, D. L.; Surewicz, W. K. Nucleation-dependent conformational conversion of the Y145Stop variant of human prion protein: structural clues for prion propagation. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** 100(21), 12069-12074.

56. Helmus, J. J.; Surewicz, K.; Nadaud, P. S.; Surewicz, W. K.; Jaroniec, C. P. Molecular conformation and dynamics of the Y145Stop variant of human prion protein in amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** 105(17), 6284-6289.

57. Zweckstetter, M. Conserved amyloid core structure of stop mutants of the human prion protein. *Prion* **2013,** 7(3), 193-197.

58. Ghetti, B.; Piccardo, P.; Spillantini, M. G.; Ichimiya, Y.; Porro, M.; Perini, F.; Kitamoto, T.; Tateishi, J.; Seiler, C.; Frangione, B.; Bugiani, O.; Giaccone, G.; Prelli, F.; Goedert, M.; Dlouhy, S. R.; Tagliavini, F. Vascular variant of prion protein cerebral amyloidosis with tau-positive neurofibrillary tangles: the phenotype of the stop codon 145 mutation in PRNP. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** 93(2), 744-748.

59. Zanusso, G.; Petersen, R. B.; Jin, T.; Jing, Y.; Kanoush, R.; Ferrari, S.; Gambetti, P.; Singh, N. Proteasomal degradation and N-terminal protease resistance of the codon 145 mutant prion protein. *J. Biol. Chem.* **1999,** 274(33), 23396-23404.

60. St George-Hyslop, P.; Lin, J. Q.; Miyashita, A.; Phillips, E. C.; Qamar, S.; Randle, S. J.; Wang, G. The physiological and pathological biophysics of phase separation and gelation of RNA binding proteins in amyotrophic lateral sclerosis and fronto-temporal lobar degeneration. *Brain Res.* **2018,** 1693, 11-23.

61. Svetoni, F.; Frisone, P.; Paronetto, M. P. Role of FET proteins in neurodegenerative disorders. *RNA biol.* **2016**, 13(11), 1089-1102.

62. Tüű-Szabó, B.; Hoffka, G.; Duro, N.; Fuxreiter, M. Altered dynamics may drift pathological fibrillization in membraneless organelles. *Biochim. Biophys. Acta. Proteins Proteomics* **2019,** 1867, 988-998.

63. Zhang, H.; Elbaum-Garfinkle, S.; Langdon, E. M.; Taylor, N.; Occhipinti, P.; Bridges, A. A.; Brangwynne, C. P.; Gladfelter, A. S. RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell* **2015,** 60, 220-230.

64. Qamar, S.; Wang, G.; Randle, S. J.; Ruggeri, F. S.; Varela, J. A.; Lin, J. Q.; Phillips, E. C.; Miyashita, A.; Williams, D.; Strohl, F.; Meadows, W.; Ferry, R.; Dardov, V. J.; Tartaglia, G. G.; Farrer, L. A.; Kaminski Schierle, G. S.; Kaminski, C. F.; Holt, C. E.; Fraser, P. E.; Schmitt-Ulms, G.; Klenerman, D.; Knowles, T.; Vendruscolo, M.; St George-Hyslop, P. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. *Cell* **2018,** 173, 720-734.

65. Hofweber, M.; Hutten, S.; Bourgeois, B.; Spreitzer, E.; NiednerBoblenz, A.; Schifferer, M.; Ruepp, M. D.; Simons, M.; Niessing, D.; Madl, T.; Dormann, D. Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **2018,** 173, 706-719.

66. Murthy, A. C.; Dignon, G. L.; Kan, Y.; Zerze, G. H.; Parekh, S. H.; Mittal, J.; Fawzi, N. L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019**, 26(7), 637-648.

67. Patel, A.; Lee, H. O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M. Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T. M.; Pozniakovski, A.; Poser, I.; Maghelli, N.; Royer, L. A.; Weigert, M.; Myers, E. W.; Grill, S.; Drechsel, D.; Hyman, A. A.; Alberti, S. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, 162(5), 1066-1077.

68. Lippert, J. L.; Tyminski, D.; Desmeules, P. J. Determination of the secondary structure of proteins by laser Raman spectroscopy. *J. Am. Chem. Soc.* **1976**, 98(22), 7075-7080.

69. Hernandez, B.; Coic, Y. M.; Pfluger, F.; Kruglik, S. G.; Ghomi, M. All Characteristic Raman Markers of Tryrosine and Tyrosinate Originate from Phenol Ring Fundamental Vibrations. *J. Raman Spectrosc.* **2016,** 47, 210-220.

70. Schlamadinger, D. E.; Daschbach, M. M.; Gokel, G. W.; Kim, J. E. UV resonance Raman study of cation-π interactions in an indole crown ether. *J. Raman Spectrosc*. **2011**, 42(4), 633-638.

71. Milan-Garces, E. A.; Mondal, S.; Udgaonkar, J. B.; Puranik, M. Intricate packing in the hydrophobic core of barstar through a CH-π interaction. *J. Raman Spectrosc*. **2014,** 45, 814-821.

72. Chen, L. X.; Strauss, H. L.; Snyder, R. G. Analysis of the delocalized Raman modes of conformationally disordered polypeptides. *Biophys. J.* **1993**, 64(5), 1533-1541.

73. Mensch, C.; Bultinck, P.; Johannessen, C. Conformational Disorder and Dynamics of Proteins Sensed by Raman Optical Activity. *ACS Omega* **2018**, 3(10), 12944-12955.

74. Sanders, D. W.; Kedersha, N.; Lee, D. S. W.; Strom. A. R.; Drake, V.; Riback, J. A.; Bracha, D.; Eeftens, J. M.; Iwanicki, A.; Wang, A.; Wei, M. T.; Whitney, G.; Lyons, S. M.; Anderson, P.; Jacobs, W. M.; Ivanov, P.; Brangwynne, C. P. Competing protein- RNA interaction networks control multiphase intracellular organization. *Cell* **2020,** 181, 306-324.

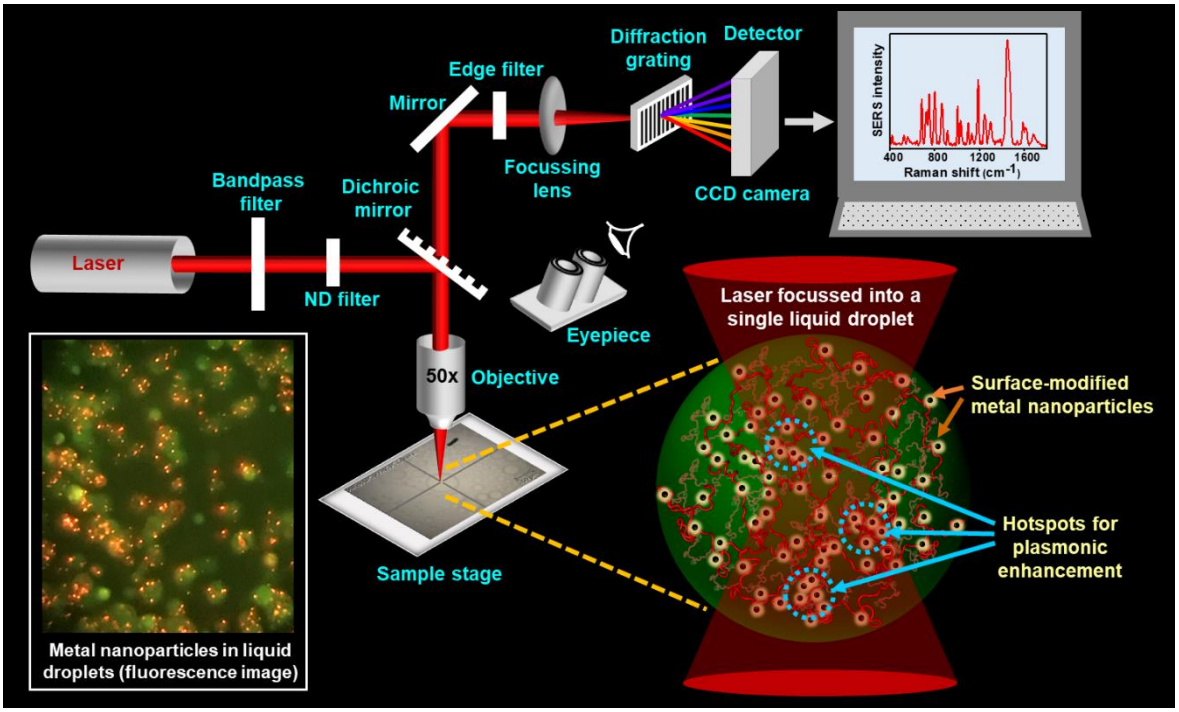
75. Devitt, G.; Crisford, A.; Rice, W.; Weismiller, H. A.; Fan, Z.; Commins, C.; Hyman, B. T.; Margittai, M.; Mahajan, S.; Mudher, A. Conformational fingerprinting of tau variants and strains by Raman spectroscopy. *RSC Adv*. **2021**, 11(15), 8899-8915.

76. Lafleur, L.; Rice, J.; Thomas, G. J. Jr. Raman studies of nucleic acids. VII. Poly A-poly U and poly G-poly C. *Biopolymers* **1972**, 11(12), 2423-2437.

77. Niaki, A. G.; Sarkar, J.; Cai, X.; Rhine, K.; Vidaurre, V.; Guy, B.; Hurst, M.; Lee, J. C.; Koh, H. R.; Guo, L.; Fare, C. M.; Shorter, J.; Myong, S. Loss of dynamic RNA interaction and aberrant phase separation induced by two distinct types of ALS/FTD-linked FUS mutations. *Mol. Cell* **2020,** 77, 82-94.

******

Single-Droplet Surface-Enhanced Raman Scattering Captures the Conformational Landscape within Phase-Separated Liquid Condensates of an RNA-Binding Protein



The work described in this chapter has been published in *Nature Communications*

Avni. A.; Joshi. A.; Walimbe. A.; Pattanashetty. S. G.; Mukhopadhyay. S. Single-Droplet Surface-Enhanced Raman Scattering Decodes the Molecular Determinants of Liquid-Liquid Phase Separation. *Nat. Commun.* **2022,** 13, 4378.

3.1 Introduction

Biomolecular condensation via liquid-liquid phase separation (LLPS) offers an exquisite mechanism for spatiotemporally-controlled organization and compartmentalization of cellular constituents into highly dynamic, permeable, liquid-like, tunable, mesoscopic, non-stoichiometric supramolecular assemblies known as membrane-less organelles[1-10]. These on-demand non-canonical organelles containing proteins and nucleic acids are present both in the cytoplasm and nucleus and include nucleoli, stress granules, P granules, nuclear speckles, and so forth. A growing body of exciting research suggests that highly flexible intrinsically disordered proteins/regions (IDPs/IDRs) containing low-complexity regions and prion-like domains that offer conformational heterogeneity, distributions, and multivalency are excellent candidates for intracellular phase separation. A unique combination of these sequence-dependent features governs the making and breaking of promiscuous and ephemeral intermolecular interactions such as electrostatic, hydrophobic, hydrogen bonding, dipole-dipole, π-π, and cation-π interactions resulting in liquid-like behavior.[3,6,11-17] While these highly dynamic condensates are proposed to be involved in a range of critical cellular functions, their transitions into less dynamic gel-like or solid-like aggregates containing more persistent interchain interactions are linked to debilitating neurodegenerative diseases. Therefore, it is imperative to decipher the fundamental molecular interactions of phase transitions involved in functions and disease.[18] A multitude of spectroscopic and microscopic methodologies have been employed to unveil the key biophysical principles of phase separation resulting in the formation of liquid droplets. For instance, high-resolution microscopic tools such as confocal, super-resolution, and high-speed atomic force microscopy can directly probe the properties within individual liquid droplets.[19,20] However, these tools do not allow us to access the wealth of molecular information in a residue-specific manner. On the contrary, high-resolution structural methods such as nuclear magnetic resonance (NMR) and small-angle X-ray scattering (SAXS) provide the atomic-resolution details of the condensed phase architecture.[21-23] However, these ensemble structural methods are not capable of yielding molecular insights from the condensed phase of individual droplets. Therefore, a method that combines these capabilities to capture residue-specific structural information at a single-droplet resolution is essential to characterize and quantify the key molecular determinants in a droplet-by-droplet manner.

Vibrational Raman spectroscopy performed in a microscopy format allows us to uniquely and elegantly combine the aforesaid capabilities to obtain the protein structural

information from a well-defined spatial location by focusing the laser beam into a sub-micron spot. Such non-invasive and label-free laser micro-Raman measurements permit us to access the wealth of structural information by monitoring a range of bond vibrational frequencies while retaining spatial resolution.[24-26] However, owing to a low Raman scattering cross-section, Raman spectroscopy is a highly insensitive technique, especially for biomolecules under physiological conditions in aqueous solutions.[27] Additionally, the high laser power required for Raman spectroscopic detection can be detrimental to soft biological samples. The low-sensitivity issue in Raman scattering can be circumvented by a near-field plasmonic enhancement by metallic nanostructured substrates giving rise to high electromagnetic/chemical enhancement of Raman signals even at extremely low analyte concentrations. This surface-sensitive technique is known as surface-enhanced Raman scattering (SERS) that can provide several orders of magnitude increase in the Raman scattering cross-section, allowing single-molecule detection and characterization even at a much lower laser power.[28-32] In the present work, we have developed an ultra-sensitive single-droplet SERS methodology that can illuminate the unique molecular details of the polypeptide chains within individual phase-separated protein liquid droplets. For our studies, we have used Fused in Sarcoma (FUS), one of the most intensely studied RNA-binding proteins containing archetypal prion-like low-complexity domains and hence one of the best prototypes of phase-separating proteins. The human genome encodes approximately 30 FUS-family proteins that are known to be involved in critical functions such as mRNA splicing, DNA damage repair, formation of stress granules as well as in deadly neurodegenerative diseases such as ALS (amyotrophic lateral sclerosis), and FTD (frontotemporal dementia).[33-38] Here, we show that upon liquid phase condensation, surface-coated, SERS-active nanosphere substrates get spontaneously encapsulated within the protein-rich condensed phase and generate plasmonic hotspots that permit us to capture the inscrutable workings of FUS condensates with unprecedented sensitivity in the absence and the presence of RNA.

3.2 Experimental details

3.2.1 Materials

Sodium phosphate monobasic dihydrate, potassium chloride, sodium phosphate dibasic dihydrate, sodium chloride, 2-mercaptoethanol, 1,4-dithiothreitol (DTT), Poly-U sodium salt, phenylmethylsulfonyl fluoride (PMSF), magnesium chloride hexahydrate, Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), ethylenediaminetetraacetic acid (EDTA),

silver nitrate, auric chloride, sodium citrate tribasic, and potassium iodide, were of MB grade purity, procured from Sigma (St. Louis, MO, USA). Luria Bertani Broth, Miller (LB), zinc chloride, nickel chloride, Lysozyme, HEPES, and maltose were procured from HiMedia laboratories (MB grade). Ampicillin, chloramphenicol, and isopropyl-β- thiogalactopyranoside (IPTG) were ordered from Gold Biocom (USA). Fluorescent probes like fluorescein-5-maleimide (F-5-M) and AlexaFluor488-C5-maleimide were purchased from Molecular Probes, Invitrogen. Ni-NTA resin and Amylose resin were purchased from Qiagen. Amicon membrane filters for concentrating protein were obtained from Merck Millipore. High-purity milli-Q water was used to prepare all the buffers in this study. A Metrohm 827 lab pH meter was used to adjust the pH (± 0.01) of all the buffer solutions prepared at 25 °C and all the buffer solutions were filtered before use. All the glassware used for nanoparticle preparation was cleaned with freshly prepared aqua regia and washed thoroughly with milli-Q water and dried overnight in a hot-air oven.

3.2.2 Recombinant protein expression and purification

The plasmid expressing MBP-*Tev*-FUS-*Tev*-His_6 was transformed into *E. coli* BL21(DE3) RIPL strain. The expression and purification protocols followed were similar as previously described with some slight modifications.[37] For overexpression, cultures were grown in LB media at 37 °C, 220 rpm till $O.D._{600}$ reached 0.8-1 and was induced with 0.1 mM isopropyl-β-thiogalactopyranoside (IPTG) at 12 °C for 22 hours. Cell pellets were obtained by centrifugation at 4°C, 4000 rpm for 40 minutes, and stored at -80 °C for further use. For purification, pellets were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 40 mM imidazole, 10 μM $ZnCl_2$, 4 mM BME, and 10% v/v glycerol, pH 8.0) and cell-lysis was done using probe sonication at 5% amplitude, 15 sec ON and 10 sec OFF for 25 minutes. The lysate was centrifuged at 4 °C, 11,400 rpm for 1 hour, followed by incubation of the supernatant with equilibrated Ni-NTA agarose beads for 1.5 hours at 4 °C. The beads were washed and protein was eluted with 250 mM imidazole followed by binding to the amylose column. Protein was eluted with 20 mM maltose elution buffer (50 mM sodium phosphate, 800 mM NaCl, 40 mM imidazole, 10 μM $ZnCl_2$, 20 mM maltose and 1 mM 1,4-dithiothreitol, pH 8.0). The concentration of the protein was estimated by measuring absorbance at 280 nm ($\varepsilon_{280\ nm}$= 1,30,670 $M^{-1}cm^{-1}$) and samples were run on an SDS-PAGE gel to confirm the purity of the protein. The purified protein was further stored at 4 °C for future use.

The plasmid containing his_6 tagged TEV protease was transformed into *E. coli* strain BL21(DE3) plysS. Cells were grown at 37 °C, 220 rpm, and overexpression was induced by 0.35 mM IPTG at 16 °C for 20 hours. Cultures were pelleted and stored at -80 °C for further use. The pellets were thawed at 30 °C and suspended in lysis buffer (25 mM HEPES, 150 mM NaCl, 20 mM KCl, and 20 mM $MgCl_2$, pH 7.4) along with phenylmethylsulfonyl fluoride and lysozyme to enhance cell lysis which was carried out by probe sonication (5% amplitude, 15 sec ON/10 sec OFF for 25 minutes). The soluble fraction was separated by centrifugation and the supernatant was passed twice through a pre-equilibrated Ni-NTA column at 4 °C. The beads were washed with wash buffer (lysis buffer + 20 mM imidazole) and protein was eluted with 300 mM imidazole and dialyzed against buffer without imidazole, overnight at 4 °C. Protein was concentrated using a 10 kDa MWCO filter and stored at -80 °C for further use.

3.2.3 Fluorescence labeling

For labeling, purified FUS was concentrated using a 50 kDa MWCO amicon filter, and incubated with 0.3 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes on ice following which reactions with 1:4.5 molar ratio of protein: dye (for AlexaFluor488-C5-maleimide) and 1:30 (for Fluorescein-5-maleimide) were set up in native buffer at 25°C and kept under shaking for 2 hours in the dark. Unreacted dye was removed by buffer exchange using 50 kDa MWCO amicon filters. Labeling efficiency was calculated by measuring absorbance at 280 nm ($\varepsilon_{280\ nm}$=1,30,670 $M^{-1}cm^{-1}$, for full-length FUS) and 494 nm (ε_{494} = 72,000 $M^{-1}cm^{-1}$, for AlexaFluor488-C5-maleimide and ε_{494} = 68,000 $M^{-1}cm^{-1}$, for F-5-M) to estimate total protein concentration and labeled protein concentration.

3.2.4 Phase separation assays

Phase separation of FUS was initiated by TEV cleavage in a 1:10 molar ratio (TEV: protein) at room temperature in 20 mM sodium phosphate, pH 7.4. Turbidity of phase-separated samples was then estimated using 96-well NUNC optical bottom plates (Thermo Scientific) on a Multiskan Go (Thermo Scientific) plate reader by recording the absorbance at 350 nm. The protein concentration for all the experiments was fixed to 20 μM along with 0.1nM of iodide-modified silver nanoparticles Ag IMNPs (for reactions in the presence of nanoparticles). For phase separation in the presence of RNA, LLPS was induced in the presence of 25 ng/μL, 50 ng/μL, and 100 ng/μL polyU RNA with or without 0.1 nM Ag IMNPs. The total sample volume

used was 100 μL for all the measurements and then background subtracted turbidity was plotted using Origin.

3.2.5 Confocal microscopy

Confocal fluorescence imaging of FUS droplets with and without Ag IMNPs was performed on ZEISS LSM 980 Elyra 7 super-resolution microscope using a 63x oil-immersion objective (Numerical aperture 1.4). For visualizing droplets of FUS, 200 nM (1 %) of Alexa488 or F-5-M labeled protein was doped with unlabeled protein, and 2-3 μL of the freshly phase-separated sample was placed into a chamber made on a glass slide (Fisher Scientific 3" × 1" × 1 mm). The chamber made by using double-sided tape was then sealed with a square coverslip to avoid evaporation of the sample. For visualization of encapsulated Ag IMNPs (0.1nM), Alexa488-labeled protein was imaged using a 488-nm laser diode (11.9 mW), and Ag IMNPs were imaged using a 405-nm laser diode (11.9 mW). For images captured through the eyepiece, a metal halide lamp was used to excite both labeled protein and nanoparticles. All the confocal images were then processed and analyzed using ImageJ (NIH, Bethesda, USA).

3.2.6 Fluorescence recovery after photobleaching (FRAP) measurements

FRAP experiments for droplets with and without Ag IMNPs were performed on ZEISS LSM 980 Elyra 7 super-resolution microscope using a 63x oil-immersion objective (Numerical aperture 1.4). All the FRAP experiments were performed using 200 nM (1%) of Alexa488-labeled protein. The recovery of the chosen region of interest (ROI) after photobleaching using a 488-nm laser was then recorded using ZEN Pro 2011 (ZEISS) software provided with the instrument. The fluorescence recovery curves were then normalized and plotted after background correction using Origin.

3.2.7 Sedimentation assays

The absence of MBP within FUS droplets was confirmed using a sedimentation assay. The MBP-FUS was cleaved by TEV protease to induce phase separation and after 20 minutes the reaction was centrifuged at 25,000 x g, 25 °C for 30 minutes to pellet down all the droplets (condensed phase). Both the supernatant and the pellet were then separated carefully and the pellet was dissolved in 8 M urea. Samples were then heated and run on 12% SDS-PAGE along with the respective controls.

3.2.8 Circular dichroism (CD) measurements

CD measurements were performed on a BioLogic scanning MOS-500 CD spectrometer using a quartz cuvette of 1 mm path length. The final protein concentration of 5 μM protein and 0.025 nM Ag IMNPs in 20 mM sodium phosphate buffer pH 7.4 was used for recording the data. The spectra were averaged over three scans and blank subtracted data were plotted using Origin software.

3.2.9 Size distribution analysis of droplets

Droplet reactions of 20 μM FUS (with 1% Alexa-488 labeled FUS) were set up in the absence and presence of 0.1 nM Ag IMNPs. Confocal imaging was performed at equal time points for both the samples using ZEISS LSM 980 Elyra 7 super-resolution microscope with a 63x oil-immersion objective (Numerical aperture 1.4). For visualization of the fluorescently labeled droplets, a 488-nm laser diode (11.9 mW) and for the Ag IMNPs, a 405-nm laser diode (11.9 mW) was used. Images were obtained from multiple reactions and were analyzed using ImageJ (NIH, Bethesda, USA). An area-wise distribution of droplets ranging from 0.3-8 μm^2 was constructed and plotted as an area distribution frequency plot. Mean area of droplets and radius was calculated.

3.2.10 Preparation of silver nanoparticles

Silver nanoparticles were prepared by the Lee-Meisel method as described previously.[48] Initially, 8.49 mg of silver nitrate was dissolved in 50 mL of filtered milli-Q water and stirred vigorously (1000 rpm) at its boiling point (~ 98 ^oC) for 30 minutes. One and five-tenths milliliters of freshly prepared 1% (w/v) aqueous trisodium citrate was added to the reaction mixture drop-wise and further stirred for additional 30 minutes till the color changed to yellow-green. The solution was cooled down to room temperature and was further characterized using UV-visible absorption spectroscopy, transmission electron microscopy, and zeta (ζ) potential measurements.

3.2.11 Preparation of iodide-modified silver nanoparticles

One milliliter of silver nanoparticles was centrifuged in a 1.5 mL microcentrifuge tube at 5000 rpm for 15 minutes at room temperature. The supernatant was discarded and the resulting colloid was resuspended in 1 mL of Milli-Q water and centrifuged again. The resulting

colloidal suspension (50 μL) was then mixed with an equal volume of 12 mM potassium iodide (KI) and incubated for 24 h at room temperature in the dark. After incubation, the resulting iodide-modified nanoparticles were centrifuged at 5000 rpm for 10 minutes at room temperature and resuspended in 100 μL of milli-Q water. Resulting Ag IMNPs were characterized using UV-visible absorption spectroscopy, transmission electron microscopy, and zeta (ζ) potential measurements.

3.2.12 UV-visible absorption spectroscopy

All the UV-vis absorption spectra were collected on a Multiskan Go (Thermo Scientific) plate reader using 96-well NUNC optical bottom plates (Thermo Scientific). The total sample volume used was 100 μL for all the measurements. Twenty micromolar full-length FUS was used whereas the concentration of Ag IMNPs used was fixed to 0.1 nM and 25 ng/μL, 50 ng/μL, and 100 ng/μL of poly-U RNA were used. Background subtracted absorption spectra from 300-800 nm were normalized and plotted using Origin.

3.2.13 Dynamic light scattering (DLS)

Zeta potential measurements for iodide-modified nanoparticles were carried out on a Malvern Zetasizer Nano ZS90 instrument (Malvern, UK) using a He-Ne laser (632 nm) as an excitation source. All the measurements were carried out at room temperature and 0.05 nM of Ag IMNPs in filtered milli-Q water was used for estimating zeta potential.

3.2.14 Transmission electron microscopy (TEM)

TEM images were obtained on a Jeol JEM-F200. Three microliters of half-diluted colloidal suspension were adsorbed on a 300-mesh carbon-coated electron microscopy grid and allowed to dry overnight. Histogram for nanoparticle size distribution was created using ImageJ (NIH, Bethesda, USA) software and plotted using Origin.

3.2.15 Single-particle surface-enhanced Raman spectroscopy (SERS)

Single-droplet SERS spectra were recorded on an inVia laser Raman microscope (Renishaw, UK) at ~ 25 ^{o}C. An NIR laser (785 nm) with an exposure time of 10 sec and 500 mW (100 %) laser power was used to excite the samples. Raman scattered light was collected and dispersed using a diffraction grating (1200 lines/mm) and was further detected by an air-cooled CCD detector whereas the Rayleigh scattered light was blocked using an edge filter of 785 nm. For

single-droplet SERS measurements, phase separation of full-length FUS (20 μM) was set up in the presence of 0.1 nM Ag IMNPs with or without RNA, and 2 μL of the droplet reaction was placed on a glass slide covered with an aluminum foil. Single droplets were focused using a 50x long-working-distance objective lens (Nikon, Japan) and an NIR laser (785 nm) with an exposure time of 10 sec and 5 mW (1 %) laser power was used to excite the samples. Experiments were repeated with different batches of freshly purified protein and freshly prepared nanoparticles. Data was acquired using Wire 3.4 software provided with the Raman spectrometer. The collected Raman spectra were baseline corrected using the cubic spline interpolation method and smoothened using Wire 3.4 software and plotted using Origin.

3.3 Results

3.3.1 Experimental design for single-droplet surface-enhanced vibrational Raman spectroscopy

Our laser micro-Raman system consists of several components, namely, an excitation source comprising of a near-infrared (NIR) laser, an integrated microscope spectrometer consisting of a combined system of lenses, mirrors, filters, and a diffraction grating, and a charge-coupled device (CCD) detector (Figure 3.1). This integrated optical setup allows us to irradiate the sample and filter out the (elastic) Rayleigh scattered light and further collimate the (inelastic) Raman scattered light onto the detector to obtain a Raman spectrum. Such a design permits us to focus the laser beam of a suitable power using an objective lens into a small sub-micron spot-size within a single protein-rich droplet and acquire (regular) normal single-droplet Raman spectra. For ultra-sensitive SERS measurements, we observed that surface-modified metal nanoparticles get spontaneously encapsulated into liquid droplets as evident by an independent confocal fluorescence imaging experiment (Figure 3.1). The single-droplet SERS methodology allows us to obtain highly enhanced Raman signals within individual liquid droplets of FUS.

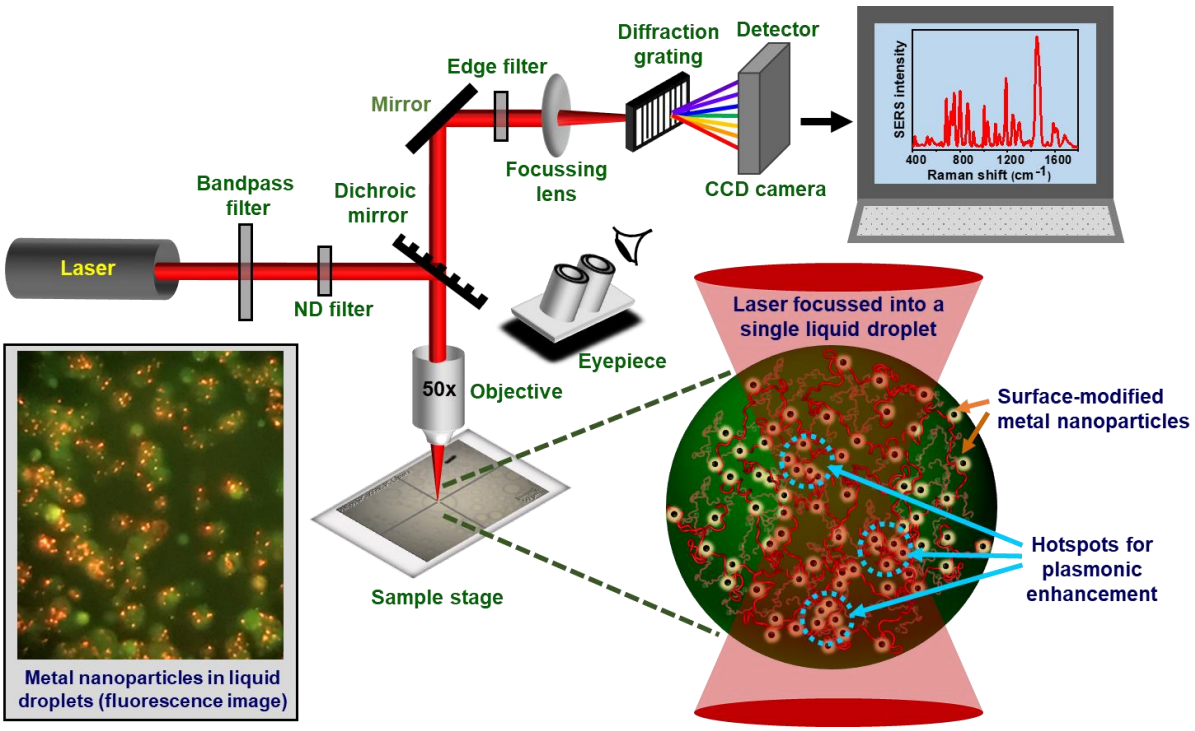


Figure 3.1: A sketch of the optical setup and a schematic of single-droplet surface-enhanced Raman spectroscopy (SERS). A near-infrared (NIR) laser gets focused within each protein-rich FUS droplet encapsulating surface functionalized silver nanoparticles through an integrated system of lenses, mirrors, and filters. Hotspots are generated within the droplets causing optical enhancement of Raman signals detected by a CCD detector. The diffraction grating shown here is for representative purpose only and does not represent the actual mechanism of dispersion. A fluorescence image taken from the eyepiece using a camera is also included to show the encapsulation of nanoparticles in the condensates.

3.3.2 Preparation of surface-modified nanoparticles for single-droplet SERS

The predicted net charge of FUS is +14.5 at physiological pH and we thus postulated that it could electrostatically interact with surface-modified negatively charged metal nanoparticles offering us an excellent system to study SERS within biomolecular condensates (amino acid sequence and domain architecture of full-length FUS is shown in chapter 2). To test our hypothesis, we started with the preparation of a suitable plasmonic SERS substrate. We chose silver nanoparticles (Ag NPs) for our experiments due to their high SERS activity and greater electromagnetic enhancements compared to other plasmonic nanomaterials. We prepared Ag NPs by a standard protocol of reduction of $AgNO_3$ by sodium citrate and characterized them

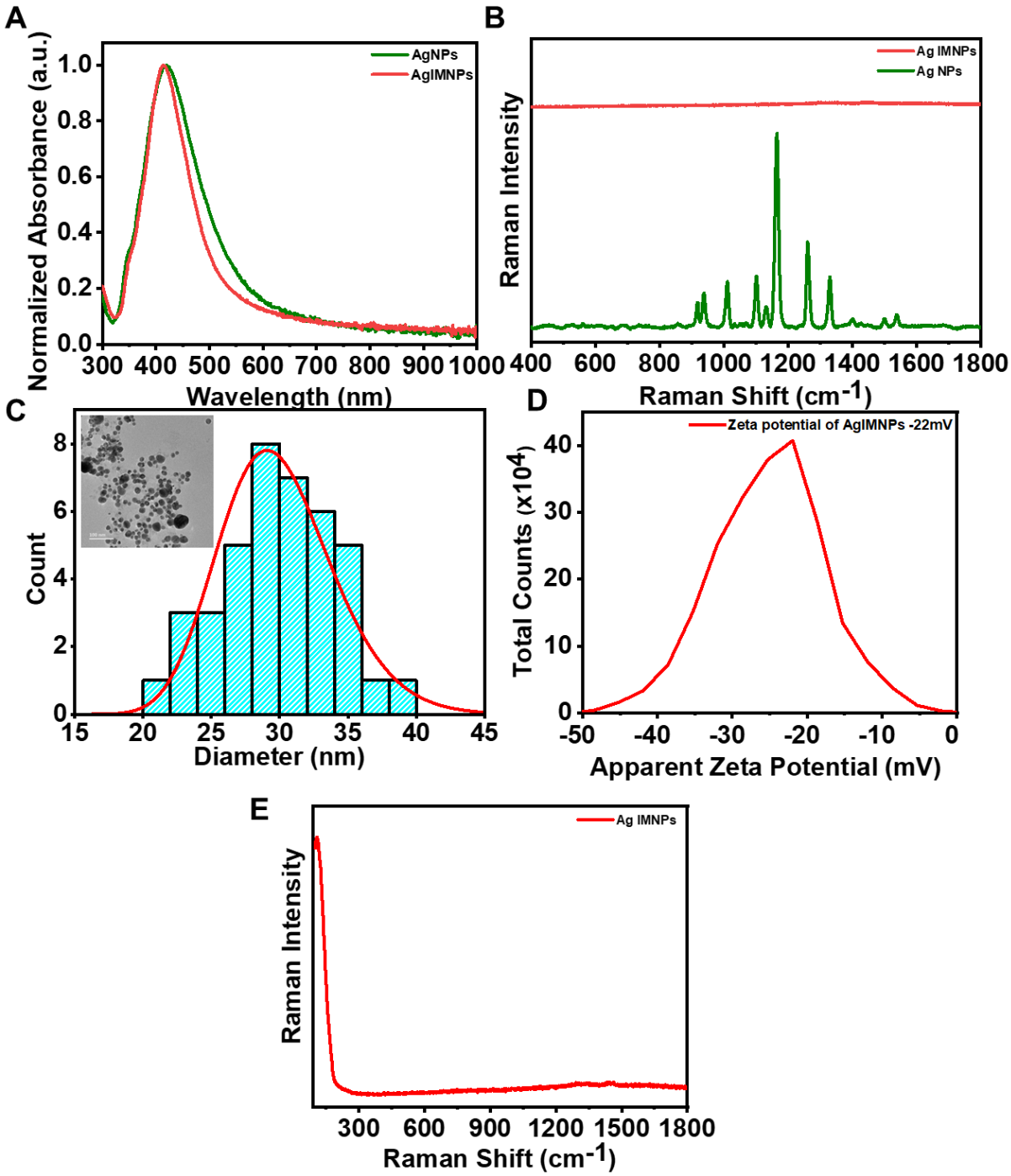


Figure 3.2: (A) UV-visible absorption spectra for silver nanoparticles (AgNPs) (olive) and iodide-modified silver nanoparticles (Ag IMNPs) (red). (B) Raman spectra for silver nanoparticles (Ag NPs) (olive) and Ag IMNPs (red). (C) Histogram for nanoparticles size distribution derived from the TEM analysis. Size analysis was performed by considering the sizes of 40 different nanoparticles. Inset shows TEM image of Ag IMNPs. Scale bar: 100 nm. (D) Zeta-potential for Ag IMNPs. Plotted here is the mean of three different measurements. (E) SERS spectrum of Ag IMNPs corresponding to the Ag-I bond at 110 cm^{-1}.

using UV-visible absorption spectroscopy, transmission electron microscopy (TEM), and zeta (ζ) potential measurements (Figure 3.2).[39] The absorption spectrum showed a single absorption band at ~ 415 nm, which corresponds to spherical nanoparticles with a diameter of ~ 30 nm (Figure. 3.2 A).[40] We next functionalized these Ag NPs with iodide to form iodide-modified silver nanoparticles (Ag IMNPs). This halide modification of nanoparticles was performed to get rid of the overwhelming citrate peaks in the Raman spectrum and aid in better attachment of the polypeptide chains to the negatively charged silver nanospheres (Figure. 3.2 B).[39] Ag IMNPs exhibited a single absorption band with λ_{max} at ~ 418 nm indicating a similar diameter of surface-modified nanoparticles (~ 30-40 nm). The size was then confirmed by TEM that revealed nanospheres with an average diameter of ~ 30 nm (Figure. 3.2 C). We next carried out zeta potential measurements to determine the effective charge on the surface of nanoparticles. The zeta potential of Ag IMNPs was -22 mV indicating an overall negative surface charge that stabilizes the nanoparticles preventing them from agglomerating into large-sized colloids (Figure. 3.2 D). Further, Raman spectra of these surface-modified nanoparticles showed a single band at 110 cm^{-1} that corresponds to Ag-I bond, indicating monolayer coating of Ag NPs (Figure. 3.2 E).[39]

We next set out to investigate the interaction of FUS with Ag IMNPs and to check their colloidal stability in the presence of the protein. We recorded UV-vis absorption spectra of Ag IMNPs in the dispersed and condensed phase of FUS at 10- and 30-minutes time points and observed a small red-shift of the absorption maxima with a slight broadening of the band as compared to only Ag IMNPs in buffer (Figure. 3.3 A, B). We chose to record the absorption spectra at these two different time points since our next set of single-droplet SERS measurements required this time interval for data acquisition. Addition of FUS monomer to Ag IMNPs causes an increase in the absorbance along with a red-shift and broadening of the absorption peak. This suggests that the NPs bind electrostatically to the monomeric protein and form smaller aggregates without altering the stability of nanoparticles.[41] In the droplet state, the broadening of the absorption peak significantly increases along with a drop in the absorbance implying extensive aggregation of NPs within the condensed phase to create hotspots necessary for plasmonic enhancements in Raman measurements. In order to test if the LLPS behavior and structure of FUS remained unaltered in the presence of nanoparticles, we performed turbidity and CD experiments (Figure. 3.3 C, D). Microscopy experiments indicated that nanoparticles do not alter the phase separation propensity of FUS and that Ag IMNPs are completely encapsulated with FUS droplets (Figure. 3.3 E, F, G). Further, size distribution

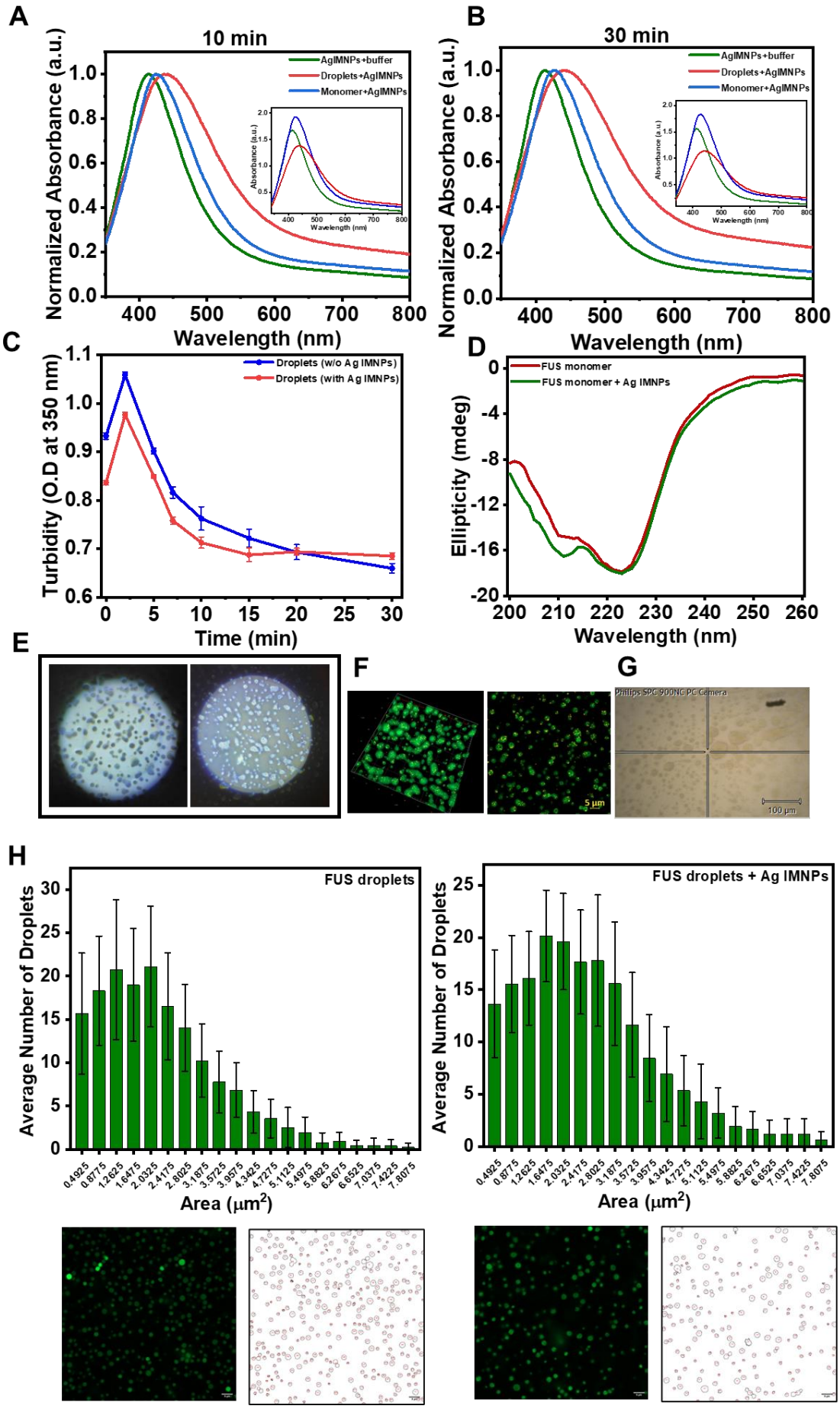

A
10 min
AgIMNPs+buffer
Droplets+AgIMNPs
Monomer+AgIMNPs
Normalized Absorbance (a.u.)
Wavelength (nm)
B
30 min
C
Droplets (w/o Ag IMNPs)
Droplets (with Ag IMNPs)
Turbidity (O.D at 350 nm)
Time (min)
D
FUS monomer
FUS monomer + Ag IMNPs
Ellipticity (mdeg)
E
F
G
H
FUS droplets
FUS droplets + Ag IMNPs
Average Number of Droplets
Area (μm²)

Figure 3.3: UV-Visible absorption spectra of Ag IMNPs in phosphate buffer (olive), monomeric FUS in the presence of Ag IMNPs (blue), FUS droplets in the presence of Ag IMNPs (red) at 10 minutes (A) and 30 minutes (B). Inset shows unnormalized UV-visible absorption spectra. (C) Turbidity plot of FUS droplets in the absence (blue) and presence (red) of Ag IMNPs (mean ± SEM; n=4). (D) Far-UV CD spectrum of monomeric FUS and FUS in the presence of Ag IMNPs (5uM protein, 0.025 nM Ag IMNPs, 20 mM sodium phosphate, pH 7.4). (E) Eye-piece image for FUS droplets without and with Ag IMNPs. (F) 3-dimensional confocal image showing the presence of Ag IMNPs within FUS droplets. Confocal microscopy image of fluorescein-5-maleimide labeled FUS and Ag IMNPs indicating encapsulation of Ag IMNPs within the droplets. (G) Encapsulation of Ag IMNPs as seen through a Raman microscope using 50x objective lens. (H) Area analysis image/projection. This is a distribution plot of middle point of class interval (0.3 to 8 μm^2) that represents area of the droplets with average number of droplets per class interval. Mean radius of FUS droplets in the absence of Ag IMNPs is 0.84 μm and in the presence of Ag IMNPs is 0.90 μm suggesting that the droplets are of similar sizes in the absence and presence of Ag IMNPs. Shown below are the confocal images and their corresponding area projection images, Scale bar: 5 μm.

analysis was performed, which showed no significant changes in the droplet dimension in the absence and presence of nanoparticles (Figure. 3.3 H). To visualize the presence of nanoparticles within the condensates, we performed two-color confocal fluorescence imaging that revealed the uptake of nanoparticles within the condensates (Figure. 3.4 A, B). Fluorescence recovery after photobleaching (FRAP) experiments on fluorescently-labeled FUS droplets revealed no significant difference in the recovery rate in the absence and the presence of nanoparticles (Figure. 3.4 C). These results indicated that the droplet interior remained mobile in the presence of nanoparticles, and the overall material property of FUS condensates remained unaltered in the presence of the SERS substrate. Together, this set of experiments suggested that FUS electrostatically interacts with surface-coated silver nanoparticles that get preferentially encapsulated into the dense phase of liquid droplets while keeping the internal mobility nearly intact. Therefore, these silver nanoparticles can act as an ideal SERS substrate for Raman enhancements within biomolecular condensates. We next directed our efforts to perform ultrasensitive SERS measurements within individual droplets.

3.3.3 Single-droplet SERS within FUS condensates

In order to record single-droplet SERS spectra, LLPS was set up in the presence of 100 pM Ag IMNPs, and a 785-nm NIR laser beam (5 mW) was focused into individual nanoparticle-

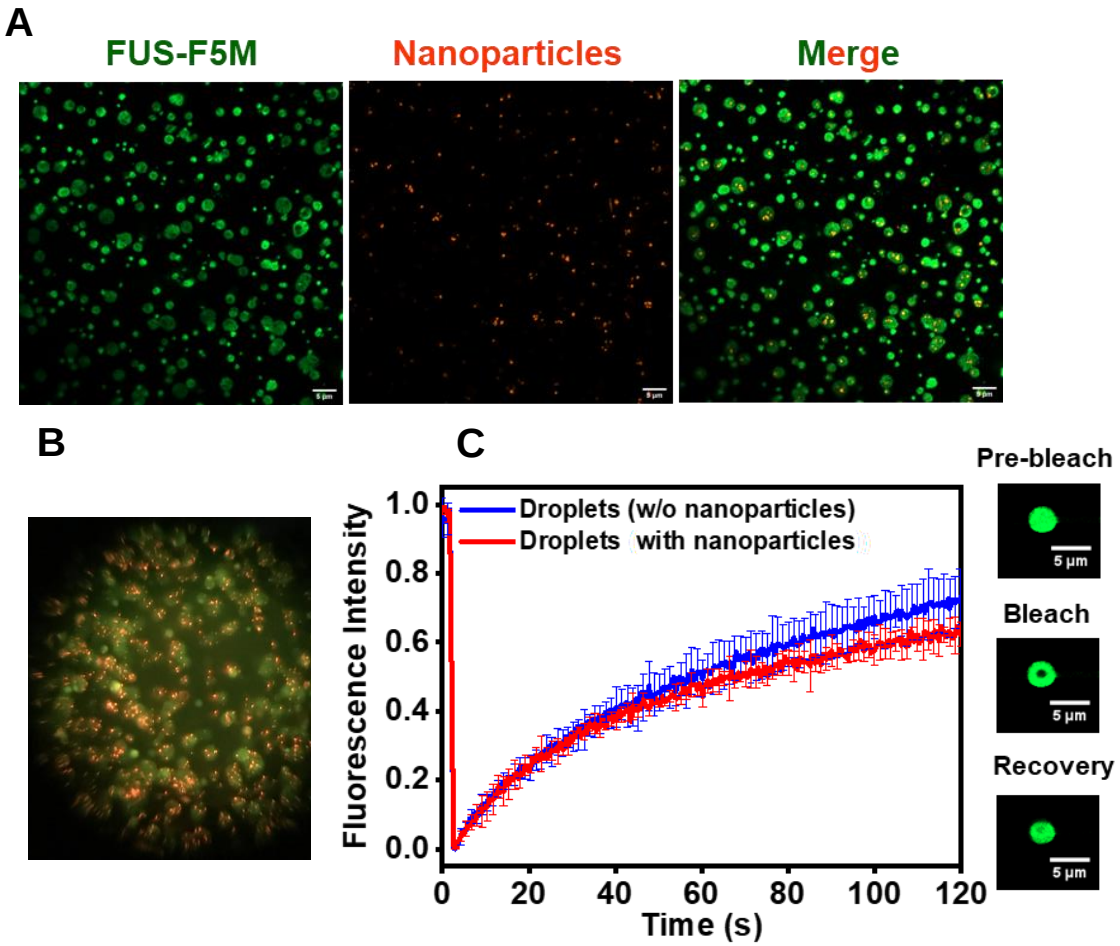


Figure 3.4: (A) Confocal images depicting encapsulation of iodide-modified silver nanoparticles (Ag IMNPs) within fluorescein-5-maleimide labeled FUS droplets. The imaging was performed thrice with similar observations. (B) Image clicked through the eyepiece using a camera indicating the encapsulation of Ag IMNPs within FUS droplets (a similar image is also shown in a scheme in Figure. 3.1. (C) FRAP kinetics of multiple droplets (~1% Alexa488-labeled protein) in the absence (blue) and presence of nanoparticles (red). The data represent mean ± s.d. for n = 3 independent experiments. Fluorescence images of droplets during FRAP measurements are shown on the right.

containing liquid droplets using a 50x objective (Figure. 3.3 G). We achieved an enhancement in the order of $\geq 10^4$ using amide III as a reference peak. We would like to note that this is an approximate (lower bound) estimation of the enhancement factor due to extremely weak signals from the droplets in the absence of nanoparticles using low power and a 50x objective (Figure. 3.5 A). The enhancement is much higher than 10^4 for the peaks that are not visible in normal Raman (Chapter 2). Interestingly, the amide I band in our SERS spectra was not enhanced and, therefore, was not visible in such a low power illumination (Figure. 3.5 B, C). According to selection rules, signal enhancement in SERS depends on the orientation of the

analyte on the surface of nanoparticles and varies inversely with the twelfth power of the distance of the target analyte from the nanoparticle surface.[42] We speculate that the interaction of bulky amino acid side chains of FUS with the negatively charged Ag IMNPs might potentially orient the backbone C=O group away from the critical near-field required for the plasmonic enhancement.[43] Nevertheless, amide III was visible at 1246 cm^{-1} and 1298 cm^{-1} corresponding to nonregular/turn structures and α-helices, respectively (Figure. 3.5 B, C). Although histidine tautomeric doublets can appear in this region, FUS being a 526-residue protein contains 525 amide bonds and only three histidine residues. Therefore, the contribution from histidine tautomeric doublets can be much lower compared to amide vibrations. Several orders of magnitude signal enhancement allowed us to observe these structures and is possibly caused by electrostatic interactions between the negatively charged surface of Ag IMNPs and arginine-rich positively charged RBD containing these secondary structural elements. In addition, a significant enhancement was also observed for backbone C-C stretch of α-helices at 938 cm^{-1}. Surprisingly, closer inspection of the tyrosine Fermi doublet showed that the lower wavenumber band at 830 cm^{-1} did not show significant enhancement, while the higher wavenumber band showed enhancement and shifted to 863 cm^{-1}. This may be attributed to the SERS selection rules, according to which the polarizability component of the 830 cm^{-1} mode may not be perpendicular to the metal surface. Further, single-droplet SERS spectra were dominated by bands at 683 cm^{-1} [methionine/δ(CH)]; 724 cm^{-1} (methionine); 749 cm^{-1} (tryptophan); 915 cm^{-1} ν(COO^-); 1003, 1032 cm^{-1} (phenylalanine); 1246, 1298 cm^{-1} (Amide III); 1588 cm^{-1} (phenylalanine/tryptophan/histidine); 1621 cm^{-1} (tyrosine) (Table 1). Additionally, we observed a highly enhanced band at 1447 cm^{-1} and a shoulder at 1465 cm^{-1} that is assigned to the bending δ(NH) of the guanidinium moiety of arginine residues and CH_2/CH_3 deformation modes, respectively.[44,45] This is in accordance with the fact that the structured C-terminal RBD of FUS contains 37 arginine residues that can facilitate its adsorption to the negatively charged SERS substrate, thereby resulting in a significant enhancement in our SERS spectra. Since, the condensates are context-dependent and the protein chain interactions vary with the solution conditions such as pH, we tested if single-droplet SERS was able to capture the changes within the condensates. These measurements showed that structural content increased upon lowering the pH from pH 8.5 to 5.5 indicating an increase in the phase separation propensity at a lower pH possibly due to increased protein-protein interactions. (Figure. 3.5 D). Such a structural transition during phase separation has also been observed for a FUS-like protein, TDP-43.[46] Taken together, this set of single-droplet SERS illuminated the inner conformational details within FUS condensates. We next asked

whether this ultra-sensitive tool can be utilized to elucidate the structural details of FUS-RNA heterotypic condensates.

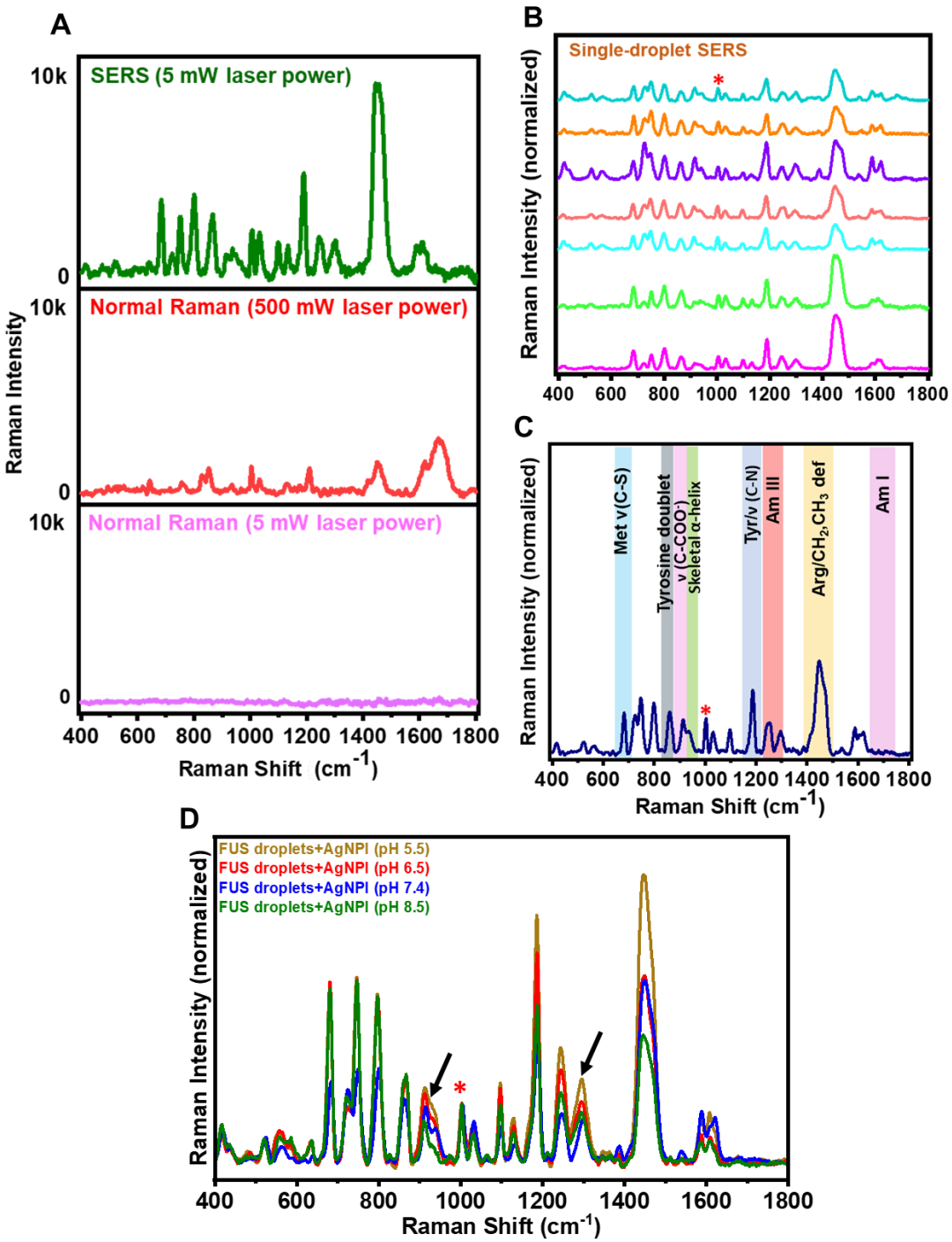


Figure 3.5: (A) Single-droplet SERS (5mW laser power), single-droplet normal Raman (500mW laser power), and single-droplet normal Raman spectra (5mW laser power) using ×50 objective lens. The observed enhancement was of the order of $\geq 10^4$ times using Amide III as a reference peak. (B) Stacked single-droplet SERS spectra of 7 FUS droplets in the presence of Ag IMNPs (spectra recorded at 5mW laser power with a ×50 objective). (C) Average single-droplet SERS spectra from individual droplets encapsulating Ag IMNPs (n = 7). All spectra are normalized with respect to the phenylalanine ring breathing band at 1003 cm^{-1} marked by

an asterisk. (D) Average single-droplet SERS spectra from individual FUS droplets encapsulating Ag IMNPs at different pH values (spectra recorded at 5 mW laser power, 50x objective, 1 accumulation; number of droplets, n = 7). All spectra are normalized with respect to the phenylalanine ring breathing band at 1003 cm^{-1} marked by an asterisk. Arrows at 940 cm^{-1} and 1296 cm^{-1} denotes greater α-helicity within FUS droplets as we go from pH 8.5 to 5.5. Peak at 940 cm^{-1} denotes backbone C-C stretch of α-helices, and peak at 1296 cm^{-1} represents Amide III band corresponding to α-helical structures.

Table 1. Raman shift values and tentative band assignments of normal Raman and SERS spectra of full-length FUS droplets.

Single-droplet Normal Raman (details in chapter 2)	**Single-droplet SERS**	**Peak assignments**[§]
1673 (s)	-	Amide I (β-sheet)
1621 (s)	1621 (s)	Tyr (R stretch)
-	1588 (s)	Phe, Trp, His
1557 (m)	-	Trp
1537 (w)	1538 (w)	Trp, Amide II
-	1447 (s)	δ(NH)-guanidinium moiety
1451 (s)	-	$\delta(CH_2/CH_3)$
-	1388 (w)	Asp, Glu $\nu_{sy}[COO^-]$
1332 (s)	-	Trp, δ(CαH)
-	1298 (s)	Amide III (α-helix)
1262 (s)	-	Amide III (nonregular/turns)
-	1246 (s)	Amide III (β-sheet)
1209 (s)	1213 (w)	Tyr [ν(C-C)]
-	1188 (s)	Tyr, Phe, ν(C-N)
1140 (s)	1132 (m)	ν_{as}(CαCN)
1101 (s)	1098 (s)	ν(C-C), ν(C-O), ν (C-N)
1042 (s)	1032 (s)	Phe [δ(R(CH))]
1006 (s)	1003 (s)	Phe R breathing
958 (w)	-	ν(N-Cα-C) skeletal
-	938 (s)	Backbone skeletal (α-helix)
-	915 (s)	$\nu(COO^-)$, C-C stretch of Pro ring

890 (s)	-	Trp (N-H bend)
857 (s)	862 (s)	Tyr Fermi doublet
830 (s)	-	Tyr Fermi doublet
798 (w)	800 (s)	ν(C-H), δ(N-H), Met [ν_{as}(C-S-C)]
767 (s)	749 (s)	Trp [$\delta(R_{breathing})$]
-	724 (s)	Met [ν(C-S)]
-	683 (s)	Met [ν(C-S)], δ(C-H)
646 (m)	640 (w)	Tyr [γ(C-C)]
623 (w)	-	Tyr, ν(C-S)
597 (m)	-	$\delta(COO^-)$
-	562 (m)	ν(S-S)
539 (s)	522 (m)	δ(skeletal), δ(N-H), ν(S-S)
461 (s)	-	ν(C-S)
-	419 (m)	Trp

[§]δ, bending; ν, stretching; R, benzene ring; as, asymmetric; sy, symmetric; w, weak; m, medium; s, strong

3.3.4 Illuminating FUS-RNA heterotypic condensates using single-droplet SERS

RNA is known to modulate the phase behavior and biophysical properties of liquid condensates formed by several RNA-binding proteins including FUS.[4,47,48] Prior to performing SERS within droplets at various concentrations of RNA, we carried UV-vis absorption spectroscopy that established the stability of nanoparticles in the presence of RNA (Figure. 3.6 A, B). Our turbidity and imaging assays showed that SERS substrate does not alter the behavior of FUS-RNA droplets (Figure. 3.6 C, D). Confocal microscopy imaging revealed complete encapsulation of nanoparticles within these droplets. We next set out to record single-droplet SERS (Figure. 3.6 E and Figure. 3.7). Figure 3.7 B depicts the stacked SERS spectra from individual droplets at varying concentrations of RNA. Interestingly, we observed a broad amide I band centered at 1682 cm^{-1} that was undetected in SERS within FUS-only droplets. This amide I peak represents disordered polypeptide conformers with some of the β structures within the FUS-RNA droplets, which is also depicted in the Raman difference plot (Figure. 3.8 A). We believe that the interactions between the negatively charged phosphate backbone of RNA

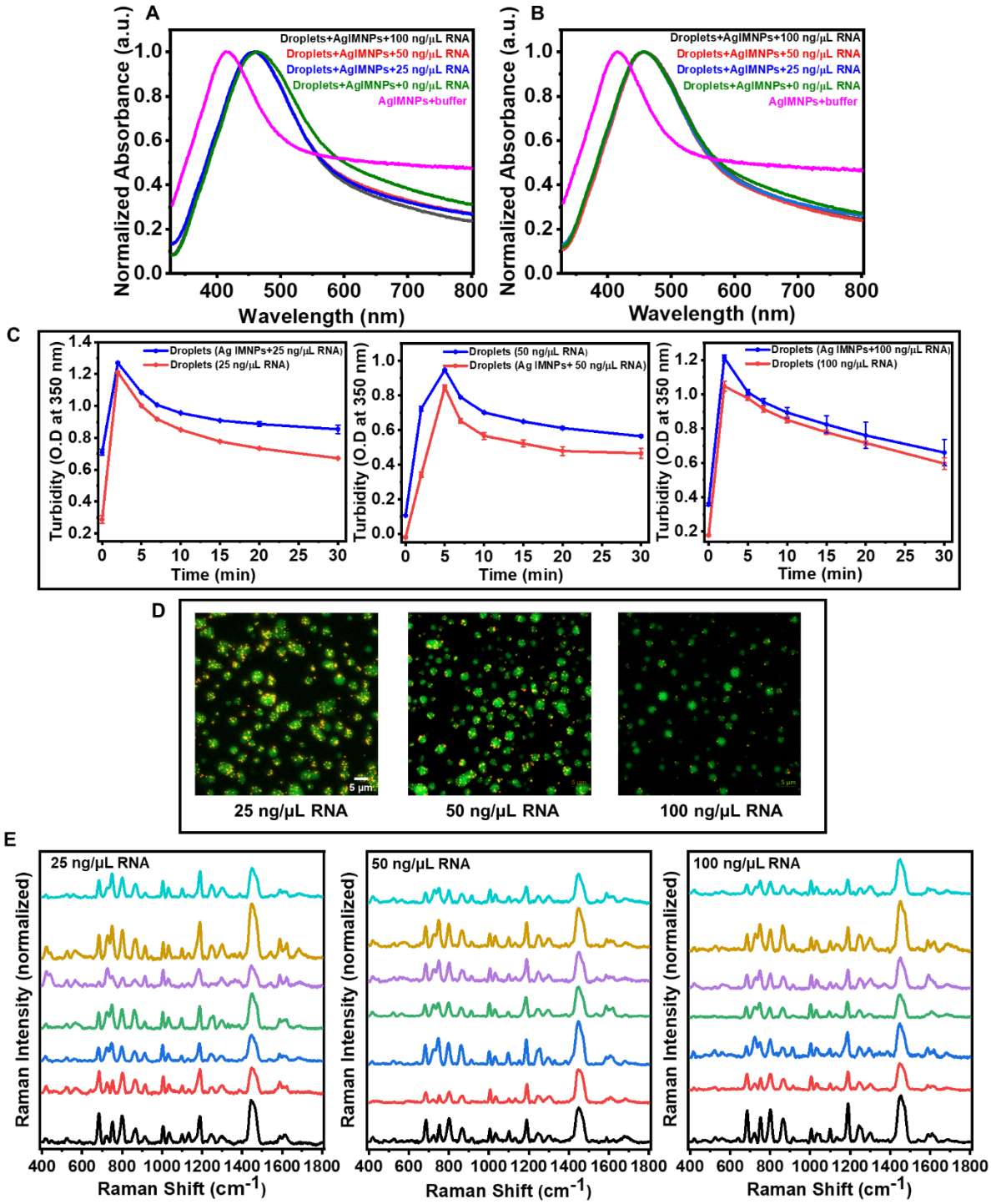


Figure 3.6: (A,B) UV-Visible absorption spectra of Ag IMNPs in phosphate buffer (magenta) and FUS droplets at different concentrations of RNA (0 ng/ µL, 25 ng/µL, 50 ng/µL, and 100 ng/µL) in the presence of Ag IMNPs at 10 minutes and 30 minutes. (C) Solution turbidity plot of FUS droplets for different concentration of RNA (25 ng/µL, 50ng/µL, and 100 ng/µL) in the absence and presence of Ag IMNPs (mean ± SEM; n= 4). (D) Confocal microscopy images of FUS droplets for different concentrations of RNA (25 ng/µL, 50 ng/µL, and 100 ng/µL) in the presence of Ag IMNPs. (E) Stacked SERS spectra of FUS droplets at different concentrations of RNA (25 ng/µL, 50 ng/µL, and 100 ng/µL) (spectra recorded at 5 mW laser power with a 50x objective; 1 accumulation and 10 sec exposure time; number of droplets, n = 7).

and positively charged C-terminal RBD alters the orientation of polypeptide chains on the surface of nanoparticles which brings the C=O groups of the polypeptide backbone in proximity to the nanoparticle surface for enhancement to occur. A closer inspection of the amide III region showed a broad band centered at 1245 cm^{-1} corresponding to β-rich and disordered/extended conformations and a band at 1300 cm^{-1} corresponding to α-helical structures for all the RNA concentrations. We observed that with an increase in RNA concentration, there is a decrease in the intensity of the amide III band at 1300 cm^{-1}, implying a reduction in the overall α-helical content indicating RNA-induced structural loss in FUS-RNA condensates (Figure. 3.8 B). Interestingly, a careful inspection of the skeletal C-C stretching mode of α-helical structures at 940 cm^{-1} showed a significant disappearance as a function of RNA (Figure. 3.8 C). This is probably because the electrostatic interactions between the protein and RNA disrupt the proposed cation-π interactions between tyrosine residues in the LC domain and arginine residues in the RBD.[14,49] We propose that the interaction between a polyanion and the FUS increases the intrinsic disorder within the polypeptide chains at the expense of α-helical structures. Moreover, we observed changes in intensities of several vibrational modes associated with aromatic residues, tyrosine, and tryptophan at 683, 724, 749, 800, 915, 1188, 1588, 1621 cm^{-1} which indicate the changes in the orientation of the aromatic ring of these residues on the nanoparticle surface in the presence of RNA. Deconvolution and analysis of the band corresponding to N-H deformations of the guanidinium moiety of arginine and shoulder band for CH_2/CH_3 deformations indicated a reduction in the enhancement of arginine residues with an increase in the RNA concentration (Figure 3.8 D, Table 2). This observation directly captures the interaction between RNA and FUS by modulating the polypeptide orientation on the SERS substrate surface.

Table 2. Percentage analysis of δ(NH)-guanidinium moiety of arginine residues and CH_2/CH_3 deformation modes obtained after deconvolution of the region 1420-1490 cm^{-1} for the single-droplet SERS spectra for various concentrations of RNA (0, 50, and 100 ng/μL).

FUS droplets with Ag IMNPs with RNA	**δNH; Guanidinium moiety (1447 cm^{-1}) (in %)**	**CH_2/CH_3 deformations (1468 cm^{-1}) (in %)**
0 ng/μL RNA	71 ± 13	44 ± 13
50 ng/μL RNA	68 ± 8	27 ± 8
100 ng/μL RNA	60 ± 10	38 ± 10

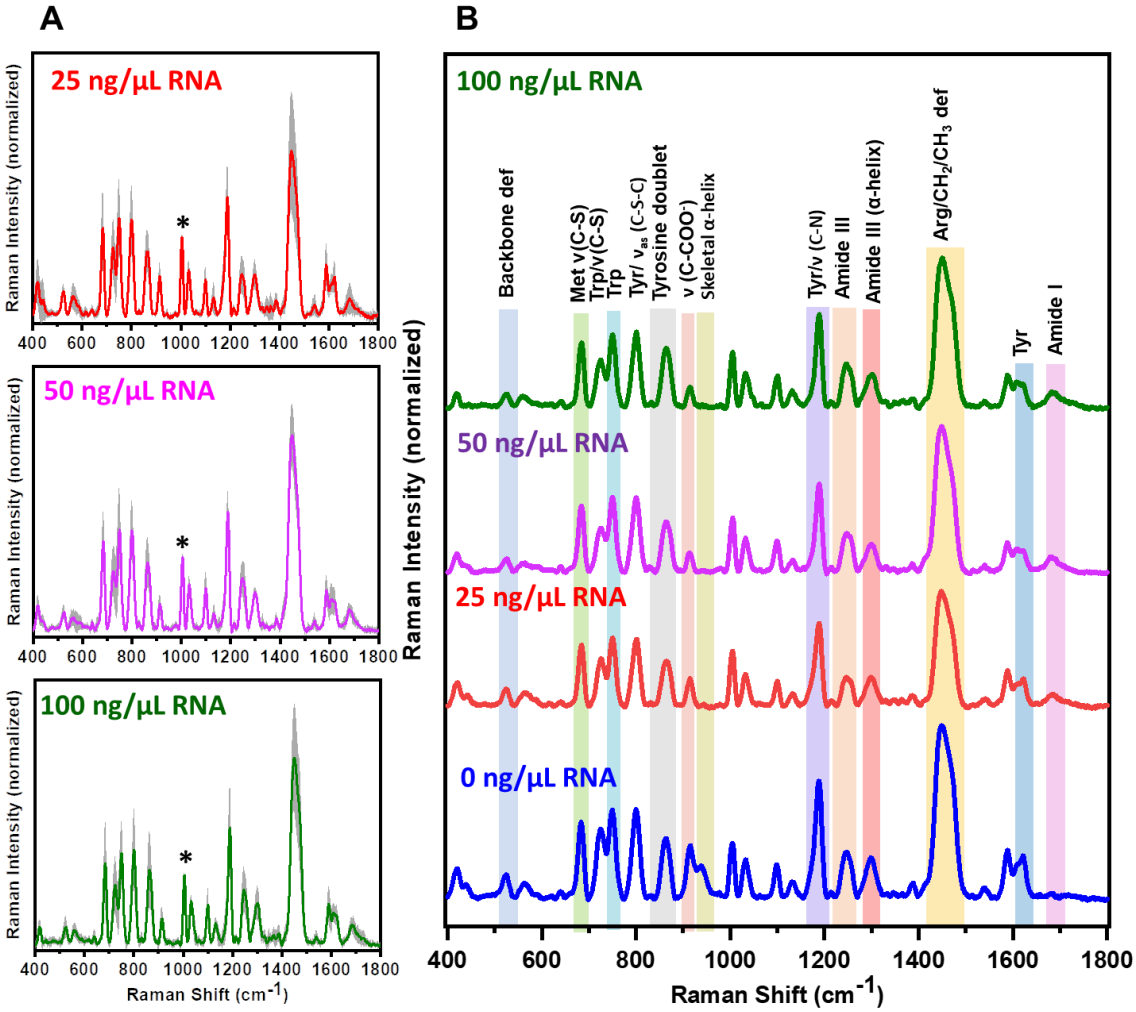


Figure 3.7: (A) Average single-droplet SERS spectra in the presence of 25 ng/µL, 50 ng/µL, and 100 ng/µL polyU RNA (spectra recorded at 5mW laser power with a ×50 objective; number of droplets, n = 7). Solid lines represent mean, whereas shaded region represents the standard deviation (n = 7). All spectra are normalized with respect to the phenylalanine ring breathing band at 1003 cm^{-1} marked by an asterisk. (B) Stacked average single droplet SERS spectra for different concentrations of RNA are shown in blue (0 ng/µL RNA), red (25 ng/µL RNA), purple (50 ng/µL RNA), and olive (100 ng/µL RNA) for comparison.

Taken together, our single-droplet SERS results illuminate some key structural details within FUS-RNA condensates and highlight RNA-mediated partial unwinding of the structured domains in the C-terminal RBD.

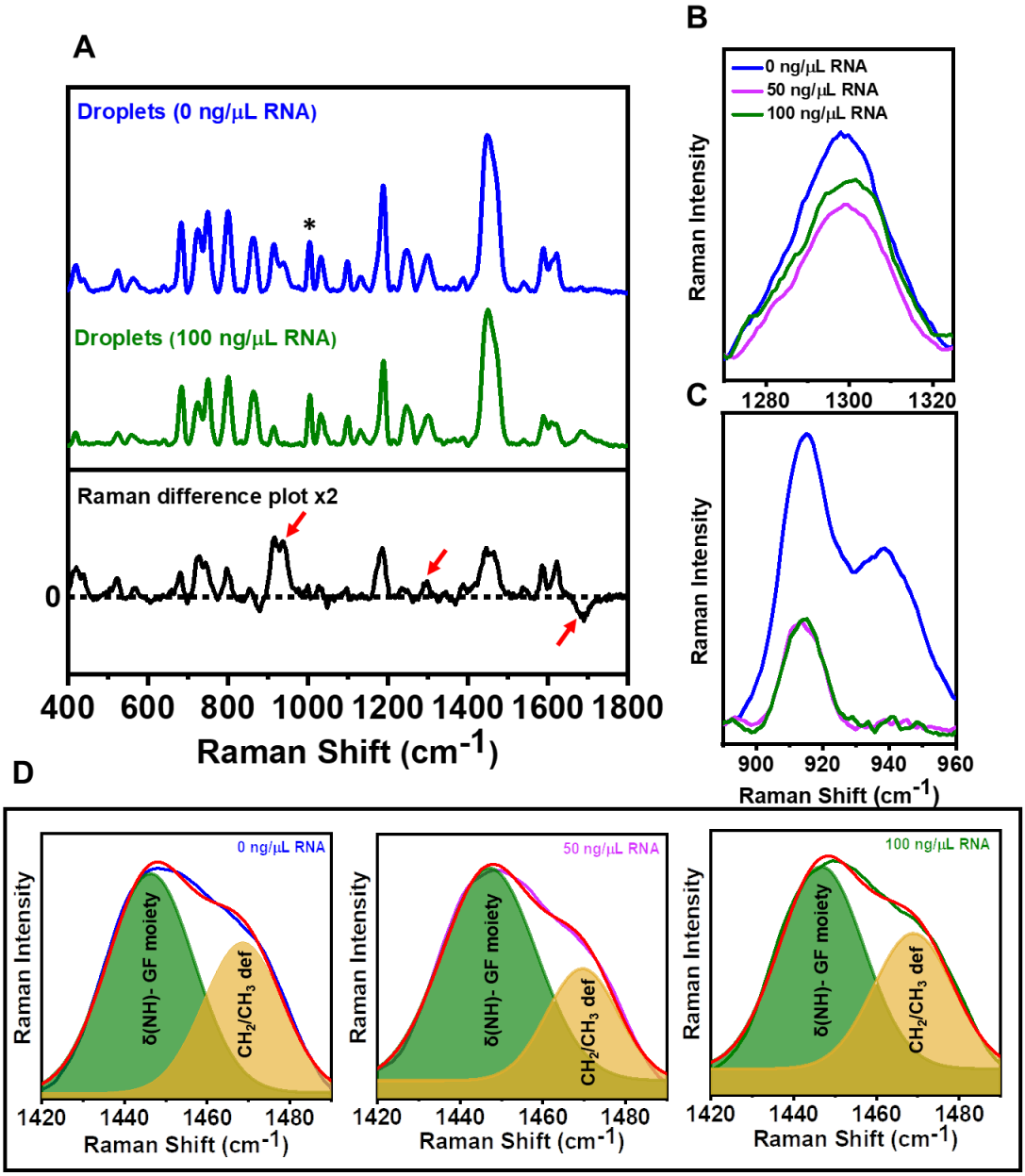


Figure 3.8: (A) Raman difference plot (between 0 ng/ μL RNA and 100 ng/ μL RNA) of single-droplet SERS spectra of droplets in the absence and presence of RNA (100 ng/μL). An arrow at 1682 cm^{-1} represents the emergence of amide I at higher RNA concentrations and arrows at 940 cm^{-1} and 1300 cm^{-1} represents greater α-helical content within droplets in the absence of RNA. (B) Zoomed in Amide III region for FUS droplets in the absence and presence of RNA (50 ng/μL, 100 ng/μL). (C) Skeletal C–C stretching mode of α-helical structures at 940 cm^{-1} that disappears at higher RNA concentrations. (D) Gaussian deconvolution of the region 1420−1490 cm^{-1}. The colored lines represent the actual data while the red line represents the cumulative fit. The Olive-green region represents the N–H deformations of the guanidinium fragment (GF) of arginine residues, while the dark-yellow region represents the CH_2/CH_3 deformations. See Table 2 for percentage analysis.

3.4 Discussion

We developed an ultrasensitive single-droplet Raman spectroscopic methodology to elucidate the inherent conformational heterogeneity and structural distribution within biomolecular condensates of FUS in a droplet-by-droplet manner. This unique methodology combines the capabilities of vibrational spectroscopy and optical microscopy offering a wealth of molecular information within the mesoscopic liquid condensed phase at the single-droplet resolution. Normal vibrational Raman spectroscopy can probe the detailed molecular structure and conformational reorganizations of the internal and external components of individual liquid droplets.[50] However, recording detailed vibrational signatures from liquid states is highly challenging due to a low Raman scattering cross-section of proteins.[27] Therefore, such measurements lack adequate sensitivity and often require unusually high concentrations, laser power, and magnifications. Such requirements can be detrimental to soft biological samples and lead to laser-induced damage and other artifacts. These limitations can be elegantly overcome by surface-engineered metal nanoparticle-induced plasmonic enhancements. The electrostatic interaction between positively charged polypeptide chains of FUS and negatively charged iodide-coated silver nanoparticles causes significant plasmonic enhancement of certain protein vibrational modes. Focusing a low-power laser beam into each droplet encapsulating surface-coated plasmonic nanostructures permitted us to record the Raman scattering bands arising due to different vibrational modes from the mesoscopic protein-rich droplets. We took advantage of the versatility of this technique to capture complex conformational characteristics of heterotypic FUS-RNA condensates at a single-droplet resolution. However, this method has certain limitations. The most critical step in our single-droplet SERS experiments is the preparation of silver plasmonic nanostructures. Since this method involves close contact between the enhancing surface and the analyte, getting rid of the surface impurities is extremely important for the homogeneity and reproducibility of the SERS signal. Additionally, different polypeptide chains carry different surface charges therefore, the same SERS substrate may not work the best for all the phase-separating systems and need to be engineered thoughtfully. The plasmonic substrates have limited re-usability and selectivity for a given analyte.

In summary, the sensitivity of our single-droplet Raman methodology can be enhanced $\geq 10^4$-fold in the SERS format. There is limited understanding regarding the mechanism of interactions between proteins and RNA and how RNA modulates the phase separation

tendency of IDRs. Our SERS spectra showed that the C-terminal RBD undergoes a partial unwinding in the presence of RNA together with a reduction in the enhancement of arginine residues. This unraveling of the ordered region in the RBD increases the polypeptide chain disorder that can promote FUS-FUS and FUS-RNA multivalent interactions within the condensed phase. It has been previously shown by smFRET that FUS-RNA interactions are mostly dynamic and multivalent, which might be the cause of enhanced phase separation as these transient interactions allow the FUS-RNA droplets to remain in the liquid state, which increases droplet fluidity and eventually dissolves with increasing RNA concentrations.[51,52] Several arginine mutations such as R244C are known to alter the phase separation propensity and bring about a conformational change in FUS that leads to extremely poor binding to RNA.[51] On the other hand, key glycine mutations such as G156E do not significantly affect the phase separation tendency of FUS rather dictates the physical properties of condensates by forming more solid-like droplets.[51] Future studies will aim at addressing such important issues through the lens of vibrational Raman spectroscopy.

Therefore, taken together single-droplet SERS allows us to zoom into the mesoscopic condensed phase to unmask the molecular determinants governing the intriguing condensate biophysics. This potent methodology also offers a unique capability and adaptability by using different surface functionalities and other metals for enhancements of unique sets of vibrational bands. Additionally, cellular uptake of these engineered nanoparticles can open new avenues to study intracellular phase transitions using vibrational spectroscopy. Such advancements will pave the way for ultrasensitive detection, characterization, and quantification of a wide range of biomolecular condensates involved in physiology and disease as well as in emerging applications in drug delivery and synthetic biology.

3.5 References

1. Alberti, S., Hyman, A. A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol*. **2021,** 22, 196–213.

2. Lyon, A. S., Peeples, W. B., Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **2021,** 22, 215–235.

3. Fuxreiter, M., Vendruscolo, M. Generic nature of the condensed states of proteins. *Nat. Cell Biol*. **2021,** 23, 587–594.

4. Roden, C., Gladfelter, A. S. RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol*. 2021, 22, 183–195.

5. Portz, B., Lee, B. L., Shorter, J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem. Sci.* **2021,** 46, 550-563.

6. Choi, J.-M., Holehouse, A. S., Pappu, R. V. Physical principles underlying the complex biology of intracellular phase transitions. *Annu. Rev. Biophys*. **2020,** 49, 107–133.

7. Shapiro, D. M., Ney, M., Eghtesadi, S. A., Chilkoti, A. Protein Phase Separation Arising from Intrinsic disorder: First-Principles to Bespoke Applications. *J. Phys. Chem. B.* **2021,** 125, 6740-6759.

8. Sabari, B. R., Dall'Agnese, A., Young, R. A. Biomolecular condensates in the nucleus. *Trends Biochem. Sci*. **2020,** 45, 961–977.

9. Nesterov, S. V., Ilyinsky, N. S., Uversky, V. N. Liquid-liquid phase separation as a common organizing principle of intracellular space and biomembranes providing dynamic adaptive responses. *Biochim. Biophys. Acta (BBA)-Molecular Cell Res*. **2021,** 1868, 119102.

10. Forman-Kay, J. D., Kriwacki, R. W., Seydoux, G. Phase Separation in Biology and Disease. *J. Mol. Biol*. **2018,** 430(23), 4603-4606.

11. Riback, J. A., Zhu, L., Ferrolino, M. C., Tolbert, M., Mitrea, D. M., Sanders, D. W., Wei, M. T., Kriwacki, R. W., Brangwynne, C. P. Composition-dependent thermodynamics of intracellular phase separation. *Nature* **2020,** 581, 209–214.

12. Ray, S., Singh, N., Kumar, R., Patel, K., Pandey, S., Datta, D., Mahato, J., Panigrahi, R., Navalkar, A., Mehra, S., Gadhe, L., Chatterjee, D., Sawner, A. S., Maiti, S., Bhatia, S., Gerez, J. A., Chowdhury, A., Kumar, A., Padinhateeri, R., Riek, R., Krishnamoorthy, G.,

Maji, S. K. α-Synuclein aggregation nucleates through liquid-liquid phase separation. *Nat. Chem*. **2020,** 12(8), 705-716.

13. Vernon, R. M., Chong, P. A., Tsang, B., Kim, T. H., Bah, A., Farber, P., Lin, H., Forman-Kay, J. D. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **2018, 7**, e31486.

14. Qamar, S., Wang, G., Randle, S. J., Ruggeri, F. S., Varela, J. A., Lin, J. Q., Phillips, E. C., Miyashita, A., Williams, D., Ströhl, F., Meadows, W., Ferry, R., Dardov, V. J., Tartaglia, G. G., Farrer, L. A., Kaminski Schierle, G. S., Kaminski, C. F., Holt. C. E., Fraser. P. E., Schmitt-Ulms, G., Klenerman, D., Knowles, T., Vendruscolo, M., St George-Hyslop, P. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation- π interactions. *Cell* **2018,** 173, 720–734.

15. Martin, E. W., Holehouse, A. S., Peran, I., Farag, M., Incicco, J. J., Bremer, A., Grace, C. R., Soranno, A., Pappu, R. V., Mittag, T. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **2020,** 367, 694–699.

16. Wang, J., Choi, J. M., Holehouse, A. S., Lee, H. O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., Poser, I., Pappu, R. V., Alberti, S., Hyman, A. A. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **2018,** 174, 688–699.

17. Uversky, V. N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Curr. Opin. Struct. Biol*. **2017,** 44, 18–30.

18. Gomes, E., Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem.* **2019,** 294(18), 7115-7127.

19. Ganser LR, Myong S. Methods to Study Phase-Separated Condensates and the Underlying Molecular Interactions. *Trends Biochem Sci*. **2020,** 45(11), 1004-1005.

20. Alberti, S., Gladfelter, A., Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019,** 176(3), 419-434.

21. Conicella, A. E., Dignon, G. L., Zerze, G. H., Schmidt, H. B., D'Ordine, A. M., Kim, Y. C., Rohatgi, R., Ayala, Y. M., Mittal, J., Fawzi, N. L. TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** 117(11), 5883-5894.

22. Rai, S. K., Savastano, A., Singh, P., Mukhopadhyay, S., Zweckstetter, M. Liquid-liquid phase separation of tau: From molecular biophysics to physiology and disease. *Protein science: a publication of the Protein Society* **2021,** 30(7), 1294–1314.

23. Emmanouilidis, L., Esteban-Hofer, L., Damberger, F. F., de Vries, T., Nguyen, C. K. X., Ibáñez, L. F., Mergenthal, S., Klotzsch, E., Yulikov, M., Jeschke, G., Allain, F. H. NMR and EPR reveal a compaction of the RNA-binding protein FUS upon droplet formation. *Nat. Chem. Biol*. **2021,** 17(5), 608-614.

24. Benevides, J. M., Overman, S. A., Thomas, G. J. Raman spectroscopy of proteins. *Curr. Protoc. Protein Sci*. **2004,** Chapter 17, Unit 17.8.

25. Chatterjee, S., Kan, Y., Brzezinski, M., Koynov, K., Regy, R. M., Murthy, A. C., Burke, K. A., Michels, J. J., Mittal, J., Fawzi, N. L., Parekh, S. H. Reversible Kinetic Trapping of FUS Biomolecular Condensates. *Adv. Sci*. **2021,** e2104247.

26. Shuster, S. O., Lee, J. C. Watching liquid droplets of TDP-43_{CTD} age by Raman spectroscopy. *J. Biol. Chem.* **2021,** 101528.

27. Rygula, A., Majzner, K., Marzec, K. M., Kaczor, A., Pilarczyk, M., Baranska, M. Raman spectroscopy of proteins: a review. *J. Raman Spectrosc*. **2013,** 44, 1061-1076.

28. Feliu, N., Hassan, M., Garcia Rico, E., Cui, D., Parak, W., Alvarez-Puebla, R. SERS Quantification and Characterization of Proteins and Other Biomolecules. *Langmuir* **2017,** 33(38), 9711-9730.

29. Langer, J., Jimenez de Aberasturi, D., Aizpurua, J., Alvarez-Puebla, R. A. *et al.* Present and Future of Surface-Enhanced Raman Scattering. *ACS Nano.* **2020,** 14(1), 28-117.

30. Zong, C., Xu, M., Xu, L. J., Wei, T., Ma, X., Zheng, X. S., Hu, R., Ren, B. Surface-Enhanced Raman Spectroscopy for Bioanalysis: Reliability and Challenges. *Chem. Rev.* **2018,** 118(10), 4946-4980.

31. Aggarwal, S., Mondal, S., Siddhanta, S., Bharat, E., Nagamalleswari, E., Nagaraja, V., Narayana, C. Divalent Ion-Induced Switch in DNA Cleavage of KpnI Endonuclease Probed through Surface-Enhanced Raman Spectroscopy. *J. Phys. Chem. B.* **2021,** 125(9), 2241-2250.

32. Siddhanta, S., Narayana, C. Surface Enhanced Raman Spectroscopy of Proteins: Implications for Drug Designing. *Nanomater. Nanotechnol.* **2012,** 2, 1-12.

33. Svetoni, F., Frisone, P., Paronetto, M. P. Role of FET proteins in neurodegenerative disorders. *RNA biol.* **2016,** 13(11), 1089–1102.

34. Tüű-Szabó, B., Hoffka, G., Duro, N., Fuxreiter, M. Altered dynamics may drift pathological fibrillization in membraneless organelles. *Biochim. Biophys. Acta, Proteins Proteomics* **2019,** 1867, 988−998.

35. Murthy, A. C., Dignon, G. L., Kan, Y., Zerze, G. H., Parekh, S. H., Mittal, J., Fawzi, N. L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019,** 26(7), 637-648.

36. Loughlin, F. E., Lukavsky, P. J., Kazeeva, T., Reber, S., Hock, E. M., Colombo, M., Von Schroetter, C., Pauli, P., Cléry, A., Mühlemann, O., Polymenidou, M., Ruepp, M. D., Allain, F. H. The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. *Mol. Cell* **2019,** 73(3), 490-504.

37. Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., NiednerBoblenz, A., Schifferer, M., Ruepp, M. D., Simons, M., Niessing, D., Madl, T., Dormann, D. Phase Separation of FUS

Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* **2018,** 173, 706−719.

38. Dhakal, S., Wyant, C. E., George, H. E., Morgan, S. E., Rangachari, V. Prion-like C-Terminal Domain of TDP-43 and α-Synuclein Interact Synergistically to Generate Neurotoxic Hybrid Fibrils. *J. Mol. Biol*. **2021,** 433(10), 166953.

39. Xu, L. J., Zong, C., Zheng, X. S., Hu, P., Feng, J. M., Ren, B. Label-free detection of native proteins by surface-enhanced Raman spectroscopy using iodide-modified nanoparticles. *Anal. Chem*. **2014,** 86(4), 2238-2245.

40. Paramelle, D., Sadovoy, A., Gorelik, S., Free, P., Hobley, J., Fernig, D. G. A rapid method to estimate the concentration of citrate capped silver nanoparticles from UV-visible light spectra. *Analyst* **2014,** 139(19), 4855-4861.

41. Wang Y, Ni Y. New insight into protein-nanomaterial interactions with UV-visible spectroscopy and chemometrics: human serum albumin and silver nanoparticles. *Analyst* **2014,** 139(2), 416-424.

42. Kneipp, K., Kneipp, H., Itzkan, I., Dasari, R. R., Feld, M. S. Surface-Enhanced Raman Scattering and Biophysics. *J. Phys.: Condens. Matter* **2002,** 14, R597-R624.

43. Kurouski, D., Postiglione, T., Deckert-Gaudig, T., Deckert, V., Lednev, I. K. Amide I vibrational mode suppression in surface (SERS) and tip (TERS) enhanced Raman spectra of protein specimens. *Analyst* **2013,** 138(6), 1665-1673.

44. Aliaga, A. E., Garrido, C., Leyton, P., Diaz, G., Gomez-Jeria, J. S., Aguayo, T., Clavijo, E., Campos-Vallette, M. M., Sanchez-Cortes, S. SERS and theoretical studies of arginine. *Spectrochim. Acta A Mol. Biomol. Spectrosc*. **2010,** 76(5), 458-463.

45. Garrido, C., Aguayo, T., Clavijo, E., Gomez-Jeria, J. S., Campos-Vallette, M. M. The effect of the pH on the interaction of L-arginine with colloidal silver nanoparticles. A Raman and SERS study. *J. Raman Spectrosc.* **2013,** 44, 1105-1110.

46. Conicella, A. E., Dignon, G. L., Zerze, G. H., Schmidt, H. B., D'Ordine, A. M., Kim, Y. C., Rohatgi, R., Ayala, Y. M., Mittal, J., Fawzi, NL. TDP-43 α-helical structure tunes liquid-liquid phase separation and function. *Proc. Natl. Acad. Sci. U. S. A.* **2020,** 117(11), 5883-5894.

47. Sanders, D. W., Kedersha, N., Lee, D. S. W., Strom. A. R., Drake, V., Riback, J. A., Bracha, D., Eeftens, J. M., Iwanicki, A., Wang, A., Wei, M. T., Whitney, G., Lyons, S. M., Anderson, P., Jacobs, W. M., Ivanov, P. & Brangwynne, C. P. Competing protein- RNA interaction networks control multiphase intracellular organization. *Cell* **2020,** 181, 306–324.

48. Silva, J. L., Cordeiro, Y. The "Jekyll and Hyde" Actions of Nucleic Acids on the Prion-like Aggregation of Proteins. *J. Biol. Chem*. **2016,** 291(30), 15482-15490.

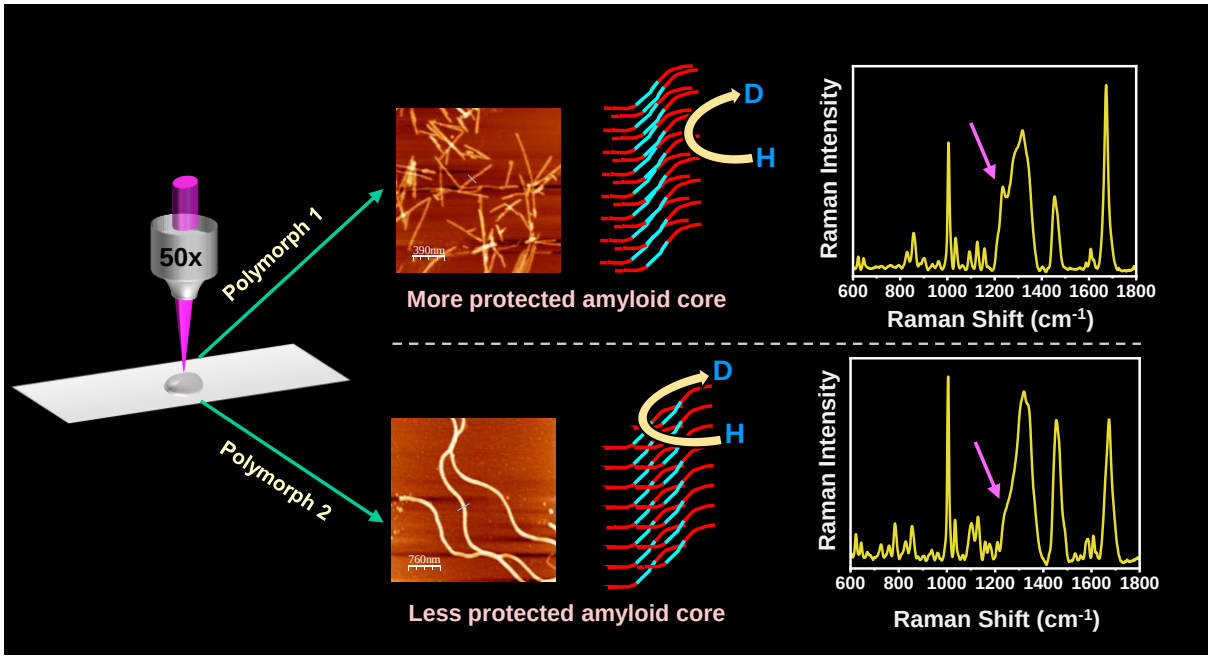
49. Kaur, T., Raju, M., Alshareedah, I., Davis, R. B., Potoyan, D. A., Banerjee, P. R. Sequence-encoded and composition-dependent protein-RNA interactions control multiphasic condensate morphologies. *Nat. Commun*. **2021,** 12(1), 872.

50. Murakami, K., Kajimoto, S., Shibata, D., Kuroi, K., Fujii, F., Nakabayashi, T. Observation of liquid-liquid phase separation of ataxin-3 and quantitative evaluation of its concentration in a single droplet using Raman microscopy. *Chem. Sci.* **2021,** 12(21), 7411-7418.

51. Rhine, K., Makurath, M. A., Liu, J., Skanchy, S., Lopez, C., Catalan, K. F., Ma, Y., Fare, C. M., Shorter, J., Ha, T., Chemla, Y. R., Myong, S. ALS/FTLD-Linked Mutations in FUS Glycine Residues Cause Accelerated Gelation and Reduced Interactions with Wild-Type FUS. *Mol. Cell.* **2020,** 80(4), 666-681.

****

Hydrogen-Deuterium Exchange Vibrational Raman Spectroscopy Distinguishes Distinct Amyloid Polymorphs Comprising Altered Core Architecture



Avni. A.; Joshi. A.; Mukhopadhyay. S. Hydrogen-Deuterium Exchange Vibrational Raman Spectroscopy Distinguishes Distinct Amyloid Polymorphs Comprising Altered Core Architecture *(Manuscript submitted).*

4.1 Introduction

Amyloids are proteinaceous fibrillar aggregates possessing a hydrogen-bonded cross-β-sheet architecture and are implicated in various neurological human proteinopathies and are also known to perform critical physiological functions in a diverse range of organisms.[1-12] Recent studies involving cell lines, animal models, and human brain extracts have revealed that these amyloid fibrils exhibit molecular-level polymorphism that is often responsible for distinct amyloid strains. Such structural variations observed within the amyloid aggregates formed from the same precursor protein can potentially account for the observed disease phenotypes.[13,14,15] A growing body of research has revealed that an intrinsically disordered presynaptic neuronal protein, α-synuclein, self-assembles into various structurally distinct fibrillar assemblies *in vitro* and *in vivo* and exhibits prion-like strain phenomema.[16-20] α-synuclein forms higher molecular weight assemblies progressively in the nerve and glial cells to form intracellular pathological lesions termed Lewy bodies and Lewy neurites which are the pathological hallmark of a diverse group of debilitating neurodegenerative disorders termed synucleinopathies such as Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy.[21,22] The exceptional protein plasticity of α-synuclein allows it to populate multiple conformation states that can self-replicate and propagate *in vivo* akin to prions which elucidate its association with various neurodegenerative diseases with distinct clinical and pathological phenotypes. *In vitro* fibril growth conditions such as temperature, incubation times, buffer compositions, posttranslational modifications, and the presence of cofactors, as well as variation in the primary amino acid sequence of the protein, can alter the molecular interactions between the polypeptide chains and influence the conformations and pathology of distinct fibrillar assemblies which can either propagate as an amyloid strain or simply form polymorphs.[23-27]

A plethora of high-resolution structural tools such as solid-state NMR, cryo-electron microscopy (Cryo-EM), X-ray, microcrystal electron diffraction, hydrogen/deuterium exchange NMR (HDX-NMR), electron paramagnetic resonance (EPR) and so on have been widely used to gain structural insights and distinguish various polymorphic structures of *in vitro* or *ex vivo* generated fibrils associated with neurodegenerative and systemic diseases as well as functional amyloids.[28-35] These structural methods can differentiate fibril polymorphs in terms of the fibril diameter, number and packing of protofilaments, presence of twists, side-chain interactions, and so forth, to provide atomic-resolution details of the supramolecular packing arrangement of the amyloid core.[36-38] However, obtaining high-resolution atomic

structures can be quite challenging since these tools either require site-specific ^{13}C and/or ^{15}N labeling of proteins or involve extensive sample preparation. Additionally, techniques such as cryo-EM demand expensive analysis software and are highly dependent on expert skills and knowledge.[39]

In this work, we utilize vibrational Raman spectroscopy coupled with H/D exchange to characterize and distinguish the fibrillar conformations of α-synuclein fibril polymorphs generated *in vitro* under different assembly conditions based on their conformation and unique Raman signatures. H/D exchange Raman spectroscopy proves to be a simple and elegant tool that can capture the molecular bond vibrations along the polypeptide backbone and its side chains, thereby characterizing the structural states of the protein in a direct, non-destructive, and label-free fashion.[40-43] Hydrogen-deuterium exchange in polypeptide causes a rapid H→D exchange in the unstructured regions lacking stable and strong hydrogen bonding while the highly structured cross-β-structural motif undergoes exchange rather slowly.[44,45] Here, we demonstrate that Raman spectroscopy coupled with H/D exchange can potentially discern the conformational differences and probe the structural heterogeneity within highly ordered cross-β-core of α-synuclein polymorphs exhibiting noticeable differences in the hydrogen bonding strength and supramolecular packing within the β-rich cores and side-chain packing.

4.2 Experimental details

4.2.1 Materials

Ampicillin, chloramphenicol, and isopropyl-thiogalactopyranoside (IPTG) were obtained from Gold Biocom (U.S.A). Thioflavin T (ThT), sodium chloride (NaCl), sodium phosphate dibasic dihydrate, Tris(hydroxymethyl)aminomethane (Trizma base), potassium chloride (KCl), 2-(N-Morpholino) ethanesulfonic acid monohydrate (MES), Deuterium oxide (D_2O) were bought from Sigma Aldrich (St. Louis, MO). EDTA (Ethylenediaminetetraacetic acid) and ammonium sulfate were procured from HIMEDIA Laboratories. Streptomycin sulfate and glacial acetic acid were purchased from MP Biomedicals and Merck, U.S.A, respectively. Membrane filters and protein concentrators were purchased from Merck, Millipore. Q-Sepharose resin was obtained from GE Healthcare Life Sciences (U.S.A). NUNC 96-well plates were purchased from ThermoFisher Scientific (Waltham, Massachusetts, U.S.A). All the buffer and salt solutions were prepared in Milli-Q water and filtered (0.22 μm filter) before use. The pH of all the buffers was adjusted on a Metrohm (Herisau Switzerland) 827 lab pH meter at room temperature.

4.2.2 Recombinant protein expression and purification

Full-length wild-type human α-synuclein was expressed in Escherichia coli BL21 (DE3) pLysS strain. Transformed cells were grown at 37 ºC, 220 rpm till optical density at 600 nm (O.D.600) reached 0.5-0.6 and then induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37 ºC. The cells were pelleted by centrifugation at 4000 rpm for 30 min at 4 ºC and lysed using lysis buffer (50mM Tris, 150mM NaCl, 10 mM EDTA, pH 8.0) and then boiled at 100 ºC for 30 min. The cell lysate was centrifuged at 11,500 rpm for 30 min at 4 ºC. To the supernatant, 136 µL/mL of 10% streptomycin sulfate and 226 µL/mL glacial acetic acid were added to precipitate out the nucleic acids. Precipitated DNA was removed by centrifugation at 11,500 rpm for 30 min at 4 ºC. The supernatant was incubated with 50% saturated ammonium sulfate at 4 ºC for 2 h with gentle agitation. The solution containing the precipitated protein was centrifuged at 11,500 rpm for 30 min at 4 ºC. The protein pellet thus obtained was washed with 100 mM ammonium acetate and chilled ethanol in a 1:1 ratio and centrifuged at 6000 rpm for 15 min at 4 ºC. The pellet was further washed with chilled ethanol (5 mL) and centrifuged at 6000 rpm for 15 min at 4 ºC. The dried protein pellet was resuspended in the equilibration buffer (10 mM Tris, pH 7.4) and poured into a pre-equilibrated Q-Sepharose column. The protein was eluted using a linear gradient of 100% final concentration of buffer B (10 mM Tris, 1 M NaCl, pH 7.4). Fractions containing the pure protein were pooled and dialyzed overnight in the buffer; 20 mM MES, 50 mM NaCl, pH 6.5 for Polymorph A, buffer; 50 mM Tris, 50 mM KCl, pH 7.5 for Polymorph B, and buffer; 10 mM Na_2HPO_4, 50 mM NaCl, pH 7.0 for Polymorph C and stored at -80 ºC for further use. Freshly purified protein was used each time for setting up the aggregation reactions to avoid multiple freeze-thaw cycles.

4.2.3 Aggregation assay

Hundred micromolar of purified α-synuclein was incubated with respective aggregation buffer (20 mM MES, 150 mM NaCl, pH 6.5 for Polymorph A; 50 mM Tris, 150 mM KCl, pH 7.5 for Polymorph B; 10 mM Na_2HPO_4, 150 mM NaCl, pH 7.0 for Polymorph C) at 37 ºC in a NUNC 96-well plate (ThermoFisher Scientific) on a POLARstar Omega Plate Reader Spectrophotometer. A total reaction volume of 200 µL was added in each well and stirred using a glass bead at 600 rpm. Aggregation was monitored by adding 2.5 µM ThT to the reaction mixtures. For Raman measurements, aggregation reactions were set up without ThT in the reaction mixture. Aggregation kinetic profiles were plotted using Origin 2022.

4.2.4 ThT fluorescence measurements

Aggregation reactions (Reaction volume: 800 µL) were pooled and centrifuged at 25,000 g for 30 min at 15 ºC. The supernatant was discarded entirely, and the pellets were suspended in 200 µL of 20 mM sodium phosphate buffer, pH 7.4. Measurements were performed on a FluoroMax-4 spectrofluorometer (Horiba JobinYvon, NJ) at room temperature using quartz cuvette of 1 mm path length. Samples were excited at 440 nm and the emission spectra were scanned from 460 nm to 520 nm. Emission maxima were plotted in Origin for the three polymorphs.

4.2.5 Atomic force microscopy

Aggregation reactions (Reaction volume: 2000 µL) were pooled and centrifuged at 25,000 g for 30 min at 15 ºC. The supernatant was discarded entirely, and the pellets were suspended in 200 µL of 20 mM sodium phosphate buffer, pH 7.4. Ten microliters (1:1 diluted) of the suspension was deposited on freshly cleaved and filtered water-washed muscovite mica (Grade V-4 mica from SPI, PA). The samples were incubated for 2 min and washed twice with 200 µL of filtered water. Samples were further dried under a gentle stream of nitrogen. Images were acquired in NanoDrive (v8.03) software at a resolution of 512 × 512 pixels. Images were further processed and analyzed using WSxM 5.0 Develop 9.1 software.

4.2.6 Circular dichroism (CD) measurements

Aggregation reactions (Reaction volume: 2000 µL) were pooled and centrifuged at 25,000 g for 30 min at 15 ºC. Pellets were suspended in 20 mM sodium phosphate buffer, pH 7.4. The far-UV CD spectra of the fibrils (15 µM) were recorded on a Chirascan CD spectrophotometer (Applied photophysics, UK) using a quartz cuvette of 1 mm path length in the scan range 200-260 nm with 1 nm step size. The spectra were averaged over 3 scans and corrected using appropriate blank spectra for all the samples. The spectra were corrected and smoothened using the ProData software provided with the Spectrometer and plotted using Origin 2022.

4.2.7 Raman spectroscopy

The Raman spectra were recorded on an inVia laser Raman microscope (Renishaw, UK). The pelleted fibrils (Reaction volume: 2000 µL) were directly deposited on a glass slide covered with aluminum foil and air-dried. An NIR laser with an exposure time of 10 sec and 50 mW

(10 %) laser power was focused on the aggregates using a 50× long-working-distance objective lens (Nikon, Japan). Rayleigh scattered light was blocked using an edge filter of 785 nm, whereas the Raman scattered light was detected by an air-cooled CCD detector following dispersion using a 1200 lines/mm diffraction grating. Raman spectra were recorded at least thrice with fibrils prepared from different batches of purified protein, and each spectrum was averaged over 20 scans. Data was acquired using Wire 3.4 software provided with the instrument, and further baseline was corrected using the cubic spline interpolation method and smoothened using Wire 3.4 software and plotted using Origin 2022.

Curve fitting analysis of amide bands and principal component analysis:

Deconvolution of the amide peaks was done to estimate the changes in the secondary structural content. Amide band curve fitting was performed using the peak analyzer option under the peak and baseline option in Origin 2022. Fitting was carried out using the Levenberg-Marquadt non-linear least squares method to obtain multiple fitted Gaussian curves and a constant baseline. Principal component analysis (PCA) was performed by the "Principal component analysis for spectroscopy" app in Origin 2022.

4.3 Results

4.3.1 Generation of distinct α-synuclein polymorphs

As a prelude, we started by generating different fibrillar polymorphs from recombinantly expressed and purified α-synuclein *in vitro* under different aggregation conditions[46,27] (Figure 4.1). Aggregation was monitored by performing an assay using Thioflavin-T (ThT), a well-known amyloid reporter. At pH 6.5, α-synuclein aggregates rapidly and exhibits a short lag phase of about 5 hours. On the contrary, at pH 7.5, α-synuclein monomers exhibit a long lag phase that takes around 3 days to saturate, while fibrils formed at pH 7.0 take around 2 days to mature (Figure 4.1 A,B). Fibrillar conformations formed at pH 6.5 are referred to as "Polymorph A", those formed at pH 7.5 as "Polymorph B", and once formed at pH 7.0 as "Polymorph C" in this study. A 14-fold higher ThT fluorescence intensity for Polymorph A was observed compared to Polymorph C, whereas Polymorph B showed approximately 8-fold greater ThT fluorescence relative to Polymorph C (Figure 4.1 C). This observation indicated weaker internal packing and H-bonding within the amyloid core for Polymorph C. Next, in order to visualize the nanoscale morphologies of fibrillar polymorphs generated, we performed

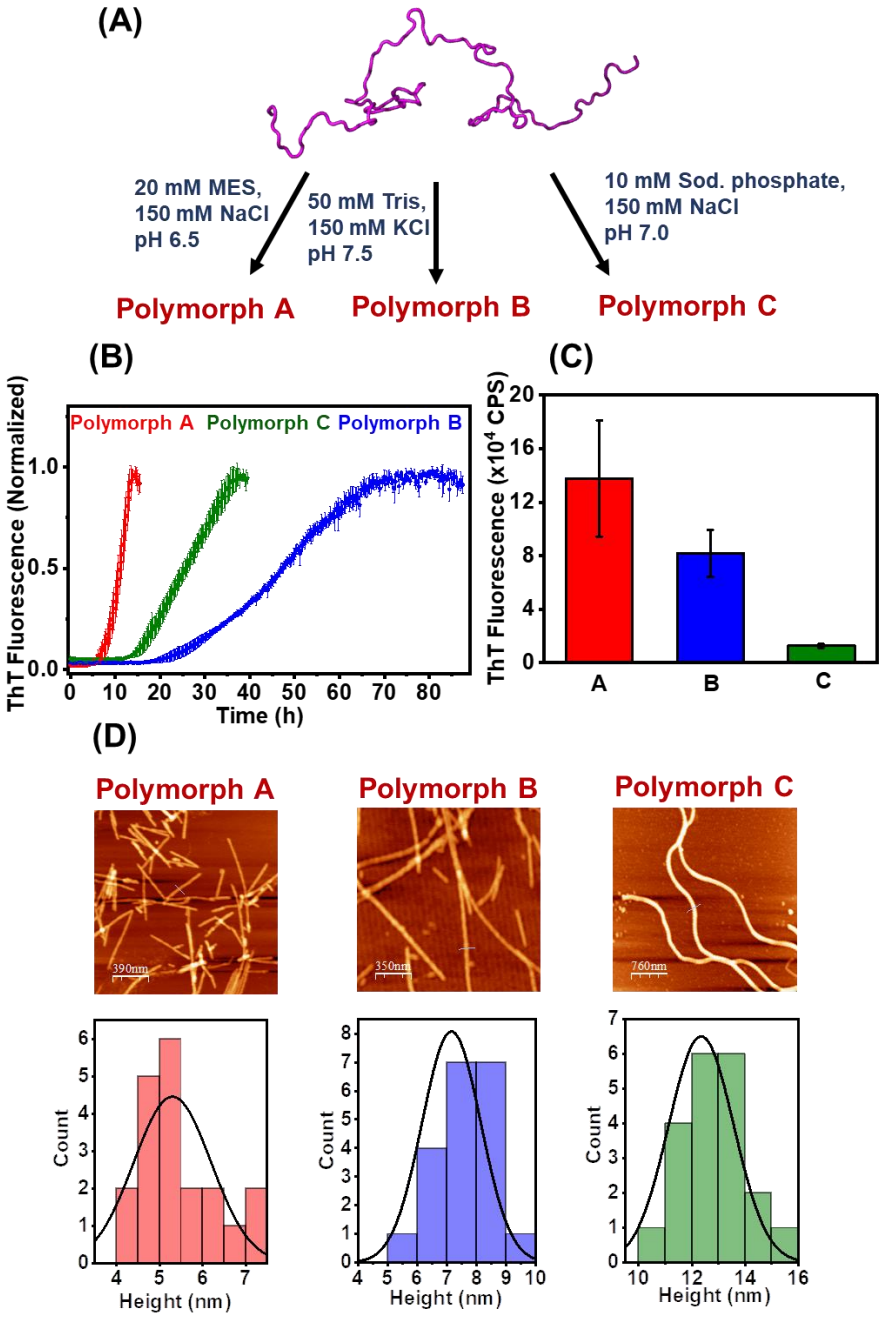


Figure 4.1: (A) Schematic representing various α-synuclein conformers used in this study. One of the NMR structures of expanded α-synuclein is shown that is taken from the protein ensemble database (PED00024e001) and generated using PyMOL (Schrodinger, LLC, New York). (B) ThT aggregation kinetics of α-synuclein at pH 6.5, pH 7.5, and pH 7.0 for generating Polymorph A, Polymorph B, and Polymorph C. The error bars denote mean ± S.D. obtained from at least four independent measurements. (C) Comparison of ThT fluorescence intensity for Polymorph A, Polymorph B, and Polymorph C. The error bars denote mean ± S.D. obtained from at least five independent measurements. (D) AFM images of different fibril polymorphs formed at different pH values. The lower panel represents the histograms depicting the heights of 20 different fibrils of each conformer.

atomic force microscopy (AFM). Our AFM images revealed that Polymorph B is long and displays cylindrical morphology with an average height of ~7-8 nm. On the contrary, the morphology of Polymorph A resembles that of Polymorph B, but they are short and stubby, while Polymorph C displayed curvy thread-like morphology with an average height of ~12 nm (Figure 4.1 D and Figure 4.2). Additionally, each of these polymorphs displayed a single minimum in the CD spectrum at ~218-220 nm with variable ellipticities suggesting distinct packing of the fibrillar core and conformations of the β-sheets (Figure 4.2 D).[47] Furthermore, it is well established that these polymorphs have quite distinct amyloid cores, as observed in their proteinase-K digestion patterns, and differential binding to the plasma membrane within the neurons where they exhibit variable seeding capacity.[27,48]

Having formed the fibril polymorphs *in vitro* with altered supramolecular packing within the amyloid core, we next aimed at capturing the conformational and structural differences within the fibril conformers through the lens of vibrational Raman spectroscopy.

4.3.2 Characterization of α-synuclein conformers using vibrational Raman spectroscopy

Following the generation of α-synuclein fibril polymorphs *in vitro,* we next set out to characterize the fibrillar structures using vibrational Raman spectroscopy. Figure 4.3 A shows Raman spectra for Polymorph A, Polymorph B, and Polymorph C, providing a unique conformational signature for each variant generated *in vitro*. The vibrational Raman spectra were dominated by secondary structural marker bands, amide I and amide III, and bands arising due to the chemical bond vibrations of the aliphatic side chains and aromatic amino acid residues.[40,49] These vibrational bands undergo a change in the band positions and intensities as a function of varying inter and intra- molecular vibrations that can be used to distinguish the polypeptide chain conformations.[49] All three fibrillar conformations showed an amide I centered at ~1673 cm^{-1} suggesting β-sheet rich conformations for each conformer. Subtle variations in the peak positions of amide I indicated slight differences in the β-sheet architecture as the band frequency is dependent on the number of β-strands in the β-sheets and their orientation. A closer look at these spectra showed a marked variation in the full width at half maxima (FWHM), suggesting distinct secondary structural composition for these polymorphs. Polymorph C showed an FWHM of ~29 cm^{-1}, while Polymorph A and B appeared to have nearly identical FWHM values of ~17 cm^{-1}, indicating a considerable degree of conformational heterogeneity within Polymorph C. Deconvolution of amide I revealed that this band contains the key peaks corresponding to different elements of secondary structures,

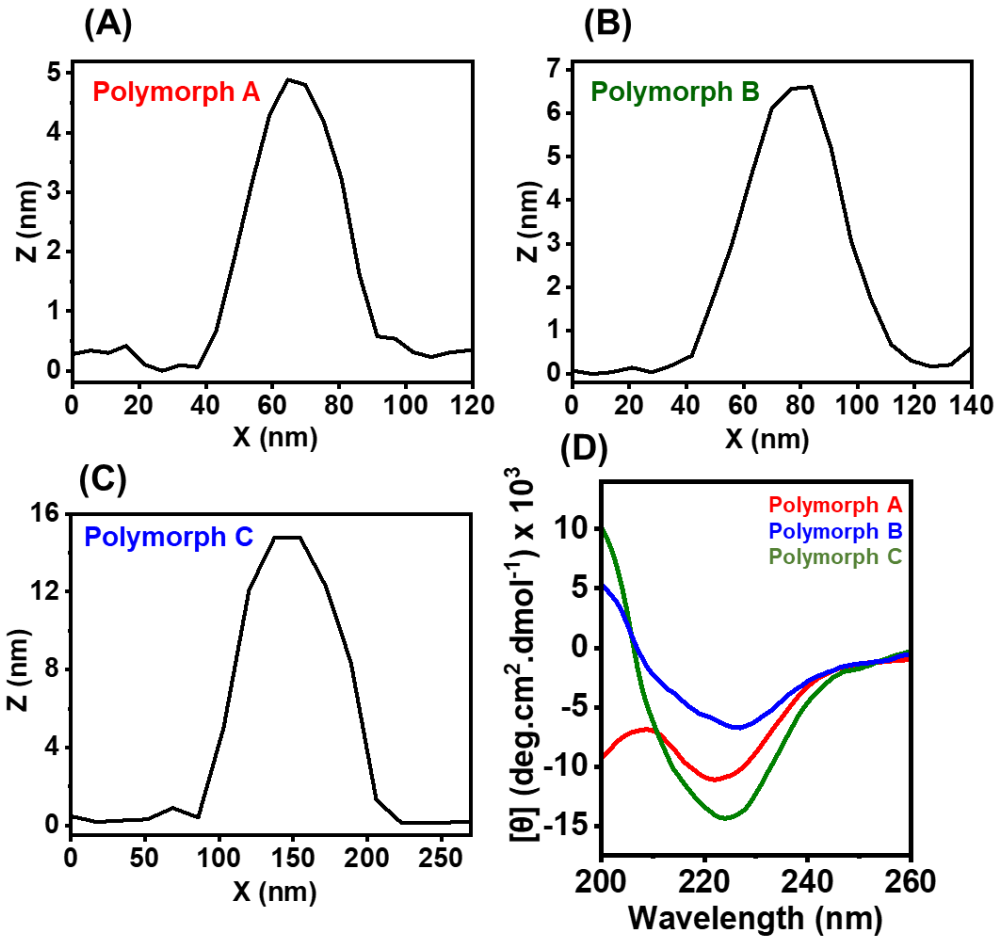


Figure 4.2: (A), (B), (C) Height profiles of the AFM images are shown in Figure 4.1. (D) Far-UV CD spectra of the three α-synuclein polymorphs formed under different aggregation conditions.

each having a discrete vibrational frequency, thereby providing a quantitative conformational fingerprinting of each fibril polymorph (Figure 4.3 B, C, D and Table 1).[50,51] Polymorph C in addition to the β-sheet and non-regular structures exhibited a small proportion of α-helical structures as opposed to Polymorph A and Polymorph B which are predominantly composed of β-rich and non-regular structures.

We next performed principal component analysis (PCA) to identify the regions of spectral variations by reducing the dimensionality within the datasets of recorded Raman spectra which helps in better visualization and analysis of the spectra (Figure 4.3 E, F).[40,41,52]

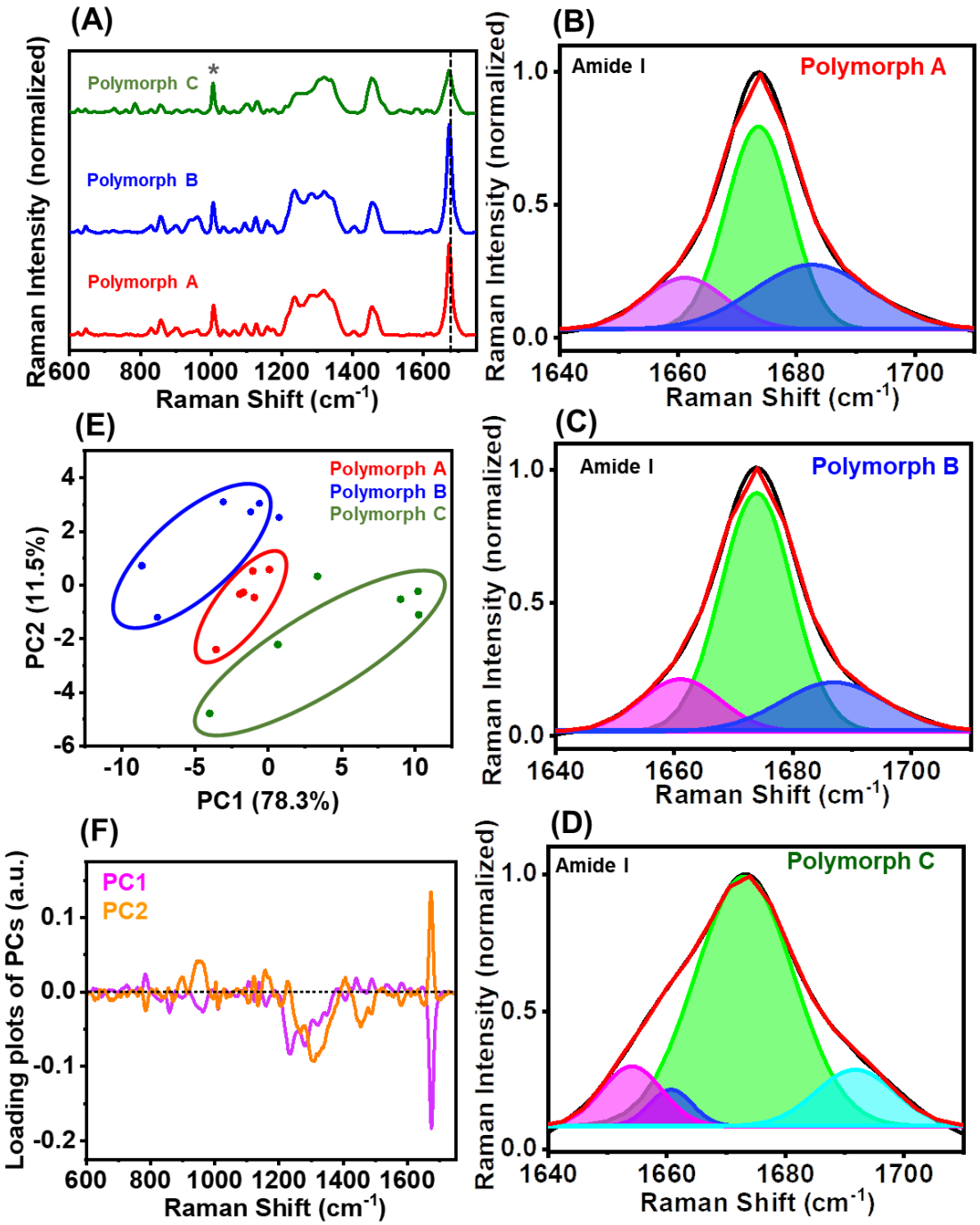


Figure 4.3: (A) Raman spectra of three α-synuclein polymorphs at an excitation wavelength of 785 nm. Spectra recorded at 50 mW laser power using 50x objective. All Raman spectra are normalized with respect to the phenylalanine ring breathing mode shown by an asterisk. Shown here is the mean of six independently recorded spectra. (B), (C), (D) Gaussian deconvolution of amide I band to analyze the percentage composition of various secondary structural elements. The black and red solid lines represent the actual data and the cumulative fit, respectively. The colored solid lines represent the Gaussian peaks obtained after deconvolution. (E) 2-Dimensional principal component analysis of Raman spectra shown in (A). PCA score plot is shown where each solid point represents the PC score of a single spectrum. (F) PC loading spectra represent the spectral variation responsible for the scores across the given PC axis.

Table 1. Percentage analyses of different secondary structural elements obtained after deconvolution of amide I region for different α-synuclein polymorphs.

Fibril Polymorph	Peak position (cm^{-1})	Assignment	FWHM (cm^{-1})	Percent of secondary structure content (%)
Polymorph A	1661.1	Nonregular/turns	16.1	16.8
	1673.5	β-sheet	12.9	52.7
	1682.6	Nonregular/turns	23.3	30.4
Polymorph B	1661.0	Nonregular/turns	15.8	15.7
	1673.9	β-sheet	14.2	64.7
	1686.9	Nonregular/turns	20.8	19.5
Polymorph C	1654.1	α-helix	11.9	10.6
	1660.7	Nonregular/turns	8.4	4.5
	1673.1	β-sheet	19.6	71.9
	1691.7	Nonregular/turns	15.6	12.9

Two principal components, PC1, and PC2 were sufficient to differentiate the spectra and represented the most variation in the data. Polymorph A and Polymorph B showed a negative PC1 score, as reflected in the score plot (Figure 4.3 C). PC1 showed strong negative coefficients on the loadings plot for β-sheet structure (1674 cm^{-1} and 1236 cm^{-1}), tyrosine ring breathing (857 cm^{-1}), and backbone/skeletal nonregular structure (960 cm^{-1}) suggesting that these peaks are stronger for Polymorph A and B (Figure 4.3 D). Along these lines, peaks representing β-sheet conformation (1673 cm^{-1} and 1228 cm^{-1}), C-C/C-N stretching (1157 cm^{-1}) showed a positive coefficient on the PC2, and Polymorph B loadings were observed to have more positive scores for PC2 highlighting the dominance of these bands in Polymorph B.

Taken together, these observations highlighted the efficacy of vibrational Raman spectroscopy in distinguishing various fibril conformations of the same protein. We next intended to dissect the structure and interactions within the extensively H-bonded β-rich amyloid core of various conformers using hydrogen-deuterium (H/D) exchange Raman spectroscopy.

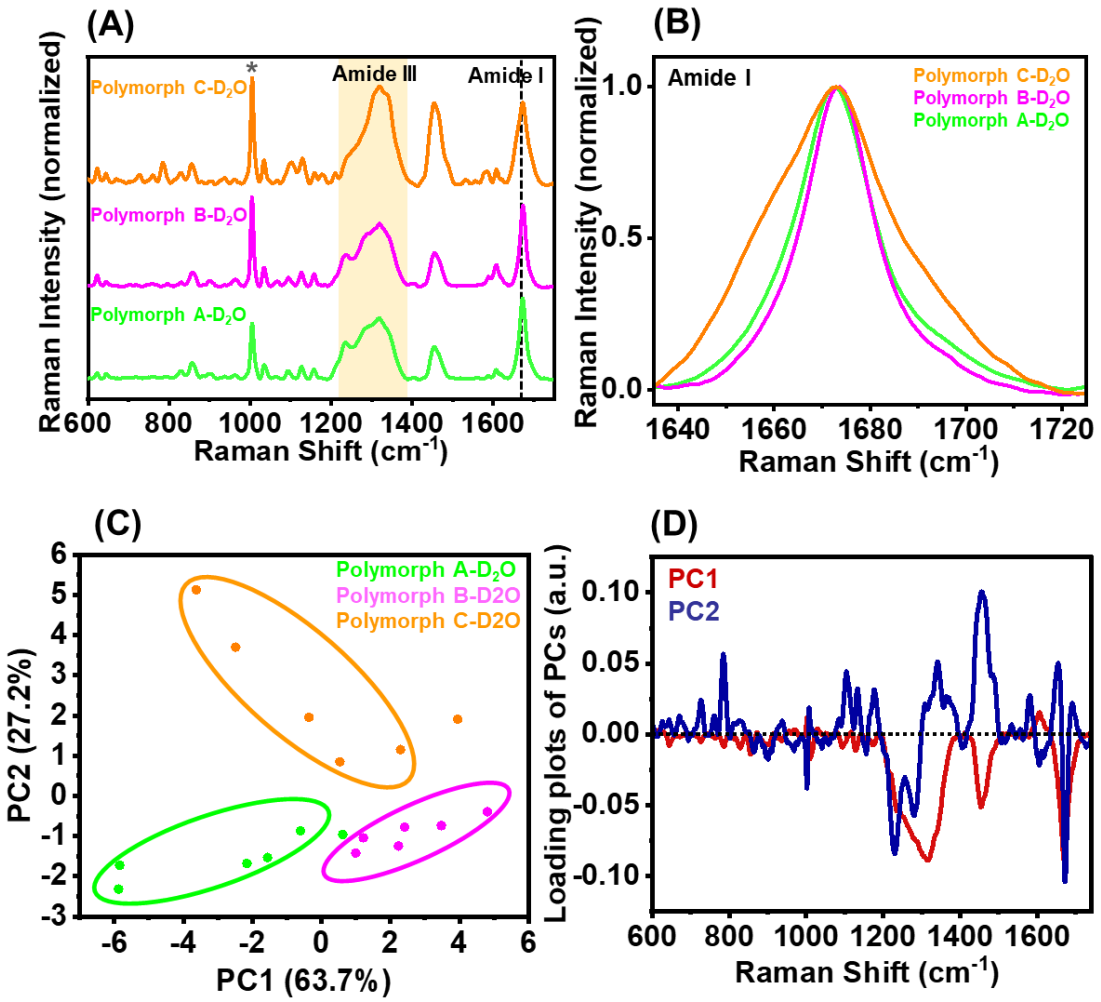


Figure 4.4: (A) Raman spectra of three α-synuclein polymorphs suspended in D_2O at an excitation wavelength of 785 nm. Spectra recorded at 50 mW laser power using 50x objective. All Raman spectra are normalized with respect to the phenylalanine ring breathing mode shown by an asterisk. Shown here is the mean of six independently recorded spectra. (B) Zoomed in amide I band of the Raman spectra showed in (A). (C) 2-Dimensional principal component analysis of Raman spectra shown in (A). PCA score plot is shown where each solid point represents the PC score of a single spectrum. (D) PC loading spectra represent the spectral variation responsible for the scores across the given PC axis.

4.3.3 Probing the amyloid core using hydrogen-deuterium exchange Raman spectroscopy

We next performed vibrational Raman experiments on deuterated fibril polymorphs post-incubation in D_2O for 24 hours. Suspending amyloid fibrils in D_2O causes a rapid H/D exchange of the backbone amide hydrogens (-NH) in the unordered parts, while those in the highly ordered strongly H-bonded β-sheet structures exchange more slowly. Upon H/D exchange, all three polymorphs showed an amide I centered around 1673 cm^{-1} having variable FWHM values (Figure 4.4 A, B). Polymorph A showed a marginally greater amide I bandwidth

than Polymorph B, suggesting varying strengths of H-bonding in their cross-β amyloid cores. We would like to point out that while the FWHM values of the amide I band for the undeuterated Polymorph A and Polymorph B were similar, there is a slight variation in the FWHM for their deuterated counterparts highlighting variable solvent exposure of the amyloid core. This is because the amide I mode consists of carbonyl C=O stretching with a very small contribution from C-N stretching and N-H bending vibrations. In general, the D_2O-treated fibrils showed a lower intensity of several Raman bands, particularly CH_2/CH_3 and Cα-H vibrations suggesting substantial changes in the polypeptide backbone conformation upon deuteration (Figure 4.5 A, C, E). The recorded Raman spectra were further distinguished using PCA. The loadings showed that peaks including α-helix (1653 cm^{-1}) and nonregular/turns (1690 cm^{-1}), CH_2/CH_3 deformations (1457 cm^{-1}), C-C/C-N/C-O stretching (1106 cm^{-1}, 1132 cm^{-1}, 1177 cm^{-1}) were more intense in Polymorph C which showed positive scores for PC2 in the score plot, whereas 1672 cm^{-1} (β-sheet), 1310 cm^{-1} (α-helix) were more intense in Polymorph A (Figure 4.4 C, D). We expect to see a disappearance of the amide III band upon H/D exchange due to the loss of coupling between C-H bending and N-H bending vibrations, while amide I from the fully protonated cross-β-core should remain unchanged. Comparing the Raman spectra of the deuterated fibrils with their corresponding undeuterated forms showed that the amide III band for Polymorph C lost its prominence while this loss was the lowest for Polymorph A, indicating a difference of H/D exchange within the fibrillar core for these conformers. Deconvolution of the amide III band showed that only 3.4 % of β-sheet content persists post-deuteration as opposed to ~19 % in Polymorph A (Figure 4.5 B, D, F and Table 2). These observations suggested that the amyloid core of Polymorph C is the least protected as compared to the core of other conformers, probably due to the presence of a disordered fuzzy coat and α-helical twists resulting in the weakening of the β-sheet hydrogen bonds corroborating our ThT binding data as well.

Collectively, our normal Raman and H/D exchange Raman experiments for the three fibril conformers suggested that each fibrillar conformation has a unique Raman fingerprint. All three conformations consist of a cross-β-sheet architecture within their amyloid cores exhibiting H-bonding of varying strengths. Polymorph C formed at pH 7.0 consists of a larger ensemble of secondary structures with an extensively hydrated amyloid core indicated by a much larger degree of H/D exchange.

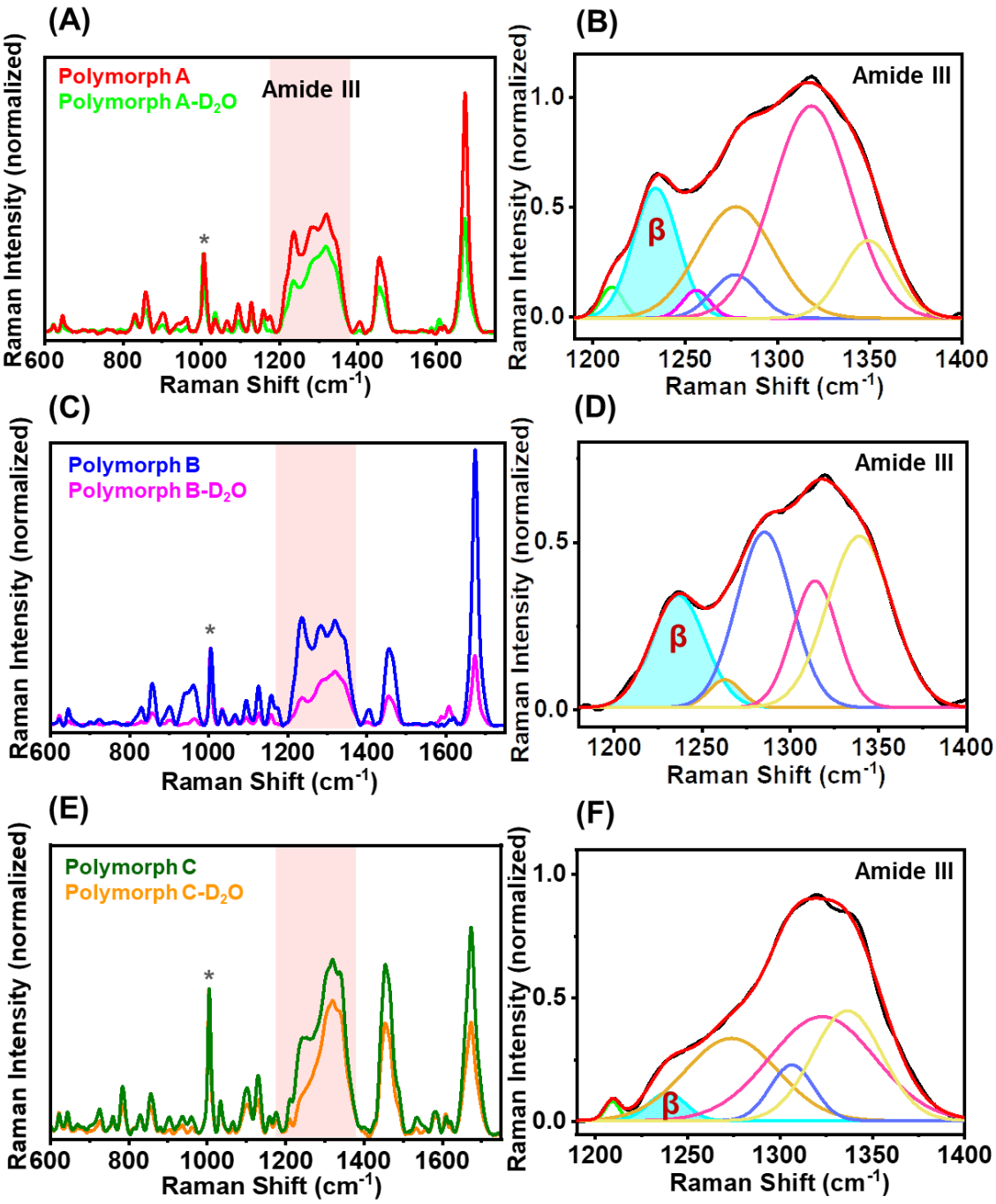


Figure 4.5: (A), (C), (E) Raman spectra of deuterated Polymorph A, Polymorph B, and Polymorph C, respectively shown along with their corresponding undeuterated spectra. Spectra recorded at 50 mW laser power using 50x objective. All Raman spectra are normalized with respect to the phenylalanine ring breathing mode shown by an asterisk. (B), (D), (F) Gaussian deconvolution of amide III band to analyze the percentage composition of various secondary structural elements. The black and red solid lines represent the actual data and the cumulative fit, respectively. The other colored solid lines represent the Gaussian peaks obtained after deconvolutions.

Table 2. Percentage analyses of different secondary structural elements obtained after deconvolution of amide III region for different D_2O treated α-synuclein polymorphs.

Fibril Polymorph	Peak position (cm^{-1})	Assignment	FWHM (cm^{-1})	Percent of secondary structure content (%)
Polymorph A	1210.2	Tyrosine [ν(C-C)]	14.1	1.8
	1233.9	β-sheet	28.6	15.1
	1256.0	β-sheet	17.5	1.9
	1276.9	Nonregular/turns	29.7	5.3
	1277.4	Nonregular/turns	48.4	21.9
	1318.4	α-helix	49.8	43.1
	1349.3	δ(CαH)	34.1	10.8
Polymorph B	1236.4	β-sheet	35.9	18.0
	1262.8	Nonregular/turns	21.8	2.7
	1285.3	Nonregular/turns	36.9	28.9
	1313.9	α-helix	30.4	17.2
	1339.2	δ(CαH)	43.2	33.1
Polymorph C	1209.3	Tyrosine [ν(C-C)]	9.1	0.9
	1238.9	β-sheet	23.5	3.4
	1273.7	Nonregular/turns	59.0	25.4
	1306.5	α-helix	27.7	8.1
	1322.9	δ(CαH)	67.1	36.6
	1336.4	δ(CαH)	44.3	25.5

4.4 Discussion

α-synuclein is known to assemble into fibrillar aggregates in the neurons and glial cells within the human brain exhibiting heterogeneous conformations potentially similar to prion strains. These fibril polymorphs generated under different cellular conditions have variable molecular-level structures that can explain the clinical and pathological variations observed in synucleinopathies.[53-55] A growing body of current research has outlined models of fibril polymorphism at the level of protofilaments and inter-protofilament interactions stabilizing the structure.[56] However, obtaining molecular insights into the fibrillar assemblies can be quite challenging, laborious, and time-consuming using several available high-resolution structural tools and methods.

In this work, we utilized vibrational Raman spectroscopy to structurally characterize three α-synuclein polymorphs generated *in vitro* under different aggregation conditions. We observe that these polymorphs have distinct ThT binding propensity and aggregation kinetics. Our AFM and CD studies revealed prominent differences in the fibril morphologies and secondary structural components. Vibrational Raman experiments successfully capture the structural heterogeneity within the polymorphs and distinguish them based on the unique and quantitative conformational fingerprint. The key Raman marker bands distinguishing these structures were analyzed using multivariate analysis, while the curve-fitting analysis of the amide bands helped to quantitatively characterize different elements of secondary structures that serve as a unique fingerprint for the pure polymorphs. Our findings highlight that Polymorph C, generated at pH 7.0 exhibit the greatest conformational heterogeneity as indicated by a broad amide I band. Furthermore, varying intensities of several backbone vibrations were observed, as indicated by the PCA plots suggesting differences in the side-chain packing arrangement for these fibril conformations. H/D exchange techniques have been widely used to study several aspects of proteins, such as conformation, dynamics, and folding mechanisms.[57,58] Incorporating a heavy atom such as deuterium causes the replacement of the easily accessible protons in the disordered fuzzy coat and loosely packed ordered segments of the protein aggregates and help us to fathom the H-bonding patterns and solvent accessibility within the extensively H-bonded amyloid core. Based on this principle, we hypothesize that Polymorph C has the least protected amyloid core as indicated by the disappearing amide III band owing to rapid H→D exchange, thereby weakening the coupling between the C-H bending and N-H bending vibrations responsible for amide III vibrations (Figure 4.6).

In summary, our studies illustrate the utility of vibrational Raman spectroscopy in capturing the amyloid fibril polymorphism that gives rise to morphological and biological strain diversities. Such structural insights on the diverse conformations adopted by the pathological aggregates of an amyloidogenic protein enable us to decipher the detailed mechanism of disease progression together with developing therapeutics that can target not only the insoluble assemblies but also the soluble species and toxic oligomers that may not be suitable for other high-resolution tools. This effortless and sensitive methodology can serve as a potent tool to study the effect of factors such as cellular conditions, posttranslational modifications, the presence of an exogenous agent, a co-factor or other amyloidogenic proteins, on the higher-order packing and nanoscale morphology of amyloid assemblies thereby delineating the process of amyloid formation and polymorphism and eventually their structure-

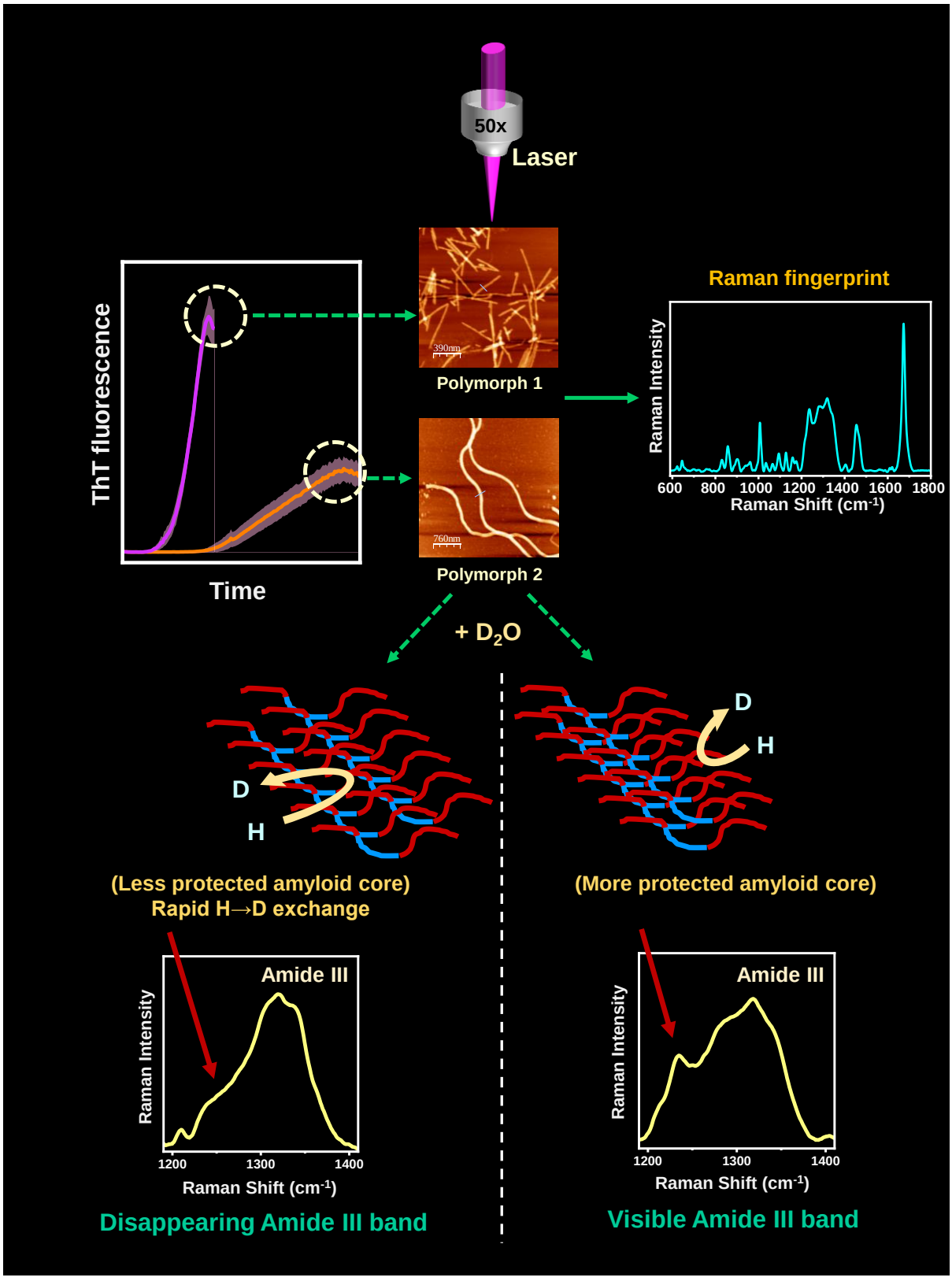


Figure 4.6: A schematic representation of characterizing α-synuclein fibrillar polymorphs using hydrogen/deuterium exchange Raman spectroscopy.

pathogenic relationship.[59,60,61] We envision that extending this tool to study the brain-derived aggregates can provide a unique quantitative fingerprint for peptide backbone conformations at all stages of fibrillation that can be highly beneficial in the early diagnosis of neurodegenerative disorders.

4.5 References

1. Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003,** 426(6968), 884-90.

2. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. The Amyloid State and Its Association with Protein Misfolding Diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** 15(6), 384-396.

3. Sipe, J.; Cohen, A. S. Review: History of the Amyloid Fibril. *J. Struct. Biol.* **2000,** 130(2-3), 88-98.

4. Gallardo, R.; Ranson, N. A.; Radford, S. E. Amyloid Structures: Much More Than Just a Cross-β Fold. *Curr. Opin. Struct. Biol.* **2020,** 60, 7-16.

5. Dobson, C. M.; Knowles, T. P. J.; Vendruscolo, M. The Amyloid Phenomenon and Its Significance in Biology and Medicine. *Cold Spring Harb. Perspect. Biol.* **2020,** 12(2), a033878.

6. Iadanza, M. G.; Jackson, M. P.; Hewitt, E. W.; Ranson, N. A.; Radford, S. E. A New Era for Understanding Amyloid Structures and Disease. *Nat. Rev. Mol. Cell. Biol.* **2018,** 19(12), 755-773.

7. Adamcik, J.; Mezzenga, R. Amyloid Polymorphism in the Protein Folding and Aggregation Energy Landscape. *Angew. Chem. Int. Ed. Engl.* **2018,** 57(28), 8370-8382.

8. Chuang, E.; Hori, A. M.; Hesketh, C. D.; Shorter J. Amyloid Assembly and Disassembly. *J. Cell Sci.* **2018,** 131(8).

9. Eisenberg, D. S.; Sawaya, M. R. Structural Studies of Amyloid Proteins at the Molecular Level. *Annu. Rev. Biochem.* **2017,** 86, 69-95.

10. Fowler, D. M.; Koulov, A. V.; Balch, W. E., Kelly, J. W. Functional Amyloid--From Bacteria to Humans. *Trends Biochem. Sci.* **2007,** 32(5), 217-224.

11. Jain, N.; Chapman, M. R. Bacterial Functional Amyloids: Order from Disorder. *Biochim. Biophys. Acta. Proteins Proteom.* **2019,** 1867(10), 954-960.

12. Otzen, D.; Riek, R. Functional Amyloids. *Cold Spring Harb. Perspect. Biol.* **2019,** 11(12), a033860.

13. Tycko, R. Physical and structural basis for polymorphism in amyloid fibrils. Protein Sci. 2014, 23(11), 1528-1539.

14. Seuring, C.; Verasdonck, J.; Ringler, P.; Cadalbert, R.; Stahlberg, H.; Böckmann, A.; Meier, B. H.; Riek, R. Amyloid Fibril Polymorphism: Almost Identical on the Atomic Level, Mesoscopically Very Different. J. Phys. Chem. B. 2017, 121(8), 1783-1792.

15. Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W. M.; Mattson, M. P.; Tycko, R. Self-Propagating, Molecular-Level Polymorphism in Alzheimer's Beta-Amyloid Fibrils. *Science* **2005,** 307(5707), 262-265.

16. Prusiner, S. B. Cell Biology. A Unifying Role for Prions in Neurodegenerative Diseases. *Science.* **2012,** 336(6088), 1511-1513.

17. Brundin, P.; Melki, R.; Kopito, R. Prion-like Transmission of Protein Aggregates in Neurodegenerative Diseases. *Nat. Rev. Mol. Cell Biol.* **2010,** 11(4), 301-307.

18. Aguzzi, A.; Rajendran, L. The Transcellular Spread of Cytosolic Amyloids, Prions, and Prionoids. *Neuron* **2009,** 64(6), 783-790.

19. Silva, J. L.; Cordeiro, Y. The "Jekyll and Hyde" Actions of Nucleic Acids on the Prion-like Aggregation of Proteins. *J. Biol. Chem.* **2016,** 291(30), 15482-15490.

20. Mehra, S.; Ahlawat, S.; Kumar, H.; Datta, D.; Navalkar, A.; Singh, N.; Patel, K.; Gadhe, L.; Kadu, P.; Kumar, R.; Jha, N. N.; Sakunthala, A.; Sawner, A. S.; Padinhateeri, R.;

Udgaonkar, J. B.; Agarwal, V.; Maji, S. K. α-Synuclein Aggregation Intermediates Form Fibril Polymorphs with Distinct Prion-like Properties. *J. Mol. Biol.* **2022,** 434(19), 167761.

21. Goedert, M. Alpha-synuclein and Neurodegenerative Diseases. *Nat. Rev. Neurosci.* **2001,** 2(7), 492-501.

22. Spillantini, M. G.; Crowther, R. A.; Jakes, R.; Hasegawa, M.; Goedert, M. Alpha-synuclein in Filamentous Inclusions of Lewy Bodies from Parkinson's Disease and Dementia with Lewy Bodies. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** 95(11), 6469-6473.

23. Bousset, L.; Pieri, L.; Ruiz-Arlandis, G.; Gath, J.; Jensen, P. H.; Habenstein, B.; Madiona, K.; Olieric, V.; Böckmann, A.; Meier, B. H.; Melki, R. Structural and Functional Characterization of Two Alpha-synuclein Strains. *Nat. Commun.* **2013,** 4, 2575.

24. Guo, J. L.; Covell, D. J.; Daniels, J. P.; Iba, M.; Stieber, A.; Zhang, B.; Riddle, D. M.; Kwong, L. K.; Xu, Y.; Trojanowski, J. Q.; Lee, V. M. Distinct α-Synuclein Strains Differentially Promote Tau Inclusions in Neurons. *Cell* **2013,** 154(1), 103-117.

25. Lau, A.; So, R. W. L.; Lau, H. H. C.; Sang, J. C.; Ruiz-Riquelme, A.; Fleck, S. C.; Stuart, E.; Menon, S.; Visanji, N. P.; Meisl, G.; Faidi, R.; Marano, M. M.; Schmitt-Ulms, C.; Wang, Z.; Fraser, P. E.; Tandon, A.; Hyman, B. T.; Wille, H.; Ingelsson, M.; Klenerman, D.; Watts, J. C. α-Synuclein Strains Target Distinct Brain Regions and Cell Types. *Nat. Neurosci.* **2020,** 23(1), 21-31.

26. Ma, M. R.; Hu, Z. W.; Zhao, Y. F.; Chen, Y. X.; Li, Y. M. Phosphorylation Induces Distinct Alpha-synuclein Strain Formation. *Sci. Rep.* **2016,** 6, 37130.

27. Shrivastava, A. N.; Bousset, L.; Renner, M.; Redeker, V.; Savistchenko, J.; Triller, A.; Melki, R. Differential Membrane Binding and Seeding of Distinct α-Synuclein Fibrillar Polymorphs. *Biophys. J.* **2020,** 118(6), 1301-1320.

28. Heise, H.; Hoyer, W.; Becker, S.; Andronesi, O. C.; Riedel, D.; Baldus, M. Molecular-Level Secondary Structure, Polymorphism, and Dynamics of Full-length Alpha-synuclein

Fibrils Studied by Solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** 102(44), 15871-15876.

29. Li, B.; Ge, P.; Murray, K. A.; Sheth, P.; Zhang, M.; Nair, G.; Sawaya, M. R.; Shin, W. S.; Boyer, D. R.; Ye, S.; Eisenberg, D. S.; Zhou, Z. H.; Jiang, L. Cryo-EM of Full-Length α-Synuclein Reveals Fibril Polymorphs with a Common Structural Kernel. *Nat. Commun.* **2018,** 9(1), 3609.

30. Cao, Q.; Boyer, D. R.; Sawaya, M. R.; Ge, P.; Eisenberg, D. S. Cryo-EM Structures of Four Polymorphic TDP-43 Amyloid Cores. *Nat. Struct. Mol. Biol.* **2019,** 26(7), 619-627.

31. Falcon, B.; Zhang, W.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Vidal, R.; Crowther, R. A.; Ghetti, B.; Scheres, S. H. W.; Goedert, M. Structures of Filaments from Pick's Disease Reveal a Novel Tau Protein Fold. *Nature*. 2018, 561(7721), 137-140.

32. Gath, J.; Bousset, L.; Habenstein, B.; Melki, R.; Böckmann, A.; Meier, B. H. Unlike Twins: An NMR Comparison of Two α-Synuclein Polymorphs Featuring Different Toxicity. *PLoS One.* **2014,** 9(3), e90659.

33. Cordeiro, Y.; Kraineva, J.; Suarez, M. C.; Tempesta, A. G.; Kelly, J. W.; Silva, J. L.; Winter, R.; Foguel, D. Fourier transform infrared spectroscopy provides a fingerprint for the tetramer and for the aggregates of transthyretin. *Biophys. J.* **2006,** 91(3), 957-967.

34. Nath, A.; Sammalkorpi, M.; DeWitt, D. C.; Trexler, A. J.; Elbaum-Garfinkle, S.; O'Hern, C. S.; Rhoades, E. The Conformational Ensembles of α-Synuclein and Tau: Combining Single-Molecule FRET and Simulations. *Biophys. J.* **2012,** 103(9), 1940-1949.

35. Fung, H. Y. J.; McKibben, K. M.; Ramirez, J.; Gupta, K.; Rhoades, E. Structural Characterization of Tau in Fuzzy Tau:Tubulin Complexes. *Structure* **2020,** 28(3), 378-384.

36. Guerrero-Ferreira, R.; Taylor, N. M.; Mona, D.; Ringler, P.; Lauer, M. E.; Riek, R.; Britschgi, M.; Stahlberg, H. Cryo-EM Structure of Alpha-synuclein Fibrils. *Elife* **2018,** 7, e36402.

37. Li, Y.; Zhao, C.; Luo, F.; Liu, Z.; Gui, X.; Luo, Z.; Zhang, X.; Li, D.; Liu, C.; Li, X. Amyloid Fibril Structure of α-Synuclein Determined by Cryo-Electron Microscopy. *Cell Res.* **2018,** 28(9), 897-903.

38. Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; Willbold, D.; Schröder, G. F. Fibril Structure of Amyloid-β(1-42) by Cryo-Electron Microscopy. *Science* **2017,** 358(6359), 116-119.

39. Zielinski, M.; Röder, C.; Schröder, G. F. Challenges in Sample Preparation and Structure Determination of Amyloids by Cryo-EM. *J. Biol. Chem.* **2021,** 297(2), 100938.

40. Devitt, G.; Crisford, A.; Rice, W.; Weismiller, H. A.; Fan, Z.; Commins, C.; Hyman, B. T.; Margittai, M.; Mahajan, S.; Mudher, A. Conformational Fingerprinting of Tau Variants and Strains by Raman Spectroscopy. *RSC Adv.* **2021,** 11(15), 8899-8915.

41. Ishigaki, M.; Morimoto, K.; Chatani, E.; Ozaki, Y. Exploration of Insulin Amyloid Polymorphism Using Raman Spectroscopy and Imaging. *Biophys. J.* **2020,** 118(12), 2997-3007.

42. Kurouski, D.; Van Duyne, R. P.; Lednev, I. K. Exploring the Structure and Formation Mechanism of Amyloid Fibrils by Raman Spectroscopy: A Review. *Analyst* **2015,** 140(15), 4967-4980.

43. Jayaraman, V.; Rodgers, K. R.; Mukerji, I.; Spiro, T. G. Hemoglobin Allostery: Resonance Raman Spectroscopy of Kinetic Intermediates. *Science* **1995,** 269(5232), 1843-1848.

44. Shashilov, V.; Xu, M.; Makarava, N.; Savtchenko, R.; Baskakov, I. V.; Lednev, I. K. Dissecting Structure of Prion Amyloid Fibrils by Hydrogen-Deuterium Exchange Ultraviolet Raman Spectroscopy. *J. Phys. Chem. B.* **2012,** 116(27), 7926-7930.

45. Lu, X.; Wintrode, P. L.; Surewicz, W. K. Beta-sheet Core of Human Prion Protein Amyloid Fibrils as Determined by Hydrogen/Deuterium Exchange. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104(5), 1510-1515.

46. Hoyer, W.; Antony, T.; Cherny, D.; Heim, G.; Jovin, T. M.; Subramaniam, V. Dependence of Alpha-synuclein Aggregate Morphology on Solution Conditions. *J. Mol. Biol.* **2002,** 322(2), 383-393.

47. Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y. H.; Goto, Y.; Réfrégiers, M.; Kardos, J. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** 112(24), E3095-103.

48. Makky, A.; Bousset, L.; Polesel-Maris, J.; Melki, R. Nanomechanical Properties of Distinct Fibrillar Polymorphs of the Protein α-Synuclein. *Sci. Rep.* **2016,** 6, 37970.

49. Tuma, R. Raman Spectroscopy of Proteins: From Peptides to Large Assemblies. *J. Raman. Spec.* **2005,** 36, 307-319.

50. Bhattacharya, M.; Jain, N.; Dogra, P.; Samai, S.; Mukhopadhyay, S. Nanoscopic Amyloid Pores Formed via Stepwise Protein Assembly. *J. Phys. Chem. Lett.* **2013,** 4(3), 480-485.

51. Chatterjee, S.; Kan, Y.; Brzezinski, M.; Koynov, K.; Regy, R. M.; Murthy, A. C.; Burke, K. A.; Michels, J. J.; Mittal, J.; Fawzi, N. L.; Parekh, S. H. Reversible Kinetic Trapping of FUS Biomolecular Condensates. *Adv. Sci. (Weinh).* **2022,** 9(4), e2104247.

52. Devitt, G.; Rice, W.; Crisford, A.; Nandhakumar, I.; Mudher, A.; Mahajan, S. Conformational Evolution of Molecular Signatures during Amyloidogenic Protein Aggregation. *ACS Chem. Neurosci.* **2019,** 10(11), 4593-4611.

53. Lashuel, H. A.; Overk, C. R.; Oueslati, A.; Masliah, E. The Many Faces of α-Synuclein: From Structure and Toxicity to Therapeutic Target. *Nat. Rev. Neurosci.* 2013, 14(1), 38-48.

54. Meade, R. M.; Fairlie, D. P.; Mason, J. M. Alpha-synuclein Structure and Parkinson's Disease - Lessons and Emerging Principles. *Mol. Neurodegener.* 2019, 14(1), 29.

55. Serpell, L. C.; Berriman, J.; Jakes, R.; Goedert, M.; Crowther, R. A. Fiber Diffraction of Synthetic Alpha-Synuclein Filaments Shows Amyloid-like Cross-beta Conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** 97(9), 4897-4902.

56. Strohäker, T.; Jung, B. C.; Liou, S. H.; Fernandez, C. O.; Riedel, D.; Becker, S.; Halliday, G. M.; Bennati, M.; Kim, W. S.; Lee, S. J.; Zweckstetter, M. Structural Heterogeneity of α-Synuclein Fibrils Amplified from Patient Brain Extracts. *Nat. Commun.* **2019,** 10(1), 5535.

57. Ortiz, C.; Zhang, D.; Ribbe, A. E.; Xie, Y.; Ben-Amotz, D. Analysis of Insulin Amyloid Fibrils by Raman Spectroscopy. *Biophys. Chem.* **2007,** 128(2-3), 150-155.

58. Pirrone, G. F.; Iacob, R. E.; Engen, J. R. Applications of Hydrogen/Deuterium Exchange MS From 2012 to 2014. *Anal. Chem.* **2015,** 87(1), 99-118.

59. Werner, T.; Horvath, I.; Wittung-Stafshede, P. Crosstalk Between Alpha-Synuclein and Other Human and Non-Human Amyloidogenic Proteins: Consequences for Amyloid Formation in Parkinson's Disease. *J. Parkinsons Dis.* **2020,** 10(3), 819-830.

60. Wittung-Stafshede, P. Crossroads Between Copper Ions and Amyloid Formation in Parkinson's Disease. *Essays Biochem.* **2022,** 66(7), 977-986.

61. Lorentzon, E.; Horvath, I.; Kumar, R.; Rodrigues, J. I.; Tamás, M. J.; Wittung-Stafshede, P. Effects of the Toxic Metals Arsenite and Cadmium on α-Synuclein Aggregation In Vitro and in Cells. *Int. J. Mol. Sci.* **2021,** 22(21), 11455.

******

Conclusions and Future Directions

Proteins are the functional workhorses in living cells that perform many vital biological functions, such as catalyzing chemical reactions, providing structure and support to the cells, transporting, and storing small molecules, immune protection, and so forth. The amino acid composition governs the well-defined, 3-dimensional structure adopted by these macromolecules, enabling them to perform these essential functions in the cells. An emerging class of proteins is termed natively unfolded or intrinsically disordered proteins (IDPs) that confront the traditional one structure-one function paradigm. Interestingly nearly 40% of the eukaryotic genome codes for IDPs or proteins containing intrinsically disordered regions (IDRs). The beauty of this class of proteins is that they can exist as dynamic ensembles of rapidly fluctuating, interconvertible structures. This conformational plasticity allows them to participate in multivalent interactions and perform various functions ranging from signaling, regulation, cell cycle control, and so on. While many studies have shown how proteins with IDRs can contribute to increased functional versatility and cellular complexity, several reports have also revealed the importance of IDRs in many human diseases, such as cancer and neurodegeneration.

This thesis elucidates the utility of vibrational Raman spectroscopy as a potent tool to obtain structural insights from a range of protein assemblies formed by amyloidogenic IDPs implicated in physiology and disease. Below, I summarize the key questions asked and the findings discussed in this thesis.

1. Probing the conformational heterogeneity and structural distribution within biomolecular condensates using single-droplet vibrational Raman spectroscopy (Chapter 2).
2. Developing a single-droplet surface-enhanced Raman scattering tool to capture the conformational landscape within phase-separated liquid condensates of an RNA-binding protein (Chapter 3).
3. Hydrogen-deuterium exchange vibrational Raman spectroscopy to distinguish distinct amyloid polymorphs comprising altered core architecture (Chapter 4).

Phase transitions are ubiquitous in nature. Like water, macromolecules also undergo liquid-liquid phase separation (LLPS). For instance, in eukaryotic cells, phase separation seems to play a vital role in the intracellular compartmentalization of various biomacromolecules through the formation of membrane-less liquid-like compartments, also known as biomolecular condensates. These condensates are dynamic, liquid-like non-stoichiometric assemblies of

several proteins and/or nucleic acids. A growing body of intense current research has revealed that many of these condensates are enriched in IDPs having low sequence complexity. These low-complexity regions (LCRs) participate in a multitude of specific but weak and transient interactions and undergo either heterotypic or homotypic self-association based on the presence and absence of binding partners, respectively. These biomolecular condensates are involved in various cellular functions and physiology. However, their maturation can cause protein misfolding or aggregation via liquid-to-solid phase transitions, thereby causing debilitating neurodegenerative diseases. Therefore, unmasking the key molecular interactions and principles regulating this process is crucial to discern the mechanism and regulation of condensate assembly and maturation. Although there exists a host of microscopic and spectroscopic tools to study LLPS, the majority of these methods fail to provide information regarding the conformational heterogeneity within the condensed phase in a single-droplet fashion. In this direction, we developed and adapted single-droplet vibrational Raman spectroscopy, which is a powerful tool to capture the complex conformational characteristics governing the LLPS of IDPs. Focusing the laser beam within each protein droplet provides a wealth of molecular information within single droplets by recording scattering bands due to different vibrational modes from the mesoscopic protein-rich droplets (Chapter 2). By knowing the internal environment of the protein-rich droplets, it is possible to regulate the factors that govern condensate assembly and maturation, as well as allow us to modulate the functional roles of condensates. Our tool allowed us to capture the conformational landscape and can be very promising for the detection, characterization, and quantification of a wide range of biomolecular condensates. Often the protein folding and aggregation propensity of a protein is governed by an intricate balance between the chain-chain and chain-solvent interactions that determine the conformational and phase equilibria for IDPs. We believe that zooming into the Raman bands corresponding to various vibrational modes of water might shed some light on the structure of water within the dense phase and the dilute phase.

Vibrational Raman spectroscopy serves as a direct, non-invasive, and label-free tool to study intermolecular interactions, intramolecular dynamics, and the 3-dimensional structure of proteins within the mesoscopic liquid-like droplets. However, the low Raman scattering cross-section of proteins makes this tool insensitive to aqueous solutions of proteins at physiological concentrations. High concentrations and laser power used can be destructive for soft biological specimens. In this direction, surface-enhanced Raman scattering (SERS) using plasmonic nanostructures has emerged as a promising tool owing to the enormous chemical and

electromagnetic enhancement of Raman signals by generating hotspots in the vicinity of metal nanoparticles. We used iodide-modified silver nanoparticles to illuminate the inner workings of the condensed phase of a well-known RNA-binding protein, FUS, in a droplet-by-droplet manner. These highly sensitive measurements offer unprecedented sensitivity to capture the crucial interactions, conformational heterogeneity, and structural distributions within the liquid droplets (Chapter 3). Our findings showed that the ordered C-terminal RNA-binding domain (RBD) of FUS undergoes partial unwinding in the presence of RNA, which increases the polypeptide chain disorder. Thoughtful engineering and surface capping of the SERS substrates can be used to obtain conformational information from the specific parts of the polypeptide chains by enhancing unique sets of vibrational bands. Recent reports have shown that unmethylated or hypomethylated FUS can undergo phase separation and gelation under physiological conditions. Several post-translational modifications tune the cation-π interactions and alter the phase transitions. Post-translational arginine methylation is known to weaken the cation-π interactions between the guanidino group in the side chain of arginine residues and the aromatic ring of tyrosine, while citrullination completely ablates it. Arginine can be methylated by the protein arginine methyl transferase enzyme (PRMT), while citrullination can be carried out by converting the ketimine moiety to a ketone by the protein arginine deiminase (PAD) enzyme. Additionally, arginine mutations (R216C, R244C, R514G) in FUS perturb the RNA-dependent phase behavior by weakening the protein-RNA interactions. On the other hand, glycine mutation (G156E) alters the material properties of condensates by making them more gel-like without significantly affecting the phase separation propensity. We believe that the ultra-sensitive SERS tool can be elegantly used to study the effect of various pathological mutations and post-translational modifications on the inter and intra-chain interactions within the phase-separated droplets.

Misfolding of IDPs leads to the extra/intracellular accumulation and deposition of misfolded proteins, termed amyloidosis. This condition leads to the formation of amorphous aggregates or amyloids. Amyloids are composed of misfolded proteins that are assembled into a highly ordered cross β-sheet rich architecture in which the beta strands are aligned perpendicular to the growing fibril axis. Deposition of amyloids is known to be the characteristic hallmark of several neurodegenerative disorders, including the plaques in Alzheimer's disease, Lewy bodies in Parkinson's disease, inclusion bodies in Huntington's disease, and so on. However, recent studies have identified the beneficial role of amyloids in a variety of organisms ranging from bacteria to humans performing an array of physiological

functions. This class of amyloids is termed functional amyloids. The highly ordered amyloid assemblies share a common core architecture exhibiting a structural diversity in their supramolecular packing arrangement within the amyloid spine. Such an altered packing results in amyloid polymorphism that is often responsible for morphological and biological strain diversities. Polymorphism is a remarkable feature of amyloids, and elucidating its mechanism is of immense importance to understanding its structure-pathology relationship. We utilized hydrogen/deuterium exchange vibrational Raman spectroscopy to discern the esoteric structural features that are responsible for yielding diverse amyloid polymorphs derived from a presynaptic intrinsically disordered neuronal protein, α-synuclein, aggregation of which is associated with Parkinson's disease and other synucleinopathies. By directly monitoring and analyzing the key vibrational Raman bands for both protein backbone and sidechains, we were able to capture the conformational heterogeneity, structural distribution, and distinguish three distinct amyloid polymorphs based on their unique Raman signatures (Chapter 4). Hydrogen-deuterium exchange Raman spectroscopy allowed us to structurally distinguish these distinct amyloid polymorphs displaying altered hydrogen bonding ability and supramolecular packing within the cross-β-structural motif. We envision that such a straightforward and convenient structural tool will pave the way for understanding the factors influencing structural polymorphism by obtaining unique conformational signatures of *in vitro* or *ex vivo* generated fibrils. Such structural insights on the diverse conformations adopted by the pathological aggregates of an amyloidogenic protein enable us to decipher the detailed mechanism of disease progression together with developing therapeutics that can target not only the insoluble assemblies but also the soluble species and toxic oligomers that may not be suitable for other high-resolution tools.

******