PHASE BEHAVIOR OF AN INTRINSICALLY DISORDERED DOMAIN OF A MELANOSOMAL PROTEIN: CONFORMATIONAL CHARACTERISTICS, AMYLOID FORMATION, AND LIQUID-LIQUID PHASE SEPARATION

PRIYANKA DOGRA

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.

May 2020

Dedicated to my family and teachers

Declaration

The work presented in this thesis has been carried out by me under the guidance of DR. 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. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement 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.

PRIYANKA DOGRA

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.

DR. SAMRAT MUKHOPADHYAY

Acknowledgements

My stay here at IISER Mohali in general and PhD, in particular, has been nothing short of a journey that transformed me from a naïve undergraduate student into a somewhat learned graduate student, a curious soul to dive deeper into this wonderful field of science.

This journey, to the outside world, appears just as a routine work but internally it has been a sinusoidal curve where the joy of crests multiplied and setbacks of troughs fractionalized with the immense support and guidance of a large number of wonderful souls to whom I will always be indebted for being there to share my joys and shelve off my lows.

I take this opportunity, here at this juncture, to pay my deepest gratitude to these people.

First comes the PI, teacher, and mentor, my PhD supervisor Dr. Samrat Mukhopadhyay. I am elated to express my deepest gratitude to Professor Mukhopadhyay for giving me an opportunity to complete my PhD thesis under his supervision. He has been a thorough source of inspiration and motivation, for he is a professional individual dedicated to his work. His sheer penchant for precision, honesty, and integrity towards research transformed me from a raw and visionless researcher I was at the beginning of my PhD into a cultivated researcher who I am today, at least what my peers testify about me. He is the one who stood beside me, used his academic excellence and experience to help me wander this multidimensional field of science with utter freedom and fearlessness. He listened with patience, my novice ideas, and then with all his alchemistry, transformed them into appropriate research thoughts. This was just because he believes in his students and gave them intellectual freedom to explore beyond tangibility. He, through his hard work, extended working hours stretching from early morning to late evenings and disciplined routine, demonstrated what a brilliant and hardworking scientist could accomplish. He has that aura characterized by his enthusiasm and energy that transfers to the people around him, and I am blessed to be one of the recipients. There were various occasions when I felt exhaustive, disheartened, and ultimately low as all my efforts of the day and at times, many days went in vain when my experiments failed to deliver the results. In that darkness used to come to my rescue, a dictum that Samrat often says, “Experiments don’t work, you have to make them work.” These inspirational words helped me to recollect myself and again embark upon the journey to excellence and success. I am extremely grateful to him for the time and effort he has invested in me in teaching me how to write manuscripts and academic/professional communications. I would also like to thank him for providing me with numerous opportunities to interact and discuss my research work with the eminent scientists in the field. One thing that I admire the most about Samrat is that he doesn’t limit his knowledge to the field of science. He has an encyclopedic thirst for knowledge gathering and has decorated many coffee sessions and group lunch/dinner outings with amazing facts and information that made the interactions even more enjoyable. My scientific outlook is all stemming from him and his guidance. Truly, I inherit my inquisitiveness from Samrat.

Though formally I was doing my PhD under the supervision of Samrat, my scientific development would have remained stunted without contributions of Dr. Mily

Bhattacharya. Mily was the one who made me adamant about pursuing a career in research. She taught me starting from minutest of things such as how to use a pipette to dexterous endeavors such as designing/performing an experiment. I'm grateful to her for rousing discussions, critical feedback on my presentations and manuscripts. Be it turbulence in my professional life or personal life, Mily was always there to hold my back and strengthen me through the advice out of her own life experience. As a mother does not need to talk to understand her daughter to find that something is going wrong with her, the same was the case with Mily and me. She could infer from my facial expressions and body language that I am running through some rough patches. Indeed, she has been nothing short of an academic mother. All that I am, or hope to be, I owe to Mily.

I want to thank the members of my doctoral committee, Professor K. S. Viswanathan, Dr. Raj Kumar Roy, and Dr. Sabyasachi Rakshit, for their kind support, encouragement and valuable suggestions on my thesis work.

Further, I would like to thank all the current and former members of the Mukhopadhyay lab. They were a guiding light and comforting cushions in the formative years. I want to express my gratitude to Dr. Shruti Arya for teaching me the analysis of time-resolved fluorescence and Dr. Dominic Narang for teaching me the basics of protein expression and purification. Dr. Karishma Bhasne, Dr. Hema M. Swasthi, Dr. Sourvav Sinha Roy, and Dr. Anupa Majumdar for healthy discussions that have helped me shape many ideas. I would also like to thank Dr. Neha Jain and Dr. Vijit Dalal for their discussions and suggestions. I owe a special acknowledgment for Ashish who though joined recently, but has helped me through constructive brainstorming discussions and was indeed a great helping hand during strenuous purification processes. Priyanka Madhu, Debapriya, Aishwarya, Sayanta, Ashish, Anamika, Lisha, Sandeep, Anusha, and Swastik have been wonderful colleagues who upheld collegiality and made my tenure as part of the lab rather a memorable journey that I will always cherish. Ashish, Anamika, and Lisha have been a sink and source of various chits and chats that often related to work and otherwise too as well, that lightened me up whenever I felt the burden too heavy, lightning, and wonderful soul they are.

I owe special thanks to Mily, Ashish, Sourav, Anupa, Anamika, and Shruti for their contributions to my thesis work.

I gratefully acknowledge Professor Anindya Datta at the Indian Institute of Technology (IIT), Bombay, and his research scholars for their help in the femtosecond experiments.

I would like to thank Bhagyashri Mahajan, Suchitra S. Prabhu, Sreelakshmi M T, Achuthan Raja Venkatesh, for helping me with purification and a few preliminary experiments.

I am also thankful to Prof. Jeffery W. Kelly and Prof. Jennifer C. Lee for providing us with Pmel17 Mα and RPT plasmids, respectively.

I am grateful to the Department of Science and Technology, Govt. of India, for the DST-INSPIRE fellowship. I express my deep gratitude to DST, IISER, the Biophysical Society, and European Micro Biology Organization (EMBO) for the travel funding and logistic support for attending the international conferences.

I would like to extend my sincere gratitude to the founding Director of IISER Mohali, Prof. N. Sathyamurthy. I did my first research project under his supervision and I'm thankful to him for inculcating the research spirit in me. I would also like to extend my gratitude to Dr. P. Visakhi (librarian IISER Mohali) and entire IISER Mohali library staff for facilitating the library services to all the students and maintaining in good health one of the best libraries of the region.

I truly acknowledge Prof. Christopher M. Dobson, Prof. Richard Kriwacki, Prof. Anthony A. Hyman, Prof. Simon Alberti, Prof. Dora Tang, Prof. Dorothee Dormann, Prof. Rohit Pappu, Prof. Songi Han and Prof. Elizabeth Komives for their stimulating discussions and encouragement during the 62nd Annual Meeting of Biophysical Society (2018), San Francisco, USA, International Conference on Intrinsically Disordered Proteins: Forms, Functions, and Diseases (2017), Indian Institute of Science Education and Research (IISER) Mohali, and Phase Transitions in Polymeric and Protein Systems (2019), Max Planck Institute for the Physics of Complex Systems (MPI-PKS), Dresden, Germany.

Last but not least, I thank the most important people of my life, my parents, my brother, my husband and in-laws for their constant love, care, and encouragement. I feel highly elated and grateful for their support and contributions.

Priyanka Dogra

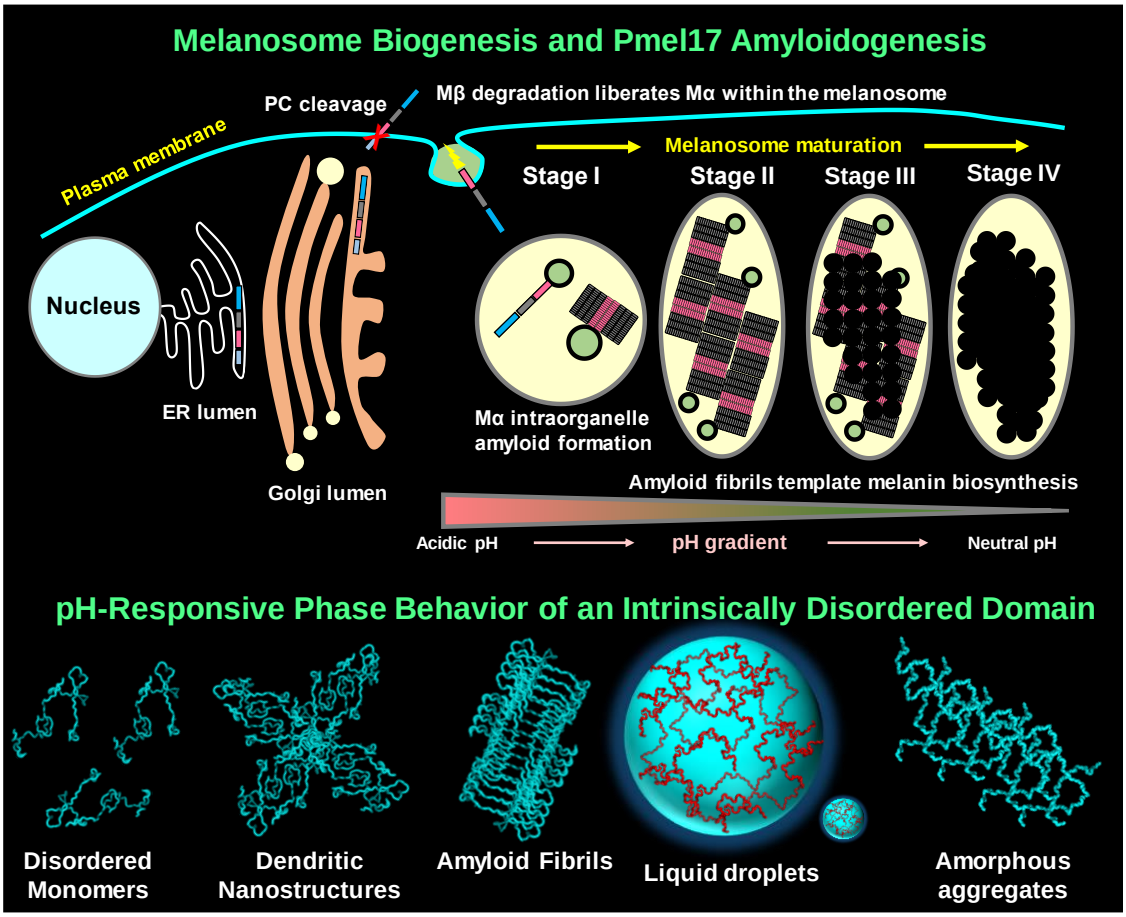
SYNOPSIS

PHASE BEHAVIOR OF AN INTRINSICALLY DISORDERED DOMAIN OF A MELANOSOMAL PROTEIN: CONFORMATIONAL CHARACTERISTICS, AMYLOID FORMATION, AND LIQUID-LIQUID PHASE SEPARATION

Chapter 1. Introduction

A growing body of current research has revealed that living cells can regulate the complex biomolecular chemistry by the spatiotemporal organization of a wide variety of different cellular components into functionally distinct intracellular liquid-like compartments or membrane-less organelles having different chemical environments. Under certain circumstances, protein misfolding occurs inside the cells, which leads to the accumulation of highly ordered cross-β sheet rich amyloid aggregates that have been implicated in many deadly neurodegenerative diseases such as Alzheimer's, Parkinson's, prion diseases, etc. However, recent studies have identified the beneficial role of amyloids in a multitude of organisms ranging from bacteria to humans' performing an array of physiological functions. Human Pmel17, a melanocyte-specific glycoprotein, forms functional amyloid that plays an essential role in melanosome development by creating a fibrillar amyloid matrix in the organelle, which acts as a template for melanin deposition underneath the skin and in the eyes. The amyloid matrix serves a beneficial role in mitigating the toxicity by sequestering and minimizing the diffusion of highly reactive quinone precursors that are required during melanin biosynthesis. It is known that an intrinsically disordered region (IDR) of Pmel17, the repeat domain (RPT) forms the amyloid core and promotes melanin formation *in vitro*. Several studies have shown that the deletion of the RPT ablates fibril formation *in vivo*. However, the molecular mechanism of amyloid formation, as well as the organization of individual protein molecules within the supramolecular assembly, remains elusive. An increasing body of work reveals that under certain physicochemical conditions, IDRs in proteins undergo liquid-liquid phase separation to form dense insoluble phases that have implications in both physiology and disease. These IDRs have an intrinsic preference for conformational disorder and are often characterized by low complexity (LC) domains. While numerous studies have discovered that LC-IDRs in proteins phase separate into mesoscopic liquid droplets, and the phase-separated state predisposes the

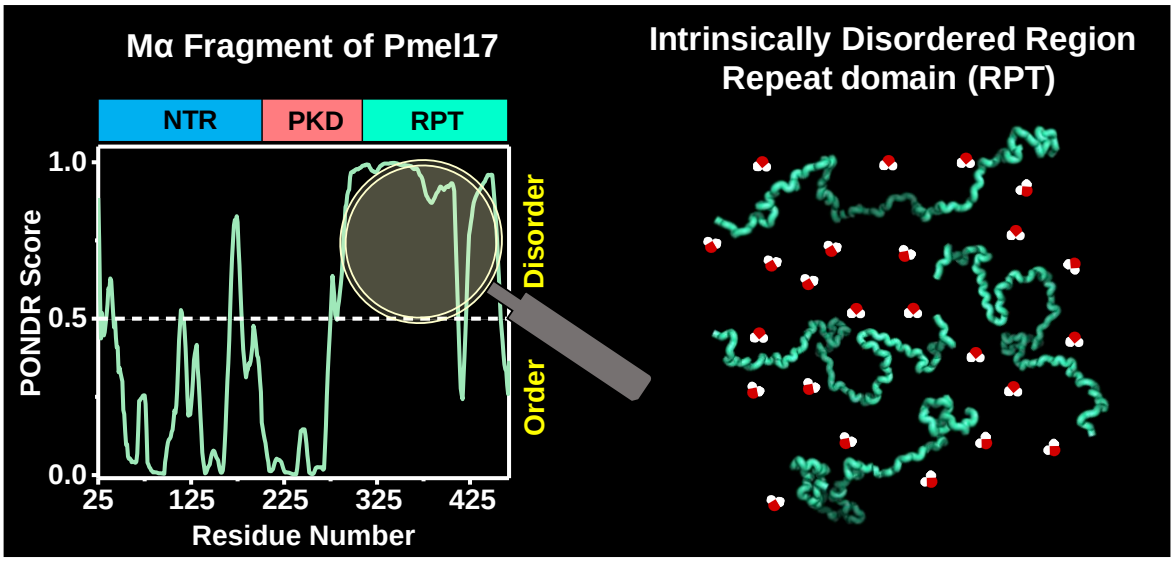
protein toward the formation of aggregates, the fundamental molecular drivers, and the sequence of events that govern the phase transitions is poorly understood. In this thesis, efforts were directed towards elucidating the molecular mechanism of amyloid formation and phase transitions of the RPT under various physicochemical conditions. The conformational dynamics, heterogeneity, and intermolecular association that drives RPT phase transitions were studied using a multidisciplinary approach involving a combination of biophysical, biochemical, molecular biology, and imaging tools.



Chapter 2. Conformational Dynamics and Solvation of a pH-Responsive Intrinsically Disordered Repeat Domain (RPT) of a Human Functional Amyloid Protein, Pmel17

Unlike globular/structured proteins, intrinsically disordered proteins (IDPs) lack the ability to undergo autonomous folding under the native condition and exist as dynamic ensembles of rapidly fluctuating interconverting structures. The dynamic nature of IDPs confers them the flexibility to adopt diverse conformational states that are typically encoded in their amino acid sequences. Protein dynamics cannot be understood without taking into account their

environment. The water molecules present in the vicinity of the protein surface is believed to have a profound influence on protein dynamics, folding, and aggregation. The entropy gain

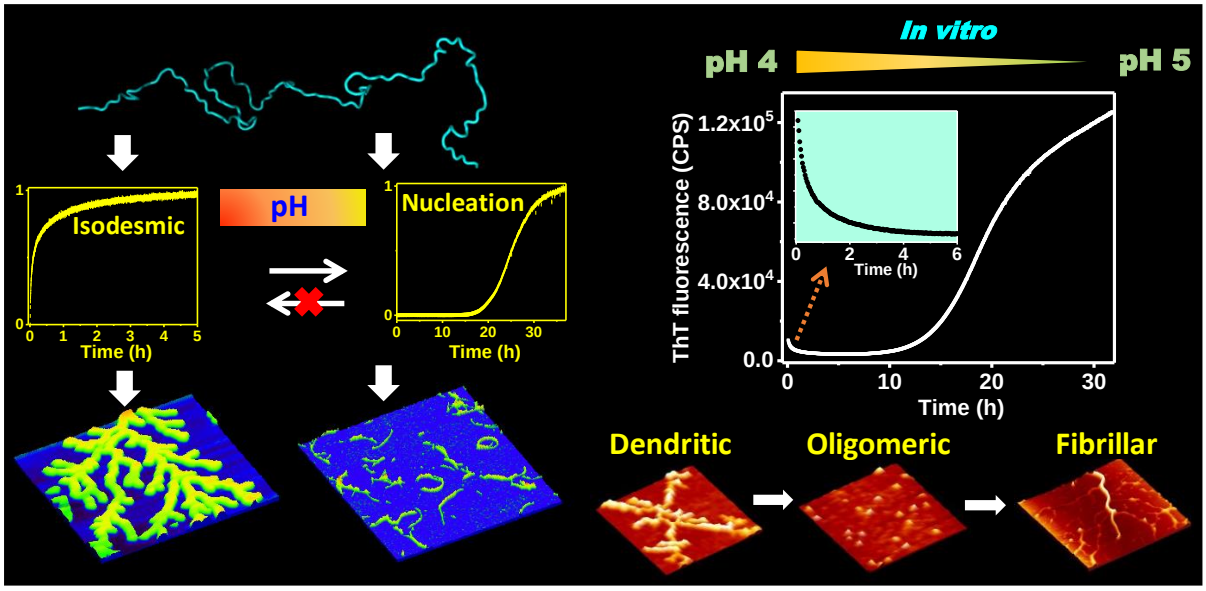


from the expulsion of water molecules residing in the hydration layer at the protein-water interface has been proposed to drive the aggregation process of IDPs. However, the behavior of water molecules in and around the disordered polypeptide chains is not fully understood. In this work, using femtosecond and picosecond time-resolved fluorescence measurements, we have probed the conformational- and hydration dynamics of the RPT in the disordered monomeric state. Our results indicate site-specific conformational changes and differences in the water arrangement within largely disordered RPT. The timescales of surface-bound water dynamics were found to be different for N-terminus and amyloid core region. Interestingly, the presence of ordered water around the amyloidogenic C-terminal core region underscores the importance of this region in the context of RPT amyloid formation. We anticipate that the differences in the water structure in distinct regions of the RPT might give rise to altered conformations that are likely to be associated with the distinct function of different regions.

Chapter 3. Mechanistic Insights into the Mechanism-to-Morphology Relationship in Amyloid Formation of a Pmel17 Fragment (Mα) and RPT

In contrast to pathological amyloids, functional amyloids are involved in crucial physiological functions. Understanding the mechanism-morphology-function relationship during the self-assembly of these amyloid fibrils is of great importance, both because they are implicated in many biological functions, and also because they have potential applications in the design of

amyloid-based novel functional nanomaterials. To design novel functional nanomaterials and to strategically control the desired nanoscale morphology, it is crucial to establish a link between the mechanism and morphology. In this work, we present a unique case of controlling the formation of dendritic and fibrillar nanostructures by switching the aggregation mechanism of functional amyloid derived from the amyloidogenic segment of the melanosomal protein. We demonstrate that the repeat domain of the melanosomal protein exhibits two distinct types of aggregation pathways that display nanoscale polymorphism in acidic pH. In the pH range of 4.5−6, the aggregation proceeds via a typical nucleation-dependent mechanism resulting in the



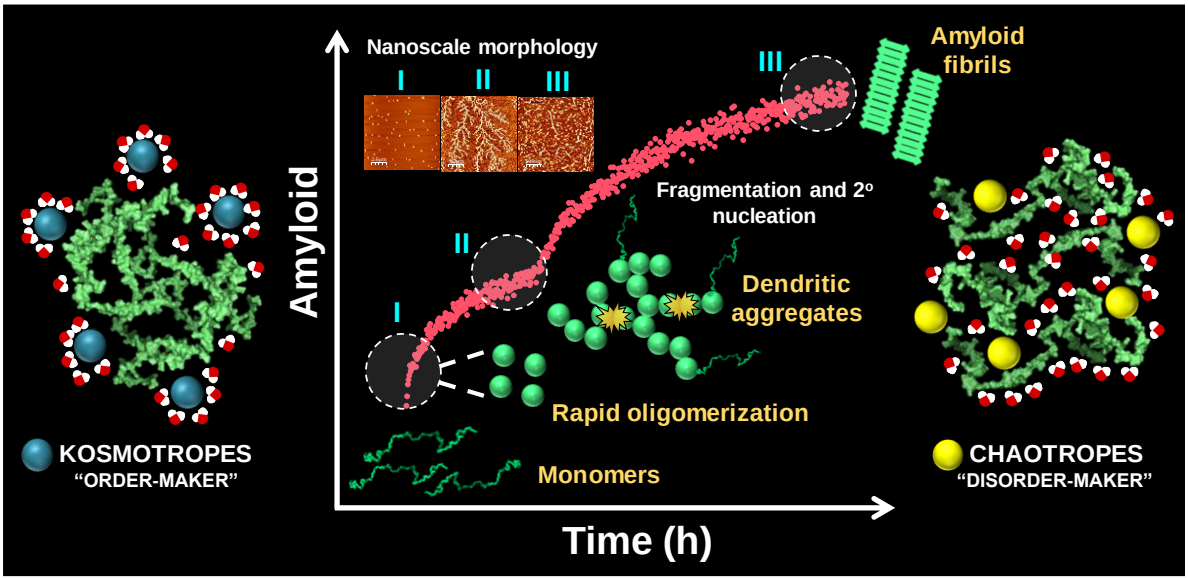
Adapted from P. Dogra et al. *J. Phys. Chem. B.* (2017)

formation of highly ordered β-rich curvy thread-like fibrils. On the contrary, at pH < 4.5 aggregation occurs through a rapid nucleation independent isodesmic polymerization process that yields dendritic aggregates having a lower degree of internal packing due to the presence of a partially ordered state. These dendritic nanostructures can be converted into more stable fibrils by switching the pH. The nanoscale polymorphism associated with the mechanistic switch is likely to be mediated by the altered conformational propensities and intermolecular interactions. Rapid aggregation without a lag phase at lower pH possibly indicates that there is little or no accumulation of toxic oligomeric species, which are efficiently recruited during aggregation. We suggest that the pH modulation within the melanosomes allows the optimal conditions for the formation of functional amyloids that dictate the template-assisted melanin biosynthesis. We propose that this striking shift in the mechanism that dictates the nanoscale morphology regulates the melanosomal maturation. In conclusion, our findings address key molecular aspects that relate the protein aggregation mechanism with the nanoscale

polymorphism and reveal a morphological change reminiscent of the amyloid transition observed during stage I to stage II of melanosome maturation.

Chapter 4. Hofmeister Ions Modulate the Autocatalytic Amplification of the Self-Assembly Process of Pmel17 RPT

Physicochemical properties of the solution such as temperature, pH, ionic strength, and the presence of cosolutes or cosolvent greatly influence protein stability, aggregation mechanism and are of great interest for materials and medical research. Among these factors, the specific ion effects on proteins predate all others. Previous reports have shown that ordered proteins are mainly influenced by hydration and solute effects where changes in solubility of protein are believed to be the outcome of the weak hydration around chaotropes, lending water molecules

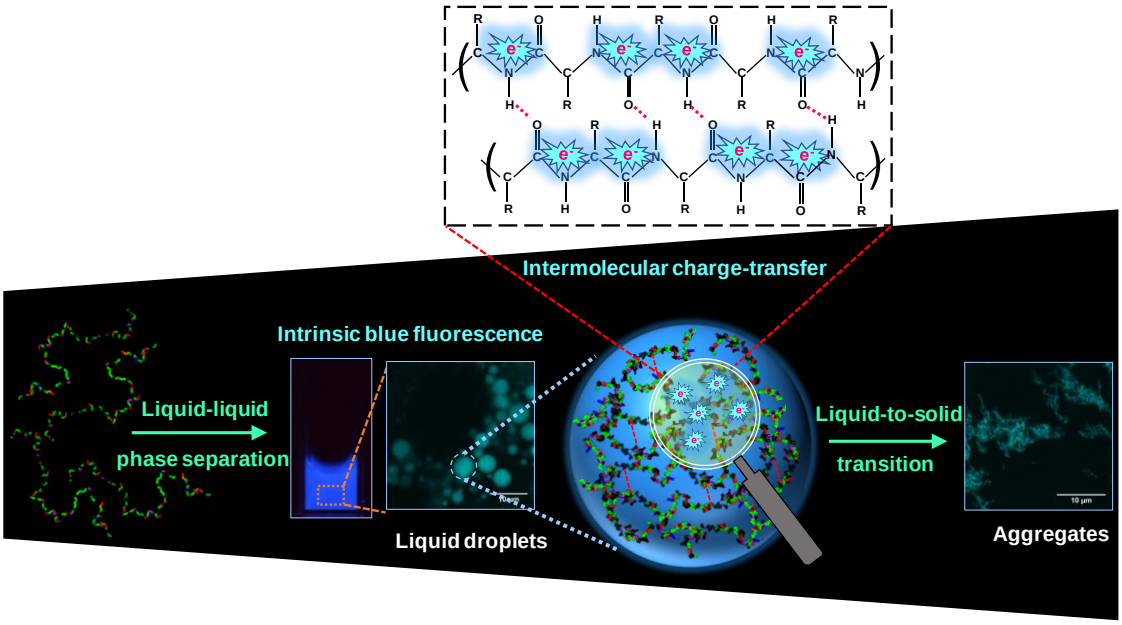


to the protein and the strong hydration around kosmotropes, removing water molecules from the immediate protein hydration shell. In contrary to the ordered proteins, IDPs/IDRs, are a complex system to show the influence of salt conditions on the aggregation and stability. With very few reports on the effect of salts on IDRs, there is still no clear understanding of the underlying molecular mechanism of salt-induced aggregation processes. We have previously shown that the RPT adopts a partially ordered state at pH 4. This study expands on previous findings and provides greater insight into how Hofmeister ions influence the oligomerization and fibrillization process of the RPT. Our results indicate that kosmotropes induces a greater extent of oligomerization of RPT in comparison with chaotropes. All salts examined, promote oligomerization, and the effect is in line with the Hofmeister series with the most pronounced

effect seen within the anions. Our results support the idea that the early intermolecular contacts leading to RPT oligomer formation are driven by both electrostatic as well as hydrophobic interactions and also provides insight into early contacts formation during oligomerization in a site-specific manner. Further, a unique biphasic behavior of the aggregation kinetics revealed a complex interplay of both fragmentation and secondary nucleation (autocatalytic amplification) processes during the self-association that is controlled by surface tension and Hofmeister ions effects, respectively. Our findings provide greater insight into salt-induced modulation in the aggregation process of the RPT and demonstrate an intriguing interplay of charge-peptide interaction, surface tension, and autocatalytic amplification in the aggregation pathway.

Chapter 5. Intermolecular Charge-Transfer Modulates Liquid-Liquid Phase Separation and Liquid-to-Solid Maturation of Pmel17 RPT

Cells regulate a diverse array of complex biochemical processes using functionally distinct membrane-bound organelles. A growing body of intense current research has revealed that cells can modulate spatiotemporal localization of cellular components, organize cellular biochemistry, and regulate critical cellular functions using nonmembrane-bound intracellular condensates termed as membrane-less organelles. These biomolecular condensates are liquid-like, mesoscopic, dynamic, non-stoichiometric supramolecular assemblies of proteins with or without nucleic acids and are thought to be formed via liquid-liquid phase separation (LLPS).



Adapted from P. Dogra et al. *J. Am. Chem. Soc.* (2019)

A majority of these condensates comprise multivalent IDPs/IDRs that are the key modulators of LLPS due to their sequence-encoded physicochemical properties, conformational heterogeneity, and flexibility. Often these proteins possess prion-like low-complexity regions (LCRs) that can participate in a multitude of specific but weak and transient intermolecular interactions resulting in liquid demixing. These condensates can mature from a functional liquid-like state to a pathological gel-like or solid-like state. In this work, we present a unique case to demonstrate that an unusual cascade of intermolecular charge-transfer coupled with a multitude of transient noncovalent interactions and conformational fluctuations can promote liquid phase condensation of a pH-responsive, intrinsically disordered, oligopeptide repeat domain of a melanosomal protein. At neutral cytosolic pH, the repeat domain forms highly dynamic, mesoscopic, permeable, liquid-like droplets possessing rapid internal diffusion and torsional fluctuations. These liquid condensates mature via pervasive intermolecular charge-transfer and persistent backbone interactions driving liquid-to-solid phase transition into heterogeneous solid-like aggregates that are structurally and morphologically distinct from typical amyloids formed at mildly acidic melanosomal pH. Our findings reveal the regulatory role of the repeat domain as a specific pH-sensor that critically controls the phase transition and self-assembly processes akin to prion-like low-complexity domains modulating intracellular phase separation.

Chapter 6. Conclusions

The work described in this thesis provides insights into the phase behavior of RPT, an IDR derived from a melanosomal protein, Pmel17. This chapter aims at constructing a bigger picture from the results described in the previous chapters. The salient features of the thesis include: (i) the conformational dynamics and solvation of an amyloidogenic IDR of Pmel17 (Chapter 2), (ii) the mechanism-morphology-function relationship in amyloid formation of Pmel17 (Chapter 3), (iii) the effect of Hofmeister ions in the autocatalytic amplification of the self-assembly process of RPT; structural sensitivity to co-solutes and its consequences on the oligomerization rates and aggregation kinetics (Chapter 4) and (iv) how intermolecular charge-transfer modulates liquid-liquid phase separation and liquid-to-solid maturation of RPT (Chapter 5). The work described in this thesis will improve our current understanding of amyloid formation and phase transitions in physiology and disease.

List of Publications

- **P. Dogra**, A. Joshi, A. Majumdar, and S. Mukhopadhyay. " Intermolecular Charge-Transfer Modulates Liquid–Liquid Phase Separation and Liquid-to-Solid Maturation of an Intrinsically Disordered pH-Responsive Domain" ***J. Am. Chem. Soc.***, **2019**, 141, 20380-20389

- **P. Dogra**, M. Bhattacharya, and S. Mukhopadhyay. "pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid" ***J. Phys. Chem. B***, **2017**, 121, 412-419.

- **P. Dogra**, S. S. Roy, A. Joshi, and S. Mukhopadhyay " Hofmeister Ions Modulate the Autocatalytic Amplification of the Self-Assembly Process of Pmel17 RPT " (manuscript in preparation)

- **P. Dogra**, S. Arya, A. Singh, A. Datta, and S. Mukhopadhyay " Conformational Dynamics and Solvation of an Intrinsically Disordered Domain of a Human Functional Amyloid Protein, Pmel17" (manuscript in preparation)

- **P. Dogra** and S. Mukhopadhyay "Dynamism of Intrinsically Disordered Proteins in Liquid-Liquid Phase Separation and Liquid-to-Solid Maturation" (manuscript in preparation)

- A. Majumdar, **P. Dogra**, S. Maity and S. Mukhopadhyay "Liquid-Liquid Phase Separation is Driven by Large-Scale Conformational Unwinding and Fluctuations of Intrinsically Disordered Protein Molecules" ***J. Phys. Chem. Lett.***, **2019**, 10, 3929−3936.

- S. Arya, A. Singh, K. Bhasne, **P. Dogra**, A. Datta, P. Das and S. Mukhopadhyay. "Femt--osecond Hydration Map of Intrinsically Disordered α-Synuclein" ***Biophys. J.***, **2018**, 114, 2540–2551.

- S. Arya, **P. Dogra**, N. Jain and S. Mukhopadhyay. "Detergent-induced Aggregation of an Amyloidogenic Intrinsically Disordered Protein" ***J. Chem. Sci***, **2017**, 129, 1817-1827.

- M. Bhattacharya and **P. Dogra** (2015) "Self-Assembly of Ovalbumin Amyloid Pores: Effects on Membrane Permeabilization, Dipole Potential, and Bilayer Fluidity" ***Langmuir***, **2015**, 31, 8911-8922.

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

- Not part of the thesis work

Presentations and abstracts

- **Priyanka Dogra** & Samrat Mukhopadhyay: "Liquid-liquid phase separation of an intrinsically disordered region of a functional amyloid-forming protein" in Phase Transitions in Polymeric and Protein Systems **(2019)**, Max Planck Institute for the Physics of Complex Systems (MPI-PKS), Dresden, Germany.

- **Priyanka Dogra**, Sourav Singha Roy, Mily Bhattacharya, Suchitra S. Prabhu & Samrat Mukhopadhyay: "Proton-induced Switching of an Intrinsically Disordered Domain of a Melanosomal Protein into a Polymorphic Functional Amyloid" at the 62nd Annual Meeting of Biophysical Society **(2018)**, San Francisco, California, USA.

- **Priyanka Dogra**, Mily Bhattacharya, Sourav Singha Roy, Suchitra S. Prabhu & Samrat Mukhopadhyay: "Mechanistic Switching in Aggregation of an Amyloidogenic Intrinsically Disordered Domain of the Melanosomal Protein" at the International Conference on Intrinsically Disordered Proteins: Forms, Functions and Diseases **(2017)**, Indian Institute of Science Education and Research (IISER) Mohali.

- **Priyanka Dogra**, Mily Bhattacharya, Sourav Singha Roy & Samrat Mukhopadhyay: "Mechanistic Switching in Protein Aggregation: Nanoscale Diversity of Amyloids Derived from Melanosomal Protein" at the Annual Symposium of the Indian Biophysical Society **(2017)**, Indian Institute of Science Education and Research (IISER) Mohali.

- **Priyanka Dogra**, Mily Bhattacharya, Sreelakshmi M.T. & Samrat Mukhopadhyay: "Mechanistic Switching in Protein Aggregation: Nanoscale Diversity of Amyloids Derived from Melanosomal Protein" at the 3rd International Symposium on Protein Folding and Dynamics **(2016)**, National Centre for Biological Sciences (NCBS), Bangalore, India.

- **Priyanka Dogra**, Mily Bhattacharya, Neha Jain & Samrat Mukhopadhyay: "Protein Conformational Changes Coupled with Nanoscale Morphological Transitions during Amyloid Formation" at the 2nd International Symposium on Protein Folding and Dynamics **(2014)**, National Centre for Biological Sciences (NCBS), Bangalore, India.

Table of Contents

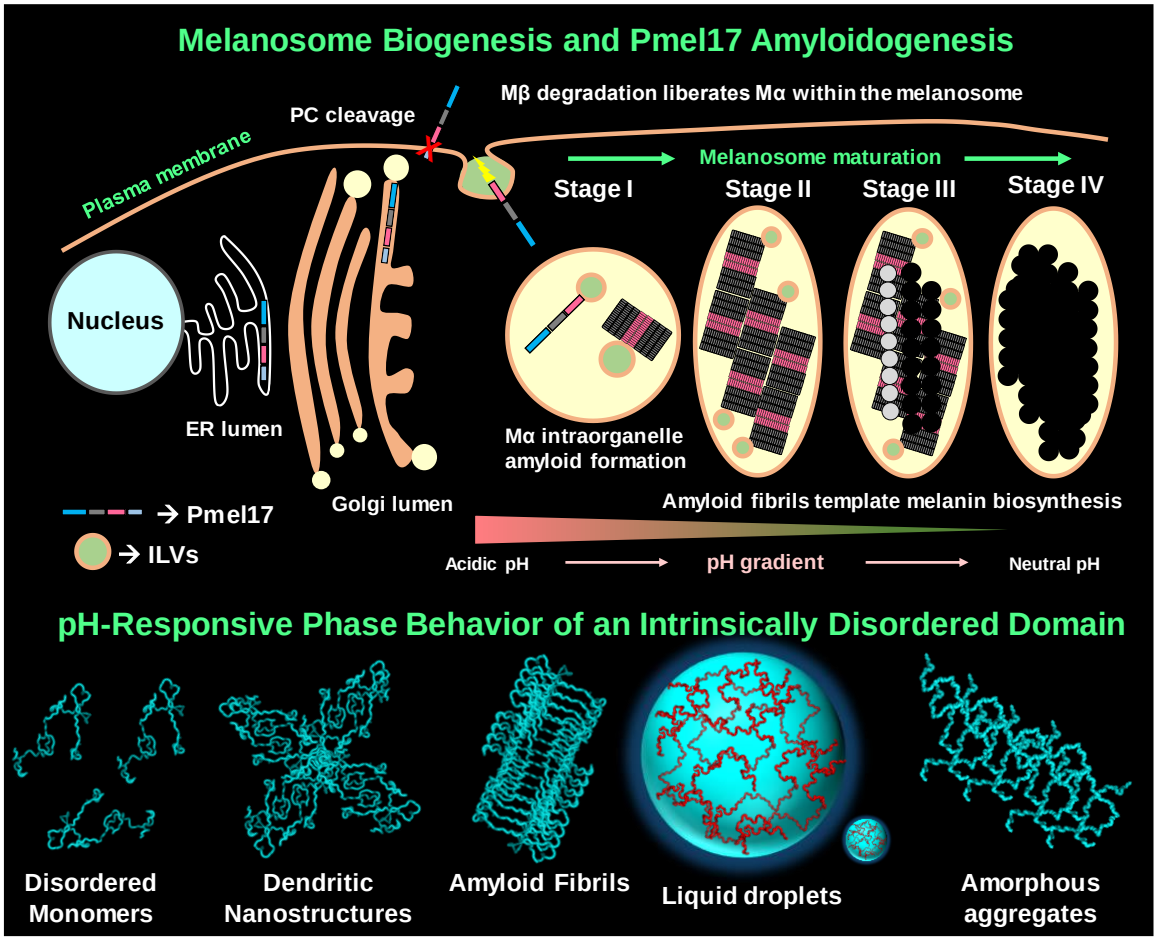
List of Abbreviations

ALS	Amyotrophic Lateral Sclerosis
AD	Alzheimer's Disease
AFM	Atomic Force Microscopy
ANS	8-Anilinonaphthalene-1-Sulfonic Acid
Acrylodan	6-Acryloyl-2-(dimethylamino)naphthalene
ADAM	A Disintegrin and Metalloproteinase
Aβ	Amyloid-β Peptide
β-BBO	Beta Barium Borate
CD	Circular Dichroism
CH	Charge-Hydropathy
CCD	Charge-Coupled Device
CR	Congo red
CTF	C-Terminal Fragment
DisProt	Database of Protein Disorder
D^2P^2	Database of Disordered Protein Predictions
DDX4	DEAD-Box Helicase 4
DOPA	L-3,4-Dihydroxyphenylalanine
DHQ	5,6-Indolequinone
DHI	5,6-dihydroxyindole
DHICA	5,6-dihydroxyindole-2-carboxylic acid
DTT	Dithiothreitol
DCVJ	9-(2,2-Dicyanovinyl)julolidine
DAPI	4′,6-Diamidino-2-Phenylindole
ELM	Eukaryotic Linear Motif
EWS	Ewing Sarcoma Protein
ELPs	Elastin-Like Polypeptides
EM	Electron Microscopy
ER	Endoplasmic Reticulum
FUS	Fused in Sarcoma
FCR	Fraction of Charged Residues

FTDs	Fronto-Temporal Dementias
FRAP	Fluorescence Recovery After Photobleaching
FWHM	Full Width at Half Maxima
FRET	Fluorescence Resonance Energy Transfer
G3BP1	GTPase-Activating Protein-Binding Protein
hnRNPA1	Heterogeneous Nuclear Ribonucleoprotein A1
hnRNPA2B1	Heterogeneous Nuclear Ribonucleoprotein A2/B1
HEPES	4-(2-hydroxyethyl)-1-Piperazineethanesulfonic Acid
HD	Huntington's Disease
HPNIR	High Power Near Infrared
IDPs	Intrinsically Disordered Proteins
IDRs	Intrinsically Disordered Regions
IUPred	Prediction of Intrinsically Unstructured Proteins
IDEAL	Intrinsically Disordered proteins with Extensive Annotations and Literature
IDPPs	Intrinsically Disordered Protein Polymers
ILVs	Intraluminal Vesicles
IRF	Instrument Response Function
IPTG	Isopropyl-thiogalactopyranoside
IAPP	Islet Amyloid Polypeptide
IAEDANS	5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1-Sulfonic Acid)
LLPS	Liquid-Liquid Phase Separation
LC	Low-Complexity
LAF-1	Lethal and Feminizing
LCRs	Regions of Low Sequence Complexity/ Low Complexity Regions
LCST	Lower Critical Solution Temperature
LARKS	Low-Complexity Aromatic-Rich Kinked Segments
MoRFs	Molecular Recognition Features
MLOs	Membrane-Less Organelles
MVBs	Multivesicular Bodies
MES	2-(N-Morpholino)EthaneSulfonic Acid Hemisodium Salt
NCPR	Net Charge per Residue
NMR	Nuclear Magnetic Resonance

NTD	N-Terminal Domain
NOE	Nuclear Overhauser Effect
NTA	Nitrilotriacetic Acid
PONDR	Predictor of Natural Disordered Regions
PTMs	Posttranslational modifications
Pmel17	Premelanosome Protein 17
PrLDs	Prion-Like Domains
PrLCDs	Prion-Like Low-Complexity Domains
PGD	Paget's Disease
PC	Proprotein Convertase
PQE	Polyglutamine Expansion Disease
Pab1	Poly-A Binding Protein
PKD	Polycystic Kidney Disease Domain
P1	Precursor 1
P2	Precursor 2
PMT	Photomultiplier Tube
RPT	Repeat Domain of Pmel17
RRMs	RNA Recognition Motif
RBPs	RNA Binding Proteins
ROI	Region of Interest
SLiM	Short Linear Motif
SOD1	Superoxide dismutase 1
SCA2	Spinocerebellar Ataxia Type 2
SMART	Simple Modular Architecture Research Tool
TDFSS	Time-Dependent Fluorescence Stokes Shift
TDP-43	TAR DNA-binding protein
TAF15	TATA-Binding Protein-Associated Factor 15
TGN	Trans-Golgi Network
TCSPC	Time-Correlated Single Photon Counting
TRES	Time-Resolved Emission Spectra
TDFSS	Time-Dependent Fluorescence Stokes Shift
ThT	Thioflavin T
UCST	Upper Critical Solution Temperature

Introduction



1.1 Intrinsically disordered proteins

Proteins are the most versatile macromolecules in the living cell that perform critical roles in various biological processes, including catalyzing biochemical reactions, immune protection, transportation and storage of molecules, etc.[1] According to the structure-function paradigm, the physiological function of a protein is achieved by its unique, well-defined 3D structure that is encoded in its amino acid sequence.[2,3,4] Another view of this paradigm is the "lock and key" analogy postulated by Emil Fischer that explains the specificity of enzymes.[5] Later, the "induced fit" model was proposed, which states that the active site of the protein/enzyme undergoes a conformational change to improve binding with the substrate.[6] Consequently, this focus on a notion that not all the native proteins or region of proteins possess a unique rigid conformation but instead can exist as conformational ensembles.[7] Later on in the 1980s, several researchers demonstrated that missing domains of electron density of numerous proteins associated with the lack of structural flexibility can be crucial for biological function.[8,9] In the late 1980s, a myriad of research articles reviewed by Sigler suggested the existence of "ill-defined structures" referred to fas "acid blobs or negative noodles" that carry out an important function in a variety of transcription factors.[8] Likewise, Holt and Sawyer proposed the term "rheomorphic protein" to describe the flexible and open conformation of the caseins.[8] Around the same time frame, it was established that the tau protein in solution resembled a Gaussian polymer that exhibited a lack of secondary structure and "compact folded conformation".[8] In 1996, another biologically relevant protein, namely α-synuclein was described to show a similar conformational behavior. α-synuclein comprises a highly charged amino acid residues and adopt a "random coil" structure as demonstrated by circular dichroism spectroscopy, possess high stability to heat denaturation, shows an uncannily high Stokes radius, and an unusual SDS binding resulting into an anomalous mobility on SDS-PAGE.[8] Furthermore, in 1998 Dunker and coworkers predicted using the Swiss Protein Database that more than 15000 proteins contain disordered regions of at least 40 consecutive amino acids.[10] In the following year (1999) Wright and Dyson uncovered that several protein domains/regions which remain unstructured in solution, become structured upon binding to their respective partners.[11] This special class of proteins is known as natively unfolded or intrinsically disordered proteins that confronts the structure-function paradigm.[11-17] In general, to

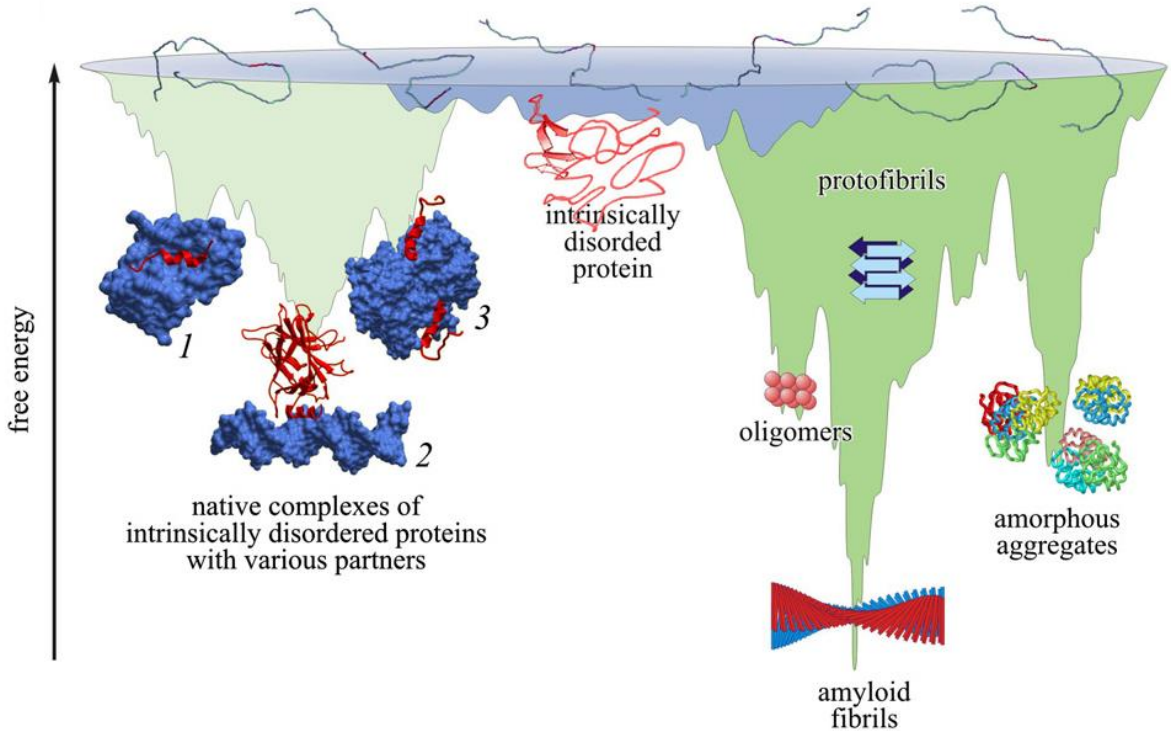


Figure 1.1 A schematic representation of the energy landscape of protein folding (light green) and aggregation (green) illustrating the formation of native globular and intrinsically disordered proteins, oligomers, amyloid fibrils, and amorphous aggregates via intermolecular contacts. The disordered segments of several native proteins interact with specific binding partners to attain ordered structures that have lower free energies than the intrinsically disordered protein and its partner. 1, 2, and 3 represent native complexes of intrinsically disordered proteins with various partners. The propensity of native completely or partially disordered proteins to interact with various partners determines their biological functions in recognition, regulation, and signal transduction. However, this conformational plasticity also allows intrinsically disordered proteins to form pathological amyloids. Reproduced with permission from Ref. 20.

achieve the functional native state, a protein molecule needs to adopt a unique conformation either via folding or by interaction with the binding partners.[3,7,18-21] As stated in the 'Levinthal paradox', due to astronomically large conformational states of the polypeptide chain, it would take an enormously long time for proteins to find the stable (lowest energy) native conformation, yet proteins fold rapidly within seconds or less.[3,18] However, the current view of the protein folding process involves the stochastic search of diverse conformational states accessible to the polypeptide chain and describes the mapping of the protein conformation to its free energy.[3,7,18,19] In this model, different pathways of protein folding emanate from the free-energy landscape/surface framework (Figure 1.1). In addition to elucidating the mechanism of globular

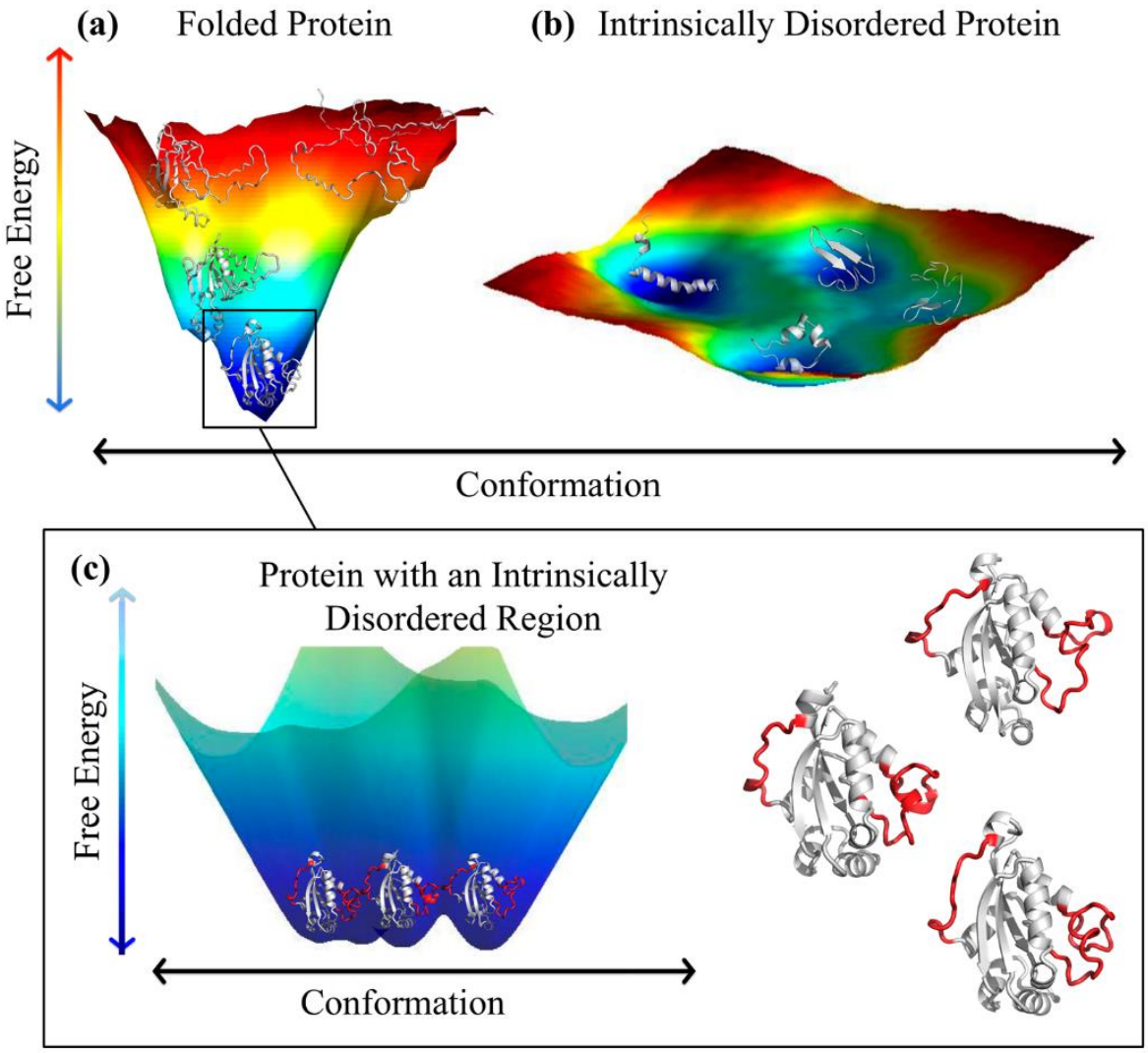


Figure 1.2 Schematic of energy landscapes for **(a)** a folded protein (human nucleoside diphosphate kinase (NDPK), PDB ID: 1nsk), and **(b)** an intrinsically disordered peptide (CcdA C-terminal, PDB ID: 3tcj). The lower free energy (dark blue) represents more probable conformations. **(c)** close-up of the minimal free energy well in (a), where IDRs are shown in red and ordered regions are shown in white. Reproduced with permission from Ref. 25.

protein folding, this energy landscape model also demonstrates the binding of native complexes of intrinsically disordered proteins (IDPs) with various partners.[20] The unfolded polypeptide chains present at the top of the so-called 'energy funnel' represents the dynamic ensembles of rapidly interconverting conformational states (Figure 1.1). The disorder-to-order transition from the unfolded state to the uniquely-folded native state often involves the formation of partially folded intermediates (metastable state) that gets transiently populated during the folding process.[21] The

accumulation of these intermediates further triggers the formation of potentially pathogenic oligomers, amorphous aggregates, protofibrils, and amyloid fibrils.[4,20,21,22] It is interesting to note that many pathological amyloid fibrils are associated with proteins that are either IDPs or have intrinsically disordered regions (IDRs).[23] IDPs also perform a wide variety of complex functions, including signaling, cell division, cell cycle control, regulation of transcription and translation, etc.[16,24] Interestingly, nearly 33% of the eukaryotic genome codes for IDPs or IDRs.[24] Globular/folded proteins can be distinguished from the IDPs based on their potential energy landscapes (Figure 1.2).[25] In the case of folded proteins, there is a "funnel-shaped" global energy minimum, which corresponds to the lowest energy native state, and it is the global energy minimum and its width that governs native state's conformational entropy .[25] On the contrary, IDPs have several local energy minima that are parted by small barriers, which result in the rapid interconversion between different local energy minima, resulting into an ensemble of a large number of conformational states having almost identical energies.[25] Therefore, IDPs exist as dynamic ensembles of rapidly interconvertible conformations in their native state.[9] Also, the energy minimum of a folded protein with an IDR has a rough surface with several smaller minima that represent distinct states sampled by the IDR within the native state (Figure 1.2).

1.1.1 Structural-sequence ensemble relationships

The dynamic nature of IDPs allows them to adopt a spectrum of conformational states, each of which is involved in performing crucial functions.[24] Many of these conformational states are akin to the non-native states of globular/structures proteins. According to the protein quartet model, the function of protein can emanate from four kinds of conformational states and the transitions between them.[24,26,27] These states comprise ordered (solid-like), pre-molten globule & molten globule (liquid-like), and random coil (gas-like) conformations, as shown in Figure 1.3a. Several studies have indicated that the pre-molten globule state has a high propensity towards binding-induced folding events.[24] Proteins present in the pre-molten state exhibit a residual secondary structure that is absent in the random coil state. One of the examples of functional significance of different conformations is the unique family of nuclear pore complex proteins termed as FG nucleoporins.[24] The intrinsic preference to remain unfolded in the native state point towards peculiar features of their amino-acid sequences.[8,26,27,28] The composition and frequency of appearance of amino acids in IDPs is quite distinct from the globular/structured proteins. The plot

in Figure 1.3b represents the average amino acid frequencies in two IDP datasets in comparison with the average frequencies found in ordered globular proteins, which is expressed as ((Disordered−Globular-3D))/(Globular-3D).[8,27,28] The negative value in the plot correspond to the amino acids that are deficient in the disordered regions compared with the ordered ones, and positive values indicate the amino acids that are enriched in the disordered segments.[27,28] This composition pattern of amino acid residues reveals a low overall hydrophobicity and high net charge in the sequence that is a characteristic of the natively unfolded proteins.[8,28] IDPs are often enriched in polar and charged residues such as G, Q, S, P, A, E, K, R, also called the "disorder promoting amino acids" and are deficient in hydrophobic and aromatic residues, I, L, V, W, F, Y, C, N that are termed as "order promoting amino acids".[28] Several computational studies coupled with experimental observations have embarked on the relationship between information coded in amino acid sequences and the conformational ensemble accessible to IDPs under different environments.[24,29,30] In light of these studies, IDPs are categorized into three distinct compositional

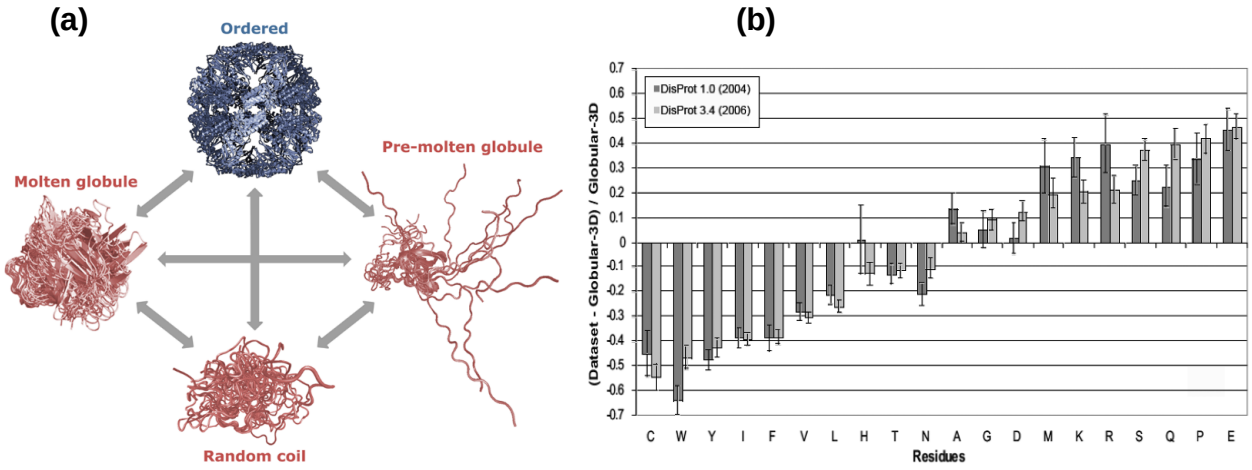


Figure 1.3 (a) The protein quartet model of protein conformational states. The function of a protein arises from any of the four conformation states of the polypeptide chain (ordered forms, molten globules, pre-molten globules, and random coils) and transitions between them. **(b)** The differences between the amino acid compositions of disordered data sets (DisProt 1.0 and Disprot 3.4) and that of an ordered data set (Globular-3D) were plotted using the DisProt database. The plot indicates that IDPs possess a tendency to be enriched in disorder-promoting amino acids and depleted in order-promoting amino acids. Reproduced with permission from Ref. 24 and 28.

classes that are based on the fraction of charged versus polar residues.[24,29] The fraction of charged residues or FCR in an IDP sequence is quantified as the sum of the fraction of positively charged

residues (f+) and fraction of negatively charged residues (f-), whereas, the net charge per residue or NCPR is defined as the difference of f+ and f-.[29,31] Figure 1.4a shows the three major distinct classes: polar tracts, polyelectrolytes, and polyampholytes.[32] Polar tracts are enriched in polar amino acids such as serine, threonine, asparagine, glycine, glutamine, and deficient in charged, hydrophobic, proline residues. Polyelectrolytes are enriched in charged residues and are biased towards one type of charged residues, i.e., f+ > f- or f- > f+, as observed in Glu/Asp-rich prothymosin α or the arginine-rich protamines.[24,29] On the other hand, polyampholytes have a roughly equal number of positive and negative charges, i.e., f+≈ f-.[29] Depending on the FCR values, polyampholytes and polyelectrolytes can be designated either weak or strong. High FCR value corresponds to strong polyelectrolytes/polyampholytes and encodes an intrinsic propensity

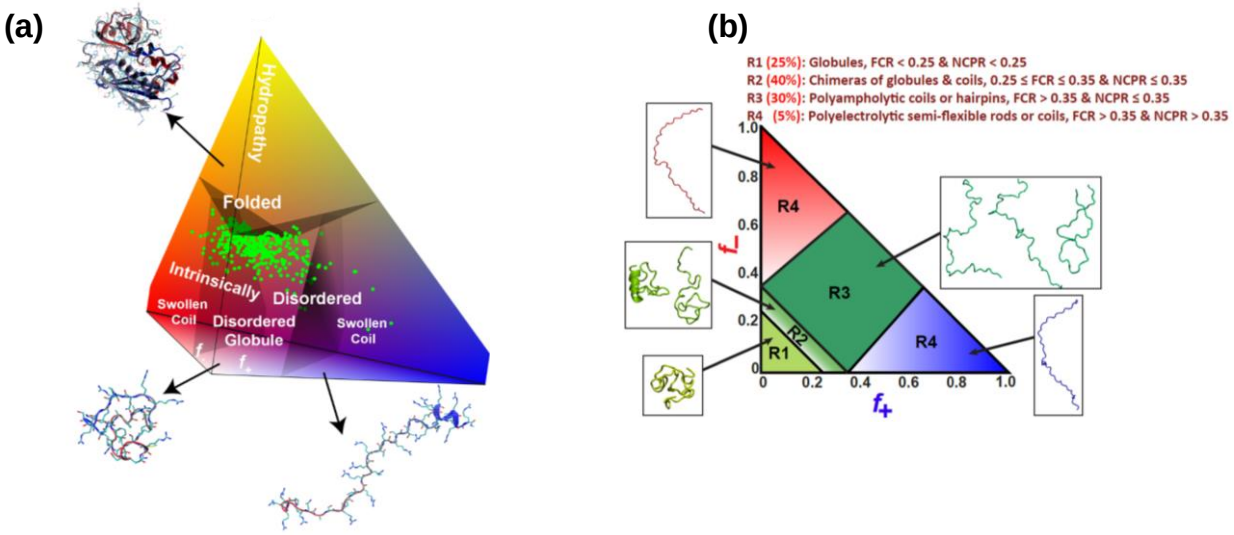


Figure 1.4 Original and modified diagram-of-states to classify predicted conformational properties of IDPs. **(a)** The original diagram constitutes three axes that represent the fraction of positively charged residues, f+, the fraction of negatively charged residues, f−, and the hydropathy. These three parameters are calculated from the amino acid composition. The green dots in the diagram represent 364 curated disordered sequences extracted from the DisProt database. **(b)** The modified diagram-of-states from panel (a) with a focus only on the bottom portion of the pyramid i.e., on f+ and f- axes since the hydropathy for IDPs is intrinsically low. The disordered globules space in panel (a) is subdivided into four distinct regions of which sequences in regions 2 and 3 are expected to be random-coil-like if oppositely charged residues are well mixed in the linear sequence. Otherwise, one can expect compact or semi-compact conformations. Statistics for different regions (percentages) are from the analysis of bona fide IDPs in DISPROT. Reproduced with permission from Ref. 24.

for populating expanded coil-like conformations due to the preference of charged residues for solvation in the aqueous environments.[29] However, an increase in the NCPR value beyond a certain threshold can reverse the propensity of polypeptide backbones to form ensembles of collapsed conformations.[24,29] Therefore, both FCR and NCPR play critical roles in predicting the conformation ensemble for an IDP. The modified diagram-of-states summarizes the relationship between the amino acid composition and conformation of IDPs (Figure 1.4b) and is valid only for sequences that have low overall hydropathy, low proline content, and at least thirty residues.[29] Based on this modified diagram-of-states, IDPs are categorized into four conformationally distinct classes referred to as R1, R2, R3, and R4.[24,29] Polar tracts and weak polyelectrolytes/polyampholytes that form disordered globules fall into the region R1 and strong polyampholytes that exist as either hairpins or coils constitute the region R3. [24,29] The region R2 have conformational characteristics of both R1 and R3. As it lies between R1 and R3, therefore, the R2 region is likely to be chimeras of globules and coils. The region R4 extends over two areas of the diagrams that correspond to acid-rich (f- > f+) and base-rich (f+ > f-) polyelectrolytes, respectively. [24,29] This region constitutes semi-flexible to worm-like chain conformational ensembles.[24] Overall, the sequence-conformation ensemble relationship allows the use of de novo sequence design as a tool for controlling the properties of IDPs and IDRs and also evaluating their influence on the functions.

1.1.2 Prediction of intrinsic disorder from sequence

Several prediction tools/methods have been developed for the recognition of disordered regions from the amino acid sequence to understand the functional role of intrinsic disorder and identify the binding sites. These include CH plot, PONDR VLXT, PONDR VSL2, PONDR VL3, PONDR FIT, IUPred, ESpritz, IDEAL, metaPrDOS, MobiDB, D^2P^2, etc.[33] The charge-hydropathy (CH) plot or Uversky plot of net charge vs. mean hydrophobicity allows us to distinguish between the disordered and ordered proteins based on the physical/chemical features of amino acid residues.[15,34] Under physiological conditions, globular/ordered proteins contain a high content of hydrophobic residues with low net charge that favors the formation of collapsed structures.[15,34] However, the presence of a high net charge and low mean hydrophobicity is an important prerequisite for the disordered/expanded structure in proteins due to charge-charge repulsion and small driving force to undergo compaction, respectively. This latter feature is a characteristic of

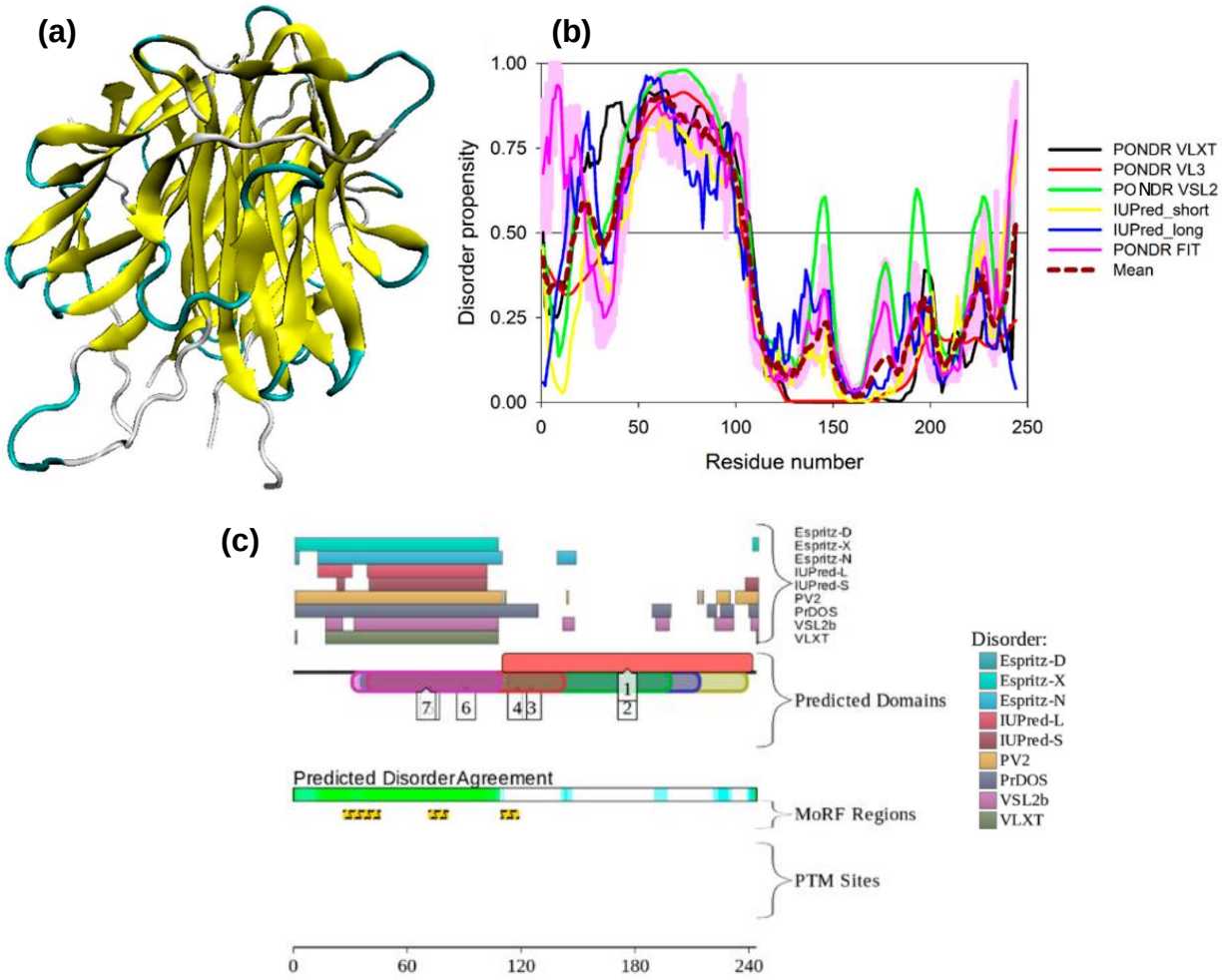


Figure 1.5 Structure and intrinsic disorder in human adiponectin (UniProt ID: Q15848). **(a)** Crystal structure of a single-chain trimer of human adiponectin globular domain (PDB ID: 4DOU). **(b)** Evaluating intrinsic disorder propensity of human adiponectin by series of per-residue disorder predictors. Disorder profiles generated by PONDR VLXT, PONDR VSL2, PONDR VL3, IUPred-short, IUPred-long, and PONDR FIT. **(c)** Analysis of the intrinsic disorder propensity and some important disorder-related functional information generated for human adiponectin by the D^2P^2 database. Reproduced with permission from Ref. 35.

IDPs/IDRs. Therefore, the CH plot is one of the simple and efficient computational tools that can be utilized for the analysis of structural disorder in proteins. Protein disorder is also associated with low sequence complexity, i.e., low variance in amino acid composition. Using machine-learning algorithms, three key databases were generated by collecting and organizing information taking into consideration the experimental analysis and the functionalities of IDRs and IDPs.[35] These databases comprise the Database of Protein Disorder (DisProt), Intrinsically Disordered proteins with Extensive Annotations and Literature (IDEAL), and MobiDB (annotation of protein mobility and disorder). One of the most widely used predictors that fall in the category of these

databases is the Predictor of Natural Disordered Regions (PONDR) that discriminates between the ordered and disordered segments from the primary amino acid sequence data using the nonlinear models (feed-forward neural networks).[36] PONDR was developed to assess whether the disorder or lack of structure can be anticipated from the amino acid sequence. Over the years, PONDR has undergone significant development, and currently, several versions of this predictor are available with the increased accuracy and reliability that includes PONDR VL-XT, VL3, and VSL2 predictors.[33] These predictors are trained to distinguish order and disorder based on several parameters, including average coordination number, net charge, and amino acid compositions (aromatic and charged residues). Additionally, the PONDR VL-XT predictor was suggested to be more sensitive than other predictors in identifying the molecular recognition features (MoRFs) and also the regions potentially undergoing disorder-to-order transition.[36] Another similar predictor of disordered proteins is IUPred that captures the fundamental difference between the ordered and disordered regions based on the energy estimation approach by determining the interresidue interaction energies.[37] For instance, globular/ordered proteins can form a large number of interresidue contacts; however, due to the unique sequence composition and lack of the ability to make stable contacts, IDPs/IDRs do not form sufficient interresidue interactions. Both in the case of IUPred and PONDR, the probabilistic score ranging from 0 to 1 represents the disorder propensity, where the value 0 and 1 corresponds to complete order and disorder, respectively. IUPred is trained for predicting long disordered segments, whereas, PONDR performs well both on short and long disordered regions.[38] Furthermore, another protein disorder predictor termed ESpritz was developed, which is a sequence-only predictor and is four orders of magnitude faster than multiple sequence alignment-based methods.[39] A unique feature of ESpritz is the explicit prediction of mobile residues in NMR structural ensembles.[39] Recently, the meta-predictors have been developed that couple numerous disorder prediction methods that are individually successful, resulting in increased prediction accuracy. For instance, metaPrDOS apply support vector machines (SVM) based machine learning from the prediction results of the seven independent predictors to arrive at a final score.[40] The IDEAL database assemble annotations of emperically verified IDPs and emphasizes on regions that experience coupled folding and binding upon interaction with other proteins. MobiDB collects experimental data on IDRs from DisProt, IDEAL, and the Protein Data Bank (structurally mobile regions in NMR ensembles and missing residues in crystal structures).[41] Finally, the Database of Disordered Protein Predictions (D^2P^2),

akin to MobiDB, stores the disorder predictions generated by a large number of independent prediction methods for proteins from completely sequenced genomes. In addition to the disorder predictions, it also contains information on MoRFs, PTM sites, and domains.[42] D^2P^2 is believed to provide greater insights into the interplay between disorder and structure, the genomic distribution of disorder, and its evolutionary history.[42] Figure 1.5 shows a typical example in which a variety of computational tools are utilized for predicting the disorder segments and disorder-related functional information of the human adiponectin protein.

1.1.3 Intrinsically disordered proteins in physiology and disease

The functionalities of disordered regions involve molecular recognition, protein modification, and molecular assembly.[24] The conformational plasticity of IDPs/IDPRs allows them to perform a wide spectrum of exceptional functions that include the ability to be involved in one-to-many binding, to fold completely or partially upon binding, and to modulate biological functions via posttranslational modifications (PTMs) (Figure 1.6a).[9] Based on their function, IDPs/IDRs can be categorized into six distinct classes that include entropic chains, display sites, chaperones, effectors, assemblers, and scavengers.[24,43] IDPs that fall in the category of entropic chains functions as linkers and spacers due to their high degree of conformational disorder. In the latter case, they bind either transiently as display sites of posttranslational modifications or associate/dissociate rapidly as chaperones, or bind permanently as effectors, assemblers, or scavengers.[43,44] IDPs often function via molecular recognition by short recognition elements (motifs) denoted as a "molecular recognition feature" (MoRF) or "eukaryotic linear motif" (ELM) or "short linear motif" (SLiM).[24,28] These short sequence motifs can target proteins to a specific subcellular location, recruit enzymes by PTMs, and promote the recruitment of binding factors to facilitate complex formation.[24] Not all IDPs undergo folding when coupled to their binding partners; some IDPs remain disordered in the bound state, forming so-called "fuzzy" complexes.[45] Such dynamic binding interactions can have functional consequences, such as enhanced target binding affinity and modulation in allosteric interactions. Furthermore, alternative splicing, that is known for enhancing protein diversity in multicellular eukaryotes were also found to be involved in the creation of the evolutionary patterns of inclusion and exclusion of IDR-encoding exons that can provide insights into the functional role of encoded IDR in protein regulation, function, and

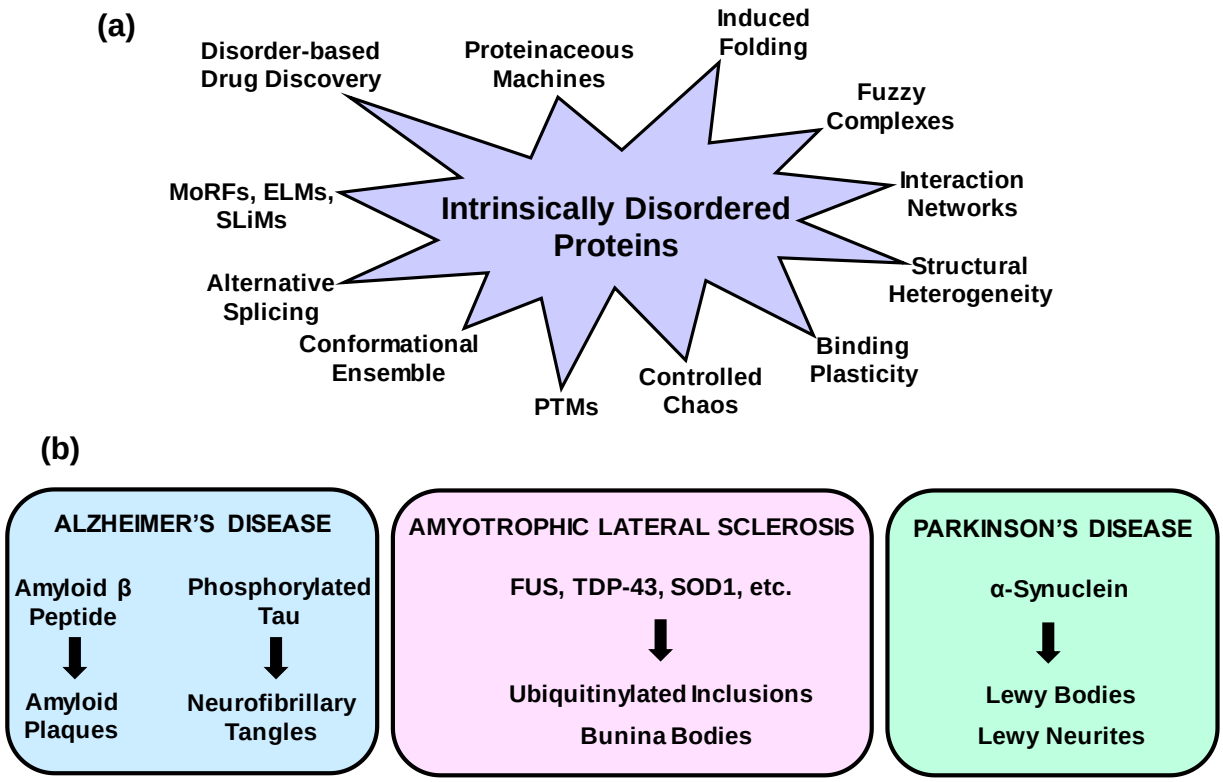


Figure 1.6 (a) Paradigm shift caused by the introduction of the protein intrinsic disorder concept opened a wide array of new directions in protein science. **(b)** Schematic representation of IDPs in human diseases.

interactions.[24,46] Finally, IDPs are also of great interest to engineer new strategies for drug design.[47] Besides playing crucial roles in numerous biological processes, IDPs/IDRs are involved in human diseases as well.[4,23,48] For instance, mutations within IDPs/IDRs that lead to the alteration in the levels of proteins in the cell can increase the aggregation propensity resulting in the formation into insoluble amyloid fibrils that accumulate in a variety of organs and tissues.[48] Moreover, gene fusions and missplicing of proteins containing IDRs have also been linked with cancer.[48,49] The deposition of filamentous protein aggregates is are known to be characteristic hallmarks of several neurodegenerative diseases such as Alzheimer's disease (deposition of misfolded amyloid-β peptide, tau-protein), amyotrophic lateral sclerosis (deposition of aggregates of Fused in Sarcoma (FUS), TAR DNA-binding protein (TDP-43), Superoxide dismutase 1 (SOD1)) and Parkinson's disease (deposition of misfolded α-synuclein).[50] The aggregates of these IDPs are present in the form of amyloid plaques, neurofibrillary tangles, Bunina bodies, Lewy bodies, etc. in brain deposits of the patients (Figure 1.6b).

1.2 Polymer physics description of IDPs

Biologically relevant forms of IDPs possess a large number of distinct conformational states. These conformational ensembles of IDPs are instrumental in regulating numerous *in vivo* functionalities.[9] In their native unfolded state, IDPs can be treated as polymers due to similarities in various properties such as wide conformational distributions, structural heterogeneity, and long-range fluctuations.[51] Long before the routine availability of high-resolution structural techniques such as crystallography and nuclear magnetic resonance (NMR), the theoretical concepts of

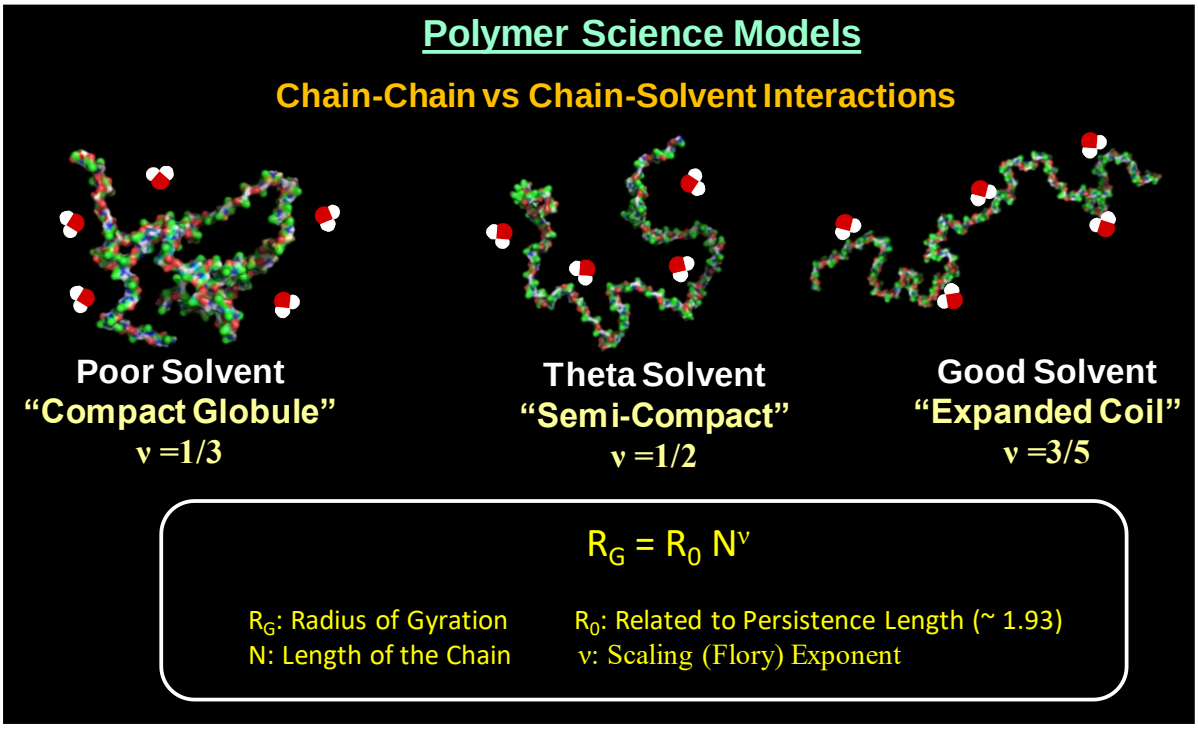


Figure 1.7 Polymer physics models that describe the effect of solvent on polymer surface structure in general, where, ν is an indicator of the solvent quantity. For IDPs, the conformational ensemble most accessible to an IDP is dictated by the solvent quantity.

polymer physics were utilized for the structural investigations of the biopolymers. The language of polymer physics provides a versatile framework for the physical understanding of structural and dynamical behavior of intrinsically disordered proteins.[52] A huge amount of experimental and theoretical work suggests that the dimension of the polypeptide chains (IDPs) is dictated by parameters such as amino acid composition, temperature, and solvent quality. Akin to polymers, IDPs also respond to changes in solvent quality and therefore, can be described by the scaling laws of polymer physics.[52-55] The intricate balance between chain-chain and chain-solvent interactions

determine the conformational as well as phase equilibria for polymers. In solution, the size of polymers has a strong dependence on the solvent quality. According to Flory's polymer theory, the effect of solvent quality on the dimension of polymers can be understood by means of scaling laws given by $R_g = R_0N^{\nu}$ that define the relationship between the length of the polymer (N) and the radius of gyration ($R_{g)}$ under various solvent conditions (Figure 1.7).[55] Also, R_0 is a constant that is related to the persistence length of the polymer, and ν is an exponential scaling factor (Flory exponent) that is a function of solvent quality.[55] For polymer chains in a poor solvent, intrachain interactions dominate over chain-solvent interactions resulting in the formation of compact globules that corresponds to $Rg \sim N^{1/3}$ (Lennard-Jones limit). However, in a good solvent, where chain-solvent interactions are preferred over chain-chain interactions, expanded/swollen coil conformations get populated with $Rg \sim N^{3/5}$ (Excluded volume limit). In a theta solvent, that lies between good solvent and poor solvent, polymer is an ideal Flory's random coil and is characterized by maximum conformational entropy. For theta solvent, $Rg \sim N^{1/2}$ (Flory random coil) because at theta point, the chain-chain and chain-solvent interactions are counterbalanced (Figure 1.7).[53-55] Thus, the scaling exponents are an elegant way of quantifying the degree of expansion or contraction of IDPs due to the fact that they are independent of the length of the individual polypeptide chain and are influenced by the effective solvent quality which in turn in dictated by the differences in amino acid composition and by changes in solution conditions.[52]

1.3 Probing water structure in proteins by solvation dynamics studies

Water performs several crucial tasks in various biological processes.[56,57] One essential role of water in biology is in the protein folding and/or unfolding process.[58] Water molecules form multiple hydrogen bonds that allow to mediate protein-protein interactions. Interestingly, water molecules present in the close vicinity of proteins is different from the bulk water and are termed as bound water or biological water.[59-63] This type of water molecules is known to maintain contact with their nearest neighbors for a longer time, unlike the water molecules in bulk, which makes transient contact. The water molecules that surround a protein surface make up the hydration layer.[61,62] The structure and dynamics of water in the hydration layer is believed to have a strong impact on the functioning of proteins.[57-61,63,64] Additionally, during different condensation processes of proteins such as crystallization and amyloid formation, the restrained water molecules present within the proteins are released into the bulk milieu, causing an entropic gain that would

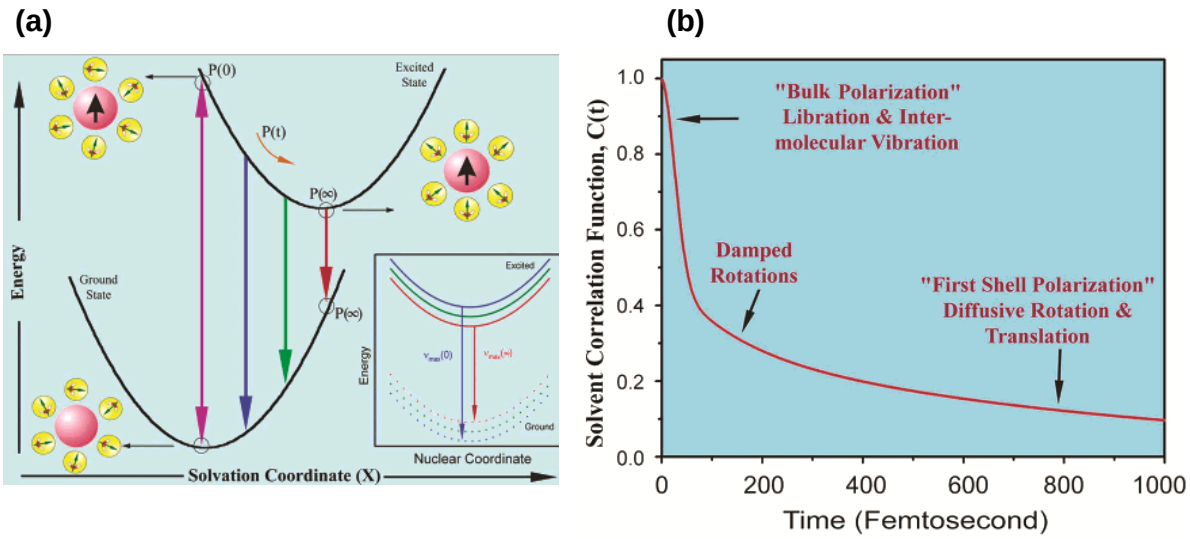


Figure 1.8 (a) Schematic illustration of the potential energy surfaces involved in solvation dynamics, showing the water orientational motions along the solvation coordinate together with instantaneous polarization P. Inset: The change in the potential energy along the intramolecular nuclear coordinate. As solvation proceeds, the energy of the solute decreases, giving rise to a red shift in the fluorescence spectrum. Note the instantaneous P, e.g., P(∞), on the two connected potentials. **(b)** A typical solvation time correlation function for water is shown. Reproduced with permission from Ref. 59.

eventually drive the formation of critical nucleus/seed required for the condensation/polymerization processes.[65] Therefore, in order to understand the behavior of proteins, it is imperative to decipher the structural and dynamical features of the water molecules. Typically, the bound/biological water exhibits slower dynamics compared to the bulk water.[57,59,63] The quantitative characterization of the dynamics remains a challenging task. Several techniques/methods including, NMR based nuclear Overhauser effect (NOE) and dielectric relaxation, have been employed to probe the behavior of water molecules on different length and time scales.[59,66,67] However, these techniques are only sensitive to either time or length scales and not to both. Therefore, the two experimental approaches that are sensitive to both length and time scales are solvation dynamics and inelastic neutron scattering.[59,60] In this thesis, we have chosen a solvation dynamics approach to probe the water molecules and estimate the solvation times in an intrinsically disordered repeat domain of a melanosomal protein, Pmel17, that has been discussed in Chapter 2. In solvation dynamics studies, the dielectric response of proteins is probed by monitoring the time-dependent fluorescence Stokes shift (TDFSS) of the fluorophore that is either part of the protein or is covalently linked to it.[57-60] Femtosecond excitation, at t = 0, leads to the creation of a dipole around the fluorophore, which gives rise to an instantaneous electric field on

the solvent (water) molecules. Due to the interaction of the permanent dipoles of water molecules with the electric field, the free energy minimum of water gets shifted to a non-zero polarization, and the water molecules at t = 0 find themselves in a high energy configuration (Figure 1.8).[59] This non-equilibrated network of water molecules around the fluorophore now has to reorient themselves to reach their new equilibrium positions (Figure 1.8).[59] The quantity measured in solvation dynamics studies is the solvation energy of the fluorescent probe due to interactions with the dipole moment of water molecules. In the linear response approximation, the solvation time correlation function denoted as C (t) is related to the solvation energy by the following relationship:

$$C(t) = \frac{\delta E(0)\delta E(t)}{\delta E2} = \frac{E(t)-E(\infty)}{E(0)-E(\infty)} \qquad (1)$$

where δE(t) is the fluctuation of solvation energy from the average, equilibrium value at time t. It is important to note that equality in the above equation indicates a direct relationship between the average of the fluctuations over the equilibrium distribution (left) and the non-equilibrium function (right). E(∞) is the result of the equilibrium term in the numerator and for normalization in the denominator. In the case of bulk water, rotational motion constitutes hindered rotation (libration), whereas translation would include the intermolecular vibration due to the extensive hydrogen bonding.[60] The initial part of the solvent relaxation is dominated by these two high-frequency specific motions. In the intermediate part, primarily, the moderately damped rotational motions of water molecules contribute.[59] The last part comprises larger amplitude rotational as well as translational motions of solvent molecules in the first solvation shell.[59] The solvation dynamics studies in the case of proteins are rather complex because, in addition to surface and bulk water molecules, it also involves contributions from the amino acid side-chains as well as ions. Typically, in the protein hydration layer, there are two kinds of water, one that is bound to the surface and others which are free. When strongly bound to protein, the water molecules can neither rotate nor translate and therefore, cannot contribute to solvation dynamics.[60,66] But there is a dynamic equilibrium between the free and the bound water molecules and the hydrogen bonding is transient in the first shell (Figure 1.8). A large number of experimental observations from the solvation dynamics of proteins also revealed the biphasic behavior of hydration, thereby suggesting the presence of bound and free water in the surface layer of proteins.

1.4 Specific-ion effects on protein stability

In addition to water, salt ions are another important parameter that influences both the structure and dynamics of proteins.[68-70] Physicochemical properties of the solution such as pH, ionic strength, and presence of cosolutes or cosolvent greatly influence protein stability, aggregation

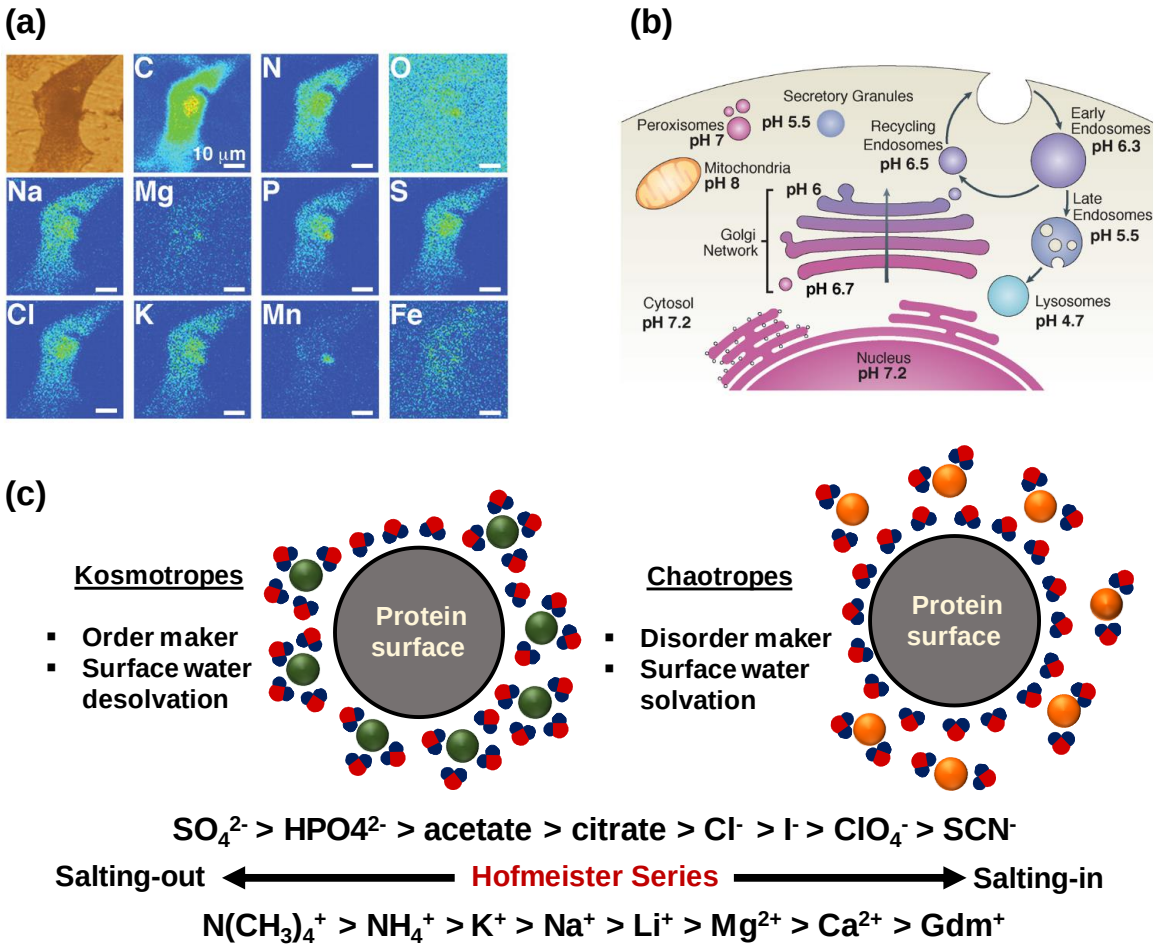


Figure 1.9 (a) Intracellular composition: Localization of various elements in a mammalian PC12 cell (exposed to MnCl2) and **(b)** average pH values of organelles and compartments in mammalian cells. Reproduced with permission from Ref. 71. **(c)** Traditional Hofmeister series and relationship between salts, protein solvation, and thermal stability.

mechanism and are of great interest for materials and medical research (Figure 1.9a, b).[71] Among these factors, the specific ion effects on proteins predate all others. Salt ions have a strong influence on the phase behavior of protein solutions such as crystallization, macroscopic aggregation, and liquid-liquid phase separation (LLPS).[72,73] In the 1870s, Franz Hofmeister, a pharmacologist, performed quantitative studies to elucidate specific ion effects on proteins that revealed varying

levels of relative effectiveness of different ions to salt-out or salt-in proteins.[74,75] Based on these observations, Hofmeister ranked ions according to their ability to precipitate proteins and created the so-called Hofmeister series (Figure 1.9c). In general, ions are believed to greatly influence the stability of proteins either by screening electrostatic interactions (Debye-Hückel effects) or by altering the protein-solvent interactions (Hofmeister effects).[72,76] The ionic strength of the solution has a strong dependence on Debye lengths that, in turn, dictates the interactions between the protein molecules. For instance, when the ionic strength of the solution increases, Debye lengths decrease that cause screening of charge resulting in reduced electrostatic repulsion between proteins, hence promoting intermolecular associations.[72,77] Furthermore, the effect of charge screening on these interactions is also modulated by the pH of the solution, which describes the charge state of the proteins. Salt ions possess the ability to alter the hydrogen bonding network around the surface of aqueous proteins. Based on this, Hofmeister categorized salt ions into two major groups, namely, kosmotropes and chaotropes.[68] The kosmotropes, also termed as 'water structure makers,' are believed to have stabilizing effects on proteins, which give rise to salting-out behavior. In contrast, chaotropes or the 'water structure breakers' are known to destabilize proteins and trigger salting-in behavior (Figure 1.9c).[68] Kosmotropes have the tendency to steal water molecules from the surrounding solute and thereby to organize several layers of water molecules around themselves. In contrast, chaotropes lend water molecules to the surrounding milieu and therefore acting as effective salting-in agents.[78] The hydration state of ions is thought to play a major role in several chemical and biological processes. Ions can be imagined as spheres of distinct sizes having a point charge at the center. The smaller, strongly hydrated ions are kosmotropes; whereas the larger, weakly hydrated ions are chaotropes.[72] Due to differences in water affinity, kosmotropes will preferentially interact with each other and the bulk water, rather than air-water or protein-water interfaces. On the other hand, chaotropes tend to accumulate at these interfaces.[79] Therefore, kosmotropes can be distinguished from chaotropes based on their affinities towards water molecules, which can be assessed from a distance between the point charge and water. IDP/IDRs rapidly exchange between conformations of marginal stability. These transient structures are much more sensitive to solution conditions than ordered proteins. Salts have a profound effect on the propensity of aggregation, amyloid formation, and on the structures of the resulting aggregates or fibrils.[72] Therefore, to elucidate the molecular mechanism of protein aggregation and amyloid formation, it is important to understand the ion-specific and water-

mediated effects on the aggregation behavior of IDPs/IDRs. In this thesis, attempts were made to decode the complex interactions among IDR, water, and ions. The role of Hofmeister ions in modulating the amyloid assembly process of an IDR, namely RPT (the repeat domain of Pmel17), has been discussed in Chapter 4.

1.5 Liquid-liquid phase separation in biology

"Phase transitions are ubiquitous in nature".[80] Like water, protein polymers also undergo liquid phase condensation. The concept of phase transitions of proteins has been well-known to the protein crystallographers for a very long time, as during the crystallization process, various futile phases such as liquid droplets, gels, aggregates, etc. form that cause nuisance in the production of protein crystals.[81] However, most recently, in the field of cell biology, liquid phase condensation is recognized as an emerging paradigm for the cellular organization.[82-84] For instance, in eukaryotic cells, liquid-liquid phase separation (LLPS) seems to play vital roles in the intracellular compartmentalization of various biomacromolecules through the formation of membrane-less organelles (MLOs) or biomolecular condensates.[83,84,85] A growing body of evidence suggests that these condensates are mesoscopic, wet, chaotic, constantly fluctuating supramolecular assemblies of several proteins and/or nucleic acids.[83-88] Such mesoscale intracellular organization of these highly dynamic structures is believed to be the consequence of weak and transient, multivalent noncovalent molecular interactions.[89] The MLOs are known to exchange their components rapidly with the surrounding milieu. Typical examples of these include nucleolus, nuclear pores, paraspeckles, Cajal bodies, nuclear speckles, chromatin, that are present in the nucleus and stress granules, P-bodies, centrosome, nuage (germline P granules), neuronal RNA bodies, in the cytoplasm.[88] Unlike the canonical membrane-bound organelles such as mitochondria, endoplasmic reticulum, Golgi apparatus, lysosomes, vacuoles, etc., these MLOs do not have well-defined boundaries yet remain coherent structures and perform a variety of different cellular functions such as ribosome biogenesis, RNA transcription and processing, regulation of protein translation and so forth.[84,90] An increasing body of work suggests many of these cellular functions are accomplished by the spontaneous liquid-liquid demixing, also known as coacervation, of IDPs and disordered regions of structured proteins that leads to the formation of spherical proteinaceous liquid-like droplets.[88,91-97] A list of proteins that are involved in the intracellular organization is

Table 1.1

Proteins involved in intracellular phase separation

Protein (Gene)	Domain/Motifs	Organelle	Biological function
Fibrillarin (FBL)	RGG box	Nucleolus	Ribosome biogenesis in nucleus
Paraspeckle component 1 (PSPC1)	RRMs, Coil	Paraspeckles	Regulation of gene expression in nucleus
Serine/arginine-rich splicing factor 1 (SRSF1)	RRMs, RS	Nuclear speckles	Regulation of gene expression via storage of splicing factors
Coilin (COIL)	Coiled-coil	Cajal bodies	Regulation of snRNP maturation
Promyelocytic leukemia (PML)	Coiled-coil	PML bodies	Regulation of transcription and protein storage
Probable ATP-dependent RNA helicase DDX4 (DDX4)	FG, RG	Germ granules	Regulation of mRNA translation in the cytoplasm of germ cells
Enhancer of mRNA-decapping protein 3 (EDC3)	LSm, FDF	P bodies	mRNA processing and decay
RNA-binding protein FUS (FUS)	RRM, RGG box, [G/S]Y[G/S]	Stress granules	Storage of translationally stalled mRNA and proteins of the translational machinery

RRMs: RNA recognition motif, LSm: Like Sm proteins

is provided in Table 1.1. Alexander Ivanovich Oparin, a Soviet biochemist, proposed a theory in one of his famous books called "The Origin of Life" according to which life has originated from the phase separation of macromolecules into liquid coacervate droplets, which could have been the first compartments on earth.[98] In line with this theory a burgeoning evidence also reveals certain potential advantages of liquid coacervates or MLOs in regulating the function of the cell via the process of LLPS.[99] Examples of such plausible functional repertoire include inactivation of certain reactions by the selective segregation of the reactants or in other words, partitioning, alteration of the enzymatic reaction rates by regulating the concentration of biomolecules inside the compartments (Figure 1.10), sensing extremely rapid responses since the biophysical response of LLPS is faster than the processes such as the initiation of transcription or translation, also in molding cellular architectures such as membranes due to their viscoelastic properties, etc.[99,100] On

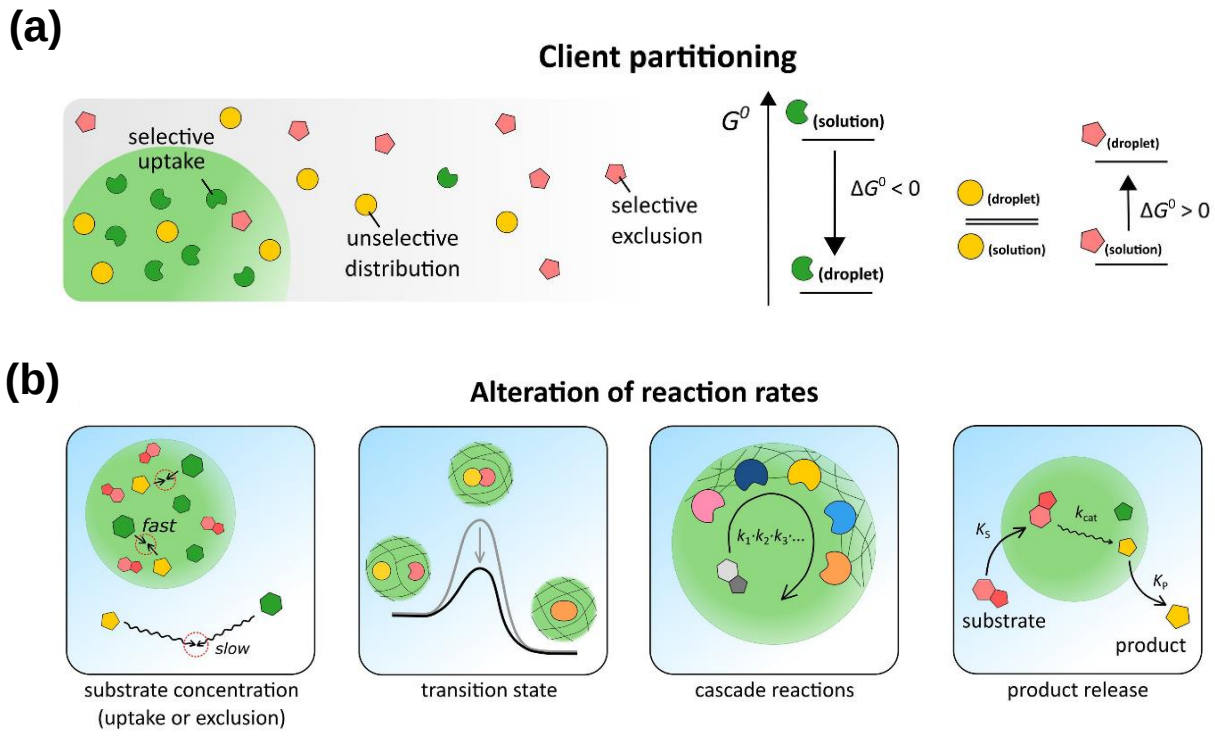


Figure 1.10 (a) Schematic representation of three scenarios for partitioning, reflective of relative free energy levels of the client molecules. **(b)** Plausible effects of coacervate-based compartments on reaction rates. Reproduced with permission from Ref. 100.

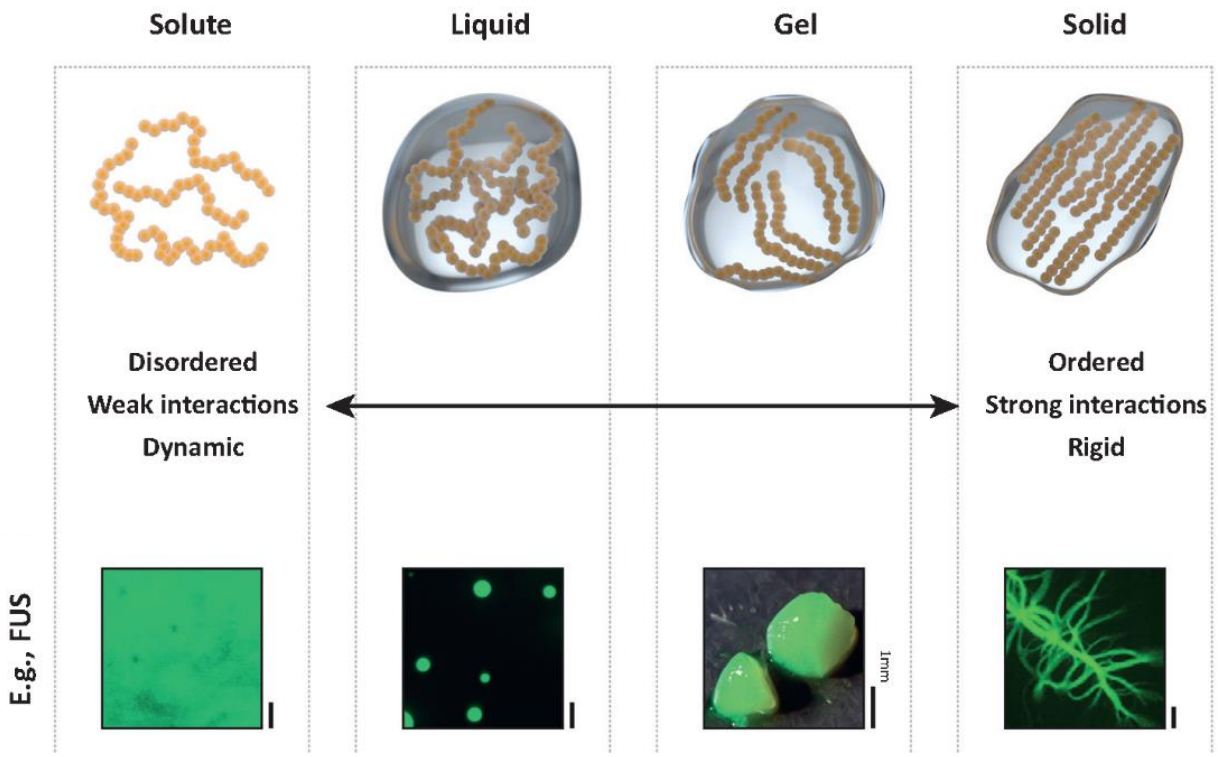


Figure 1.11 Distinct characteristics of phase transitions of a protein. Both material states and dynamics of protein alters and spans a wide range from liquid-like to solid-like states. FUS protein is among one of the phase separating proteins that shows a variety of different material states (liquid droplets, hydrogels, aggregates) in vitro. Reproduced with permission from Ref. 82.

the contrary, a current flurry of research has also indicated that the aberrant liquid-to-solid phase transitions of several IDPs/IDRs are implicated in a variety of debilitating diseases such as familial amyotrophic lateral sclerosis (ALS) and fronto-temporal dementias (FTDs) (Figure 1.11).[101-104] These IDPs/IDRs typically contain low-complexity (LC) domains often comprising repetitive amino acid sequences also termed as intrinsically disordered protein polymers (IDPPs) and prion-like domains (PrLDs) that offer high conformational flexibility.[105,106] Liquid-like dynamic assemblies are proposed to be crucial for functions, whereas, maturation and hardening resulting into less dynamic aberrant states favor formation of protein-trapping cross-linked hydrogels and neurotoxic ordered protein aggregates (Figure 1.11).[107] Table 1.2 shows a list of proteins that are involved in (aberrant) liquid-to-gel and liquid-to-solid phase transitions. However, the underlying molecular mechanisms of the formation of condensates and their roles in physiology as well as disease are currently under scrutiny. One approach that has been advantageous for delineating the

Table 1.2

Proteins involved in aberrant phase transitions

RNA-binding protein	Granule	Association with pathology
Fused in Sarcoma (FUS)	Stress granule	Mutations and inclusions in ALS, FTD, and PQE
TAR DNA-binding protein (TDP-43)	Stress granule, processing body, transport ribonucleoprotein particle	Mutations in ALS, FTD, and inclusions in AD and HD
Ewing sarcoma protein (EWS)	Stress granule	Mutations in ALS, and inclusions in FTD
TATA-binding protein-associated factor 15 (TAF15)	Stress granule, processing body, transport ribonucleoprotein particle	Mutations in ALS, and inclusions in ALS, FTD
Ataxin-2 (ATXN2)	Stress granule	PolyQ expansions in ALS and SCA2
Survival of motor neuron (SMN)	Stress granule, transport ribonucleoprotein particle	Mutations in ALS, SMA
Heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1)	Stress granule	Mutations in ALS, PGD
Heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2B1)	Stress granule	Mutations in ALS, FTD, PGD

ALS: Amyotrophic lateral sclerosis, FTD: Frontotemporal Dementia, PQE: Polyglutamine expansion disease, AD: Alzheimer's disease, HD: Huntington's disease, SCA2: Spinocerebellar ataxia type 2, SMA: Spinal muscular atrophy, PGD: Paget's disease.

compositional characteristics of the condensate/coacervate is their characterization as scaffolds and clients.[99,100] In general, protein polymers that drive the phase separation process are referred to as scaffolds (or hosts), whereas, biomacromolecules that directly interact with scaffolds and preferentially partition into condensates depending on their requirement in the biochemical reaction are known as clients.[100,108] The scaffold proteins can be envisioned as the biological counterparts of associative polymers. 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 interactions between the polymers or protein molecules, whereas, spacer molecules furnish the additional heterogeneous conformational flexibility and mobility so as to maintain the liquid-like nature of the physical cross-links that are formed within the condensates.[108] Several studies have revealed that the scaffold proteins that include IDPs/IDRs form the basic structure of biomolecular condensates.[82] Unlike scaffolds, clients are not crucial players in condensate formation, but if present at high concentrations inside the condensates can affect the phase diagram of the scaffolds.[100,109] Moreover, the reactivity of clients is also affected by the inner environment of the condensates such as crowding, local polarity, and presence of other scaffolds. In natural condensates, the distinction between scaffold and client may be blurred for some molecules and could vary with cellular conditions.[100] Nevertheless, the scaffold-client distinction is useful in considering condensate composition.[109] Therefore, understanding the principles that govern the phase behavior of scaffolds and partitioning of clients is essential to elucidate the functions of biomolecular condensates. Figure 1.10a represents a schematic model demonstrating three distinct cases of client partitioning that includes selective uptake, unselective distribution, and selective exclusion. Further, the difference in the partitioning patterns of clients corresponds to distinct values of the relative standard molar Gibbs free energy (ΔG^0). In the case of complex processes comprising reaction networks, the condensates/coacervates are known to affect the rate of kinetics (k), organize enzymatic cascades, and enhance overall processivity (Figure 1.10b).[100] Hence, deciphering the physical and chemical

principles underlying the intracellular phase separation can provide mechanistic insights into the complex behavior of biomolecular condensates and their associated functions.

1.5.1 Physical chemistry principles of condensate formation

Phase separation of protein polymers is governed by the minimization of the free energy of mixing that is associated with the critical balance between the enthalpy and the entropy terms as described by the Flory−Huggins theory.[110,111] The protein backbones in an aqueous solvent can be envisioned as a chain of dipoles that primarily possess three types of dipolar interactions, namely, protein-solvent, protein-protein, and solvent-solvent.[110] The entropic term comprises degrees of freedom of protein and solvent molecules. In contrast, the pairwise interactions are included in the enthalpic term, and these terms are derived from the Flory−Huggins theory by employing lattice models with mean-field approximations.[55,110] In the lattice model of a protein-solvent system, the volume of the system is defined N_p and N_s, where N_p denotes the number of lattice sites occupied by the polymer and solvent molecules, respectively. The volume fraction of the protein polymer is represented as ϕ_p and that of the solvent as ϕ_s. If ϕ_p is denoted as ϕ then $\phi_s = 1- \phi$, since $\phi_p + \phi_s = 1$. The Flory-Huggins free energy of mixing per lattice site ($F(\phi)$) derived using the mean-field approximation is given by the following relationship:

$$\frac{F(\phi)}{k_B T} = \frac{\phi}{N_p}\ln\phi + \frac{1-\phi}{N_s}\ln(1-\phi) + \chi\phi(1-\phi) \qquad (2)$$

where k_B is the Boltzmann constant, T is the temperature and χ is the Flory interaction parameter, given by Equation (3) that quantifies the competition among the three types of dipolar interactions mentioned above.

$$\chi = \frac{z}{k_B T}\left[u_{ps} - \frac{1}{2}\left(u_{pp} + u_{ss}\right)\right] \qquad (3)$$

where, u_{ps}, u_{pp} and u_{ss} are the mean-field energies per site represents pairwise interaction between protein-solvent, protein-protein, and solvent-solvent, respectively. z is the coordination number of the lattice.[110,111,112]

The demixing of protein polymers into protein-rich and solvent-rich phases is described by the Flory interaction parameter (χ) where, $\chi > 0$ indicates protein polymers in a "poor solvent" and as the polymer concentration is increased in a poor solvent, we enter the two-phase regime.

Whereas, in the one-phase regime, $\chi < 0$ implies polymer chains in a "good solvent".[55,110] This mean-field Flory-Huggins theory has been increasingly utilized as a conceptual framework for understanding the intracellular phase separation. Proteins having low-complexity disordered regions are often are enriched in charged and/or aromatic residues.[106] Several studies have shown the effect of charge patterning on the phase behavior of several IDPs.[112] Therefore, a more generalized approach that takes into account the long-range electrostatic interactions is the Overbeek–Voorn theory, which describes complex coacervation in polyelectrolytes. This theory is an extension of the Flory-Huggins mean-field theory. Here, the mean-field free energy of mixing per lattice site is given by the following relationship:

$$\frac{F(\phi)}{k_B T} = \frac{\phi}{N}\ln\frac{\phi}{2} + (1-\phi)\ln(1-\phi) - \alpha[\sigma\phi]^{3/2} \qquad (4)$$

This above formalism applies to a system that is comprised of polycations and polyanions, each occupying N lattice sites, and ϕ denotes the total volume fraction. In equation (4), the first and second terms represent the mixing entropy akin to equation (2). However, in this case, the third term is an electrostatic free energy term instead of the χ term and is described using a Debye-Hückel theory. Also, α is determined by the charge per lattice site, the partial molar volume of the solvent, thermal energy, and σ denotes the linear charge density per polyion.[110,111,112] Several

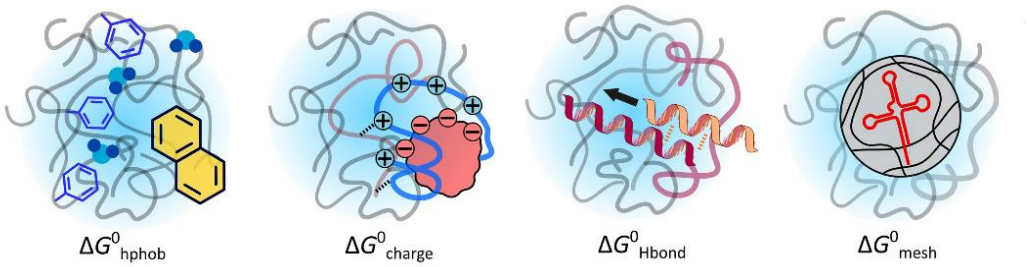


Figure 1.12 Schematic representations of different molecular interactions in the condensate/coacervate. The standard molar Gibbs free energy difference of the solute between the two phases (ΔG^0) is the sum total of individual energies having contributions from different molecular interactions. Reproduced with permission from Ref 100.

forces are responsible for driving the phase separation process, which is expressed in terms of a combinatorial effect of the difference in the molar Gibbs free energy that has contributions from multiple interactions, as shown in Figure 1.12.[100] These specific interactions are discussed in great detail in section 1.4.4. Phase diagrams are the visual representation of phase transitions. Figure

1.13 shows a schematic of phase diagrams under various physicochemical conditions. A simple and symmetric phase diagram generated from a first-order phase transition gives rise to

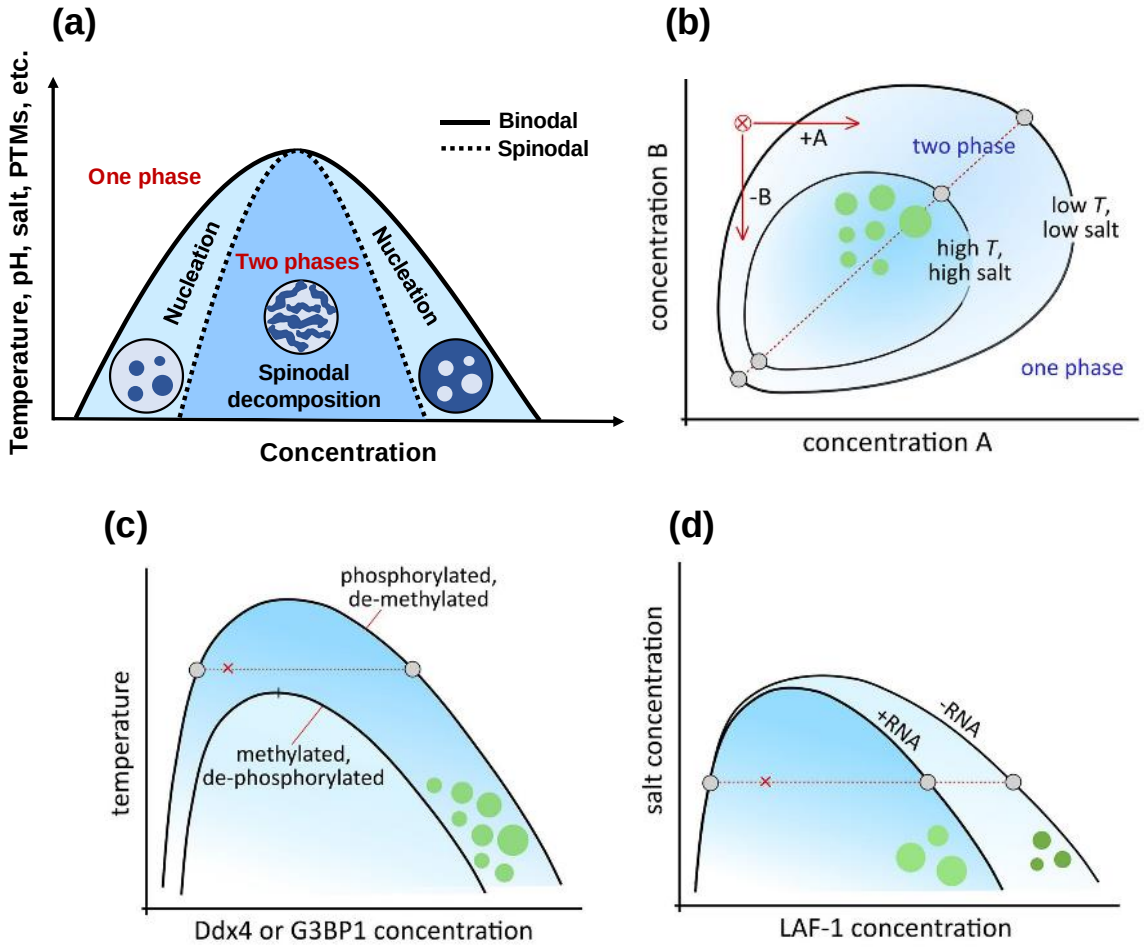


Figure 1.13 Schematic phase diagrams. **(a)** The phase diagram depicted is obtained by systematically changing two conditions (e.g., temperature, concentration, salt, pH) and determining the regions in which the system switches from a one-phase state to a two-phase state. In the two-phase state, a dense phase and light phase stably coexist. The binodal (coexistence line) separates the one-phase state from the two-phase state, and it describes the region where phase separation is thermodynamically favorable. The spinodal is the line that separates nucleation-limited (discrete) phase separation (the region where nucleation takes place) from diffusion-limited (uniform) phase separation (the region where spinodal decomposition takes place). Light blue and dark blue indicate the dilute phase and the dense phase of the phase-separated system, respectively. **(b)** Non-symmetric complex coacervates. Increasing the concentration of A, or reducing the concentration of B can induce condensation. **(c)** Control over coacervation by reversible post-translational modifications, like in Ddx4 and G3BP1. **(d)** Effect of RNA on the density and viscosity of the coacervate phase of LAF-1. Reproduced with permission from Ref 100.

coexistence or binodal line and a spinodal line, as shown in Figure 1.13a.[99] The width of the two-phase region is determined by the physicochemical properties of the solution, such as temperature, pH, salt concentration, and also, by posttranslational modifications (Figure 1.13a). The area between the binodal and the spinodal line is the nucleation and growth region, where the system demixes only when nucleated. Spinodal decomposition in the phase diagram indicates a region of global instability in which the system must undergo phase separation.[99] One of the key features of the phase diagram that plays a prominent role in controlling the condensate formation is the saturation concentration.[99,100] Inside the cells, a large number of IDPs are believed to exist close to their respective saturation concentration, and a small variation in the cellular milieu can alter their properties or shift their phase boundaries.[100] *In vitro* studies also suggest that changes in the concentration of the scaffold, mixing ratio of scaffold and client, and posttranslational modifications, have a great impact on the phase behavior of condensates. Figure 1.13b shows a non-symmetric phase diagram where condensates are formed toward higher salt concentration and temperature by increasing the concentration of A or reducing the concentration of B. Further, posttranslational modifications (PTMs) are also known to affect the condensation propensities by changing the charge distribution of amino acid residues that vary the multivalent interactions. For instance, in vivo studies have shown that a large number of PTMs such as serine, threonine phosphorylation, arginine methylation, and lysine acetylation drastically affect the phase behaviors.[113-115] Figure 1.13c shows the effect of methylation and phosphorylation on the condensates of Ddx4 and G3BP1, where at a given temperature, due to methylation or de-phosphorylation, the saturation concentration increases two-folds which leads to the dissolution of the droplets.[100] Moreover, protein-RNA interactions are also known to modulate the physical properties of the condensates. The addition of RNA to the LAF-1 protein reduces the viscosity of the condensates without altering the saturation concentration that results into dilution of condensates (Figure 1.13d).[96,100] Together, these phase diagrams allow for a more quantitative investigation of condensate formation and also provide insights into the modulators of phase separation that might have therapeutic implications.

1.5.2 Thermoresponsive phase behavior of IDPPs

Intrinsically disordered protein polymers (IDPPs) are a special class of proteins that contain long stretches of repetitive amino acid patterns. IDPPs are known to undergo “stimulus-responsive

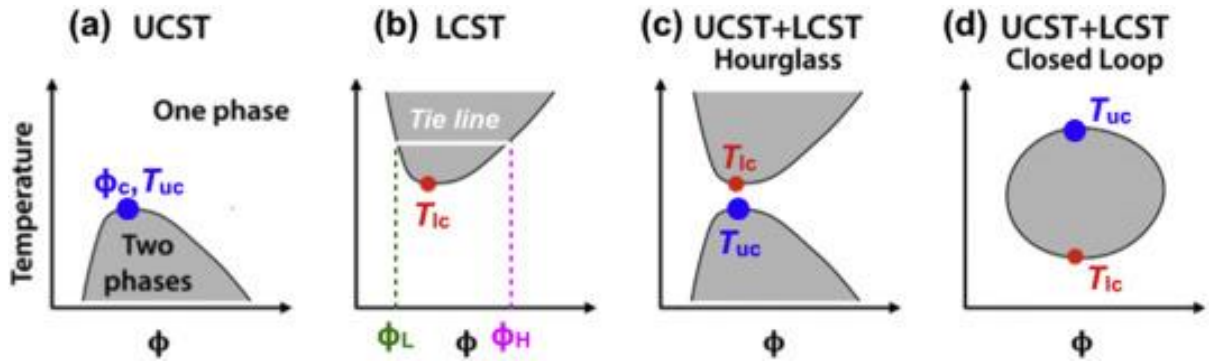


Figure 1.14 Four types of thermoresponsive phase transitions for IDPPs. **(a)** The upper critical solution temperature (UCST) behavior is shown by the coexistence curve (binodal) for systems. **(b)** The lower critical solution temperature (LCST) is shown by the binodal. **(c and d)** UCST and LCST behavior is shown by the binodals for systems. Two phase-regimes are shown by the gray shaded areas, whereas the one-phase regimes are denoted by white areas. The upper critical and lower critical solution temperatures are denoted by T_{uc} and T_{lc}, respectively, below T_{uc} and above T_{lc}, a system separates into two phases. The volume fraction of the IDPP is denoted by ϕ, the volume fraction of the IDPP at the critical temperature is denoted by ϕ_c, and volume fractions of the coexisting dilute and dense phases at a given temperature, T, are denoted by ϕ_L and ϕ_H, respectively. Reproduced with permission from Ref. 116.

phase transitions".[116] Examples of these polymers include elastomeric proteins such as resilins, tropoelastins, spider silk proteins, fibrillin, and titin.[116] In order to understand the stimulus-responsive phase behavior, several studies have been performed by designing polymers that are akin to the IDRs present in the IDPPs.[116,117] In general, the polymers that are insoluble, in other words, undergo demixing upon cooling below a critical temperature exhibit a so-called upper critical solution temperature (UCST) phase behavior, whereas polymers that demix upon heating above a critical temperature have the lower critical solution temperature (LCST).[118] Certain polymers tend to exhibit dual UCST-LCST behavior. The description of each of these thermoresponsive transitions is given below.

UCST behavior: Figure 1.14a represents a typical phase diagram of ϕ (volume fraction) vs. T (temperature) showing UCST transition for a thermoresponsive polymer. For a fixed temperature below T_{uc} (upper critical solution temperature), an increase in ϕ results into a metastable state giving rise to two coexisting phases, namely, mixed (dilute) and demixed (dense).[116,119] The term ϕ_c imply the volume fraction of the system at temperature T_{uc}. An increase in the temperature above T_{uc} mitigates the driving forces for phase separation and leads to the formation of a well-mixed one phase. In Figure 1.14, ϕ_L and ϕ_H denotes the volume fraction of a dilute and a dense

phase, respectively. At $T < T_{uc}$, the width of the two-phase regime is defined as $w_T = (\phi_H - \phi_L)/\phi_c$.[116] For polymers possessing UCST behavior, with an increase in temperature, ϕ_L will increase, whereas ϕ_H will decrease. In general, UCST transitions occur at lower temperatures and involve molecular interactions that increase in strength due to a decrease in the temperature. Examples of such polymers exhibiting UCST phase behavior include disordered proteins having polar LCRs, comprising numerous asparagine, serine, glycine, and glutamine residues.[106] Apart from polar residues, the protein sequences rich in aromatic amino acids such as tyrosine, phenylalanine, and charged residues like arginine, lysine also promote UCST phase transitions. For instance, the LCR of FUS protein contains a lot of tyrosine residues along with a few polar residues and have been shown to undergo a UCST phase transition.[120] Using mutagenesis, recent studies have revealed that aromatic residues are critical for the UCST transition of FUS, whereas aliphatic hydrophobic residues weaken/inhibit the phase separation indicating the dominant role of aromatic character over hydrophobic character. Similar to FUS, several other RNA binding proteins (RBPs) such as hnRNPA1, hnRNPA2, TDP-43, and EWS, are also known to undergo UCST phase transitions.

LCST behavior: Figure 1.14b shows the coexistence curve for the LCST transition of a thermoresponsive polymer. At temperatures below T_{lc} (lower critical solution temperature), polymers exist in a dispersed phase, however for temperatures above T_{lc}, an increase in ϕ results into the saturation of the miscible dispersed phase leading to the formation of demixed or two phases.[116] For polymers possessing LCST behavior, with an increase in temperature, ϕ_L will decrease, whereas ϕ_H will increase. LCST is an entropically-driven transition that involves the release of water molecules from the hydration shell, leading to a gain of entropy of the water molecules.[118] Examples of biopolymers that exhibits LCST phase behavior includes elastin-like polypeptides (ELPs) that are rich in hydrophobic residues comprising Val-Pro-Gly-Xaa-Gly motifs, where Xaa are guest amino acids.[121] In particular, tropoelastins show canonical LCST transition due to the presence of nonpolar and hydrophobic amino acids such as glycine, valine, proline, and leucine. Around 82% of the primary sequence of tropoelastins contains hydrophobic residues.[122] Generally, the driving forces that are responsible for an LCST phase transition are sensitive to the conformational state as well as dynamics. However, irrespective of the expanded or collapsed conformation of the chains in the dense phase, there is an increase in the entropy of the system during phase separation due to the release of water molecules.[106,116] Additionally, the phase separation of ELPs also involves the enthalpic effects that emanate from the stabilization of

the dense protein phases primarily by hydrogen bonding and also by other interactions. Furthermore, several studies have revealed that the transition temperature of ELPs is highly sensitive to the hydrophobic (guest) amino acids and can be controlled by varying the extent of hydrophobicity in the protein sequence.[123] An example of this effect has been recently observed in a yeast heat shock granule-associated protein, namely, poly-A binding protein (Pab1), where the hydrophobicity of the disordered P-domain has been utilized to tune the transition temperature.[116,124]

Dual UCST and LCST behavior: A large number of sequences have been designed to understand the unique dual UCST and LCST behavior of IDPPs.[116,123] Such sequences of IDPPs are designed by incorporating the regions that exhibit UCST transition interspersed by regions that show LCST transition. This type of system gives rise to two types of coexistence curves, in which if the upper critical temperature is less than the lower critical temperature i.e., $T_{uc} < T_{lc}$, then the phase diagram will have an hourglass shape as shown in Figure 1.14c. However, if $T_{lc} < T_{uc}$, then the two coexistence curves will have a closed-loop profile on the phase diagram, as shown in Figure 1.14d. Both tropoelastins and resilins comprise repetitive motifs that are rich in proline (Pro) and glycine (Gly) residues.[117] The regions containing these residues are believed to promote disorder as Gly residues impart flexibility, and Pro residues tend to disrupt the formation of secondary structures and hence, disfavors the formation of highly ordered assemblies such as amyloids. In fact, the composition of these residues is known to dictate the functional properties of the elastomeric proteins. An interesting feature of resilins and tropoelastins is that regardless of having similar content of Pro and Gly residues, they exhibit distinct phase behavior.[117] Resilins have been shown to exhibit a dual UCST and LCST phase behavior, whereas tropoelastins exhibit LCST phase behavior.[116,117] This difference in the phase behavior suggests that apart from these two residues, other residues are also involved in modulating the phase transition properties of these proteins. A careful look at the amino acid composition of the two revealed that resilins contain charged and polar residues, whereas tropoelastins are enriched in nonpolar hydrophobic residues.[117,122] Additionally, resilins have a preference for arginine over lysine, and in contrast, tropoelastins have a preference for lysine over arginine. Taken together, these observations unveil how different amino acids and their compositional variation drive distinct thermoresponsive phase behaviors. Hence, systematic designs of the IDPPs sequence will provide insights into understanding the role

of sequence composition responsible for regulating the thermoresponsive phase transitions of the IDPs involved in physiology and disease.

1.5.3 Physical characterization of condensates

Intrinsically disordered proteins comprising regions of low sequence complexity (LCRs) are strikingly abundant in eukaryotic proteomes. LCRs are known to play a variety of functional roles, including protein subcellular localization, modulation of the activity of several transcriptional factors, protein translation, protein-protein interactions, etc.[125] A large number of prion-like LCRs are also associated with debilitating disorders such as cancer and neurodegeneration.[101] However,

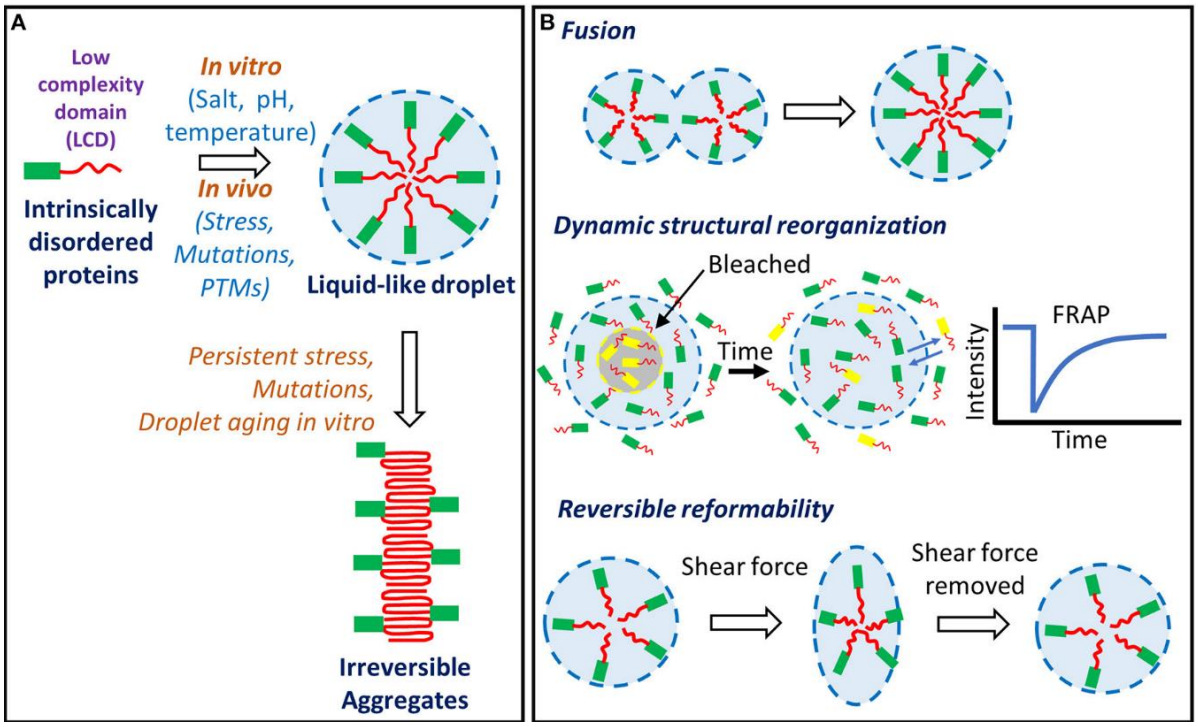


Figure 1.15 A) Low complexity intrinsically disordered proteins undergo liquid-liquid phase separation (liquid-like droplets) and liquid-to-solid transition (irreversible aggregates) under various physicochemical conditions. B) The phase separated droplets demonstrate liquid-like properties such as the ability of the smaller droplets to freely fuse into a larger droplet. Dynamic rearrangement of the internal structural components that can be captured using fluorescence recovery after photobleaching (FRAP); and reversible reformability upon removal of the external shear forces. Reproduced with permission from Ref. 131.

recent studies have shown that these LCRs contain certain sequences having chaperone-like activities that regulate protein phase behavior and inhibit the formation of aberrant phases.[105] The

process of phase separation is extremely sensitive to changes in physicochemical conditions such as temperature, pH, ionic strength, the concentration of proteins as well as nucleic acids (Figure 1.15A). Several PTMs have been shown to act as key regulators of phase separation of multiple IDPs having LC-domains. Certain PTMs are known to inhibit the phase separation process while some favor and promote liquid-to-solid phase transitions (Figure 1.15A).[114,126,127,128] Furthermore, mutations and prolonged stress conditions also lead to the formation of solid-like irreversible pathological aggregates (Figure 1.15A).[129] A variety of methods have been employed to characterize biomolecular condensates and understand their material properties.[86,130] The liquid-like behavior of the dense protein-rich phases of these condensates is reflected in their properties, such as spherical shape, fusion, and deformability upon shear force (Figure 1.15B).[131] One of the most widely utilized technique to assess the mobility, as well as the dynamic internal rearrangements of molecules localized within the biomolecular condensates or phase-separated liquid droplets, is the fluorescence recovery after photobleaching (FRAP) (Figure 1.15B). In general, FRAP assays provide insights into the rate of diffusion of fluorescently labeled molecules present within the condensates/liquid droplets. For performing FRAP, a region of interest (ROI) is photobleached using a high-intensity laser, which is typically the excitation wavelength of the fluorophore. The kinetics of the fluorescence intensity from the photobleached and unbleached molecules is then monitored to estimate the diffusion constant.[130] Also, based on the recovery profiles or kinetics, different material states of the condensates can be determined. In the liquid-like state, condensates possess a rapid fluorescence recovery due to the rapid diffusion of the fluorescent molecules. However, the gel-like state shows very slow and impartial recovery of the fluorescence signal, and the solid-like protein aggregates do not recover after photobleaching.[99]

1.5.4 Molecular interactions underlying protein phase separation

The process of phase separation occurs due to an entropic as well as the energetic cost of concentrating the protein molecules to a confined space.[106,110,111] A symphony of weak and noncovalent interactions plays a crucial role in counterbalancing the entropic and energetic costs. These interactions include cation-π interactions, π-π interactions, electrostatic interactions, transient cross-β contacts, and intermolecular charge-transfer interactions (Figure 1.16 ab).[85] Due to the variation in the composition of the amino acid sequence, different interactions are

responsible for driving the phase separation of different proteins. The examples of such sequence-specific interactions are as follows:

1) *Cation- π interactions*

Cation-π interactions are involved in various physiological processes, such as molecular recognition and catalysis.[132] These interactions also play critical roles in the formation of protein-

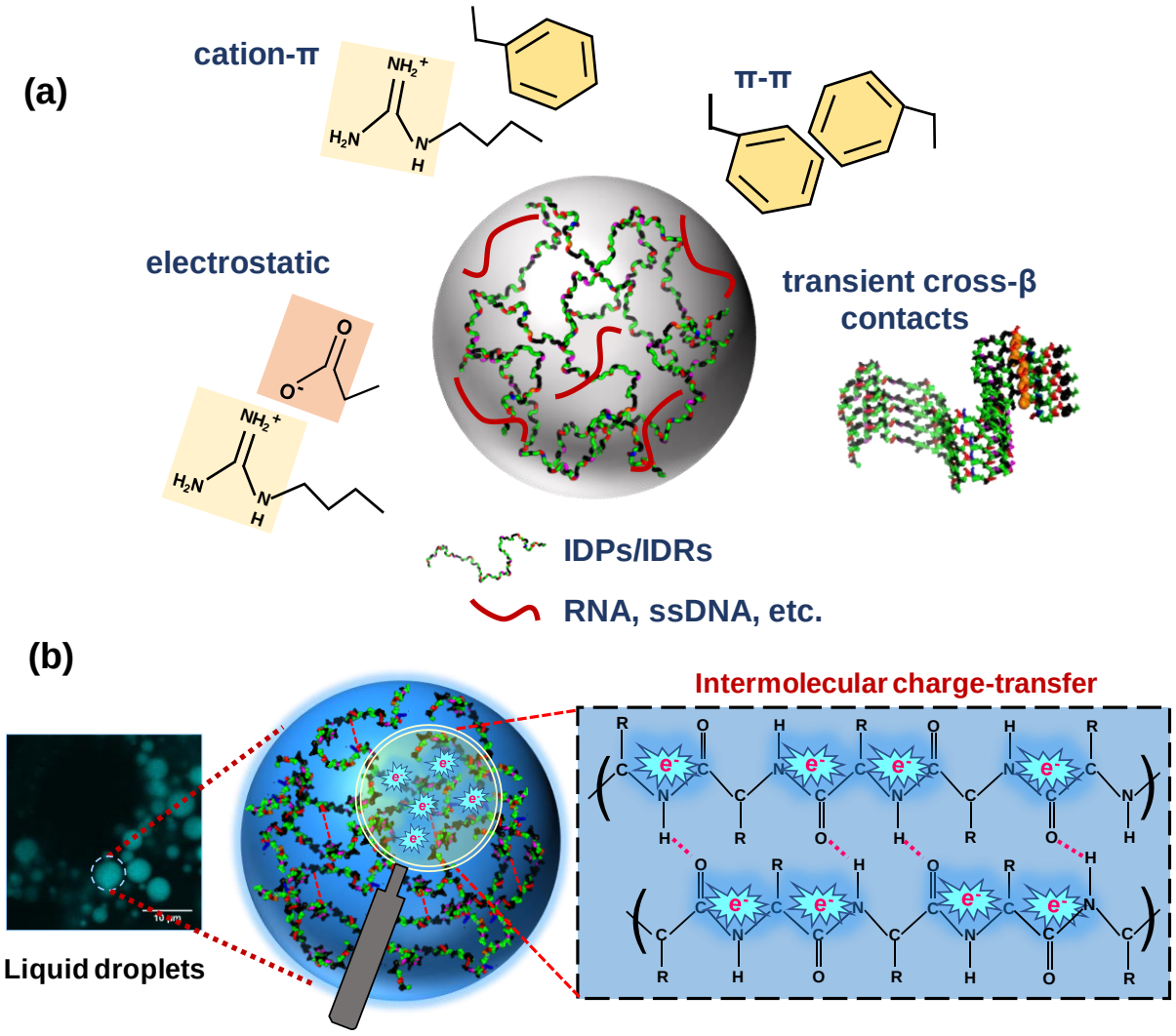


Figure 1.16 (a) Critical interactions that drive LLPS of IDPs/IDRs. The interactions important in LLPS include cation–π, π – π, electrostatic, and transient cross–β-contacts. **(b)** A schematic of intermolecular charge-transfer linked to the delocalization of electrons on an extended hydrogen-bonded backbone. A confocal fluorescence micrograph shows intrinsic blue fluorescence in liquid-like droplets.

protein and protein-DNA complexes. In proteins, these interactions are known to occur between the positively-charged amino acid groups such as arginine, lysine, and the electron-rich aromatic cluster e.g., tyrosine, phenylalanine, tryptophan.[132] Several recent studies have shown the importance of cation-π interactions in the liquid-liquid phase separation of an RNA binding protein, namely FUS and also in the RNA-helicase DDX4.[95,114] In the case of FUS, cation-π interactions between the N-terminal tyrosine-rich intrinsically disordered low complexity domain and structured C-terminal arginine-rich domain is known to drive the phase separation process, whereas, in case of DDX4, cation-π interactions between phenylalanine and arginine residues are thought to play significant roles in stabilization of the condensed droplets.[95,114] Interestingly, the strong short-range cation-π interactions have also been shown to promote the coacervation of two likely charged polymers by counteracting the long-range repulsive electrostatic interactions.[133]

2) *π- π interactions*

π-π interactions are crucial noncovalent interactions in structural biology that are primarily responsible for structural stability in protein and DNA as well as for the formation of recognition *motifs* in proteins and enzymes.[134] Typically, π-π interactions occur between aromatic rings. In case of proteins, these interactions emanate from the delocalized π electrons present in the polypeptide backbone (amide group) and side groups (amide, carboxyl or guanidinium) with π bonds which include tyrosine (Tyr), phenylalanine (Phe), tryptophan (Trp), histidine (His), glutamine (Gln), asparagine (Asn), glutamic acid (Glu), aspartic acid (Asp), and arginine (Arg).[135] In addition to these side groups, other residues such as glycine (Gly), serine (Ser), threonine (Thr), and proline (Pro) have relatively exposed backbone peptide bonds.[135] A flurry of current research uncovered the role of π-π interactions in the liquid phase condensation of several proteins comprising low complexity IDRs such as FUS, TDP-43, EWS, and hnRNPA1.[85,90,131,136] These interactions also govern the phase separation of elastin-like polypeptides that contain Val-Pro-Gly-Xaa-Gly repeats, rich in Pro and Gly residues, and lack the sidechain π groups.[137] Furthermore, in the nuclear-pore complex, the π-π interactions between phenylalanine residues in the FG repeats of nucleoporins results into the formation of a gel-like states.[138]

3) *Electrostatic interactions*

Electrostatic or charge-charge interactions play key roles during folding, binding, and liquid phase condensation of proteins.[139] These interactions are now recognized as the predominant drivers for

the phase separation of IDPs. Clusters of opposite charges within the same protein or two different proteins often promote phase separation. It has been shown that several IDPs/IDRs enriched in positively charged residues undergo LLPS in the presence of RNA, and this complex coacervation is believed to occur through charge neutralization.[85,87] Typically, RNA reduces the saturation concentration of proteins (concentration above which the protein phase separates); however, excessive RNA has been shown to inhibit the formation of condensates.[82,99,140] Also, several *in vitro* and *in vivo* studies have revealed that both length and the type of RNA modulate the phase separation properties of proteins.[141] Additionally, the concentration of RNA inside the condensates is known to determine the protein dynamics that are responsible for controlling distinct material states of proteins.[141] An emerging body of evidence suggests that charge patterning rather than just the presence of charged residues is crucial for the LLPS process.[106]

4) *Transient cross-β contacts*

The prion-like low-complexity domains (PrLCDs) are known to govern the phase transitions of several RBPs.[85,105] For instance, LCD in the RBP FUS, phase separates *in vitro* as well as in cells into liquid and gel phases.[101] FUS LCD is enriched in polar and aromatic residues such as serine, glycine, glutamine, and tyrosine. This domain is involved in the formation of cytoplasmic aggregates that are associated with neurodegenerative diseases. A recent study revealed that the condensed phase of FUS LCD is stabilized by transient hydrogen bonds formed by the glutamine and tyrosine residues.[142] Interestingly, fibrils formed by the short fragments of PrLCDs of RBPs that undergo LLPS are shown to be structurally different from the typical amyloid fibrils.[143] Unlike the classical amyloid fibrils that form highly stable cross-β sheets steric zippers, the fibrils formed by PrLCDs have kinked cross-β sheets that are known to be thermodynamically less stable. These kinked cross-β sheets formed by PrLCDs are termed as low-complexity aromatic-rich kinked segments (LARKS).[85,143] These findings suggest that weak and transient cross-β contacts maintain the fluidic condensed phases of PrLCDs, whereas strong and persistent cross-β contacts lead to the formation of solid-like phases or pathological aggregates.

5) *Intermolecular charge-transfer*

Generally, the aromatic amino acid residues, such as tryptophan and tyrosine, are responsible for the intrinsic fluorescence in proteins. However, several recent studies have identified novel charge transfer transitions that give rise to intrinsic fluorescence, which is observed in the near UV-visible

region of the electromagnetic spectrum.[144,145] These unusual low energy transitions are often observed in highly charged proteins. For instance, these transitions have been found to occur in the monomeric native state of a protein, namely α_3C, due to the presence of a large number of lysine and glutamic acid residues.[145] Similar to the charged proteins, recent reports have shown the intrinsic blue fluorescence emerging from protein crystals as well as fibril-forming peptides.[144] In another study, charge transport through the hydrogen-bonded networks of amyloid fibrils of the elastin-like polypeptide has been proposed to be responsible for the intrinsic blue fluorescence.[146] Furthermore, an octapeptide, GVGVAGVG, has been shown to exhibit blue autofluorescence in

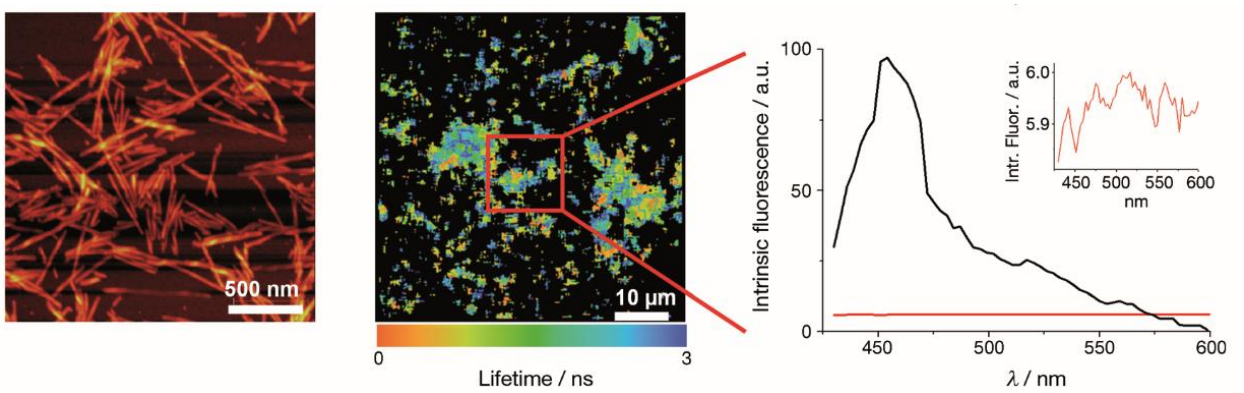


Figure 1.17 Intrinsic fluorescence in the visible range detected from amyloid fibrils formed from human α-synuclein in vitro. AFM images of fibrils formed along with the intrinsic fluorescence lifetime images of the same amyloid fibrils obtained by confocal microscopy. Black traces show the emission spectra of the intrinsic fluorescence from the aggregates, and the red traces show a fluorescence signal from the corresponding monomeric protein. Reproduced with permission from ref 149.

the cross-β-sheet rich amyloid state due to extensive delocalization of peptide bond electrons through the hydrogen bond networks of amyloid structure.[147] Interestingly, proton transfer along with charge delocalization between the N- and C-termini of hydrogen bond-rich protein structures have been associated with the structure-specific visible fluorescence.[148] Another example of this phenomenon is observed in Parkinson's disease-associated protein, namely human α-synuclein. In its monomeric form α-synuclein, which shows negligible charge-transfer fluorescence, whereas, in the aggregated state, it exhibits huge intrinsic fluorescence, as shown in Figure 1.17.[149] The phenomenon of charge-transfer fluorescence has been shown to be extremely sensitive to various physicochemical conditions such as pH, salt concentration, and temperature. Although a wealth of information is available on intrinsic blue fluorescence from an amyloid perspective, the role of charge transfer fluorescence in protein condensates remains elusive. Recently, we have discovered

that intermolecular charge-transfer fluorescence is an important molecular driver during the liquid-liquid phase separation and liquid-to-solid maturation of a pH-responsive intrinsically disordered domain of a melanosomal protein, Pmel17 (Figure 1.16b).[150] The details of this study are described in Chapter 5 of this thesis.

1.6 Amyloid structure: Conformational diversity and consequences

The folding of proteins into their native conformations is tightly assisted by highly orchestrated chaperone machinery.[3,21,22] However, under certain circumstances, proteins escape the cellular quality control machinery and form partially folded/unfolded intermediates.[21] The accumulation of these intermediates results in the formation of intractable aggregates that are either amorphous or thermodynamically stable structures called amyloids.[151-156] Deposition of amyloids are known to be characteristic hallmarks of several neurodegenerative disorders.[4,155-157] Amyloids are composed of misfolded proteins that are assembled into a highly ordered cross-β sheet rich architecture in which the β strands are aligned perpendicular to the fibril axis. This cross-β sheet rich architecture was observed by X-ray diffraction pattern, where amyloid fibrils show two major reflections arising from H-bonding distances between two β strands in each sheet that is 4.7 Å and between two β- sheets is 6-11 Å (Figure 1.18a).[153-159] After this observation, multiple structural traits have been characterized as components of the common amyloid fold. Amyloid fibrils are often composed of protofilaments (typically 2- 6 fibrillar subunits) that are entwined around each other with regular periodicity imparting a twisted ultrastructural appearance (Figure 1.18b).[154] The protofilament's structure comprises multiple β-sheets that can be arranged either in an antiparallel or parallel fashion (Figure 1.18b). In an antiparallel arrangement, the two β strands are present in alternate positions within the sheet, resulting in a distance greater than 4.7 Å, whereas, in case of parallel in-register sheet arrangement, the β strands are directly adjacent to each other giving rise to a distance of ~4.7 Å. In a β-solenoid, a single monomer forms two layers of the fiber structure resulting in a distance of ~9.4 Å between the strands.[154] The β-solenoid can be further categorized into two types (Figure 1.18b). The first one is the β-helix, where three β-strands often form a triangular layer and loop back to form an additional layer. The second is the β-roll, where a single layer is composed of two β-strands and resembles the β-sandwich but loops back to form a second identical layer.[154] Furthermore, by performing dilution experiments, one can distinguish the intramolecular distances from intermolecular distances since dilution increases intermolecular

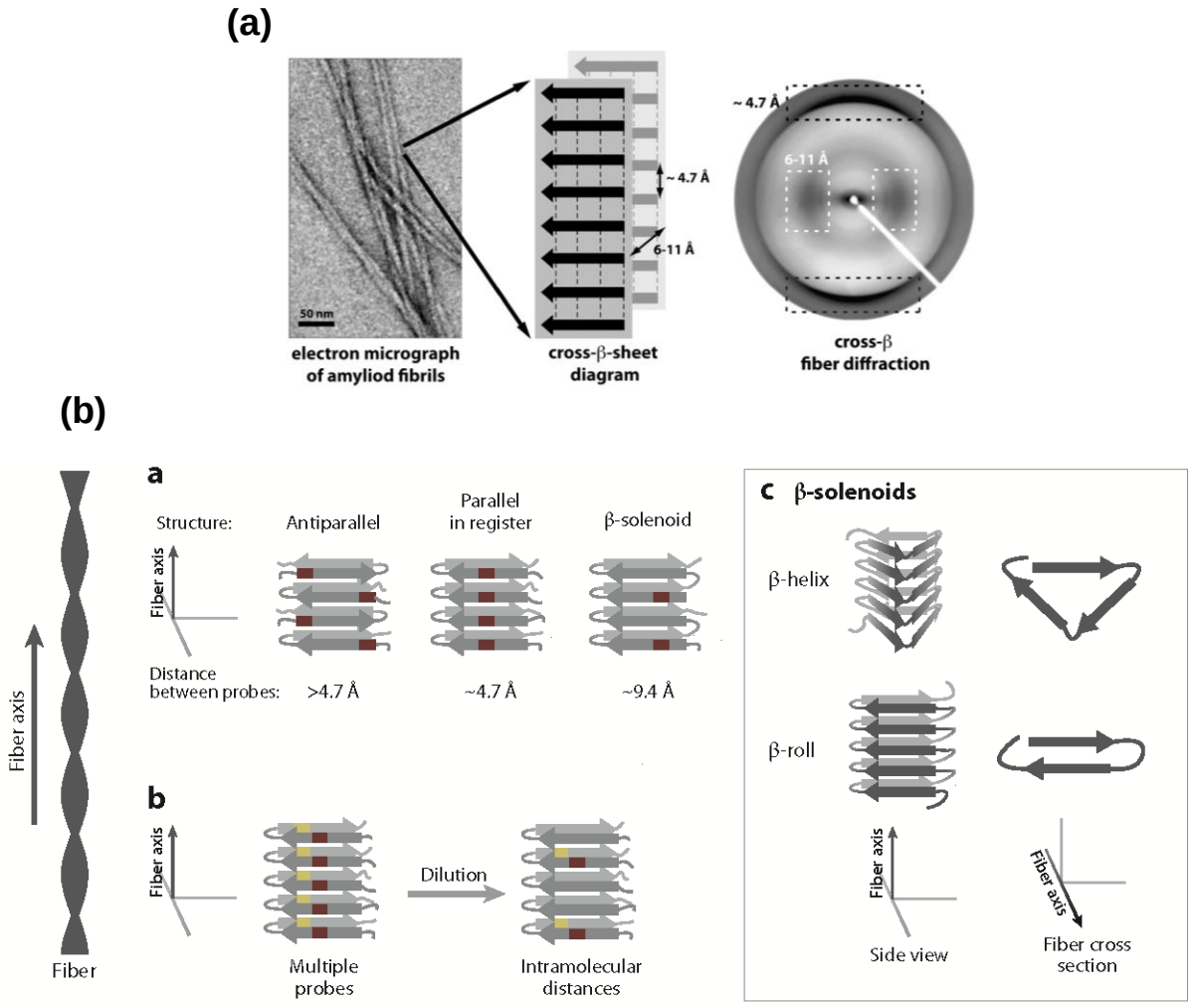


Figure 1.18 (a) Negatively stained transmission electron micrographs are used to show amyloid fibrils and its long filaments. The repetitive spacings that give rise to the typical fiber diffraction pattern with a meridional reflection at 4.7 Å (black dashed box) and an equatorial reflection at 6–11 Å (white dashed box) are represented by the schematic diagram of the cross-β sheets in a fibril, with the backbone hydrogen bonds represented by dashed lines. Reproduced with permission from ref 153. **(b)** Classification of different amyloid folds. Amyloid structures have largely fallen into three broad categories: antiparallel arrangement (left), parallel in-register arrangement (middle), and β-solenoid (right). All of which can be distinguished from each other by the introduction of a single probe (red rectangle) in identical positions for every monomer and by measuring the distance between probes. In systems in which two probes have been introduced (yellow and red rectangles), intramolecular distances can be distinguished from intermolecular distances by diluting labeled protein with unlabeled protein. Types of β-solenoids are also depicted schematically. Reproduced with permission from Ref. 154.

Table 1.3

Disease-related Amyloids

Protein/peptide	Structural information	Related pathology	Affected tissues
Amylin	Intrinsically disordered	Type II diabetes	Pancreas
α-synuclein	Intrinsically disordered	Parkinson's disease	Brain
Huntingtin fragments	Mostly intrinsically disordered	Huntington's disease	Brain
Amyloid-β peptide and Tau	Intrinsically disordered	Alzheimer's disease	Brain
Prion protein or its fragments	Intrinsically disordered and α-helical	Spongiform encephalopathies	Brain
Superoxide dismutase	β-sheet and Ig-like	Amyotrophic lateral sclerosis	Brain and spinal cord
Apolipoprotein A1 fragments	Intrinsically disordered	Apolipoprotein A1 amyloidosis	Eyes
β2-microglobulin	β-sheet and Ig-like	Haemodialysis-related amyloidosis	Musculoskeletal tissues, GI tract, heart
Transthyretin mutants	β-sheet	Familial amyloidotic polyneuropathy	Peripheral nervous system
Calcitonin	Intrinsically disordered	Medullary carcinoma of the thyroid	Thyroid

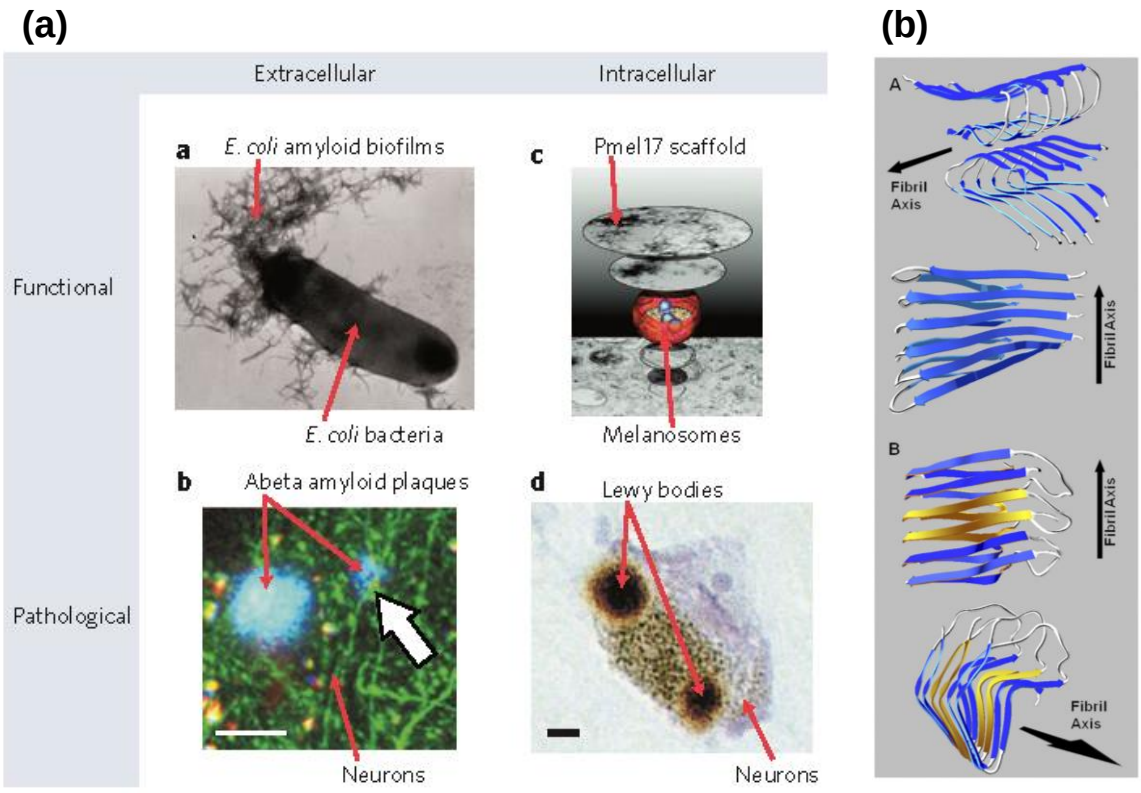


Figure 1.19 (a) Classification of amyloids. Amyloids can be extracellular or intracellular, and functional or pathological. a: Functional amyloid in biofilms produced by bacterial species (*E. coli*). b: Amyloid plaques, as seen in a mouse model of Alzheimer's disease. c: Transmission electron microscope image of Pmel17 scaffolds in melanosomes involved in the biosynthesis of melanin. d: Lewy bodies, pathological protein aggregates that develop in neurons in Parkinson's disease. Reproduced with permission from ref 164. **(b)** Structures of pathological and functional amyloids. A: schematic representation of a parallel in-register β-sheet structure of $A\beta_{1-40}$. The fibril is composed of two protofibrils (a single protofibril is shown in the lower half of A), which are each composed of stacked $A\beta_{1-40}$ peptides in-register with the preceding and following peptides. B: schematic representation of HET-s β-helical amyloid (residues 223–283 from Protein Data Bank code 2KJ3). The monomers alternate between blue and yellow, revealing that each polypeptide provides two β-strands that encircle the long axis of the fiber. Reproduced with permission from Ref. 165.

distances while preserving intramolecular distances (Figure 1.18b).[154] Amyloids are often nonbranched fibrous structures, typically 5-15 nm in width and several micrometers in length.[153,154,160] In addition, amyloid can be distinguished from other disordered aggregates by various properties, including insolubility in ionic detergent, protease resistance, and recognition by several dyes such as Congo red (CR) and thioflavin T (ThT).[161] Additionally, the information

Table 1.4

Functional Amyloids

Polypeptide	Organism	Function of the resulting amyloids
Curlin	*Escherichia coli* (bacterium)	Involved in biofilm formation and host invasion
Chaplins	*Streptomyces coelicolor* (bacterium)	Regulate surface tension of water and promote the development of aerial hyphae
HET-s (prion)	*Podospora anserine* (fungus)	Stimulate a complex programmed cell death phenomenon
Sup35p (prion)	*Saccharomyces cerevisiae* (yeast)	Promote read-through of stop codon which give rise to new phenotypes
Neuron-specific isoform of CPEB (prion)	*Aplisia californica* (marine snail)	Facilitate long-term maintenance of synaptic changes linked with memory storage
Spidroin	*Nephila edulis* (spider)	Helps in the formation of silk fibers of the spider web
Proteins of the chorion of the eggshell	*Bombyx mori* (silkworm)	Protect the oocyte and the developing embryo from environmental hazards
Semenogelin1 and 2; SEVI 1 and 2	*Homo sapiens*	Sperm selection, clearance, and antimicrobial activity
Intralumenal domain of Pmel17	*Homo sapiens*	Form fibrous striations that templates melanin biogenesis

obtained from atomic-level structural studies on microcrystals of short amyloidogenic peptides or segments of amyloid-forming proteins revealed the presence of two different kinds of β-sheet stacking interfaces that were termed as dry- and wet interface, where, the dry interface, also known as 'dry steric zipper', involves complementary side-chain interactions and wet interface is mostly of inter-strand hydrogen bonds.[153] Recently, methods like solid-state NMR and cryo-electron microscopy have further provided greater insights into characterizing the fibril structure from a number of proteins.[162] A growing body of evidence from neuropathological, cellular, and biochemical studies have shown the involvement of several proteins and peptides in triggering a plethora of pathological abnormalities that are responsible for human disease.[163] A list of some of these amyloid-related human diseases is given in Table 1.3. In contrast to the disease-associated amyloids, several amyloids are also utilized by nature in performing an array of physiological functions. This class of amyloids is termed as functional amyloids.[160] The manifestations of these are observed in a wide variety of organisms ranging from bacteria to humans. Two examples of these functional amyloids are curli in *Escherichia coli* that is important for biofilm formation and the melanosomal protein, Pmel17 in humans responsible for melanin biogenesis, which is important for skin and eye pigmentation (Figure 1.19a).[164] In this thesis, we have studied the molecular mechanism of amyloid formation of Pmel17. A representative list of functional amyloids, along with their respective biological involvement, is provided in Table 1.4. Figure 1.19b shows the structure of pathological as well as functional amyloid. In the case of pathological amyloid, the Aβ (1–40) fibril is composed of two protofibrils that are arranged in a parallel in-register β-sheet structure. In contrast, the functional amyloid HET-s forms the β-helical structure.[165] The fundamental question that arises is why functional amyloids are benign while their disease-related counterparts are harmful. The potential reason could be that the self-assembly process of functional amyloids takes place at exceedingly rapid rates and is efficiently regulated by the cellular machinery that might circumvent the production of cytotoxic prefibrillar oligomers. Furthermore, in the case of certain functional amyloids, the assembly reactions occur within the membrane-bound compartments, thereby preventing unwanted interactions with other cellular components. Since functional amyloids avoid damaging their cellular environment, deciphering their structures and elucidating mechanisms of their assembly would provide cues to mitigate the toxicity of their pathological counterparts.

1.6.1 Polymorphism of amyloid fibrils

Polymorphism is one of the remarkable features of amyloid fibril structures.[153,166-168] Polymorphs are the consequence of a plethora of possible 3D arrangements acquired by protofilaments that results in distinct modes of alternative packing of polypeptide chains.[153] Structural polymorphism occurs due to the difference in the number of protofilaments forming amyloid fibrils from a particular polypeptide chain. In addition, variation in the orientations and conformations of the protofilaments can further contribute to the enrichment of polymorphism.[166] Interestingly, the manifestations of polymorphism can be observed in a wide range of fibrils that are either formed under *in vitro* conditions or isolated from the patients. A spectrum of fibril morphologies is linked

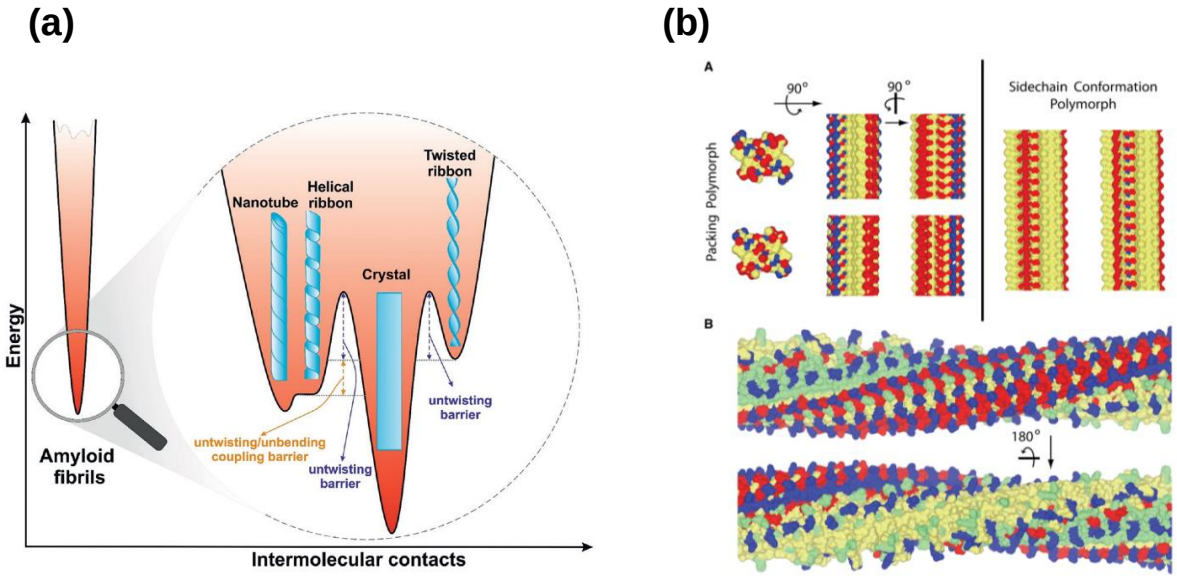


Figure 1.20 (a) Suggested energy landscape for the main amyloid polymorphs: twisted ribbons, helical ribbons, nanotubes, and crystals. Amyloid crystals occupy the ground state in the landscape, whereas all other polymorphs are metastable and occupy relative minima. Reproduced with permission from ref 168. **(b)** Amyloids Polymorphism (A) Packing polymorphism (left panel) and side-chain polymorphism (right panel). The packing of the polymorphic structures of the peptide NVGSNTY (PDB 3FTK top and 3FTL bottom) with three orthogonal views is compared by the left panel. A single view of the peptide LVEALYL (PDB 3HYD) in the true crystal structure conformation on the left and with the Glu side chain rotated to its most favored conformation while maintaining its hydrogen bond network on the right is shown by the right panel. (B) The two broken panels show the HET-s PFD surface showing a complete 360° twist of the fibrils. The amino acid residues are colored blue and red for charged, green for polar and yellow for nonpolar. Reproduced with permission from Ref. 153.

with different variants of amyloid disorders *in vivo*. Therefore, deciphering their specific structural features remains challenging. However, *in vitro* recapitulation of these morphologies might provide structural information of their *in vivo* counterparts. The three-dimensional packing of the protofilaments determines the final morphology of the amyloid structure.[156] A schematic of the energy landscape of amyloid polymorphs is shown in Figure 1.20a. These include twisted ribbon, helical ribbon, nanotube, and crystal morphologies. Among these, amyloid crystals occupy the absolute minimum, and all other polymorphs possess a metastable state.[168] It turns out that the twisted ribbon polymorph is the most generic morphology of amyloid fibrils. This polymorph can further undergo two possible polymorphic transitions.[168] In the first transition, it transforms into a helical ribbon through the lateral addition of protofilaments resulting in increased growth in width. The helical ribbon polymorphs can then further reduce its free energy by closing into nanotubes. In the second transition, the twisted ribbon untwists through thermal fluctuations and enters into the deepest point in the energetic funnel of the amyloid crystal polymorph (Figure 1.20a).[168] Furthermore, several studies have indicated that the pathological amyloids are often polymorphic, and in contrast, the functional amyloid structures like HET-s are isomorphic (Figure 1.20b), albeit structural data on functional amyloids is limited.[153] Akin to the energy landscape of IDPs, local minima in this energy landscape of amyloids represent different polymorphs that are formed due to subtle variations in the aggregation conditions. Moreover, polymorphs are classified into different types, such as segmental, packing, side chain, and assembly.[153] Figure 1.20b shows packing and segmental polymorphs that emanates from the distinct steric zipper packings in the amyloid peptide crystal structures. In the case of packing polymorphs, the nature and position of the side chains present on the outer side of the fibrils differ. Therefore, the sequence involved in the formation of the cross-β core remains the same, but packing differs. However, segmental polymorphs differ in terms of the two or more segments from the same amyloid-forming protein that gives rise to different morphologies due to the involvement of distinct sequence composition in the formation of the cross-β spine, and this polymorphism has been observed for Pmel17 fibrils.[153,167] Another example of a polymorph is the side-chain conformers in which the side chain conformations are distinct. Furthermore, the tight binding between protofilaments via the repeat-induced cooperativity gives rise to diverse supramolecular structures that are termed as assembly polymorphs. In addition to the energetic description of polymorphism, several techniques, including electron microscopy (EM), atomic force microscopy (AFM), and NMR provide insights

into the structural origins of polymorphism. For instance, both EM and AFM images can differentiate polymorphs based on the degree of twisting and the diameter or mass per length of the fibrils. Similarly, the solid-state NMR spectra show signal with multiple NMR signals per atom or very broad resonance lines indicating structural heterogeneity due to polymorphisms. The phenomenon of amyloid polymorphism is believed to play a crucial role in the biology prion propagation and the strain phenomenon.[169] In this thesis, we have utilized AFM to investigate polymorphism in a functional amyloid, which has been discussed in chapters 3 and 4.

1.6.2 Molecular mechanism of amyloid formation

Elucidating the molecular mechanisms of the conversion of proteins from its soluble functional state to the amyloid state is an essential prerequisite for deciphering and controlling the onset and progression of human disorders.[155] During the aggregation/fibrillation process, multiple precursor species/intermediates get populated, and these intermediates are believed to play key roles in the amyloid assembly. Upon aggregation or amyloid formation, a structural transformation occurs that leads to the formation of β-sheet rich architecture, irrespective of the initial fold of the protein.[170,171] There is an emerging consensus that these intermediate species might perhaps be more cytotoxic than the mature amyloid fibrils.[172] Therefore, in vitro, studies of amyloid formation are critical for understanding the role of these intermediates. Both polymerization and aggregation have been used to describe the mechanism of amyloid formation.[173] However, polymerization is more specific, while aggregation implies a non-specific process. As amyloid formation involves both specific and non-specific interactions, therefore, these terms can be interchangeably used to describe the amyloid assembly process.[173] Typically, two models have been proposed to describe the mechanism of amyloid formation, namely, nucleation-dependent polymerization (NDP) and isodesmic (linear) polymerization (Figure 1.21a). In the nucleation-dependent polymerization model, monomeric (native) proteins exist in a dynamic equilibrium with their partially folded/unfolded states until a thermodynamically unfavorable oligomeric species, called the critical nucleus or seed, is formed.[174] The kinetic phase of NDP, where these highly dynamic, transient, and heterogeneous oligomeric structures are formed, is termed as the nucleation/lag phase (Figure 1.21a). The length of the lag time is governed by the probability of the formation of a critical nucleus. The oligomeric species formed in the nucleation phase further associate to produce higher-order species in the growth/elongation phase, which is characterized by the rapid

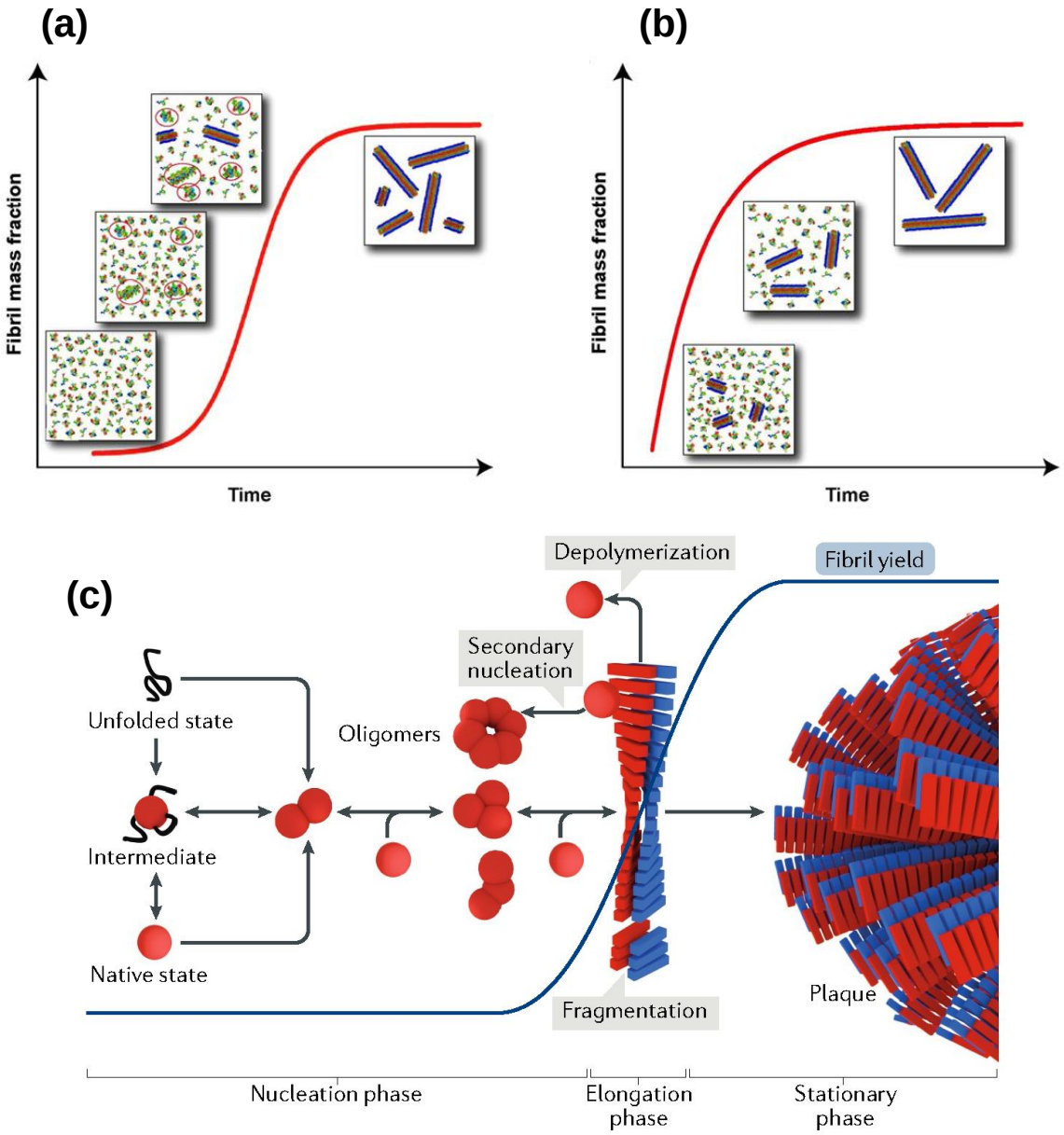


Figure 1.21 Mechanism of formation of amyloid aggregates. Examples of the kinetics of formation of amyloid aggregates, showing hypothetical snapshots of the ensemble of protein species present at different phases of the aggregation process. **(a)** Formation of aggregates largely by primary nucleation and elongation, **(b)** Formation of aggregates by seeding processes that bypass the primary nucleation step. Reproduced with permission from ref 176. **(c)** Aggregation precursors (native, partially folded or unfolded) form oligomeric species which then assemble further to form higher-order oligomers or a fibril nucleus. This phase is known as the nucleation phase of assembly. Further onto the fibril nucleus, monomers rapidly recruit in the elongation phase and enters into the stationary phase, yielding amyloid fibrils. Two important processes take place during the elongation phase are secondary nucleation and fragmentation. Reproduced from Ref. 156.

polymerization into amyloid fibrils. The rate of growth of the fibrils becomes slower once equilibrium is reached between the monomeric and fibrillar state. The growth rate eventually ceases indicating the onset of the stationary phase.[175] A typical NDP reaction has the following characteristics: (i) The kinetics of aggregation comprises a lag phase in which the dissociation rate constant of monomers is greater than the association rate constant. Also, the duration of the lag time is dependent on the size of the critical nucleus as well as the concentration of the monomers (protein concentration). The lag time gets shortened with an increase in the protein concentration (ii) A critical protein concentration is associated with polymer formation which is defined as the lowest concentration of monomeric protein, below which the polymerization would not take place. This concentration varies depending on the type of proteins and the amino acid composition. It is typically estimated from a plot of the rate/amount of polymer formed versus monomeric protein concentration. (iii) The addition of pre-formed nuclei (fibrillar seeds) at the beginning of an NDP reaction mitigates and often abolishes the lag phase that is akin to the crystallization process. In contrast, in an isodesmic or linear polymerization, there are no discrete nucleation and elongation phases (Figure 1.21b).[176] This type of polymerization is devoid of a lag phase, and also the rate constants are independent of the size of the polymer. Here, the concentration of monomers is highest at the onset of the reaction that eventually gets slowed down as the equilibrium is attained. Furthermore, no critical concentration barrier is there for polymer formation.

In the NDP polymerization, the fibrils formed in the elongation phase can fragment, producing new fibril ends that can recruit free monomers resulting in exponential growth of the fibrillar material. Furthermore, the dynamic nature of the fibrils can also cause depolymerization via the release of oligomers from the fibrillar species. Typically, there are three processes that generate new aggregates, namely, primary nucleation, fragmentation, and secondary nucleation,[177] in case of primary nucleation, the rate of formation of the new aggregates depends only on the concentration of free monomeric species (Figure 1.21c). Whereas, in fragmentation, the rate depends only on the concentration of existing fibrils (i.e., the aggregate mass). In secondary nucleation, the rate is dependent on both the concentration of free monomers and existing fibrils. Moreover, the surface of pre-existing fibrils can enhance the nucleation of new aggregates from the same kind of monomeric species that may lead to autocatalytic amplification of the amyloid assembly process (Figure 1.21c).[177,178] One another type of process in the amyloid formation is seeding, where the addition of pre-formed fibrils abolishes the nucleation step, and the aggregation

profile then follows a single exponential function.[179] This phenomenon is also termed as templating. For elucidating the seeding mechanisms, pre-formed fibrillar seeds rich in cross-β structure are introduced into a solution containing monomeric protein. These pre-formed seeds act as templates for monomers and predispose the monomers towards amyloid formation. This type of process has been proposed to be a possible basis for the propagation of infectious prion diseases and also other amyloid-related disorders. Finally, the most common methods that are used to monitor amyloid formation kinetics are ThT fluorescence, laser light scattering, and circular dichroism (CD) spectroscopy.

1.7 Structure and biogenesis of human functional amyloid protein, Pmel17

Pmel17 is a 668-residue melanocyte-specific type I transmembrane glycoprotein composed of a short N-terminal signal peptide (signal sequence), a large luminal subunit (Mα fragment), and an integral membrane subunit (Mβ fragment) (Figure 1.22a).[180] The protein name was given 'Pmel17' because while screening a λ-**p**hage library containing inserts from **mel**anocyte cDNA, the authors found that the longest of the derived clones was the **17**th isolated phage.[180] Each discoverer gave different names to Pmel17 that include Silver, SILV, gp100, ME20, etc. The fibrillar amyloid matrix of Pmel17 within acidic melanosomes (lysosome-related organelles) act as templates for melanin deposition underneath the skin and in the eyes.[181] Pmel17 plays an essential role in the structural organization of premelanosomes. The amyloid matrix formed by Pmel17 serves a beneficial role in mitigating the toxicity by sequestering and minimizing the diffusion of highly reactive quinone precursors that are required during melanin biosynthesis.[182] The proteolytic cleavage of Pmel17 by furin or a related protease of the proprotein convertase (PC) family of enzymes generates a highly amyloidogenic Mα fragment and a transmembrane Mβ fragment.[183] The luminal domain of Pmel17 is divided into distinct subdomains based on sequence homologies: an N-terminal domain (NTD) harbors three highly conserved N-glycosylation sites and shares sequence similarity to a homologous glycoprotein nonmetastatic *melanoma* B (GPNMB). It also also harbors three cysteine residues that might participate in disulfide bonding; the polycystic kidney disease domain (PKD) with homology to a repeat region in the polycystic kidney disease-associated protein, polycystin 1 (PKD1), is devoid of glycosylation sites and adopts a β-sheet conformation like other characterized PKD domains; an intrinsically disordered repeat domain (RPT) consisting of 10 imperfect repeats of a 13-residue each having a significant enrichment with

specific amino acid types such as proline, serine, threonine, valine, glycine, and glutamate (Figure 1.22a).[181] Pmel17 undergoes multiple sequential posttranslational proteolytic processing steps to generate a highly amyloidogenic Mα fragment composed of NTD, PKD, and RPT. It is anchored to membranes by NTD, facing the extracellular space and the short C-terminus, facing the cytoplasm.[184] In order to avert cytotoxicity, the formation of Pmel17 amyloid fibrils must occur in a directed pathway under tightly controlled conditions inside the cells. To understand the spatio-

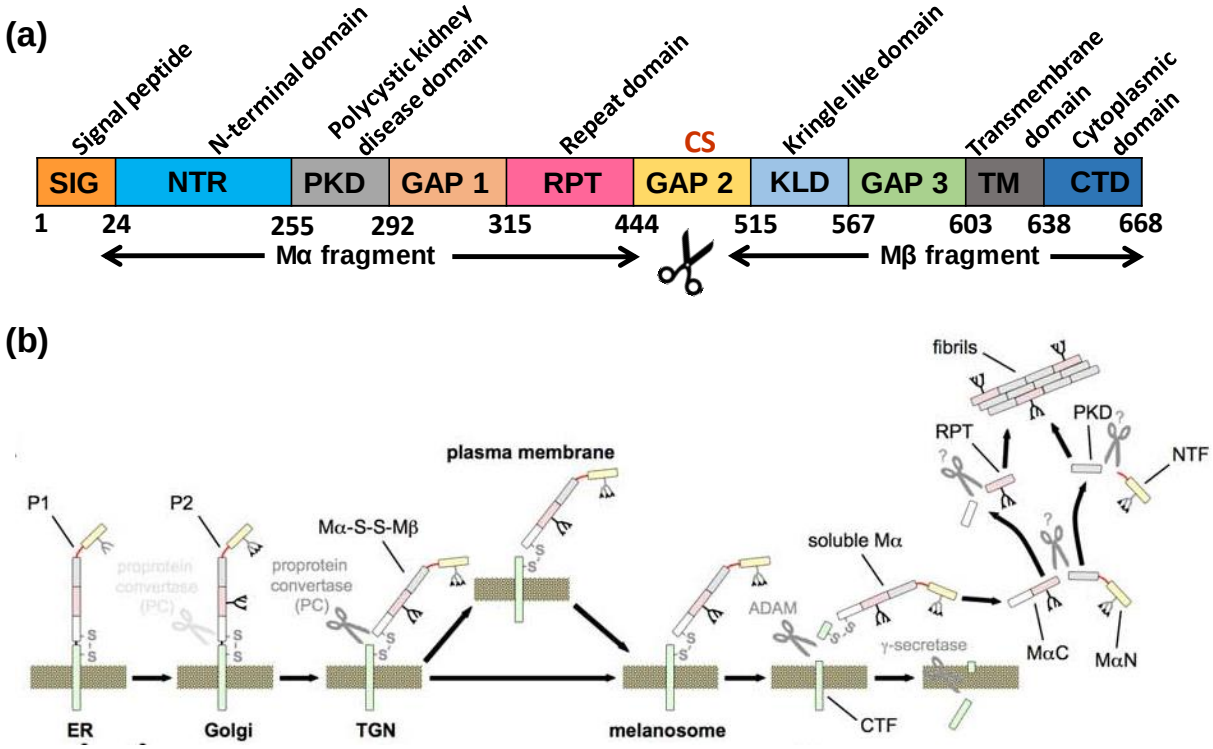


Figure 1.22 (a) Domains within pre-melanosomal protein (Pmel17). SIG, signal peptide; NTD, N-terminal domain; PKD, polycystic kidney disease-like domain; RPT, repeat domain (see below); KRG, Kringle-like domain; TM, transmembrane domain; CTD, C-terminal domain; GAP1, GAP2, and GAP3, undefined domains. **(b)** Pmel17 forms functional amyloid in melanin biogenesis. PMEL fibril formation is highly regulated by post-translational cleavage into its amyloidogenic form and compartmentalization within melanosomes during melanosome maturation. The Furin-like cleavage C-terminal to RPT at "CS" generates Mα fragment that is necessary for fibril formation Pmel17 fibrils catalyze the formation of melanin, concentrate melanin, and facilitate bulk transport of melanin. Reproduced with permission from Ref. 184.

-temporal transformation of Pmel17 from a transmembrane glycoprotein to an insoluble amyloid, it is imperative to elucidate the mechanisms regulating the biosynthesis, maturation, and trafficking of Pmel17. After the translocation into the endoplasmic reticulum (ER), Pmel17 gets

modified by removal of the signal peptide and by addition of four N-linked core oligosaccharides, generating a 'precursor 1' (P1) form (Figure 1.22b).[181,184] It is then slowly exported from the ER and released to the Golgi apparatus, where further modification takes place, in which the core N-linked oligosaccharides matures, and RPT gets modified extensively by terminally sialylated O-linked oligosaccharides that are known to be critical for Pmel17 recognition by the monoclonal antibody HMB45.[181] This highly modified full-length form of Pmel17 is termed as the precursor 2 or P2 form. Further, starting in the mid-Golgi or in the trans-Golgi network (TGN), Pmel17 undergoes cleavage by a proprotein convertase at a dibasic Lys-Arg sequence, that separates the large luminal fragment Mα from the small membrane-standing fragment Mβ.[181] These fragments, however, remain temporarily tethered by a disulfide bridge (Figure 1.22b). In this form (Mα-S-S-Mβ), the protein is either directly or via the plasma membrane, delivered to specialized multivesicular early endosome (stage I melanosomes) by the limiting membrane of this organelle that eventually buds into the interior in a process dependent on the tetraspanin CD63.[184] Followed by this, a second proteolytic cleavage in the luminal juxtamembrane region by a 'site 2 protease' and/or a metalloproteinase such as a disintegrin and metalloproteinase (ADAM) liberates soluble Mα fragment from the membrane- bound luminal region of Mβ (Mβ N). The remaining truncated portion of Mβ, known as the C-terminal fragment (CTF), is degraded by γ-secretase.[181,184] In addition, this cleavage leads to the segregation of the CTF from the melanosomal pathway into the degradative late endosomal/lysosomal system (Figure 1.22b). After this, the fibrillogenic Mα fragment is further cleaved between the PKD and the RPT domain by an unknown protease that releases the N-terminal fragment (MαN) and the C-terminal fragment (MαC). Moreover, both of these fragments undergo further processing to liberate the NTR/NTF, PKD, and RPT fragments. The PKD and RPT fragments then associate to form the ladder of fibrils that subsequently constitute the Pmel17 fibrillar assembly; however, the remaining NTR fragment contributes only to a minor extent (Figure 1.22b).

1.7.1 Pmel17 and generation of early-stage melanosomes

The most prominent features of early-stage melanosomes are the intraluminal fibrils that are primarily composed of fragments derived from a melanosomal protein, Pmel17. A key to understanding the formation of early-stage melanosomes is deciphering the crucial steps that are involved in the transformation of Pmel17 from transmembrane glycoprotein to a parallel sheet-

like fibrillar architecture.[185] Although the Pmel17 gene (PMEL) was first discovered in mice in 1930, however, it was named and mapped to the silver locus only in 1991.[180] Typically, the expression of Pmel17 is regulated by the pigment cells. It was reported that the overexpression of Pmel17 in non-pigmented cells results in late endosomes with fibrous striations akin to those found in melanosomes that revealed Pmel17 as the main protein component in the melanosomal fibrils. In addition, it has also been shown that the absence of fibrils resulted in a silver genotype in mice. This observation sheds light on the physiological role of Pmel17 in the biosynthesis of melanin.

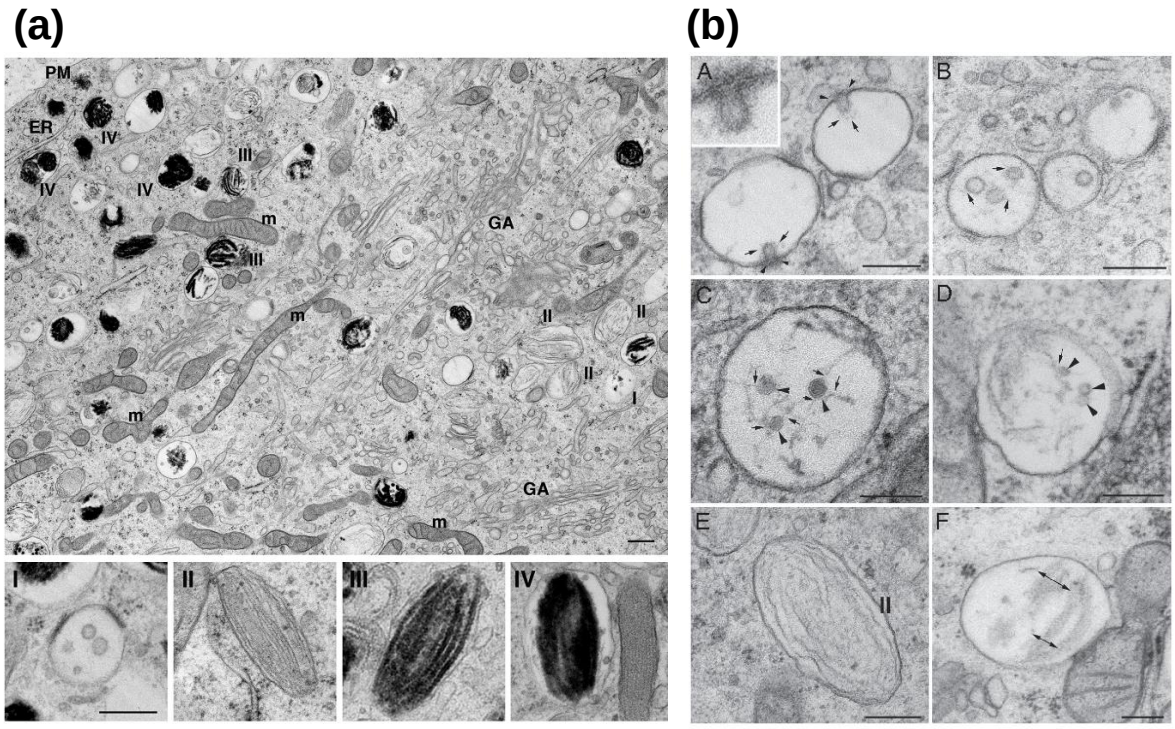


Figure 1.23. (a) Ultrastructure of MNT-1 cells represented in the electron micrographs captured using high-pressure freezing and freeze substitution (upper). GA, Golgi Apparatus; PM, plasma membrane; ER, endoplasmic reticulum; m, mitochondria. The lower image represents stage I, II, III, and IV melanosomes. **(b)** On intraluminal vesicles (ILVs), fibrils begin to form within early multivesicular bodies (MVBs). The arrows indicate small fibrils that are closely linked to nascent buds. The inset (A) shows higher magnification. B, C, D show vesicles of early MVBs from which thin fibrils appear emerging from 60–70 nm ILVs, refer arrows in B. The arrowheads and arrows, in D show examples of organelles bearing smaller, 40-nm vesicles and more organized fibrils, respectively. E shows a stage II ellipsoidal premelanosome with organized arrays of fibrils. The double arrows in F show the sheet-like appearance of the fibrils lying parallel to the plane of the section. The scale bar in each image is 200 nm. Reproduced with permission from Ref. 185.

Several groups have investigated the biogenesis of melanosomes by using MNT-1 cell line, a pigmented melanoma cell line. The cytoplasm of MNT-1 cells is composed of numerous mitochondria, an extended Golgi apparatus, and distinct melanosomal structures as shown in Figure 1.23a.[185] The melanosomal structures that comprise unpigmented (immature) melanosome is referred to as stages I and II melanosomes, whereas, the pigmented (mature) melanosomes are categorized as stages III and IV melanosomes. A myriad of studies reflects the following picture of the association of Pmel17 with melanosome maturation.[185,186] In the stage I melanosomes, upon Pmel17 cleavage, initiation of the formation of very short and thin fibrillar structures (1.3–2.4 nm in diameter) occurs on the spherical intraluminal vesicles (ILVs) that are present within numerous early multivesicular bodies (MVBs) (Figure 1.23a,b). Electron micrographs of MNT-1 cells generated using high-pressure freezing and freeze substitution revealed MVBs/stage-I melanosomes bearing coats on the limiting membrane undergo inward budding and producing 1-4 ILVs with diameter between 50–90 nm bearing as shown in Figure 1.23b.[185] The fibrillar structures have been observed to emanate from multiple sites on the ILVs, indicating that ILVs might be seeding the formation of these fibrillar structures. Furthermore, both the size and shape of melanosomes change in late-stage I due to the maturation of the fibrils that are accompanied by a decrease in the size of ILVs to 30-45 nm (Figure 1.23b).[185] The melanosome fibrils begin to elongate in association with the ILVs, and as the fibrils grow, the ILVs become 'pushed' to the periphery of the organelle. Figure 1.23a shows an ellipsoidal shape of stage II melanosome comprising an organized sheet-like array of longer and thicker intraluminal fibrils having a diameter of 6-10 nm.[185] Stage II melanosomes are found to be enriched in the luminal fragment of Pmel17 that is detected using monoclonal antibodies.[180] In stage III melanosomes, due to melanin deposition, the thickness of the fibrils increases 2-fold to a diameter of 14-20 nm (Figure 1.23a).[185] However, the detection of fibrils becomes challenging in stage III as well as stage IV melanosomes, most likely due to the masking of epitopes by the melanin deposits. Finally, in stage IV melanosomes, the fibrillar sheets are completely masked by the melanin (Figure 1.23a).[185]

1.7.2 Pmel17 amyloidogenesis and melanin biogenesis

The transformation of Pmel17 from an integral membrane protein to insoluble amyloid aggregates occurs through the secretory pathway within acidic endosomal compartments.[181] This endogenously synthesized protein assembles into amyloid fibrils only at a given stage of the

endosomal pathway that is tightly regulated by multiple sequential proteolytic cleavage steps. It was shown that the luminal fragment of Pmel17 (Mα fragment) is responsible for the amyloid formation, which was further detected in stage II melanosomes by immunocytochemistry. The amyloid fibrils are responsible for producing the elliptical shape of melanosomes that are proposed to be important for the transfer of melanosomes from melanocytes to keratinocytes.[187] Also, the formation of round and enlarged melanosomes in stage II of maturation has been observed in silver mice, which is due to the truncation of Pmel17.[180] Under in vitro conditions, the amyloid fibrils of Pmel17 fibrils is known to accelerate melanin polymerization, indicating that the fibrils might play a kinetic role in the biosynthesis of melanin or in minimizing the diffusion of highly toxic oxidative melanin intermediates.[182] After getting synthesized as transmembrane glycoprotein in the

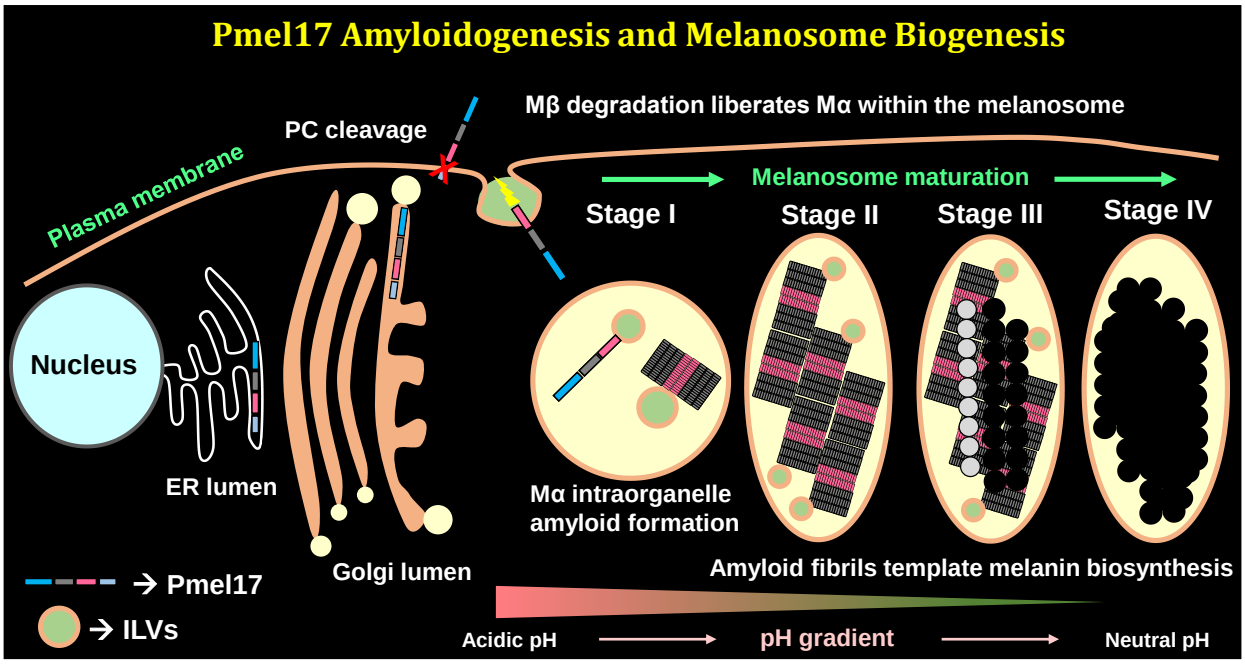


Figure 1.24. Schematic representation showing the formation of functional amyloid Pmel17 and melanin biogenesis. The formation of amyloid is highly regulated by post-translational modifications and several proteolytic processing steps. Stage I contains pre-fibrillar aggregates (grey and magenta) of Pmel17 emanating from intraluminal vesicles (ILVs). Stage II comprises long fibrillar striations spanning the length of the melanosome giving rise to an ellipsoidal shape. Stage III contains an enzyme, namely tyrosinase, along with melanin precursors represented as light grey that catalyze the formation of melanin, concentrate melanin deposits (black circles) on the fibrils. This stage indicates the scaffolding role of Pmel17 on melanin polymerization. In stage IV, the fibrils are completely masked by the melanin. Melanosome maturation follows a pH gradient ranging from acidic to neutral.

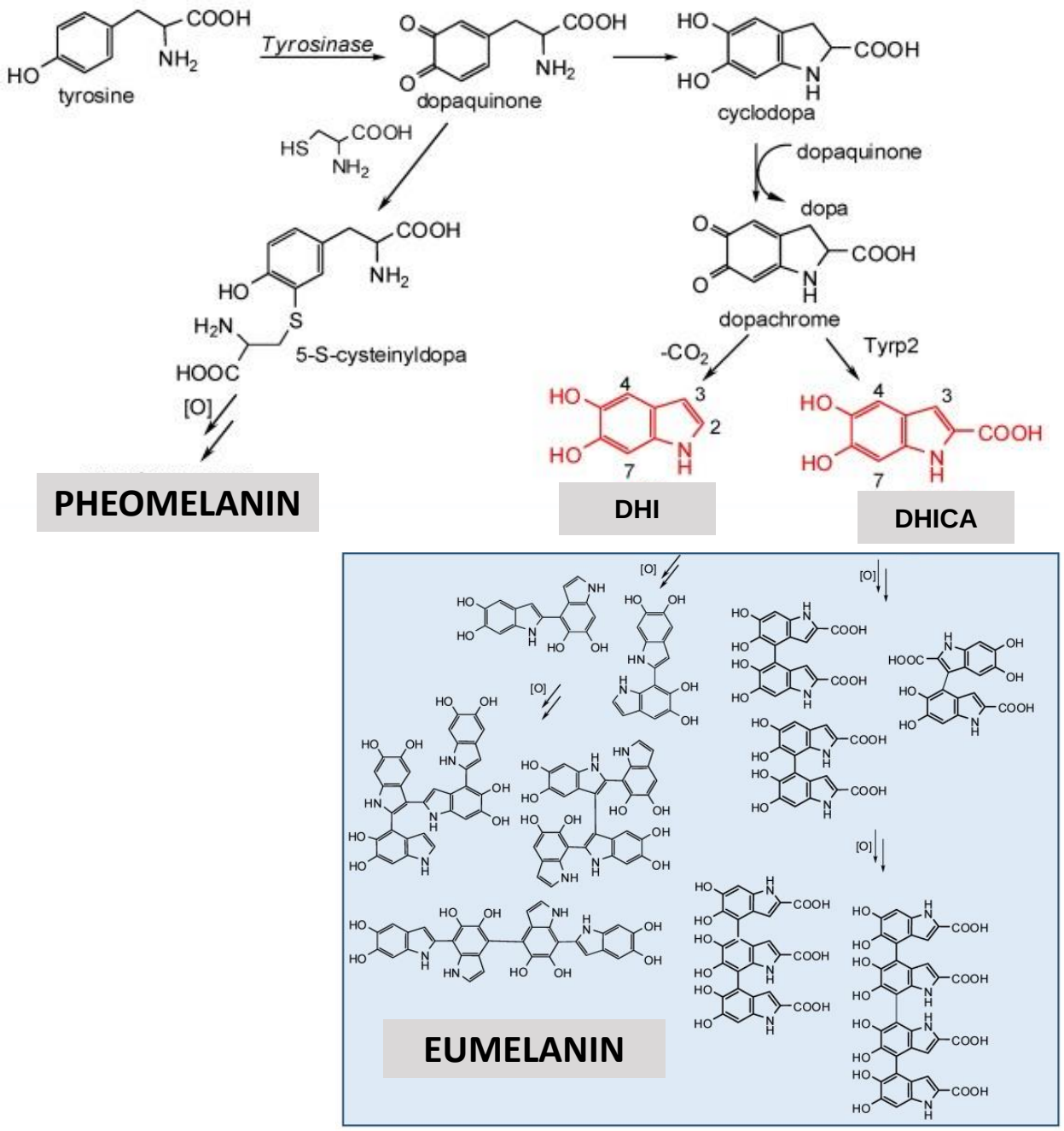


Figure 1.25 Schematic overview of the chemistry of melanin biosynthesis pathway. Tyrosinase related protein 2 (Tyrp2), 5,6-dihydroxyindole (DHI), 5,6-dihydroxyindole-2-carboxylic acid (DHICA). Reproduced with permission from Ref. 190.

endoplasmic reticulum (ER) lumen, Pmel17 enters the secretory pathway where it gets posttranslationally modified and after that undergoes multiple proteolytic cleavage (Figure

1.24).[184] The fibrillogenic Mα fragment is released into the lumen and gets endocytosed into an endosomal compartment that mature into melanosomes (Figure 1.24). Since the fibrils are formed within acidic compartments, it is imperative to understand the role of pH in amyloid formation. Also, it is known that the maturation of melanosomes is accompanied by a pH gradient.[188] Under acidic environment conditions, the amyloid assembly is initiated, which gets completed by stage II. Melanin is produced in the mature melanosomes near-neutral pH and is deposited during later stages, III and IV.[188] The rate-limiting step in melanin synthesis is catalyzed by a key enzyme, namely, tyrosinase. It has been reported that *in vitro*, the activity of tyrosinase is pH-dependent, with high activity close to neutral pH, which decreases under acidic conditions.[189] During melanin synthesis, tyrosine is converted to dihydroxyphenylalanine (DOPA), which is then transformed into dopaquinone.[190] The oxidation of dopaquinone results in the formation of 5,6-dihydroxyindole followed by 5,6-indolequinone (DHQ), which further polymerizes into an insoluble polymer called melanin (Figure 1.24 and Figure 1.25).[190] All the melanin precursors including, DHQ are known to be highly cytotoxic and are also potent oxidizers that can react both as electrophiles and nucleophiles. Therefore, diffusion and leakage of these precursors out of melanosomes into the cytosol would have appalling effects on cellular functions. Importantly, Mα fibrils play a crucial role in protecting from the toxic side effects of generating melanin by binding and sequestering these precursors and thereby, preventing their diffusion out of the melanosome.[182] One of the possible reasons for the binding/affinity of DHQ for amyloid fibrils might be due to the fact that DHQ shares a core that is isostructural with the benzothiazole substructure of a well-known amyloid marker, thioflavin T.[182] Although several studies have shed light on the ultrastructure of melanosomes and its maturation, yet, the molecular phase behavior of Pmel17 and the effect of intramelanosomal solution conditions on amyloid formation remain to be elucidated.

1.8 Thesis motivation and perspective

Amyloid structure has been considered notorious for causing a variety of human disorders. However, it has been found that their benign counterparts, also known as functional amyloids, play crucial roles in many biological processes. Despite possessing all the characteristics of pathological amyloid structures such as extensive β-sheet rich conformation, fibril diameter of 6-12 nm, and binding to CR and ThT, an intriguing question that still remains enigmatic is why pathological amyloids are highly toxic whereas functional amyloids are not. Therefore, detailed

studies on elucidating the amyloid formation kinetics, as well as the structural folds, are required. Interestingly, there are many similarities between the biology of Pmel17 and pathological amyloids. For instance, the proteolytic processing steps of the Pmel17 amyloid fragment are akin to those of disease-associated Aβ as well as BRI2-mutant peptides. A common set of proteases are involved in the processing steps of these amyloids, that includes proprotein convertases, α, β, and γ-secretases.[191] Moreover, many of these proteins/peptides form amyloid fibrils inside the multivesicular compartments. Also, ILVs and low pH conditions have been shown to promote the aggregation of both pathological and functional amyloids. Therefore, these remarkable similarities indicate that elucidating the molecular mechanisms of Pmel17 amyloid formation might provide important cues to mitigate the toxicity of their disease-related counterparts.[191]

In this thesis, the efforts were directed towards understanding the phase behavior of a low-complexity intrinsically disordered domain of a melanosomal protein under various physicochemical conditions. The dynamism, heterogeneity, and intermolecular association that drives RPT phase transitions and amyloid formation were studied using a multidisciplinary approach involving a combination of biophysical, biochemical, molecular biology, and imaging tools. In Chapter 2, using ultrafast spectroscopy, we embarked upon studies aimed at characterizing the conformational dynamics and behavior of water molecules around the amyloidogenic intrinsically disordered RPT. Our time-resolved fluorescence measurements revealed subtle alterations in the conformational- and solvation dynamics in a site-specific manner. RPT forms the amyloid core of Pmel17, that promotes the biosynthesis of UV-shielding pigment melanin, and its fibril stability is particularly shown to be sensitive to pH. Furthermore, the production and deposition melanin are entirely based on the morphology of the amyloid fibrils. Chapter 3 of this thesis provides mechanistic insights into the mechanism-to-morphology relationship in amyloid formation of RPT. Specific-ion interactions are known to modulate the aggregation kinetics of a variety of proteins that are associated with several human disorders. In Chapter 4, attempts were made to decode the complex interactions among ions, disordered protein molecules, and water by revisiting the Hofmeister series in the context of amyloid formation from the RPT. Our results indicate that Hofmeister ions modulate the autocatalytic amplification of the self-assembly process of RPT. A flurry of recent work suggests that IDPs/IDRs can also assemble into highly dynamic and reversible states. The self-assembly of these proteins into dense liquid phases is known to be a key for the formation of numerous intracellular compartments. However,

under stressed conditions or certain circumstances, these dense phases are also precursors of insoluble amyloid-like states. Therefore, it is important to decipher the critical balance between function and dysfunction of these dense phases. In Chapter 5, we shed light on the liquid-liquid phase separation and liquid-to-solid transition of RPT. We show how intramolecular charge-transfer promotes the conformational maturation of highly dynamic liquid-like droplets to solid-like aggregates. I believe that the work described in this thesis will improve our current understanding of the molecular mechanism of amyloid formation as well as phase transitions of mammalian functional amyloid, Pmel17, and will also be useful for studying the proteins that are involved in aberrant phase transitions.

1.9 References

(1) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Molecular biology of the cell, 4th ed.; Garland Science, Taylor & Francis Group: New York, 2002.

(2) Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973,** 181, 223-230.

(3) Dobson, C. M.; Šali, A.; Karplus, M. Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem. Int. Ed.* **1998,** 37, 868-893.

(4) Dobson, C. M. The structural basis of protein folding and its links with human diseases. *Phil. Trans. R. Soc. Lond. B* **2001,** 356, 133-145.

(5) Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. Berichte der deutschen chemischen Gesellschaft **1894,** 27, 2985-2993.

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

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

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

(9) Vladimir, N. Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics. *Front. Phys.* **2019,** 7, 1-18.

(10) 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.

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

(12) 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.

(13) 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. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001,** 19, 26-59.

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

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

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

(17) 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. *Intrinsically Disordered Proteins* **2013,** 1, e24157.

(18) Dill, K. A.; Chan, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Mol. Biol.* **1997,** 4, 10-19.

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

(20) Turoverov, K. K.; Kuznetsova, I. M.; Uversky, V. N. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggre-

-gation. *Prog. Biophys. Mol. Bio*, **2010,** 102, 73-84.

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

(22) Dobson, C. M. Protein folding and misfolding. *Nature* **2003,** 426, 884-890.

(23) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D^2 concept. *Annu. Rev. Biophys.* **2008,** 37, 215-246.

(24) 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, 6589-6631.

(25) Burger, M. V.; Gurry, T.; Stultz, M. C. Intrinsically disordered proteins: Where computation meets experiment. *Polymers* **2014,** 6, 2684-2719.

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

(27) 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, 1-26.

(28) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing Protein Intrinsic Disorder. *Chem. Rev.* **2014,** 114, 6561-6588.

(29) 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.

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

(31) 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, 8183-8188.

(32) 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, 13392-13397.

(33) He, Bo.; Wang, K.; Liu, Y.; Xue, Bin.; Uversky, V. N.; Dunker, A. K. Predicting intrinsic disorder in proteins: an overview. Cell Res. **2009,** 19, 929-949.

(34) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Structure, Function, and Bioinformatics* **2000,** 41, 415-427.

(35) Du, Z.; Uversky, V.N. A Comprehensive Survey of the Roles of Highly Disordered Proteins in Type 2 Diabetes. *Int. J. Mol. Sci.* **2017,** 18, 1-44.

(36) Xue, B.; Dunbrack, R. L.; Williams, R. W.; Dunker, A. K.; Uversky, V. N. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **2010,** 1804, 996-1010.

(37) Dosztányi, Z.; Csizmok, V.; Tompa, P.; Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **2005,** 21, 3433-3434.

(38) Li, J.; Feng, Y.; Wang, X.; Li, J.; Liu, W.; Rong, L.; Bao, J. An Overview of Predictors for Intrinsically Disordered Proteins over 2010–2014. *Int J Mol Sci.* **2015,** 16, 23446-23462.

(39) Walsh, I.; Martin, A. J. M.; di Domenico T.; Tosatto, S.C.E. ESpritz: Accurate and fast prediction of protein disorder. *Bioinform. Oxf. Engl.* **2012,** 28, 503-509.

(40) Ishida, T.; Kinoshita, K. Prediction of disordered regions in proteins based on the meta approach. *Bioinform. Oxf. Engl.* **2008,** 24, 1344-1348.

(41) Di Domenico, T.; Walsh, I.; CE Tosatto, S. Analysis and consensus of currently available intrinsic protein disorder annotation sources in the MobiDB database. *BMC Bioinformatics* **2013,** 14, 1-11.

(42) Oates, M. E.; Romero, P.; Ishida, T.; Ghalwash, M.; Mizianty, M. J.; Xue, B.; Dosztányi, Z.; Uversky, V. N.; Obradovic, Z.; Kurgan, L.; Dunker, A. K.; Gough, J. D^2P^2: database of disordered protein predictions. *Nucleic Acids Res.* **2013,** 41, D508-D516.

(43) Tompa, P.; Fersht, A. Structure and Function of Intrinsically Disordered Proteins. CRC Press: 2009.

(44) Gsponer, J.; Futschik, M. E.; Teichmann, S. A.; Babu, M. M. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **2008,** 322, 1365-1368.

(45) Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and Structural Disorder in Protein-Protein Interactions. *Trends Biochem. Sci.* **2008,** 33, 2-8.

(46) Romero, P. R.; Zaidi, S.; Fang, Y. Y.; Uversky, V. N.; Radivojac, P.; Oldfield, C. J.; Cortese, M. S.; Sickmeier, M.; LeGall, T.; Obradovic, Z.; Dunker, A. K. Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** 103, 8390-8395.

(47) Wang, J.; Cao, Z.; Zhao, L.; Li, S. Novel Strategies for Drug Discovery Based on Intrinsically Disordered Proteins (IDPs). *Int. J. Mol. Sci.* **2011,** 12, 3205-3219.

(48) 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, 432-440.

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

(50) Uversky, V. N. Intrinsically disordered proteins and their (disordered) proteomes in neurodegenerative disorders. *Front. Aging Neurosci.* **2015,** 7, 1-6.

(51) Pappu, R. V.; Wang, X.; Vitalis, A.; Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **2008,** 469, 132-141.

(52) Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu. Rev. Biophys.* **2016,** 45, 207-231.

(53) Rubinstein, M.; Colby, R. H. *Polymer Physics*. OUP Oxford: 2003.

(54) Raos, G.; Allegra, G. Macromolecular clusters in poor-solvent polymer solutions. *J. Chem. Phys.* **1997,** 107, 6479-6490.

(55) Flory, P. J. *Principles of Polymer Chemistry*. Cornell University Press: 1953.

(56) Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **2008,** 108, 74-108.

(57) Bagchi, B. Water in Biological and Chemical Processes: From Structure and Dynamics to Function. Cambridge University Press: Cambridge, 2013.

(58) Kamal, J. K.; Zhao, L.; Zewail, A. H. Ultrafast hydration dynamics in protein unfolding: human serum albumin. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 13411-13416.

(59) Pal, S. K.; Zewail, A. H. Dynamics of water in biological recognition. *Chem. Rev.* **2004,** 104, 2099-2124.

(60) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological water: femtosecond dynamics of macromolecular hydration. *J. Phys. Chem. B* **2002,** 106, 12376-12395.

(61) Bagchi, B. Water dynamics in the hydration layer around proteins and micelles. *Chem. Rev.* **2005,** 105, 3197-3219.

(62) Zhong, D.; Pal, S. K.; Zewail, A. H. Biological water: A critique. *Chem. Phys. Lett.* **2011,** 503, 1-11.

(63) Pal, S. K.; Peon, J.; Zewail, A. H. Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** 99, 1763-8.

(64) Qin, Y..; Wang, L.; Zhong, D. P. Dynamics and mechanism of ultrafast water-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** 113, 8424-8429.

(65) Pavlova, A.; Cheng, C. Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S. Protein structural and surface water rearrangement constitute major events in the earliest aggregation stages of tau. Proc. Natl. Acad. Sci. U. S. A. **2016,** 113, E127-E136.

(66) Nandi, N.; Bagchi, B. Dielectric relaxation of biological water. *J. Phys. Chem. B* **1997,** 101, 10954-10961.

(67) Nandi, N.; Bhattacharyya, K.; Bagchi, B. Dielectric relaxation and solvation dynamics of water in complex chemical and biological systems. *Chem. Rev.* **2000,** 100, 2013-2046.

(68) Zhang, Y.; Cremer, P. S. Interactions between macromolecules and ions: The Hofmeister series. *Current opinion in chemical biology* **2006,** 10, 658-663.

(69) Arosio, P.; Jaquet, B.; Wu, H.; Morbidelli, M. On the role of salt type and concentration on the stability behavior of a monoclonal antibody solution. *Biophysical chemistry* **2012,** 168-169, 19-27.

(70) Munishkina, L. A.; Henriques, J.; Uversky, V. N.; Fink, A. L. Role of protein-water interactions and electrostatics in alpha-synuclein fibril formation. *Biochemistry* **2004,** 43, 3289-3300.

(71) Theillet, F.; Binolfi, Andres.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; Elcock, A. H.; Gershenson, A.; Selenko, P. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **2014,** 114, 6661-6714.

(72) Lo Nostro, P.; Ninham, B. W., Hofmeister phenomena: an update on ion specificity in biology. *Chem. Rev.* **2012,** 112, 2286-2322.

(73) Okur, H. I.; Hladilkova, J.; Rembert, K. B.; Cho, Y.; Heyda, J.; Dzubiella, J.; Cremer, P. S.; Jungwirth, P., Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **2017,** 121, 1997-2014.

(74) Hofmeister, F. Zur Lehre von der Wirkung der Salze. II. *Arch. Exp. Pathol. Pharmakol.* **1888,** 24, 247-260.

(75) Kunz, W.; Henle, J.; Ninham, B. W. 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. *Curr. Opin. Colloid Interface Sci.* **2004,** 9, 19-37.

(76) Salis, A.; Ninham, B. W. Models and mechanisms of Hofmeister effects in electrolyte solutions, and colloid and protein systems revisited. *Chem. Soc. Rev.* **2014,** 43, 7358-7377.

(77) Marek, P. J.; Patsalo, V.; Green, D. F.; Raleigh, D. P. Ionic strength effects on amyloid formation by amylin are a complicated interplay among Debye screening, ion selectivity, and Hofmeister effects. *Biochemistry* **2012,** 51, 8478-8490.

(78) Zhang, Y. J.; Cremer, P. S. Chemistry of Hofmeister Anions and Osmolytes. *Annu. Rev. Phys. Chem.* **2010,** 61, 63-83.

(79) Ou, S.; Cui, Di.; Patel, S. Molecular modeling of ions at interfaces: exploring similarities to hydrophobic solvation through the lens of induced aqueous interfacial fluctuations. *Phys. Chem. Chem. Phys.* **2016,** 18, 30357-30365.

(80) Weber, S. C.; Brangwynne, C. P. Getting RNA and protein in phase. *Cell* **2012,** 149, 1188-1191.

(81) Dumetz, A. C.; Chockla, A. M.; Kaler, E. W.; Lenhoff, A. M. Protein Phase Behavior in Aqueous Solutions: Crystallization, Liquid-Liquid Phase Separation, Gels, and Aggregates. *Biophys. J.* **2008,** 94, 570-583.

(82) 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.

(83) 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.

(84) Shin, Y.; Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **2017,** 357, 1-11.

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

(86) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A. A. Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* **2009,** 324, 1729-1732.

(87) 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.* **2017,** 56, 11354−11359.

(88) Darling, A. L.; Liu, Y.; Oldfield, C. J.; Uversky, V. N. Intrinsically Disordered Proteome of Human Membrane-Less Organelles. Proteomics, **2017,** 18, 1-12.

(89) Hyman, A. A.; Brangwynne, C. P. Beyond Stereospecificity: Liquids and Mesoscale Organization of Cytoplasm. *Dev. Cell* **2011,** 21, 14-16.

(90) Mitrea, D. M.; Kriwacki, R. W. Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.* **2016,** 14, 1-20.

(91) Protter, D. S. W.; Rao, B. S.; Van Treeck, B.; Lin, Y.; Mizoue, L.; Rosen, M. K.; Parker, R. Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep.* **2018,** 22, 1401-1412.

(92) 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.

(93) Nott, T. J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T. D.; Bazett-Jones, D. P.; Pawson, T.; Forman-Kay, J. D.; Baldwin, A. J. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **2015,** 57, 936-947.

(94) 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, 1118-1125.

(95) 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, 8194-8203.

(96) 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, 7189−7194.

(97) 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, 72−85.

(98) Oparin, A. I. The Origin of Life. *McMillan, New York,* 1938.

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

(100) Nakashima, K. K.; Vibhute, M. A.; Spruijt, E. Biomolecular Chemistry in Liquid Phase Separated Compartments. *Front. Mol. Biosci.* **2019,** 6, 1-9.

(101) 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.

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

(103) 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.

(104) Aguzzi, A.; Altmeyer, M. Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol.* **2016,** 26, 547-558.

(105) Franzmann, T.; Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* **2019,** 294, 7128−7136.

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

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

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

(109) 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, 651–663.

(110) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015,** 11, 899−904.

(111) Bentley, E. P.; Frey, B. B.; Deniz, A. A. Physical Chemistry of Cellular Liquid-Phase Separation. *Chem. Eur. J.* **2019,** 25, 5600 – 5610.

(112) Lin Y.; Forman-Kay, J. D., Chan, H. S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry* **2018,** 57, 2499−2508.

(113) Reineke, L. C.; Tsai, W. C.; Jain, A.; Kaelber, J. T.; Jung, S. Y.; and Lloyd, R. E. Casein kinase 2 is linked to stress granule dynamics through phosphorylation of the stress granule nucleating protein G3BP1. *Mol. Cell. Biol.* 2017, 37:e00596–16.

(114) 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.

(115) Saito, M.; Hess, D.; Eglinger, J.; Fritsch, A. W.; Kreysing, M.; Weinert, B. T.; Choudhary, C.; Matthias, P. Acetylation of intrinsically disordered regions regulates phase separation. *Nat. Chem. Biol.* **2019,** 15, 51–61.

(116) Ruff, K. M.; Roberts, S.; Chilkoti, A.; Pappu, R. V. Advances in Understanding Stimulus-Responsive Phase Behavior of Intrinsically Disordered Protein Polymers. *J. Mol. Biol.* **2018,** 430, 4619−4635.

(117) Quiroz, F. G.; Chilkoti, A. Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat. Mater.* **2015,** 14, 1164−1171.

(118) Zhang, Qilu.; Weber, Christine.; Schubertcd, U. S.; Hoogenboom, Richard. Thermoresponsive polymers with lower critical solution temperature: from fundamental aspects

and measuring techniques to recommended turbidimetry conditions. *Mater. Horiz.*, **2017,** 4, 109-116.

(119) Niskanen, J.; Tenhu, H. How to manipulate the upper critical solution temperature (UCST)? *Polym. Chem.*, **2017,** 8, 220-232.

(120) Burke, K. A.; Janke, A. M.; Rhine, C. L.; Fawzi, N. L. Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol Cell.* **2015,** 60, 231–241.

(121) Urry, D. W., Gowda, D. C., Parker, T. M., Luan, C. H., Reid, M. C., Harris, C. M., Pattanaik, A., and Harris, R. D. Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* **1992,** 32, 1243−1250.

(122) Wise S. G.; Yeo, G. C.; Hiob, M. A.; Rnjak-Kovacina, J.; Kaplan, D. L.; Ng, M. K. C.; Weiss, A. S. Tropoelastin - a versatile, bioactive assembly module. *Acta Biomater.* **2014,** 10, 1532–1541.

(123) Simon, J. R.; Carroll, N. J.; Rubinstein, M.; Chilkoti, A.; López, G. P. Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. *Nat. Chem.* **2017,** 9, 509-515.

(124) Riback, J.A.; Katanski, C.D.; Kear-Scott, J. L.; Pilipenko, E.V.; Rojek, A. E.; Sosnick, T. R.; Drummond, D. A. Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **2017,** 168, 1028-1040.

(125) Kumari, B.; Kumar, R.; Kumar, M. Low complexity and disordered regions of proteins have different structural and amino acid preferences. *Mol. BioSyst.* **2015,** 11, 585-594.

(126) Hofweber, M.; Dormann, D. Friend or foe—Post-translational modifications as regulators of phase separation and RNP granule dynamics. *J. Biol. Chem.* 2019, 294, 7137-7150.

(127) 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.

(128) Monahan, Z.; Ryan, V. H.; Janke, A. M.; Burke, K. A.; Rhoads, S. N.; Zerze, G.; O'Meally, R.; Dignon, G. L.; Conicella, A. E.; Zheng, W.; Best, R. B.; Cole, R. N.; Mittal, J.; Shewmaker, F.; Fawzi, N. L. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.* **2017,** 36, 2951-2967.

(129) Murakami, T.; Qamar, S.; Lin, J. Q.; Schierle, G. S.; Rees, E.; Miyashita, A.; Costa, A. R.; Dodd, R. B.; Chan, F. T.; Michel, C. H.; Kronenberg-Versteeg, D.; Li, Y.; Yang, S. P.; Wakutani, Y.; Meadows, W.; Ferry, R. R.; Dong, L.; Tartaglia, G. G.; Favrin, G.; Lin, W. L.; Dickson, D. W.; Zhen, M.; Ron, D.; Schmitt-Ulms, G.; Fraser, P. E.; Shneider, N. A.; Holt, C.; Vendruscolo, M.; Kaminski, C. F.; St George-Hyslop, P. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015,** 88, 678-90.

(130) Wang, Z.; Zhang, G.; Zhang, H. Protocol for analyzing protein liquid−liquid phase separation. *Biophys. Rep.* **2019,** 5, 1−9.

(131) Prasad, A.; Bharathi, V.; Sivalingam, V.; Girdhar, A.; Patel, B. K. Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. *Front. Mol. Neurosci.* **2019,** 12, 1-36.

(132) Dougherty, D. A. The Cation-π Interaction. *Acc Chem Res.* **2013,** 46, 885-893.

(133) Kim, S.; Huang, Jun.; Lee, Y.; Dutta, S.; Yoo, H. Y.; Jung, Y. M.; Jho, Y.; Zeng, H.; Hwanga, D. S. Complexation and coacervation of like-charged polyelectrolytes inspired by mussels. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** 113, E847-E853.

(134) Sivasakthi, V.; Anitha, P.; Kumar, K. M.; Bag, S.; Senthilvel, P.; Lavanya, P.; Swetha, R.; Anbarasu, A.; Ramaiah, S. Aromatic-aromatic interactions: analysis of π-π interactions in interleukins and TNF proteins. *Bioinformation* **2013,** 9, 432-439.

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

(136) Kang, J.; Lim, L.; Song, J. ATP enhances at low concentrations but dissolves at high concentrations liquid-liquid phase separation (LLPS) of ALS/FTD-causing FUS. *Biochem. Biophys. Res. Commun.* **2018,** 504, 545-551.

(137) Yeo, G. C.; Keeley, F. W.; Weiss, A. S. Coacervation of tropoelastins. *Adv. Colloid Interface Sci.* **2011,** 167, 94-103.

(138) Dölker, N.; Zachariae, U.; Grubmüller, H. Hydrophilic Linkers and Polar Contacts Affect Aggregation of FG Repeat Peptides. *Biophys J.* **2010,** 98, 2653-2661.

(139) Zhou, H.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2019,** 118, 691-1741.

(140) Maharana, S.; Wang, Jie.; 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.; Alberti, S. RNA buffers the phase separation behavior of prion-like RNA-binding proteins. *Science* **2018,** 360, 918-921.

(141) Navarro, M. G.; Kashida, S.; Chouaib, R.; Souquere, S.; Pierron, G.; Weil, D.; Gueroui, Z. RNA is a critical element for the sizing and the composition of phase-separated RNA–protein condensates. *Nat. Commun.* **2019,** 10, 1-13.

(142) 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, 637-648.

(143) Feng Luo Xinrui Gui Heng Zhou Jinge Gu Yichen Li, Xiangyu Liu, Minglei Zhao, Dan Li Xueming Li and Cong Liu. Atomic structures of FUS LC domain segments reveal bases for reversible amyloid fibril formation. *Nat. Struct. Mol. Biol.* **2018,** 25, 341-346.

(144) Shukla, A.; Mukherjee, S.; Sharma, S.; Agrawal, V.; Radha Kishan, K. V.; Guptasarma, P. A novel UV laser-induced visible blue radiation from protein crystals and aggregates: scattering artifacts or fluorescence transitions of peptide electrons delocalized through hydrogen bonding? *Arch. Biochem. Biophys.* **2004,** 428, 144-153.

(145) Prasad, S.; Mandal, I.; Singh, S.; Paul, A.; Mandal, B.; Venkatramani, R.; Swaminathan, R. Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **2017,** 8, 5416-5433.

(146) Del Mercato, L. L.; Pompa, P. P.; Maruccio, G.; Della Torre, A.; Sabella, S.; Tamburro, A. M.; Cingolani, R.; Rinaldi, R. Charge transport and intrinsic fluorescence in amyloid-like fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104, 18019-18024.

(147) Sharpe, S.; Simonetti, K.; Yau, J.; Walsh, P. Solid-State NMR characterization of autofluorescent fibrils formed by the elastin-derived peptide GVGVAGVG. *Biomacromolecules* **2011,** 12, 1546-1555.

(148) Pinotsi, D.; Grisanti, L.; Mahou, P.; Gebauer, R.; Kaminski, C. F.; Hassanali, A.; Kaminski Schierle, G. S. Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **2016,** 138, 3046-3057.

(149) Pinotsi, D.; Buell, A. K.; Dobson, C. M.; Schierle, G. S. K.; Kaminski, C. F. A Label-Free, Quantitative Assay of Amyloid Fibril Growth Based on Intrinsic Fluorescence. *ChemBioChem* **2013,** 14, 846-850.

(150) Dogra, P.; Joshi, A.; Majumdar, A.; Mukhopadhyay, S. Intermolecular Charge-Transfer Modulates Liquid–Liquid Phase Separation and Liquid-to-Solid Maturation of an Intrinsically Disordered pH-Responsive Domain. *J. Am. Chem. Soc.* **2019,** 141, 20380-20389.

(151) Sipe, J. D.; Cohen, A. S. Review: history of the amyloid fibril. *J. Struct. Biol.* **2000,** *130*, 88-98.

(152) Baldwin, A. J.; Knowles, T. P. J.; Tartaglia, G. G.; Fitzpatrick, A. W.; Devlin, G. L.; Shammas, S. L.; Waudby, C. A.; Mossuto, M. F.; Meehan, S.; Gras, S. L.; Christodoulou, J.; Anthony-Cahill, S. J.; Barker, P. D.; Vendruscolo, M.; Dobson, C. M. Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* **2011,** *133*, 14160-14163.

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

(154) Toyama, B. H.; Weissman, J. S. Amyloid Structure: Conformational Diversity and Consequences. *Annu. Rev. Biochem.* **2011,** 80, 557–85.

(155) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** 15, 384-396.

(156) 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, 755-773.

(157) Eisenberg, D.; Jucker, M. The amyloid state of proteins in human diseases. *Cell* **2012,** 148, 1188-1203.

(158) Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E. "Cross-β" conformation in proteins. J. Mol. Biol. 1968, 32, 343-358.

(159) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** 435, 773-778.

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

(161) Ow, S.; Dunstan, D. E. A brief overview of amyloids and Alzheimer's disease. *Protein Sci.* **2014,** 23, 1315-1331.

(162) Gallardo, R.; Ranson, N. A.; Radford, S. E. Amyloid structures: much more than just a cross-β fold. *Curr. Opin. Cell Biol.* **2020,** 60, 7-16.

(163) Soto, C.; Pritzkow, S. Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nat. Neurosci.* **2018,** 21, 1332-1340.

(164) Knowles, T. P.; Buehler, M. J. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* **2011,** 6, 469−479.

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

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

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

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

(169) Mukhopadhyay, S.; Dalal, V.; Arya, S. Nanoscale Optical Imaging of Amyloid Fibrils in "Bionanoimaging: Insights into Protein Misfolding and Aggregation. Bio-Nanoimaging: Protein Misfolding & Aggregation Elsevier 2013.

(170) Powers, E. T.; Powers, D. L. Mechanisms of protein fibril formation: nucleated polymerization with competing off-pathway aggregation. *Biophys. J.* **2008,** 94, 379-391.

(171) Serio, T. R.; Cashikar, A. G.; Kowal, A. S.; Sawicki, G. J.; Moslehi, J. J.; Serpell, L.; Arnsdorf, M. F.; Lindquist, S. L. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **2000,** 289, 1317.

(172) Stefani, M.; Dobson, C. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **2003,** 81, 678-699.

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

(174) Kumar, S.; Udgaonkar, J. B. Mechanisms of amyloid fibril formation by proteins. *Curr. Sci.* **2010,** 98, 639-656.

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

(176) Cremades, N.; Dobson, C. M. The contribution of biophysical and structural studies of protein self-assembly to the design of therapeutic strategies for amyloid diseases. *Neurobiol. Dis.* **2018,** 109,178-190.

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

(178) 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.* **2018,** 54, 8667-8684.

(179) Harper, J. D.; Lansbury, P. T. Models of amyloid seeding in alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* **1997,** 66, 385-407.

(180) Theos, A. C.; Truschel, S. T.; Raposo, G.; Marks, M. S. The Silver Locus Product Pmel17/gp100/Silv/ME20: Controversial in Name and in Function. *Pigm. Cell Res.* **2005,** 18, 322-336.

(181) Watt, B.; van Niel, G.; Raposo, G.; Marks, M. S. PMEL: A Pigment Cell-Specific Model for Functional Amyloid Formation. *Pigm. Cell Melanoma Res.* **2013,** 26, 300-315.

(182) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional amyloid formation within mammalian tissue. *PLoS Biol.* **2005,** 4, No. e6.

(183) Berson, J. F.; Theos, A. C.; Harper, D. C.; Tenza, D.; Raposo, G.; Marks, M. S. Proprotein Convertase Cleavage Liberates a Fibrillogenic Fragment of a Resident Glycoprotein to Initiate Melanosome Biogenesis. *J. Cell Biol.* **2003,** 161, 521-533.

(184) Leonhardt, R. M.; Vigneron, N.; Hee, J. S.; Graham, M.; Cresswell, P. Critical Residues in the PMEL/Pmel17 N-terminus Direct the Hierarchical Assembly of Melanosomal Fibrils. *Mol. Biol. Cell* **2013,** 24, 964-981.

(185) Hurbaina, I.; Geerts, W. J. C.; Boudier, T.; Marco, S.; Verkleij, A. J.; Marks, M. S.; Raposo, G. Electron tomography of early melanosomes: Implications for melanogenesis and the generation of fibrillar amyloid sheets. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** 105, 19726-19731.

(186) Berson, J. F.; Harper, D. C.; Tenza, D.; Raposo, G.; Marks, M. S. Pmel17 Initiates Premelanosome Morphogenesis within Multivesicular Bodies. *Mol. Biol. Cell* **2001,** 12, 3451-3464.

(187) Bissig, C.; Rochin, L.; van Niel, G. PMEL Amyloid Fibril Formation: The Bright Steps of Pigmentation. *Int J Mol Sci.* **2016,** 17, 1-14.

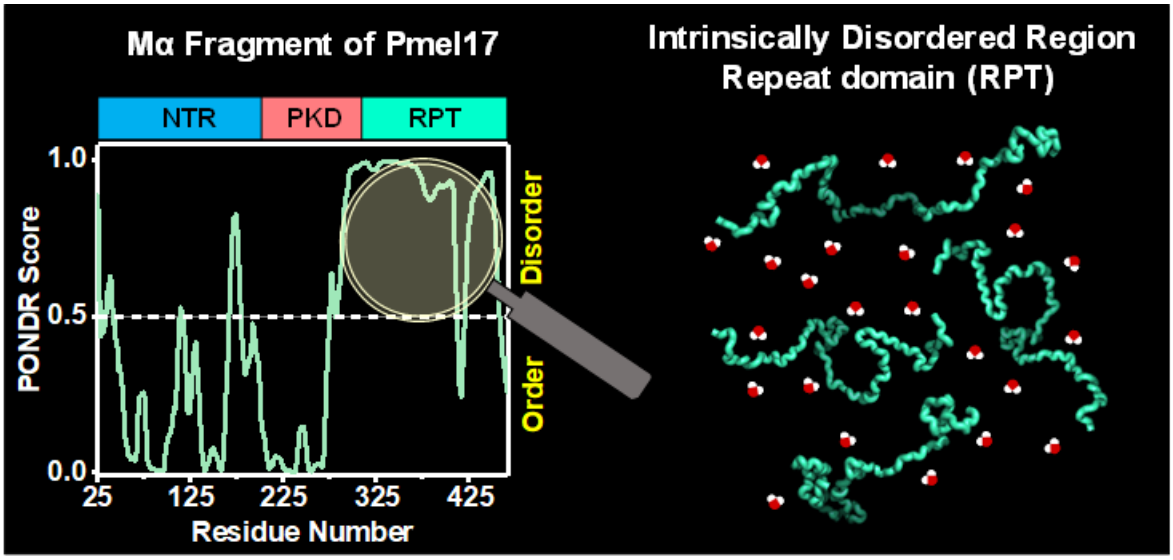
(188) Pfefferkorn, C. M.; McGlinchey, R. P.; Lee, J. C. Effects of pH on Aggregation Kinetics of the Repeat Domain of a Functional Amyloid, Pmel17. *Proc. Natl. Acad. Sci. U.S.A.* 2010, 107, 21447-21452.

(189) Devi, C. C.; Tripathi, R. K.; Ramaiah, A. pH-dependent interconvertible allosteric forms of murine melanoma tyrosinase—physiological implications. *Eur. J. Biochem.* **1987,** 166, 705-711.

(190) Panzella, L.; Ebato, A.; Napolitano, A.; Koike, K. The Late Stages of Melanogenesis: Exploring the Chemical Facets and the Application Opportunities. *Int. J. Mol. Sci.* **2018,** 19, 1-16.

(191) Hee, J. S.; Mitchell, S. M.; Liu, X.; Leonhardt, R. M. Melanosomal formation of PMEL core amyloid is driven by aromatic residues. *Sci. Rep.* **2017,** 7, 1-15.

Conformational Dynamics and Solvation of a pH-Responsive Intrinsically Disordered Repeat Domain (RPT) of a Human Functional Amyloid Protein, Pmel17



Dogra, P.; Arya, S.; Singh, A.; Datta, A.; Mukhopadhyay, S. Conformational Dynamics and Solvation of an Intrinsically Disordered Domain of a Human Functional Amyloid Protein, Pmel17 (manuscript in preparation).

2.1 Introduction

Unlike globular/structured proteins, intrinsically disordered proteins or regions (IDPs or IDRs) exist as highly dynamic ensemble of structures due to significant fluctuations in the polypeptide backbone in the Ramachandran ϕ-ψ conformational space.[1-17] This conformational flexibility of IDPs allow them to perform a wide variety of physiological functions like recognition, signaling, cell division, cell cycle control, regulation of transcription and translation etc.[2,3,6,9,11,13,15] However, certain promiscuous and unwanted interactions in IDPs/IDRs due to altered expression and processing can lead to numerous debilitating disorders such as cardiovascular disease, diabetes, cancer, and neurodegenerative diseases.[2,18-21] Several studies have revealed that the amino acid composition dictates the conformational state of IDPs and determines whether they will adopt an extended or a compact conformation.[1,2,4] In other words, the conformational flexibility and structural instability are encoded in their amino acid sequences.[2,22,23] IDPs that are rich in uncharged, polar amino acid residues have a tendency to form compact/collapsed globules in water.[24-26] In addition, the net charge per residue governs the intrinsic preference of the polypeptide chain to undergo collapse or expansion.[27] A high net charge per residue in the sequence of IDPs favors globule-to-coil transition.[27,28] In order to understand the function and conformational behavior of IDPs, it is imperative to take into consideration its surrounding environment. The polypeptide chain and solvent (water) interactions are believed to strongly influence a myriad of critical processes like folding, assembly, misfolding, and aggregation.[29-35] The water molecules present at the protein-water interface constitute the hydration layer and possess distinct dynamical behavior from the bulk water. These interfacial water molecules are termed as "biological water", which exhibit unique dielectric properties.[36-38] Several studies have indicated the importance of biological water in various phenomena like molecular recognition, DNA-protein, and protein-protein interactions.[39-41] The intricate balance between chain-solvent and chain-chain interactions defines the conformational states of IDPs and controls the self-assembly of highly ordered amyloid structures.[42-53] Therefore, to understand different conformations of the IDPs and their roles in physiology and disease, it is essential to monitor the distribution and dynamics of water molecules present in the vicinity of the polypeptide chains. In this work, we have chosen an intrinsically disordered repeat domain (RPT) of a human functional amyloid protein, Pmel17.[54-59] RPT is

known to form the amyloid core that promotes melanin biogenesis inside specialized skin cells called melanocytes.[55,56,58,59] However, the behavior of water molecules around RPT and the role of water in RPT aggregation remains elusive. Here, using ultrafast laser spectroscopy, we have been able to delineate the differences in the conformations and surface water dynamics within different regions of natively unfolded RPT. In order to probe the conformational states and behavior of water molecules within RPT, we have employed two fluorescence-based techniques, namely, femtosecond up-conversion and time-correlated single photon counting, that offers high sensitivity and selectivity to monitor the reorientation dynamics of water molecules both in bulk phase and at the protein-water interface and allows monitoring the dynamic motions of water molecules with femtosecond time resolution.[39,52] This timescale is significantly crucial for the detailed understanding of water present at the protein interfaces since hydrogen bond rearrangements in water occur on this time scale. We propose that the differences in water arrangement are likely to translate distinct roles of different regions of RPT in aggregation and amyloid formation.

2.2 Experimental Section

2.2.1 Materials

Guanidinium hydrochloride and urea were purchased from Ameresco. Isopropyl-thiogalactopyranoside (IPTG) was obtained from Gold Biocom. Dithiothreitol (DTT), tris HCl, sodium chloride, and sodium hydrogen phosphate (monobasic) were procured from Sigma (St. Louis, MO). 6-Acryloyl-2-(dimethylamino)naphthalene (acrylodan) was purchased from Molecular Probes, Invitrogen. The pH of the buffers was adjusted using on Metrohm (Herisau, Switzerland) 827 pH meter at ~25°C.

2.2.2 Bioinformatic analyses of RPT (315-444) using various prediction tools

The intrinsic disorder in RPT was predicted using the disorder prediction tool PONDR[1]. PONDR-VLXT plot was generated using PONDR (http://www.pondr.com/) and was then plotted using the Origin 2018 software. For identifying the low-complexity regions in RPT, we used SMART[3] (http://smart.embl-heidelberg.de/).

2.2.3 RPT expression, purification, and labeling

The repeat domain of Pmel17 (RPT) (315-444) with C-terminal His-tag present in the pET-21a(+) vector was expressed in E. coli BL21 (DE3) RIPL cells/strain (Agilent Technologies, Stratagene). Single point mutation at S350C (Cys350) and A415C (Cys415) were created using a site-directed mutagenesis kit, QuikChange (Stratagene). The following primers were used for generating the mutations.

S350C Forward: 5' -GGAACCACATGTGTGCAGGTGCC-3'

S350C Reverse: 5' -GGCACCTGCACACATGTGGTTCC-3'

A415C Forward: 5' -CTGGAACCACATGTGCACAGGTAAC-3'

A415C Reverse: 5' -GTTACCTGTGCACATGTGGTTCCAG-3'

and purified using the protocol as described previously[56] with slight modification. Briefly, the transformed cells were grown at 37 °C with shaking to an optical density at 600 nm (A_{600}) equal to 0.4−0.5 and then induced using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 8 h at 30 °C. The cells obtained after centrifugation (4000 rpm) for 60 min at 4 °C were lysed using the lysis buffer (6 M guanidinium hydrochloride, 100 mM NaCl, 100 mM K_2HPO_4, pH 7.5, and 10 mM imidazole) and then incubated for 2 h at room temperature with gentle agitation. The cell lysate obtained was centrifuged at 11,500 rpm for 60 min at 4 °C. The supernatant obtained was allowed to mix with nickel-nitrilotriacetic acid agarose [Ni-NTA (Qiagen); 5 mL per 1 L culture] resin (pre-equilibrated with the lysis buffer) at 4 °C for 60 min with gentle agitation. The mixture was poured into a 20 mL polypropylene column and washed with 4-column volumes of buffer containing 6 M guanidinium hydrochloride, 100 mM NaCl, 100 mM K_2HPO_4, pH 7.5, and 10 mM imidazole, followed by washing with 10-column volumes of buffer containing 6 M guanidinium hydrochloride, 100 mM NaCl, 100 mM K_2HPO_4, pH 7.5, and 30 mM imidazole. RPT was eluted using the buffer (8 M urea, 10 mM NaCl, 100 mM K_2HPO_4, pH 7.5, and 250 mM imidazole). The eluted protein was then diluted using 20 mM Tris, pH 8.0 buffer and further purified by FPLC using a linear NaCl gradient on anion exchange column (Q-Sepharose; GE Healthcare Life Sciences). The protein fractions were then desalted/buffer-exchanged by gel filtration chromatography (Superdex G-75; GE Healthcare Life Sciences) using 20 mM Tris buffer, pH 8.0. The fractions containing pure protein were assayed using 12% sodium dodecyl sulfate-

polyacrylamide gel electrophoresis. The molar extinction coefficient used for estimating the protein concentration was 5500 $M^{-1}cm^{-1}$.[56] Typical concentration of the protein obtained was ~10 mg/mL. Finally, purified RPT in 20 mM tris buffer, pH 8.0, was stored at -80 °C for further use.

The labeling of the cysteine mutants was carried out in native buffer (20 mM NaH_2PO_4 buffer, 1 mM DTT, pH 7.4). Purified protein was mixed in a molar ratio of 5:1 (Acrylodan:RPT[S422C]). The reaction mixtures were allowed to stir in the dark for 2-3 h at room temperature and the labeled proteins were then passed through PD-10 column to remove excess dye using 20 mM NaH_2PO_4 buffer, pH 7.4 containing 50 mM NaCl. The concentration of the labeled protein was estimated using ε_{365} = 12800 M^{-1} cm^{-1}.[48] The total protein concentration was estimated using ε_{280} = 5500 M^{-1} cm^{-1} [56], taking correction factor of acrylodan into account. The labeling efficiency was >85% for all the mutant proteins. The concentration of the labeled proteins used for all our measurements was 20-40 μM.

2.2.4 Circular dichroism (CD) measurements

For far-UV CD measurements were performed on a Chirascan Spectrophotometer (Applied Photophysics, UK) in a 1 mm pathlength quartz cell. The CD spectra were collected in the range of 200-260 nm with a bandwidth of 1 nm. The final protein concentration used for the measurements was 15 μM. The spectra obtained were corrected for the buffer background and were further smoothened using ProData software provided with the instrument. Finally, the spectra were plotted in the Origin 2018 software.

2.2.5 Steady-state fluorescence measurements

All the steady-state fluorescence data were acquired in Fluoromax-4 (Horiba Jobin Yvon, NJ). For recording tryptophan fluorescence, the excitation wavelength (λex) was set at 280 nm, and the emission was collected in the range of 315-500. For monitoring acrylodan fluorescence, the following parameters were adjusted: λ_{ex} = 375 nm, emission range = 440-600 nm. The steady-state fluorescence anisotropies were measured at emission maxima. The steady-state fluorescence anisotropy (r_{ss}) is given by the following equation:

$$r_{ss} = (I_{\parallel} - I_{\perp}G) / (I_{\parallel} + 2I_{\perp}G) \qquad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor.

2.2.6 Time-correlated single photon counting (TCSPC) measurements:

Time-resolved emission spectra (TRES): The time-resolved fluorescence intensity decay profiles were recorded at the magic angle (54.7°) using a time-correlated single photon counting (TCSPC) setup (Fluorocube, Horiba Jobin Yvon, NJ). A 375-nm laser diode was used as an excitation source for acrylodan. The peak count was fixed at 10,000. The instrument response function (IRF) was collected using colloidal silica (Ludox) solution and had a full width at half maxima (FWHM) of ~ 250 ps. The fluorescence decay profiles were collected from 440-600 nm at 10-nm interval. The fluorescence decays were then analyzed and fitted using DAS6 software provided with the instrument.

Fluorescence anisotropy decay: The time-resolved fluorescence anisotropy decay measurements were performed using a TCSPC setup coupled with a diode laser as described above. The orientation of emission polarizer was set at 0° and 90° with respect to the excitation polarizer for collecting the parallel fluorescence intensities ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission monochromator was fixed at 530 nm. Data analyses were performed using our previously reported protocol.[48-52,60] Briefly, the anisotropy decays were analyzed by globally fitting $I_{\parallel}(t)$ and $I_{\perp}(t)$ as follows:

$$I_{\parallel}(t) = I(t)[1+2r(t)]/3 \qquad (2)$$

$$I_{\perp}(t) = I(t)[1-r(t)]/3 \qquad (3)$$

$I(t)$ is the fluorescence intensity collected at the magic angle (54.7°) at time t, and $I_{\perp}(t)$ was always corrected using a *G*-factor. The anisotropy decays were then analyzed using a biexponential decay model describing fast and slow rotational correlation times as follows:

$$r(t) = r_0\left[\beta_{fast}\, exp\left(-\frac{t}{\phi_{fast}}\right) + \beta_{slow}\, exp\left(-\frac{t}{\phi_{slow}}\right)\right] \qquad (4)$$

where r_0: time-zero intrinsic fluorescence anisotropy; ϕ_{fast} & ϕ_{slow}: fast and slow rotational correlation times, respectively; β_{fast} & β_{slow}: amplitudes associated with fast and slow rotational time, respectively. See Table 2 for recovered parameters.

2.2.7 Construction of TRES and data analysis from TCSPC and up-conversion measurements

The time-resolved emission spectra (TRES) analysis was done using the methods described in our previous articles.[48,50,51,52] Briefly, time-resolved emission spectra (TRES) were constructed using:

$$I(\lambda, t) = I_{\lambda}^{SS} I_{\lambda}(t) / \sum_i \alpha_i \tau_i \qquad (5)$$

where $I_{\lambda}(t) = \sum_i \alpha_i \exp(-t/\tau_i)$; α_i and τ_i represent the amplitude and lifetime of the different lifetime components, respectively, I_{λ}^{SS} represents the steady-state fluorescence intensity at a fixed emission wavelength λ, and $I_{\lambda}(t)$ is same as mentioned above. The fluorescence decay parameters recovered were used to simulate the fluorescence decays at different emission wavelengths using Origin 2018 software. The spectra obtained at different time points were fitted using the log-normal or Gaussian function to recover the peak frequency $\nu(t)$ and were then normalized for the estimation of overall time-dependent fluorescence Stokes shift (TDFSS). The solvation correlation function was then constructed using:

$$C(t) = [\nu(t) - \nu(\infty)] / [\nu(0) - \nu(\infty)] \qquad (6)$$

where $\nu(0)$, $\nu(t)$, and $\nu(\infty)$ are the emission peak frequencies at time 0, t, and ∞, respectively. The recovered time-zero frequency ($\nu(0)$) from our femtosecond experiments was 21,000 ± 270 cm^{-1}, which is close to the time zero reported previously for acrylodan using the method described for theoretical estimation of time-zero spectrum.[61,65] The correlation function was finally fitted using a sum of exponentials as the following:

$$C(t) = \sum_i \alpha_{si} \exp(-t/\tau_{si}) \qquad (7)$$

where α_{si} and τ_{si} represent the amplitude and solvation time of the different solvation components, respectively. The solvation correlation function constructed from femtosecond up-conversion data was fitted till 200 ps.

2.2.8 Femtosecond fluorescence up-conversion measurements

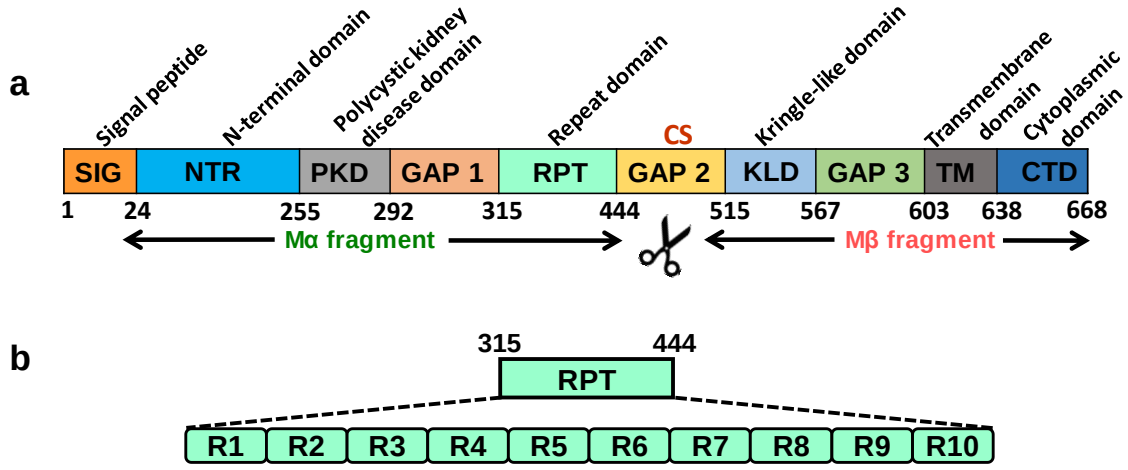
The fluorescence up-conversion measurements and data analyses were performed using the methods described in our previous articles.[51,52] Briefly, a second harmonic light of mode-locked Ti:Sapphire laser (Tsunami, Spectra Physics, USA) with pulses of 100 fs (1 W, 80 MHz repetition rate), was used as the excitation source. The Ti:Sapphire laser was then pumped by a continuous wave second harmonic light of the output of 4.5 W Nd:YVO_4 laser (Millenia Pro, Spectra Physics, USA) which was further pumped by an electrically pumped diode laser to generate 800 nm light. For exciting the samples, ~25 mW of 400 nm light was used. The polarization of the excitation light was fixed at 54.7° (magic angle) with respect to the gate light. The up-conversion of emission light was done using 800 nm gate light in 1 mm thick β-BBO (type I) crystal. The resultant up-converted light was passed through a UV bandpass filter, which cuts excitation, emission, and gate light and then passed through a double monochromator (DM2100, CDP, Russia). For the detection of the fluorescence signal, a photo-multiplier tube (PMT) having dark noise <5 cps was used. The IRF was estimated by measuring the cross-correlation between the gate light and Raman signal of buffer at 400 nm (excitation). The resulting profile, when fitted using a Gaussian equation yielded a full width at half maxima (FWHM) of ~ 250 fs. The fluorescence up-conversion transients were collected from 440-560 nm at 10-nm intervals till 300 ps. The emission decay traces were fitted using two/three exponentials in a homebuilt program in Igor Pro 6.37 (Burai et al. *J. Chem. Phys.* 2009, *131*, 034504). The raw- and the fitted data were then plotted in Origin 2018 software.

2.3 Results and Discussion

2.3.1 Conformational characterization of the native state of RPT

Pmel17 is a 668-residue melanocyte-specific transmembrane protein that undergoes multiple sequential proteolytic processing steps to generate a highly amyloidogenic luminal fragment (Mα) and a transmembrane fragment (Mβ). The fibrillogenic Mα fragment consists of three domains: an N-terminal domain (NTD) that shares sequence similarity to a homologous glycoprotein nonmetastatic melanoma B (GPNMB); the polycystic kidney disease domain (PKD) with homology to a repeat region in the polycystic kidney disease-associated protein, polycystin 1 (PKD1); and the repeat domain (RPT) having a low complexity region (LCR) with little diversity in their amino acid composition and significant enrichment with specific amino acid types located

in the N-terminal (Figure 2.1a,b). RPT is an intrinsically disordered region (IDR) composed of 10 pseudo-repetitive sequence motifs enriched in threonine, proline, serine, alanine, valine, glycine, and glutamic acid residues (Figure 2.1 b). It is known that under *in vitro* conditions, RPT forms the amyloid core that promotes melanin biogenesis and forms amyloid fibrils under mildly acidic



R1 PTAEAPNTTAGQV R2 PTTEVVGTTPGQA R3 PTAEPSGTTSVQV R4 PTTEVISTAPVQM
R5 PTAESTGMTPEKV R6 PVSEVMGTTLAEM R7 STPEATGMTPAEV R8 SIVVLSGTTAAQV
R9 TTTEWVETTARER L R10 PIPEPEGPDASSI

(Hydrophobic, Polar, Negative, Positive)

Figure 2.1 Schematic representation of the primary amino acid sequence of Pmel17. (a) Proprotein convertase cleavage of full-length Pmel17 protein yields two fragments, namely, a luminal fragment Mα and a transmembrane fragment Mβ. The Mα fragment (residues 25–467) and Mβ (residues 468–668) fragments along with cleavage site (CS) are shown. The Mα fragment of Pmel17 consists of three amyloidogenic domains-NTR, PKD, and RPT. (b) RPT (315-444) contains 10 imperfect repeats, each containing 13 amino acid residues that are rich in hydrophobic (black), polar (olive), and acidic (red) residues. The basic residues are shown in blue, and a single native tryptophan (W423) is highlighted in purple. The underlined sequence shows a low complexity region (LCR) within RPT predicted using the Simple Modular Architecture Research Tool (SMART).

pH (pH 4-5.5) typical of melanosomes but remains highly soluble in the monomeric form at cytosolic pH.[55,56] The PONDR score plot indicates that RPT is predominantly unstructured and has local order in the C-terminal amyloid core region, which is likely to be due to the absence of prolines in this region (Figure 2.2a). We found that RPT, whose size is ~14 kDa, confirmed by mass spectrometry shows an anomalous migration on a 12% SDS/PAGE and migrates at approximately > 40 kDa (Figure 2.2b). Next, we recorded fluorescence spectra and anisotropy of the single tryptophan (Trp) present in the C-terminal region of RPT. Highly red-shifted emission

maxima (~ 345 nm) and low anisotropy indicated the solvent-exposed environment as expected for the natively unfolded/disordered state (Figure 2.2c). Further, we took advantage of the fact that RPT is devoid of cysteines and created two single cysteine mutations: Cys350 in N-terminus and Cys415 in the amyloid core region in the C-terminus (Figure 2.1b, highlighted in yellow) to probe

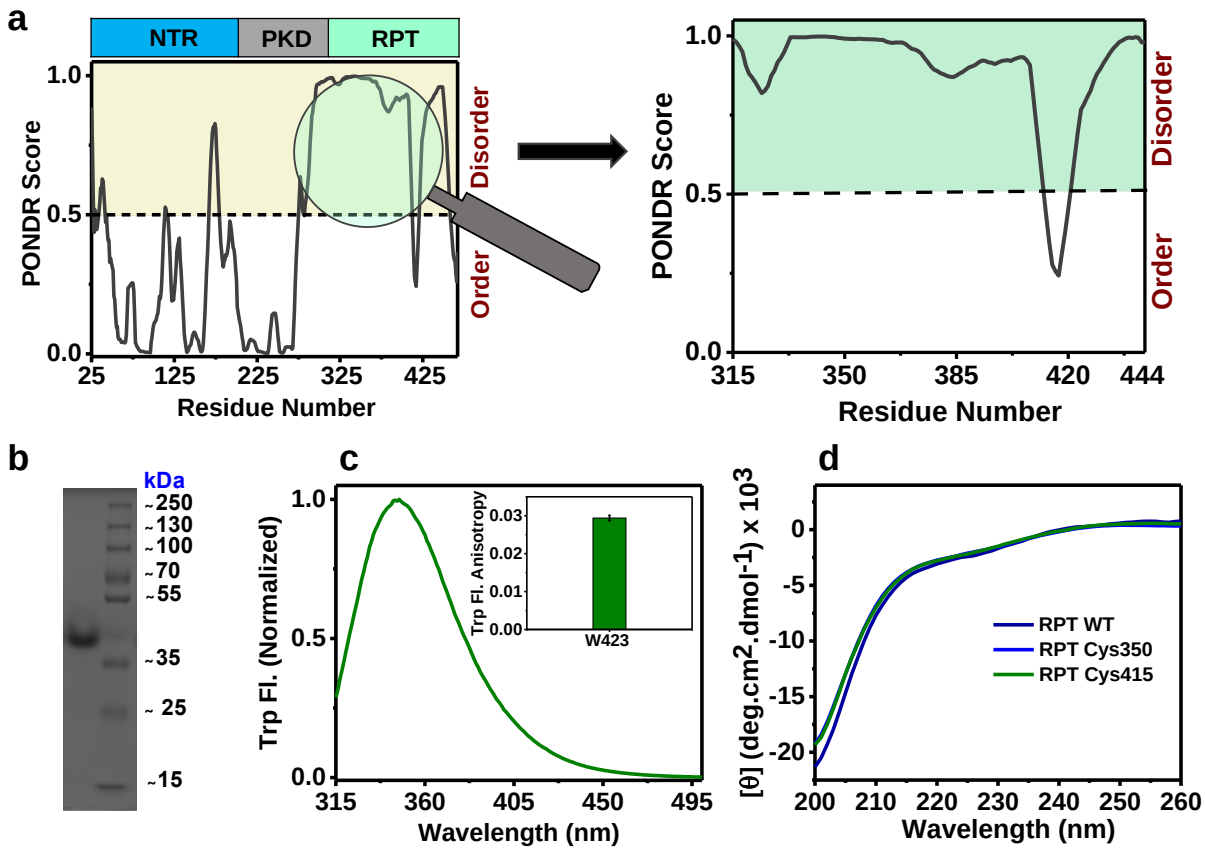


Figure 2.2 (a) Predictor of Natural Disordered Regions (PONDR)-VLXT plot of the Mα fragment showing disorder propensity of RPT. (b) SDS-PAGE (12%) analysis of purified RPT under native conditions. (c) Steady-state fluorescence spectra of single native Trp present in the C-terminal of RPT. The inset shows Trp fluorescence anisotropy of RPT in 20 mM sodium phosphate (pH 7.4) buffer, 50 mM NaCl. (d) Far-UV CD spectra for wild-type (WT) and cysteine variants (Cys350 and Cys415, highlighted in Figure 2.1 b) of monomeric RPT.

the local structural differences and solvation dynamics of the polypeptide chain. In order to ensure that there was no perturbation in the secondary structure due to mutations in disordered monomeric RPT, we performed circular dichroism (CD) spectroscopy for the mutants. The CD spectra of the mutants overlapped with the wild type RPT and showed a typical random coil signature (Figure 2.2d). We next covalently labeled the cysteines with a highly environment-sensitive fluorophore, namely acrylodan that is an ideal fluorophore for solvation dynamics studies in proteins due to its

high sensitivity towards the polarity of the medium and a large Stokes shift.[62] The emission maximum for Cys350 was found to be ~530 nm, and that of Cys415 was ~526 nm, which indicates that acrylodan is highly solvent-exposed (Figure 2.3a). The steady-state fluorescence anisotropy value in the case of Cys415 (C-terminus) was higher than Cys350 (N-terminus) (Figure 2.3b). These observations suggest a significant hydration of acrylodan by the solvent molecules as expected for the expanded IDPs. Together, the steady-state fluorescence attributes of acrylodan labeled mutants suggested that the microenvironment is slightly different for the probed residues

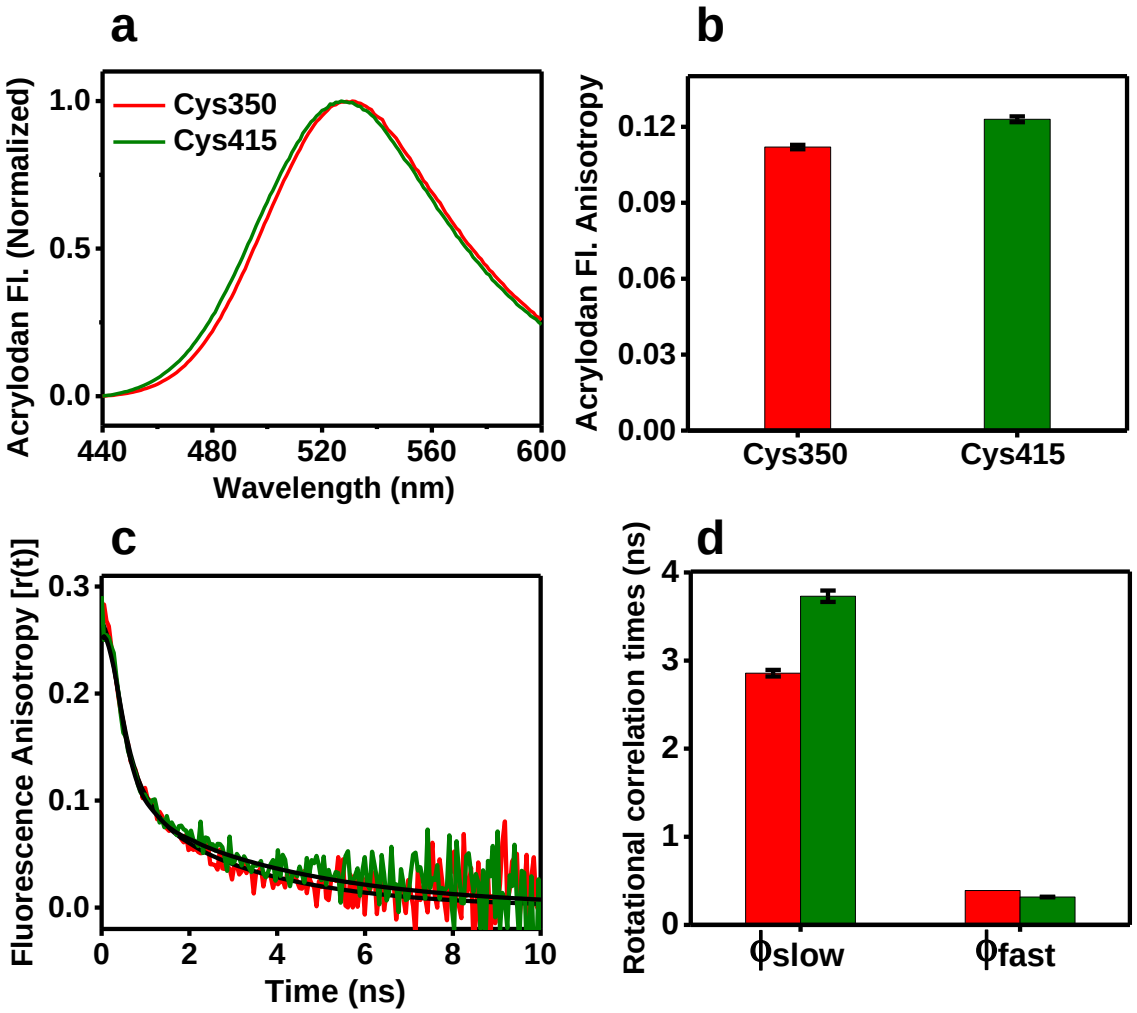


Figure 2.3 The steady-state fluorescence spectra (a) and (b) fluorescence anisotropy for acrylodan labeled cysteine positions in the monomeric state of RPT. (c) Picosecond time-resolved anisotropy decays of acrylodan labeled cysteine positions, Cys350 (red), Cys 415 (olive) of RPT. The black lines are fits using the biexponential decay kinetics. (d) Comparison of the slow and fast rotational correlation times of Cys350 (red), Cys 415 (olive).

Table 2

The typical parameters recovered from the fluorescence anisotropy decay analyses shown in Figure 2.3c:

Sample	ϕ_{fast} **(ns)**	β_{fast}	ϕ_{slow} **(ns)**	β_{slow}
Cys350	0.39	0.63	2.86	0.37
Cys415	0.32	0.64	3.73	0.36

in the monomeric state of RPT. Next, to monitor the rotational dynamics of the fluorophore (acrylodan), we performed fluorescence depolarization measurements in the monomeric state of RPT by using time-correlated single photon counting (TCSPC) technique. A typical bi-exponential fluorescence anisotropy decay represents the (fast) c motion of the fluorophore attached to the protein and the (slow) global tumbling of the entire protein.[48,49,51,60] The observed increase in the longer rotational correlation time for Cys415 is suggestive of the dampening of the segmental motion resulting from the slightly compact local structure in the amyloidogenic C-terminus (Figure 2.3c,d) corroborating PONDR results.[49,60] In addition, a significant amplitude (~ 60%) of local motion was observed, which suggests that acrylodan is flexible and is certainly exposed to solvent (Table 2). Therefore, taken together, these results revealed altered conformations in different regions of disordered monomeric RPT that might be associated with the distinct function of these regions in specific interactions during aggregation and amyloid formation. Next, in order to delineate the differences in the surface water dynamics within different regions of RPT, we embarked upon probing the water structures in the intrinsically disordered monomeric state of RPT in a site-specific manner using time-resolved fluorescence spectroscopic measurements.

2.3.2 Solvation dynamics in the intrinsically disordered RPT probed by TCSPC and femtosecond up-conversion

In order to directly monitor the reorientation dynamics of water, we have utilized picosecond time-resolved spectroscopy. This technique allows us to probe the dielectric response of proteins as well as the organization and dynamics of water molecules around a fluorophore that is either part of the protein or is covalently attached to it.[63] When a fluorophore is excited to higher vibrational energy levels of the first excited S_1 state, there is a rapid loss of excess vibrational energy to the surround-

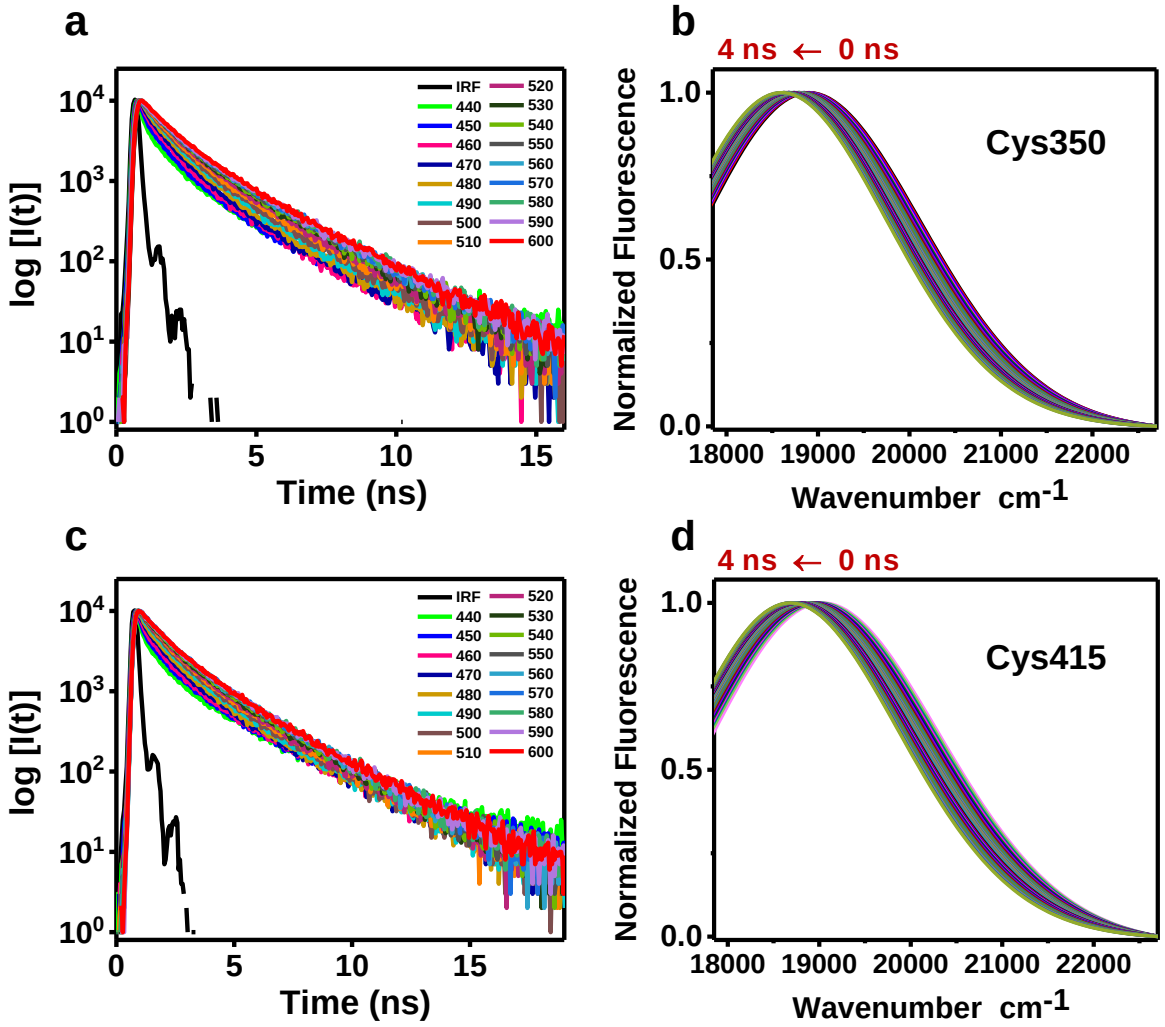


Figure 2.4 (a) Time-resolved fluorescence decays of acrylodan at Cys350 of RPT obtained at different emission wavelengths from 440 to 600 nm. (b) Time-resolved emission spectra (TRES) constructed from the time-resolved fluorescence decays at different wavelengths. The arrow indicates the total time-dependent fluorescence Stokes shift (TDFSS) within 4 ns. (c) Time-resolved fluorescence decays of acrylodan at Cys415 of RPT obtained at different emission wavelengths from 440 to 600 nm. (d) Time-resolved emission spectra (TRES) constructed from the time-resolved fluorescence decays at different wavelengths. The arrow indicates the total TDFSS within 4 ns.

-ing solvent molecules as the fluorophore slowly relax to the lowest vibrational energy state. The solvent molecules further lower the energy level of the excited state by re-orienting via hydrogen bond rearrangement around the excited fluorophore via a process called solvent relaxation.[64] This reorganization of solvent molecules around the instantaneously created excited state dipole of the fluorophore results in an emission shift to longer wavelength and this shift in the emission spectrum that occurs during reorganization process is termed as time-dependent fluorescence

Stokes shift (TDFSS).[64] The extent of this shift is an indicator of how much solvent relaxed the fluorophore is with respect to the initial unrelaxed state that is in turn dependent on the organization of water molecules and the rate of hydrogen bond rearrangements around the fluorophore. To assess TDFSS and the solvation timescales, we constructed time-resolved emission spectra (TRES) from the collected time-resolved fluorescence decays at different emission wavelengths encompassing the entire spectral range of the acrylodan fluorescence. The TRES profiles for both Cys350 and Cys415 were indistinguishable and did not exhibit significant TDFSS (Figure 2.4b,d). The insignificant (230-280 cm^{-1}) TDFSS from TCSPC TRES suggests that there is very little/no nanosecond solvation component. Since the differences in the water structure within RPT were inaccessible by picosecond/nanosecond measurements, we next embarked upon the solvation dynamics studies using the femtosecond up-conversion technique. These measurements allow us to distinguish and estimate the time scale of solvent relaxation of bulk water and surface-bound biological water with femtosecond resolution.[39,41,51,52] Next, we collected the femtosecond-resolved fluorescence decays as a function of wavelength across the emission spectrum of acrylodan and constructed TRES to estimate the TDFSS. Figure 2.5a,c shows four representative femtosecond transients spanning blue to red edge of the emission spectra. We observed a decay component in the fluorescence signal on the blue edge and a rise component having negative on the red edge which is a typical signature of solvent relaxation.[66] Further, we estimated the TDFSS (see Experimental Section for the details of data analysis) for the acrylodan-labeled Cys350 and Cys415 that were found to be ~1927 cm^{-1} and ~1972 cm^{-1}, respectively. A large number of theoretical and experimental investigations have revealed that the bulk water solvent relaxation occurs on the timescale of 100 femtoseconds to picoseconds,[39] whereas, "biological water" relaxation time spans from 10 and 100 ps.[38,39] We next analyzed the solvation correlation functions [C(t)][39,66,67] and estimated the solvent relaxation times. The recovered parameters of solvation dynamics and the C(t) for Cys350 and 415 are shown in Figure 2.6a,b. Here, the ultrafast component (~1 ps) represents the bulk water, and the slower component (of tens of picosecond) represents "biological water". The solvation correlation plots revealed that the solvation was nearly 95% complete with respect to $\nu(\infty)$ in the Cys350 (N-terminus) and was found to be almost 100% complete in Cys415 (C-terminus) within 200 ps (Figure 2.6a,b). Similar timescale of bulk water relaxation is indicative of homogeneous solvation of the polypeptide chain for both of the mutants.

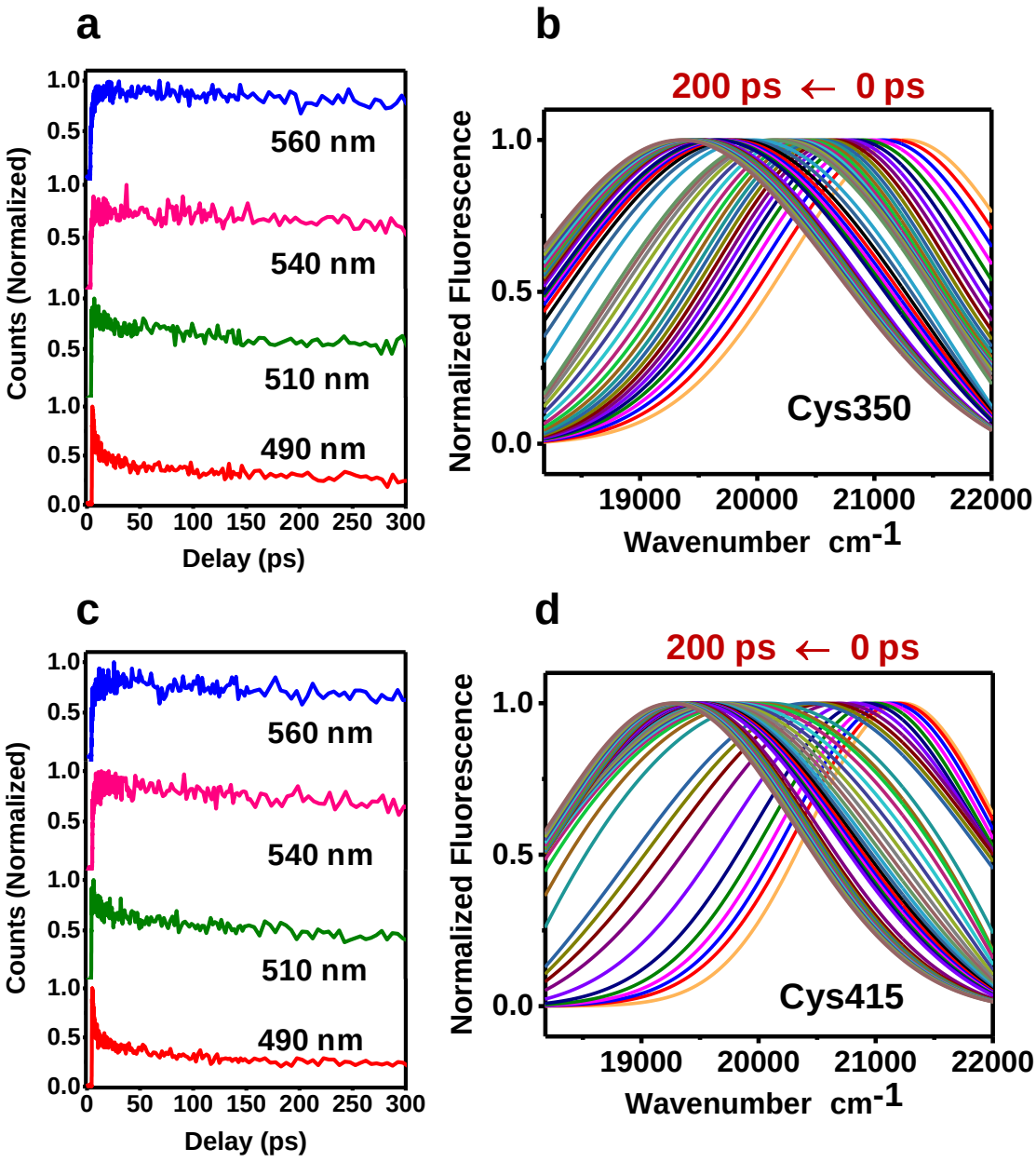


Figure 2.5 (a) Femtosecond transients for monomeric acrylodan-Cys350 of RPT at four emission wavelengths: 490 nm (red), 510 nm (olive), 540 nm (magenta), and 560 nm (blue). (b) Time-resolved emission spectra (TRES) of monomeric RPT from 0 to 200 ps constructed from the time-resolved fluorescence decay at different wavelengths. The arrow indicates the total TDFSS within 200 ps. (c) Femtosecond transients for monomeric acrylodan-Cys415 of RPT at four emission wavelengths: 490 nm (red), 510 nm (olive), 540 nm (magenta), and 560 nm (blue). (b) Time-resolved emission spectra (TRES) of monomeric RPT from 0 to 200 ps constructed from the time-resolved fluorescence decay at different wavelengths. The arrow indicates the total TDFSS within 200 ps.

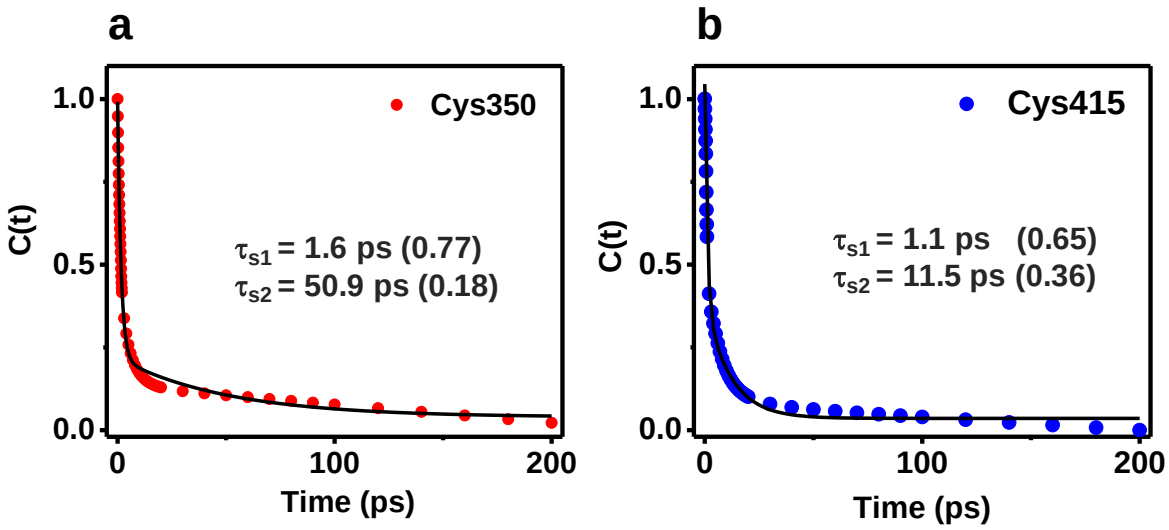


Figure 2.6 Solvation correlation function [C(t)] plots are shown for both (a) Cys350 (red filled circle) and (b) Cys415 (blue filled circle) from 0 to 200 ps. The black lines are fits using the biexponential decay kinetics. The recovered solvation times are τs_1 = 1.58 ns, amplitude = 0.77 and τs_2 = 50.9 ns, amplitude = 0.18 for Cys350 and τs_1 = 1.1 ns, amplitude = 0.65 and τs_2 = 11.5 ns, amplitude = 0.36 for Cys415. See Experimental Section for data analysis.

A careful look at the recovered parameters indicated the differences in the amplitude/contribution of solvation times for bulk water as well as "biological water" at distinct Cys positions present in the N- and C-terminus. Additionally, the difference in the timescale and contribution of bulk water and "biological water" relaxation are indicative of some unique conformational differences within distinct regions of RPT. Taken together, our results from the TCSPC and femtosecond up-conversion experiments revealed an increase in the contribution of "biological water" in the case of Cys415 that corroborate our fluorescence depolarization results and reveal more ordering of water molecules around local collapsed structure indicating the crucial role of ordered water molecules in the amyloid core region during the self-association of the polypeptide chains to form amyloid fibrils via entropic release of the water molecules.

2.4 Conclusion

In this work, we elucidate the behavior of water in intrinsically disordered RPT that forms the amyloid core of the Pmel17 protein, involved in the intracellular organization of melanosomes and melanin biogenesis. According to our results, the C-terminal proline-free amyloid core region is anticipated to have a slightly compact local structure. Additionally, by probing the water structure

in different regions of RPT, we have been able to identify the plausible differences in the secondary structure propensities within different regions. Our ultrafast measurements revealed the presence of ordered water molecules in the C-terminal amyloid core region of RPT. The specificity and selectivity of different domains in RPT amyloid formation may be tailored to the differences in the water organization. Cys415 (C-terminus) exhibited faster-bound water relaxation dynamics compared to Cys350 (N-terminus). However, the contribution of bound water relaxation dynamics was more in the case of Cys415 than Cys350. The difference in the contribution of the bound water solvation time can be explained to a certain extent based on the amino acid composition, which is appreciably different in the case of C-terminus due to the proline free stretch. On the contrary, there is a significant number of proline residues in N-terminus. The proline residues are known to participate in strong hydrogen bonding with water[68], and therefore, this might account for the lower contribution of the bound water solvation time in the case of Cys350. Together, our findings provide novel insights into the potential role of water in RPT aggregation and amyloid formation. We propose that the contrasting nature of water molecules within different regions of RPT might have significant biological consequences, such as in specific ion-protein interactions, protein-protein interactions, and amyloid formation.

2.5 References:

(1) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V.N. Introducing Protein Intrinsic Disorder. *Chem. Rev.* **2014,** 114, 6561−6588.

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

(3) Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008,** *18*, 756-764.

(4) Kriwacki, R. W.; Hengst, L.; Tennant, L.; Reed, S. I.; Wright, P. E. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93*, 11504-11509.

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

(6) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Structure, Function, and Bioinformatics* **2000,** *41*, 415-427.

(7) 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.

(8) 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. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001,** *19*, 26-59.

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

(10) Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and Structural Disorder in Protein-Protein Interactions. *Trends Biochem. Sci.* **2008**, *33*, 2–8.

(11) Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004,** *337*, 635-645.

(12) Fink, A. L. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* **2005,** *15*, 35-41.

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

(14) Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **2012,** *37*, 509-516.

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

(16) 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. *Intrinsically Disordered Proteins* **2013,** *1*, e24157.

(17) Burger, M. V.; Gurry, T.; Stultz, M. C. Intrinsically disordered proteins: Where computation meets experiment. *Polymers* **2014,** *6*, 2684-2719.

(18) 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*, 384-396.

(19) Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.* **2014,** *114*, 6844-6879.

(20) Tycko, R.; Wickner, R. B. Molecular Structures of Amyloid and Prion Fibrils: Consensus versus Controversy. *Acc. Chem. Res.* **2013,** *46*, 1487-1496.

(21) Eisenberg, D.; Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **2012,** *148*, 1188-1203

(22) Mao, Albert H.; Lyle, N.; Pappu, Rohit V. Describing sequence–ensemble relationships for intrinsically disordered proteins. *Biochem. J.* **2012,** *449*, 307.

(23) 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.

(24) Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2649–2654.

(25) Jain, N.; Bhattacharya, M.; Mukhopadhyay, S. Chain Collapse of an Amyloidogenic Intrinsically Disordered Protein. *Biophys. J.* **2011**, *101*, 1720–1729.

(26) Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence Correlation Spectroscopy Shows that Monomeric Polyglutamine Molecules Form Collapsed Structures in Aqueous Solutions. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16764–16769.

(27) 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*, 8183–8188.

(28) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14609–14614.

(29) Bagchi, B. Water in Biological and Chemical Processes: From Structure and Dynamics to Function. Cambridge University Press: **2013**.

(30) Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A Unified Model of Protein Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 5129-5134.

(31) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008,** *108*, 74-108.

(32) Chaplin, M. Do We Underestimate the Importance of Water in Cell Biology? *Nat. Rev. Mol. Cell Biol.* **2006,** *7*, 861-866.

(33) Bagchi, B. Water Dynamics in the Hydration Layer around Proteins and Micelles. *Chem. Rev.* **2005,** *105*, 3197-3219.

(34) Pal, S. K.; Zewail, A. H. Dynamics of Water in Biological Recognition. *Chem. Rev.* **2004,** *104*, 2099-2124.

(35) Bhattacharyya, K.; Bagchi, B. Slow Dynamics of Constrained Water in Complex Geometries. *J. Phys. Chem. A* **2000,** *104*, 10603-10613.

(36) Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A unified model of protein dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 5129-5134.

(37) Pethig, R. Protein-water interactions determined by dielectric methods. *Annu. Rev. Phys. Chem.* **1992,** *43*,177-205.

(38) Nandi, N.; Bagchi, B. Dielectric relaxation of biological water. *J. Phys. Chem. B* **1997,** *101*, 10954-10961.

(39) Pal, S. K.; Peon, J.; Zewail, A. H. Biological Water at the Protein Surface: Dynamical Solvation Probed Directly with Femtosecond Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1763–1768.

(40) Zhong, D.; Pal, S. K.; Zewail, A. H. Biological Water: A Critique. *Chem. Phys. Lett.* **2011,** *503*, 1-11.

(41) Qin, Y.; Wang, L.; Zhong, D. Dynamics and Mechanism of Ultrafast Water–protein Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 8424-8429.

(42) Chong, S. H.; Ham, S. Distinct Role of Hydration Water in Protein Misfolding and Aggregation Revealed by Fluctuating Thermodynamics Analysis. *Acc. Chem. Res.* **2015,** *48*, 956-965.

(43) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012,** *45*, 83-92.

(44) Pavlova, A.; Cheng, C. Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S. Protein Structural and Surface Water Rearrangement Constitute Major Events in the Earliest Aggregation Stages of Tau. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113*, E127-36.

(45) Schwierz, N.; Frost, C. V.; Geissler, P. L.; Zacharias, M. Dynamics of Seeded Abeta40-Fibril Growth from Atomistic Molecular Dynamics Simulations: Kinetic Trapping and Reduced Water Mobility in the Locking Step. *J. Am. Chem. Soc.* **2016,** *138*, 527-539.

(46) Fichou, Y.; Schiro, G.; Gallat, F. X.; Laguri, C.; Moulin, M.; Combet, J.; Zamponi, M.; Hartlein, M.; Picart, C.; Mossou, E.; Lortat-Jacob, H.; Colletier, J. P.; Tobias, D. J.; Weik, M. Hydration Water Mobility is Enhanced Around Tau Amyloid Fibers. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112*, 6365-6370.

(47) Despa, F.; Fernandez, A.; Scott, L. R.; Berry, R. S. Hydration Profiles of Amyloidogenic Molecular Structures. *J. Biol. Phys.* **2008,** *34*, 577-590.

(48) Arya, S.; Mukhopadhyay, S. Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein. *J. Phys. Chem. B* **2014,** *118*, 9191-9198.

(49) Jain, N.; Bhattacharya, M.; Mukhopadhyay, S. Chain Collapse of an Amyloidogenic Intrinsically Disordered Protein. *Biophys. J.* **2011,** *101*, 1720-1729.

(50) Dalal, V.; Arya, S.; Mukhopadhyay, S. Confined Water in Amyloid-Competent Oligomers of the Prion Protein. *Chemphyschem* **2016**, *17*, 2804-2807.

(51) Arya, S.; Singh, A. K.; Khan, T.; Bhattacharya, M.; Datta, A.; Mukhopadhyay, S. Water Rearrangements upon Disorder-to-Order Amyloid Transition. *J. Phys. Chem. Lett.* **2016,** *7*, 4105-4110.

(52) Arya, S.; Singh, A. K.; Bhasne, K.; Dogra, P.; Datta, A.; Das, P.; Mukhopadhyay, S. Femtosecond Hydration Map of Intrinsically Disordered α-Synuclein. *Biophys. J.* **2018,** 114, 2540–2551.

(53) Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone−Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008,** *130*, 7380-7392.

(54) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional Amyloid Formation within Mammalian Tissue. *PLoS Biol.* **2006,** 4, e6.

(55) McGlinchey, R. P.; Shewmaker, F.; McPhie, P.; Monterroso, B.; Thurber, K.; Wickner, R. B. The Repeat Domain of the Melanosome Fibril Protein Pmel17 Forms the Amyloid Core Promoting Melanin Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** 106, 13731−13736.

(56) Dogra, P.; Bhattacharya, M.; Mukhopadhyay, S. pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid. *J. Phys. Chem. B* **2017,** 121, 412−419.

(57) Watt, B.; van Niel, G.; Raposo, G.; Marks, M. S. PMEL: A Pigment Cell-Specific Model for Functional Amyloid Formation. *Pigm. Cell Melanoma Res.* **2013,** 26, 300−315.

(58) McGlinchey, R. P.; Jiang, Z.; Lee, J. C. Molecular Origin of pH Dependent Fibril Formation of a Functional Amyloid. *ChemBioChem* **2014**, 15, 1569−1572.

(59) Hu, K. N.; McGlinchey, R. P.; Wickner, R. B.; Tycko, R. Segmental Polymorphism in a Functional Amyloid. *Biophys. J.* **2011,** 101, 2242−2250.

(60) Majumdar, A.; Mukhopadhyay, S. Fluorescence Depolarization Kinetics to Study the Conformational Preference, Structural Plasticity, Binding, and Assembly of Intrinsically Disordered Proteins. *Methods Enzymol.* **2018,** 611, 347−381.

(61) Fee, R. S.; Maroncelli, M. Estimating the Time-Zero Spectrum in Time-Resolved Emission Measurements of Solvation Dynamics. *Chem. Phys.* **1994,** *183*, 235-247.

(62) Prendergast, F. G.; Meyer, M.; Carlson, G. L.; Iida, S.; Potter, J. D. Synthesis, spectral properties, and use of 6-acryloyl-2-dimethylaminonaphthalene (Acrylodan). A thiol-selective, polarity-sensitive fluorescent probe. *J. Biol. Chem.* **1983,** *258*, 7541-7544.

(63) Changenet-Barret, P.; Choma, C. T.; Gooding, E. F.; DeGrado, W. F.; Hochstrasser, R. M. Ultrafast Dielectric Response of Proteins from Dynamics Stokes Shifting of Coumarin in Calmodulin. J. Phys. Chem. B 2000, 104, 9322-9329.

(64) Lakowicz, J. R. Principles of Fluorescence Spectroscopy. Springer US: **2007**.

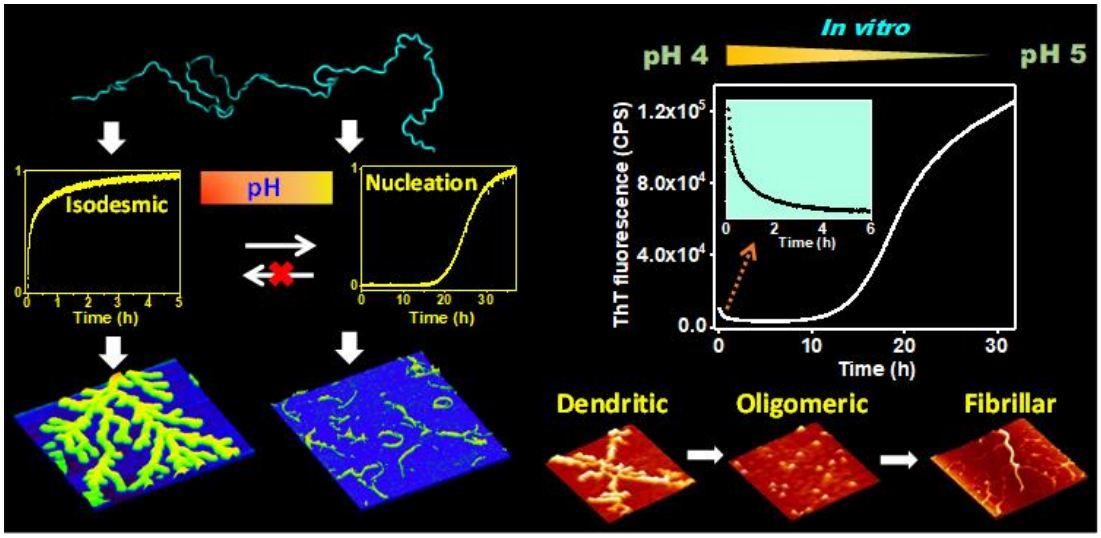
(65) Jha, A.; Ishii, K.; Udgaonkar, J. B.; Tahara, T.; Krishnamoorthy, G. Exploration of the Correlation between Solvation Dynamics and Internal Dynamics of a Protein. *Biochemistry* **2011,** *50*, 397-408.

(66) Kamal, J. K.; Zhao, L.; Zewail, A. H. Ultrafast Hydration Dynamics in Protein Unfolding: Human Serum Albumin. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 13411-13416.

(67) Maroncelli, M.; Fleming, G. R. *J. Chem. Phys.* **1987,** *86*, 6221.

(68) McLain, S. E.; Soper, A. K.; Terry, A. E.; Watts, A. Structure and Hydration of l-Proline in Aqueous Solutions. *J. Phys. Chem. B* **2007,** *111*, 4568-4580.

Mechanistic Insights into the Mechanism-to-Morphology Relationship in Amyloid Formation of a Pmel17 Fragment (Mα) and RPT



The work described in this chapter has been published in the *Journal of Physical Chemistry*.

Reference: Dogra, P.; Bhattacharya, M.; Mukhopadhyay, S. pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid. *J. Phys. Chem. B* **2017,** 121, 412−419.

3.1 Introduction

Protein misfolding and aggregation has been associated with a plethora of human diseases such as Parkinson's, Alzheimer's etc. A variety of factors are involved in the misfolding and aggregation of proteins viz. mutation in the polypeptide chain or changes in solution conditions such as ionic strength, pH, and temperature. During the aggregation process the protein goes to a thermodynamically more stable state known as the amyloid state, composed of cross-β sheet rich architecture.[1-5] Deposition of amyloids are known to be characteristic hallmarks of several neurodegenerative disorders.[1,5-8] However, the beneficial role of amyloids in a multitude of physiological functions is now gaining prominence.[9–16] This emerging class of amyloids is termed as functional amyloids, which challenge the common belief that amyloids are cytotoxic disease-causing agents.[10,15] The manifestations of functional amyloids are observed across various organisms ranging from bacteria to humans.[10] For instance, biofilm-forming curli in E. coli,[12] yeast prions in fungi,[16] peptide amyloid hormones in the endocrine system,[11]. In humans, Pmel17, a melanocyte-specific glycoprotein, forms functional amyloid that plays an essential role in melanosome development by creating a fibrillar amyloid matrix in the organelle, which acts as a template for melanin deposition underneath the skin and in the eyes.[9] The amyloid matrix serves a beneficial role in mitigating the toxicity by sequestering and minimizing the diffusion of highly reactive quinone precursors that are required during melanin biosynthesis. It is known that during melanosome maturation, the pH environment changes from more acidic to neutral with different stages of melanosome development. Stage I is acidic, composed of short fibrils that mature into fibrillar sheets in stage II, and by melanized stage III/IV, the pH reaches to neutral.[17-21] Pmel17 is a 668-residue polypeptide that undergoes multiple sequential proteolytic processing steps to generate a highly amyloidogenic luminal fragment (Mα) and a transmembrane fragment (Mβ). Further, the luminal fragment comprises an intrinsically disordered region (IDR), the repeat domain (RPT), that aggregates under mildly acidic conditions, typical of melanosomes.[17-21] It is known that RPT promotes melanin formation in vitro and has other properties in common with Pmel17 filaments of melanosomes. Several studies have shown that the deletion of the RPT ablates fibril formation in vivo.[21] However, the molecular mechanism of amyloid formation, as well as the organization of individual protein molecules within the supramolecular assembly, remains elusive. Both pathological and functional amyloids are believed to share some common structural features; however, the underlying molecular principles that distinguish the (toxic) pathological amyloids

from the (benign) functional amyloids remain elusive.[22-29] In addition to the roles in diseases and functions, amyloids represent exquisite proteinaceous nanomaterials for a variety of future bionanotechnological applications.[25-28] Decoding the amyloid formation mechanism has inspired the advent and utility of amyloid-based functional nanomaterials in recent years.[25-28] To design novel functional nanomaterials as well as to strategically control the desired nanoscale morphology, it is important to establish a link between the mechanism and morphology. The mechanism of amyloid aggregation is broadly classified into two types: nucleation-dependent and nucleation-independent isodesmic polymerization.[30-37] The nucleation-dependent assembly process comprises a lag phase that can be bypassed using preformed amyloid seeds. In contrast, the isodesmic mechanism is devoid of a lag phase and exhibits a monotonic growth kinetics. Such competing pathways as a function of pH have been shown to govern the fibril morphology.[32-34] Amyloid fibrils from a variety of proteins are shown to display nanoscale polymorphism.[38-41] In this work, we present a unique case of controlling the formation of dendritic and fibrillar nanostructures by switching the aggregation mechanism of functional amyloid derived from the amyloidogenic segment of the melanosomal protein. We demonstrate that the repeat domain of the melanosomal protein exhibits two distinct types of aggregation pathways that display nanoscale polymorphism in acidic pH. Rapid aggregation without a lag phase at lower pH would likely to indicate that there is little or no accumulation of toxic oligomeric species, which are efficiently recruited during the course of aggregation. We speculate that the transition in the aggregation mechanism could occur due to the protonation/deprotonation of glutamic acid residues, which has been previously shown to influence the course of aggregation. These dendritic nanostructures can be converted into more stable fibrils by switching the pH. The nanoscale polymorphism associated with the mechanistic switch is likely to be mediated by the altered conformational propensities and intermolecular interactions due to the protonation/deprotonation of critical glutamate residues. We suggest that the pH modulation within the melanosomes allows the optimal conditions for the formation of functional amyloids that dictate the template-assisted melanin biosynthesis. In conclusion, our findings address key molecular aspects that relate the protein aggregation mechanism with the nanoscale polymorphism and reveal a morphological change reminiscent of the amyloid transition observed between stage I and stage II melanosomes.

3.2 Experimental Section

3.2.1 Materials

Guanidinium hydrochloride and urea were purchased from Ameresco. Isopropyl-thiogalactopyranoside (IPTG) was obtained from Gold Biocom. 8-Anilinonaphthalene-1-sulfonic acid (ANS), thioflavin T (ThT), dihydroxyphenylalanine (DOPA), tyrosinase (from mushroom) and other chemicals used for buffer preparation were purchased from Sigma (St. Louis, MO) and used as received.

3.2.2 Expression and purification of the Mα fragment of Pmel17

The Mα fragment of Pmel17 (25-467) with C-terminal his-tag present in pET3c vector was expressed in E. coli BL21 (DE3) pLysS cells/strain and purified using the protocol as described previously with slight modifications.[9] Briefly, for 1 L culture, the transformed cells were grown at 37 °C with shaking to an optical density at 600 nm (A_{600}) equal to 0.4−0.5 and then induced using 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C. The cells obtained after centrifugation (4000 rpm) for 60 min at 10 °C were dissolved in 25 mL of TBS (50 mM tris-base and 150 mM NaCl, pH 7.50) buffer and frozen at -80 °C. The cell suspension was then lysed using probe sonication (Amplitude:15 Pulse on-time: 15 s Pulse off-time: 10 s Process time: 30 min). The sonicated suspension was centrifuged at 11500 rpm, 30 min, 10 °C. The inclusion bodies obtained were washed twice using buffer containing 1.5 M NaCl, 50 mM KH_2PO_4/K_2HPO_4 (pH 7.4), 1% Triton-X 100, and the homogeneous suspension was then centrifuged at 11500 rpm, 30 min, 10 °C. Similar washing step was performed by using the TBS buffer. Finally, the inclusion body pellet was dissolved in the extraction buffer containing 8 M urea, 50 mM KH_2PO_4/K_2HPO_4 (pH 7.4), 100 mM KCl, 5 mM EDTA, and the solution was kept at 4 °C on stirring for 48 h. The resulting solution was centrifuged, filtered through a 0.22 μM syringe filter (Merck Millipore) and was frozen at -80 °C. After thawing, the protein was purified by an anion exchange chromatography using 8 M urea, 50 mM KH_2PO_4/K_2HPO_4 (pH 7.4), 5 mM EDTA and the flow through obtained was further purified using gel filtration chromatography using 6 M GdmCl, 50 mM K_2HPO_4/KH_2PO_4, 100 mM KCl, 5 mM EDTA, pH 7.4 with a Superdex 200 26/60 column. The molar extinction coefficient used for estimating the protein concentration was 80400 $M^{-1}cm^{-1}$. Finally, purified Mα in 6 M GdmCl, 50 mM K_2HPO_4/KH_2PO_4, 100 mM KCl, 5 mM EDTA, pH 7.4, was stored at -20 °C for further use.

3.2.3 Expression and purification of the repeat domain (RPT) of Pmel17

The repeat domain of Pmel17 (RPT) (315-444) with C-terminal his-tag present in pET-21a (+) vector was expressed in E. coli BL21 (DE3) RIPL cells/strain (Agilent Technologies, Stratagene) and purified using the protocol that is described in great detail in chapter 2.

3.2.4 Steady-state fluorescence measurements

RPT aggregation kinetics was recorded on a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon, NJ) at room temperature. All of the measurements were carried out in a quartz cuvette (10 × 10 mm) with a reaction volume of 1.5-2.0 mL. For recording ThT fluorescence, the samples were excited at 440 nm and the emission was collected at 480 nm. For tryptophan fluorescence studies, the samples were excited at 280 nm and the spectra were collected in the emission range between 310 and 410 nm. For ANS fluorescence, the samples were excited at 375 nm and the emission was collected at 475 nm. The concentration of ANS was 20 μM. The steady-state fluorescence anisotropies were measured at 525 nm. The steady-state fluorescence anisotropy (r_{ss}) is given by the following equation:

$$r_{ss} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \qquad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are the fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor.

3.2.5 Aggregation studies

For setting up the aggregation reaction of the Mα fragment, the concentrated protein solution was fast diluted using 50 mM of sodium acetate buffer containing 100 mM KCl to a final concentration of 5 μM. The resultant solution was kept at 37 °C and the reaction was monitored under quiescent conditions. For setting up RPT aggregation reactions, 25 μM of purified RPT was added to 20 mM sodium citrate buffer, pHs 3, 4, and 5 containing 150 mM NaCl. The temperature of the reaction was maintained at 25°C and the reaction was continuously stirred at 600 rpm using a small magnetic bead. The reaction was monitored by using different fluorescent probes. The parameters for steady-state fluorescence measurements were same as mentioned above. The pH titration studies were performed on a POLARstar Omega plate reader spectrophotometer (BMG LABTECH, Germany) using 96-well plates. The sample volume used for pH titration studies was 200 μL, and reactions were agitated using tiny glass beads (3 mm diameter; Fisher Scientific). The final ThT concentration used for all of the experiments was 20 μM. For the pH range between 4 and 5, the buffer used was 20 mM sodium

citrate with 150 mM NaCl, whereas for pH 6, 20 mM MES buffer containing 150 mM NaCl was used. The aggregation reactions were repeated at least three times to determine the kinetic parameters. All of the isodesmic aggregation reactions were fitted using a biexponential function with k_1 and k_2 as rate constants (see Table 3). The nucleation-dependent aggregation reactions with lag time (t_{lag}) and aggregation half-time ($t_{1/2}$) were fitted using the nucleation-dependent polymerization model (Table 3). For pH transfer studies (pH 4 → pH 5), the fibrils at pH 4 were concentrated using a 3 kDa AMICON filter (Millipore) and then transferred to a solution of pH 5 such that the final protein concentration remains 25 μM and the pH of the resultant solution is 5. A similar type of protocol was followed in studying the switch from pH 5 to pH 4. For the steady-state tryptophan fluorescence anisotropy, the samples were excited at 280 nm and the anisotropy was measured at 330 nm. The fluorescence data at different points were then plotted and fitted with the following equation described previously for nucleation-dependent polymerization.[42]

$$y = y0 + \frac{a}{1+\exp(-\frac{x-x0}{b})} \qquad (2)$$

where y is the fluorescence at time x, $y0$ is the initial fluorescence value, $x0$ is the time when fluorescence reaches 50% of its maximum value, and a is the maximum fluorescence at stationary phase. The lag time is given by $(x0-2b)$.

The data analysis and plotting were performed using Origin 8.5 software.

3.2.6 Atomic force microscopy (AFM)

AFM images of RPT aggregates were acquired on an Innova atomic force microscope (Bruker). The AFM was operated in the tapping mode. For imaging, the silicon nitride cantilever probe with radius ~8 nm was used. Aliquots of 20 μL were withdrawn from the reaction mixture at regular intervals and were deposited on freshly cleaved, buffer-washed muscovite mica (Grade V-4 mica from SPI, PA). The samples were allowed to incubate for 5 min at room temperature and were then washed with 40 μL of the aggregation buffer. Furthermore, the samples were kept under a gentle stream of nitrogen for 5 min prior to scanning. All of the buffers used for preparing the samples were filtered through a 0.22 μm membrane filter (Millipore). The images were collected in NanoDrive (v8.03) software at a resolution of 1024 × 1024 pixels. The collected AFM images were further processed and analyzed using WSxM version 4 software.[43]

3.2.7 Circular dichroism (CD) spectroscopy

The far-UV CD spectra were recorded on a Chirascan spectrophotometer (Applied Photophysics, U.K.) in a 1 mm path length quartz cell with a scan range of 200−260 nm and step size of 1 nm. The concentration of the Mα fragment used for the experiments was 5 μM and of RPT was 10−15 μM. The spectra were averaged over 3 scans and were corrected against the buffer signal using the ProData software provided with the CD instrument. The spectra were plotted using Origin 8.5 software.

3.2.8 Dynamic light scattering

Dynamic light scattering measurements were carried out on a Malvern Zetasizer NanoZS90 instrument (Malvern, U.K.) at room temperature. A He−Ne laser (632 nm) was used for exciting the samples. All of the buffers, 20 mM sodium citrate and 150 mM NaCl (pHs 3−5), were filtered through 0.02 μm filters (Anatop 10 filter; Whatman). The concentration of the RPT used for these studies was 30 μM. The measurements were performed thrice, and the time taken for each measurement was ∼15 min. The mean hydrodynamic diameter and the size distributions were plotted using Origin 8.5 software.

3.2.9 Raman spectroscopy

Raman spectra of the samples were recorded on an inVia Raman microscope (Renishaw, UK). The aggregates were deposited on a glass slide covered with aluminium foil and dried using a gentle stream of nitrogen. A HPNIR laser with 785 nm wavelength was used as an excitation source and focused into the sample spot using a 100 × objective lens (Nikon, Japan). A 1200 l per mm grating was used for the dispersion of the scattered light and the signal was detected using an air-cooled CCD detector. The data acquisition was done using Wire 3.1 software provided with the Raman spectrometer. All the spectra were averaged over 50 scans with an exposure time of 10 s for the spectral range 400-1200 cm^{-1}, 1200-1600 cm^{-1} and 1600-1800 cm^{-1}. All the data were corrected for tilt in the baseline using the cubic spline interpolation method and smoothened in Wire 3.1 software. The baseline corrected and smoothened Raman spectra were plotted in Origin 2018 software. The deconvolution of Raman spectra was performed in Origin 2018 software as described in our previous publication.[44]

3.2.10 Synthetic melanogenesis.

The ability of Pmel17 RPT fibrils to hasten melanin biosynthesis was evaluated using a modified protocol.[9] Briefly, 2 mg of lyophilized RPT fibrils was added to 5 mM L-3,4-

dihydroxyphenylalanine (DOPA) in 4 mL of 125 mM sodium citrate buffer, pH 5.0. Tyrosinase (16 µL from 2 mg/mL stock) was added to initiate the reaction. The reaction was incubated at room temperature for 24 h. The reaction was stopped by centrifugation (15,000 rpm, 20 min) at different time points. The pellets, which contain essentially all of the melanin product, were resuspended in 1 M NaOH and heated at 60°C for 5 min. Absorbance was recorded at 350 nm since melanin absorbs at this wavelength.

3.3 Results

3.3.1 Insights into the amyloid formation of Pmel17 Mα fragment

The 442-residue long luminal fragment of Pmel17, namely Mα, was purified in denaturing buffer to preserve the non-aggregated state (Figure 3.1a,b,c). To monitor the amyloid formation of Mα fragment, we used Thioflavin T, a well-known amyloid marker. Dilution of Mα fragment into mildly acidic buffers that mimics the acidic environment of melanosomes, resulted in exceedingly rapid amyloidogenesis (Figure 3.1d). Next, we recorded fluorescence spectra and anisotropy of tryptophan (Trp) present in the Mα fragment. In the denatured state, a red shift

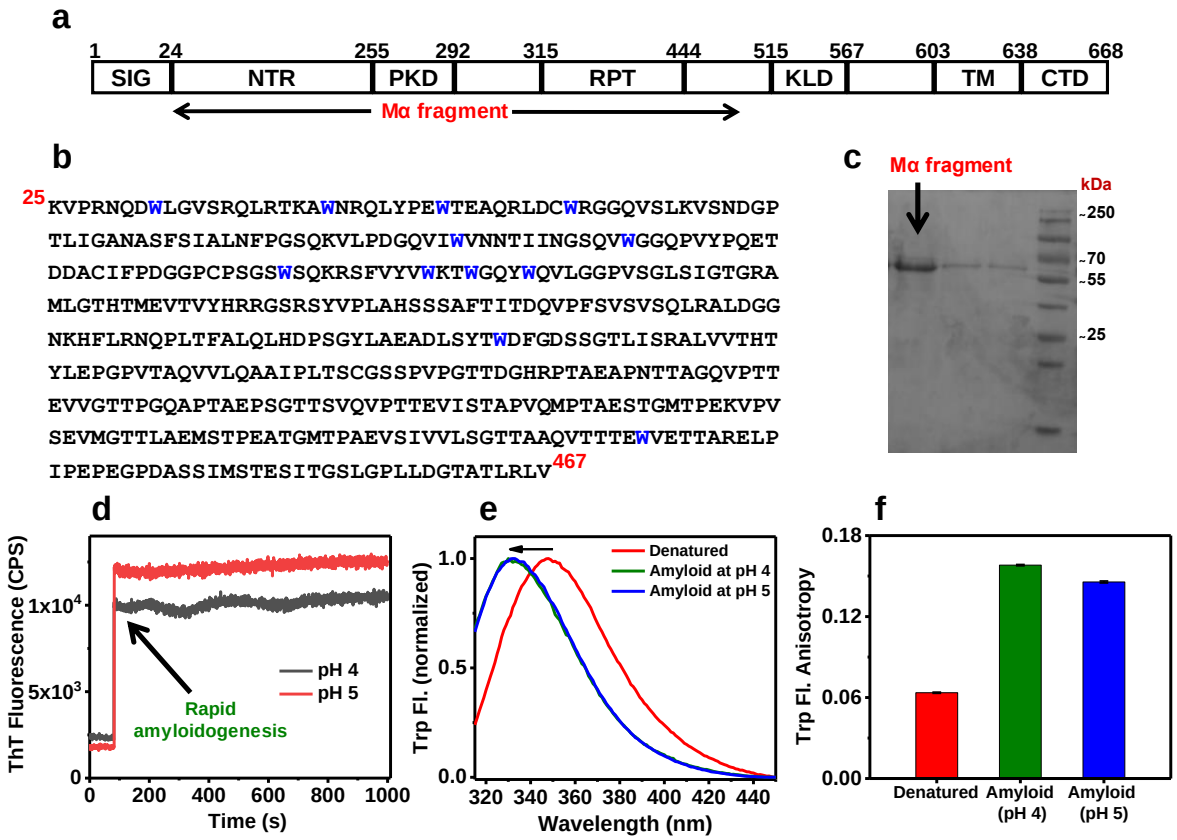


Figure 3.1 (a) Full-length Pmel17 is composed of 10 domains. SIG, signal peptide; NTD, N-terminal domain; PKD, polycystic kidney disease-like domain; RPT, repeat domain; KRG, kringle-like domain; TM, transmembrane domain; CTD, C-terminal domain. (b) Amino acid sequence of the Mα fragment (25-467) of Pmel17. (c) SDS-PAGE (12%) analysis of purified Mα fragment under denaturing

conditions. (d) Aggregation kinetics of Mα fragment (concentration: 5 μM, 37 °C) monitored by thioflavin T fluorescence. (e) Trp steady-state fluorescence spectra, (f) The changes in the Trp fluorescence anisotropy of Mα fragment in the denatured state (red), amyloid state at pH 4 (olive), and at pH 5 (blue).

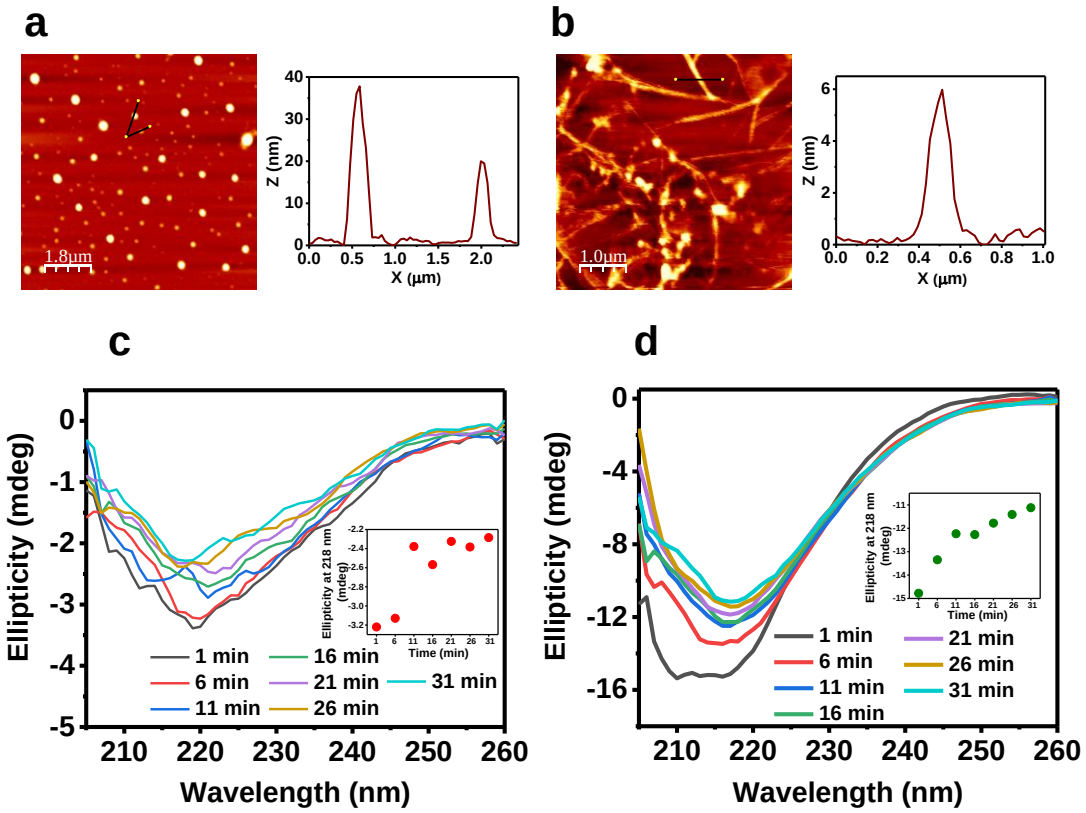


Figure 3.2 AFM images of Mα fragment aggregates (a) at pH 4 (b) at pH 5 imaged after overnight incubation at room temperature. The height profiles are shown on the right side of each AFM image. Far-UV CD spectra of the Mα fragment (c) at pH 4 and (d) at pH 5 in the presence of 100 mM KCl at 37 °C monitored as a function of time.

in the emission maxima and lower anisotropy indicated that the Trp residues are solvent exposed in the free unfolded/disordered state of the Mα fragment at pH 7.4. However, in the amyloid state (at pH 4 and pH 5), a blue shift in emission maxima of ~18 nm and higher anisotropy suggested that Trp gets buried into the hydrophobic environment during aggregation (Figure 3.1e,f). Next, we performed atomic force microscopy (AFM) imaging to discern the nanoscale morphology of these aggregates. We observed oligomers of 20-40 nm in height for pH 4 sample (Figure 3.2a), however at pH 5 thread-like amyloid fibrils of height ~ 6-7 nm was observed, after overnight incubation at room temperature (Figure 3.2b). We next monitored the secondary structural changes by using far-UV circular dichroism (CD) spectroscopy. At pH 4, CD spectra showed a rapid conformational conversion of the Mα fragment into β-sheet structure within 1 min of mixing (Figure 3.2c). On the contrary, at pH 5 a complete transition

to β-sheet was observed after 20 min of incubation at 37 °C (Figure 3.2d). The presence of β-sheet structure was further supported by the Raman spectrum that showed two major amide I peaks at 1634 cm^{-1} and 1670 cm^{-1} which correspond to β-sheet rich conformation[45] (Figure 3.3a,b). In addition to the protein backbone conformational changes, we also investigated the side chain vibrations at the residue-specific level. We observed multiple peaks in 500-600 cm^{-}

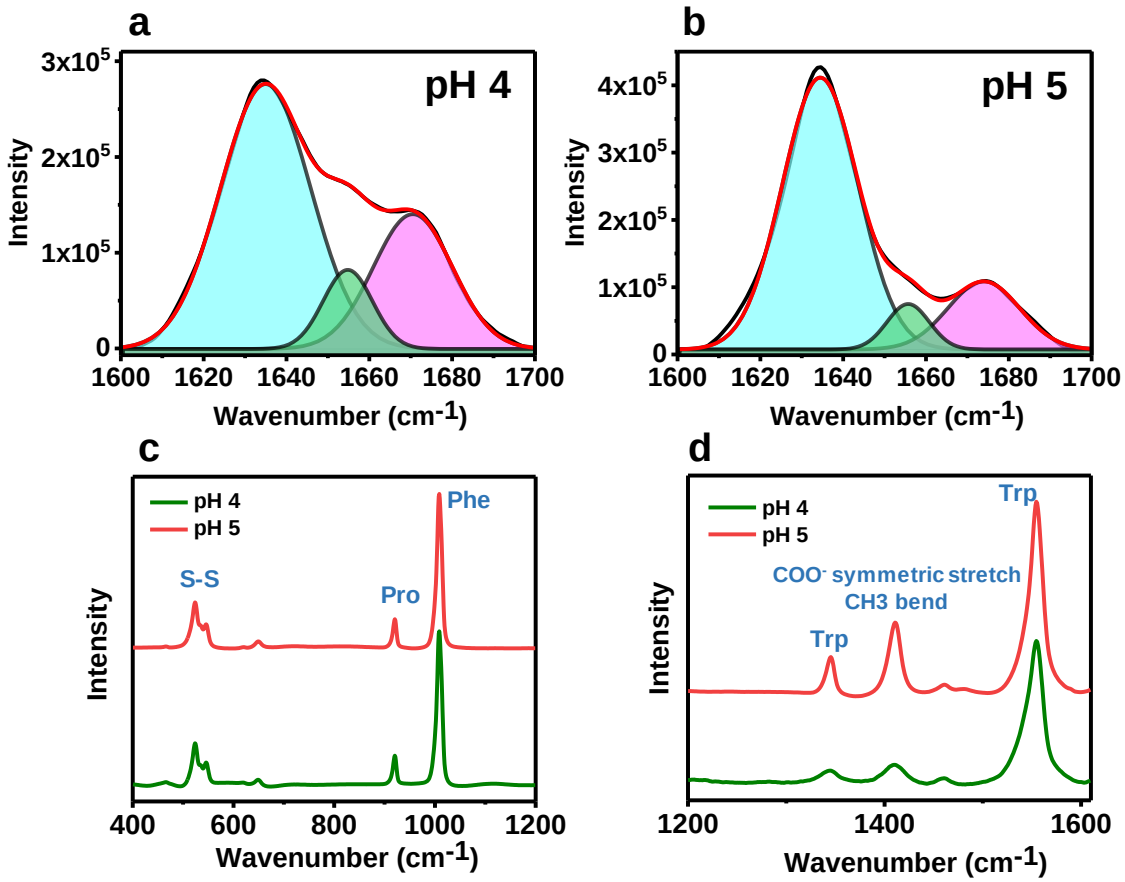


Figure 3.3 The deconvolution of amide I Raman region of the Mα fragment aggregates formed at pH 4 (a) and pH 5 (b). The individual peaks obtained upon fitting are shown in cyan, green and magenta. The cumulative peaks are shown in red. Raman spectra (c) (400-1200 cm^{-1}) and (d) (1200-1600 cm^{-1}) of the Mα fragment aggregates.

^{1} for the disulfide bonds with two major peaks positioned at ~524 cm^{-1} and ~545 cm^{-1} that are attributed to gauche–gauche–trans (χ: ±80-900) and trans–gauche–trans conformations, respectively that originates from a strong dependence of S-S wavenumber position on the torsional angle χ(C–S–S–C) which provides information about the internal rotation about C–S and C–C bonds in Cα–Cβ–S–S–Cβ'–Cα' conformations[44] (Figure 3.3c). The NTD and PKD present in the Mα fragment contains three and one cysteine residue, respectively. We believe that the cysteines present in these domains are involved in the formation of disulfide bonds during the amyloid formation. We also observed peaks at 920 cm^{-1}, 1008 cm^{-1}, 1340 cm^{-1},

1410 cm^{-1}, and 1550 cm^{-1} that are assigned to proline, phenylalanine, tryptophan, COO^- stretching, CH_3 bending and the dihedral angle between the indole ring and the peptide bond plane in tryptophan, respectively[45] (figure 3.3c,d). Moreover, we observed differences in the intensities as well as width of the Raman bands that indicate the formation of two distinct species at two different pHs (pH 4 and pH 5) which also corroborates our AFM results. Together, these findings provide structural and morphological insights into the amyloid formation of the Mα fragment and reveals a functionally controlled aggregation pathway of Mα fragment amyloid assembly at melanosomal pH that involves rapid oligomerization at pH 4 and formation of long fibrils at pH 5 akin to the amyloid transition during stage I to stage II melanosome maturation in vivo. We next embarked upon the studies aimed at elucidating the mechanism of amyloid formation of an intrinsically disordered pH-responsive repeat domain (RPT) of Mα fragment that is known to form the amyloid core that promotes melanin biogenesis.[46]

3.3.2 pH-responsive mechanistic switch in the aggregation pathway of RPT

RPT is a disordered domain that resides in the Mα fragment of Pmel 17 that comprises 10 copies of an imperfect repeat sequence (Figure 3.4a), remains in the natively unfolded monomeric form at neutral pH, and aggregates only under acidic conditions.[17,21,41,46-53] First, to elucidate the aggregation mechanism of RPT as a function of pH, we carried out aggregation reactions in acidic to mildly acidic pH range. The amyloid formation kinetics was monitored by fluorescence of an amyloid reporter dye, ThT. At pH 3/4, RPT rapidly aggregates without a lag phase, indicating an isodesmic polymerization kinetics (Figure 3.4b). On the contrary, at pH 5, aggregation of RPT exhibits a typical nucleation-dependent polymerization kinetics that displays a long lag phase followed by a rapid assembly phase (Figure 3.4b). Also, the final ThT intensity was ~ 10 times lower for the aggregates formed at pH 3/4 compared to that at pH 5, suggesting weaker internal packing at pH 3/4. The ThT fluorescence kinetics is also supported by a similar kinetic profile of tryptophan fluorescence anisotropy that reports the changes in the rotational mobility of the single tryptophan (W423) during aggregation (Figure 3.4c). This result also indicated that the C-terminal segment of RPT harboring the W423 residue constitutes the amyloid core, as suggested previously.[48,51] A careful look at the anisotropy kinetics suggested that the RPT undergoes a rapid oligomerization during aggregation at pH 3/4. We next set out to determine the specific pH at which the switch between the two mechanisms occurs. We carried out aggregation reactions by varying the pH between 4 and 6 (Figure 3.4d). We observed that at pH < 4.5, RPT readily aggregates without a lag phase,

whereas at pH > 4.5, the lag phase emerges. It should be noted that the isoelectric point (pI) of the RPT is ~4.5, at which presumably both nucleation and isodesmic mechanisms are operative. Next, to ascertain whether instantaneous oligomerization occurs after transferring the RPT from pH 7 to pH 3/4 as inferred from the tryptophan anisotropy kinetics, we carried out dynamic light scattering experiments to monitor changes in the average hydrodynamic sizes. As we expected, initial oligomers were indeed formed immediately after mixing at pH 3/4 ((hydrodynamic radius ~12 nm). In contrast, RPT remained monomeric immediately after

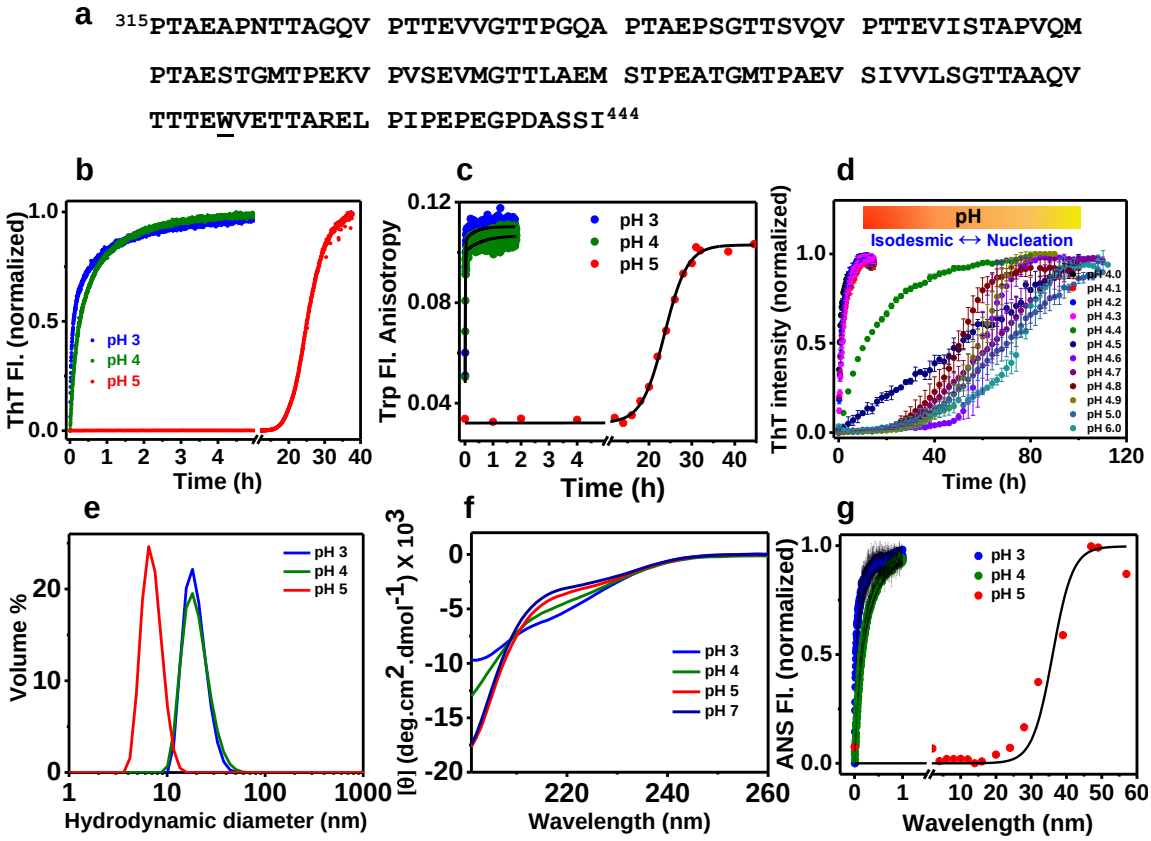


Figure 3.4 (a) Amino acid sequence of the RPT (W423 is underscored). (b) ThT fluorescence kinetics of RPT aggregation (25 μM, 25 °C) when incubated in aggregation buffer (20 mM sodium citrate, 150 mM NaCl buffer), at pHs 3 (blue), 4 (olive), and 5 (red), with agitation at 600 rpm. (c) Trp fluorescence anisotropy kinetics during aggregation at pHs 3 (blue), 4 (olive), and 5 (red). The fitted lines were obtained using a biexponential rate equation (pHs 3 and 4) and a nucleation-dependent model (pH 5). (d) pH titration of RPT aggregation (25 μM, 25 °C, 600 rpm) from pH 4 to pH 6 (20 mM citrate buffer, for pH 6, MES buffer, 150 mM NaCl) showing the mechanistic switch at ~pH 4.5. The kinetic experiments were performed at least three times to obtain the standard error. (e) The size distribution of RPT immediately after mixing at pH 3 (blue), pH 4 (olive), and 5 (red). (f) Far-UV CD spectra collected immediately after mixing RPT (15 μM) in pH 3 (blue), pH 4 (olive), pH 5 (red) buffers (20 mM sodium citrate and 150 mM NaCl) and pH 7 (deep blue) (20 mM HEPES buffer, 150 mM NaCl) at 25°C. (g) Aggregation kinetics of RPT at pHs 3 (blue) and 5 (red) monitored using ANS fluorescence. Reprinted/adapted with permission from Dogra et al. (Ref. 57).

transferring into pH 5 (hydrodynamic radius ~4 nm) (Figure 3.4e). This is also accompanied by a secondary structural change monitored by CD (Figure 3.4f) that showed an intrinsically disordered state at pH 5/7 and partially ordered state at pH 3/4. Our ANS binding experiments also revealed that the formation of hydrophobic pockets during the aggregation followed a non-nucleation mechanism at pH 3/4 and a nucleation mechanism at pH 5 (Figure 3.4g).

Table 3: Kinetics parameters obtained from the analysis of aggregation kinetics studied using ThT fluorescence (mean ± SEM).

pH	k_1 (h^{-1})	k_2 (h^{-1})	t_{lag} (h)	$t_{1/2}$ (h)
4.0	3.58 ± 0.50	0.40 ± 0.07	-	-
4.1	2.17 ± 0.13	0.30 ± 0.07	-	-
4.2	1.87 ± 0.25	0.42 ± 0.08	-	-
4.3	1.05 ± 0.41	0.35 ± 0.08	-	-
4.4	0.22 ± 0.07	0.04 ± 0.01	-	-
4.5	-			
4.6			53.67 ± 3.11	61.72 ± 4.98
4.7			31.19 ± 2.61	68.24 ± 10.77
4.8			34.14 ± 4.97	48.66 ± 3.24
4.9			44.19 ± 3.85	60.01 ± 4.26
5.0			35.97 ± 4.40	67.98 ± 7.43
6.0			56.12 ± 4.89	75.62 ± 1.76

Kinetic traces were fitted using isodesmic (pH 4-4.4) and nucleation-dependent polymerization (pH 4.6-6) models. Kinetics at pH 4.5 could not be fitted with either of the models presumably because of the coexistence of both the aggregation mechanisms.

3.3.3 Distinct conformations and nanoscale morphologies of RPT aggregates

Next, to investigate the changes in the secondary structure during aggregation, we utilized far-UV CD spectroscopy (Figure 3.5a). At pH 5, RPT remained largely disordered in the lag phase. However, upon aggregation into amyloids, a sharp peak at 215 nm was observed, indicating the formation of a characteristic β-sheet-rich structure (Figure 3.5a). On the contrary, at pHs 3 and 4, upon commencement of aggregation, a sharp drop in the random coil content and a concomitant increase in the β-sheet content indicated a fast transition from an unordered to a

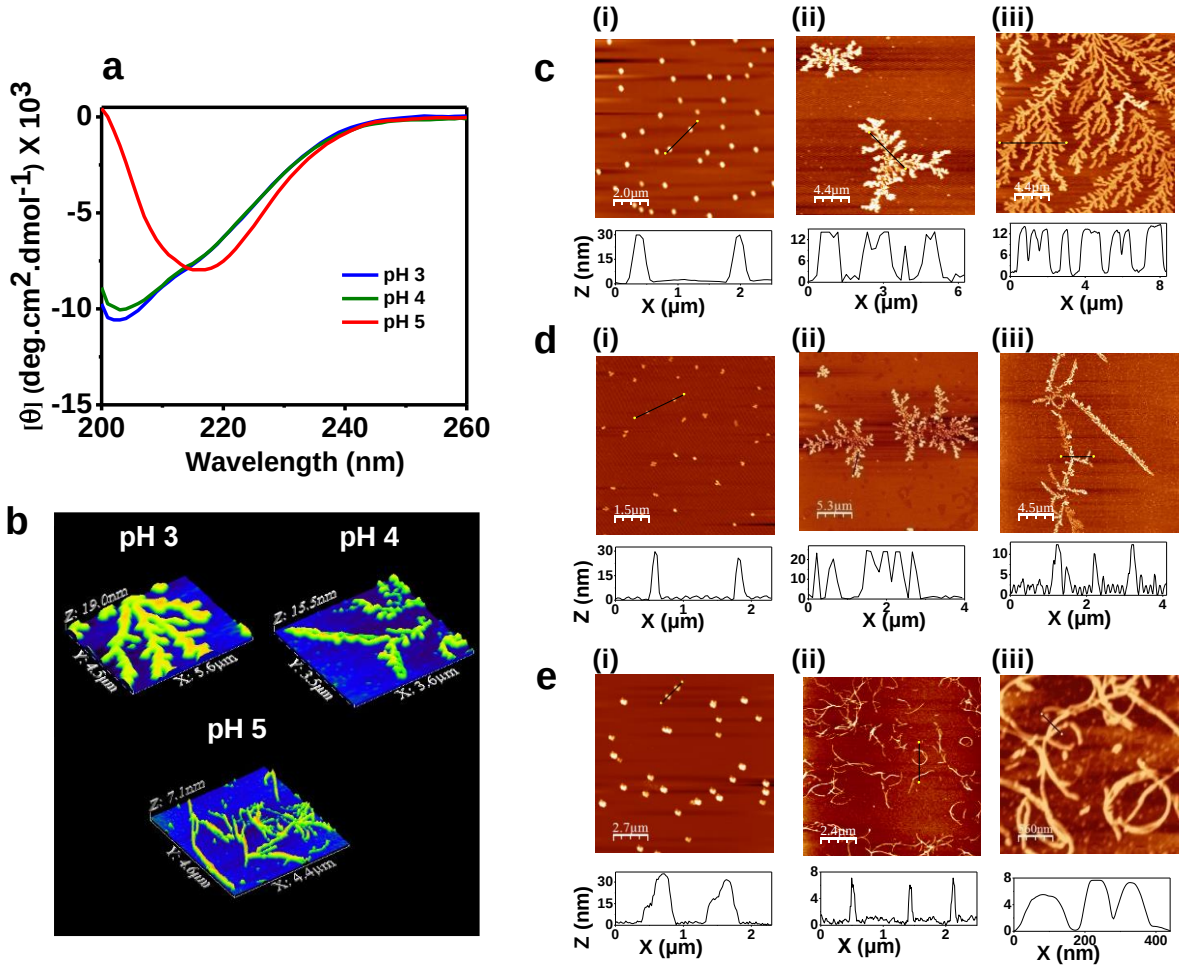


Figure 3.5 (a) Far-UV CD spectra of the aggregates formed at pHs 3 (blue), 4 (olive), and 5 (red) after 60 h, at 25 °C. (b) AFM images of RPT aggregates showing fractal dendrites and curvy thread-like fibrils formed at pHs 3, 4 (after 48 h) and 5 (after 25 h), respectively. Time-dependent AFM images of RPT aggregates with their respective height profiles at 25 °C: (c(i), d(i)) oligomers imaged immediately (25 s) after mixing the protein at pHs 3 and 4, respectively. (c(ii), d(ii)) The star-like structure formed after 5 h of shaking at 600 rpm at pHs 3 and 4, respectively. (c(iii), d(iii)) Fractal dendrites formed when the aggregation reaction was incubated further for 2 days at pHs 3 and 4, respectively. (e(i)) Oligomers imaged after 5 h of shaking at 600 rpm at pH 5, (e(ii)) Curvy thread-like protofilaments formed after 25 h at pH 5. (e(iii)) Long entangled fibrils formed after 60 h at pH 5. Reprinted/adapted with permission from Dogra et al. (Ref. 57).

partially ordered state. Upon saturation, the CD spectra indicated the presence of both unordered and β-sheet structures within the aggregates (Figure 3.5a). Therefore, our CD data revealed that aggregates formed at pH 5 by the nucleation dependent mechanism exhibit a more ordered β-sheet-rich structure compared to that of the aggregates formed at pH 3/4. We next performed AFM, the morphology of these aggregates revealed that the distinct mechanisms resulted into widely diverse nanoscale morphologies. RPT formed curvy thread-like fibrils at pH 5 and fractal dendritic superstructures at pH 3/4 (Figure 3.5b). Next, to directly watch the time evolution of aggregate morphologies at different pH values, we performed time-dependent AFM imaging. At both pH 3 and 4, RPT rapidly assembled to form spherical oligomers, having an average height of ∼25 nm (Figure 3.5c,d(i)). The early oligomerization corroborated our light scattering results. These oligomers slowly transformed into star-like morphologies (Figures 3.5c,d(ii)). These star-like aggregates finally matured into fractal dendritic nanostructures with an average height of ∼10 nm presumably by multiple steps of dissociation and association (Figure 3.5c,d(iii)). On the contrary, at pH 5, RPT did not undergo rapid oligomerization that is in agreement with our light scattering results. Oligomers were observed later in the lag phase after several hours and transformed into curvy and straight fibrils, having average height of 7−8 nm and length of several microns (Figure 3.5e(i,ii,iii)). Upon prolonged incubation, presumably both monomer addition and end-to-end association of the shorter fibrillar precursors formed long and entangled fibrils with unaltered height profiles (Figure 3.5e(iii)). Overall, the nanoscale diversity in the fibrillar morphological transformations under different pH values reaffirms that the aggregation mechanism plays a critical role in dictating the RPT amyloid polymorphism.

3.3.4 Amyloid formation of RPT is regulated in a directed pathway

We next directed our efforts to study whether the final nanoscale morphologies are interconvertible upon change in the pH. We hypothesized that this set of experiments will allow us to identify as to how the fibril formation is regulated during melanosome maturation.[41] Upon transfer from pH 4 to pH 5, the ThT fluorescence kinetics showed an initial drop and a lag phase before RPT reassembled into fibrillar amyloids having much higher ThT binding affinity (Figure 3.6a). Our time-dependent AFM imaging experiments revealed that this conversion process was mediated by initial disintegration of the dendritic structure to oligomeric intermediates that slowly transformed into fibrillar morphology (Figure 3.6b). During the

conversion process, CD results also indicated a change from a partially ordered structure to a highly ordered characteristic β-structure (Figure 3.6c). These studies revealed that the dendritic morphology formed via the isodesmic polymerization mechanism at pH 4 retains the capability to convert into highly ordered amyloid fibrils through dissociation, followed by the nucleation-

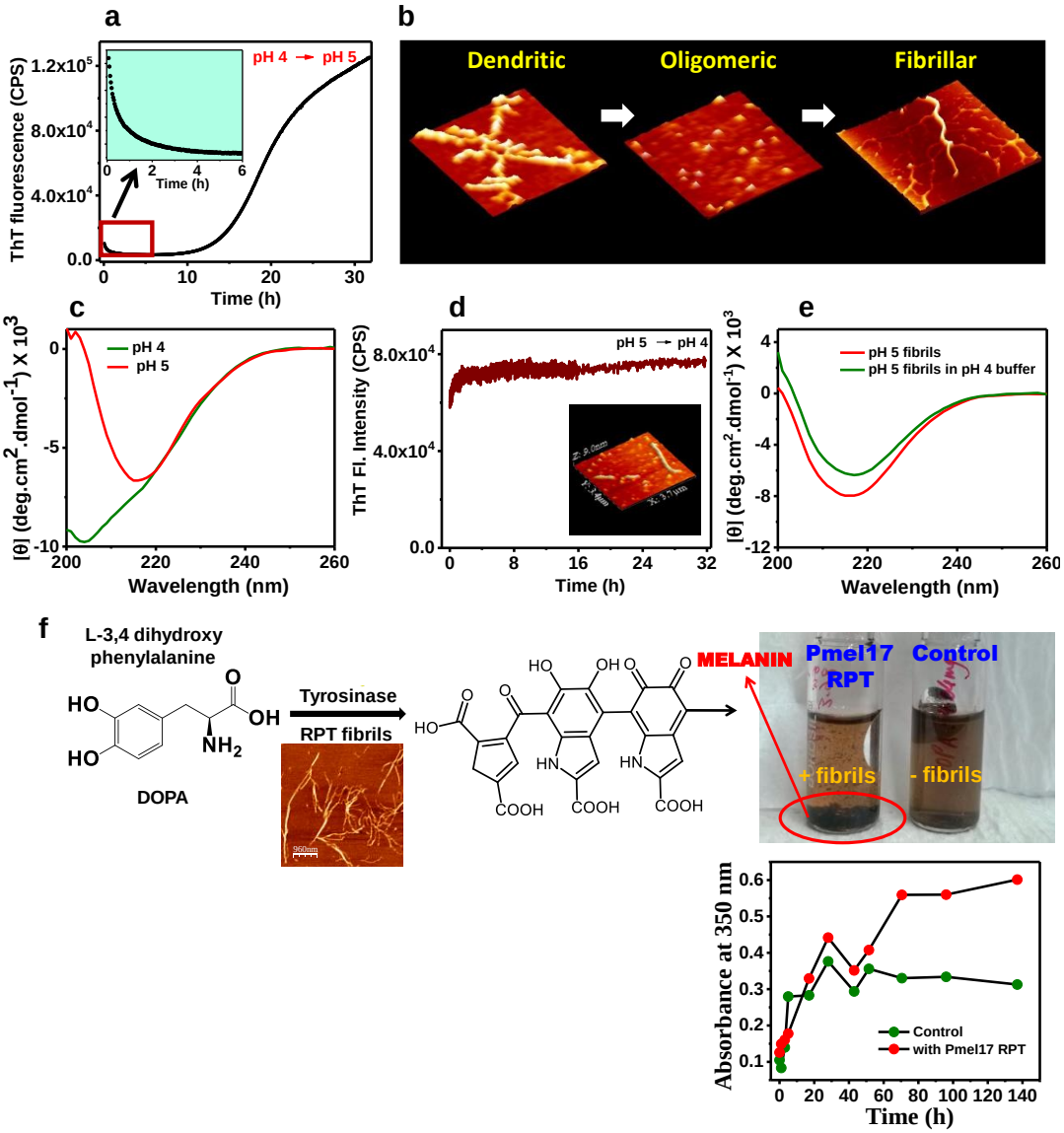


Figure 3.6 (a) ThT fluorescence kinetics upon transferring the aggregates from pH 4 to pH 5. The arrow pointing toward the inset indicates an early drop in the ThT fluorescence. (b) AFM images obtained during transfer of aggregates from pH 4 to pH 5, showing morphological transition from a dendritic to a fibrillar nanostructure through spherical oligomeric intermediates. (c) Changes in far-UV CD upon transition from pH 4 to pH 5. (d) ThT fluorescence kinetics of amyloid fibrils formed at pH 5 showed no significant change in the fluorescence intensity when transferred to pH 4 buffer. An AFM image (represented as 3D) collected after 40 h of incubation at 25°C with agitation is shown in the inset. (e) Far-UV CD spectra of pH 5 fibrils (red) when transferred to pH 4 buffer showed no change in the conformation (olive). Reprinted/adapted with permission from Dogra et al. (Ref. 57). (f) A schematic representation of *in vitro* melanin biosynthesis shows that fibrils of Pmel17 RPT enhance melanin formation.

dependent polymerization mechanism. In sharp contrast, highly ordered fibrils formed at pH 5 did not convert into dendritic morphology even after prolonged incubation at pHs 3 or 4 (Figure 3.6d,e). These results indicated that dendritic nanostructures internal packing can be readily switched to more ordered fibrils by changing the pH, whereas highly ordered fibrils exhibit resistance to the change in possessing lower pH and are inconvertible at lower pH. We propose that such a pH-dependent mechanism−morphology relationship could be a strategic quality control mechanism employed during pH-dependent melanosome maturation, which might be critical for melanin biosynthesis. We next performed *in vitro* melanogenesis assay to observe the scaffolding role of Pmel17 RPT fibrils during melanin biosynthesis. We added L-3,4-dihydroxyphenylalanine (DOPA) to a vial containing 125 mM of sodium citrate buffer pH 5.0 and initiated the reaction using an enzyme, namely tyrosinase. Equal volume of this solution was then transferred into two vials. We then added Pmel17 RPT fibrils to the first vial. Within 24 h we observed the formation of insoluble melanin polymer, that settled at the bottom of the vial. These results tested the templating ability of Pmel17 RPT fibrils and also recapitulated melanin formation within the melanosome. Furthermore, we set out to decipher the thermodynamic and kinetic stability of RPT aggregates.

3.3.4 Kinetic and thermodynamic stability of RPT aggregates

RPT fibrils possess a unique property of undergoing rapid disaggregation upon transferring into a solution of neutral or mildly alkaline pH.[17] This is a unique aspect of RPT aggregation in which protonation/deprotonation acts as a reversible switch between aggregated/monomeric states. We took advantage of this phenomenon to monitor the kinetic stability of different aggregates. For carrying out the time-resolved fibril dissociation assay for RPT amyloid aggregates formed at pHs 4 and 5, we concentrated the fibrils and finally transferred it to 20 mM HEPES buffer, 150 mM NaCl, pH 7.0, at 25 °C, with agitation. Both dendrites and fibrils demonstrated exponential decay of ThT fluorescence upon jumping the pH to neutrality. Fibrillar aggregates showed a much slower disaggregation kinetics compared to dendritic aggregates, indicating higher kinetic stability (Figure 3.7a). The kinetic traces were fitted using exponential decay functions. The time-resolved dissolution kinetics at pH 4 was fitted using a monoexponential decay function, and the rate recovered was 1.28 min^{-1}. However, in the case of pH 5, the reaction kinetics was fitted using a biexponential decay function and the rates recovered were as follows: k_{fast} was 0.66 min^{-1}, k_{slow} was 0.11 min^{-1}, and kavg was 0.36 min^{-1}. We also monitored the thermodynamic stability by denaturing the aggregates using urea

(Figure 3.7b). For thermodynamic stability experiments, the fibrils were incubated with urea for 1 h at room temperature prior to data acquisition. For estimating the stability, the tryptophan fluorescence ratio at 350 nm (denatured) and 330 nm (aggregated) was plotted as a function of urea concentration. The data were fitted using the sigmoidal Boltzmann function, and the midpoints of transitions (C_m) were obtained. In the case of pH 4 aggregates, C_m was 3.0 M, and at pH 5, C_m was 3.5 M. The minor yet significant differences in the dissociation rates and midpoints of transitions of these aggregates indicated that fibrillar aggregates have more kinetic and thermodynamic stability compared to dendritic aggregates. Taken together, our results demonstrate that fibrils formed via the nucleation mechanism at pH 5 have higher-order

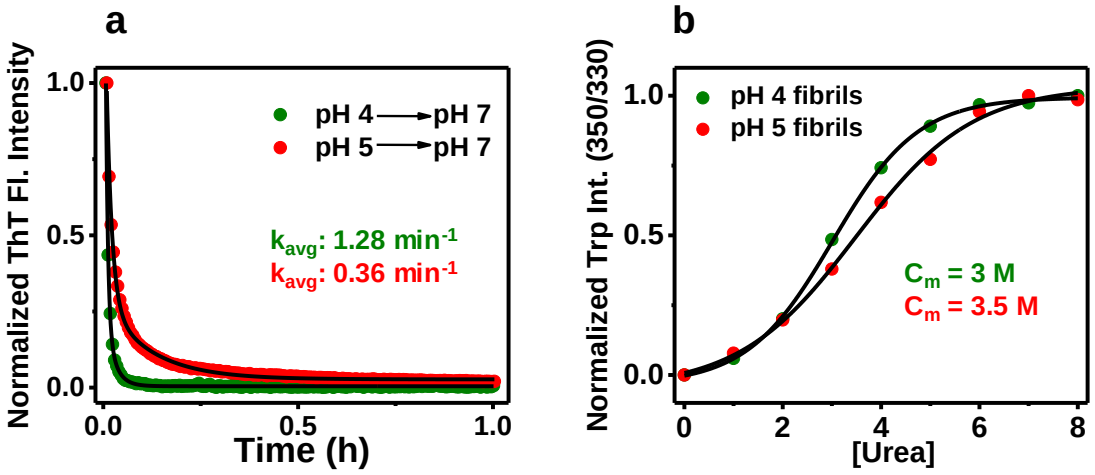


Figure 3.7 (a) The time course of fibril dissolution assay for RPT amyloid aggregates formed at pHs 4 (olive) and 5 (red) in 20 mM HEPES buffer, 150 mM NaCl, at pH 7 and 25 °C, with agitation. The average rates of dissolution were 1.28 and 0.36 min^{-1} for fibrils formed at pH 4 and 5, respectively. (b) The denaturation of the two types of aggregates monitored using the ratio of Trp fluorescence at 350 and 330 nm. Reprinted/adapted with permission from Dogra et al. (Ref. 57).

structure and stability compared to those of dendritic nanostructures formed via the isodesmic mechanism at pH 4.

3.4 Discussion

In summary, our results reveal a striking switch in the mechanism of aggregation of the RPT of Pmel17 from an isodesmic polymerization model to a nucleation-dependent model as a function of pH. Consequently, alteration in the aggregation pathway results in a nanoscale diversity of aggregates, ranging from fractal dendritic to fibrillar morphology having varied internal packing and stability. We speculate that the transition in the aggregation mechanism could occur due to the protonation/deprotonation of glutamic acid residues, which has been previously shown to influence the course of aggregation.[51] At pH 3/4, the protonation of

glutamates leads to a minimal charge repulsion that reinforces intermolecular RPT association mediated via noncovalent interactions. For instance, hydrogen bonding and hydrophobic interactions will facilitate fast oligomerization that can potentially create collapsed and less ordered supramolecular structures possessing multiple sticky surfaces responsible for branched dendritic morphology upon maturation. On the contrary, at pH 5, electrostatic repulsions between the polypeptide chains prevent instantaneous oligomerization. Therefore, the polypeptide chain has to overcome the electrostatic repulsions during the nucleation process that eventually results in the formation of a critical nucleus. Then, this process is presumably driven by the nucleated conformational conversion model that allows the conformational sequestration within the oligomers that mature into canonical amyloid fibrils.[54−56] Additionally, the formation of the interbackbone hydrogen-bonding network arising due to the variation in side-chain interactions may be different for the two types of amyloids.[7] These altered intermolecular interactions can lead to the formation of varied amyloid cores as seen previously under different aggregation conditions.[41] We suggest that the pH modulation within the melanosomes allows the optimal conditions to avoid the formation of toxic intermediates prominent in pathogenic amyloidogenesis. In conclusion, our findings address key molecular aspects that relate the protein aggregation mechanism with the nanoscale polymorphism. Our spectroscopic readouts allowed us to delineate the structural transition of protein molecules within the aggregates formed by two distinct mechanisms. Additionally, our time-dependent AFM imaging allowed us to decipher the morphological transition of the early oligomeric intermediates into higher-order (matured) fibrillar and dendritic aggregates. We believe that the pH-responsive regulation of nanostructures will open new avenues for the fabrication of novel functional nano-biomaterials with a wide range of applications.

3.5 References:

(1) 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, 384−396.

(2) Tycko, R.; Wickner, R. B. Molecular Structures of Amyloid and Prion Fibrils: Consensus vs. Controversy. *Acc. Chem. Res.* **2013,** 46, 1487−1496.

(3) Greenwald, J.; Riek, R. Biology of Amyloid: Structure, Function, and Regulation. *Structure* **2010,** 18, 1244−1260.

(4) Eichner, T.; Radford, S. E. A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Mol. Cell* **2011,** 43, 8−18.

(5) Tipping, K. W.; van Oosten-Hawle, P.; Hewitt, E. W.; Radford, S. E. Amyloid Fibres: Inert End-Stage Aggregates or Key Players in Disease? *Trends Biochem. Sci.* **2015,** 40, 719−727.

(6) Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.* **2014,** 114, 6844−6879.

(7) 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, 1900−1903.

(8) Bhattacharya, M.; Mukhopadhyay, S. Nanophotonics of Protein Amyloids. *Nanophotonics* **2014,** 3, 51−59.

(9) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional Amyloid Formation within Mammalian Tissue. *PLoS Biol.* **2006,** 4, e6.

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

(11) Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P.; Simon, R.; Schubert, D.; et al. Functional Amyloids as Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009,** 325, 328−332.

(12) Barnhart, M. M.; Chapman, M. R. Curli Biogenesis and Function. *Annu. Rev. Microbiol.* **2006,** 60, 131−147.

(13) Frederick, K. K.; Debelouchina, G. T.; Kayatekin, C.; Dorminy, T.; Jacavone, A. C.; Griffin, R. G.; Lindquist, S. Distinct Prion Strains are Defined by Amyloid Core Structure and Chaperone Binding Site Dynamics. *Chem. Biol.* **2014,** 21, 295−305.

(14) Otzen, D. Functional Amyloid: Turning Swords into Plowshares. *Prion* **2010,** 4, 256−264.

(15) Otzen, D.; Nielsen, P. H. We Find Them Here, We Find Them There: Functional Bacterial Amyloid. *Cell. Mol. Life Sci.* **2008,** 65, 910−927.

(16) Shorter, J.; Lindquist, S. Prions as Adaptive Conduits of Memory and Inheritance. *Nat. Rev. Genet.* **2005,** 6, 435−450.

(17) McGlinchey, R. P.; Lee, J. C. Why Study Functional Amyloids? Lessons from the Repeat Domain of Pmel17 *J. Mol. Biol.* **2018,** 430, 3696-3706.

(18) Berson, J. F.; Harper, D. C.; Tenza, D.; Raposo, G.; Marks, M. S. Pmel17 Initiates Premelanosome Morphogenesis within Multivesicular Bodies. *Mol. Biol. Cell* **2001,** 12, 3451−3464.

(19) Berson, J. F.; Theos, A. C.; Harper, D. C.; Tenza, D.; Raposo, G.; Marks, M. S. Proprotein Convertase Cleavage Liberates a Fibrillogenic Fragment of a Resident Glycoprotein to Initiate Melanosome Biogenesis. *J. Cell Biol.* **2003,** 161, 521−533.

(20) Theos, A. C.; Truschel, S. T.; Raposo, G.; Marks, M. S. The Silver Locus Product Pmel17/gp100/Silv/ME20: Controversial in Name and in Function. *Pigm. Cell Res.* **2005,** 18, 322−336.

(21) Hoashi, T.; Muller, J.; Vieira, W. D.; Rouzaud, F.; Kikuchi, K.; Tamaki, K.; Hearing, V. J. The Repeat Domain of the Melanosomal Matrix Protein PMEL17/GP100 is Required for the Formation of Organellar Fibers. *J. Biol. Chem.* **2006,** 281, 21198−21208.

(22) Knowles, T. P.; Buehler, M. J. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* **2011,** 6, 469−479.

(23) Shewmaker, F.; McGlinchey, R. P.; Wickner, R. B. Structural Insights into Functional and Pathological Amyloid. *J. Biol. Chem.* **2011,** 286, 16533−16540.

(24) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006,** 75, 333−366.

(25) Li, C.; Mezzenga, R. The Interplay between Carbon Nanomaterials and Amyloid Fibrils in Bio-Nanotechnology. *Nanoscale* **2013,** 5, 6207−6218.

(26) Bolisetty, S.; Mezzenga, R. Amyloid−Carbon Hybrid Membranes for Universal Water Purification. *Nat. Nanotechnol.* **2016,** 11, 365−371.

(27) Cherny, I.; Gazit, E. Amyloids: Not only Pathological Agents but also Ordered Nanomaterials. *Angew. Chem., Int. Ed. Engl.* **2008,** 47, 4062−4069.

(28) Knowles, T. P.; Mezzenga, R. Amyloid Fibrils as Building Blocks for Natural and Artificial Functional Materials. *Adv. Mater.* **2016,** 28, 6546−6561.

(29) Li, C.; Bolisetty, S.; Mezzenga, R. Hybrid Nanocomposites of Gold Single-Crystal Platelets and Amyloid Fibrils with Tunable Fluorescence, Conductivity, and Sensing Properties. *Adv. Mater.* **2013,** 25, 3694−3700.

(30) Frieden, C. Protein Aggregation Processes: In Search of the Mechanism. *Protein Sci.* **2007,** 16, 2334−2344.

(31) Ferrone, F. A. Assembly of Aβ Proceeds via Monomeric Nuclei. *J. Mol. Biol.* **2015,** 427, 287−290.

(32) Gosal, W. S.; Morten, I. J.; Hewitt, E. W.; Smith, D. A.; Thomson, N. H.; Radford, S. E. Competing Pathways Determine Fibril Morphology in the Self-Assembly of beta2-microglobulin into Amyloid. *J. Mol. Biol.* **2005,** 351, 850−864.

(33) Kumar, S.; Udgaonkar, J. B. Mechanisms of Amyloid Fibril Formation by Proteins. *Curr. Sci.* **2010,** 98, 639−656.

(34) Jung, J. M.; Savin, G.; Pouzot, M.; Schmitt, C.; Mezzenga, R. Structure of Heat-Induced beta-lactoglobulin Aggregates and Their Complexes with Sodium-Dodecyl Sulfate. *Biomacromolecules* **2008,** 9, 2477−2486.

(35) Hurshman, A. R.; White, J. T.; Powers, E. T.; Kelly, J. W. Transthyretin Aggregation under Partially Denaturing Conditions is a Downhill Polymerization. *Biochemistry* **2004,** 43, 7365−7381.

(36) Xue, W. F.; Homans, S. W.; Radford, S. E. Systematic Analysis of Nucleation-Dependent Polymerization Reveals New Insights into the Mechanism of Amyloid Self-Assembly. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** 105, 8926−8931.

(37) Meisl, G.; Kirkegaard, J. B.; Arosio, P.; Michaels, T. C.; Vendruscolo, M.; Dobson, C. M.; Linse, S.; Knowles, T. P. Molecular Mechanisms of Protein Aggregation from Global Fitting of Kinetic Models. *Nat. Protoc.* **2016,** 11, 252−272.

(38) Pedersen, J. S.; Andersen, C. B.; Otzen, D. E. Amyloid Structure− One but not the Same: The Many Levels of Fibrillar Polymorphism. *FEBS J.* **2010,** 277, 4591−4601.

(39) Annamalai, K.; Guhrs, K. H.; Koehler, R.; Schmidt, M.; Michel, H.; Loos, C.; Gaffney, P. M.; Sigurdson, C. J.; Hegenbart, U.; Schonland, S.; et al. Polymorphism of Amyloid Fibrils In Vivo. *Angew. Chem., Int. Ed. Engl.* **2016,** 55, 4822−4825.

(40) Paravastu, A. K.; Leapman, R. D.; Yau, W. M.; Tycko, R. Molecular Structural Basis for Polymorphism in Alzheimer's beta-Amyloid Fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** 105, 18349−18354.

(41) Hu, K. N.; McGlinchey, R. P.; Wickner, R. B.; Tycko, R. Segmental Polymorphism in a Functional Amyloid. *Biophys. J.* **2011,** 101, 2242−2250.

(42) Arosio, P.; Knowles, T. P.; Linse, S. On the Lag Phase in Amyloid Fibril Formation. *Phys. Chem. Chem. Phys.* **2015,** 17, 7606−7618.

(43) Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. WSXM: A Software for Scanning Probe Microscopy and a Tool for Nanotechnology. *Rev. Sci. Instrum.* **2007,** 78, No. 013705.

(44) Arya, S.; Kumari, A.; Dalal, V.; Bhattacharya, M.; Mukhopadhyay, S. Appearance of Annular Ring-like Intermediates during Amyloid Fibril Formation from Human Serum Albumin. *Phys. Chem. Chem. Phys.* **2015,** 17, 22862-22871.

(45) de la Cuesta, R. G.; Goodacre, R.; Ashton, L. Monitoring Antibody Aggregation in Early Drug Development Using Raman Spectroscopy and Perturbation-Correlation Moving Windows. *Anal. Chem.* **2014,** 86, 11133−11140.

(46) McGlinchey, R. P.; Shewmaker, F.; McPhie, P.; Monterroso, B.; Thurber, K.; Wickner, R. B. The Repeat Domain of the Melanosome Fibril Protein Pmel17 Forms the Amyloid Core Promoting Melanin Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** 106, 13731−13736.

(47) Ancans, J.; Hoogduijn, M. J.; Thody, A. J. Melanosomal pH, Pink Locus Protein and Their Roles in Melanogenesis. *J. Invest. Dermatol.* **2001,** 117, 158−159.

(48) Pfefferkorn, C. M.; McGlinchey, R. P.; Lee, J. C. Effects of pH on Aggregation Kinetics of the Repeat Domain of a Functional Amyloid, Pmel17. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** 107, 21447−21452.

(49) McGlinchey, R. P.; Shewmaker, F.; Hu, K. N.; McPhie, P.; Tycko, R.; Wickner, R. B. Repeat Domains of Melanosome Matrix Protein Pmel17 Orthologs Form Amyloid Fibrils at the Acidic Melanosomal pH. *J. Biol. Chem.* **2011,** 286, 8385−8393.

(50) McGlinchey, R. P.; Gruschus, J. M.; Nagy, A.; Lee, J. C. Probing Fibril Dissolution of the Repeat Domain of a Functional Amyloid, Pmel17, on the Microscopic and Residue Level. *Biochemistry* **2011,** 50, 10567−10569.

(51) McGlinchey, R. P.; Jiang, Z.; Lee, J. C. Molecular Origin of pH Dependent Fibril Formation of a Functional Amyloid. *ChemBioChem* **2014**, 15, 1569−1572.

(52) Watt, B.; van Niel, G.; Raposo, G.; Marks, M. S. PMEL: A Pigment Cell-Specific Model for Functional Amyloid Formation. *Pigm. Cell Melanoma Res.* **2013,** 26, 300−315.

(53) Jiang, Z.; Lee, J. C. Lysophospholipid-Containing Membranes Modulate the Fibril Formation of the Repeat Domain of a Human Functional Amyloid, pmel17. *J. Mol. Biol.* **2014,** 426, 4074−4086.

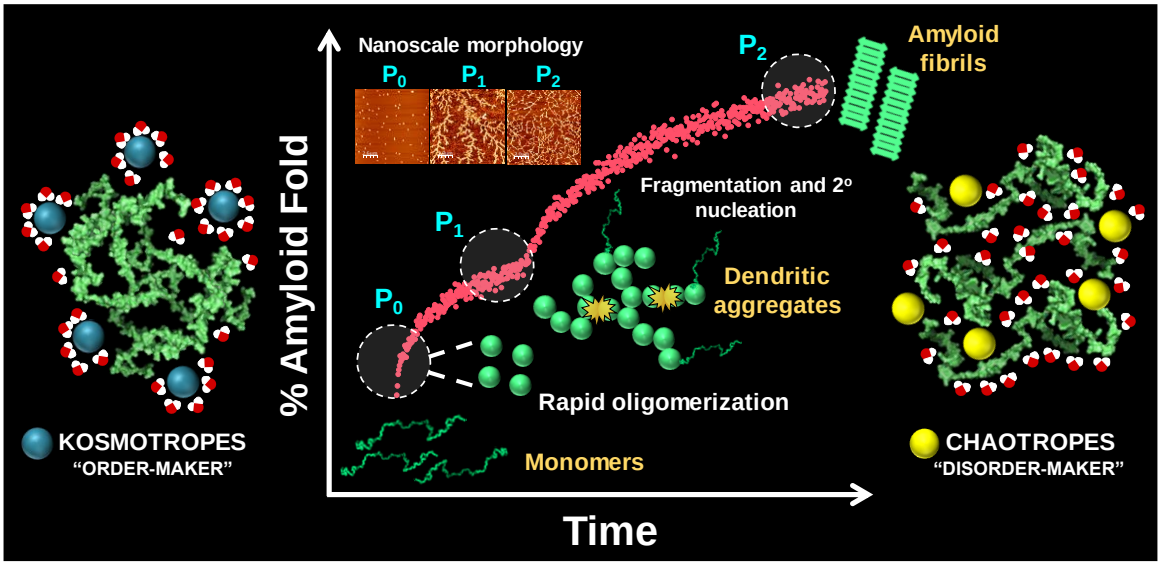
(54) Serio, T. R.; Cashikar, A. G.; Kowal, A. S.; Sawicki, G. J.; Moslehi, J. J.; Serpell, L.; Arnsdorf, M. F.; Lindquist, S. L. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. *Science* **2000,** 289, 1317−1321.

(55) Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-β Forms Fibrils by Nucleated Conformational Conversion of Oligomers. *Nat. Chem. Biol.* **2011,** 7, 602−609.

(56) Vitalis, A.; Pappu, R. V. Assessing the Contribution of Heterogeneous Distributions of Oligomers to Aggregation Mechanisms of Polyglutamine Peptides. *Biophys. Chem.* **2011,** 159, 14−23.

(57) Dogra, P.; Bhattacharya, M.; Mukhopadhyay, S. pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid. *J. Phys. Chem. B* **2017,** 121, 412−419.

Hofmeister Ions Modulate the Autocatalytic Amplification of the Self-Assembly Process of Pmel17 RPT



Dogra, P.; Roy, S. S.; Joshi, A.; Mukhopadhyay, S. Hofmeister Ions Modulate the Autocatalytic Amplification of the Self-Assembly Process of an Intrinsically Disordered Amyloidogenic Domain of a Melanosomal Protein (manuscript in preparation).

4.1 Introduction

Intrinsically disordered proteins/regions (IDPs/IDRs) are associated with vital physiological functions and devastating protein misfolding diseases.[1,2] IDPs are highly sensitive to the changes in the environment. Physicochemical properties of the solution such as pH, ionic strength, and temperature, are believed to greatly influence their stability, conformational state, and aggregation propensity.[3] Specific salt-protein interaction plays a crucial role in governing different phenomena like protein folding, stability, and aggregation .[4-6] In the 1870s, Franz Hofmeister, a pharmacologist, did enormous work to study the relative efficacy of different salts on the precipitation of proteins. Based on their relative qualitative ability to salt-out or salt-in proteins, Hofmeister ions are generally ranked as follows: $CO_3^{2-} > SO_4^{2-} > S_2O_3^{2-} > H_2PO_4^- > F^- > Cl^- > Br^- \approx NO_3^- > I^- > ClO_4^- > SCN^-$ (anions) and $NH_4^+ > Cs^+ > Rb^+ > K^+ > Na^+ > Li^+ > Ca^{2+} > Mg^{2+}$ (cations).[5,7,8,9] The anions and cations present on the left are called kosmotropes, which are believed to be 'water structure makers' have stabilizing, salting-out effects on proteins and promote precipitation. In contrast, ions present on the right are known as chaotropes, which are considered as 'water structure breakers', increase solubility, and promote protein unfolding. Typically, Cl^- from the anionic series and Na^+ from the anionic series are considered as the dividing line between these two types of ions. [5,7,8,9] Subsequently these studies have been found to be of great importance in understanding many complex processes such as protein folding and aggregation in which water-accessible surface area changes significantly.[7] Salt-specific aggregation studies from various IDPs revealed that salt ions have a strong impact on the kinetics and thermodynamics of amyloid formation.[10-16] However, due to their complex kinetic behavior with salts, there is still no clear understanding of the molecular mechanism of salt-induced aggregation processes. Typically, protein aggregation processes occur either via a nucleation dependent polymerization (NDP) or an isodesmic polymerization mechanism.[17,18,19] The NDP mechanism is often associated with the autocatalytic proliferation of the aggregates and generation of cytotoxic molecular species.[20,21] The nucleation events that occur during the NDP include primary nucleation, secondary nucleation, and fragmentation.[20] Several disease-associated IDPs such as α-synuclein, islet amyloid polypeptide (IAPP), and amyloid-β peptide (Aβ) have been shown to undergo monomer-dependent secondary nucleation that is known to cause the autocatalytic amplification of aggregate species.[22-25] Moreover, this type of process has been proposed to be a possible basis for the propagation of infectious prion diseases.[26] However, unlike mammalian prion protein which is associated with a pathogenic process, yeast prion proteins serve

functional roles.[26] Several studies have revealed that the prion propagation in yeast prion proteins occurs via chaperone-mediated fragmentation process.[27] Therefore, to obtain deeper insight into such a phenomenon as well as to understand the roles of secondary nucleation and fragmentation in disease and physiology, it is imperative to study the intrinsic and environmental factors that influence such processes. Although recent studies on the aggregation mechanisms of Aβ and α-synuclein have shown that electrostatic interactions and salt concentrations play a key role in regulating the secondary nucleation processes,[28,29] the role of different Hofmeister ions in modulating these processes remains elusive.

In the present work, attempts were made to understand the effect of Hofmeister ions on the aggregation mechanism/nucleation events of an intrinsically disordered region, the repeat domain (RPT) present in the human Pmel17 protein, a melanoma-specific glycoprotein that plays a critical role in the development of melanosome by forming functional amyloid matrix under mildly acidic pH (4-5.5).[30-32] In order to decode the complex interactions among RPT, water, ions, and the role of Hofmeister effect in the nucleation events, we have used a wide range of salts spanning from kosmotropes to chaotropes. At pH 4, all salts examined, promote oligomerization, and the effect is in line with the Hofmeister series. Interestingly, we observed an unusual biphasic aggregation kinetics of RPT in the presence of different salts that arises due to the dual forward and inverse Hofmeister effects. We demonstrate that this behavior is rooted in the structural sensitivity of the RPT to co-solutes, which in turn affects the oligomerization and self-assembly processes. We show that the early intermolecular contact formation is driven by both electrostatic as well as hydrophobic interactions, and rapid oligomerization initiates in the proline free repeat present in the C-terminal region of RPT. Furthermore, fragmentation, coupled with secondary nucleation processes, dictates the amyloid formation of RPT. Together, our findings reveal the role of different Hofmeister ions in regulating the autocatalytic amplification process involved in the aggregation pathway of RPT.

4.2 Experimental Section

4.2.1 Materials

Guanidinium chloride and urea were purchased from Ameresco. Ampicillin, chloramphenicol and isopropyl-thiogalactopyranoside (IPTG) were obtained from Gold Biocom. 8-Anilinonaphthalene-1-sulfonic acid (ANS), 9-(2,2-Dicyanovinyl)julolidine (DCVJ), thioflavin T (ThT), dimethyl sulfoxide (DMSO), sodium citrate tribasic dihydrate ($C_6H_5Na_3O_7$. $2H_2O$), sodium chloride (NaCl), magnesium sulphate anhydrous ($MgSO_4$), magnesium

chloridehexahydrate ($MgCl_2$. $6H_2O$), ammonium chloride (NH_4Cl), magnesium perchlorate ($Mg(ClO_4)_2$), and sodium sulfate anhydrous (Na_2SO_4) used for buffer preparation were purchased from Sigma (St. Louis, MO). 5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1-Sulfonic Acid) (IAEDANS) was purchased from Invitrogen. PD-10 columns were procured from GE Healthcare Life Sciences (USA). Other chemicals were obtained from Sigma (St. Louis, MO), unless stated otherwise. All of the salt solutions and buffers were prepared in filtered water. The pH of all the buffers were adjusted (± 0.01) on a Metrohm (Herisau, Switzerland) 827 lab pH meter at room temperature.

4.2.2 Mutagenesis, expression, purification and labeling of recombinant RPT

The repeat domain of Pmel17 (RPT) with C-terminal his-tag was expressed in E. coli BL21 (DE3) RIPL cells/strain (Agilent Technologies, Stratagene). Single point mutation at S350C (Cys350), A415C (Cys415) and S442C (Cys442) were created using site-directed mutagenesis kit, QuikChange (Stratagene). Primers used for creating the mutations are as following:

Mutant	Primer	Sequence (5' to 3')
S350C	Forward	GGAACCACATGTGTGCAGGTGCC
	Reverse	GGCACCTGCACACATGTGGTTCC
A415C	Forward	CTGGAACCACATGTGCACAGGTAAC
	Reverse	GTTACCTGTGCACATGTGGTTCCAG
S442C	Forward	GGTCCAGATGCCTGCTCAATCCATCAC
	Reverse	GTGATGGATTGAGCAGGCATCTGGACC

The repeat domain of Pmel17 (RPT) with C-terminal his-tag was expressed in E. coli BL21 (DE3) RIPL cells/strain (Agilent Technologies, Stratagene) and purified using the protocol as described previously.[33] For details please see chapter 2.

The labeling of the cysteine mutants was performed in native buffer (20 mM NaH_2PO_4 buffer, 1 mM DTT, pH 7.4). A 200 mM stock of IAEDANS was prepared in DMSO and mixed in a molar ratio of 20:1 (IAEDANS:RPT) with the RPT mutants. The reaction mixtures were kept in the dark for 2-3 h at room temperature on a rotospin test tube rotator disk. In order to remove excess dye, the labeled proteins were then passed through PD-10 column using 20 mM NaH_2PO_4 buffer, pH 7.4. The concentration of the labeled protein was estimated using ε_{340} = 6100 M^{-1} cm^{-1}. The total protein concentration was estimated using ε_{280} = 5500 M^{-1} cm^{-1}, taking correction factor of IAEDANS into account. The labeling efficiency was >90% for all the

mutant proteins. The concentration of the labeled proteins used for all our measurements was 20 μM.

4.2.3 Circular dichroism (CD) measurements

The far-UV CD spectra were recorded on a Chirascan CD Spectrometer (Applied Photophysics, UK) at room temperature. All the spectra were collected in the scan range of 200–260 nm with 1 nm step size using a quartz cuvette of 1 mm path length. The final concentration of RPT used for all of the experiments was 15 μM. The spectra were averaged over 5-10 scans and the buffer signal was corrected. All of the spectra were then smoothened using the ProData software provided with the Chirascan CD Spectrometer and finally, the mean residue molar elipticity [θ] was calculated and plots were generated using Origin 2018 software.

4.2.4 Dynamic light scattering

Dynamic light scattering measurements were performed on a Malvern Zetasizer NanoZS90 instrument (Malvern, U.K.) at room temperature. For exciting the samples, a He−Ne laser (632 nm) was used. The concentration of the protein used for the measurements was 30 μM. All of the buffers were filtered through 0.02 μm filters (Anatop 10 filter; Whatman). The experiments were repeated thrice and the size distributions were plotted using Origin 2018 software.

4.2.5 Atomic force microscopy (AFM)

The AFM imaging was performed on an Innova atomic force microscope (Bruker) operating in the tapping mode. The scanning was done in sample scanning configuration on NanoDrive (v8.03) software using a slicon nitride cantilever probe of radius ∼8 nm. AFM images were captured at resolution of 256 × 256 or 1024 × 1024 pixels. At different time points, aliquots of 10 μL were withdrawn from the reaction mixture and were deposited on a freshly cleaved muscovite mica (Grade V-4 mica from SPI, PA), that was washed with the appropriate aggregation buffer filtered using 0.22 μM membrane (Millipore) prior to sample deposition. The samples were incubated for 5 min at room temperature and were then washed with the aggregation buffer. Furthermore, the samples were dried under a gentle stream of nitrogen for ∼5 min prior to scanning. All of the AFM images were processed further using WSxM version 4.0 Beta 9.0 software.[34]

4.2.6 Steady-state fluorescence measurements

All of the steady-state fluorescence measurements were performed on a FluoroMax-4 spectrofluorometer (Horiba JobinYvon, NJ) at room temperature using a quartz cuvette of 1 mm path length. The final concentration of RPT used in all the experiments was 20 μM. For DCVJ fluorescence experiments, samples were excited at 453 nm and the emission spectra were scanned from 470 nm to 550 nm. For ANS fluorescence, λ_{ex} = 375 nm and the emission spectra were collected at λ_{em} (range) = 420-580 nm. The final concentration of DCVJ and ANS used for all of the experiments was 20 μM. For tryptophan (Trp) fluorescence measurements following parameters were adjusted: λ_{ex} = 280 nm and λ_{em} (range) = 310-450 nm.
The steady-state fluorescence anisotropy (r_{ss}) is defined by the following relation:

$$r_{ss} = (I_{\parallel} - I_{\perp}G)/(I_{\parallel} + 2I_{\perp}G)$$

where $I_{\parallel}$ and $I_{\perp}$are the parallel and perpendicular fluorescence intensities components, respectively. The perpendicular components were corrected using a G-factor. For Trp fluorescence anisotropy measurements, λ_{ex} = 280 nm and λ_{em} = 345 nm, and for recording AEDANS fluorescence anisotropy following parameters were set: λ_{ex} = 340 nm and λ_{em} = 500. The experiments were repeated at least three times to obtain the standard deviation.

The aggregation kinetic studies of RPT were recorded on a POLARstar Omega plate reader spectrophotometer (BMG LABTECH, Germany) using black 96-well plates (Thermo Fisher Scientific) at 37 °C. In order to avoid evaporation of the solution, the plates were sealed using a clear polyolefin sealing tape (Thermo Fisher Scientific). The final concentration of RPT used for all of the fluorescence studies was 20 μM in 20 mM sodium citrate buffer at pH 4. The final concentration of ThT used for all of the experiments was 10 μM. The reaction volume used for all of the aggregation kinetic studies was 200 μL and the samples, whenever required were agitated using a tiny glass bead (3 mm diameter; Fisher Scientific) at 600 rpm. For recording ThT fluorescence, the samples were excited at 440 nm and the emission was collected at 485±12 nm. Each experiment was repeated at least three times and the data were analyzed and plotted using Origin 2018 software.

4.2.7 Time-resolved fluorescence measurements

The time-resolved fluorescence lifetime measurements were acquired using a time-correlated single photon counting (TCSPC) setup (Fluorocube, Horiba JobinYvon, NJ). The fluorescence intensity decays were collected at a magic angle of 54.7°. The peak count was fixed to 10,000

and a 375 nm laser diode was used to excite ANS. The emission monochromator was fixed at 475 nm with a slit of 8 nm. The fluorescence intensity traces were deconvoluted with respect to the instrument response function (IRF) which was collected using Ludox (colloidal silica). The width (FWHM) of IRF was ~ 250 ps. The traces were analysed by a sum of exponentials using nonlinear least square method to obtain the lifetime components. All of the intensity decays were fitted into the following tri-exponential equation:

$$I(t) = I_0 [\alpha_1 \exp (t/\tau_1) + \alpha_2 \exp (t/\tau_2) + \alpha_3 \exp(t/\tau_3)]$$

where I_0 is the intensity at time zero, α_1, α_2, α_3 represents the amplitude associated with different lifetimes and τ_1, τ_2, τ_3 are the respective lifetime components. The average fluorescence lifetime was calculated using the following equation:

$$\tau_{avg} = (\alpha_1\tau_1 + \alpha_2\tau_2 + \alpha_3\tau_3) / (\alpha_1 + \alpha_2 + \alpha_3)$$

All of the data were analyzed and fitted using DAS6 software provided with the instrument and were finally plotted using Origin 2018 software.

4.2.8 Raman spectroscopy

The Raman spectra were recorded on an inVia laser Raman microscope (Renishaw, UK). Freshly prepared RPT aggregates formed after the incubation of 80 μM of RPT in 20 mM sodium citrate, pH 4.0, 150 mM NaCl incubated at 37 °C, with agitation at 600 rpm for 12 h were then directly deposited on glass slides covered with aluminum foil and subsequently dried under a stream of nitrogen. The samples were then focused using a 100X objective lens (Nikon, Japan). An NIR laser (785 nm) was used to excite the samples with an exposure time of 10 s and 100 % laser power. The Raman scattered light was collected, and the Rayleigh scattered light was removed by using an edge filter of 785 nm. The Raman scattered light was then dispersed using a 1200 lines/mm diffraction grating and was detected by an air-cooled CCD detector. The data was acquired using Wire 3.1 software provided with the instrument. All the spectra were averaged over 100 scans. The Raman spectra were baseline corrected using the cubic spline interpolation method for eliminating the tilt and were then smoothed using Wire 3.1 provided with the instrument. The baseline corrected and smoothed Raman spectra were finally plotted using Origin.

4.2.9 Ultracentrifugation

The concentration of RPT in the soluble fractions (monomer pool) was measured by subjecting 500 μL samples to ultracentrifugation (Hitachi Micro Ultracentrifuge CS150FNX) for 45 min

100,000 rpm at 25 °C. After centrifugation, the supernatant was immediately removed using a pipette and was transferred to an autoclaved microcentrifuge tube. In order to avoid any contamination from the aggregate material, the supernatant was carefully removed from the top. The concentrations of the soluble fractions (monomer pool) were then estimated by absorbance at 280 nm, using molar extinction of 5500 M^{-1} cm^{-1}.

4.3 Results

4.3.1 Hofmeister salts modulates the oligomerization of RPT

Our previous studies showed that an intrinsically disordered domain of Pmel17, namely, RPT, composed of 10 imperfect repeats (Figure 1a) undergo rapid oligomerization and adopts a partially ordered state at pH 4 as described in Chapter 3.[33] In this work we aim at obtaining detailed molecular insight into the effect of Hofmeister ions on the secondary structure and oligomerization/aggregation of RPT. In order to investigate the changes in the secondary structural content as a function of different salts, we utilized far-UV circular dichroism (CD) spectroscopy. Upon addition of RPT to the sodium citrate buffer, pH 4, containing different Hofmeister salts, we observed a profound change in the secondary structural elements in the presence of kosmotropes as compared to chaotropes (Figure 1b). Further, we plotted the ratiometric mean residue ellipticity ($\theta_{218}/\theta_{200}$) as a function of salts to check the efficacy of these salts to form β-sheets at the expense of random coil (inset in Figure 1b). We observed that the extent of increase in the β-sheet content was greater in the case of kosmotropes than chaotropes (inset in Figure 1b). However, at pH 5, the conformation of RPT remains unaltered in the presence of different salts (Figure 2a). Next, to gain further insight into the oligomerization process and to characterize the initial oligomers, we performed steady-state fluorescence measurements. For our studies, we used 9-(2,2-Dicyanovinyl) julolidine (DCVJ), a molecular rotor that is known to bind to early amyloidogenic oligomers.[35] We observed an increase in the DCVJ intensity that scales with the forward Hofmeister series. The oligomers formed in the presence of kosmotropes exhibited a much stronger increase in the DCVJ fluorescence as compared to the ones formed in the presence of chaotropes (Figure 1c). On the contrary, DCVJ fluorescence was absent in the case of pH 5 (Figure 2b). We next performed steady-state fluorescence measurements and recorded the intensity as well as anisotropy of W423 present in the C-terminal amyloid core region of RPT at pH 4 (Figure 1d). In the presence of kosmotropes, W423 showed a blue-shift in the emission and exhibited higher anisotropy (due to rotational constrain), indicating that the C-terminal region harboring W423 gets buried

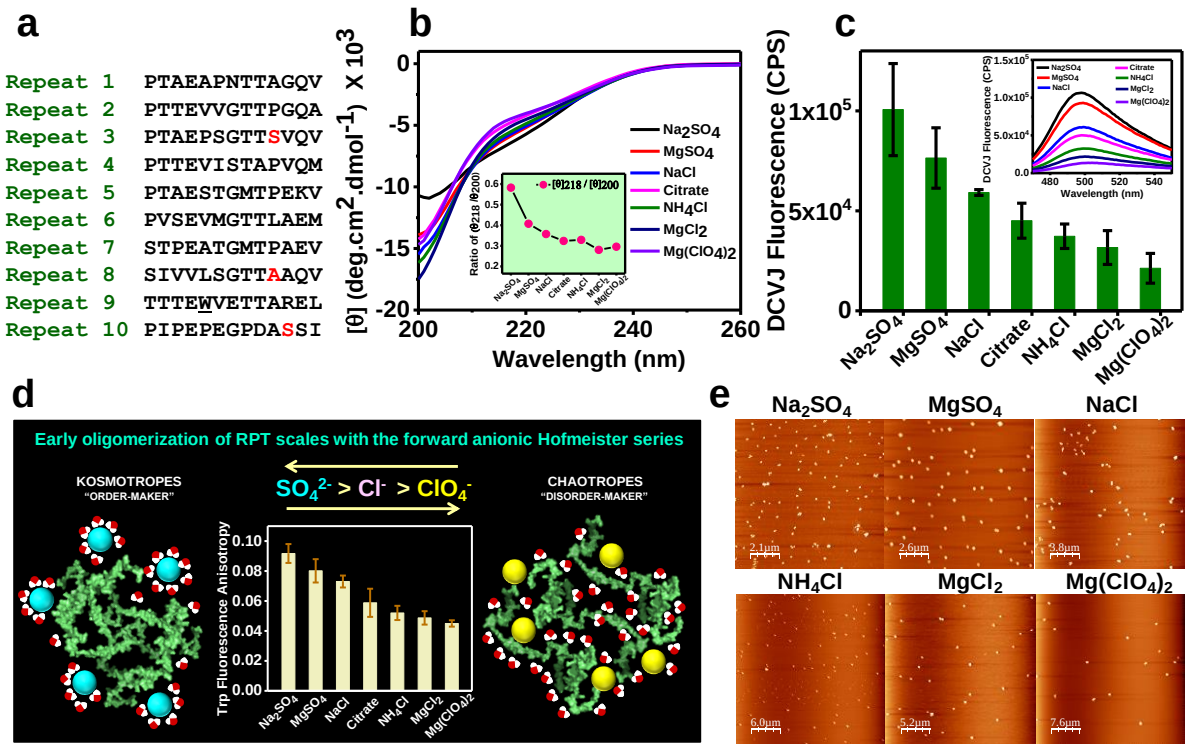


Figure 4.1 Effect of Hofmeister ions on the oligomerization of RPT at pH 4. (a) Amino acid sequence of the RPT. Mutated amino acid residues used in this study, S350, A415, and S442, are colored red. The single native tryptophan (W423) is underlined. (b) Far-UV CD spectra of 15 μM RPT when added immediately (mixing time is 1 min) to 20 mM sodium citrate buffer containing different salts at a concentration of 150 mM. The inset shows a plot of ratiometric ellipticity ($\theta_{218}/\theta_{200}$) as a function of different salts. (c) The extent of oligomerization from DCVJ fluorescence intensity at 500 nm as a function of different salts. The inset shows DCVJ fluorescence spectra as a function of salts. (d) Schematic showing the effect of kosmotropes and chaotropes on the oligomerization of RPT. The changes in the steady-state fluorescence anisotropy of W423 as a function of salt type. Error bars are obtained from at least three independent measurements. (e) AFM images of RPT oligomers in the presence of different salts imaged immediately after mixing RPT into the respective buffers.

during oligomerization, whereas, in the case of chaotropes, a red-shift in the emission and lower anisotropy values (due to rotational freedom) suggested that W423 is solvent-exposed (Figure 1d,2c).[36,37] Similar to chaotropes, RPT at pH 5 showed a highly red-shifted emission maxima and lower anisotropy values akin to the monomeric state of RPT (Figure 2d). Additionally, the extent of the burial of W423 upon oligomerization was revealed by the spectral shift ($\Delta\lambda$), the difference between emission maxima of the aggregated and monomeric state of RPT (Figure 2e). Next, we utilized atomic force microscopy, to gain insights into the 3D nanoscale morphology of the early aggregating species (Figure 1d). The AFM images revealed the formation of spherical oligomers, that were found to be greater in number for kosmotropes than

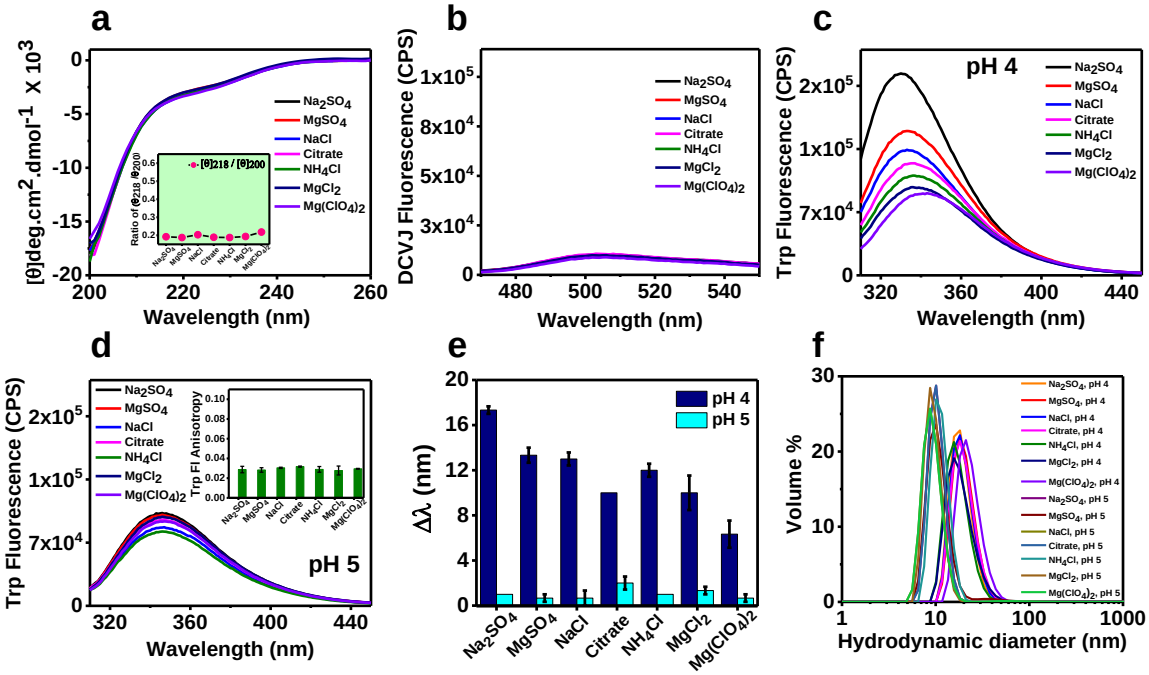


Figure 4.2 Structural investigation of the RPT at pH 4 and pH 5. (a) Far-UV CD spectra of 15 μM RPT in the presence of different salts at a concentration of 150 mM at pH 5.0, acquired immediately after mixing. The inset shows a plot of ratiometric ellipticity ($\theta_{218}/\theta_{200}$) as a function of different salts. Fluorescence spectra as a function of different salts, (b) DCVJ at pH 5 [note that the scale on the y-axis is kept same as that of pH 4 for comparison], (c)Tryptophan (W423) at pH 4, and pH 5 (d). The inset of panel d shows W423 fluorescence anisotropy in the presence of different salts at pH 5. (e) The spectral shift Δλ (difference in the W423 emission maxima between aggregated and monomeric state of the RPT) at pH 4 and pH 5. All data are shown as mean ± S.D. (n = 3). (f) The size distributions of RPT at pH 4.0 and pH 5.0 in the presence of 150 mM of different salts immediately after mixing.

chaotropes (Figure 1d). In order to monitor the size of early-stage aggregates (oligomers), we used dynamic light scattering. The result suggests a higher size distribution (hydrodynamic radius of ~8.5 nm) for oligomers formed at pH 4, whereas at pH 5, the hydrodynamic radius was ~4-5 nm (Figure 2f). Together, these findings revealed that all salts examined, promote oligomerization, and the effect is in line with the Hofmeister series with the most pronounced effect seen within the anions. These results suggest that the partial disordered state of RPT at pH 4 has a preferential binding for different ions that promotes oligomerization than the completely disordered state at pH5. We next asked: What are the critical interactions responsible for RPT oligomerization at pH 4?

4.3.2 Electrostatic and hydrophobic interactions govern RPT oligomerization

Electrostatic interactions are known to affect protein structure, folding, binding, and aggregation.[38] Alteration in charge of proteins influence protein-protein interactions and are often responsible for the mechanistic switch in protein aggregation pathways.[33] In order to

understand the role of electrostatic interaction in the oligomerization of the RPT, we carried out multi-parametric fluorescence measurements in conjunction with CD spectroscopy as a function of salt concentration. Using CD spectroscopy, we monitored the effect of ionic strength on the secondary structure upon the oligomerization of the RPT (Figure 3a). The increase in NaCl concentration was accompanied by a considerable increase in the β-sheet content and a substantial decrease in the disordered content at pH 4, as revealed by the ratiometric plot of mean residue ellipticity as a function of NaCl concentration (inset in Figure 3a). These results indicate that the rate of formation of the β-sheet is affected by increasing salt concentration. However, in the case of pH 5, no change in the secondary structural element was observed upon varying the ionic strength (Figure 4a). To compare the extent of formation of β-sheet among different Hofmeister salts, we have varied the salts at two different concentrations i.e., 100 and 400 mM. The results, as depicted by the 3 D plot (Figure 3b), reveal that the ionic strength- dependent change in conformation was more predominant in the case of kosmotropes than chaotropes. Since W423 is an amyloid reporter of RPT, we next performed steady-state fluorescence anisotropy measurements of W423 by varying ionic strength. We observed an increase in the anisotropy values as a function of ionic strength, indicating that higher ionic strength accelerates the oligomerization of RPT, most likely due to charge screening (Figure 3c). This set of experiments indicates that electrostatic interaction plays a key role in triggering the oligomerization process. The presence of hydrophobic patches on the surface of proteins is known to promote the oligomerization process. Since the primary sequence of RPT is composed of a large number of hydrophobic residues, we conjectured that the oligomerization could be hydrophobically driven. In order to establish this, we used 8-anilino-1-naphthalene sulfonic acid (ANS), a highly hydrophobic environment-sensitive fluorescent probe that exhibits high quantum yield when bound to hydrophobic patches/clusters of protein.[35] ANS in the free form is non-fluorescent in aqueous solutions (λ_{em} ~525 nm); however, upon binding to hydrophobic pockets, it becomes highly fluorescent and undergoes a significant blue shift (λ_{em} ~475 nm). Due to these unique properties, ANS has been widely used for the investigation of initial oligomerization, and subsequent growth steps that are involved in the formation of amyloid fibrils.[38] We recorded the steady-state fluorescence of ANS as a function Hofmeister salts and observed an increase in the ANS fluorescence intensity along with a huge blue shift in the presence of kosmotropes. On the contrary, the emission maxima of ANS was red-shifted in the presence of chaotropes and showed a very little increase in the intensity (Figure 3d). Additionally, RPT in the presence of different salts at pH 5 did not show any ANS binding, indicating the absence of rapid oligomerization at pH 5 (Figure 4b).

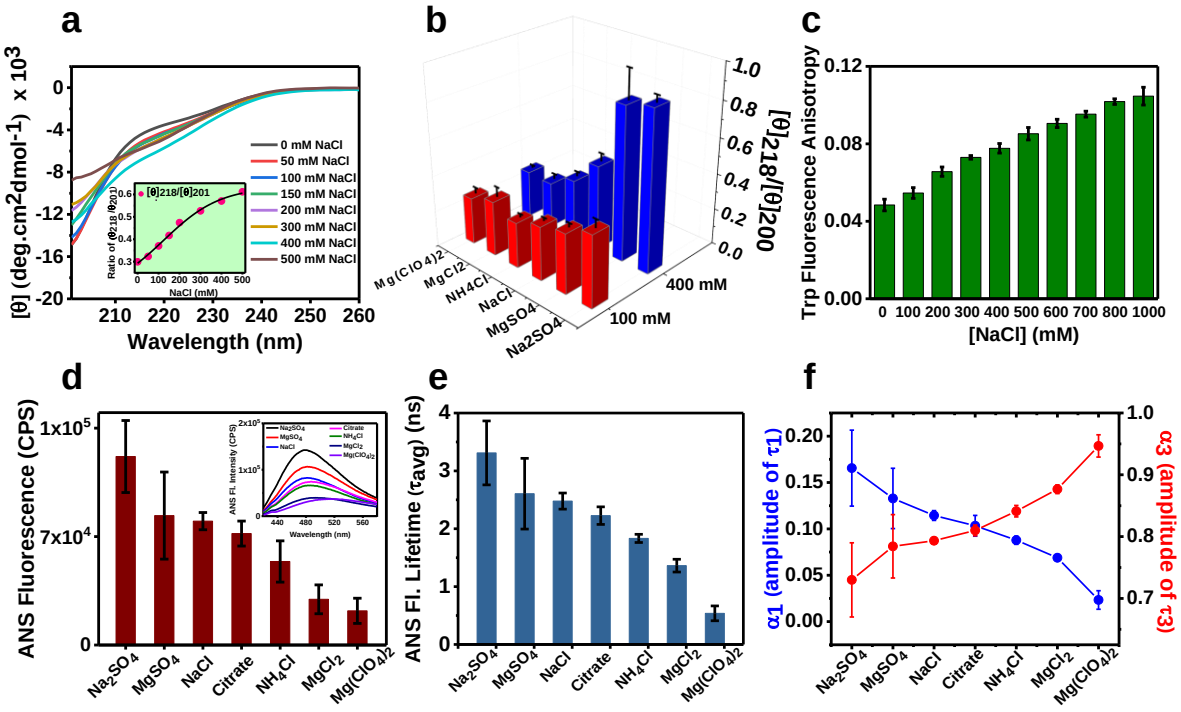


Figure 4.3 Molecular interactions that governs RPT oligomerization at pH 4. (a) Far-UV CD spectra of 15 μM RPT as a function of NaCl concentration: 0, 50, 100, 150, 200, 300, 400, and 500 mM upon mixing RPT with 20 mM sodium citrate buffer containing respective NaCl concentration at pH 4.0. The inset shows a plot of ratiometric ellipticity ($\theta_{218}/\theta_{201}$) as a function of ionic strength. (b) The steady-state fluorescence anisotropy of W423 as a function of salt concentration. (c) A 3D column plot for the comparison of ratiometric ellipticity as a function of different salts at two different salt concentrations of 100 and 400 mM. (d) Comparison of the extent of formation of hydrophobic pockets as a function of different salts from the steady-state ANS fluorescence at 475 nm. The inset shows ANS fluorescence spectra as a function of salts. (e) The average lifetime of ANS fluorescence recovered from the time-resolved fluorescence intensity decays, indicating the salt-dependent formation of hydrophobic pockets. (f) Changes in the contribution of the longer- and shorter-lifetime components of ANS. The error bars denote mean ± S.D. obtained from at least three independent measurements.

Besides fluorescence intensity, the changes in the fluorescence lifetime of ANS provide further information regarding the solvent-exposed or buried hydrophobic patches of the oligomeric aggregates. We next performed time-resolved fluorescence measurements and recorded the fluorescence lifetime of ANS bound to the distinct conformation of RPT formed using different Hofmeister salts. The recovered fluorescence lifetime components obtained after fitting the fluorescence intensity decays using the tri-exponential function are shown in Table 4. The average lifetime of ANS fluorescence as a function of salts at pH 4 and pH 5 is shown in Figure 3e and Figure 4c, respectively. In the case of pH 4, upon moving from kosmotropes to chaotr-

Table 4.

Comparison of the fluorescence lifetime components of ANS bound to RPT in the presence of different salts at pH 4.0

Salt type	τ_1 (α_1)	τ_2 (α_2)	τ_3 (α_3)	τ_{avg}
Na_2SO_4	16.02 (0.16)	4.96 (0.11)	0.21 (0.73)	3.3
$MgSO_4$	15.89 (0.13)	4.95 (0.09)	0.22 (0.78)	2.6
NaCl	15.88 (0.11)	4.91 (0.09)	0.24 (0.79)	2.5
Citrate	15.59 (0.10)	4.74 (0.09)	0.24 (0.81)	2.2
NH_4Cl	15.18 (0.09)	4.24 (0.07)	0.22 (0.84)	1.8
$MgCl_2$	14.10 (0.07)	3.76 (0.05)	0.22 (0.88)	1.4
$Mg(ClO_4)_2$	11.18 (0.02)	2.24 (0.03)	0.21 (0.95)	0.5

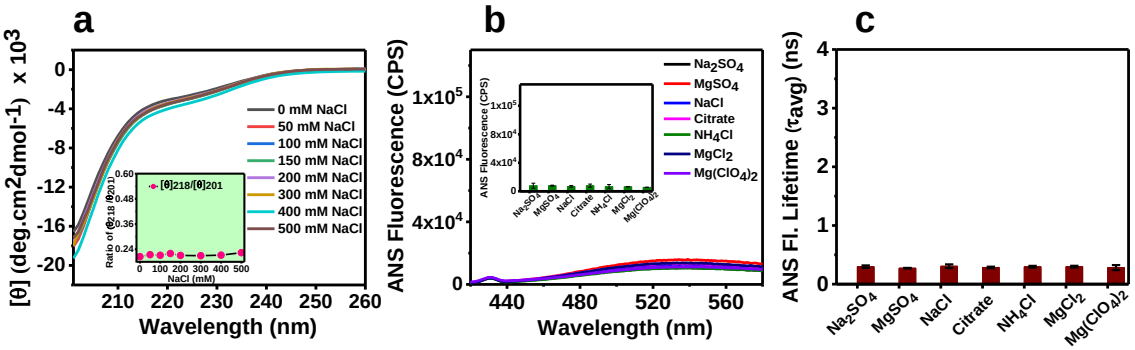


Figure 4.4 The initial conformational state of RPT at pH 5 is insensitive to both electrostatic and hydrophobic interactions. (a) Far-UV CD spectra of 15 μM RPT as a function of NaCl concentration. The inset shows a plot of ratiometric ellipticity ($\theta_{218}/\theta_{201}$) as a function of ionic strength. (b) ANS fluorescence spectra as a function of salt concentration. The inset shows ionic strength dependent steady-state fluorescence intensity of ANS at 475 nm. (c) The average lifetime of ANS fluorescence recovered from the time-resolved fluorescence intensity decays. Note that the scale on the y-axis for panel b and c is kept same as that of pH 4 for comparison. The error bars denote mean ± S.D. obtained from at least three independent measurements.

-otropes, a sharp decrease in the average fluorescence lifetime indicates that either the oligomeric species formed in the presence of chaotropes possess low hydrophobicity and are exposed to the solvent molecules or there is little population of the oligomeric species. In

contrast, no change in the average fluorescence lifetime of ANS was observed at pH 5. Typically, ANS in an aqueous solvent (water) exhibits a monoexponential decay with a fluorescence lifetime of 0.25 ns (τ_3).[39] However, when bound to proteins, it shows two longer lifetime components (τ_1 and τ_2), where, the longer lifetime component (τ_1, τ >10 ns) corresponds to the solvent protected hydrophobic cores and the shorter lifetime component (τ_2, τ < 10 ns) indicates the presence of solvent-exposed hydrophobic patches.[39] A careful look at the recovered fluorescence lifetimes of ANS revealed that upon varying the salts (kosmotropes to chaotropes), there was a significant decrease in the lifetime of these two components from ~ 16 ns to ~11 ns and ~5 ns to ~ 2 ns, indicating that the hydrophobic pockets are becoming more solvent-exposed. Additionally, the contribution (α_1) of the longest lifetime component (τ_1) decreases from ~ 16 % to ~ 2%, whereas, the contribution shortest lifetime (α_3) increases from ~ 73 % to ~ 95 % (Figure 3f). These results suggest that RPT oligomers in the presence of kosmotropes form solvent protected hydrophobic patches, whereas the oligomers formed in the presence of chaotropes are exposed to solvent molecules. Together, these findings revealed that there is a formation of hydrophobic patches during RPT oligomerization, and the nature of these patches varies as a function of Hofmeister salts that modulate the solvation properties around the oligomers. We next directed our efforts in decoding the hierarchical order of distinct regions of RPT that drives the oligomerization process and the influence of different salts on the early contact formation?

4.3.3 Region-specific early contact formation during oligomerization

In order to decipher the early contacts formed during RPT oligomerization in a site-specific manner, we took advantage of the fact that RPT is devoid of cysteines and created single cysteine mutations at various positions encompassing the distinct regions including the N-terminal region (residue position 350), the amyloid core region (residue position 415), and the C-terminal region (residue position 442). We then covalently label the single-cysteine mutants with an environment-sensitive thiol-reactive fluorescent dye, namely, (5-((((2-Iodoacetyl)amino)ethyl) amino)Napthalene-1-Sulphonic acid) (IAEDANS). In order to gain region-specific information about the rapid oligomerization process and the role of salt in modulating these early associations, we recorded the steady-state fluorescence anisotropy of AEDANS labeled RPT in the presence of different salts (Figure 5). As expected, we observed a higher anisotropy in the of case Cys415 compared to Cys350 and Cys442, irrespective of different salts. The comparatively low anisotropy values at the N- and C-terminal were suggestive of a higher degree of rotational flexibility as compared to the core region. However,

the anisotropy values at all of these positions scaled with the forward Hofmeister series, with the most profound effect seen in case Cys350. Therefore, the high specificity of ions for the N-terminal region indicates that this region of RPT might be sensitive to the associations that are electrostatically driven. Additionally, we measured the fluorescence anisotropy at these different regions in the monomeric state of RPT at pH 7.4 in the presence of different salts (Figure 4). We did not observe any change in the anisotropy values as a function of different

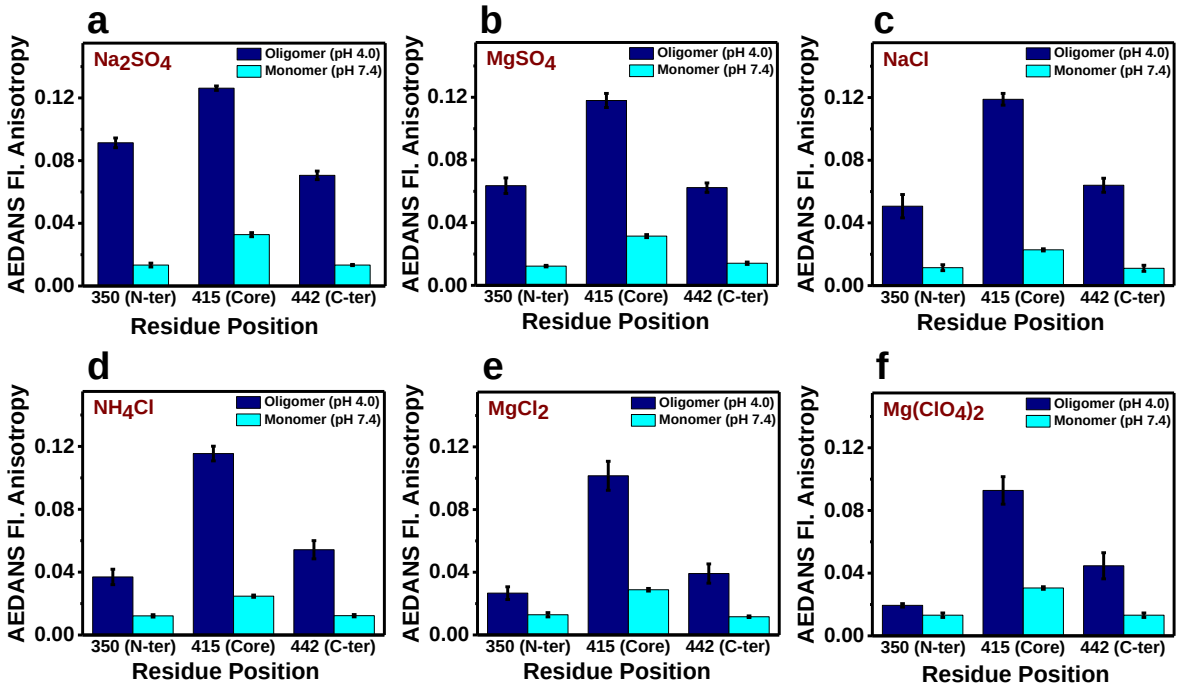


Figure 4.5 Site-specific early contact formation during the oligomerization process at pH 4. Steady-state fluorescence anisotropy of AEDANS labeled RPT at three different positions, namely, Cys350 (N-terminal), Cys415 (Core), and Cys442 (C-terminal) in the presence of 150 mM (a) Na_2SO_4, (b) $MgSO_4$, (c) NaCl, (d) NH_4Cl, (e) $MgCl_2$, and (f) $Mg(ClO_4)_2$, both in the oligomeric (navy blue) and monomeric state (cyan). The error bars denote mean ± S.D. obtained from at least three independent measurements.

salts; however, Cys415 showed a bit higher anisotropy value even in the native monomeric state, and we believe that it's due to lack of disordered promoting proline residue in this region. Overall, the rotational mobility map depicted in Figure 4 reveals that the oligomerization of RPT initiates in the proline-free amyloid-core region followed by the association of the N- and C-terminal regions. Together, these results reveal the hierarchical association of different regions of RPT during the oligomerization process and also indicate that salt ions greatly influence early contact formation in the N-terminal region. We next asked the following

question: What effect does the salt-induced distinct oligomerization behavior have on the aggregation kinetics of RPT?

4.3.4 Forward and inverse effect of Hofmeister ions on the aggregation pathway of RPT

To delineate the effect of Hofmeister ions on the aggregation pathway of RPT, we initiated our studies by monitoring the aggregation kinetics of RPT in the presence of different ions, using Thioflavin T (ThT), a well-known amyloid marker. For these studies, we selected three different salts (Na_2SO_4, NaCl, $Mg(ClO_4)_2$) from the Hofmeister series. Interestingly, the ThT fluorescence kinetic profiles revealed the biphasic nature of the aggregation process, where the first phase (P_1) of the kinetics follows an isodesmic polymerization mechanism and the rate of formation of the initial aggregate species scales with the forward Hofmeister series (Figure 6a). We observed that in the presence of Na_2SO_4, the aggregation reaction reaches the first saturation phase faster than NaCl, whereas, $Mg(ClO_4)_2$ took the longest time (inset in Figure 6a). Surprisingly, $Mg(ClO_4)_2$ enters into phase 2 (P_2) faster than NaCl and Na_2SO_4, and this trend is in line with the reverse Hofmeister series (Figure 6a). Therefore, these results suggest that the biphasic behavior of the aggregation kinetics is a consequence of dual forward and reverse Hofmeister effect. Next, to investigate the changes in the secondary structure at different phases, i.e., the onset, phase 1 and phase 2 of the aggregation kinetics, we utilized far-UV CD spectroscopy (Figure 6b,c,d). We then plotted the ellipticity ratio ($\theta_{218}/\theta_{200}$) to estimate the β-sheet content at the expense of random-coil. The plot showed an interesting trend, where both at the onset of the reaction and in phase 1 (P_1) kosmotrope (Na_2SO_4) showed higher β-sheet content than the chaotrope ($Mg(ClO_4)_2$). However, in the second phase, this trend was reversed (Figure 6e). Additionally, we carried out vibrational Raman spectroscopy studies at these different phases. We observed an increase in the intensity of amide I (1670 cm^{-1}) band as the evolution of amyloid formation, which suggests the formation of a typical cross-β amyloid structure (Figure 6f). This piece of data corroborated our CD results. Next, we performed atomic force microscopy (AFM) to visualize the nanoscale morphologies of phase 1 and phase 2 aggregates as a function of different salts. The AFM images acquired in phase 1 revealed the formation of dendritic nanostructures or fractal networks in the presence of different salts. However, in the case of Na_2SO_4, a large number of oligomers were also observed along with the structures mentioned above (Figure 6g,h,i). In the case of phase 2, the AFM images showed thread-like morphologies of the amyloid fibrils (Figure 6g,h,i). Since stirring introduces secondary processes such as fragmentation, we conjectured that the second phase in the aggregation kinetics could be due to fragmentation. In order to establish this, we performed

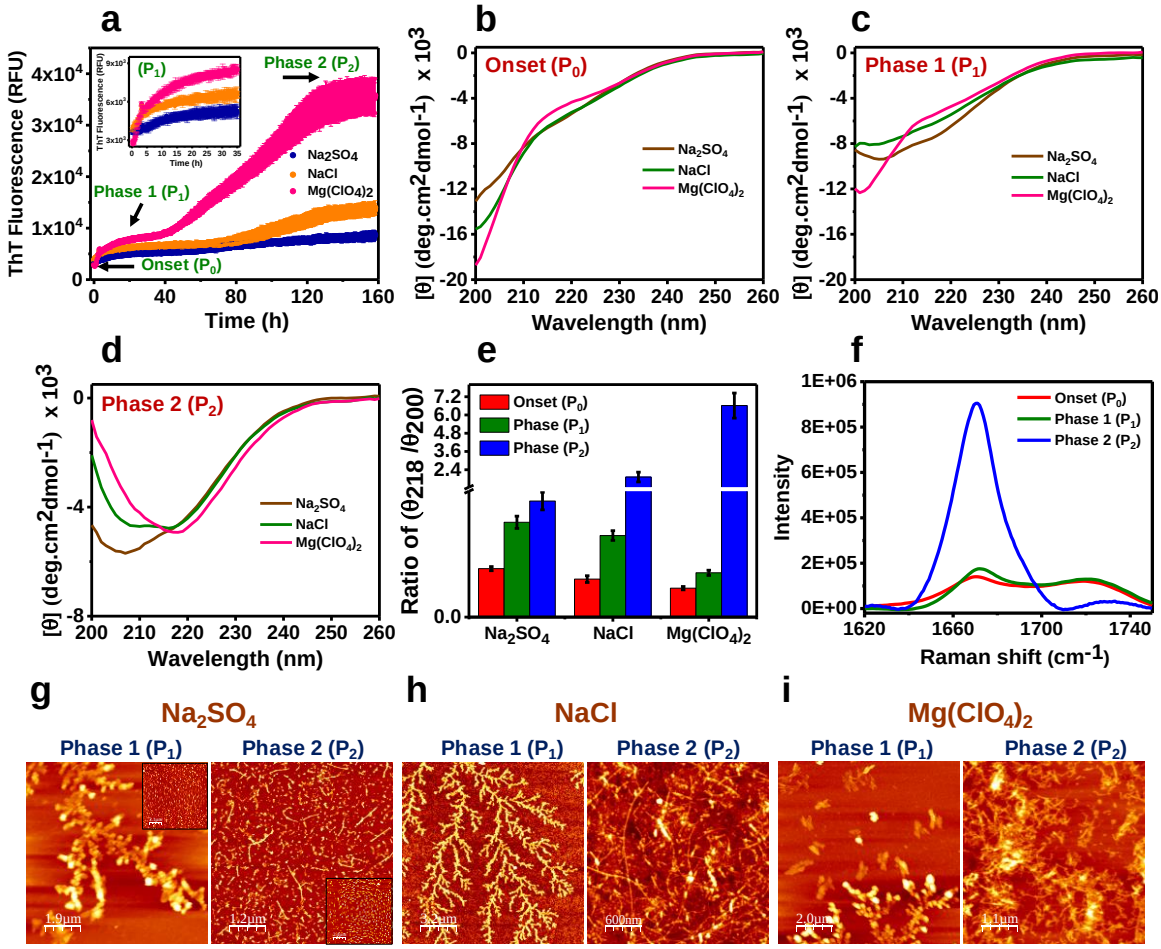


Figure 4.6 Aggregation kinetics of RPT investigated at pH 4 in the presence of Na_2SO_4, NaCl, and $Mg(ClO_4)_2$ at 37 °C and 600 rpm. (a) ThT fluorescence kinetics of RPT aggregation (20 μM) in the presence of different salts shows a complex biphasic behavior. The inset shows the data till 35 h for better visualization. Far-UV CD spectra of RPT collected at (b) t_{0h}, onset (P_0), (c) t_{12h}, phase 1 (P_1), and (d) t_{160h}, phase 2 (phase 2) of RPT aggregation as a function of salt type. (e) Ratiometric ellipticity ($\theta_{218}/\theta_{200}$) as a function Hofmeister ions indicate salt-dependent evolution of highly ordered β-sheet rich conformation. (f) Raman spectra showing amide I (1620-1750 cm^{-1}) region at the onset (P_0), phase 1 (P_1), and phase 2 (phase 2) of RPT aggregation pathway, in the presence of 150 mM NaCl indicates a transition from partially ordered state to completely ordered β-sheet rich state. AFM images of RPT polymorphic amyloid nanostructures at pH 4 upon incubation at 37 °C and 600 rpm. (g) with 150 mM Na_2SO_4, the insets show the co-existence of oligomers, (h) with 150 mM NaCl, and (i) with 150 mM $Mg(ClO_4)_2$. The AFM images reveal diversity in amyloid fibrillar morphologies in the presence of different salts. See the experimental section for details of AFM image acquisition and processing. The error bars denote mean ± S.D. obtained from at least three independent measurements.

the aggregation studies under quiescent conditions (Figure 7a) and indeed observed that the second phase was absent in the aggregation kinetics. Next, we set out to determine the role of other secondary processes, such as secondary nucleation in the aggregation process. Since

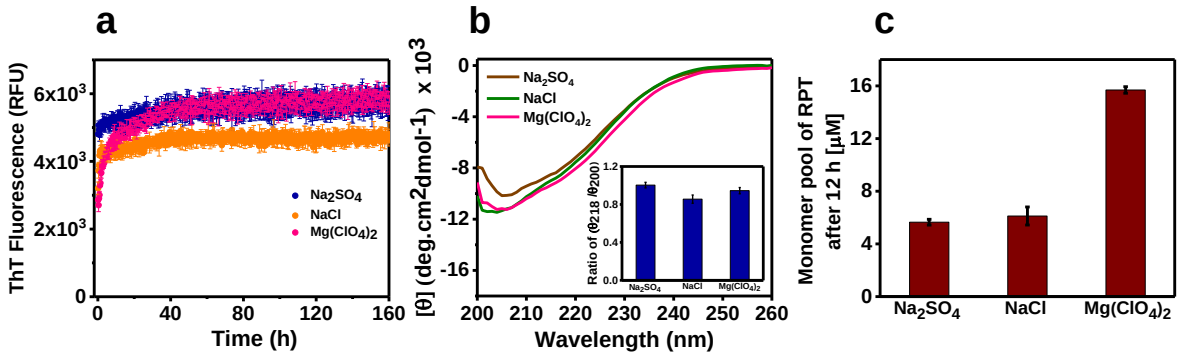


Figure 4.7 Aggregation kinetics of RPT investigated at pH 4 in the presence of Na_2SO_4, NaCl, and $Mg(ClO_4)_2$ at 37 °C under quiescent conditions. (a) ThT fluorescence kinetics of RPT aggregation (20 μM) in the presence of different salts. (b) Far-UV CD spectra of RPT collected at t_{160h} of RPT aggregation as a function of salt type. The inset shows ratiometric ellipticity ($\theta_{218}/\theta_{200}$) as a function Hofmeister ions. (c) Estimation of monomer pool of RPT present in the reaction mixtures containing different salts, at phase 1 (P_0) using ultracentrifugation. See the experimental section for details. The error bars denote mean ± S.D. obtained from at least three independent measurements.

secondary nucleation is dependent on the monomer pool of protein, we next estimated the soluble fraction of RPT from the reaction mixture of Phase 1 (Figure 7c). The monomer pool observed was higher in the case of chaotropes than kosmotropes. Together, these findings reveal a unique dual Hofmeister effect that is responsible for the biphasic nature of the aggregation process.

4.4 Discussion

In this work, we have explored the effect of ions on the oligomerization and aggregation process of the RPT at acidic pH, typical of melanosomes. In the presence of kosmotropes, RPT forms oligomers at a much faster rate compared to the chaotropes. Our multiparametric fluorescence studies in conjunction with CD spectroscopy and AFM revealed that the oligomerization of RPT in the presence of salts scales with the forward Hofmeister series: SO_4^{2-} > Cl^- > ClO_4^-. Moreover, the anions have a stronger effect on the oligomerization (Na_2SO_4 > NaCl, $MgSO_4$ > $MgCl_2$), which could be due to the fact that the anions screen and minimize/neutralize the charge repulsion of positively charged RPT at pH 4 (pI ~4.5), thereby promoting the oligomerization process. However, in the case of pH 5, RPT carries a net negative charge, and therefore due to charge repulsion, it does not undergo rapid oligomerization. The electrostatic interactions, in conjunction with the hydrophobic interactions, appears to drive the early oligomerization process of RPT. At pH 4, the positive charge on RPT is neutralized by the anions, and due to this charge neutralization, the hyropho-

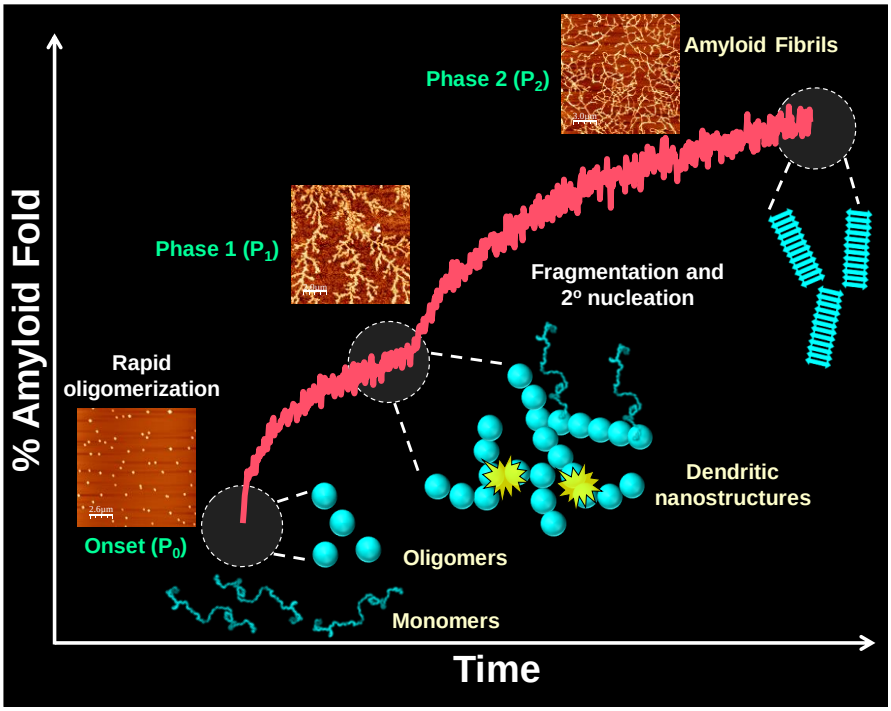


Figure 4.8 Schematic representation of the proposed RPT aggregation pathway at pH 4 in the presence of Hofmeister ions. See the discussion for more details.

-bic interactions become highly favorable for the self-assembly process. Similar molecular interactions are also observed in the oligomer formation of a peptide, namely, peptibody A.[40] Furthermore, the initiation of RPT oligomerization occurs in the proline-free amyloid core region. A plethora of studies have revealed that the salts perturb the water structure. In general, if in a system, the salts alter the water structure, then the effect is expected to follow the Hofmeister series. We observed that for all salts examined, the anions followed the same order as it is ranked in the Hofmeister series, which reaffirms the role of Hofmeister effect in RPT oligomerization. Additionally, ANS fluorescence lifetimes studies also showed a distinct solvent environment around the oligomers formed in the presence of kosmotropes and chaotropes. Unlike the quiescent condition in which the aggregation was solely driven by the charge peptide interaction, stirring the solution resulted in a significant modification of the route of aggregation due to fragmentation, which leads to the proliferation of the total number of aggregates and secondary nucleation. The salt-specific effect modulates the kinetics by introducing a biphasic behavior. Different ions modulated the autocatalytic proliferation of aggregates, which further affected the rate of secondary nucleation. In the case of $Mg(ClO_4)_2$ multiplication of aggregates occurs at a much faster than Na_2SO_4, largely as a consequence of much more rapid secondary nucleation. We propose that due to prolonged stirring there might be the dissociation of dendritic nanostructures into the fragmented aggregate species that can

act as a seed, onto which monomers in the solution can bind and initiate the self-assembly of RPT into thread-like fibrils. On the basis of these observations, we propose a model that describes the possible nucleation events involved in the aggregation pathway of RPT at pH 4 in the presence of Hofmeister ions (Figure 8). For certain proteins, the structure and morphology of the amyloid fibrils formed does not depend on the different salts. However, in the case of RPT, salt modulates the aggregation pathway and can induce structural changes. The morphology of fibrils also changed during aggregation revealed by the AFM images. Salts have a profound effect on the propensity of aggregation and on the structures of the resulting aggregates or fibrils. These effects are specifically relevant because salts occur ubiquitously in physiological environments and under in vivo conditions of amyloid fibril formation. In summary, our findings illustrate an intriguing interplay of Debye-Huckel screening, Hofmeister effect, fragmentation, and secondary nucleation during the RPT aggregation process. After rapid oligomerization, RPT forms dendritic nanostructures or a fractal network, which upon fragmentation via the process of secondary nucleation, organizes into an amyloid structure. A similar kind of morphological transition has been observed in the case of Silicatein protein.[41] We believe that these mechanistic underpinning will be of importance in understanding the complex kinetics of a variety of other IDPs/IDRs.

4.5 References:

(1) 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, 6589-6631.

(2) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D^2 concept. *Annu. Rev. Biophys.* **2008,** 37, 215-246.

(3) Theillet, F.; Binolfi, Andres.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.; Crowley, P. B.; Gierasch, L.; Pielak, G. J.; Elcock, A. H.; Gershenson, A.; Selenko, P. Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chem. Rev.* **2014,** 114, 6661-6714.

(4) Okur, H. I.; Hladilkova, J.; Rembert, K. B.; Cho, Y.; Heyda, J.; Dzubiella, J.; Cremer, P. S.; Jungwirth, P. Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **2017,** 121, 1997-2014.

(5) Zhang, Y.; Cremer, P. S. Chemistry of Hofmeister anions and osmolytes. *Annu. Rev. Phys. Chem.* **2010,** 61, 63-83.

(6) Zhang, Y.; Cremer, P. S. Interactions between macromolecules and ions: The Hofmeister series. *Curr. Opin. Chem. Biol.* **2006,** 10, 658-63.

(7) Lo Nostro, P.; Ninham, B. W. Hofmeister phenomena: an update on ion specificity in biology. *Chemical reviews* **2012,** 112, 2286-322.

(8) Gurau, M. C.; Lim, S. M.; Castellana, E. T.; Albertorio, F.; Kataoka, S.; Cremer, P. S. On the mechanism of the hofmeister effect. *J. Am. Chem. Soc.* **2004,** 126, 10522-3.

(9) Okur, H. I.; Kherb, J.; Cremer, P. S. Cations Bind Only Weakly to Amides in Aqueous Solutions. *J. Am. Chem. Soc.* **2013,** 135, 5062-5067.

(10) Klement, K.; Wieligmann, K.; Meinhardt, J.; Hortschansky, P.; Richter, W.; Fandrich, M. Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils. *J. Mol. Biol.* **2007,** 373, 1321-1333.

(11) Munishkina, L. A.; Henriques, J.; Uversky, V. N.; Fink, A. L. Role of protein-water interactions and electrostatics in alpha-synuclein fibril formation. *Biochemistry* **2004,** 43, 3289-3300.

(12) Jain, S.; Udgaonkar, J. B. Salt-induced modulation of the pathway of amyloid fibril formation by the mouse prion protein. *Biochemistry* **2010,** 49, 7615-7624.

(13) Rubin, J.; Khosravi, H.; Bruce, K. L.; Lydon, M. E.; Behrens, S. H.; Chernoff, Y. O.; Bommarius, A. S. Ion-specific effects on prion nucleation and strain formation. *J. Biol. Chem.* 2013, 288, 30300-8.

(14) Sengupta, I.; Bhate, S. H.; Das, R.; Udgaonkar, J. B. Salt-Mediated Oligomerization of the Mouse Prion Protein Monitored by Real-Time NMR. *J. Mol. Biol.* **2017,** 429, 1852-1872.

(15) Yeh, V.; Broering, J. M.; Romanyuk, A.; Chen, B.; Chernoff, Y. O.; Bommarius, A. S. The Hofmeister effect on amyloid formation using yeast prion protein. *Pro. Sci.* **2010,** 19, 47-56.

(16) Marek, P. J.; Patsalo, V.; Green, D. F.; Raleigh, D. P. Ionic strength effects on amyloid formation by amylin are a complicated interplay among Debye screening, ion selectivity, and Hofmeister effects. *Biochemistry* **2012,** 51, 8478-90.

(17) Krishnan, S.; Raibekas, A. A. Multistep Aggregation Pathway of Human Interleukin-1 Receptor Antagonist: Kinetic, Structural, and Morphological Characterization. *Biophys. J.* **2009,** 96, 199-208.

(18) Wood, S. J.; J. Wypych; S. Steavenson; J. C. Louis; M. Citron; Biere, A. L. 1999. α-synuclein fibrillogenesis is nucleation-dependent. Implications for the pathogenesis of Parkinson's disease. *J. Biol. Chem.* **1999,** 274, 19509-19512.

(19) Ferrone, F. Analysis of protein aggregation kinetics. *Methods Enzymol.* **1999,** 309, 256-274.

(20) Meisl, G.; Kirkegaard, J. B.; Arosio, P.; Michaels, T. C.; Vendruscolo, M.; Dobson, C. M.; Linse, S.; Knowles, T. P. Molecular Mechanisms of Protein Aggregation from Global Fitting of Kinetic Models. *Nat. Protoc.* **2016,** 11, 252−272.

(21) 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.*, **2018,** 54, 8667-8684.

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

(23) Buell, A. K.; Galvagnion, C.; Gaspar, R.; Sparr, E.; Vendruscolo, M.; Knowles, T. P. J.; Linse, S.; Dobson, C. M. Solution conditions determine the relative importance of nucleation and growth processes in α-synuclein aggregation. *Proc. Natl. Acad. Sci. U.S.A.*, **2014,** 111, 7671-7676.

(24) Ruschak, A. M.; Miranker, A. D. Fiber-dependent amyloid formation as catalysis of an existing reaction pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104, 12341–12346.

(25) Cohen, S. I. A.; Linse, S.; Luheshi, L. M.; Hellstrand, E.; White, D. A.; Rajah, L.; Otzen, D. E.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. Proliferation of amyloid-β42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** 110, 9758–9763.

(26) Cobb, N. J.; Surewicz, W. K. Prion diseases and their biochemical mechanisms. *Biochemistry* **2009,** 48, 2574-85.

(27) Higurashi, T.; Hines, J. K.; Sahi, C.; Aron, R.; Craig, E. A. Specificity of the J-protein Sis1 in the propagation of 3 yeast prions. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** 105, 16596-16601.

(28) Meisl, G., Yang, X.; Dobson, C. M.; Linse, S.; Knowles, T. P. J. Modulation of electrostatic interactions to reveal a reaction network unifying the aggregation behaviour of the Aβ42 peptide and its variants. *Chem Sci.* **2017,** 8, 4352-4362.

(29) de Oliveira, G. A. P.; Silva, J. L. Alpha-synuclein stepwise aggregation reveals features of an early onset mutation in Parkinson's disease. *Commun. Biol.* 2019, 2, 1-13.

(30) McGlinchey, R. P.; Lee, J. C. Why Study Functional Amyloids? Lessons from the Repeat Domain of Pmel17 *J. Mol. Biol.* **2018,** 430, 3696-3706.

(31) Pfefferkorn, C. M.; McGlinchey, R. P.; Lee, J. C. Effects of pH on aggregation kinetics of the repeat domain of a functional amyloid, Pmel17. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** 107, 21447-21452.

(32) McGlinchey, R. P.; Jiang, Z.; Lee, J. C. Molecular origin of pH-dependent fibril formation of a functional amyloid. *ChemBioChem* **2014,** 15, 1569-1572.

(33) Dogra, P.; Bhattacharya, M.; Mukhopadhyay, S. pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid. *J. Phys. Chem. B* **2017,** 121, 412−419.

(34) Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. WSXM: A Software for Scanning Probe Microscopy and a Tool for Nanotechnology. *Rev. Sci. Instrum.* **2007,** 78, No. 013705.

(35) Lindgren, M.; Sörgjerd, K.; Hammarström, P. Detection and Characterization of Aggregates, Prefibrillar Amyloidogenic Oligomers, and Protofibrils Using Fluorescence Spectroscopy. *Biophys. J.* **2005,** 88, 4200-4212.

(36) Lakowicz, J. R. Principles of Fluorescence Spectroscopy, 3rd ed.; 745 Springer: New York, 2006.

(37) Majumdar, A.; Mukhopadhyay, S. Fluorescence Depolarization Kinetics to Study the Conformational Preference, Structural Plasticity, Binding, and Assembly of Intrinsically Disordered Proteins. *Methods Enzymol.* **2018,** 611, 347-381.

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

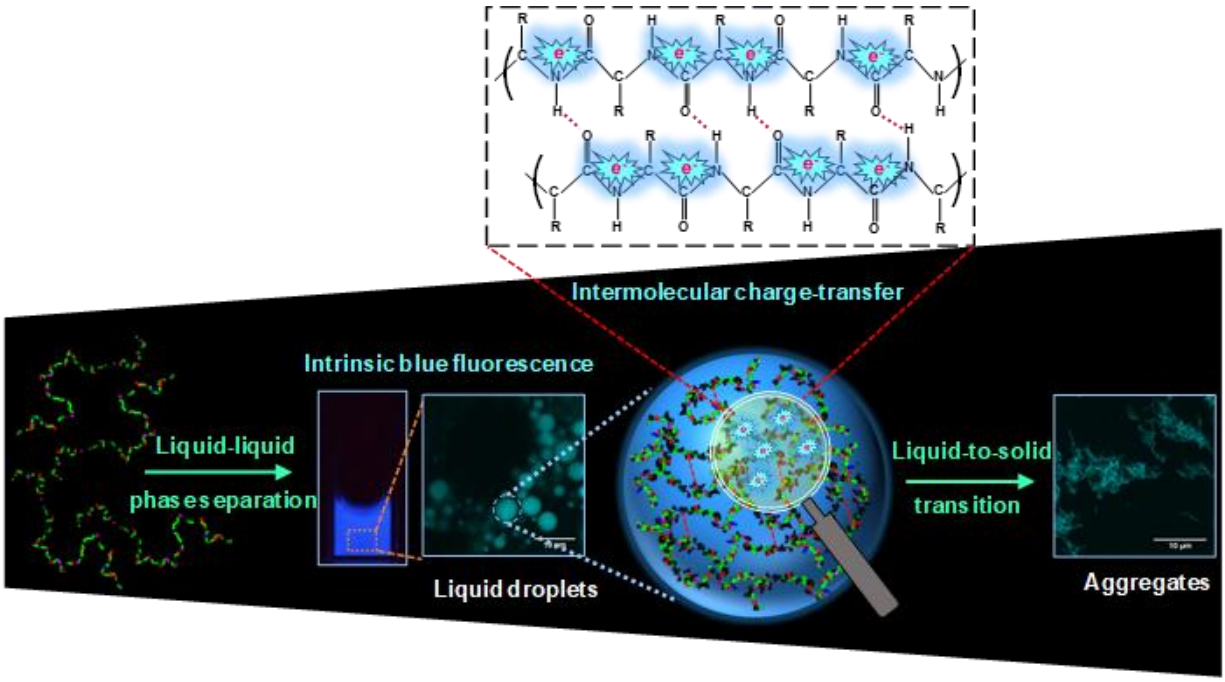
(38) Gasymov, O. K; Glasgow, B. J. ANS Fluorescence: Potential to Augment the Identification of the External Binding Sites of Proteins. *Biochim. Biophys. Acta* **2007,** 1774, 403-11.

(39) Uversky, V. N.; Winter, S.; Löber, G. Use of Fluorescence Decay Times of 8-ANS-Protein Complexes to Study the Conformational Transitions in Proteins which Unfold Through The Molten Globule State. *Biophys. Chem.* **1996,** 60, 79-88.

(40) Gokarn, Y. R.; Fesinmeyer, R. M.; Saluja, A.; Cao, S.; Dankberg, J.; Goetze, A.; Remmele, R. L. Jr.; Narhi, L. O.; Brems, D. N. Ion-specific modulation of protein interactions: Anion-induced, reversible oligomerization of a fusion protein. *Protein Sci.* **2009,** 18, 169-179.

(41) Murr, M. M.; Morse, D. E. Fractal intermediates in the self-assembly of silicatein filaments. *Proc. Natl. Acad. Sci. U.S.A.*, **2005,** 102, 11657-11662.

Liquid-Liquid Phase Separation and Liquid-to-Solid Maturation of Pmel17 RPT



The work described in this chapter has been published in the *Journal of the American Chemical Society*.

Reference: Dogra. P.; Joshi, A.; Majumdar, A.; Mukhopadhyay, S. Intermolecular Charge-Transfer Modulates Liquid–Liquid Phase Separation and Liquid-to-Solid Maturation of an Intrinsically Disordered pH-Responsive Domain. *J. Am. Chem. Soc*. **2019,** 141, 20380-20389.

5.1 Introduction

Living cells regulate the complex biochemical processes by the spatiotemporal organization of cellular components into functionally distinct membrane-bound compartments or organelles. However, there exists another important class of intracellular compartments that do not have the delimiting membrane yet provide a suitable environment for cellular biochemistry.[1-12] These membraneless organelles or biomolecular condensates are non-stoichiometric supramolecular assemblies of proteins and/or ribonucleic acids and are thought to be formed by the process of liquid-liquid phase separation (LLPS) or condensation.[1-12] Several recent studies have shown that these condensates have liquid-like physical properties as reflected by their spherical shape, ability to fuse, drip, wet, and rapid exchange of components with the surrounding environment.[9-11] An increasing body of work suggests that majority of these condensates/liquid droplets comprises multivalent or intrinsically disordered proteins/regions (IDPs/IDRs) which are believed to be the key modulators of LLPS due to their sequence-encoded physicochemical properties, conformational heterogeneity, and flexibility.[13-23] Often these proteins possess prion-like low-complexity regions (LCRs) that can participate in a multitude of specific but weak and transient intermolecular interactions resulting in liquid demixing.[24–31] The phase transition of LCRs under physiological conditions is highly dependent on the composition and patterning of amino acid residues that govern the thermodynamics of phase separation.[12,30,32] A flurry of experiments has revealed that the pathological maturation of liquid droplets of many RNA-binding proteins having low-complexity regions (LCRs) such as FUS, TDP-43, and hnRNPA1 into gel-like and solid-like assemblies are implicated in neurodegenerative diseases such as amyotrophic lateral sclerosis and fronto-temporal dementia.[33-43] However, the interplay of the key molecular drivers of LLPS and the liquid-to-solid transition is poorly understood.

In addition to more common LCR-containing RNA-binding proteins, there is a class of extracellular matrix proteins, including elastin, spider silks proteins, resilin, abductin, insect cuticle, etc. that phase separates into functional supramolecular assemblies.[16,17,31,44] In humans, a functionally controlled amyloid structure is generated by the protein Pmel17 that act as scaffolds for melanin deposition, sequestering toxic intermediates in melanin biogenesis.[45-50] Pmel17 is a 668-residue melanocyte-specific transmembrane protein that undergoes multiple sequential

proteolytic processing steps to generate a highly amyloidogenic luminal fragment (Mα) and a transmembrane fragment (Mβ). The Mα fragment (25−467) consists of three domains: the N-terminal domain (NTD); the polycystic kidney disease domain (PKD); and the oligopeptide repeat domain (RPT) (Figure 1a). The RPT (315−444) is a pH-responsive IDR comprising 10 imperfect repeats [PTXEXGTTPXQVX] that are enriched in threonine, proline, serine, alanine, valine, glycine, and glutamate residues (Figure 1b). It is known to form the amyloid core that promotes melanin biogenesis and forms amyloid fibrils under mildly acidic pH (pH 4−5.5) in melanosomes but remains soluble at cytosolic pH.[46,48,49] While several studies have shown the crucial role of RPT in providing a fibrillar matrix in vivo, the coupled regulatory and structural roles of RPT at cytosolic pH in vitro remain elusive. Here, we show that at cytosolic pH, the RPT undergoes LLPS in vitro and forms highly dynamic liquid-like droplets containing highly expanded, solvated, and rapidly fluctuating chains. We demonstrate that the phase separation of the RPT is driven by an intriguing interplay of a gamut of nonvalent interactions coupled with an unusual cascade of intermolecular charge-transfer through the backbone hydrogen-bonded network that controls the liquid-to-solid phase transition.

5.2 Experimental Section

5.2.1 Materials

Sodium phosphate monobasic dihydrate, sodium chloride, 1,6-Hexanediol, sodium acetate, potassium iodide, sodium citrate tribasic dihydrate, Thioflavin T, 8-anilino-1-naphthalenesulfonic acid (ANS), DTT (1,4-dithiothreitol), β-mercaptoethanol, and TWEEN 20 were purchased from Sigma Aldrich (St. Louis, MO). The fluorescent probes, namely, AlexaFluor 488 C5-maleimide, AlexaFluor 594 C5-maleimide, 5-((((2-Iodoacetyl)amino)ethyl)amino)Naphthalene-1-Sulfonic Acid) (IAEDANS), and fluorescein-5-maleimide were obtained from Molecular Probes, Invitrogen. The free fluorescein dye was procured from Fluka Analytical. The protein concentrators, filters, A11 (anti-amyloid oligomer antibody), OC (anti-amyloid fibril antibody), HRP-conjugated goat-anti-rabbit antibody, and nitrocellulose membrane, 0.45 μm were obtained from Merck Millipore. Bovine serum albumin and enhanced chemiluminescence substrate were purchased from HiMedia and Thermofisher Scientific, respectively. Guanidinium chloride and urea were procured from Amresco. Ni-NTA Agarose resin was purchased from Qiagen. Resins Q-sepharose, superdex G-75, and PD-10 columns were obtained from GE Healthcare Life Sciences

(USA). All the buffer solutions prepared were filtered before use. A Metrohm 827 lab pH meter was used to adjust the final pH (± 0.01) at 25 °C.

5.2.2 Recombinant protein expression, purification, and labeling

Human RPT (315-444) was recombinantly expressed in E.coli BL21(DE3) RIPL and purified using previously reported methods.[47] The C-terminal single point cysteine mutant (S442C) was generated using the Stratagene Quickchange kit (Agilent Technologies), and the mutation was confirmed by DNA sequencing. The primer sequence is as follows:

5'-GGTCCAGATGCCTGCTCAATCCATCAC-3' (forward)

5'-GTGATGGATTGAGCAGGCATCTGGACC-3' (reverse)

For labeling RPT, purified RPT mutant eluted in native buffer (50 mM NaH2PO4 buffer, 1 mM DTT, pH 7.0) was mixed in a molar ratio of 3:1 (AlexaFluor 488 C5-maleimide:RPT[S422C]), 2:1 (AlexaFluor 594 C5-maleimide:RPT[S422C]), 4:1 (Fluorescein-5-maleimide (F-5-M):RPT[S422C]), and 20:1 (IAEDANS:RPT[S422C]). The reaction mixtures were allowed to stir in the dark for 2-3 h at room temperature. After the completion of the reactions, the labeled proteins were passed through a PD-10 column to remove any unreacted dye. The concentration of the labeled proteins were estimated using $\varepsilon_{493\ nm}$ = 72000 $M^{-1}cm^{-1}$, for AlexaFluor 488 C5-maleimide, $\varepsilon_{588\ nm}$ = 96000 $M^{-1}cm^{-1}$, for AlexaFluor 594 C5-maleimide, $\varepsilon_{490\ nm}$ = 68000 $M^{-1}cm^{-1}$, for Fluorescein-5-maleimide (F-5-M), and $\varepsilon_{340\ nm}$ = 6100 $M^{-1}cm^{-1}$, for IAEDANS.

5.2.3 Phase separation assays

Purified RPT was concentrated using 3 kDa MWCO AMICON filter (Millipore) at 4 °C in a 20 mM Tris buffer pH 8.0 to a final concentration ~ 800 μM. The LLPS was initiated by a rapid dilution of concentrated RPT into the droplet buffer (50 mM NaH_2PO4, 50 mM NaCl, pH 7.0). The final concentration of RPT used for droplet formation was 100 μM. All the samples were prepared in autoclaved 1.5 mL micro-centrifuge tubes with a typical reaction volume of 250-300 μL. The formation of the RPT droplets was monitored under different conditions by varying pH (4.5, 5.0, 5.5 [50 mM sodium acetate], and pH7.0 [50 mM sodium phosphate]), salt (NaCl) concentration (25 mM, 50 mM, 150 mM, 300 mM, 600 mM, 800 mM, and 1000 mM) and temperature (4 ˚C, 15 ˚C, 25 ˚C, 37 ˚C, 42 ˚C, and 50 ˚C). Prior to all the spectroscopic studies,

we performed turbidity assays and confocal microscopy to confirm phase separation and liquid droplet formation. The turbidity of protein samples was estimated by recording the optical density at 350 nm on a Multiskan Go (Thermoscientific) plate reader using 96-well optical bottom NUNC plates (Thermoscientific). The sample volume used for all the measurements was 150 μL. We would like to note that the turbidity values reported are raw and uncorrected. The mean and standard error was obtained from at least three independent sets of experiments.

5.2.4 Phase-contrast and confocal fluorescence microscopy

For imaging of RPT droplets by phase-contrast microscopy, a 60x oil-immersion objective (Numerical aperture: 1.35) was used on an Olympus FLUOVIEW confocal laser scanning microscope (Model No. FV10i). For visualization of droplets at different time-points of the reaction, 4-6 μL of an aliquot was withdrawn from the reaction mixture and placed onto a glass slide (Fisher Scientific 3" x 1" x 1 mm), which was then immediately covered with a circular coverslip. Both the edges of the coverslip were sealed to avoid evaporation of the solution during image acquisition. We would like to note that we observed a larger number of droplets at the outer edge of the coverslip. For visualizing the intrinsic blue fluorescence of RPT droplets, confocal microscopy was performed using the DAPI channel (excitation 359 nm, emission 461 nm). Additionally, droplets formed using IAEDANS labeled RPT were also imaged using the DAPI channel. RPT-Alexa488 and RPT-fluorescein droplets were imaged using an excitation source at 499 nm with the corresponding emission at 520 nm. RPT-Alexa594 droplet images were acquired using excitation at 590 nm and emission at 618 nm. All the images were analyzed and processed using ImageJ software. (NIH, Bethesda, MD, USA).

5.2.5 Fluorescence recovery after photobleaching (FRAP) measurements

FRAP measurements were performed on a Confocal Laser Scanning ZEISS 710 Microscope equipped with a Plan Apochromat 63x oil-immersion objective (Numerical aperture 1.4) along with a high-resolution monochrome cooled AxioCamMRm Rev. 3 FireWire(D) camera. For carrying out FRAP experiments, 10% of AlexaFluor 594 C5-maleimide labeled protein was used. The FRAP profiles and fluorescence images were acquired using the ZEN Pro 2011(ZEISS) software provided with the instrument. For FRAP measurements, a region of interest (ROI) was chosen inside a single liquid droplet, and the region was bleached to approximately 50% of the initial intensity using a 561 nm laser (20 mW laser power). Fluorescence recovery curves were

constructed from the normalized fluorescence intensity values in the ROI (corrected for the background noise) and were plotted using the Origin 2018 software. The fluorescence images were analyzed and processed using ImageJ software.

5.2.6 Far-UV circular dichroism (CD) spectroscopy

Far-UV CD measurements were performed on a Chirascan spectrophotometer (Applied Photophysics, U.K.) using a 1 mm path length quartz cell. The typical protein concentration used for the measurement was 20 μM. The spectra were recorded with a scan range of 200-260 nm with a step size of 1 nm and 5 accumulations per scan. All measurements were recorded in triplicate. The spectra were averaged over 5 scans and were baseline corrected against the buffer signal using the ProData software provided with the instrument. All the spectra were plotted using Origin 2018 software.

5.2.7 Dot-blot assay

Aliquots were withdrawn from the reaction mixture after 168 h, and 2 μL of each aliquot was spotted on the nitrocellulose membrane. The blots were allowed to dry, and the membrane was then blocked by 3 % BSA in PBST (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.05% Tween 20) for 1 h at room temperature. The membranes were probed with primary antibody overnight at 4 °C (A11, 1:500; OC, 1:1000). The blots were washed six times with PBST and incubated with appropriate HRP-conjugated goat anti-rabbit secondary antibody for 1 h at room temperature. Then the blots were washed thrice with PBS buffer and visualized using an ECL kit.

5.2.8 Steady-state fluorescence measurements

All the steady-state fluorescence measurements were carried out on a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon, NJ) using a 1-mm-pathlength quartz cuvette. The final concentration of RPT used for all of the fluorescence studies was 100 μM. For recording tryptophan fluorescence intensity, the samples were excited at 280 nm, and the spectra were collected in the emission range between 310 and 550 nm. For collecting tryptophan anisotropy, the following parameters were used: λ_{ex} = 280 nm, λ_{em} = 340 nm. The steady-state fluorescence anisotropy (r_{ss}) is given by the following relation:

$$r_{ss} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}} \qquad \text{(Eq. 1)}$$

where $I_{\parallel}$ and $I_{\perp}$ are the fluorescence intensities collected when emission polarizer is oriented parallel and perpendicular to polarized excitation, respectively, and the measured intensities were corrected using the corresponding G-factor. For monitoring fluorescein fluorescence, the following parameters were adjusted: λ_{ex} = 485 nm, Emission range = 491-540 nm, and the r_{ss} was recorded at the emission maximum. For recording charge-transfer fluorescence, the samples were excited at 340 nm, and the emission spectra were collected in the range of 360-500 nm. The excitation spectra were collected in the range of 260-390 nm by keeping the emission wavelength fixed at 400 nm. The AEDANS fluorescence anisotropy was recorded by exciting the samples at 375 nm and collecting the emission at λ_{max}. For the ThT fluorescence assay, the samples were excited at 450 nm, and emission was collected in the range of 460-600 nm. The final ThT concentration used for all of the experiments was 20 μM. The ANS fluorescence measurements were performed in a similar manner as that of ThT fluorescence experiments. For recording ANS fluorescence, the excitation wavelength (λ_{ex}) was set at 375 nm, and the emission was collected in the range of 440-580 nm.

5.2.9 Fluorescence quenching experiments

Fluorescence quenching experiments of fluorescein-labeled RPT were performed as described previously.[21] Briefly, droplet reactions were set up with 200 nM fluorescein-labeled RPT with an excess of unlabeled wild-type RPT (99.8 μM). Fluorescence lifetimes were measured in the absence and the presence of increasing amounts of quencher (KI). The time-resolved fluorescence intensity decay profiles were recorded at the magic angle (54.7º) using a time-correlated single-photon counting (TCSPC) setup (Fluorocube, Horiba Jobin Yvon, NJ). The samples were excited using a 485-nm laser diode. The instrument response function (IRF) was ~ 270 ps. The ratio of the average fluorescence lifetimes in the absence and in the presence of quencher (τ_0/τ) was plotted against the quencher concentration ([Q]). The Stern-Volmer quenching constants (K_{sv}) and the bimolecular quenching rate constants (k_q) were obtained using the following relationships (equations 2 & 3). See Table 5.1 for recovered parameters.

$$[\tau_0/\tau] = 1 + K_{sv}[Q] \qquad \text{(Eq. 2)}$$

$$k_q = K_{sv}/\tau_0 \qquad \text{(Eq. 3)}$$

5.2.10 Time-resolved fluorescence anisotropy measurements

The time-resolved fluorescence anisotropy decay measurements were performed using a TCSPC setup coupled with a diode laser as described above. Data acquisition and analyses were performed using the methods described in our previous articles.[21,51] Briefly, the polarized emission intensities were collected at 0° and 90° with respect to the excitation polarization and the anisotropy decays were analyzed by globally fitting $I_{\parallel}(t)$ and $I_{\perp}(t)$ as follows:

$$I_{\parallel}(t) = I(t)[1+2r(t)]/3 \qquad \text{(Eq. 4)}$$

$$I_{\perp}(t) = I(t)[1-r(t)]/3 \qquad \text{(Eq. 5)}$$

$I(t)$ is the fluorescence intensity collected at the magic angle (54.7°) at time t and $I_{\perp}(t)$ was always corrected using a *G*-factor. The anisotropy decays were then analyzed using a biexponential decay model describing fast and slow rotational correlation times as follows[51]:

$$r(t) = r_0\left[\beta_{fast}\, exp\left(-\frac{t}{\phi_{fast}}\right) + \beta_{slow}\, exp\left(-\frac{t}{\phi_{slow}}\right)\right] \qquad \text{(Eq. 6)}$$

where r_0: time-zero intrinsic fluorescence anisotropy; ϕ_{fast} & ϕ_{slow}: fast and slow rotational correlation times, respectively; β_{fast} & β_{slow}: amplitudes associated with fast and slow rotational time, respectively. See Table 5.2 for recovered parameters.

5.2.11 Atomic Force Microscopy (AFM)

For preparing samples, 10 µL of aliquots were taken from the reaction mixture and were deposited onto freshly cleaved, buffer-washed muscovite mica (Grade V-4 mica from SPI, PA). The samples were then allowed to incubate for 5 min at room temperature. Following the incubation, the samples were washed with 100 µL of the appropriate filtered [0.22 µm membrane filter (Merck Millipore)] buffer and were kept under a gentle stream of nitrogen for 10 min prior to scanning. AFM images of RPT aggregates were acquired on an Innova atomic force microscope (Bruker) operating in the tapping mode. For imaging, a silicon nitride cantilever probe with a radius of ~8 nm was used. The images were collected using NanoDrive (v8.03) software at a resolution of 1024

× 1024 pixels and 256 × 256 pixels. The collected AFM images were further processed and analyzed using WSxM version 4 software.[57]

5.2.12 Raman Spectroscopy

The Raman spectra were recorded on an inVia laser Raman microscope (Renishaw, UK). For preparing samples, RPT aggregates formed under two different pH conditions were lyophilized. The samples were deposited onto a glass slide and were focused using a 50X objective lens (Nikon, Japan). An NIR laser (785 nm) was used to excite the samples with an exposure time of 10 s and 10 % laser power. The Raman scattered light was collected, and the Rayleigh scattered light was removed by using an edge filter of 785 nm. The Raman scattered light was then dispersed using a 1200 lines/mm diffraction grating and was detected by an air-cooled CCD detector. The data was acquired using Wire 3.1 software provided with the instrument. All the spectra were averaged over 25 scans. The Raman spectra were baseline corrected using the cubic spline interpolation method for eliminating the tilt and were then smoothed using Wire 3.1 provided with the Raman microscope. The baseline corrected and smoothed Raman spectra were finally plotted using Origin. The Raman peaks were interpreted as described previously.[58,59]

5.3 Results

5.3.1 RPT Undergoes Liquid−Liquid Phase Separation at Cytosolic pH.

As a prelude, we first performed bioinformatic analyses using various prediction tools that revealed that the RPT is an IDP possessing an LCR (325−339) (Figure 5.1a,b,c). As several IDPs/IDRs containing LCRs have been shown to phase separate, we postulated that the RPT can undergo LLPS. To test this hypothesis, we purified the recombinantly expressed RPT and tested its ability to phase-separate by performing a turbidity assay and phase-contrast microscopy (Figure 5.1d,e). For our studies, we incubated 100 μM RPT under cytosolic conditions (pH 7; 37 °C). Initially, the solution was clear indicating the highly soluble dispersed state of the RPT, and upon longer incubation, mesoscopic spherical droplets were formed (Figure 5.2a). Further, the smaller droplets fused rapidly (<10 s) to form larger droplets (Figure 5.2b,c). These droplets also exhibited other liquid-like properties such as dripping and surface-wetting[11] (Figure 5.2b). We next estimated the protein concentration inside the liquid droplets and found that the concentration in

the droplet (dense) phase was in the millimolar range which is ~45- fold excess of the dispersed (light) phase (Figure 5.2d). We would like to note that this is an approximate estimation and may

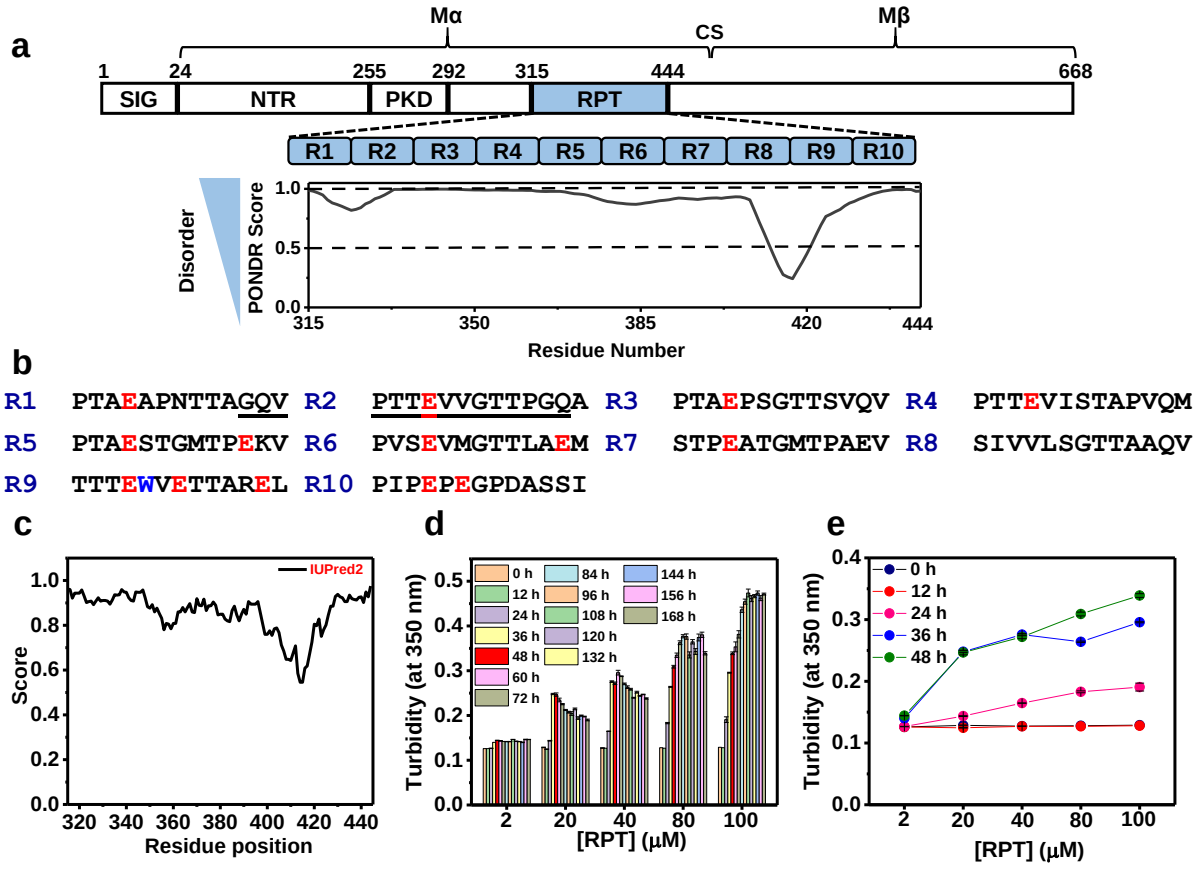


Figure 5.1 Liquid−liquid phase separation of the RPT at cytosolic pH. (a) Schematic representation of human Pmel17. Full-length Pmel17 with Mα and Mβ fragments separated by a cleavage site (CS). Mα fragment is composed of three highly amyloidogenic domains: the NTR, N-terminal region; the PKD, polycystic kidney disease protein homology domain; and the RPT, repeat domain. RPT comprises 10 13-residue repeats. Predictor of Natural Disordered Regions (PONDR)[54] showing disorder propensity of the RPT. (b) Primary sequence of the RPT showing imperfect repeats rich in hydrophobic, polar, and acidic (in red) residues. The single native tryptophan (W423) is highlighted in blue. The underlined sequence shows a low complexity region within the RPT predicted using the Simple Modular Architecture Research Tool (SMART).[55] (c) Intrinsic disorder in the primary sequence of RPT predicted using IUPred2[56] (plot generated using https://iupred2a.elte.hu/) (d) The changes in the turbidity during the phase transition of RPT as a function of concentration (50 mM NaH_2PO_4, 50 mM NaCl, pH 7.0, 37 °C). The turbidity data are shown as mean ± SEM (n ≥ 3). The increase in the turbidity values indicating the initiation of the LLPS process after 24 h upon incubation of RPT (≥ 20 uM) at 37 °C. (e) Select data from (d) are shown for clarity. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

not be absolutely correct due to high protein concentrations within the droplets. Together, these findings revealed the liquid-like properties of the condensates and suggest a condensed protein-rich environment within the liquid droplets. After establishing the liquid-like properties of the RPT droplets, we next set out to characterize the inner environment and dynamic nature of the droplets.

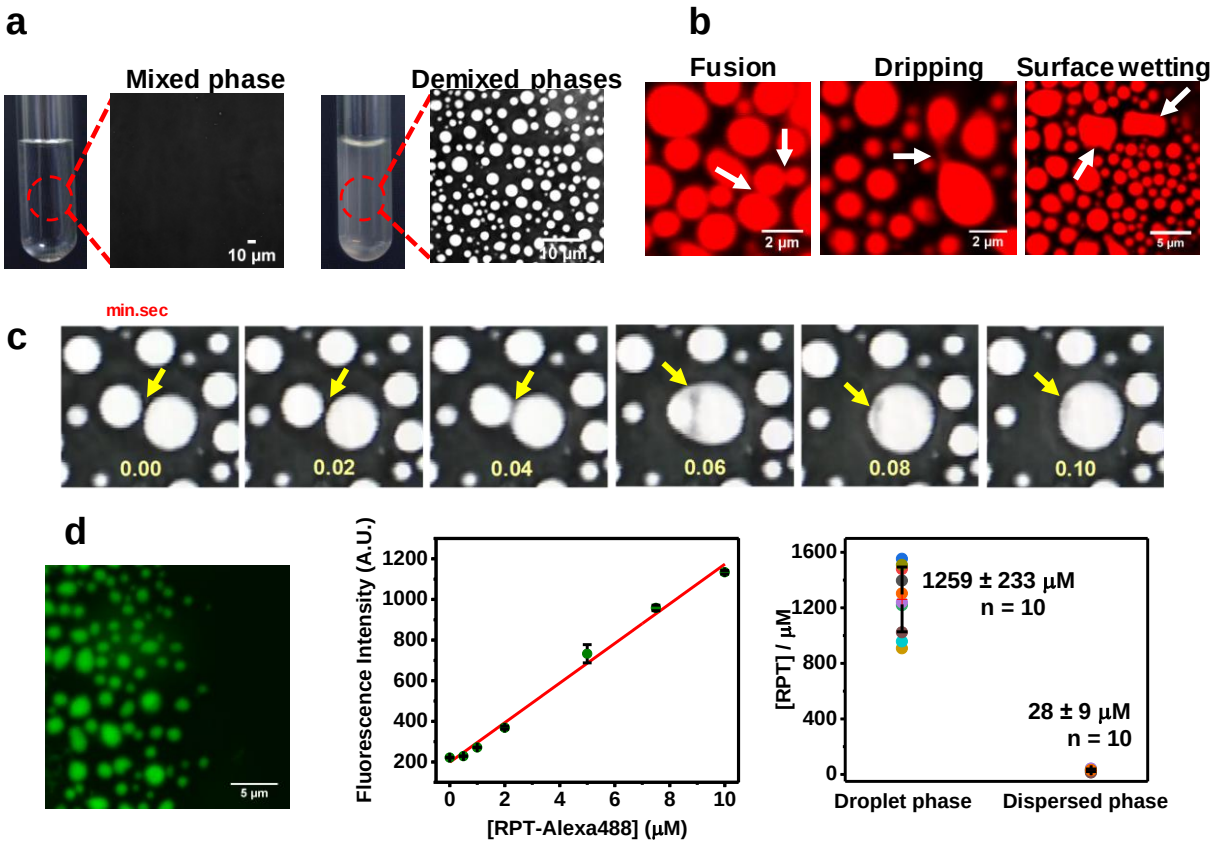


Figure 5.2 (a) Phase transition of the RPT (100 µM, pH 7.0, 50 mM NaCl, 37 °C, 48 h) from mixed dispersed phase to demixed droplet phase. (b) Liquid-like properties of the RPT droplets (labeled with Alexa594) showing fusion, dripping, and surface-wetting (indicated using arrows). (c) Time-lapse imaging showing the fusion of two liquid droplets (indicated using arrows). (d) A confocal image of fluorescently labeled RPT-Alexa488 liquid droplets after 48 h. We estimated the concentration of RPT inside the droplet (dense) phase and dispersed (light) phase by using RPT-Alexa488 (labeled:unlabeled = 2:98). Only 2% of labeled protein was used to minimize the effects from high fluorophore concentrations. The fluorescence intensity was calibrated against different concentrations of RPT-Alexa488 at 0 h (mixed phase). The mean (n = 10 droplets) estimated concentration was 1259 ± 233 µM in the droplets and 28 ± 9 µM in the dispersed phase. This is an approximate concentration estimation. Data are represented as mean ± SEM of three independent measurements. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

5.3.2 Inner Environment of Liquid Droplets.

Internal Diffusion. For these studies, we took advantage of the fact that the RPT is devoid of any Cys and created a single-Cys variant at 442 using site-directed mutagenesis. We then covalently labeled Cys-442 with AlexaFluor594 and verified that labeled RPT retained the ability to undergo LLPS (Figure 5.3a). To characterize the internal diffusion of the RPT molecules inside these droplets, we carried out fluorescence recovery after photobleaching (FRAP) experiments (Figure 5.3b,c). Rapid recovery of the fluorescence signal indicated that the interior of the droplet is highly mobile, which is expected for a liquid phase. Next, we intended to investigate the chain solvation in the liquid droplets.

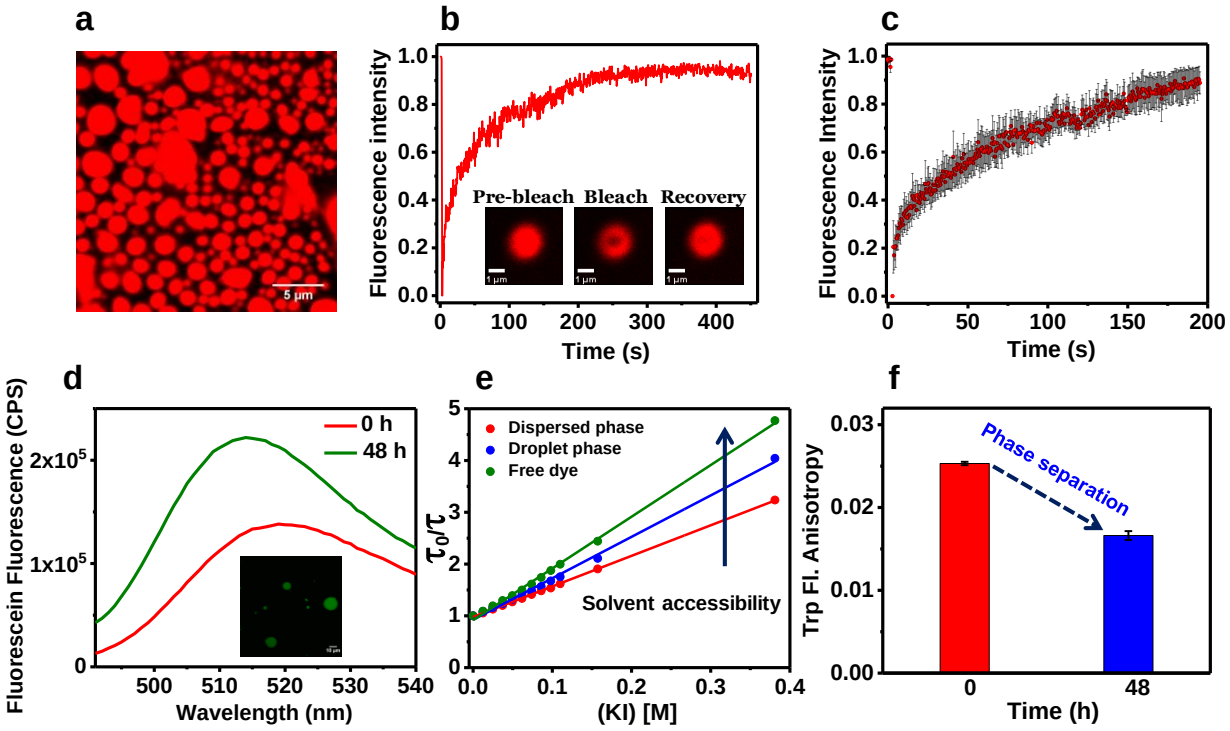


Figure 5.3 Internal diffusion and solvation in RPT condensates. (a) Confocal fluorescence image of Alexa594 labeled RPT droplets. (b) FRAP kinetics of RPT (single-Cys variant is labeled with Alexa594 maleimide) showing rapid and complete recovery. The inset shows fluorescence images of droplets during FRAP measurements. (c) The FRAP recovery timescales within droplets. The kinetic profiles shown were obtained from three different droplets. (d) Fluorescence emission spectra of F5M-labeled RPT at 0 h (red) in the mixed phase and 48 h (olive) in the droplet phase. The inset shows a confocal image of F5M-labeled RPT. (e) The Stern−Volmer plots of free fluorescein (green), the dispersed phase of the RPT (red) and droplet phase (48 h) (blue) from fluorescein-labeled (0.2%) RPT by plotting the ratio of mean fluorescence lifetimes in the absence and presence of KI (quencher). (f) The changes in the steady-state fluorescence anisotropy of Trp. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

Table 5.1

Recovered Stern-Volmer quenching constants (K_{SV}) and bimolecular quenching rate constants (k_q):

Sample	τ_0 (ns)	K_{sv} (M^{-1})	k_q ($M^{-1}s^{-1}$)
Free Fluorescein	3.93 ± 0.01	10.02 ± 0.03	(2.55 ± 0.01) x 10^9
RPT Droplet Phase	3.94 ± 0.01	7.47 ± 0.45	(1.90 ± 0.12) x 10^9
RPT Disperse Phase	3.93 ± 0.01	5.98 ± 0.02	(1.52 ± 0.01) x 10^9

Solvation. To directly probe the chain solvation in the condensed phase, we covalently labeled Cys-442 with sulfhydryl-reactive fluorescein-5-maleimide (F5M) and formed RPT droplets using 0.2% of labeled protein (Figure 5.3d). For measuring solvent accessibility, we estimated the biomolecular quenching rate constants by performing time-resolved fluorescence experiments using a water-soluble quencher. The droplet state exhibited much higher water accessibility compared to the dispersed state (Figure 5.3e, Table 5.1). These results indicated recruitment of water and significant chain solvation due to conformational expansion of the RPT within the liquid droplets as described previously.[21]

Chain Fluctuations. We next directed our efforts to characterize the intrinsic chain dynamics of the RPT in two distinct phases. We carried out steady-state fluorescence (polarization) anisotropy measurements and recorded the anisotropy of the single tryptophan (W423) that resides in the amyloidogenic C-terminal region.[46,48,49] Interestingly, we observed a sharp drop in the fluorescence anisotropy value from the dispersed phase to the droplet phase indicating enhanced rotational mobility in the droplet phase (Figure 5.3f). We next used both F5M and IAEDANS [5-((((2- iodoacetyl)amino)-ethyl)amino)naphthalene-1-sulfonic acid] labeled RPT at Cys-442 to confirm that the decrease in the anisotropy is not specific to either the site or the fluorophore but is a general characteristic due to enhanced chain mobility upon LLPS (Figure 5.4a,c). We also ruled out the possibility of any turbidity-induced artifact by performing dilution experiments and observed no change in the anisotropy upon dilution (Figure 5.4b). Next, to discern the origin of the chain mobility, we performed picosecond time-resolved fluorescence anisotropy measurements using the RPT labeled with IAEDANS possessing longer fluorescence lifetime (>10 ns) that allows us to follow the dynamical events on a wide range of time scales. Typically, for a biomolecule covalently labeled with a fluorophore, the depolarization kinetics is described by a

biexponential decay function comprising a (fast) local- and a (slow) global rotational correlation time that is related to size in the Stokes−Einstein relationship.[51–53] Our previous studies demonstrated that for an archetypal expanded IDP, the (characteristic) slower nanosecond correlation time corresponds to collective backbone torsional fluctuations in the Φ−Ψ dihedral angle space.[53] Picosecond depolarization kinetics recapitulated the inherent backbone dihedral dynamics upon phase separation (Figure 5.4d,e,f, and Table 5.2). These rapid nanosecond torsional fluctuations characterize the expanded polypeptide chains within the droplet environment that can

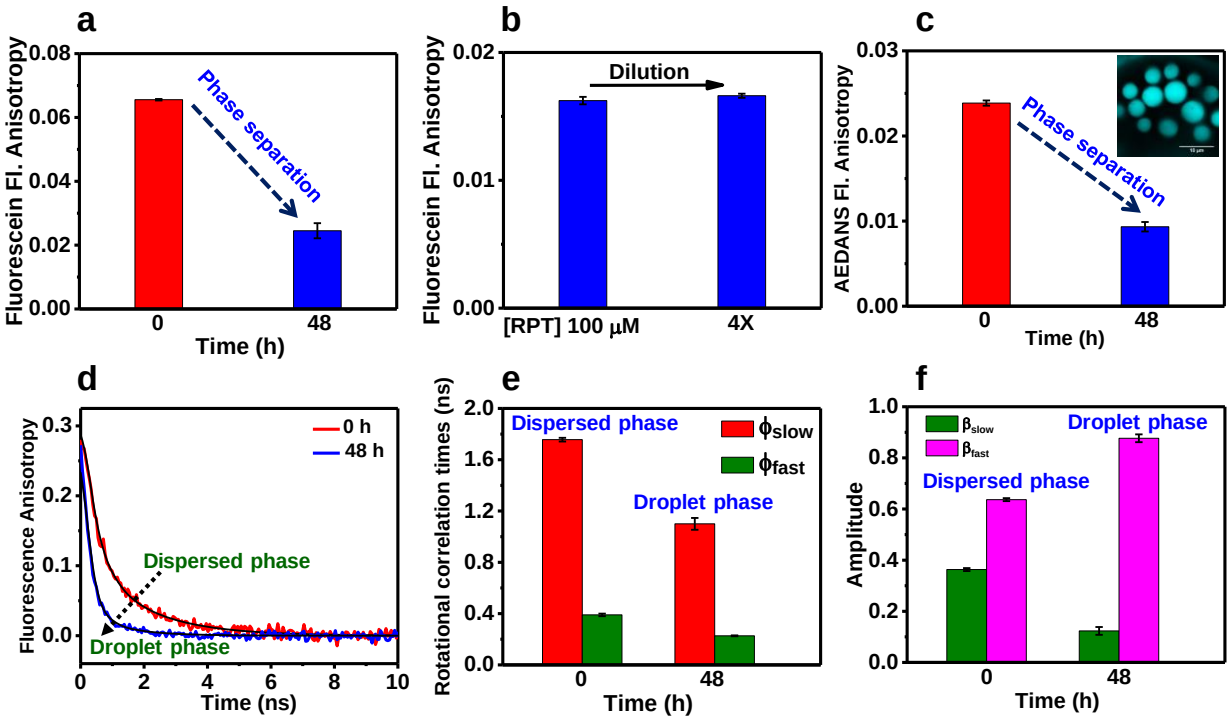


Figure 5.4 Chain fluctuations in RPT condensates. (a) Steady-state fluorescence anisotropy of F5M-labeled RPT (0.2 % labeled) at 0 h (red) and 48 h (blue). (b) The steady-state fluorescence anisotropy values of F5M-labeled RPT droplets formed after 120 h after dilution eliminated the turbidity-induced artifact in the anisotropy measurements. (c) The changes in the steady-state fluorescence anisotropy of IAEDANS-labeled RPT. The inset shows a confocal image of IAEDANS-labeled RPT. (d) Picosecond time-resolved anisotropy decays of IAEDANS-labeled RPT in dispersed (red) and droplet (blue) phases. The black lines are fits using the biexponential decay kinetics. (e) Changes in the slow and fast rotational correlation times during phase separation. (f) Changes in the amplitude of rotational correlation times during LLPS. All data are shown as mean ± SEM (n = 3). Reprinted/adapted with permission from Dogra et al. (Ref. 66).

Table 5.2

The typical parameters recovered from fluorescence anisotropy decay analyses:

Sample	ϕ_{fast} (ns)	β_{fast}	ϕ_{slow} (ns)	β_{slow}
RPT Droplet Phase	0.23 ± 0.01	0.88 ± 0.02	1.10 ± 0.05	0.12 ± 0.02
RPT Disperse Phase	0.39 ± 0.01	0.64 ± 0.01	1.76 ± 0.01	0.36 ± 0.01

provide a better solvent quality than bulk water as described previously.[21] We would like to note that the information we obtained from all of our fluorescence studies is the ensemble average of light and dense phase properties. Taken together, this set of studies revealed that in the liquid droplet state, the expanded and solvated polypeptide chains exhibit rapid internal diffusion and large-scale intrinsic conformational fluctuations that have earlier been proposed to maintain the liquid-like interior.[18−26,60] We next asked: What are the critical noncovalent interactions responsible for RPT phase separation?

5.3.3 Molecular Drivers of RPT Phase Separation.

Electrostatic Interactions. To answer whether electrostatic interactions are essential for LLPS of the RPT, we performed the phase separation assays as a function of salt concentration. All of our experiments described so far were performed using 50 mM of NaCl. We observed that in the absence of salt (NaCl), the RPT showed only a slight increase in the turbidity, whereas, upon increase in the salt concentration up to 150 mM of NaCl, we observed a sharp increase in the turbidity (Figure 5.5a). Phase-contrast microscopy also revealed the presence of spherical droplets with an increase in the droplet number and the size in the concentration range 0−150 mM of NaCl (Figure 5.5b). No droplets were observed at salt concentrations ≥ 300 mM. To establish that the LLPS caused the turbidity increase, we incubated the RPT at high salt (1 M NaCl) in which phase separation does not occur.. By using different total protein concentrations, there was no increase in the turbidity values indicating that the increase in the turbidity is indeed due to LLPS (Figure 5.5c). Additionally, to verify that the increase in the turbidity values is due to LLPS and not due to protein aggregation, we next recorded fluorescence spectra and anisotropy of the single

tryptophan (W423) present in the RPT. As stated above, LLPS is associated with solvation and rapid chain fluctuations that are indicated by a red-shifted emission and a lower anisotropy, respectively.[21] On the contrary, aggregation results in the blue-shift in the emission due to less solvent accessibility and impedes the chain dynamics that is typically revealed by a sharp increase in the anisotropy.[51]

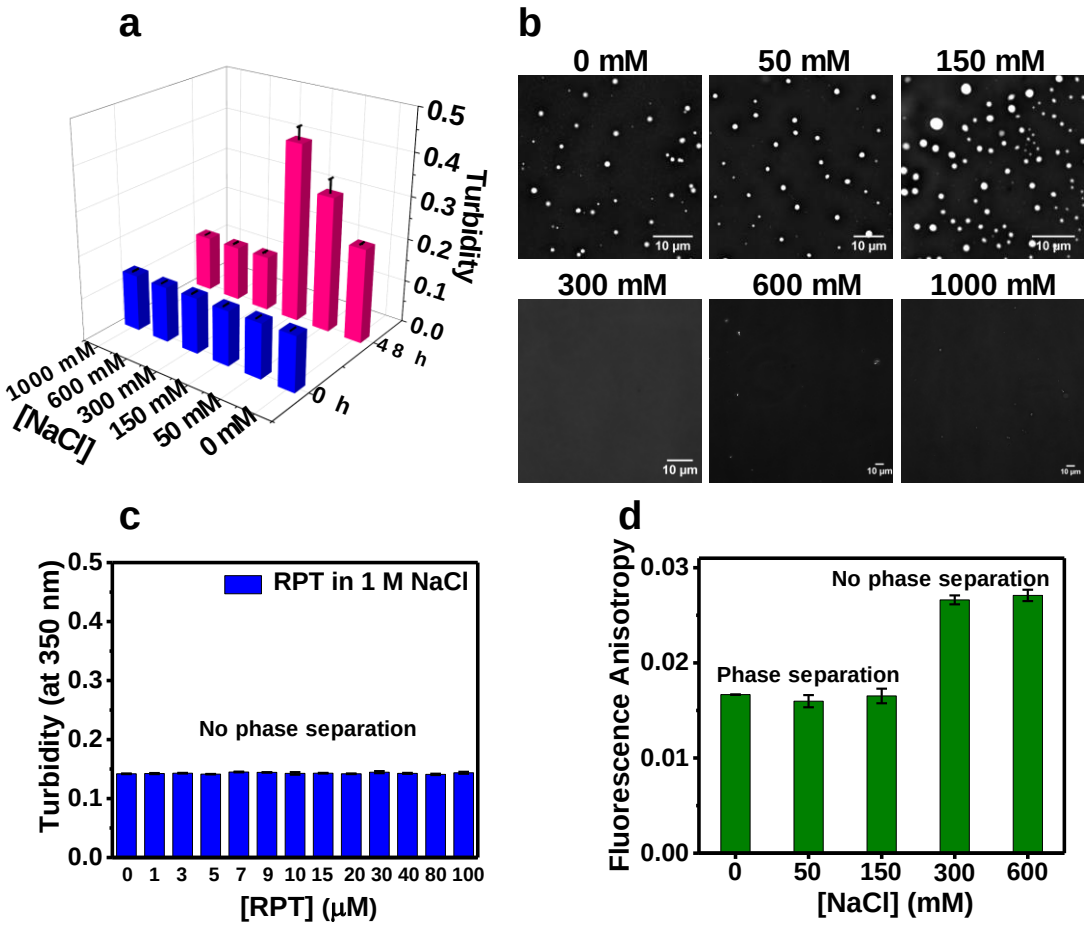


Figure 5.5 Role of electrostatic interactions in RPT phase separation. (a) A 3D column plot showing the effect of salt on turbidity. (b) Phase-contrast microscopy reveals salt-dependent phase separation of RPT. (c) No change in the turbidity values was observed as a function of total protein concentrations after incubating RPT at high salt (1 M NaCl), in which phase separation does not occur. (d) Steady-state fluorescence anisotropy values of (W423) RPT as a function of salt concentration. Data are represented as mean ± SEM of three independent measurements. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

Highly red-shifted emission maxima and lower anisotropy values ruled out the possibility of any aggregation of the RPT at various ionic strengths (Figure 5.5d, 5.6a). Together, these results

suggested a strong influence of ionic strength on the RPT phase behavior, and therefore, underscore the importance of optimal intermolecular interactions in the LLPS process.

Intermolecular Charge-Transfer. While investigating LLPS using Trp fluorescence, we serendipitously observed an additional long-wavelength emission shoulder for condensed liquid droplets (Figure 5.6a). This blue emission band (~400 nm) was absent in the monomeric dispersed phase and grew as a function of time during the phase transition. This unusual long-wavelength intrinsic blue fluorescence band is known to arise due to an extensive intermolecular charge-transfer through the hydrogen-bonded network of the polypeptide backbone mediated by trapped

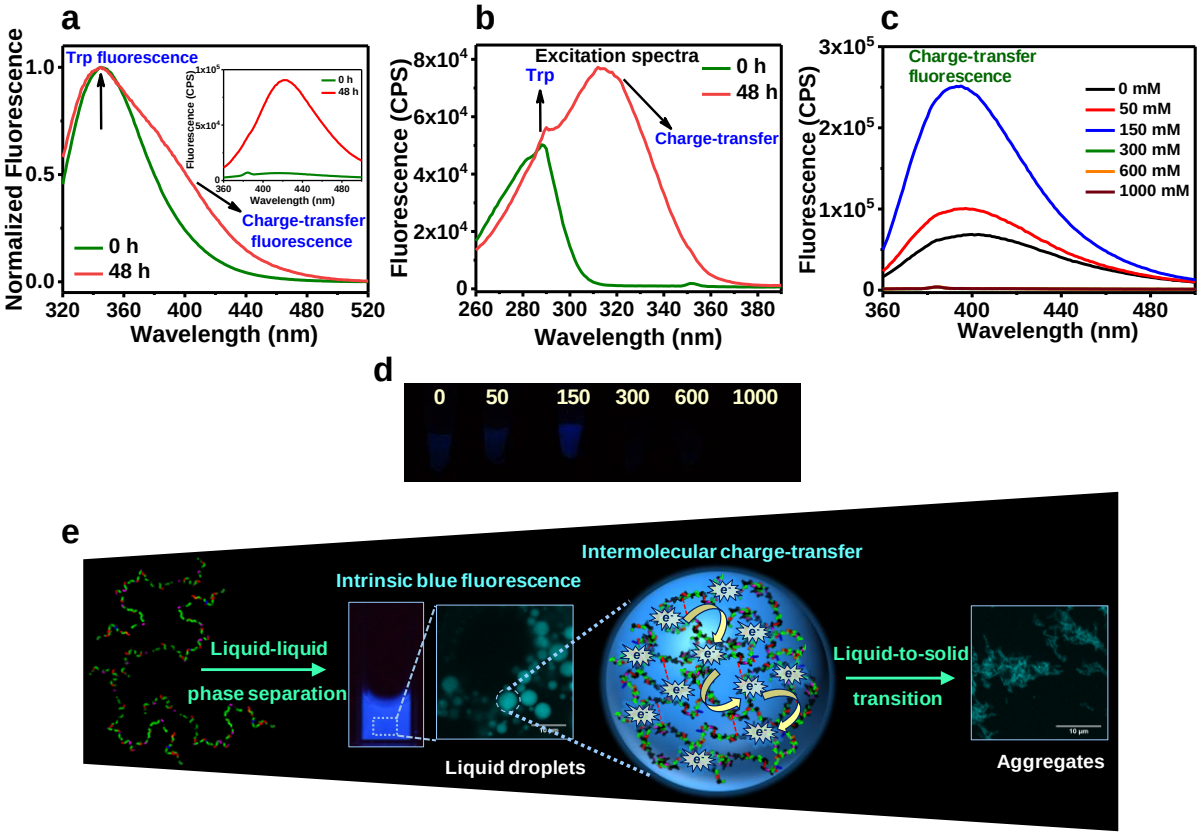


Figure 5.6 Intermolecular charge transfer modulates RPT phase transitions. (a) Fluorescence emission spectra of (W423) RPT at 0 h (olive) in the dispersed phase and 48 h (red) in the droplet phase. The inset shows charge-transfer fluorescence (λex=340 nm) spectra at 0 h (olive) and 48 h (red) (b) Fluorescence excitation spectra at 0 h and 48 h. (c) Charge-transfer blue fluorescence spectra as a function of salt concentration (d) Picture shows the intrinsic blue fluorescence upon UV illumination at 302 nm. (e) A schematic of intermolecular charge-transfer linked to the delocalization of electrons on an extended hydrogen-bonded backbone. A cuvette containing the RPT liquid droplets illuminated with UV light and confocal fluorescence micrographs of liquid-like and solid-likes states are also shown in the schematic. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

water molecules.[61–65] To confirm this, we recorded the excitation spectra that revealed the presence of a distinct ground-state species (λ_{max} ~ 320 nm) corresponding to the charge-transfer band in the droplet state (Figure 5.6b). To selectively monitor the (visible) blue fluorescence, we directly excited the charge-transfer absorption band at 340 nm that is well-separated from Trp absorption (280 nm) (inset in Figure 5.6a). Furthermore, the charge-transfer fluorescence intensity was found to be strongly dependent on the ionic strength that governs the phase transition. We observed an increase in the blue fluorescence when the salt concentration was increased from 0 to 150 mM, whereas, no blue fluorescence was observed at concentrations ≥ 300 mM where no phase separation was observed (Figure 5.6c). To directly visualize the intrinsic blue fluorescence from the condensed liquid-droplets as a function of ionic strength, we illuminated the samples under UV light (Figure 5.6d). A schematic of intermolecular charge-transfer linked to the delocalization of electrons on an extended hydrogen-bonded peptide supramolecular architecture is depicted in Figure 5.6e. These studies suggest that the protein-rich liquid phase comprises transient backbone interactions via which the intermolecular charge-transfer can occur from a charged side chain such as glutamate giving rise to intrinsic blue fluorescence that grew upon maturation (see later).

Hydrophobic Interactions. Since the amino acid sequence of the RPT has several hydrophobic residues, we next investigated the role of hydrophobics in the RPT phase transition. For our studies,

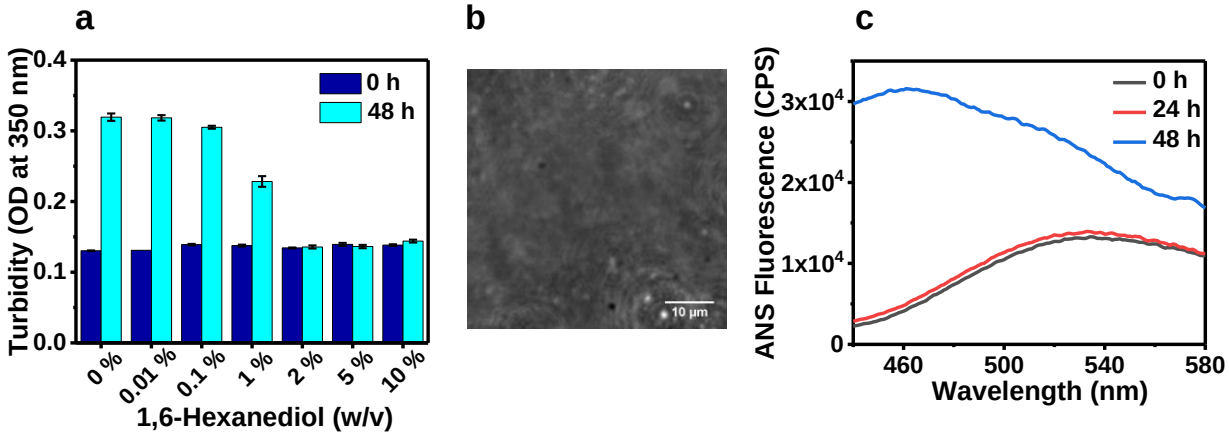


Figure 5.7 Role of hydrophobic interactions in RPT LLPS. (a) Turbidity measurements shows that the addition of ≥ 2 % (w/v) of 1,6-hexanediol completely inhibits droplet formation. Data are shown as mean ± SEM (n ≥ 3). (b) Phase-contrast micrograph shows no phase separation in the presence of 2 % 1,6-hexanediol). (c) ANS fluorescence 0 h (grey), 24 h (red), and 48 h (blue) during phase separation in the absence of 1,6-hexanediol. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

we used 1,6-hexanediol, an aliphatic alcohol, that is known to disrupt weak hydrophobic interactions.[39] We found that the RPT droplet formation was perturbed with a low fraction of 1,6-hexanediol and was eliminated at ≥2% (w/v) (Figure 5.7a,b). Additionally, we used a dye, 8-anilino-1-naphthalenesulfonic acid (ANS), the fluorescence of which is highly sensitive to the environment polarity and detects hydrophobic patches. Upon RPT phase separation, ANS exhibited an enhanced fluorescence with a moderate blue-shift indicating the presence of local hydrophobic patches within the droplets (Figure 5.7c). Together, these results indicate that LLPS of the RPT is driven by a multitude of dynamic multivalent interactions involving an interplay of the hydrophobic effect, electrostatic interactions, and a cascade of intermolecular charge-transfer through the hydrogen-bonded network. The relative contributions of these intermolecular interactions can potentially be crucial in determining the material property of the condensates. To understand how these interactions, contribute to the thermodynamics of mixing/demixing, we next studied the temperature-dependent phase behavior of the RPT.

5.3.4 Thermoresponsive Phase Transition.

The amino acid composition of IDPs/IDRs governs their temperature-dependent phase behavior.[16,17] For instance, the high content of aromatic and charged residues is believed to promote electrostatic/cation−π/π−π interactions that favor LLPS below a critical temperature and undergo upper critical solution temperature (UCST) transition, as seen for FUS.[30,32] On the contrary, elastin-like peptides having a large number of hydrophobic residues phase-separate above a critical temperature due to enhanced hydrophobic interactions that encode the lower critical solution temperature (LCST) behavior.[16,30,32] Several recent studies have shown that increased hydrophobicity and/or modulation in the number of polar/aromatic residues affect the demixing temperature.[31] To study the thermoresponsive phase behavior of the RPT, we performed the phase separation experiments at temperatures ranging from 4 °C to 50 °C. At 4 °C, no evidence for liquid demixing was found using both turbidity measurements and microscopy (Figure 5.8a,b). Phase separation of RPT was observed ≥ 15 °C, reached a maximum at 37 °C, and then sharply dropped at ≥42 °C. These observations are also corroborated by a sharp increase in the charge-transfer fluorescence in the temperature range in which we observed droplet formation (Figure 5.8c). We constructed a phase diagram of RPT by varying the temperature and the ionic strength (Figure 5.8d). A closed-loop profile on the phase diagram reveals an interesting interplay of

competing electrostatic and hydrophobic interactions. This behavior reflects the fact that at a lower temperature, electrostatic salt-bridges and charge-transfer interactions are more favored, and at a higher temperature, hydrophobic interactions are more preferred. These results show that the RPT exhibits an intriguing dual LCST and UCST phase behavior that is reminiscent of the phase behavior of a proline/glycine-rich IDP namely, resilin.[16]

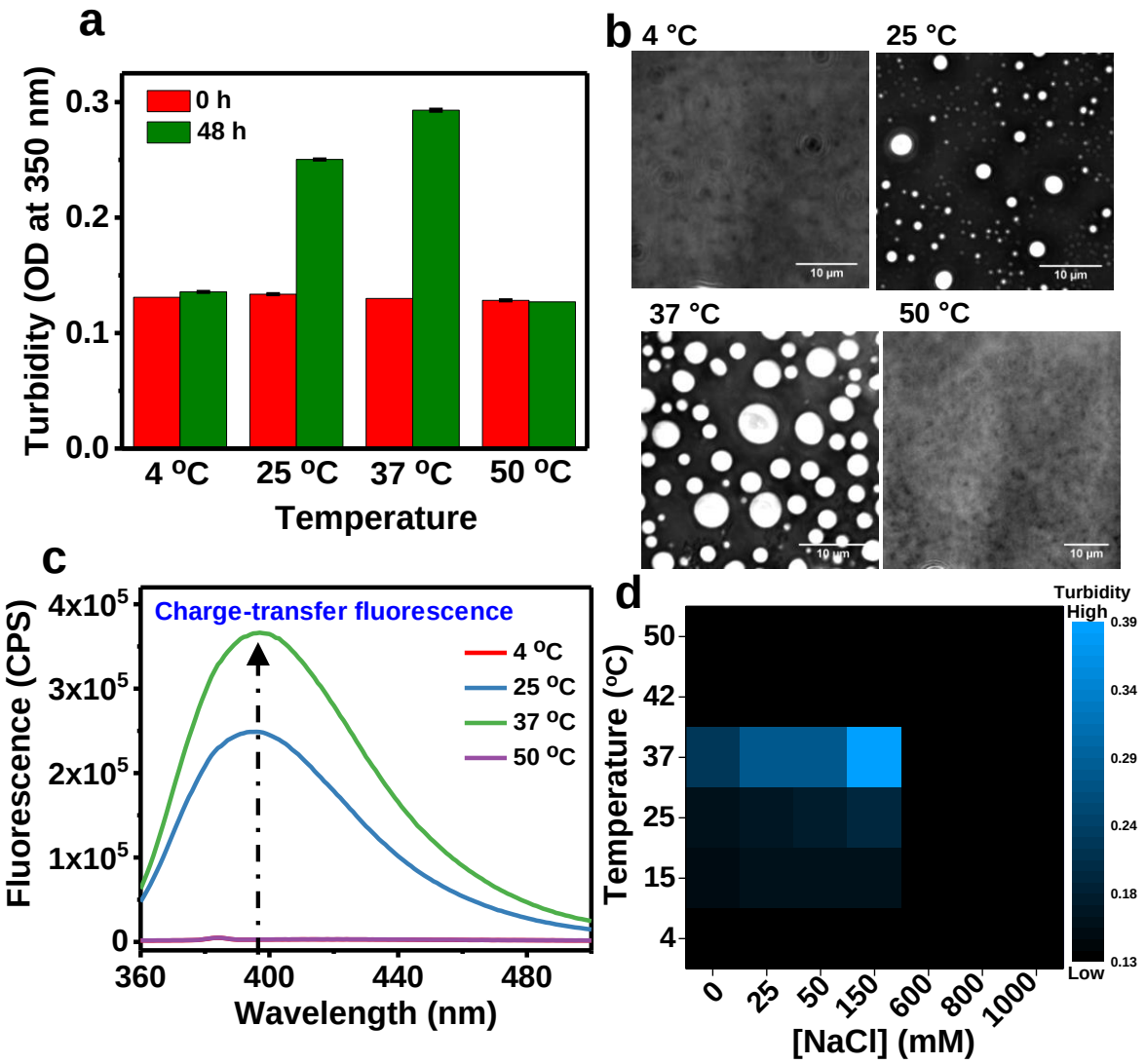


Figure 5.8 Thermoresponsive phase transition of the RPT. (a) Plot showing an effect of temperature on turbidity. The raw turbidity data are shown as mean ± SEM (n ≥ 3). (b) Phase contrast micrographs shows temperature-dependent droplet formation. (c) Charge-transfer fluorescence spectra as a function of temperature after 48 h of incubation. (d) Phase diagram of RPT in the presence of different salt concentrations and at different temperatures. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

5.3.5 pH Controls the Formation of Liquid-like and Amyloid Assemblies.

Competing Pathways. Since the pH is known to play a crucial role during different stages of melanosome development, we directed our efforts to study the effect of pH on the phase separation of the RPT. Stage I of the melanosome maturation is known to occur at an acidic pH (pH ~ 4.0) and comprises short amyloid fibrils that mature into fibrillar sheets in stage II. By the melanized

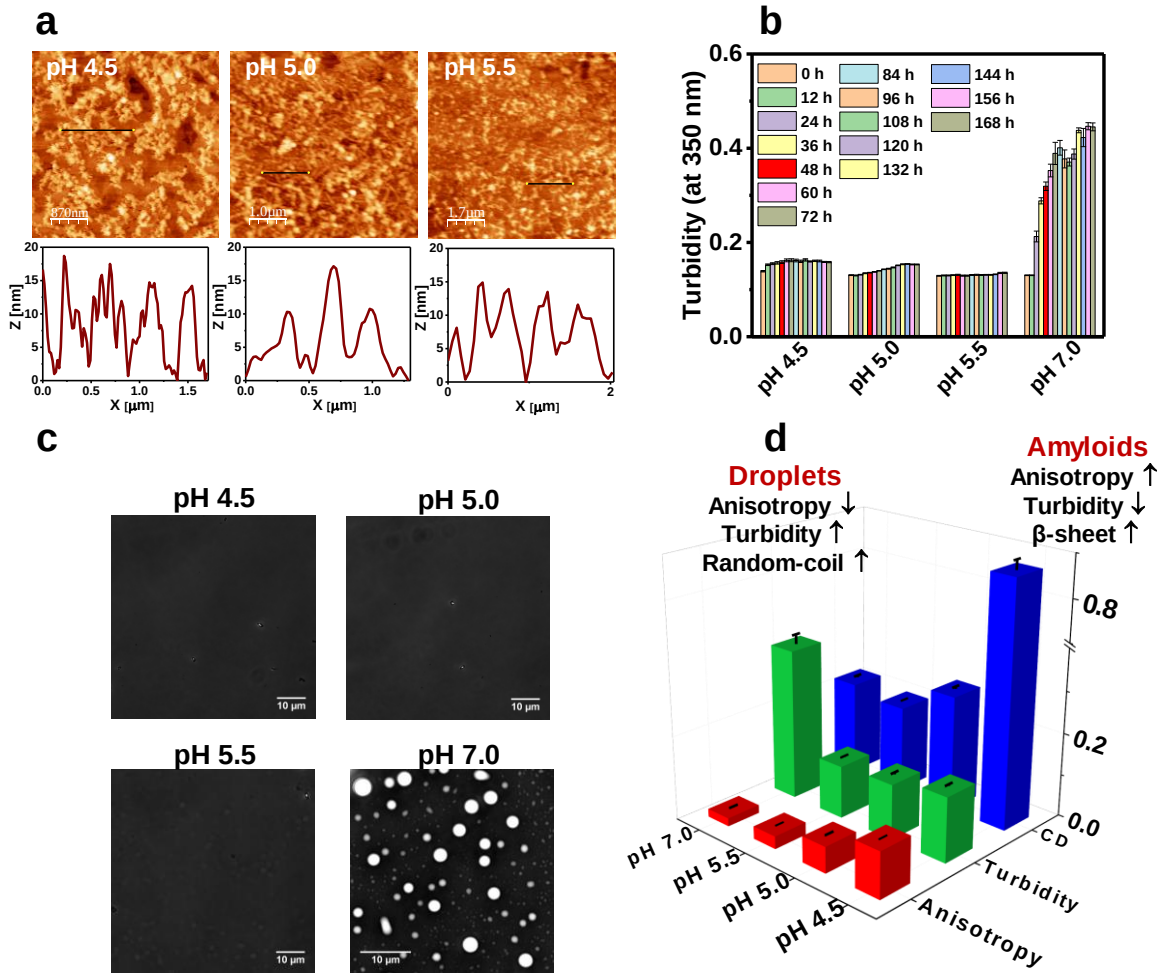


Figure 5.9 pH controls the formation of droplet-like assemblies and amyloids. (a) AFM images of the RPT aggregates at different pHs (with height profiles) collected after 168 h under quiescent conditions when incubated at 37°C. (b) pH-dependent changes in turbidity (at 350 nm) of solutions as a function of time containing 100 μM of RPT in 50 mM sodium acetate (pH 4.5, pH 5.0, and pH 5.5), sodium phosphate (pH7.0), 50 mM NaCl at 37 °C. The turbidity data are shown as mean ± SEM (n ≥ 3). (c) Phase-contrast micrographs show pH-dependent droplet formation. (d) A 3D column plot for the comparison of the CD ellipticity ratio ($\theta_{218}/\theta_{200}$) (blue), turbidity (green), and tryptophan fluorescence anisotropy (red) as a function of pH. Data are shown as mean ± SEM (n = 3). Reprinted/adapted with permission from Dogra et al. (Ref. 66).

stage III/IV the pH reaches close to neutral.[46,49] The RPT aggregates only under mildly acidic pH (Figure 5.9a) and remains soluble at neutral pH.[49] Our previous studies showed that at pH ≤ 4.5, the RPT rapidly aggregates via a nucleation-independent pathway, whereas, at pH > 4.5, it follows a typical nucleation-dependent mechanism.[47] The pH titration studies (pH 4 → pH 5) revealed a morphological change reminiscent of the transition observed from stage I → II melanosomes.[46,49] A number of glutamate residues present in the RPT have been shown to control the assembly/disassembly of amyloid fibrils through protonation/deprotonation.[50] To obtain the mechanistic insights into the pH-regulated condensation, we incubated 100 μM of RPT at 37 °C for 48 h using buffers ranging from (low) melanosomal pH to cytosolic pH. The RPT formed droplets only at the cytosolic pH but not at the melanosomal pH (Figure 5.9b,c). The ellipticity ratio ($\theta_{218}/\theta_{200}$) in our circular dichroism (CD) studies revealed a sharp drop in the random-coil content with a concomitant increase in the β-sheet content of amyloids formed at pH 4.5. However, the β-sheet content decreased from pH 4.5 to 7 as the RPT switched from the amyloid-state to liquid droplets (Figure 5.9d). Thioflavin T (ThT), a well-known amyloid marker, showed a much higher binding for amyloid species formed at mildly acidic pH and corroborated our CD results (Figure 5.10a). Additionally, these low-pH amyloid species can be converted into liquid droplets via a dispersed phase upon a pH jump to the neutral pH (Figure 5.10b-d).

Flexibility and Rigidity. Next, to distinguish the conformational dynamics of liquid droplets and amyloids, we compared the fluorescence anisotropy, a measure of the rotational hindrance that increases upon aggregation.[51,52] Upon increasing the pH, we observed a sharp decrease in the steady-state anisotropy as shown in Figure 5.9d. Additionally, we observed a progressive red-shift in the Trp emission spectra that suggested exposure to the aqueous environment (Figure 5.10e). We observed charge-transfer fluorescence only under neutral pH conditions that yielded liquid droplets (Figure 5.10f). The low-pH amyloid-state did not exhibit charge-transfer fluorescence presumably because of the protonation of charged glutamate residues into neutral glutamic acids and the altered conformational packing. Next, to directly visualize the intrinsic fluorescence within the droplets, we performed fluorescence microscopy using the DAPI channel (excitation 359 nm, emission 461 nm) that showed the presence of visible blue fluorescent droplets (inset in Figure 5.10f). Taken together, these results revealed two pH-regulated competing and mutually exclusive pathways of LLPS and amyloid formation. The phase-separated liquid droplets are characterized by high flexibility (low anisotropy), conformational disorder (CD), and higher charge-transfer

emission arising due to both charged amino acid side chains and hydrogen-bonded networks. Weak and transient side chain-side chain- and/or side chain-backbone contacts are believed to be respo-

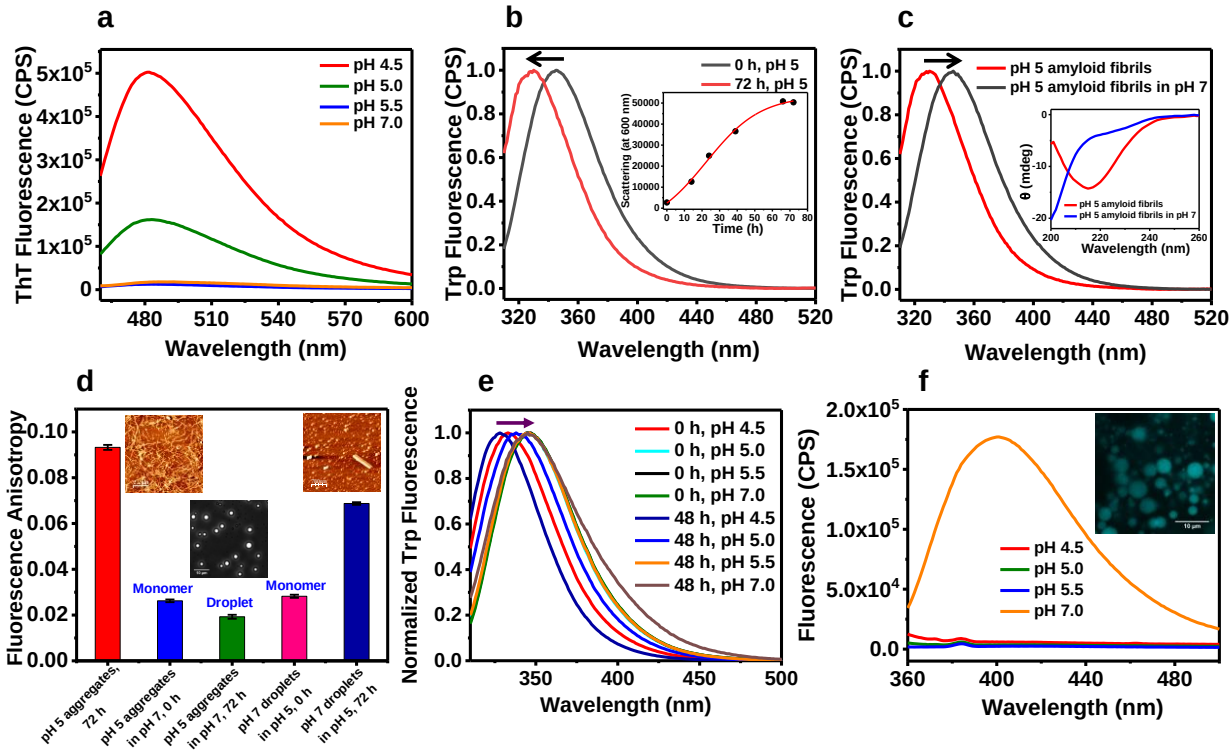


Figure 5.10 Phase transitions of RPT as a function of pH. (a) ThT fluorescence recorded after 48 h at pH 4.5 (red), pH 5.0 (olive), pH 5.5 (blue), and pH 7.0 (orange). (b) Steady-state fluorescence spectra of W423 RPT in monomeric (dark grey) and amyloid (red) state. An arrow indicates blue-shift in the fluorescence emission upon amyloid formation. The inset shows aggregation kinetics of RPT at pH 5 under stirring conditions obtained by recording scattering at 600 nm. (b) Steady-state fluorescence spectra of W423 RPT amyloid fibrils formed at pH 5 (red) when dissolved in pH 7 (dark). An arrow indicates red-shift in the fluorescence emission upon dissolution. Inset shows Far-UV CD spectra of RPT amyloid (red) and disassembled amyloids (blue). (c) Steady-state fluorescence anisotropy of W423 RPT amyloids (red) formed at melanosomal pH dissociate at cytosolic pH into monomers (blue) that forms liquid droplets (olive) when subjected to LLPS conditions which then recycle into monomers (pink), heterogeneous aggregates (navy blue) upon changing the solution pH to mildly acidic. Phase-contrast micrographs show pH-dependent droplet formation. (d) Trp fluorescence spectra (λex 280 nm) at different pH values at 0 and 48 h. The pH-dependent red-shift is shown by an arrow. (e) Charge-transfer fluorescence spectra (λex 340 nm) as a function of pH after 48 h of incubation. Inset: a confocal fluorescence image of the RPT droplets show intrinsic blue fluorescence. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

-nsible for maintaining the liquid-like nature of the droplets/condensates; however, persistent interactions can result in hardening of these droplets into gel-like or solid-like aggregates.[34] To

delineate the role of these interactions, next, we asked the following question: How do the side chains and backbones modulate phase separation and maturation?

5.3.6 Liquid-to-Solid Phase Transition of RPT.

Participation of Backbones and Side Chains. Prolonged incubation of RPT droplets at pH 7 resulted in an irreversible liquid-to-solid phase transition into gel-like/solid-like aggregates as observed on a microscope by monitoring the intrinsic blue fluorescence (Figure 5.11a). The solid

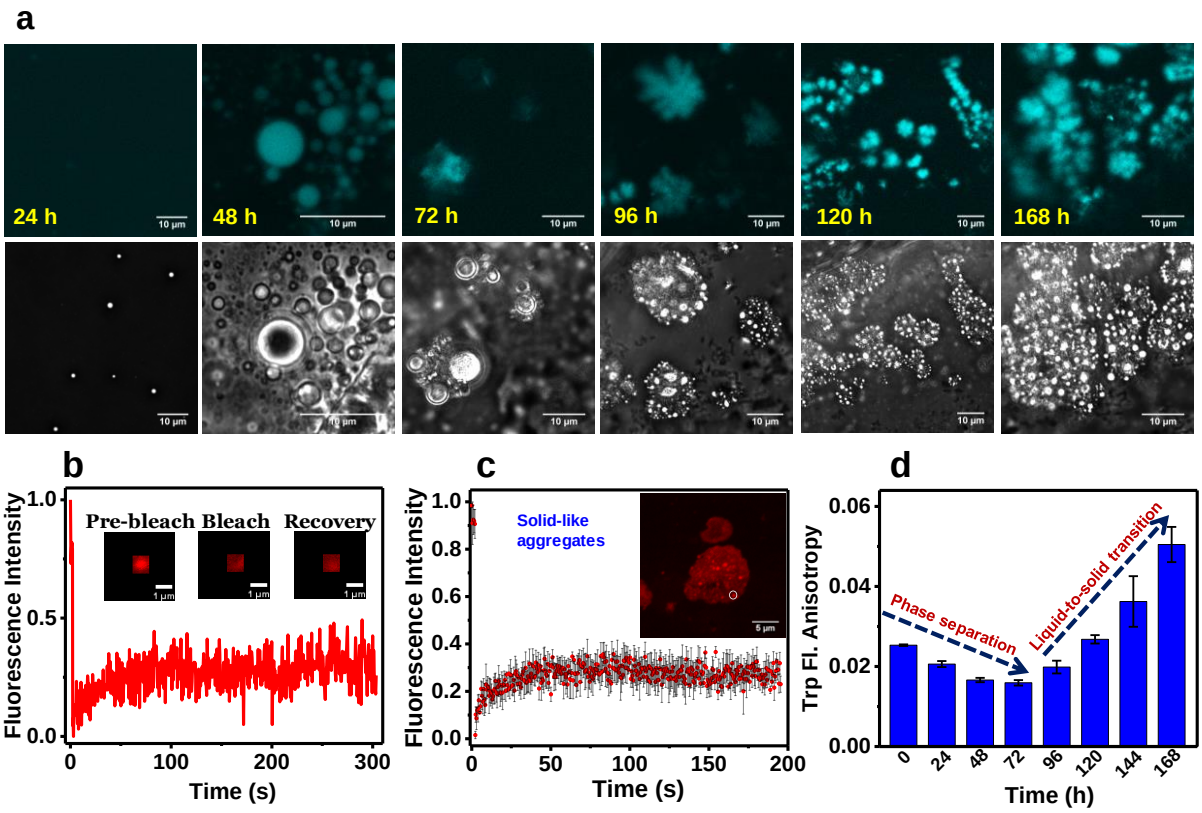


Figure 5.11 Liquid-to-solid transition of the RPT. (a) Confocal fluorescence- and phase contrast images showing aging and morphological changes of the RPT droplets. (b) A representative FRAP kinetics of the RPT showing a slow and partial recovery. The inset shows fluorescence images of droplets during FRAP measurements. (c) FRAP kinetic profiles of aged RPT droplets (n = 3 droplets) obtained after 168 h of incubation at 37 °C. Inset shows the confocal image of aged droplets of Alexa594 labeled RPT. (d) Initial decrease in the steady-state fluorescence anisotropy due to enhanced chain dynamics upon liquid phase condensation and the increase in the anisotropy due to liquid-to-solid phase transition. Data are shown as mean ± SEM (n = 3). Reprinted/adapted with permission from Dogra et al. (Ref. 66).

-like characteristic is also supported by a slow and incomplete FRAP recovery kinetics (Figure 5.11b,c). Additionally, a dip followed by a rise in the fluorescence anisotropy values suggested a

restriction in the chain dynamics arising due to a liquid-to-solid transition (Figure 5.11d). Since the more intimate intermolecular association is believed to yield solidification, we postulated that the maturation can be characterized by the proximal crosstalk between the backbones and side chains. To follow the appearance of these proximal interactions, we took advantage of fluorescence resonance energy transfer (FRET) between the indole side chain of Trp (donor) and the charge-transfer species (acceptor). As described above, we serendipitously discovered charge-transfer

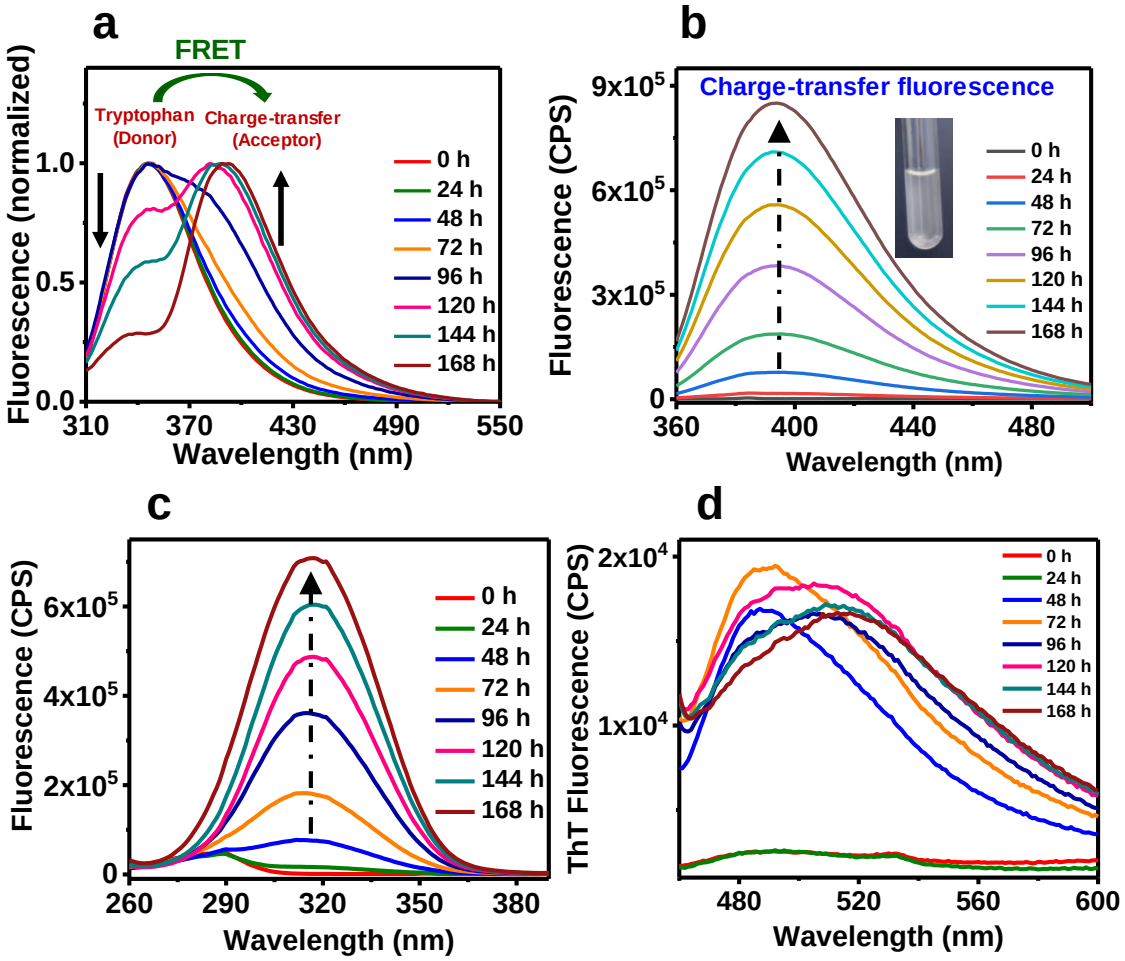


Figure 5.12 Side chains and backbones interaction during phase separation and maturation. (a) Fluorescence spectra of RPT showing FRET between side chain (Trp) and backbone (charge-transfer species) as a function of time at various stages of phase transition. (b) Enhancement in the charge-transfer fluorescence over the course of time. Inset shows a photograph of aged RPT droplets (solid-like aggregates). (c) Fluorescence excitation spectra during liquid-to-solid transition. (d) ThT fluorescence of 100 μM of RPT in 50 mM sodium phosphate buffer pH 7, 50 mM NaCl at 37 °C as a function of time. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

fluorescence by exciting Trp and obtaining a sensitized long-wavelength blue fluorescence which appeared due to FRET. In this set of experiments, we utilized Trp → charge-transfer FRET readouts to monitor the chain sequestration into a densely packed assembly leading to liquid-to-solid maturation. During the course of phase transition, we observed a drop in the Trp fluorescence (donor) with an increase in sensitized charge-transfer fluorescence (acceptor) due to strong FRET that grew further upon a liquid-to-solid transition (Figure 5.12a). A strong sensitized FRET-induced charge-transfer fluorescence indicated proximal and persistent backbone−backbone/backbone−side chain interactions upon the liquid-to-solid phase transition. We also observed an increase in the intensity of visible blue fluorescence by directly exciting the charge-transfer excitation band indicating more extensively networked polypeptide chains within the solid-like state (Figure 5.12b,c).

Protein Conformation. To test whether these solid-like aggregates contained β-rich amyloid-like structures, we carried out the ThT assay that indicated the presence of β-rich species (Figure 5.12d). Although these aggregates exhibited ThT fluorescence, the intensity was significantly lower compared to that of the amyloid-state formed at pH 5 (inset in Figure 5.13b). Additionally, a dot-blot assay and CD spectroscopy further confirmed the presence of β-sheet species (Figure 5.13a). Our vibrational Raman spectroscopic studies indicated distinct structural packing of RPT within the aggregates. In the case of pH 5 amyloids, we observed sharp amide I (1670 cm^{-1}) and amide III (1230−1240 cm^{-1}) bands that are the hallmarks of typical cross-β amyloid architecture. On the contrary, broad amide I and III bands for the solid-like state at pH 7 indicated the presence of heterogeneous packing at pH 7 (Figure 5.13b). Our AFM studies revealed the existence of fibrillar and nonfibrillar amorphous aggregates suggesting that the matured form of the phase-separated state contains heterogeneous supramolecular assemblies having lower β-content than typical amyloids (Figure 5.13c). Together this set of studies revealed that the liquid droplet state of the RPT is a metastable state that matures into a solid-like aggregated state via backbone−backbone/backbone−side chain interactions. The highly ordered amyloid aggregates formed at mildly acidic pH have distinct secondary structure and backbone packing compared to the heterogeneous aggregates formed at cytosolic pH via phase separation.

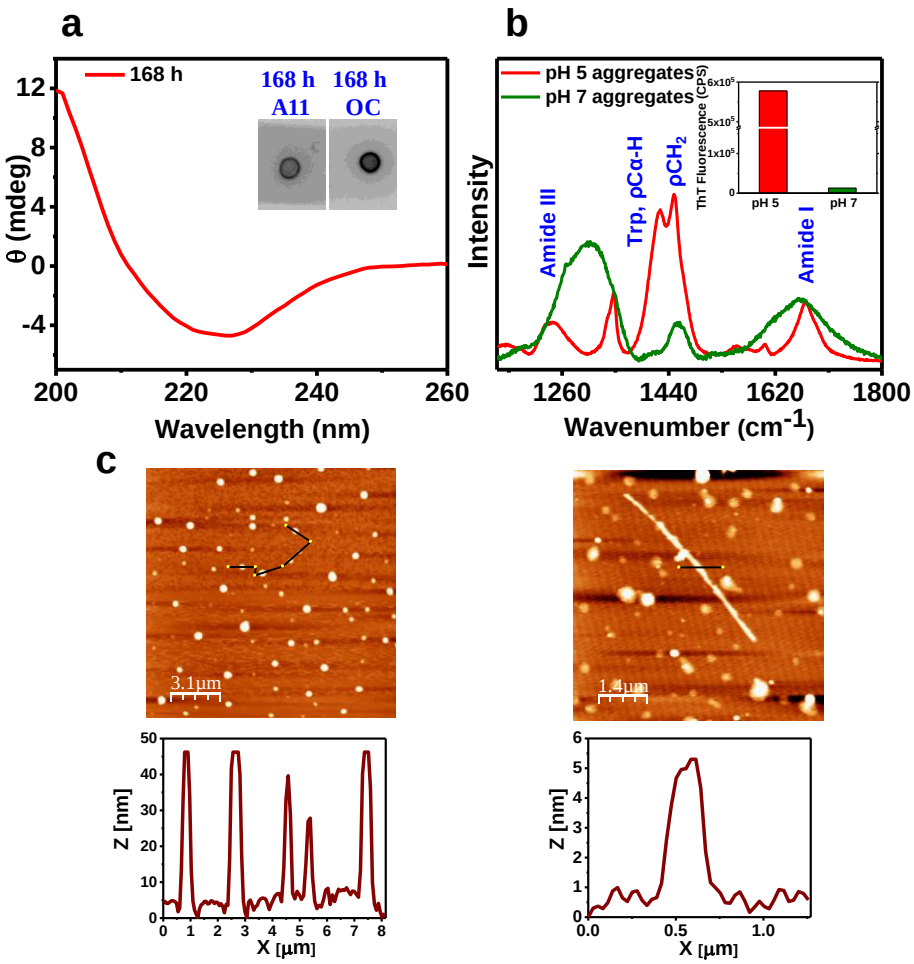


Figure 5.13 Conformation and nanoscale morphology of RPT aggregates (a) Far-UV CD spectra of RPT recorded at 168 h. Inset shows dot-blot analysis of 168-h sample when probed with A11 and OC antibody.(b) Raman spectra of RPT aggregates at pH 5 (red) and pH 7 (olive) at an excitation wavelength of 785 nm formed using 100 μM of RPT in 50 mM sodium acetate (pH 5.0) and sodium phosphate (pH7.0) buffer, 50 mM NaCl at 37 °C. The Raman peak at (1230-1240) cm^{-1} corresponds to amide III, tryptophan/Cα-H deformation (1348 cm^{-1}), CH_2 deformation (1450 cm^{-1}), and amide I (1670 cm^{-1}). (c) AFM images recorded after 168 h of the aged RPT sample (with height profile) showing the formation of heterogeneous aggregates. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

5.4 Discussion

In this work, we showed that the RPT, a pH-responsive disordered domain of a melanosomal protein, Pmel17, containing an LCR undergoes LLPS at cytosolic pH. The slow kinetics of LLPS may arise presumably due to the nucleation and growth mechanism in the RPT liquid phase condensation. These liquid-like condensates are characterized by fast internal diffusion, chain solvation, and rapid torsional fluctuations that in turn dictate the relay of making-and-breaking of

transient noncovalent contacts giving rise to the liquid-like interior. In addition to interchain electrostatic and hydrophobic stickers arising due to the associative motifs and segments that enable specific physical cross-links via non- covalent interactions, these assemblies are stabilized by extensive intermolecular charge-transfer interactions. These noncovalent interactions together with chain reorientation and solvation contribute to the enthalpic and the entropic terms in the modified Flory−Huggins theory.[12,19,21,22,30,32] The electrostatic and charge-transfer interactions can contribute to the enthalpy and rapid chain fluctuations can introduce the virial coefficient terms.[19] Whereas, the local hydrophobic contacts can cause the entropic release of water molecules. A complex interplay of these thermodynamic parameters allows RPT to display a unique dual UCST-LCST thermoresponsive phase behavior.[16,17] Liquid droplets upon aging undergo a liquid-to-solid

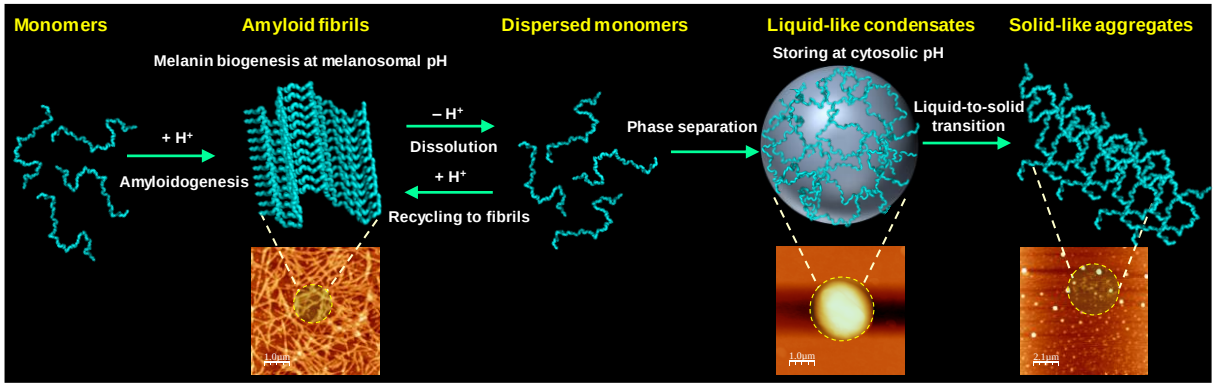


Figure 5.14 Proposed model for distinct pH-dependent pathways of phase separation/maturation and amyloid formation. Reprinted/adapted with permission from Dogra et al. (Ref. 66).

phase transition that is promoted by strong and persistent interactions between side chains and/or backbones. The solid-like state is structurally and morphologically distinct from typical amyloid fibrils and is characterized by intense blue fluorescence arising due to extensive intermolecular charge transfer via the hydrogen-bonded network. Our results reveal two distinct and competing pathways of pH-responsive RPT. Amyloidogenesis at mildly acidic melanosomal pH yields highly ordered β-rich fibrils, whereas, LLPS at neutral cytosolic pH promotes the formation of liquid droplets that are characterized by rapid chain fluctuations and lack of persistent secondary structures. Therefore, our findings highlight the regulatory role of RPT as a pH-sensor that strictly controls the amyloid formation and phase separation pathways. We postulate that LLPS can allow the melanosomal fragment to be stored as condensed liquid compartments at cytosolic pH. Upon

a pH change, the protein can undergo a rapid disorder-to-order amyloid transition that allows the melanin biosynthesis using fibrillar templates. On the basis of these observations, we propose a model that describes how amyloids formed at melanosomal pH dissociate at cytosolic pH into monomers that can recycle into fibrils upon switching the solution pH to mildly acidic or can readily convert into liquid droplets at cytosolic pH (Figure 5.14). These findings also highlight the fact that functionally controlled amyloid formation can occur only under a mildly acidic condition within melanosomes, and when the pH reaches close to neutrality, a competing pathway can potentially maintain the assembly integrity by liquid phase condensation retaining the melanin biogenesis ability. In summary, our studies illuminate an intriguing interplay of molecular drivers of phase separation and the potential regulatory role of a low-complexity disordered region as a specific pH-sensor similar to other prion-like low-complexity domains that regulate phase separation in vivo.[27] Our results providing the mechanistic underpinning of biological phase transition will be of importance to other IDPs/IDRs that constitute membraneless organelles with distinct cellular functions. Highly dynamic and permeable liquid organelles possessing a myriad of transient interchain contacts can undergo maturation into gel-like or solid-like aggregates via chain dehydration and sequestration. This solid-like state comprises extensive and persistent backbone interactions that can be characterized by the novel charge-transfer emission coupled with highly suppressed chain dynamics. The condensed liquid-like states can allow nucleation by lowering the free-energy barrier of self-association and by altering the conformational dynamics that can facilitate sequestration of polypeptide chains into highly ordered solid-like aggregates. Our intriguing observation suggests that the charge-transfer-mediated phase transition can potentially modulate the mesoscale material property of a multitude of biological supramolecular assemblies derived from highly charged biopolymers. Such unique dynamical readouts will also be useful for the study of other physiologically important intracellular phase separations of a variety of proteins and nucleic acids as well as for pathological liquid-to-solid phase transitions involved in deadly neurodegenerative diseases.

5.5 References:

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

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

(3) 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.

(4) 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.* **2017,** 56, 11354−11359.

(5) 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.

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

(7) Feric, M.; Vaidya, N.; Harmon, T. S.; Mitrea, D. M.; Zhu, L.; Richardson, T. M.; Kriwacki, R. W.; Pappu, R. V.; Brangwynne, C. P. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* **2016,** 165, 1686−1697.

(8) Mitrea, D. M.; Cika, J. A.; Stanley, C. B.; Nourse, A.; Onuchic, P. L.; Banerjee, P. R.; Phillips, A. H.; Park, C. G.; Deniz, A. A.; Kriwacki, R. W. Self-interaction of NPM1 modulates multiple mechanisms of liquid-liquid phase separation. *Nat. Commun.* **2018,** 9, 1−13.

(9) Shin, Y.; Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **2017,** 357, 4382.

(10) Weber, S. C.; Brangwynne, C. P. Getting RNA and protein in phase. *Cell* **2012,** 149, 1188−1191.

(11) Wang, Z.; Zhang, G.; Zhang, H. Protocol for analyzing protein liquid−liquid phase separation. *Biophys. Rep.* **2019,** 5, 1−9.

(12) Brangwynne, C. P.; Tompa, P.; Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015,** 11, 899−904.

(13) Protter, D. S. W.; Rao, B. S.; Van Treeck, B.; Lin, Y.; Mizoue, L.; Rosen, M. K.; Parker, R. Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep.* **2018,** 22, 1401−1412.

(14) 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.

(15) Lemke, E. A. The Multiple Faces of Disordered Nucleoporins. *J. Mol. Biol.* **2016,** 428, 2011−2024.

(16) Quiroz, F. G.; Chilkoti, A. Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat. Mater.* **2015,** 14, 1164−1171.

(17) Ruff, K. M.; Roberts, S.; Chilkoti, A.; Pappu, R. V. Advances in Understanding Stimulus-Responsive Phase Behavior of Intrinsically Disordered Protein Polymers. *J. Mol. Biol.* **2018,** 430, 4619−4635.

(18) Nott, T. J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T. D.; Bazett-Jones, D. P.; Pawson, T.; Forman-Kay, J. D.; Baldwin, A. J. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **2015,** 57, 936−947.

(19) 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, 1118−1125.

(20) 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, 8194−8203.

(21) Majumdar, A.; Dogra, P.; Maity, S.; Mukhopadhyay, S. Liquid-Liquid Phase Separation Is Driven by Large-Scale Conformational Unwinding and Fluctuations of Intrinsically Disordered Protein Molecules. *J. Phys. Chem. Lett.* **2019,** 10, 3929−3936.

(22) 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, 7189−7194. 631

(23) 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, 72−85.

(24) Reichheld, S. E.; Muiznieks, L. D.; Keeley, F. W.; Sharpe, S. Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** 114, 4408−4415.

(25) Wu, H.; Fuxreiter, M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **2016,** 165, 1055−1066. 642

(26) 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, 637−648.

(27) Franzmann, T. M.; Jahnel, M.; Pozniakovsky, A.; Mahamid, J.; Holehouse, A. S.; Nüske, E.; Richter, D.; Baumeister, W.; Grill, S. W.; Pappu, R. V.; Hyman, A. A.; Alberti, S. Phase separation of a yeast prion protein promotes cellular fitness. *Science* **2018,** 359, 5654.

(28) Franzmann, T.; Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* **2019,** 294, 7128−7136.

(29) 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.

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

(31) Simon, J. R.; Carroll, N. J.; Rubinstein, M.; Chilkoti, A.; Lopez, G. P. Programming molecular self-assembly of intrinsically disordered proteins containing sequences of low complexity. *Nat. Chem.* **2017,** 9, 509−515. 665

(32) Lin, Y. H.; Forman-Kay, J. D.; Chan, H. S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry* **2018,** 57, 2499−2508.

(33) 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.

(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) 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, 1066−1077.

(36) Molliex, A.; Temirov, J.; Lee, J.; Coughlin, M.; Kanagaraj, A. P.; Kim, H. J.; Mittag, T.; Taylor, J. P. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **2015,** 163, 123−133.

(37) Ambadipudi, S.; Biernat, J.; Riedel, D.; Mandelkow, E.; Zweckstetter, M. Liquid−liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat. Commun.* **2017,** 8, 1−13.

(38) 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.

(39) 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, 98049.

(40) Zhang, X.; Lin, Y.; Eschmann, N. A.; Zhou, H.; Rauch, J. N.; Hernandez, I.; Guzman, E.; Kosik, K. S.; Han, S. RNA stores tau reversibly in complex coacervates. *PLoS Biol.* **2017,** 15, No. e2002183.

(41) Lin, Y.; Protter, D. S.; Rosen, M. K.; Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* **2015,** 60, 208−219.

(42) 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.

(43) 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.

(44) Muiznieks, L. D.; Sharpe, S.; Pomes, R.; Keeley, F. W. Role of Liquid-Liquid Phase Separation in Assembly of Elastin and Other Extracellular Matrix Proteins. *J. Mol. Biol.* **2018,** 430, 4741−4753.

(45) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional amyloid formation within mammalian tissue. *PLoS Biol.* **2005,** 4, No. e6.

(46) McGlinchey, R. P.; Lee, J. C. Why Study Functional Amyloids? Lessons from the Repeat Domain of Pmel17. *J. Mol. Biol.* **2018,** 430, 3696−3706.

(47) Dogra, P.; Bhattacharya, M.; Mukhopadhyay, S. pH-Responsive Mechanistic Switch Regulates the Formation of Dendritic and Fibrillar Nanostructures of a Functional Amyloid. *J. Phys. Chem. B* **2017,** 121, 412−419.

(48) McGlinchey, R. P.; Shewmaker, F.; McPhie, P.; Monterroso, B.; Thurber, K.; Wickner, R. B. The repeat domain of the melanosome fibril protein Pmel17 forms the amyloid core promoting melanin synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** 106, 13731−13736.

(49) Pfefferkorn, C. M.; McGlinchey, R. P.; Lee, J. C. Effects of pH on aggregation kinetics of the repeat domain of a functional amyloid, Pmel17. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** 107, 21447−21452.

(50) McGlinchey, R. P.; Jiang, Z.; Lee, J. C. Molecular origin of pH-dependent fibril formation of a functional amyloid. *ChemBioChem* **2014,** 15, 1569−1572.

(51) Majumdar, A.; Mukhopadhyay, S. Fluorescence Depolarization Kinetics to Study the Conformational Preference, Structural Plasticity, Binding, and Assembly of Intrinsically Disordered Proteins. *Methods Enzymol.* **2018,** 611, 347−381.

(52) Lakowicz, J. R. Principles of Fluorescence Spectroscopy, 3rd ed.; Springer: New York, 2006.

(53) Jain, N.; Narang, D.; Bhasne, K.; Dalal, V.; Arya, S.; Bhattacharya, M.; Mukhopadhyay, S. Direct Observation of the Intrinsic Backbone Torsional Mobility of Disordered Proteins. *Biophys. J.* **2016,** 111, 768−774.

(54) Xue, B.; Dunbrack, R. L.; Williams, R. W.; Dunker, A. K.; Uversky, V. N. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **2010,** *1804*, 996-1010.

(55) Letunic, I.; Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **2018,** *46*, D493-D496.

(56) Meszaros, B.; Erdos, G.; Dosztanyi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **2018,** *46*, W329-W337.

(57) Horcas, I.; Fernandez, R.; Gomez-Rodriguez, J. M.; Colchero, J.; Gomez-Herrero, J.; Baro, A. M. WSXM: a software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* **2007,** *78*, 013705.

(58) Flynn, J. D.; Lee, J. C. Raman fingerprints of amyloid structures. *Chem. Commun.* **2018,** *54*, 6983-6986.

(59) 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.

(60) Zaslavsky, B. Y.; Uversky, V. N. In Aqua Veritas: The Indispensable yet Mostly Ignored Role of Water in Phase Separation and Membrane-less Organelles. *Biochemistry* **2018,** 57, 2437−2451.

(61) Shukla, A.; Mukherjee, S.; Sharma, S.; Agrawal, V.; Radha Kishan, K. V.; Guptasarma, P. A novel UV laser-induced visible blue radiation from protein crystals and aggregates: scattering artifacts or fluorescence transitions of peptide electrons delocalized through hydrogen bonding? *Arch. Biochem. Biophys.* **2004,** 428, 144−153.

(62) Del Mercato, L. L.; Pompa, P. P.; Maruccio, G.; Della Torre, A.; Sabella, S.; Tamburro, A. M.; Cingolani, R.; Rinaldi, R. Charge transport and intrinsic fluorescence in amyloid-like fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** 104, 18019−18024.

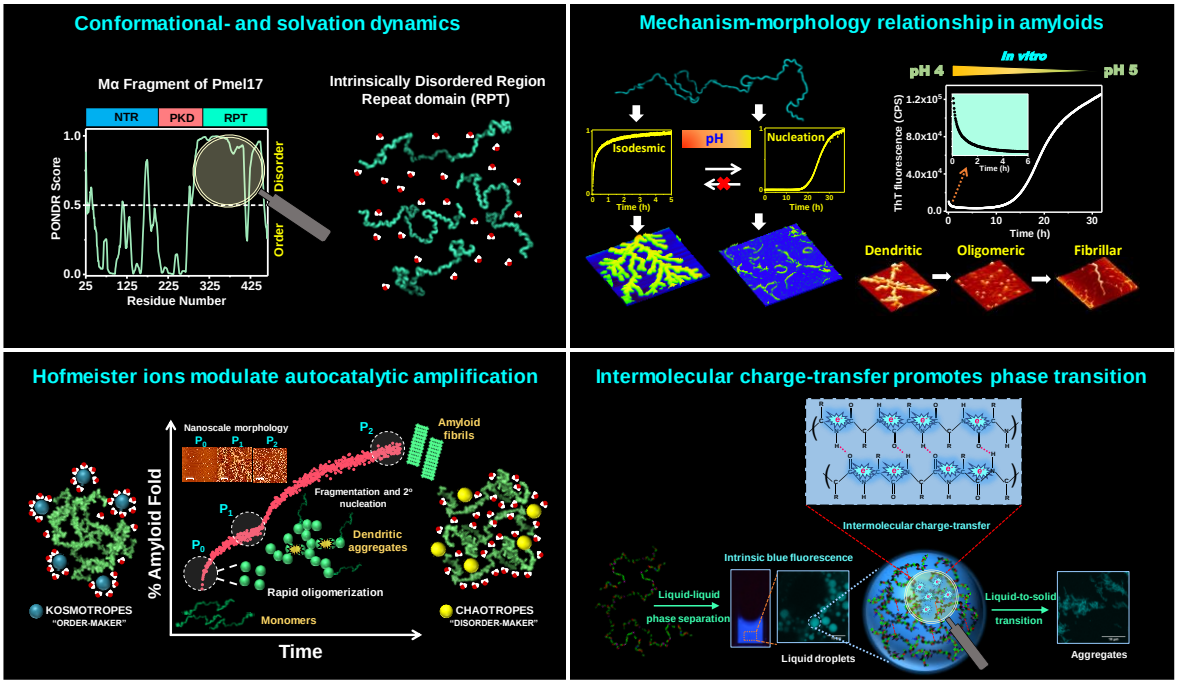
(63) Pinotsi, D.; Grisanti, L.; Mahou, P.; Gebauer, R.; Kaminski, C. F.; Hassanali, A.; Kaminski Schierle, G. S. Proton Transfer and Structure-Specific Fluorescence in Hydrogen Bond-Rich Protein Structures. *J. Am. Chem. Soc.* **2016,** 138, 3046−3057.

(64) Prasad, S.; Mandal, I.; Singh, S.; Paul, A.; Mandal, B.; Venkatramani, R.; Swaminathan, R. Near UV-Visible electronic absorption originating from charged amino acids in a monomeric protein. *Chem. Sci.* **2017,** 8, 5416−5433.

(65) Sharpe, S.; Simonetti, K.; Yau, J.; Walsh, P. Solid-State NMR characterization of autofluorescent fibrils formed by the elastin-derived peptide GVGVAGVG. *Biomacromolecules* **2011,** 12, 1546−1555.

(66) Dogra, P.; Joshi, A.; Majumdar, A.; Mukhopadhyay, S. Intermolecular Charge-Transfer Modulates Liquid–Liquid Phase Separation and Liquid-to-Solid Maturation of an Intrinsically Disordered pH-Responsive Domain. *J. Am. Chem. Soc.* **2019,** 141, 20380-20389.

Conclusions and future directions



The luminal fragment of a transmembrane glycoprotein, Pmel17, self-assembles to form the functional amyloid matrix within the early endosomes and melanosomes. These amyloid fibrils are generated in a regulated way in a directed pathway that can avert cytotoxicity. Pmel17 contains a highly amyloidogenic polypeptide repeat domain (RPT) that forms the amyloid core, which stimulates melanin biogenesis within the specialized skin cells called melanocytes. A growing body of evidence suggests that RPT aggregates into amyloid fibrils only under mildly acidic conditions, typical of melanosomes, and the fibrils dissociate at neutral pH. However, the molecular mechanism of amyloid formation, as well as the organization of individual protein molecules within the supramolecular assembly, remains elusive. The work described in this thesis addresses the underlying molecular mechanism of amyloid formation, liquid-liquid phase separation, and liquid-to-solid transition of the RPT. Below, I briefly summarize:

(i) The C-terminal proline-free amyloid core region has a slightly compact local structure, and the presence of ordered water molecules in this region reveals its crucial role in the aggregation of RPT. (Chapter 2),

(ii) pH-responsive mechanism-morphology relationship could be a strategic quality control mechanism employed during the melanosome maturation (pH 4 → pH 5), which might be critical for melanin biosynthesis. (Chapter 3)

(iii) Fragmentation, coupled with secondary nucleation processes, dictates the amyloid assembly of RPT in the presence of Hofmeister ions. (Chapter 4)

(iv) RPT undergoes liquid-liquid phase separation (LLPS) at cytosolic pH and forms highly dynamic liquid-like droplets under *in vitro* conditions that further matures to form solid-like aggregates. (Chapter 5)

For performing these studies, we used a diverse array of biophysical and biochemical tools, that include ultrafast spectroscopy, steady-state and time-resolved fluorescence spectroscopy, circular dichroism (CD) spectroscopy, Raman spectroscopy, atomic force microscopy (AFM), confocal microscopy, ultracentrifugation, chromatography, etc

The intricate balance between protein-protein and protein-solvent interaction governs several vital processes, such as protein folding and aggregation. IDPs/IDRs exist as dynamic ensembles of rapidly fluctuating conformations. The conformational preference of IDPs/IDRs is strongly coupled to the behavior of water molecules around the surface of the proteins. Therefore, the conformational preference of IDPs/IDRs can be characterized based on

solvation dynamic studies. In order to delineate the differences in the conformational states of intrinsically disordered RPT and discern the behavior of water molecules around the RPT, we employed time-resolved fluorescence studies. Using ultrafast laser spectroscopy coupled with picosecond time-resolved fluorescence depolarization kinetics measurements, we have been able to capture the structural and dynamical changes as well as surface water dynamics within different regions of natively unfolded RPT. We observed that the contribution of bound water relaxation dynamics was more in the case of the C-terminal amyloid core region of RPT than the N-terminal region. We propose that these differences in water arrangement are likely to translate distinct roles of different regions of RPT in aggregation, specific ion interactions, and amyloid formation (Chapter 2). We believe that this approach can be further extended to several other aggregation-prone IDPs/IDRs and might prove very promising for the designing of amyloid detection tools.

RPT is a pH-sensitive amyloidogenic domain that is known to aggregate under mildly acidic pH. The amino acid composition of RPT is enriched in carboxylic acid residues, which are primarily responsible for controlling the pH-responsive behavior of RPT. In vivo studies have revealed that during melanosome development, the pH changes from more acidic (pH ~ 4.0)

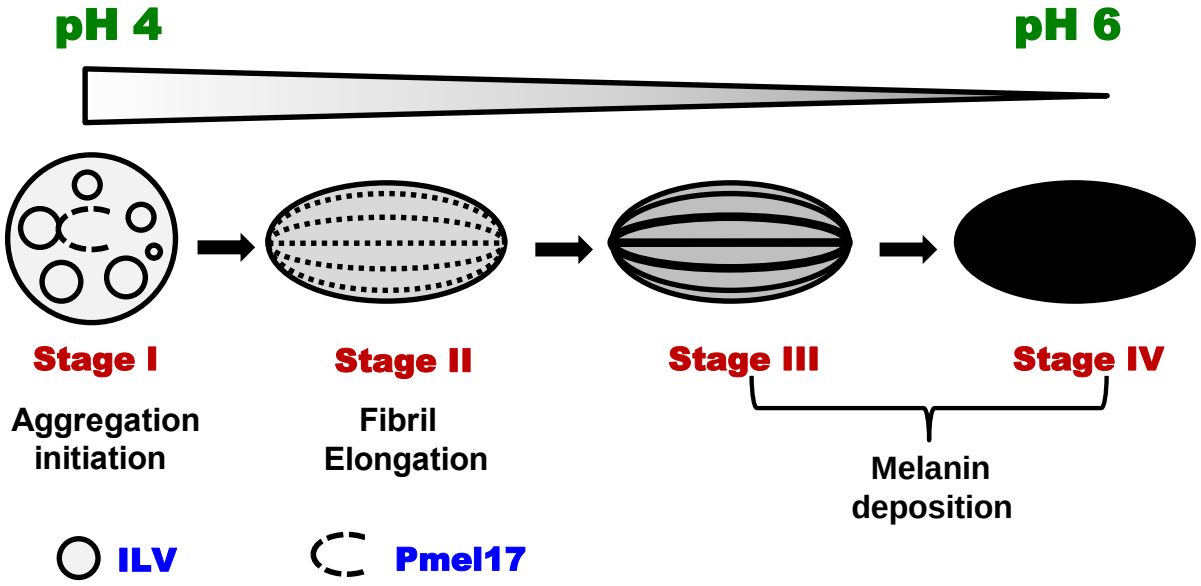


Figure 6.1. Schematic representation showing the four morphological stages of melanosome maturation. In stage I Pmel17 interacts with the intraluminal vesicles (ILVs), undergoes a conformational change, and generates pre-fibrillar aggregates. Stage II consists of long fibrillar striations that run in parallel arrays spanning the length of the melanosome. Stages III and IV contain melanins, synthesized from tyrosine by the enzyme, tyrosinase. The maturation of the melanosome is accompanied by a pH change, with a gradient ranging from 4.0–6.0.

to neutral (pH ~ 6) with different stages of development (Figure 6.1). Also, the amyloid fibrils mature during the developmental stages and have different morphologies in different stages. Stage I is composed of pre-fibrillar aggregates, while stage II is rich is long fibrous striations. We utilized a wide variety of biophysical and imaging tools to recapitulate the Stage I to Stage II morphological transition under *in vitro* conditions. Our results demonstrate that the RPT form different types of aggregates at pH 4 and pH 5 having distinct nanoscale morphology (Chapter 3). The aggregation kinetics at these two pH conditions reveal two different mechanisms of polymerization of the RPT. Furthermore, fibrils formed via the nucleation mechanism at pH 5 have higher-order structure and stability compared to those of dendritic nanostructures formed via the isodesmic mechanism at pH 4. Additionally, a pH-jump, akin to Stage I → Stage II, reveals a mechanistic switch in the aggregation pathway and conversion of less-ordered aggregates (seen in Stage I) into highly ordered amyloid fibrils (seen in Stage II). These results suggest that the pH modulation within the melanosomes allows the optimal conditions for the formation of functional amyloids that dictate the template-assisted melanin biosynthesis. We believe that these types of mechanistic studies involving pH-responsive regulation of nanostructures will also open new avenues for the fabrication of novel functional nano-biomaterials with a wide range of applications.

Specific salt-protein interaction plays key roles in governing different processes like protein folding, stability, and aggregation. We next embarked upon studies aimed at elucidating the role of different Hofmeister ions on the aggregation behavior of the RPT (Chapter 4). We observed that all ions examined, promote oligomerization, and the effect scales with the forward Hofmeister series. Interestingly, we observed an unusual biphasic aggregation kinetics of RPT in the presence of kosmotropes and chaotropes, that arises due to the dual forward and inverse Hofmeister effects. We demonstrate that this behavior is rooted in the structural sensitivity of the RPT to co-solutes, which in turn affects the oligomerization and self-assembly processes. Together, our findings reveal the role of different Hofmeister ions in regulating the autocatalytic amplification process involved in the aggregation pathway of RPT. We believe that these mechanistic underpinnings will be of significance in discerning the complex kinetics of several other IDPs/IDRs in the presence of ions.

Unlike the pathological amyloids, the amyloid fibrils of RPT dissociates at cytosolic pH. This peculiar property of RPT fibrils is proposed to be of functional significance in recycling the amyloids within melanocytes. However, pH-dependent phase transitions of RPT, ranging from melanosomal pH to cytosolic pH and the recycling mechanism, still remain elusive. We

discovered that at cytosolic pH, the RPT undergoes (liquid-liquid phase separation) LLPS *in vitro* and forms highly dynamic mesoscopic liquid-like droplets. Further, we show two distinct and competing pathways of pH-responsive RPT, where at mildly acidic melanosomal pH, RPT forms highly ordered β-rich fibrils, and at neutral cytosolic pH, it undergoes LLPS to form liquid droplets that further mature into solid-like aggregates. We next show that the amyloids formed at melanosomal pH dissociate at cytosolic pH into monomers. These monomers can then either recycle into fibrils under mildly acidic conditions or can readily convert into liquid droplets at cytosolic pH. We postulate that LLPS can allow the melanosomal fragment to be stored as condensed liquid compartments at cytosolic pH. Our findings highlight the regulatory role of RPT as a specific pH-sensor that stringently controls the amyloid formation and phase separation pathways. In Chapter 5, we present a unique case to demonstrate that an unusual cascade of intermolecular charge-transfer coupled with a multitude of transient noncovalent interactions and conformational fluctuations promote LLPS of RPT. Our intriguing observation suggests that the charge-transfer-mediated phase transition can potentially modulate the mesoscale material property of a multitude of biological supramolecular assemblies derived from highly charged biopolymers. We believe that these findings will be useful for the study of other physiologically relevant intracellular phase-separating proteins involved in physiology and disease.